Referee 3 (Report 1 04 Mar 2024)

We are grateful to the referee for devoting time to our manuscript.
We will here respond to comments made:

(Comment 1) The reviewer acknowledges the authors' efforts in addressing the review comments. However, the current format of their responses lacks point-by-point clarification, posing challenges in evaluating their responses. Therefore, it is recommended to reformat the responses for clarity……..…2

(Comment 2) Furthermore, considering the authors' assertion that the proposed 05 system simulates "5 to 7 main highs and lows that correspond to planetary waves (Rossby wave)," it would be advantageous to discuss whether the proposed system, without the Coriolis force, could replicate key features of the Rossby wave, including phase speeds. Historically, experiments such as dishpan experiments aimed to "simulate" weather features, yielding diverse outcomes like chaotic solutions and vacillation (e.g., limit cycle)……..……17

(Comment 3) This study extends from the authors' previous research. The reviewer acknowledges the related efforts. However, after examining their earlier studies, the reviewer proposes the following: 1. Document and report the calculation of Lyapunov exponents (LEs) within the proposed 05 system. For instance, employing the 1963 model with common parameters, the largest LE (LE1) is 0.906, as exemplified in the link provided (https://sprott.physics.wisc.edu/chaos/lorenzle.htm). This task holds significant importance. 2. Develop the error growth model, e.g., dE/dt = sigma E (1 - E/Es), and furnish a mathematical expression for sigma and LE1 of the proposed system. It should be noted that the long-time average of (1/E dE/dt) is not precisely equal to sigma……..19

(Comment 1) - Referee 3 (Report 1 04 Dec 2023) - reformated

We are grateful to the referee for devoting their time to our manuscript. The valuable comments and suggestions will help us to improve the paper.
We will here respond to comments made:

(Major Comment A) Different two-scale models in Lorenz (1996) and Lorenz (2005)
Figure R1 displays the two-scale model proposed by Lorenz (1996, 2006), including Eqs. (3.2)-(3.3) of Lorenz (2006). It is worth mentioning that Lorenz (1966) and Lorenz (2006) are the same article. Eq (3.2) for the large-scale flow does not include the explicit forcing term "F", which appears in his one-scale model. This is a typo. For the small-scale flow in Eq. (3.3), where F is not explicitly included, the coupling term acts as the forcing to derive the small scale process. Within the two-scale model, the grid system was illustrated in Figure R2 derived from Wilks (2005). Such a grid system is similar to the grid system of the multiscale modeling framework (MMF, e.g., Tao et al. 2008; Shen et al. 2011), consisting of a general circulation model (GCM, e.g., Lin et al. 2003; Lin 2004; Shen et al. 2006) for large-scale flows, and multiple copies of a cloud model (e.g., Tao 2003) for small-scale flows. Specifically, a copy of the cloud model at fine resolutions is embedded within each grid of the GCM.
Within the 2005 models, Lorenz first included additional nonlinear terms in the 1996 one-scale model (e.g., Eq. 8 in Figure R3) for slow variables (represented as Xn). Based on the 1996 one-scale model with coefficients of ("b2 ", "b", "0") for nonlinear terms, dissipative terms, and forcing term, respectively, a subsystem for fast variables (represented as Yn) was deployed and coupled with the subsystem for the slow variables. The coupled system with a 3 coefficient of "c" for coupling terms is referred to as the two-scale system (Eqs. 12a and 12b in Figure R4). The coupling terms were established based on a one-to-one relationship between Xn and Yn. Thus, the Lorenz 2005 two-scale model is different from the 1996 two-scale model. Will it be feasible for providing a diagram for illustrating the grid system within the 2005 two-scale model?

How good are such naive estimates? We can demonstrate some simple systems where they describe the situation rather well, at least on the average. One system is one that I have been exploring in another context as a one-dimensional atmospheric model, even though its equations are not much like those of the atmosphere. It contains the K variables $X_1, \ldots, X_K$, and is governed by the K equations

$$dX_k/dt = -X_{k-2}X_{k-1} + X_{k-1}X_{k+1} - X_k + F, \tag{3.1}$$

where the constant F is independent of k. The definition of X_k is to be extended to all values of k by letting X_{k-K} and X_{k+K} equal X_k, and the variables may be thought of as values of some atmospheric quantity in K sectors of a latitude circle. The physics of the atmosphere is present only to the extent that there are external forcing and internal dissipation, simulated by the constant and linear terms, while the quadratic terms, simulating advection, together conserve the total energy $(X_1^2 + \cdots + X_K^2)/2$.

distinct time scales. The model has been constructed by coupling two systems, each of which, aside from the coupling, obeys a suitably scaled variant of Eq. (3.1). There are K variables X_k plus JK variables $Y_{j,k}$, defined for $k = 1, \ldots, K$ and $j = 1, \ldots, J$, and the governing equations are

$$dX_k/dt = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k - (hc/b)\sum_{j=1}^{J} Y_{j,k}, \tag{3.2}$$

$$dY_{j,k}/dt = -cbY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - cY_{j,k} + (hc/b)X_k. \tag{3.3}$$

The definitions of the variables are extended to all values of k and j by letting X_{k-K} and X_{k+K} equal X_k, as in the simpler model, and letting $Y_{j,k-K}$ and $Y_{j,k+K}$ equal $Y_{j,k}$, while $Y_{j-J,k} = Y_{j,k-1}$ and $Y_{j+J,k} = Y_{j,k+1}$. Thus, as before, the variables X_k can represent the values of some quantity in K sectors of a latitude circle, while the variables $Y_{j,k}$, arranged in the order $Y_{1,1}, Y_{2,1}, \ldots, Y_{J,1}, Y_{1,2}, Y_{2,2}, \ldots, Y_{J,2}, Y_{3,1}, \ldots$, can represent the values of some other quantity in JK sectors. A large value of J implies that many of the latter sectors are contained in one of the former, and we may think of the variables $Y_{j,k}$ as representing a convective-scale quantity, while, in view of the form of the coupling terms, the variables X_k should represent something that favours convective activity, possibly the degree of static instability.

Figure R1: Lorenz 1996 one-scale (top) and two-scale (bottom) systems (Lorenz 2006). Since the 1996 model was applied to represent an atmospheric variable in K sectors of a latitude circle. Thus, the value of K indicates the number of grid points within the large-scale system, while the value of J represents the number of grid points within the small-scale system. Compared to the 2005 version, (1) the last term in Eq. (3.2) represents a feedback term that is a summation of small scale modes and (2) both Eqs. (3.2) and (3.3) contain two nonlinear terms, involving three neighboring grid points (at k-1, k+1, k+2).

The Lorenz '96 system (Lorenz 1996) is given by:

$$\frac{dX_k}{dt} = -X_{k-1}(X_{k-2} - X_{k+1}) - X_k + F - \frac{hc}{b}\sum_{j=J(k-1)+1}^{kJ} Y_j; \quad k = 1, \ldots, K \tag{1a}$$

$$\frac{dY_j}{dt} = -cbY_{j+1}(Y_{j+2} - Y_{j-1}) - cY_j + \frac{hc}{b}X_{\text{int}[(j-1)/J]+1}; \quad j = 1, \ldots, JK. \tag{1b}$$

It is is used here to define 'truth,' i.e. the quantities to be predicted. This system has been used in several previous studies as a metaphor for the atmosphere (Lorenz 1996; Palmer 2001; Smith 2001; Orrell 2002, 2003; Vannitsem and Toth 2002; Roulston and Smith 2003), although with slightly different notation. Equation (1a) describes the linked dynamics of a set of K slow, large-amplitude variables X_k, each of which is associated with J fast, small-amplitude variables Y_j whose dynamics are described by Eq. (1b). Here $K = 8$ and $J = 32$, so that there are $JK = 256$ Y variables in total, as illustrated in Fig. 1. The scaling constants h, c, and b are taken to be 1, 10, and 10, respectively, as is conventional; and F is a forcing taken in the following to be either 18 or 20. The subscripts are cyclic so, for example, $X_0 = X_K$, $X_{-1} = X_{K-1}$, etc. and likewise for the Y variables.

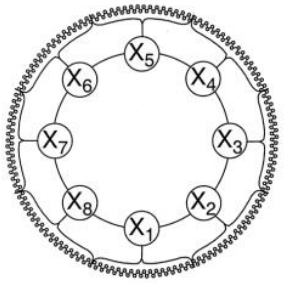

Figure 1. Schematic illustration of the Lorenz '96 system (Eq. (1)) with $K = 8$ resolved variables X_k (large circles), each associated with $J = 32$ unresolved variables Y_j (unlabeled small circles grouped according to the X variable to which they belong), so that there are $JK = 256$ Y variables in total. The forecast model (Eq. (2)) represents explicitly only the X variables, with contributions to each tendency that are due to the unresolved scales being parametrized in terms of the local resolved variable only.

Figure R2: Mathematical equations (top) and grid systems (bottom) within the Lorenz 1996 two-scale model (e.g., Wilks 2005). Eight large-scale variables (denoted as X) are selected at 8 data points within the large-scale system. Each large-scale variable acts as a force to drive a small-scale system consisting of thirty-two variables (denoted as Y). Based on the linear stability analysis, local growth rates should display a dependence on the number of data points in both the large-scale and small-scale systems and the coupling terms between the two-scale systems.

2. One chooses a number K, much smaller than N and let $J = K/2$ if K is even and $J = (K - 1)/2$ if K is odd. Then, for any two sets of variables X and Y, one defines

$$[X, Y]_{K,n} = \sum_{j=-J}^{J}{}' \sum_{i=-J}^{J}{}' (-X_{n-2K-i}Y_{n-K-j} + X_{n-K+j-i}Y_{n+K+j})/K^2 \tag{7}$$

if K is even, with Σ' replaced by Σ if K is odd. The equation for Model II, where the only set of variables is X, will be

$$dX_n/dt = [X, X]_{K,n} - X_n + F. \tag{8}$$

Note that setting $K = 1$ makes $J = 0$; hence $[X, X]_{1,n}$ represents the single pair of products appearing in Eq. (1). Model II then reduces to Model I.

Figure R3: Equation (8) the above excerpt represents a revised one-scale model, proposed as the uncoupled version of the Model II in Lorenz (2005). The notation of [X, X] defined in Equation (7) indicates nonlinear terms. Compared to the original one-scale model in Figure R1 that contains a pair of nonlinear terms, a value of K > 1 in Eq. (8) suggests more than one pair of nonlinear terms. From a perspective of scale interactions (e.g., Lorenz 1969b), additional nonlinear terms (at different grid points) may improve the representation of scale interaction.

entirely by the coupling. One obtains the system

$$dX_n/dt = [X, X]_{K,n} - X_n - cY_n + F, \tag{12a}$$

$$dY_n/dt = b^2[Y, Y]_{1,n} - bY_n + cX_n, \tag{12b}$$

where, like b, the coupling coefficient c is an additional parameter of the model.

Figure R4: Equation (12) in the above excerpt represents a revised two-scale model, proposed as Model II in Lorenz (2005). Here, Eq. (12a) is a revised large-scale system with more than one pair of nonlinear terms (when K > 1). Eq. (12b) indicated a revised small-scale system with one pair of nonlinear terms. In the coupled system, there exists a one-to-one relationship between the large-scale variable Xn and the small-scale variable Yn within the coupling terms.

Response: Figure RR1 shows the similarity of the 1996 (Eqs. (1a) and (1b) in Figure R1) and 2005 (Eqs. (12a) and (12b) in Figure R4) two-scale systems in the attempt to maintain 5 to 7 main highs and lows and several smaller waves for large scales X_n. While for the 1996 two-scale system, this is ensured by a number of *N* large scale variables X_n close to 30 (and a number of *JN* variables for the small scales), for the 2005 system, it is ensured by linking the X_n variables as described in Eq. (8) in Figure R3 (with the same number of small scale variables, however, determined from Eq. (3.1) in Figure R1, see Figure RR2). The 2005 two-scale system thus produces a smoother and more realistic evolution of the large-scale variable while maintaining properties similar to the 1996 system.

The systems used in this manuscript, which are described in Appendix A (of the manuscript), address one more condition that brings them closer to real systems. This condition is the fact that the large scale and small scale features in Eqs. (1a) – (1b) in Figure R2 and Eqs. (12a) – (12b) in Figure R4 are represented by separate sets of variables instead of appearing as superimposed features of a single set. To satisfy this condition, the coupling of one small-scale variable and one large-scale variable is more realistic than the coupling that is present in the 1996 system (Eqs. (1a) and (1b) in Figure R2).

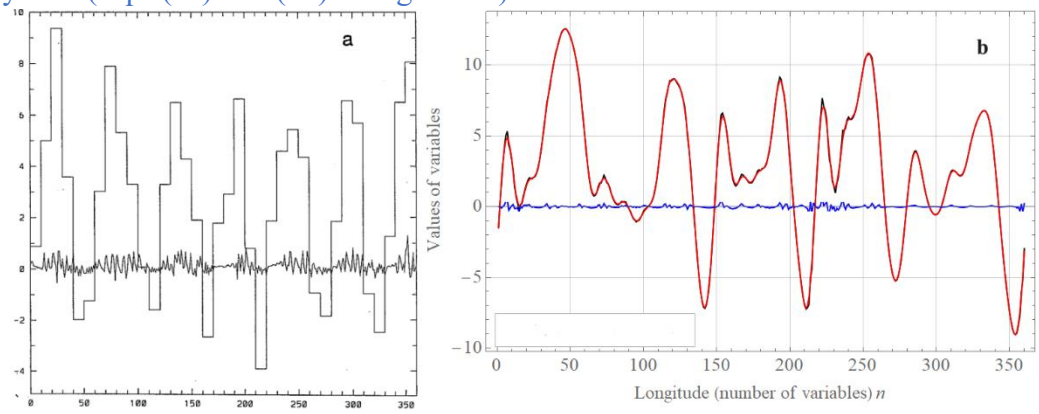

Figure RR1: Comparison of longitudinal profiles at one time of two-scale Lorenz systems (a) from 1996 (Eqs. (1a) and (1b) in Figure R2) and (b) from 2005 (Eqs. (12a) and (12b) in Figure R4).

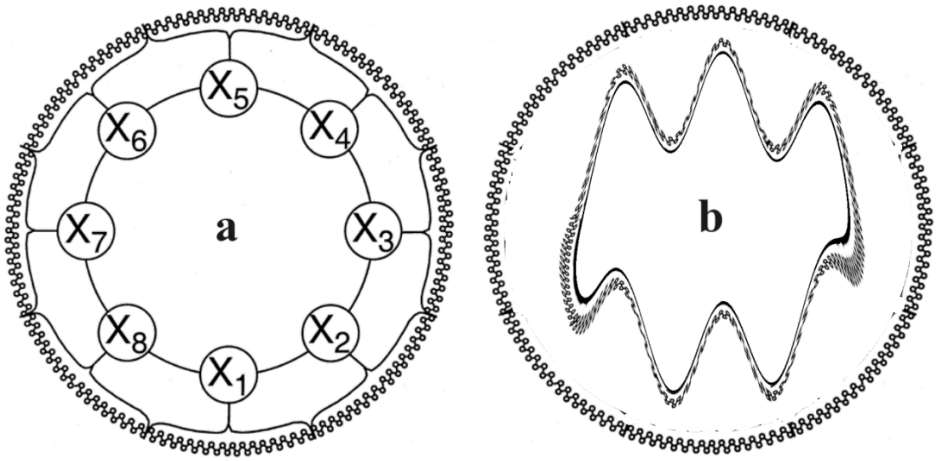

Figure RR2: Comparison of schematic illustrations of two-scale Lorenz systems (a) from 1996 (Eqs. (1a) and (1b) in Figure R2) and (b) from 2005, where the inner wave curve represents the large-scale variables described by Eq. (12a) in Figure R4, which produce 5-7 main waves, and where the outer curve represents the small-scale variables described by Eq. (12b) in Figure R4, which are not limited by the number of waves. In contrast to (a), one large scale variable is coupled to one small scale variable.

(Major Comment B) Dependence of findings on temporal spacing (i.e., Δt) and "spatial" spacing (e.g., the number of sectors, N)

As an analogy, the CFL condition, requiring $c\Delta t/\Delta x < 1$, here c is the space speed, suggests the importance of selecting temporal and spatial spacings for solution's stability. In this study, it is important to explore the impact of Δt and N.

Similarly, the concept of computational chaos (Lorenz 1989) also suggests the importance of wisely choosing Δt. Computational chaos appears "when the exact solution varies periodically with time, there is sometimes a range of time increment where the computed solution is chaotic" (Lorenz 2006). Computational chaos can be illustrated by a comparison of the Logistic differential equation and the Logistic map (i.e., difference equation). While the former has analytical, regular solutions, the latter produces chaotic solutions when a control parameter is sufficiently large. A dependence of the control parameter on a temporal spacing (i.e., Δt) can be shown by deriving the Logistic map from the Logistic differential equation (Shen et al. 2023). In this study, Δt is 1/240 ~ 4.2×10^{-3} unit, N = 360 (indicating a "spatial" spacing), and L = 12 (i.e., indicating complexities of scale interaction). It would be ideal for additional tests with a smaller $\Delta t = 10^{-5}$ (or $\Delta t = 10^{-4}$). Additionally, the choice of N and L should be explored since N = 960 and L = 32 were used in Lorenz (2005).
As discussed below, the values of the coefficients for the coupling terms could impact the growth rate of the system as well.

Response: The choice of the variable $N = 360$ was made because the value of the largest Lyapunov exponent λ^{L05} of the system described by Eq. (8) in Figure R3 ($F = 15$, time unit = 5 days) does not change for $N = 360$ and $N = 960$ (Table RR1) and therefore we chose the lower of the two values for computational efficiency.

N	λ^{L05}
30	0.70
60	0.29
90	0.35
120	0.32
150	0.33
360	0.33
960	0.33

Table RR1: Values of the largest Lyapunov exponent λ^{L05} for selected numbers of variables N in the 2005 Lorenz system (Eq. (8) in Figure R3, $F = 15$, time unit = 5 days).

Figure RR3 compares the time evolution of the average value of the variables for the 2005 Lorenz system (Eq. (8) in Figure R3) with time step Δt=1/240 and Δt=1/2400. It can be seen that the values are similar. Given this, we use the larger time step dt=1/240 for faster computations.

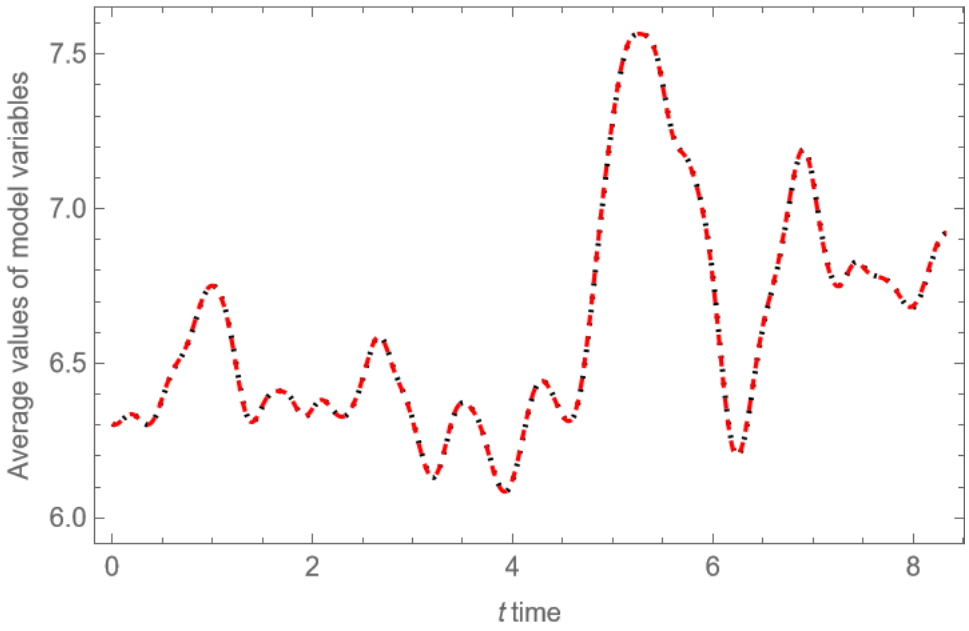

Figure RR3: Comparison of the time evolution of the mean value of the variables ($N = 360$) for the 2005 Lorenz system (Eq. (8) in Figure R3) based on the same initial conditions with time step Δt=1/240 (red dashed curve) and Δt=1/2400 (black dotted curve).

(Major Comment C) Impact of model's configuration and complexity on critical points (equilibrium points)

Based on the linearization theorem, critical points of the Lorenz systems could roughly indicate the local behavior of the solutions. As a result, initial error growth should display a dependence on the equilibrium state. Please consider identifying the appearance of the critical points and perform stability analysis using the Jacobian matrix of the linearized system at each of the critical points.

Below, a simple illustration for the linear stability analysis is provided using the 1996 one-scale model with N = 5. Based on the Figure R5 and Table R1, it is suggested that a larger F may produce a larger eigenvalue (a larger real part of the eigenvalue), suggesting a larger growth rate.
Based on the following preliminary analysis of the one- and two-scale models with the same value of the forcing parameter F, the effective forcing parameter for the two-scale model is smaller, yielding a smaller leading eigenvalue (i.e., a smaller real part of the eigenvalue). This is consistent with the finding that Figures 5 and 6 display larger growth rates (λ) within the one-scale system (e.g., L05-1) than the two-scale system (e.g., L05-2). [Such a finding is supported by the so-called aggregated negative feedback reported by Shen 2014, 2019.]

Consider Eqs. (A2) and (A3). From the nonlinear terms of Eq. (A2) and (A3), we expect that $X_{1,1} = X_{1,2} = X_{1,3} = \cdots X_{1,c}$ and $X_{2,1} = X_{2,2} = \cdots X_{2,c}$ may be a critical point. Here, $X_{1,c}$ and $X_{2,c}$ represent the value of steady state solutions for the slow and fast variables, respectively. From Eq. (A3), we have $X_{2,c} = cX_{1,c}/b$. Plugging the above into Eq. (A2), the right hand side of Eq. (A2) contains two dissipative terms, $-X_{1,n}$ and $-c^2X_{1,c}/b$, yielding $X_{1,c} = bF/(b + c^2) < F$. Namely, the effective forcing for slow variables is weaker, indicating a smaller growth rate within the two-scale model, as compared to the one-scale model.

On the other hand, the above along with Figure R5 and Table R1 only provide a preliminary, qualitative, analysis. The authors may want to further verify or comment the above since the Jacobian matrix for the two-scale system that includes fast variables is larger, as compared to the Jacobian within the corresponding one-scale system.

For example, with the two-scale or three-scale system, the value of parameter "b1" (b1 > 1) determine the (temporal) scale as well as the magnitude of the fast variables. Please provide justifications for the choice of b1 = 10 for the two-scale system but b1 = 1 for the three-scale system. Additionally, within the three-scale system, are nonlinear terms (e.g., c1 and c2 in Eq. A9) applied for coupling the "sub-systems" for the small- and medium-scale variables with the large-scale system? Please comment on the impact of c1 and c2 on system's stability.

$$J_{L96} = \begin{pmatrix} -1 & X_c & 0 & -X_c & 0 \\ 0 & -1 & X_c & 0 & -X_c \\ -X_c & 0 & -1 & X_c & 0 \\ 0 & -X_c & 0 & -1 & X_c \\ X_c & 0 & -X_c & 0 & -1 \end{pmatrix}.$$

Figure R5: A Jacobian matrix for the linearized version of the Lorenz 1996 model with N = 5 from Eq. 3.1 in Figure R1. Here, Xc indicates a critical point solution and is equal to F.

Table R1: An eigenvalue analysis of the Lorenz 1995 one-scale model. The corresponding Jacobian matrix is shown in Figure R5. Here, Xc = F.

Xc	eigenvalues
0.5	-0.4410 + 0.7694i -0.4410 - 0.7694i -1.0000 + 0.0000i -1.5590 + 0.1816i -1.5590 - 0.1816i
1	0.1180 + 1.5388i 0.1180 - 1.5388i -1.0000 + 0.0000i -2.1180 + 0.3633i -2.1180 - 0.3633i
10	10.1803 +15.3884i 10.1803 -15.3884i -1.0000 + 0.0000i -12.1803 + 3.6327i -12.1803 - 3.6327i
20	21.3607 +30.7768i 21.3607 -30.7768i -1.0000 + 0.0000i -23.3607 + 7.2654i -23.3607 - 7.2654i
30	32.5410 +46.1653i 32.5410 - 46.1653i -1.0000 + 0.0000i -34.5410 +10.8981i -34.5410 -10.8981i

Response: While for analytical studies the instability of fixed points (critical points) is certainly of high interest, we are interested in the typical error growth and therefore focus on the Lyapunov exponent on the chaotic attractor. Since the phase space is so high dimensional, we are not even sure that unstable fixed points are embedded in the chaotic attractor or whether they are outside, as they are in the Lorenz 1963 low dimensional model. We therefore calculate the maximal LE numerically in the following way: a reference trajectory (considered the "truth" or verification) and a trajectory which is the numerical solution of the systems with a given error, are repeatedly generated. For this scheme to be meaningful, we have to ensure that the reference trajectory is on the system's attractor and that the repetition of this scheme samples the whole attractor with correct weights (the invariant measure). We solve this issue in the following way: We first integrate the system over ten years (175200 steps), starting from arbitrary initial conditions, and assume that after discarding this transient, the trajectory is on the attractor. We continue to integrate this single trajectory and consider segments of it as reference trajectories for error growth, i.e., the many reference trajectories are simply segments of one very long trajectory, which ensures not only that all these segments are located on the attractor but that in addition, they sample the attractor according to the invariant measure.

Figure RR4 compares the error growth rates of the L05-1 (Eq. (A1) in manuscript), L05-2 (Eq. (A8) in manuscript), and L05-3 (Eq. (A9) in manuscript) systems. In contrast to the reviewer's findings, the figure shows the smallest growth rate for the L05-1 system and the largest for the L05-3 system. We confirm that the effective forcing for slow variables is weaker, indicating a smaller growth rate within the two-scale model, as compared to the one-scale model. However, it should be noted that in Figure RR4 the values of the single-scale system (L05-1) are not compared with the large-scale values of the multi-scale systems (L05-2 and L05-3), but are compared with the total values of the L05-2 and L05-3 systems, where the large-scale and small-scale features are appearing as superimposed features of a single set.

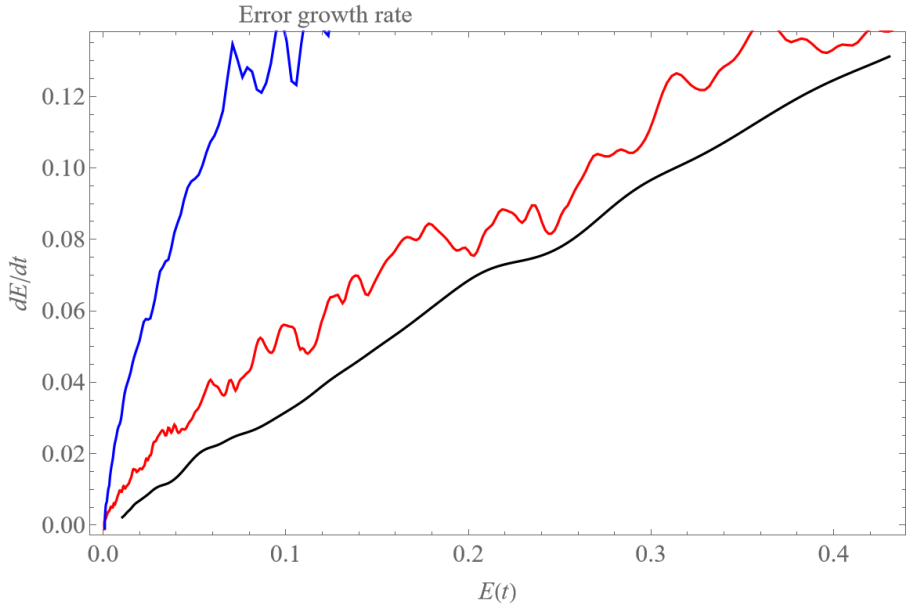

Figure RR4: Initial error growth tendency (rate) dE/dt as a function of the error magnitude E for L05-1 system (Black, Eq. (A1) in manuscript), for L05-2 system (Red, Eq. (A8) in manuscript), and for L05-3 system (Blue, Eq. (A9) in manuscript).

A justification for the use of the L05-2 (Eq. (A8) in manuscript) and L05-3 (Eq. (A9) in manuscript) systems as the "reality" and the L05-1 system as the "model." is presented in the manuscript (Lines 220-228 in revised manuscript):

"This approach is justified by the fact that the L05-2 and L05-3 systems can be viewed as a variant of the L05-1 system:

$$dX_{tot,n}/dt = [X_1, X_1]_{L,n} - X_{1,n} + \tilde{F}_n(t), \tag{12}$$

where $\tilde{F}_n(t) = b^2[X_2, X_2]_{1,n} + c[X_2, X_1]_{1,n} - bX_{2,n} + F$ for the L05-2 system and $\tilde{F}_n(t) = b_1^2[X_2, X_2]_{1,n} + b_2^2[X_3, X_3]_{1,n} + c_1[X_2, X_1]_{1,n} + c_2[X_3, X_2]_{1,n} - b_1X_{2,n} - b_2X_{3,n} + F$ for the L05-3 system are treated as a forcing, which varies in a complicated manner with time. We parameterize these small-scale phenomena contained in $\tilde{F}_n(t)$ by the average value of these phenomena, which is close to zero, and therefore we can write:

$$\langle \tilde{F}_n(t) \rangle \approx F = 15, \tag{13}$$

where $\langle \ldots \rangle$ represents the mean calculated over a long orbit on the L05-2 and L05-3 systems attractors."

The parameters of any multi-level Lorenz's system (L96-2, L05-2, L05-3) should be set so that all levels behave chaotically (the largest Lyapunov exponent of each level is positive) and that all levels have a significant difference in amplitudes and fluctuation rates. For the L-96 system (Eq. (3.1) in FIgure R1), the chaotic behavior is determined by the value of F, and the number of variables N. Lorenz (2005) states that as long as $N \geq 12$ chaos is found when $F > 5$ (for $N = 4$ it is when $F > 12$ and for $N > 6$ when $F > 8$). In cases such as the L96-2 system (Eqs. (1a) and (1b) in Figure R2), where the forcing F acts only on the largest scale, the chaotic behavior of smaller scales is created by coupling. The size of the coupling is cascaded from the largest scale to the smaller ones. Because the values of the largest scale variables are determined by the forcing F, the F value indirectly affects the smaller scales' chaotic behavior and must be chosen large enough to ensure chaotic behavior through coupling for all scales (levels). For the L05-2 system (Eq. (A8)), variables are superposed features of a single set calculated by Eqs (A4) and (A5). In addition to those mentioned above, this procedure affects the chaotic behavior, amplitude, and fluctuation rate of the levels, and the choice of I between 10 and 20 may be optimal (Lorenz, 2005). In order to maintain the required properties of the two scales L05-2 system, Lorenz (2005) chose $N = 960$, $L = 32$, $I = 12$, $F = 15$, $b = 10$, and $c = 2.5$ (**note that for L05-2 and L05-3 systems it is not possible to directly determine the amplitude and fluctuation rate of smaller scales using spatiotemporal scaling factors *b*, because these values are mainly determined by the procedure for expressing variables and the length of the intervals** $[-I, I]$).

For the L05-3 system (Eqs. (A9) – (A12)), it is necessary to specify eight parameters. We tested that the values of coupling coefficients c_1 and c_2 do not affect the L05-3 system compared to the values of other parameters, and therefore for simplification $c_1 = 1$ and $c_2 = 1$. The parameter $F = 15$ is set the same as for other L05 systems. For the medium scale amplitude to be approximately ten times smaller than the large scale amplitude and the small scale amplitude to be approximately ten times smaller than the medium scale amplitude and for the scales to have different oscillation rates, the spatiotemporal scale factors are chosen $b_1 = 1$ and $b_2 = 10$ and interval lengths $I_1 = 20$, and $I_2 = 10$. $N = 360$ turned out to be most suitable for the chaotic behavior of all three levels (found experimentally).

(Major Comment D) Separations of initial and model errors

Based on the linearization theorem, a locally linearized system may represent the local feature of the corresponding nonlinear system (for a hyperbolic critical point). The stability of the linearized system depends on locations of the critical points that depend on model's complexity (i.e., nonlinear terms in the system). Thus, the model complexity (i.e., nonlinear terms) could impact the critical points and thus the growth of the initial errors. As a result, it is not easy to separate the initial errors and model errors. (For example, given the same initial error for a large-scale variable, the time varying difference between two nearby trajectories are different in two different models.)

Response: We fully agree with the comment. We simulate the initial error growth in the same systems (perfect model assumption), and the model error growth with zero initial error (perfect initial conditions assumption). Combination of both is studied in section 3.3 of the manuscript.

(Major Comment E) Validity of error saturation for periodic attractors and coexisting attractors

Earlier studies suggest that the Lorenz 1996 two-scale model could produce nonlinear periodic solutions. In your ensemble runs, have you observed periodic solutions? Can you comment on the validity of error saturation for periodic solutions?

Additionally, recent studies reported the appearance of multistability (for coexisting attractors) within the 1996 model (e.g., Van Kekem and Sterk 2018a,b, 2019; Pelzer et al. 2020). Have you observed multistability in your ensemble runs?

Response: In our research, we focused only on the average value of error growth (over variables and number of runs). We set all the scales through the parameters of the Lorenz systems to behave chaotically (details can be found in Bednar and Kantz (2022)) and the evolution of the average error growth did not show signs of periodic solution or multistability.

Specific Comments:

(Specific Comments 1) Please check consistency in the capitalization of the initial letters of words within a title.

Response: We checked and fixed it. Thank you for pointing this out. (Lines 1-2 in revised manuscript):

"Analysis of model error in forecast errors of ~~E~~extended ~~A~~atmospheric Lorenz' 05 ~~S~~systems and the ECMWF system"

(Specific Comments 2) Lines 45-50, the application of the Lyapunov exponent (LE) is not accurate. A global LE represents a long-term average of "local" growth rates (determined by the separations of two nearby trajectories). Initial separations should remain small. Local growth rates may vary with time. As a result, Eq. (1) with a constant growth rate is valid only for a finite time interval. During different time internals, different growth rates may appear. Note that in addition to one positive LE, solution's boundedness is another important feature that defines a chaotic system.

Response: We have added information about boundedness and validity for a finite time interval. (Lines 47-48 in revised manuscript)

"In low-dimensional bounded chaotic systems with at least one positive Lyapunov exponent, the growth of infinitesimal errors is exponential for a finite time interval, given by a linear time derivative:"

(Specific Comments 3) Lines 45-55, please consider referring to the growth rates in Eqs. (1) and (2) as the exponential growth rate (with a J-shaped curve) and logistic growth rate (with a S-shaped curve), respectively.

Response: We changed the description of Eqs. (1) and (2). (Lines 808-810, 820-822, 833-835, 846-848, 861-863 in revised manuscript).

"the early part of the growth by exponential growth rate dE_{ex} (Eq. (1), green, dashed), exponential growth rate with model error dE_r (Eq. (5), blue, dashed), power law dE_p (Eq. (3), red, dashed) and approximation of the full curve by growth rate of quadratic hypothesis dE_{qu} (Eq. (2), green), growth rate of quadratic hypothesis with model error dE_q (Eq. (6), blue) and extended power law"

"the early part of the growth by exponential growth rate dE_{ex} (Eq. (1), green, dashed), exponential growth rate with model error dE_r (Eq. (5), blue, dashed), power law dE_p (Eq. (3), red, dashed) and approximation of the full curve by growth rate of quadratic hypothesis dE_{qu} (Eq. (2), green), growth rate of hypothesis with model error dE_q (Eq. (6), blue) and extended power law dE_{ep} "

"the early part of the growth by exponential growth rate dE_{ex} (Eq. (1), green, dashed), exponential growth rate with model error dE_r (Eq. (5), blue, dashed), power law dE_p (Eq. (3), red, dashed) and approximation of the full curve by growth rate of quadratic hypothesis dE_{qu} (Eq. (2), green), growth rate of quadratic hypothesis with model error dE_q (Eq. (6), blue) and extended power law"

"black, dot-dashed for $E(0)=0.2$), approximation of the early part of the model growth by exponential growth rate dE_{ex} (Eq. (1), green, dashed), exponential growth rate with model error dE_r (Eq. (5), blue, dashed), power law dE_p (Eq. (3), red, dashed) and approximation of the full curve by growth rate of quadratic hypothesis dE_{qu} (Eq. (2), green), growth rate of quadratic hypothesis"

"black, dot-dashed for $E(0)=0.2$), approximation of the early part of the model growth by exponential growth rate dE_{ex} (Eq. (1), green, dashed), exponential growth rate with model error dE_r (Eq. (5), blue, dashed), power law dE_p (Eq. (3), red, dashed) and approximation of the full curve by growth rate of quadratic hypothesis dE_{qu} (Eq. (2), green), growth rate of quadratic hypothesis"

(Specific Comments 4) Line 80, the term "error growth laws" should be rephrased since they are not necessarily physical laws.

Response: We replaced the term law with the term hypothesis. (Lines 81, 309 in revised manuscript)

"While the above-listed error growth ~~laws~~approximations are supposed to approximate the effectively observed average error"

"numerical error growth curves using the hypotheses ~~or laws~~ Eqs. (1) - (6) and try to identify the most appropriate description. "

(Specific Comments 5) Lines 122, statements are not accurate. Unless additional forcing terms are introduced, improving model's spatial or temporal resolution does not necessarily enhance instability. (Please think of a convergent Taylor series.)

Response: We added to the introduction:

"Buizza (2010), Magnusson and Kallen (2013) or Jacobson (2001) show that improving the model's spatial and temporal resolution will improve the ability to predict, especially for short forecast range (Buizza, 2010). However, the cited studies work with models that do not model small spatiotemporal phenomena (they are parameterized) and whose initial condition error magnitude is larger than the magnitude of these phenomena. We have verified the fact that the high resolution model (that models small scales) is less stable than the low resolution model (that doesn't model small scales) against initial condition errors (Bednar and Kantz, 2022; Budanur and Kantz, 2022), and that therefore the issue of omitting small scales has another facet. Our new approach models and omits small spatiotemporal scales using…"

(Lines 129-135 in revised manuscript)

(Specific Comments 6) Lines 128-130: it is wired that the two-scale system contains large- and small-scale systems while the three-scale system adds a medium scale, in addition to large- and small-scale flows. Any justifications?

Response: It would be more natural to take the L05-2 and L05-1 systems as the model and the L05-3 system as the reality.). A variant where the L05-2 system was used as the model and the L05-3 system as the "reality" was also tested. The resulting model error growth is approximately identical to the previous variant (L05-1 system as the model and L05-3 system as the "reality"). That's why we chose the settings we present. Further, it would be more natural for the L05-2 system to have a small scale comparable to the medium scale of the L05-3 system. However, our intention was to be close to the L05-2 system presented by Lorenz (2005), whose small scale is equivalent to the small scale of our L05-3 system.

(Specific Comments 7) Lines 160-165, have you observed coexisting attractors (e.g., more than one attractors) in your ensemble runs? (e.g., see multistability in Van Kekem and Sterk 2018a,b, 2019; Pelzer et al., 2020).

Response: In our research, we focused only on the average value of error growth (over variables and number of runs) and we did not observe signs of multistability.

(Specific Comments 8) Line 170, does the statement "errors might even shrink in short times" indicates the existence of a stable manifold?

Response: Yes, the Lorenz L05-systems possess rather high dimensional stable manifolds, along which trajectories are attracted towards the attractor. Calculation of the Lyapunov-dimension done by us for L05-2 show this very clearly, the attractor dimension is much smaller than the phase space dimension, where the attractor is the unstable manifold. But the statement on line 170 does not indicate the existence of a stable manifold but the fact that initial perturbations might not point into the locally most unstable direction.

(Specific Comments 9) Lines 194, while N=360 was used in this study, N=960 was appied in Lorenz (2005).

Response: Thank you for pointing this out. The problem is already discussed in comment **(Major Comment B)**.

(Specific Comments 10) line 186, how many time steps for the transfer of error to the small-scale variables?

Response: The error would immediately (one time step) propagate into the small-scale variables.

(Specific Comments 11) Section 3.1, please confirm whether the leading LE in the L05-1 system is larger (smaller) than that in the L05-2 (L05-03) system.

Response: Figure RR4 compares the error growth rates of the L05-1 (Eq. (A1) in manuscript), L05-2 (Eq. (A8) in manuscript) and L05-3 (Eq. (A9) in manuscript) systems. The figure shows the smallest growth rate for the L05-1 system and the largest for the L05-3 system (therefore also for LE). It should be noted that for the L05-2 and L05-3 systems, the error growth rate is scale dependent.

(Specific Comments 12) Line 382-394: The key point that higher resolution model produces better predictability is acceptable. However, it is not clear whether Figure 10 is sufficient to support this point. Please see details in the last specific comment below.

Response: Please see the discussion at: **(Specific Comments 17)**

(Specific Comments 13) Line 656: The statement "Based on the fact that scale-dependent error growth implies an intrinsic predictability limit" is not accurate. A finite growth rate may indicate a limit for practical predictability. By comparison, a finite intrinsic predictability is established by the feature of chaos (e.g., sensitive dependence on initial condition, SDIC; e.g., Shen, Pielke Sr., and Zeng, 2023)

Response: Our statement really refers to the finite intrinsic predictability that is established by the features of chaos. The statement is based on Brisch and Kantz (2019), Bednar and Kantz (2022), and Budanur and Kantz (2022).

(Specific Comments 14) Lines 612 - 623, discussions are duplicated; they are the same as those in Lines 600-611.

Response: We deleted the duplicated part. Thank you for pointing this out.

(Specific Comments 15) Line 715, the parameter "K" should be replaced by "L".

Response: We replaced K by L. Thank you for pointing this out.

(Specific Comments 16) Line 716, Lorenz (2005) did not explicitly suggest the ratio of N/L = 30 nor provide justification for the choice of N = 960 and L = 32.

Response: We assume the requirement for a model to have 5 to 7 main highs and lows that correspond to planetary waves (Rossby waves) and several smaller waves corresponding to synoptic-scale waves, and we follow the text of Lorenz (2005) on the pages 1579 (Fig. RR5) and 1580 (Fig RR6).

Figure 4a shows typical profiles produced by Eq. (8) when $N = 240$ and $F = 10$ for selected values of K. When $K = 2$, there are nearly as many waves as if Eq. (1) had been retained. Increasing K to 4 decreases the number, but there are still too many compared with Fig. 1a. When $K = 8$, one evidently succeeds in producing an acceptable number of major waves, although weaker smaller-amplitude waves are superposed. In drawing the curve I have, as usual, connected the successive values of X_n with straight-line segments, but these are hard to detect. Any other reasonable interpolation procedure would have produced an indistinguishable curve. Increasing K to 16, 32, or 64 lengthens the waves still more, and, evidently, one can produce any wavenumber desired by choosing K judiciously.

Since the ratio N/K is 30 in the third profile, whose dominant wavenumber agrees most closely with Fig. 1a, where $N = 30$, there is a suggestion that the appearance of a profile may depend largely upon N/K. Figure 4b is constructed with $F = 10$ and $N/K = 30$ in each profile, and with N successively doubling from 30 in the leading profile to 960 in the final one. The conjecture seems to be well supported; the profiles in Fig. 4b show little resemblance to any profile in Fig. 4a except the third one.

With $N = 960$ and again with $K = 32 = N/30$ and $F = 10$, Fig. 5a has been constructed in the manner of Fig. 1a; it shows profiles produced by Eq. (8) at 6-h intervals for five days. Again, at least for the five days, the major crests and troughs retain their identities, while minor ones come and go. One can conclude that Model II is ready for some applications for which Model I would have been inadequate.

For Model I the doubling time for small errors, as seen in Fig. 2a, depends strongly upon F, but is nearly independent of N if N is not too small. For Model II, with $K > 1$, it also proves to depend strongly upon F while being nearly independent of N and K if N/K is not too small, but, for a given value of F, it is much smaller when $K > 1$ than when $K = 1$. Thus, for the values used in Fig. 5a, the doubling time is about four days—considerably longer than expected in the atmosphere. It can be restored to a more nearly atmospheric value by increasing F.

Figure 5b is constructed like Fig. 5a, again with $N =$

Figure RR5: Page 1579 in Lorenz (2005).

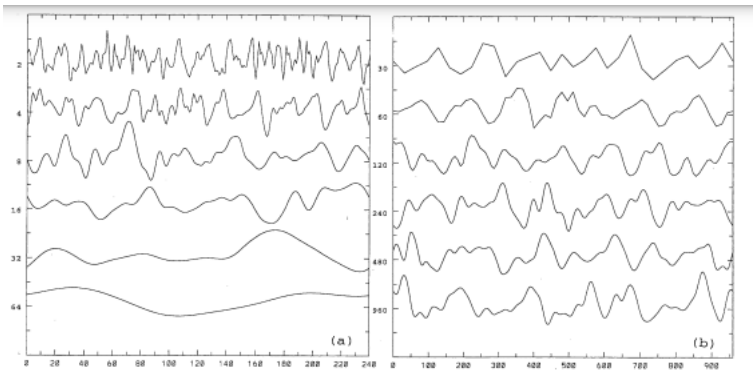

FIG. 4. (a) Profiles of X produced by Eq. (8) with $N = 240$, $F = 10$, and values of K indicated by numbers at left. Scale at bottom is gridpoint number. (b) Profiles of X produced by Eq. (8) with values of N indicated by numbers at left and with $K = N/30$ and $F = 10$. Scale at bottom is gridpoint number for bottom curve.

960 and $K = 32$, but with $F = 15$. There is still a suggestion of six or seven longer waves, but the shorter waves are more in evidence. Note that, with N so large, even these shorter waves are 30 or more grid intervals long—the point-to-point variations are very smooth. The doubling time has been reduced to about two days.

Apparently, in trying to make the curves produced by Model II look like reasonable spatial interpolations of the kind of curve produced by Model I, one must choose between too long a doubling time (smaller F) or unanticipated shorter waves (larger F). The value $F = 15$ is a compromise.

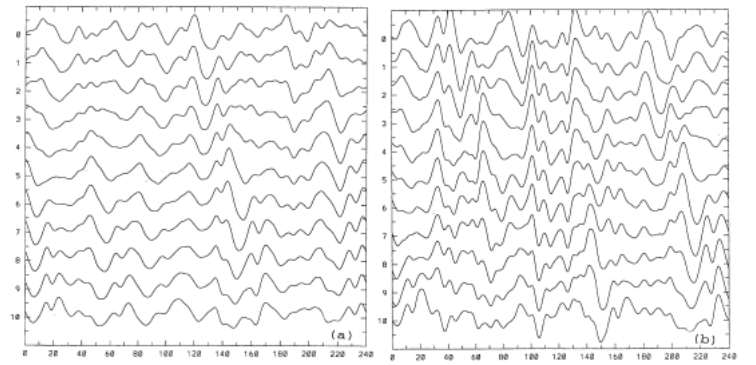

FIG. 5. (a) Profiles of X produced by Eq. (8) with $N = 240$, $K = 8$, and $F = 10$, at 12-h intervals for 5 days. Scale at bottom is gridpoint number. Numbers at left indicate chronological order of profiles. (b) Same as (a) but with $F = 15$.

Figure RR6: Page 1580 in Lorenz (2005).

(Specific Comments 17) page 40, line 870-875, Figure 10. Figure's title and captions are confusing. Since L05-02 and L05-03 systems were used to provide the "ground true" (or reference) for computing errors, these errors do not represent the errors of the L05-02 and L05-03 systems, respectively, the growth of initial errors within the L05-02 or L05-03 system does contribute to the growth of differences of the solutions between the L05-1 and L05-02 (or L05-03) systems.

For a comparison in Figures 5-7, let's simply choose λ+, = 0.33, 0.29, and 0.46 for the L05-1, L05-2, and L05-3 systems, respectively. The comparison of the above selected growth rates produces a consistent finding that larger differences (in error growths) are reported in Figure 10b than in Figure 10a. However, on the other hand, considering differences between the L05-02 and L05-03 systems, the differences may produce the largest growth rates as compared to those in Figure 10a and Figure 10b.

Response: The question under investigation in this paper is whether omitting small scale atmospheric phenomena, which contribute little to the final value, will improve the predictability of the resulting value. In other words, how does the average forecast error growth change in a model where small-scale phenomena are omitted but where model errors are therefore introduced, compared to a model where all phenomena are present but the average forecast error growth is scale-dependent. So if we use L05-02 and L05-03 systems to provide

the "ground true" (or reference) then, when searching for an answer to the research question, it is reasonable to use the results presented in Figure 10.

Figures 5-7 show that, the L05-1 system is a classical chaotic system with the largest Lyapunov exponent of about $\lambda \approx 0.33$ 1/ day. The data of the L05-2 and L05-3 are best approximated by the power law . For a power law: $\lambda_p(E) := \frac{d\ln(E)}{dt} = \frac{\dot{E}}{E} = aE^{-\sigma}$, with an exponent σ and a coefficient $a > 0$, the error growth rate $\lambda(E) \approx \frac{1}{\Delta t}\ln(E(t+\Delta t)/E(t))$ is expected to be a function of the error magnitude E, and is not constant as for classical chaotic systems. For exponential growth (classical chaos) $E_{exp}(t) = E_0 e^{\lambda_{exp} t}$ and for an initial error E_0 going to zero, the time t_{lim} at which the error reaches a limiting value E_{lim}, goes to infinity: $t_{lim} = \frac{\ln E_{lim} - \ln E_0}{\lambda_{exp}} \to \infty \ for \ E_0 \to 0.$

However, a strict predictability limit t_{lim} exists for scale-dependent error growth even when the initial error E_0 vanishes. For a description by a power law dE_p, the predictability limit t_{lim} is: $t = \left(E^b(t) - E_0^b\right)/(a \cdot b) \ \to \ t_{lim} = E_{lim}^b/(a \cdot b) < \infty \ for \ E_0 \to 0.$

It is true that if we show the growth of the model and initial error in Figure 10, this is the initial error of the L05-1 system, but this is consistent with the question under investigation. At the same time, Figure 10 compares the strictly model error growth (no initial error) with the strictly initial error growth (L05-2, L05-3 systems), where the initial error is limiting towards zero and is then a strict predictability limit.

References:

Bednář, H., and Kantz, H.: Prediction error growth in a more realistic atmospheric toy model with three spatiotemporal scales, Geosci. Model Dev., 15, 4147–4161, https://doi.org/10.5194/gmd-15-4147-2022, 2022.

Brisch, J., and Kantz, H.: Power law error growth in multi-hierarchical chaotic system-a dynamical mechanism for finite prediction horizon, New J. Phys., 21, 1–7, https://doi.org/10.1088/1367-2630/ab3b4c, 2019.

Budanur, N. R., Kantz, H.: Scale-dependent error growth in Navier-Stokes simulations, Phys. Rev. E, 106, 1–7, https://doi.org/10.1103/PhysRevE.106.045102, 2022.

Buizza, R.: Horizontal resolution impact on short- and long-range forecast error, Quarterly Journal of the Royal Meteorological Society, 136, 1020–1035, https://doi.org/10.1002/qj.613, 2010.

Jacobson, M. Z.: GATOR-GCMM: 2. A study of day- and nighttime ozone layers aloft, ozone in national parks, and weather during the SARMAP field campaign, J. Geophys. Res., 106, 5403-5420, https://doi.org/10.1029/2000JD900559, 2001.

Lorenz, E. N.: Predictability: a problem partly solved, in: Predictability of Weather and Climate, edited by: Palmer, T., and Hagedorn, R., Cambridge University Press, Cambridge, UK, 1–18, https://doi.org/10.1017/CBO9780511617652.004, 1996.

Lorenz, E. N.: Designing chaotic models, J. Atmos. Sci., 62, 1574–1587, https://doi.org/10.1175/JAS3430.1, 2005.

Magnusson, L., and Kallen, E.: Factors Influencing Skill Improvements in the ECMWF Forecasting System, Mon. Wea. Rev., 141, 3142–3153, https://doi.org/10.1175/MWR-D-12-00318.1, 2013.

(Comment 2) Furthermore, considering the authors' assertion that the proposed 05 system simulates "5 to 7 main highs and lows that correspond to planetary waves (Rossby wave)," it would be advantageous to discuss whether the proposed system, without the Coriolis force, could replicate key features of the Rossby wave, including phase speeds. Historically, experiments such as dishpan experiments aimed to "simulate" weather features, yielding diverse outcomes like chaotic solutions and vacillation (e.g., limit cycle).

Response: Lorenz and Emanuel (1998) showed that the initial wave of the L96 system has a westward phase velocity and an eastward group velocity, which is in agreement with the evolution of Rossby waves. We show a description of the evolution of the incipient waves of the L96 system in Figure RR7. Lorenz and Emanuel (1998) also showed a numerical calculation of the evolution of the L96 system, which is presented in Figure RR8. In the same manner, we present in Figure RR9 the numerical calculation of the evolution of the L05 system for $N = 30$ (left column) and $N = 360$ (right column). From Figures RR8 and RR9, we can see the agreement and confirmation of the theoretical calculation.

$$dX_j/dt = (X_{j+1} - X_{j-2})X_{j-1} - X_j + F, \qquad (1)$$

As with typical nonlinear systems, any solutions found analytically are likely to be rather specialized, but certain properties may be deduced without solving the equations at all. First, if a bar ($\overline{\ }$) over a quantity denotes an average over all values of j and over a long enough time to make average time derivatives negligibly small, it follows from multiplying (1) by X_j and averaging that

$$\overline{X^2} = F\overline{X}, \qquad (2)$$

whence

$$\overline{X^2} - \overline{X}^2 = \overline{X}(F - \overline{X}). \qquad (3)$$

Since the variance $\sigma^2 = \overline{X^2} - \overline{X}^2$ of X is nonnegative, it follows from (3) that the mean $\overline{X}$ of X lies in the interval $[0, F]$, whence the standard deviation σ lies in the interval $[0, F/2]$. In the obvious steady solution where $X_j = F$ for each j, $\overline{X} = F$ and $\sigma = 0$.

Small perturbations x_j about the steady solution obey the equation

$$dx_j/dt = F(x_{j+1} - x_{j-2}) - x_j. \qquad (4)$$

If we let

$$x_j = \sum_k p_k \exp(ikj), \qquad (5)$$

$$dp_k/dt = [(e^{ik} - e^{-2ik})F - 1]p_k. \qquad (6)$$

The steady state is therefore unstable if $(\cos k - \cos 2k)F > 1$ for some k. The factor $\cos k - \cos 2k$ assumes its maximum positive value 9/8 when $\cos k = 1/4$; thus the steady solution becomes unstable with respect to waves of length $2\pi/\cos^{-1}(1/4) = 4.77$ zones when F exceeds 8/9. Since the actual number of zones in a wave must be a divisor of J, we find, for $J = 40$, that waves of five zone lengths, or wavenumber 8, will begin to grow when F exceeds $(\cos 2\pi/5 - \cos 4\pi/5)^{-1} = (4/5)^{1/2} = 0.894$, a value only slightly exceeding 8/9.

The incipient waves will move with velocity $c = -(\sin k + \sin 2k)(F/k)$, which, when $k = 2\pi/5$ and $F = 0.894$, equals -1.09; thus they will drift westward. The group velocity $c_g = -(\cos k + 2\cos 2k)F$, on the other hand, equals $+1.17$, implying eastward propagation of regions of enhanced activity; this feature has been an important consideration in our decision to adopt the model.

Figure RR7: Description of the evolution of the incipient waves of the L96 system (Pages 400-401 in (Lorenz and Emanuel, 1998)).

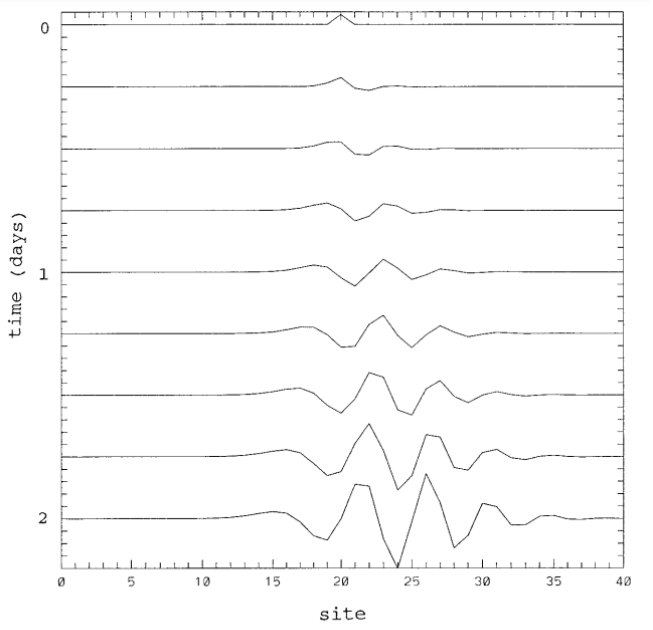

FIG. 1. Longitudinal profiles of X_j at 6-h intervals, as determined by Eq. (1) with $N = 40$ and $F = 8.0$, when initially $X_{20} = F + 0.008$ and $X_j = F$ when $j \neq 20$. On horizontal portion of each curve, $X_j = F$. Interval between successive short marks at left and right is 0.01 units.

Figure RR8: Numerical documentation of the evolution of the incipient waves of the L96 system (Page 401 in (Lorenz and Emanuel, 1998)).

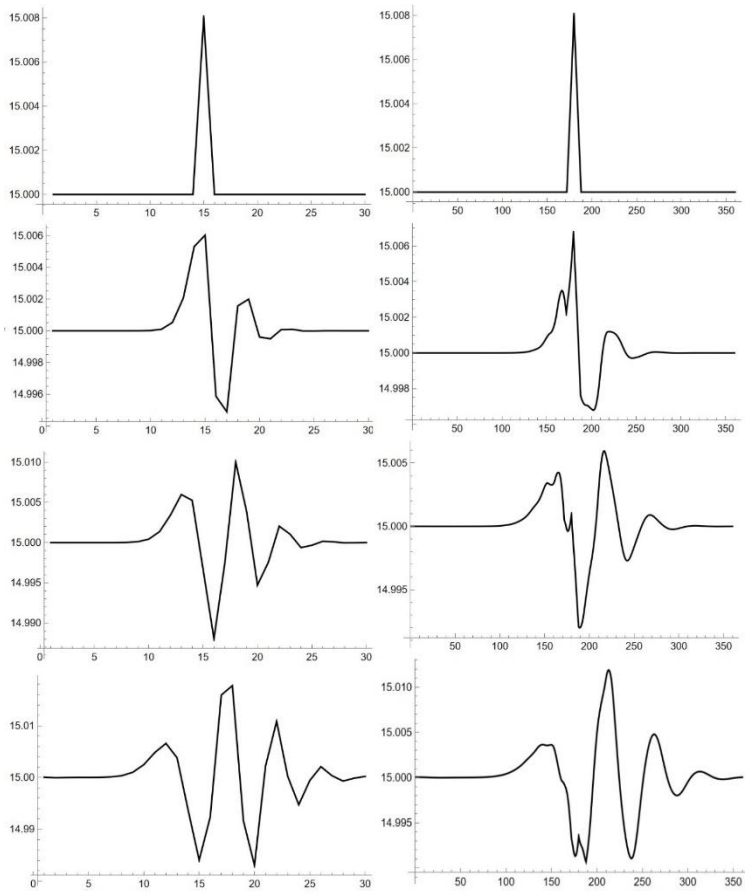

Figure RR9: Numerical documentation of the evolution of the incipient waves of the L05 system for $N = 30$ (left column) and $N = 360$ (right column). Longitudinal profiles of Xj at 6-h intervals, with $F = 15$, when initially $X_{15} = F + 0.008$ or $X_{180} = F + 0.008$ and $X_j = F$ when $j \neq 15$ or $j \neq 180$.

(Comment 3) This study extends from the authors' previous research. The reviewer acknowledges the related efforts. However, after examining their earlier studies, the reviewer proposes the following:

(Comment 3.1) Document and report the calculation of Lyapunov exponents (LEs) within the proposed 05 system. For instance, employing the 1963 model with common parameters, the largest LE (LE1) is 0.906, as exemplified in the link provided (https://sprott.physics.wisc.edu/chaos/lorenzle.htm). This task holds significant importance.

Response: Table RR1 shows the values of the largest Lyapunov exponent *LE1* of the L05 system described by Eq. (8) in Figure R3 for selected numbers of variables *N* (*F* = 15, time unit = 5 days) calculated by Sprott's (2006) method.

N	*LE1*
30	0.70
60	0.29
90	0.35
120	0.32
150	0.33
360	0.33
960	0.33

Table RR1: Values of the largest Lyapunov exponent *LE1* for selected numbers of variables *N* in the 2005 Lorenz system (Eq. (8) in Figure R3, *F* = 15, time unit = 5 days).

We also calculated the largest Lyapunov exponent *LE1* in the L05-3 system (three scales, N = 390, *F* = 15, time unit = 5 days) using the method of Sprott (2006). We determined the maximal Lyapunov exponents in all four cases and find the values *LE1* = 2.5 $(\text{day})^{-1}$ for overall and small scale and *LE1* = 2 $(\text{day})^{-1}$ for medium and small scale. The similarity of the values for all levels indicates that they are coupled, so that the maximal Lyapunov exponent when calculated in the double limit $E_0 \to 0$ and $t \to \infty$ shows up in arbitrary subsystems. The evolution of the errors *E* can always be studied in a way to see the largest exponent of the system (done here), but also in a way to see a value which would be the exponent of the corresponding sub-system if one were able to isolate this, but this cannot be calculated using standard methods for calculating the largest Lyapunov exponent.

(Comment 3.2) Develop the error growth model, e.g., dE/dt = sigma E (1 - E/Es), and furnish a mathematical expression for sigma and LE1 of the proposed system. It should be noted that the long-time average of (1/E dE/dt) is not precisely equal to sigma.

For a chaotic L05 system with average initial error growth E(t), the largest Lyapunov exponent is defined as: $LE1 = \lim_{t\to\infty}\lim_{\epsilon\to 0}(1/t)\ln(E(t)/\epsilon)$. This exponential growth is associated with single scale systems, infinitesimal initial error ϵ, and the early part of the error growth. For a not infinitesimally small initial error and the entire evolution of the error, Lorenz (1982) defined the quadratic hypothesis: dE/dt = sigma E (1 - E/Es). The error growth rate, for comparison with LE1, can be determined as: 1/E dE/dt = sigma (1 - E/Es). Thus, the sigma determines the value at the beginning of the decrease in the error growth rate (dE/dt = sigma*E(t) in the limit E/Es << 1 with sigma ≈ LE1, Figure RR10). Sigma is an approximation of LE1, which is biased by the error due to the approximation of the data, not the infinitesimal initial error, and the use of data from the entire development period.

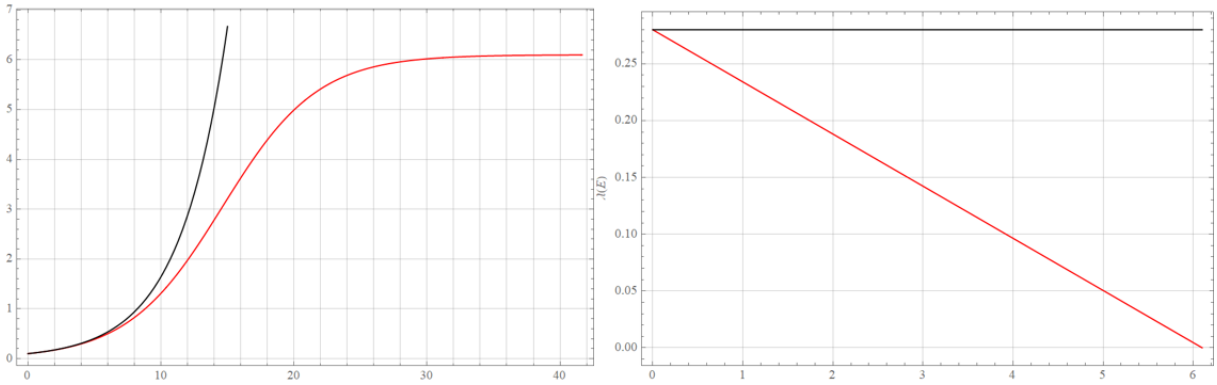

Figure RR10: Exponential growth E(t) = E(0) exp(sigma*t) (left figure, black curve) and sigma determined from 1/E(dE/dt) of E(t) = E(0) exp(sigma*t) as a function of E (right figure, black curve). Growth of E(t) determined from the quadratic hypothesis (left figure, red curve) and linear decline determined from 1/E dE/dt = sigma (1 - E/Es) as a function of E (right figure, red curve).

References:

Lorenz, E. N.: Atmospheric predictability experiments with a large numerical model, Tellus, 34, 505–513, https://doi.org/10.1111/j.2153-3490.1982.tb01839.x 1982.

Lorenz, E. N. and Emanuel, K. A.: Optimal sites for supplementary weather observations: Simulation with a small model, Journal of the Atmospheric Sciences, 55, 399–414, https://doi.org/10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2 1998.

Sprott, J. C.: Chaos and Time-series Analysis, Oxford University Press, New York, USA, 2006.