NOTES AND CORRESPONDENCE

On the Existence of Extended Range Predictability[1]

Edward N. Lorenz

Dept. of Meteorology, Massachusetts Institute of Technology, Cambridge, Mass. 02139

15 November 1972 and 9 January 1973

ABSTRACT

On the basis of five years of Northern Hemisphere isobaric height data, states of the atmosphere separated by 12 days or less are found on the average to resemble each other more closely than randomly selected states, even after adjustment for seasonal trend has been made. The existence of partial predictability of instantaneous weather patterns at least 12 days in advance is thereby confirmed.

During the present century there has been no scarcity of proposed schemes for predicting various aspects of the weather a week or more in advance. Although the proponents have often attributed only marginal skill to their respective schemes, the claims in some instances have been rather spectacular. Recent theoretical studies of atmospheric predictability, which indicate that small but inevitable uncertainties in estimating the state of the atmosphere tend to double every three days during the forecast interval, have sometimes appeared to contradict the more optimistic claims. Indeed, even though it has been stressed that the growth rate of uncertainties subsides once the uncertainties cease to be small, it has often been inferred that there is a range of predictability, of perhaps a few weeks, beyond which virtually no useful forecasting is possible. Appreciable predictability of time-averaged weather patterns is sometimes assumed to exist at ranges where instantaneous patterns are considered almost completely unpredictable.

Regardless of what may be indicated by theory, a conclusive proof that partial predictability exists at a given range would be afforded by any demonstration that at least one forecasting procedure exhibits skill at that range. Since even pure guesses will be correct at times, it goes without saying that to establish skill one must put the procedure to use enough times to obtain a substantial statistical sample. It is the purpose of this note to provide documentation for partial predictability of instantaneous weather patterns at least 12 days in advance.

[1] Research supported by the Atmospheric Sciences Section, National Science Foundation, under Grant GA-28203X.

This note is a by-product of a study by the writer (1969), hereafter referred to as A, which sought to determine the predictability of the atmosphere by examining its behavior following the occurrence of analogues. The data used in A consisted of heights of the 200-, 500-, and 850-mb surfaces, at a grid of 1003 points covering about three-fourths of the Northern Hemisphere, twice daily for the five years 1963–67. As a measure X_{kl}^2 of the difference between the states of the atmosphere at the two times t_k and t_l (the subscripts measure the number of half-days from the start of the data), we chose the ratio of a suitably weighted mean-square height difference to the estimated climatological normal of this difference (for the times of year of t_k and t_l). We also introduced the alternative measure

$$E_{kl}=8\log_2 X_{kl}^2, \tag{1}$$

after which we rounded off E_{kl} to the closest integer, so that effectively the values of E_{kl} represented categories of mean-square height difference, the largest negative values designating the best analogues. For the complete procedure the reader is referred to A.

Values of E_{kl} were determined when the *times of year* of t_k and t_l were separated by a month or less. Only when t_k and t_l themselves were separated by 11 months or more was further use of these values made in A. The number [denoted by $N_\alpha(\infty)$] of occurrences of each value α of E_{kl} is shown in the final column of Table 1 (and also in A). The distribution possesses a mean value [denoted by $\bar{\alpha}(\infty)$] of -0.17 and a standard deviation $\sigma(\alpha)=2.10$. The small negative rather than zero mean occurs because the values were normal-

TABLE 1. Number of occurrences $N_\alpha(m)$ of each observed value α of $E_{k,\,k+2m}$, column totals $N(m)$, distribution means $\bar{\alpha}(m)$, and corresponding lag correlations r_m. Values of $m \geqslant 335$ are classified together as ∞.

	m															
α	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	∞
−24	1															
−23	4															
−22	20															
−21	38															
−20	103															
−19	158															
−18	244															
−17	290															
−16	275	2														
−15	249	6														
−14	181	36	1													
−13	99	62	5													
−12	44	124	6	2												
−11	7	174	29	3	3											2
−10	8	235	67	12	5	3	3			1					2	13
−9	5	277	115	40	20	12	7	3	3	0	2	1	2	1	0	62
−8	2	239	152	83	32	17	7	17	7	7	2	8	5	4	3	230
−7		277	247	125	61	36	20	16	14	7	9	8	11	14	6	631
−6		152	294	202	120	70	58	36	25	30	25	32	28	23	21	2053
−5		77	283	249	200	152	90	75	71	52	53	54	45	43	44	5739
−4		34	225	282	254	210	196	151	140	122	100	80	86	78	66	14602
−3		19	151	304	302	291	274	228	210	187	195	176	144	146	146	30988
−2		3	82	200	282	297	290	317	310	307	241	237	224	231	208	52606
−1		4	43	121	224	249	298	315	306	288	310	277	300	277	273	72980
0		1	18	55	125	211	234	244	261	262	283	314	306	315	329	79154
1			4	30	51	100	126	159	208	232	248	257	275	270	277	69000
2				9	26	35	72	95	93	127	144	153	148	166	187	47785
3				2	8	23	28	46	50	69	59	68	88	82	82	25098
4					1	10	8	8	14	18	33	33	29	43	45	10443
5							3	4	1	2	4	9	15	10	13	3636
6									1		1	3	4	4	4	894
7																190
8																34
9																4
$N(m)$	1728	1722	1722	1719	1716	1716	1714	1714	1714	1711	1709	1710	1710	1707	1706	416144
$\bar{\alpha}(m)$	−16.46	−8.73	−5.72	−4.10	−3.05	−2.32	−1.85	−1.50	−1.28	−1.01	−0.83	−0.72	−0.59	−0.52	−0.39	−0.17
r_m	0.756	0.524	0.382	0.289	0.221	0.170	0.135	0.109	0.092	0.070	0.056	0.047	0.036	0.030	0.019	0.000

ized to give X_{kl}^2 a mean value of unity, and the mean logarithm of a positive variable falls short of the logarithm of the mean.

In this note our concern is with values of E_{kl} when t_k and t_l are separated by 15 days or less. Table 1 also contains the number $N_\alpha(m)$ of occurrences of each value α of E_{kl} when $l=k+2m$, i.e., when t_l follows t_k by m days, for $m \leqslant 15$. The column totals $N(m)$ and the distribution means $\bar{\alpha}(m)$ are also given.

It requires no refined statistical test to decide that the values of E_{kl} summarized in the $m=1$ column are not a sample drawn at random from the population represented by the $m=\infty$ column, i.e., that isobaric height fields separated by one day resemble each other more closely than height fields at two randomly chosen times. The same conclusion is evidently true for separations up to five days, although to a lesser degree. On the other hand, the distributions of E_{kl} within the columns $m=12$–15 are not obviously different from the population distribution, and in each case the most frequently occurring value is zero.

To determine just how far apart t_k and t_l may become while the corresponding states still retain more than chance resemblance, we may, for each m, test the plausibility of the null hypothesis that the population, of which the values of α in the mth column constitute a sample, is identical with the population represented by the final column, and that the departure of $\bar{\alpha}(m)$ from $\bar{\alpha}(\infty)$ is due only to chance selection. For this purpose we require the standard deviation $\sigma[\bar{\alpha}(m)]$ of the distribution of $\bar{\alpha}(m)$ which would be obtained by re-performing the computations many times with different five-year data sets. It would be most inappropriate simply to divide $\sigma(\alpha)$ by the square root of $N(m)$, since the $N(m)$ values of E_{kl} comprising a sample correspond to consecutive values of $k/2$ or $l/2$ (consecutive days), and are therefore not independent. Instead,

$$\sigma^2[\bar{\alpha}(m)] = N^{-1}(m)\sigma^2(\alpha) \sum_{n=-\infty}^{\infty} R_n, \qquad (2)$$

where

$$R_n = \sigma^{-2}(\alpha)\overline{[E_{kl}-\bar{\alpha}(\infty)][E_{k+2n,\,l+2n}-\bar{\alpha}(\infty)]} \qquad (3)$$

is a special serial correlation of E_{kl}. Here the bar denotes an average with respect to k and l.

Data sufficient for evaluating R_n for $0<n\leqslant 8$ are found in Tables 1 and 3 of A. We find that $R_1=0.833$,

while, to a close approximation $R_n=(0.75)^{n-1}R_1$. Assuming that the latter relation holds also for $n>8$, we find that $\Sigma R_n=7.66$, whereupon $\sigma[\bar{\alpha}(m)]=0.14$, approximately, for each value of m. Effectively, the 1700 or more members of each sample comprise only about 225 independent members.

Thus, for $m=12$, 13, 14, 15, the values of $\bar{\alpha}(m)$ depart from the hypothesized normal value -0.17 by 3.9, 3.0, 2.5, 1.6 standard deviations, respectively. The corresponding probabilities of obtaining departures as large as these by chance are 0.0001, 0.003, 0.012, 0.110. With conventional criteria of significance, there is insufficient evidence for concluding that states separated by 15 days resemble each other more closely than randomly selected states. A 14-day separation is marginal, while the result for 12 days is highly significant.

Strictly speaking, the 416,144 values of E_{kl} summarized in the final column of Table 1 form only a sample, although a very large one, which is, in fact, composed of many partly independent samples similar in structure to those in the earlier columns; therefore, $\bar{\alpha}(\infty)$ possesses a distribution, although with a standard deviation $\sigma[\bar{\alpha}(\infty)]$ smaller than $\sigma[\bar{\alpha}(m)]$. As a result a preferable procedure would have been to determine the standard deviation of $\bar{\alpha}(m)-\bar{\alpha}(\infty)$, rather than simply of $\bar{\alpha}(m)$. If anything, $\bar{\alpha}(m)$ and $\bar{\alpha}(\infty)$ should be positively rather than negatively correlated from one five-year data set to another, whereupon

$$\sigma^2[\bar{\alpha}(m)-\bar{\alpha}(\infty)]\leqslant\sigma^2[\bar{\alpha}(m)]+\sigma^2[\bar{\alpha}(\infty)]. \quad (4)$$

It follows that use of the revised procedure would reduce the quoted departures of 3.9, 3.0, 2.5 and 1.6 only by a fairly small percentage, and at most would shift the marginal separation time from 14 to 13 days. It therefore appears that we have processed a large enough sample of data to show that for the Northern Hemisphere as a whole, in terms of isobaric heights, two instantaneous states of the atmosphere separated by 12 days or less resemble each other more closely than two randomly selected states, even after adjustment for the effect of normal seasonal variations has been made.

We have noted that the existence of *predictability* at a given range may be established by demonstrating a forecasting procedure which exhibits skill. It is evident that anyone can make a plausible forecast at any range by pure guesswork, if this consists of choosing a past atmospheric state at random and using this state for the future predicted state. One can make a considerably better prediction, at least in terms of mean-square error, by "climatology," i.e., by choosing the climatological mean state as the predicted state. It follows that if any more elaborate procedure is to justify the effort needed for its development and use, it must be superior in some sense not only to guesswork but also to climatology. Pure persistence will not serve this purpose (beyond two days), even though we have shown it to be superior to guesswork up to 12 days.

TABLE 2. Relative mean-square prediction errors at a range of m days yielded by selected prediction procedures.

Procedure	Mean-square error
Guesswork	1
Climatology	$\frac{1}{2}$
Persistence	$1-r_m$
Damped persistence	$\frac{1}{2}(1-r_m^2)$

A simple procedure which in the long run is superior or equal to both climatology and persistence is *damped persistence*, given by

$$\hat{Z}(k+2m)-\bar{Z}(k+2m)=r_m[Z(k)-\bar{Z}(k)], \quad (5)$$

where Z, $\hat{Z}$ and $\bar{Z}$ denote respectively the observed, predicted and climatological normal isobaric height fields at the indicated times, and

$$r_m=1-\overline{X_{k,k+2m}^2} \quad (6)$$

is a space-and-time averaged correlation between height fields separated by m days (and is not to be confused with R_m). Table 2 shows mean-square errors yielded by climatology, persistence, and damped persistence, relative to that yielded by guesswork. It is evident that damped persistence is superior to climatology and persistence whenever it can be established that $r_m\neq 0$. One should also note that for damped persistence to be superior in practice, a data sample large enough to determine r_m with reasonable accuracy must be processed.

There is always room for some difference of opinion as to what constitutes skill in prediction, and one might argue that damped persistence rather than climatology ought to serve as the zero mark on the skill scale. In this event, however, one could probably by repeated use of similar reasoning attribute zero skill to successively more elaborate statistical forecasting schemes. We prefer to regard damped persistence as possessing some forecasting skill, since, unlike climatology, its application does require some implicit knowledge of the chronological order in which states of the atmosphere occur.

Values of r_m for lags up to 15 days occur in the final row of Table 1. They have been computed from the values of $\bar{\alpha}(m)$, under the assumption that, as in the final column, $8\,\overline{\log_2 X_{kl}^2}$ falls short of $8\log_2\overline{X_{kl}^2}$ by 0.17 units. Statistically significant values of r_m are those which correspond to significant values of $\bar{\alpha}(m)$; these occur for lags up to 12 days.

Obviously at a range of 12 days the superiority of damped persistence over climatology is barely measurable, and, indeed, the individual forecasts produced by the two methods are barely distinguishable. It is also obvious that simple formulas superior to (5) could easily be constructed. For example, r_m could be made to depend upon geographical position, or the time of year.

The latter possibility could, in fact, be explored with the same data set, by repeating the work separately for different times of year, say for four seasons. We have made a few spot calculations. At the outset it would appear more difficult to establish predictability for a single season at a given range, even if it exists, because of the smaller sample size. Moreover, for winter we encountered larger values of $\sigma(\alpha)$ and R_n in Eq. (2), which made $\sigma[\bar{\alpha}(m)]=0.35$, as opposed to 0.14 for the whole year. For summer, when $\sigma(\alpha)$ and R_n were smaller, $\sigma[\bar{\alpha}(m)]=0.24$. Significant values of r_m at the 1% level, with our data sample, are therefore 0.075 for winter and 0.052 for summer, as opposed to 0.031 for the whole year.

Nevertheless, we found significant predictability in winter not only at 12 but also at 15 days, with $r_{12}=0.105$ and $r_{15}=0.087$. Summer isobaric height patterns appeared to be considerably less predictable, with $r_9=0.045$ just failing to meet the significance criterion, and $r_{12}=0.001$. Our earlier statement to the effect that weather situations separated by 12 days or less bear more than random resemblance to one another therefore tacitly assumes that the time of year is randomly chosen.

Presumably we could obtain still better prediction formulas at extended range by not restricting them to simple variants of damped persistence. However, the purpose of this note is not to establish a working procedure for extended prediction; it is merely to show that, at least up to 12 days, some predictability of instantaneous patterns, although perhaps not very much, is present. This result could not have been established with comparable confidence from a smaller sample of data. Meanwhile, we might note that at lags of five days or so, r_m is high enough so that in assessing currently used methods one would do well to compare their performance with that of damped persistence.

REFERENCE

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