REVIEW OF RECENT PROGRESS MADE IN MEDIUM RANGE WEATHER FORECASTING

Lennart Bengtsson
ECMWF

1. INTRODUCTION

The first ECMWF Seminar in 1975 (ECMWF, 1975) considered the scientific foundation of medium range weather forecasts. It may be of interest as a part of this lecture, to review some of the ideas and opinions expressed during this seminar.

It was generally recognised, and still is, that the mediumm range forecasting problem, defined as predictions within a time scale from 2 days to 2 weeks, was both a data- and a model problem. For this reason substantial efforts and resources were to be allocated to global analysis and initialization in addition to the modelling work. There was also a consensus of opinions that "the brute force approach" namely to use a high resolution general circulation type of model, was the only feasible method, in particular with respect to the Centre's time schedule and operational objective.

The practical experience in medium range weather prediction was very limited at this time and only a limited set of real data predictions of 2 weeks length had at that time been carried out at GFDL, Princeton (Miyakoda et al, 1972). As we will see later, these forecasts indicated in the average positive skill beyond a week, although the level of useful skill was small after 3-4 days. Nevertheless, these integrations represent a remarkable achievement and served an important catalytic function. They stimulated the planning and later the successful implementation of the First GARP Global Experiment in 1978 - 79 and they constituted a very important impetus in setting up the ECMWF. They also set the benchmark for future efforts in medium range weather forecasting.

2. PROBLEM AREAS IN NWP IN 1974/75

There were a large number of obstacles to medium range weather prediction in 1975. Some of them still remain but some have been overcome at least partially. Table 1 lists a selection of the most serious problems as seen 6 years ago. They are divided into five different areas.

The data assimilation (analysis and initialization) has been discussed extensively by Bengtsson (1975) and we will here only highlight a few facts. The large increase of non-synoptic quantitative observations, mainly from satellites, which started in the 1970's necessitated a continuous assimilation or alternatively a discontinuous assimilation with high temporal resolution (at least 6 hours). However, if this was going to be possible, ways were to be found to suppress the high frequency noise created by lack of a proper initialization before the next set of observations were inserted into the model. If this was not possible the noise level would successively increase to a level where new observations could not be assimilated any longer. The non-linear normal mode initialization developed simultaneously by Machenhauer (1977) and Baer (1977) presented a solution to this problem at least at middle and high latitudes. Problems still remain in the Tropics where it seems necessary to incorporate the effect of non-adiabatic processes. This is less important in the subtropics.

Non-conventional data such as winds from aircraft and satellites need to be interpolated in the vertical as well as in the horizontal. In order to do this in a systematic way, a 3-dimensional analysis system based on optimum interpolation of observed increments was developed, Lorenc (1981).

Although this method has its limitations in rapidly changing situations (structure functions are statistical) it provides nevertheless an efficient method to analyse observations which in principle can have a random distribution in space. The method is also very useful in automatic control

Table 1 Problem areas in medium range weather prediction in 1974/1975.

Analysis	. analysis methods in the Tropics
	. usage of non-synoptic data
	. usage of non-conventional data
	. automatic control of observations in data sparse areas.
Initialization	. initialization in the Tropics
	. initialization of the ultra-long waves
	. initialization in the presence of mountains
	. the "spin-up" problem
Numerical methods	. computational stability in long term integration
	. grid representation vs. spectral
	. grid representation in polar regions
	. semi-implicit algorithms
	. computational economy
Parameterization	. turbulent fluxes
	. radiation and clouds
	. heat balance
Surface representation	. orography and roughness
	. soil moisture and surface heat balance
	. air-sea interaction and representation of the oceans in the
	. surface albedo

of observations which is carried out by the same technique.

The computational stability of the numerical model is provided by a finite different scheme which conserves potential enstrophy for a non-divergent flow. It uses a staggered grid of variables known as the C-grid (Arakawa and Lamb, 1977). This grid was selected because of its low computational noise and the ease of implementation of a semi-implicit scheme. It has excellent stability properties in extended integration as has been demonstrated by Sadourny (1975). Further details of the scheme can be found in Burridge and Haseler (1977), and Burridge (1979).

There was not enough practical experience in 1975 to select a spectral model. However, considerable experience has been obtained since then and the general opinion is now that spectral models are superior to grid point models for a comparative amount of computational cost.

A problem which was carefully considered 6 years ago, was the level of complexity or sophistication of the parameterization of sub-grid scale processes. Miyakoda and Sirutis (1977) evaluated 3 different levels of parameterization characterized by an increasing complexity of the description of vertical turbulent fluxes and convection. These results indicated an advantage in using the so-called higher order closure schemes (Mellor and Yamada 1974), while one of the most elaborate schemes to describe convection (Arakawa and Schubert, 1979) did not show the same advantage. ECMWF selected after some evaluation a scheme of moderate complexity. The parameterization of surface and turbulence fluxes was based on Monin-Obukov similarity theory and K-theory (first order closure) respectively, while Kuo's scheme (Kuo, 1974) was selected to describe convection. Louis (1979) has shown that this parameterization performed well in cases of intense air-sea interaction. The scheme also simulated in satisfactory details the structure of the boundary layer from the O'Neill experiment (Lettau and Davidson, 1957) and other similar experiments.

The radiative processes have been incorporated in great detail and the scheme is fully interactive with the large scale cloud-distribution (given as an empirical function of relative humidity and pressure). This scheme has been found to provide positive skill (Geleyn, 1980), as compared with a non-interactive cloud/radiation model (clouds geographically prescribed).

The physical processes at the surface and in the soil are described in a very simple heuristic manner. At this time the specification of the surface condition was not believed to be essential for medium range forecasting. However, recent studies, Shukla and Mintz (1982) and van Maanen, (pers.inform.) show that the description of the evapotranspiration is of fundamental importance, in particular in the Tropics and over the Northern Hemisphere in the summer, due to the strong feedback between convective precipitation and evapotranspiration from the surface. For medium range prediction it is probably more important to know the soil water or the potential evapotranspiration than to know the water vapour in the atmosphere!

3. OPERATIONAL ASPECTS

To establish an operational forecasting system implies a number of practical/technical constraints which must be considered carefully. Table 2 summarises some criteria which are crucial for model selection and model development and modification.

From the operational point of view the ECMWF plans called for an operational forecast a day, which by and large has to be produced in less than 8 hours including all operational aspects (decoding, analyses prediction, post-processing and dissemination). Of this time a 10-day prediction can approximately occupy about 5 hours which means that the prediction must

Table 2 Model selection and model modificiation criteria.

Practical/ Operational	. timing (50/1); 1 forecast a day
	. robustness (/failure/month)
Model formulation	. generality in formulation
	. lucidity (simplicity)
	. minimum of tuning
Programming aspects	. efficiency
	. flexibility toward future changes
Quality aspects	. objective evaluation
	. subjective evaluation

progress about 50 times faster than the evolution of the real weather. Given the constraints of the CRAY-1 computer this limited the resolution of the model to about 200 km and 15 vertical levels. The vertical resolution was selected because of the necessity to describe accurately enough the vertical structure of developing cyclones and the boundary layer. The horizontal resolution is still unsatisfactory and theoretical and synoptic investigations suggest that a resolution of around 100 km would be more satisfactory.

The model has to be designed in a way that calculations under no circumstances would create computational instability. Furthermore the reliability of the computer as well as the overall program structure must be such as to guarantee a very stable operational performance with less than 1 failure/month. This condition has been satisfied by a wide margin and the number of operational failures has been 5 since the start of operational forecasts in August 1979 and none at all in the last operational year 1981.

A very important factor in designing and maintaining forecasting systems of the complexity of that of the Centre is a clear and well defined structure both in the physical formulation and in the programming design.

From the very beginning the Centre has strived to maintain a generality in the physical formulation and to keep "tuning" to a minimum. Modifications of the model must therefore on the first hand be justified on dynamical and physical grounds and not only by the fact that a change happens to give better results. The reason is simply that it is extremely difficult to prove a better result on a limited set of numerical experiments only. Furthermore, conceptually simpler formulations have been selected in favour of more complex ones unless obvious improvements can be demonstrated.

The programming aspects have been constrained by computational efficiency and programming flexibility. While the first of these conditions is obvious, the second may need some justification. The Centre's forecasting system (data-asimilation and prediction model) contains the order of 100 000 lines of code. The additional support system such as pre- and post processing as well as verification and diagnostic evaluation is of a similar size. It is obvious that well-organized and well-documented programs must be set up if such a system would be fully useful. There are examples of large program systems of this kind which are impossible to use efficiently because the complexity is beyond a clear understanding - one has passed over the "threshold of complexity".

The careful design in programming implemented from the beginning at ECMWF has paid off and so far about 30 changes have taken place in the operational forecasting system during 2 years of operation. These changes often associated with major model and system changes, have been straightforward to implement and the number of support progammers have been very few.

4. IMPROVEMENT OF NUMERICAL FORECASTS

A very substantial improvement has taken place in numerical weather prediction since the very first forecasts were made more than 30 years ago. These improvements are in essence due to model improvements and to a better utilization of available observations. The observing system has undergone some improvements during the last 30 years, but it is not very likely that these have had any substantial effect on the quality of weather prediction at middle and high latitudes. Results of 24 operational and quasi-operational 24 h predictions by the barotropic model (Staff Members, University of Stockholm, 1954) show an average standard deviation error of 76 m when verified over an area covering Northwestern Europe and Northeastern Atlantic. The integrations were done over a limited area $(5700\ \text{km})^2$ extending about 1200 km beyond the area of verification.

In order to compare the performance of the barotropic model with the Centre's operational model, we have carried through daily barotropic forecasts for the month of Janury 1981. The horizontal resolution was 381 km . In this case the integration domain was almost hemispheric and we used the initial state produced operationally. It is to be expected that these initial states are more accurate tha those of 30 years ago, due to a more systematic use of observations at other levels and to a better first guess provided with the ECMWF model. The importance of an accurate initial state was in fact demonstrated (Ibid) and better forecasts were obtained when more carefully analysed data were used. Table 3 shows the standard deviation error for the barotropic model of 1954 and 1981 as well as the result of the operational ECMWF model. The verification for the latter experiment has been carried out over a domain covering Europe, 72N - 36N and 12W - 42E. This verification area is somewhat larger and in a more easterly position than the one used in 1954. The figures in the table demonstrates the achievements in numerical weather prediction over the last 30 years. It is also of interest to compare the computer time for a 24 hour forecast. The Staff Members, University of Stockholm were using the Swedish constructed BESK computer which about this time was one of the fastest if not the fastest computer in the world. BESK had an internal William-type memory; 512 words of 40 bits. No external memory was available at this time. Table 4 shows the necessary computer time (elapsed time) to produce a 1 day forecast. It should be pointed out that while both the BESK-code and the operational code on CRAY-1 are optimised, the barotropic code on CRAY-1 is non-vectorized standard Fortran.

Another example of the improvement in numerical weather prediction can be found from a recent study by Bengtsson and Lange (1982). In this study operational 3-day forecasts from different centres are compared. It is found that the forecasts by high resolution ECMWF model for instance have errors which are less than 60% compared to those based on less advanced filtered

Table 3 Standard deviation error for 500 mb forecasts over Europe.

Time	Jan. 81 (31 cases)		Nov. 51 - April 54 (24 cases)
	ECMWF Operational model	Barotropic model	Barotropic model
24 h	22 m	47 m	76 m
48 h	41 m	97 m	
72 h	62 m	151 m	

Table 4 Computer time for a 24 h forecast

CRAY-1 ECMWF model	CRAY-1 Barotropic model	BESK Barotropic model
18500 grid points	2350 grid points	400 grid points
1100 sec	0.5 sec*	2400 sec

*(This time could be reduced to about 0.1 sec by redesigning the code)

models.

We will next study the improvement of medium range prediction at ECMWF since operational forecasts started in August 1979. An operational model normally goes through a process of improvements due to refinements in the treatment of physical parameterization and due to identification and correction of erroneous formulations and program errors. Such improvements have taken place both in the model and in the data-assimilation and are successively inserted into the operational forecasting system. Figure 1 shows the anomaly correlation for the 500 mb prediction for the winter of 1979/80 and 1980/81. It can be clearly seen that the scores for the second winter are significantly higher, reflecting a genuine improvement of the model. We have also incorporated as a benchmark for this comparison the very early results by Miyakoda et al, 1972, which constitute the first comprehensive trial of medium range predictions. Their ensemble mean anomaly correlation for the extra-tropical Northern Hemisphere was based on 12 January cases taken from the years 1964 to 1969.

5. POTENTIAL IMPROVEMENTS OF MEDIUM RANGE FORECASTS

The possibility of a further improvement of the numerical forecasts and the possibility of extending the length of useful forecasts is a fundamental issue of this seminar. As Professor Lorenz has pointed out in his lecture, there is an inherent limitation in the predictability of the atmosphere which is possibly of the order of a few weeks. However, there are also clear indications, Lorenz (1982) by comparing the internal error growth of the ECMWF model, that there is scope for extending the range of useful forecasts by 3 - 4 days even with today's incomplete and inaccurate observations. These improvements are likely to fall into the largest scale of motion where the predictions have considerable errors of a systematic nature, Bengtsson and Simmons (1982). We will illustrate the characteristic features of the systematic error by simply correcting the forecast from January 1981 by the

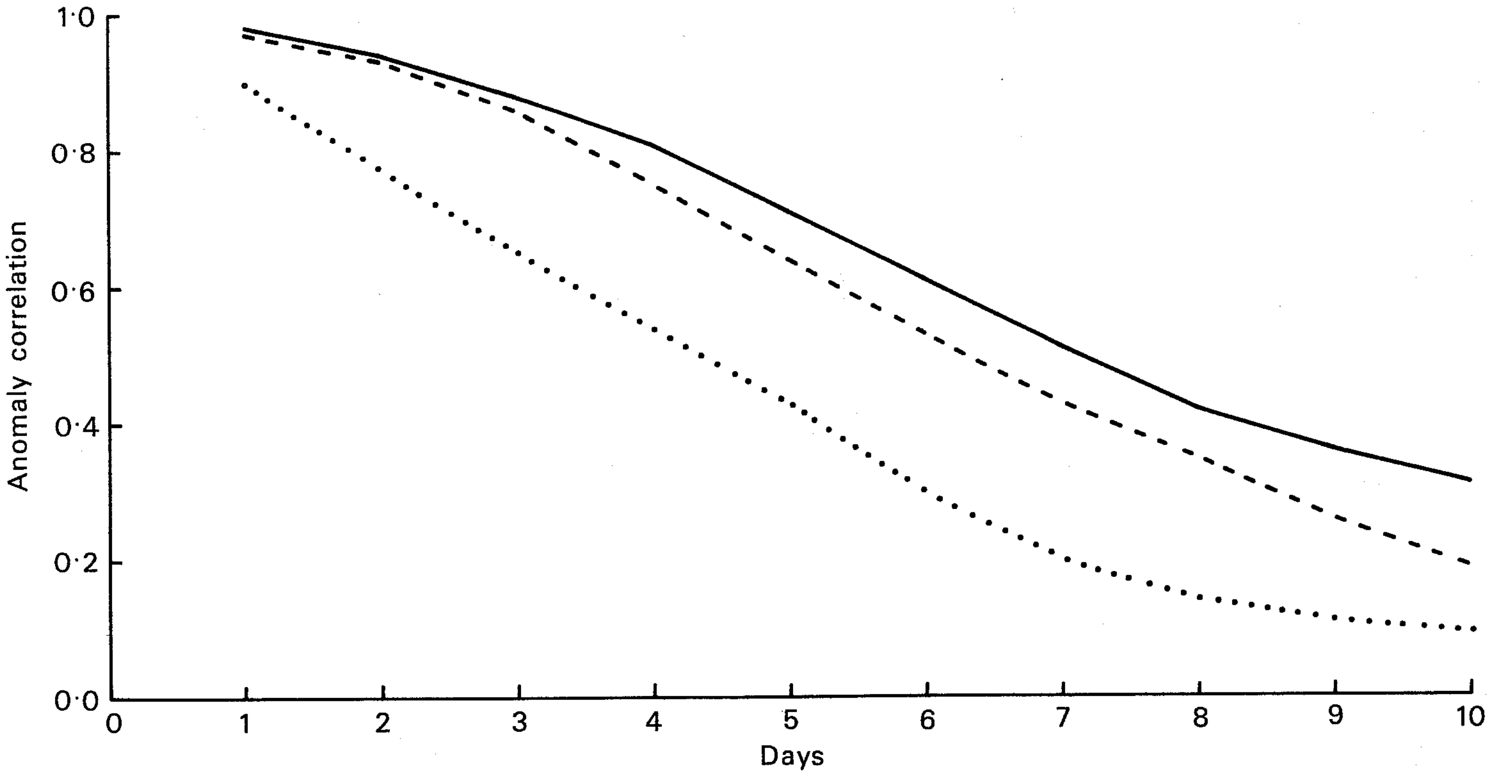

Fig. 1 Anomaly correlation for the 500 mb flow (Northern Hemisphere extra-tropics) for the winter 1980/81 (Dec., Jan., Feb.), full lines, for the winter 1979/80, dashed lines and Miyakoda et al. (1972) dotted lines.

monthly mean error of January 1980.The result is shown in Table 5. The table shows the extension of useful predictability in days as defined by 2 different levels of anomaly correlation. It is interesting to note the particular improvement for the ultralong waves, although the whole spectrum has in fact improved.

These experiments strongly suggest that there are still fundamental errors in the maintenance of the large scale quasi-stationary flow and the "climate" of the model differs from the real climate. The predicted "quasi-stationary" flow has a smaller amplitude than the observed as can be seen from Figure 2 which shows the observed 500 mb geopotential for January 1981 and the averages of all the 10-day forecasts for January 1981.

The reduced poleward transport of heat of momentum by the quasi-stationary flow is compensated by an increased transportation undertaken by the transient eddies which will develop into more active disturbances. This is perhaps best illustrated by Figure 3, where we study the performance of the operational model with respect to the energy cycle and to the kinetic energy budget. The figure shows the global energy budget for January 1981 as calculated from 12Z analyses as well as the ensembles of 2, 5 and 10-day forecasts. All standard levels between 1000 and 30 mb have been used. There is an overall increase of around 10% in total available and in zonal kinetic energy. The increase of the zonal part is greatest during the last five days. The eddy kinetic energy on the other hand is decreasing by a similar percentage. The energy cycle as measured by the dissipation rate is successively intensifying from an initial value of 1.7 Watts/m^2 towards a value of 3.3 Watts/m^2. The calculation from the analyses is clearly an underestimation due to the spin-up process, while the values towards the end of the forecast are very likely too large. Probably the figures valid for day 2 (2.6 Watts/m^2) is the most correct. It is interesting to note the

Table 5 Predictability in days for 500 mb for January 1981 for the ECMWF model with and without statistical correction (for further information see text).

	Anomaly correlation score (A.C.)	Number of days required to reach A.C. value Before Statistical correction	 After Statistical correction
Long waves	80%	4.3	4.8
(1 - 3)	50%	8.9	>10
Medium waves	80%	3.5	3.7
(4 - 9)	50%	6.0	6.4
Short waves	80%	1.4	1.5
(10 - 20)	50%	3.2	3.3
Total field	80%	3.9	4.3
(all waves)	50%	6.8	8.3

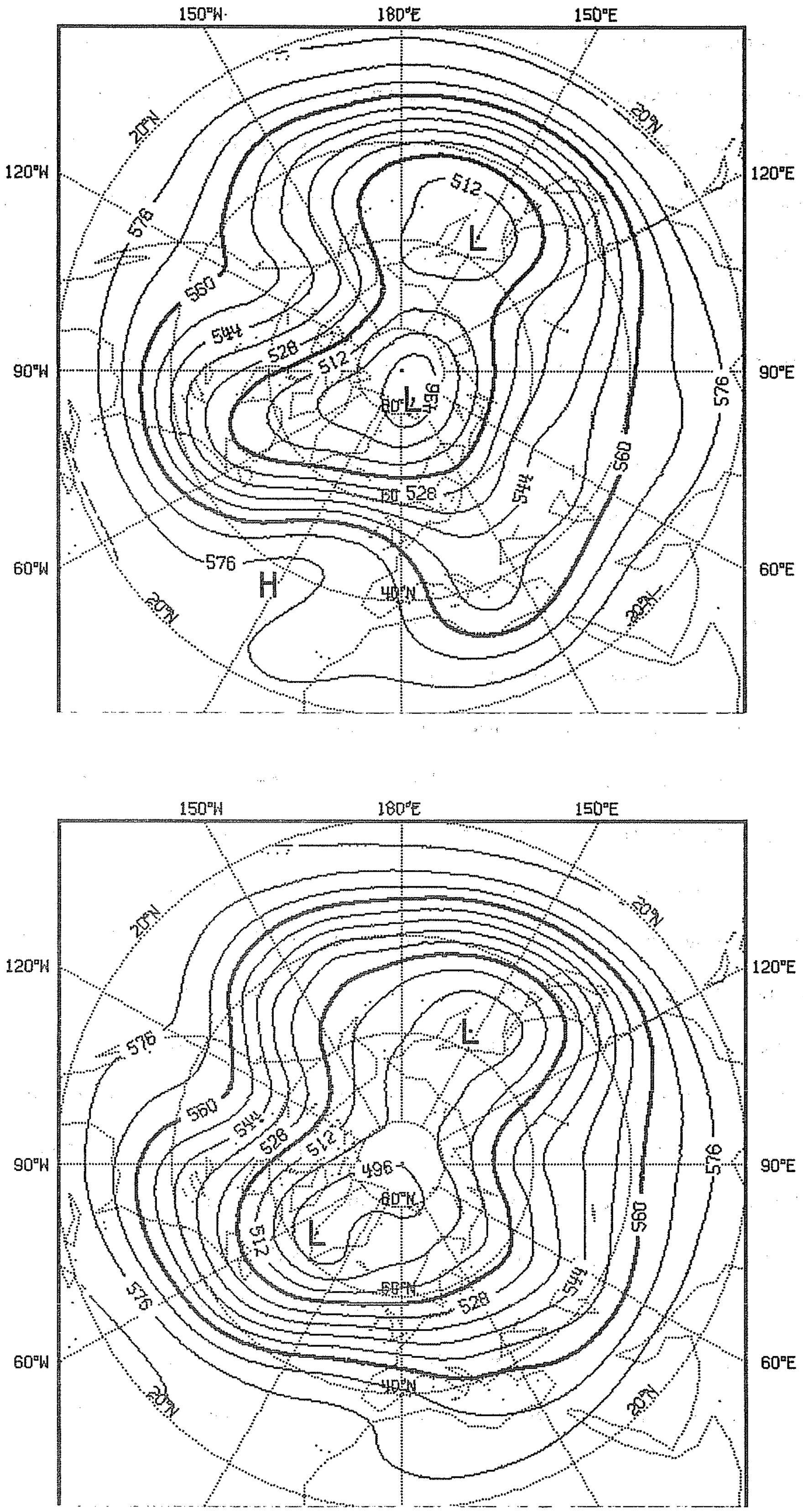

Fig. 2 500 mb flow January 1981 (top) and ensemble average of all 10-day forecasts for January 1981 (bottom).

figures which describe the generation of eddy available potential energy. When calculating this term from initial data or from short range forecasts, it is found that heating processes destroy eddy available potential energy mainly due to the heating of deep cold air masses over oceans which destroy eddy available potential energy. However, heating due to release of latent heat in the warm air masses increases the eddy available potential energy and hence counteracts this destruction. Because of the fact that the spin-up time for latent heat release is a few days, this contribution to the eddy available potential energy is underestimated and cannot be accurately assessed from the analyses.

Figure 3 illustrates an interesting aspect of the model and its deviation from nature. The intensification of the energy cycle is caused by a too high generation of available potential energy by non-adiabatic processes. This is in its turn possibly related to an incomplete destruction of available potential energy by turbulence and convection in the cold air masses. That this is perhaps the case, is suggested from Figure 4 which shows the heat balance over a 10-day forecast averaged over the whole globe. The radiative flux is cooling the atmosphere more than is compensated for by the surface fluxes of latent and sensible heat. This gives a cooling of the whole atmosphere by 18.8 Watts/m^2 or 1.6°C in 10 days. The heat balance is negative up to about 50 days, at which time the average global temperature has fallen by about 4°. No further cooling appear to take place hereafter. The gradual cooling and adjustment of the heat balance is shown in Figure 5.The cooling of the atmosphere has a typical vertical profile with a maximum around 500 mb and in the lower troposphere. The reason for this is not yet fully understood and intense research effort is under way to understand this problem.

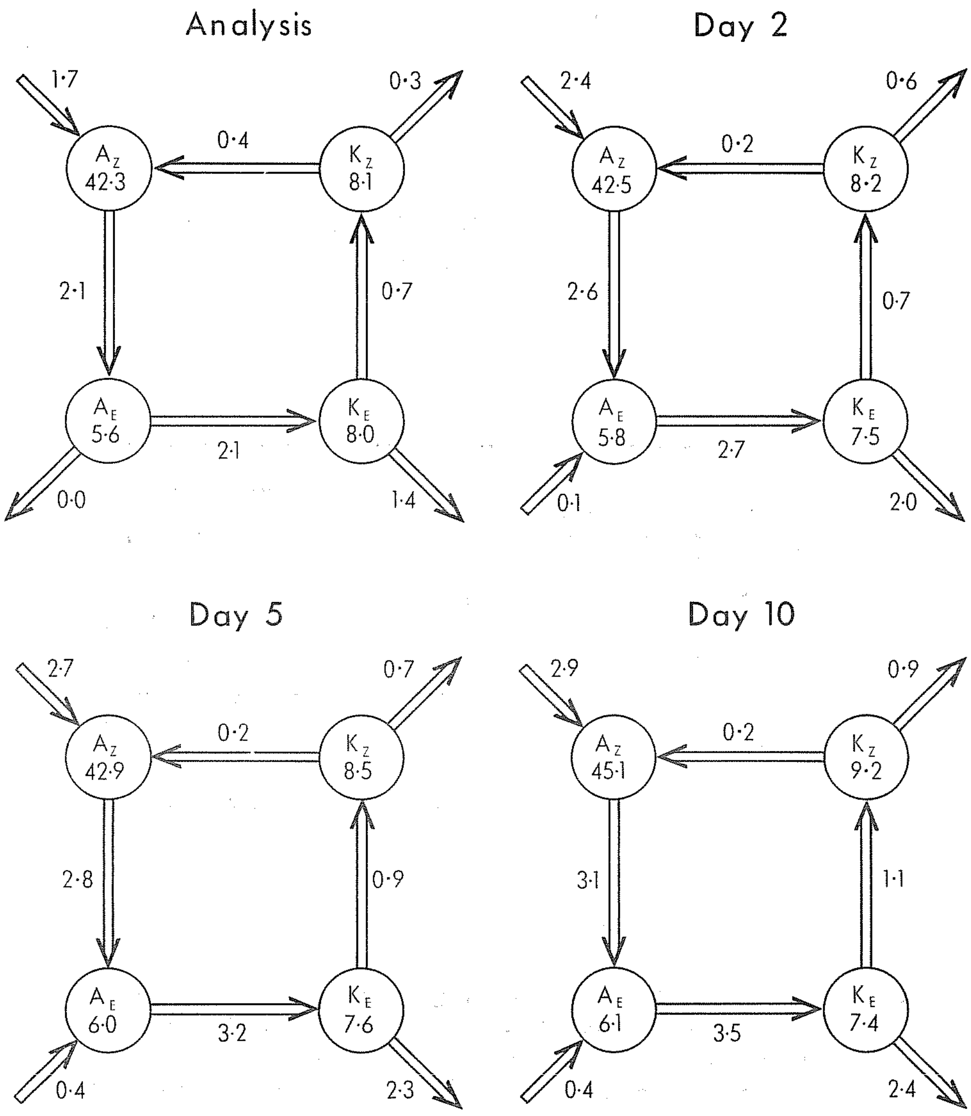

Fig. 3 Global energy diagram for January 1981 as calculated from operational analyses and forecasts. The energies (inside the circles) are given in 10^2 kj/m^2 while the energy transformations are given in watts/m^2.

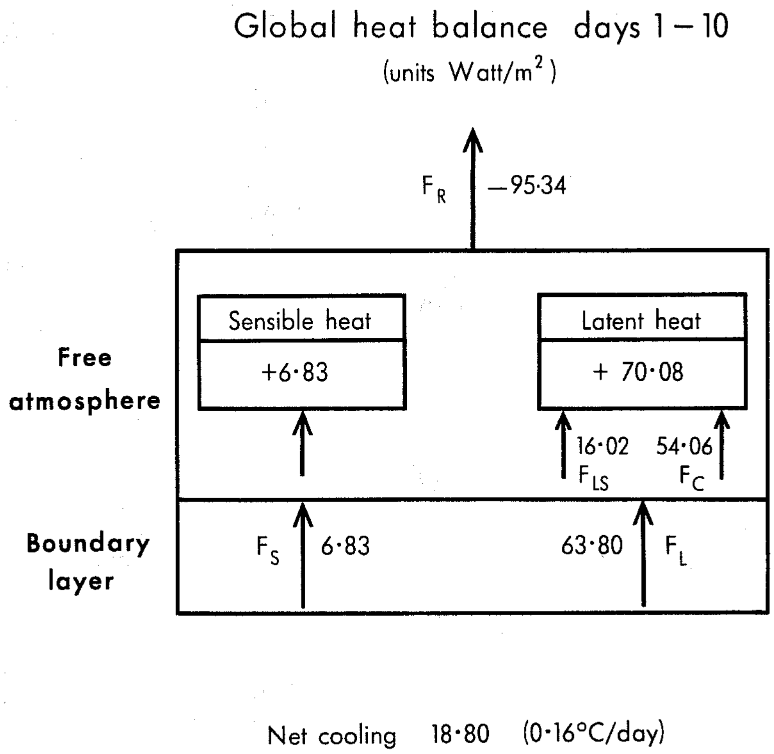

Fig. 4 Global heat balance averaged over a 10-day forecast. F_R indicates the total radiation flux, F_S the sensible heat flux and F_L the latent heat flux. F_{LS} and F_C denotes heating by large scale precipitation and by convective precipitation respectively.

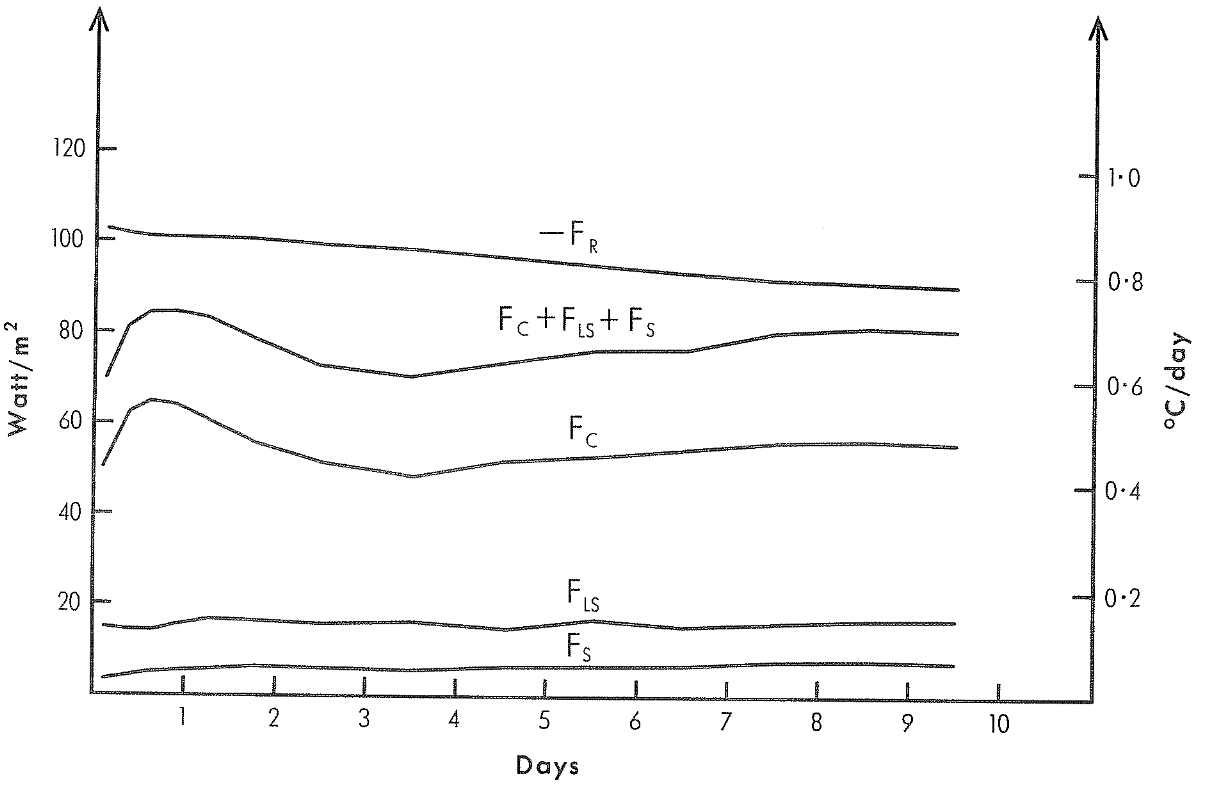

Fig. 5 Global fluxes of F_R (total radiation flux), F_S (sensible heat flux), F_{LS} (heating by large scale precipitation) and F_C (heating by convection) over a 10-day forecast. Note the spin-up effect (in particular for F_C) and the tendency to over-shooting during the first day.

6. CONCLUSIONS

We have demonstrated that substantial improvements have taken place in numerical forecasting since the start of numerical weather prediction in 1950. The improvements can be seen as successively better short range forecasts and a gradual extension of the length of useful forecasts up to about 6 to 7 days for the northern hemisphere winter. There are indications that this perhaps can be extended with another couple of days when the systematic deficiencies of present models have been eliminated. However, this may take a considerable time. There is still no realistic alternative to the "brute force approach" and future models are likely to improve through a systematic and meticulous improvement of all aspects of the forecasting system. We can therefore recommend the same general strategy as six years ago.

An alternative approach can possibly be considered for extended range forecasts where the predictability of the individual weather systems has ceased. As has been shown by Miyakoda (pers.inform.) and Shukla (1982) there are indications that the useful predictions of time averages can be extended even further, possibly up to a month. It seems that the most logical approach to such very long integrations would be to apply a simple kind of Monte-Carlo prediction using a sequence of initial states, say over a few days, to produce an ensemble of predictions.

7. ACKNOWLEDGEMENT

The author acknowledges the assistance of H.Savijärvi in undertaking the prediction experiments with the barotropic model.

REFERENCES

Arakawa, A. and V.R.Lamb 1977 Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics*, *Vol.17* J.Chang, Ed., Academic Press, 337 pp.

Arakawa, A. and W.H.Schubert 1974 Interaction of a cumulus cloud ensemble with the large-scale environment. *J.Atmos.Sci.*, *31*, 5674-701.

Baer, F. 1977 Adjustment of initial conditions required to suppress gravity oscillations in non-linear flows. *Beitr.Phys.Atmos.*, *50*, 350-366.

Bengtsson, L. 1975 4-dimensional assimilation of meteorological observations, GARP Publiction Series No.15, World Meteorological Organizationn, 76pp.

Bengtsson, L. and A.Lange 1982 Results of the WMO/CAS NWP data study and intercomparison project for forecasts for the Northern Hemisphere in 1979-80. CAS Working group on weather prediction research. WMO

Burridge, D.M. 1979 Some aspects of large scale numerical modelling of the atmosphere. *Proceedings of ECMWF Seminar on Dynamical Meteorology and Numerical Weather Prediction*, *Vol.2*, 1-78.

Burridge, D.M. and J.Haseler 1977 A model for medium range weather forecasting - Adiabatic formulation. *ECMWF Tech.Rep.No.4*, 46pp.

ECMWF, 1975 Seminars on scientific foundation of medium range weather forecasts, ECMWF 713pp.

Geleyn, J.-F. 1980 Some diagnostics of the cloud/radiation interaction in ECMWF forecasting model. *Workshop on radiation and cloud-radiation interaction in numerical modelling*, *ECMWF*, 135-162.

Kuo, H-L. 1974 Further studies of the parameterization of the influence of cumulus convection in large-scale flow. *J.Atmos.Sci.*, *31*, 1232-1240.

Lettau, H.H. and Davidson, B. 1957 *Exploring the Atmosphere's First Mile*, Vol.2, Pergamon Press, New York, 578pp.

Lorenc, A.C. 1981 A global three-dimensional multivariate statistical interpolation scheme. *Mon.Wea.Rev.*, *109*, 701-721.

Lorenz, E. 1982 Atmospheric predictability experiments with a large numerical model. Submitted to *Tellus*.

Louis, J.-F. 1979 A parametric model for vertical eddy fluxes in the atmosphere. *Bound.Lay.Met.*, *17*, 187-202.

Machenhauer, B. 1977 On the dynamics of gravity oscillations in a shallow-watermodel with application to normal mode initialisation. *Beitr.Phys.Atmos.*, *50*, 253-271.

Mellor, G.L. and Yamada, T. 1974 A hierarchyof turbulence closure models for planetary boubdary layers. *J.Atm.Sci.*, *31*, 1791-1806.

Miyakoda, K., G.D.Hembree, R.F.Strickler and I.Shulman 1972 Cumulative results of extended forecast experiments. I. Model performance for winter cases. *Mon.Wea.Rev.*, *100*, 836-855.

Miyakoda, K. 1975 Weather forecasts and the effects of the sub-grid scale processes. Seminar on scientific foundation of medium range weather forecasts, ECMWF, 380-593.

Miyakoda, K. and Sirutis, J. 1977 Comparative integration of global models with various parameterized processes of sub-grid scale vertical transports: Description of parameterization. Beitr.Phys.Atmos., 50, 445-487.

Sadourny, R. 1975 The dynamics of finite difference models of the shallow-water equations. J.Atmos.Sci., 32, 680-689.

Shukla, J. 1982 Predictability of monthly means. Part I: Dynamical predictability. Part II: The influence of the boundary forcing. Proceedings of ECMWF Seminar: Problems and prospects in long and medium range weather forecasting.

Shukla, J. and Mintz, Y. 1982 The influence of land-surface evapotranspiration on the earth's climate. Submitted to Science.

Staff Members, University of Stockholm 1954 Results of forecasting with the barotropic model on an electronic computer (BESK) Tellus, 6, 139-149.