We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

4,200
Open access books available

116,000
International authors and editors

125M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us? Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
1. Introduction

Since the early 20th century numerical weather prediction (NWP) has increasingly become one of the most important and complicated problems of modern science. With the advent of computers, increased observations, and progress in theoretical understanding, numerical models were developed. Since then, such models are playing an increasing role in understanding and predicting weather and climate and have been a driving force in the advancement of the meteorological sciences. Numerical models are a mathematical representation of the earth’s climate system including the atmosphere, ocean, cryosphere and land, among others. The models divide the area of interest into a set of grids and then make use of observations of variables such as surface pressure, winds, temperature and humidity at numerous locations throughout the globe. The observed values are then assimilated and used by the model to predict future evolution of the earth’s weather and climate. In the mid 20th century, models evolved from a simple model with a single atmospheric layer to a multi-layer primitive equation model capable of predicting cyclone development [1].

Due to the amount of computer processor time, memory, and disk storage required to run numerical models, the atmosphere cannot be represented perfectly by the model and thereby is approximated by a finite data set. The atmosphere is represented in a model by a three-dimensional set of points, called grids that cover the region of interest. Figure 1 demonstrates the importance of the number of discrete grid points in order for the model to best represent the atmospheric structures. Figure 1a uses a grid spacing of 1 in. whereas Figure 1b uses a grid spacing of 0.5 in. Comparison of the two figures shows how increasing the number of grid points allows the model to better represent the actual wave function. As the number of discrete grid points increases, so increases the representation of the atmos-
The NWP models in the 1950s had grid points every few hundred kilometers in the horizontal, whereas today, models used in operational forecasting have grid points every 10-100 km. The horizontal distance between the adjacent grids are often known as grid-spacing or resolution. Many models have the ability to nest finer grids within a coarse grid resulting in a nested grid with much higher resolution. The vertical resolution of numerical models improved alongside horizontal resolution, such that models today can have more than 50 vertical layers.

The resolution of a model also depends on the area coverage. General circulation models (GCMs), which are global, typically have coarse resolutions and are necessary for long-range forecasts. Regional models often called limited area models, have finer resolution and are used for short-range forecasts. Limited area models can be run closer to real time because they can be initialized using local observations, but still depend on other model output for boundary conditions. On the other hand, global models must wait for observations from around the globe for their initializations. With the increase in computer power, high-resolution regional models are covering larger area, whereas, global models are having finer resolutions. It is expected that cloud-system-resolving global models will be common in the coming decades.

Figure 1. Idealized grid point approximation (gray) of a function z (black) plotted on the interval \([0, 8]\). Grid points every 1 in (a) and every 0.5 in. (b) are indicated by black dots. (Source: [1])

An integral part of any weather and climate model is the numerical schemes that are designed to convert the set of partial differential equations (PDE’s) that represent geophysical equations into a set of algebraic expressions. This is essential because computers cannot
solve PDE’s directly; but can perform algebraic operations like addition and multiplication. The conversion of PDE’s to simple algebraic equations involves primarily two steps; computation of a derivative and representation of the solution by a finite data set. This can be achieved using two different methods:

i. Series expansion method: In this method, an unknown function is approximated by a linear combination of a finite set of continuous expansion functions. When the expansion functions are orthogonal, the method is known as the spectral method. When the expansion functions are non-zero in only a small part of the local domain, the method is called the finite element method. The finite element method is generally not as efficient as the spectral method because it usually requires the solution of implicit algebraic equations and is mostly time-independent. However, there are certain situations when the finite element method is useful, such as approximating the vertical structure of the atmosphere or in steady state situations [2]. The spectral method is used extensively for global medium-range forecasting models and global climate models. Gates [3] noted that out of 29 models that participated in AMIP (Atmospheric Model Inter-comparison Project), 19 were spectral.

ii. Grid-point method: In this method, the equations are approximated on grid points using a finite set of data. Finite difference is an example of the grid-point method. Series expansion and grid-point methods are used extensively for spatial derivatives [4]. However, the time derivative is almost always approximated using the finite difference method. Most short-range and regional models are grid-point models.

The development of modern discretization techniques like that of semi-implicit and the semi-Lagrangian schemes, have kept the cost of numerical modeling in check by having less stringent stability conditions on the time-step and more accurate space discretization [4, 5].

The chapter is organized as follows: Section 2 documents a brief history of the model grids. Section 3 describes different shapes of grids, followed by staggering of grids in section 4. Further discussions are presented in section 5, followed by conclusions in section 6.

2. A brief history of model grids

Earlier days of real time numerical weather prediction were far from what is now considered “the greatest intellectual achievement and scientific advancement in twentieth century atmospheric science” [4]. Bjerknes [6] recognized for the first time that NWP is an initial value problem. Prior to Bjerknes, however, Cleveland Abbe, who was the first chief meteorologist of the US weather bureau, prepared and issued the first official weather forecast in the US on February 19, 1871. He introduced standard forecast verification procedures and wrote a paper on “The Physical basis of Long-range Weather Forecasts” [7], where he correctly wrote the set of primitive equations that govern the atmospheric motions.

The idea of forecasting weather by numerical process using physically based models was first proposed by Lewis Fry Richardson [8]. Following the basic idea of an initial value prob-
lem, if the values of certain environmental variables are known, then the physical equations can be used to calculate their values at a time in the future. Richardson was able to simplify the equations of motions presented by Abbe and later by Bjerknes using his knowledge of meteorology and mathematics. Richardson proposed to divide the earth’s surface into a grid, with each grid cell the base of a vertical column of the atmosphere. Each vertical column was then divided into several layers, resulting in a three-dimensional grid of atmospheric boxes. In Figure 2, each column (or grid box) extends 3° in the east-west direction and 200 km in the north-south direction. This resulted in 12,000 columns to cover the entire globe. To test his technique and provide an example of how to use it, Richardson performed the calculations for two adjacent columns [8]. Each column was divided into five layers at heights of 2, 4.2, 7.2, and 11 km (or about 800, 600, 400 and 200 hPa) above sea-level and the values of the variables were fixed at the center of each box (Figure 3). He computed only the initial tendency at a single point for pressure at the base of each layer, temperature at the stratosphere, water content at the lower four layers and the two components of horizontal wind. His calculation of change of pressure at the point considered was 145 hPa in 6 hrs., an obviously unrealistic value. His forecast failed as a result of short-period oscillations called gravity waves that created “noise” in the observed data set, thereby causing error in the initial conditions used in his forecast.

Figure 2. Grid used by Richardson in 1922 to calculate the pressure change in central Germany. X-axis shows the longitude and Y-axis shows the distance (in km) from the equator. Each grid box is 3° in the longitudinal direction and 200 km in the latitudinal direction. (Source: [9])
3. Shapes of grids

Richardson’s effort of predicting weather using grid points (as seen in Figures 2, 3) set the stage for future development of grids in different shapes. In order to accommodate the spherical shape of the earth and represent the equations more accurately and efficiently, there are different grid shapes used in numerical models; e.g., rectangular, triangular, and hexagonal (Figure 4). A brief description of these grids along with their relative advantages and disadvantages are presented next with an emphasis on the following criteria: (i) Suitability for cloud-scale to global-scale; (ii) Efficiency on different computer architectures and scalability on massively parallel computers; (iii) Conservation of mass and other quantities; and (iv) Capability of local grid refinement and regional domains. However, the method of distribution of grid points over the sphere is yet to be solved in a fully satisfactory manner.

![Figure 3. Vertical grid used by Richardson (1922).](image1)

![Figure 4. Examples of (a) rectangular or latitude-longitude grid, (b) triangular grid and (c) hexagonal grid.](image2)
3.1. Rectangular/square grids

The rectangular/square (or latitude-longitude) grid is the most commonly used grid in the NWP models (e.g., [10]). The rectangular grid is simple in nature but suffers from “the polar problem” where the lines of equal longitude, known as meridians, converge to points at the poles (Figure 5a). The poles are unique points and may cause violation of global conservation laws within the model. To maintain computational stability near the poles, small integration time-steps could be used, but at great expense. The high resolution in the east-west direction near the poles would be wasted because the model uses lower resolution elsewhere ([11]).

![Figure 5. (a) Demonstration of the polar problem in which the meridians converge to a single point at the poles and (b) its remedy using the Kurihara grid.](image)

A rotated grid can overcome the polar problem for limited area models ([4]), whereas for global models, other grid shapes are used. For example, Kurihara [12] proposed to use ‘skipped’ or ‘Kurihara’ grid (Figure 5b). Unfortunately use of the Kurihara grid causes spuriously high pressure to develop at the poles. As a result, their use has been severely limited or abandoned in finite difference models. However, problems due to the use of the Kurihara grid can be resolved by using more accurate numerical schemes [13]. In the late 60’s and early 70’s, the application of quasi-uniform grids was proposed as a method to avoid the polar problem of the grid-point models [14]. For example, the Global Forecasting System (GFS) model has roughly a square grid near the equator, a more rectangular grid in the mid-latitudes, and a triangular grid near the poles, eventually converging to a point at the poles. Another example of a model that uses the rectangular grid type is the North American Mesoscale Model (NAM). When compared to the resolution of the GFS, the NAM does not have a grid stretching problem since the model calculates variables close to the poles. This is due to the NAM not relying on a latitude-longitude system for creating its grid bounds and deferring to a more precise horizontal measurement system. The other problem with the latitude-longitude grid is the need for special filters to deal with the pole singularities. They also do not scale well on massively parallel computers.
3.2. Triangular grids

Triangular grids are not used as often in models as are rectangular grids. One form of quasi-uniform grid whose base element is a triangle is the spherical geodesic grid. Icosahedral grids, first introduced in the 1960s, give almost homogeneous and quasi-isotropic coverage of the sphere. The grid is made by dividing the triangular faces of an icosahedron into smaller triangles, the vertices of which are the grid points. Each point on the face or edge of one of the faces of the icosahedron is surrounded by six triangles making each point the center of a hexagon. The triangular faces of the icosahedrons are arranged into pairs to form rhombuses, five around the South Pole and five around the North Pole. The poles are chosen as two pentagonal points where the five rhombuses meet. The main advantage of the geodesic grid is that all the grid cells are nearly the same size. The uniform cell size allows for computational stability even with finite volume schemes.

3.3. Hexagonal grids

Similar to triangular grids, hexagonal grids are also not used as often as the rectangular/square grids. In this method, variables are calculated at each grid intersection between different hexagons, in addition to being calculated in the center of the hexagonal grid. Sadourny et al. [15] describes in detail how the spherical icosahedral-hexagonal grid is constructed. They solved the non-divergent barotropic vorticity equation with finite difference methods on the icosahedral-hexagonal grids. Majewski et al. [16] utilizes an approach that uses local basis functions that are orthogonal and conform perfectly to the spherical surface. A study done by Thuburn [17] shows a method of creating a global hexagonal grid, but then using a finite differencing method to calculate the rate-of-change of different variables without having to create triangles within the hexagonal grids. The space differencing scheme using the icosahedral-hexagonal grid gives a satisfactory approximation to the analytical equations given an initial condition and remains nonlinearly stable, for any condition.

Thuburn [17] also noted that his method may not be as accurate as those which included an additional point in the center of the hexagonal grid, but his method was computationally faster, and was able to accurately depict the polar regions since there was no need of stretching the grid in that region. The other advantages of the hexagonal grid are: (i) Removes the polar problem. (ii) Permits larger explicit time steps. (iii) Most isotropic compared to other grid types. (iv) Conservation of quantities in finite volume formulation. (v) Can be generalized easily to arbitrary grid structures.

4. Grid staggering

After the choice of distribution of grid points (i.e., rectangular, triangular, hexagonal, etc.), the next step is to arrange the prognostic variables on the grid. When all the prognostic variables are defined at the same point in a grid, it is called an unstaggered grid. On the other hand, when prognostic variables are defined at more than one point in a grid, it is called a staggered grid. Characteristically a staggered grid has the values of the wind components at
separate points than the thermodynamic variables within the grid cell. Consequently, a model’s resolution is defined as the average distance between adjacent grid points with the same variables.

Staggering is not only performed in the horizontal direction, but also in the vertical direction, as well as in time, and any combinations thereof.

4.1. Horizontal staggering

There are a variety of different methods in which models calculate temperature, winds, surface pressure, geopotential, and other meteorological variables within their grids. Five different types of grids were introduced by Arakawa and Lamb [18] and are shown in Figure 6. Of these grids, A is an unstaggered grid where the variables are defined at the same points, e.g., at the center or at the corners of the grid. Grids B through E are all staggered grids where the variables are defined at different points. Since all variables are defined at all the grid points in the A grid, it is easy to construct a higher-order accurate scheme. The main disadvantage is that the differences are computed over a distance of 2Δx, and the adjacent points are not coupled for the pressure and convergence terms. In grid B, evaluation of the two sets of variables are at different points, e.g., one might evaluate the velocities at the center of a grid and masses at the grid corners. Since the B and E grids have wind components at the same point, they are often called semi-staggered. In grid C, velocities are calculated at the mid-point between grid cells and h is calculated at the corners (or intersections of grid cells). The main advantage of the C grid is that the pressure and convergence terms are computed over a distance Δx, which is half of that in the A grid indicating a doubling of the resolution compared to the A grid. Most non-hydrostatic mesoscale models like fifth generation mesoscale model (MM5) and Weather Research and Forecasting (WRF) use the C grid. The B grid was chosen in the UK Met Office (UKMO) unified model [19, 20] for climate simulation as well as numerical weather prediction.

From the grids presented in Figure 6 it can be seen that the resolution of models does not depend solely on the size or shape of the grid but also the location in which the model calculates various atmospheric variables. As the number of grid cells increases, the differences in effective resolution between different grid types eventually goes to zero as the number of grid cells goes to infinity [21].

In Figure 6, grid D is a slight variation of grid C, with the u and v variables being oriented with a rotation of 90°. This variation allows for a simple evaluation of the geostrophic wind. This is done by creating better averages for variables such as pressure gradient, mass convergence/divergence, and the Coriolis terms [21]. The D grid was used (with time staggering) in National Meteorological Center’s (NMC’s) nested grid model [22]; however, this grid is no longer used in any popular atmospheric model because of no added benefit.

The staggered E grid is rotated 45° relative to the B grid, but has an increased grid-spacing (Figure 6e) compared to the B grid. One problem with this grid is when the domain is small and one-dimensional, this grid is equivalent to grid A, but with less computational efficiency. National Centers for Environmental Prediction (NCEP) eta model uses grid E [23].
Although staggered grids have higher equivalent resolution than unstaggered grids, they are also more complex. Overall, the C grid is becoming more popular in recent times with the E grid its closest competitor.

Figure 6. Grid types based on horizontal staggering, namely, (a) the A grid, (b) the B grid, (c) the C grid, (d) the D grid and (e) the E grid. Δx is the east-west resolution and Δy is the north-south resolution.
4.2. Vertical staggering

Staggering of grids in the vertical direction also provides certain advantages. For example, vertical staggering introduced by Lorenz [24] maintains the requirement of boundary conditions of no flux at the top and bottom (Figure 7a). However, the Lorenz grid allows the formation of a spurious computational mode [25]. This problem does not exist in the Charney-Phillips grid ([26]; Figure 7b) in which the vertical staggering, being more consistent (compared to Lorenz grid) with hydrostatic equation, does not allow the additional computational mode [27]. Most of the present state-of-the-art numerical models have staggered grids in the vertical direction with prognostic variables at the center of the layer and the vertical velocity at the boundary of the layers. Such an example is shown in Figure 8 from the WRF model [28].

![Figure 7. Staggered grids in the vertical following (a) Lorenz [24] & (b) Charney-Phillips [26].](image)

![Figure 8. Vertically staggered grid from the WRF model. (from [28])](image)
4.3. Time staggering

Staggering of grids is not confined in space, staggering can also be in time. For example, for atmospheric flow using the leapfrog scheme grid D is ideal when staggered in time (Figure 9). Time staggering was first introduced by Eliassen [29], which involves defining variables at every second time step on an offset D grid. A slight variation of this approach performed by Bratseth [30] used a higher-order interpolation to transfer values back from the offset grid to the original grid. All the differences are calculated on a distance $\Delta x$. Despite this advantage, such time staggering is not used due to the complexities that arise from the additional staggering and need of special procedures for starting the leapfrog scheme.

Figure 9. Staggered grids in space and time with (a) even time steps (b) odd time steps

5. Discussions

The ability of the NWP model to accurately represent atmospheric phenomena is based on three conditions; scientific knowledge, the availability of observational data, and computer processing abilities. If enough observational data is available and enough scientific knowledge is present then the limiting factor of an accurate forecast is the power of the processing computer. There are other issues of relevance that warrant further discussion.

5.1. Grid splitting

The B and E grids discussed in section 4.1 can be considered as being made up of two C grids. An example is shown in Figure 10 for the B grid. Ignoring the distinction between variables in upper- and lower-case characters, the figure represents a B grid. Considering only the lower-case characters, the figure represents a C grid whose axes are rotated $45^\circ$ counterclockwise relative to that of the B grid. On the other hand, considering only the upper-case characters, the figure represents a second C grid that is shifted by one grid-length along the x-axis of the B grid. Precaution must be taken in formulating B- and E-grid models to avoid solutions splitting into two separate distributions on the two C grids [31].
5.2. Nested grids

Some models are run with finer-resolution grids nested inside coarser-resolution grids within the same model. Grid nesting is used when computational limitations prohibit fine-resolution grids from covering the entire model domain. Nesting can be one-way or two-way. Information in a two-way nested grid is shared both ways, from the coarse-grid to the fine-grid and from the fine-grid to the coarse-grid. In Figure 11, where the fine-grid covers the coarse-grid, the forecast variables for the coarse-grid are updated based on the fine-grid prediction. The coarse-grid prediction provides boundary conditions on the nest interface for use in the fine-grid prediction. Advantages of the two-way nested grid include, fine-scale processes resolved on the finer grid are allowed to affect the larger-scale flow on the coarse-grid. This is important for numerical weather prediction because the small-scale processes in the atmosphere greatly influence the large-scale processes in the atmosphere. Since the predictions on coarse-resolution grids take less computer time and memory compared to fine-resolution grids, the outermost boundary of the model can be moved far from the forecast region, while the fine-resolution domain remains small enough to run in real time. Moving nests are also common in the present models where a higher resolution nest can move with the phenomenon of interest (e.g., hurricane) to provide details that wouldn’t be possible in a coarse resolution simulation.

An example of staggered C grid with nesting is also shown in Figure 12.
Figure 11. Example of a two way nested grid with coarse resolution outer domain and finer resolution inner domains. Staggering type is typically same for all domains. The arrows indicate direction of information exchange. (Source: www.comet.ucar.edu)

Figure 12. A portion of a nested grid with C grid staggering with 3:1 grid size ratio. The solid lines denote coarse-grid cell boundaries, and the dashed lines are the boundaries for each fine grid cell. The bold typeface variables along the interface between the coarse- and the fine-grid define the locations where the specified lateral boundaries for the nest are in effect. (Taken from [28]).
5.3. Mesh refinement

Adaptive mesh refinement is a method where model resolution is refined and the solution has fine-scale structure, rather than in a fixed region as done in a conventional nesting. This is being used in the model for prediction across scales [32, 33, 34]. It has all the advantages of a hexagonal grid as described in section 3.3. It uses centroidal Voronoi tessellations with a C grid staggering and can be used for global (Figure 13a) and regional (Figure 13b) applications. Compared to traditional grid nesting, Voronoi meshes can cleanly incorporate both downscaling and upscaling effects. It can also handle variable resolution at any region of interest even using other grid shapes. For example, Figure 14 shows selective mesh refinement based on terrain height. Note that the high terrains are accompanied by higher resolutions.

![Figure 13](image1.png)

*Figure 13.* Variable resolution mesh at (a) global and (b) regional scale [32].

![Figure 14](image2.png)

*Figure 14.* Selective mesh refinement based on centroidal Voronoi terrain height (courtesy Michael Duda).

5.4. Parametrization and grid-spacing

Parameterizations approximate the combined effects of physical processes (e.g., cumulus convection, radiation, microphysics, planetary boundary layer) that are too small, too com-
plex, or too poorly understood to be explicitly represented in numerical models. Models with a grid spacing of 10 km or larger, which is roughly 10 to 20 times the size of the cumulus cloud, needs much finer resolution to resolve small cumulus clouds well. Parameterization schemes in a model are often optimized for a certain range of grid-spacing in the model. Such dependence on the grid-spacing makes a parameterization not suitable when the grid-spacing of the model decreases or increases. Most mesoscale models that can be used over a wide range of grid-spacing typically have multiple parameterization schemes for the same process to be used at multiple resolutions. Such dependence of parameterization schemes on the grid-spacing often works as an obstacle towards the use of higher resolution.

6. Summary

Since the first forecast, the complexity and sophistication of numerical weather prediction models have increased tremendously [35]. The continued improvement in data assimilation and numerical models, and the continued availability of ever larger and faster computers have allowed numerical weather prediction to become more accurate [36]. Most of the improvements in numerical models that occurred over the past 50 years can be categorized as either improved numerical techniques, improved model resolution, or improved model physics. Development of different types of grids, grid-staggering and grid-spacing are intimately related to the improvement in the weather and climate prediction.

There are three major shapes of grids, namely latitude-longitude, triangular and hexagonal. Although latitude-longitude grid is the most common, it suffers from the polar problem. This problem can be overcome by using a triangular or hexagonal grid. Variables in a grid can be defined at the same point (“unstaggered grid”) or at different points (“staggered grid”). The staggered grid often has higher effective resolution than the unstaggered grid and is used in almost all popular models. Staggering can also be implemented in the vertical direction and in time. Physics parameterization schemes in a model are often optimized for a certain range of grid-spacing. Such dependence on the grid-spacing makes a parameterization not suitable when the grid-spacing of the model decreases or increases. Further development of the grids is needed to better represent the variables in a model for weather and climate prediction.

Author details

Sarah N Collins, Robert S James, Pallav Ray*, Katherine Chen, Angie Lassman and James Brownlee

Department of Marine and Environmental Systems, Florida Institute of Technology, Melbourne, FL, USA
References


