

Convex Error Growth Patterns in a Global Weather Model

John Harlim,^{*} Michael Oczkowski,[†] James A. Yorke,[‡] Eugenia Kalnay,[§] and Brian R. Hunt[¶]
University of Maryland, College Park, Maryland 20742

(Dated: October 8, 2004)

We investigate the error growth, that is, the growth in the distance E between two typical solutions of a weather model. Typically E grows until it reaches a saturation value E_s . We find two distinct broad *log-linear regimes*, one for E below 2% of E_s and the other for E above. In each, $\log(E/E_s)$ grows as if satisfying a linear differential equation. When plotting $d\log(E)/dt$ vs $\log(E)$, the graph is convex. We argue this behavior is quite different from other dynamics problem with saturation values that yield concave graphs.

The main characteristic of a temporal chaotic dynamical system is having a positive leading Lyapunov exponent. This quantity measures the long time average exponential growth rate of two solutions initially separated by infinitesimal distance. One simple example of a chaotic system is the logistic map: $x_{n+1} = 4x_n(1 - x_n)$ with Lyapunov exponent equal to $\log(2)$. This quantity suggests that the difference between two nearby solutions doubles on each iterate until it reaches saturation. From the point of view of weather forecasting, the logistic map has an infinite limit of predictability: we can always extend the forecast for one iterate by halving the uncertainty of the initial conditions. Lorenz [1] suggested that the Lyapunov exponents may be unbounded for the partial differential equation representing a global atmospheric model. He also argued that at the latest stages, the error growth rate behaves differently from Lyapunov exponents.

The atmosphere includes multiple scales of motions, which suggests that different scales of motion grow with different rates. In [2], Lorenz observed that small scales tend to grow at a fast rate and the larger scales of motion grow with a slower rate. He concluded that as one refines the accuracy of the initial states, smaller increments of forecast skill are obtained, and there appears to be a finite limit of predictability. Numerous studies have been devoted to predictability of different types of motion that mimic the atmosphere. For example, Aurell et al. [3] shows that in the case of turbulence the growth rate is determined by the cumulative effects of multiple characteristic times.

A hallmark of chaos is the exponential growth of errors, where by **error** we mean the distance $E(t)$ between two trajectories that are close to each other at time $t = 0$. When trajectories are bounded, the exponential growth of $E(t)$ can not continue indefinitely; $E(t)$ saturates near a value E_s that is representative of the size of the chaotic attractor. We consider the dependence of the exponential growth rate of the error on the size of $E(t)/E_s$ for a global weather forecast model, and we contrast our results with those for some simpler models. The **exponential growth rate** is E'/E where E' is either the time derivative of E or a finite-time approximation of the time derivative. Throughout this Letter, we approximate the

exponential growth rate as follows:

$$\frac{E'}{E} = \frac{d \log E(t)}{dt} \approx \frac{\log E(t + \Delta t) - \log E(t - \Delta t)}{2\Delta t}. \quad (1)$$

This approximation is used to mask the rapid fluctuation of the derivative. In some cases, again to suppress fluctuations, we may take E to be an average distance, averaging over several pairs of solutions.

We call an interval J of values of E a **log-linear regime**, if for some $\lambda < 0$ and $C \leq E_s$ and all $E \in J$, E'/E approximately satisfies the linear differential equation [13]:

$$\frac{E'}{E} = \lambda \log(E/C). \quad (2)$$

The exponent λ describes how E approaches C . A typical dynamical system has one log-linear regime with $C = E_s$; then λ is called the **saturation exponent**. In this Letter, we investigate a realistic weather forecast model and find that this model has two distinct broad log-linear regimes (Figure 1).

Weather model: We are reporting a striking behavior for a global weather model, the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS). This formerly operational global atmospheric model is a pseudo-spectral model, described in detail in [4, 5]. The resolution considered here is chosen such that it has a maximum zonal wavenumber of 62, and so is referred to as T62, where “T” stands for triangular truncation. There are 192×94 horizontal grid points at each of 28 vertical levels. In addition to the surface pressure, there are five variables defined at every grid point: 1) vertical component of the vorticity of the horizontal wind, 2) divergence of the horizontal wind, 3) a generalized temperature that reflects humidity, 4) relative humidity, and 5) ozone. Overall the state has dimension $N \approx 3 \times 10^6$.

We select a point \mathbf{x} in the state space from a trajectory after transients have died away. We also choose at random an N -vector $\delta\mathbf{p}$ and renormalize it so that $\|\delta\mathbf{p}\| = E_s$. Then for $k = 3, \dots, 7$, we choose perturbed initial points

$$\mathbf{p}^k(0) = \mathbf{x}(0) + 10^{-k}\delta\mathbf{p}. \quad (3)$$

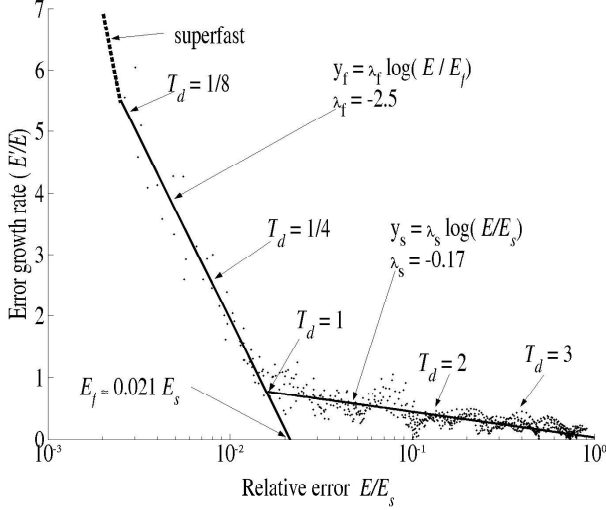


FIG. 1: Convex exponential error growth rate of the NCEP GFS model as a function of relative error E/E_s . The graph reveals two linear regimes. For $E < 0.02E_s$, E'/E lies close to line $y_f = \lambda_f \log(E/E_f)$ with $\lambda_f \approx -2.5$ and intercept at $E_f \approx 2.1\%$ of the saturation level E_s . For $E > .02E_s$, E'/E lies close to the line $y_s = \lambda_s \log(E/E_s)$ with $\lambda_s \approx -0.17$. The doubling time of errors are $1/8$, $1/4$, 1 , 2 , and 3 days when E is 0.23% , 0.7% , 1.6% , 13% and 26% of saturation level, respectively. Small errors grow in amplitude, moving from left to right, first along the left line and then along the right line until they saturate at E_s . Each point plotted here is for an individual pair of initial conditions. The dashes represent the “superfast” regime for $E < 2 \times 10^{-3} E_s$ (see text).

We do not show the cases for $k = 1, 2$ because we are interested in the behavior of perturbations that are initially small. Both the reference state \mathbf{x} and perturbations \mathbf{p}^k are integrated from $t = 0$ for 14 days with a 20 minute time step. Now, define $E^k(t)$ to be the root-mean-square (rms) error of \mathbf{x} and \mathbf{p}^k :

$$E^k(t) = \|\mathbf{p}^k(t) - \mathbf{x}(t)\|. \quad (4)$$

We focus our study on midlatitude tropospheric wind prediction. That is, calculations are restricted to the midlatitude bands in the Northern and Southern Hemispheres ($22.5^\circ - 70^\circ$ N/S) where the model is considered most accurate. We report the rms errors calculated for the atmospheric level where the pressure equals half of the surface pressure. We show only the rms errors of the vorticity, but the other variables behave in a similar way.

In Figure 1, the exponential growth rate E'/E is plotted as a function of relative error size E/E_s in logarithmic coordinates. The error growth rate E'/E decays along two broad log-linear regimes (for time $t > 18$ hours). In what we call the “fast regime” ($0.002E_s < E < 0.02E_s$), the errors double in less than a day. Here the error growth rates move along a straight line (2) with $\lambda = -2.5$ ($= \lambda_f$ in Figure 1) as if to saturate at $C = E_f \approx 0.021E_s$, that is, 2.1% of the actual saturation size. After E reaches

about $0.02E_s$, the growth rate enters the “slow regime.” We refer to the point $E = 0.02E_s$ as the **KT boundary** [14]. After E passes the KT boundary, E'/E follows the straight line (2) with $\lambda = -0.17$ ($= \lambda_s$) and $C = E_s$. Here the growth rate E'/E slows to zero as E approaches its saturation level E_s . In creating Figure 1, we approximate E'/E using (1) with $\Delta t = 6$ hours, with t in steps of 3 hours. The finite differences show some oscillation about the lines. We plot E'/E starting from $t = 18$ hours because the NCEP model exhibits a third regime: superfast growth of extremely small perturbations. Errors of size less than $10^{-4}E_s$ climb rapidly to $10^{-4}E_s$, usually in about 1 hour or less, even when beginning with size $10^{-7}E_s$ (not shown here). Thus, the errors grow to approximately $0.002E_s$ after 18 hours independently of $E(0)$ provided $E(0) < 10^{-4}E_s$. Toth and Kalnay [6, 7] investigated weather prediction using a reference state plus multiple perturbations as initial conditions. They report “enormous” growth of errors, more than a factor of 5 per day, when perturbations have amplitude less than $0.001E_s$. They attribute such growth mostly to tropical convection which they say saturates at less than $0.01E_s$. They see slow growth for amplitude between $0.01E_s$ and $0.1E_s$ and attribute this behavior to baroclinic instabilities. They do not discuss the transition between these behaviors.

In contrast with the striking results of Figure 1, simpler models often have a single log-linear regime. We illustrate the typical behavior of simpler models with the “Lorenz-40” model [1] of differential equations and the Quasi-Geostrophic (QG) model of Marshall and Molteni [8] with resolution T21.

Lorenz-40 model: The Lorenz-40 model [1, 9] represents an “atmospheric variable” with values x_j at N -equally spaced points around a circle of constant latitude:

$$5 \frac{dx_j}{dt} = (x_{j+1} - x_{j-2})x_{j-1} - x_j + F \quad (5)$$

where $j = 1, \dots, N$ represent the spatial coordinates (“longitude”). Periodic boundary conditions are imposed by identifying $x_{-1} \equiv x_{N-1}$, $x_0 \equiv x_N$, and $x_{N+1} \equiv x_1$. This model is designed to satisfy three basic properties: it has linear dissipation (the $-x_j$ term) that decreases the total energy defined as $V = \frac{1}{2} \sum_{j=1}^N x_j^2$, an external forcing term F that can increase or decrease the total energy, and a quadratic advection-like term that conserves the total energy (i.e., it does not contribute to $\frac{d}{dt}V$). Following [1, 9], we choose the external forcing to be $F = 8$ and the number of spatial elements to be $N = 40$. The “5” in (5) scales the unit time to correspond to one day in real time. We also use a fourth-order Runge-Kutta scheme for time integration of (5) with time step $\Delta t = 1/4$ days. With these parameters, the solution to (5) has a behavior reminiscent of the midlatitude atmosphere. It has 13 positive Lyapunov exponents, with the leading Lyapunov exponent corresponding to a doubling time of 2.1 days,

and a Kaplan-Yorke dimension of 27.1 [1].

QG model: The QG approximation describes weather-like slow atmospheric motion. Steady state flow in a rotating sphere satisfies the geostrophic balance, i.e., the horizontal pressure gradient balances the Coriolis force. Slow Rossby waves satisfy a quasi-geostrophic balance as represented by the conservation of quasi-geostrophic potential vorticity (QGPV), see Holton [10].

We use the QG model developed by Marshall and Molteni [8] in our study. This global three-level model (200hPa, 500hPa, and 800hPa) can be described as follows:

$$\frac{\partial q_i}{\partial t} + J(\psi_i, q_i) = -D_i(\psi) + S_i, \quad i = 1, 2, 3, \quad (6)$$

where i denotes each layer, and $q = (q_1, q_2, q_3)$ and $\psi = (\psi_1, \psi_2, \psi_3)$ denote the potential vorticities and streamfunctions, respectively, on a spherical planet. The left hand side of (6) is the vertically discretized version of the QGPV equation, and J represents the advection terms. In the right-hand side of (6), D is a linear operator for dissipation, and S is a time-independent source of potential vorticity chosen to realistically simulate the northern hemisphere winter climate. The resolution of this spectral model is T21, which yields 64×32 grid points at each of the three atmospheric levels. Hence the streamfunction ψ has a state space of $N = 6144$ variables and the potential vorticity can be derived from it.

In the NCEP model, we examined the rms difference between a pair of trajectories \mathbf{x} and \mathbf{p}^k . For the Lorenz-40 and QG models, we examine L such pairs of trajectories and average the resulting errors in order to mask the rapid fluctuation of E . The rms of the i -th pair of $\mathbf{x}(t)$ and $\mathbf{p}^k(t)$ at time t , denoted as $E_i^k(t)$, is calculated from (4) for $i = 1, \dots, L$. Then, we let $E^k(t)$ be the geometric mean

$$E^k(t) = \exp(\langle \log E_i^k(t) \rangle), \quad (7)$$

where the average $\langle \cdot \rangle$ is computed over L pairs. Hence the previous $E^k(t)$ in Figure 1 is an average of $L = 1$ pair of trajectories. In contrast, we use $L = 1000$ for the Lorenz-40 model and $L = 100$ for the QG model. Both the reference state \mathbf{x} and perturbations \mathbf{p}^k , with $k = 3, \dots, 6$, are integrated with 6 hour time steps from time $t = 0$ for 60 days for Lorenz-40 model and for 360 days for QG model.

In Figure 2 ((a) for Lorenz-40 and (b) for QG model), the exponential growth rates are plotted as functions of the relative rms difference between each \mathbf{p}^k and \mathbf{x} . In creating these plots, the growth rate is approximated using (1) with $\Delta t = 1/2$ days for the Lorenz-40 and $\Delta t = 10$ days for the QG model. Our experiments with the single time scale Lorenz-40 and the QG models show small averaged errors growing exponentially. Close to saturation (i.e., as $E \rightarrow E_s$), the exponential growth rate slows down as with differential equation (2) with $C = E_s$ and enters

the log-linear regime. Specifically, the log-linear regime of Lorenz-40 model is approximately $0.4E_s < E < E_s$ with saturation exponent $\lambda \approx -0.22$. We obtain a similar range of log-linear behavior with $\lambda \approx -0.035$ for the QG model. This result suggests that for these two models, the error growth is a single homogeneous phenomenon. Next, we offer a scalar linear differential equation that justifies the typical error growth as observed in Lorenz-40 or QG model.

A simple error growth model: In this Letter, we modify the logistic differential equation [15] and model the growth of the error $E(t)$ by a scalar differential equation:

$$E' = \frac{dE}{dt} = aE(1 - E^{-\lambda/a}). \quad (8)$$

Near the steady state $E = 0$, we have $E'/E \approx a$. For $0 < E < 1$, we have $E \rightarrow 1$ as $t \rightarrow \infty$, and for $E \approx 1$, $E'/E \approx \lambda \log(E)$. With this variant of the logistic equation, two parameters a and λ can be selected independently. Figure 3 shows the exponential growth rate of

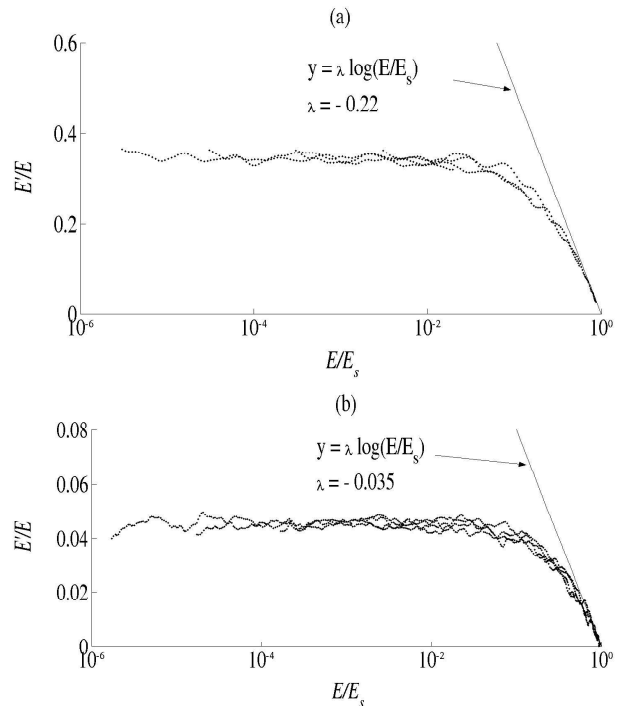


FIG. 2: Concave error growth rates, estimated by finite-differences in (1), as functions of relative error E/E_s . (a) The error growth rate of the Lorenz-40 model fluctuates about 0.35 for $E < 10^{-3}E_s$ and decays in accordance to (2) where $\lambda = -0.22$ (see the line tangent to the growth rate) and $C = E_s$. (b) In the QG model, the growth rate fluctuates about 0.044 when $E < 10^{-3}E_s$ and decays asymptotically to (2) where $C = E_s$, with $\lambda = -0.035$ as $E > 0.4E_s$. As time increases, errors move from left to right, asymptoting at E_s . Each point plotted here is averaged over $L = 1000$ pair of trajectories for the Lorenz-40 model and $L = 100$ for the QG model.

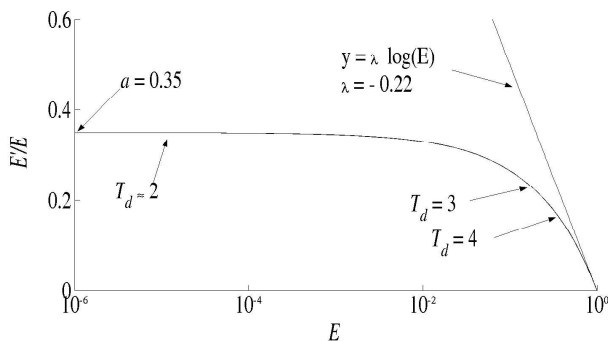


FIG. 3: Concave exponential growth rate, E'/E for (8) as a function of E in logarithmic scale. We set $a = 0.35$, $\lambda = -0.22$, $E(0) = 10^{-6}$, and $\Delta t = 1$. When $E < 10^{-3}$, we get $E'/E \approx a = 0.35$. As the error grows, the instantaneous doubling time T_d increases. We show $T_d = 2, 3$ and 4 corresponding to exponential growth rate of $0.34, 0.23$ and 0.17 , respectively. The line $y = \lambda \log(E)$ with $\lambda = -0.22$ is tangent to the E'/E curve at the asymptotic value $E = 1$.

the error, E'/E compared to E . We choose a and λ to fit the rate of Lorenz-40 model in Figure 2(a). For $E > 0.7$ [16], E'/E is close to the line (2) with $C = 1$. Similarly, the error growth rate of the QG model can be illustrated by (8) by choosing $a = .044$ and $\lambda = -0.035$.

Summary: Convex error growth in the NCEP global atmospheric model behaves completely differently from concave simpler models. In the realistic weather model, we do not see any exponential growth where E'/E is constant. Instead, E'/E is strictly decreasing. We see an initial super rapid growth, followed by two log-linear regimes. This initial superfast regime takes place because even a single bit difference in the initial conditions can trigger cumulus convection in the model, not unlike the way a small irregularity in the terrain can trigger a summer storm at a given point in Oklahoma. The fact that in this superfast period of growth, “butterfly”-sized model perturbations become finite, supports the insight of Lorenz [2] and is consistent with the results of Zhang et al. [11] and Toth and Kalnay [6].

The overall convex graph in Figure 1 reflects the smallest scale saturating very quickly at very small levels, intermediate scales saturating less quickly at intermediate levels, and large scales saturating slowly at large levels. Although these results were obtained for the NCEP global model, we believe that the results are relevant to all realistic models of the atmosphere that include convective phenomena.

Acknowledgements: The authors thank D.J. Patil and

I. Szunyogh for many insightful discussions and criticisms. This research was partially supported by the NSF grants DMS0104087 and ATMO328402, NOAA grant NA040AR4310103, NASA grant NNG04GK78A, the Army Research Office and the James S. McDonell Foundation.

* email: jharlim@math.umd.edu; Institute for Physical Science and Technology; Department of Mathematics

† Current address: Department of Physics and Astronomy, Francis Marion University, Florence, SC 29501

‡ Institute for Physical Science and Technology; Department of Mathematics; Department of Physics

§ Institute for Physical Science and Technology; Department of Meteorology

¶ Department of Mathematics; Institute for Physical Science and Technology

- [1] E. Lorenz, in *Proceedings on predictability, held at ECMWF on 4-8 September 1995* (1996), pp. 1–18.
- [2] E. Lorenz, *Tellus* **21**, 289 (1969).
- [3] E. Aurell, G. Boffetta, A. Crisanti, G. Paladin, and A. Vulpiani, *Physical Review E* **52** (1996).
- [4] J. Derber and Coauthors, NOAA/NWS Technical Procedures Bulletin **449** (1998), available from office of meteorology, NWS, 1325 East-West Hwy., Silver Spring, MD 20910.
- [5] J. Sela, *Monthly Weather Review* **108**, 1279 (1980).
- [6] Z. Toth and E. Kalnay, *Bulletin of the American Meteorological Society* **74**, 2317 (1993).
- [7] Z. Toth and E. Kalnay, *Monthly Weather Review* **125**, 3297 (1997).
- [8] J. Marshall and F. Molteni, *Journal of the Atmospheric Sciences* **50**, 1792 (1993).
- [9] E. Lorenz and K. Emmanuel, *Journal of the Atmospheric Science* **55**, 399 (1998).
- [10] J. R. Holton, *An Introduction to Dynamic Meteorology* (Academic Press, 1992), 3rd ed.
- [11] F. Zhang, C. Snyder, and R. Rotunno, *Journal of the Atmospheric Science* **60**, 1173 (2003).
- [12] E. Lorenz, *Tellus* **34**, 505 (1982).
- [13] All logs are natural logarithms; generally we want E'/E to be within 10% of $\lambda \log(E/C)$
- [14] Named after Toth and Kalnay [6, 7] and the famous paleontological demarcation point.
- [15] Lorenz [12] modeled the error growth by the one-parameter logistic equation $E' = aE(1 - E)$. This equation, however, has an initial exponential growth rate equal to its saturation rate; that is $a = -\lambda$ in (8). Such a constraint, as we see in Figure 2, is not necessarily satisfied.
- [16] [note: We say E is in the log-linear regime when $\frac{a(1-E^{-\lambda/a})}{\lambda \log E} > 0.9$, namely for $E \in (0.7, 1)$.]