A review report for the manuscript, titled "***Analysis of model error in forecast errors of extended atmospheric Lorenz' 05 systems and the ECMWF system*** by Bednář and Kantz"

**Review Summary:**

The study applies the Lorenz 2005 models to investigate the forecast error growth in atmospheric predictions, attributing it to initial and model errors. It focuses on the impact of small-scale phenomena on predictability, questioning whether omitting them would enhance forecast accuracy. Using the Lorenz 2005 one-, two-, and three-scale systems, the research reveals that excluding small spatiotemporal scales diminishes predictability more than modeling them. Contrary to expectations, omitting phenomena does not improve predictability; instead, it results in increased model error. The study proposes a hypothesis explaining this behavior, linking model error to the differences between systems at each time step. Fit parameters are used to compare the hypothesis with approximations of average forecast error growth, interpreting them in the context of model error. The findings were applied to the ECMWF system to reveal the reduction of model error from 1987 to 2011 based on the hypothesis, despite a concurrent growth in system instability related to initial condition errors.

This study is interesting and has the potential to improve our understanding of predicting capabilities within idealized and real-world models. However, there are some major issues that require further clarification. The reviewer suggests major revisions, and specific comments are provided below.

**Major Comments:**

Most of Lorenz's models were developed to effectively illustrate the chaotic and unstable nature of weather and climate and/or estimate the growth rates of the systems (which include the natural system in Lorenz 1969a and/or numerical models). While some of Lorenz's models (e.g., the Lorenz 1963 model) have been extensively studied by researchers in various fields for over 50 years, the Lorenz 1996/2005 model is still relatively young and the 1996 and 2005 versions are not exactly the same. Below, major features of the Lorenz 1996 and 2005 models are provided.

- The 1996 and 2005 models were developed to illustrate the growth of errors for chaotic responses that contain one or two scales.
- However, these models were not derived from physics-based systems (i.e., partial differential equations).
- These models contain constant coefficients for nonlinear terms, dissipative terms, and forcing terms. Thus, these models lack some realistic features (e.g., differences between ocean and land; Lorenz and Emanuel 1998).

Regarding the analysis of errors, the following assumptions were applied in most studies:

- Saturation for error growth,

- Quadratic hypothesis for error growth, and
- Monostability for single type of chaotic solutions (in contrast to multistability for coexisting attractors).

Thus, by considering the above features, assumptions, and the following quote from Lorenz himself:

"

> *I have not developed anything resembling a general theory of model design. What successes I have enjoyed have resulted from trial and error, but not, however, from random trial and error. Each satisfactory attempt has been guided by the detailed analysis of previous failures. I make no claim to have discovered the ideal equations (Lorenz 2005),*

"

findings obtained using these idealized models should be explained with caution. For example, what can we learn when the 2005 model produces a comparable growth rate (i.e., doubling time), as compared to the real world model?

The 1996 versions of the models, including one-scale and two-scale models, were first proposed in a report (Lorenz, 1996). Later, Prof. Lorenz made an attempt to propose improved models in 2005. While the 1996 and 2005 versions of the models include the same one-scale model, they contain different two-scale models. The 2005 version of the two-scale model has a shorter history. As a result, analysis of stability within the 2005 two-scale models and sensitivities of findings on parameters should be explored further to support their conclusion.

Please provide discussions and/or responses to clarify or address the following:

(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 ("$b^2$", "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 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?

(B) Dependence of findings on temporal spacing (i.e., $\Delta t$) 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 $\Delta t$ and N.

Similarly, the concept of computational chaos (Lorenz 1989) also suggests the importance of wisely choosing $\Delta t$. 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., $\Delta t$) can be shown by deriving the Logistic map from the Logistic differential equation (Shen et al. 2023).

In this study, $\Delta t$ is 1/240 ~ $4.2 \times 10^{-3}$ 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.

(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 ( $\lambda$) 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 $X_{1,1} = X_{1,2} = X_{1,3} = \cdots X_{1c}$ and $X_{2,1} = X_{2,2} = \cdots X_{2c}$ may be a critical point. Here, $X_{1c}$ and $X_{2c}$ represent the value of steady state solutions for the slow and fast variables, respectively. From Eq. (A3), we have $X_{2c} = cX_{1c}/b$. Plugging the above into Eq. (A2), the right hand side of Eq. (A2) contains two dissipative terms, $-X_{1,n}$ and $-c^2 X_{1c}/b$, yielding $X_{1c} = bF/(b + c^2) < F$. 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.

(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.)

(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?

**Specific Comments:**

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

(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.

(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.

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

(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.)

The impact of additional dissipative terms and/or additional heating term has been previously examined using the Lorenz 1963 model (e.g., Shen 2014, 2015, 2019).

(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?

(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)

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

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

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

(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.

(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.

(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).

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

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

(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.

(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 $\lambda_{ex} = 0.33$, 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.

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

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:

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

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

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.

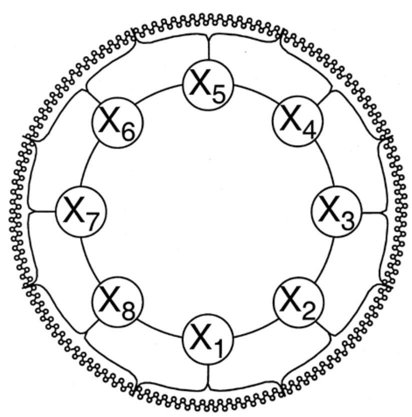


Figure 1. Schematic illustration of the Lorenz ’96 system (Eq. (1)) with $K = 8$ resolved variables $X_k$ (large circles), each associated with $J = 32$ unresolved variables $Y_j$ (unlabeled small circles grouped according to the $X$ variable to which they belong), so that there are $JK = 256$ $Y$ variables 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

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

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

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

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

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

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

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.

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

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.

**References:**


- Lin, S.-J., B.-W. Shen, W. P. Putman, J.-D. Chern, 2003: Application of the high-resolution finite-volume NASA/NCAR Climate Model for Medium-Range Weather Prediction Experiments. EGS - AGU - EUG Joint Assembly, Nice, France, 6 - 11 April 2003
- Lin, S.-J., 2004: A vertically Lagrangian finite-volume dynamical core for global models, Mon. Weather Rev.,132, 2293–2307.
- Lorenz, E.N., 1969b:The predictability of a flow which possesses many scales of motion. *Tellus*, *21*, 289–307.
- Lorenz, E.N., 1969a: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, *26*, 636–646.
- Lorenz, E. N., 1989: Computational chaos: a prelude to computational instability. Physica, 35D, 299-317.
- Lorenz, E. N., 1996: Predictability: A problem partly solved. Proc. Seminar on Predictability, Vol. 1, ECMWF, Reading, Berkshire, UK, 1–18.
- Lorenz, E. N.,, and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. J. Atmos. Sci., 55, 399–414.
- Lorenz, E., 2005a: Designing chaotic models. J. Atmos. Sci., 62, 1574-1587.
- Lorenz, E., 2006: Regimes in simple systems. J. Atmos. Sci., 63, 2056–2073.
- Pelzer, Anouk F. G., and Alef E. Sterk. 2020. "Finite Cascades of Pitchfork Bifurcations and Multistability in Generalized Lorenz-96 Models" Mathematical and Computational Applications 25, no. 4: 78. https://doi.org/10.3390/mca25040078
- Shen, B.-W., 2019: Aggregated Negative Feedback in a Generalized Lorenz Model. International Journal of Bifurcation and Chaos, Vol. 29, No. 3 (2019) 1950037 (20 pages). https://doi.org/10.1142/S0218127419500378
- Shen, B.-W., 2015: Nonlinear Feedback in a Six-dimensional Lorenz Model. Impact of an additional heating term. Nonlin. Processes Geophys., 22, 749-764, doi:10.5194/npg-22-749-2015, 2015. (link) (pdf)

- Shen, B.-W., 2014: Nonlinear Feedback in a Five-dimensional Lorenz Model. J. of Atmos. Sci., 71, 1701–1723. doi:http://dx.doi.org/10.1175/JAS-D-13-0223.1
- Shen, B.-W., R. A. Pielke Sr., and X. Zeng 2023: The 50th Anniversary of the Metaphorical Butterfly Effect since Lorenz (1972): Special Issue on Multistability, Multiscale Predictability, and Sensitivity in Numerical Models. [Editorial] Atmosphere 2023, 14(8), 1279; https://doi.org/10.3390/atmos14081279
- Shen, B.-W., W.-K. Tao, and B. Green, 2011: Coupling Advanced Modeling and Visualization to Improve High-Impact Tropical Weather Prediction(CAMVis), IEEE Computing in Science and Engineering (CiSE), vol. 13, no. 5, pp. 56-67, Sep./Oct. 2011, doi:10.1109/MCSE.2010.141
- Shen, B.-W., R. Atlas, O. Oreale, S.-J Lin, J.-D. Chern, J. Chang, C. Henze,and J.-L. Li, 2006b: Hurricane Forecasts with a Global Mesoscale-Resolving Model: Preliminary Results with Hurricane Katrina(2005). Geophys. Res. Lett., L13813, doi:10.1029/2006GL026143.
- Van Kekem, D.; Sterk, A. Travelling waves and their bifurcations in the Lorenz-96 model. Phys. D Nonlinear Phenom. 2018a, 367, 38–60.
- Van Kekem, D.; Sterk, A. Wave propagation in the Lorenz-96 model. Nonlinear Process. Geophys. 2018b, 25, 301–314.
- Van Kekem, D.; Sterk, A. Symmetries in the Lorenz-96 model. Int. J. Bifurc. Chaos 2019, 29, 1950008.
- Tao, W.-K. 2003. Goddard Cumulus Ensemble (GCE) Model: Application for Understanding Precipitation Processes Meteorological Monographs 29 (51): 107-107 [10.1175/0065-9401(2003)029<0107:CGCEGM>2.0.CO;2]
- Tao, W.-K. Tao, D. Anderson, J. Chern, J. Entin, A. Hou, P. Houser, R. Kakar, S. Lang, W. Lau, C. Peters-Lidard, X. Li, T. Matsui, M. Rienecker, M. R. Schoeberl, B.-W. Shen, J. J. Shi, and X. Zeng, 2009: A Goddard Multi-Scale Modeling System with Unified Physics. Special Issue dedicated to The 1st International Conference on From Desert to Monsoons. Ann. Geophys., 27, 3055-3064
- Wilks, D., 2005: Effects of stochastic parametrizations in the Lorenz '96 system. Q. J. Roy. Meteor. 335 Soc., 131 (606), 389–407, doi:10.1256/qj.04.03.