History of Numerical Weather Prediction at the National Meteorological Center

FREDERICK G. SHUMAN
Fort Washington, Maryland

(Manuscript received 22 February 1988, in final form 24 April 1989)

ABSTRACT

The first modern numerical weather prediction (NWP) models were developed for the computer that was announced in 1952 at the Institute for Advanced Study, in Princeton, New Jersey. Within 3 yr three agencies of the United States Government jointly created a numerical weather prediction service, but it was quickly discovered that current models had very serious defects. After considerable research, the first operationally effective model was achieved in 1958—a barotropic model covering most of the Northern Hemisphere. Over the years, models have evolved through multilevel filtered equation models and several primitive equation models. Analysis and data assimilation systems necessary for timeliness were also developed, and have likewise evolved. The result has been a revolution in weather forecasting.

1. Introduction

The idea of numerical weather prediction (NWP) was first recorded in a paper by Vilhelm Bjerknes (1904), in which he discussed the application of physical laws to the problem of predicting the atmosphere. More than a decade later, Lewis Fry Richardson (1922) described in great detail the tasks required to acquire and process data, and to disseminate forecasts. His fundamental procedures remarkably resemble those of today, even though he did not look beyond the data sources, data processing tools and methods, and communications devices of his time. The single 6-h time step that he calculated, however, indicated a grossly failing forecast, which discouraged further attempts. Not until 1948 was another attempt proposed.

Some of the problems with Richardson's prediction model were related to the state of knowledge during that time. In particular, he did not understand the necessity for delicate geostrophic balance between mass and motion in the initial conditions, and was unaware of the computational stability criterion of Courant et al. (1928). His most basic problem, however, was the gross inadequacy of computational facilities. His computing tools were a 10-in. slide rule and a table of five-place logarithms, with little else being available at that time. In order to keep up with the weather, **64 000** human computers would be required, according to his estimate. He simply did not have the facilities to experiment and become familiar with the computational behavior of the equations he was dealing with. Had he access to a modern electronic computer, his brilliance as a scientist would probably have allowed him to overcome the other problems.

In the mid-1940s electronic computers were invented, and in 1946 John von Neumann organized the Electronic Computer Project at the Institute for Advanced Study (IAS) in Princeton, New Jersey. The goal of the project was to design and build an electronic computer that would by far exceed the power of other early electronic computers. In 1948 Jule G. Charney established the Meteorology Group within the project. The job of the group was to apply dynamical laws to the problem of forecasting the weather. They were to use the electronic computer that was to be developed by the project. Weather forecasting was one of the three thrusts of the Electronic Computer Project; the other two were numerical mathematics and engineering (Goldstine 1972). The two most important distinctive features of the computer being built were

1) Computer programs were stored in internal memory
2) The computer was parallel

The first feature allowed programs to modify themselves, and Paul Armer (1962) did not exaggerate when he said that it is "one of the great milestones in man's advance." In laying out the preliminary logical design of the computer, Burks et al. (1946) provided the first conceptual paper on an internally programmed computer. The second feature was more architectural; processing was done using entire numbers at a time rather than bit by bit. These two features are now common to virtually all modern general-purpose electronic computers. On 10 June 1952 the institute announced the successful development of the new computer.

Charney et al. (1950) had earlier run a successful forecast on the ENIAC (electronic numerical integrator

Corresponding author address: Dr. Frederick Shuman, 212 Inverness Lane, Fort Washington, MD 20744.

and computer), which is generally recognized as the first general-purpose electronic computer. The ENIAC was developed by the Moore School of Electrical Engineering at the University of Pennsylvania. With the simplest model—the barotropic—it had taken 24 h to make a 24-h forecast on the ENIAC. On the IAS machine the same forecast took less than 5 min, although timely acquisition and preparation of the initial data were not feasible at that time and place. Both the ENIAC and the IAS computer now have their place in history in the Smithsonian Institution.

The barotropic model, with only one level of information in the vertical, can represent potential energy only by the height of a free surface or by surface pressure, depending on how the equations are derived. The barotropic model, therefore, does not have the capability to convert significant potential energy to kinetic energy, and therefore cannot explicitly predict storm development. The so-called Princeton n-level model was soon developed and run with $n = 3$ beginning with an analysis of observations taken just before the Thanksgiving Day 1950 storm over the northeastern United States. It did, indeed, predict a developing storm.

Impetus was thus provided to the ongoing plans of the U.S. Weather Bureau (later the National Weather Service, NWS), the Air Weather Service of the U.S. Air Force, and the Naval Weather Service to acquire a commercial computer with capabilities similar to the IAS machine, and to mount a numerical weather prediction service. On 1 July 1954 the Joint Numerical Weather Prediction Unit (JNWPU) was organized, funded and staffed equally by the three weather services. An IBM 701 was installed in March 1955 and by the following summer the unit began issuing numerical predictions twice daily. The heads of the three weather services are due credit for these decisions. They were Dr. Francis W. Reichelderfer, Chief of the U.S. Weather Bureau; Brig. Gen. Thomas S. Moorman, Commanding General of the Air Weather Service, U.S. Air Force; and Capt. Robert O. Minter, Commander of the Naval Weather Service, U.S. Navy. The principal advisers were Dr. Harry Wexler of the U.S. Weather Bureau, and Dr. George P. Cressman of the Air Weather Service. The JNWPU was organized within the U.S. Weather Bureau, but was under the authority of the Joint Meteorology Committee (JMC), whose members were the heads of the weather services.

In the beginning, 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.

2. The Joint Numerical Weather Prediction Unit

When the JNWPU was organized there were two ways that it could have gone. One was to study the mechanics of the atmosphere, to look for tools to be used by forecasters, and at the same time to develop and ultimately demonstrate the utility of NWP as an operational product in itself. The second was to proceed immediately to operational NWP production. The JMC directed JNWPU to pursue the latter course. As it turned out, this was based on false optimism about the short term, but had the decision been otherwise, the effort would likely have failed. The operational environment was the appropriate environment for the early problems to be quickly encountered and solved. A substantial portion of the resources of the nation's weather services were being devoted to the effort; the necessary support could not long be sustained without significant positive results.

The models that were considered for operations at the time were Charney's (1954) Princeton three-level model, and Thompson and Gates' (1956) two-level thermotropic model. The former had been run on 14 cases, the latter on 60, and based on results, it was assumed that either would serve well operationally. The problem at the time seemed to be one of choice between the two. Since the testing and evaluation were done independently from each other, and not on the same cases, much comparability was lacking, and the choice was not a straightforward one.

The JMC decided in favor of the Princeton three-level model. When the model was programmed for the IBM 701 in 1955, run on an operational schedule, and subjected to the critical eye of the practicing synoptician in real time, it was discovered that it not only was unable to predict reliably and accurately, but it also provided little or no useful information to the forecaster.

Although disappointing at the time, this was the key to future success, and was the immediate and essential result of the decision by JMC to go operational. This "baptism by fire" immediately utilized the talents of the modeler, the judgment of the practicing synoptician, and the skills of the computer programmer. It established patterns of motivation with attention centered on accuracy and timeliness of delivery. This concentration of effort was a product of the operational environment.

The payoff came in 1958 when the problems were identified and solved, a suitable automatic analysis system was invented, and automatic data handling was developed. The model at that time was not the three-level, nor the two-level, but the single-level barotropic model. Skillful, timely numerical predictions were delivered to central forecasters, who in turn used them as guidance for their own manually prepared prognostic charts. The impact on centrally prepared guidance can be seen in Fig. 1 by the drop in the S_1 score beginning in 1958. About 2 years following this success, the JNWPU divided 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).

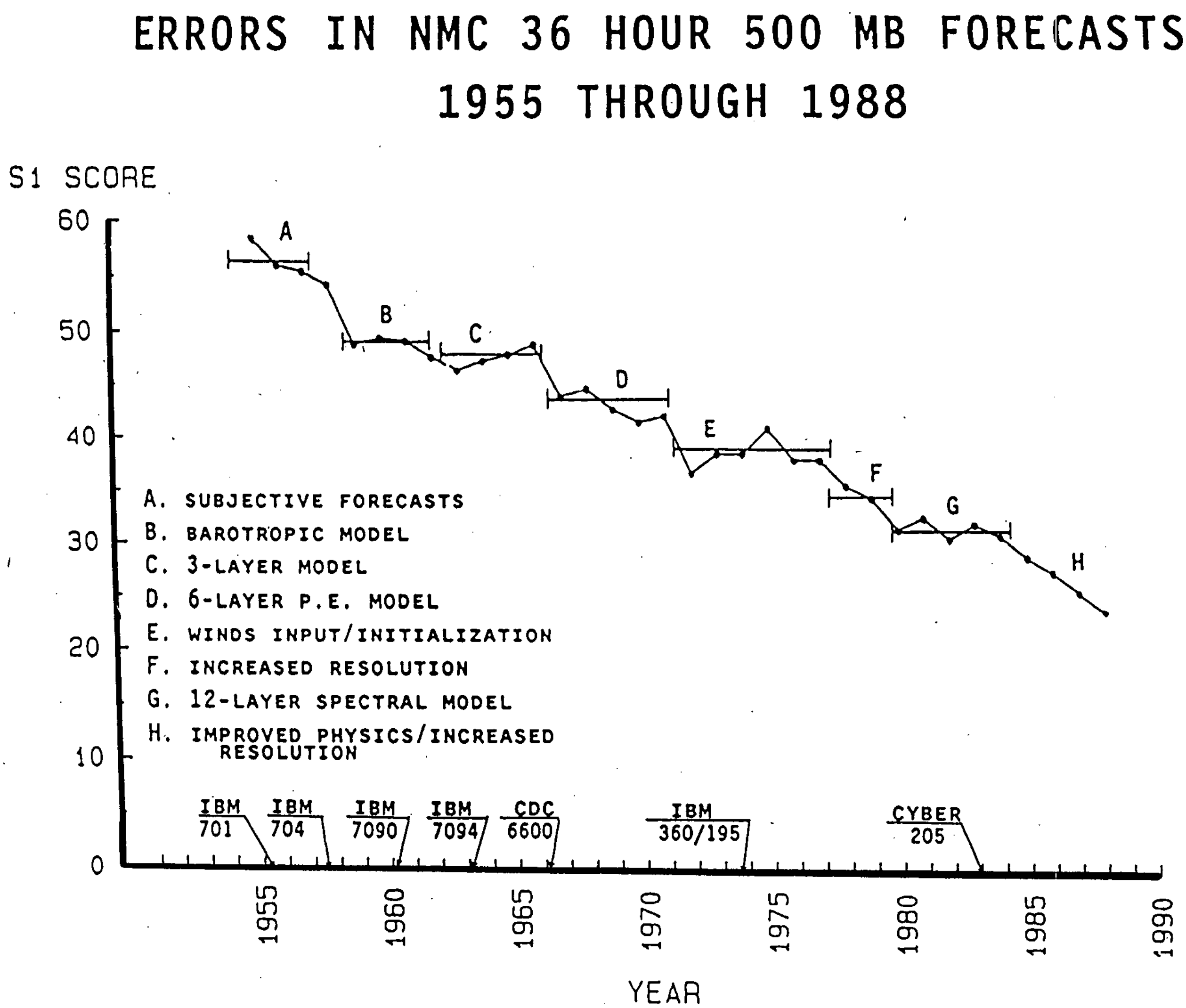

FIG. 1. Record of S_1 score (Teweles and Wobus 1954) for 36-h predictions of geopotential height at 500 mb. The S_1 score is roughly a measure of normalized rms vector error of geopotential height gradient. The area of verification is North America and adjacent waters. To calibrate the S_1 score in terms of practical skill, a forecast with a score of 20 is virtually perfect, and one with a score of 70 is worthless, e.g., skill (percent) = $2 \times (70 - S_1)$.

By 1960, NWP had progressed to the extent that some of the products not only surpassed in quality those made manually at the National Meteorological Center (NMC), but also could not successfully be improved by manual methods. The first such products were predictions of the height of the 500 mb surface out to 36 h. Automated products began replacing manual products at that time. About 95% of NMC products are now produced automatically. The remaining 5% consist primarily of predictions of weather itself: rain, snow, heavy precipitation, etc. These remain intractable to automatic methods alone, although the analysts producing them rely heavily on numerical weather predictions of circulation, temperature, humidity, and other parameters, including precipitation itself.

It should be understood that at NMC the replacement of one product or method with another has not been based on statistics alone. In this connection, statistical measures of quality are not as objective as they are sometimes taken to be. Significant subjective factors are considered when a set of statistics is selected to measure quality of predictions—except for highly specialized predictions. Selection of statistics to measure quality of general purpose weather predictions is explicitly or implicitly based on expectations of how well they will agree with expert opinion. In fact, the S_1 score (Teweles and Wobus 1954) was designed, tested, and later modified to correlate with expert opinion. Although the opinions of juries of expert analysts have consistently been included at NMC in making the important decisions on product replacements, to the best of my recollection, serious disagreements between statistical measures and human judgement have not occurred.

3. The barotropic model

More than back-breaking labor was required to achieve operational NWP. Essential knowledge in mathematics and fluid dynamics of the atmosphere had first to be discovered. Operations began with the most sophisticated model of the time, the Princeton three-level model (Charney 1954). Because of lack of skill it was soon abandoned in favor of the two-level thermotropic model (Thompson and Gates 1956), which for similar reasons was abandoned, in turn, in favor of the single-level barotropic model (Charney 1949). The barotropic model, too, was found to be lacking in essential skill, but its relative simplicity established a tractable problem for the researcher. This was a pattern often repeated later, the moral being that when confronted with a problem, try to capture it for study in the simplest possible system.

The theory behind all of these models was basically sound, and their frameworks were later proven in operation. It is fair to describe the situation not as a case of faulty theory, but rather as a case of incomplete theory. In other words, the ideas leading to these models were appropriate and correct, indeed ingenious, but some necessary things were overlooked. In light of the prejudgment of the models' operational suitability, the operational environment must be considered as a necessary milieu for completion of the theory. Through repetition during more than 30 yr, this has become a well-established principle.

Three things were wrong with the barotropic model, two of which called for fundamental research. These were also wrong with the two- and three-level models, but in their complexity the multilevel models had additional errors. The symptoms of the two fundamental problems with the early barotropic model were spurious anticyclogenesis in low middle latitudes, and excessive retrogression of the planetary-scale waves. The spurious anticyclogenesis was due to the divergence of the geostrophic wind, which was used explicitly in the early barotropic model. It was resolved by replacing the geostrophic wind with a nondivergent wind defined by a streamfunction derived from the balance equation (Shuman 1957a). Operational use of the balance equation depended on successful applied mathematical research (Miyakoda 1956; Shuman 1957b).

The retrogression of the planetary-scale waves was, at first, inadvertently suppressed by the small area of integration on the IBM 701, but appeared when the area was enlarged on the IBM 704. The error was due to lack of an adjustment mechanism between mass and motion and was at first resolved by holding wave numbers-1, -2, and -3 fixed during the prediction (Wolff 1958). Later a more natural solution was applied that was derived from free-surface models (Cressman 1958).

The third deficiency of the barotropic model was its restricted area of integration, although generally not for periods of 24 h and less. The operational model at the end of the era of the IBM 701 was being run on a 1020 point 30 × 34 grid with grid interval of 381 km on a polar stereographic projection true at 60°N. The grid was centered at about 70°N 105°W, between the North Pole and the United States, and covered North America and adjacent waters. On the IBM 704 the area of integration was enlarged virtually to cover the Northern Hemisphere. The new grid was the so-called JNWP octagonal grid of 1977 points, centered on the North Pole, with grid intervals of 381 km on a polar stereographic projection true at 60°N. The octagonal boundary lay between 9° and 15°N. The more powerful IBM 704 was essential for the resulting increase in number of calculations.

Another problem arose when NMC became hemispheric. A very large error in the streamfunction became painfully evident, of the same scale as the grid itself. It was soon traced to an inconsistent finite-difference formulation of the nonlinear terms in the balance equation. Small but systematic truncation errors integrated over the larger area into prohibitive errors in the streamfunction. Like the physical error of wave retrogression, the numerical truncation error had been suppressed by the small area of the model run on the IBM 701.

In the press of operational events, the error was never reported, nor was its correction. It was first described by Bolin (1956) as one of inconsistent truncation errors for the nonlinear terms in the balance equation. These terms contain products of all of the second derivatives of the streamfunction in the two horizontal dimensions. Specifically, they are factors in the difference of two quadratic terms, i.e.,

$$(\ )_{xx}(\ )_{yy} - (\ )^2_{xy}.$$

The initial impulse has often been to approximate the two factors in the first term by the simplest of second differences in x and y using a space difference of one grid interval, and the two factors in the last term using a space difference of two grid intervals. This disparity in space differences is the inconsistency that Bolin pointed out. He recommended turning the x, y coordinates 45°, then forming the differences for $(\ )_{xx}$ and $(\ )_{yy}$ using the central point and the corners of the eight-point box surrounding the central point, and forming the difference for $(\ )_{xy}$ using the grid points on the sides of the box. This results in a consistent space difference of the grid interval multiplied by the square root of 2.

Inconsistency of differencing intervals in itself is not necessarily bad; it is the result of the inconsistency that counts. The nonlinear terms in the balance equation arise from the differentiation of the nonlinear terms in the equations of motion. Therefore, their integral over an area will not depend on values of anything within the area, but only on values at the boundary of the area. The finite-difference terms should have the corresponding characteristic; they should sum over an area to boundary values. Analysis shows that the "inconsistent" forms do not, but Bolin's forms do. In fact, one can derive Bolin's forms in the balance equation by differencing the equations of motions stated in an appropriate finite difference form, and thereby ensuring that they properly sum. This is what was done at NMC.

The same type of inconsistent error was contained in work reported by Charney et al. (1956). In that case, five second derivatives of geopotential height in isobaric coordinates appeared in the nonlinear terms that constituted the potential vorticity. The quadratic terms in which the derivatives appeared as factors may be grouped into two pairs to appear like the nonlinear terms in the balance equation, i.e.,

$$(\ )_{xx}(\ )_{pp} - (\ )^2_{xp} \quad \text{and} \quad (\ )_{yy}(\ )_{pp} - (\ )^2_{yp},$$

where p is pressure, the vertical coordinate. It should be noted that there are other quadratic terms involving first derivatives, and the same principles must be applied to them. The result of the inconsistency was excessive cyclogenesis in the predictions. I later reran one of the forecasts with the error corrected, and it did indeed eliminate the excess cyclogenesis.

With the research on the barotropic model and balance equation completed, and the results applied, the first successful operational barotropic model was inaugurated. Operational success, however, depended not only on NWP research advances but also on coincident advances in technology. As the first numerical predictions were not reliably accurate, so were they not timely. Automatic processing of the incoming data, numerical weather analysis, and automatic graphical output had to be developed.

The first operation with the IBM 701 in 1955 depended on hand methods of analysis, and primitive communications methods. Incoming data were manually plotted from teletype pages and analyzed manually by traditional methods, then a transparency with the model's grid points was overlaid, and values at the grid points were interpolated and read by an analyst to a card punch operator. The collection and preparation of data, punching of cards, and manual checking of the input required about 10 h from the nominal data time, which meant that products were not available for use until 12 h after the nominal data time. Early in the first operation a method of machine graphical output was developed at NMC for a line printer—a method that survives today—but mainly for research purposes. Experiments were made with machine methods of input data processing and machine analysis, but the problems of error control, running time, and costs were not adequately solved on the first machine. This had to wait for the more powerful IBM 704.

The first analysis method used was a sectionalized fit of second-degree polynomials (Gilchrist and Cressman 1954). Because the method had difficulties with uneven data distribution, such as around silent and data-sparse areas, and was expensive in machine time, it was abandoned in favor of a scheme invented by Bergthórsson and Döös (1955) and further developed at NMC by Cressman (1959). Considerable effort was made to protect the analysis against gross errors and early efforts were directed to allow manual modification of the result, particularly over oceans and other data-sparse areas. Early in this development the benefits of adjustment of the previous forecast by new data were shown, and the concept of a *forecast-analysis cycle* came into being (*data assimilation system* in current terminology).

Parallel developments for input of data occurred. A start was made with hand punching of data cards from teletype copy, and was followed by automatic conversion of paper tape to cards. This introduced a new concept called automatic data processing (Bedient and Cressman 1957) before that terminology began to have a wider meaning in computer science. Automatic data processing is the reading of remotely manually prepared teletype texts into computer-quality databases. It has many of the qualities of reading natural languages, although the forms are fortunately more restricted. The input text contains observations in a dozen or so formats, with variations and errors normally found in language that must be recognized in context amid extraneous material. Development of the method, together with the introduction of high-speed paper-tape readers on an IBM 1401, enabled an earlier start of the analysis and prediction procedures from 10 h after data time in 1955 to 6 h after data time in 1959. In addition, the number of fields and levels analyzed increased.

The advances so far described enabled the first NWP operation that successfully exceeded the minimum requirements of quality and timeliness. The result on quality of central guidance can clearly be seen in Fig. 1, as the drop in S_1 score beginning in 1958.

4. Baroclinic models

The barotropic model does not account for significant conversion of potential and internal energy to kinetic energy and therefore does not predict the development of storms. This was the next big problem to be tackled. Initial success was achieved with the NMC three-level filtered-equation model (Cressman 1963). It not only utilized the theory and structure of the Princeton three-level model (Charney 1954) along with the new "corrections" contained in the barotropic model, but also used theoretically derived factors essential for filtered baroclinic models. The principal departures of the NMC three-level model from its predecessors were an additional term to account for advection of vorticity by the divergent component of the wind, use of the balance equation to maintain the proper relation of the mass field to the wind field during the integration, and the careful construction of finite-difference forms and numerical procedures to prevent systematic accumulation of truncation error.

This first successful baroclinic model became operational in 1962. Acquisition of the more powerful

IBM 7094 was an essential enabling factor. The three-level model established a new lower plateau of error, shown in Fig. 1.

The behavior of the various models at 500 mb is an important indicator of the general skill of models, and because of the high autocorrelation in the vertical of winds from 700 mb well into the stratosphere, it is almost a direct indication of skill in forecasting winds for aviation. Most of the forecast service, however, is directed to conditions at the surface of the earth. For this the quality of central guidance at sea level (or earth's surface) is perhaps NMCs most important product. Early on, the analyst learned to use the 500 mb numerical predictions to improve the central guidance at sea level. There was not a precipitous drop in error at sea level in 1958, as there was at 500 mb, because of the time factor involved in the learning process. Rather, there was a steady decline in error.

The first useful numerical prediction at sea level was achieved by Reed (1963) during a 1-yr visit to NMC. Reed's model went into operation in 1962. In framework, it drew from the early two-level thermotropic models (Thompson and Gates 1956), the thickness equation of the thermotropic model being used to predict the 1000 to 500 mb thickness. To obtain the prediction at 1000 mb, the thickness was combined with predictions of 500 mb height made independently by the operational models already discussed. Although Reed's model could not compete directly with manual prognoses, not even those made prior to 1958, it did provide the analyst with much useful information about the development and placement of systems at sea level. The error at sea level continued to decline, and for the 5 years from 1962 to 1966 the decline was attributed largely to Reed's model.

George Cressman (in undocumented work) developed the NMC three-level model into a four-level model in which the balance equation was incorporated into each time step to find the explicit mass field. This amounted to a successful computation of Charney's (1962) balance equations. It was not pursued further because the focus of effort at the time was on development of a primitive equation model. The balance equations are an extension of the filtered equations, in which the gravity waves are filtered out. It was clear at the time that as the filtered equations were made more complete, the calculations that would be required would surpass those of the primitive equations, and thus their principal advantage would be lost.

Not only advances in modeling were occurring at NMC; important improvements in other areas were being made as well. For example, developments around the mechanical curve-plotter at NMC were unique at the time. Curve-plotters had existed before, but the transistorized models that appeared in 1958 had plotting speeds which could meet weather service requirements. The National Meteorological Center took advantage of the state of the art and developed algorithms for finding contours, following them with acceptable accuracy, locating centers, and drawing a production map able to meet critical standards. These developments were undertaken and completed by 1959 in conjunction with the speedup of the National Facsimile Circuit to 120 rpm. Expansion occurred rapidly with the addition of facsimile circuits. In little more than a decade, daily production by this method for facsimile transmission increased to about 500 charts.

5. Primitive equations

Research with primitive equations began at NMC in 1959 at a time when no satisfactory finite difference system existed for them. The problem at the time was an overwhelming accumulation of a truncation error early in the integration, intimately connected with the nonlinearity of the equations, and usually referred to as "nonlinear instability" to distinguish it from the problem that Courant et al. (1928) analyzed. Experiments with numerical systems borrowed from others' work quickly proved that no acceptable numerical system existed. Furthermore, no theory existed for stability of nonlinear systems, and even today, theory is rudimentary. Perhaps the most complete theory in this area was developed in 1969 by NMC workers in collaboration with Canadian, A. Robert (Robert et al., 1970, and was built on earlier work by Phillips (1959) and Richtmyer (1963). Later Shuman (1974) extended Phillips' and Richtmyer's analyses somewhat.

Numerical research at NMC was confined to time-centered difference systems because of their characteristic of conserving energy in finite-difference analogs to linear energy-conserving differential equations. To my knowledge, no difference system, time-centered or otherwise, will remain *perfectly* stable indefinitely when used in finite-difference analogs to nonlinear equations as complex as those of NWP. By *perfect* stability, I mean the existence of neither unlimited growth nor unlimited decay of available energy (or an appropriate square of dependent variables) in a difference analog to an energy conserving system of differential equations. Time-forward difference systems that damp, however, have been devised that allow atmospheric models to run well beyond present limits of predictability with little or no significant loss of energy. The Lax-Wendroff (1964) difference system used in the currently operational nested grid model (NGM) is an example.

The approach to the problem was empirical. First, a workable number of difference systems were selected for experimentation, for without a guiding theory, the possible number of forms is virtually infinite. Second, the instability was captured in the simplest possible physical system, so that large numbers of experiments could be run in a reasonable time. Forms passing tests with very simple physical systems were subjected to tests with more complex physical systems, and so on up the hierarchy of complexity to a full-scale primitive equation model. By mid-1961, two acceptable, relatively stable forms had been discovered (Shuman 1962). A relative of one of the two is still being used today (Gerrity et al. 1972). Not until 1969 was the relative stability of the two forms at least partly explained (Robert et al. 1970).

The discovery in 1961 of relatively stable difference systems for primitive equations was indeed good news, but the bad news was that the IBM 7094 was not powerful enough to run a multilevel primitive equation model. Primitive equation models require far more calculations than the then operational filtered equation model did, and lack of timeliness delayed a primitive equation operation for 5 yr. The six-layer primitive equation model became operational on 6 June 1966 (Shuman and Hovermale 1968). Enabling factors at that time were acquisition by NOAA of the more powerful CDC 6600, and the appearance of the U.S. Air Force Automated Weather Network (AWN). The AWN provided quicker data collection, and allowed an earlier start of the operation.

A drop in S_1 score at 500 mb, beginning in 1966, can clearly be seen in Fig. 1. This drop is attributed to the introduction of the NMC six-layer primitive equation model.

After a few years of development in operation, the NMC six-layer primitive equation model became the first to produce a prediction at sea level that was directly competitive with the manual predictions. In fact, the model produced, by itself, a more skillful forecast than a human could produce without NWP guidance. For instance, the "raw" NWP at sea level in 1971 averaged only five S_1 points higher than the score for the manual product, still well below the scores prior to 1958. The five S_1 points contributed manually were an important five points, however. They can be translated into 5 yr of progress. Without the analysts' skill, the product in 1971 would have been only at the 1966–67 level of skill. The analyst contributed other essential skills to NMC's products—and still does—especially in quantitative precipitation and cloudiness forecasting and in frontal analysis. Man's part in the man-machine mix has been essential to overall quality, and will remain so for the foreseeable future.

As with all operational models, the six-layer primitive equation model underwent continuous development. It was initialized with data from the 1977-point JNWP octagonal grid (on a polar stereographic projection true at 60°N and with an interval of 381 km), but was run on a rectangular 53 × 57-point grid. The latter being the larger in all dimensions, data was created in the nonoverlapping area with a stable extrapolation procedure, and then the boundary region was smoothed. At its inception the model had the effects of skin friction, transfer of heat from warm oceans, and topographical effects. A few of the major subsequent improvements were the introduction of a calculation of water vapor and latent heat beginning with one level of resolution (precipitable water—Weather Analysis and Prediction Division 1968a), and later with three levels of resolution (Weather Analysis and Prediction Division 1970a). Additional improvements included the introduction of both long- and shortwave radiation effects (Weather Analysis and Prediction Division 1967), refinement of the description of topography to as accurate a representation as the grid could carry (Weather Analysis and Prediction Division 1968b), and introduction of the feedback effects of convective rain (Weather Analysis and Prediction Division 1971a). The reduction of pressure to sea level was also improved in the output (Weather Analysis and Prediction Division 1970b). A reduction of about 50% in computer time was achieved through a technique of time-averaging the pressure force terms in the equations of motion (Shuman 1971; Brown 1978).

At first, the balance equation was used to relate the initial winds to the mass field, but was later replaced by direct wind analyses (Weather Analysis and Prediction Division 1971b). Much work on the initialization problem had been done at NMC by Nitta and Hovermale (1969), and by Okland (1970). These efforts, like most work elsewhere, used the dynamics of the model to obtain the initial relationship between mass and motion, but none worked well in tests. Meanwhile, direct wind analyses worked very well indeed, as can be seen by the decline in S_1 in 1972.

In September 1971 the limited-area fine-mesh (LFM) model (Howcroft 1971) was introduced as the first operational regional model. In its essentials it was the same as the six-layer model described earlier, but with half the mesh size and time step, and a smaller area. In August 1977 the grid interval was reduced from 190.5 to 127 km, but on 11 June 1981 it was increased again to 190.5 km simultaneously with the introduction of fourth-order difference systems (Gerrity et al. 1972). The LFM covers roughly an octant of the globe, having 53 × 45 grid points with the larger interval, 79 × 67 with the smaller, and is still being run as a first, quick prediction in each 12-h cycle.

A direct analysis of amplitudes of a set of 24 Hough functions was developed at NMC (Flattery 1971) and introduced into operations on 18 September 1974, replacing the Cressman (1959) analysis for the six-layer model and for the data assimilation system. Cressman's analysis remained in operation for the regional LFM. Hough functions are global solutions to the tidal equations, and they contain a relationship between wind and pressure fields. They are especially suitable for analysis of a mix of wind and pressure over the globe, since observations are, and will continue to be, of winds in low latitudes, and a mix of winds and pressures in high latitudes.

On 19 January 1978 the so-called seven-layer primitive equation model was introduced. It, too, in its essentials was the same as the six-layer model. An inactive seventh layer in the six-layer model was activated, but more importantly, the mesh size and time step were halved, resulting in more accurate estimates of derivatives. The increase in number of calculations required by the new model was enabled by the IBM 360/195. The accuracy of NWP reacted very substantially to the increased resolution, as can be seen in Fig. 1 by the lower S_1 scores beginning in 1978. I believe that truncation error had been the most significant limiting factor before the seven-layer model, and that little if any improvement could have been obtained by improving the vertical resolution or the physics of the model without reducing horizontal truncation error.

In June 1978 a model for tracking hurricanes was declared operational, but it had been running since 1975 on a quasi-operational basis. It was not run on an operational schedule, but was on call not only for hurricanes, but also for threats of heavy precipitation. It had ten levels in the vertical and a grid of 50 × 50 points, with intervals of 60 km for hurricanes, and 100 km for heavy precipitation. When run on hurricanes, the center of the grid moved with the hurricane. It was therefore named the moveable fine-mesh (MFM) model.

Global models have been used since 1974, when a finite-difference model was introduced with a grid interval of 2.5 lat–long degrees and nine layers (Stackpole 1978). This model was similar to the six-layer primitive equation model in some ways, but differed in the number of layers, the method of resolving the horizontal, and its global extent. Computational instability in high latitudes, where the east–west grid interval is small, was avoided by smoothing the differences. The weighting function was approximately triangular, designed to avoid negative responses anywhere in the spectrum while providing computational stability. A method of treating central time-differences with a time filter was borrowed from work done in Canada (Robert 1966; Asselin 1972). The model was first implemented on 18 September 1974 using the Hough analysis for 6-h projections in the global data assimilation system, but was soon extended to 5, and later to 10 days, and used as guidance for medium-range forecasts and the early part of monthly forecasts.

The global spectral model (Sela 1982) replaced the seven-layer primitive equation model in August 1980. Earlier, in June 1980, it had replaced the global nine-layer model in the global data assimilation system. When implemented, its highest resolution was 12 layers in the vertical and 30 spherical harmonic modes for horizontal resolution. It had rhomboidal truncation, and a Gaussian grid in physical space for easy and accurate conversion from physical to spectral space and vice versa. The Machenhauer (1977) normal mode initialization was used. To save computer time, at 48 h the model's horizontal resolution was reduced to 24 modes, at 144 h the model's vertical resolution was reduced to six layers, and at 8 days its number of modes was halved by reducing the effective area of prediction to the Northern Hemisphere. The progressive reduction of resolution during the run is consistent with the general decay of predictive skill. The model was run to 10 days at 0000 UTC and to 60 h at 1200 UTC.

A method of multivariate optimum interpolation (Bergman 1979; McPherson et al. 1979; DiMego 1988) was developed at NMC and first placed into operation on 14 August 1982. A unique feature of the method as developed at NMC is that the update with data is performed *in the prediction model's vertical coordinate,* rather than on standard isobaric surfaces. The model's vertical coordinate was a modified form of Phillips' (1957) sigma. At the time, it was not fully competitive with the Hough analysis, and was therefore restricted to the global data assimilation system, where its influence on the forecast products was minimal, but where valuable operational experience and feedback could be obtained. On 6 December 1982 it was introduced as the analysis system for the barotropic model, which is less sensitive to initial error than the other models. It gained full operational status in 1985, when it was implemented as the analysis system for the nested grid model.

On 27 March 1985 the Regional Analysis and Forecast System (RAFS) was implemented. The prediction model was Phillips' (1979) nested grid model and the analysis method was multivariate optimum interpolation. The nested grid model will be described in detail in other articles in this issue, so will only be mentioned briefly here. Its innermost third grid has a mesh length in the horizontal of about half that of the seven-layer primitive equation model and it now has 16 levels in the vertical. As with other large-scale models, a new computer was required to run it operationally, particularly with its third grid. That computer is the Cyber 205.

6. Computers and models

Modern NWP began with the invention of the modern electronic computer, and subsequent improvements of NWP have been paced primarily by advances in computer technology. Other requirements are essential for improvements, but they generally have not been the principal limiting factors since 1950. The other requirements include greater knowledge of atmospheric physics, better meteorological observations, and quicker communications to gather data and disseminate results.

Computers have been playing an ever expanding role in NWP and its supporting technologies. Today, computers are ubiquitous—a part of virtually all modern technology. Not only does increasing supercomputer power enable operation of more sophisticated and highly resolved NWP models, but also computers of many kinds have become inextricably a part of supporting technologies such as observations and communications. The powerful state of the art supercomputers have advanced hand in hand with all of the other electronic computers; they all generally benefit from the same technology. In the case of satellite observations, computers are essential for launch and control, for on-board operation, recording, and programming of the various instruments, for data acquisition on the ground, and for reduction and interpretation of data. Not only do satellite observations depend on computers, all paths to more accurate and more extensive observations will rely heavily on computers. Even research and development of new NWP models and techniques are to a great extent inspired by the power, or expected power, of supercomputers.

Of the many supporting technologies necessary for an NWP operation, the increasing supercomputer power to run the operational models has been directly and proximately responsible for the decreasing error shown in Fig. 1, through the introduction and enhancements of NWP models during the past 35 yr. Other centers around the world have had similar experiences. This is not to say, however, that improvements in NWP immediately and automatically follow the availability of increased supercomputer power to run the models. All technologies supporting NWP must advance along with, or closely follow, supercomputer power, and vice versa. What I am saying here is that they have in the past, and I expect them to in the future.

If the skill of numerical predictions increases appropriately when supercomputer power increases, then supercomputer power is an index of skill, in the sense that one could predict skill if one could predict supercomputer power. It is not unreasonable to suppose that the state of NWP reflects the state of a technology such as computers, which is so intertwined with all other supporting technologies. If there is some factor other than supercomputer power limiting forecast improvement when a more powerful supercomputer becomes available, pressure would be generated to relieve that limitation. The relief might even depend (at least in part) on the technology that produced the more powerful supercomputer in the first place. The very high correlation between supercomputer power and skill in the past supports this idea, and in itself suggests a predictive relationship between the two (Shuman 1982).

Today we may be at least as limited by our observational capability as we are by the power of supercomputers available to run the models. Uniform coverage over the globe, including the oceans and other uninhabited areas of the earth, is required, and weather satellites have made such coverage economically feasible. Satellite observations of wind, temperature, and humidity, however, are of lower quality than rawinsonde observations, which remain the standard for accuracy. For improved short range forecasting over the continental United States we need high quality observations off both the Atlantic and Pacific coasts. Many "surprise" winter storms in the northeastern and Middle Atlantic states would be better forecast with improved observations over the western Atlantic Ocean, and a similar situation exists for the western United States. Forecast errors are consistently higher over the western half of the contiguous 48 United States, due in large part to the low observational quality and coverage over the eastern Pacific Ocean, which is upstream to the West. As stated previously, all of the factors that make for a successful NWP operation must be improved for continuing improvement in the predictions. Still, advances in NWP have been paced by the increases in the power of supercomputers, so some details are provided here about the historical relationship. In spite of the existence of such a close relationship, bureaucratic impedimenta have made United States government computer acquisitions more and more difficult and time consuming.

Since 1952, breakthroughs in computer technology have been the norm. The National Meteorological Center has undergone six state-of-the-art supercomputer acquisitions in 33 yr—IBM 701, IBM 704, IBM 7094, CDC 6600, IBM 360/195, and Cyber 205. Each new supercomputer has been about six times more powerful than its predecessor, so that the Cyber 205, in both speed and storage capacity, is about 10 000 times more powerful than the IBM 701. Computer power is continuing to grow; there are now machines on the market more powerful than the Cyber 205.

Accuracy of NWP increased over the years, but not by several orders of magnitude, as computer power did. By most measures, skill has more than doubled in the last 33 yrs. This "discrepancy" between computer power and skill comes from the nature of the problem as we now understand it. For example, model resolution (density of data-carrying points) is closely related to skill. An order of magnitude increase in computer power will permit only a doubling of resolution. Doubling of resolution, on the other hand, will not double skill, but increase it around 15%, and even then only if other characteristics of the model are suitably enhanced.

Each new acquisition of more powerful computers enabled the introduction of a more sophisticated NWP model. Operational NWP began on the IBM 701. The IBM 704 enabled the barotropic model to be enlarged to cover most of the Northern Hemisphere. The IBM 7094 enabled the introduction of a three-level baroclinic model. The CDC 6600, along with earlier collection of global data, enabled a six-layer primitive equation model to be inaugurated. The quicker data collection was provided by the U.S. Air Force Automated Weather Network. The IBM 360/195 allowed a doubling of the horizontal resolution in the six-layer model. At the same time a seventh inactive layer at the top of the model was activated, and the model was renamed the seven-layer primitive equation model. The IBM 360/195 also enabled the introduction of a 30-mode spectral, spherical harmonics model that replaced the seven-layer model, and a 50% increase in the resolution of the operational regional model. The Cyber 205 enabled an increase in the number of modes of the spectral model, and the introduction of the Phillips' (1979) nested grid model with three nested grids. The operational repertoire of programs run on the Cyber 205 are discussed in other articles in this issue. This is a bare-bones list of the ways increasing computer power has been used. Other uses include more sophisticated data assimilation systems, graphical output, quality control of data, and communications.

7. Conclusions

All the meteorological world was watching the work of JNWPU in the late 1950s. Our job was no less than to revolutionize weather forecasting, which had begun almost a century earlier as a centralized operation, and which had not changed much since then in its fundamental processes. Other nations took a wait-and-see attitude regarding whether or not the dynamical, more scientific approach could or would succeed. After we demonstrated that NWP was feasible, other nations began to establish NWP operations, and now all industrial and some developing nations have NWP operations. The initial failures were great disappointments, but made the hard earned success in 1958 all the sweeter. The failures also taught a lesson that advances in this field would come slowly and only with great difficulty.

Research in the operational environment is often at the same level as in the academic and laboratory environments, but when successful always includes the extra step of application. This last step is seldom easy and, in a large project like the development of a new model, ties the researcher to the operation for months at a time. Applied research in NWP is demanding work, and the operational environment is a hard task master, but has its advantages. One important advantage is that no researcher at an operational center can entertain false theories for long, for theories are critically tested in operations. The researcher at a meteorological center lives very much in the real world.

The first 35 yr (1954–89) were exciting and productive for everyone concerned. As in all endeavors of this kind, the easy problems were solved first. Scientific and technological advances are more difficult to obtain now, but there are challenges and prizes still to be won. Bright young people should continue to consider numerical weather prediction as a productive and satisfying career.

REFERENCES

Armer, P., 1962: An introduction to a condensed version of Burks et al. (1946). *Datamation,* **8,** issue #9, 24.
Asselin, R., 1972: Frequency filter for time integration. *Mon. Wea. Rev.,* **100,** 487–490.
Bedient, H. A., and G. P. Cressman, 1957: An experiment in automatic data processing. *Mon. Wea. Rev.,* **85,** 333–340.
Bergman, K. H., 1979: Multivariate analysis of temperatures and winds using optimum interpolation. *Mon. Wea. Rev.,* **107,** 1423–1444.
Bergthórsson, P., and B. Döös, 1955: Numerical map analysis. *Tellus,* **7,** 329–340.
Bjerknes, V., 1904: Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik. *Meteor. Z.,* **21,** 1–7.
Bolin, B., 1956: An improved barotropic model and some aspects of using the balance equation for three-dimensional flow. *Tellus,* **8,** 61–75.
Brown, J. A., and K. Campana, 1978: An economical time-differencing system for numerical weather prediction. *Mon. Wea. Rev.,* **106,** 1125–1136.
Burks, A. W., H. H. Goldstine and J. von Neumann, 1946: Preliminary discussion of the logical design of an electronic computing instrument. A Report to Ordinance Department, U.S. Army (Electronic Computer Project, Institute for Advanced Study, Princeton, NJ, June 28, 1946) A two-part condensed version is published in *Datamation,* **8:** Part 1 in issue #9, 24–31; Part 2 in #10, 36–41.
Charney, J. G., 1949: On a physical basis for numerical prediction of large-scale motions in the atmosphere. *J. Meteor.,* **6,** 371–385.
——, 1954: Numerical prediction of cyclogenesis. *Proc. Nat. Acad. Sci. U. S.,* **40,** 99–110.
——, 1962: Integration of the primitive and balance equations. *Proc. Intern. Symp. Numerical Weather Prediction,* Meteor. Soc. Japan, Tokyo, 131–152.
——, R. Fjortoft and J. von Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus,* **6,** 309–318.
——, B. Gilchrist and F. G. Shuman, 1956: The prediction of general quasi-geostrophic motions. *J. Meteor.,* **13,** 489–499.
Courant, R., K. Friedrichs and H. Lewy, 1928: Uber die partiellen Differenzgleichungen der mathematischen Physik. *Math. Ann.,* **100,** 32–74.
Cressman, G. P., 1958: Barotropic divergence and very long atmospheric waves. *Mon. Wea. Rev.,* **86,** 293–297.
——, 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87,** 367–374.
——, 1963: A three-level model suitable for daily numerical forecasting. NMC Tech. Memo. No. 22, Weather Bureau, ESSA, U.S. Department of Commerce, 22 pp.[1]
DiMego, G. J., 1988: The National Meteorological Center regional analysis system. *Mon. Wea. Rev.,* **116,** 977–1000.
Flattery, T. W., 1971: Spectral models for global analysis and forecasting. Tech. Report 243, Air Weather Service, U.S. Air Force, 42–54.
Gerrity, J. P., R. D. McPherson and P. D. Polger, 1972: On the efficient reduction of truncation error in numerical weather prediction models. *Mon. Wea. Rev.,* **100,** 637–643.
Gilchrist, B., and G. P. Cressman, 1954: An experiment in objective analysis. *Tellus,* **8,** 61–75.
Goldstine, H. H., 1972: *The Computer from Pascal to von Neumann.* Princeton University Press, 378 pp.
Howcroft, J. G., 1971: Local forecast model: Present status and preliminary verification. NMC Office Note 50, National Weather Service, NOAA, U.S. Department of Commerce, 22 pp.[1]
Lax, P., and B. Wendroff, 1964: Difference schemes for hyperbolic equations with high order of accuracy. *Commun. Pure App. Math.,* **17,** 381–398.
Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Beitrage zur Physik der Atmosphare,* **50,** 253–271.
McPherson, R. D., K. H. Bergman, R. E. Kistler, G. E. Rasch and D. S. Gordon, 1979: The NMC operational global data assimilation system. *Mon. Wea. Rev.,* **107,** 1445–1461.

[1] Upon request, the National Oceanic and Atmospheric Administration (NOAA) will undertake to acquire these publications for loan or for reproduction by office copier. Call (301)443-8330 or write NOAA Central Library, 6009 Executive Boulevard, Rockville, MD 20852.

Miyakoda, K., 1956: On a method of solving the balance equation. *J. Meteor. Soc. Japan,* **34,** 364–367. (Reprinted in *Vortex No. 2,* June 1957.)

Nitta, T., and J. B. Hovermale, 1969: A technique of objective analysis and initialization for the primitive forecast equations. *Mon. Wea. Rev.,* **97,** 652–658.

Okland, H., 1970: On the adjustment toward balance in primitive equation weather prediction models. *Mon. Wea. Rev.,* **98,** 271–279.

Phillips, N. A., 1957: A coordinate system having some special advantages for numerical forecasting. *J. Meteor.,* **14,** 184–185.

——, 1959: An example of nonlinear computational instability. *The Atmosphere and the Sea in Motion,* Rockefeller Institute Press, 501–504.

——, 1979: The nested grid model. NOAA Tech. Report NWS 22, U.S. Department of Commerce, Washington, DC, 80 pp.

Reed, R. J., 1963: Experiments in 1000 mb prognosis. NMC Tech. Memo. No. 26, Weather Bureau, ESSA, U.S. Department of Commerce, 31 pp.[1]

Richardson, L. F., 1922: *Weather Prediction by Numerical Process.* Cambridge University Press, 236 pp. (Reprint, with a new introduction by Sidney Chapman, Dover Publications, 1965, 236 pp.)

Richtmyer, R. D., 1963: A survey of difference methods for nonsteady fluid dynamics. NCAR Tech. Notes 63-2, Boulder, 25 pp.

Robert, A., 1966: The integration of a low-order spectral form of the primitive equations. *J. Meteor. Soc. Japan,* **44,** 237–245.

Robert, A. J., F. G. Shuman and J. P. Gerrity, Jr., 1970: On partial difference equations in mathematical physics. *Mon. Wea. Rev.,* **98,** 1–6.

Sela, J. G., 1982: The NMC spectral model. NOAA Tech. Report NWS 30, U.S. Department of Commerce, Washington, DC, 36 pp.

Shuman, F. G., 1957a: Predictive consequences of certain physical inconsistencies in the geostrophic barotropic model. *Mon. Wea. Rev.,* **85,** 229–234.

——, 1957b: Numerical methods in weather predication: I. The balance equation. *Mon. Wea. Rev.,* **85,** 329–332.

——, 1962: Numerical experiments with the primitive equations. *Proc. Intern. Symp. Numerical Weather Prediction,* Meteor. Soc. Japan, Tokyo. 85–107.

——, 1971: Resuscitation of an integration procedure. NMC Office Note 54, National Weather Service, NOAA, U.S. Department of Commerce, 15 pp.[1]

——, 1974: Analysis and experiment in nonlinear computational stability. *Proc. of the Symposium on Difference and Spectral Methods for Atmospheric and Ocean Dynamics Problems, Novosibirsk, 1973, Part I.* USSR Academy of Sciences, Siberian Branch, Computing Center, Novosibirsk, 51–81.

——, 1982: Numerical weather prediction. *The Federal Plan for Meteorological Services and Supporting Research for Fiscal Year 1983.* NOAA, U.S. Department of Commerce, 1–13.[1]

——, and J. B. Hovermale, 1968: An operational six-layer primitive equation model. *J. Appl. Meteor.,* **7,** 525–547.

Stackpole, J., 1978: The NMC 9-layer global primitive equation model on a latitude-longitude grid. NMC Office Note 178, National Weather Service, NOAA, U.S. Department of Commerce, 31 pp.[1]

Teweles, S., and H. Wobus, 1954: Verification of prognostic charts. *Bull. Amer. Meteor. Soc.,* **35,** 455–463.

Thompson, P. D., and W. L. Gates, 1956: A test of numerical prediction methods based on the barotropic and two-parameter baroclinic models. *J. Meteor.,* **13,** 127–141.

Weather Analysis and Prediction Division, 1967: External and internal heat sources and sinks in the 6-layer (primitive equation) numerical prediction model. Tech. Procedures Bull. No. 2, Office of Meteorological Operations, Weather Bureau, ESSA, U.S. Department of Commerce, 5 pp.[1]

——, 1968a: Precipitation forecasting in the 6-layer (PE) model. Tech. Procedures Bull. No. 12, Office of Meteorological Operations, Weather Bureau, ESSA, U.S. Department of Commerce, 6 pp.[1]

——, 1968b: Surface topography used in the 6-layer (PE) numerical prediction model. Tech. Procedures Bull. No. 15, Office of Meteorological Operations, Weather Bureau, ESSA, U.S. Department of Commerce, 16 pp.[1]

——, 1970a: Latest change and summary of current procedures in precipitation forecasting in the 6-layer (PE) numerical prediction model. Tech. Procedures Bull. No. 55, Office of Meteorological Operations, National Weather Service, NOAA, U.S. Department of Commerce, 9 pp.[1]

——, 1970b: Revised method of 1000 mb height computation in the PE model. Tech. Procedures Bull. No. 57, Office of Meteorological Operations, National Weather Service, NOAA, U.S. Department of Commerce, 6 pp.[1]

——, 1971a: Convective rain in the PE model. Tech. Procedures Bull. No. 64, Office of Meteorological Operations, National Weather Service, NOAA, U.S. Department of Commerce, 3 pp.[1]

——, 1971b: New initialization procedure for the 6-layer (PE) numerical weather prediction model. Tech. Procedures Bull. No. 65, Office of Meteorological Operations, National Weather Service, NOAA, U.S. Department of Commerce, 5 pp.[1]

Wolff, P. M., 1958: The error in numerical forecasts due to retrogression of ultra-long waves. Tech. Memo. No. 13, Joint Numerical Weather Prediction Unit, Weather Bureau, U.S. Department of Commerce, 22 pp.[1]