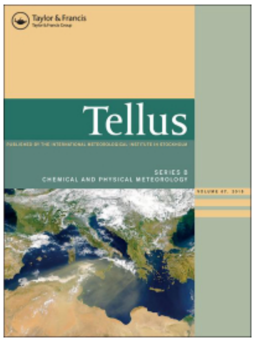


Tellus B: Chemical and Physical Meteorology

ISSN: (Print) 1600-0889 (Online) Journal homepage: https://www.tandfonline.com/loi/zelb20

From short-range barotropic modelling to extended-range global weather prediction: a 40-year perspective

Lennart Bengtsson

To cite this article: Lennart Bengtsson (1999) From short-range barotropic modelling to extended-range global weather prediction: a 40-year perspective, Tellus B: Chemical and Physical Meteorology, 51:1, 13-32, DOI: 10.3402/tellusb.v51i1.16257

To link to this article: https://doi.org/10.3402/tellusb.v51i1.16257

Tellus (1999), 51A-B, 13–32
Printed in UK – all rights reserved

TELLUS
ISSN 0280–6495
ISSN 0280–6509

From short-range barotropic modelling to extended-range global weather prediction: a 40-year perspective

By LENNART BENGTSSON, *Max-Planck-Institute für Meteorologie, Bundesstrasse 55, D-20146 Hamburg, Germany*

(Manuscript received 13 October 1998; in final form 30 October 1998)

ABSTRACT

At the end of the 20th century, we can look back on a spectacular development of numerical weather prediction, which has, practically uninterrupted, been going on since the middle of the century. High-resolution predictions for more than a week ahead for any part of the globe are now routinely produced and anyone with an Internet connection can access many of these forecasts for anywhere in the world. Extended predictions for several seasons ahead are also being done — the latest El Niño event in 1997/1998 is an example of such a successful prediction. The great achievement is due to a number of factors including the progress in computational technology and the establishment of global observing systems, combined with a systematic research program with an overall strategy towards building comprehensive prediction systems for climate and weather. In this article, I will discuss the different evolutionary steps in this development and the way new scientific ideas have contributed to efficiently explore the computing power and in using observations from new types of observing systems. Weather prediction is not an exact science due to unavoidable errors in initial data and in the models. To quantify the reliability of a forecast is therefore essential and probably more so the longer the forecasts are. Ensemble prediction is thus a new and important concept in weather and climate prediction, which I believe will become a routine aspect of weather prediction in the future. The limit between weather and climate prediction is becoming more and more diffuse and in the final part of this article I will outline the way I think development may proceed in the future.

1. Introduction

There are few individuals who have had a larger influence on the development of meteorology and atmospheric science than Carl Gustaf Rossby. As is abundantly demonstrated in his monumental review on current problems in meteorology for the Swedish Natural Research Council (Rossby, 1957), his grasp on the key issues and his insight into the fundamentals of physics, dynamics and chemistry of the atmosphere were quite extraordinary not least in the retrospective of the actual developments during the last 40 years. Rossby was one of the principal initiators of numerical weather prediction (NWP) and of the establishment of NWP as a major tool for weather forecasting by the meteorological services.

Rossby realized very early the importance of numerical weather prediction and was instrumental to start the very first experiments at the Institute for Advanced Studies in Princeton in the late 1940s. When he returned to Sweden after the years in United States, he established the International Meteorological Institute in Stockholm and set up a group of highly talented young scientists. In a short time the institute became the leading centre for numerical weather prediction in the world (Wiin-Nielsen, 1990; Bolin, 1999), and practically all leading meteorologists visited for a shorter or longer period of time Rossby's Institute in Stockholm.

From an exclusive research activity at a few leading institutes around the world in the 1950s, numerical weather prediction is today a routine operational function at practically all weather services. The predictions are done for all parts of the world. The integration domain covers a wide range of areas from smaller regions to the whole globe and from the surface boundary layer to the upper stratosphere. The quality of the predictions has undergone a spectacular improvement and useful predictions are now routinely being done for more than a week ahead.

There are several reasons behind this development, the most important being our advances in the scientific understanding of the prediction and data assimilation problem. A necessary condition has also been the establishment of a global observation system and the rapid development of increasingly powerful computers. These factors have made it possible to develop increasingly realistic models of the atmosphere and to construct advanced methods for the assimilation of observations from a multitude of observational systems of the atmosphere and from the surface of the earth.

I will here highlight some key aspects of the spectacular development in numerical weather prediction starting with the early real-time experiments in Stockholm in the 1950s. I will commence this presentation by shortly describing the type of models which were used in the first numerical weather prediction experiments in the 1950s. The experience from these prediction experiments set the scene for a more comprehensive formulation of numerical forecasting and data-assimilation. The 1960s and 1970s were characterized by deepened understanding of the dynamics and physics of the atmosphere, and it was very much during these decades that the major foundation was laid for the development which was to evolve in the following years, with the establishment of operational global prediction and the extension of the length of the forecasts into medium range.

The extension of the prediction domain to the whole earth necessitated the use of models which could simultaneously handle the prediction of extra-tropical depressions and tropical phenomena such as variations in the monsoon circulation, atmospheric response to sea surface temperature anomalies such as El Niño, and the developments of hurricanes and typhoons. The extension of the forecasting range from a few days to a week or longer also required an extension of the models to incorporate a major part of the stratosphere.

As a central part of this study I will retrospectively try to determine the main factors that contribute to the large improvements in predictive skill. This is difficult to do for many reasons, such as the strong non-linearities of atmospheric processes and frequent changes in the operational forecasting system, including usage of new types of observations, extension of the integration domain and numerous modifications of numerical methods, physical parameterization and data-assimilation. The analysis will therefore only be tentative.

Finally, I will outline some circumstances in the present ongoing evolution which I anticipate will determine the way numerical prediction will be shaped in the beginning of the next century. Two major points will be mentioned: The establishment of ensemble prediction and the efforts to build forecasting systems for the coupled ocean-land-atmosphere for the combined prediction of weather, climate anomalies and environmental changes.

2. The early years

Wiin-Nielsen (1991) has given a comprehensive account of the early years of numerical weather prediction including the work which took place at the International Meteorological Institute in Stockholm. During a period of a few years, mainly 1952–56, the development of a forecasting system based on the equivalent barotropic model took place in cooperation between the Institute and the Weather Service of the Swedish Air Force. The forecasts were carried out over a major part of the extratropical region of the northern hemisphere (a rectangular grid consisting of 32 × 41 points with a gridsize of 300 km at 50°N). It included objective analysis based on the correction method (Bergthorsson and Döös, 1955) and the necessary pre- and postprocessing including the decoding of radiosonde data and graphical presentation of the results (Staff Members, 1954; Bergthorsson et al., 1955; Döös, 1956). The full calculation was ultimately carried out in less than two hours in order to make the results operationally useful. Fig. 1 illustrates the numerical solution of the barotropic model on the Swedish-built BESK computer. The first integration area comprised only 20 × 20 gridpoints, but was later on extended to the larger grid of 32 × 41 points. With an 1 h time step the total number of calculations

First real-time operational NWP

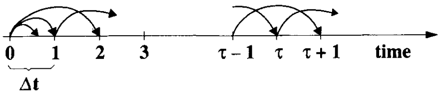
The Barotropic model

$$\frac{D}{Dt}(\zeta+f)=0 \qquad f = 2\Omega\sin\varphi$$

$$\frac{\partial}{\partial t}\nabla^2\phi = J(\zeta+f,\phi)$$

ϕ **= geopotential height** **m = mapfactor**

$$\zeta^{\tau+1} = \zeta^{\tau-1} + 2\Delta t J\left(m^2\zeta^{\tau}+f\ ,\phi^{\tau}\right) \qquad \zeta = \frac{1}{f_0}\nabla^2\phi$$



		ϕ^{τ}	$\zeta^{\tau-1}$	
a) Compute	$\zeta^{\tau} = \frac{1}{f_0}\nabla^2\phi^{\tau}$	ϕ^{τ}	$\zeta^{\tau-1}$	ζ^{τ}
b) Compute	J^{τ}	J^{τ}	$\zeta^{\tau-1}$	ζ^{τ}
c) Extrapolate	$\zeta^{\tau+1} = \zeta^{\tau-1} + 2\Delta t J^{\tau}$	J^{τ}	$\zeta^{\tau+1}$	ζ^{τ}
d) Solve the Poisson eq.	$\zeta^{\tau+1} = \frac{1}{f_0}\nabla^2\phi^{\tau+1}$	$\phi^{\tau+1}$	ζ^{τ}	
		14 bits	**13 bits**	**13 bits**

Integration area:
20 x 20 grid points; Δs= 300 km; Δt=1hr.

Fig. 1. The numerical solution of the barotropic vorticity equation. The finite difference form is valid for a polar-sterographic projection where *m* is the map scale factor. The time-integration is with a leap-frog scheme with an initial short forward timestep. For an optimum usage of the limited memory (originally only 512 words of 40 bit each) every bit has to be used efficiently. At a start of a given cycle geopotential at the present timestep and vorticity, ξ, at the previous timestep is assumed to be known. In the first step (a) vorticity for the present timestep is calculated and in step (b) the Jacobian determinant (advection of vorticity). In step (c) the vorticity for the next timestep is calculated. Finally in step (d) the geopotential for the next timestep is calculated by means of Liebmann relaxation. Both geopotential and vorticity have now been forwarded one timestep and the process is repeated. Courtesy of B. Döös.

for the larger area was some 3×10^7 and took 20–25 min computer time. With the very limited memory size, originally only 512 words, with each word having 40 bits, the programmers had to be very skilful. The programming was done in machine language. There was no floating point unit and every step of the calculations had to be be carefully scaled in order not to lose accuracy. Fig. 3 is an example of one of the very first real-time predictions carried out from 1 October 1954 and valid for 72 h. This forecast was operationally used by the Swedish Air Force during a military field exercise. The quality of the forecast was poor by today's standard, but succeeded in predicting the building-up and strengthening of the ridge over Scandinavia. (Figs. 3, 4). Wiin-Nielsen (1991) has pointed out that on 23 and 24 March 1954, two 24-h forecasts were already made which were also finished in time to be of use in operation. This was probably the first real-time numerical forecast ever.

The Institute undertook series of experimental forecasts. Results of these early forecasts can be found in papers by Staff Members, Institute of Meteorology (1954), Bergthorsson et al. (1955) and Döös (1956). The results from the first series of 24 24 h forecasts had an average correlation coefficient of 0.77 when validated over Europe. The results reported by Döös (ibid) for 36 forecasts for November and December 1955 are summarized in Table 1.

Evaluation of the forecast results indicated different sources of the errors. Some were clearly related to observational errors and to boundary errors caused by the use of a limited domain. Other errors were related to baroclinic processes (not possible to handle by the barotropic model) and to developments occurring at scales which were not resolved by the grid of the model. However, one of the most intriguing errors was related to the inability to predict the phase speeds of the ultralong waves. As can be seen in Wiin-Nielsen (1991), this was partly overcome by introducing a large scale divergence by allowing the barotropic flow to have a free surface (compare

Table 1. *Prediction results of the 500 hPa height field for 36 forecasts Nov.–Dec. 1995; from Döös (1956)*

forecast length	24 h	48 h	72 h
RMS	54	120	175
corr. coeff.	0.85	0.74	0.62
highest and lowest value of the daily corr. coeff.	0.97–0.63	0.94–0.46	0.94–0.02

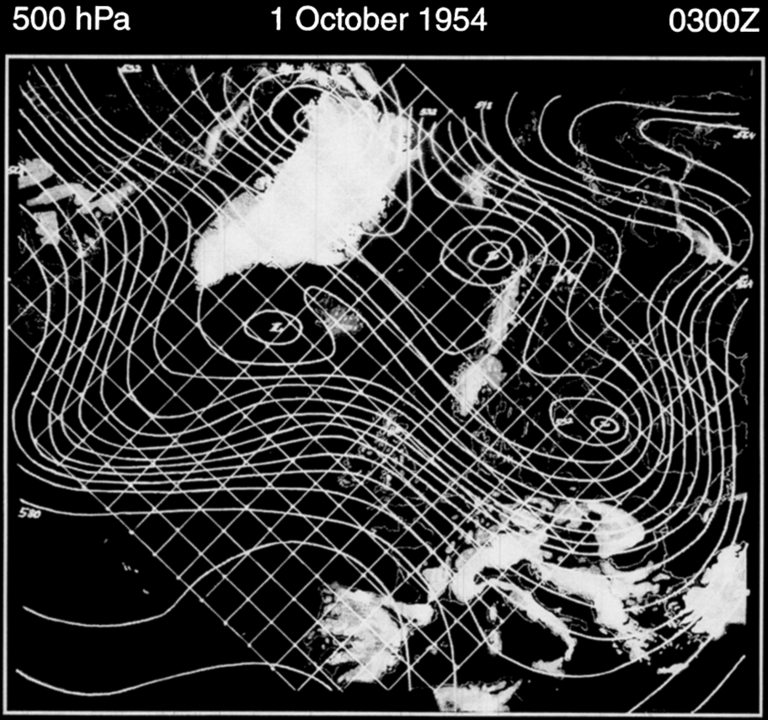


Fig. 2. 500 hPa geopotential height for 1 October 1954, 0300Z. The field is the initial state for a 72 h barotropic forecast.

with shallow water equation). The coefficients of the divergent term were determined empirically from observational data.

3. Improvements in numerical weather prediction

The accuracy of numerical prediction is a priori limited due to the inevitable growth of all errors in the initial state. This will set a limit to useful prediction of the order of a few weeks, even if we have a perfect model at our disposal. However, at the end of the 1950s, useful forecast skills were only of the order of 1–2 days, so it was evident that there was considerable scope for significant improvement in predictive skill.

Now, 40 years later, the situation is very different. As can be seen from Fig. 5, predictive skill has been extended by more than 5 days, the major improvement taking place as we went from limited area or hemispheric models, many of them based on the quasi-geostrophic equations to global comprehensive models. The speed of improvement has diminished, suggesting that we are slowly approaching the predictability limit. Another important point to highlight is the vast increase in forecast products now including many hundred forecast variables. One of the necessary conditions to this development has been the practically continuous evolution of increasingly powerful computers. Table 2 illustrates this by providing some key data for the barotropic model using the Swedish-built BESK computer and the ECMWF operational model in early 1998 on the Fujitsu VPP 700 supercomputer. The evolutionary development of more advanced models is shown in Table 3, which gives some data for the operational models at ECMWF during the last 15 years.

While more powerful computers have been necessary conditions for the improvements of NWP they are of course not sufficient conditions. By analysing the evolution over the past decade, we shall try to identify what could have been the

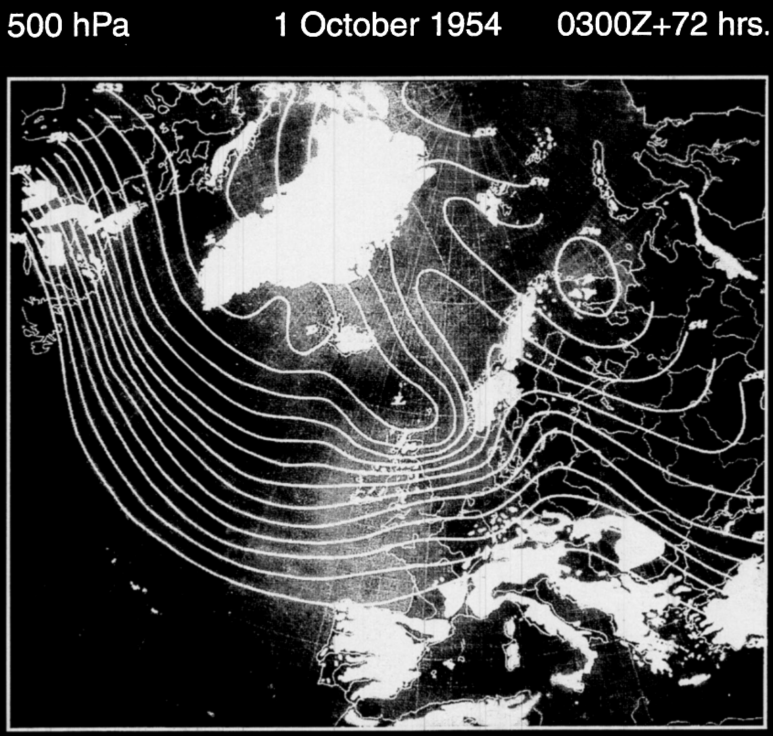


Fig. 3. 500 hPa geopotential height for a 72 h forecast from 1 October 1954, 0300Z.

Table 2. *Supercomputer and NWPP 1954–1998*

Comparison	Besk, Sweden	Fujitsu VPP 700 (16 proc)
	MISU/MVC (1954)	ECMWF (1998)
model	1 level, regional	31 level, global
parameter	1	~200
no. variables	1200	$3 \cdot 10^7$
calculations/day	$3 \cdot 10^7$ (40 bit fix number)	$5 \cdot 10^{12}$ (64 bit float number)
calculation time	25 min	6 min*
useful forecast	1 day	6–9 days

*With 115 proc. 66 s.

Examples of two typical models and associated computational requirements in 1954 and 1998 which at the actual time happens to be the most powerful computers used for operational numerical weather prediction in the world. The barotropic model was run jointly by the Meteorological Institute at Stockholm University (MISU) and the Swedish Air Force (MVC) at the Swedish-built BESK computer. The data from ECMWF refers to the Fujitsu VPP700 computer with 16 processors. The model data refer to the operational model used in the first half year of 1998, T213 with 31 vertical levels. Courtesy of ECMWF.

Table 3. *Development of spectral transform models at ECMWF during the period 1983–1998*

T63	17 levels	top 25 hPa
T106	19 levels	top 10 hPa
T213	31 levels	top 10 hPa
T319	31 levels	top 10 hPa
T319	60 levels	top 0.1 hPa

The latest version having 60 vertical levels with the top at 0.1 hPa has not yet been operationally implemented (October 98). At the same time as this evaluation has taken place the available computing power has increased by a factor of ca. 4000. Courtesy of ECMWF.

main contributing factors to the improvements in predictive skill and thereby try to consider the following possibilities.

(i) What improvements are due to better observations?
(ii) What is the importance of using a more general form of the basic equations?
(iii) What has been the effect of the numerical methods used?

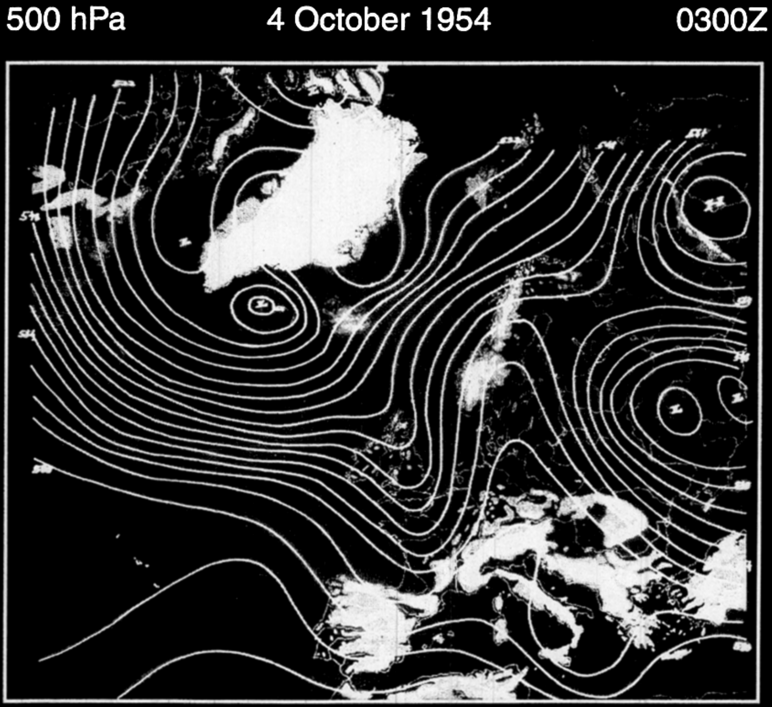


Fig. 4. The same as Fig. 2 but for the 4 October 1954, 0300Z.

(iv) What has been the effect of increased resolution?
(v) What has been the importance of an advanced physical parameterization?
(vi) What has been the role of data assimilation?

3.1. *Improvements due to better observation*

A major improvement in the global atmospheric observing system took place in the late 1970s in relation to the establishment of World Weather Watch and the first global weather experiment (FGGE) in 1979. For the first time a global atmospheric observing system was now set up. It consisted of a combination of surface observations, radiosondes and pilot balloons, aircraft reports, vertical temperature and moisture soundings from polar orbiting satellites and cloud drift winds from geostationary satellites. Other observational improvements included anchored and drifting ocean buoys and automatic surface stations in remote land areas. The global observing system made it practically feasible to consider operational numerical weather prediction on a global scale and was one of the justifications to establish an organisation such as ECMWF. Although the global observing system has continued to evolve since 1979, there have not been any major changes of relevance for numerical weather prediction. Satellite systems have gradually improved and new sensors, such as active scatterometers, have been employed. The network of buoys has also been improved, particularly in the tropical Pacific under the auspices of the TOGA program. However, the conventional system has simultaneously deteriorated, with a reduced number of radiosondes and surface observations from commercial ships as a consequence.

We are now in the very fortunate situation that, with reasonable confidence, we are able to quantitatively estimate the effect the different changes in the global observing system have had on the quality and accuracy of numerical weather predic-

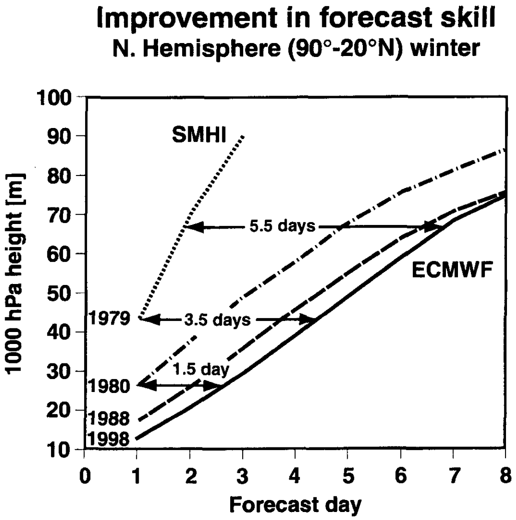


Fig. 5. The evolution in forecast skill for the northern hemisphere winter (90°–20°N) from 1979 until present. The dotted curve shows the error growth of the operational forecast model at SMHI (Swedish Meteorological and Hydrological Institute) in 1979. This was a 3-level quasi-geostrophic model typical of the prediction models used by many weather services in the 1960s and early 1970s. At day 1, the RMS error has reached over 40 m and grows to well over 90 m at day 3. The three additional curves show the results of the operational ECMWF model in 1980, 1988 and 1998. The ECMWF models were global in extent with different quality having undergone major improvements in modelling physics, resolution and in data assimilation. A 4.5 day forecast from ECMWF in the winter of 1998 has the same RMS error as a 1 day operational forecast at SMHI in 1979. Similarly, a 2.5 day forecast from ECMWF in 1998 has the same RMS error as a 1 day operational forecast at ECMWF in 1980. Note, however, that the speed of improvement is decreasing, suggesting that we are approaching the predictability limit. Courtesy of SMHI and ECMWF.

tion during the period since 1979. Both ECMWF and NCEP have in recent years undertaken so-called re-analysis experiments whereby an identical data-assimilation system has been used to re-assimilate the observations for a longer period of time. The main objective of such a re-analysis exercise (Bengtsson and Shukla, 1988) was primarily to obtain a homogenous data record for diagnostic studies on the general circulation of the atmosphere. The re-analysis is nothing but a repetition of the operational numerical weather prediction cycle which was carried out previously but using today's models and data-assimilation systems. ECMWF used the operational forecasting system available in 1994 (Gibson et al., 1997) and re-analysed the 15-year period 1979–1993. The model was run at a horizontal resolution of T106 with 31 vertical levels. The data assimilation was done in steps of 6 h. Every day a 24 h forecast and twice a month a 10 day forecast were made. NCEP procedure was rather similar (Kalnay et al., 1996), but they actually started their forecasts already in 1958 and they still continue to run this system on a daily basis. They have by now generated a homogenous record of global analyses for more than 40 years. The NCEP model employs a T62 resolution using 28 vertical levels. The two model- and data-assimilation systems for ECMWF and NCEP are qualitatively rather similar and represent the state of art of atmospheric modelling of the mid-1990s. Fig. 6 shows the forecast skill by the actual ECMWF operational model at the time compared to the forecast from the re-analysed data set. The lower curves show in the same way the number of days with "accurate forecasts" (taken as the number of days the model needs to reach an anomaly correlation of 0.95). The upper curves show the number of days with useful predictive skill (as expressed in the length of time it takes for the model to reach an anomaly correlation of 0.6) for the two series of forecasts. Although the number of forecast experiments is rather small (24 for every year), it is clearly seen that the operational forecasts have improved by about a day for the period 1979–93. Surprisingly enough this is true both for the short range as for the extended range. This may at first reflection indicate an improvement in the observing system and hence a more accurate initial state. However, if we then examine the same forecasts but now those from the re-analysed data, *the quality of the forecasts stays the same,* about 2.5 days for the "accurate forecasts" and some 7 days for the "useful forecasts" *for the whole period.* An examination of the NCEP forecasts shows more or less the same results. Table 4 shows the anomaly correlation for a 5 day forecast in 1972, 1980 and 1990 for NCEP and ECMWF. It is interesting to note that even in 1972 the forecast skill is only slightly smaller than in 1990 but significantly improved compared to the operational model at the time (0.7 instead of 0.45).

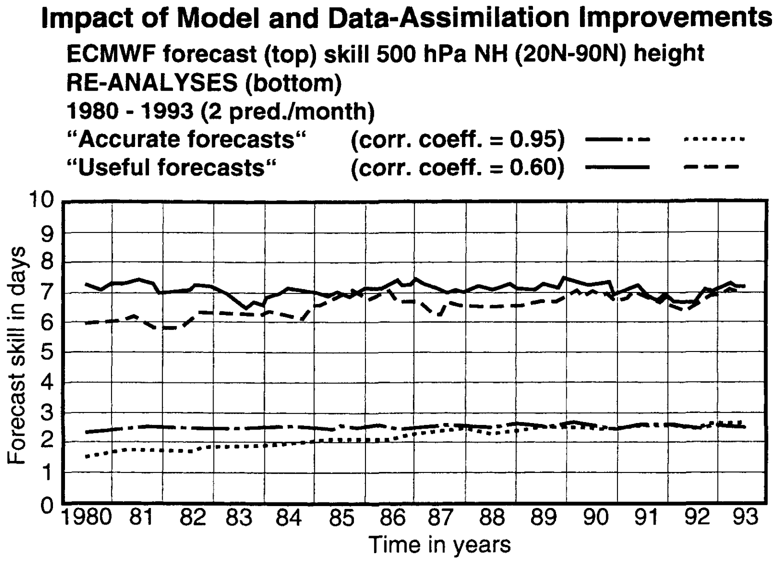


Fig. 6. Impact of model and data assimilation improvements at the northern hemisphere (20°N–90°N) for the 500 hPa geopotential height. The full line shows the length of "useful forecasts" (anomaly corr. coeff. = 0.60) from the re-analysis data at ECMWF and the dashed line the same from the actual forecasts at the time. The dash-dotted and the dotted lines, respectively, show the same for "accurate forecasts" (anomaly corr. coeff. = 0.95) Courtesy P. Kållberg and S. Uppala, ECMWF.

Table 4. *Impact of model and data-assimilation at NCEP and ECMWF for the 500 geopotential height at the northern hemisphere for the years 1972 (only NCEP), 1980 and 1990; courtesy E. Kalnay, P. Kållberg and S. Uppala*

Improved weather predictions (summary of reanalysis/forecast results); skill for 5-day weather prediction (anomally correlation: NH, 500 mb)

NCEP	1972	1980	1990
using models of that year	0.45	0.55	0.8
using models of 1995 (T62)	0.7	0.7	0.8
ECMWF	1972	1980	1995
using models of that year	—	0.65	0.8
using models of 1995 (T106)	—	0.8	0.8

We may conclude from this comparison that the improvement in forecast skill from 1979 is not due to better observations, but to improvements in the overall forecasting system including data-assimilation and the way observations are actually used in the forecasting system.

3.2. The importance of using a more general form of the basic equations

The early models were based on the quasi-geostrophic equations which essentially restricted numerical weather prediction to middle and high latitudes. Furthermore, the quasi-geostrophic application is not applicable for small-scale systems, although successful attempts were made (Bengtsson, 1973) to incorporate the small scale terms in the vorticity equation through an iterative procedure. However, the main reason why quasi-geostrophic modelling was abandoned was the development of more efficient ways of integrating the primitive equations through a semi-implicit integration (Robert, 1982) or through a split-explicit integration technique (Marchuk, 1974).

In both these cases, the same time-step could be used as for the quasi-geostrophic equations (in the case of the split-explicit for the bulk of the calculation) and hence the advantage in using the quasi-geostrophic models came to an end. The incorporation of physical processes, radiation, clouds, precipitation processes etc. are also more complicated to implement in the quasi-geostrophic models, and this was an additional reason not to use them any longer in NWP.

A problem which remained was the initialization of the primitive equation. Due to initial imbalances, gravity waves are generated which require a couple of days of integration before they are damped out. This does not really have any effect on the quality of the forecast beyond a couple of days, but creates serious difficulties in data-assimilation where short-range forecasts (3–6 h) are used as a first guess of the next assimilation step. The original proposal by Machenhauer (1977) to set the initial tendency of the gravity waves to zero effectively solved this problem.

3.3. *The effect of the numerical methods used*

At most operational and research institutes the spectral transform technique has replaced finite difference except for limited area modelling. The advantage of the spectral transform technique is essentially the care and clarity by which one can constrain the model to a specific resolution and thus easily, for example, incorporate a resolution-dependent parameterization of geostrophic turbulence (Laursen and Eliasen, 1989) and the ease by which one could handle the semi-implicit integration and flow over the poles. The incorporation of physical parameterization and calculation of non-linear interaction terms was furthermore done in grid point space, and was thus handled in the same way as in the grid point models.

Jarraud and Girard (1984) undertook a detailed intercomparison between a grid point model and a series of spectral transform models. The result was slightly in favour of the spectral model (for the same amount of computing time), but generally insignificant for the same grid size as the Gaussian grid used in the spectral transform model.

A special problem occurs when using high vertical resolution in combination with steep terrain. This requires the use of short time steps to satisfy the CFL-criteria. The use of semi-Lagrangian approach (Robert et al., 1985) or implicit solution in the vertical generally leads to more efficient integration schemes making it possible to use a higher horizontal resolution at the same computational cost. ECMWF was able to increase the timestep of its present operational model (T319 and 31 vertical levels) from 2 min to 20 min at the same forecast quality by using a semi-Lagrangian approach.

3.4. *The effect of increased resolution*

As computer capabilities have increased, so have the possibilities to use higher horizontal and vertical resolution. As has been discussed by Wiin-Nielsen (1985), the high resolution is not required for resolving the three-dimensional flow as such including the phase-speed of synoptic scale system, but it is required in order to resolve the non-linear interactions between small scale systems and the large scale flow. Series of experiments to predict intense extra-tropical storms show that a resolution around 1° or T106 is required to predict such development accurately (Bengtsson, 1990, Figs. 11.13, 11.20).

Bengtsson (1981) has also demonstrated the importance of horizontal resolution for the prediction of blocking. An example of blocking prediction can be seen in Fig. 7, where we show the results from forecasts of blocking pattern over Scandinavia with three different horizontal resolutions, T42, T63 and T106. In spite of the fact that we depict the 5-day mean the effect of small scale systems has played an important part in creating the characteristic omega pattern of the blocking. This pattern can only be seen in the T106 experiment.

Another important example of resolution concerns the ability to resolve lee-wave cyclogenesis where it is of primary importance to properly resolve orographic obstacles. This will normally require a horizontal resolution of 1° or less. (Tibaldi et al., 1990).

Finally, the ability to predict hurricanes and taifuns with large scale models has gradually evolved into an operationally feasible approach. To resolve the very intense inner core of a tropical storm and associated areas of intense precipitation we will probably require a resolution of around 10 km or less, while the more large scale features are resolvable already at a resolution of 1° or so. Fig. 8 shows an example of a prediction of hurricane Hugo as obtained by the operational ECMWF models at T106 resolution.

3.5. *The importance of an advanced physical parameterization*

A major challenge in modelling of climate and weather is the handling of the physical processes

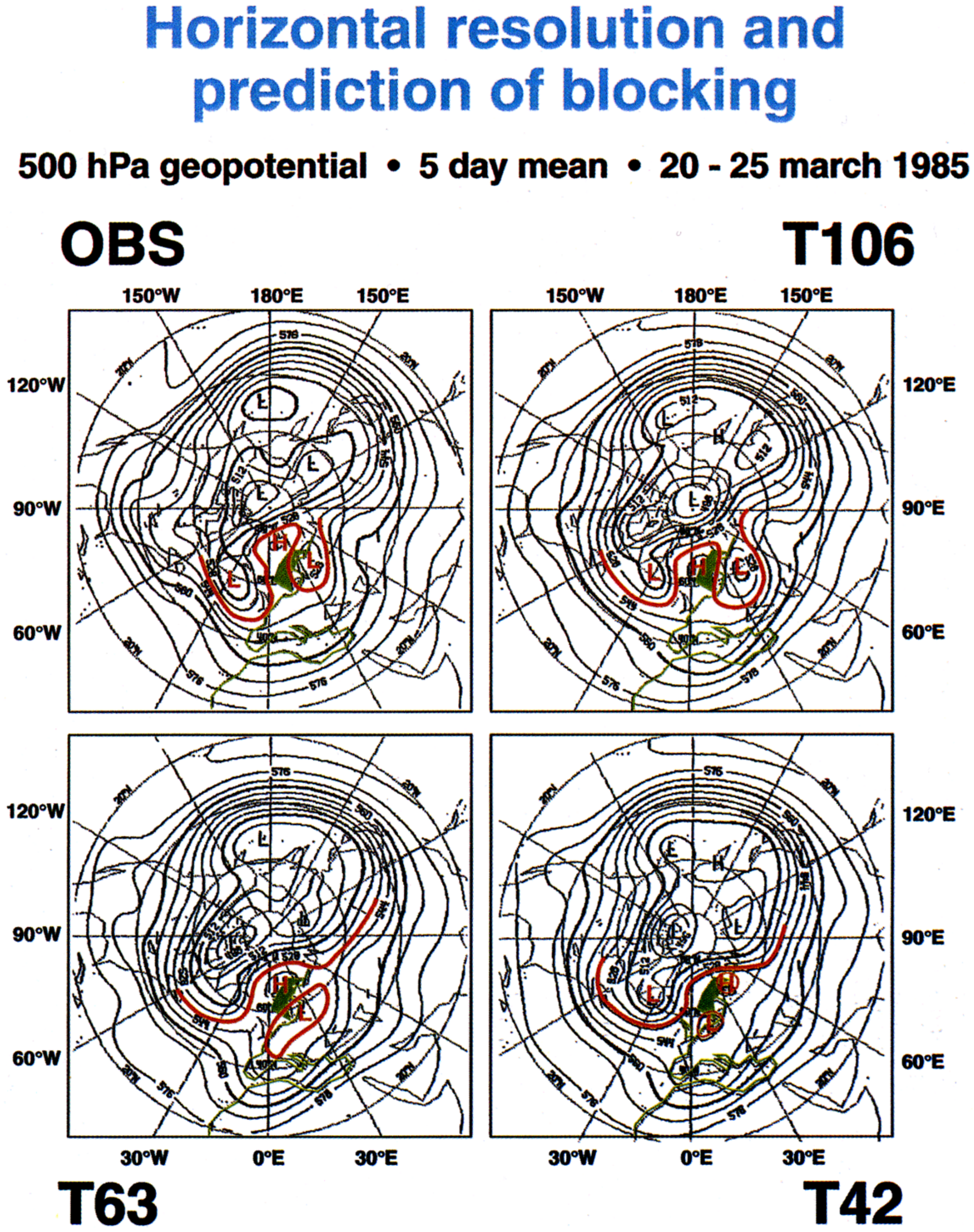


Fig. 7. Horizontal resolution and blocking as observed (20–25 March 1985) and predicted at three different resolutions T106, T63 and T42. 5-day means of the 500 hPa geopotential height. Resolution 8 dm.

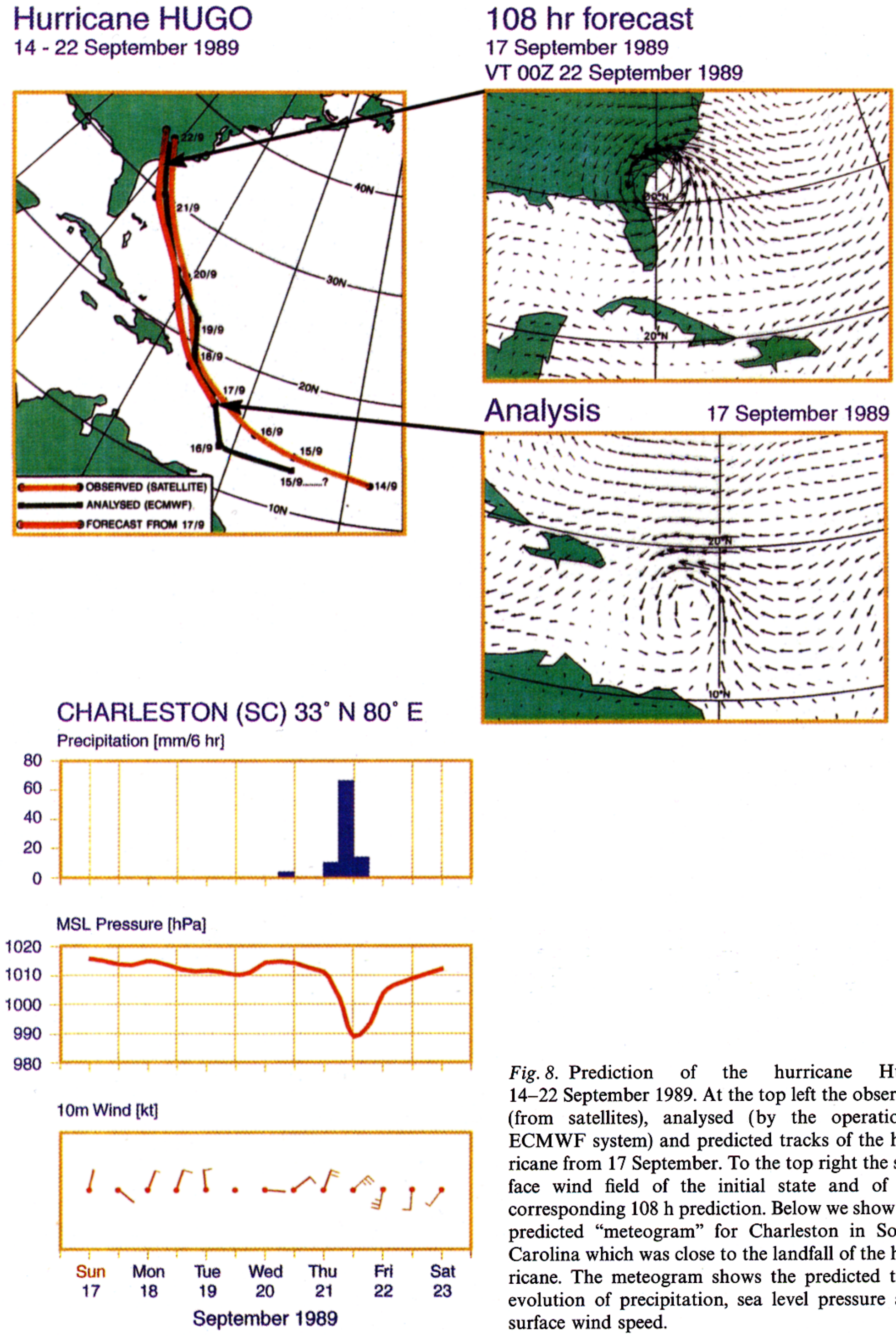


Fig. 8. Prediction of the hurricane Hugo 14–22 September 1989. At the top left the observed (from satellites), analysed (by the operational ECMWF system) and predicted tracks of the hurricane from 17 September. To the top right the surface wind field of the initial state and of the corresponding 108 h prediction. Below we show the predicted "meteogram" for Charleston in South Carolina which was close to the landfall of the hurricane. The meteogram shows the predicted time evolution of precipitation, sea level pressure and surface wind speed.

which determine the sources and sinks of heat, momentum and water. The relative importance of different physical processes depends on the time scale and whether we are dealing with extra-tropical or tropical systems.

Generally it is very difficult, if at all possible, to identify the relative importance of individual physical processes due to numerous feedback mechanisms which are activated when a certain model modification is introduced. This modifies rapidly the initial change, and there are many examples when an obvious improvement in representing, say, deep convection may interact incorrectly with the cloud scheme and hence generate an overall deterioration of the forecast. It may therefore be more important to apply some general principles in improving the physical parameterization of models for weather and climate.

Firstly, it is important that the models are capable of generating a correct time-space spectrum of the large scale atmospheric flow. This requires a sufficient resolution and an accurate treatment of basic processes such as barotropic and baroclinic instability, realistic boundary conditions and an accurate representation of orography and land-sea contrasts.

Secondly, a consistent treatment of the sources and sinks for heat, momentum and water is needed. This will require internal consistency in the parameterization of sub-grid scale processes.

Thirdly, it must be recognized that the handling of sub-grid scale parameterization is strongly empirical and can only partly be done from first principle. The exchange of heat, momentum and water between the atmosphere and a highly heterogeneous land-surface, which even at very high resolutions are averaged over regions of several thousand km^2 cannot be represented by single numbers. Similar representation difficulties concern the model's representation of the highly heterogeneous three-dimensional cloud distribution.

Model improvements thus include careful tuning of the different constants and coefficients in the equations within the range of empirical data. The final decision which values to select requires long series of systematic numerical experiments.

Between 1980 and 1988 ECMWF reduced the systematic error variance in 10 day forecasts by more than a factor of 5 (Bengtsson, 1990). Typically, this improvement could not be traced to any specific causes but was the result of an overall development of the forecasting system at large.

3.6. The rôle of data-assimilation

As was discussed in Subsection 3.1, it was clear from the forecast experiments from the re-analysis data sets that a significant improvement also was evident in the short range prediction. This may appear somewhat puzzling since, apparently, there have been no visible improvements in the observational data. Following Dalcher and Kalnay (1987), let us assume that the forecast error variance can be expressed by an error growth equation of the form

$$\frac{\mathrm{d}E}{\mathrm{d}T} = (\alpha E + S)\left(1 - \frac{E}{E_\infty}\right), \qquad (1)$$

where E is the forecast error variance of the 500 hPa geopotential height field, α is daily rate of growth of the forecast error variance, S is the amount of error variance introduced by model deficiencies in one day and E_∞ is the asymptotic value of the error variance.

In order to examine the effect of data-assimilation on the forecast error, we will follow a procedure proposed by Leith (1983). The initial state is usually determined by two virtually independent assessments: the first guess and the actual observations. If two such independent determinations, z_1 and z_2 of, say, the 500 hPa geopotential height with error variances E_1 and E_2 are combined into a better final determination $(E_2 z_1 + E_1 z_2)/(E_1 + E_2)$ then the inverses of the variances, which measure information content, are summed:

$$E^{-1} = E_1^{-1} + E_2^{-1}. \qquad (2)$$

This general statistical principle is used in the objective analysis. By combining eqns. (1) and (2), it is now possible to calculate the initial error of a particular forecast as a result of a series of data-assimilation steps.

The overall principle is explained in Fig. 9. Let us assume that we start the data-assimilation cycle from a given initial state with initial error variance E_0. This first initial state can consist of an analysis using all available observations at $t = t_0$ using, say, climatology as the first guess.

The initial error variance grows in time during the first data-assimilation step in accordance with

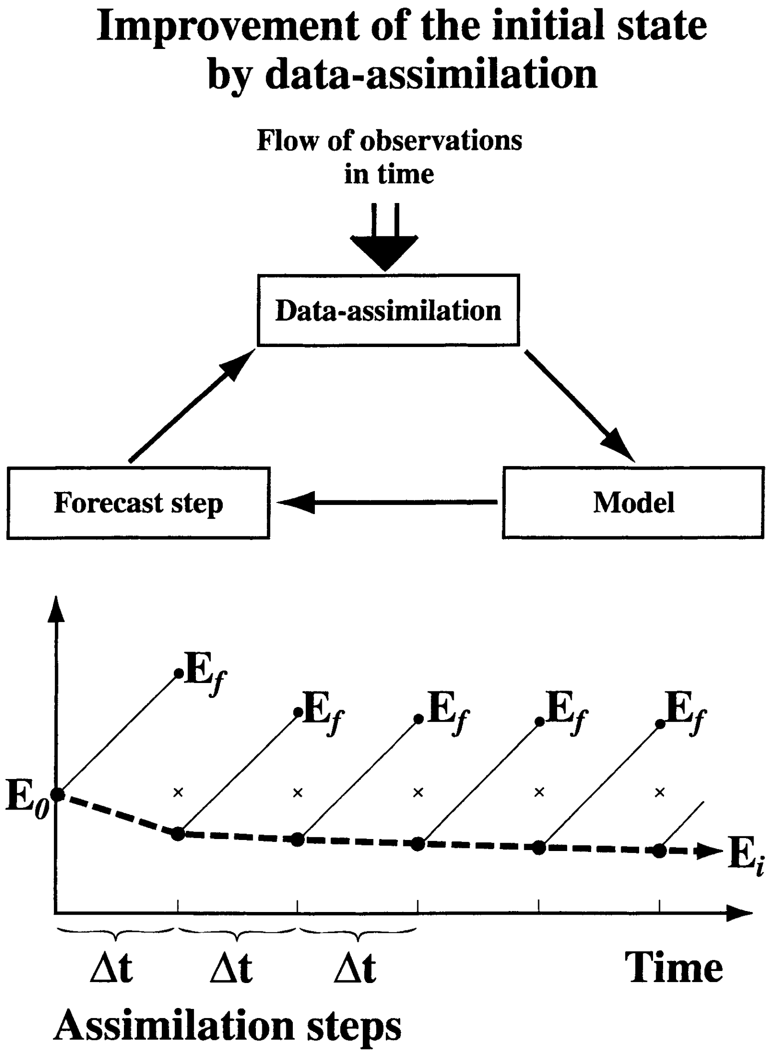


Fig. 9. The successive reduction of the error of the initial state as a function of a data-assimilation process. E_0 is the observational error variance. E_f indicates the error variance at each data-assimilation step. E_n is the asymptotic error which will develop after a series of data-assimilation steps.

eq. (1). After Δt hours, a new initial state is formed by combining the new observations at this time with the first guess obtained from the prediction eq. (1). The sum of the inverses of the two variances provides the new estimated initial error variance E_n in accordance with eq. (2).

Following Leith we can easily estimate the *asymptotic* value of E as it will emerge after a series of data-assimilation steps. In order to simplify the analytic solution, we will replace eq. (1) by

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \alpha E + S. \tag{3}$$

This simplification is justified since we will only use eq. (1) for short time scales when the ratio of E/E_∞ is very small. We shall now let Δt be a fixed time interval of the data-assimilation cycle. We shall also let E_n be the error variance after the n'th data-assimilation cycle.

Simple algebraic calculations easily show

$$E_n = E_0\left[ab + \left\{\frac{a(b-1)}{2}\right\}^2\right]^{1/2} - \frac{E_0}{2}\{a(b-1)\}, \tag{4}$$

where E_n is the asymptotic error variance of the initial state, $a = \alpha\Delta t/(1+\alpha\Delta t)$ and $b = S/\alpha E_0$.

It is easily seen that E_n is proportional to E_0 (which can be interpreted as a measure of the variance error of the observational state at a given time). E_n further diminishes when Δt and S are reduced, which means that a higher data-assimilation frequency is beneficial as well as an accurate model (small value of S).

In Fig. 10 we show how the error variance reduction ratio E_n/E_∞ varies as a function of $(S/E_0)^{1/2}$ for different values of the data-assimilation frequency of 3, 6, 12 and 24 h. α, the inherent error growth, has been put equal to 0.55 corresponding to an error variance doubling time of 1.25 days. In order to simplify the interpretation, Fig. 10 presents the result in terms of relative *error reduction* instead of *error variance reduction.*

Having so determined the asymptotic value E_n after a large number of assimilation steps we can now use eq. (1) to calculate predictive skills in days as a function of S and Δt. We will here simply assume that the limit of useful predictive skill is reached when E is equal to $0.75 \times E_\infty$. Figs. 11, 12 show the predictive skill in days for the 500 hPa height for the northern and southern hemisphere.

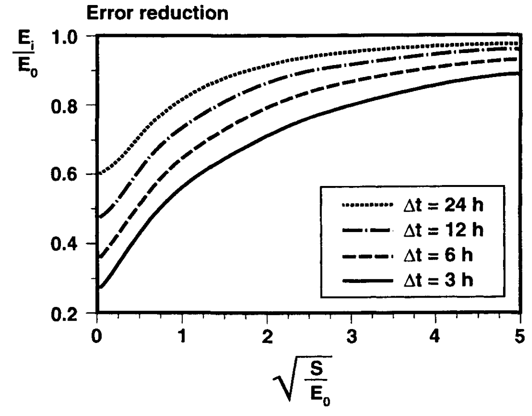


Fig. 10. Assimilation equilibrium error reduction as a function of the ratio between the model error source rate and the observational error. Note the dependence of the data-assimilation frequency on the error reduction. For further information see text.

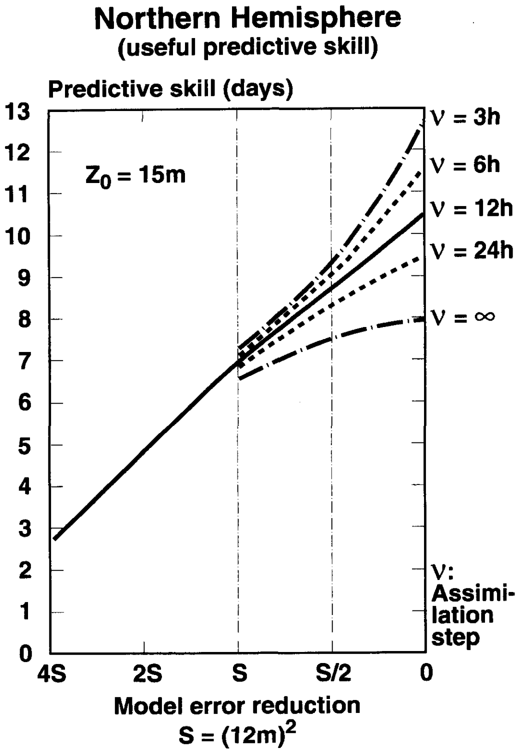


Fig. 11. Useful predictive skill for the northern hemisphere as expresses as $0.75E/E_\infty$ where E is the actual variance of the forecast and E_∞ the daily climate variance. The predictive skill is given as a function of the actual forecast skill S. For further information see text.

Bengtsson (1989) has estimated the value of S from the actual error of the ECMWF operational forecasts for the 500 hPa from the two hemispheres and concluded that the model error growth was reduced from $S = (18\text{ m})^2$/day in the winter 1980/81 to $S = (12\text{ m})^2$/day in the winter 1987, meaning the S has virtually halved during this period. Assuming a perfect data assimilation this would correspond to an improvement of prediction skill by about 1.25 days. It is interesting to note that the effect of a higher frequency of data assimilation will only become efficient at a further reduction of the model error variance S. At the southern hemisphere, the improvement in predictive skill is about the same, but here a higher frequency of data-assimilation will have impact due to larger errors in the observing system.

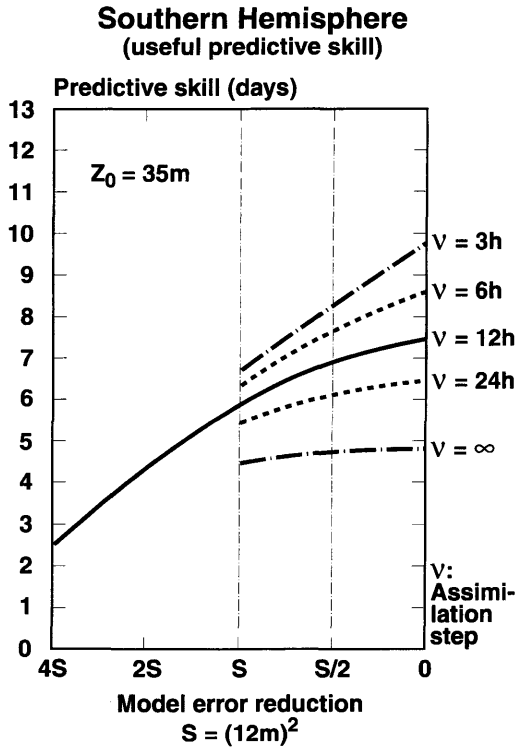


Fig. 12. The same as Fig. 11 but for the southern hemisphere. Note that the only difference between Figs. 11, 12 is the estimated observational accuracy (expressed as the analysis error variance with climatology as the first guess.

We may conclude from this simple analysis that an efficient data-assimilation is indeed essential in order to fully realize the improvements in the model. It further suggests that additional forecast improvements are feasible even with the same model, if the observing systems are able to provide observations at higher time frequencies. The relative impact of such observations is likely to be more quickly realized the more accurate the data-assimilation model is. Although the analysis of the last section indicated that there is still scope for improving predictive skill from the present 7–8 days with another 2–3 days, the experience from practical forecasting suggests that this may be somewhat optimistic. We may conclude that our simple analysis indicates an upper bound of the predictive skill, since it assumes a perfect assimilation procedure and statistically independent errors in the first guess and in the observations. The error growth corresponds to an approximate doubling of the error variance in 1.25 days. This may be a reasonable global or hemispheric mean value, but clearly too small in active baroclinic zones where error growth is much faster. Furthermore, the accuracy of the initial state also varies since the global observing system is far from homogeneous.

The large variability in forecast skill from day to day is demonstrated from Fig. 13, which shows the daily ECMWF forecasts at 1000 hPa for Europe for the period January–March 1998 in the form of a percentage cumulative frequency distribution. In 50% of the cases there is useful predictive skill (0.6 of the anomaly correlation) 7 days ahead, while for 5% of the worst cases the limit is reached only after 4 days. 5% of the best cases, on the other hand, have useful skill beyond 10 days. A comparison with another period in January–March 1998 shows that the forecast quality spread is about the same, but with a superimposed overall improvement which has taken place between 1988 and 1998. The large variation in predictive skill from day to day especially, for smaller regions, are characteristic of the atmosphere. It is simply due to the fact that the growth of errors, for example, is much faster in active baroclinic zones than in the inner regions of large anticyclones. The issue which we have to address is instead whether it is feasible to assess predictive skill in advance and hence obtain information of the expected forecast reliability. This will of course not be as useful as a really improved forecast, but could nevertheless for many applications provide valuable information. In the next section, I will present the approach which has been taken to address this issue and the principal value this will have for the future in weather and climate modelling and prediction.

4. Ensemble prediction

As we have seen in Subsection 3.6, the accuracy of a given forecast depends on the internal error growth of the model, the model accuracy and the errors which exist in the initial state. Ideally we would therefore wish to be able not only to provide an accurate forecast but also its potential accuracy as a function of area, the forecast's length, and this for every individual forecast. Furthermore, it would

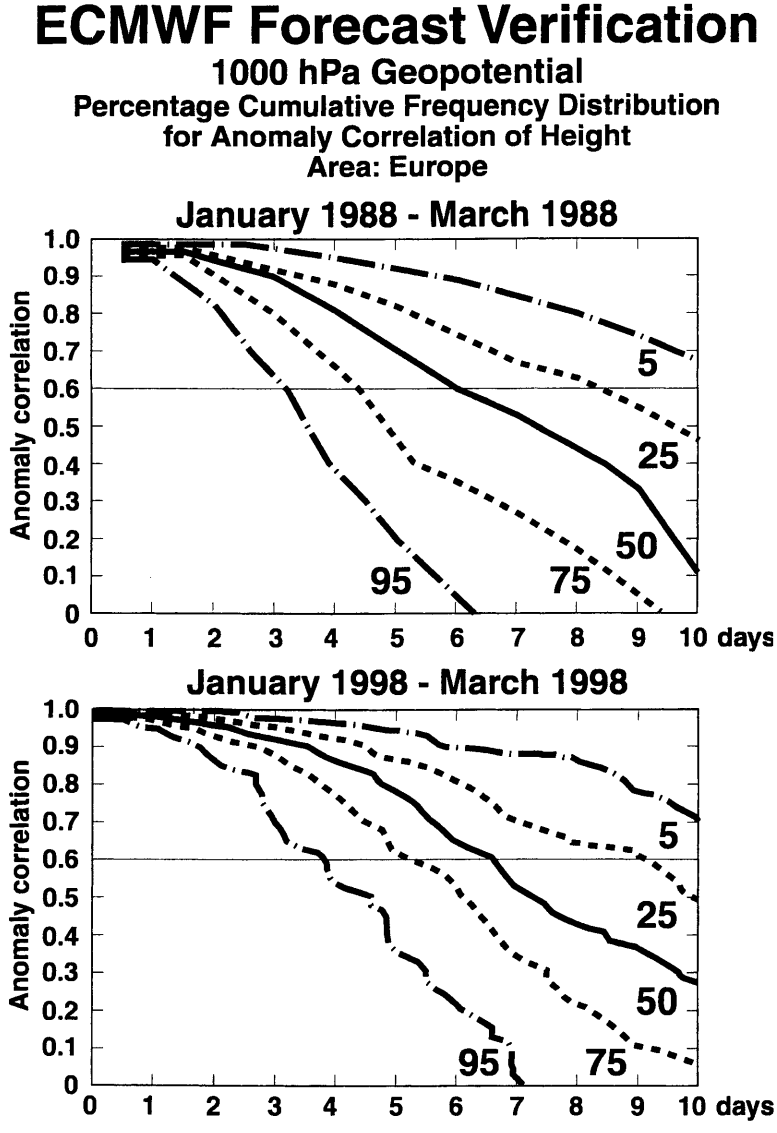


Fig. 13. Above: anomaly correlation for the ECMWF 1000 hPa geopotential height forecasts for January–March 1988 over Europe. The different curves show the percentage of the cumulative frequency distribution. 95 means that 95% of all the daily forecasts are better and 5 means that only 5% of the forecasts are better. The line for 50% is close to the average result. Below: The same, but for January–March 1998. Courtesy of ECMWF; cf. Table 1.

also be valuable to indicate the probability of certain extreme events, such as intense cyclones or situations of extreme temperatures and precipitation. Meteorological forecasters are under great pressure to provide such information, but so far the only assistance they have had has been access to forecasts from different centres. It is certainly very important to provide such probability estimates by which forecasters in a more systematic way can assess the probability of certain weather developments and thereby a measure of the reliability of a given forecast.

The first attempt to address this problem has been to calculate how errors of the initial state are likely to grow in particular meteorological situations.

This can be accomplished by undertaking series of forecasts with the same model but starting the forecasts from slightly different states. Small perturbations are added to the reference model, with amplitudes selected to be within the accuracy of the initial state. The practical issue to face is what error patterns should be assigned and in what areas they should be inserted. Another question to decide is the size of the ensemble and how best to present the results of a relatively large number of forecasts to the users. The ECMWF has made major contributions to this novel forecasting technique and has over the last years also developed an operational system for ensemble prediction (Molteni et al., 1996).

It is clear that the ensemble size has to be much smaller than the number of possible perturbations (which is related to the number of degrees of freedom given by the model equations) and consequently will suffer from severe sampling errors. The strategy chosen by ECMWF has therefore been to select those perturbations that are most dynamically active for explicit integration in the ensemble. A calculation of these dynamically active perturbations is made using the forward and adjoined tangent equations (Courtier et al., 1991), linearised about a forecast phase-space trajectory taken from the operational analysis. This is presently too demanding to calculate from the high resolution operational model and instead the calculation is done at a T42 resolution with 19 vertical levels.

The perturbations selected, called singular or optimal vectors (Buizza and Palmer, 1995), are those that grow most rapidly over the first 48 h of the forecast, when error growth can be adequately described by linear theory. The 25 fastest growing modes are selected from which 25 perturbations are calculated by taking linear combinations of these singular vectors. The amplitude of the perturbations is chosen to be consistent with the local amplitude of a typical analysis error. The 25 perturbations are then added to and subtracted from the operational analysis to produce 50 perturbed initial conditions.

Numerical experiments (Palmer et al., 1998) show that the ensemble forecasting system is capable of generating a broad range of skill suggesting that the daily forecast variability in Fig. 13 essentially can be explained as the result of errors in the initial state. Additional variability would presumably be added if empirical parameters in the parameterization (which may be equally insufficiently known) would be perturbed in an analogue way.

The experience with the ensemble forecasting system has demonstrated that future forecasting systems need to incorporate this aspect. In principle any weather and climate predictions are probabilistic forecasts, since a calculation will contain a certain amount of unpredictable noise. To estimate this and to identify the actual forecast signal and this in probabilistic terms, is expected to become a standard procedure in future prediction of climate and weather.

5. Discussion

We have in this paper tried to address some of the major developments in large-scale numerical weather prediction over the last decades. We have not specifically discussed the development of limited area models which has led to excellent improvements of short range prediction but concentrated on the improvements of the global models. It has been found that useful predictive skill over most parts of the northern hemisphere extratropics has been extended to presently more than a week from a day or two in the 1950s and about three days through the 1960s and 1970s. The main factors have been the improvement of the observing system until the late 1970s but even more so improvements in model physics, dynamics and numerical methods. From the late 1970s, forecast improvements are essentially due to better forecasting systems. No single factor can be clearly identified, but the increased resolution appears to have created the necessary conditions for a model improvement. The importance of data assimilation has been fundamental, since without that it would not have been possible to fully *realize* the model improvements. The reason is that with the combination of an accurate model and an advanced data-assimilation scheme it is possible to significantly *reduce* the error of the initial state. Numerical experiments using today's models and data-assimilation system show that past forecasts could be significantly improved even by only using observations which were available at the time. This justifies the ongoing efforts to reanalyse past observations by today's forecasting systems and demonstrates clearly the great importance of using advanced systems for data-assimilation in order to realize the value of global observations.

Forecast skill varies considerably from day to day, suggesting that the length of useful forecasts can vary by more than 6 days for 90% of the predictions. The remaining 10% will have an even larger span. Ensemble prediction studies at ECMWF suggest that the major part of this variability is related to the inaccuracy of the initial state and the way these inaccuracies are exposed to rapidly growing perturbations.

This fact is the reason why ensemble prediction has become a central corner stone in weather prediction. Weather predictions as well as climate predictions will always be incorrect due to the inherent unpredictability of the atmosphere and any users of weather forecasts need a measure of the accuracy as well as a measure of the probability to estimate the potential risks of certain events which can be both damaging and dangerous. Prediction of intense tropical and extra-tropical storms is here such an example.

6. Anticipated development of NWP in the next 10 years

How may numerical weather prediction evolve in the future? It appears that the computational capabilities continue to evolve as they have done almost uninterruptedly since the days of the early experiments in Stockholm. This indicates that the computational speed will increase by a factor of ten for every fifth or sixth year at least a few times more. This means that computers capable of carrying out more than 10^{12} operations/s (1 Teraflop) will be generally available and probably also affordable for medium size organizations 5–10 years from now.

Concerning the improvements of the global observing system, we may be less optimistic. Due to economical and political problems in many parts of the world the present radiosonde network and other parts of meteorological observing systems are deteriorating. The free exchange of data is also tending to be more restricted due to commercial activities of many meteorological services. The improvements of observing systems in the future is likely to come through the development of more advanced satellite based observing systems. However, this is not a fast process due to the complexity of the present satellite system and the comparatively long time it takes to plan, develop and put such systems into operation. It is hoped that this process could be accelerated and that user-oriented systems could be developed in parallel and not be done after many years due to insufficient resources and incomplete planning.

It is unavoidable that exchange of NWP know-how, software and probably also exchange of observational data will become more widely accessible than at present thanks to Internet and the ongoing globalisation. This will accelerate the development and encourage and stimulate new ideas.

Climate models and NWP models will converge as already happens and prediction will be extended to include environmental prediction in general. This will mean the development and use of coupled ocean/atmosphere-models. This development will accelerate the establishment of agencies having an overall responsibility for environmental prediction. Some meteorological services are already in the process of such a transformation. Ensemble prediction will become a necessary tool, in particular for the longer predictions.

Due to falling computer costs and the efficient exchange of know-how and observational data even smaller NWP units will be capable of undertaking advanced modelling and forecasting. How to best organise operational NWP in a region, say like Europe, will thus be a rather delicate issue. Perhaps we will see a combination of very high resolution global models and regional models over small areas capable of predicting local weather.

The operational agencies such as ECMWF will gradually extend their activities to include interannual predictions and later, if found feasible, decadal prediction.

Forecast services will become strongly user-oriented. Internet provides here extraordinary possibilities. The question what should be openly available to the public and what should be commercially available is not yet clearly defined. The present system with vastly different procedures, as between Europe and USA, must urgently find a solution in order to avoid damages to the international cooperation.

7. Acknowledgement

The author wishes to acknowledge staff members of ECMWF who have been most helpful in providing information on the performance of the ECMWF forecasting systems. The author also wishes to thank Dr. E. Kalnay and Dr. J. Shukla for providing information on the re-analysis system at NCEP. Professor Bo Döös was very helpful to provide information on the performance of the forcast experiments at the International Meteorological Institute in Stockholm in the 1950s. Ms. Kornelia Müller and Mr. Norbert Noreiks have been helpful as always in providing their professional services.

REFERENCES

Bengtsson, L. 1973. *A three-parameter model for limited area forecasting*. Environmental Research Facility, Naval Postgraduate School, Monterey, California. Technical Paper No. 73, 88 p.

Bengtsson, L. 1981. Numerical prediction of atmospheric blocking. A case study. *Tellus* **33**, 19–42.

Bengtsson, L. and J. Shukla, 1988. Integration of space and in situ observations to study global climate change. *Bul. Amer. Met. Soc.* **69**, 1130–1143.

Bengtsson, L. 1989. On the growth of errors in data assimilation systems. In: *Data assimilation and the use of satellite data.* European Centre for Medium Range Weather Forecasts, Seminar, Reading, 5–9 September 1988, vol. 1, pp. 3–16.

Bengtsson, L. 1990. Advances in the numerical prediction of the atmospheric circulation in the extra-tropics. Extratropical Cyclones. The Erik Palmén Memorial Volume, ed. C. Newton and E. Holopainen. *Am. Met. Soc.* **107–127**, 194–217.

Berghtorsson, P., B. R. Döös, S. Fryklund, O. Haug and R. Lindquist, 1955. Routine forecasting with the barotropic mode. *Tellus* **7**, 272–274.

Berghtorsson, P. and B. R. Döös, 1955. Numerical weather map analysis. *Tellus* **7**, 329–340.

Bolin, B. 1999. Carl-Gustaf Rossby. The Stockholm period 1947–1957. *Tellus* **51A/B**, this issue.

Buizza, R. and T. N. Palmer, 1995. The singular vector structure of the atmospheric general circulation. *J. Atmos. Sci.* **52**, 1434–1456.

Courtier, P., E. Anderson, W. Heckley, G. Kelly, J. Pailleux, F. Rabier, J.-N. Thépaut, P. Undén, D. Vasiljevic, C. Cardinali, J. Eyre, M. Hamrud, J. Haseler, A. Hollingsworth, A. McNally and A. Stoffelen, 1993. *Variational assimilation at ECMWF.* ECMWF Research Department Technical Memorandum 194.

Dalcher, A. and E. Kalnay, 1987. Error growth and predictability in operational ECMWF forecasts. *Tellus* **39A**, 474–491.

Döös, B. R. 1956. Automation of 500 mb forecasts through successive numerical map analyses. *Tellus* **8**, 76–81.

Gibson, J. K., P. Kållberg, S. Uppala, A. Hernandez, A. Nomura and E. Serrano, 1997. *ECMWF reanalysis project report series* (1). *ERA description.* European Centre for Medium-Range Weather Forecasts, Reading/UK, 72 pp.

Jarraud, M. and C. Girard, 1984. Extensive quasi-operational comparison between a spectral and a grid-point model. In: Numerical methods for weather prediction. European Centre for Medium Range Weather Forecasts, Seminar, Reading, 5–9 September 1983, vol. 2, pp. 61–111.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa and R. Reynolds, 1996. The NCEP/NCAR 40-year reanalysis project. *Bul. Amer. Met. Soc.* **77**, 437–471.

Laursen, L. and E. Eliasen, 1989. Effects of the damping mechanism in an atmospheric general circulation model. *Tellus* **41A**, 385–400.

Leith, C. E. 1983. Predictability in theory and practice. In: *Large-scale dynamical processes in the atmosphere*, B. J. Hoskins and R. P. Pearce (eds.). Academic Press, 365–383.

Machenhauer, B. 1977. On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Contributions to Atmospheric Physics* **50**, 253–271.

Marchuk, G. I. 1974. *Numerical methods in weather prediction.* N.Y., Academic Press, 277 pp.

Molteni, F., R. Buizza, A. Lanzinger and T. N. Palmer, 1996. Potential use of the ECMWF Ensemble Prediction System: methodology and validation. *Q. J. R. Meteorol. Soc.* **122**, 73–119.

Palmer, T. N., R. Buizza and F. Lalaurette, 1998. *Performance of the ECMWF ensemble prediction system.* ECMWF Workshop on Predictability, 1998. In press. Available from ECMWF.

Robert, A. J. 1982. A semi-Lagrangian and semi-implicit numerical integration scheme for the primitive meteorological equations. *J. Meteor. Soc. Japan* **60**, 319–324.

Robert, A. J., Yee, T.-L. and H. Ritchie, 1985. A semi-Lagrangian and semi-implicit numerical integration scheme for multi-level atmospheric models. *Mon. Wea. Rev.* **113**, 388–394.

Rossby, C. G. 1957. Aktuella meteorologiska problem.

Svensk Naturvetenskap 1956. Swedish Natural Science Research Council. Translated into English: *Current problems in meteorology*. The Rossby Memorial Volume. Rockefeller Institute Press in association with Oxford University Press, New York 1959.

Staff Members, Inst. of Meteorology, Univ. of Stockholm, 1954: Results of forecasting with the barotropic model on an electronic computer (BESK). *Tellus* **6**, 139–149.

Tibaldi, S., A. Buzzi and A. Speranza, 1990. Orographic cyclogenesis. Extratropical cyclones. The Erik Palmén Memorial Volume, eds. C. Newton and E. Holopainen. *Am. Met. Soc.* 107–127.

Wiin-Nielsen, A. 1985. *The scientific problem of medium-range weather prediction. Medium range weather forecasts — the first 10 years.* Proceedings of a Seminar to commemorate the 10th anniversary of ECMWF. 22 November 1985, 31–53.

Wiin-Nielsen, A. 1991. The birth of numerical weather prediction. *Tellus* **43**, AB, 36–52.