Atmospheric predictability experiments with a large numerical model

By E. N. LORENZ,[1] *European Centre for Medium Range Weather Forecasts, Reading RG2 9AX, England*

(Manuscript received January 28, 1982)

ABSTRACT

The instability of the atmosphere places an upper bound on the predictability of instantaneous weather patterns. The skill with which current operational forecasting procedures are observed to perform determines a lower bound. Estimates of both bounds are obtained by comparing the ECMWF operational forecast for each day of a 100-day sequence at one range with the operational forecast for the same day at another range, and with the analysis for that day. The estimated bounds are reasonably close together.

Predictions at least ten days ahead as skilful as predictions now made seven days ahead appear to be possible. Additional improvements at extended range may be realized if the one-day forecast is capable of being improved significantly.

1. Introduction

Although many years ago Richardson (1922) formulated a rather sophisticated procedure for numerical weather prediction, the first moderately successful 24-hour numerical forecast, which had to await the advent of the computer, was based on nothing more complicated than the barotropic vorticity equation (Charney *et al.*, 1950). During the three decades which have subsequently elapsed, as computers have become more and more powerful, and the equations to which they have been applied have been made more and more realistic, numerical weather forecasting has progressed from an experimental to an operational procedure, and the range of operational forecasts has been lengthened several fold. At the European Centre for Medium Range Weather Forecasts (ECMWF), numerical forecasts from one to ten days in advance are now prepared every day for operational use. It is the ECMWF analyses and forecasts which will form the basis of the present study.

During these same decades it has become reasonably well established that prediction of instantaneous weather patterns at sufficiently long range is impossible. This state of affairs arises because of the instability of the atmosphere with respect to perturbations of small amplitude; i.e., two or more slightly different states, each evolving according to the same physical laws, may in due time develop into appreciably different states. Since meterological observations can never determine the state of the entire atmosphere exactly, we cannot tell which of a multitude of nearly identical states is the true present state, and we therefore lack a basis for predicting which of a multitude of considerably different states will occur at some reasonably distant future time.

The lack of complete periodicity in the atmosphere's behavior is sufficient evidence for instability (Lorenz, 1963), but it does not reveal the range at which the uncertainty in prediction must become large. Most estimates of this range have been based on numerical integrations of systems of equations of varying degrees of complexity, starting from two or more rather similar initial states. It has become common practice to measure the error which would be made by assuming one of these

[1] Present Address: Department of Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

 0040-2826/82/060505-09$02.50/0

states to be correct, when in fact another is correct, by the root-mean-square difference between the two fields of wind, temperature, or some other element, and to express the rate of amplification of small errors in terms of a doubling time.

The first systematic study of error growth was performed with a "low-order" model (Lorenz, 1965), in which a state of the atmosphere was represented by only 28 numbers, and its evolution was governed by 28 ordinary differential equations. The growth rate proved to be highly dependent upon the synoptic situation, but, on the long-term average, small errors in wind or temperature doubled in about four days. With present-day accuracy in observations, this would imply that reasonably good one-week forecasts should be attainable, while one-month forecasts would be out of the question.

The results of such a simple model could not be considered definitive, but the implications were important enough to imply that similar studies ought to be made with the most realistic models possible. The available models then were those of Smagorinsky (1963), Mintz (1964), and Leith (1965); predictability studies which they subsequently performed with these models were described by Charney *et al.* (1966), who concluded that a reasonable estimate of the doubling time was five days.

As more refined models were developed and applied to the predictability problem, estimates of the doubling time tended to become smaller. A landmark study was that of Smagorinsky (1969), who used a nine-level primitive-equation model containing moist processes and other features which earlier models had omitted. His numerical integrations indicated a three-day doubling time for the smallest errors. As was the case with other models, the growth rate subsided as the magnitude of the errors increased; obviously the systematic growth must cease when the separate solutions lose all resemblance to one another, since, from that time onward, they are effectively solutions chosen at random.

It might have appeared at this point that the range of acceptable weather forecasts could be extended by three days simply by reducing the observational errors to half their present size—a rather costly but not impossible task. However, the models which have indicated a doubling time of several days do not explicitly contain smaller-scale features ranging in size from squall lines and thunderstorms to dust whirls, whose amplitides should double in hours or minutes or less. The effects of these features upon the larger scales appear in the models, in parameterized form, but the uncertainties in these features do not. It is hardly to be expected that the details of the smaller-scale features will ever be revealed on a global basis by a regular observing network. A study by Lorenz (1969a) indicated that even if the larger scales could be observed perfectly, the inevitable uncertainties in the smaller scales would after a day or so induce errors in the larger scales, comparable to the larger-scale initial errors which presently result from inadequate observations. The induced errors would then grow as if they had been present initially.

It therefore seems likely that the possible accuracy of forecasts at some short range, say one day, is strictly bounded by the existence of smaller-scale features, although presumably we are still a long way from the day when no further improvements can be made. The quoted doubling times of three days or so may therefore be logically interpreted as the doubling time after the first day, but before the errors have become too great.

If the accuracy of one-day forecasts is really bounded, a doubling time effectively places an upper bound upon the extent to which prediction a few days or weeks in advance is possible. Many predictability studies have, in fact, implicitly been concerned only with upper bounds. Lower bounds, although often neglected, are perhaps of equal interest. It is possible to establish a lower bound to predictability by determining how well a particular forecasting procedure regularly performs.

The purpose of this study is to use the output of the ECMWF operational model to obtain estimates of upper and lower bounds to atmospheric predictability, at ranges between a day and about two weeks. We shall not be concerned with climatic or other long-range predictability, which should exist if certain features of the atmosphere can still be predicted when most of the atmosphere cannot. Neither shall we deal explicitly with predictability at ranges shorter than a day.

Since the ECMWF model is not perfect, and since we do not possess the ultimate observational system, our estimated upper and lower bounds may be expected to differ considerably. Future studies should aim at establishing a smaller upper bound and a larger lower bound. When the established bounds have essentially converged upon each other, another chapter in the study of atmospheric predictability will have been completed.

2. The data

Forecasts from one up to ten days in advance are prepared daily at ECMWF. The operational model used for these forecasts is a 15-level global primitive-equation model with moisture and orography. Fields of various meterological elements at various levels are contained in the output, but in our study we have used only the analyzed and predicted 500-mb height fields. We shall refer to these as the analyses and prognoses; we may also refer to an analysis as a zero-day prognosis.

The model is a grid-point model, but, before the 500-mb data are archived, they are transformed into global spherical-harmonic sequences, triangularly truncated at wave number 40. Each height field $z(\lambda, \phi)$, where λ is longitude and ϕ is latitude, is therefore represented by a set of $41 \times 42 = 1722$ spherical-harmonic coefficients A_{mn} or B_{mn}, according to the formula

$$z(\lambda,\phi) = \sum_{m=0}^{40} \sum_{n=m}^{40} (A_{mn} \cos m\lambda + B_{mn} \sin m\lambda) \times P_n^m(\sin \phi), \tag{1}$$

where P_n^m is the associated Legendre function of degree n and order m, suitably normalized.

For our data set we have chosen the 100-day period from 1 December 1980 to 10 March 1981, and we have used the 0,1, . . ., 10-day prognoses *for* the above dates. Our complete data set therefore consists of $100 \times 11 \times 1722 = 1{,}894{,}200$ numbers. Prior to our computations these were placed on a single tape in the form of 1100 records each containing two indicator numbers and 1722 data values.

We may normalize the spherical harmonics so that the average square of P_n^0 is unity, while, for $m > 0$, the average squares of $P_n^m \cos m\lambda$ and $P_n^m \sin m\lambda$ are each unity. The coefficients B_{0n} are not defined by the field of z, and may be set equal to zero. The data are then especially convenient for statistical studies in which global mean squares are to be evaluated, since it follows from (1) that

$$S^{-1}\int_S z^2(\lambda, \phi)\,\mathrm{d}S = \sum_{m=0}^{40} \sum_{n=m}^{40} (A_{mn}^2 + B_{mn}^2), \tag{2}$$

where S is the area of the earth, $\mathrm{d}S$ is an element of area, and the integration extends over the earth's surface. An analogous result holds if the height field $z(\lambda,\phi)$ in (2) is replaced by the difference between two height fields.

We may question the accuracy of the analyses, since there are large regions of sparse data, particularly in the Southern Hemisphere. In such regions the analyses are not only questionable, but the initialization procedure is likely to bias them toward the previous prognoses. Estimates of the model's performance must be viewed with these considerations in mind. The prognoses, on the other hand, do not need to constitute accurate forecasts; it is sufficient that they tell us accurately what the model has predicted.

3. The first experiment

Our initial experiment represents an attempt to update the results of earlier predictability studies, using the ECMWF operational model, which seems to be as up-to-date as any model available. As in the earlier studies, we compare two or more solutions of the same system of equations having somewhat similar initial conditions, and observe the growth rate of the difference between the solutions. It turns out that most of the computing has already been done for us, in the course of preparing the ECMWF operational forecasts.

Specifically, the analysis for a given day, regardless of its accuracy, and the one-day prognosis for the same day represent two states which do not differ too greatly. One-day forecasts made from these two states are simply the one-day and two-day prognoses for the following day. Thus, by comparing the average difference between one-day and two-day prognoses for the same day with the average difference between analyses and one-day prognoses, we can obtain an estimate of the average one-day amplification of moderately small errors.

To obtain the average two-day amplification we need only compare two-day and three-day prognoses for the same day, and, in fact, we can continue this procedure up to nine-day amplifications. Likewise, to obtain amplifications of somewhat larger initial errors we may, for example, compare the average difference between one-day and three-day prognoses with the average difference between analyses and two-day prognoses, or more generally, the difference between j-day and k-day prognoses with the difference between analyses and $(k-j)$-day prognoses.

We have made such comparisons with our data set, If $z_{ij}(\lambda, \phi)$ is the j-day forecast for the value of $z(\lambda, \phi)$ on the ith day of the sample, and E_{jk} is the root-mean-square difference between j-day and k-day prognoses for the same day, averaged over the globe and over all N (=100) days of the sample,

$$E_{jk}^2 = N^{-1} S^{-1} \sum_{i=1}^{N} \int_S [z_{ij}(\lambda, \phi) - z_{ik}(\lambda, \phi)]^2 \, \mathrm{d}S. \quad (3)$$

Recalling eq. (2) we see that

$$E_{jk}^2 = N^{-1} \sum_{i=1}^{N} \sum_{m=0}^{40} \sum_{n=m}^{40} [(A_{mn,ij} - A_{mn,ik})^2 + (B_{mn,ij} - B_{mn,ik})^2] \quad (4)$$

where $A_{mn,ij}$ and $B_{mn,ij}$ are the coefficients in the spherical-harmonic sequence for z_{ij}.

Fig. 1 shows the results. Root-mean-square differences E_{jk}, in meters, for $j < k$, are plotted against k; values of (j, k) are shown beside some of the points. The heavy curve connects values of E_{0k}, and its upward slope as k increases represents the rate of increase of the model's forecast error with increasing range. This is the rate at which solutions of two different systems of equations—the true atmospheric equations and those of the model—diverge as time progresses. The thin curves, on the other hand, connect values of E_{jk} having like values of $k - j$, and their upward slope as k increases represents the rate at which two solutions of the same system of equations—those of the model—diverge. This is the rate which is generally sought in predictability experiments.

We observe that the smallest value of E_{jk}, namely 25 m, requires about 3.5 days to double. Larger errors amplify less rapidly, and show signs of leveling off. Apparently the growth rate is determined by the magnitude; i.e., to a fairly close degree, the thin curves differ from one another only by a horizontal displacement.

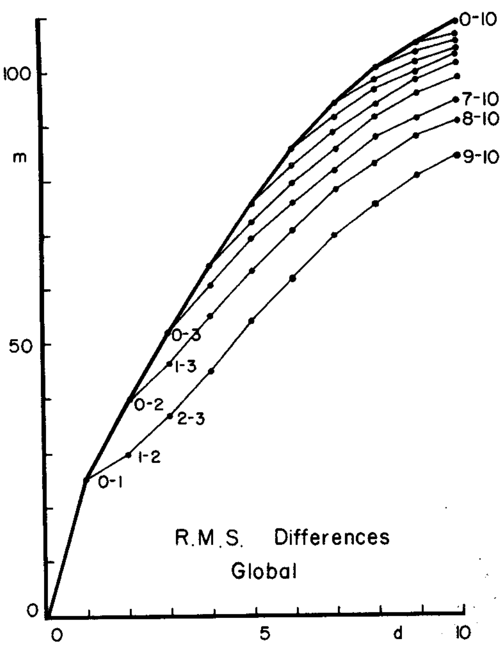

Fig. 1. Global root-mean-square 500-mb height differences E_{jk}, in meters, between j-day and k-day forecasts made by the ECMWF operational model for the same day, for $j < k$, plotted against k. Values of (j,k) are shown beside some of the points. Heavy curve connects values of E_{0k}. Thin curves connect values of E_{jk} for constant $k - j$.

If the analyses and prognoses possessed the same time means, and the same variances about these means, the heavy and the thin curves ought to level off at the same value, which would equal the root-mean-square difference between randomly chosen analyses or prognoses. Actually the prognoses and analyses have different means, the discrepancy increasing with the range of the forecast, and consequently the thin curves level off below the heavy curve, at roughly 110 m.

The doubling time quoted in most predictability studies is the doubling time for errors of very small amplitude. To estimate this time from Fig. 1 we should have to extrapolate the thin curves backward until they approached zero—a rather difficult task. The task becomes easy, however, if we introduce one assumption, In an earlier study of predictability (Lorenz, 1969b), based entirely upon analyses, we found that reasonable, although not readily verifiable, results could be obtained by assuming that the nonlinear terms in the equation governing the growth of E_{jk}, for constant $k - j$, were quadratic. That is, we assumed that

$$dE/dt = aE - bE^2, \tag{5}$$

where t is time. We shall make the same assumption now.

The constant a measures the growth rate of small errors. The quadratic term must be negative if a is positive, since it is the only thing that can halt the growth. If E is normalized so that the value which it approaches as $t \to \infty$ is unity, $b = a$. The solution of (5) is then

$$E/(1 - E) = \exp\,[a(t - t_0)], \tag{6}$$

where t_0 is the time at which $E = \frac{1}{2}$, or, equivalently,

$$E = \tfrac{1}{2} + \tfrac{1}{2}\tanh\,[\tfrac{1}{2}a(t - t_o)]. \tag{7}$$

The thin curves in Fig. 1 do seem to resemble segments of hyperbolic tangent curves. According to (6), the time required for $E/(1 - E)$ to double is independent of t, so that the doubling time for small errors is also the time required for E to increase from $\frac{1}{3}$ to $\frac{1}{2}$, or $\frac{1}{2}$ to $\frac{2}{3}$. From Fig. 1, we find that about five days are required for E_{jk}, with $k - j = 1$, to increase from $\frac{1}{3}$ to $\frac{2}{3}$ of its limiting value, whence our preliminary estimate of the doubling time for small errors is 2.5 days. This is entirely consistent with the results of earlier studies.

The rate at which separate solutions of the model diverge is supposed to approximate the rate at which separate solutions of the true atmospheric equations diverge. If it does, and if, at some time during the forecast, the model could suddenly be replaced by the true equations, the remainder of the heavy curve would follow one of the thin curves. The excess slope of the heavy curve over that of an intersecting thin curve may therefore be regarded as a measure of the maximum amount by which the model may still be improved. Even without further improvement in one-day prediction, the performance of the perfect model should then be given by the lowest thin curve in Fig. 1 and its extrapolation to the right, and skilful forecasts more than two weeks ahead should ultimately be expected.

This is the optimistic view. The pessimistic view is that, as the model is continually made more realistic, the estimate of the doubling time which it yields will continue to decrease. In that event, as the top curve in Fig. 1 drops, the bottom curve will move upward, possibly approaching it at a level closer to the present top curve than the present bottom curve. Further improvements in extended range forecasting will then not be spectacular.

In any event, the two curves can never completely coincide. The small-scale weather features which assure us that by one day there will be a considerable error also assure us that beyond one day a model cannot perform perfectly.

In Fig. 2, we present the relevant material of Fig. 1 in an alternative form. For each of the 45 one-day segments of the thin curves in Fig. 1, we have plotted, as large dots, the increase in root-mean-square error $y = E_{j+1,k+1} - E_{jk}$ against the average root-mean-square error $x = E_{jk} + y/2$. Thus we have plotted a finite difference estimate of dE/dt against E. Likewise, for each of the nine one-day segments of the heavy curve we have plotted, as crosses, $y' = E_{j+1,0} - E_{j0}$ against $x' = E_{j0} + y'/2$.

We see first of all that there is no tendency at all for the crosses to fall among the cluster of dots; in fact, they would come closer to doing so if the values of y' were divided by two. We conclude, as before, that considerable further improvement in forecasting is possible.

If E were really governed by (5), the dots would lie on a parabola. We feel that, although the vertical spread of the dots is obvious, they come remarkably close to doing so, in view of the limited size of the data set.

There is probably no unequivocal definition of the parabola $y = ax - bx^2$ which best fits the dots. We have chosen to minimize a weighted mean

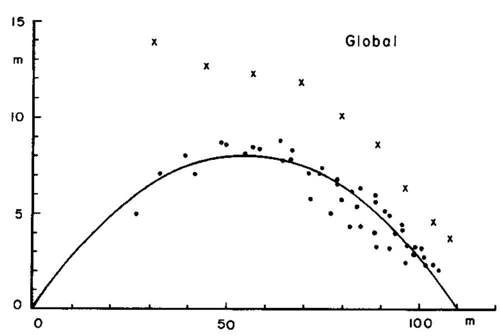

Fig. 2. Increases in global root-mean-square 500-mb height differences, $E_{j+1,k+1} - E_{jk}$, plotted against average height differences $(E_{j+1,k+1} + E_{jk})/2$, in meters, for each one-day segment of each thin curve in Fig. 1 (large dots), and increases $E_{0,k+1} - E_{0k}$ plotted against average differences $(E_{o,k+1} + E_{0k})/2$, for each one-day segment of heavy curve in Fig. 1 (crosses). Parabola of "best fit" to large dots is shown; see text.

square of $y - (ax - bx^2)$; our weighting function is the product of two factors. The first factor, $1/(ax - bx^2)$, is included because we feel that the *ratio* of the heights of the dots and the parabola is important even when both heights are small. The second factor, $a - bx$, is included to give greater weight to individual dots in the left portion of the figure, where the density of dots is smaller. The net effect is a weighting function $1/x$. The resulting parabola, shown in Fig. 2, corresponds to a limiting error of 109.9 m and a doubling time of 2.40 days. We might add that the parabola is not very sensitive to the weighting function; without it the limiting error and the doubling time would be 110.2 m and 2.42 days.

4. Further experiments

In the coming years, as new operational models with additional refinements replace the present ones, and as these new models are applied to new data samples, it should be possible to construct new diagrams similar to Fig. 1, with the slopes of the heavy and thin curves more nearly equal, or to Fig. 2, with the crosses closer to the parabola. In the meantime, we can even now make some changes in the ECMWF model which will improve its performance, and which will enable us to construct the desired diagrams without further numerical integration. We can do this by replacing the prediction X for a predictand Y, where X stands for any spherical-harmonic coefficient A_{mn} or B_{mn} in a prognosis, and Y stands for the same coefficient in an analysis, by the linear function

$$X' = A + BX \tag{8}$$

of X, where A and B are to be chosen so that X' possesses the same temporal mean and standard deviation as Y. A considerable sample of data is needed to obtain good estimates of A and B, and in the present study we have used the same sample used subsequently to evaluate the root-mean-square errors.

It is easily seen that if we should set $B = 1$, choosing A to make only the mean of X' equal to that of Y, the model would perform better in the root-mean-square sense. On the other hand, when we make the variance and hence the standard deviation of X' equal to those of Y, we may actually increase the mean square of $Y - X'$; this is especially likely to be so if X has a smaller variance than Y, and if X and Y are not highly correlated.

We do not believe that in this event we would make a poorer forecast by making X' have the proper variance. We believe instead that the possibility of an increased mean square points to a serious shortcoming of the mean-square error, or root-mean-square error, as a general measure of the goodness of a forecasting procedure.

One property of numerical-weather-prediction models which is considered desirable and possibly essential is that the prognostic maps which they produce should look like real weather maps. In order for these maps to possess migratory synoptic features of the proper intensity, the variables must possess the proper variances. One could eliminate the migratory systems by choosing, as a prediction, the climatological normal weather map plus some very weak superposed random pattern. At a range of six days, such a prediction would be superior to the ECMWF model, in the root-mean square sense, but it is doubtful that any serious numerical modeler would accept it as a replacement for the ECMWF model.

We suggest that a more satisfactory measure of the goodness of a forecasting procedure is afforded by the root-mean-square error *after* the predictions have been modified by replacing the predicted value of each element by the linear function of itself which possesses the correct temporal mean and variance. Evaluated in this manner, the ECMWF model, at ranges up to ten days, is definitely superior to "climatology" plus infinitesimal superposed noise. (Climatology without superposed noise possesses a zero variance, and attempts to choose A and B to correct the variance would result in attempts to divide by zero.) A persistence forecast, incidentally, is superior to climatology plus noise and inferior to the ECMWF model.

Having suggested this measure of goodness, we must point out that correcting the variance of each spherical-harmonic coefficient is not equivalent to correcting the variance at each grid point. We suspect that either procedure would produce a prognostic map resembling a real weather map. Correcting the variance of each coefficient *and* the covariance of each pair of coefficients is equivalent to correcting the variances and covariances of grid-point values. With a much larger data sample it should be feasible to do this after first replacing the set of coefficients, or grid-point values, by a set of uncorrelated linear combinations.

Strictly speaking, the modified ECMWF model is not a suitable model for evaluating error growth. What one should do is to integrate the equations for one day, modify the one-day prognosis according to (8), integrate for a second day from the modified conditions, modify the prognosis again, etc. The advantages and disadvantages of such a procedure have been discussed by Leith (1978). In the present study, in order to obtain answers without performing additional costly numerical integrations, we shall have to assume that the result of integrating for j days, and then modifying once according to (8), does not differ too greatly from the result of making j modifications at one-day intervals.

Fig. 3 is like Fig. 2, for the modified model. The parabola of best fit corresponds to a limiting error of 104.1 m and a doubling time of 2.16 days. The crosses are noticeably lower than in Fig. 2, indicating that the modified model does indeed perform better. On the right side of the figure the crosses are especially close to the dots, implying that any further improvement in the model in the seven-to-ten-day range must be something which improves it at shorter range.

At the same time, it appears that the original model has over-estimated the doubling time. The optimistic view mentioned in the previous section is therefore a bit overoptimistic; if a figure similar to Fig. 1 were constructed, the upper curve would become less steep, but the lower curve would become slightly steeper.

Incidentally, we do not recommend the correction procedure as a permanent step in the continual improvement of the ECMWF or some other model; it is too likely to lead to a dead end. We feel instead that it provides an excellent means of demonstrating that improvements are possible; presumably these can ultimately be realized by representing the physical processes more realistically, or using superior mathematical techniques for solving the equations. Improving the physics or mathematics may very well yield improvements beyond those which the correction procedure would reveal to be possible, If a correction procedure is to be built into a model, it would seem best to remove it completely before introducing any physical or mathematical refinements, possibly reintroducing it afterward.

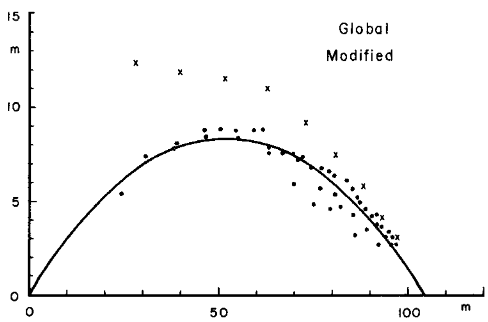

Fig. 3. Same as Fig. 2, but for the modified ECMWF model.

In a final attempt to bring the crosses and the parabola closer together, we have effectively improved the ECMWF model without actually altering it at all; we have simply evaluated its performance for the Northern Hemisphere (NH) alone, again applying the correction (8). Since a good prognosis requires a good analysis, the model behaves like a better model in regions where the data are more plentiful, i.e., extratropical regions of the NH, provided that these are not too closely coupled with the regions where the data are sparser. It seems possible that during forecast intervals of ten days or less, the initial errors in one hemisphere do not contaminate the forecast in the other hemisphere too greatly. At the same time, the contribution of possible poor prognoses in the tropics has been minimized by measuring the error in terms of the 500-mb height, whose variance in the tropics is rather small.

It is a simple matter to verify for the NH alone without transforming from spherical harmonics to latitudes. We replace each Legendre function P_n^m, where $n - m$ is odd, by

$$Q_n^m = \sum_j \overline{P_n^m P_j^m} P_j^m, \tag{9}$$

where the summation is over values of j for which $j - m$ is even, and the bar indicates an average over the NH, P_n^m and P_j^m having been renormalized so that their mean squares are unity. Thus Q_n^m and P_n^m will be equal in the NH, but will have opposite signs in the Southern Hemisphere (SH). When Q_n^m, given by (9), is substituted for P_n^m in (1), new coefficients of the even functions may be evaluated and used in the subsequent computations. Effectively we replace the analyses and prognoses by analyses and prognoses which are unchanged in the NH, but which treat the SH as a mirror image of the NH; this is allowable since we are verifying for the NH only.

The results appear in Fig. 4, which is similar in format to Figs. 2 and 3. The parabola corresponds to a limiting error of 112.6 m and a doubling time of 1.85 days. We believe that the further lowering of the doubling time occurs not because the 2.16-day value is an over-estimate, but because the period of the data is NH winter, when the NH weather systems are most active. We would anticipate slower NH error growth in NH summer, again assuming that the two hemispheres are not too strongly coupled.

Again the crosses have moved closer to the dots. They would fit well in the cluster of dots if their heights were reduced by only 25%.

To further compare the performances of the three "models", relative to the best possible performance, we have constructed Fig. 5 by superposing the parabolas and the points marked by crosses in Figs. 2–4, after first altering the horizontal and vertical scales so that the three parabolas coincide. Effectively we use the limiting error as the unit for measuring the error, and the doubling time as the unit of time. We see that each model represents an improvement over the previous one.

Accepting the modified ECMWF model, applied to the NH, as a state-of-the-art model, we find that we have established upper and lower bounds to atmospheric predictability which are reasonably close together. Assuming that we have correctly estimated the doubling time, we find that, even without further improvement in one-day forecasting, we may eventually make ten-day forecasts as good as present seven-day forecasts, and 13.5-day forecasts as good as present ten-day forecasts. Cutting the one-day root-mean-square error in half should add another two days to the range of predictability; possibly we may cut this error in half more than once.

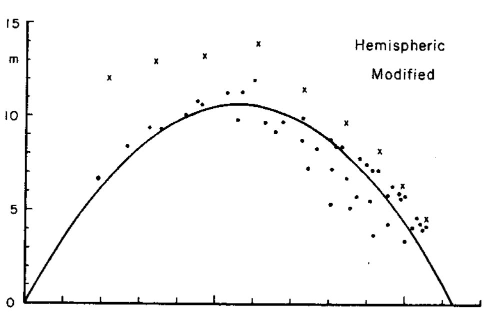

Fig. 4. Same as Fig. 2, but for the modified ECMWF model, for the Northern Hemisphere only.

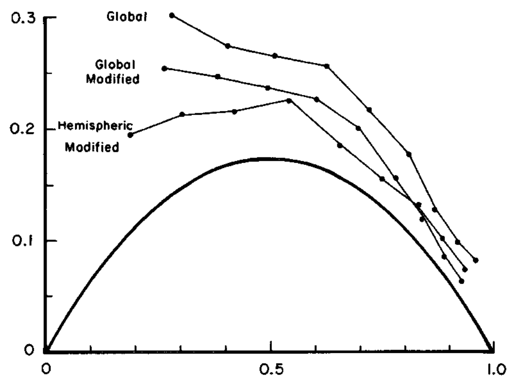

Fig. 5. Superposition of points marked by crosses in Figs. 2–4, after horizontal and vertical scales have been altered so that parabolas coincide. Curves labelled "global", "global modified", and "hemispheric modified" connect points from Figs. 2, 3 and 4 respectively.

5. Conclusions

The instability of the atmosphere with respect to small-amplitude perturbations places an upper bound upon the atmosphere's predictability. A lower bound is afforded by the established skill of forecasting procedures which have already seen operational use. The current state of the art places the two bounds reasonably close together.

For example, on the average, in Northern Hemisphere winter, the time during which the root-mean-square error in predicting the 500-mb height field, with the best possible forecasting procedure, will remain below one seventh of its limiting value is no less than one day; the *additional* time during which this error will remain below six sevenths of its limiting value is no less than seven days nor more than ten days. Unfortunately our study does not yield an upper bound for the time required for the error to reach the level which it presently reaches in one day.

There are a number of reasons why our results cannot be looked upon as the final word. First, we have examined only 500-mb height data. It seems probable that strongly coupled elements will have similar ranges of predictability, so that our conclusions may well also apply to middle-latitude tropospheric temperature and wind prediction, but they may be quite unrealistic for such elements as tropical cloudiness and rainfall.

Second, the analyses by which the performance of the ECMWF model has been evaluated are far from perfect. The points on the upper curve in Fig. 1, and the crosses in Figs. 2–4, are perhaps not properly located. Next, the initial errors whose growth we have studied are rather specialized. They presumably are somewhat like typical prediction errors, and we might have preferred typical analysis errors. Next, we may have relied too heavily upon the assumption that the growth of root-mean-square errors satisfies a quadratic equation. There are curves other than parabolas which would fit the dots in Figs. 2–4 about as well.

Finally, and perhaps most importantly, our conclusions are based upon a rather small data sample, consisting of only 100 consecutive days. Other winters, not to mention summers, may be marked by more, or less, predictable weather.

Nevertheless, we believe that our conclusions are sufficiently well founded to be regarded as indicators of the most promising avenues for future forecasting research. Better-than-guesswork forecasts of instantaneous weather patterns nearly two weeks in advance appear to be possible, and efforts to establish numerical-prediction models which are potentially capable of making such forecasts, and observing systems which enable the models to realize their potentialities, should continue. Skilful forecasts of instantaneous patterns a month or more ahead still appear to be out of the question, and attempts to predict at these longer ranges should be confined mainly to predictions of properties which appear to be predictable, such as weekly, monthly, and longer-period averages and other statistics.

6. Acknowledgments

I am especially grateful to ECMWF for giving me the opportunity to work there during my sabbatical leave from M.I.T. I wish to thank Reinhard Strüfing for preparing the necessary data tape and for assisting with the computations.

REFERENCES

Charney, J. G., Fjörtoft, R. and von Neumann, J. 1950. Numerical integration of the barotropic vorticity equation. *Tellus 2*, 237–254.

Charney, J. G., Fleagle, R. G., Riehl, H., Lally, V. E. and Wark, D. Q. 1966. The feasibility of a global observation and analysis experiment. *Bull. Amer. Meteorol. Soc. 47*, 200–220.

Leith, C. E. 1965. Numerical simulation of the earth's atmosphere. In *Methods in Computational Physics*, Vol. 4, New York, Academic Press, 1–28.

Leith, C. E. 1978. Objective methods for weather prediction. *Ann. Rev. Fluid Mech. 10*, 107–128.

Lorenz, E. N. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci. 20*, 130–141.

Lorenz, E. N. 1965. A study of the predictability of a 28-variable atmospheric model. *Tellus 17*, 321–333.

Lorenz, E. N. 1969a. The predictability of a flow which possesses many scales of motion. *Tellus 21*, 289–307.

Lorenz, E. N. 1969b. Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci. 26*, 636–646.

Mintz, Y. 1964. Very long term global integration of the primitive equations of atmospheric motion. *WMO-IUGG Sympos. Res. Dev. Aspects of Long-Range Forecasting*, World Meteor. Org., Tech. Note No. 66, 141–155.

Richardson, L. F. 1922. *Weather prediction by numerical process*. London, Cambridge Univ. Press, 236 pp.

Smagorinsky, J. 1963. General circulation experiments with the primitive equations. I. The basic experiment. *Mon. Wea. Rev. 91*, 99–164.

Smagorinsky, J. 1969. Problems and promises of deterministic extended range forecasting. *Bull. Amer. Meteorol. Soc. 50*, 286–311.