DETERMINISTIC AND STOCHASTIC ASPECTS OF ATMOSPHERIC DYNAMICS

Edward N. Lorenz
Center for Meteorology and Physical Oceanography
Massachusetts Institute of Technology
Cambridge, Massachusetts, U.S.A.

ABSTRACT. We raise the question as to whether the atmosphere should be treated as deterministic or stochastic, for the purpose of investigating atmospheric dynamics most effectively. Because the atmospheric equations are nonlinear, all but special solutions must be sought numerically. The range of scales which numerical models can handle explicitly is limited, and the influence of smaller scales must be introduced through parameterization. The most realistic parameterizations contain stochastic terms in addition to the deterministic ones. However, since realistic atmospheric models ordinarily possess aperiodic general solutions with or without their stochastic terms, they tend to yield similar results in either event. The choice between a deterministic and a stochastic formulation of the equations can therefore be dictated by convenience.

1. INTRODUCTION

Among the many questions which have inspired considerable debate among meteorologists, one in particular has also attracted some prominent mathematicians: Should the weather be treated as a deterministic or a stochastic process, for the purpose of making the best attainable weather forecasts? The differences of opinion have led to the development of two rather different objective methods of weather forecasting, popularly known as numerical and statistical weather prediction. In the former method one attempts to predict future atmospheric states by integrating formally deterministic systems of differential or integro-differential equations which represent the governing physical laws, using observed values of atmospheric variables as initial conditions. In the latter one attempts to establish formulas which minimize the expected mean-square error in prediction, using observations of past weather to determine the numerical values of the coefficients in the formulas. Former champions of the two methods include John von Neumann and Norbert Wiener [1].

One should not conclude that practitioners of numerical weather prediction believe that the atmosphere is deterministic. The assumption

C. Nicolis and G. Nicolis (eds.), Irreversible Phenomena and Dynamical Systems Analysis in Geosciences, 159–179.

is simply that, despite any possible randomness, a formally deterministic approach will produce acceptable forecasts. Likewise, the use of statistical weather prediction does not presuppose any randomness. The assumption is simply that the laws, even if deterministic, may not be perfectly known or may be too difficult to apply, and that empirical procedures offer an acceptable alternative.

It might be added that the equations generally used in numerical weather prediction, even though formally deterministic, are not derived exclusively from the physical laws, but contain some empirically determined functions and coefficients. Likewise, the selection of predictors to be used in a statistical forecasting scheme is often guided by a knowledge of the physical laws.

Since we shall be dealing with the general topic of atmospheric dynamics, let us pose the following more general question: Should the atmosphere be treated as deterministic or stochastic, for the purpose of investigating atmospheric dynamics most effectively? Our ensuing discussion will be directed toward reaching a suitable answer to this question.

We must immediately note that the system in whose deterministic or stochastic nature we are interested is actually not restricted to the atmosphere, but includes also those portions of the underlying oceans and continents which influence the atmosphere, and which in turn are significantly influenced by the atmosphere. It therefore includes at least the upper layers of the ocean and land, and the sea ice and continental snow and ice cover and soil moisture. We shall nevertheless find it convenient to refer to this system as the atmosphere.

We should also point out that neither question is equivalent to asking whether the atmosphere actually is deterministic or stochastic. We shall not dwell at length on the determinism of the atmosphere, and simply note that it is influenced to some extent by human activity, particularly when that activity consists of clearing large forests or building dams to create large lakes. Even on rather short time scales, intentional or inadvertent weather modification through cloud seeding or setting large fires sometimes occurs. Any claim that the future of the atmosphere is predetermined would therefore imply a claim that human activity is predetermined. However, our concern in this discussion is not whether the behavior of the atmosphere involves some randomness, but whether it is important to take any such randomness into account.

2. OBSERVATIONS AND PHYSICAL LAWS

To deal effectively with the dynamics of any time-dependent system, whether it is a spiral galaxy, a planetary atmosphere, a glacier, or a small waterfall, we need a set of observations and a set of governing physical laws. Observations are needed first of all to make us aware of the system's existence, and subsequently with some degree of precision to reveal the system's typical structure and behavior. The goal of dynamical studies is to explain the observed features in terms of the physical laws, and sometimes to anticipate or predict additional features which have not yet been observed.

At this point we may ask why the roles of observations and laws should not be reversed, i.e., why the goal of dynamical studies should not be to deduce the physical laws from the observations. The answer is that this might well be the goal in dealing with certain systems. In historical times the discovery of the laws of motion was facilitated by observations of the motions of the planets. In the case of present-day studies of systems like the atmosphere, we can still deduce rules, some of which may be useful for weather forecasting or other practical tasks, but we tend to think of the physical laws are something more basic than specialized rules; possibly an extensive set of rules could be analyzed into basic laws. We also assume that the most basic laws--the laws of motion and thermodynamics--are already known with sufficient precision.

Turning to the atmosphere, let us enumerate a few observed features. Some of the qualitative features are revealed by casual observation; quantitative measurements may require sophisticated instrumentation.

First, the atmosphere consists mostly of a gas, with small amounts of liquid and solid matter. Readily noticed properties are the wind and the temperature; these vary from one location to another, and at any location they vary from one time to another. The pressure and density vary similarly, although the changes at one elevation might go unnoticed in the absence of instruments. Most of the gaseous constituents occur in nearly constant proportions, the most notable exception being water vapor. At high levels, variations in ozone content are significant, while over long periods the carbon dioxide content appears to undergo progressive changes. Liquid and solid water occur in the form of droplets and small crystals which are suspended as clouds, and larger drops and flakes which fall out as precipitation. Dust and other solid matter also occur in variable concentrations. The state of the atmosphere may be expressed in terms of the spatial and temporal distributions of wind components, temperature, pressure, density, mixing ratios of the various phases of water, and concentrations of other substances such as dust.

Observations also show that the atmospheric variables are not randomly distributed, but that certain spatial and temporal distributions are highly favored over others, so that the atmosphere tends to be organized into identifiable structures, each having a typical size and shape, and life span and life history. A partial listing of these structures, arranged in order of decreasing size, could include circumpolar westerly wind belts, migratory extratropical cyclones, tropical cyclones, squall lines, thunderstorms, fair-weather cumulus clouds, tornado funnels, individual wind gusts, hailstones, snow crystals, and cloud droplets. Separate occurrences of a particular structure, other than a cloud droplet, are generally not exact repetitions, but they tend to have much in common.

The laws governing the atmosphere include the basic laws of motion and thermodynamics, and some more specialized laws involving such processes as the absorption, emission, and scattering of radiation by atmospheric constituents and the changes of phase of water. The latter are complicated by the occurrence of water in the form of cloud droplets and ice crystals, and the presence of hygroscopic particles. The laws are commonly expressed as a system of mathematical equations. A typical equation may be written

$$dx/dt = F, \quad (1)$$

where t represents time, x represents the value of an atmospheric variable, such as temperature or a wind component, and F represents the sum of the physical processes which change x.

As originally formulated the laws apply to a fixed mass. Because the atmosphere is a fluid whose different parts move at different velocities, an initially concentrated mass tends to be stretched and twisted, so that tracing a particular mass may prove difficult. It is thus more convenient to introduce a coordinate system which is fixed in the atmosphere, and to rewrite eq. (1) to apply to fixed locations. The result is

$$\partial x/\partial t = - \underline{v} \cdot \underline{\nabla} x + F, \quad (2)$$

where $\underline{v}$ represents the three-dimensional wind vector.

3. APPLICATION OF THE PHYSICAL LAWS

The difficulties encountered in applying eq. (2) directly to the atmosphere become apparent when we ask what we mean by a point, at which eq. (2) is to be applied. Certainly we do not mean a geometrical point, which would be smaller than a molecule, whence the velocity and temperature at the point would not even be defined. The use of the gradient operator in eq. (2) implies that we are treating the atmosphere as a continuum, and any "point" must actually be large enough to contain many molecules. Similarly, except in clear air, any point must be large enough to include many cloud droplets. Values of the variables at such a point must actually be averages over a region with a diameter of at least a centimeter.

Further difficulties appear when we note that we do not have observations spaced at one-centimenter intervals, and, except in data sets gathered for special studies, we do not even have observations at ten-kilometer intervals. To apply the equations to globally distributed observational data, we must therefore interpret "values at a point" as meaning averages over regions with horizontal extents comparable to 100 kilometers. For dealing with the internal dynamics of individual structures, such as thunderstorms, the regions may be considerably smaller.

These requirements might seem to disappear when we apply the equations to idealized rather than observed distributions of the variables, but here another practical difficulty arises. The advective term $-\underline{v} \cdot \underline{\nabla} x$ in eq. (2) contains the product of one variable, $\underline{v}$, with the gradient of another variable, x, and is therefore inherently nonlinear. Of course, the function F might also be nonlinear, as in the case when it represents the effect on temperature of radiative heating and cooling. Analytic solutions of nonlinear equations generally represent specialized cases, with different properties from the general solutions, and, particularly since the advent of high-speed computers, it has been customary to seek approximate solutions using numerical methods. Even the largest computers have their limitations, and at present it is not practical to solve the equations when the state of the atmosphere, or of the individual atmospheric structure which is being studied, is represented by more than about one million numbers. Even restriction of the data to four physical variables and ten elevations would allow only 25000 regions at each elevation, and for global coverage the diameter of a region would still have to exceed 100 kilometers.

It appears, then, that eq. (2) should be replaced by an equation governing changes of the average values of the variables over rather extensive regions--regions which may actually contain many of the smaller individual structures. One way to do this is to use the equation of continuity of mass,

$$\partial\rho/\partial t = -\underline{\nabla} \cdot (\rho\underline{v}), \tag{3}$$

where ρ represents density. When combined with eq. (3), eq. (2) becomes

$$\partial(\rho x)/\partial t = -\underline{\nabla} \cdot (\rho x\underline{v}) + \rho F \; ; \tag{4}$$

the quantities ρx and ρF are values per unit volume when x and F are values per unit mass. When averaged over a region, eq. (4) becomes

$$\partial\, \overline{\rho x}/\partial t = -\underline{\nabla} \cdot (\overline{\rho x}\; \overline{\underline{v}}) - \nabla \cdot \overline{(\rho x)'\underline{v}'} + \overline{\rho F}, \tag{5}$$

where a bar over a quantity denotes a regional average, and a prime denotes a local departure from a regional average. In addition to terms obtained simply by replacing quantities in eq. (4) by their averages, we find a new term which depends on variations within the region.

The averaging process which produces eq. (5) entails two new practical difficulties. First, the regional averages which may be computed from observational data are generally averages of rather small statistical samples, and may therefore contain sampling errors. If the observing stations in a particular region had been established at slightly different locations, the computed averages at any time would presumably be somewhat different. A region might contain a single thunderstorm, and the computed average temperature, wind, and water content will depend upon whether the thunderstorm coincides with an observing station. Thus, for example, a realistic processing of a data set, instead of concluding that the average temperature over a region is 18°C, might more realistically conclude that it is 18°C plus an error, whose expected absolute value is 1.2°C.

The second difficulty involves the term $-\underline{\nabla} \cdot \overline{(\rho x)'\underline{v}'}$ in eq. (5). This term includes the transport of the property represented by x across the boundary of the averaging region by the circulations associated with structures of smaller spatial scale than the region itself. If, for example, the regions are 200 kilometers square and 2 kilometers deep, the term includes the exchange, between vertically adjacent regions, of heat and water and possibly momentum by cumulus-cloud circulations. If the region is considerably smaller, it may still include transports by individual wind gusts.

Since these transports often account for important fractions of the total change of a variable, it is important that the equation which we finally use in our investigations should not disregard them altogether. It is presently standard practice to include the effects through parameterization [2]. That is, we assume that the effects can be expressed reasonably well as functions of the averaged variables which now appear as dependent variables in our equations.

For example, we might assume that the number and size distribution and typical structure of cumulus clouds in a region is fairly well determined by the average temperature, moisture, and wind velocity in the region, and the manner in which these averages vary from this region to the regions immediately above and below. From the assumed cumulus-cloud statistics we could evaluate the amounts of heat, moisture, and momentum carried upward or downward by the cloud circulations. We would then include expressions for these amounts, in terms of the averaged quantities, as additional terms in our equations.

It has been found through experience that systems of equations which parameterize the effects of unresolved processes can perform considerably better than those which merely disregard the effects. Nevertheless, the distribution of cumulus clouds or other small-scale structures contained in any region or influencing the region at any instant constitutes at best a statistical sample drawn from the set of distributions which could conceivably have been present. A cloud which occupies a region may move out of the region within a few minutes, or it may alter its shape considerably, during which time the average properties of the region may not have detectably changed. We therefore face a sampling problem again; in addition to a deterministic term, which represents the "expected" or most probable effect of the unresolved structures, the equations should contain a stochastic term, representing the distribution of departures from the expected effect. It appears that this uncertainty in the equations far outweighs any possible uncertainty due to unpredictable human behavior.

We thus find that averaging produces two types of uncertainty, one in estimating the initial state, and one in formulating the governing laws. The effects of these uncertainties on operational weather forecasts appear to be far from negligible.

Instead of introducing regional averages we may expand the fields of the variables in series of orthogonal functions, such as spherical harmonics if the investigation is global, or multiple Fourier series if it is local. The coefficients in the series then become the new dependent variables. However, in order to retain a finite system, we must discard all but a finite number of coefficients in each series; ordinarily these represent features of small spatial scale. We find that the difficulties introduced by averaging, although perhaps slightly alleviated, are by no means eliminated. After all, a Fourier coefficient is nothing more than an average of the product of a variable and a trigonometric function, and statistical sampling problems remain.

It would thus seem that the atmosphere might best be treated as a stochastic system. That this is not necessarily the case will become evident after we consider in detail the phenomenon of chaos.

4. CHAOS

The term "chaos" has been used in mathematical and physical works with a number of meanings [3]. Often it is used as a synonym for randomness or lack of complete determinism, so that, in this sense, any stochastic process would be chaotic. More recently the term has been used to describe any system which varies aperiodically, or perhaps more often any system of equations where, in some sense, almost all solutions are aperiodic [4]. Under the category of periodic solutions we include not only those which exactly repeat themselves, but also those which eventually acquire a state arbitrarily close to some previous state, provided that the evolution following the near repetition remains arbitrarily close to the evolution following the original occurrence. We also include any other solutions which asymptotically approach those solutions which we have already included as periodic.

The distinction between the concepts of chaos more or less disappears if we confine our attention to finite systems of linear ordinary differential equations, since the solutions of these equations are generally periodic if the equations are deterministic, and not exactly periodic if stochastic terms are added. The concepts differ when we turn to nonlinear equations, which often have aperiodic general solutions even if they are deterministically formulated.

Some investigators prefer to reserve the term "chaos" for those aperiodically varying systems which are governed by formally deterministic equations. Others liberalize the definition to include stochastic systems, provided that it appears that the system would remain aperiodic even if the stochastic part of the governing equations were eliminated. This modification makes it possible to include real physical systems, whose actual determinism is likely to be in doubt.

The feature of aperiodically varying systems which has earned them the designation of "chaos" is their sensitive dependence on initial conditions [5]. If a system possesses a finite number of variables, and if each variable continues to oscillate between fixed upper and lower bounds, the system will in due time necessarily assume a state arbitrarily close to some previously encountered state. By definition, if the system is aperiodic, the evolution following the near repetition cannot forever remain arbitrarily close to the evolution following the original occurrence. If there is no semblance of periodicity, the evolutions following the two occurrences will ultimately go their own ways. Thus two states which are nearly alike will ultimately evolve into two states which lack any resemblance. If, for example, the system is a chaotic atmosphere and the observations are anything but exact, there will be no basis for choosing among a number of possible evolutions, and weather forecasting at some sufficiently distant range will be impossible.

The idea that deterministic equations may have aperiodic general solutions is not particularly new. For a long time, investigators of fluid turbulence have worked with deterministically formulated systems, and have assumed that the turbulent motion which satisfies these equations is non-repetitive. What is relatively new is the general realization that systems consisting of very few simple nonlinear ordinary differential equations may have aperiodic solutions.

The possibility of studying chaos with small systems, together with the general availability of high-speed computers, has made it feasible to examine the attractors of chaotic systems [4]. An attractor is actually a kind of multidimensional graph, and it is most easily described in terms of the phase space of a system. This is a Euclidean space with as many dimensions as the number of variables, and these varibles serve as coordinates. An instantaneous state is thus represented by a point in the phase space, while a time-variable solution is represented by an orbit.

A particular point which is approached arbitrarily closely, arbitrarily often, by a point traversing a given orbit is an attracting point for that orbit. A point which has a greater-than-zero probability of being an attracting point for a randomly selected orbit is a point of the attractor set. This set may be connected, or it may consist of a number of disjoint connected sets, in which case each of these is an attractor.

If almost all solutions of a system of equations approach a single repeating solution asymptotically, the attractor is simply the closed orbit representing this solution. If the system is chaotic, the attractor set is generally more complicated. When the system in question is the atmosphere, points on the attractor set represent states which are likely to be approximated again and again as the weather continues to evolve, i.e., states which are compatible with the climate. For example, hypothetical states where the poles are warm and the equator is cold, where the surface winds are everywhere of hurricane strength, or where the winds blow the wrong way around most of the high and low pressure centers are represented by points which are <u>not</u> on the attractor set.

To obtain an approximate picture of an attractor, we may select an arbitrary initial state and perform an extended numerical integration. We discard the leading part of the solution as possibly representing transient conditions, and assume that the remaining part lies as close to the attractor as the resolution will allow [6]. Unless the system consists of only two equations, an actual picture is likely to be the projection of the attractor on a plane, or the intersection with a plane.

A procedure which is equally good in concept although more difficult to approximate in practice consists of taking a small sphere centered at an arbitrary point, and finding the successive shapes into which the interior of the sphere is deformed as each point in the interior moves along its orbit. Ultimately the deformed sphere should look like an attractor, or perhaps several attractors connected by infinitesimal threads. If the system is dissipative, the volume of the deformed sphere will shrink toward zero. If it is also chaotic, with sensitive dependence on initial conditions, the maximum diameter will grow. If the sphere is initially very small, it will for a while be deformed into an approximate ellipsoid. The long-term average rates of stretching or compression of the axes of the ellipsoid are called the Lyapunov exponents, and the condition for chaos is that at least one exponent should be positive [4]. If the equations are differential rather than difference equations, one exponent will also be zero, indicating that two points within the sphere moving along the same orbit will tend to maintain their initial separation. If the system is dissipative, the sum of the exponents will be negative.

As the deformation continues, the ellipsoidal shape will be lost. In the case of a dissipative chaotic system of three ordinary differential equations, the deformed sphere will come to resemble a strip of paper, which is continually becoming thinner but increasing in area. As the strip is stretched, it is bent and twisted so that it continues to fit within a reasonably confined volume. In due time different parts of the strip will be brought close to one another, so that locally two and then several sheets of paper will appear to be pressed together, although they will never actually merge. In the limit there will be an infinite number of sheets, which an ordinary picture might resolve into several. A transverse line will intersect these sheets in a Cantor set. An attractor with such a Cantor-set structure is called a strange attractor [7].

The above arguments may be generalized to systems of more than three equations. It would be difficult to draw a picture of a chaotic attractor, or even visualize its shape, in a high-dimensional system, so mathematicians who are principally interested in the topology of attractors have tended to use small systems as illustrative examples. Obtaining a picture of an obviously strange attractor is often an effective way of convincing oneself that a given system is chaotic.

5. EXAMPLES OF CHAOS

For a first example we shall choose one of the simplest possible nonlinear systems--the quadratic difference equation

$$x_{n+1} = (x_n - 2)^2 \tag{6}$$

in the single variable x. Starting with an initial value x_0 of x, with $0 \leq x_0 \leq 4$, we let x_n be the value of x after n applications of eq. (6). It is evident that $0 \leq x_n \leq 4$ for all values of n.

If x_0 is an even integer, the sequence $x_0, x_1, \ldots$ soon becomes a repetition of 4's, while if x_0 is an odd integer, it becomes a repetion of 1's, but, for most non-integer values of x_0, the sequence is aperiodic. Special values of x_0 where the sequence is periodic but not steady include $x_0 = (3 + \sqrt{5})/2 = 2.618$, when x alternates between $(3 + \sqrt{5})/2$ and $(3 - \sqrt{5})/2$.

Table I shows values of x when $x_0 = 0.5$. The lack of periodicity is apparent. These values are compared with values when $x_0 = 0.5001$, and the differences are also tabulated. The continual although non-uniform increase in the difference, until it becomes large, is typical of the sensitive dependence on initial conditions exhibited by aperiodically varying systems.

To demonstrate that the solutions are truly aperiodic, we let y_n be one of the solutions of the equation

$$x_n = 2(1 + \cos(2\pi y_n)), \tag{7}$$

with $0 \leqq y_n < 1$. It follows from eq. (6) that

$$x_{n+1} = 2(1 + \cos(4\pi y_n)), \tag{8}$$

so that eqs. (6) and (7) can be satisfied for all values of n, with $0 \leqq y_n < 1$, if

$$y_{n+1} = 2y_n (\mathrm{mod}\ 1) . \tag{9}$$

Table I. Values x_n' and x_n'' of x_n determined by successive iterations of eq. (6) from initial values x_0' and x_0'', and difference $\varepsilon_n = x_n''-x_n'$.

n	x_n'	x_n''	ε_n
0	0.5000	0.5001	0.0001
1	2.2500	2.2497	-0.0003
2	0.0625	0.0624	-0.0001
3	3.7539	3.7545	0.0006
4	3.0762	3.0782	0.0020
5	1.1582	1.1626	0.0044
6	0.7087	0.7013	-0.0074
7	1.6676	1.6866	0.0191
8	0.1105	0.0982	-0.0123
9	3.5701	3.6169	0.0468
10	2.4653	2.6143	0.1490
11	0.2165	0.3774	0.1609
12	3.1809	2.6329	-0.5480
13	1.3945	0.4006	-0.9939
14	0.3667	2.5582	2.1915
15	2.6677	0.3116	-2.3561
16	0.4458	2.8506	2.4048
17	2.4155	0.7236	-1.6919
18	0.1726	1.6292	1.4566
19	3.3394	0.1375	-3.2019
20	1.7940	3.4689	1.6750

If y_n is expressed in binary notation, eq. (9) simply shifts the bits of y_n one place to the left and drops the leading bit. The sequence $y_0, y_1, \ldots$, and hence the related sequence $x_0, x_1, \ldots$, is therefore periodic or aperiodic according to whether the bits of y_0 are arranged periodically or aperiodically, i.e., according to whether or not y_0 is a rational fraction. Thus almost all choices of y_0, and hence of x_0, lead to aperiodicity. For example, the value $x_0 = 1/2$ of Table I corresponds to $y_0 = (1/2\pi) \cos^{-1}(-3/4)$, which is not a rational fraction.

It will not surprise anyone to learn that one can obtain aperiodic solutions of a discontinuous equation simply by choosing an aperiodic infinite sequence of 0's and 1's and continually shifting left and removing the leading member. What is not so obvious until the above analysis is performed is that the possibility of doing so implies also that one can find aperiodic solutions of simple continuous equations, with simple rational numbers as initial values.

Eq. (6) does not possess an attractor with a Cantor-set structure, because it is not dissipative. In fact, the attractor is the entire interval $0 \leqq x \leqq 4$, and it is only because two distinct values of x_n can produce the same value of x_{n+1} that a small one-dimensional sphere, i.e., a segment, is not stretched to infinite length. The single Lyapunov exponent is log 2, i.e., on the average, the length of a small segment doubles with each iteration.

The more general equation

$$x_{n+1} = (x_n - c)^2 \tag{10}$$

possesses aperiodic solutions for some values of c between 1.4 and 2.0, although in most cases it cannot be converted to an equation like eq. (9) by a trigonometric transformation. Identification of the transitions from periodic to chaotic and from chaotic to periodic behavior which occur as c continually increases constitutes an interesting problem in dynamical-systems theory.

For examples of chaos more closely related to the atmosphere we turn first to a system of 12 ordinary differential equations which we introduced about 25 years ago for the specific purpose of obtaining a meteorological model with an aperiodic general solution [8]. As we have noted, even a million numbers governed by a million equations would give a somewhat incomplete picture of the atmosphere, so a 12-variable model must be very crude indeed. The model was obtained by representing the horizontal wind components by a stream function, expanding the fields of the vertically averaged stream function and the vertically averaged temperature in orthogonal functions, and then truncating each series to six terms. Vertical variations of the wind were identified with horizontal temperature gradients through the geostrophic relation. Other variables were inferred implicitly or disregarded altogether. The equations assumed the general form

$$dx_i/dt = \Sigma\, a_{ijk}x_jx_k + \Sigma\, b_{ij}x_j + c_i \ , \tag{11}$$

with x_i representing a stream function if $0 < i \leq 6$ and a temperature if $6 < i \leq 12$. In each equation at most four coefficients a_{ijk} and at most two coefficients b_{ij} differed from zero.

We integrated the equations numerically, using six-hour time steps, for a total of about 20 years. Fig. 1 shows the variations of x_1, representing the strength of the globally averaged westerly wind current, during a typical 18-month interval. Although no true periodicity is apparent, we see a succession of episodes, each lasting a month or somewhat longer, and each bearing a fair resemblance to the others. Each episode is marked by a rapid rise from very weak to very strong westerlies, followed by a somewhat less rapid fall to weak westerlies. The episode is completed by oscillations with periods of a week or two, generally about low values of x_1, but occasionally, as from days 140 to 170 and 380 to 410, about rather high values. The appearance of pronounced regularities, which, however, fall short of exact repetitions, is a typical feature of chaos which is deterministically generated. The succession of episodes may be regarded as a model of the atmosphere's index cycle, although the true index cycle is less regular [9].

Some of the properties of the 12-variable model are more easily illustrated by turning to a 3-variable model, which may be derived from

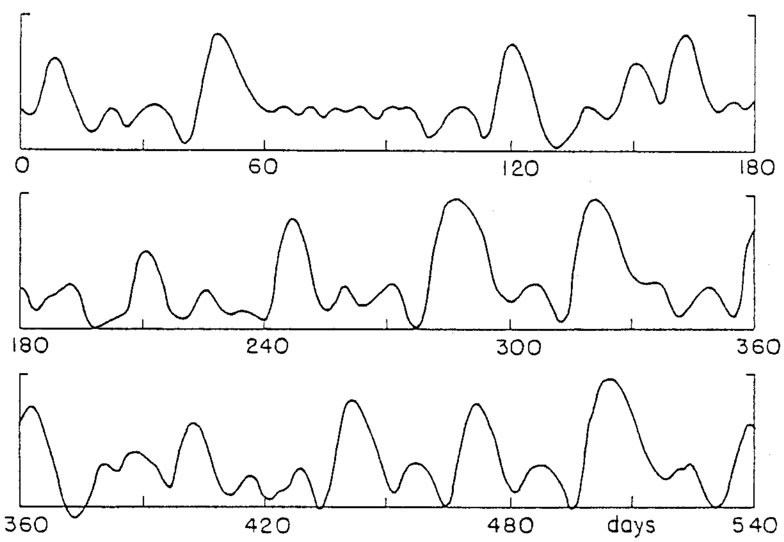

Fig. 1. Variations of the variables x_1 in the 12-variable model governed by eq. (11), representing the strength of the globally averaged westerly wind, during a particular 18-month interval.

the larger model by replacing the 12 variables by 12 linear combinations, and then performing additional truncations [10]. As might be expected, the new model is an even cruder approximation to the real atmosphere than the old one, yet it retains some of the real atmosphere's properties.

The equations of the new model are

$$dx/dt = -y^2 - z^2 - ax + aF , \tag{12}$$

$$dy/dt = xy - bxz - y + G , \tag{13}$$

$$dz/dt = bxy + xz - z . \tag{14}$$

Here x denotes the strength of the globally averaged westerly current, which is identified through the geostrophic relation with the cross-latitude temperature contrast, while y and z denote the cosine and sine phases of a chain of superposed waves, whose troughs and ridges are constrained to tilt westward with increasing elevation. The waves transport heat poleward, thus reducing the temperature contrast, as indicated by the $-y^2$ and $-z^2$ terms in eq. (12). The energy thus removed from the zonal current is added to the waves, as indicated by the xy and xz terms in eqs. (13) and (14). The waves are also carried along by the current, as indicated by the -bxz and bxy terms. The linear terms represent mechanical and thermal damping, while the constant terms

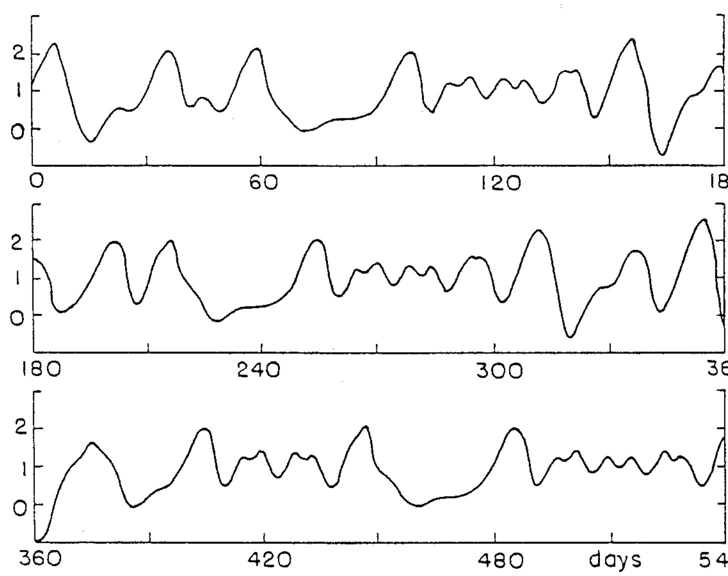

Fig. 2. Variations of the variable x in the 3-variable model governed by eqs. (12)-(14) with a = 1/4, b = 4, F = 8, and G = 5/4, representing the strength of the globally averaged westerly wind, during a particular 18-month interval

represent thermal forcing. The time unit equals the damping time for the waves, which is assumed to be five days.

For suitable choices of a, b, F, and G, eqs. (12)-(14) produce chaos. Such choices include a = 1/4, b = 4, and F = 8, with G some number between 0.85 and 1.3. Fig. 2, which is like Fig. 1, shows the variations of x for a typical 18-month interval, when G = 1.25. Again we see aperiodic variations, but with certain preferred types of behavior. The oscillations may again be regarded as modeling the atmospheric index cycle, but a six-month sequence produced by eq. (11) would probably not be mistaken for one produced by eqs. (12)-(14), nor, presumably, would a sequence produced by either model be mistaken for a real atmospheric index-cycle sequence.

A simple measure of the difference between two states is the distance in phase space. Fig. 3 shows the growth of such a difference, during a 12-month interval. The two initial states are the initial state of Fig. 2 and the same state with a small perturbation added. Eventually the difference becomes large, but the significant increases seem to be confined to the phase of the index cycle when the westerlies are approaching a maximum, and there are intervals as long as three months with no growth at all. The Lyapunov exponents prove to be 0.18, 0.00, and -0.52; the first exponent indicates that, on the average, small differences double in about 3.6 time units, or 18 days. We might add that by real atmospheric standards this growth is unreasonably slow.

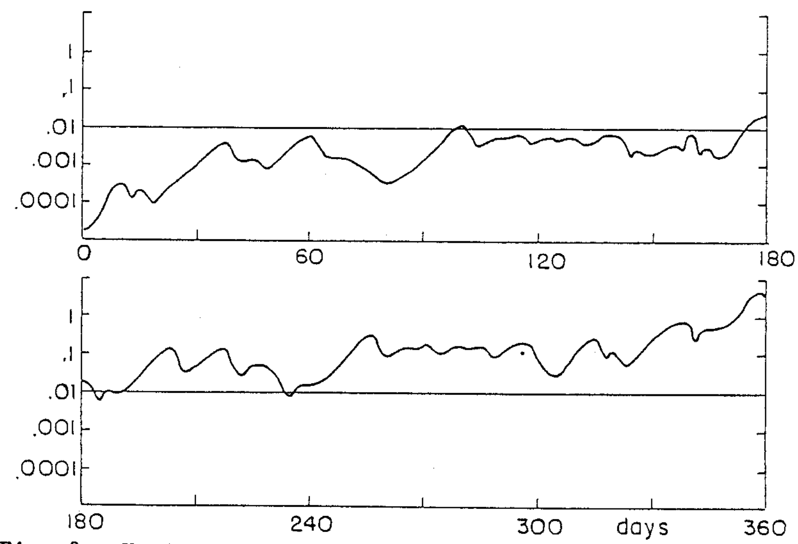

Fig. 3. Variations of the root-mean-square difference between the solution given in Fig. 2 and a second solution obtained by adding 0.00001 to the initial value of each variable, during the first 12 months of the 18 month interval of Fig. 2.

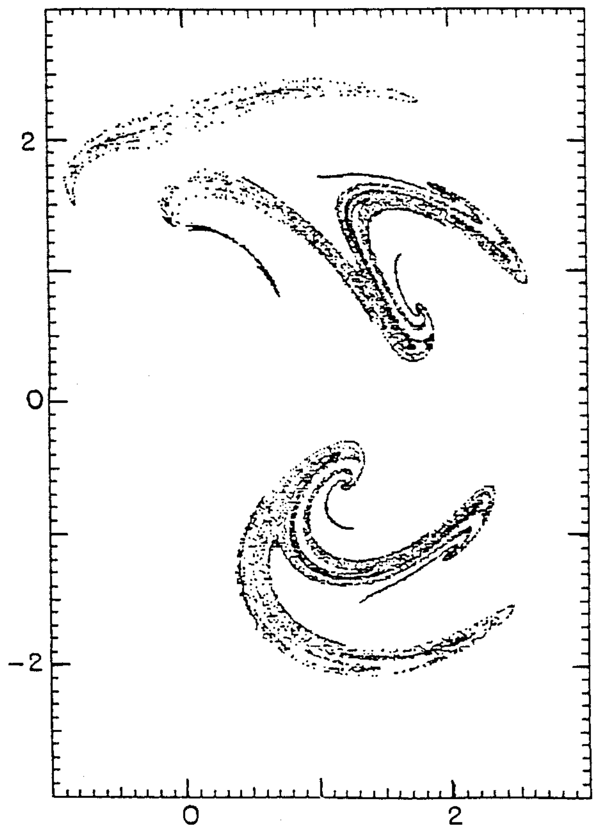

Fig. 4. The intersection of the attractor of eqs. (12)-(14) with the plane $z = 0$, when $a = 1/4$, $b = 4$, $F = 8$, and $G = 5/4$, as represented by 15000 successive intersections of a single orbit with the plane $z = 0$.

One feature of the 3-variable model which is easily examined is its attractor set. Fig. 4 shows the intersection of the single attractor with the plane $z = 0$, as approximated by 15000 successive intersections of a single orbit with the plane; these took place during about 160 years. The points appear to be concentrated on a few dozen curves, and nearby curves are approximately parallel. Between the curves are large areas which are avoided. In particular, the line $y = 0$ is avoided, indicating that states where $y = z = 0$, i.e., where the flow is independent of longitude, are never approached.

Fig. 5 shows an enlargement of a portion of Fig. 4, while Fig. 6 is an enlargement of part of Fig. 5. Additional curves are resolved. Further enlargements, not shown, reveal further curves, and it seems evident that a line cutting across all the curves would intersect them in a Cantor set, i.e., that the attractor is strange.

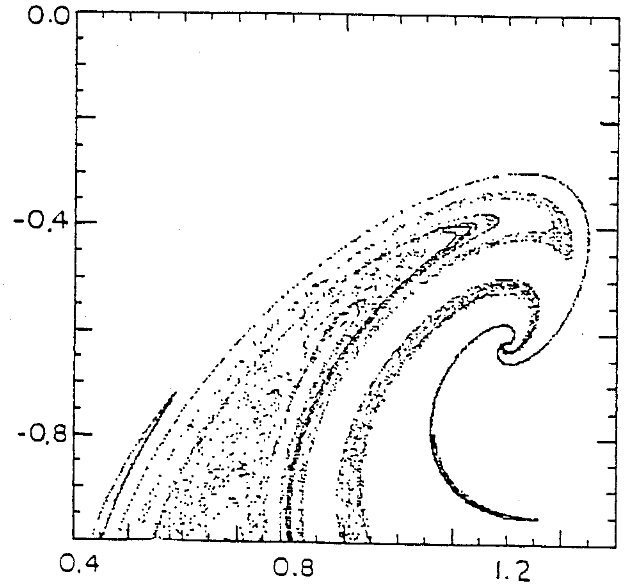

Fig. 5. An enlargement of a portion of Fig. 4, as represented by 8000 successive intersections of a single orbit with the included portion of the plane z = 0.

For our next example of chaos we proceed from one of the smallest possible "global circulation models" to one of the largest yet constructed. This is the operational forecasting model of the European Centre for Medium Range Weather Forecasts (ECMWF). The principal dependent variables of the model are horizontal wind components, temperature, and water-vapor mixing ratio; other variables are determined from these by auxiliary diagnostic formulas. The variables are independently defined at 15 elevations, and, in a recent version of the model, each horizontal field is represented by more than 10000 spherical-harmonic coefficients. The model thus consists effectively of more than 600000 ordinary differential equations in as many variables.

The model contains such physical features as orography. The effects of structures which are unresolved by the model, such as cumulus clouds, are included via parameterization. The intent is to make the model as good an approximation to the real atmosphere as is practical, in view of today's observation and computation systems. Diagnostic studies are regularly performed to determine how closely the climate produced by the model resembles the real atmosphere's climate, and significant differences generally lead to further research aimed at eliminating the discrepancies.

As the name of the Centre might imply, the principal purpose of the model is to produce weather forecasts at the "medium range" extending from a few days to a week or two. The present operational routine

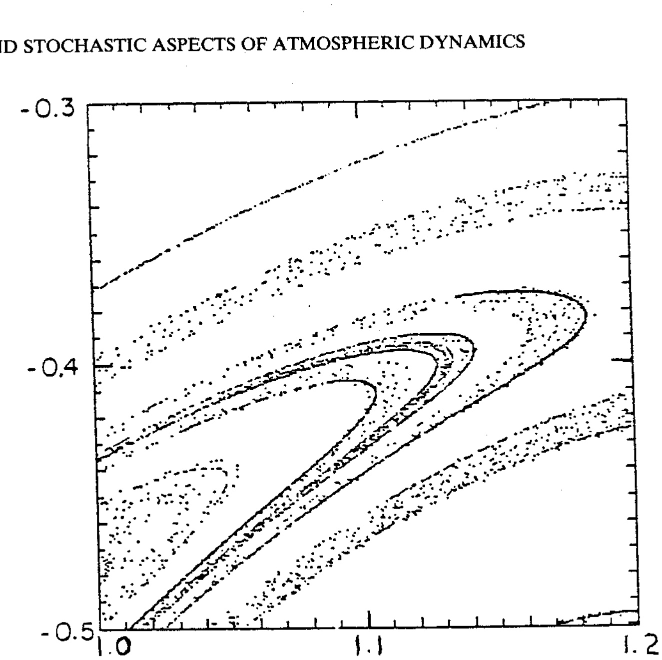

Fig. 6. An enlargement of a portion of Fig. 5, as represented by 6000 successive intersections of a single orbit with the included portion of the plane $z = 0$.

involves preparing, every day, a ten-day forecast of the global atmospheric state, using the present day's state as initial conditions. Since the equations are solved by stepwise integration, forecasts for intermediate ranges are automatically produced, and one-day, two-day, . . ., ten-day forecasts are routinely archived and made available for further research. However, forecasts more than ten days in advance are not generally prepared, and anything like an 18-month time series, comparable to Fig. 1 or 2, is unavailable.

Since the climate of the model differs from that of the real atmosphere, initial states determined from the real atmosphere need not lie on the model's attractor, and, since transient effects may well take more than ten days to die out, not even one point on the model's attractor set is known, let alone an entire attractor. That the model behaves chaotically rather than periodically is best determined by examining it for sensitive dependence on initial conditions.

We have performed a detailed examination of this sort. It would have been computationally expensive to perform many additional runs, in which the operationally used initial states were slightly modified. Instead we have capitalized on the fact that the model produces rather good one-day forecasts, so that the state predicted for a given day, one day in advance, may be regarded as equal to the state subsequently observed on the given day, plus a relatively small error. By comparing the one-day and two-forecasts for the following day, the two-day and three-day forecasts for the day after that, etc., we can determine how rapidly the error grows. Moreover, there are no practical barriers to averaging the results over a large sample of forecasts.

Fig. 7 presents the principal results. Points labeled i,j, where i and j are integers, indicate the globally averaged root-mean-square temperature difference at the 500-millibar level between i-day and j-day forecasts for the same day, averaged over 100 consecutive days beginning 1 December 1984. A 0-day forecast is simply an initial analysis.

The upper curve, connecting points labeled 0,j, for different values of j, therefore measures the model's performance, and indicates how rapidly the difference between two states, one governed by the model and one by the real atmospheric equations, will amplify. The lower curve, connecting points labeled i,j, with j - i = 1, indicates how rapidly the difference between two states, both governed by the model, will amplify.

The lower curve clearly indicates sensitive dependence on initial conditions. Extrapolation of the curve to very small differences suggests a doubling time of about 2.5 days. Detailed forecasting of weather states at sufficiently long range is therefore impractical. However, the difference between the slopes of the two curves indicates

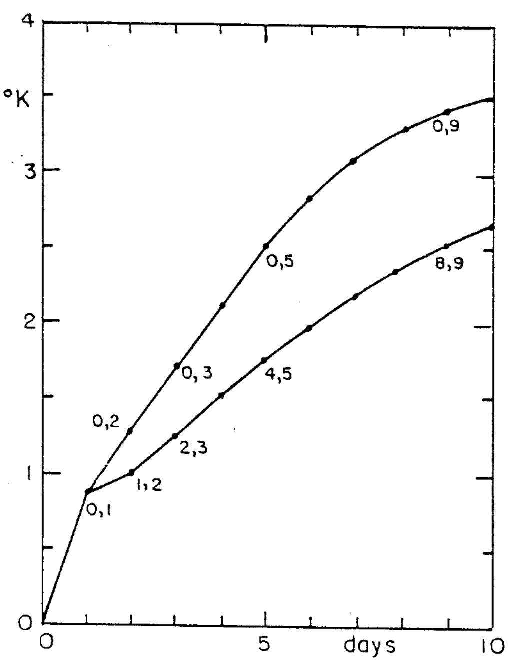

Fig. 7. Root-mean-square differences between i-day and j-day forecasts of the 500-millibar temperature for the same day, made by the ECMWF operational model, averaged over 100 days beginning 1 December 1984. Numbers i, j appear beside selected difference values, which are plotted against values of j.

that there is still considerable room for improvement in forecasting, and implies that we may, for example, some day produce one-week forecasts as good as today's three-day forecasts. Fig. 7 closely resembles a figure constructed from an earlier version of the ECMWF model [11], and both studies tend to confirm the results of earlier studies performed with less elaborate models [12].

Our final example of chaos is the weather itself. In contrast to the case of large atmospheric models, our evidence for chaotic behavior is mainly the absence of any tendency for exact repetitions, and the accompanying presence of continua in the many available variance spectra. We cannot perturb the atmosphere and observe what happens, and at the same time know what would have happened if we had not introduced the perturbation. In principle we could wait for an atmospheric state which closely resembles a previous state, and regard the new state as equal to the old state plus a small perturbation, but in practice we would have to wait too long. We recently estimated that we would have to wait 140 years to obtain one pair of states with a difference of one half of the difference between randomly chosen states [13].

Frequently we observe atmospheric states which closely resemble one another over limited regions; for example, two extratropical cyclones may look very much alike. After a few days the local resemblance will be much weaker, but it is not certain whether this is so because of local amplification or because of the influence of more distant regions where the states are quite different.

Probably our confidence in the chaotic nature of the atmosphere is fortified by the fact that the various large global models exhibit behavior resembling that of the real atmosphere fairly closely, and all of these models show sensitive dependence on initial conditions and agree fairly well as to the rate of error growth. We may also be influenced by our familiarity with baroclinic instability, where perturbed states will depart from unperturbed states.

6. CONCLUSIONS

We may now return to our question as to whether, in investigating atmospheric dynamics, we ought to treat the atmosphere as a deterministic or a chaotic system. The possibly surprising answer is that for most investigations it does not matter. The system of equations which we will be using to study the atmosphere will necessarily involve some approximations, and it may be regarded as a model. Provided that the model is realistic enough to produce a chaotic atmosphere with essentially correct gross features, its behavior will be about the same whether or not it contains some stochastic terms. Here we are assuming that the magnitude of these terms is not completely out of proportion with the actual randomness present in the laws governing the atmosphere.

Our choice between a formally deterministic and a stochastic model will therefore be one of convenience. If our reasoning can be facilitated by the knowledge that our equations contain no randomness, we should use a deterministic formulation. If explicit randomness will aid our investigation, we should introduce it.

As with most general conclusions, there are particular exceptions. If we are studying the growth of the difference between two atmospheric states, using a model in which the smaller scales have been parameterized, and if the initial difference is very small, it will grow quasi-exponentially and require a number of days to become appreciable, if the parameterization is deterministic. With a stochastic parameterization the difference, even if it is initially zero, will quickly become appreciable, possibly during the first day. The latter type of behavior seems more realistic, since it appears that if the small scales could be carried explicitly, uncertainties in these scales would rapidly spread to the larger scales [14], [15]. Once the differences in the resolved scales have become appreciable, it matters little whether the parameterization is deterministic or stochastic.

We are not maintaining that a system of equations with no random terms, and the same system with random terms added, can produce quantitatively identical results. Qualitatively the results may be nearly indistinguishable, or they may be quite different if some of the constants in the system are close to their bifurcation values. In the latter event, the addition of small random terms may still be nearly equivalent to making small alterations in the numerical values of the constants.

ACKNOWLEDGMENT. This work has been supported by the GARP Program of the Atmospheric Sciences Section, National Science Foundation, under Grant 82-14582 ATM.

REFERENCES

1. Wiener, N., 1956: I am a Mathematician. New York, Doubleday, 380 pp (see p. 259).
2. Haltiner, G., 1971: Numerical Weather Prediction. New York, Wiley, 317 pp.
3. Prigogine, I. and Stengers, I., 1983: Order out of Chaos. New York Bantam, 349 pp.
4. Guckenheimer, J. and Holmes, P., 1983: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. New York, Springer-Verlag, 453 pp.
5. Lorenz, E. N., 1963: 'Deterministic nonperiodic flow' J. Atmos. Sci., **20** , 130-141.
6. Ruelle, D., 1981: 'Small random perturbations of dynamical systems and the definition of attractors' Comm. Math. Phys., **82** , 137-151.
7. Ruelle, D. and Takens, F., 1971: 'On the nature of turbulence' Comm. Math. Phys., **20** ,167-192; **23** , 343-344.
8. Lorenz, E. N., 1962: 'The statistical prediction of solutions of dynamic equations' Proc. Internat. Sympos. Numerical Weather Prediction, Meteorol. Soc. Japan, 629-635.

9. Namias, J., 1950: 'The index cycle and its role in the general circulation' J. Meteorol., **7** , 130-139.
10. Lorenz, E. N., 1984: 'Irregularity: a fundamental property of the atmosphere' Tellus, **36A** , 98-110.
11. Lorenz, E. N., 1982: 'Atmospheric predictability experiments with a large numerical model' Tellus, **34** , 505-513.
12. Smagorinsky, J., 1969: 'Problems and promises of deterministic extended range forecasting' Bull. Amer. Meteorol. Soc., **50** , 286-311.
13. Lorenz, E. N., 1969: 'Atmospheric predictability as revealed by naturally occurring analogues' J. Atmos. Sci., **26** , 636-646.
14. Lorenz, E. N., 1969: 'The predictability of a flow which possesses many scales of motion' Tellus, **21** , 289-307.
15. Leith, C. E., 1971: 'Atmospheric predictability and two-dimensional turbulence' J. Atmos. Sci., **28** , 148-161.