Numerical Weather Prediction

Peter Lynch

1 Introduction

The development of computer models for numerical simulation and prediction of the atmosphere and oceans is one of the great scientific triumphs of the past fifty years. Today, numerical weather prediction (NWP) plays a central and essential role in operational weather forecasting, with forecasts now having accuracy at ranges beyond a week. There are several reasons for this: enhancements in model resolution, better numerical schemes, more realistic parametrizations of physical processes, new observational data from satellites, and more sophisticated methods of determining the initial conditions. In this article we focus on the fundamental equations, the formulation of the numerical algorithms, and the variational approach to data assimilation. We present the mathematical principles of NWP and illustrate the process by considering some specific models and their application to practical forecasting.

2 The Basic Equations

The atmosphere is governed by the fundamental laws of physics, expressed in terms of mathematical equations. They form a system of coupled nonlinear partial differential equations (PDEs). These equations can be used to predict the evolution of the atmosphere and to simulate its long-term behavior.

The primary variables are the fluid velocity $\boldsymbol{V}$ (with three components, u eastward, v northward, and w upward), pressure p, density ρ, temperature T, and humidity q. Using Newton's laws of motion and the principles of conservation of energy and mass, we can obtain a system whose solution is well determined by the initial conditions.

The central components of the system are the NAVIER–STOKES EQUATIONS [**??**] governing fluid motion. We write them in vector form:

$$\frac{\partial \boldsymbol{V}}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla} \boldsymbol{V} + 2\boldsymbol{\Omega} \times \boldsymbol{V} + \frac{1}{\rho}\boldsymbol{\nabla} p = \boldsymbol{F} + \boldsymbol{g}.$$

The equations are relative to the rotating Earth and $\boldsymbol{\Omega}$ is the Earth's angular velocity. In order, the terms of this equation represent local acceleration, nonlinear advection, Coriolis term, pressure gradient, friction, and gravity. The friction term $\boldsymbol{F}$ is small in the free atmosphere but is crucially important in the boundary layer (roughly, the first 1 km above the Earth's surface). The apparent gravity $\boldsymbol{g}$ includes the centrifugal force, which depends only on position.

The temperature, pressure, and density are linked through the equation of state

$$p = R\rho T,$$

where R is the gas constant for dry air. In practice, a slight elaboration of this is used that takes account of moisture in the atmosphere.

Energy conservation is embodied in the first law of thermodynamics,

$$c_{\mathrm{v}} \frac{\mathrm{d}T}{\mathrm{d}t} + RT\boldsymbol{\nabla} \cdot \boldsymbol{V} = Q,$$

where c_{v} is the specific heat at constant volume and Q is the diabatic heating rate. Conservation of mass is expressed in terms of the continuity equation:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho\boldsymbol{\nabla} \cdot \boldsymbol{V} = 0.$$

Finally, conservation of water substance is expressed by the equation

$$\frac{\mathrm{d}q}{\mathrm{d}t} = S,$$

where q is the specific humidity and S represents all sources and sinks of water vapor.

Once initial conditions, appropriate boundary conditions, and external forcings, sources, and sinks are given, the above system of seven (scalar) equations provides a complete description of the evolution of the seven variables $\{u, v, w, p, \rho, T, q\}$.

For large-scale motions the vertical component of velocity is very much smaller than the horizontal components, and we can replace the vertical equation by a balance between the vertical pressure gradient and gravity. This yields the hydrostatic equation

$$\frac{\partial p}{\partial z} + g\rho = 0.$$

Hydrostatic models were used for the first fifty years of NWP but nonhydrostatic models are now coming into widespread use.

3 The Emergence of NWP

The idea of calculating the changes in the weather by numerical methods emerged around the turn of the twentieth century. Cleveland Abbe, an American meteorologist, viewed weather forecasting as an application of hydrodynamics and thermodynamics to the atmosphere. He also identified a system of mathematical equations, essentially those presented in §2 above, that govern the evolution of the atmosphere. This idea was developed in greater detail by the Norwegian Vilhelm Bjerknes, whose stated goal was to make meteorology an exact science: a true physics of the atmosphere.

3.1 Richardson's Forecast

During World War I, Lewis Fry Richardson, an English Quaker mathematician, calculated the changes in the weather variables directly from the fundamental equations and presented his results in a book, *Weather Prediction by Numerical Process*, in 1922. His prediction of pressure changes was utterly unrealistic, being two orders of magnitude too large. The primary cause of this failure was the inaccuracy and imbalance of the initial conditions. Despite the outlandish results, Richardson's methodology was unimpeachable, and is essentially the approach we use today to integrate the equations.

Richardson was several decades ahead of his time. For computational weather forecasting to become a practical reality, advances on a number of fronts were required. First, an observing system for the troposphere, the lowest layer of the atmosphere, extending to about 12 km, was established to serve the needs of aviation; this also provided the initial data for weather forecasting. Second, advances in numerical analysis led to the design of stable and accurate algorithms for solving the PDEs. Third, progress in meteorological theory, especially the development of the quasigeostrophic equations and improved understanding of atmospheric balance, provided a means to eliminate the spurious high-frequency oscillations that had spoiled Richardson's forecast. Finally, the invention of high-speed digital computers enabled the enormous computational task of solving the equations to be undertaken.

3.2 The ENIAC Integrations

The first forecasts made using an automatic computer were completed in 1950 on the ENIAC (Electronic Numerical Integrator and Computer), the first programmable general-purpose computer. The forecasts used a highly simplified model, representing the atmosphere as a single layer and assuming conservation of absolute vorticity expressed by the barotropic vorticity equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\zeta + f) = 0,$$

where ζ is the vorticity of the flow and $f = 2\Omega \sin\phi$ is the Coriolis parameter, with Ω the angular velocity of the Earth and ϕ the latitude. The Lagrangian time derivative

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla}$$

includes the nonlinear advection by the flow. The equation was approximated by finite differences in space and time with a grid size of 736 km (at the North Pole) and a time step of three hours. The resulting forecasts, while far from perfect, were realistic and provided a powerful stimulus for further work.

Baroclinic, or multilevel, models that enabled realistic representation of the vertical structure of the atmosphere were soon developed. Moreover, the simplified equations were replaced by more accurate primitive equations, that is, the equations presented in §2 but with the hydrostatic approximation. As these equations simulate high-frequency gravity waves in addition to the motions that are important for weather, the initial conditions must be carefully balanced. Techniques for ensuring this were developed. Most notable among these was the normal-mode initialization method: the flow is resolved into normal modes and modified to ensure that the tendencies, or rates of change, of the gravity wave components vanish. This suppresses spurious oscillations.

4 Solving the Equations

Analytical solution of the equations is impossible, so approximate methods must be employed. We consider methods of discretizing the spatial domain to reduce the PDEs to an algebraic system, and of advancing the solution in time.

4.1 Time-Stepping Schemes

Let Q denote a typical dependent variable, governed by an equation of the form

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = F(Q).$$

We replace the continuous-time domain t by a sequence of discrete times $\{0, \Delta t, 2\Delta t, \dots, n\Delta t, \dots\}$, with the solution at these times denoted by $Q^n = Q(n\Delta t)$. If this solution is known up to time $t = n\Delta t$, the right-hand term $F^n = F(Q^n)$ can be computed. The time derivative is now approximated by a centered difference

$$\frac{Q^{n+1} - Q^{n-1}}{2\Delta t} = F^n,$$

so the "forecast" value Q^{n+1} may be computed from the old value Q^{n-1} and the tendency F^n:

$$Q^{n+1} = Q^{n-1} + 2\Delta t F^n.$$

This is called the *leapfrog scheme*. The process of stepping forward from moment to moment is repeated a large number of times, until the desired forecast range is reached.

The leapfrog scheme is limited by a stability criterion that restricts the size of the time step Δt. One way of circumventing this is to use an implicit scheme such as

$$\frac{Q^{n+1} - Q^{n-1}}{2\Delta t} = \frac{F^{n-1} + F^{n+1}}{2}.$$

The time step is now unconstrained by stability, but the scheme requires the solution of the equation

$$Q^{n+1} - \Delta t F^{n+1} = Q^{n-1} + \Delta t F^{n-1},$$

which is prohibitive unless $F(Q)$ is a linear function. Normally, implicit schemes are used only for particular (linear) terms of the equations.

4.2 Spatial Finite Differencing

For the PDEs that govern atmospheric dynamics we must replace continuous variations in space by discrete variables. The primary way to do this is to substitute finite-difference approximations for the spatial derivatives. It then transpires that the stability depends on the *relative* sizes of the space and time steps. A realistic solution is not guaranteed by reducing their sizes independently.

We consider the simple one-dimensional wave equation

$$\frac{\partial Q}{\partial t} + c\frac{\partial Q}{\partial x} = 0,$$

where $Q(x,t)$ depends on both x and t, and where the advection speed c is constant. We consider the sinusoidal solution $Q = Q^0 \exp[\mathrm{i}k(x - ct)]$ of wavelength $L = 2\pi/k$. We use centered difference approximations in both space and time:

$$\frac{Q_m^{n+1} - Q_m^{n-1}}{2\Delta t} + c\left(\frac{Q_{m+1}^n - Q_{m-1}^n}{2\Delta x}\right) = 0,$$

where $Q_m^n = Q(m\Delta x, n\Delta t)$. We seek a solution of the form $Q_m^n = Q^0 \exp[\mathrm{i}k(m\Delta x - Cn\Delta t)]$. For real C, this is a wave-like solution. However, if C is complex, this solution will behave exponentially, quite unlike the solution of the continuous equation. Substituting Q_m^n into the finite-difference equation, we find that

$$C = \frac{1}{k\Delta t}\sin^{-1}\left[\left(\frac{c\Delta t}{\Delta x}\right)\sin k\Delta x\right].$$

If the argument of the inverse sine is less than unity, C is real. Otherwise, C is complex, and the solution will grow with time. Thus, the condition for stability of the solution is

$$\left|\frac{c\Delta t}{\Delta x}\right| \leqslant 1.$$

This is the Courant–Friedrichs–Lewy criterion, discovered in 1928. It imposes a strong constraint on the relative sizes of the space and time grids. The limitation on stability can be circumvented by means of implicit finite differencing. Then

Typesetter's note: author wondered whether the CFL criterion is being discussed elsewhere, and whether a cross-reference could therefore be added here.

$$C = \frac{2}{k\Delta t}\tan^{-1}\left[\left(\frac{c\Delta t}{2\Delta x}\right)\sin k\Delta x\right].$$

The numerical phase speed C is always real, so the implicit scheme is unconditionally stable, but the cost is that a linear system must be solved at each time step.

4.3 Spectral Method

In the spectral method, each field is expanded in a series of spherical harmonics:

$$Q(\lambda,\phi,t) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} Q_n^m(t) Y_n^m(\lambda,\phi),$$

where the coefficients $Q_n^m(t)$ depend only on time, and where $Y_n^m(\lambda,\phi)$ are the spherical harmonics $Y_n^m(\lambda,\phi) = \exp(\mathrm{i}m\lambda)P_n^m(\phi)$ for longitude λ and latitude ϕ. The coefficients Q_n^m of the harmonics provide an alternative to specifying the field values $Q(\lambda,\phi)$ in the spatial domain. When the model equations are transformed to spectral space they become a coupled set of equations (ordinary differential equations) for the spectral coefficients Q_n^m. These are used to advance the coefficients in time, after which the new physical fields may be computed.

In practice, the series expansion must be truncated at some point:

$$Q(\lambda_i,\phi_j,t) = \sum_{n=0}^{N}\sum_{m=-n}^{n} Q_n^m(t) Y_n^m(\lambda_i,\phi_j).$$

This is called *triangular truncation*, and the value of N indicates the resolution of the model. There is a computational grid, called the Gaussian grid, corresponding to the spectral truncation.

5 Initial Conditions

Numerical weather prediction is an initial-value problem: to integrate the equations of motion we must specify the values of the dependent variables at an initial time. The numerical process then generates the values of these variables at later times. The initial data are ultimately derived from direct observations of the atmosphere.

The *optimal interpolation* analysis method was, for several decades, the most popular method of automatic analysis for NWP. This method optimizes the combination of information in the background (forecast) field and in the observations, using the statistical properties of the forecast and observation errors to produce an analysis that, in a precise statistical sense, is the best possible analysis.

An alternative approach to data assimilation is to find the analysis field that minimizes a *cost function*. This is called *variational assimilation* and it is equivalent to the statistical technique known as the maximum-likelihood estimate, subject to the assumption of Gaussian errors. When applied at a specific time, the method is called three-dimensional variational assimilation, or 3D-Var for short. When the time dimension is also taken into account, we have 4D-Var.

5.1 Variational Assimilation

The cost function for 3D-Var may be defined as the sum of two components:

$$J = J_{\mathrm{B}} + J_{\mathrm{O}}.$$

We represent the model state by a high-dimensional vector $\boldsymbol{X}$. The term

$$J_{\mathrm{B}} = \tfrac{1}{2}(\boldsymbol{X} - \boldsymbol{X}_{\mathrm{B}})^{\mathrm{T}}\boldsymbol{B}^{-1}(\boldsymbol{X} - \boldsymbol{X}_{\mathrm{B}})$$

represents the distance between the model state $\boldsymbol{X}$ and the background field $\boldsymbol{X}_{\mathrm{B}}$ weighted by the background error covariance matrix $\boldsymbol{B}$. The term

$$J_{\mathrm{O}} = \tfrac{1}{2}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X})^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X})$$

represents the distance between the analysis and the observed values $\boldsymbol{Y}$ weighted by the observation error covariance matrix $\boldsymbol{R}$. The observation operator $\boldsymbol{H}$ is a rectangular matrix that converts the background field into first-guess values of the observations. More generally, the observation operator is nonlinear but, for ease of description, we assume here that it is linear.

The minimum of J is attained at $\boldsymbol{X} = \boldsymbol{X}_{\mathrm{A}}$, where

$$\nabla_{\boldsymbol{X}} J = 0.$$

that is, where the gradient of J with respect to each of the analyzed values is zero. Computing this gradient, we get

$$\nabla_{\boldsymbol{X}} J = \boldsymbol{B}^{-1}(\boldsymbol{X} - \boldsymbol{X}_{\mathrm{B}}) + \boldsymbol{H}^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X}).$$

Setting this to zero we can deduce the expression

$$\boldsymbol{X} = \boldsymbol{X}_{\mathrm{B}} + \boldsymbol{K}(\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X}_{\mathrm{B}}).$$

Thus, the analysis is obtained by adding to the background field a weighted sum of the difference between observed and background values. The matrix $\boldsymbol{K}$, the *gain matrix*, is given by

$$\boldsymbol{K} = \boldsymbol{B}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{R} + \boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^{\mathrm{T}})^{-1}.$$

The analysis error covariance is then given by

$$\boldsymbol{A} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{B}.$$

The minimum of the cost function is found using a descent algorithm such as the CONJUGATE GRADIENT METHOD [**??**]; 3D-Var solves the minimization problem directly, avoiding computation of the gain matrix.

The 3D-Var method has enabled the direct assimilation of satellite radiance measurements. The error-prone inversion process, whereby temperatures are deduced from the radiances before assimilation, is thus eliminated. Quality control of these data is also easier and more reliable. As a consequence, the accuracy of forecasts has improved markedly since the introduction of variational assimilation. The accuracy of medium-range forecasts is now about equal for the two hemispheres (see figure 1). This is due to better satellite data assimilation. Satellite data are essential for the Southern Hemisphere as conventional data are in such short supply. The effort to extract useful information from satellite soundings has been one of the great research triumphs of NWP over the past forty years.

5.2 Inclusion of the Time Dimension

Whereas conventional meteorological observations are made at the main synoptic hours, satellite data are distributed continuously in time. To assimilate these data, it is necessary to perform the analysis over a time interval rather than for a single moment. This is also more appropriate for observations that are distributed inhomogeneously in space. Four-dimensional variational assimilation, or 4D-Var for short, uses all the observations within an interval $t_0 \leqslant t \leqslant t_N$. The cost function has a term J_{B} measuring the distance to the background field $\boldsymbol{X}_{\mathrm{B}}$ at the initial time t_0, just as in 3D-Var. It also contains a summation of terms measuring the distance to observations at each time step t_n in the interval $[t_0, t_N]$:

$$J = J_{\mathrm{B}} + \sum_{n=0}^{N} J_{\mathrm{O}}(t_n),$$

where J_{B} is defined as for 3D-Var and $J_{\mathrm{O}}(t_n)$ is given by

$$J_{\mathrm{O}}(t_n) = (\boldsymbol{Y}_n - \boldsymbol{H}_n\boldsymbol{X}_n)^{\mathrm{T}}\boldsymbol{R}_n^{-1}(\boldsymbol{Y}_n - \boldsymbol{H}_n\boldsymbol{X}_n).$$

The state vector $\boldsymbol{X}_n$ at time t_n is generated by integration of the forecast model from time t_0 to t_n, written

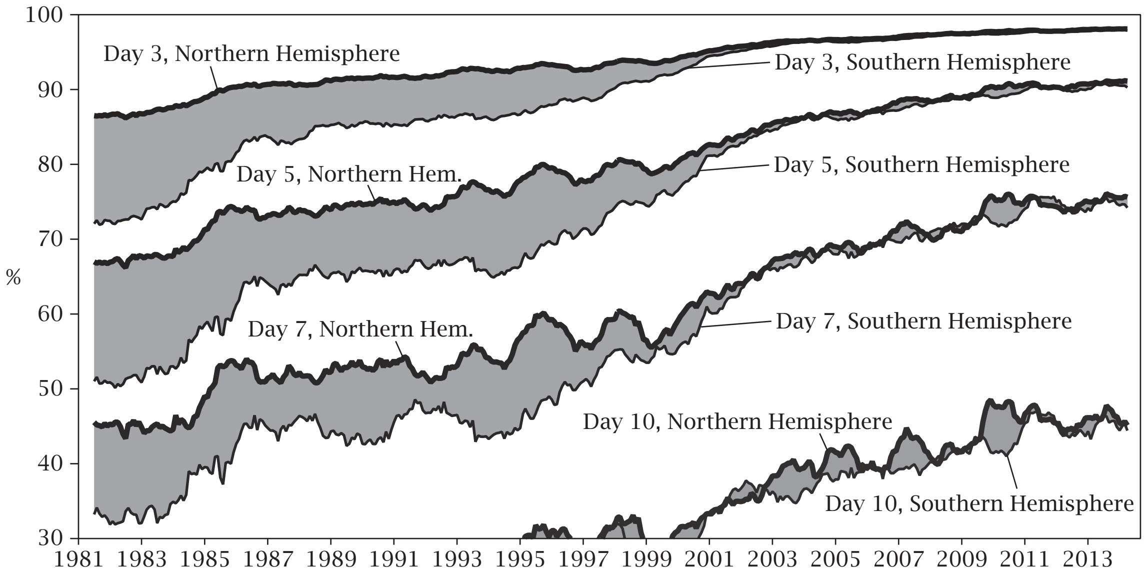

Figure 1 Anomaly correlation (%) of 500 hPa geopotential height: twelve-month running mean. For each pair of graphs, the upper one is for the Northern Hemisphere and the lower for the Southern Hemisphere (©ECMWF).

$\boldsymbol{X}_n = \mathcal{M}_n(\boldsymbol{X}_0)$. The vector $\boldsymbol{Y}_n$ contains the observations valid at time t_n.

Just as the observation operator had to be linearized to obtain a quadratic cost function, we linearize the model operator $\mathcal{M}_n$ about the trajectory from the background field, obtaining what is called the *tangent linear model* operator $\boldsymbol{M}_n$. Then we find that 4D-Var is formally similar to 3D-Var with the observation operator $\boldsymbol{H}$ replaced by $\boldsymbol{H}_n\boldsymbol{M}_n$. Just as the minimization of J in 3D-Var involved the transpose of $\boldsymbol{H}$, the minimization in 4D-Var involves the transpose of $\boldsymbol{H}_n\boldsymbol{M}_n$, which is $\boldsymbol{M}_n^{\mathrm{T}}\boldsymbol{H}_n^{\mathrm{T}}$. The operator $\boldsymbol{M}_n^{\mathrm{T}}$, the transpose of the tangent linear model, is called the *adjoint model*. The *control variable* for the minimization of the cost function is $\boldsymbol{X}_0$, the model state at time t_0, and the sequence of analyses $\boldsymbol{X}_n$ satisfies the model equations, that is, the model is used as a strong constraint.

The 4D-Var method finds initial conditions $\boldsymbol{X}_0$ such that the forecast best fits the observations within the assimilation interval. This removes an inherent disadvantage of optimal interpolation and 3D-Var, where all observations within a fixed time window (typically of six hours) are assumed to be valid at the analysis time. The introduction of 4D-Var at the European Centre for Medium-Range Weather Forecasts (ECMWF) led to a significant improvement in the quality of operational medium-range forecasts.

6 Forecasting Models

Operational forecasting today is based on output from a suite of computer models. Global models are used for predictions of several days ahead, while shorter-range forecasts are based on regional or limited-area models.

6.1 The ECMWF Global Model

As an example of a global model we consider the integrated forecast system (IFS) of the ECMWF (which is based in Reading, in the United Kingdom). The ECMWF produces a wide range of global atmospheric and marine forecasts and disseminates them on a regular schedule to its thirty-four member and cooperating states. The primary products are deterministic forecasts for the atmosphere out to ten days ahead, based on a high-resolution model, and probabilistic forecasts, extending to a month, made using a reduced resolution and an ensemble of fifty-one model runs.

The basis of the NWP operations at the ECMWF is the IFS. It uses a *spectral representation* of the meteorological fields. The IFS system underwent major resolution upgrades in 2006 and in 2010. Table 1 compares the spatial resolutions of the three model cycles, indicating the substantial improvements in model resolution in recent years. The truncation of the deterministic model is now T1279, that is, the spectral expansion is terminated at total wave number 1279. This is equivalent to a

Table 1 Upgrades to the ECMWF IFS in 2006 and 2010. The spectral resolution is indicated by the triangular truncation number, and the effective resolution of the associated Gaussian grid is indicated. The number of model levels, or layers used to represent the vertical structure of the atmosphere, is also given.

	Before 2006	2006–9	After 2009
Spectral truncation	T511	T799	T1279
Effective resolution	39 km	25 km	16 km
Model levels	60	91	137

spatial resolution of 16 km. The number of model levels in the vertical has recently been increased to 137. The new Gaussian grid for the IFS has about 2×10^6 points. With 137 levels and five primary prognostic variables at each point, about 1.2×10^9 numbers are required to specify the atmospheric state at a given time. That is, the model has about a billion degrees of freedom. The computational task of making forecasts with such high resolution is truly formidable. The ECMWF carries out its operational program using a powerful and complex computer system. At the heart of this system is a Cray XC30 high-performance computer, comprising some 160 000 processors, with a sustained performance of over 200 teraflops (2×10^{14} floating-point operations per second).

6.2 Mesoscale Modeling

Short-range forecasting requires detailed guidance that is updated frequently. Many national meteorological services run limited-area models with high resolution to provide such forecast guidance. These models permit a free choice of geographical area and spatial resolution, and forecasts can be run as frequently as required. Limited-area models make available a comprehensive range of outputs, with a high time resolution. Nested grids with successively higher resolution can be used to provide greater local detail.

The Weather Research and Forecasting Model is a next-generation mesoscale NWP system developed in a partnership involving American national agencies (the National Centers for Environmental Prediction and the National Center for Atmospheric Research) and universities. It is designed to serve both operational-forecasting and atmospheric-research needs. The Weather Research and Forecasting Model is suitable for a broad range of applications, from meters to thousands of kilometers, and it is currently in operational use at several national meteorological services.[1]

6.3 Ensemble Prediction

The chaotic nature of atmospheric flow is now well-understood. It imposes a limit on predictability, as unavoidable errors in the initial state grow rapidly and render the forecast useless after some days. As a result of our increased understanding of the inherent difficulties of making precise predictions, there has been a paradigm shift in recent years from deterministic to probabilistic prediction. A forecast is now considered incomplete without an accompanying error bar, or quantitative indication of confidence.

The most successful way of producing a probabilistic prediction is to run a series, or ensemble, of forecasts, each starting from a slightly different initial state, and each randomly perturbed during the forecast to simulate model errors. The ensemble of forecasts is used to deduce probabilistic information about future changes in the atmosphere. Since the early 1990s this systematic method of providing an a priori measure of forecast accuracy has been operational at both the ECMWF and at the National Centers for Environmental Prediction in Washington. In the ECMWF's ensemble prediction system, an ensemble of fifty-one forecasts is performed, each having a resolution half that of the deterministic forecast. Probability forecasts for a wide range of weather events are generated and disseminated for use in the operational centers, and they have become the key tools for medium-range prediction.

7 Verification of ECMWF Forecasts

Forecast accuracy has improved dramatically in recent decades. This can be measured by the *anomaly correlation*. The anomaly is the difference between a forecast value and the corresponding climate value, and the agreement between the forecast anomaly and the observed anomaly is expressed as the anomaly correlation. The higher this score the better; by general agreement, values in excess of 60% imply skill in the forecast. In figure 1, the twelve-month running mean anomaly correlations (in percentages) of the three-, five-, seven-, and ten-day 500 hPa height forecasts are shown for the extratropical Northern Hemisphere and Southern Hemisphere. The lines above each shaded region are for the Northern Hemisphere and the lines below are

1. Full details of the system are available at www.wrf-model.org.

for the Southern Hemisphere, with the shading showing the difference in scores between the two.

The plots in figure 1 show a continuing improvement in forecast accuracy, especially for the Southern Hemisphere. By the turn of the millennium, the accuracy was comparable for the two hemispheres. Predictive ability has improved steadily over the past thirty years, and there is now accuracy out to eight days ahead. This record is confirmed by a wealth of other data. Predictive skill has been increasing by about one day per decade, and there are reasons to hope that this trend will continue for several more decades.

8 Applications of NWP

NWP models are used for a wide range of applications. Perhaps the most important purpose is to provide timely warnings about weather extremes. Great financial losses can be caused by gales, floods, and other anomalous weather events. The warnings that result from NWP guidance can greatly diminish losses of both life and property. Transportation, energy consumption, construction, tourism, and agriculture are all sensitive to weather conditions. There are expectations from all these sectors of increasing accuracy and detail in short-range forecasts, as decisions with heavy financial implications must continually be made.

NWP models are used to generate special guidance for the marine community. Predicted winds are used to drive wave models, which predict sea and swell heights and periods. Prediction of road ice is performed by specially designed models that use forecasts of temperature, humidity, precipitation, cloudiness, and other parameters to estimate the conditions on the road surface. Trajectories are easily derived from limited-area models. These are vital for modeling pollution drift, for nuclear fallout, smoke from forest fires, and so on. Aviation benefits significantly from NWP guidance, which provides warnings of hazards such as lightning, icing, and clear-air turbulence.

9 The Future

Progress in NWP over the past sixty years can be accurately described as revolutionary. Weather forecasts are now consistently accurate and readily available. Nevertheless, some formidable challenges remain. Sudden weather changes and extremes cause much human hardship and damage to property. These rapid developments often involve intricate interactions between dynamical and physical processes, both of which vary on a range of timescales. The effective computational coupling between the dynamical processes and the physical parametrizations is a significant challenge. *Nowcasting* is the process of predicting changes over periods of a few hours. Guidance provided by current numerical models occasionally falls short of what is required to take effective action and avert disasters. Greatest value is obtained by a systematic combination of NWP products with conventional observations, radar imagery, satellite imagery, and other data. But much remains to be done to develop optimal nowcasting systems, and we may be optimistic that future developments will lead to great improvements in this area.

At the opposite end of the timescale, the chaotic nature of the atmosphere limits the validity of deterministic forecasts. Interaction between the atmosphere and the ocean becomes a dominant factor at longer forecast ranges, as does coupling to sea ice. Also, a more accurate description of aerosols and trace gases should improve long-range forecasts. Although good progress in seasonal forecasting for the tropics has been made, the production of useful long-range forecasts for temperate regions remains to be tackled by future modelers. Another great challenge is the modeling and prediction of climate change, a matter of increasing importance and concern.

Typesetter's note: cross-reference to Golden article here Nick?

Further Reading

Lynch, P. 2006. *The Emergence of Numerical Weather Prediction: Richardson's Dream.* Cambridge: Cambridge University Press.