A reexamination of methods for evaluating the predictability of the atmosphere

J. L. Anderson, V. Hubeny

Nonlinear Processes in Geophysics (1997) 4: 157–165

Nonlinear Processes in Geophysics

A reexamination of methods for evaluating the predictability of the atmosphere

J. L. Anderson and V. Hubeny

Geophysical Fluid Dynamics Laboratory, P.O. Box 308, Princeton University, Princeton, NJ 08542 USA.

Received: 20 August 1996 – Accepted: 29 October 1997

Abstract. Pioneering work by Lorenz (1965, 1968, 1969) developed a number of methods for exploring the limits of predictability of the atmosphere. One method uses an integration of a realistic numerical model as a surrogate for the atmosphere. The evolution of small perturbations to the integration are used to estimate how quickly errors resulting from a given observational error distribution would grow in this perfect model context.

In reality, an additional constraint must be applied to this predictability problem. In the real atmosphere, only states that belong to the atmosphere's climate occur and one is only interested in how such realizable states diverge in time. Similarly, in a perfect model study, only states on the model's attractor occur. However, a prescribed observational error distribution may project on states that are off the attractor, resulting in unrepresentative error growth. The 'correct' error growth problem examines growth for the projection of the observational error distribution onto the model's attractor.

Simple dynamical systems are used to demonstrate that this additional constraint is vital in order to correctly assess the rate of error growth. A naive approach in which this information about the model's 'climate' is not used can lead to significant errors. Depending on the dynamical system, error doubling times may be either underestimated or overestimated although the latter seems more likely for more realistic models. While the magnitude of these errors is not large in the simple dynamical systems examined, the impact could be much larger in more realistic forecast models.

1 Introduction

Queries about how far in advance the atmosphere could be successfully predicted are a natural outgrowth of the developments in numerical weather prediction that have occurred during recent decades. Some studies have attempted to reach conclusions about fundamental physical limitations which prevent the atmosphere from ever being forecast beyond a certain time into the future (Lorenz, 1968). This study is concerned with a more mundane but perhaps more practical definition of predictability: given some distribution of observational error, how far into the future can this distribution be integrated before it is indistinguishable from a climatological distribution? This definition can be applied to either a numerical model or to the real atmosphere if one thinks of the atmosphere as simply a large, extremely complicated model. With this definition the predictability is a function of the initial observational distribution which is the only source of error. The predictability may also vary depending upon the initial state around which the observational error is placed. One can define the predictability of a model as the average predictability over all possible model states. Since it is a difficult task to determine when the integrated observational distribution and a climatological distribution become indistinguishable, a rough measure of the predictability that is used in the following is the time taken for the mean error of an initial observational distribution to double.

This paper presents a study of the predictability of several simple low order dynamical systems which are used to gain insight into the behavior of more realistic models and, perhaps, the atmosphere. A hypothetical probability distribution is taken to represent the observational error. The observational distribution for a particular point on the attractor, referred to as the control point, is represented by placing the observational error distribution around the control point. Because there are no simple efficient methods for integrating probability distributions in the dynamical systems studied here (Epstein, 1969), the observational distribution at later times is approximated using Monte Carlo techniques. The predictability in this perfect model context can be studied by examining the change in time of the mean distance between the integration of the control point and the integrated observational distribution. In an ergodic system like the ones studied here, the mean error eventually asymptotes to the expected value of the 'error' between two randomly selected points on the model's attractor.

In the scenario just described, the distribution of the

Correspondence to: J. Anderson

observational error alone is used to generate the observational distribution. However, there are additional constraints placed on the error growth because only points on the model's attractor can ever occur (Anderson, 1996). The actual question of interest for predictability is the evolution in time of the mean error between the control point and a projection of the observational distribution onto the model's attractor. In other words, how fast do model states that can actually occur separate from one another as they are integrated in time.

This paper examines the impacts of using this additional information about the model's attractor (climate) on the computed predictability. Failure to project observational error distributions onto the attractor is related to behavior observed in much more realistic atmospheric prediction models. While the differences in computed predictability found in the simple models studied here are relatively small, it is possible that the impacts could be considerably larger for models with higher dimensional phase spaces.

It is convenient to examine an additional question in the study of predictability which can be addressed with similar numerical techniques. There have been attempts (Lorenz, 1969) to determine the predictability of the atmosphere by examining the growth in error between observations subsequent to a pair of observed atmospheric states that are relatively similar by some measure; such states being referred to as 'analogs'. Because the atmosphere exists in a huge phase space, and because the observational record is quite short, even the best analogs found tend to be separated by some significant fraction of the average distance between randomly chosen observed states (Van den Dool, 1994). One can mimic this situation in simple models by examining the error growth between a control point and a set of other points on the attractor selected so that the initial error is relatively large compared to the distance between randomly selected points on the attractor. The error growth and predictability inferred from the evolution of the mean error between these 'poor analog points' and the control point can be compared to the error growth found using 'good analog points'.

Section 2 describes the design of the experiments used to address the first question above. This method is applied to the three variable model of Lorenz (1963) in Sect. 3 and the results are interpreted in the light of information about the model's attractor. The method is applied to several additional low order chaotic models in Sect. 4. Section 5 examines the impact of using poor analogs in computing predictability in the same three dynamical systems and conclusions are presented in Sect. 6.

2 Experimental design

The method used to study error growth is described in terms of a generic dynamical system with an m-dimensional phase space. An initial control point, $A^0 = A(t=0)$, on the model attractor is generated by integrating from arbitrary initial conditions for a very long time. The initial control point is then integrated to produce a time series of control points A^n where the superscript represents the number of timesteps that A^0 has been advanced in the model.

A set of points that randomly sample (Epstein, 1969) the model attractor within a radius r of A^0 in phase space is selected by integrating the model for a very long time and conditionally selecting 1000 points close to A^0 in phase space using the algorithm described in appendix 1. In the results shown here, r is 0.05 times the average distance between randomly selected states on the attractor (a value selected to be roughly consistent with the ratio between observational error and climatological variance in the present day atmospheric observing system). This set of 1000 points, C_i^0 (i = 1,..., 1000), is referred to as the 'correct' initial condition distribution since it contains only points that actually occur in the model's climate and samples them in a fashion consistent with their climatological frequency of occurrence. The 1000 points, C_i^0, are integrated to produce 1000 point samples, C_i^n, of the 'correct' distribution after n timesteps in the model.

A second set of 1000 points, U_i^0 (i = 1,..., 1000), called the 'unconstrained' initial condition distribution, is generated from the correct initial condition distribution. For each point C_i^0 the corresponding point U_i^0 is the same distance from the initial control point A^0 ($|C_i^0 - A^0| = |U_i^0 - A^0|$ where the vertical bars represent a Euclidean distance in phase space). However, the direction between the unconstrained initial point and A^0 is randomly selected from a uniform distribution spanning the m-dimensional phase space of the dynamical system. This direction can be computed easily using the algorithm outlined in Knuth (1981) in which m independent standard normally distributed random numbers are used to define the m coordinates of the random vector. Each member of the unconstrained initial condition distribution is integrated in the model to produce a sample of the unconstrained distribution after n timesteps, U_i^n. Figure 1 is an idealized depiction of the process of selecting the correct and unconstrained initial condition distributions for a model with an two-dimensional attractor embedded in a 3-dimensional phase space.

The mean error between the correct distribution and the control point as a function of timestep is $EC^n = \sum_{i=1}^{1000} L(C_i^n - A^n)$ where L represents a norm measuring the error; a similar expression defines the error of the unconstrained distribution, EU^n. The mean square error in phase space is selected as the norm in order to be compatible with classical works on error growth (Lorenz, 1965) and because of the ease of deriving analytical results such as the

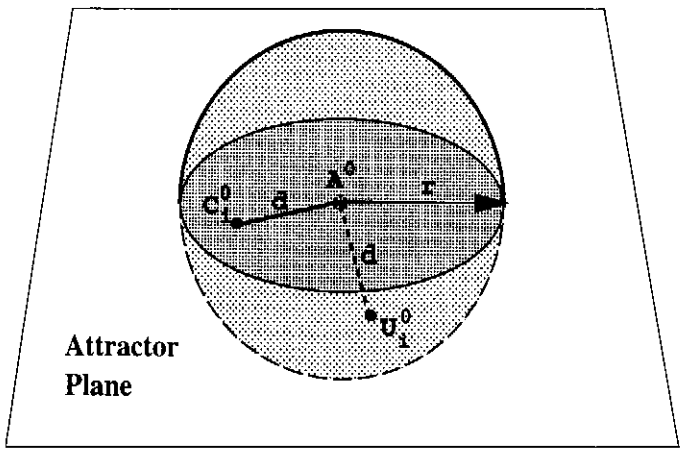

Fig. 1. Idealized depiction of the correct and unconstrained initial condition distributions for an initial control point A^0 on an approximately two-dimensional attractor embedded in a three-dimensional phase space. The correct distribution is represented by the circle of radius r on the attractor while the unconstrained distribution is represented by the sphere. The sample unconstrained point U_i^0 is the same distance, d, from A^0 as is C_i^0 but the direction between U_i^0 and A^0 is randomly selected.

one in Appendix 2 using this norm. The use of other norms could have significant quantitative, and possibly qualitative, impacts on the results. All experiments described here were repeated with the root mean square error with no qualitative impact on the results. Results are plotted as the ratio of the mean square error to the mean square error from the initial condition distribution (EC^n / EC^0 and EU^n / EU^0) in order to examine error growth.

Although the error growth for a single control point is of interest, it may be more revealing to examine the average error growth for a large number of independent control points in order to make statements about the predictability of the dynamical system as a whole. For the results discussed below, the error growth experiments described above were repeated for 500 independent initial control points randomly selected on the model attractor by changing the length of the long integration used to find A^0 in each case. The results displayed here are the 500 case average of the error growth. Considerable variability in the error growth rates was found for the individual control points, however, the vast majority of the points had behavior qualitatively similar to that found in the mean.

For each dynamical system considered, the entire experiment was repeated for a number of different randomly selected initial points to start the long integration used to define the model attractor. In all cases, the selection of the initial point had no qualitative impact on the results. This increases confidence that the results presented are representative of the behavior of the model attractors.

3 Lorenz-63 Model

The first dynamical system examined with the methods described in Sect. 2 is the 3 variable convective model of Lorenz (1963), referred to here as the Lorenz-63 model, which has become one of the mainstays for the study of chaotic systems (Palmer, 1993; Pasmanter, 1995). The model's 3 equations are:

$$\dot{x} = -\sigma x + \sigma y \quad (1)$$
$$\dot{y} = -xz + rx - y \quad (2)$$
$$\dot{z} = xy - bz \quad (3)$$

where the dot represents a derivative with respect to time. The model is integrated using the standard values of 10, 28 and 8/3 for the parameters σ, b and r respectively and the time step described in the original Lorenz paper resulting in a system with chaotic dynamics.

Figure 2 shows the mean error growth for the correct and unconstrained distributions for the Lorenz-63 model. During the early stages of the integrations (Fig. 2a) the correct distribution shows an approximately exponential error growth with instantaneous growth very nearly 0 at the initial time (Thompson, 1984). The error doubles after about 0.13 time units for the correct distribution. The unconstrained distribution shows an initial decrease in error to about 0.8 times the initial error at time 0.04, followed by an approximately exponential growth in error at later times. The unconstrained error doubles after about 0.165 time units, about 25 percent later than for the correct case. For longer integrations times (Fig. 2b), both correct and unconstrained distributions show continued error growth at an approximately exponential rate for times out to about 2 units. This exponential growth is modulated by a number of local 'ripples' which are related to times at which portions of the distributions get 'bifurcated' (in the 500 case mean) by being stretched onto different parts of the two lobed Lorenz-63 attractor (Anderson, 1996). These ripples are a fundamental part of the behavior and would not be smoothed out by increasing the sample size. Eventually, the error growth for both distributions asymptotes to the 'climatological' value for this model.

If one attempts to place a value on the predictability of the Lorenz-63 system by examining the error doubling time, significant differences are found between the results for the unconstrained and correct distributions. The unconstrained distribution appears to be predictable for a longer time, primarily because of the initial error decrease before the initiation of exponential growth. The existence of a mean decrease in error would be inconsistent with the fact that the Lorenz-63 model is in a chaotic regime for the parameter range chosen if all the perturbed points were themselves on the attractor.

The difference between the unconstrained and correct distribution error growth at early integration times demonstrates that failing to utilize constraints from the dynamical system can lead to fundamentally erroneous conclusions

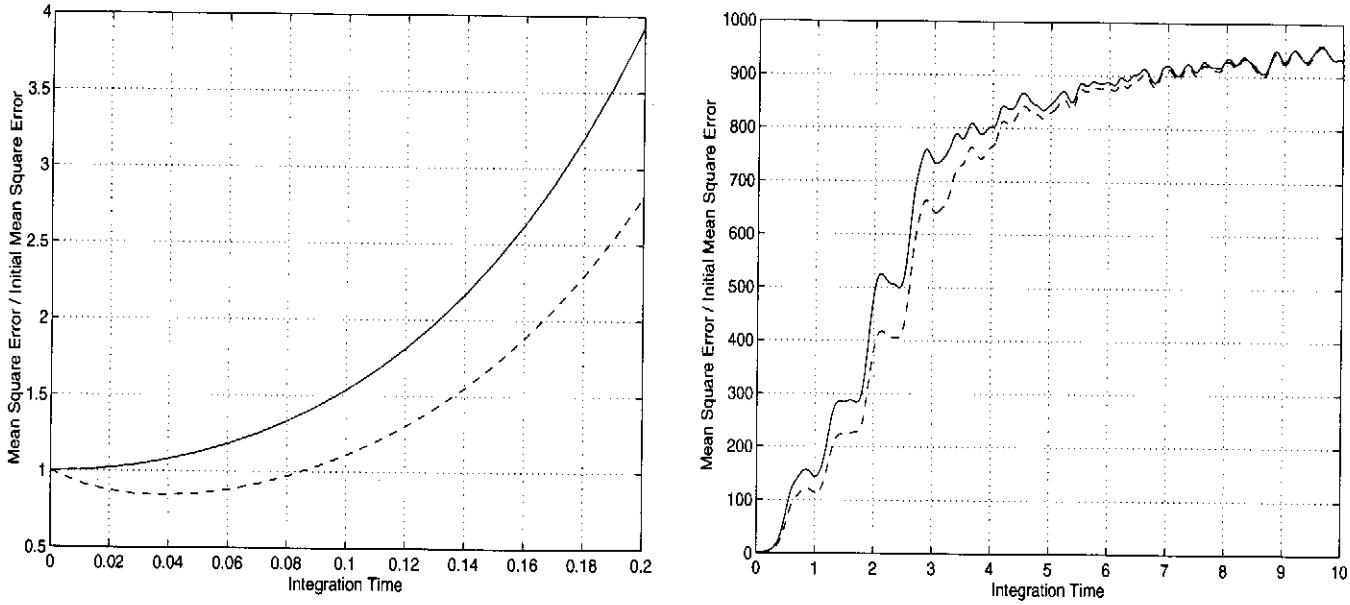

Fig. 2. Mean square error growth in Lorenz-63 model for correct distribution (solid) and unconstrained distribution (dashed). Fig. 2a is an increased resolution view of the early times of Fig. 2b.

about the predictability. In the case of the Lorenz-63 model, the attractor is simple enough that one can offer a qualitative explanation of the differences in error growth. As noted in Anderson (1996) and a host of other publications in many fields, the attractor of the Lorenz-63 model for the standard parameters is locally a nearly 2-dimensional flat structure embedded in the 3-dimensional model phase space. For the initial condition distribution radius used here, the correct initial condition distribution is approximately a circle, or a pair of nearly parallel circles with the initial control point approximately at the center of the single circle in the former case and at the center of the larger of the two circles in the latter case (slightly more complicated loci can occur near the edge of one of the attractor lobes or near areas where the two lobes bifurcate, but this has no impact on the qualitative mechanism being described). As the correct distribution is integrated, the circles are initially stretched along the attractor becoming ellipses, and eventually become long distorted filaments that are found on both of the nearly flat lobes of the attractor.

The unconstrained initial distribution is spherical in this 3-dimensional phase space (see Fig. 1 for an idealized depiction of the correct and unconstrained distributions). As the unconstrained distribution is integrated in time, points that are not on the attractor are rapidly pulled onto one of the foliation of sheets in the closest lobe of the attractor. In general, the evolution of the unconstrained distribution in directions parallel to the nearby attractor lobe is quite similar to the evolution of the correct distribution. However, as the 3-dimensional unconstrained distribution is compressed to 2-dimensions during the early stages of the integration, the mean square error is reduced compared to the value for the correct case. If the growth of the mean square error for the correct case (growth parallel to the attractor) is relatively slow compared to the rate at which points off the attractor are pulled to the attractor, the initial evolution of the unconstrained distribution will result in a reduction of the mean square error. This happens for most individual control points in the Lorenz-63 model, and dominates the behavior in the mean seen in Fig. 2a.

An analytic result derived in Appendix B quantifies the maximum decrease in the mean square error that could be expected from the mechanism described above. If the unconstrained distribution is in an n-dimensional phase space, and the model attractor has a structure that is locally approximately m-dimensional, then the mean square error can be reduced by a factor m/n by the collapse of the unconstrained distribution to the attractor. For the Lorenz-63 model this limiting error reduction would lead to a mean square error of 2/3 times the initial mean square error. Reductions of nearly this magnitude are seen for a number of individual control points.

Behavior similar to that seen for the unconstrained distribution has been seen in more realistic numerical models (Vukicevic and Errico, 1990; Vannitsem and Nicolis, 1994). A naive interpretation might lead to conclusions that predictability is actually increasing with lead time in such a system. However, for the Lorenz-63 system shown here, the correct distribution always shows error growth in the mean [although Smith(1995) has established that for infinitesimal perturbations there are regions of the Lorenz-63 attractor for which all perturbations decrease with time]. The fact that the unconstrained distributions contain many points that cannot occur in the 'climate' of the model leads to an apparent increased predictability in time; in fact, this is just a loss of uncertainty that has been inappropriately included in the

initial condition distributions. The initial reduction of mean error for unconstrained distributions could become more significant for more realistic models if the ratio of the phase space dimension to the approximate dimension of the attractor were to increase. However, more complicated attractors which do not display integral dimension locally might lead to different behavior.

4 Other dynamical systems

It is not clear how the results for the Lorenz-63 model will generalize to models with more complicated attractors or higher dimensional phase spaces. These questions are investigated here, first by studying another 3-dimensional model with a more complicated attractor, then by studying a nine variable model with an attractor similar to that of the Lorenz-63 model.

4.1 Lorenz-84 model

The Lorenz-84 model (Lorenz, 1984) can be regarded as a highly truncated representation of the large scale atmospheric circulation. Although this model has not received as much attention as the Lorenz-63 model, its more complicated attractor structure has led to a recent surge of interest (Lorenz, 1990; Leonardo, 1995). The model can be represented by the equations:

$$\dot{x} = -y^2 - z^2 - ax + aF \tag{4}$$

$$\dot{y} = xy - bxz - y + G \tag{5}$$

$$\dot{z} = bxy + xz - z \tag{6}$$

Values for the parameters a, b, F, and G are 0.25, 4.0, 8.0 and 1.25 respectively as in Lorenz's original description.

Figure 3 shows the initial evolution of the mean square error in the Lorenz-84 model for the correct and unconstrained distributions. Both distributions demonstrate similar exponential growth starting at the initial time. The unconstrained distribution shows slightly greater error growth with a mean error doubling at time 0.31 compared to 0.33 for the correct distribution. In this case, one might slightly underestimate the predictability of the dynamical system by naively examining the unconstrained distribution evolution.

The lack of initial error decay in the Lorenz-84 model is consistent with the attractor of this model which has a local dimension that is considerably greater than 2, for instance, the Kaplan-Yorke dimension is approximately 2.41 (Leonardo, 1995). The attractor is not of significantly lower dimension than the entire phase space, so there is no opportunity for reduction of error through the mechanism described for the Lorenz-63 model. The slightly enhanced error growth for the unconstrained distribution is somewhat more difficult to explain. Apparently, points on the fully 3-dimensional unconstrained distribution are somewhat more likely to end up being initially attracted to portions of the attractor that rapidly diverge from the control point than are the equitably sampled points in the correct distribution.

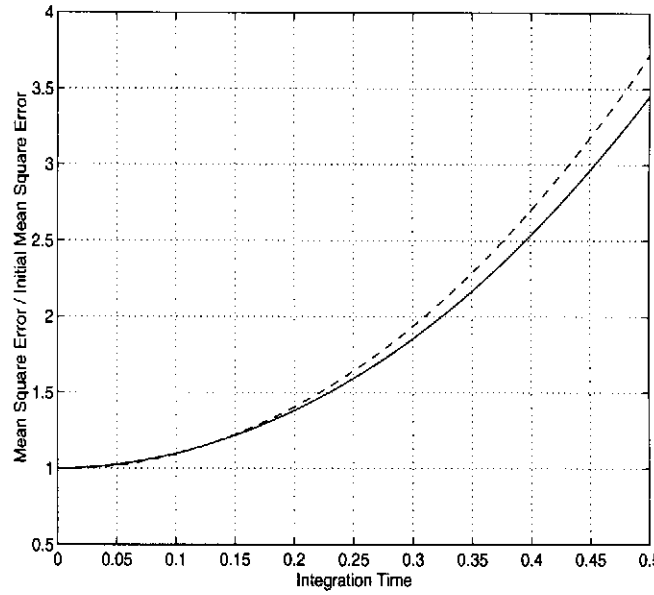

Fig. 3. Mean square error growth in Lorenz-84 model for correct distribution (solid) and unconstrained distribution (dashed).

It is not clear how the Lorenz-84 results might scale to dynamical systems with more degrees of freedom because its attractor is nearly the same dimension as the embedding phase space. The occurrence of initial error reduction in realistic models (Sect. 3) hints that the Lorenz-63 case may be more relevant than the Lorenz-94 case, although great care must be taken in extending any of the results here to more realistic systems. Whatever the case, caution must be exercised when attempting to bound model predictability without using information about the model's attractor.

4.2 9-variable model

A simple dynamical system with a slightly higher dimensional phase space is the 9-variable model of Lorenz (1980). This model is a truncated version of the primitive equations which has been used to study the behavior of gravity waves (Lorenz and Krishnamurthy, 1987). The 9-variable model is represented by the equations:

$$\dot{X}_i = U_j U_k + V_j V_k - v_0 a_i X_i + Y_i + a_i z_i \tag{7}$$

$$\dot{Y}_i = U_j Y_k + Y_j V_k - X_i - v_0 a_i Y_i \tag{8}$$

$$\dot{z}_i = U_j(z_k - h_k) + (z_j - h_j)V_k - g_0 X_i - K_0 a_i z_i + F_i \tag{9}$$

$$U_i = -b_j x_i + c y_i \tag{10}$$

$$V_i = -b_k x_i - c y_i \tag{11}$$

$$X_i = -a_i x_i \tag{12}$$

$$Y_i = -a_i y_i \tag{13}$$

where each equation is defined for cyclic permutations of the indices (i, j, k) over the values (1, 2, 3). The X, Y and z variables can be thought of as representing divergence, vorticity and height respectively while the subscripts can be viewed as representing a zonal mean plus two wave components for each of the three fields. All the parameters in Eqs. (7) through (13) are selected as in Lorenz (1980) in order to produce a chaotic system.

The mean square error evolution for the correct and unconstrained distributions in the 9-variable model are displayed in Fig. 4. The total mean square error and the mean square error of only the 3 height variables (referred to as the height error) are plotted as separate curves for each of the distributions. For the correct distribution, the difference between the total and height error curves is barely distinguishable at the plotting resolution. This is because the correct distribution consists of states on the model's attractor which are a subset of the model's 'slow manifold' (Vautard and Legras, 1986); the height and vorticity fields are nearly in balance (Warn and Menard, 1986) with only a very small divergence (by the classical definition, this is not a slow manifold since there is not an exact balance).

For the unconstrained distribution, the height error evolution is quite different from that of the total error. Because they are not confined to the attractor, most of the states in the unconstrained distribution are not nearly in balance. The result is that large gravity waves are generated during the early phases of the integration; these waves have significant projections on the height variables but huge projections on the divergence variables.

For the mean height error, these gravity waves appear as vacillations with period of about 2 time units in Fig. 4. For the total mean square error, there are much larger amplitude vacillations leading to a very large short term error growth that masks any other effects. It is quite apparent that these short term gravity waves result from an imbalance and are not part of the model's climate (attractor). It is also obvious that the error doubling time of about 0.1 time units is not representative of the true predictability of the system. This paper argues that more subtle balances are also part of the climate and can lead to similar errors in evaluating the error doubling time and predictability.

The height mean square error curve for the unconstrained distribution shows an initial decay in the mean square error with lowest mean values of about 0.7 at 1 time unit. This is a reflection of the same collapse to the attractor as was seen in the Lorenz-63 model. The attractor of the 9-variable model is quite similar to that for the Lorenz-63 model (Moritz and Sutera, 1981). It is nearly flat locally and globally consists of a pair of nearly flat lobes that intersect at one edge. In this case, however, the attractor is embedded in a 9-dimensional

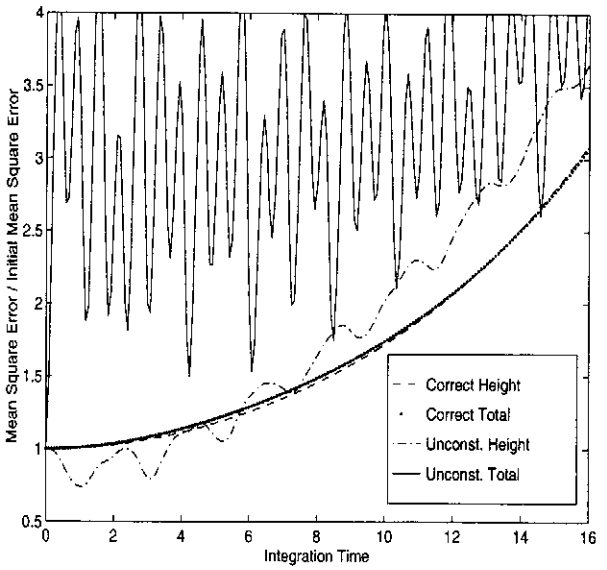

Fig. 4. Mean square error growth in 9-variable model evaluated using the total 9-variable mean square error and using only the three height variables.

phase space. The result of appendix B gives a lower bound of 2/9 for the mean square error ratio in this case, but this bound is not approached closely for any individual control point. The very rapid error growth associated with the appearance of gravity waves masks the collapse of the 9-dimensional distribution to the 2 dimensional attractor so that this limit is not approached as closely as in the Lorenz-63 case. The gravity wave growth dominates relatively quickly even in the height only error so that the unconstrained mean error doubles after approximately 10 time units, 2 units before the correct distribution. In this case, one would underestimate the predictability of the system by using the naive unconstrained distribution.

One could ask if a simple 'initialization' procedure, designed to remove gravity waves from the states in the unconstrained distribution, would remove the differences between the error growth in the unconstrained and correct distributions. Examination of the local Lyapunov vector structure of the 9-variable model reveals that the attractor is an approximately 2-dimensional sub-manifold of a 3-dimensional slow manifold on which gravity wave amplitude is small. If an initialization is performed on the unconstrained distribution, the resulting 'initialized' distribution is an approximately 3-dimensional structure in the 9-dimensional phase space while the attractor is nearly 2-dimensional; other types of 'balance' that are not associated with gravity waves are enforced on the attractor. When an initialization is done, the error growth for the initialized distribution is qualitatively indistinguishable from the growth of the unconstrained distribution for the Lorenz-63 model.

5 Predictability from poor analogs

Another method for evaluating the predictability of the atmosphere (Lorenz, 1969) makes use of pairs of similar observed states referred to as analogs. The growth of error between the observations that followed the analog pair is used to evaluate predictability. Since both of the analog observed states are in the atmosphere's climate (attractor), this method is analogous to the correct distribution method for evaluating model predictability outlined in previous sections. The fact that the observations are contaminated by observational error does not have the same relevance as it did in the discussion of previous sections since the state being 'integrated' by the atmosphere does not contain any error. Unfortunately, only very poor analogs exist in the available observational record (Lorenz, 1969; van den Dool, 1994) with the best analog pairs being separated by distances that are a significant fraction of the climatological variance.

The method for selecting the correct initial condition distribution outlined in Sect. 2 can be readily adapted to examine the impact of using poor analogs to evaluate predictability. The difference is that for this 'poor analog' case, only points with distances from the control point bounded below by 0.62 and above by 0.95 times the average separation between two randomly selected points on the model attractor are chosen. These bounds are somewhat arbitrarily selected in an attempt to mimic the distribution of analog distances found by Lorenz (1969).

The error evolution of the poor analog case can then be compared to that for the correct distribution, also referred to as the 'good analog' distribution. Figure 5 shows the evolution of the mean square error for the Lorenz-63 model. At very early stages of the integration, the error growth curves for the poor and good analog cases are nearly indistinguishable; this corresponds to the linear growth regime mentioned by Lorenz(1969) and many subsequent studies. After time 0.1, the poor analog distribution shows slower error growth; by time 0.2, the concavity of the poor analog error growth curve has changed as the effects of the limits on growth imposed by the climatological variance of the attractor begin to take effect. These results suggest that in the Lorenz-63 model, using poor analogs to investigate the effects of error growth may produce nearly correct results for very short integration periods. If longer periods are examined, for instance the time to error doubling, the poor analog case tends to overestimate the predictability considerably.

Results for the poor analog cases in the Lorenz-84 and 9-variable model (not shown) are qualitatively indistinguishable from those for the Lorenz-63. The poor and good analog distributions have similar error growth at the earliest phases of the integration, but the error doubling time is somewhat shorter in the poor analog cases. It is not immediately clear how these results would scale to much larger dynamical systems; until that problem can be investigated caution must be exercised in interpreting predictability results derived from the evolution of atmospheric analogs.

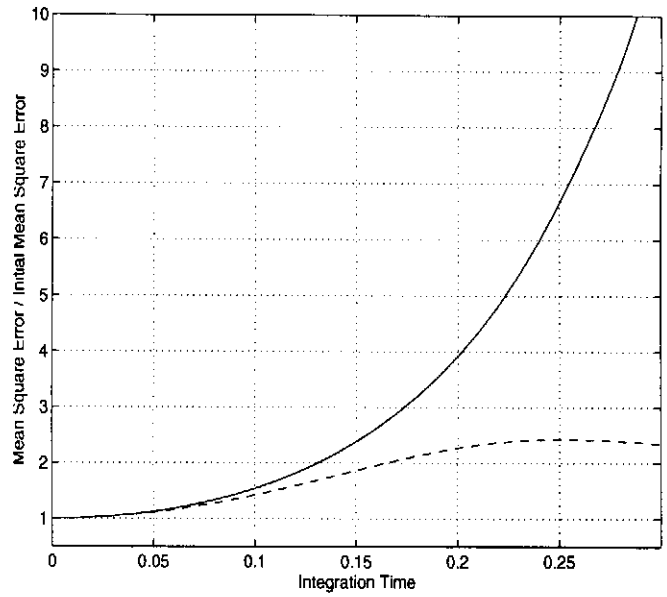

Fig. 5. Mean square error growth in Lorenz-63 model for good analog (correct) distribution (solid) and poor analog distribution (dashed).

6 Summary and Conclusions

Three simple low-order dynamical systems have been used to examine methods for computing the predictability of forecast models. In this perfect model study, the only source of error in a model integration has been assumed to come from observational errors in the integration's initial condition. The predictability is measured by the length of time required for the average error of an observational error distribution to double when integrated in the model.

This method as generally applied fails to take into account that not all points in the phase space of the model can actually occur once the impact of arbitrary initial conditions has become negligible; only points on the model's attractor are possible. However, the observational error distribution may project on states that are not on the attractor, resulting in unrepresentative error growth results. The 'correct' error growth problem examines the error growth for the projection of the observational error distribution onto the model's attractor.

The 'correct' error growth has been compared to the unconstrained error growth that results from ignoring the existence of the model attractor for a trio of simple dynamical systems. In the 3-variable Lorenz-63 model, the unconstrained problem was found to give error doubling times that are longer than those for the correct problem. This behavior was explained in terms of the collapse of the initially 3-dimensional unconstrained error distribution onto the nearly 2-dimensional attractor of the model. This type of behavior could be considerably more significant in models with higher dimensional phase space.

Results for a 9-variable primitive equation model were more complicated but revealed the same type of behavior as

for the Lorenz-63 model. Unconstrained initial states off the attractor led to extremely rapid growth of gravity waves which result in very short error doubling times. Even if initialization is performed to remove the gravity waves, more subtle balances are still not enforced in the unconstrained distribution and the results look very much like those for Lorenz-63.

A third model, the Lorenz-84 three variable model, does not allow the collapse of the unconstrained distribution to the attractor because the attractor does not have a significantly smaller dimension than the whole phase space. In this case the results were quite different with the unconstrained distributions showing faster error growth. It seems likely that the attractors of large forecast models are much smaller than the embedding phase space (Selten, 1993), so this result may not have as much relevance for larger models. However, it points to the difficulties involved in knowing the exact impacts on predictability of using information about the model states that can actually occur.

The errors that arise from failing to use information about the model's attractor in the simple systems examined here suggest that the results from the many unconstrained distribution studies [for instance Farrell (1990) and Houtekamer (1991) which do not directly consider model attractor structure] in higher order models must be regarded as somewhat tentative. In general, it seems likely that the type of errors encountered in large models are more likely to be consistent with those found for the Lorenz-63 type model. Failing to use information about the attractor would lead to an overestimate of the predictability in such systems. It is important to note that in almost all such studies, it would be prohibitively expensive to attempt to determine much information about the model attractor.

Some previous studies with realistic models have attempted to avoid unbalanced perturbed initial conditions by using information from the model's long term climate in the selection of the perturbations (Schubert and Suarez, 1989). In the Lorenz-63 model, such an approach would be more similar to the correct distribution than to the unconstrained distribution because the attractor is globally relatively flat. However, this method would have little impact in the Lorenz-84 or 9-variable models where the attractors curve throughout the phase space. Again, it is not immediately clear if this type of approach gives a reasonable estimate of the predictability in realistic models or not. Even if the attractors of large models are of relatively low dimension compared to the model phase space, these attractors are unlikely, to be 'flat' globally.

The great expense of finding reasonable analogs in higher order models precludes the type of comprehensive study that was performed here. However, even a small sample of moderately good analogs might be sufficient to see if states on the attractor separate less rapidly than states selected randomly from the entire phase space. The results of Sect. 5, that suggest that relatively low quality analogs may still give reasonable information about error doubling times, lend some support to the notion that relatively poor (and therefore more easily computed) analogs could be used successfully in such a study.

As pointed out in Anderson (1996), the predictability results here may also have relevance to the ensemble forecasting problem (Houtekamer and Derome, 1995). If the error growth for points off the attractor are truly unrepresentative then an ensemble forecast that samples an observational error distribution without taking into account the constraints from the model attractor may produce incorrect results. Unfortunately, addressing this problem would require the existence of relatively inexpensive algorithms for finding the local structure of the model's attractor. Research on establishing whether such algorithms are possible would be a useful contribution to the development of operational ensemble forecasts. Results for ensemble forecasting experiments are also complicated by non-equilibrium dynamics since the observed point around which the observational error distribution is 'centered' is unlikely to be on the model's attractor.

Appendix A Generating analog distributions

The 'correct' distributions used to evaluate predictability are hypothetical representations of observational error. The correct distributions are chosen to be a random sample (Epstein, 1969) of the attractor structure within a radius, r, of a control point, A^0. To generate this equitable sample, the model is first integrated for a long time starting from the control point. The integration is continued and a record is kept of each temporally contiguous sequence of points that passes through the hypersphere of radius r surrounding A^0; such a sequence is referred to as an analog segment. Only one point on a given analog segment can be chosen as a member of the final correct distribution because points on the same segment are not independent and will not undergo any exponential separation when integrated. However, if one randomly selects a single point from each analog segment, points on short segments are more likely to be selected than those on long segments resulting in an inequitable sample. To avoid this, an analog segment of length n is only sampled with probability n/m, where m is the number of points on the analog segment with the most points.

Appendix B Error ratio for compressed hyperspheres

The question of interest is the ratio of mean square error between an error distribution that is directionally uniformly distributed around a control point in an n-dimensional phase space and the mean square error when that distribution is compressed onto an m-dimensional subspace that also contains the control point. One can compute this ratio by looking at the expected value of the ratio for a single point in the error distribution.

Let $S = \{x_1, x_2, ..., x_n\}$ be a vector randomly selected from

a distribution that is directionally uniformly distributed on an n-dimensional sphere where each of the x_i is the projection of S onto one of a set of orthogonal vectors that span the n-dimensional phase space. S can be generated by independently sampling each of the x_i from a standard normal distribution (Knuth, 1981). Without loss of generality, let the first m basis vectors lie in the m-dimensional subspace and let the remaining n-m basis vectors be perpendicular to this subspace. The expected value of the ratio of mean square error in the m-dimensional subspace to the mean square error in the full n-dimensional space is:

$$E\left[\frac{(x_1^2 + x_2^2 + \ldots + x_m^2)}{(x_1^2 + x_2^2 + \ldots x_n^2)}\right] = \frac{mE[x_1^2]}{nE[x_1^2]} = \frac{m}{n} .$$

This result is independent of the magnitude of S so the ratio applies to arbitrary distributions of error radius in the n-dimensional space.

Similar results can be derived for other measures of error growth. For instance, a much more involved derivation reveals that the ratio of expected values for the root mean square error is:

$$\frac{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{m}{2}\right)}$$

where Γ is the gamma function. This result is related to techniques used to derive confidence bounds for the traditional chi-square test (Knuth, 1981); the ratio can also be expressed in terms of a ratio of beta functions.

Acknowledgments. The first author is grateful to Theo Opsteegh, Jan Barkmeijer, and KNMI for providing the opportunity to begin this study. Thanks to John Lanzante and Jerry Mahlman for comments on early versions of this manuscript, four anonymous reviewers for helping to clarify the discussion, and Lenny Smith for help, encouragement and lots of insight.

References

Anderson, J. L., Selection of initial conditions for ensemble forecasts in a simple perfect model framework. *J. Atmos. Sci., 53*, 22-36, 1996.

Epstein, E. S., Stochastic dynamic prediction. *Tellus, 21*, 739-759, 1969.

Farrell, B., Small error dynamics and the predictability of atmospheric flows. *J. Atmos. Sci., 47*, 2191-2199, 1990.

Houtekamer, P. L., Variation of the predictability in a low-order spectral model of the atmospheric circulation. *Tellus, 43A*, 177-190. 1991.

Houtekamer, P. L. and Derome, J., Methods for ensemble prediction. *Mon. Wea. Rev., 123*, 2181-2196, 1995.

Knuth, D. E., *Seminumerical Algorithms*, Addison Wesley, 688 pp, 1981.

Leonardo, A., *Numerical studies on the Lorenz-84 atmospheric model.* Department of Mathematics Technical Report, State University of Utrecht, 49 pp., 1995.

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

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

Lorenz, E. N., The predictability of a flow which possesses many scales of motion. *Tellus, 21*, 289-307, 1968.

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

Lorenz, E. N., Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci., 37*, 1685-1699, 1980.

Lorenz, E. N., Irregularity: A fundamental property of the atmosphere. *Tellus, 36A*, 98-110, 1984.

Lorenz, E. N., Can chaos and intransitivity lead to interannual variability? *Tellus, 42A*, 378-389, 1990.

Lorenz, E. N. and Krishnamurthy, V., On the nonexistence of a slow manifold. *J. Atmos. Sci., 44*, 2940-2950, 1987.

Moritz, R. E. and Sutera, A., The predictability problem: Effects of stochastic perturbations in multiequilibrium systems. *Advances in Geophysics, 23*, 345-383, 1981.

Palmer, T. N., Extended-range atmospheric prediction and the Lorenz model. *Bull. Amer. Meteor. Soc., 74*, 49-66, 1993.

Pasmanter, R. A., Searching for periodic motions in long time series. Predictability and nonlinear modeling in natural sciences and economics. *Chaos, Solitons and Fractals, 6*, 447-454, 1995.

Schubert, S. D. and Suarez, M., Dynamical predictability in a simple general circulation model: Average error growth. *J. Atmos. Sci., 46*, 353-370, 1989.

Selten, F. M., Toward an optimal description of atmospheric flow. *J. Atmos. Sci., 50*, 861-877, 1993.

Smith, L. A., Accountability and error in ensemble prediction in baroclinic flows. In Predictability, volume 1 of Seminar Proceedings, pages 351-368. European Centre for Medium-Range Weather Forecasts, September, 1995.

Thompson, P. D., A review of the predictability problem. Predictability of Fluid Motions, American Institute of Physics Conference Proceeding No. 106, 1-10, 1984.

Vukicevic, T. and Errico, R. M., The influence of artificial and physical factors upon predictability estimates using a complex limited-area model. *Mon. Wea. Rev., 118*, 1460-1481, 1990.

Van Den Dool, H. M., Searching for analogues, how long must we wait? *Tellus, 46A*, 314-324, 1994.

Vannitsem, S., and Nicolis, C., Predictability experiments on a simplified thermal convection model: the role of spatial scales. *J. Geophys. Res., 99*, 10377-10385, 1994.

Vautard, R. and Legras, B., Invariant manifolds, quasi-geostrophy and initialization. *J. Atmos. Sci., 43*, 565-584, 1986.

Warn, T. and Menard, R., Nonlinear balance and gravity-inertial wave saturation in a simple atmospheric model. *Tellus, 38A*, 285-294, 1986.