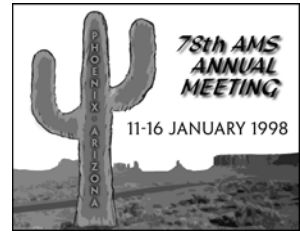


# Maturity of Operational Numerical Weather Prediction: Medium Range



Eugenia Kalnay,\* Stephen J. Lord, and Ronald D. McPherson  
National Centers for Environmental Prediction, Washington, D.C.

**Editor's note:** *This is the third of a series of papers about Carl-Gustaf Rossby that will be published in the Bulletin. All papers were presented at the 1998 Annual Meeting of the American Meteorological Society as part of the Special Session Honoring the Centennial of the Birth of Carl-Gustaf A. Rossby. Other papers in the series will appear in upcoming issues of the Bulletin.*

## ABSTRACT

In 1939 Rossby demonstrated the usefulness of the linearized perturbation of the equations of motion for weather prediction and thus made possible the first successful numerical forecasts of the weather by Charney et al. In 1951 Charney wrote a paper on the science of numerical weather prediction (NWP), where he predicted with remarkable vision how NWP would evolve until the present. In the 1960's Lorenz discovered that the chaotic nature of the atmosphere imposes a finite limit of about two weeks to weather predictability. At that time this fundamental discovery was "only of academic interest" and not really relevant to operational weather forecasting, since at that time the accuracy of even a 2-day forecast was rather poor. Since then, however, computer-based forecasts have improved so much that Lorenz's limit of predictability is starting to become attainable in practice, especially with ensemble forecasting, and the predictability of longer-lasting phenomena such as El Niño is beginning to be successfully exploited.

The skill of operational weather forecasts has at least doubled over the last two decades. This improvement has taken place relatively steadily, driven by a large number of scientific and computational developments, especially in the area of NWP. It has taken place in all the operational NWP centers, as friendly competition and information sharing make scientific improvements take place faster than they would in a single center. Because the improvements have occurred steadily, rather than suddenly, the overall increase in forecast skill due to NWP has not been clearly recognized by the media and the public despite the impact that improved forecasts have on the national economy and on the lives of every American.

In this paper the authors review several measures of operational forecast skill that quantify improvements in NWP at the National Centers for Environmental Prediction (NCEP, formerly the National Meteorological Center) of the National Weather Service, although they are representative of improvements in all major NWP operational centers. The authors point out that there are three major requirements for improved numerical weather prediction: better atmospheric models, better observational data, and better methods for data assimilation. These improvements are generally very computer intensive and can only be made operational with the availability of more powerful supercomputers. Operational forecasts are compared with "re-forecasts" from the NCEP-NCAR 40-Year Reanalysis, showing that, if the present-day NWP systems had been available many decades ago, skillful 5-day forecasts would have been possible in the Northern Hemisphere with the upper-air network of the late 1950s. The authors discuss new approaches in the use of observations (variational assimilation of remote observations) and of numerical weather prediction guidance (ensemble forecasting) that have allowed the recent extension of operational predictions into longer ranges and the possibility of adaptive observing systems. The extension of operational forecast skill into seasonal predictions of the El Niño-Southern Oscillation phenomena using coupled ocean-atmosphere models is also discussed. In the last section the authors attempt to "forecast" the future of NWP.

---

\*Additional affiliation: School of Meteorology, University of Oklahoma, Norman, Oklahoma.

Corresponding author address: Eugenia Kalnay, Energy Center, 100 East Boyd, Room 1310, Norman, OK 73019-0470.

E-mail: ekalnay@ou.edu

In final form 5 August 1998.

©1998 American Meteorological Society

## 1. Introduction

In 1939, Carl-Gustaf Rossby demonstrated the usefulness of the linearized perturbation of the equations of motion for weather prediction. Rossby assumed in his simple model that the atmospheric velocity was horizontal and nondivergent, so that its evolution was governed by the conservation of potential vorticity about the vertical axis, and he obtained the famous formula for the phase speed of the atmospheric large-scale waves that have been named after him. The first successful numerical forecasts of the weather were made by Charney et al. (1950), using Rossby's model, except that they included nonlinear advection of vorticity. Phillips (1990) pointed out that Rossby not only influenced the design of the numerical model, but his introduction in 1945 of the concept of group velocity also affected the choice of domain for the experiment: it allowed Charney et al. (1950) to estimate the area of influence of the information coming from the initial conditions (see also Phillips 1998).

Only one year after the successful forecasts, Charney (1951) wrote what can be considered a truly visionary text on numerical weather prediction (NWP). In this paper, which has not become dated even after half a century, Charney discussed the importance of Richardson's (1922) experiment and the influence that Rossby's (Rossby et al. 1939; Rossby 1940) model had on the development of the quasigeostrophic model used on the ENIAC, the first electronic computer. He also reviewed the theoretical basis of the numerical model and boundary conditions and their numerical formulation. He described with amazing insight the evolution that he foresaw in NWP, including the replacement of the quasigeostrophic equations by the primitive equations, the balancing of initial conditions, the need for realistic parameterizations of small-scale physical processes, and the development of automatic data assimilation and quality control methods. He even anticipated that there would be a finite limit to the length of weather predictability, but he attributed it to unavoidable inaccuracies in models and estimates of the atmospheric initial conditions.

In a series of remarkable papers, Lorenz (1963, 1965, 1968) made the fundamental discovery that even with perfect models and perfect observations, the chaotic nature of the atmosphere would impose a finite limit of about two weeks to the predictability of the weather. In the 1960s Lorenz's discovery, which started the theory of chaos, was "only of academic

interest" and not really relevant to operational weather forecasting, since at that time the accuracy of even a two-day forecast was quite poor. Since then, however, computer-based forecasts have improved so much that Lorenz's limit of predictability is starting to become occasionally attainable in practice,<sup>1</sup> especially with ensemble forecasting, and predictability of longer-lasting phenomena such as El Niño is beginning to be exploited.

The skill of operational weather forecasts has increased very substantially over the last two decades. This improvement has taken place slowly and relatively steadily, driven by a large number of scientific and computational developments, especially in the area of NWP. It has taken place in all the operational NWP centers, as friendly competition and information sharing make scientific improvements take place faster than they would in a single center. Because the improvements have occurred steadily, rather than suddenly, the dramatic overall increase in forecast skill due to NWP has not been clearly recognized by the media and the public, despite the impact that the forecasts have on the national economy and on the lives of every American.

Nevertheless, evidence suggests that the media and the public have become aware of the fact that weather forecasts are currently much more reliable than they used to be. As an example, there was a striking contrast between the reaction of commercial airlines to the forecasts of the "storm of the century" (13 March 1993) and the Blizzard of 1996 (6 January 1996) that closed down the government on the East Coast. Both storms were first anticipated about 5 days in advance, and forecasts for extreme conditions were made a few days before the blizzards. In 1993 the airlines kept flying until their planes were grounded by the weather, and it took them several days to dig the planes out. By 1996, they decided that the forecasts meant what they said and kept their planes out of the path of the storm; substantial savings resulted (W. Hooke 1998, personal communication).

The media have also reacted to the availability of more reliable forecasts in the medium range (beyond two days) and are using the National Weather Service (NWS) NWP and Model Output Statistics (MOS) forecasts for days 2–7 and issuing local and national

---

<sup>1</sup>An example of an operational forecast (initial conditions 10 December 1995) where the anomaly correlation remained above 70% for 15 days has been shown in Toth and Kalnay (1996).

“value-added” forecasts. In recent years the Weather Channel and all major television networks have started issuing daily forecasts for 5 days in advance, and it is commonly accepted that people use them routinely for their daily work, leisure, and travel plans (and express their displeasure when the forecasts are not correct). In Oklahoma City, for example, the public has access to at least five major presentations for 5-day local forecasts of maximum and minimum temperature (three TV networks, one morning newspaper, and the local NWS office). Two of those five media also issued numerical forecasts of probability of precipitation for days 1–7, whereas the others used symbols or qualitative language (Brooks et al. 1997).

In this paper we review some measures of operational forecast skill that quantify improvements in NWP at the National Centers for Environmental Prediction (NCEP, formerly the National Meteorological Center, NMC) of the National Weather Service, although they are representative of improvements in all major NWP operational centers. They are based on scientific research and development. Some of the improvements were developed at NCEP, others were implemented through collaboration with national research organizations, and still others originated from international collaboration. The collaboration with the Geophysical Fluid Dynamics Laboratory (GFDL) and the National Environmental Satellite, Data and Information Systems (NESDIS), the National Center for Atmospheric Research (NCAR), the National Air and Space Administration (NASA), the National Science Foundation/NCEP Joint Program for Research in NWP, and the University Corporation for Atmospheric Research/NCEP Visiting Scientist Program were essential in this process. Friendly competition and open sharing of advances in NWP have accelerated progress in all national and international operational centers. The review is made possible by the availability of very long records of operational forecast skill verification maintained at NCEP and by the completion of the NCEP–NCAR 40-year global reanalysis, in which only the observing system changes with time, thus allowing the impact of observations to be isolated.

We point out that there are three major requirements for improved numerical weather prediction: better atmospheric models, better observational data, and, very importantly, better methods for data assimilation. These improvements are generally very computer intensive and can only be made operational with the availability of more powerful supercomputers. We discuss some recent new approaches in the use of ob-

servations (variational assimilation of remote observations) and of numerical weather prediction guidance (ensemble forecasting) that have allowed the recent extension of operational predictions into longer ranges and the possibility of adaptive observing systems.

In section 2 we review the evolution of several objective measures of skill, indicating that over the last two decades the forecast skill has at least doubled. Section 3 compares operational forecasts with “reforecasts” from the reanalysis. Section 4 discusses ensemble forecasting and adaptive observations, and the extension of operational forecast skill into seasonal predictions of El Niño–Southern Oscillation (ENSO) phenomena using coupled ocean–atmosphere models. Section 5 is a brief discussion of what we “forecast” for the future.

## 2. Review of the evolution of operational forecast skill at NCEP

The history of operational numerical weather prediction at the NMC (now NCEP) from its beginning to 1985 has been carefully reviewed by Shuman (1989). It started with the organization of the Joint Numerical Weather Predictions Unit (JNWPU) on 1 July 1954, staffed by members of the U.S. Weather Bureau (later the NWS), the Air Weather Service of the U.S. Air Force, and the Naval Weather Service.<sup>2</sup> Shuman pointed out that in the first few years, numerical predictions could *not* compete with those produced manually. They had several serious flaws, among them overprediction of cyclone development. Far too many cyclones were predicted to deepen into storms. With time, and with the joint work of modelers and practicing synopticians, major sources of model errors were identified, and operational NWP became the central guidance for operational weather forecasts.

Shuman (1989) included a chart with the evolution of the S1 score (Teweles and Wobus 1954), the first measure of error in a forecast weather chart, which, according to Shuman (1989), was designed, tested, and modified to correlate well with expert forecasters’ opinions on the quality of a forecast. The S1 score measures the average relative error in the pressure gradient (com-

---

<sup>2</sup>In 1960 the JNWPU divided back into three organizations: the National Meteorological Center (National Weather Service), the Global Weather Central (U.S. Air Force), and the Fleet Numerical Oceanography Center (U.S. Navy).

pared to a verifying analysis chart). Experiments comparing two independent subjective analyses of the same data-rich North American region made by two experienced analysts suggested that a “perfect” forecast would have an S1 score of about 20%. It was also found empirically that forecasts with an S1 score of 70% or more were useless as synoptic guidance.

Shuman pointed out some of the major system improvements that enabled NWP forecasts to overtake and surpass subjective forecasts. The first major improvement took place in 1958 with the implementation of a barotropic (one level) model, which was actually a simplification from the three-level model first tried, but that included better finite differences and initial conditions derived from an objective analysis scheme (Bergthorsson and Doos 1954; Cressman 1959). It also extended the domain of the model to an octagonal grid covering the Northern Hemisphere down to 9°–15°N. These changes resulted in numerical forecasts that for the first time were competitive with subjective forecasts, but in order to implement them, JNWPU (later NMC) had to wait for the acquisition of a more powerful supercomputer, an IBM 704, replacing the previous IBM 701. This pattern of forecast improvements that depend on a combination of better use of the data and better models and would require more powerful supercomputers in order to be executed in a timely manner has been repeated throughout the history of operational NWP. Table 1 (adapted from Shuman 1989) summarizes the major improvements in the first 30 years of operational numerical forecasts at the NWS. The first primitive equations model (Shuman and Hovermale 1968) was implemented at NMC in 1966. The first regional system (limited fine mesh or LFM model; Howcroft 1971) was implemented in 1971. It remained in use for over 20 years, and it was the basis of MOS. Its

development was frozen in 1986. A more advanced model and data assimilation system, the Regional Analysis and Forecasting System (RAFS) was implemented as the main guidance for North America in 1982. The RAFS was based on the multiple Nested Grid Model (NGM; Phillips 1979) and on a regional Optimal Interpolation (OI) scheme (DiMego 1988). The global spectral model (Sela 1980) was implemented in 1980.

Table 2 summarizes the major improvements implemented in the global system, starting in 1985 with the implementation of the first comprehensive package of physical parameterizations from GFDL. Other major improvements in the physical parameterizations were made in 1991, 1993, and 1995. The most important changes in the data assimilation were an improved OI formulation in 1986, the first opera-

TABLE 1. Major operational implementations and computer acquisitions at NMC between 1955 and 1985 (adapted from Schuman 1989).

Year	Operational model	Computer
1955	Princeton three-level quasigeostrophic model (Charney 1954); not used by the forecasters	IBM 701
1958	Barotropic model with improved numerics, objective analysis initial conditions, and octagonal domain	IBM 704
1962	Three-level quasigeostrophic model with improved numerics	IBM 7090 (1960) IBM 7094 (1963)
1966	Six-layer primitive equations model (Shuman and Hovermale 1968)	CDC 6600
1971	Limited area fine mesh (LFM) model (Howcroft 1971) (first regional model at NMC)	
1974	Hough functions analysis (Flattery 1971)	IBM 360/195
1978	Seven-layer primitive equation model (hemispheric)	
1978	Optimal interpolation (Bergman 1979)	Cyber 205
Aug 1980	Global spectral model, R30/12 layers (Sela 1980)	
March 1985	Regional Analysis and Forecast System based on the NGM (Phillips 1979) and optimal interpolation (DiMego 1988)	

TABLE 2. Major changes in the NMC/NCEP global model and data assimilation system since 1985 (from a compilation by P. Caplan 1998, personal communication).

Year	Operational model	Computer acquisition
Apr 1985	GFDL physics implemented on the global spectral model with silhouette orography, R40/18 layers	
Dec 1986	New Optimal Interpolation code with new statistics	
1987		Second Cyber 205
Aug 1987	Increased resolution to T80/18 layers, Penman–Montieth evapotranspiration and other improved physics (Caplan and White 1989; Pan 1990)	
Dec 1988	Implementation of Hydrostatic Complex Quality Control (Gandin 1988)	
1990		Cray YMP/8 CPU/ 32 megawords
Mar 1991	Increased resolution to T126/18 layers and improved physics, mean orography (Kanamitsu et al. 1991)	
Jun 1991	New 3D Variational Data Assimilation (Parrish and Derber 1992; Derber et al. 1991)	
Nov 1991	Addition of increments, horizontal and vertical OI checks to the CQC (Collins and Gandin 1990; Collins 1998)	
7 Dec 1992	First ensemble system: one pair of bred forecasts at 0000 UTC to 10 days, extension of AVN to 10 days (Toth and Kalnay 1993; Tracton and Kalnay 1993)	
Aug 1993	Simplified Arakawa–Schubert cumulus convection (Pan and Wu 1995), resolution T126/28 layers	
Jan 1994		Cray C90/16 CPU/ 128 megawords
Mar 1994	Second ensemble system: five pairs of bred forecasts at 0000 UTC, two pairs at 1200 UTC, extension of AVN, a total of 17 global forecasts every day to 16 days	
10 Jan 1995	New soil hydrology (Pan and Mahrt 1987), radiation, clouds, improved data assimilation; reanalysis model	
25 Oct 1995	Direct assimilation of TOVS cloud-cleared radiances (Derber and Wu 1998); new PBL based on nonlocal diffusion (Hong and Pan 1996); improved CQC	Cray C90/16 CPU/ 256 megawords
5 Nov 1997	New observational error statistics; changes to assimilation of TOVS radiances and addition of other data sources	
13 Jan 1998	Assimilation of non-cloud-cleared radiances (Derber et al. 1998, personal communication); improved physics	
Jun 1998	Resolution increased to T170/40 layers (to 3.5 days); improved physics; 3D ozone data assimilation and forecast; nonlinear increments in 3D-VAR	

tional three-dimensional variational data assimilation (3D-VAR) in 1991, the replacement of the satellite retrievals of temperature with the direct assimilation of cloud-cleared radiances in 1995, and the use of “raw” (not cloud cleared) radiances in 1998. The model resolution was increased in 1987, 1991, and 1998. The first operational ensemble system was implemented in 1992 and enlarged in 1994.

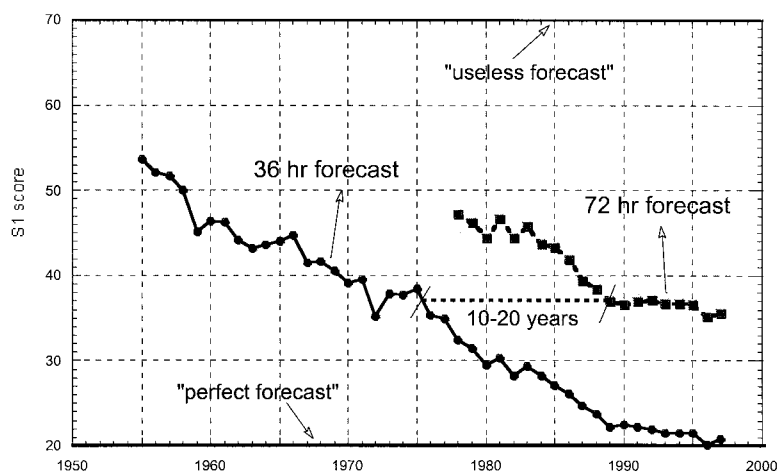
Table 3 contains a summary for the regional systems used for short-range forecasts (to 48 h). The RAFS (triple-nested NGM and OI) were implemented in 1985, and the last upgrades were made in 1991 (DiMego et al. 1992). The Eta Model, designed with advanced finite differences, step-mountain coordinates, and physical parameterizations, was implemented in 1993, with the same 80-km horizontal resolution as the NGM (Mesinger et al. 1988; Janjic 1994; Black et al. 1993; Black 1994). It was denoted “early” because of a short data cutoff. The resolution was increased to 48 km, and a first “mesoscale” version with 29 km and reduced coverage was implemented in 1995. A cloud prognostic scheme was implemented in 1995, and a new land surface parameterization was incorporated in 1996. The OI data assimilation was replaced by a 3D-VAR analysis in 1998, and at this time the early and Meso Eta Models were unified into a 32-km/45-level version. Many other less significant changes were also implemented in the global and regional systems and are not listed here for the sake of brevity. The Rapid Update Cycle (RUC), which provides frequent updates of the analysis and very short range forecasts over the continental United States (CONUS), developed at NOAA’s Forecast System Laboratory, was implemented in 1994 and upgraded in 1998 (Benjamin et al. 1996).

The 36-h S1 forecast verification scores constitute the longest record of forecast verification available anywhere. They were started for mean sea level pressure (MSLP) in the late 1940s for subjective surface forecasts, before operational computer forecast guidance, and for 500 hPa in 1954, with the first

numerical forecasts. Figure 1a includes the forecast scores for 500 hPa from 1954 until the present, as well as the scores for the 72-h forecasts. It is clear that the forecast skill has improved substantially over the years and that the current 36-h 500-hPa forecasts are close to a level that in the 1950s would have been considered perfect (Shuman 1989). The 72-h forecasts have also improved and are now as accurate as the 36-h forecasts were about 15 years ago. This doubling of the skill over 10–20 years can be observed in other types of forecast verifications as well.

(a)

**NMC/NCEP operational S1 scores over North America (500 hPa)**



(b)

**NMC/NCEP operational forecasts: S1 scores for MSLP over North America**

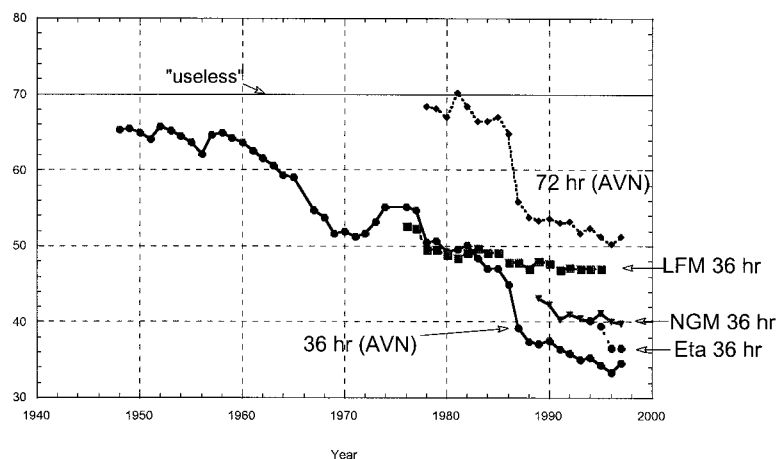


FIG. 1. (a) Record of the S1 score (Teweles and Wobus 1954), which measures the relative error in the pressure gradient for the 36-h and 72-h operational predictions of geopotential heights at 500 hPa. (b) As in (a) but for sea level pressure. The scores of several model (including two models, the LFM frozen in 1986 and the NGM frozen in 1991) are also included. Data courtesy of C. Vlack

TABLE 3. Major changes in the NMC/NCEP regional modeling and data assimilation since 1985 (from compilations by F. Mesinger and G. DiMego 1998, personal communication).

Year	Operational model	Computer
Mar 1985	RAFS based on the triply NGM (Phillips 1979) and OI (DiMego 1988). Resolution: 80 km/16 layers.	Cyber 205
Aug 1991	RAFS upgraded for the last time: NGM run with only two grids with inner grid domain doubled in size. Implemented Regional Data Assimilation System with 3-hourly updates using an improved OI analysis using all off-time data including Profiler and ACARS wind reports (DiMego et al. 1992) and complex quality control procedures (Gandin et al. 1993).	Cray YMP 8 processors 32 megawords
Jun 1993	First operational implementation of the Eta Model in the 0000 and 1200 UTC early run for North America at 80-km and 38-layer resolution (Mesinger et al. 1988; Janjic 1994; Black et al. 1993).	
Sept 1994	The RUC (Benjamin et al. 1994) was implemented for CONUS domain with 3-hourly OI updates at 60-km resolution on 25 hybrid (sigma-theta) vertical levels.	Cray C-90 16 processors 128 megawords
Sept 1994	Early Eta analysis upgrades (Rogers et al. 1995).	
Aug 1995	A mesoscale version of the Eta Model (Black 1994) was implemented at 0300 and 1500 UTC for an extended CONUS domain, with 29-km and 50-layer resolution and with NMC's first predictive cloud scheme (Zhao and Carr 1997) and new coupled land-surface-atmosphere package (two-layer soil).	Cray C-90 16 processors 256 megawords
Oct 1995	Major upgrade of early Eta runs: 48-km resolution, cloud scheme, and Eta Data Assimilation System (EDAS) using 3-hourly OI updates (Rogers et al. 1996).	
Jan 1996	New coupled land-surface-atmosphere scheme put into early Eta runs (Chen et al. 1997; Mesinger 1996).	
Jul-Aug 1996	Nested capability demonstrated with twice-daily support runs for Atlanta Olympic Games with 10-km 60-layer version of Meso Eta.	
Feb 1997	Upgrade package implemented in the early and Meso Eta runs.	
Feb 1998	Early Eta runs upgraded to 32 km and 45 levels with four soil layers. OI analysis replaced by 3D-VAR with new data sources. EDAS now partially cycled (soil moisture, soil temperature, cloud water/ice, and turbulent kinetic energy) (Parrish et al. 1996; Rogers et al. 1998)	
Apr 1998	RUC (3-hourly) replaced by hourly RUC II system with extended CONUS domain, 40-km and 40-level resolution, additional data sources, and extensive physics upgrades.	
Jun 1998	Mesoscale Eta runs 4 times a day for North American domain at 32 km and 45 level resolution, using new snow analysis. All runs started from EDAS which is fully cycled for all variables.	

The 36-h forecasts of 500 hPa, which indicate the position and intensity of the large-scale atmospheric waves and centers of high and low pressure, are generally excellent, as suggested by the nearly perfect S1 score.

However, smaller-scale atmospheric structures, such as fronts, mesoscale convective systems, etc., which are more apparent in surface maps than at 500 hPa, are still difficult to forecast in detail, and hence they have

a poorer S1 score (Fig. 1b). The solid line with circles starts in 1947 with scores from subjectively made surface forecasts, then barotropic and baroclinic quasigeostrophic models (Table 1), the LFM and, since 1983, the global spectral model (denoted aviation or AVN). Other model forecasts are also presented separately in Fig. 1b. Note that the global spectral model (AVN) and

the Eta Model, which continue to be developed, show the most improvement. The development of the LFM was frozen in 1986, and that of the NGM in 1991, when more advanced systems were implemented and therefore their forecasts show no further improvement with time (except for the effect of improved global forecasts used as a first guess for the LFM).

### 500MB RMS FITS TO RAWINSONDES 6 HR FORECASTS

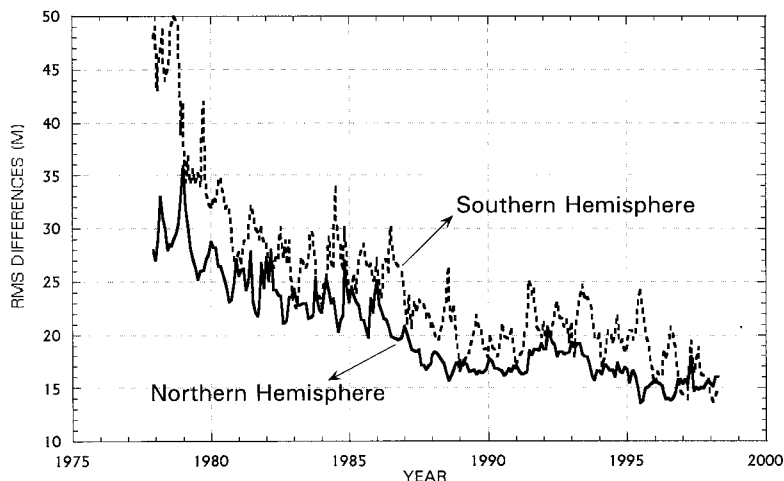


FIG. 2. Rms difference between 6-h forecast of the 500-hPa heights and collocated rawinsonde measurements for the NH and the SH (data courtesy of C. Vlecek and L. Morone).

### THREAT SCORES: day 1 and day 2 for 0.5' and 1' 24 hr accumulation, annual ave. (from NCEP/HPC)

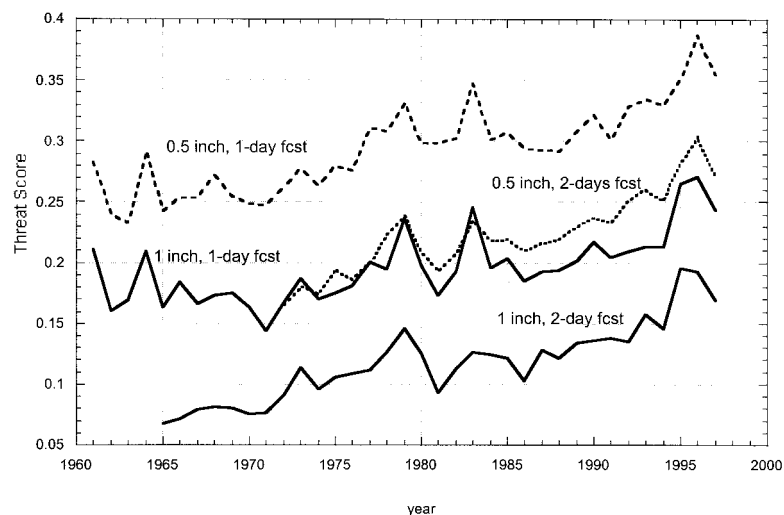


FIG. 3. Annual average of threat scores of forecasts made by expert forecasters from the NCEP Hydrometeorological Prediction Center. Forecast lengths: 1- and 2-day; threshold: 0.5 and 1 in. (adapted from data courtesy of D. Reynolds).

There are many other scores that also measure forecast skill improvement. Figure 2 presents the rms difference between the 6-h forecast (which is used as a first guess) and the rawinsonde observations from 1978 to the present. It should be noted that the rms differences are not necessarily forecast errors, since the observations also contain errors. In the Northern Hemisphere (NH) the rms differences have been halved from about 30 m in the late 1970s to about 15 m at present, equivalent to a mean temperature error of about 0.75 K, not much larger than rawinsonde observational errors. In the Southern Hemisphere (SH) the improvements are even larger, with the differences decreasing from about 47 m to about 17 m, close to the present forecast error in the NH.

Figure 3 shows threat scores for precipitation predictions made by expert forecasters from the NCEP Hydrometeorological Prediction Center (HPC, the Meteorological Operations Division of the former NMC). The threat score (TS) is defined as the intersection of the predicted area of precipitation exceeding a particular threshold (P), in this case 0.5 in. in 24 h, and the observed area (O), divided by the union of the two areas:  $TS = (P \cap O) / (P \cup O)$ . The bias (not shown) is defined by  $P/O$ . In other words, the threat score is given by the area where precipitation was correctly predicted, divided by the area where either the model or the observations or both indicate precipitation above a given threshold. The TS, also known as critical success index (CSI), is a particularly useful score for quantities that are relatively rare. Figure 3 indicates that the forecaster's skill in predicting accumulated precipitation has been increasing with time and that the

current average skill in the 2-day forecast is as good as the 1-day forecasts were in the 1970s. Although the TS are better for the winter than for the summer, both warm and cold season scores have shown similar improvements (not shown). Beyond the first 6–12 h, the forecasts are based mostly on numerical guidance, so that the improvement reflects to a large extent improvements of the numerical forecasts, which the human forecasters in turn improve upon based on their knowledge and expertise. The forecasters also have access to several model forecasts, and they use their judgment in assessing which one is more accurate in each case. This constitutes a major source of the value added by the human forecasters.

The relationship between the evolution of human and numerical forecasts is clearly shown in a record compiled by the late F. Hughes (Hughes 1987), reproduced in Fig. 4. It is the first operational score maintained for the “medium range” (beyond the first two days of the forecasts). The score used by Hughes was a standardized anomaly correlation (SAC), which accounted for the larger variability of sea level pressure at higher latitudes compared with lower latitudes. Unfortunately the SAC is not directly comparable to other scores such as the anomaly correlation (discussed in the next section). The fact that until 1976 the 3-day forecast scores from the model were essentially constant is an indication that the rather low skill of human forecasts was more based on synoptic experience than on model guidance. The forecast skill started to improve after 1977 for the 3-day forecast and after 1980 for the 5-day forecast. Note that the human forecasts are on the average significantly more skillful than the numerical guidance, but it is the improvement in NWP forecasts that drives the improvements in the subjective forecasts.

### 3. Operational and reanalysis forecasts: Impact of the observing systems

Numerical weather prediction is an initial value problem, in which the computer models of the atmosphere start from the best estimate of the state of the atmosphere and are integrated forward in time, thus simulating (predicting) the

evolution of the atmosphere. Therefore, the numerical forecasts can be improved by either improving the models or the estimate of the initial conditions. Both types of improvements are necessary and are made possible only by the availability of more powerful computers. For example, increasing the spatial resolution (Tables 1–3) significantly increases the accuracy of atmospheric models, but every time the grid sizes are halved the computational requirement increases by a factor of  $2^4 = 16$ , since the number of grid points is doubled in each space direction, and the number of time steps also has to be doubled. Similarly, in order to be realistic, the model has to have accurate representations of the net effect of small-scale physical processes, such as clouds, convection, precipitation, radiation, and turbulent transports of heat and moisture (see Tables 1–3 for the major implementations of representation of physical processes). When these parameterizations of the “physics” are made more realistic, and more processes are included, they usually become more computationally expensive.

The same is true about the estimation of the initial state of the atmosphere (denoted atmospheric analysis). In modern analysis schemes the state of the atmosphere is estimated by a statistical interpolation between a short-range forecast (first guess) and the at-

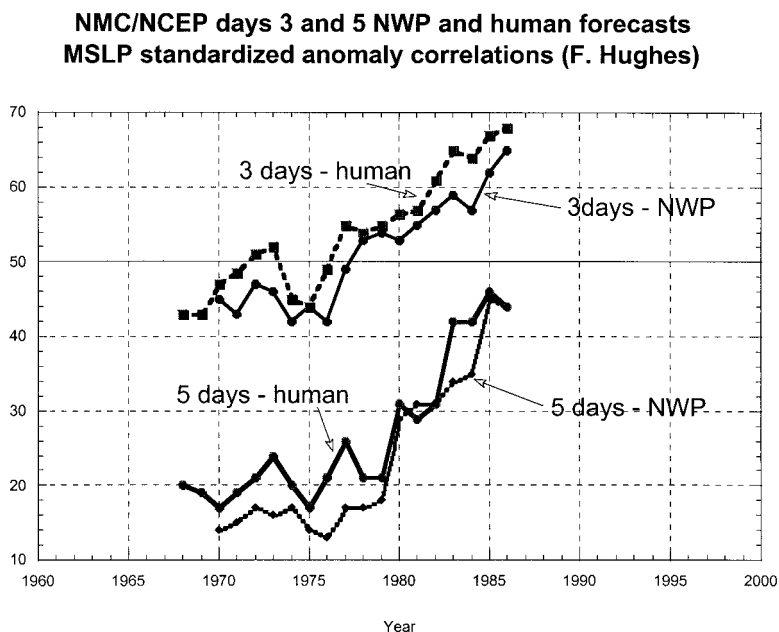


FIG. 4. Comparison of the skill of numerical weather prediction guidance and the forecasts made by expert members of the NCEP/HPC using the numerical guidance, adapted from a compilation by Hughes (1987). The score is the sea level pressure anomaly correlation over North America, where the correlations are standardized to account for the latitudinal dependence of the variability.

mospheric observations. The 3D-VAR method operational in both the regional and the global NCEP systems requires the solution of systems of equations with matrices of the order of the number of degrees of the models (about  $10^6$ ). The 4D-VAR system implemented at the European Centre for Medium-Range Weather Forecasts (ECMWF) is much costlier and is made possible by their access to a larger super-computer.

The four-dimensional operational analysis of the atmosphere is a very convenient tool for climate researchers, who can use these uniform gridded fields instead of manually interpolating observations. However, every time the model or the data assimilation system is improved, a change in the estimate of the state of the atmosphere is also introduced, and for climate researchers this appears as a spurious climate change. For this reason, "reanalysis" projects (in which the data assimilation system is kept unchanged) were carried out at NCEP (jointly with NCAR), ECMWF, and NASA/Goddard (Schubert et al. 1993; Kalnay et al. 1996; Gibson et al. 1997). The NCEP-NCAR reanalysis used a data assimilation/model system identical to the one implemented for the global system in January 1995 (Table 3) but with half its horizontal resolution, equivalent to about 200 km. Forty years of data (1958-97) have been reanalyzed at NCEP, producing an estimate of the state of the atmosphere four times a day<sup>3</sup> (Kalnay et al. 1996; R. Kistler et al. 1998, personal communication). Eight-day-long forecasts were also recomputed every 5 days of reanalysis. In these reforecasts, the model and analysis scheme are kept constant for 40 years, so systematic changes in the forecast skill (beyond those produced by natural interannual variations in atmospheric predictability) must be due to changes in the observing systems.

Figure 5 compares the anomaly correlation (AC) for the 5-day forecasts over the NH north of 20°N for the operational and reanalysis forecasts. Anomalies are defined as the difference between an analysis of forecast field and the corresponding climatology for that month. The AC is the pattern correlation between the forecast anomalies and the verifying analysis anomalies. Experience indicates that a forecast is useful only if the AC is greater than about 60%. The SAC that Hughes defined over North America, although not strictly comparable because of the standardization and

the use of a different verification domain, has been plotted as well to provide a qualitative extension to the 1970s, for which regular NH ACs are not available. (Note that the left scale in Fig. 5 corresponds to NH AC and the one on the right to the North American SAC).

The 5-day forecasts from the reanalysis are skillful (have AC above 60%) as early as 1958. The improvements in the observing systems (the only factor affecting the reanalysis) increase the AC from about 65% in the 1950s and 1960s to about 70% in the present day. The operational 5-day forecasts, meanwhile, were of little use until the 1980s, and since then their skill increased to a level of about 75% at the present time. The operational forecasts are better than those of the reanalysis starting in 1991, when the operational resolution was increased to T126 (equivalent to about 100 km), compared to 200-km resolution used in the reanalysis.

Figure 6 includes the AC scores for both the NH and SH reanalysis and operational 5-day forecasts. It is clear that the impact of the changes in the observing systems (especially the introduction of the satellite observing system) is much larger in the SH than in the NH, because in the SH there are very few rawinsondes compared to the NH. The reanalysis used satellite temperature retrievals (VTPR) produced by NOAA/NESDIS for the SH starting in 1975. The first

**NCEP 5-day 500hPa anomaly correlation**  
**NH Reanalysis, NH Operational and N. America standardized**

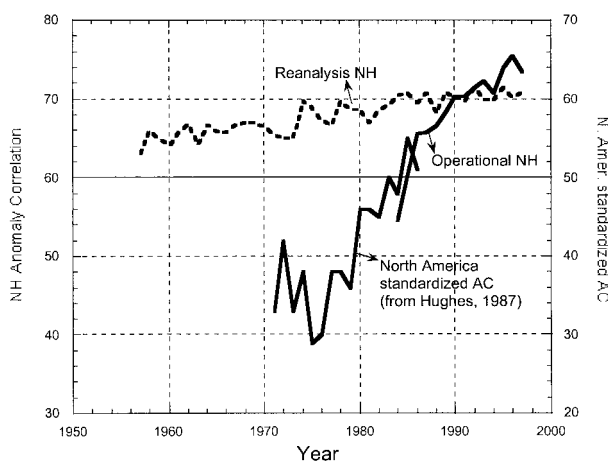


FIG. 5. Comparison of the NH anomaly correlation from operational and reanalysis forecasts (scale on the left). The standardized anomaly correlation scores of Hughes (1987), although not exactly comparable, are also displayed for comparison with earlier periods (scale on the right). Adapted from Hughes (1987) and data courtesy of P. Caplan.

<sup>3</sup>In 1998, the reanalysis of the previous decade (1948-57) was also completed (R. Kistler et al. 1998, personal communication).

Television Infrared Observational Satellite (TIROS) Operational Vertical Sounder (TOVS) retrievals started in December 1978 and have continued to the present. The marked improvements in the reanalysis forecasts for the SH from 1974 to the present are mostly due to better operational retrievals from NESDIS, as well as some improvements in the cloud-tracked winds. As a result, the reanalysis 5-day forecasts start to have useful skill (AC above 60%) around 1987, but the operational SH 5-day forecasts (made in real time) reached 60% starting in 1990.

Although significant improvements in observing systems have taken place, by far the largest gains in forecast skill have come from data usage. In modern methods of data assimilation (3D-VAR, 4D-VAR, and Kalman filtering), there has been a major shift from using “retrieved” atmospheric parameters, such as satellite temperature and moisture soundings, to using the observed variables directly, such as satellite channel radiances. At NCEP, a 3D-VAR data assimilation system was introduced into operations in June 1991 (Parrish and Derber 1992; Derber et al. 1991), and this allowed the implementation, in October 1995, of the first assimilation of TOVS radiances within a 3D-VAR approach (Derber and Wu 1998). In 1994 ECMWF implemented a 1D-VAR system of assimilation of TOVS radiances, a 3D-VAR data assimilation system in January 1995, and a 4D-VAR system in late 1997. The assimilation of radiances is, indeed, a much more effective use of the TOVS data, and the large improvements observed in 1996–97 operational forecasts compared to the reanalysis forecasts, in both the NH and the SH, are primarily due to this implementation (Fig. 6). It should be noted that this was the first time in which large positive impacts from satellite data were obtained in the Northern Hemisphere (17 years after the launch of *TIROS-N* in 1978, which carried the first operational TOVS).

#### 4. Ensemble forecasting and seasonal forecasting

Another revolution in operational NWP in recent years has been the implementation of ensemble fore-

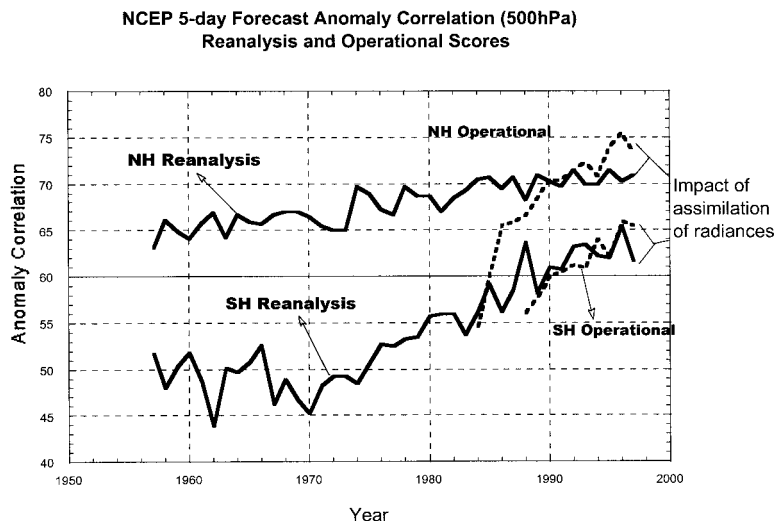


FIG. 6. Comparison of operational and reanalysis 5-day forecast anomaly correlations for the NH and the SW. The large improvement in operational forecasts observed in 1996–97 is due, to a large extent, to the direct assimilation of TOVS radiances (data courtesy of R. Kistler).

casting in December 1992 at both NCEP and ECMWF (Tracton and Kalnay 1993; Toth and Kalnay 1993; Palmer et al. 1993; Toth et al. 1997; Molteni et al. 1996; Toth and Kalnay 1997). In 1992, NCEP started operational ensemble forecasting with a set of five daily global forecasts integrated to 10 days. Since 1994 NCEP has been running 17 global forecasts per day, each out of 16 days, with initial perturbations obtained using the “breeding” method. This method ensures that the initial perturbations contain leading Lyapunov vectors, the fastest naturally growing dynamical perturbations in the atmosphere, which are also present in the analysis errors. The length of the forecasts will allow the generation of “outlooks” for week 2, instead of the present 6–10 days. The NCEP ensemble forecasts can be accessed on the World Wide Web at the Environmental Modeling Center Web site ([nic.fb4.noaa.gov:8000](http://nic.fb4.noaa.gov:8000)), linking to the ensemble Web site. All the NCEP centers can also be accessed through this link.

Ensemble forecasting has accomplished two main goals. The first one is to provide an ensemble average forecast that beyond the first few days is more accurate than individual forecasts, because the components of the forecast that are most uncertain tend to be averaged out. The second and most important goal is to provide forecasters with an estimation of the reliability of the forecast, which, because of changes in atmospheric predictability, varies from day to day and from region to region. Additional applications of the en-

Anomaly Correlation for the winter of 1997/98 (from 00Z)  
Controls (T126 and T62) and 10 perturbed ensemble forecasts

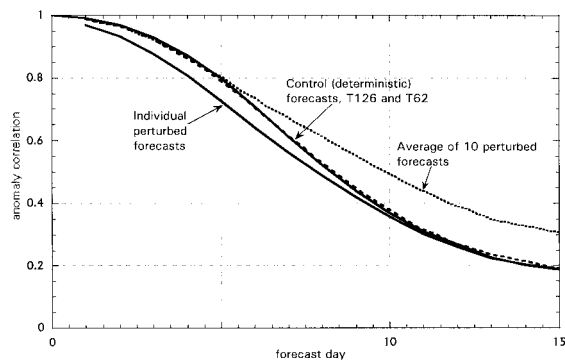


FIG. 7. Anomaly correlation as a function of time, averaged for the winter of 1997/98 (all forecasts started at 0000 UTC). It shows the control deterministic forecast at T126 and T62 resolution, the average AC of the individual perturbed forecasts, and the AC of the ensemble average of the 10 perturbed forecasts (data courtesy of J. Schemm and H. van den Dool).

semble system are estimates of the uncertainty of the forecast used in the analysis (Pu et al. 1997a) and guidance for adaptive observations.

The first goal is illustrated in Fig. 7, prepared at the Climate Prediction Center (CPC, the Climate Analysis Center of the former NMC) for the verification of the NCEP ensemble during the winter of 1997/98. This was an El Niño winter with major anomalies in the atmosphere circulation, and the operational forecasts had excellent skill. The control “deterministic” forecast had an AC in the 5-day forecast of 80%, which is quite good. Individually, the 10 perturbed ensemble members have a poorer verification with an average AC of about 73% at 5 days. This is because in the initial conditions, the control starts from the best estimate of the state of the atmosphere (the analysis), but growing perturbations are added to this analysis for each additional ensemble member. However, the ensemble average forecast tends to average out uncertain components, and, as a result, starting at day 5, it has much better skill than the control forecast. Note that the ensemble extends by one day the length of the useful forecast (AC greater than 60%), from about 7 days in the control to about 8 days in the ensemble average.

The second and most important goal of the ensemble forecasting, to provide human forecasters with guidance on the uncertainty of each forecast, is accomplished best by the use of two types of plots. The “spaghetti” plots show a single contour line for all 17 forecasts, and the probabilistic plots show, for example, what percentage of the ensemble predicts 24-h

accumulated precipitation of more than 1-in. at each grid point (for probabilistic quantitative precipitation forecasts or pQPF). Both types of plots provide an easy-to-understand depiction of the reliability of the forecasts. The use of the ensembles has provided the U.S. NWS forecasters with the confidence to issue storm forecasts 5–7 days in advance when the spaghetti plots indicate good agreement in the ensemble. Conversely, the spaghetti plots also indicate when a short-range development may be particularly difficult to predict, so that the users should be made aware of the uncertainty of the forecast. Figure 8a shows an example of the 5-day forecast for 15 November 1995, the first East Coast winter storm of 1995/96: the fact that the ensemble showed good agreement provided the forecasters with the confidence to issue a storm forecast many days in advance. By contrast, Fig. 8b shows a 2.5-day forecast for a storm verifying on 21 October 1995, and it is clear that even at this shorter range the atmosphere is much less predictable and there is much more uncertainty in the location of the storm.

The use of ensembles has also led to another major development, the possibility of an adaptive observing system. As an example, consider a case in which the lack of agreement among the ensemble members indicates that a 3-day forecast in a certain region is exceedingly uncertain, as in Fig. 8b. Recently three new techniques have been developed to trace such a region of uncertainty backward in time, for example, 2 days (Bishop and Toth 1996; Rabier et al. 1996; Rohaly et al. 1998; Pu et al. 1997b; Toth et al. 1997; Pu et al. 1998). These techniques will point to a region or regions where additional observations would be especially useful. The additional observations could be dropwindsondes launched from a reconnaissance or a pilotless airplane, additional rawinsondes, or especially intensive use of satellite data such as a Doppler wind lidar (D. Emmitt 1998, personal communication). If additional observations are available within 24 h after the start of the originally critically uncertain 3-day forecast, they can substantially increase the usefulness of the 2-day forecast. Similarly, a few additional rawinsondes could be launched in critical areas where short-range ensemble forecasts (12–24 h) indicate that they are most needed. Preliminary tests of this approach of targeted observations have been successfully performed within an international Fronts and Storm Track Experiment (FASTEX) in the North Atlantic during January and February 1997, and in the North Pacific Experiment (NORPEX) in January–February 1998.

Ensemble forecasting also provides the basic tool to extend forecasts beyond Lorenz's 2-week limit of weather predictability. Slowly varying surface forcing, especially from the tropical ocean and from land surface anomalies, can produce atmospheric anomalies that are longer lasting and more predictable than individual weather patterns. The most notable of these is the ENSO produced by unstable oscillations of the coupled ocean-atmosphere system, with a frequency of 3–7 yr. Because of their long period, the ENSO oscillations can be predicted a year or more in advance, as shown first by Cane et al. (1986). The warm phases of ENSO (El Niño) episodes are associated with warm sea surface temperature (SST) anomalies in the equatorial central Pacific Ocean, and cold phases (La Niña episodes) with cold anomalies. A single atmospheric forecast forced with the SST anomalies would not be useful beyond the first week or so, when unpredictable weather variability would mask the forced atmospheric anomalies. Ensemble averaging many forecasts made with atmospheric models forced by SST anomalies (and by other slowly varying anomalies over land such as soil moisture and snow cover) allows the unpredictable components of the forecast to be filtered out and more of the forced predictable components to be retained. This filtering is reflected in the fact that the ensemble average for the second week of the forecasts for the winter of 1997/98 (Fig. 7) had a high AC of 57%, much higher than previously obtained. Researchers at the Japanese

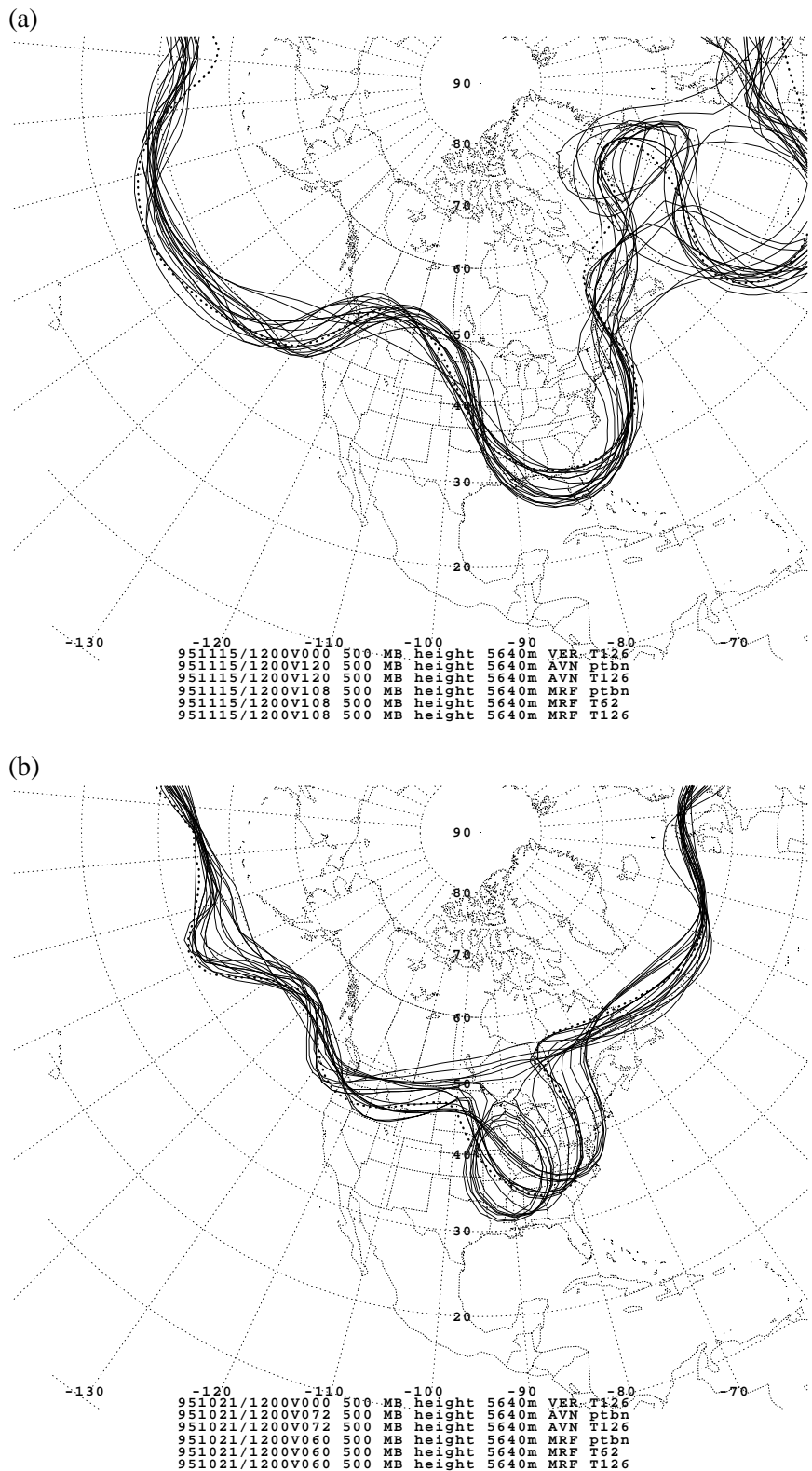


FIG. 8. (a) Spaghetti plot for the 4.5- and 5.0-day-long forecasts verifying on 15 November 1995. (Figure courtesy of D. Wobus.) (b) Spaghetti plot for the 2.5- and 3.0-day-long forecasts verifying on 21 October 1995. (Figure courtesy of D. Wobus.)

Meteorological Agency have performed forecasts for the 28-day average and also found that ensemble average substantially increased the information on the second week and the last 2 weeks of the forecast. The very successful operational forecasts of the ENSO episode of 1997/98 performed both at NCEP and at ECMWF have been substantially based on the use of ensembles to extract the useful information on the impact of El Niño from the “weather noise.”

## 5. Summary and discussion

NWP in the medium range has experienced a long period of steady progress, beginning before the first operational 72-h hour forecast was made at NMC in 1975. This period can be characterized as an era of discovery, during which progress has stemmed from three major ingredients: ever-increasing computer power (Table 4), improvements to observing systems, and, most importantly, the human dedication to scientific advances in numerical models and usage of available observations. By the same token, future improvements will require continued dedication to development of new scientific techniques and appropriate resources for all ingredients (Eliassen et al. 1954).

It seems to us unlikely that the era of discovery is over. There are still improvements likely in all three areas, and no end in sight to the advances in computational capability that have enabled past improvements in NWP and are essential for future improvements. In the area of observations, the greatest need is for wind profiles over the world's oceans and sparsely observed landmasses such as Africa. Recent experiments at NCEP and other centers have confirmed that for the NH extratropics, the component of the observing system that contributes the most to the forecast skill is the rawinsonde wind profiles, which are very sparse in other regions of the world. Not far behind is the need for improved measurements of atmospheric moisture, especially for short-range forecasting. The great spatial and temporal variability of moisture, and our lack of observations, are major obstacles to improved precipitation prediction. Global positioning system (GPS) measurements that are affected by the humidity may provide moisture estimates with the temporal and horizontal resolution required for short-range forecasting. Extensive experimentation will further define the contributions of each observing system (e.g., rawinsondes, aircraft, satellites) over

longer periods of time. This input can be used to design and implement future observing systems on a cost-effective basis.

In the area of data assimilation, the advent of 4D-VAR and other advanced techniques will enable still further extraction of useful information from the observing systems. And, in modeling, there is still room for improvement through increased resolution, especially in the shorter ranges, and in more advanced representation of smaller-scale phenomena such as turbulent and cloud microphysical processes. For the medium and longer ranges, however, increased computer capacity is more likely to be deployed in larger ensembles, including ensembles of different models as well as perturbed initial conditions. The needs of society will continue to demand useful forecasts extended to longer ranges, including one month to seasonal, and improved accuracy for special weather threats such as hurricanes and floods. For example, the introduction of the GFDL hurricane forecasting system at NCEP and the U.S. Navy Fleet Numerical Meteorology and Oceanography Center (FNMOC) in 1995 produced marked gains in hurricane track prediction, but much progress is needed to produce useful intensity forecasts.

We foresee the next major extension of NWP science to be in the area of operational prediction of atmospheric constituents. As human society places more demands on the atmosphere, it will become increasingly important to provide decision makers in government and industry with forecasts of chemical and particulate loading of the atmosphere. NCEP has taken its first step in this direction by including ozone as a predictive variable in its global spectral model. Atmospheric chemistry must now be incorporated in the models along with atmospheric physics, and this has important implications for both observing and data assimilation.

Another area of discovery that remains to be exploited is the complete coupling of land (including biological) and hydrological processes with oceanic and atmospheric processes. The prediction of the full hydrological cycle, including evaporation, transpiration, atmospheric transport, precipitation, and runoff, is becoming a reality, but poses a formidable scientific challenge. We envision that in the next decades, river flow forecasting will become fully coupled with the NWP forecasts. The hydrological applications of precipitation forecasts to floods and to water resources management will continue to grow and produce societal benefits. The coupling of NWP models must also

TABLE 4. Characteristics of the operational computers at NMC/NCEP (courtesy of M. McCracken). Note that the peak performance is several times larger than the actual performance attained with optimized codes.

Year	Computer characteristics	Computer
1955	Vacuum tube computer of less than 1Kflop designed exclusively for business applications	IBM 701
1958	Vacuum tube computer of about 8Kflops	IBM 704
1960	This model was capable of about 67Kflops	IBM 7090
1963	The 7094 was transistorized and capable of 100Kflops	IBM 7094
1966	The CDC 6600 had multiple functional arithmetic processing units and were capable of 3 Mflops peak performance	CDC 6600
1972–1973	Two systems were installed in 1972, and a third transferred in 1973 from GFDL (which moved to Texas Instruments ASC); the performance of each system was about 10Mflops	IBM 360/195
1978	Two-pipe system with a theoretical peak performance of 400Mflops in single precision; the well-optimized NMC spectral model achieved little over 100Mflops	CDC Cyber 205
1987	A second identical system was acquired, doubling the attainable performance to about 200Mflops	Second CDC Cyber 205
1990	The eight-processor vector machine had a peak performance of 2.6 Gflops	CrayYMP, 8CPU/32MB memory
1992	The memory was increased to 64 MB	CrayYMP, 8CPU/64MB memory
1994	The 16-processor C90 is rated at a peak performance of 15.3 Gflops	Cray C90, 16CPU/128MB memory

be extended upward in the atmosphere to include processes of the near-space environment. Again, this has important implications for observing, data assimilation, and modeling.

While we are confident that the era of discovery will continue unabated for the foreseeable future, we also observe that numerical weather prediction in the medium range has now achieved a level of maturity that is ushering in a new era of applications. We believe that NWP science and practice will be the basis for an explosion of applications in daily decision making not only by individual citizens but also by operating officers in weather-vulnerable economic sectors. Many of the latter will be engaged in very high stakes decisions, with a cumulative total of hundreds of billions of dollars at stake each year. For example, operational forecasts are already being used for decisions in commodities futures exchanges.

It seems likely to us that, because of these activities with extremely large economic value, operational NWP will evolve from what has been almost exclusively a government activity to one that also has an important private sector component. This, we believe, will further stimulate both discovery and applications.

*Acknowledgments.* In this paper we have reported progress in operational numerical weather prediction at NCEP. We would like to express our gratitude to all the people at NCEP and elsewhere whose work contributed to this progress, and especially to those who started and maintained statistics of verification of skill. We gratefully acknowledge Chuck Vlcek, Lauren Morone, James Hoke, David Reynolds, Robert Kistler, Jae Schemm, Huug van den Dool, Richard Wobus, and Michiko Masutani for providing us with the NCEP forecast verifications statistics and examples presented here. Pete Caplan, Fedor Mesinger, and Geoffrey DiMego provided information about operational changes in the global and regional systems.

## References

- Benjamin, S. G., J. M. Brown, K. J. Brundage, D. Devenyi, B. Schwartz, T. G. Smirnova, T. L. Smith, and F.-J. Wang, 1996: The 40-km 40-level version of MAPS/RUC. Preprints, *11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., 161–163.
- Bergman, K. H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.*, **107**, 1423–1444.
- Bergthorsson, P., and B. Doos, 1955: Numerical map analysis. *Tellus*, **7**, 329–340.
- Bishop, C., and Z. Toth, 1996: Using ensembles to identify observations likely to improve forecasts. Preprints, *11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., 72–74.
- Black, T. L., 1994: The new NMC mesoscale Eta model: Description and forecast examples. *Wea. Forecasting*, **9**, 265–278.
- Black, T., D. Deaven, and G. DiMego, 1993: The step-mountain eta coordinate model: 80 km ‘Early’ version and objective verifications. NWS Tech. Procedures Bull. 412, National Oceanic and Atmospheric Administration/National Weather Service, 31 pp. [Available from National Weather Service, Office of Meteorology, 1325 East–West Highway, Silver Spring, MD 20910.]
- Brooks, H. E., A. Witt, and M. D. Ellis, 1997: Verification of public weather forecasts available via the media. *Bull. Amer. Meteor. Soc.*, **78**, 2167–2177.
- Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature*, **321**, 827–832.
- Caplan, P., and G. White, 1989: Performance of the National Meteorological Center’s medium-range model. *Wea. Forecasting*, **4**, 391–400.
- Charney, J. G., 1951: Dynamical forecasting by numerical process. *Compendium of Meteorology*, T. F. Malone, Ed., Amer. Meteor. Soc., 470–482.
- , 1954: Numerical prediction of cyclogenesis. *Proc. Natl. Acad. Sci. U.S.*, **40**, 99–110.
- , R. Fjørtoft, and J. von Neuman, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237–254.
- Chen, F., Z. Janjic, and K. Mitchell, 1997: Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta Model. *Bound.-Layer Meteor.*, **85**, 391–421.
- Collins, W. G., 1998: Complex quality control of significant level radiosonde temperatures. *J. Atmos. Oceanic Technol.*, **15**, 69–79.
- , and L. S. Gandin, 1990: Comprehensive hydrostatic quality control at the National Meteorological Center. *Mon. Wea. Rev.*, **118**, 2752–2767.
- Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374.
- Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299.
- , D. F. Parrish, and S. J. Lord, 1991: The new global operational analysis system at the National Meteorological Center. *Wea. Forecasting*, **6**, 538–547.
- DiMego, G. J., 1988: The National Meteorological Center regional analysis system. *Mon. Wea. Rev.*, **116**, 977–1000.
- , K. E. Mitchell, R. A. Petersen, J. E. Hoke, J. P. Gerrity, J. J. Tuccillo, R. L. Wobus, and H.-M. H. Juang, 1992: Changes to NMC’s regional analysis and forecast system. *Wea. Forecasting*, **7**, 185–198.
- Eliassen, A., J. S. Sawyer, and J. Smagorinsky, 1954: Upper air network requirements for numerical weather predictions. World Meteorological Organization Tech. Note 29, 16 pp.
- Flattery, T. W., 1971: Spectral models for global analysis and forecasting. Tech. Rep. 243, Air Weather Service, U.S. Air Force, 42–54. [Available from the National Centers for Environmental Prediction, 5200 Auth Rd., Rm. 100, Camp Springs, MD 20746.]
- Gandin, L. S., 1988: Complex quality control of meteorological observations. *Mon. Wea. Rev.*, **116**, 1137–1156.
- , L. L. Morone, and W. G. Collins, 1993: Two years of operational comprehensive hydrostatic quality control at the NMC. *Wea. Forecasting*, **8**, 57–72.
- Gibson, J. K., P. Kallberg, S. Uppala, A. Nomura, A. Hernandez, and E. Serrano, 1997: ECMWF reanalysis description. ECMWF Reanalysis Report Series 1, 50 pp. [Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom.]
- Hong, S.-Y., and H.-L. Pan, 1996: Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.*, **124**, 2322–2339.
- Howcroft, J. G., 1971: Local forecast model: Present status and preliminary verification. NMC Office Note 50, NWS, NOAA, U.S. Dept. of Commerce, 22 pp. [Available from the National Centers for Environmental Prediction, 5200 Auth Rd., Rm. 100, Camp Springs, MD 20746.]
- Hughes, F. D., 1987: Skill of medium-range forecast group. NMC Office Note 326, NWS, NOAA, U.S. Dept. of Commerce.
- Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Kanamitsu, M., and Coauthors, 1991: Recent changes in the global forecast system at NMC. *Wea. Forecasting*, **6**, 425–435.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- , 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333.
- , 1968: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307.
- Mesinger, F., 1996: Improvements in quantitative precipitation forecasts with the eta regional model at the National Centers for Environmental Prediction: The 48-km upgrade. *Bull. Amer. Meteor. Soc.*, **77**, 2637–2649; Corrigendum, **78**, 506.
- , Z. I. Janjic, S. Nickovic, D. Gavrilov, and D. G. Deaven, 1988: The step-mountain coordinate: Model description and performance for cases of Alpine lee cyclogenesis and for a case of an Appalachian redevelopment. *Mon. Wea. Rev.*, **116**, 1493–1518.
- Molteni, F., and T. N. Palmer, 1993: Predictability and finite time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.*, **119**, 269–298.
- , R. Buizza, T. N. Palmer, and T. Petroliaigis, 1996: The ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119.

- Palmer, T. N., F. Molteni, R. Mureau, R. Buizza, P. Chapelet, and J. Tribbia, 1993: Ensemble Prediction. *Proc., ECMWF Seminar on Validation of Models over Europe*, Vol. 1, 285 pp. [Available from ECMWF, Shinfield Park, Reading, RG2 9AX, United Kingdom.]
- Pan, H.-L., 1990: A simple parameterization scheme of evapotranspiration over land for the NMC Medium-Range Forecast Model. *Mon. Wea. Rev.*, **118**, 2500–2512.
- , and L. Mahrt, 1987: Interaction between soil hydrology and boundary-layer development. *Bound.-Layer Meteor.*, **38**, 185–202.
- , and W.-S. Wu, 1995: Implementing a mass flux convection parameterization package for the NMC Medium-Range Forecast Model. NMC Office Note 409, NWS, NOAA, U.S. Dept. of Commerce, 43 pp.
- Parrish, D. F., and J. D. Derber, 1992: The National Meteorological Center spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- , J. Purser, E. Rogers, and Y. Lin, 1996: The regional 3D-variational analysis for the Eta Model. Preprints, *11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., 454–455.
- Phillips, N. A., 1979: The nested grid model. NOAA Tech. Rep. NWS 22, U.S. Dept. of Commerce, 80 pp. [Available from National Centers for Environmental Prediction, 5200 Auth Rd., Rm. 100, Camp Springs, MD 20746.]
- , 1990: Dispersion processes in large-scale weather. World Meteorological Organization, 700, 126 pp. [Available from World Meteorological Organization, Avenue Giuseppe-Motta, 41, CP2300, 1211 Geneva 2, Switzerland.]
- , 1998: Carl-Gustaf Rossby: His times, personality, and actions. *Bull. Amer. Meteor. Soc.*, **79**, 1097–1112.
- Pu, Z.-X., E. Kalnay, D. Parrish, W. Wu, and Z. Toth, 1997a: The use of the bred vectors in the NCEP operational 3-dimensional variational system. *Wea. Forecasting*, **12**, 689–695.
- , —, J. Sela, and I. Szunyogh, 1997b: Sensitivity of forecast errors to initial conditions with a quasi-inverse linear model. *Mon. Wea. Rev.*, **125**, 2479–2503.
- , S. J. Lord, and E. Kalnay, 1998: Forecast sensitivity with dropwindsonde data and targeted observations. *Tellus*, **50A**, 391–410.
- Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.*, **122**, 121–150.
- Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, 236 pp.
- Rogers, E., D. G. Deaven, and G. J. DiMego, 1995: The regional analysis system for the operational “early” Eta Model. *Wea. Forecasting*, **10**, 810–825.
- , T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. Baldwin, N. W. Junker, and Y. Lin, 1996: Changes to the operational “early” eta analysis forecast system at the National Centers for Environmental Prediction. *Wea. Forecasting*, **11**, 319–413.
- , D. Parrish, and G. J. DiMego, 1998: Data assimilation experiments with the regional 3-D variational analysis at the National Centers for Environmental Prediction. Preprints, *12th Conf. on Numerical Weather Prediction*, Phoenix, AZ, Amer. Meteor. Soc., 32–33.
- Rohaly, G. D., R. H. Langland, and R. Gelaro, 1998: Identifying regions where the forecast of tropical cyclone tracks is most sensitive to initial condition uncertainty using adjoint methods. Preprints, *12th Conf. on Numerical Weather Prediction*, Phoenix, AZ, Amer. Meteor. Soc., 337–340.
- Rossby, C.-G., 1940: Planetary flow patterns in the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **66** (Suppl.), 68–87.
- , 1945: On the propagation of frequencies and energies in certain types of oceanic and atmospheric waves. *J. Meteor.*, **2**, 187–204.
- , and Coauthors, 1939: Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semi-permanent centers of action. *J. Mar. Res.*, **2**, 38–55.
- Schubert, S., R. Rood, and J. Pfaendtner, 1993: An assimilated data set for earth science applications. *Bull. Amer. Meteor. Soc.*, **74**, 2331–2342.
- Sela, J. G., 1980: Spectral modeling at the National Meteorological Center. *Mon. Wea. Rev.*, **108**, 1279–1292.
- Shuman, F. G., 1989: History of numerical weather prediction at the National Meteorological Center. *Wea. Forecasting*, **4**, 286–296.
- , and J. B. Hovermale, 1968: An operational six-layer primitive equation model. *J. Appl. Meteor.*, **7**, 525–547.
- Teweles, S., and H. Wobus, 1954: Verification of prognostic charts. *Bull. Amer. Meteor. Soc.*, **35**, 455–463.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330.
- , and —, 1996: Ensemble forecasting at NCEP. *Seminar Proc. a Workshop on Predictability*, Vol. 2, Reading, United Kingdom, 39–60. [Available from ECMWF, Shinfield Park, Reading, RG2 9AX, United Kingdom.]
- , and —, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3318.
- , —, S. Tracton, R. Wobus, and J. Irwin, 1997: A synoptic evaluation of the NCEP ensemble. *Wea. Forecasting*, **12**, 140–153.
- Tracton, M. S., and E. Kalnay, 1993: Ensemble forecasting at NMC: Practical aspects. *Wea. Forecasting*, **8**, 379–398.
- Zhao, Q., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models. *Mon. Wea. Rev.*, **125**, 1931–1953.

