Advances and prospects in numerical weather prediction

By L. BENGTSSON*

European Centre for Medium Range Weather Forecasts, Reading

(= Presidents' Day Lecture delivered to IAMAP, July 1989)

1. INTRODUCTION

The complexity and variability of atmospheric motion and its manifestation in all kinds of weather phenomena present a continuous challenge. The weather affects human affairs and influences most aspects of society, both directly and indirectly. Prediction of the weather requires a fundamental understanding of the forces which govern the atmosphere, and of the many complex feedback processes involving the oceans and land surfaces.

Weather prediction is a well defined deterministic problem. Starting from a given initial state, any future state can be obtained by integrating the classical Navier–Stokes equations forward in time. Therefore a weather forecast can, in principle, be calculated in the same way as the motion of the planets or the trajectory of a missile.

However, a more thorough analysis reveals an enormous complexity. Atmospheric processes span a large range of scales. The largest scale encompasses the whole atmosphere, while the smallest components, such as cloud droplets, have sizes of a few micrometres. The sizes of the cloud droplets and their distribution are of importance for the calculation of radiative processes, and are crucial components in the mechanism for release of precipitation. This range of atmospheric motion covers 13 orders of magnitude ($10^{-6}$ to $10^7$ m). There are thunderstorms with a horizontal extension of a few kilometres, travelling frontal depressions in middle latitudes stretching over 1000–2000 km and, finally, the largest circulation systems of global scale forced by the large mountain ranges and by the distribution of oceans and continents. Every weather system has its characteristic time-scale, which generally increases with the spatial scale. This time-scale is about an hour for a thunderstorm and up to a few months for the largest circulation system such as the monsoon.

There is continuous interaction between all these scales of motion. In the tropical belt the trade winds provide heat and moisture to feed the convective system in the equatorial zone. The latent heat released by all the many convective systems in this zone in turn maintains the trade winds. In a similar way the microscale processes in the form of heat and moisture fluxes over land and sea are the driving forces for the large-scale circulation systems of the earth.

This non-linear interaction between different scales of motion is the fundamental
reason for the difficulty in predicting atmospheric flow. Inevitable errors in observing the smallest scales of motion sooner or later contaminate the larger scales and finally destroy the accuracy of any prediction. Weather prediction can therefore be seen as an unstable problem in that small initial differences can have large final effects. Although the prediction problem as such is deterministic, it is for practical reasons non-deterministic, since the initial stage can never be perfectly known. Theoretically, weather prediction has much in common with more general non-deterministic problems involving economic and social systems. Atmospheric prediction models may therefore serve as useful prototypes for better understanding of a more general class of problems where the dynamical laws are not yet well understood.

Numerical prediction using mathematical models of the atmosphere has been in existence for almost 40 years. During this period a spectacular development has taken place, encompassing the complete field of numerical weather prediction. A global observing system has been established and operational models provide daily routine predictions for the whole global atmosphere from the surface of the earth and high into the stratosphere. The effect on weather prediction has been dramatic; today 5-day forecasts for the 500 hPa geopotential are as accurate as the 1-day forecasts produced in the early 1950s (Bengtsson 1990). By far the greatest advances in forecast skill have taken place in the extra-tropics.

This development has been possible through advances in the understanding of atmospheric dynamics and physics and the very rapid progress in computer technology (Fig. 1) which, together with more economical integration methods, has made it possible to significantly increase horizontal and vertical resolution and hence make more accurate calculations of dynamical and physical processes.

Another necessary condition for the large improvement in numerical prediction has been the establishment of global observing systems, so successfully tested during the Global Weather Experiment, and the development of data-assimilation methods for the assimilation of non-synoptic observations, such as data from satellites, aircraft and drifting buoys.

---

**Figure 1.** History of supercomputer performance. Performances are given in MIPS (millions of instructions per second) to 1975, thereafter in MFLOPS (millions of floating point operations per second).
As will be demonstrated in this article, the improvements in forecast skill manifest themselves as more accurate short-range forecasts, an extension in time of useful forecasts, and a significant increase in the number of predicted quantities.

In section 2 we will discuss some of the more important advances which have taken place in numerical modelling, such as efficient time integration schemes and the use of the spectral transform technique. The incorporation of orographic effects and the parametrization of the hydrological cycles in numerical weather prediction will be described. The emphasis will be put on progress which has taken place over the last 10–15 years. Section 3 will be devoted to the observational problems including initialization of numerical models and data-assimilation. Sections 4 and 5 will provide examples of forecast skill. Of particular importance is the improvement in predicting intense cyclogenesis on time-scales up to five days ahead. Although model errors have been very much reduced, numerical models still have characteristic systematic deficiencies. These problems are discussed in section 6. Studies of atmospheric predictability clearly indicate that further improvements in numerical weather prediction are possible even within the constraints of the present observing system. Different ideas by which further improvements can be achieved will be discussed in section 7.

The presentation in the review will concentrate on the prediction of the large-scale synoptic flow by hemispheric or global models. An extensive review of progress in mesoscale modelling and prediction has recently been carried out by Anthes (1990).

2. NUMerical MODELLing

One of the principal difficulties in atmospheric modelling is the very general character of the equations of motion. Present models used for numerical weather prediction make use of the hydrostatic approximation, thus restricting the description of the flow to Rossby waves and gravity waves. This set of equations, known as the primitive equations, was first explored by Richardson (1922).

As discussed by Phillips (1973), the governing dynamic, thermodynamic and conservation equations for mass and moisture are transformed to a spherical geometry. The reduced set of primitive equations is obtained by assuming the height scale of the motion to be small compared with its horizontal length scale. The basic predicted variables in such a set are the horizontal wind components, $u$ and $v$, the temperature, $T$, water vapour, normally represented by the specific humidity $q$, and the surface pressure, $p_s$.

We may illustrate the form of the primitive equations using the vertical coordinate system in most common use, namely the ‘sigma’ coordinate system proposed by Phillips (1957). In the sigma system the vertical coordinate is given by

$$ a = \frac{p}{p_s} $$

where $p$ is the pressure and $p_s$ the surface pressure. In this case the governing equations become:

**momentum**

$$ \frac{DV}{Dt} + f k x \nabla + \nabla \varphi + RT \nabla (\ln p_s) = F_m $$

(2.2)

**thermodynamic**

$$ \frac{DT}{Dt} - \frac{RT \omega}{c_p p_s} = F_T $$

(2.3)
mass conservation

\[ \frac{Dp_s}{Dt} + p_s \left( \nabla \cdot \mathbf{v} + \frac{\partial \mathbf{v}}{\partial \sigma} \right) = 0 \]  
(2.4)

moisture conservation

\[ \frac{Dq}{Dt} = F_q \]  
(2.5)

hydrostatic

\[ \frac{\partial q}{\partial \sigma} = -\frac{RT}{\sigma} \]  
(2.6)

Here \( t \) is time and \( D/Dt \) denotes the rate of change for a particle moving with the fluid. In \( \sigma \) coordinates this derivative takes the form

\[ \frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla + \mathbf{v} \cdot \frac{\partial}{\partial \sigma}, \]  
(2.7)

Here \( \mathbf{v} \) is the horizontal velocity vector, \( \mathbf{V} = (u, v, 0) \) and \( \nabla \) the two-dimensional gradient operator on a constant \( \sigma \)-surface, \( f \) is the Coriolis parameter, \( k \) the unit vertical vector, \( \varphi \) the geopotential, \( R \) the gas constant and \( c_p \) the specific heat of air at constant pressure. \( F_x \) denotes the rate of change of variable \( x \) due to the parametrized processes of radiation, convection, turbulent vertical mixing and large-scale precipitation. It also represents the rate of change of \( x \) due to the explicit horizontal smoothing that is usually included in models in order to prevent an unrealistic growth of the smallest scales which can be represented by the model.

A predictive equation for surface pressure is obtained by integrating Eq. (2.4) from \( \sigma = 0 \) to \( \sigma = 1 \), using the boundary conditions \( \dot{\sigma} = 0 \) at \( \sigma = 0 \) and \( \sigma = 1 \):

\[ \frac{\partial p_s}{\partial t} = -\int_0^1 \nabla \cdot (p_s \mathbf{V}) \, d\sigma. \]  
(2.8)

Vertical velocities are not explicitly predicted but they can easily be obtained from Eq. (2.2).

The form of the primitive equations given above neglects the local mass of water vapour compared with that of dry air, an approximation that can introduce a small, but not always negligible, error in moist and warm conditions. A more accurate form can be obtained by replacing the temperature, \( T \), appearing in equation (2.2) and (2.6), by the virtual temperature, \( T_v \), which is defined by:

\[ T_v = T \left[ 1 + (R_v/R - 1)q \right] \]  
(2.9)

where \( R_v \) is the gas constant for water vapour. Equation (2.3) must be corrected in a corresponding way.

\( (a) \) Numerical formulation and resolution

Three different techniques, finite difference, finite element and spectral transform, are being used for the horizontal discretization of global and hemispheric numerical models. The numerical methods in the late 1950s and 1960s for simplified models (Arakawa 1966; Lilly 1965) led to the design of stable finite difference schemes for multi-level primitive equation models. At the beginning of the 1970s almost all large-scale
models were founded on these techniques. Since then, the performance of finite difference schemes has been greatly improved by a clever staggering of dependent variables and the use of approximations with higher-order accuracy (Arakawa 1972; Arakawa and Lamb 1977). The extensions of finite difference techniques to global modelling have required the development of satisfactory treatments of polar singularities in spherical coordinates and methods of east–west filtering at high latitudes (e.g. Burridge and Haseler 1977).

Perhaps the most important contribution has been the development of spectral methods for the horizontal representation. This was made possible by the use of a spectral transform technique first proposed by Eliasen et al. (1970) and Orzag (1970). Hereby the linear terms in the forecasting equation are calculated in the spectral form, while non-linear terms and the physical forcing terms are calculated in the grid-point domain. Transform between grid-point representation and spectral representation takes place every time-step.

A particular advantage with the spectral model is the ease with which a semi-implicit time-stepping algorithm can be implemented (see below). As a result during the 1980s most operational forecasting centres carrying out global or hemispheric forecasting have switched from grid-point to spectral models. These models largely follow the approach of the first multi-level spectral model of Bourke (1972), Hoskins and Simmons (1975) and Baede et al. (1979) and use vorticity and divergence as predictive variables.

These equations are easily derived from Eq. (2.2). The predicted variables are represented in terms of truncated expansions of spherical harmonics

\[ X(\lambda, \theta, \sigma, t) = \sum_{m=-M}^{M} \sum_{n=-|m|}^{N} X^m_n(\sigma, t) p^n_m(\sin \theta) e^{im\lambda} \]  

(2.10)

where \( X \) is any variable, \( \theta \) is latitude, \( \lambda \) is longitude and the \( p^n_m \) are the associated Legendre polynomials.

For equivalent resolution (e.g. the finite difference grid has the same spacing as the Gaussian grid used by the spectral model) and for the same physical parametrization and the same initial data there are generally small differences between a finite difference model and a spectral transform model.

Girard and Jarraud (1982) undertook an extensive intercomparison between the ECMWF grid-point model (horizontal resolution of 1.875° latitude and longitude) and the ECMWF spectral transform model with triangular truncation at T63. The spectral model gave slightly better results for the same computational cost. However, it should be pointed out that differences in mathematical representation (spectral or grid-point for instance) generally have less impact on predictive skill than most other aspects of the forecasting system such as resolution, sub-grid-scale parametrization and the initial state.

Of particular importance has been the development of efficient time-integration schemes. The use of explicit leapfrog time-steps necessitates the use of a time-step which is inversely proportional to the horizontal grid size and to the maximum phase speed. Since the primitive equations allow the existence of gravity waves, which may move 5–6 times faster than the fastest Rossby waves, the time-step must be accordingly reduced in order to avoid computational instability (Courant–Friedrichs–Lewy instability). Robert (1969) and Robert et al. (1972) proposed an efficient so-called semi-implicit scheme, whereby the terms in the equations controlling the Rossby waves are treated explicitly, while the terms controlling gravity waves are solved by an implicit scheme. Since the implicit scheme is computationally stable, this technique makes it possible to use a 5–6 times larger time-step or the same time-step as for the filtered equations. An alternative way to speed up the numerical solution of the primitive equations was proposed by
Marchuk (1974). In this scheme, which uses only two time-levels, the calculations are split into two parts, whereby the advective terms are calculated using a longer time-step and the adjustment term, controlling the gravity waves, with a shorter time-step. The split-explicit scheme has been successfully implemented for operational numerical prediction at the UK Meteorological Office (Gadd 1978).

Table 1 outlines in a qualitative sense the importance of different atmospheric processes with respect to the prediction of the synoptic flow in the extra-tropics. In the short time-range (1–2 days) a detailed description of the adiabatic processes is of utmost importance. This includes the rôle of orographic obstacles which may influence the flow via the lower boundary condition. Jet streams and baroclinic zones have sharp structures (see e.g. Defant 1959), and resolutions of 50–100 km in the horizontal and 25–50 hPa in the vertical are required to keep truncation errors comfortably small. Simmons et al. (1989) have demonstrated the importance of a high horizontal resolution convincingly. Further remarks will be made in section 6.

Perhaps the most important effect of an increased horizontal resolution is the benefit it brings in resolving land–sea contrasts and steep orography. Dell’Osso (1984) and Wu and Chen (1985) have demonstrated the effect of resolution in predicting lee-wave cyclogenesis. For such developments, proper resolution of orographic obstacles is essential; hardly any positive effect of increased atmospheric resolution is obtained if there is no simultaneous improvement in resolving the orography.

(b) Parametrization

Observational and computational restrictions make it necessary to confine the numerical models to a description of phenomena larger than a certain scale. Although this scale has gradually become smaller following computer developments, the minimum scale is still significantly larger than that of many weather systems. These so-called sub-grid-scale processes must be related to the macroscale currently resolved by the models. This step is known as parametrization, since the sub-grid-scale processes are described in terms of the parameters resolved and predicted by the model.

We will not give a comprehensive presentation of the parametrization schemes presently used in numerical models. These schemes are being successively modified. The reader is advised to study the latest model documentation from the large forecasting centres. Instead, we will focus on a few important areas which have been found to be of major importance in weather prediction in the extra-tropics in the short and medium range.
(i) Cumulus convection. A proper parametrization of deep cumulus convection is of primary importance for numerical weather prediction in the tropics and plays an important rôle at middle and high latitudes in the medium and extended range (e.g. Simmons and Miller 1988). Several schemes have been developed that relate the occurrence of convection to the existence of an unstable stratification and a net conveyance of moisture by the large-scale flow and surface friction. For an extensive review see Arakawa and Chen (1987). One of the most common schemes is the one proposed by Kuo (1974) and now used by several models.

In the scheme proposed by Kuo a partition parameter, $b$, was introduced in such a way that moistening of the environmental air was proportional to $b$ while heating due to release of latent heat was proportional to $1 - b$. However, there is no universally accepted method at the present time of specifying $b$. Anthes (1977) has proposed an empirical formulation for $b$ depending on the relative humidity. Krishnamurti et al. (1980) have given an expression for $b$ depending on the large-scale vertical velocity and vorticity, derived from multiple regression analysis on the GATE upper-air dataset. Geleyn (1985) has proposed a modified version where the partition is determined locally through the use of the wet-bulb characteristics.

In recent years new conceptual ideas have been proposed. Betts (1986) and Betts and Miller (1986) have developed a convective adjustment scheme where adjustment is also taking place towards a prescribed moisture profile typical of convective situations. Recent evaluation carried out at ECMWF (Simmons and Miller 1988) indicates better performance with this scheme than with the Kuo scheme presently used. Edelman (1984) and Bougeault (1985) have proposed simplified mass flux schemes with a parametrization of the detrainment and a closure assumption similar to those of the Kuo scheme. The schemes have been implemented in the German and French operational models respectively (Müller et al. 1987; Coiffier et al. 1987). A further development of the mass flux scheme has recently been proposed by Tiedtke (1989) (Fig. 2). It has recently been implemented in the ECMWF operational model.

(ii) Stratiform precipitation. The treatment of non-convective precipitation appears to be one of the simpler elements of the overall parametrization scheme. Stratiform precipi-

![Figure 2. Schematic illustration of a mass flux scheme for the parametrization of convection (Tiedtke 1989).](image-url)
tation is normally calculated after the computation of other dynamical and physical processes that change temperature and water vapour content. It generally allows for condensation, with associated latent heat release, of sufficient vapour to keep the relative humidity below a fixed threshold value. The threshold value is normally set to a value in the range of 80–100% relative humidity. Partial evaporation of raindrops falling through unsaturated layers is normally incorporated following the proposal by Kessler (1969).

(iii) Dissipation of momentum. The term $F_m$ representing the dissipation of momentum in the momentum equation (2.2) is a comparatively small term outside the boundary layer. Very little observational evidence exists of its distribution in time and space. So far, the main evidence for the existence of a frictional force comes from the residual term in the longitudinally averaged budget of zonal momentum (Swinbank 1985). Studies of the local momentum budget (e.g. Holopainen et al. 1980) are inconclusive owing to lack of data on vertical velocity.

In numerical models we must distinguish between dissipation which is introduced for numerical reasons, and dissipation which is due to real physical sub-grid-scale processes. In practice there is no clear way to separate these. The dissipation terms are determined by systematic numerical experiments with selected synoptic cases (e.g. Machenhauer 1988). It is occasionally found that the forecasts are very sensitive to the diffusion factor (Simmons and Miller 1988).

Recent evaluation at ECMWF, for example, of the role of vertical diffusion has indicated that previously used vertical dissipation has been too large and led to an excess reduction of eddy kinetic energy. From January 1988, vertical dissipation was removed in the free atmosphere except in areas of dry static instability. Although there are no indications that the vertical dissipation is zero in the real atmosphere, there are observational indications that dissipation takes place in narrow bands which during the course of integration are smoothed by the vertical finite difference scheme. The effect of reduced vertical diffusion is an improved maintenance of eddy kinetic energy and a better prediction of intense synoptic features (section 5).

A common systematic error in numerical integration is a marked tendency to intensify the westerlies over the continents in middle and high latitudes. The error is most pronounced during the winter. It can be identified clearly in short-range forecasts, but becomes serious beyond 4–5 days. The erroneous flow regime is accompanied by unrealistically low values of surface pressure, geopotential height and temperature in high latitudes. Two traditionally neglected processes have recently been investigated as being of importance in alleviating this problem in large-scale numerical models. Both are associated with the occurrence of stable stratified flow over unresolved orography—a common situation in the extra-tropics in winter.

The first of these is the damming and trapping of cold air by ridges and valleys in mountainous regions. Wallace et al. (1983) suggested a procedure whereby the unresolved orography was accounted for by adding an amount proportional to the standard deviation of the unresolved sub-grid-scale orography—a procedure which is equivalent to partially filling unresolved valleys. This so-called ‘envelope orography’ was found to improve the skill of the ECMWF model by reducing the magnitude of the westerly error. Later studies by other modellers, e.g. Chouinard et al. (1986), have verified the results obtained at ECMWF. Figure 3 shows the effect of envelope orography on a particular forecast.

The second process, which may be of importance in alleviating the westerly flow error, is associated with the occurrence of mesoscale orographically excited gravity-wave systems in stable stratified flow. The idea that such drag forces due to mountain waves
Figure 3. Analysed 500 hPa height field, averaged for the period 20–25 March 1984 (upper left) and corresponding day 5–10 average T106 forecasts from 15 March using the following orographies: Upper right: Mean orography, Mid left: Envelope orography, Mid right: Envelope, but mean over Rockies, Lower left: Envelope, but mean over Greenland and northern Canada, Lower right: Envelope, but mean over Alps and Pyrenees, (Jarraud et al. 1987).
may have an important impact on the synoptic and large-scale flow was proposed by Sawyer (1959) and Bretherton (1969). Observational studies by Lilly (1972) and Lilly et al. (1982) reinforced the earlier studies.

The first use of a wave drag parametrization was in the Canadian Atmospheric Environment Service's general circulation model (Boer et al. 1984). It has recently been incorporated in the UK Meteorological Office operational model (Palmer et al. 1986) and in the ECMWF model (Miller et al. 1989). In the model implementation, the drag is proportional to the surface wind speed, the variance of the sub-grid-scale orography and the Brunt–Väisälä frequency. In the ECMWF implementation, the variance of the sub-grid-scale orography is also dependent on the wind direction relative to the orographic obstacle. The principle of gravity-wave drag is illustrated in Fig. 4. Numerical experiments have shown that a combination of envelope orography and gravity-wave drag is the best solution (McFarlane et al. 1987). Figure 5 shows a similar result from ECMWF.

\[
\tau_{gw} \propto N_L |V_{lw}||H|^2
\]

\[|H|^2 = \text{Variance of sub-gridscale orography (directionally dependent)}\]

Figure 4. Schematic illustration of gravity-wave drag as presently implemented in the ECMWF model. Middle figure shows the potential temperature from a high resolution simulation of a Boulder wind storm (Pelletier and Clark 1979).

3. Observations and Their Assimilation

Numerical weather prediction is an initial value problem and the accuracy of any forecast is directly related to the accuracy of the initial state. The radiosonde network established after the Second World War is still the cornerstone of the atmospheric observing system, but it has been greatly enhanced with satellite and aircraft measurements. Since the end of the 1970s there has in essence been a global observing system (Fig. 6). However, in spite of the impressive development of the observing system over the last 15 years, the present observing system cannot satisfy the observational requirements that ideally would be needed at every grid point of the model, consisting of the basic parameters: horizontal wind, temperature, moisture and surface pressure. Additionally, sea surface temperature, soil moisture, and snow observations are required.
Figure 5. Anomaly correlation of 500 hPa height forecasts for the extra-tropical northern hemisphere averaged over 11 forecasts with the ECMWF T106-L19 model. Full line shows the score using a mean orography, dot-dashed line with an envelope orography and dashed line with both envelope orography and gravity-wave drag. For further information, see text.

Fortunately, the very demanding requirement can be relaxed, owing to the strong dynamical and physical coupling between meteorological observations in space and time. This was pointed out by Smagorinsky et al. (1970) who demonstrated that a dynamical model, for example, was able to reconstitute the humidity field in most of its details within a day or two in an experiment in which humidity initially was prescribed as a zonal mean. It should be stressed, however, that this essentially is valid only in the extratropics when the dynamical forcing is dominating. In the tropics, humidity observations often play a crucial rôle and, as has been demonstrated by Krishnamurti et al. (1983), are required for an accurate prediction of tropical disturbances.

Numerical experiments (e.g. Arpe et al. 1985; Hollingsworth et al. 1985) have shown that an accurate specification of the coupled wind and mass fields in the baroclinic zones is crucial for forecast quality, while a detailed analysis of the boundary layer appears to be less important. The explanation is again that a numerical model is able to generate missing information, indicating a considerable degree of redundancy among the large-scale atmospheric variables. There is also increasing evidence that a realistic high-resolution dynamical model can even simulate the formation and growth of small-scale weather systems (smaller than can be resolved by available observations) even if only the relatively large-scale circulation features are explicitly described by global observations (Anthes 1972; Dell’Osso and Bengtsson 1985; Tuleya 1987). In this respect, a realistic global model can be viewed as a unique and independent observing system that can generate information at a scale finer than that of the conventional observing system.

This is possible only because the models are able to predict the short-term evolution of atmospheric circulation reasonably accurately.

We will illustrate this by the following example taken from the operational ECMWF prediction system. Figure 7 shows the surface observations at 1200 UTC on 29 August 1985 over north Africa. Owing to a technical problem at the regional telecommunication centre in Paris during the period 25 to 30 August 1985, data were lost for the major part of north Africa. Figure 8 shows the wind field at 10 m above ground analysed at ECMWF
Figure 6. Coverage of observational data received at ECMWF between 0900 UTC and 1500 UTC on 19 September 1986. The data distribution is typical for the period 1986-88. SYNOP: Observations from surface stations including observations from ships of opportunity, TEMP: wind, temperature and moisture observations from radiosondes, PILOT: wind observations by pilot balloons, TOVS: temperature and moisture soundings by polar-orbiting satellites (NOAA), SATOB: wind measurements from geostationary satellites, AI REP: wind measurements from commercial airlines, SEA: pressure observations from drifting buoys.
Figure 7. Surface observations for 1200 UTC on 29 August 1985.

Figure 8. ECMWF analysis (10 m winds) for 1200 UTC on 29 August 1985.
(the lowest level of the ECMWF model). As can be seen, an intense surface vortex has been analysed with the centre near 20°N, 5°W. Since there are no available data whatsoever to analyse this feature at 1200 UTC on 29 August (satellite temperature and wind data are not used over land) the vortex must have been present in the first-guess field obtained by integrating the model. That this indeed was the case can be seen from Fig. 9, which shows the 24 h forecast from 1200 UTC on 28 August 1985. A careful inspection of preceding forecasts shows that the vortex was generated near the Hoggar Mountains, farther to the east, a few days earlier. The reality of the strong vortex is supported by the dust signature seen in the Meteosat infra-red image for 1145 UTC on 29 August (Fig. 10). This example lends further support to the expectation that a good dynamical model is apparently capable of generating synoptic-scale disturbances that form and amplify due to the dynamical instabilities of the large-scale flow, if the latter can be described adequately by the observational network. For a more comprehensive discussion on this and other similar cases see Reed et al. (1986).

Consequently the assimilation of observations in four dimensions using a comprehensive numerical model as a tool is a necessary requirement for an accurate determination of the initial state. The development of such systems during the past 10 years has played a very important part in the improvement which has taken place in numerical weather prediction. Moreover, synoptic data, such as observations from satellites and aircraft, cannot be adequately used without four-dimensional data-assimilation.

Most of the operational data-assimilation systems developed so far are not fully four-dimensional but instead arrange data in time slots ranging from a few hours up to 12 hours. There are several advantages in such a procedure, since it creates a possibility for a rational monitoring and checking of the data. As an example, we will describe the

![Figure 9. Twenty-four-hour forecast of 10 m winds from initial condition of 1200 UTC on 28 August 1985.](image)
data-assimilation system used at ECMWF. This system, which is outlined in Fig. 11, organizes the data in 6 h time windows. The global atmosphere is analysed at 17 levels, 1000–10 hPa, four times a day.

A data-assimilation step consists of a first guess (a 6 h forecast from a previous state), an analysis procedure using all available observations within ±3 h and finally an initialization step. After the initialization, integration continues for another 6 h, whereafter the process is repeated and so on. A numerical forecast for any length of time can be started from any of these initial states.

Objective analysis methods for numerical weather prediction have recently been thoroughly reviewed by Hollingsworth (1987). Most methods are based on an extension of optimum interpolation (Eliassen 1954; Gandin 1963) to a multivariate three-dimensional interpolation of deviation from forecast fields (e.g. Lorenc 1981). The multivariate technique allows for a consistent use of observations with different error characteristics and takes into full account the irregular distribution of these observations. Multivariate methods are consequently convenient for analysing the mixture of different observing systems that we have today.

As has been demonstrated by Leith (1983) and Bengtsson (1989) the accuracy of the initial state as obtained through a data-assimilation process is determined by: (i) the observational error variance at a given time, (ii) the data-assimilation frequency and (iii) the accuracy of the model. It can be demonstrated (Bengtsson 1989) that an extension of predictive skill of at least one day by the ECMWF model is due to the use of an accurate high-resolution model in the data-assimilation process. The smaller the forecast
errors are, the more linear will be the multivariate relation between the mass field and the wind components, the smaller will be the changes made by the analysis, and the smaller will be the changes made by the initialization. It is essential to use the same model for data-assimilation as for the forecast production, in order to minimize the shocks to the model. The reasons are evident; the background field for the analysis is as accurate as possible, the effects of the forward interpolation are minimized, and the adjustment between physics and dynamics at the start of the forecast, the ‘spin-up’, is reduced or eliminated entirely.

If analysed data are used directly as the initial condition for a forecast, imbalance between the mass and wind field will cause the forecast to be contaminated by spurious, high-frequency gravity-wave oscillations of much larger amplitudes than are observed in the real atmosphere (Fig. 12).

Although these oscillations tend to die away slowly because of various dissipation mechanisms in the model, they may be quite detrimental to the data-assimilation cycle in that the 6 h forecast is used as a first-guess field for the next analysis.

The synoptic changes over the 6 h period may thus be swamped by spurious changes due to these oscillations, with the consequence that at the next analysis time good data may be rejected as being too different from the first-guess field. Initialization may have little direct effect on the forecast skill but the indirect rôle of initialization is of the greatest importance.

Of particular importance in the development of efficient initialization procedures has been the idea to utilize the normal modes for separating Rossby and gravity waves. Through such a process, gravity waves can be eliminated (Dickinson and Williamson 1972). However, as has been shown by Leith (1980), this does not work particularly well and the concept of normal mode initialization was first brought to a practical solution by Machenhauer (1977) and Baer and Tribbia (1977), who independently incorporated the effect of non-linear terms. The physical meaning of this non-linear normal mode initialization was to put the initial tendency of the gravity-wave modes equal to zero.
Fig. 12. The change of surface pressure (hPa) as a function of time in a particular point. Full line: uninitialized forecast; dotted line: forecast initialized by non-linear normal mode (Andersen 1977).

Andersen (1977), Daley (1979) and Temperton and Williamson (1979) demonstrated that the method also worked well in multi-level models. Briere (1982) modified the technique to be used for limited-area models and Wergen (1987) generalized it in order to incorporate the effect of diabatic processes, tidal effects, etc. Non-linear normal mode initialization, now widely used, has played a major part in the improvement of numerical weather prediction.

An interesting by-product of data-assimilation is the possibility for data-control and monitoring. Hollingsworth et al. (1986) have developed very powerful systems for routine monitoring of the global observing system. Figure 13 shows the perceived wind and height forecast error for 6 h forecasts verified against 1200 UTC North American radiosonde data for the winter 1984/85. The perceived forecast error has been partitioned into spatially correlated errors (prediction errors) and spatially uncorrelated errors (observational errors).

The size of the estimated observational error agrees by and large with independent evaluations carried out by WMO/CIMO. It is interesting to note that the magnitude of the 6 h forecast error and the analysis error are approximately of the same order. A corresponding study for Europe shows in fact that the observational error is somewhat larger than the forecast error, presumably owing to the many different types of radiosonde equipment in Europe compared with North America.

As a consequence of the high performance of the data-assimilation system, ECMWF has started on a systematic basis to inform operators of the global observing system on the likely deficiencies in the observations. This exercise has revealed several radiosonde stations with systematic errors in the data caused by technical errors in the equipment or in operational procedures. Many of these stations have subsequently been corrected to the benefit of all users of these data.
Figure 13. The total perceived height and wind forecast error over North America. The perceived error has been divided into a correlated part (forecast error) and an uncorrelated part (observation error) as indicated in the legend. The calculations are based on 6 h forecasts for the first quarter of 1984 with the ECMWF system. Note that the sizes of the prediction error and the observational error are comparable.

4. OPERATIONAL FORECAST SKILL

As a consequence of the continuous development of numerical models and data-assimilation systems, the quality of the numerical forecasts has undergone a significant improvement. This advancement can be seen as: (i) an increased quality and accuracy of short-range prediction, (ii) an extension in time of useful predictive skill and (iii) a vast increase in the number of forecast products.

A systematic evaluation of the quality of short-range forecasts has been carried out by the WMO/CAS Working Group on Weather Prediction Research for the last 10 years (Bengtsson and Lange 1982; Lange and Hellsten 1983). Under this intercomparison project, operational forecasts from several centres have been verified daily. Figure 14 presents the standard deviation of the monthly averaged forecast error for 72 h forecasts for the height field at 500 hPa and 1000 hPa. It can be noted that the best available model in 1987 has a standard deviation error of only 55% of the best model available in 1979.

In the following presentation we will show results obtained by the ECMWF model. Results could equally well have been chosen from the operational model of any other large forecasting centre. However, the forecast from ECMWF provides a unique record over a period of almost 10 years of global forecasts on a time-scale as long as 10 days. In addition, as is demonstrated in Fig. 14, objective verification has shown that the forecasts for the time range of 3 days produced by the ECMWF system are more accurate than those produced by other current operational systems. The results described here thus indicate present practical limits of predictability. ECMWF has undertaken daily operational forecasts up to 10 days since 1 August 1980 (5 days a week in the period 1
August 1979 to 31 July 1980). As a consequence of successive improvements of the model and the data-assimilation system, shown in Tables 2 and 3, a gradual extension of predictive skill has taken place. Figure 15 shows the monthly means for 1980–88 as well as the 12-month running averages. The result is given for the northern and southern hemispheres separately. The skill score consists of a weighted average of anomaly correlation and normalized root-mean-square error for extra-tropical temperature and height fields.

Two periods of major improvements can be noted, 1980–81 and 1985–86. The largest contribution during the first period stemmed from an implementation of data checking (elimination of erroneous and non-representative observations) and changes in the

<table>
<thead>
<tr>
<th>Date</th>
<th>Model changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>27 July 1982</td>
<td>SST climatology replaced by analysed SST</td>
</tr>
<tr>
<td>21 Apr. 1983</td>
<td>Spectral model—T63, 16 vertical levels, hybrid vertical coordinates, envelope orography</td>
</tr>
<tr>
<td>2 May 1984</td>
<td>Diurnal radiation cycle</td>
</tr>
<tr>
<td>4 Dec. 1984</td>
<td>New parametrization of radiative transfer, incorporation of stratospheric drag</td>
</tr>
<tr>
<td>1 May 1985</td>
<td>Spectral model—T106, incorporation of shallow convection, modification of the Kuo convective scheme, new interactive cloud parametrization (incorporation of boundary-layer clouds)</td>
</tr>
<tr>
<td>13 May 1986</td>
<td>19 vertical levels—3 additional stratospheric levels</td>
</tr>
<tr>
<td>10 July 1986</td>
<td>Gravity-wave drag</td>
</tr>
<tr>
<td>7 Apr. 1987</td>
<td>New parametrization of land surface processes and vegetation</td>
</tr>
<tr>
<td>5 Jan. 1988</td>
<td>Removal of vertical diffusion in the free atmosphere</td>
</tr>
<tr>
<td>2 May 1989</td>
<td>New parametrization of radiation, convection and gravity-wave drag</td>
</tr>
</tbody>
</table>
TABLE 3. MAJOR CHANGES IN THE ECMWF DATA-ASSIMILATION SYSTEM

<table>
<thead>
<tr>
<th>Date</th>
<th>Data-assimilation changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>21 Sept. 1982</td>
<td>Operational implementation of first system September 1979: Multivariate O/I analysis—</td>
</tr>
<tr>
<td></td>
<td>nonlinear normal mode initialization</td>
</tr>
<tr>
<td>22 May 1984</td>
<td>Diabatic initialization</td>
</tr>
<tr>
<td>26 Feb. 1985</td>
<td>Revised analysis scheme (refined structure functions)</td>
</tr>
<tr>
<td>4 Mar. 1986</td>
<td>Use of high-resolution satellite soundings (TOVS, 250 km)</td>
</tr>
<tr>
<td>11 Mar. 1986</td>
<td>Special handling of tidal signals in initialization</td>
</tr>
<tr>
<td>9 Sept. 1986</td>
<td>Use of precipitable water derived from polar-orbiting satellites</td>
</tr>
<tr>
<td>26 Jan. 1988</td>
<td>Analysis at models levels (hybrid coordinates), further refined structure functions</td>
</tr>
<tr>
<td>12 July 1988</td>
<td>Use of observed divergent wind component</td>
</tr>
<tr>
<td>31 Jan. 1989</td>
<td>Future refined vertical and horizontal structure functions</td>
</tr>
<tr>
<td></td>
<td>Enhanced control of satellite temperature soundings</td>
</tr>
</tbody>
</table>

physical parametrization. The major reason for the improvement in 1985 was the implementation of a high-resolution spectral model, T106, and further improvements in the physical parametrization, deep and shallow convection and cloud description (Jarraud and Simmons 1985; Tiedtke et al. 1988). The positive interaction between the increased horizontal resolution and the physical parametrization had the interesting result that the combined effect of the two changes was larger than the sum of the two changes taken independently.

The skill of prediction models varies considerably in space and time. An example of the variation in predictive skill as a function of the spatial scale may be gained from Fig. 16(a) which shows the anomaly correlation for three separate groups of zonal wave numbers for the geopotential field in the troposphere (see also Fig. 29). The spectral decomposition demonstrates vividly that the larger scales are more accurately predicted, in contrast to earlier experience reviewed by Leith (1978) when medium scales were reported to be better predicted than planetary scales. A poorer forecast of shorter synoptic scales is represented here by zonal wave numbers 10 to 20. The latter result may be associated with the wrong timing of individual weather events within an overall weather situation, which is better predicted. Examples can also be found in which the erroneous forecast of a small-scale feature is followed by a deterioration of the forecast over a larger scale.

The spatial variability of predictive skill has also been examined by comparing objective measures over more limited areas. In the extra-tropics predictive skill is generally lower in areas of high dynamic activity where the error growth is strong. It is also lower downstream of data-sparse regions. During the winter the scores over eastern Asia and North America are often higher than those for the European area and western USA. The predictive skill also varies considerably with height. The lower troposphere is strongly influenced by small-scale features, while the upper troposphere and lower stratosphere are dominated by large planetary waves. Figure 16(b) shows the anomaly correlation of the 50 hPa, 500 hPa and 1000 hPa height fields calculated for most of the northern hemisphere and averaged over all operational forecasts for the winter 1987/88. This shows that 50 hPa has the most accurate forecast and also that the forecast at 500 hPa is generally more accurate than at 1000 hPa, a result in general agreement with synoptic assessment. Using 60% as a measure of useful predictive skill, it is found that the predictive skill at the three levels is more than 10 days, 7 days and 6.5 days respectively. The high predictive skill at 50 hPa during the winter is interesting to note. Bengtsson et al. (1982) have shown an example of an excellent 10-day forecast correctly predicting a split of the polar night vortex into two separate vortices.
Figure 15. A measure of skill of the ECMWF extra-tropical forecasts from January 1980 to April 1989. Above: northern hemisphere. Below: southern hemisphere. The number of days of predictive skill is derived from monthly means of daily averages of the anomaly correlation and standard deviation of the errors of geopotential height and temperature forecasts for the levels 850 to 200 hPa. Thick line, 12-month running mean; thin line, monthly mean values.

Examples of successful forecasts up to a week ahead are shown in Figs. 17 and 18. During the winter 1988/89 the weather over most of Europe was anomalously mild, with drought in the Mediterranean and excessive precipitation in the Scandinavian mountains. In addition to correctly maintaining this flow pattern more than a week ahead, the model also correctly predicted major changes in the dominating flow pattern which took place at the end of February 1989. During this episode the northerly position of the jet
was replaced by a large-scale trough extending deep into the Mediterranean. This change, which was associated with adverse weather in southern Europe, was well predicted one week ahead, as was the return to the anomalously mild weather in early March.

As can be seen from Fig. 15, there are considerable temporal variations in the predictive skill of the forecasts. There is clear annual variation in the northern hemisphere with a minimum in the summer. There is also significant interannual variation in predictive skill, as well as considerable variations within a month or a season.

Figure 19 shows the anomaly correlation scores for every individual day during the winter 1987/88 expressed in the form of a cumulative frequency diagram. As can be seen, useful predictive skill for 90% of the cases varies between 5 and 10 days. An inspection of the daily scores indicates that there are periods of high and low scores lasting from a week to a few weeks. Superimposed on this pattern, there are occasional low scores occurring in a random fashion. Careful investigation reveals that the isolated cases of low scores are mostly related to data problems; such problems have increased somewhat with the latest version of the ECMWF model owing to reduced data-filtering.
Figure 17. ECMWF analysis 500 hPa heights, 1200 UTC on 19 February 1989 (top); seven-day forecast from 1200 UTC 19 February 1989 (middle) valid 26 February, and verifying analysis 500 hPa heights, 1200 UTC on 26 February 1989 (bottom).
Figure 18. ECMWF analysis 500 hPa, 1200 UTC on 26 February 1989 (top); seven-day forecast from 1200 UTC on 26 February 1989 (middle) valid 5 March, and verifying analysis 500 hPa heights, 1200 UTC on 5 March 1989 (bottom).
in the data-assimilation. It is expected that such problems can again be reduced by improved data-checking.

Of more interest from a scientific point of view are the more systematic variations of the forecast skill. Removing cases of high or low scores, which are simply artefacts of the verification procedure (there is normally a strong correlation between an anomaly correlation and the size of the anomaly), there still remains a relation between predictive skill and flow pattern. This can, to some extent, be explained either as a data or a model problem.

If the intense baroclinic zones, for instance, where error growth is rapid, are found in data-sparse regions, say over the oceans and the Arctic, this is likely to have a more detrimental effect than if the baroclinic zones are in data-rich regions. Similarly, there are certain weather situations where a particular model is performing worse owing to limitations in the physical parametrization or where one is more exposed to resolution limitations, such as cases with a double jet structure, where both horizontal and vertical features are very sharp.

However, as has been demonstrated recently by Palmer (1988), there are also clear indications that there are circulation patterns which are inherently more predictable. An example of such a pattern is the so-called Pacific–North Atlantic (PNA) pattern (Wallace and Gutzler 1981), which has been found to give rise to more skilful forecasts over the USA and eastern Pacific than the reverse PNA pattern. By inserting a perturbation off the east coast of China in a barotropic 500 hPa flow pattern, the error growth was significantly larger in the reverse PNA flow pattern than the error growth of the same perturbation in the normal PNA pattern. In other parts of the northern hemisphere, such as over Europe, the situation is less clear, with more complex relations.

In view of the large variability in predictive skill from episode to episode, there is a considerable practical interest in having an estimate of the probable reliability of a given forecast—a kind of confidence factor. On a trial basis, ECMWF has started such quasi-operational experiments (Palmer and Tibaldi 1988). So far the practical value of these experiments has been limited. In the longer term a type of Monte Carlo forecast is
anticipated, in which several forecasts are run from slightly different but equally likely initial states. The difference of such forecasts in space and time will provide users with reliability factors.

5. PREDICTION OF INTENSE EXTRA-TROPICAL CYCLONES

The westerlies are generally baroclinically unstable all through the year and the development of cyclonic eddies of different scales and intensity is a fundamental feature of the general circulation of the atmosphere. Cyclones are associated with adverse and rapidly changing weather, such as strong winds and precipitation, and can in their most intense forms constitute a major hazard. Accurate and timely prediction of extra-tropical cyclones accordingly constitutes a major objective for the meteorological services. As numerical models have developed, so has their ability to predict intense extra-tropical depressions improved (Gadd and Kruze 1988).

In this section we will describe the prediction of an intense cyclogenesis which took place over the North Atlantic Ocean on 7–8 February 1988. The period 28 January–8 February was characterized by a very strong westerly flow over the North Atlantic, associated with sharp temperature contrast between north and south. This condition created a very unstable flow in which three intense cyclogenoses took place. Figure 20 shows the observed and predicted surface pressure in the centre of these depressions as well as the geographical position for the beginning of the development and the position where the cyclone reached its maximum intensity. The latest of these cyclones was the most intense and reached a maximum deepening rate of 3 hPa h⁻¹. The corresponding forecasts with the operational ECMWF T106-L19 model were quite successful and predicted all the three cyclones with considerable accuracy. The operational forecasts from 1200 UTC on 4 February 1988 predicted a cyclogenesis which started at the eastern US seaboard around 38°N. The storm move swiftly along the US and Canadian coastline, rapidly deepening. After 60 hours it had reached the southern part of the Davis Strait where it slowly started to fill. On its southern flank very cold arctic air from Canada was advected east and south-east over the Atlantic Ocean with a widespread temperature fall of 15–20 degC in 12 hours. The intense cold air advection sharpened the frontal zone over the western Atlantic. This in turn enhanced the conditions for a second cyclogenesis which started to develop around 0000 UTC on 7 February in the area around 35°N and 65°W. This cyclone developed with extreme intensity with a central pressure fall of more than 30 hPa in 12 hours (Fig. 20(c)). The accuracy of the prediction was quite remarkable, as can be seen from Fig. 21, showing the surface pressure and 850 hPa temperature as predicted and observed from T + 60 h up to T + 120 h.

In many ways this storm was a typical ‘textbook’ development and showed all the characteristic features which we associate with an intense extra-tropical depression. We will illustrate this by showing a series of predicted quantities for the most intense phase of the development, which took place between T + 72 h and T + 96 h. Figure 22 shows the wind field at the lowest model level, 30 m above the surface. The well developed discontinuity in the wind field indicates the position of the surface warm and cold fronts; the rapid shrinkage of the warm sector between T + 84 h and T + 108 h is evident. Twelve hours later the occlusion is completed and a warm sector can no longer be identified at the surface.

Figure 23 shows the predicted cloud cover. Low, middle, high and convective clouds are shown separately. The representation of clouds in the ECMWF model is shown in Fig. 24. The cloud forecast appears realistic. Infra-red cloud images from Meteosat valid for the same time are shown in Fig. 25. The cloud prediction and their evaluation are in
agreement with synoptic experience. It is interesting to see the band of clear airmass gradually increasing in depth; behind the cold front deep convective cloud systems are developing. The transfer of sensible and latent heat in the cold airmass to the north and west of the cyclone is intense; the total maximum heat flux amounts to more than 1 kW m$^{-2}$. The magnitude and distribution of sensible and latent heat flux is in excellent agreement with diagnostic calculations (e.g. Petterssen et al. 1962). In Fig. 26 we show the instantaneous precipitation expressed in mm (2 h)$^{-1}$ for the stratiform and the convective precipitation respectively. The stratiform precipitation dominates and amounts to around 4 mm (h)$^{-1}$ during the most intense phase of the development. Since the stratiform precipitation takes place in the centre of the cyclone, it contributes significantly to the rapid deepening process. However, this is not the case with the surface
Figure 21. Surface pressure and 850-hPa temperature prediction and verification for every 12 hours from T + 48 h to T + 120 h for the forecast starting 1200 UTC on 4 February 1988.
Figure 22. Predicted wind vectors at the lowest σ-level of the ECMWF model (30 m above the surface) for (a) T + 84 h and (b) T + 108 h in the same forecast as Fig. 21. Wind vectors are given a unique colour for every 4 deg C as indicated by the 850 hPa temperature field. Observe the very rapid occlusion process.
Figure 23. Predicted cloud cover at (a) $T + 84h$ and (b) $T + 108h$ in the same forecast as Fig. 21. Low (yellow), middle (green), high (blue) and convective clouds (red) are shown. Cloud description in the ECMWF model is shown in Fig. 24.
Figure 24. Cloud description in the ECMWF model (Tiedtke et al. 1988).

Figure 25. Observed clouds from Meteosat (infra-red) for 0000 UTC on 9 February 1988 (corresponding to the T + 108 h forecast; see Figs. 22 and 23).
flaxes which occur in the cold airmass. Their effect is likely to be the reverse with respect to the cyclogenesis process; it will heat the cold airmass and thereby reduce the available potential energy of the storm.

As discussed by Kuo and Reed (1988), it is expected that the accuracy of the adiabatic calculation (including high resolution) and the release of latent heat are of significance for the prediction of intense cyclogenesis, while other processes such as the detailed description of the surface friction are of minor importance. In order to demonstrate this, a series of numerical experiments have been carried out. The result of these studies is as follows:
(i) **Horizontal resolution.** For the case from 28 January 1988 (Fig. 20(a)) separate forecasts were made with four different horizontal resolutions: T21, T42, T63 and T106. The physical parametrization and the initial state were the same in all cases, but orography and land–sea contrasts were consistent with the respective horizontal resolution. The result is given in Fig. 27, which shows the predicted pressure in the centre of the cyclone.

![Graph showing pressure as a function of time in the Centre of the low, forecast from 1200 UTC on 28 January 1988.](image)

Figure 27. Surface pressure as a function of time in the centre of the low, forecast from 1200 UTC on 28 January 1988 (Fig. 20(a)). Forecasts for T21, T42, T63 and T106 are shown separately (see legend).

While the T106 forecast is in good agreement with observation, a steady deterioration can be seen as resolution becomes coarser. The T63 reaches a maximum depth of only 972 hPa compared with 950 hPa for the T106 and has no clearly defined closed circulation between T+66 and T+78h. The T42 calculation has no clearly defined centre after T+60h but a trough reaching a maximum depth of only 984 hPa. The T21 does not develop any separate cyclone centre at all, except a noticeable weak centre around T+36h. The actual forecasts at T+96h (not shown here) differ considerably; the T106 has positioned the cyclone correctly over northern Scotland, the T63 has its centre over south-eastern England.

(ii) **Hydrological cycle.** In another experiment the T106 model was run with the release of latent heat switched off. This forecast failed to develop the storm. Figure 28 shows the comparison between the ‘dry’ and ‘moist’ integration.

(iii) **Modification of surface friction.** In this experiment the Charnock formula was replaced by a modified scheme which enhanced the surface friction over ocean areas as a function of the intensity of the surface wind. No noticeable difference in the forecast could be observed.
Figure 28. 96-hour forecast for 1200 UTC on 28 January 1988. Left: operational 500 hPa height field (top) and surface pressure (bottom). Right: the forecast obtained by switching off the release of latent heat (dry model). The deep low over Scotland is described in Fig. 20(a).
6. Systematic Model Errors

Owing to the general instability of the flow of the atmosphere at middle latitudes, the error growth is large. Several studies indicate an error doubling time of around two days for the synoptic-scale flow as determined by the 500 hPa height field (Lorenz 1982; Dalcher and Kalnay 1987). Actual atmospheric models have a larger error growth due to model error. Bengtsson (1989) has recently estimated the error growth by the ECMWF model for the winter 1986/87 (DJF). Using an error growth equation of the form

\[ \frac{dV}{dt} = (aV + S)(V - V_\infty) \]  

(6.1)

where \( V \) is the forecast error variance of the 500 hPa height field, \( a \) the rate of growth of the forecast error variance, \( S \) the amount of error variance introduced by model deficiencies in one day and \( V_\infty \) the asymptotic value of the error variance. The value of \( S \) has been empirically determined to be 144 m^2. Comparing the result from previous years it was found that \( S \) had gradually been reduced; for the winter 1980/81 the corresponding value was 324 m^2.

A more detailed inspection shows that the error of the forecast can be split into a 'systematic part' or average error for a particular month or season, and a 'transient part' which changes from day to day. The sum of the systematic and transient error variance constitutes the total error variance. Figure 29 (from Mureau 1990) shows the growth of error variance for different scales of motion: T40, T10, T7 and T5 for the 500 hPa height field for the winter 1980/81 and the winter 1987/88 respectively. The stippled area between the pair of curves indicates the systematic error. The figure shows clearly that the predictability increases when the field is truncated, indicating that the largest scales of motion are the most predictable; this was certainly not the case with forecast models from the 1970s or earlier. The improvement in prediction skill during the seven years is evident with a considerable reduction in the systematic error, in particular for the largest scales of motion.

Systematic errors exist in all models. A summary of typical errors is given in Table 4 (ECMWF 1988). These errors are common for most models. The reason for the systematic errors is not known but they are sensitive to the parametrization of certain processes. These processes are indicated in the table.

As shown in Table 4, a serious deficiency of present models is their reduced ability to predict blocking for more than about five days ahead. Atmospheric blocking has for a long time been recognized as a process of profound dynamical interest and is of great practical relevance to operational weather prediction. The crucial effects of the onset, development and decay of block-like structures on atmospheric low-frequency variability, and hence on forecasting on different time-scales, have made blocking one of the most studied atmospheric processes.

Tibaldi and Molteni (1988) have recently undertaken a diagnostic evaluation of the capability of the ECMWF model to predict blocking. In order to arrive at a simple and yet reasonably realistic definition of blocking, they have used a definition essentially derived from a study by Lejenäs and Økland (1983). Following this idea, the geopotential height gradients are computed between 60°N and 40°N for every latitude line. When this gradient is positive, i.e. the geopotential height at 60°N is larger than at 40°N, and where there at the same time exists a negative gradient between 80°N and 40°N (more than 10 m/deg lat) a state of blocking is defined. This definition agrees conceptually rather well with the original phenomenological definition of blocking as proposed by Rex (1950).

Figure 30 shows the percentage frequency of blocking as a function of longitude,
calculated for the period 1 December to 10 March for the seven years 1980/81–1986/87 from the ECMWF operational analyses. Figure 31 shows the corresponding prediction of blocking as calculated from the ensemble of forecasts valid at day 1, day 3, day 6 and day 10, respectively. The reduction in the frequency of blocking as the forecast is extended can clearly be seen. Although a gradual improvement has taken place in the prediction of blocking during the seven-year period, the characteristic deficiency still remains, in particular in the Pacific sector.

Tibaldi and Molteni (1988) also evaluated the model’s ability to predict blocking correctly. It was found that the prediction of blocking was quite successful when the onset signal of blocking did exist initially, as previously demonstrated in case-study evaluations by Bengtsson (1981), and Grönås (1982). On the other hand, the performance deteriorated rapidly when the blocking took place 4–5 days into the forecasts. While the prediction of blocking per se may be limited owing to the unstable nature of the
TABLE 4. SYSTEMATIC MODEL ERRORS

<table>
<thead>
<tr>
<th>Phenomenon</th>
<th>Known to be sensitive to</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Extra-tropics</td>
<td>Momentum fluxes*, upper boundary condition, radiation in stratosphere</td>
</tr>
<tr>
<td>Zonalization of mean mid-latitude circulation, i.e. too weak diffuent flow over Europe and Northern Pacific. Dipole pattern of 500 hPa height error</td>
<td></td>
</tr>
<tr>
<td>Reduced variability of standing waves from month to month (too small low frequency variability)</td>
<td></td>
</tr>
<tr>
<td>Poleward and upward shift of subtropical jets</td>
<td>Convection, momentum fluxes</td>
</tr>
<tr>
<td>Reduced ability to predict blocking more than 5 days ahead</td>
<td>?</td>
</tr>
<tr>
<td>Increase of eddy momentum flux together with increased tilting of troughs and ridges</td>
<td>?</td>
</tr>
<tr>
<td>(ii) Tropics</td>
<td></td>
</tr>
<tr>
<td>Excessive easterlies in upper tropical troposphere</td>
<td>Convection, momentum fluxes</td>
</tr>
<tr>
<td>Weakening of the trade winds</td>
<td>Radiation, convection</td>
</tr>
<tr>
<td>Upper branch of Hadley circulation too weak and spread over too many layers</td>
<td>Radiation, convection</td>
</tr>
<tr>
<td>Weakening of divergent mean flow over main tropical convection areas in connection with reduced precipitation</td>
<td>Radiation, convection</td>
</tr>
<tr>
<td>Overestimation of ITCZ in eastern Pacific</td>
<td>?</td>
</tr>
<tr>
<td>Weakening of transients in tropics</td>
<td>Convection, momentum fluxes</td>
</tr>
</tbody>
</table>

* Momentum fluxes including gravity waves, vertical and horizontal diffusion and mountain blocking.

Figure 30. Percentage frequency of blocking as a function of longitude, computed from all ECMWF daily objective analyses for the 500 hPa height field (all winters 1980/81-1986/87).

phenomenon, the development of a model which can generate blocking with the same frequency as observed in nature is certainly a tractable problem. Consequently, before this has been achieved, there is no reason to be overly pessimistic concerning the prediction of blocking.
Figure 31. Percentage frequency of blocking as a function of longitude. Dotted lines as in Fig. 30. Full lines: blocking frequency computed on the forecast fields, for Day 1, 5, 6 and 10 (panels (a) to (d), respectively).
7. Future prospects

Numerical weather prediction has undergone an exceptional development since its inception almost 40 years ago. It has developed from a limited experimental activity into a major function of the meteorological services, without which present-day forecasting services to the public and different special customers would be impossible. The quality of the forecasts has improved in a way which could hardly be foreseen in the early years; useful forecasts in the extra-tropics of the northern hemisphere have been extended to up to a week ahead and forecast products are now made available for a wide range of products covering the whole globe from the surface of the earth to high into the stratosphere.

The basic conditions which have made this evolution possible are the spectacular developments which have taken place in observational technology and computer technology.

Global observations are regularly made available from polar-orbiting and geostationary satellites; these data, although not of the same quality as radiosonde observations, are unique in the sense that they provide data from previously data-empty regions; without them global weather prediction would not be achievable. Other important contributions are aircraft observations (accurate but sparse), observations from drifting buoys and a slow but steady increase of observations from ships of opportunity; here data collection by satellites has made a great contribution. Furthermore, observations are now better utilized through improved analysis methods, more accurate initialization, and a more accurate and consistent use of the prediction model to provide the first guess.

Computer development has been and continues to be extraordinary. At the leading edge are the supercomputers which by now can perform well over 10^9 floating-point operations (1 GigaFlop) per second averaged over a complete forecast; this is more than a million times faster than the early electronic calculating machines used in the early 1950s. The increase in storage capacity has been more or less equivalent. The supercomputers have made it possible to develop successively more and more realistic models of the atmosphere and to undertake the time-consuming data processing required for the assimilation of observations.

Significant scientific progress has taken place in conjunction with the technical achievements. Efficient numerical integration algorithms (such as semi-implicit and semi-Lagrangian time integration) have made it possible to increase the time-step, thus speeding up the calculation further. Other contributions have eliminated non-linear instabilities and made extended integrations possible.

The parametrization of physical (sub-grid-scale) processes has developed in an impressive way, although the data required for a deeper understanding of these processes and the necessary computer resources to undertake the calculations of them in a more accurate way are still not available.

How is this development likely to continue? Predictability studies (Lorenz 1982; Dalcher and Kalnay 1987) indicate that present forecast skill can be extended by about two days, solely through the perfecting of the forecasting model. A further improvement by perhaps another day is feasible by making efficient use of the same perfect model in the data-assimilation process (Bengtsson 1989).

However, further improvement of numerical modelling will require a deeper understanding of sub-grid-scale processes and their interaction with the large-scale flow. Examples of such processes are cloud and radiation, convective precipitation, land surface evapotranspiration and momentum flux in convective systems. We do not yet
have a comprehensive understanding of how to parametrize these processes in numerical models. Progress is expected to be slow, because of lack of relevant observations.

The other way to improve numerical forecasts is by more accurate observations. According to Lorenz (1982) a halving of the initial error is expected to extend predictive skill by about two days. The development of more advanced observing systems progresses at a disappointingly slow pace, and practically no progress has been made since the Global Weather Experiment ten years ago. Present short-range forecasts (<48 hours) over and downstream of the continental areas of the northern hemisphere are generally more accurate than present satellite temperature and wind observations. This has very much reduced the overall value of satellite observations for numerical weather prediction of the northern hemisphere. Although it is expected that more advanced retrieval methods will enhance the use of the data from present satellite systems, urgent efforts are required to speed up the development of more accurate three-dimensional measurements of wind and temperature on a global scale.

However, in the short term, there are many ways to improve the present global observing system. Presently, owing to deficiencies in observational technique and careless procedures, almost 10% of all radiosonde observations are useless because of gross errors. Another 10% of the radiosonde observations are not made available in time or arrive too late for operational use because of deficiencies in the telecommunication network. Minor investment in more efficient management, training and better use of modern communication techniques would make these data available for numerical weather prediction.

High-quality wind observations are regularly being recorded by commercial aircraft, but only a fraction of these observations are made available to the meteorological services for operational weather prediction. Similar deficiencies exist with respect to data from ships, in spite of rapid developments of satellite communications systems. There is no question that major efforts are required in order to improve the operational database both in the short and long term. If this is not done, the large improvement which we have seen in numerical forecasting over the last decade will come to a halt.

While the progress to improve the database is slow, the opposite is true for computers. The construction of faster processors as well as the building of computer systems having more and more processors working in parallel means that the overall computing capacity continues to increase at a rapid pace. Present supercomputers already offer impressive peak performances of several Gigaflops. Such results can only be obtained for specific algorithms which have been specifically adjusted to a particular machine. For a code as complex as a numerical model the average performance is generally significantly lower. The average performance of the present operational ECMWF model on the CRAY X-MP/48 computer, using all four processors simultaneously, is about 0.4 Gigaflops. The time it takes to run the T106-L19 out to 10 days is about 3 hours. Table 5 gives the time,

<table>
<thead>
<tr>
<th>Model</th>
<th>Horizontal grid</th>
<th>Computers</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cray 1</td>
<td>Cray X-MP/48</td>
<td>'X'</td>
</tr>
<tr>
<td>T62-L16</td>
<td>200 km</td>
<td>4 h</td>
<td>30 min</td>
<td>2.5 min</td>
</tr>
<tr>
<td>T106-L19</td>
<td>125 km</td>
<td>---</td>
<td>3 h</td>
<td>15 min</td>
</tr>
<tr>
<td>T213-L31</td>
<td>60 km</td>
<td>---</td>
<td>---</td>
<td>3 h</td>
</tr>
</tbody>
</table>

T63-L16 refers to a spectral model with 63 wave numbers in triangular truncation and 16 vertical levels. 'X' represents a projected supercomputer from 1992 with an average processing speed of 5 Gigaflops.
observed and projected, that it will take to run a global spectral model with different horizontal and vertical resolutions. The estimates are given using a semi-implicit time-integration scheme, where the time-step is inversely proportional to the spectral resolution. Using a semi-Lagrangian time-integration it appears feasible to increase the time-step by a factor of two to three, leading to a corresponding saving of computer time.

Computers with sustained speeds of 5 Gigaflops, which appears feasible during the next two to three years, will make it possible to run global models with a horizontal resolution of the order of 50–60 km and vertical resolution of 30–40 levels. Plans for further improvements in computing speed and performance of up to 40 Gigaflops are already being planned by supercomputer manufacturers. Such computers may be available during the second half of the 1990s.

However, increase of resolution is not only a technical matter but also to a high degree a scientific problem. As has been found in recent experiments with high-resolution limited area models (Lambert et al. 1988) the parametrization of convective precipitation in particular is sensitive to resolution; indications are in fact that parametrization may have to be retuned and even perhaps be reformulated when resolution is changed significantly. Other problems are aliasing of frontal systems and mesoscale phenomena which may take place as horizontal resolution increases beyond certain critical scales. Consequently, major scientific efforts will be required to develop such high-resolution models properly.

In addition, the increasing difference in the resolution of models and the resolution of the observing system is creating severe data-assimilation problems. A model of the resolution discussed above will certainly both resolve and predict the evolution of lee-wave cyclogenesis, polar lows and tropical cyclones; phenomena which it will at best be possible to verify only qualitatively by satellite images. It will be very difficult to set up a data-assimilation system whereby sparsely distributed observations will adjust and correct synoptic features generated by the high-resolution model in a dynamically and physically consistent way.

Given the conditions in computers and observations, what are the ways whereby we can accomplish further improvements in numerical predictions?

Firstly, we may assume as a reasonable proposition that it is feasible to build a forecasting model which will have the same statistical behaviour as the real atmosphere, i.e. with the same mean state and with the same variability in time and space. As has been discussed in section 6, this has not yet been achieved and there are, in spite of significant improvements in recent years, still important differences between the behaviour of the real atmosphere and that of the most advanced models. The systematic errors in the large-scale flow and deficiencies in properly simulating low-frequency variability, such as blocking patterns, are examples of such model limitations.

There does not seem to be any realistic alternative to the ‘brute force approach’ and future models are likely to be improved through the systematic and meticulous development of all aspects of the forecasting system. An alternative approach can possibly be considered when the forecasts are extended beyond the predictability of the individual weather systems. As has been shown by Miyakoda and Chao (1982) and Shukla (1984) there are indications that useful prediction of time averages can be extended even further, perhaps up to a month or more. The fact that considerable anomalies can exist for this length of time and that present general circulation models can simulate such long-lasting anomalies even in the extra-tropics (Lau 1981; Volmer et al. 1983) gives us hope that we may be successful.

However, there are other interesting aspects of weather prediction. As was discussed in section 4, predictive skill varies considerably from period to period. Individual forecasts
verified over the northern hemisphere extra-tropics vary in useful skill over a range of more than five days (Fig. 19). If the skill of a given forecast can be estimated in advance and, so to say, be given a confidence factor, this would add considerably to its value for the forecasters. Alternatively, a whole series of forecasts, presumably equally likely, could also be made available to the forecaster, who himself could judge the likelihood of a particular weather event occurring. In fact this is a situation which many forecasters to some extent face already when they receive forecasts from different forecasting centres around the world.

8. SUMMARY AND CONCLUSION

We have reviewed advances in numerical weather prediction with the emphasis on the development which has taken place over the last 10–15 years. Considerable progress has taken place and useful forecasts, 6–7 days ahead, over the extra-tropical latitudes of the northern hemisphere are being made daily. For the southern hemisphere the corresponding skill is 4–5 days. Progress has also taken place in the prediction of intense cyclogenesis, which has led to great improvements in providing proper warnings to the public and to weather-sensitive industries. Major cyclogeneses are now being predicted up to five days in advance. Further improvements in weather prediction are possible but lack of accurate observations on a global scale is hampering the development. Substantial efforts are required, preferably on the same level as during FGGE, in order to bring observational technology to a more advanced level.

In spite of all the improvements, numerical models are still afflicted with systematic deficiencies. These are particularly serious in the tropics and probably related to incorrect treatment of physical processes. The problems are partly the same as those for climate models, and it is expected that weather prediction models and climate models will coincide even more in the future. Parametrization of clouds, land surface processes and air–sea interactions are as essential for weather prediction as for climate simulation and prediction.

REFERENCES

Andersen, J. H. 1977 'A routine for normal mode initialization with non-linear correction for multi-level spectral model with triangular truncation.' ECMWF Internal Report No. 15


1972 'Design of the UCLA General Circulation Model.' Technical Report No. 7, Department of Meteorology. University of California, Los Angeles, California


Daley R. 1979 The application of non-linear normal mode initialization to an operational forecast model. Atmos.-Ocean, 17, 97–124

Defant, F. 1959 On hydrodynamical instability caused by an approach of subtropical and polar jet stream in northern latitudes before the onset of strong cyclogenesis. In The atmosphere and the sea in motion, Ed. B. Bolin, New York, 305–325


Dell'Oso, L. and Bengtsson, L. 1985 Prediction of a typhoon using a fine-mesh NWP model. Tellus, 37A, 97–105

ECMWF 1988 Proceedings from the ECMWF Workshop on Diabatic Forcing, 30 November—2 December 1987, ECMWF

Edelmann, W. 1984 A convection scheme used by the German forecast model. LAM Newsletter, Number 8 (May 1984). European Working Group on Limited Area Modelling.


Eliaissen, A. 1954 Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Rap. No. 5 Videnskapsakademisk Institutte for Vær og Klimaforskning, Oslo, Norway


Gandin, L. S. 1963 Objective analysis of meteorological fields. Translated from Russian by the Israel Program for Scientific Translations (1965)


Kuo, H. L. 1974 Further studies of the parameterization of the influence of cumulus convection on large-scale flow. J. Atmos. Sci., 22, 40–63


Lorenz, E. N. 1982 Atmospheric predictability with a large numerical model. Tellus, 34, 505–513


Machenhauer, B. 1988 Objective analysis and numerical forecasting of an explosive deepening cyclone using pre-operational HIRLAM systems. ECMWF Seminar on the nature and prediction of extra-tropical weather systems, 7–11 September 1987, Volume 2, 193–230


<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phillips, N. A.</td>
<td>1957</td>
<td>A coordinate system having some special advantages for numerical forecasting. J. Meteorol., 14, 184–185</td>
</tr>
<tr>
<td>Rex, D. R.</td>
<td>1950</td>
<td>Blocking action in the middle troposphere and its effect upon regional climate. II: The climatology of blocking actions. Tellus, 2, 275–302</td>
</tr>
<tr>
<td>Richardson, L. F.</td>
<td>1922</td>
<td>Weather prediction by numerical process. Cambridge University Press</td>
</tr>
<tr>
<td>Smagorinsky, J., Miyakoda, K. and Strickler, R. F.</td>
<td>1970</td>
<td>The relative importance of variables in initial conditions for dynamical weather prediction. Tellus, 22, 141–157</td>
</tr>
</tbody>
</table>


