Supplementary Materials:

Bo-Wen Shen [1,*]

In this report, Sections 2 and 3 provide comments on the analysis in Palmer et al., 2014 [1] and related mathematical concepts, respectively. Key points include the following: (1) An analysis of the Lorenz 1963 model (Lorenz, 1963 [2]) by Palmer et al., 2014 [1], which did not take the sensitive dependence of initial conditions (SDIC) into consideration, is inaccurate. (2) The newly invented "real butterfly effect" is fundamentally based on the findings of Lorenz, 1969 [3] who proposed a closure-based, physically multiscale, mathematically linear, and numerically ill-conditioned model. The Lorenz 1969 model is not a turbulence model and, therefore, the model's features should be interpreted with caution. As a result, (3) the mechanism of finite predictability associated with the "real butterfly effect" is different from that associated with the original butterfly effect (i.e., SDIC) within the Lorenz 1963 model.

2. Comments on the Analysis of Palmer et al. (2014)

In Section 2 of the Supplemental Materials, we provide illustrations to show that: (1) the analysis of the Lorenz 1963 (L63) model [2] by Palmer et al., 2014 [1] is inaccurate (e.g., SDIC is not considered), and (2) the newly invented "real butterfly effect" is fundamentally based on the findings of Lorenz,1969 [3] who proposed a closure-based, physically multiscale, and mathematically linear model that is ill-conditioned.

Palmer et al., 2014 [1] introduced the term "the real butterfly effect" based on the following feature of Lorenz,1969 [3]:
"Two states of the system differing initially by a small observational error will evolve into two states differing as greatly as randomly chosen states of the system within a finite time interval, which cannot be lengthened by reducing the amplitude of the initial error".

As a result, the view of Palmer et al., 2014 [1] suggested that the concept of finite predictability (in weather) is solely derived from the Lorenz 1969 study but not the Lorenz 1963 study. Palmer et al., 2014 [1] performed a mathematical analysis for the L63 model, which is not accurate, and did not discuss the ill-conditioning and numerical instability, presented in the main text, within the Lorenz 1969 model. Below, their analysis of the L63 model (e.g., in Lists S1 and S2) is examined. To facilitate discussions, we have added labels (A) and (B) in List S2 to refer to specific equations.

Low-order chaos cannot have this property of finite-time loss of predictability: for example, even though the Lorenz 1963 system exhibits sensitive dependence on initial conditions, it nevertheless exhibits continuous dependence on initial conditions. Put simply, one can predict as far ahead as one wants in the Lorenz 1963 system, providing uncertainty in the initial conditions is small enough. To show that the Lorenz 1963 system has the property of continuous dependence on initial conditions, consider first the energy identity:

$$\frac{1}{2}\frac{\partial}{\partial t}|X|^2 + \sigma x^2 + y^2 + bz^2 = (\sigma + r)xy,$$

where $|X|$ is the length of the state vector X. Since the rate of growth of $|X|$ is determined by terms which are no more than quadratic in the state variables, very rough arguments show that there exists a positive constant c_1, depending on σ, r, and b such that

$$|X(t)| \leqslant |X_0| \exp(c_1 t)$$

for all $t \geqslant 0$, where $X_0 = (x_0, y_0, z_0)$ denote initial conditions. Of course, Lorenz showed something much stronger—i.e. that trajectories cannot escape to infinity as $t \to \infty$. More

List S1: Boundedness but no SDIC [1].

In List S1, Palmer et al., 2014 [1] discussed the continuous dependence of solutions on initial conditions (CDIC) and solution boundedness within the L63 model. As discussed in Section 3 and in Shen, 2021 [4], CDIC is one of the major features of smooth dynamic systems, and boundedness is one of two features that define a chaotic system. However, as illustrated below, List S1 did not take into consideration the sensitive dependence of solutions on initial conditions (SDIC).

The first equation does not necessarily lead to boundedness in the second equation provided in List S1. When the length of a state variable, $|X|$, is considered, the time evolution of length cannot represent the rapid divergence of initial nearby trajectories at the saddle point (i.e., at the origin within the L63 model). Specifically, within the X-Y phase space, eastward and westward movement near the saddle point may produce the same time evolution of $|X|$. For example, consider control and parallel runs with two starting points at $(X,Y,Z) = (2\epsilon, \epsilon, constant)$ and $(X,Y,Z) = (-2\epsilon, -\epsilon, constant)$, respectively, where ϵ is a small number. The (initial) time varying trajectories for the control and parallel runs display "rapid" divergence within the L63 model but yield the same $\frac{\partial}{\partial t}|X|^2$ for the above equation in List S1. In other words, the above equation for $\frac{\partial}{\partial t}|X|^2$ only illustrates partial features of the L63 model. In fact, the above analysis by Palmer et al., 2014 [1] may be extended to illustrate the time evolution of the "bounded" "volume", as compared to a similar analysis from Strogatz, 2015 (pp. 323, [5]), as summarized in Section 3 for the reader's convenience. The dependence of finite predictability on the initial condition is also presented in Section 3.

and b only. In the language of dynamical systems, (2.1) has a bounded absorbing set. Now let $X(t) + \Delta X(t)$ be a solution to the Lorenz system with perturbed initial data $X_0 + \Delta X_0$ so that $\Delta X(t)$ satisfies the perturbed system:

$$\begin{aligned} \Delta\dot{x} &= \sigma(\Delta y - \Delta x), \\ \Delta\dot{y} &= \Delta x(r - \Delta z) - x\Delta z - z\Delta x - \Delta y, \\ \Delta\dot{z} &= x\Delta y + y\Delta x + \Delta x\Delta y - b\Delta z. \end{aligned} \tag{A}$$

The energy identity for the perturbed system can be written as

$$\frac{1}{2}\frac{\partial}{\partial t}|\Delta X|^2 + \sigma(\Delta x)^2 + (\Delta y)^2 + b(\Delta z)^2 = (\sigma + r)\Delta x\Delta y + (y\Delta z - z\Delta y)\Delta x. \tag{B}$$

Consider a time t in the interval $[0, T]$, then, taking into account the first energy estimate, we can state that there exists a positive constant $c_2(\sigma, r, b, X_0, T)$ such that

$$|\Delta X(t)| \leqslant |\Delta X_0| \exp(c_2 t) \tag{2.2}$$

for all $t \in [0, T]$. The latter means that solutions to the Lorenz system depend continuously on the initial data on the closed interval $[0, T]$.

List S2: An incomplete derivation for Eq. (2.2) in [1].

In List S2, the system of three equations in Eq. (A) describes the time evolution of perturbations (or the differences between two solutions with an initial difference of ΔX_0). This system along with the L63 model, containing "six" time-varying variables $(x, y, z, \Delta x, \Delta y, \Delta z)$, is similar to (but not the same as) the so-called variational equations (e.g., Shen, 2014 [6]). Obtaining a solution to Eq. (A) requires solving a system of six, first-order ordinary differential equations (ODEs), including Eq. (A), as well as the original Lorenz model. As a result, the time varying features of $x, y,$ and z (e.g., with SDIC) have an impact on time varying $\Delta x, \Delta y,$ and Δz.

With an assumption of boundedness, that is valid for the Lorenz model, and additional effort, Eq. (B) "may" lead to Eq. (2.2) under special conditions (e.g., $r \leq 1$), as discussed in Section 3. Eq. (2.2) suggests that error may grow but grows no faster than exponential growth. However, *depending on the choice of* c_2, *Eq. (2.2) may only reveal the features of CDIC and boundedness over a finite time interval for* $t \in [0, T]$. Details are provided below. From Eq. (B) to Eq. (2.2), the authors implicitly assume that a sufficiently large c_2 leads to:

$$c_3(t) \equiv \frac{1}{|\Delta X|}\frac{\partial |\Delta X|}{\partial t} \leq c_2 \qquad \text{(C)}$$

Here, $c_3(t)$ is referred to as the local time varying growth rate. Since the Lorenz model has bounded solutions, as well as bounded errors (e.g., Figure 2 of Shen 2019b [7]), c_2 can be determined based on the maximum of $c_3(t)$. However, Eq. (2.2), with such a choice of c_2, only indicates boundedness within a finite time interval. (In general, the notation of $\exp(c_2 T)$ with a positive c_2 cannot be applied for $T \to \infty$.) Specifically, although $\Delta x, \Delta y$, and Δz within the chaotic Lorenz system may irregularly oscillate with time, they are bounded with growth rates equal to or smaller than c_2 (= max (c_3)). In other words, *Eq. (2.2.) with a sufficiently large* c_2 $(= max(c_3))$ *cannot reveal information regarding whether "errors" irregularly oscillate.*

On the other hand, Eqs. (A-B) and Eq. (2.2) can have non-chaotic solutions, such as steady-state solutions (say $r = 20$). For steady-state solutions that decay with time, Eq. (2.2), with the choice of a positive c_2 that is mathematically valid, cannot properly describe the time evolution of steady-state solutions. When a (small), negative c_2 is properly selected, the time interval for the solution in Eq. (2.2) can be extended without a limit (i.e., for $t \in [0, T]$ and $T \to \infty$) due to the finite value of $\exp(c_2 T)$. As a result, unlimited predictability can be obtained.

With the above being said, *additional mathematical derivations are needed for revealing both the CDIC and SDIC of chaotic orbits* (e.g., Figure 2 of Shen, 2019b [7]). Here, we provide a very simple illustration using Eq. (2.2). With a choice of $c_2 = \lambda = 0.5 * \max(c_3)$ for the entire time interval during which both CDIC and SDIC appear, Eq. (2.2) can reveal both CDIC (with local time varying growth rates no larger than λ) and SDIC (with local time varying growth rates larger than λ) in addition to boundedness. For example, as shown in *Figure S2*, we may consider determining a $\max(c_3)$ using Eq. (C) over a shorter time interval (e.g., $\tau \in [0, 25]$ in *Figure S2*) in order to determine c_2 in Eq. (2.2).

The above approach for determining c_2 is based on the features of CDIC (prior to the onset of SDIC). However, as show in *Figures S6 and S7* (taken from Figure 1 of Slingo and Palmer, 2011 [8] and Figure 2 of Nese,1989 [9]), the time interval of CDIC or a predictability horizon displays a dependence on initial values, suggesting the dependence of c_2 on initial conditions.

Based on Eq. (2.2), with a proper choice of c_2 for CDIC within a finite time interval for $t \in [0, T]$, we can say that improving the accuracy of initial conditions (ICs) may improve accuracy predictions within the time interval (i.e., prior to the onset of SDIC).

As a brief summary, we reiterate that (1) since the smooth L63 model possesses both CDIC and SDIC, a proper choice of c_2 in Eq. (2.2) of Palmer et al., 2014 [1] is needed for revealing both the CDIC and SDIC, the latter of which displays the rapid divergence of initial, nearby trajectories; and (2) the original butterfly effect and the so-called real butterfly effect share a common feature of "finite predictability", derived from the Lorenz 1963 and Lorenz 1969 models, respectively.

3. Important Concepts and Terminology in Support of the Analysis in Section 2

To provide support for the analysis in Section 2 regarding the "real butterfly effect of Palmer et al., 2014" [1] and accurate implications, we provide detailed discussions and mathematical analyses in order to illustrate the following concepts:

A. Existence and Uniqueness

B. An Illustration of a Finite Predictability Horizon

C. Continuous Dependence on Initial Conditions (CDIC)

D. Sensitive Dependence on Initial Conditions (SDIC vs. CDIC)

E. Volume Contraction and Global Stability with (Energy-like) Lyapunov Functions

F. The Dependence of Finite Predictability on ICs Within the L63 Model

Note that related discussions are widely documented in textbooks. To properly acknowledge the efforts by textbook authors, below, we directly cite (some of) their discussions as Lists S3-S12. These concepts are based on additional references (Alligood et al., 1996 [10]; Hirsch et al., 2013 [11]; Drazin, 1992 [12]; Strogatz, 2015 [5]).

3A. Existence and Uniqueness

The Existence and Uniqueness Theorem. *Consider the initial value problem*

$$X' = F(X),\ X(t_0) = X_0,$$

where $X_0 \in \mathbb{R}^n$. Suppose that $F : \mathbb{R}^n \to \mathbb{R}^n$ is C^1. Then, first, there exists a solution of this initial value problem, and second, this is the only such solution. More precisely, there exists an $a > 0$ and a unique solution,

$$X : (t_0 - a, t_0 + a) \to \mathbb{R}^n,$$

of this differential equation satisfying the initial condition $X(t_0) = X_0$.

Without dwelling on the details here, the proof of this theorem depends on an important technique known as *Picard iteration*. Before moving on, we illustrate how the Picard iteration scheme used in the proof of the theorem works in several special examples. The basic idea behind this iterative process is to construct a sequence of functions that converges to the solution of the differential equation. The sequence of functions $u_k(t)$ is defined inductively by $u_0(t) = x_0$, where x_0 is the given initial condition, and then

$$u_{k+1}(t) = x_0 + \int_0^t F(u_k(s))\,ds.$$

List S3: The fundamental local theorem of ODEs [11].

Some important notes in List S3 are: (1) the system is smooth with C[1] (i.e., F and its first partial derivatives with respect to state variable X are continuous); (2) *a solution exists over a finite interval*; and (3) constructing "a sequence of functions" to prove the above theorem is required.

3B. An Illustration of a Finite Predictability Horizon

Here, we apply a simple nonlinear system in order to illustrate a finite predictability horizon and to discuss its dependence on initial conditions. Although we may assume that a unique solution can be defined on a maximal time domain, *there is no guarantee that a solution X(t) can be defined for all time, no matter how "nice" F(X) is* (Alligood et al., 1996 [10]). An example is given by:

$$x' = 1 + x^2 \tag{1}$$

The solution, $x(t) = \tan(t + c)$ is defined over a finite interval $-c - \frac{\pi}{2} < t < -c + \frac{\pi}{2}$, where c is a constant for the initial condition of $x(0) = \tan(c)$. The relationship provides an example that: (1) the predictability horizon is finite; the "lead time" of prediction cannot be extended while the accuracy of solutions may be continuously improved within the above time interval; and (2) the length of the time interval for the solution depends on the initial condition (i.e., c). (Note that the system is nonlinear but not chaotic).

By comparison, when the method (e.g., $x = x_c + \Delta x$, here, x_c is a critical point solution) that leads to Eq. (A) in List S2 is applied, the above equation yields:

$$\Delta\dot{x} = 2\Delta x x + (\Delta x)^2 \tag{2}$$

A simple analysis of the equation with an assumption of solution boundedness (e.g., bounded x and Δx) cannot help determine the predictability horizon ($-c - \frac{\pi}{2} < t < -c + \frac{\pi}{2}$) for the solution.

3C. Continuous Dependence on Initial Conditions (CDIC)

In addition to solution existence and uniqueness, as listed below, continuous dependence is an important feature.

Theorem. *Consider the differential equation $X' = F(X)$ where $F : \mathbb{R}^n \to \mathbb{R}^n$ is C^1. Suppose that $X(t)$ is a solution of this equation that is defined on the closed interval $[t_0, t_1]$ with $X(t_0) = X_0$. Then there is a neighborhood $U \subset \mathbb{R}^n$ of X_0 and a constant K such that if $Y_0 \in U$, then there is a unique solution $Y(t)$ also defined on $[t_0, t_1]$ with $Y(t_0) = Y_0$. Moreover $Y(t)$ satisfies*

$$|Y(t) - X(t)| \le |Y_0 - X_0| \exp(K(t - t_0))$$

for all $t \in [t_0, t_1]$. ■

This result says that, if the solutions $X(t)$ and $Y(t)$ start out close together, then they remain close together for t close to t_0. Although these solutions may separate from each other, they do so no faster than exponentially. In particular, since the right side of this inequality depends on $|Y_0 - X_0|$, which we may assume is small, we have

Corollary. (Continuous Dependence on Initial Conditions) *Let $\phi(t, X)$ be the flow of the system $X' = F(X)$, where F is C^1. Then ϕ is a continuous function of X.* ■

List S4: Continuous Dependence on Initial Conditions [11].

Some important notes from the above excerpt are: (1) two (nearby) trajectories may diverge but do not separate faster than exponentially *over a finite time interval*; and (2) CDIC appears in smooth systems with C^1. (3) In some textbooks, the above constant "K" is further discussed, as listed below, using the concept of the "Lipschitz constant".

Let $U \subset \mathbb{R}^n$ be open set. A function $F : U \to \mathbb{R}^n$ is said to be Lipschitz on U if there exists a constant L such that

$$|F(\mathbf{y}) - F(\mathbf{x})| \le L|\mathbf{y} - \mathbf{x}|$$

for all $\mathbf{x}, \mathbf{y} \in U$. We call L a Lipschitz constant for F.

List S5: Lipschitz Constant (e.g., [10]).

Let $\mathbf{G}$ be defined on the open set U in $\mathbb{R}^n$, and assume that $\mathbf{G}$ has Lipschitz constant L in the variables $\mathbf{V}$ on U. Let $\mathbf{V}(t)$ and $\mathbf{W}(t)$ be solutions of $\frac{d\mathbf{V}(t)}{dt} = \mathbf{G}(\mathbf{V})$, and let $[t_o, t_1]$ be a subset of the domains of both solutions. Then

$$|\mathbf{V}(t) - \mathbf{W}(t)| \leq |\mathbf{V}(t_o) - \mathbf{W}(t_o)|\, e^{L(t-t_o)},$$

for all $t \in [t_o, t_1]$.

List S6: CDIC and the Lipschitz Constant (e.g., [10]).

Several important notes in List S6 that are the same as the above in List S4 include: (1) "G" in List S6 is comparable to the above "F" in List S4; (2) "G" has bounded first partial derivatives; and (3) *two trajectories display CDIC over a finite time interval for $t \in [t_0, t_1]$*.

Based on the above, the relationship between c_2 of Eq. (2.2) in Palmer al., 2014 [1] and the Lipschitz constant "L" is illustrated in *Figure S1*, suggesting that Eq. (2.2) with a sufficiently large c_2 indicates the boundedness of solutions. However, as discussed in Section 2, we reiterate that *a proper choice of c_2 in Eq. (2.2) of Palmer et al.,2014 [1] is needed for revealing both CDIC and SDIC, the latter of which displays the rapid divergence of initial, nearby trajectories*.

3D. Sensitive Dependence on Initial Conditions (SDIC, as compared to CDIC)

As known, there is no universal definition for chaos. Several definitions of chaos were documented in a presentation by the lead author (e.g., Shen 2021 [4]). Amongst definitions, as shown in List S7 from page 8 of Lorenz, 1993 (pp. 8, [13]), sensitive dependence on ICs is one of the major features for chaos.

With a few more qualifications, to be considered presently, sensitive dependence can serve as an acceptable definition of chaos, and it is the one that I shall choose.

List S7: SDIC and Chaos (pp. 8, [13]).

To improve our understanding for the dependence of predictability on ICs, as well as initial errors, it is important to understand differences between the CDIC and SDIC. *Table S1* briefly summarizes both CDIC and SDIC, and presents major features for the L63 model. The two terms are defined as follows (details are provided in the Shen, 2017 [14] lecture notes):

- CDIC: *solutions through nearby ICs remain close over short time intervals*. Mathematically, $|X(t) - Y(t)| < |X(t = t_o) - Y(t = t_o)|e^{K(t-t_o)}$.
- SDIC: *"The property characterizing an orbit if most other orbits that pass close to it at some point do not remain close to it as time advances" (Lorenz, 1993 [13]). Mathematically, given δ, we may find n such that $|X(t = t_n) - Y(t = t_n)| > \delta$ (e.g., Drazin, 1992 [12]).*

Table S1. The definitions of CDIC and SDIC.

	CDIC	**SDIC**
Definition	solutions through nearby ICs remain close over short time intervals	*The property for an orbit when most other orbits that pass close to it at some point do not remain close to it as time advances.*

An example	the "gradual divergence" or convergence of initial nearby trajectories	"rapid divergence" of initial nearby trajectories

The lead author gave a virtual, invited talk to Prof. Palmer's group at Oxford University on 11 October 2021 (Shen, 2021 [4]). Based on these discussions, the lead author feels that it may be easier to illustrate differences between the two terms by first quoting discussions from the textbook by Alligood, Sauer, and Yorke, 1996 [10].

(a) On page 103, Alligood et al.,1996 (pp. 103, [10]) presented the following, which suggests finite predictability for a chaotic system (e.g., [15]):

"While sensitive dependence on initial conditions causes unpredictability at large time scales, it can provide opportunity at shorter, predictable time scales".

(b) On page 281, Alligood et al., 1996 (pp.281, [10]) clarify differences between the continuous and sensitive dependence on initial conditions, as follows:

"… The third concept is referred to as "continuous dependence of solutions on initial conditions". All solutions of sufficiently smooth differential equations exhibit the property of continuous dependence, which is a consequence of continuity of the slope field. Theorems pertaining to existence, uniqueness, and continuous dependence are discussed in Section 7.4.

The concept of continuous dependence should not be confused with "sensitive dependence" on initial conditions. This latter property describes the behavior of unstable orbits over longer time intervals. Solutions may obey continuous dependence on short time intervals and also exhibit sensitive dependence, and diverge from one another on longer time intervals. This is the characteristic of a chaotic differential equation".

We additionally realized that it may be effective to provide the following excerpt from page 325 of Hirsch et al, 2013 [11] to show SDIC within the L63 model:

Theorem. (Dynamics of the Lorenz Model) *The Poincaré map Φ restricted to the attractor A for the Lorenz model has the following properties:*

1. *Φ has sensitive dependence on initial conditions*
2. *Periodic points of Φ are dense in A*
3. *Φ is transitive on A*

List S8: A Definition of Chaos: Dynamics of the L63 Model (e.g., [11]).

Calling a continuously differentiable function a C^1 function is traditional. Features for the CDIC and SDIC within the L63 model are provided in *Figure S2*. Control and parallel runs were performed using the same model and the same parameters. The only difference between the two runs is the inclusion of an initial tiny perturbation, $\varepsilon = 10^{-10}$, within the parallel run. In *Figure S2*, two runs initially produced almost the same result (for $\tau \in [0, 25]$), as shown with the red curve. This feature is called CDIC. During longer time integrations, the appearance of two curves (in red and blue) indicates significant differences (i.e., "rapid divergence") in solutions for the control and parallel runs. Such a feature is then called SDIC, due to the initial tiny perturbation, $\varepsilon = 10^{-10}$. As a brief summary, the L63 model within the chaotic regime (say $r = 28, \sigma = 10$, and $b = 8/3$) displays CDIC within an initial finite time interval and SDIC at a later time.

As a result of CDIC and SDIC, a finite predictability (i.e., predictability at finite time scales; Lighthill, 1986 [15]) can be obtained. On the other hand, determining the exact interval for continuous dependence on initial conditions is challenging. In fact, as a result of time varying growth rates, leading to different predictability, such an interval, if found, should, as discussed below, display a dependence on the initial conditions.

Figures S3 and S4 present a "simple illustration" in order to reveal the role of a saddle point in producing SDIC. In *Figure S3*, a control and two runs with initial perturbations, $\varepsilon = 10^{-10}$ and $\varepsilon = -10^{-1}$, can produce very different results. In the bottom panels of

Figure S4, a parallel run with the smallest initial perturbation (panels (e) and (f)) clearly does not produce the best prediction as compared to the other two parallel runs in panels (a)-(d).

3E. Volume Contraction and Global Stability with Lyapunov Functions

Here, we discuss (1) volume contraction and (2) global stability using equations that may be viewed as special kinds of the single, first-order ODE. The corresponding mathematical approaches may help provide an idea of how to derive an improved Eq. (2.2) for Palmer al., 2014 [1] in List S2 for revealing the features of various types of solutions (within the L63 model). (Current Eq. (2.2) is very limited and cannot reveal the major feature of SDIC.).

List S9 displays a solution of $V = V_0 e^{-(\sigma+1+b)t}$, indicating that the volume element exponentially contracts in time for positive parameters. The flow contracts the volume element in some directions, but stretches it along others. As a result, to remain confined to a bounded domain, the volume element is folded at the same time. Stretching, squeezing, and folding processes are associated with the SDIC.

Consider $\Delta V = \Delta X \Delta Y \Delta Z$, which represents the volume of the solution.

$$\begin{aligned}\frac{d\Delta V}{dt} &= \frac{d\Delta X}{dt}\Delta Y \Delta Z + \frac{d\Delta Y}{dt}\Delta X \Delta Z + \frac{dZ}{dt}\Delta X \Delta Y \\ &= \left(\frac{1}{\Delta X}\frac{dX}{dt} + \frac{1}{\Delta Y}\frac{dY}{dt} + \frac{1}{\Delta Z}\frac{dZ}{dt}\right)\Delta X \Delta Y \Delta Z \\ &= (\nabla \cdot F)(\Delta V) \\ &= -(\sigma + 1 + b)\Delta V\end{aligned}$$

Here, F represents the forcing terms of the Lorenz 1963 model (e.g., the terms on the right-hand side of Eqs. A1-A3 in the main text).

$$\Delta V = V_o e^{-(\sigma+1+b)t}$$

List S9: Volume Contraction Within the L63 Model (e.g., [16]).

Below, we provide a brief introduction to the Lyapunov function, constructed as an energy-like function, in order to provide a global approach towards determining the asymptotic behavior of solutions (e.g., Alligood et al., 1996 [10]). The Lyapunov function may be viewed as a generalization for a potential energy function. Let $E(\vec{X})$ be a Lyapunov function and $\vec{X}_*$ be a point such as $E(\vec{X}_*) = 0$. $\vec{X}$ is a vector that consists of all state variables, $\vec{X} = (X, Y, Z)$, for the L63 model. A theory of Lyapunov functions is provided in List S10 (as derived from Strogatz, 2015 [5]).

The Lyapunov function $E : R^n \to R$ for $\vec{X}_*$ in some neighborhood D of $\vec{X}_*$ has the following properties:

(1) $E(\vec{X}_*) = 0$ and $E(\vec{X}) > 0$ for all $X \neq X_*$ in D, and

(2) $\dot{E}(\vec{X}) \leq 0$ for all X in D,

here, $\dot{E}(\vec{X})$ represents the rate of change of E along a solution trajectory. Therefore, the above properties suggest that the Lyapunov function (E) is positive definite and decreases along the trajectory. Then, X_* is globally asymptotically stable for all ICs at $t \to \infty$.

List S10: A Definition of the Lyapunov Function (e.g., [5]).

The above suggests that *if we can find an energy-like function that is positive definite where its time derivative is non-positive, we can show a global stability for point* $\vec{X}_*$. For all X, Y, and Z when $r \leq 1$, we may have the following Lyapunov function for the L63 model (e.g., pp.238, [12]), showing global stability ($\frac{dE}{dt} \leq 0$) for $r \leq 1$ (in List S11):

$$E(X,Y,Z) = \frac{1}{2}(X^2 + \sigma Y^2 + \sigma Z^2).$$

$$\frac{dE}{dt} = -\frac{1}{2}\sigma(1+r)(X-Y)^2 - \frac{1}{2}\sigma(1-r)(X^2+Y^2) - \sigma b Z^2 \leq 0 \quad \text{for all } r \leq 1$$

List S11: Existence of the Lyapunov Function and Global Stability Within the L63 Model for $r \leq 1$.

By introducing $V = \frac{2E}{\sigma}$, Strogatz turned the above equation (for dE/dt) in List S11 into the last equation (dV/dt) in List S12, displaying the same global stability for $r < 1$. (Note that the conclusion remains true for $r = 1$ in List S12, suggesting the condition of $r \leq 1$.)

Here, consider $V(x,y,z) = \frac{1}{\sigma}x^2 + y^2 + z^2$. The surfaces of constant V are concentric ellipsoids about the origin (Figure 9.2.3).

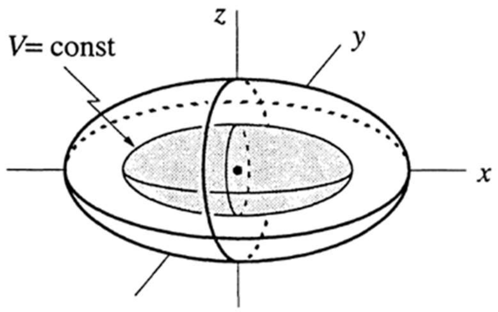

Figure 9.2.3

The idea is to show that if $r < 1$ and $(x,y,z) \neq (0,0,0)$, then $\dot{V} < 0$ along trajectories. This would imply that the trajectory keeps moving to lower V, and hence penetrates smaller and smaller ellipsoids as $t \to \infty$. But V is bounded below by 0, so $V(\mathbf{x}(t)) \to 0$ and hence $\mathbf{x}(t) \to 0$, as desired.

Now calculate:

$$\begin{aligned} \tfrac{1}{2}\dot{V} &= \tfrac{1}{\sigma}x\dot{x} + y\dot{y} + z\dot{z} \\ &= (yx - x^2) + (ryx - y^2 - xzy) + (zxy - bz^2) \\ &= (r+1)xy - x^2 - y^2 - bz^2. \end{aligned}$$

Completing the square in the first two terms gives

$$\tfrac{1}{2}\dot{V} = -\left[x - \tfrac{r+1}{2}y\right]^2 - \left[1 - \left(\tfrac{r+1}{2}\right)^2\right]y^2 - bz^2.$$

We claim that the right-hand side is strictly negative if $r < 1$ and $(x,y,z) \neq (0,0,0)$. It

List S12: Global Stability Within the L63 Model for $r < 1$ [5].

In List S12 where $V(x,y,z) = \frac{1}{\sigma}x^2 + y^2 + z^2$ represents a family of concentric ellipsoids, the right hand side of the last equation is negative as long as r is less than one. *Although the above theory of Lyapunov functions seems promising, there is no systematic way to construct a Lyapunov function. As a result, the right hand side of Eq. (2.2) may be positive or* *negative. Thus, additional research is required in order to analyze and refine Eq. (2.2) of Palmer et al., 2014 [1].*

3F. The Dependence of Finite Predictability on ICs Within the L63 Model

Chaotic solutions with one or more positive global Lyapunov exponents possess time-varying local Lyapunov exponents (e.g., *Figure S5*) or time varying local growth rates. As shown in *Figure S6* (as derived from Slingo and Palmer, 2011 [8]), time varying growth rates, indeed, indicate the dependence of predictability on "location" within the phase space. In fact, as shown in *Figure S7*, such a dependence of predictability was previously documented in Nese, 1989 [9]. *In summary, the chaotic L63 model displays the dependence of finite predictability on ICs.*

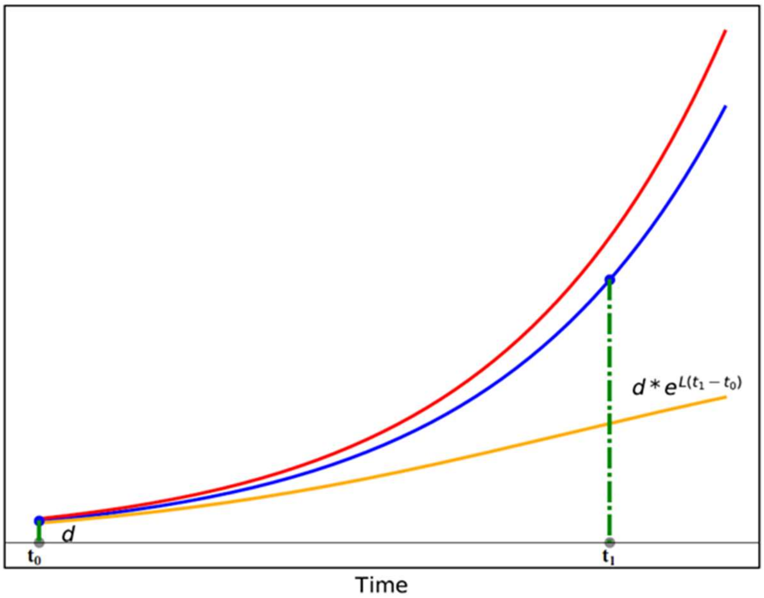

Figure S1. A relationship between c_2 of Eq. (2.2) of Palmer et al, 2014 [1] and the Lipschitz constant "L" (e.g., Alligood et al., 1996 [10]). Nearby solutions can diverge no faster than an exponential rate determined by the Lipschitz constant of the system (in blue). A sufficiently large c_2 leads to growth rates (in red) larger than the exponential rate determined by the Lipschitz constant (in blue). Eq. (2.2), with such a choice of c_2, indicates the boundedness of solutions with CDIC and SDIC. See Figure 7.13 of Alligood et al, 1996 [10] for details.

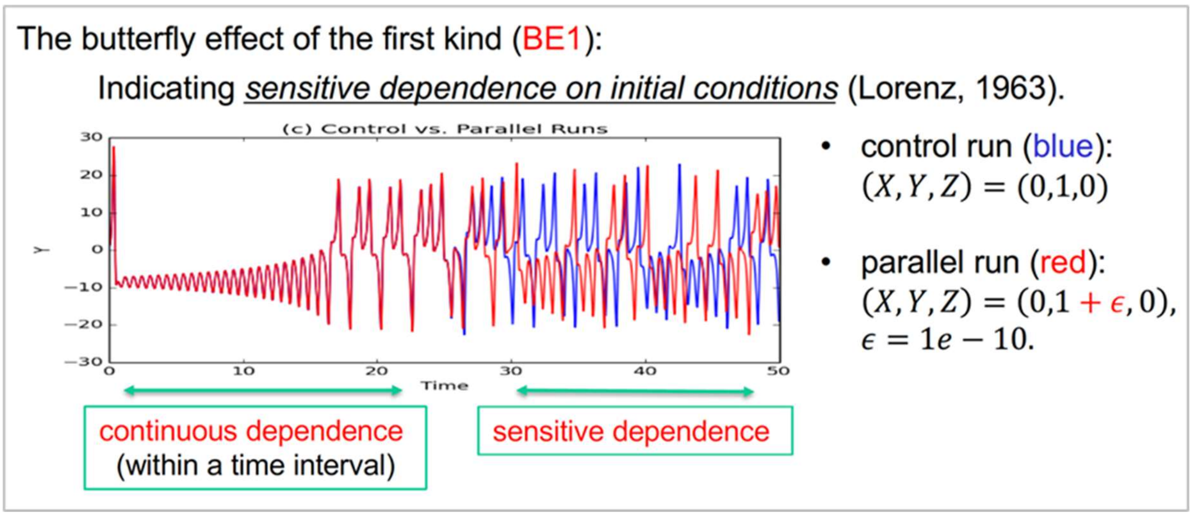

Figure S2. An illustration of continuous and sensitive dependence on initial conditions (CDIC and SDIC).

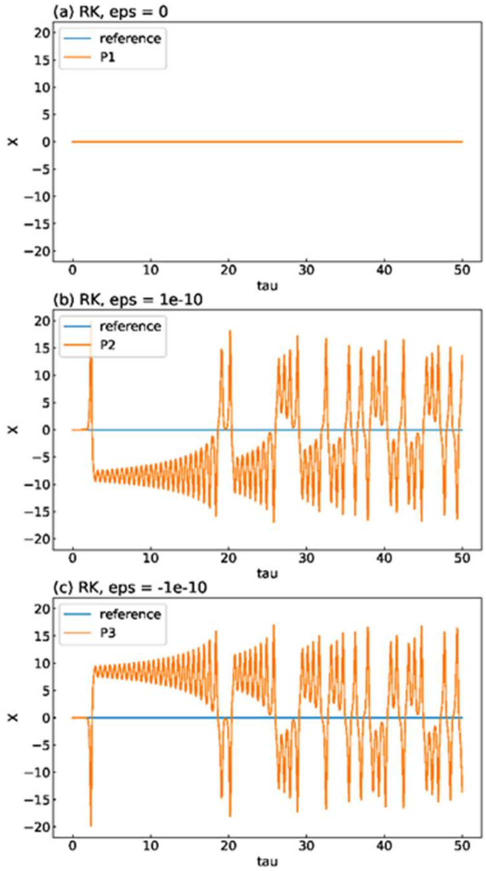

Figure S3. Numerical experiments for revealing the role of the saddle point in producing SDIC. Three runs with ICs of $(\varepsilon, 0, 0)$ apply $\varepsilon = 0,\ 10^{-10},\ and\ -10^{-10}$, respectively.

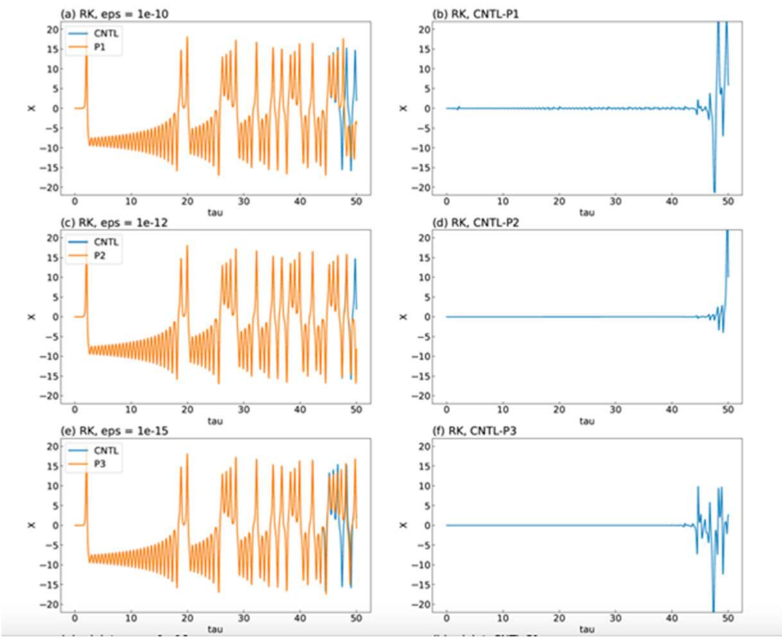

Figure S4. Experiments with various tiny perturbations for revealing the continuous and sensitive dependence of initial conditions. Left panels compare the control run (in blue) with one of the parallel runs that contain an initial tiny perturbation of $\varepsilon = 10^{-10}, 10^{-12}$, and 10^{-15}, respectively. Right panels display the differences of the results from the control and one of the parallel runs. In the bottom panel with the smallest tiny perturbation, rapid divergence of the trajectories for the control and parallel runs makes the earliest appearance.

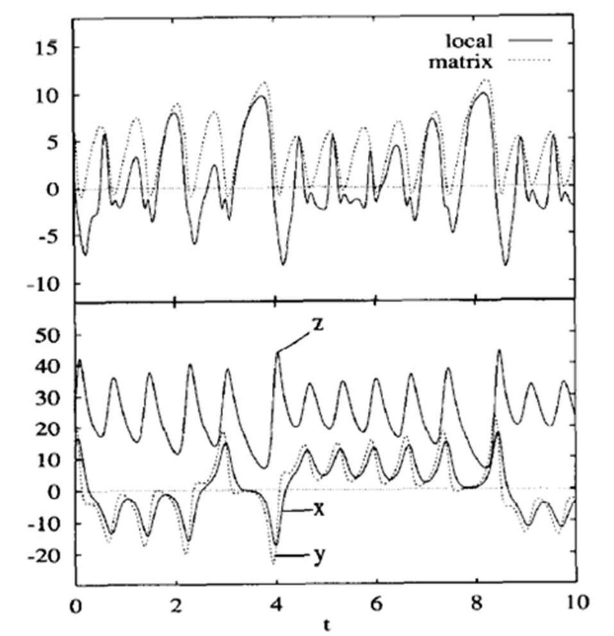

Fig. 1. Local Lyapunov exponents for the Lorenz system. The top figure shows the local Lyapunov exponent as computed from eq. (4) along a segment of a longer trajectory (full line). For comparison we also show the root of the maximal eigenvalue of the symmetrized matrix $(\mathcal{L}^{\mathrm{T}} + \mathcal{L})/2$ (dashed line). In the lower half the x-, y- and z-components of the trajectory are shown. One notes a correlation between large Lyapunov exponents and approaches to the origin, though in the local Lyapunov exponent this is not as closely correlated as in the eigenvalues of the matrix.

Figure S5. Local Lyapunov exponents that change with time (Figure 1 of Eckhardt and Yao, 1993 [17]).

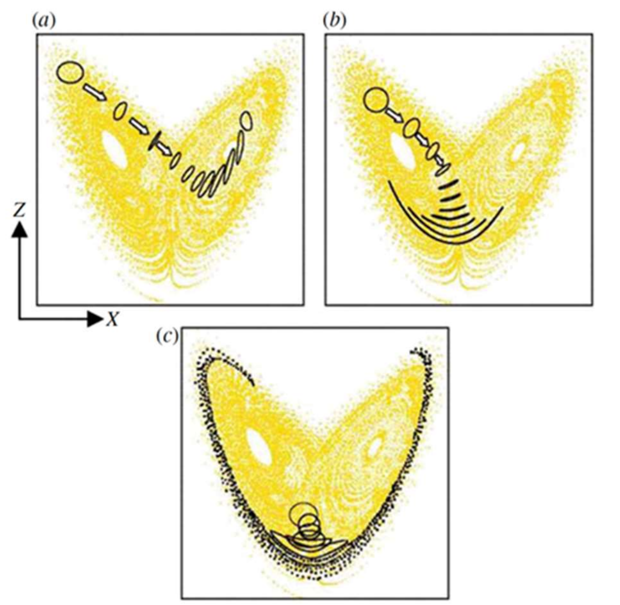

Figure S6. "Examples of finite-time error growth on the Lorenz attractor for three probabilistic predictions starting from different points on the attractor. (a) High predictability and therefore a high level of confidence in the transition to a different 'weather' regime. (b) A high level of predictability in the near term but then increasing uncertainty later in the forecast with a modest probability of a transition to a different ‘weather’ regime. (c) A forecast starting near the transition point between regimes is highly uncertain.” (Figure 1 of Slingo and Palmer, 2011 [8]).

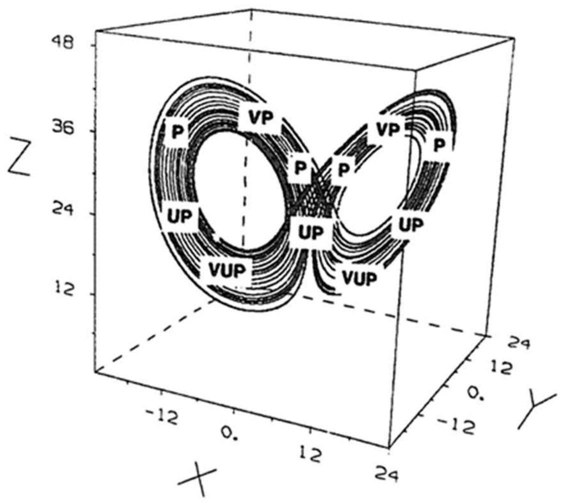

Figure S7. “Qualitative description of local predictability on the Lorenz attractor. The abbreviations VP, P, UP, and VUP denote very predictable, predictable, unpredictable, and very unpredictable” (Figure 7 of Nese, 1989 [9]). Note that Figure 2 of Nese, 1989 [9] discussed the temporal variation of predictability.

References:

1. Palmer, T. N., A. Doring and G. Seregin [2014] The real butterfly effect. Nonlinearity 27 (2014) R123–R141 doi:10.1088/0951-7715/27/9/R123
2. Lorenz, E. N. [1963] Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130–141.
3. Lorenz, E. N. [1969] The predictability of a flow which possesses many scales of motion. Tellus, 21, 289-307.
4. Shen, B.-W. [2021] Is Weather Chaotic? Multistability, Multiscale Instability, and Predictability within Lorenz Models. Oxford University. 11 October 2021. https://doi.org/10.13140/RG.2.2.13991.88485
5. Strogatz, S. H. [2015] Nonlinear Dynamics and Chaos. With Applications to Physics, Biology, Chemistry, and Engineering. Westpress view, Boulder, CO, 513 pp.
6. Shen, B.-W. [2014] Nonlinear feedback in a five-dimensional Lorenz model, J. Atmos. Sci., 71, 1701-1723. https://doi.org/10.1175/JAS-D-13-0223.1
7. Shen, B.-W., [2019b] On the predictability of 30-day global mesoscale simulations of African easterly waves during summer 2006: A view with the generalized Lorenz model. Geosciences, 9, 281, https://doi.org/10.3390/GEOSCIENCES9070281
8. Slingo, J., and T. Palmer [2011]: Uncertainty in weather and climate prediction. Philos. Trans. Roy. Soc, 369A, 4751–4767. https://doi.org/10.1098/RSTA.2011. 0161.
9. Nese, J. M. [1989]: Quantifying local predictability in phase space. Physica D, 35, 237–250, https://doi.org/10.1016/0167-2789(89)90105-X.
10. Alligood K., T. Saucer, J. Yorke [1996] Chaos An Introduction to Dynamical Systems. Springer-Verlag New York.
11. Hirsch, M., S. Smale, and R. L. Devaney [2013] Differential Equations, Dynamical Systems, and an Introduction to Chaos. 3rd edition. Academic Press, 432 pp.
12. Drazin, P. G. [1992] Nonlinear Systems. Cambridge University. Press, New York, pp. 333.
13. Lorenz, E. N. [1993] The essence of chaos. University ofWashington Press, Seattle, 227pp.
14. Shen, B.-W. [2017] Lecture Slides for Linearization Theorems. Course Materials for Math537, Fall, 2017. http://doi.org/10.13140/RG.2.2.26204.90247
15. Lighthill, J. [1986] The recently recognized failure of predictability in Newtonian dynamics. Proc. Roy. Soc. Lond. A 407, 35-50.
16. Schuster, H. G. and W. Just [2005] Deterministic Chaos: An Introduction, Fourth Edition. Wiley-VCH Verlag GmbH & Co. KGaA. DOI:10.1002/3527604804. 287pp.
17. Eckhardt, B., and D. Yao [1993] Local Lyapunov exponents in chaotic systems. Physica D, 65, 100–108.