Mesoscale Dynamics

By Yuh-Lang Lin
[Lin, Y.-L., 2007: *Mesoscale Dynamics*. Cambridge University Press, 630pp.]

13. Numerical modeling of geophysical fluid systems

Chapter 13 Numerical modeling of geophysical fluid systems

In the previous chapter, we discussed various numerical approximations of the advection equation. However, to simulate a geophysical fluid system, such as the atmosphere and ocean, within a finite region, we need to choose the domain size, grid size, time interval, total integration time, and consider other factors, such as the initial condition and boundary conditions. In addition, when we deal with a real fluid system, the governing equations are much more complicated than the one-dimensional, linear advection equation, as considered in the previous chapter. For example, we have to integrate three-dimensional nonlinear governing equations with several dependent variables, instead of a one-dimensional advection equation with only one variable. When a nonlinear equation is being approximated by numerical methods, one may face new problems such as nonlinear computational instability and nonlinear aliasing. Special numerical techniques are required to avoid these types of problems. Once optimal approximate forms of the equations are selected, it is still necessary to define the domain and grid structure over which the partial differential equations will be approximated. In this chapter, we will also briefly describe on how to build a basic numerical model based on a set of partial differential equations governing a shallow water system, and a hydrostatic or nonhydrostatic continuously stratified fluid system.

13.1 Grid systems and vertical coordinates

The first step in developing a mesoscale numerical model is to determine appropriate domain size, grid intervals, time interval, and total integration time of the model. Selection of the domain size, grid interval, total integration time and time interval in a mesoscale model is usually based on both physical and numerical factors, such as: (1) spatial scales and dimensionality of the forcing and physical processes, (2) time scales of the forcing and the fluid responses to the forcing, (3) stability criterion of the adopted numerical scheme, (4) limitations of predictability of the atmospheric phenomena, and (5) availability of computer resources. To represent mesoscale atmospheric systems properly, it is required that: (a) the meteorologically significant variations in the dependent variables caused by the mesoscale forcing and fluid responses be contained within the model domain, and that (b) the averaging volume used to define the model grid spacing be small enough for the mesoscale forcing and responses to be accurately represented.

In making mesoscale numerical weather prediction (NWP), one needs to accurately represent multiscale processes in a finite domain. On the one hand, it is important to capture the smaller-scale weather systems, requiring the grid and time intervals to be fine enough to resolve small-scale processes. On the other hand, it is equally important to capture the larger-scale environment, which is responsible for the formation and modification of the smaller-scale weather systems. Therefore, the model domain should be large enough to contain the essential large-scale environment. Problems arise if a domain is inappropriately selected. For example, Fig. 13.1 shows two identical simulations for a hydrostatic flow over a bell-shaped mountain except with different

domain sizes. With a smaller domain (Fig. 13.1b), the simulated horizontal wind field is more horizontally oriented and the wave weakens more rapidly with height. The difference in the horizontal orientation is due to that the upstream flow conditions at the left boundary, such as the wind speed and Brunt-Vaisala frequency of the basic flow, have been affected by the upstream propagating waves generated by the mountain. The flow field in the interior of the domain has thus been affected artificially by the wave reflection from the lateral boundaries, which means part of the waves cannot propagate upward, as the hydrostatic waves are supposed to do.

Once the domain interval is chosen, the next natural step is to choose an appropriate grid size (interval). The choice of grid interval used in a numerical model depends on spatial scales of the dominative forcing and flow responses. The grid interval should represent the forcing and fluid responses well; otherwise the simulated results will not be accurate. For example, in simulating a stratified fluid flow over a region with orographic or thermal forcing, the grid interval should be fine enough to properly represent the geometry of the forcing. Normally, this can be done by inspecting the smoothness of the mountain shape, but it can also be carried out more rigorously by performing power spectrum analysis which gives a plot of the portion of a signal's power (energy per unit time) falling within given frequency range. Numerous examples can be found for the impacts of the grid interval on simulations of weather systems in the literature.

After choosing the domain size, grid interval and grid structure, the next step is to choose the time interval, which is normally determined by the time variation of the forcing and stability criterion. The latter depends on the numerical method adopted. The linear stability criterion, as discussed in the last chapter, can serve, as a first guess of the

time interval needed to guarantee the computational stability of the model. As will be discussed later, nonlinear equations have a stricter criterion on computational stability, which is related to the time scale and the predictability of the weather phenomenon interested, as well as the computing resources available.

13.1.1 Grid systems

Different types of grid meshes have been adopted in mesoscale NWP models, such as (a) uniform grid mesh, (b) stretched grid mesh, (c) nested grid mesh, (d) movable grid mesh, (e) adaptive grid mesh, and (f) staggered grid mesh. Some of these grid meshes can be combined in a model.

For the *uniform grid mesh*, grid intervals are set to be equal in horizontal or in vertical. The advantages of this type of grid mesh are that it is relatively easy to code onto the computer, and it is relatively simple to input geographic features into the model. The disadvantages of the constant grid mesh include that it is difficult to properly incorporate both large and small features within the same model domain. For example, if one uses the same grid interval in vertical, it will be difficult to properly resolve the boundary layer circulation, while it is more than enough to resolve circulation in the free atmosphere, especially in the upper troposphere.

In order to improve the computing efficiency, *stretched grid mesh* is used, in which the grid intervals vary in space. The advantage of the stretched grid mesh is that with the same computing time, a much larger domain than the uniform grid mesh with the same number of grid points can be adopted for numerical simulations. Adopting a grid mesh stretched from a finer resolution near the surface to a relatively coarse resolution in the

upper layer allows a model to resolve some smaller-scale disturbances present in the planetary boundary layer. If the grid intervals are stretched too abruptly, internal reflection may occur. In many mesoscale numerical simulations, the grid mesh is often stretched. The vertical turbulent mixing provides a measure of the needed grid resolution. The vertical stretch of the grid intervals may be defined by a known function, such as a logarithmic function, or specified by particular heights or horizontal distances. A two-way stretching may also be used. For example, in order to better represent the waves near a critical level, one may need to specify the grid mesh with very fine resolution near the critical level and stretches to coarser resolution both upward and downward.

An alternative approach to the stretched grid mesh is to insert a fine-mesh grid mesh, i.e. a *nested grid mesh*, within a coarse grid, which allows one to simulate smaller-scale features in the nested grid meshes. The smaller-scale features are not resolvable by the coarse grid mesh. During the simulation, the coarse mesh provides the boundary conditions for the fine mesh. There are two types of grid nesting techniques: (i) one-way nesting which only permits disturbances in the outer grid mesh to enter the finer nested grid mesh, but not the other way around, and (ii) two-way nesting in which the boundary values of the inner grid mesh are passed back to the outer grid mesh. Note that under the nested grid mesh system there is a discontinuity across the boundaries of the fine and coarse grid meshes. Nested-grid mesh techniques have been used in many research and operational models. The above-mentioned nested grid mesh can also be designed to move with a weather system, such as a thunderstorm, mesoscale convective system, front, midlatitude cyclone or tropical cyclone. The advantage of a *moving grid mesh* is

that it can follow the weather system and always provides a finer resolution needed in the vicinity of it.

To improve the nested and moving grid mesh techniques, *adaptive grid mesh* techniques have been developed. Grid points can be added in a structured manner through the placement of multiple and perhaps overlapping finer-scale grids in the domain. Figure 13.2 shows one example of using 4 adaptive grid meshes to simulate a cold pool collapse. Regions of strong potential temperature gradients along the gust front and Kelvin-Helmholtz billows are well simulated by using the adaptive grid mesh. In the adaptive grid mesh, a fixed number of grid points or collocation points may also be redistributed in a predetermined manner to provide locally increased resolution and thus an improved solution in certain regions of the domain. Figure 13.3 shows a simulation of kinematic frontogenesis, which is similar to a smoothed Rankine vortex being advected by a steady, nondivergent field, using this type of structured *continuous dynamic grid adaptation* (CDGA). With both 31x31 grid points, the simulated kinematic frontogenesis being advected by a steady, nondivergent flow field with structure similar to that of a smoothed Rankine vortex (Fig. 13.3c) by using CDGA (Fig. 13.3d) is much better than that of uniform grid mesh (Fig. 13.3b), when comparing with the exact solution (Fig. 13.3a). One limitation of this type of structured adaptive grid mesh is that it is not suitable for dynamic grid adaptation because the grid generation requires a high degree of user interaction and user expertise. Thus, it is not an easy task to apply this type of method for real-case simulations. In order to resolve the problem, the *unstructured adaptive grid mesh* has been proposed in dealing with both large- and small-scale features without having to use a nested grid mesh (e.g., Bacon et al. 2000).

In order to preserve some conservation laws, a *staggered grid mesh* is used. In a staggered grid mesh, variables of a system of differential equations are defined at different grid points which are staggered with respect to each other. To elucidate the formulation of a staggered grid mesh, let us consider the two-dimensional incompressible continuity equation:

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 . \tag{13.1.1}$$

For an unstaggered grid mesh (Fig. 13.4a), a simple finite difference form can be written

$$w_{i,k} = w_{i,k-1} - \frac{u_{i+1,k-1/2} - u_{i-1,k-1/2}}{2\Delta x} \Delta z , \tag{13.1.2}$$

where

$$u_{i+1,k-1/2} = (u_{i+1,k} + u_{i+1,k-1})/2; \;\; u_{i-1,k-1/2} = (u_{i-1,k} + u_{i-1,k-1})/2 . \tag{13.1.3}$$

For a staggered grid mesh (Fig. 13.4b), a simple finite difference form can be written

$$w_{i,k} = w_{i,k-1} - \frac{u_{i+1/2,k-1/2} - u_{i-1/2,k-1/2}}{\Delta x} \Delta z , \tag{13.1.4}$$

where u is defined at half-way between the grid points at which w is defined. Thus, staggering the dependent variables as given by (13.1.4) increases the effective resolution by a factor of two, since derivatives are defined over an increment Δx, for instance, rather than $2\Delta x$, yet without requiring averaging as in (13.1.2).

To examine the computational stability and phase velocity associated with a staggered grid mesh, we can consider applying the leapfrog in time and second-order centered difference scheme to the two-dimensional, linear shallow water equations, (3.4.7) and (3.4.9), with $U = 0$ on a staggered grid mesh as shown in Fig. 13.5:

$$\frac{u_i^{\tau+1} - u_i^{\tau-1}}{2\Delta t} + g\frac{h_{i+1/2}^{\tau} - h_{i-1/2}^{\tau}}{\Delta x} = 0, \tag{13.1.5}$$

$$\frac{h_{i+1/2}^{\tau+1} - h_{i+1/2}^{\tau-1}}{2\Delta t} + H\frac{u_{i+1}^{\tau} - u_i^{\tau}}{\Delta x} = 0. \tag{13.1.6}$$

The numerical dispersion relationship for the above system is

$$\sin \omega\Delta t = \pm\frac{2c\Delta t}{\Delta x}\sin\frac{k\Delta x}{2}, \tag{13.1.7}$$

where $c = \pm\sqrt{gH}$ are the shallow-water phase speeds. Since the solutions of the two-dimensional shallow water wave system are neutral, a real ω is required, which implies that a stable solution requires $|C| = |c\Delta t/\Delta x| \le 1/2$. The stability criterion for an unstaggered grid mesh can be derived to be $|C| \le 1$. Therefore, the maximum time interval for a staggered grid system is half of the corresponding unstaggered mesh system, which implies that the computational time is almost doubled. The more stringent requirement on integration time interval may be compensated for by an improved computational phase speed in using a staggered grid mesh, as shown in Fig. 13.6. Figure 13.7 shows five grid meshes proposed by Arakawa and Lamb (1977). For staggered grid meshes, grid meshes B and C, often referred to as *Arakawa-B grid* and *Arakawa-C grid*, respectively, can better preserve the phase speed and group velocity.

13.1.2 Vertical coordinates

In simulating a mesoscale flow within a finite domain, the height vertical coordinate may propose problems. For example, it may intercept the terrain in a mountainous area and thus create problems in dealing with the lower boundary condition. Similar problems

happen to the pressure coordinate and isentropic coordinate when isobaric surfaces and isentropic surfaces intercept the lower boundary, respectively, which may occur when there is strong orographic blocking.

To avoid the problem, a vertical *sigma coordinate*, which matches the lowest coordinate surface with the bottom topography, has been proposed (Phillips 1957). In this type of *vertical* $\sigma - p$ *coordinates*, the pressure coordinate is normalized by the surface pressure, p_s,

$$\sigma = \frac{p}{p_S}. \tag{13.1.8}$$

Thus, $\sigma = 1$ at the surface, and $\sigma = 0$ at the top of the atmosphere. The σ vertical velocity, $\dot{\sigma} = D\sigma / Dt$, is 0 at both the surface and top of the atmosphere. In (13.1.8), σ can also be defined as $\sigma = (p - p_T)/(p_S - p_T)$, where p_T is the pressure at the top of the numerical domain. The same concept may also be applied to the *isentropic coordinates*, in which σ can be defined as

$$\sigma = \frac{\theta - \theta_S}{\theta_T - \theta_S}. \tag{13.1.9}$$

The coordinate system in (13.1.9) is called $\sigma - \theta$ *coordinates*. One of the advantages of the vertical isentropic coordinate is that it can better resolve the vertical structure of weather systems, such as tropopause folding and upper-level frontogenesis. When the sigma coordinate is applied to the height coordinates, it is called $\sigma - z$ or the *terrain-following coordinates* in which the σ can be defined as

$$\sigma = \frac{z_T(z - z_S)}{z_T - z_S}, \tag{13.1.10}$$

where z_S is the height of the lower surface in the $\sigma - z$ coordinates, which is independent of time, and z_T is the constant domain height or the constant height of the terrain-following part of the domain. In the general σ coordinates, the pressure, *p*, can be written as

$$p(x, y, z, t) = p[x, y, \sigma(x, y, z, t), t]. \tag{13.1.11}$$

The pressure gradient in *x* direction in the *z* coordinates can be obtained by performing the chain-rule,

$$\left(\frac{\partial p}{\partial x}\right)_z = \left(\frac{\partial p}{\partial x}\right)_\sigma + \frac{\partial p}{\partial \sigma}\left(\frac{\partial \sigma}{\partial x}\right)_z. \tag{13.1.12}$$

The pressure gradient ($\partial p / \partial x$) in the σ coordinates can be obtained by deriving $(\partial \sigma / \partial x)_z$ from (13.1.10) and substituting it into (13.1.12):

$$\left(\frac{\partial p}{\partial x}\right)_\sigma = \left(\frac{\partial p}{\partial x}\right)_z - \left[\frac{\sigma - z_T}{z_T - z_S}\frac{\partial z_S}{\partial x}\right]\frac{\partial p}{\partial \sigma}. \tag{13.1.13}$$

The sigma coordinate transformation may also be applied to the *mass* (*hydrostatic pressure*) *coordinates* (Skamarock et al. 2005), in addition to the $\sigma - z$ coordinates. If one replaces *p* by a general variable *A*, then the above transformation may be used to derive the gradient of *A* in *x* direction.

One problem of the sigma vertical coordinate systems is that errors in two terms of the pressure gradient force do not cancel out (Smagorinski et al. 1967). To avoid this problem, the *step-mountain* or *eta coordinates* have been proposed (Mesinger et al. 1988). In that model, *eta* (η) is defined as

$$\eta = \frac{p - p_T}{p_S - p_T}\eta_S \tag{13.1.14}$$

with

$$\eta_S = \frac{p_r(z_S) - p_T}{p_r(0) - p_T}. \qquad (13.1.15)$$

In the above equations, p is pressure; the subscripts T and S denote the top and surface values of the model atmosphere; z is geometric height, and $p_r(z)$ is a suitably defined reference pressure as a function of z. The boundary conditions assumed are p=constant at the top boundary ($\eta = 0$), and $\dot{\eta} = 0$ at $\eta = 0$ and at the horizontal parts of the ground surface ($\eta = \eta_S$). The advantage of this approach is that it does not transform the governing equations into complicated forms. However, the disadvantage of this coordinate is that it is of the first order of accuracy in representing the terrain, which is less accurate compared with the terrain-following coordinates, which is of the second-order of accuracy. A third method is to adopt a *finite-element scheme*, which approximates the mountain surface by one side of the finite elements. It has the same advantage as not having to transform the governing equations into complicated forms as well as having a higher-order accuracy compared to the step-mountain coordinates.

So far, we have discussed about how to set the domain size, vertical and horizontal grid intervals, grid mesh and vertical coordinate. However, we still need to find (a) the temporal values of variables at the beginning of the integration, and (b) the boundary values of a finite domain. These are required by the fact that we are solving the initial-boundary value problems mathematically.

13.2 Boundary conditions

To make the mathematical problem well-posed, appropriate boundary conditions need to be specified in any limited-area models, such as mesoscale NWP models. If the model domain represents only part of the atmosphere in every direction, then boundary conditions are needed at the top, lateral, and lower boundaries of the model domain. The number of boundary conditions depends on the order of the differential equations in a particular direction involved.

13.2.1 Lateral boundary conditions

Basically, there are five types of lateral boundary conditions adopted in mesoscale NWP models: (1) closed boundary condition, (2) periodic boundary condition, (3) time-dependent boundary condition, (4) sponge boundary condition, and (5) open (radiation) boundary condition.

In a *closed lateral boundary condition*, variables at the lateral boundaries are specified as constant values. Under such a situation, the waves or disturbances generated within the domain cannot propagate out of the domain; instead they are reflected back into the domain once they reach the lateral boundaries. The reflection of waves and disturbances back into the domain gives the name "closed boundary condition", since the model domain retains such disturbances. The use of the closed lateral boundary conditions works if the actual physical condition is closed, such as a solid wall. In addition, the boundary conditions will also work if the lateral boundaries are located far away so that the generated disturbances or waves do not reach the lateral boundaries at the desired integration time.

A *periodic boundary condition* assumes all the variables at the right boundary are equal to the left boundary, i.e. $\phi(x_N) = \phi(x_0)$, where ϕ is a field variable, or vice versa. Period boundary condition is an appropriate choice for a fluid flow system which does repeat itself at the lateral boundaries, such as sinusoidal waves propagating around a latitudinal circle of the Earth. A *time-dependent lateral boundary condition* is often adopted when numerical integration are performed concurrently at both the inner and outer domains, the lateral boundary values of the inner domain need to be specified by the updated values predicted by the outer domain. In this way, the weather systems, waves, or disturbances are able to propagate into the inner domain and produce the desired weather systems or fluid motion. Otherwise, the simulations of the inner domain cannot reflect the larger-scale environmental changes with time, which is called *one-way nesting*. If the lateral boundary values of the inner domain are passed back to the outer domain, then it is called *two-way nesting*.

The *sponge* or wave-*absorbing layer boundary condition* uses an enhanced filtering near the lateral boundaries to damp the waves or disturbances generated within the domain out of the lateral boundaries. For example, a sponge region can be formulated the as follows (Perkey and Kreitzberg 1976):

$$\frac{\partial \phi_i}{\partial t} = W_i \left(\frac{\partial \phi_i}{\partial t} \right)_m + (1 - W_i) \left(\frac{\partial \phi_i}{\partial t} \right)_{ls}, \qquad (13.2.1)$$

where m denotes the model calculated tendency of variable ϕ, ls denotes the larger scale specified tendency and W_i is a weighting factor which is given as follows:

W_i = 0.0 for the boundaries for i = 0 or N

= 0.4 for the boundaries for i = 1 or N - 1

= 0.7 for the boundaries for i = 2 or N - 2

= 0.9 for the boundaries for i = 3 or N - 3

= 1.0 for all the interior points of $4 \le i \le N-4$. (13.2.2)

Even though the wave-absorbing layer oversimplifies the boundary conditions, it has considerable practical applications (Davies 1983).

For pure gravity waves, the horizontal phase velocity is directed in the same sense as the horizontal *group velocity*, such as those shown in (3.5.14) and (3.5.15): thus it is possible to use the advection equation to advect the *wave energy* out of the lateral boundaries. Based on this concept, an *open* (*radiation*) *boundary condition* for a hyperbolic flow in a numerical model was proposed (Orlanski 1976). For the outflow boundary, the radiation boundary condition may be written

$$\frac{\partial \phi}{\partial t} + (U + c_o^*)\frac{\partial \phi}{\partial x} = 0, \qquad \text{at } x = L, \tag{13.2.3}$$

where $U + c_o^*$ is the propagation speed at the outflow boundary ($x = L$), which is yet to be determined. The leapfrog finite difference representation for time step τ - 1 of the above equation may be written

$$\frac{\phi_{b-1}^{\tau} - \phi_{b-1}^{\tau-2}}{2\Delta t} = \frac{-(U + c_o^*)}{\Delta x}\left(\phi_{b-1}^{\tau-1} - \phi_{b-2}^{\tau-1}\right). \tag{13.2.4}$$

Based on the above approximation, the phase speed can be estimated by

$$U + c_o^* = -\frac{\Delta x}{2\Delta t}\left(\frac{\phi_{b-1}^{\tau} - \phi_{b-1}^{\tau-2}}{\phi_{b-1}^{\tau-1} - \phi_{b-2}^{\tau-1}}\right) \qquad \text{at } b-1, \tag{13.2.5}$$

$$U + c_o^* = 0 \qquad \text{if R.H.S. of (13.2.5) } < 0,$$

$$U + c_o^* = \frac{\Delta x}{\Delta t} \qquad \text{if R.H.S. of (13.2.5)} > \Delta x / \Delta t,$$

where subscript *b* denotes the boundary point and R.H.S. means the right-hand side. For a hydrostatic and incompressible fluid system, since *w* is coupled with *u*, one can use the estimated phase speed of *u* for *w*. The phase speed estimation can also be applied to the coupled variables of potential temperature and pressure. Once the phase speed is estimated, then the boundary value at time step $\tau+1$ can be determined

$$\phi_b^{\tau+1} = \phi_b^{\tau-1} - \frac{2\Delta t}{\Delta x}(U + c_o^*)\left(\phi_b^{\tau} - \phi_{b-1}^{\tau}\right), \tag{13.2.6}$$

where $U + c_o^*$ is estimated by (13.2.5). A similar formula can be formed for the inflow boundary:

$$\phi_b^{\tau+1} = \phi_b^{\tau-1} - \frac{2\Delta t}{\Delta x}(U - c_i^*)\left(\phi_{b+1}^{\tau} - \phi_b^{\tau}\right). \tag{13.2.7}$$

Note that the specification of ϕ at both boundaries will lead to an overdetermined problem for the first-order advection equation. In fact, this renders the problem ill-posed (Oliger and Sundström 1978). Nevertheless, wave-absorbing layers have considerable practical utility even when they require overspecification of the boundary conditions (Davies 1983).

In practice, the *constant gradient lateral boundary condition* has also been adopted in mesoscale NWP models, which specifies a constant gradient, such as zero gradients (e.g., $\partial\phi / \partial x = 0$, ϕ is the variable concerned), at the lateral boundaries. The effectiveness of the constant gradient lateral boundary condition to propagate the waves out of the domain depends on how close the specified constant gradients are to the advection speed of the physical waves. The constant gradient lateral boundary conditions can be viewed as a special case of the radiation boundary condition. For example, a *zero gradient lateral*

boundary condition, such as $\phi_N^{\tau+1} = \phi_{N-1}^{\tau}$, assumes the wave propagating out of the right boundary at a speed of $c = \Delta x / \Delta t$. If the real physical wave speed is very different from this numerical phase speed, then a large reflection from the boundary may occur.

13.2.2 Upper boundary conditions

The upper boundary of a mesoscale NWP model should be placed high enough above the region with active mesoscale waves and disturbances. Ideally, it should be placed at the top of the atmosphere, i.e. p = 0. However, practically, this is impossible due to the restriction of computing resources. Depending on the weather phenomena simulated by the model, it may require the top boundary of a numerical model domain be placed at deep within the stratosphere, at the tropopause, or within the stable layer of the troposphere. For example, a sea breeze circulation in a stable boundary layer normally does not penetrate to a high altitude, thus allowing a top boundary of a mesoscale model to be placed at the mid-troposphere. On the other hand, in simulating a flow over a mesoscale mountain, the mountain waves often can propagate to a very high altitude, therefore a much higher vertical domain is required. No matter how high the model domain extends in vertical, an appropriate upper boundary condition is still needed. For example, the disturbances or waves generated by a mesoscale mountain in the lower stratosphere may look like very weak, but the energy per unit area associated with them could be very large because it is proportional to N (e.g. see (13.2.31)) and N is much larger in the stratosphere compared to that in the troposphere.

Upper boundary conditions that have been adopted in mesoscale NWP models can be categorized as (1) rigid lid, (2) sponge layer (Klemp and Lilly 1978; Anthes and Warner

1978; Mahrer and Pielke 1978), and (3) radiation condition (Klemp and Durran 1983; Bougeault 1983). A *rigid lid* upper boundary condition can be implemented by simply setting the desired variables to be constants. The sponge layer and radiation conditions are found to be effective in radiating the wave energy out from the interior of the numerical domain. Basically, these two approaches are taken to numerically approximate the *Sommerfeld* (1949) *radiation boundary condition*, which allows the energy associated with disturbances generated in the interior of a physical system to propagate out of the domain.

The addition of a *sponge* (*wave-absorbing*) *layer* to the top of the *physical domain* is a simple way to mimic the Sommerfeld (1949) radiation boundary condition in a numerical model. The sponge layer is designed to damp out disturbances generated in the physical layer out of the upper boundary. To elucidate the formulation of a sponge layer, we consider the following two-dimensional, steady-state, linear, hydrostatic, nonrotating, Boussinesq flow with Rayleigh friction and Newtonian cooling added to the momentum and thermodynamic equations, respectively:

$$U\frac{\partial u'}{\partial x}+\frac{1}{\rho_o}\frac{\partial p'}{\partial x}=-\nu u', \tag{13.2.8}$$

$$\frac{\partial p'}{\partial z}-\left(\frac{g\rho_o}{\theta_o}\right)\theta'=0, \tag{13.2.9}$$

$$\frac{\partial u'}{\partial x}+\frac{\partial w'}{\partial z}=0, \tag{13.2.10}$$

$$U\frac{\partial \theta'}{\partial x}+\left(\frac{N^2\theta_o}{g}\right)w'=-\nu\theta'. \tag{13.2.11}$$

The above equations may be obtained from (3.5.1)-(3.5.4). To minimize reflections caused by rapid increases in viscosity, one may consider the following function, which gradually increase ν from 0 at z_1 to ν_T at z_T,

$$\nu(z) = \nu_T \sin^2\left(\frac{\pi}{2}\frac{z - z_1}{z_T - z_1}\right), \qquad z_1 \le z \le z_T . \tag{13.2.12}$$

To investigate the properties of wave reflection from the wave-absorbing layer, we assume a wave-like solution in x direction,

$$(u', w', p', \theta') = [\hat{u},\ \hat{w},\ \hat{p},\ \hat{\theta}]\, e^{ikx} . \tag{13.2.13}$$

Substituting the above equation into (13.2.8) - (13.2.11) yields

$$\frac{\partial^2 \hat{w}}{\partial z^2} + \frac{N^2}{U^2}\hat{w} = 0 , \qquad z \le z_1, \tag{13.2.14a}$$

$$\frac{\partial^2 \hat{w}}{\partial z^2} + \frac{N^2}{U^2(1 - i\nu/kU)^2}\hat{w} = 0 , \qquad z_1 < z \le z_T . \tag{13.2.14b}$$

The general solutions of the above equations may be written

$$\hat{w} = c_1 e^{il_1 z} + c_2 e^{-il_1 z}, \qquad \text{for } z \le z_1, \tag{13.2.15a}$$

$$\hat{w} = d_1 e^{il_2 z} + d_2 e^{-il_2 z}, \qquad \text{for } z_1 \le z \le z_T , \tag{13.2.15b}$$

where

$$l_1 = \frac{N}{U} \quad \text{and} \quad l_2 = \frac{N}{U\sqrt{1 - i\nu/kU}} \tag{13.2.16}$$

are the Scorer parameters for uniform basic flow (U) in the physical and sponge layers, respectively. The four coefficients in (13.2.15) are determined by the upper boundary condition, lower boundary condition, and two interface conditions at z_1. According to

Eliassen and Palm theorem (Section 4.4), the term c_1 (c_2) represents the upward (downward) propagation of the wave energy. Thus, the ratio

$$r = \left| \frac{c_2}{c_1} \right|, \tag{13.2.17}$$

represents the reflectivity produced by the upper viscous layer. Note that r can be obtained by applying the interface conditions at $z = z_1$ and the boundary condition $\hat{w} = 0 \ at \ z = z_T$. To minimize the reflection from the upper boundary, it is suggested that the depth of the sponge layer should be greater than the hydrostatic vertical wavelength ($\lambda = 2\pi U/N$) of the mesoscale disturbance (Klemp and Lilly 1978). Figure 13.8 shows the reflectivity from the sponge layer as a function of the nondimensional inverse Reynolds number, $1/\mathrm{Re} = k\nu_T / U$, where k is the horizontal wave number. In practice, $\nu_T / \nu_1 < 6$ is a better choice to avoid the reflection due to the rapid increase of the coefficient of viscosity (Klemp and Lilly 1978). If the physical layer is assumed to be inviscid ($\nu_1 = 0$), one may choose $2 \le \left| \nu_T / kU \right| \le 5$, where k is the horizontal wave number. For example, we may choose $\nu_T = 0.002 \ s^{-1}$ for a basic flow with $U = 10 \ ms^{-1}$ over a bell-shaped mountain with $a = 20 \ km$. Figure 13.9 shows the results from a hydrostatic numerical model (panels a and b) using a sponge layer for flow over a bell-shaped mountain and compared with those calculated from Long's (1953) nonlinear theory. A vertical domain of 3.4λ is used, in which the upper half is the sponge layer, and thus, the vertically propagating hydrostatic waves are effectively absorbed by the sponge layer.

Since the addition of a sponge layer increases the computational time significantly, a direct application of the Sommerfeld (1949) radiation condition has been proposed. To elucidate the numerical radiation boundary condition, we may consider the two-dimensional, linear, hydrostatic, Boussinesq equations for a uniform basic state in the absence of Coriolis force:

$$\frac{\partial u'}{\partial t} + U\frac{\partial u'}{\partial x} + \frac{1}{\rho_o}\frac{\partial p'}{\partial x} = 0, \tag{13.2.18}$$

$$\frac{\partial p'}{\partial z} - \left(\frac{g\rho_o}{\theta_o}\right)\theta' = 0, \tag{13.2.19}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0, \tag{13.2.20}$$

$$\frac{\partial \theta'}{\partial t} + U\frac{\partial \theta'}{\partial x} + \left(\frac{N^2\theta_o}{g}\right)w' = 0. \tag{13.2.21}$$

We assume a wave-like solution,

$$(u', w', p', \theta') = (u_o, w_o, p_o, \theta_o)\ e^{i(kx+mz-\omega t)}, \tag{13.2.22}$$

and substitute it into (13.2.18) - (13.2.21) to yield the dispersion relation,

$$m^2(\omega - kU)^2 = N^2k^2. \tag{13.2.23}$$

From the above equation, the horizontal phase speed and the horizontal group velocity characterize the horizontal propagation of hydrostatic gravity waves:

$$c_{px} = \frac{\omega}{k} = U \pm \frac{N}{m}, \tag{13.2.24}$$

$$c_{gx} = \frac{\partial \omega}{\partial k} = U \pm \frac{N}{m}. \tag{13.2.25}$$

Thus, for each wavenumber pair (*k, m*), the horizontal propagation speeds of the phase lines and energy are identical. Consequently, the outward propagating wave energy can be transmitted through a lateral boundary by numerically advecting disturbances out of the boundary based on their horizontal phase speed, as discussed in Section 13.2.2. For nonhydrostatic waves, c_{px} and c_{gx} are not identical: however, they are propagating in the same direction. Thus, the radiation or open lateral boundary condition is still able to advect the energy out by a simple advection equation.

In the vertical direction, the situation is completely different. For simplicity, we may assume $U = 0$. The phase speed and group velocity can be derived,

$$c_{pz} = \frac{\omega}{m} = \pm \frac{Nk}{m^2}, \tag{13.2.26}$$

$$c_{gz} = \frac{\partial \omega}{\partial m} = \mp \frac{Nk}{m^2}. \tag{13.2.27}$$

Note that c_{pz} and c_{gz} have opposite signs, which implies that a positive c_{pz} corresponds to downward energy propagation. Thus, the advection equation is unable to advect wave energy generated within the domain out of the upper boundary, as adopted for the radiation (open) lateral boundary condition.

To identify wave modes with upward energy propagation, we consider a wave-like solution of the form

$$\phi(x, z, t) = \hat{\phi}(k, z, \omega) e^{i(kx - \omega t)}, \tag{13.2.28}$$

where ϕ may represent any dependent variables, $u', w', p',$ or θ'. Substituting the above equation into the governing equation, (13.2.18)-(13.2.21), yields

$$\frac{\partial^2 \hat{w}}{\partial z^2} + \frac{N^2}{(U-\omega/k)^2}\hat{w} = 0\,. \tag{13.2.29}$$

By assuming a positive k, the above equation has the following general solution:

$$\hat{w} = Ae^{iNz/(U-\omega/k)} + Be^{-iNz/(U-\omega/k)}\,. \tag{13.2.30}$$

A similar argument may also be made easily for a negative k. The horizontally-averaged vertical energy flux can then be obtained,

$$\overline{p'w'} = \frac{\rho_o N}{2k}\left(|A|^2 - |B|^2\right), \tag{13.2.31}$$

where terms of A and B represent the upward and downward propagation of wave energy, respectively. Thus, to avoid the wave reflection from the top boundary, we require $B = 0$. For upward propagating waves, we choose

$$\hat{w} = Ae^{iNz/(U-\omega/k)}, \tag{13.2.32}$$

which implies

$$\frac{\partial \hat{w}}{\partial z} = \frac{iN}{U-\omega/k}\hat{w}\,. \tag{13.2.33}$$

From the continuity equation and momentum equation, we have

$$\frac{\partial \hat{w}}{\partial z} = -ik\hat{u} = \frac{ik/\rho_o}{U-\omega/k}\hat{p}\,. \tag{13.2.34}$$

If both positive and negative k are taken into account, then the above two equations lead to

$$\hat{p} = \frac{\rho_o N}{|k|}\hat{w} \quad \text{at} \quad z = z_T\,. \tag{13.2.35}$$

Since the above equation has no frequency dependence, we can write the above upper radiation condition in the wave number or Fourier space,

$$\tilde{p}(z,t) = \frac{\rho_o N}{|k|}\tilde{w}(z,t), \tag{13.2.36}$$

where $\tilde{p}$ and $\tilde{w}$ are defined as

$$(p,w) = (\tilde{p},\tilde{w})e^{ikx} = (\hat{p},\hat{w})e^{i(kx-\omega t)}. \tag{13.2.37}$$

The numerical implementation of the upper radiation boundary condition to the geophysical fluid system of (13.2.18) - (13.2.21) is sketched in Fig. 13.10:

(1) Integrate $w^{\tau+1}$ upward to z_T based on the continuity equation, (13.2.20).

(2) Make the Fourier transform of $w^{\tau+1}(z_T)$ to obtain $\tilde{w}^{\tau+1}(z_T)$. A Fast Fourier Transform (FFT) numerical software may accelerate the computation.

(3) Apply (13.2.36) to obtain $\tilde{p}^{\tau+1}(z_T)$.

(4) Make the inverse Fourier transform of $\tilde{p}^{\tau+1}(z_T)$ to obtain $p^{\tau+1}(z_T)$.

(5) Integrate hydrostatic equation downward based on the upper boundary condition of $p^{\tau+1}(z_T)$ to obtain $p^{\tau+1}(z\ )$ at every height level in the domain.

Although the numerical radiation boundary condition is based more solidly on gravity wave theory, other factors, such as nonlinearity, can play roles in a more complicated fluid flow system (Klemp and Durran 1983). In addition, the flow response is sensitive to the domain height when an upper radiation boundary condition is implemented.

13.2.3 Lower boundary conditions

The lower boundary condition for an inviscid flow over a flat surface is that the flow near the surface is allowed to flow over it freely, which is often referred to as the *free-slip lower boundary condition*. Since the normal velocity is required to be 0 at a rigid

surface, the inviscid flow is always tangential to the surface. For an inviscid flow over a mountainous terrain, the free-slip lower boundary condition requires that the flow be parallel to the surface. For a two-dimensional flow, this requires

$$\frac{w}{u} = \frac{dh}{dx} \quad \text{at } z = h(x), \tag{13.2.38}$$

where $h(x)$ is the mountain profile. Thus, the *linear lower boundary condition* can be derived by making a linear approximation of the above equation,

$$w' = (U + u')\frac{dh}{dx} \approx U\frac{dh}{dx} \text{ at } z = 0. \tag{13.2.39}$$

The above condition is often adopted in mountain wave theories. In deriving the above equation, two nonlinearities have been neglected by assuming: (1) $u' << U$, and (2) the lower boundary condition is applied at $z = 0$, instead of $z = h(x)$. Similarly, the linear lower boundary condition for a three-dimensional flow can be derived,

$$w' = U\frac{\partial h}{\partial x} + V\frac{\partial h}{\partial y}, \text{ at } z = 0. \tag{13.2.40}$$

Equations (13.2.39) and (13.2.40) are only valid for linear flows over small-amplitude mountains. In mesoscale NWP models, the inviscid lower boundary condition for flow over mountains is implicitly incorporated in the terrain-following (sigma) coordinates.

With the planetary boundary layer considered, the frictional effects have to be taken into account. Therefore, the free-slip lower boundary condition is no longer valid. Instead, the *no-slip lower boundary condition* is applied,

$$u(z_o) = v(z_o) = w(z_o) = 0, \tag{13.2.41}$$

where z_o is the *roughness length*, which is defined as

$$u = (u_* / k)\ln(z / z_o). \tag{13.2.42}$$

In the above equation, k is a universal constant called the *von Karman constant*, which has a value of about 0.4 based on measurements, and u_* is the *friction velocity*, which can be obtained from the vertical momentum fluxes at surface,

$$u_*^2 = \sqrt{\left(\overline{u'w'}\right)^2 + \left(\overline{v'w'}\right)^2}\,, \tag{13.2.43}$$

where $\overline{u'w'}$ and $\overline{v'w'}$ are the turbulent momentum fluxes. Measurements indicate that the magnitude of the surface momentum flux is on the order of 0.1 m^2s^{-2}. Thus, the friction velocity is typically on the order of 0.3 ms^{-1}. In addition to the specifications of the velocities, we also have to specify pressure and potential temperature. The surface pressure can be specified based on hydrostatic balance, while the potential temperature can be prescribed as a periodic heating function,

$$\theta(z_o) = \theta_o(z_o) + \Delta\theta_{\max}\sin(2\pi t / 24h), \tag{13.2.44}$$

where t is the time in hours after sunrise, $\theta_o(z_o)$ is the potential temperature at z_o at sunrise, and $\Delta\theta_{\max}$ the maximum temperature attained during the day. To permit interactions between ground and the atmosphere, calculations of surface heat energy budget are needed.

For flow over water surfaces, the *air-sea interaction* processes need to be considered. Basically, the water affects the atmospheric circulation through sensible and latent heat fluxes directly related to its sea surface temperature (SST), including its time and space variability, while the atmosphere feeds back to the water or ocean through the wind stress to produce a deepening of the *ocean mixed layer*, induces water or ocean currents, and alters the *upwelling-downwelling* pattern.

13.3 Initial conditions and data assimilation

Mathematically, NWP can be viewed as solving an initial-boundary value problem in which the governing equations of geophysical fluid system are integrated forward in time in a finite region. Therefore, in addition to the boundary conditions as discussed in Section 13.2, we must provide suitable initial conditions for the model. For idealized numerical simulations, the initial conditions can be prescribed by known functions or values. If the Coriolis force is included in the model, then the initial basic state should be in geostrophic balance. If shear is included in a rotating atmosphere, then the initial basic state should be in thermal wind balance. Otherwise, the initial state will be adjusted to reach a new balanced state by the model, which may not be desirable. For real data mesoscale NWP, the observational data must be modified to be dynamically consistent with the governing equations of the model. The process in producing initial conditions includes the following four components: (i) *quality control*, (ii) *objective analysis*, (iii) *initialization*, and (iv) *initial guess from a short-range forecast by an NWP model*. These components have been taken to form a continuous cycle of *data assimilation*, often called *four-dimensional data assimilation* (4DDA). Many of the methods described below are still current topics of research, thus only a brief review is appropriate. Although some of the techniques might be rarely used in today's NWP models, they still provide useful information in helping understand NWP technique development.

The necessity of performing *quality control* on meteorological data was recognized long ago, which is especially important when the data are used to initialize a NWP model because the errors associated with the data may be misrepresented (*nonlinear aliasing*)

and amplified by the model. To reduce the errors in the sounding data, the following steps of quality control have been taken in NWP: (a) *plausibility check*, (b) *contradiction check*, (c) *gross check*, and (d) *buddy check* (Gandin 1988). In *plausibility check*, data values that cannot possibly occur in the real atmosphere or extremely exceed climatological mean are rejected. For example, positive temperatures in Celsius at 300 hPa are rejected. In *contradiction check*, data values of two or more parameters at the same location contradicting each other are removed. For example, the occurrence of rain in the absence of clouds is removed. In *gross check*, observations with large deviations from the first guess field forecast by an operational model are removed. In *buddy check*, observations not agreeing with neighboring observations are removed.

Observational data are often not regularly spaced, and thus are not ready for use as initial fields for a mesoscale NWP model because they do not match the model grid mesh. In some areas, such as over ocean, observational data are sparse. Therefore, in order to use the observational data as initial fields for a mesoscale NWP model, one needs to interpolate or extrapolate the data to fit into the grid mesh of the model. This procedure is called *objective analysis,* and one example is the *Barnes objective analysis* (Barnes 1964). In an objective analysis, it is desirable to do the following: (1) filter out scales of motion that cannot be resolved by the mesoscale model, (2) use a first guess field or background field provided by an earlier forecast from the same model, which will help avoid the extrapolation of observation data in data sparse areas and introduce dynamically consistency, and (3) make use of our knowledge of the probable errors associated with each observation, which can be weighted based on past records of accuracy. When the maximum information from data sources, including the

observations, climatological records, space correlation among the meteorological variables, etc., are extracted statistically, the approach is called *optimal interpolation* (e.g., Kalnay 2003). The optimal interpolation often requires knowledge of the statistical structure of the fields of the variables. The variables may be analyzed separately or simultaneously, which is referred to as *univariate analysis* or *multivariate analysis*, respectively.

The objective analysis procedure generally does not provide fields of mass and motion that are consistent with model dynamics to initiate a forecast. Thus, the use of such objectively analyzed data to initialize an NWP model may generate large, spurious inertial-gravity wave modes. Theoretically, these inertial-gravity wave modes will be dispersed, dissipated or propagate out of the domain, after sufficient time due to redistribution of mass and wind fields. The inertia gravity-wave modes or noise, as often referred to by NWP modelers, however, cannot be dissipated locally because of the relatively low resolution in NWP models. Therefore, an additional procedure, called *initialization*, is required to force the data after objective analysis to be dynamically consistent with the model dynamics, and to allow the model to integrate forward in time with a minimum of noise and the maximum accuracy of the forecasts. To improve the NWP, a number of initialization techniques have been developed in the past, such as: (a) *damping method*, (b) *static initialization*, (c) *variational method*, (d) *normal mode initialization*, and (e) *dynamic initialization*.

A simple and straightforward way to reducing the *gravity wave mode* is to dampen or filter the *inertial-gravity wave 'noise'* by adding a divergence damping term to the

horizontal momentum equation (Talagrand 1972). In this approach, the local rate of change of the divergence is diffused according to

$$\frac{\partial D}{\partial t} = -\boldsymbol{V}\cdot\nabla D - \frac{1}{\rho}\nabla^2 p + \nu\nabla^2 D + f\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) - \left(\frac{\partial u}{\partial x}\right)^2 - \left(\frac{\partial v}{\partial y}\right)^2 - 2\frac{\partial v}{\partial x}\frac{\partial u}{\partial y} - \left(\frac{\partial w}{\partial x}\frac{\partial u}{\partial z} + \frac{\partial w}{\partial y}\frac{\partial v}{\partial z}\right) + \frac{1}{\rho^2}\left(\frac{\partial p}{\partial x}\frac{\partial \rho}{\partial x} + \frac{\partial p}{\partial y}\frac{\partial \rho}{\partial y}\right) \tag{13.3.1}$$

The approach identified in (13.3.1) in initializing the data is called *damping method*. Another way to adjust the data at a single time level, usually to conform to some dynamical constraints in order to reduce or eliminate the generation of *inertial-gravity wave 'noise',* is the *static initialization*. For example, in an isobaric model, one may (i) estimate the geopotential field (ϕ) from the pressure-height data and the geostrophic wind relations; (ii) calculate the streamfunction (ψ) from analyzed ϕ fields on the isobaric surfaces; and then (iii) compute the rotational wind component from the following relationship,

$$\boldsymbol{V}_\psi = k \times \nabla\psi \,. \tag{13.3.2}$$

The above equation can be written as an elliptic function of ϕ, which may become hyperbolic in some areas. In order to ensure ellipticity in these areas so that the numerical method for the elliptic equations will apply, the geopotential fields must be altered (Haltiner and Williams 1980). In addition to this difficulty, the gravitational modes still exist even using the balance equation to determine a rotational wind for initialization. Another approach to initializing the data is to adopt the *variational method*, in which one or more of the conservation relations are applied to minimize the

variance of the difference between the observations and the objectively analyzed fields. In performing the variational method, the principles of variational calculus is applied (Sasaki 1970). For example, the difference can be minimized in a least-square sense subjected to one or more dynamical constraints, such as the balance equation, hydrostatic relation, and steady state momentum equation.

The *static initialization* described above is based on the distinction between gravity wave modes with relatively high divergence and other meteorological modes of the quasi-geostrophic type with small divergence and relatively high vorticity. However, in reality the separation is far less clear cut in some instances. Thus, it has been proposed to keep some normal modes, if they can be represented by the model grid resolution. Retaining these gravity wave modes are important since some severe weather has been found to be induced by gravity waves. Unlike applying the balance equation constraint, the *normal mode initialization* produces a divergent component as well. The normal mode initialization makes an optimal use of the observed data by adjusting both mass and motion fields while achieving dynamical consistency through appropriate constraints. In the *linear normal mode initialization* the original objectively analyzed fields are adjusted to the linearized versions of the model equations, and the undesirable gravity wave modes are removed. However, the disadvantage of this type of method is that nonlinear terms tend to regenerate the high-frequency wave modes, and also the curvature in the flow is neglected so that the fit with the original data may suffer, which is overcome by taking the *nonlinear normal mode initialization* technique. In the nonlinear normal mode initialization, the tendency of the undesirable wave modes, instead of the amplitude, is set to zero. The nonlinear normal mode initialization may also be applied to the vertical

direction, too. Figure 13.11 shows the time evolution of a height field after applying two iterations of the *implicit normal mode initialization* scheme. In comparison with the time evolution of the same height field with no initialization, the implicit nonlinear normal mode initialization appears to be able to remove high-frequency oscillations.

Since the normal mode initialization is performed separately right after the objective analysis, the initialized fields may no longer fit the observations as closely as possible. Therefore, the *dynamic initialization* was proposed (Miyakoda and Moyer 1968). The basic idea of dynamic initialization is to let the NWP model do the job by itself because any primitive equation models are supposed to inherently possess the mechanism for the geostrophic adjustment process. Indeed, the mass and velocity fields do mutually adjust to one another toward a quasi-geostrophic state when they are executed in an NWP model. In this way, observations are inserted intermittently or continuously over a period of time. In this type of initialization, the model is integrated forward and backward around the initial time, and this lets the model adjust itself before starting the forecast. During this process, it is desirable to use an integration scheme with selective damping technique. For example, the following iterative scheme consists of a forward step, then a backward step and finally an averaging:

$$
\begin{aligned}
u_*^{\tau+1} &= u^{\tau} + \Delta t(\partial u / \partial t)^{\tau}, \\
u_{**}^{\tau} &= u_*^{\tau+1} - \Delta t(\partial u_* / \partial t)^{\tau+1}, \\
\bar{u}^{\tau} &= 3u^{\tau} - 2u_{**}^{\tau}.
\end{aligned}
\tag{13.3.3}
$$

Figure 13.12 shows an example of gravity wave activity after normal mode initialization and dynamic initialization for five vertical modes of a baroclinic model, as compared to

that with no initialization. Gravity wave activity is dramatically reduced by the dynamic initialization. The disadvantages of dynamic initialization scheme are that: (i) each iteration requires the equivalent of two prognostic steps, thus taking considerable computer time, (ii) it is unable to distinguish between large-scale gravity wave modes and small-scale Rossby modes, and (iii) backward integration may not be applicable to some irreversible physical processes.

In order to incorporate the invaluable asynoptic data, such as Next Generation Weather Radar Doppler radar (NEXRAD), wind profilers, acoustic sounders, high-resolution dropsondes, satellite and aircraft, observed at nonstandard time (i.e., not at 0000 and 1200 UTC) into an NWP model system, the quality control, objective analysis, initialization, and initial guess forecast from the same model are combined into a 4DDA cycle. This 4DDA cycle can be carried out in an intermittent or continuous fashion. In the *intermittent* 4DDA, the data are assimilated intermittently at specified time intervals. The background or first guess fields forecast by the model plays a very important role, especially in data sparse regions. In data rich regions, usually the analysis is dominated by the information contained in the observations. The boundary conditions of regional NWP are provided by global model forecast. The intermittent 4DDA technique, such as three-dimensional variational data assimilation (3DVAR), is used in many global and regional operational NWP systems due to its computational efficiency. Figure 13.13 illustrates a 32-km data assimilation cycle adopted by NCEP's ETA model by using 3DVAR technique. Data types used in the 3DVAR of ETA model are rawinsonde mass and wind, pibal winds, dropwindsondes, wind profilers, surface land temperature and moisture, oceanic surface data (ships and buoys), aircraft winds, satellite cloud-drift

winds, oceanic TOVS thickness retrievals, GOES and SSM/I precipitable water retrievals, ACARS temperature data, surface winds over land, VAD winds from NEXRAD, SSM/I oceanic surface winds, and tropical cyclone bogus data (see Roger et al. 1998 for detailed information). Note that more advanced techniques have been developed, including the use of the adjoint model in intermittent data assimilation systems (e.g. Huang 1999).

The intermittent updating process is appropriate as long as most available data are taken at a fixed time period, which may vary from 3 to 12 h in practice. However, in order to take advantage of the asynoptic data, which comes in much more frequently than the synoptic data, methods of *continuous* or *dynamic 4DDA* are desired. In these methods, the observational data are essentially introduced into the assimilation system at each time step of the model integration during the assimilation time period. Examples of this type of continuous 4DDA are: (1) *nudging* or *Newtonian relaxation* (Hoke and Anthes 1976), (2) *variational assimilation* (also called *4DVAR*; Stephens 1970), and (3) *Kalman-Bucy filtering* (Kalman and Bucy 1961).

In the *nudging* or *Newtonian relaxation method*, there is pre-forecast integration period during which the model variables are driven toward the observations by adding extra forcing terms in the equations. When the actual initial time is reached, the extra terms are dropped from the model equations, and the forecast proceeds without any forcing. For example, a forcing term is added to the *x*-momentum equation,

$$\frac{\partial u}{\partial t} = -\boldsymbol{V} \cdot \nabla u + fv - \frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{u - u_{obs}}{\tau}. \qquad (13.3.4)$$

The time scale for the relaxation, τ depends on the variable and is chosen to slowly increase (decrease) prior to (after) the time of the observation to prevent any shocks to the model during the assimilation time period. Nudging has been tested for use with the new generation of observing systems, such as dropsondes, wind profilers, and surface data. Compared to the variational assimilation and Kalman-Bucy filtering techniques, to be discussed below, the nudging or Newtonian relaxation technique is less elegant mathematically but very practical. In the *four-dimensional variational data assimilation* (*4DVAR*), one tries to create the best possible fit between the model and the observational data such that the adjusted initial conditions are optimal for use in subsequent model forecasts. For example, Fig. 13.14 shows that the value produced by the first analysis is *A*. Although it fits the data well at *T*-3h, it leads to a forecast that does not match the observations well by $T = 0$ h. Note that even data collected at the same time do not necessarily agree with each other. One way to resolve the problem is to take the *adjoint method*, in which an iterative approach is used to adjust the initial analysis so that it is optimal for prediction, as one type of the 4DVAR. In other words, the adjusted analysis, A_{adj}, leads to a model trajectory (heavy curve) that produces a better 3-h forecast for T = 0 h, even though it may not be the best fit at T-3h. In the *Kalman-Bucy filtering* technique, the data sequentially adjusts the assimilated fields as the model is integrated forward in time. The Kalman-Bucy filter minimizes the analysis error variance not only at every time step, but also over the entire assimilation period in which data are provided. In addition, the filter is able to extract all useful information from the observational increment or residual at each time step, thus allowing observations to be discarded as soon as they are assimilated.

In order to involve the standard or nonstandard data to reduce or eliminate the spin-up error caused by the lack, at the initial time, of the fully developed vertical circulation required to support regions of large rainfall rates, one may adopt the *diabatic or physical normal mode initialization*, which may improve quantitative precipitation forecasts, especially early in the forecast. Two key issues can be raised here: (1) choice of technique and (2) sources of hydrologic/hydrometeor data. Diabatic heating information in nonlinear normal mode initialization can be either from the model estimates or from observed rainfall data. Various methods have been developed to infer three-dimensional fields of latent heating, moisture and divergence from two-dimensional rainfall data, such as static methods, dynamic methods, adjusting the convective parameterization scheme to match the observed rainfall, and latent heat nudging. A major problem in all of the techniques presented is the need for accurate vertical distribution of the heating and moistening rates. For example, (i) surface rain gauge data are not available on hourly basis, (ii) rawinsonde data are sparse horizontally (may be overcome by combining infrared and microwave satellite estimates), or (iii) cloud and hydrometeor data to be deduced from the network of Doppler radars are not complete (retrieval techniques are critical). To utilize the radar reflectivity measurements and satellite cloud observations, as well as surface-based cloud observations, complex cloud analysis packages have been developed that define the three dimensional cloud and hydrometeor fields as well as the associated temperature perturbations, and such analyses have been successfully applied to the initialization of storm-scale NWP models (e.g., Hu et al. 2006). With the 4DVAR method, observed measurements of rainfall rates can be incorporated into a NWP model in a more dynamically and physically consistent manner to derive improved initial

conditions (Zou and Kuo 1996). Through the model precipitation calculation and adjoint model, information from the space of the model variables, such as wind, temperature, and humidity, can be projected to that of the measured variables (i.e. rainfall rates), and back, in a consistent manner.

A numerical forecast model and a mesoscale or local analysis system can be combined as an operational atmospheric prediction system to initialize the model, such as the *Rapid Update Cycle* (*RUC*; Benjamin et al. 2004). The *Mesoscale Analysis and Prediction System* (*MAPS*) is the research counterpart to the *RUC*. The RUC has been developed to serve users needing short-range weather forecasts, including those in the US aviation community. In MAPS, a mesoscale model is employed to make a 3h data assimilation in $\sigma-\theta$ coordinates. Advances in remote sensing from earth- and space-borne systems, expanded in situ observation network, and increased low-cost computer capability allow an initialization for meso- and convective scale models. *Local Analysis and Prediction System* (*LAPS*; Shaw et al. 2004) uses data from local *mesonetwork* (*mesonet*) of surface observing systems, Doppler radars, satellites, wind and temperature profilers, as well as aircraft, which are incorporated every hour into a three-dimensional grid covering a 1000km x 1240km area. The prediction component of the LAPS is configured using a model chosen from a suite of mesoscale NWP models. Any or all of these models, usually being initialized with LAPS analyses, is run to provide 0-18 h forecasts. Another example of this type of operational atmospheric prediction system is the ARPS Data Analysis System (ADAS) and its 3DVAR system which are often used in the intermittent assimilation cycle mode, at up to 5 minute assimilation intervals (e.g.,

Xue and Martin 2006). Surface mesonet data available as frequently as every 5 minutes are also routinely assimilated into the model system.

Many of the data assimilation methods developed for larger-scale models cannot be applied to the storm-scale models. For example, storm-scale phenomena are highly ageostrophic and divergent, so that the constraints between the mass and momentum field applied at larger scale (geostrophic and thermal wind balances) cannot be applied to the storm scale. Furthermore, the mass field typically has to be inferred from the reflectivity and radial velocity measured by Doppler radars, instead of being measured directly. Thus, the retrieval techniques become very critical. The adjoint-based 4DVAR method can improve the accuracy in retrieving the thermodynamic fields, compared to more conventional method that retrieves the thermodynamic fields from the retrieved wind fields (Sun and Crook 1996). Fig. 13.15 shows an example of a vertical velocity field retrieved by an adjoint method.

Ensemble-based data assimilation, a collection of flexible state-estimation techniques that use short-term ensemble forecasts to estimate the flow-dependent background error covariance, has recently been implemented in various atmospheric models. These experimental studies demonstrated the feasibility and effectiveness of the ensemble-based techniques for different scales and flows of interest (e.g., Houtekamer et al. 2005; Zhang et al. 2006a,b) and for parameter estimation, which offers hope to the treatment of different sources of parametric model error (e.g., Anderson 2001; Aksoy et al. 2006). The best-known form of ensemble-based assimilation is the ensemble Kalman filter (EnKF) (e.g., Evensen 2003; Snyder and Zhang 2003; Hamill 2006; Xue et al. 2006). Major advantages of using the EnKF over existing data assimilation schemes include its

use of flow-dependent error covariance, its simplicity in implementation and maintenance, and its automatic generation of ensembles consistent with the analysis error covariance for the subsequent ensemble forecasts. The EnKF method is also capable of better handling nonlinear physical processes and/or nonlinear observational operators, such as that for the radar reflectivity (Tong and Xue 2005).

13.4 Nonlinear aliasing and nonlinear instability

In discussing numerical instabilities in Chapter 12, we have neglected the nonlinear effects. However, in the real atmosphere, kinetic energy generated at large scale or mesoscale tends to transfer to smaller scales. When it is transferred to the *inertial subrange*, the kinetic energy is neither produced nor dissipated, but handed down to smaller and smaller scales. The inertial subrange is an intermediate range of turbulent scales or wavelengths such that the kinetic energy is independent of original forcing of the motion and molecular dissipation. When the kinetic energy is transferred to an even smaller scale on the order of O(cm), which is called the *dissipation range*, the kinetic energy is converted into internal energy by molecular interaction.

In a numerical mesoscale model, this cascade of energy to smaller scales cannot occur because the smallest feature that can be resolved has a wavelength of $2\Delta x$. For example, let us consider

$$\phi_1 = \phi_o \cos k_1 \Delta x, \text{ and} \tag{13.4.1}$$

$$\phi_2 = \phi_o \cos k_2 \Delta x,$$

which represent 2 waves with the same amplitude ϕ_o and different wave numbers, k_1 and k_2. A *nonlinear interaction* between these two waves produces

$$\phi_1\phi_2 = (\phi_o^2/2)\left[\cos(k_1+k_2)\Delta x + \cos(k_1-k_2)\Delta x\right]. \tag{13.4.2}$$

From the above equation, one can determine that two waves with wave numbers, k_1+k_2 and k_1-k_2 have resulted from this *wave-wave interaction*. Assume k_1 and k_2 represent the following $2\Delta x$ and $4\Delta x$ waves:

$$k_1 = 2\pi/(2\Delta x), \text{ and } k_2 = 2\pi/(4\Delta x), \tag{13.4.3}$$

then we have

$$\phi_1\phi_2 = \left(\frac{\phi_o^2}{2}\right)\left\{\cos\left[2\pi\Delta x\left(\frac{3}{4\Delta x}\right)\right] + \cos\left[2\pi\Delta x\left(\frac{1}{4\Delta x}\right)\right]\right\}. \tag{13.4.4}$$

The second cosine term of the above equation is a $4\Delta x$ wave, which can be appropriately represented by the grid mesh. However, the first term is a $1.33\Delta x$ wave, which cannot be resolved by the grid mesh. The wave will be fictitiously represented by a $4\Delta x$ wave because the first integer multiple of $4\Delta x/3$ is $4\Delta x$. This phenomenon is called *nonlinear aliasing*. Figure 13.16 shows a schematic that illustrates of how a physical solution with a wavelength of $4\Delta x/3$, caused by the nonlinear interaction of $2\Delta x$ and $4\Delta x$ waves, is seen as a computational $4\Delta x$ wave in the numerical grid mesh. In the real world, we have the large-scale disturbance generated by forcing, which then cascades to a mesoscale disturbance, small-scale disturbance, and then dissipates at an even smaller scale. However, it does not seem to happen in the same way in the numerical model, in which waves with wavelength shorter than $2\Delta x$ will be represented as larger scale waves. Therefore, even if a numerical method is linearly stable, the results can degrade into

computational noise. The erroneous accumulation of energy can cause the model dependent variables to increase in magnitude abruptly without bound, which is called *nonlinear instability*.

Two methods can be applied to avoid the nonlinear instability: (1) proper parameterization of the *subgrid-scale correction* terms, such as $\overline{u'w'}$, $\overline{v'w'}$, and $\overline{\theta'w'}$, so that energy is extracted from the averaged equations, or (2) the use of a spatial *numerical smoother* or *filter* to remove the shorter waves, which leaves the longer waves relatively unaffected. The first approach is better than the second one because it is based on physical principle. However, it requires a good knowledge about the subgrid-scale correlation terms. The second approach can be accomplished in a relatively easier manner (e.g, Shapiro 1975). To understand numerical smoothing, we may consider a simple one-dimensional, *three-point operator*,

$$\bar{\phi}_j = (1-s)\phi_j + (s/2)(\phi_{j-1} + \phi_{j+1}), \tag{13.4.5}$$

where $x = j\Delta x$ and s is a constant that can be negative. If this operator is applied to the harmonic form of a wave

$$\phi = Ae^{ikx}, \tag{13.4.6}$$

where $k = 2\pi/L$ is the wave number, and A is a constant that may be a complex, then the result may be written as

$$\bar{\phi} = R\phi, \quad R = 1 - s(1-\cos k\Delta x) = 1 - 2s\sin^2(\pi\Delta x/L). \tag{13.4.7}$$

In the above equation, R is referred to as the *response function*. If $R \geq 0$, then the wave number and phase are not affected but only the wave amplitude. If $|R| > 1$, then the wave is amplified by the operator. On the other hand, if $|R| < 1$, then the wave is damped by

the operator. If $R<0$, then the phase of the wave is shifted by 180°, which is undesirable. With $s=1/2$, we obtain the *second-order smoother*:

$$\overline{\phi}_j = \frac{\phi_{j-1} + 2\phi_j + \phi_{j+1}}{4}, \tag{13.4.8}$$

and

$$R(1/2) = 1 - \frac{1-\cos k\Delta x}{2} = \cos^2(\pi \Delta x / L). \tag{13.4.9}$$

From above, if $L = 2\Delta x$, then $R = 0$. Hence, for a $2\Delta x$ wave, the smoother will eliminate it immediately.

Since a *three-point smoother*, (13.4.8), damps the shorter waves too strongly, it is less desirable. A *five-point smoother* can be obtained by applying 2 successive three-point smoother with $s = 1/2$ and $-1/2$:

$$\begin{aligned} \overline{\overline{\phi}}_j &= \frac{1}{16}\left[10\phi_j + 4\left(\phi_{j-1} + \phi_{j+1}\right) - \left(\phi_{j-2} + \phi_{j+2}\right)\right] \\ &= \phi_j - \frac{1}{16}\left[6\phi_j - 4\left(\phi_{j-1} + \phi_{j+1}\right) + \left(\phi_{j-2} + \phi_{j+2}\right)\right] \end{aligned}. \tag{13.4.10}$$

The above smoother will also remove the $2\Delta x$ wave immediately but will preserve more of the longer waves. In fact, the above *five-point smoother* is analogous to the finite difference form of the *fourth-order diffusion equation*,

$$\frac{\partial \phi}{\partial t} + c\frac{\partial^4 \phi}{\partial x^4} = 0, \tag{13.4.11}$$

which has a finite difference form,

$$\phi_j^{\tau+1} = \phi_j^{\tau} - \gamma_1[6\phi_j^{\tau} - 4(\phi_{j-1}^{\tau} + \phi_{j+1}^{\tau}) + (\phi_{j-2}^{\tau} + \phi_{j+2}^{\tau})], \tag{13.4.12}$$

where $\gamma_1 = c\Delta t / \Delta x^4$. If we choose $\gamma_1 = 1/16$, then the above equation is analogous to (13.4.10). Thus, in applying the five-point smoother, it has a similar effect as the fourth-

order diffusion, which is why *numerical smoothing* has also been referred to as *numerical diffusion*. In order to retain the amplitude of longer waves, the coefficient 1/16 in (13.4.10) or γ_1 is often reduced. Testing is needed to find out the most appropriate coefficient of the numerical smoothing or diffusion. In practice, smoothing is not applied to the boundary points. For the grid points adjacent to the boundaries, we may need to apply the three-point smoother or second-order diffusion

$$\overline{\phi_j} = \phi_j - \gamma_2[2\phi_j - (\phi_{j-1} + \phi_{j+1})]. \tag{13.4.13}$$

In order to make (13.4.13) consistent with (13.4.12), we require

$$\gamma_2 = 4\gamma_1. \tag{13.4.14}$$

Notice that the leapfrog scheme also produces a computational mode with 2Δt wave. To suppress this, we may apply the time smoother (Asselin, 1972)

$$\phi^{\tau+1} = \overline{\phi}^{\tau-1} + 2\Delta t(\partial\phi/\partial t)^{\tau}, \tag{13.4.15}$$

where

$$\overline{\phi}^{\tau-1} = \phi^{\tau-1} + \gamma(\phi^{\tau} - 2\phi^{\tau-1} + \overline{\phi}^{\tau-2}). \tag{13.4.16}$$

Based on numerical testing, a choice of $\gamma < 0.25$ has been recommended.

In general, high-order filters or smoothers are more *scale selective,* i.e., they filter shorter waves at or close to two-grid interval wavelength most effectively while leaving longer waves less affected. High-order filters can have the effect of generating undesirable overshooting and undershooting in the filtered field, however. The erroneous effects in the filtered field can be avoided by a simple treatment that ensures that the diffusion flux is always down-gradient (Xue 2000).

13.5 Modeling a stratified fluid system

To elucidate how to model a stratified fluid flow system, we consider the nonlinear, hydrostatic, incompressible fluid system similar to that governed by (2.2.14) - (2.2.18) with $V = 0$,

$$\frac{\partial u'}{\partial t} + (U + u')\frac{\partial u'}{\partial x} + v'\frac{\partial u'}{\partial y} + w'\left(U_z + \frac{\partial u'}{\partial z}\right) - fv' + \frac{1}{\bar{\rho}}\frac{\partial p'}{\partial x} = \nu\nabla^2 u', \tag{13.5.1}$$

$$\frac{\partial v'}{\partial t} + (U + u')\frac{\partial v'}{\partial x} + v'\frac{\partial v'}{\partial y} + w'\frac{\partial v'}{\partial z} + fu' + \frac{1}{\bar{\rho}}\frac{\partial p'}{\partial y} = \nu\nabla^2 v', \tag{13.5.2}$$

$$\frac{1}{\bar{\rho}}\frac{\partial p'}{\partial z} = g\frac{\theta'}{\bar{\theta}}, \tag{13.5.3}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} = 0, \tag{13.5.4}$$

$$\frac{\partial \theta'}{\partial t} + (U + u')\frac{\partial \theta'}{\partial x} + v'\frac{\partial \theta'}{\partial y} + w'\left(\frac{N^2\bar{\theta}}{g} + \frac{\partial \theta'}{\partial z}\right) = \frac{\theta_o}{c_p T_o}q' + \kappa\nabla^2\theta', \tag{13.5.5}$$

where q' is the diabatic heating rate in J kg^{-1}s^{-1}, which may represent surface sensible and/or elevated heating, ν the *eddy viscosity*, κ *eddy thermal diffusivity*, and the Brunt-Vaisala frequency N is defined as $N^2 = (g/\bar{\theta})(d\bar{\theta}/dz)$. Other symbols are defined in Chapter 2. The basic state is assumed to be in geostrophic and hydrostatic balances. The basic and perturbation quantities have been separated in the above system, which allows one to examine the nonlinear effects by comparing with the corresponding linear simulation.

To elucidate how to model the nonlinear system of (13.5.1)-(13.5.5), we consider relatively simple and straightforward schemes, i.e. the leapfrog in time and second-order centered in space to the prognostic equations. The variables $u^{\tau+1}$, $v^{\tau+1}$ and $\theta^{\tau+1}$ at time

step $\tau+1$ are obtained from other variables at time steps τ and $\tau-1$, based on (13.5.1), (13.5.2) and (13.5.5), respectively. The vertical velocity $w^{\tau+1}$ can then be obtained by integrating the continuity equation, (13.5.4), upward. The upper boundary condition can be approximated by a sponge layer with ν increasing from the top of the physical domain to the top of the model domain, or following the flow chart of Fig. 13.10 to apply the upper radiation boundary condition numerically. Following Fig. 13.10, we can obtain $p^{\tau+1}$ by integrating the hydrostatic equation (13.5.3) downward. In this approach, all variables at time step $\tau+1$ are then obtained numerically.

In some mesoscale models, the hydrostatic assumption is relaxed to more properly simulating *deep convection* and effects of steep topography. A set of fully compressible fluid system can be written as,

$$\frac{Du}{Dt} = fv - \frac{1}{\rho}\frac{\partial p}{\partial x} + \nu\nabla^2 u, \tag{13.5.6}$$

$$\frac{Dv}{Dt} = -fu - \frac{1}{\rho}\frac{\partial p}{\partial y} + \nu\nabla^2 v, \tag{13.5.7}$$

$$\frac{Dw}{Dt} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g + \nu\nabla^2 w, \tag{13.5.8}$$

$$\frac{D\theta_v}{Dt} = S_\theta + \kappa\nabla^2\theta \tag{13.5.9}$$

$$\frac{D\phi}{Dt} = S_\phi + \kappa\nabla^2\phi, \quad \phi = q_v, q_c,\ q_i,\ q_r,\ q_s,\ q_g, \tag{13.5.10}$$

$$\frac{D\rho}{Dt} + \rho\nabla\cdot\boldsymbol{V} = 0, \tag{13.5.11}$$

$$p = \rho RT \tag{13.5.12}$$

$$\theta_v = T_v \left(\frac{p_o}{p} \right)^{R_d / c_p}, \tag{13.5.13}$$

$$T_v = T(1 + 0.61 q_v), \tag{13.5.14}$$

where T_v is called *virtual temperature:* θ_v is the *virtual potential temperature:* p_o is the basic state pressure at the ground, usually taken as 1000 hPa: S_θ is any source or sink of θ_v, such as surface long-wave radiation and elevated latent heating: S_ϕ is any source or sink of the hydrometeor ϕ, such as mixing ratios of *water vapor* (q_v), *cloud water* (q_c), *rain* (q_r), *cloud ice* (q_i), *snow* (q_s), and *graupel/hail* (q_g). The virtual (potential) temperature is the (potential) temperature that a dry air parcel would have if its pressure and density were equal to those of a given sample of moist air. The virtual temperature is a fictitious temperature of a moist air parcel that satisfies the equation of state for dry air. More realistic and sophisticated parameterizations of planetary boundary layer processes can be adopted.

The above equations may also be represented in terms of the *Exner funtion*, to be defined below. Including the moisture, the equation of state can be written in the form,

$$p = \rho R_d T_v . \tag{13.5.15}$$

In order to avoid an explicit treatment of the density, an *Exner function* has been adopted in some mesoscale and cloud-scale models,

$$\pi = c_p \left(\frac{p}{p_o} \right)^{R_d / c_p} = \frac{c_p T_v}{\theta_v} . \tag{13.5.16}$$

The pressure gradient force terms can then be approximately represented by

$$\frac{1}{\rho}\frac{\partial p}{\partial x_i} = \bar{\theta}_v \frac{\partial \pi}{\partial x_i}, \quad i = 1,\ 2,\ 3, \tag{13.5.17}$$

where $\bar{\theta}_v$ is the initial undisturbed state θ_v and a function of z only, defined as $\theta_v = \bar{\theta}_v + \theta_v'$. The Exner function can be partitioned into $\pi = \bar{\pi} + \pi'$, where $\bar{\pi}$ is the initial basic state, and π' is the perturbation from the initial state $\bar{\pi}$. The initial basic state is assumed to be in geostrophic balance in horizontal and hydrostatic balance in vertical,

$$\frac{\partial \bar{\pi}}{\partial x} = fv_g; \quad \frac{\partial \bar{\pi}}{\partial y} = -fu_g, \tag{13.5.18}$$

$$\frac{\partial \bar{\pi}}{\partial z} = -\frac{g}{\bar{\theta}_v}. \tag{13.5.19}$$

The advantages of using π, instead of p are that: (a) ρ is not treated explicitly in the governing equations, (b) π does not present in the buoyancy term even if the vertical scale of the motion L_z is equivalent to the scale height H; (c) there is no need to compute the density perturbation; and (d) less truncation error is introduced since $\partial\bar{\pi}/\partial z$ is much less than $\partial\bar{p}/\partial z$ (Pielke 2002). One disadvantage of using the Exner function is that the anelastic continuity equation for π is much more complicated and need a Poisson equation solver, which is very tedious in the terrain-following coordinates (Huang 2000).

Using the Exner function and approximating θ_v by $\bar{\theta}_v$ in pressure gradient forces, the momentum equations become

$$\frac{Du}{Dt} = f(v - v_g) - \bar{\theta}_v \frac{\partial \pi'}{\partial x} + \nu\nabla^2 u, \tag{13.5.20}$$

$$\frac{Dv}{Dt} = -f(u-u_g) - \bar{\theta}_v \frac{\partial \pi'}{\partial y} + \nu \nabla^2 v, \tag{13.5.21}$$

$$\frac{Dw}{Dt} = -\bar{\theta}_v \frac{\partial \pi'}{\partial z} + g\left[\frac{\theta_v'}{\bar{\theta}_v} - q_H\right] + \nu \nabla^2 w, \quad q_H = q_c + q_i + q_r + q_s + q_g \tag{13.5.22}$$

where $D/Dt \equiv \partial/\partial t + u\partial/\partial x + v\partial/\partial y + w\partial/\partial z$. The pressure equation can be written as

$$\begin{aligned}\frac{D\pi'}{Dt} = &-\frac{1}{\bar{\theta}_v}(\boldsymbol{V}\cdot\nabla\bar{\pi}) - \left(\frac{R_d\bar{\pi}}{c_v\bar{\rho}\bar{\theta}_v^2}\right)\nabla\cdot\left(\bar{\rho}\bar{\theta}_v\boldsymbol{V}\right) - \left(\frac{R_d\pi'}{c_v}\right)\nabla\cdot\boldsymbol{V} \\ &+ \frac{R_d(\bar{\pi}+\pi')}{c_v\bar{\theta}_v}\dot{\theta} + \kappa\nabla^2\pi',\end{aligned} \tag{13.5.23}$$

where $\dot{\theta} \equiv D\theta/Dt$ is the diabatic heating rate (J s^{-1}). The thermodynamic equation and the equations governing hydrometeors can be written as

$$\frac{D\theta}{Dt} = S_\theta + \kappa\nabla^2\theta, \tag{13.5.24}$$

$$\frac{D\phi}{Dt} = S_\phi + \kappa\nabla^2\phi, \quad \phi = q_v,\ q_c,\ q_i,\ q_r,\ q_s, \text{ and } q_g. \tag{13.5.25}$$

The advantage of using the fully-compressible fluid system of (13.5.20) - (13.5.25) is that every equation can be integrated alone, numerically, to obtain its own value at the next time step without having to couple with other equations, such as the hydrostatic, incompressible fluid system of (13.5.1) - (13.5.5). However, this set of equations contains sound waves, which propagate at much higher speeds than the gravity waves and require a very small time step to insure the numerical stability. In practice, it is almost impossible to adopt such a small time step, even at the research mode of numerical simulations. In order to improve numerical efficiency of the above compressible atmospheric system, a *time-splitting scheme* was proposed (Klemp and Wilhelmson 1978). In the time-splitting scheme, equations with no sound wave modes,

i.e. (13.5.24) and (13.5.25), embedded are marched with a large time step, Δt, while equations with sound waves embedded, i.e. (13.5.20) - (13.5.23), are integrated with a small time step Δt from time $t-\Delta t$ to $t+\Delta t$.

Scale analysis indicates that the only important term in (13.5.23) for representing convection is the second term on the right hand side, which leads to the anelastic continuity equation (Ogura and Philips 1962)

$$\nabla \cdot (\bar{\rho}\boldsymbol{V}) = 0. \tag{13.5.26}$$

Taking the time derivative of the above equation and using the momentum equations yield an elliptic equation

$$\nabla^2 \bar{\theta}_v \pi' + \nabla \bar{\theta}_v \cdot \nabla \pi' + \frac{\partial}{\partial z}(\ln \bar{\rho}) \bar{\theta}_v \frac{\partial \pi'}{\partial z} = \text{source terms}, \tag{13.5.27}$$

where the source terms include acceleration terms. One of the disadvantages in adopting the anelastic approximation is that (13.5.27) becomes very complicated and computationally expensive when it is transformed into terrain-following coordinates.

13.6 Predictability and ensemble forecasting

One major challenge in NWP is whether the weather phenomena of concern are predictable or not. If they are intrinsically unpredictable, then the improvements in more accurate initial and boundary conditions, numerical methods, and subgrid-scale parameterizations of a NWP model will be useless. If they are predictable, it is important to know how long we can make numerical prediction with a "perfect" NWP model, if there were one. As discussed briefly in Chapter 1, in reality, most of the weather systems more or less have limited predictability. Thus, it leaves us some room to make

improvements in the accuracy of NWP models, which is still a topic of current research, especially for mesoscale NWP. Thus we will only make a very brief summary of the *predictability* problem and the *ensemble forecasting*.

In the early 1950s, some meteorologists started to apply statistical methods to weather prediction to cope with the uncertainties encountered in forecasting (Gleeson 1961). The weather forecasting problem has been viewed as evolving probabilities. Even with improved model techniques, the weather prediction has its own inherent limitations due to the inevitable model deficiencies and errors in the initial conditions, or the predictability problem. The atmosphere, like any other dynamical systems with instabilities, has an inherit time limit of predictability (Lorenz 1963). Based on the Saltzman's (1962) convective equations, Lorenz found that two complete different solutions were predicted by the same model with slightly different initial conditions. He later found that errors of different spatial scales grow at different rates (Lorenz 1969). On average, the fastest error growth occurs at smallest scales, which would have significant impacts on the mesoscale NWP.

Predictability can have two fundamentally different meanings. Intrinsic predictability can be defined as "the extent to which prediction is possible if an optimum procedure is used" in the presence of infinitesimal initial errors (Lorenz 1969). Practical predictability, on the other hand, can be specified as the ability to predict based on the procedures that are currently available. Practical predictability is limited by realistic uncertainties in both the initial states and the forecast models, which in general are not infinitesimally small (Lorenz 1996). Compared to typical synoptic-scale flows, recent studies showed that predictability of mesoscale weather systems, particularly the associated quantitative

precipitation, which are of most concern to the public, can be very limited (e.g., Ehrendorfer 1997). The predictability of quantitative precipitation is different from earlier results which indicated that the mesoscale enjoyed enhanced predictability. Through high-resolution mesoscale simulations, recent studies demonstrated that mesoscale predictability depends strongly on the background flow regime and dynamics. The simulations also show that moist convection is a primary mechanism for forecast-error growth at sufficiently small scales, and that convective-scale errors contaminate the mesoscale within lead times of interest to NWP, thus effectively limiting the predictability of the mesoscale (e.g., Zhang et al. 2006). Understanding of the limit of mesoscale predictability and the associated error growth dynamics is essential for setting up expectations and priorities for advancing deterministic mesoscale forecasting, and for providing guidance on the design, implementation and application of short-range ensemble prediction systems (e.g., Tracton and Kalnay 1993; Stensrud et al. 1999).

Over the past decade, *ensemble forecasting* has emerged as a powerful tool for operational numerical weather prediction (Molteni et al. 1996). Ensemble forecast is a collection of different forecasts all valid at the same forecast time. Ensemble forecasts start from different initial conditions, boundary conditions, parameter settings, or entirely independent NWP models. The various forecasts all represent possibilities given the uncertainties associated with forecasting. From these possibilities, one can estimate probabilities of various events as well as an averaged or consensus forecast. An ensemble of forecasts can be used to: (a) composite into a single forecast by means of a weighted average, (b) provides an excellent probabilistic alternative to explore the dynamics, predictability and background error covariance for mesoscale weather systems,

(c) estimate the reliability of the composite forecast, (d) suggest where additional special observations might be targeted to improve forecast accuracy (Kalnay 2003), and (e) determine the sensitivity of forecasts to the model input parameters, including the initial and boundary conditions (Martin and Xue 2006).

In order to take advantage of forecasts from different models and analyses, the multi-analysis-multimodel forecasts along with the aforementioned statistics are used to provide the superensemble forecasts. The superensemble has a higher accuracy compared to that of the ensemble mean because the superensemble is selective in assigning weights and the past history of performance of models from the past statistics. A real-time multianalysis-multimodel superensemble forecasts can make a significant improvement in precipitation forecasts (Krishnamurti et al. 2001). As indicated in Fig. 13.17, during the training period, the observed fields provide statistics that are then passed on to the area on the right, where $t > 0$. Compared to medium-range ensemble forecasting with global models, limited-area short-range ensemble prediction is less widely used, which is at least partially due to our limited knowledge of mesoscale predictability (Stensrud et al. 1999). Still, limited studies on even high-resolution storm-scale ensemble predictions have been undertaken (Kong et al. 2006) and such research is expected to expand significantly in the future.

References

Aksoy A., F. Zhang, J. W. Nielsen-Gammon, 2006: Ensemble-based simultaneous state and parameter estimation with MM5. *Geophys. Res. Lett.*, **33**, L12801, doi:10.1029/2006GL026186, 2006.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884-2903.

Anthes, R. A., and T. T. Warner, 1978: Development of hydrostatic models suitable for air pollution and other mesometeorological studies. *Mon. Wea. Rev*., **106**, 1045-1078.

Arakawa, A., and V. R. Lamb, 1977: Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Compu. Phys*., 17, Academic Press, 174-265, 337pp.

Asselin, R. A., 1972: Frequency filter for time integration. *Mon. Wea. Rev.*, **100**, 487-490.

Bacon, D. P., N. N. Ahmad, Z. Boybeyi, T. J. Dunn, M. S. Hall, P. C. S. Lee, R. A. Sarma, M. D. Turner, K. T. Waight III, S. H. Young, and J. W. Zack, 2000: A dynamically adapting weather and dispersion model: The operational multiscale environment model with grid adaptivity (OMEGA). *Mon. Wea. Rev*., **128**, 2044-2076.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor*., **3**, 396-409.

Benjamin, S. G., D. Dévényi, S. S. Weygandt, K. J. Brundage, J. M. Brown, G. A. Grell, D. Kim, B. E. Schwartz, T. G. Smirnova, T. L. Smith and G. S. Manikin, 2004: An Hourly Assimilation–Forecast Cycle: The RUC. *Mon. Wea. Rev*., **132**, 495–518.

Bougeault, P., 1983: A non-reflective upper boundary condition for limited-height hydrostatic models. *Mon. Wea. Rev*., **111**, 420-429.

Davies, H., 1983: Limitations of some common lateral boundary conditions used in regional NWP models. *Mon. Wea. Rev*., **111**, 1002-1012.

Dietachmayer, G. S., and K. K. Droegemeier, 1992: Application of continuous dynamic grid adaptation techniques to meteorological modeling. Part I: Basic formulation and accuracy. *Mon. Wea. Rev*., **120**, 1675-1706.

Doswell, C. A. III, 1984: A kinematic analysis of frontogenesis associated with a nondivergent vortex. *J. Atmos. Sci*., **41**, 1242-1248.

Durran, D. R., 1998: *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics.* Springer-Verlag Inc., 465pp.

Ehrendorfer, M., 1997: Predicting the uncertainty of numerical weather forecast: A review. *Meteor. Z.*, **6**, 147-183.

Evensen, G., 2003: The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, **53**, 343-367.

Gandin, L. S., 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev*., **116**, 1137-1156.

Gleeson, T. A., 1961: A statistical theory of meteorological measurements and predictions. *J. Meteor*., **18**, 192-198.

Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology*. 2nd Ed., John Wiley & Sons, Inc., 477pp.

Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate* (Ch. 6), Cambridge Press, 124-156.

Hoke, J. E., and R. A. Anthes, 1976: The initialization of numerical models by a dynamical initialization technique. *Mon. Wea. Rev*., **104**, 1551-1556.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.,* **133**, 604-620.

Hu, M., M. Xue, J. Gao, and K. Brewster, 2006: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part II: Impact of radial velocity analysis via 3DVAR. *Mon. Wea. Rev.,* **134,** 699-721.

Huang, C.-Y., 2000: A forward-in-time anelastic nonhydrostatic model in a terrain-following coordinate. *Mon. Wea. Rev*., **128**, 2108-2134.

Huang, X.-Y. 1999: A generalization of using an adjoint model in intermittent data assimilation systems. *Mon. Wea. Rev*., **127**, 766-787.

Kalman, R. E., and R. S. Bucy, 1961: New results in linear filtering and prediction theory. Trans. ASME. *J. Basic Eng*., **83D**, 95-108.

Kalnay. E., 2003: *Atmospheric Modeling, Data Assimilation and Prediction*. Cambridge University Press, 341pp.

Klemp, J. B., and D. R. Durran, 1983: An upper boundary condition permitting internal gravity wave radiation in numerical mesoscale models. *Mon. Wea. Rev*., **111**, 430-444.

Klemp, J. B., and D. K. Lilly, 1978: Numerical simulation of hydrostatic mountain waves. *J. Atmos. Sci.,* **35**, 78-107.

Klemp, J. B., and R. B. Wilhelmson, 1978: The simulation of three-dimensional convective storm dynamics. J. Atmos. Sci., 35, 1070-1096.

Kong, F., K. K. Droegemeier, and N. L. Hickmon, 2006: Multi-resolution ensemble forecasts of an observed tornadic thunderstorm system, Part I: Comparison of coarse and fine-grid experiments. *Mon. Wea. Rev.*, **134,** 807-833.

Krishnamurti, T. N., S. Surendran, D. W. Shin, R. J. Correa-Torres, T. S. V. V. Kumar, E. Williford, C. Kummerow, R. F. Adler, J. Simpson, R. Kakar, W. S. Olson, and F. J. Turk, 2001: Real-time multianalysis-multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev*., **129**, 2861-2883.

Lin, Y.-L., and T.-A. Wang, 1996: Flow regimes and transient dynamics of two-dimensional stratified flow over an isolated mountain ridge. *J. Atmos. Sci*., **53**, 139-158.

Long, R. R., 1953: Some aspects of stratified fluids. I. A theoretical investigation. *Tellus*, **5**, 42-58.

Lorenz, E. N., 1963: Deterministic non-periodic flow. *J. Atmos. Sci*., **20**, 130-141.

Lorenz, E. N. 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289-307.

Lorenz, E. N., 1996: Predictability - A problem partly solved. *Seminar on Predictability, 1995 ECMWF Seminar Proceedings*, ECMWF, Reading, UK. Volume **I**, 1-19.

Mahrer, Y., and R. A. Pielke, 1978: A test of an upstream spline interpolation technique for the advecgive terms in a numerical model. *Mon. Wea. Rev*., **106**, 818-830.

Martin, W. J. and M. Xue, 2006: Initial condition sensitivity analysis of a mesoscale forecast using very-large ensembles. *Mon. Wea. Rev.*, **134,** 192–207.

Mesinger, F., Z. I. Janjić, S. Nickovic, D. Gavrilov, and D. G. Deaven, 1988: The step-mountain coordinate: model description and performance for cases of Alpine lee

cyclogenesis and for a case of an Appalachian redevelopment. *Mon. Wea. Rev*., **116**, 1493-1518.

Miyakoda, K., and R. Moyer, 1968: A method of initialization for dynamical weather forecasting. *Tellus*, **20**, 115-128.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system. Methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122,** 73–120.

Ogura, Y., and N. A. Philips, 1962: Scale analysis of deep and shallow convection in the atmosphere. *J. Atmos. Sci.,* **19**, 173–179

Orlanski, I., 1976: A simple boundary condition for unbounded hyperbolic flows. *J. Compu. Phys*., **21**, 251-269.

Oliger, J., and A. Sundström, 1978: Theoretical and practical aspects of some initial boundary value problems in fluid dynamics. *Int'l J. Numer. Methods Fluids*, **21**, 183-204.

Perkey, D. J., and C. W. Kreitzberg, 1976: A time-dependent lateral boundary scheme for limite-area primitive equation models. *Mon. Wea. Rev*., **104**, 744-755.

Pielke, R. A., 2002: *Mesoscale Meteorological Modeling.* 2nd Ed., Academic Press, Inc., 676pp.

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

Rogers, E., M. Baldwin, T. Black, K. Brill, F. Chen, C. DiMego, J. Gerrity, G. Manikin, F. Mesinger, K. Mitchell, D. Parrish, and Q. Zhao, 1998: Changes to the NCEP

operational “early” Eta analysis/forecast system. NWS Tech. Proc. Bull., Ser. No. 447, NWS, NOAA, 10pp. (http://www.nws.noaa.gov/om/tpb/447.htm)

Saltzman, B., 1962: Finite amplitude free convection as an initial value problem - I, *J. Atmos. Sci.*, **19**, 329-341.

Sasaki, Y., 1970: Some basic formulations in numerical variational analysis. *Mon. Wea. Rev*., **98**, 875-883.

Schoenstadt, A. L., 1978: A transfer function analysis of numerical schemes used to simulate geostrophic adjustment. NPS Rept. NPS-53-79-001, 44pp.

Shapiro, R., 1975: Linear filtering. *Math. Compu*., **29**, 1094-1097.

Shaw, B. L., S. Albers, D. Birkenheuer, J. Brown, J. McGinley, P. Schultz, J. Smart, and E. Szoke, 2004: Application of the Local Analysis and Prediction System (LAPS) diabatic initialization of mesoscale numerical weather prediction models for the IHOP-2002 field experiment. Preprint, 20th Conf. on Wea. Anal. and Forec. And 16th Conf. on Numer. Wea. Predic., Amer. Meteor. Soc., Seattle, WA.

Skamarock, W. C., 1989: Truncation error estimates for refinement criteria in nested adaptive models. *Mon. Wea. Rev*., **117**, 872-886.Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang and J. G. Powers, 2005: A Description of the Advanced Research WRF Version 2. NCAR Technical Note, NCAR/TN–468+STR. [Available at http://wrf-model.org/wrfadmin/docs/arw_v2.pdf]

Smagorinski, J., J. L. Holloway, Jr., and G. D. Hembree, 1967: Prediction experiments with a general circulation model. Proceed.. Int’l. Sympo. Dynamics Large Scale Atmospheric Processes, Nauka, Moscow, U.S.S.R., 70-134.

Snyder, C., and F. Zhang, 2003: Tests of an ensemble Kalman filter for convective-scale data assimilation. *Mon. Wea. Rev.*, **131**, 1663-1677.

Sugi, M., 1986: Dynamic normal mode initialization. *J. Meteor. Soc. Japan*, **64**, 623-632.

Sommerfeld, A., 1949: Partial Differential Equations in Physics. Academic Press, 335pp.

Stensrud, D. J., H. E. Brooks, J. Du, M. S. Tracton, E. Rogers, 1999: Using ensembles for short-range forecasting. *Mon. Wea. Rev*., **127**, 433–446.

Stephens, J., 1970: Variational initialization of the balance equation. *J. Appl. Meteor*., **9**, 732-739.

Sun, J., and N. A. Crook, 1996: Comparison of thermodynamic retrieval by the adjoint method with the traditional retrieval method. *Mon. Wea. Rev*., **124**, 308-324.

Talagrand, O., 1972: On the damping of high-frequency motions in four-dimensional assimilation of meteorological data. *J. Atmos. Sci*., **29**, 1571-1574.

Temperton, C., 1988: Implicit normal mode initialization. *Mon. Wea. Rev*., **116**, 1013-1031.

Tong, M. and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS Experiments. *Mon. Wea. Rev.*, **133,** 1789-1807.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Weather and Forecasting,* 8, 379–398.

Xue, M., 2000: High-order monotonic numerical diffusion and smoothing, *Mon. Wea. Rev*. **128**, 2853-2864.

Xue, M. and W. J. Martin, 2006: A high-resolution modeling study of the 24 May 2002 case during IHOP. Part I: Numerical simulation and general evolution of the dryline and convection. *Mon. Wea. Rev.*, **134,** 149–171.

Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. J. Atmos. Ocean Tech., 23, 46–66.

Zhang, F., Z. Meng, and A. Aksoy, 2006a: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation, Part I: Perfect model experiments. *Mon. Wea. Rev.*, **134**, 722-736.

Zhang, F., A. Odins, and J. W. Nielsen-Gammon, 2006b: Mesoscale predictability of an extreme warm-season rainfall event. *Weather and Forecasting*, **21**, 149-166.

Zou, X. and Y.-H. Kuo, 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev*., **124**, 859-2882.

Problems

13.1 From (13.1.5) and (13.1.6) derive the computational dispersion relationship of the two-dimensional shallow water system with $U = 0$, (13.1.7).

13.2 Derive the computational dispersion relationship corresponding to (13.1.7) but for an unstaggered grid mesh.

13.3 Substituting $u(x,t) = \hat{u}(k,\omega)\exp[i(kx - \omega t)]$ into a particular finite difference form of the advection equation gives

$e^{-2i\omega\Delta t} + i\beta e^{-i\omega\Delta t} - 1 = 0$.

Find the stability criterion for this scheme if $0 \le \beta \le 2$.

13.4 Derive (13.1.13) from (13.1.12)

13.5 Applying the forward-in-time and backward-in-space scheme to the advection equation, (12.3.1), at the right boundary to show that the zero gradient boundary condition assumes the waves propagate out of the right boundary at a speed of $\Delta x / \Delta t$.

13.6 Derive a numerical radiation boundary condition at the inflow boundary, (13.2.7), and $U - c_i^*$.

13.7 Derive (13.2.23) from (13.2.18) - (13.2.21) by assuming a wave-like solution.

13.8 Derive the complete equation of (13.3.1).

13.9 Prove the response function (R) in (13.4.7) by applying the three-point smoother (13.4.5) to a sinusoidal wave (13.4.6).

13.10 Derive (13.4.10).

13.11 Prove $\gamma_2 = 4\gamma_1$ by considering a $2\Delta x$ wave, $\phi_i = (-1)^{i+1}$.

13.12 Derive the pressure gradient forces, (13.5.17), by using the Exner function.

Modeling Project

13.A1 Set smoothing coefficient as 0 and different values in the advection model. Discuss the results.

13.A2 Replace the 5-point smoother of the shallow water model by a three-point smoother in the shallow water model and discuss the smoothing effects.

13.A3 Modify your 2D Tank Model to simulate a two-layer fluid system with H = 1 km and $\Delta\rho/\rho_o$ = 0.1. Rerun the model for a flow over a bell-shaped mountain with U = 20, 30, and 40 ms^{-1}. Describe and explain the results.

13.A4 (i) Extend the 2D Tank Model to a 3D Tank Model.

(ii) Now run this 3D Tank Model with $U = 0$ m s^{-1} and $h_m = 0$ m by giving a three-dimensional initial field in h'.

(iii) Simulate a basic flow with U=20 ms^{-1} over a 3D bell-shaped mountain,

$h_s = h_m /[1+(x/a)^2 + (y/a)^2]^{3/2}$.

Use the same H and $\Delta\rho/\rho_o$ as in project 13.A3 for (ii) and (iii). You may use *surface*, *contour* and *vector* plotting subroutines to plot u' and h' fields. Explain your results. The governing equations are expressed as (3.4.4)-(3.4.6).

13.A5 Change the current finite difference scheme (leapfrog, fourth-order central difference) to (i) forward-in-time and upstream-in-space scheme, and (ii) leapfrog-in-time and second-order-central-difference-in-space scheme.

Using the above schemes to rerun project 13.A4 (ii) and compare the results.

13.A6 Model the three-dimensional, hydrostatic, Boussinesq system of (13.5.1) -(13.5.4).

Figure captions

Fig. 13.1: Sensitivity of flow responses to the numerical domain size for a hydrostatic flow over a bell-shaped mountain. Displayed are the horizontal wind fields for different domain sizes: (a) 22.4*a* and (b) 12*a*, where *a* is the half-width of the mountain. The abscissa and ordinate are non-dimensionalized by *a* and $2\pi U/N$ (vertical wavelength), respectively. The Froude number is 1.2 for both cases.

Fig. 13.2: Potential temperature fields at (a) 450 s and (b) 900 s in a cold pool collapse simulation using adaptive grid meshes. (After Skamarock 1989)

Fig. 13.3: (a) Exact solution of a passive scalar Q at t = 4 s for a kinematic frontogenesis being advected by a steady, nondivergent flow field with structure similar to that of a smoothed Rankine vortex (Doswell 1984). (b) Numerical solution for Q at t = 4 s on a fixed 31x31 uniform grid mesh. (c) Numerical solution for Q at t = 4 s using a continuous dynamic grid adaptation (CDGA). (d) Gridpoint distribution at t = 4 s with default parameters. (Adapted after Dietachmayer and Droegemeier 1992)

Fig. 13.4: A schematic of (a) an unstaggered grid mesh and (b) a staggered grid mesh for the computation of *u* and *w* with the two-dimensional incompressible continuity equation.

Fig. 13.5: A schematic of a staggered grid mesh for the shallow water system.

Fig. 13.6: Phase speed as a function of spatial resolution for the exact solution (E), for second- (2U) and fourth-order (4U) spatial derivative on an unstaggered grid mesh, and for second-order spatial derivatives on a staggered grid mesh (2S). Symbol Δ denotes one grid interval. (After Durran 1998, with kind permission of Springer Sciences and Media.)

Fig. 13.7: (a) Five grid meshes proposed by Arakawa and Lamb (1977). The computational phase velocity ($c_p = \omega/k$) and the group velocity ($\partial\omega/\partial k$) analyzed as functions of kd/π for the four grids as shown in panels (b) and (c), respectively (Schoenstadt 1978). The differential equation solution is also included. These results use the following values: $\sqrt{gH} = 10^2 ms^{-1}$, f=10^{-4} s^{-1}, and d=500 m, where d is the grid interval depicted in panel (a). (Adapted after Arakawa and Lamb 1977)

Fig. 13.8: Reflectivity, from the sponge layer as a function of a nondimensional inverse Reynolds number, 1/*Re* for several nondimensional sponge layer depths (d=0.5, 1.0, 1.5, and 2.0). A value of *r*=1.0 corresponds to complete reflection from the top boundary of the computational domain. The nondimensional numbers are defined as $\mathrm{Re} = U/k\nu_T$ and $d = (z_T - z_1)/\lambda$, where $\lambda = 2\pi U/N$. The viscosity coefficient is defined

as $\nu = \nu_T \sin^2[\pi \ln(\bar{\theta}/\theta_1)/2\ln(\theta_T/\theta_1)]$, where θ_T and θ_1 are the potential temperatures at the top and bottom of the sponge layer, respectively. (Adapted after Klemp and Lilly 1978)

Fig. 13.9: Potential temperature ((a) and (c)) and total horizontal velocity fields ((b) and (d)) for a two-dimensional, continuously stratified, uniform flow over a bell-shaped mountain predicted at the nondimensional time $Ut/a = 100$ by a hydrostatic numerical model [(a) and (b)] and calculated by Long's steady state hydrostatic solution [(c) and (d)]. The Froude number ($F = U/Nh$) and hydrostatic parameter (*Na/U*) associated with the basic flow are 1.3 and 7.7, respectively. The dimensional flow and orographic parameters are U = 13 ms^{-1}, N = 0.01 s^{-1}, h = 1 km, and a = 10 km. The vertical coordinate is nondimensionalized by the hydrostatic wavelength $\lambda = 2\pi U/N$. (After Lin and Wang 1996)

Fig. 13.10: A flow chart for modeling the fluid flow system of (13.2.18) - (13.2.21) and implementing the numerical upper radiation boundary condition at the top of computational domain.

Fig. 13.11: Time evolution of height field after two iterations of the implicit nonlinear normal mode initialization scheme (thin curve) and with no initialization (bold curve). (After Temperton 1988)

Fig. 13.12: Gravity wave activity after normal mode initialization, dynamic initialization for five vertical modes of a baroclinic model, and compared to that with no initialization. Note that the gravity wave modes of large equivalent depths (small j) are dramatically reduced by dynamic initialization. (After Sugi 1986)

Fig. 13.13: An example of data assimilation cycle adopted by NCEP's ETA model by using 3DVAR technique. (Adapted after Rogers et al. 1998)

Fig. 13.14: A sketch of four-dimensional variational data assimilation (4DVAR). (a) The value produced by the first analysis is A, which fits the data well at T-3h, but leads to a forecast that does not match the observations well by T = 0 h. The shaded band is the observations. Note that even data collected at the same time do not necessarily agree

with each other. (b) An iterative approach, i.e. the adjoint method, is taken to adjust the initial analysis so that it is optimal for prediction. (Courtesy of F. H. Carr)

Fig. 13.15: Vertical velocity field from (a) control run, (b) vertical integration of the continuity equation, and (c) adjoint retrieval. (After Sun and Crook 1996)

Fig. 13.16: Schematic illustration of nonlinear aliasing. A physical solution with a wavelength of $1.33\Delta x$, caused by the nonlinear interaction of waves of $2\Delta x$ and $4\Delta x$, is seen as a $4\Delta x$ wave in a numerical grid mesh. (After Pielke 2002, reproduced with permission from Elsvier.)

Fig. 13.17: A flow chart of multianalysis-multimodel superensemble forecasting. The vertical line in the center denotes the initial time ($t = 0$), and the area to the left denotes the training period where a large number of forecast experiments are carried out by the multianalysis-multimodel system. During the training period, the observed fields provide statistics that are then passed on to the period on the right, where $t > 0$. Here the multianalysis-mulimodel forecasts along with the statistics provide the superensemble forecasts. (Adapted after Krishnamurti et al. 2001)

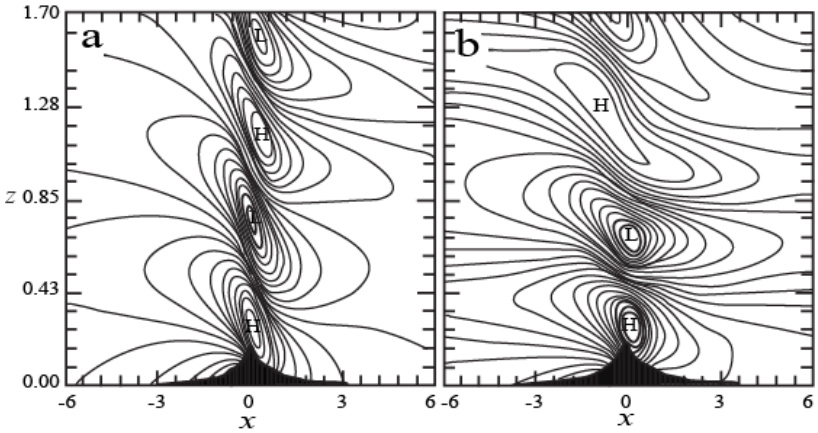

Fig. 13.1: Sensitivity of flow responses to the numerical domain size for a hydrostatic flow over a bell-shaped mountain. Displayed are the horizontal wind fields for different domain sizes: (a) 22.4a and (b) 12a, where a is the half-width of the mountain. The abscissa and ordinate are non-dimensionalized by a and $2\pi U / N$ (vertical wavelength), respectively. The Froude number is 1.2 for both cases.

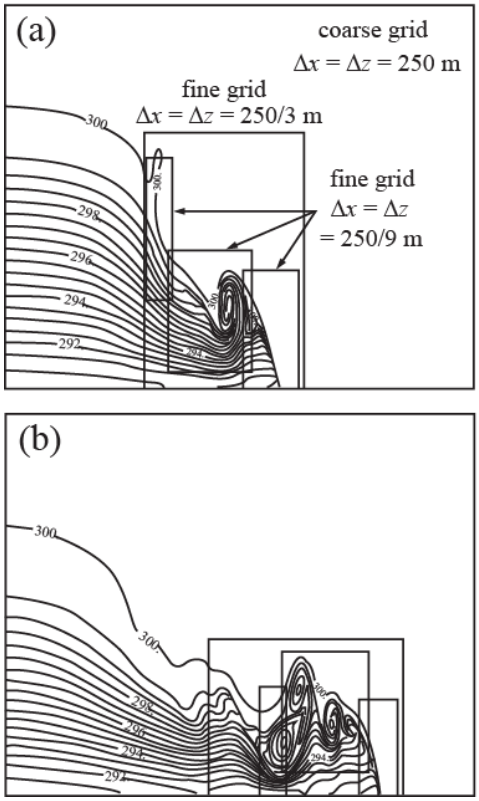

Fig. 13.2: Potential temperature fields at (a) 450 s and (b) 900 s in a cold pool collapse simulation using adaptive grid meshes. (After Skamarock 1989)

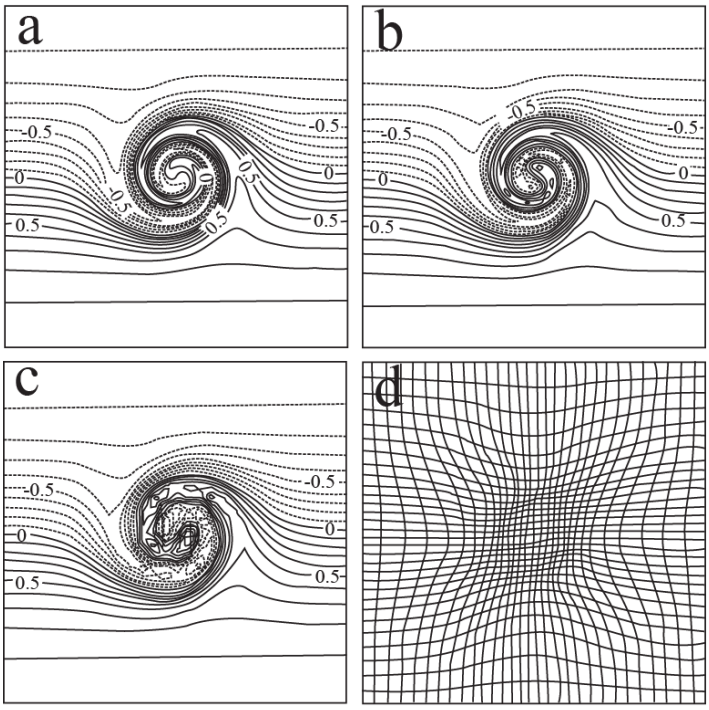

Fig. 13.3: (a) Exact solution of a passive scalar Q at t = 4 s for a kinematic frontogenesis being advected by a steady, nondivergent flow field with structure similar to that of a smoothed Rankine vortex (Doswell 1984). (b) Numerical solution for Q at t = 4 s on a fixed 31x31 uniform grid mesh. (c) Numerical solution for Q at t = 4 s using a continuous dynamic grid adaptation (CDGA). (d) Gridpoint distribution at t = 4 s with default parameters. (Adapted after Dietachmayer and Droegemeier 1992)

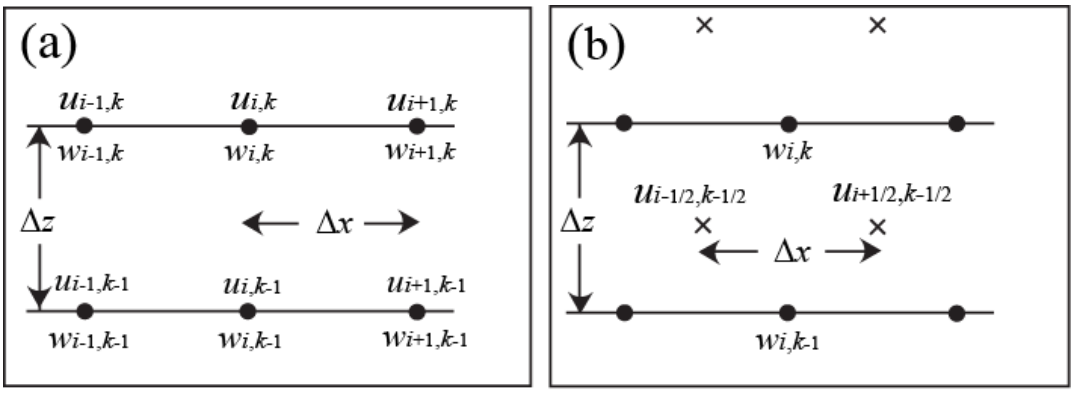

Fig. 13.4: A schematic of (a) an unstaggered grid mesh and (b) a staggered grid mesh for the computation of *u* and *w* with the two-dimensional incompressible continuity equation.

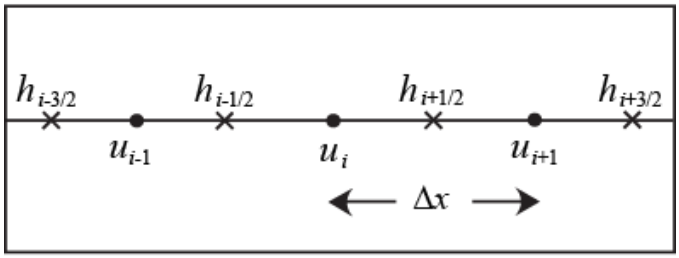

Fig. 13.5: A schematic of a staggered grid mesh for the shallow water system.

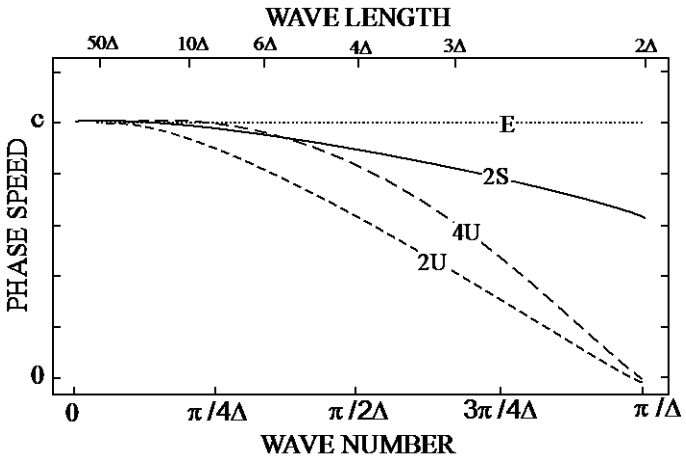

Fig. 13.6: Phase speed as a function of spatial resolution for the exact solution (E), for second- (2U) and fourth-order (4U) spatial derivative on an unstaggered grid mesh, and for second-order spatial derivatives on a staggered grid mesh (2S). Symbol Δ denotes one grid interval. (After Durran 1998, with kind permission of Springer Sciences and Media.)

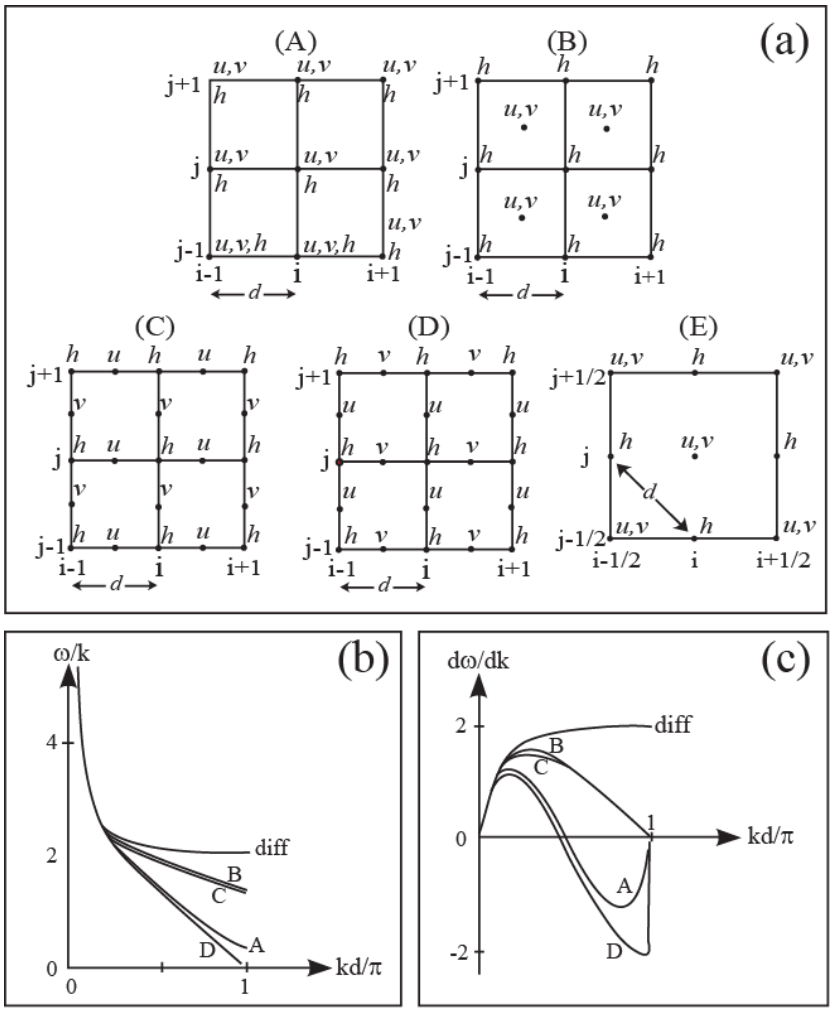

Fig. 13.7: (a) Five grid meshes proposed by Arakawa and Lamb (1977). The computational phase velocity ($c_p = \omega / k$) and the group velocity ($c_g = \partial\omega / \partial k$) analyzed as functions of kd/π for the four grids as shown in panels (b) and (c), respectively (Schoenstadt 1978). The differential equation solution is also included. These results use the following values: $\sqrt{gH} = 10^2 ms^{-1}$, f=10^{-4} s^{-1}, and d=500 m, where d is the grid interval depicted in panel (a). (Adapted after Arakawa and Lamb 1977)

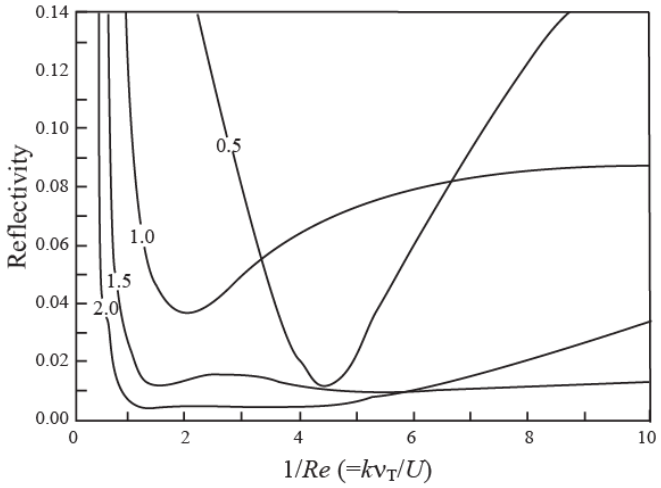

Fig. 13.8: Reflectivity, from the sponge layer as a function of a nondimensional inverse Reynolds number, 1/*Re* for several nondimensional sponge layer depths (*d* = 0.5, 1.0, 1.5, and 2.0). A value of *r*=1.0 corresponds to complete reflection from the top boundary of the computational domain. The nondimensional numbers are defined as $Re = U / k\nu_T$ and $d = (z_T - z_1)/\lambda$, where $\lambda = 2\pi U/N$. The viscosity coefficient is defined as $\nu = \nu_T \sin^2[\pi \ln(\bar{\theta}/\theta_1)/2\ln(\theta_T/\theta_1)]$, where θ_T and θ_1 are the potential temperatures at the top and bottom of the sponge layer, respectively. (Adapted after Klemp and Lilly 1978)

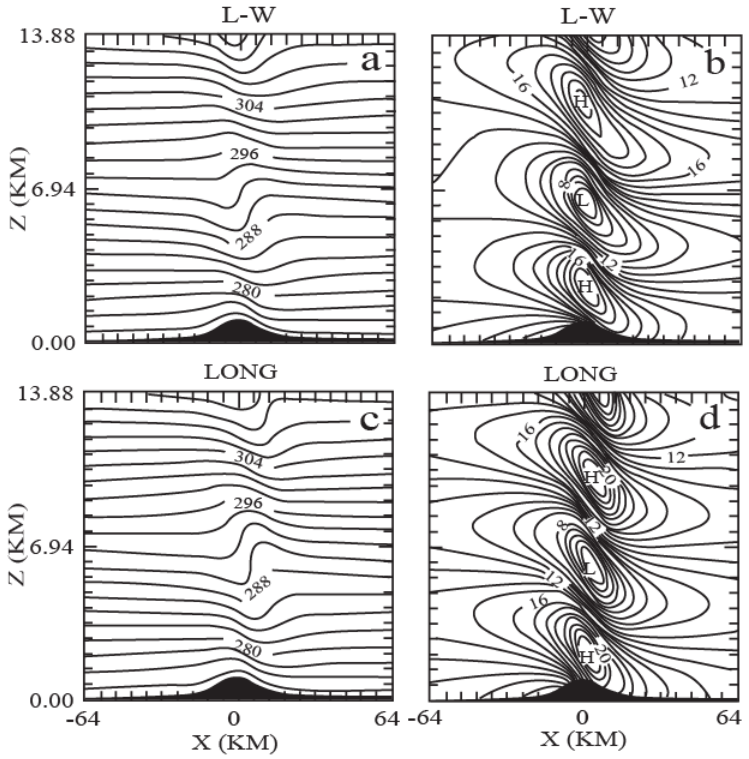

Fig. 13.9: Potential temperature ((a) and (c)) and total horizontal velocity fields ((b) and (d)) for a two-dimensional, continuously stratified, uniform flow over a bell-shaped mountain predicted at the nondimensional time $Ut/a = 100$ by a hydrostatic numerical model [(a) and (b)] and calculated by Long's steady state hydrostatic solution [(c) and (d)]. The Froude number ($F = U/Nh$) and hydrostatic parameter (Na/U) associated with the basic flow are 1.3 and 7.7, respectively. The dimensional flow and orographic parameters are U = 13 ms^{-1}, N = 0.01 s^{-1}, h = 1 km, and a = 10 km. The vertical coordinate is nondimensionalized by the hydrostatic wavelength $\lambda = 2\pi U/N$. (After Lin and Wang 1996)

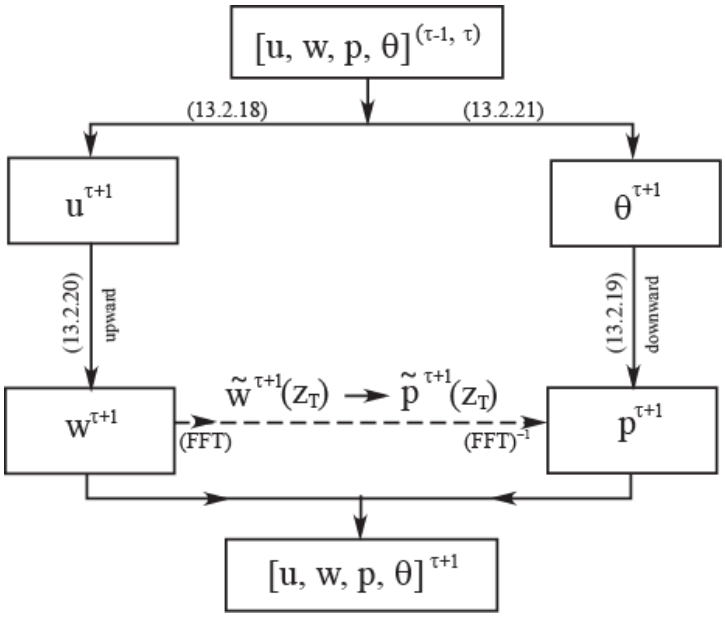

Fig. 13.10: A flow chart for modeling the fluid flow system of (13.2.18) - (13.2.21) and implementing the numerical upper radiation boundary condition at the top of computational domain. FFT and FFT^{-1} denote the Fast Fourier Transform and the inverse Fast Fourier Transform.

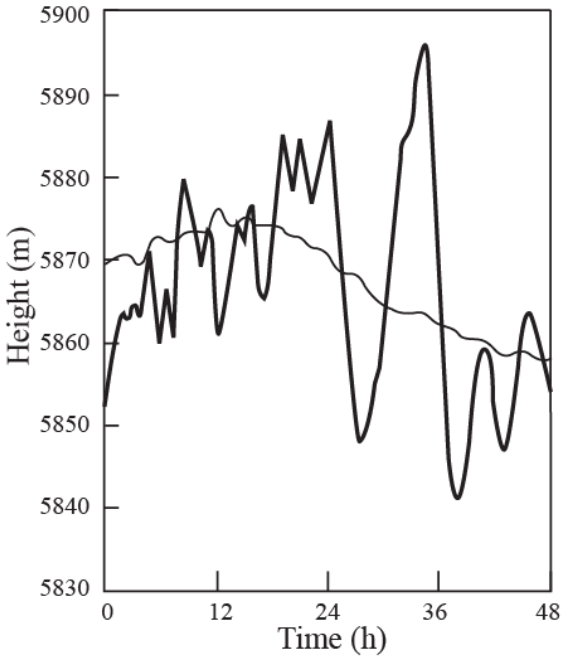

Fig. 13.11: Time evolution of height field after two iterations of the implicit nonlinear normal mode initialization scheme (thin curve) and with no initialization (bold curve). (After Temperton 1988)

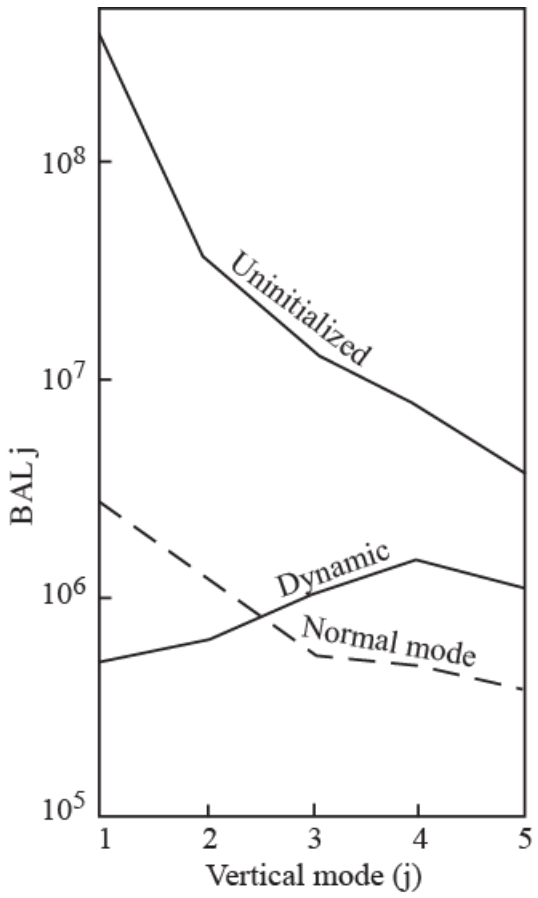

Fig. 13.12: Gravity wave activity after normal mode initialization, dynamic initialization for five vertical modes of a baroclinic model, and compared to that with no initialization. Note that the gravity wave activity is dramatically reduced by dynamic initialization. (Adapted after Sugi 1986)

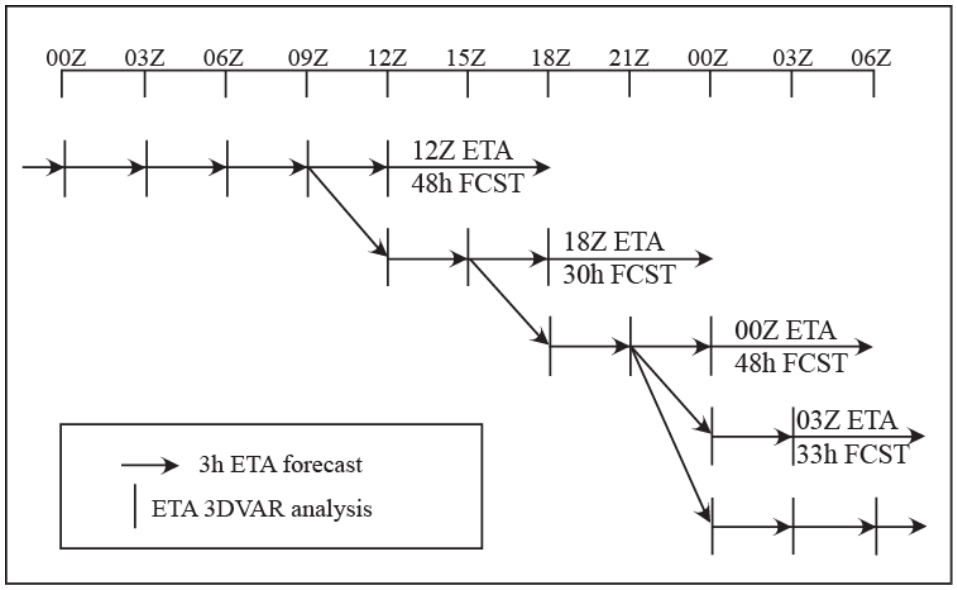

Fig. 13.13: An example of data assimilation cycle adopted by NCEP's ETA model by using 3DVAR technique. (Adapted after Rogers et al. 1998)

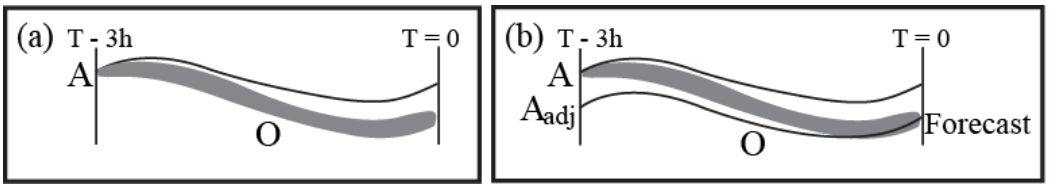

Fig. 13.14: A sketch of four-dimensional variational data assimilation (4DVAR). (a) The value produced by the first analysis is A, which fits the data well at *T*-3h, but leads to a forecast that does not match the observations well by $T = 0$ h. The shaded band is the observations. Note that even data collected at the same time do not necessarily agree with each other. (b) An iterative approach, i.e. the adjoint method, is taken to adjust the initial analysis so that it is optimal for prediction. (Courtesy of F. H. Carr)

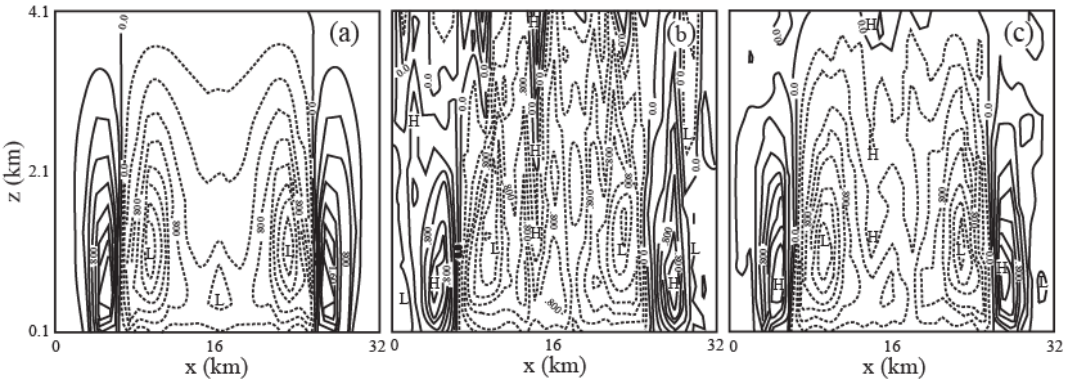

Fig. 13.15: Vertical velocity field from (a) control run, (b) vertical integration of the continuity equation, and (c) adjoint retrieval. (After Sun and Crook 1996)

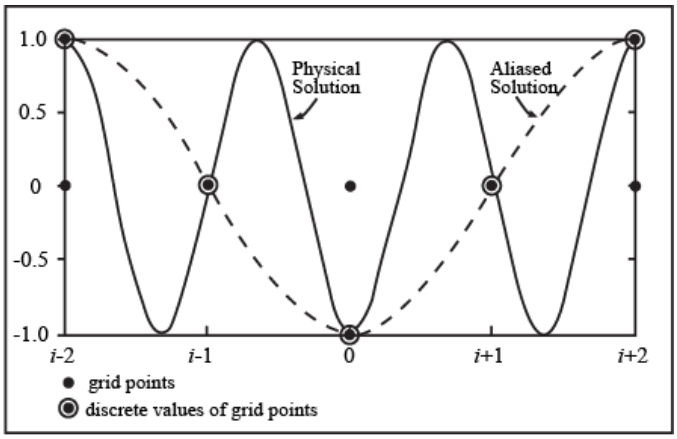

Fig. 13.16: Schematic illustration of nonlinear aliasing. A physical solution with a wavelength of $1.33\Delta x$, caused by the nonlinear interaction of waves of $2\Delta x$ and $4\Delta x$, is seen as a $4\Delta x$ wave in a numerical grid mesh. (After Pielke 2002, reproduced with permission from Elsvier.)

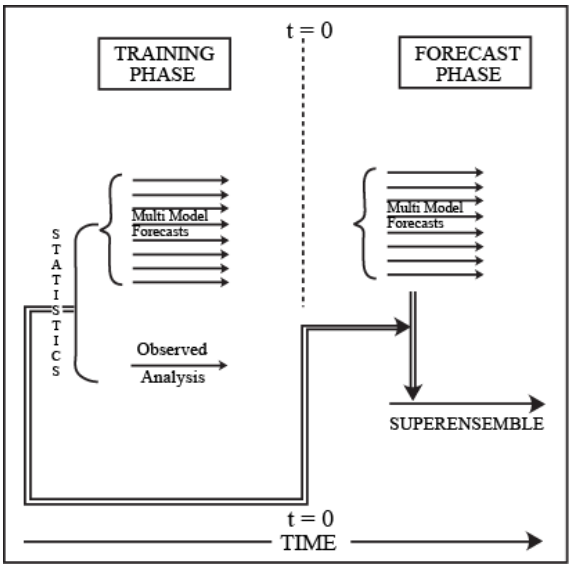

Fig. 13.17: A flow chart of multianalysis-multimodel superensemble forecasting. The vertical line in the center denotes the initial time ($t = 0$), and the area to the left denotes the training period where a large number of forecast experiments are carried out by the multianalysis-multimodel system. During the training period, the observed fields provide statistics that are then passed on to the period on the right, where $t > 0$. Here the multianalysis-mulimodel forecasts along with the statistics provide the superensemble forecasts. (Adapted after Krishnamurti et al. 2001)