

• Original Paper •

The Relationship between Deterministic and Ensemble Mean Forecast Errors Revealed by Global and Local Attractor Radii

Jie FENG¹, Jianping LI^{*2,3}, Jing ZHANG⁴, Deqiang LIU^{5,6}, and Ruiqiang DING^{7,8}

¹*School of Meteorology, University of Oklahoma, Norman, OK 73072, USA*

²*College of Global Change and Earth System Science (GCESS), Beijing Normal University, Beijing 100875, China*

³*Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China*

⁴*Cooperative Institute for Research in the Atmosphere, GSD/ESRL/OAR/NOAA, Boulder, CO 80305, USA*

⁵*Fujian Meteorological Observatory, Fuzhou 350001, China*

⁶*Wuyishan National Park Meteorological Observatory, Wuyishan 354306, China*

⁷*State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China*

⁸*Plateau Atmosphere and Environment Key Laboratory of Sichuan Province, Chengdu University of Information Technology, Chengdu 610225, China*

(Received 6 June 2018; revised 7 September 2018; accepted 10 October 2018)

ABSTRACT

It has been demonstrated that ensemble mean forecasts, in the context of the sample mean, have higher forecasting skill than deterministic (or single) forecasts. However, few studies have focused on quantifying the relationship between their forecast errors, especially in individual prediction cases. Clarification of the characteristics of deterministic and ensemble mean forecasts from the perspective of attractors of dynamical systems has also rarely been involved. In this paper, two attractor statistics—namely, the global and local attractor radii (GAR and LAR, respectively)—are applied to reveal the relationship between deterministic and ensemble mean forecast errors. The practical forecast experiments are implemented in a perfect model scenario with the Lorenz96 model as the numerical results for verification. The sample mean errors of deterministic and ensemble mean forecasts can be expressed by GAR and LAR, respectively, and their ratio is found to approach $\sqrt{2}$ with lead time. Meanwhile, the LAR can provide the expected ratio of the ensemble mean and deterministic forecast errors in individual cases.

Key words: attractor radius, ensemble forecasting, ensemble mean, forecast error saturation

Citation: Feng, J., J. P. Li, J. Zhang, D. Q. Liu, and R. Q. Ding, 2019: The relationship between deterministic and ensemble mean forecast errors revealed by global and local attractor radii. *Adv. Atmos. Sci.*, **36**(3), 271–278, <https://doi.org/10.1007/s00376-018-8123-5>.

1. Introduction

The atmospheric model, as a nonlinear chaotic system, shows sensitivity to initial and model related errors. In other words, any arbitrarily small errors in initial conditions will grow with time, finally causing a loss of most forecast information (Thompson, 1957; Lorenz, 1963, 1965). The deterministic (or single) forecast, as the most widely used product in weather forecasting, can provide useful and informative prediction within a certain range of predictability limit, but is unable to give quantitative estimation of the prediction

reliability. Instead, ensemble forecasting is a feasible approach to supply quantitative reliability information for forecasts in the form of probability (Leith, 1974; Toth and Kalnay, 1993, 1997; Molteni et al., 1996). Another general advantage of ensemble forecasting over deterministic forecasts is the forecast errors can be efficiently reduced by nonlinear filtering, in which the arithmetic mean of an ensemble of forecasts is taken (Leith, 1974; Szunyogh and Toth, 2002).

Over the past several decades, numerous studies have demonstrated the overall higher forecasting skill of the ensemble mean over that of deterministic forecasts by using numerical models with different degrees of complexity (Houtekamer and Derome, 1995; Toth and Kalnay, 1997; Buizza et al., 1999; Wang and Bishop, 2003; Wei et al.,

* Corresponding author: Jianping LI
Email: ljp@bnu.edu.cn

2008; Zheng et al., 2009; Feng et al., 2014; Duan and Huo, 2016). However, quantitative estimations and comparisons of sample-mean deterministic and ensemble mean forecast errors are subject to sampling errors as a result of limited numbers of forecast samples and ensemble members, especially in complex numerical weather prediction models. It is even more challenging to compare the deterministic and ensemble mean forecasting skill for specific weather and climate events, due to the sampling uncertainties from the day-to-day variation of the underlying flow (Toth and Kalnay, 1997; Corazza et al., 2003).

The forecasts and the corresponding verifying references (generally the analysis states) are the evolving states of the model and reference attractors, respectively. Therefore, the forecast errors essentially are the distances between states in attractor space. Li et al. (2018) proposed two statistics, i.e., the global and local attractor radii (GAR and LAR, respectively), with regard to the average distances between states on attractors. GAR measures the average distance between two randomly selected states on an attractor, while LAR quantifies the average distance of all states on the attractor from a given state. For complex nonlinear dynamical systems, e.g., the atmosphere, GAR and LAR can be estimated simply using a long time series of their observed states. Moreover, GAR is found to be a more accurate criterion to measure the predictability limit than the traditional saturated value of the sample mean deterministic forecast errors. The latter, due to model errors, usually overestimates the actual error size that totally chaotic forecasts should have on average (Li and Ding, 2015; Li et al., 2018). In our study, GAR and LAR will be further used to interpret the differences of deterministic and ensemble mean forecast errors in both sample-mean and single-case contexts without running the numerical forecasts. It is expected to supply a reference for the verification and assessment of the skill of deterministic and ensemble mean forecasts.

The paper is organized as follows: Section 2 briefly introduces the definitions of GAR and LAR used in this study and the relevant theories. The experimental setup is presented in section 3. Section 4 displays and analyzes the roles of the attractor statistics in interpreting the relationship between deterministic and ensemble mean forecast errors. A discussion and conclusions are provided in section 5.

2. Definitions of GAR and LAR

The definitions of GAR and LAR are based on the premise that a compact attractor has an invariant probability density function and marginal density function. Following Li et al. (2018), consider $\mathbf{x} = (x_1, x_2, \dots, x_n)$ to be the state vector on a compact attractor \mathcal{A} . $R_{L,i}$ is the LAR of one given state \mathbf{x}_i on attractor \mathcal{A} , defined by

$$R_{L,i} = R_L(\mathbf{x}_i) = \sqrt{E(\|\mathbf{x}_i - \mathbf{x}\|^2)}, \quad \mathbf{x}_i, \mathbf{x} \in \mathcal{A}, \quad (1)$$

where $E(\cdot)$ represents the mathematical expectation and $\|\cdot\|$ is the L_2 norm of a vector. Geometrically, LAR measures the

average root-mean-square (RMS) distance of all states on the attractor from a given state. Assuming that \mathcal{A} does not vary with time, for a specific state \mathbf{x}_i , $R_{L,i}$ is an invariant quantity. Particularly, if \mathbf{x}_i is chosen to be the mean state \mathbf{x}_E of \mathcal{A} , R_E is calculated by $\sqrt{E(\|\mathbf{x}_E - \mathbf{x}\|^2)}$ and is defined as the attractor radius. It has the same form as the standard deviation (SD) in statistics, measuring the variability of a variable.

Theorem 1: Let d_i and d_j denote the RMS distances of two states \mathbf{x}_i and \mathbf{x}_j on \mathcal{A} from the mean state \mathbf{x}_E . Let $R_{L,i}$ and $R_{L,j}$ represent the LAR of \mathbf{x}_i and \mathbf{x}_j , respectively. Then, they satisfy the following relationship:

$$R_{L,i} > R_{L,j}, \text{ if } d_i > d_j. \quad (2)$$

This means that the minimum value of LAR is exactly the attractor radius. The proof of Theorem 1 can be referred to in the Appendix.

The RMS of LARs over all states on \mathcal{A} is defined as the GAR:

$$R_G = \sqrt{E(R_L^2)} = \sqrt{E(\|\mathbf{x} - \mathbf{y}\|^2)}, \quad (3)$$

where \mathbf{x} and \mathbf{y} are two randomly selected state vectors from \mathcal{A} . GAR is an estimate of the average RMS distance between any two states on the same attractor space.

Theorem 2: A constant proportional relationship between R_G and attractor radius R_E of a compact attractor \mathcal{A} exists as:

$$R_G = \sqrt{2}R_E. \quad (4)$$

The two statistics, GAR and LAR, and their relevant theorems will be applied to the quantitative estimation of deterministic and ensemble mean forecast errors.

3. Experimental setup

In our experiments, the simple Lorenz96 model (Lorenz, 1996) is used so that a large sample of ensemble and forecasts can be generated with low computational cost to significantly reduce the sampling noise of the estimation of forecast errors. The Lorenz96 model is a 40-variable model and has been widely used to investigate theorems and methods of ensemble prediction and data assimilation (e.g., Lorenz and Emanuel, 1998; Ott et al., 2004; Basnarkov and Kocarev, 2012; Feng et al., 2014). The model is given by:

$$dx_k/dt = (x_{k+1} - x_{k-2})x_{k-1} - x_k + F, \quad (5)$$

where x_k ($k = 1, 2, \dots, 40$) represents the state variable and F is a forcing constant. It assumes that $x_{-1} = x_{39}$, $x_0 = x_{40}$ and $x_{41} = x_1$. In the case of $F = 8$, the model has chaotic behavior. The solutions can be solved by a fourth-order Runge–Kutta scheme with a time step of 0.05 time units (tu).

After an initial spin-up stage of 1000 tu, the model is naturally run for a sufficiently long time (10^4 tu, i.e., 2×10^5 time steps) to generate the true states used as references for the forecasts. There are a total of 2×10^5 cases initiated from each true state. If the initial true state is denoted by \mathbf{x}_t , the initial analysis state \mathbf{x}_a is given by superposing analysis errors $\boldsymbol{\delta}$ on \mathbf{x}_t :

$$\mathbf{x}_a = \mathbf{x}_t + \boldsymbol{\delta}. \quad (6)$$

For simplicity, each element of the analysis error δ is arbitrarily generated from the Gaussian distribution with expectation 0 and SD 1. The RMS size of δ is then rescaled to 0.1, which is about 3% of the climatic SD of x_t (3.63). Each ensemble perturbation is generated with the same approach as the analysis error but using different realizations of noise and the RMS size of each perturbation is rescaled to 0.1 as well. There are a total of 2.5×10^5 ensemble perturbations produced in each case and added and subtracted from the analysis x_a to generate $N = 5 \times 10^5$ initial ensemble members (2.5×10^5 pairs), making their mean still equal to x_a . The deterministic and ensemble forecasts in each case are derived by integrating analysis states and initial ensemble members for 10 tu using the same model generating the truth (i.e., perfect model scenario).

Although the analyses in our study are not generated through generally used data assimilation approaches, the initial ensemble perturbations have the same probability distribution as the analysis errors and thus are expected to optimally sample the analysis errors. Moreover, the ensemble member number (5×10^5) is significantly larger than the model dimension (40). The above two designs eliminate the possible effects from suboptimal initial ensemble members and a limited number of ensembles on the ensemble mean skill.

4. Results

4.1. GAR and LAR of the Lorenz96 model

Due to the ergodicity of attractors, the evolving states of chaotic dynamical systems have stable probability to visit different regions of the attractor in the long run (Farmer et al., 1983; Eckmann and Ruelle, 1985; Zou et al., 1985; Li and Chou, 1997). Initially, the probability distribution of the attractor of the Lorenz96 system is given to display the long-term behavior of the system. Figure 1 shows the probability distribution of variable x_1 (the selection of x_i has no effect on the results, because of their homogeneous properties in the Lorenz96 model). It is evident that the probability distribution of variable x_1 tends to be invariant with the evolving time increased to sufficiently long (2.5×10^6 tu here). The mean value and the SD of the attractor are 2.22 and 3.63, respectively.

Figure 2 shows the variation of LAR (red solid line) as a function of the value of x_1 calculated with a 2.5×10^6 tu time series. The probability distribution of the system (black solid line) is also given as a reference. It is found that LAR is dependent on the specific state on the attractor. The states with longer distances to the mean state have smaller probability to occur and larger LAR, as in Theorem 1. When x_1 moves to the mean state, the minimal value of LAR—namely, the attractor radius R_E —is exactly reached and equal to the SD (3.63). Additionally, the R_G of variable x_1 (5.13), calculated through the RMS of R_L over all given states on the attractor, is exactly $\sqrt{2}$ times the R_E , as revealed by Theorem 2.

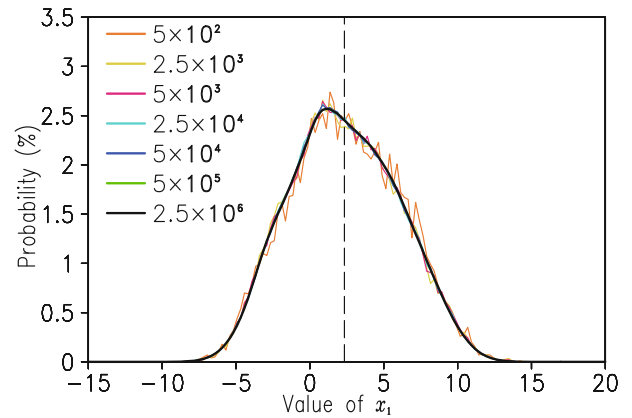


Fig. 1. Probability (%) distribution of variable x_1 of the Lorenz96 model over the attractor set with different lengths of time series (5×10^2 , 2.5×10^3 , 5×10^3 , 2.5×10^4 , 5×10^4 , 5×10^5 , 2.5×10^6 tu, in order).

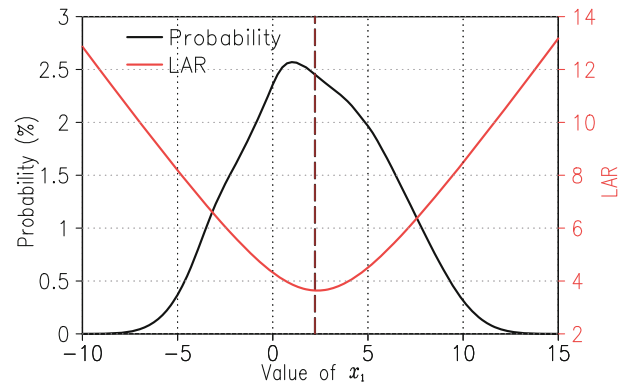


Fig. 2. Variation of LAR (red solid line) as a function of the value of x_1 and the probability distribution of variable x_1 (black solid line). The x_1 value with the lowest LAR (red dashed line) is exactly the same as the mean state (black dashed line) 2.22.

4.2. Evolution of ensemble mean and deterministic forecast states

The differences between deterministic and ensemble mean forecast errors are essentially associated with their differing forecast states. Therefore, the statistical characteristics of deterministic and ensemble mean forecast states are analyzed before comparing their forecast errors. Each panel of Fig. 3 illustrates the probability distribution of the deterministic (black line) and ensemble mean (blue line) forecast states over all cases at the same lead time. It shows that the probability distribution of deterministic forecasts is always consistent with that of the reference (red line) from 0.5 to 6 tu, since they are from the same attractor. In contrast, the probability distribution of ensemble mean states appears to have a narrower range and a higher peak as time increases. In other words, the ensemble mean forecasts tend to, on the whole, move toward the climatic mean value (2.22) with lead time because of the nonlinear smoothing effect of the arithmetic mean of the forecast ensemble (Toth and Kalnay, 1997). Finally, when all

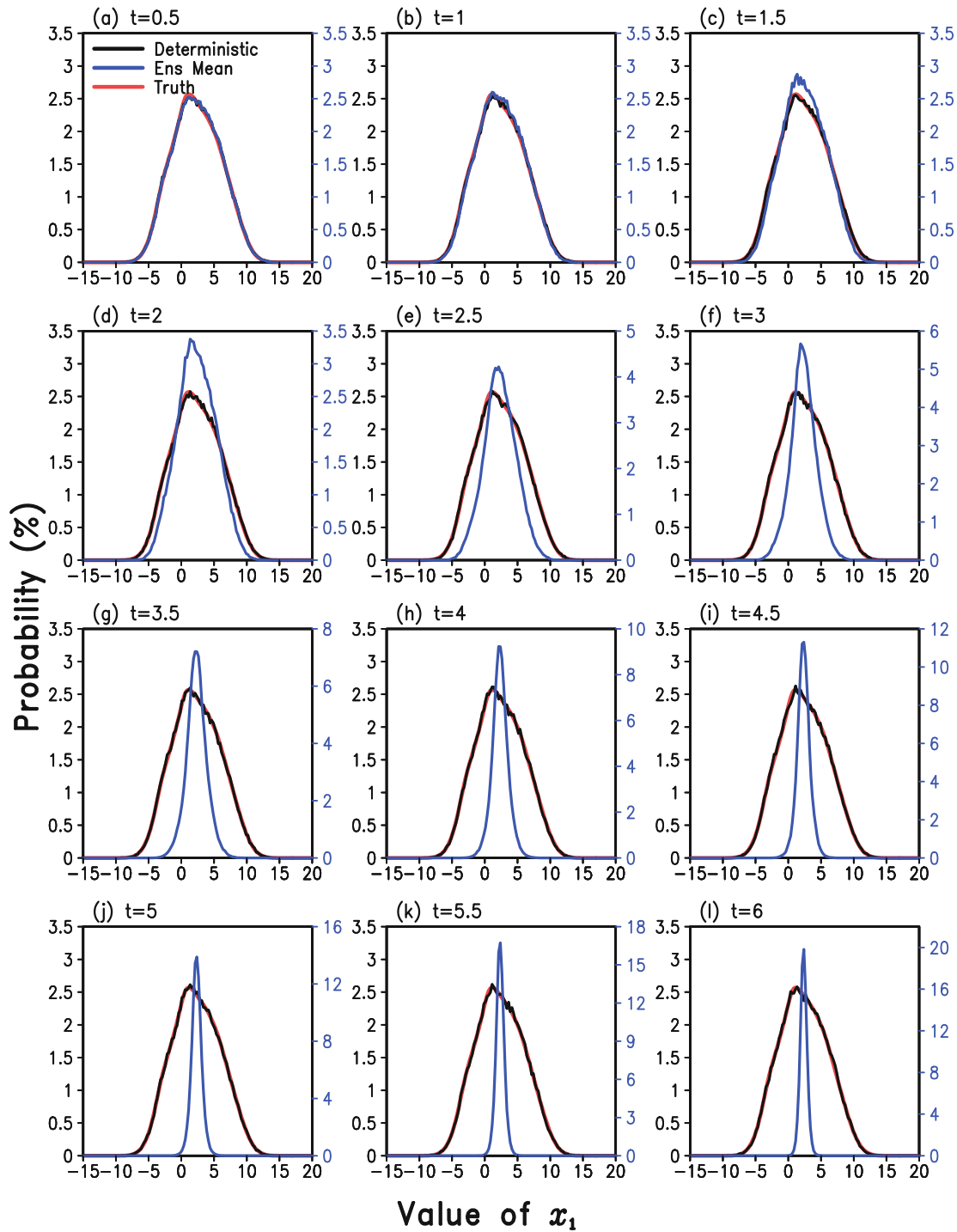


Fig. 3. Probability (%) distribution of the ensemble mean (blue line; right-hand scale) and deterministic (black line; left-hand scale) forecast states and true states (red line; left-hand scale) over all 2×10^5 samples as a function of lead time.

forecast members become chaotic with sufficiently long lead time, their ensemble mean without exception would equal the climatic mean in any individual case. It indicates that the ensemble mean reduces the forecast error compared to deterministic forecasts, but at the expense of losing information and variability in forecasts. On the other hand, according to

the characteristics of forecast states, it could be expected that the saturation value of sample mean ensemble mean forecast errors would be consistent with the attractor radius, while deterministic forecast errors will saturate at the level of GAR. The conclusion is verified through the results of forecast experiments in section 4.3.

4.3. Sample mean forecast errors

Figure 4 shows the RMS error of x_1 for deterministic and ensemble mean forecasts as a function of lead time averaged over all cases. Within the initial 1 tu, the deterministic and ensemble mean forecasts have similar errors due to the offset of the approximate linear growth of the positive and negative initial ensemble perturbations. After 1 tu, ensemble mean forecasts retain smaller errors compared to deterministic ones, and their difference continuously increases with lead time. Finally, deterministic and ensemble mean forecast errors both enter the nonlinear saturation stage and reach 5.13 and 3.63, respectively. The former is the same as the GAR and the latter equals the attractor radius. Their ratio of the saturation values is $\sqrt{2}$, as is derived in section 4.2. It is also consistent with the conclusions in Leith (1974) and Kalnay (2003).

4.4. Forecast errors in individual cases

In comparison with the sample mean forecasts, the forecasts of a specific weather or climate event is strongly influenced by the evolving dynamics (Ziehmann et al., 2000; Corazza et al., 2003), and it is thus difficult to estimate the expected values of both deterministic and ensemble mean forecast errors. LAR is a feasible statistic to estimate the expected value of deterministic and ensemble mean forecast errors in individual cases without running practical forecasts. As the nonlinearity in forecasts intensifies, the ensemble mean approaches the mean state (see Fig. 3), while the deterministic forecast tends to be a random state on the attractor. Referring to the definition of LAR in Eq. (1), the ratio r of the expected values of deterministic and ensemble mean forecast errors for a specific predicted state x_i can be expressed by:

$$r = \frac{R_{L,i}}{\|x_i - x_E\|} = \frac{\sqrt{\|x_i - x_E\|^2 + R_E^2}}{\|x_i - x_E\|}. \quad (7)$$

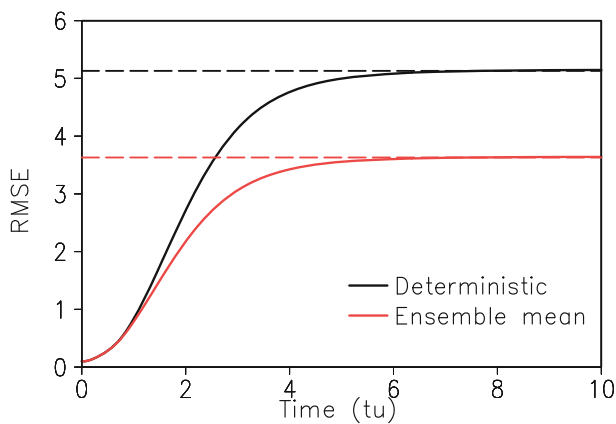


Fig. 4. RMS error averaged over 2×10^5 samples for the deterministic (black solid line) and ensemble mean (red solid line) forecasts as a function of time. The dashed lines are the saturation values, 5.13 and 3.63, of deterministic and ensemble mean forecasts errors, respectively.

Figure 5 shows the variation of r as a function the true state x_1 . It can be seen that the ensemble mean has the maximum advantage over the deterministic forecast if the truth (or the observed state) is close to the climatic mean state. When the truth gradually deviates from the mean state, the superiority of the ensemble mean over deterministic forecasts diminishes fast. For an event within 1 to 2 SD, r ranges approximately from 0.7 to 0.9. Once the event is out of 2 SD, r is almost 0.95, which means the ensemble mean and deterministic forecasts perform very similarly. This indicates that the ensemble mean has no advantage over deterministic forecasts in predicting the variabilities of extreme events, and the overall better performance of the former (see Fig. 4) originates from its higher skill for neutral events. With a long-term series of a variable, its distribution of r can be estimated in advance and used as a reference for deterministic and ensemble mean forecast skill in individual cases, especially for long-range forecasts.

To verify the above result, the practical errors of deterministic and ensemble mean forecasts are compared. The forecast skills are assessed against the truth divided into three categories—namely, the neutral (within 1 climatic SD), weak extreme (within 1–2 SD), and strong extreme (beyond 2 SD) events. Figure 6 compares the deterministic and ensemble mean forecast errors for the three groups of events at lead times of 1, 2, 3 and 4 tu. It can be seen that at 1 tu the deterministic and ensemble mean forecast errors are within similar ranges; at later times, the range of the ensemble mean errors, due to the nonlinear filtering, is evidently smaller than that of the deterministic forecast errors. After 1 tu, for both the deterministic and ensemble mean forecasts, the forecast errors of an extreme event are overall larger than those of a neutral

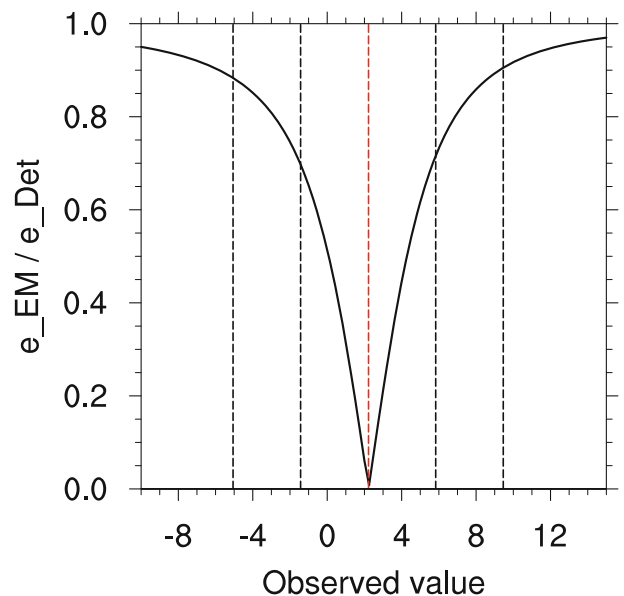


Fig. 5. Ratio between the expected values of the ensemble mean (e_{EM}) and deterministic (e_{Det}) forecast errors of x_1 as a function of the observed value of x_1 . The red dashed line represents the mean state and the black dashed lines are 1 and 2 SD, respectively.

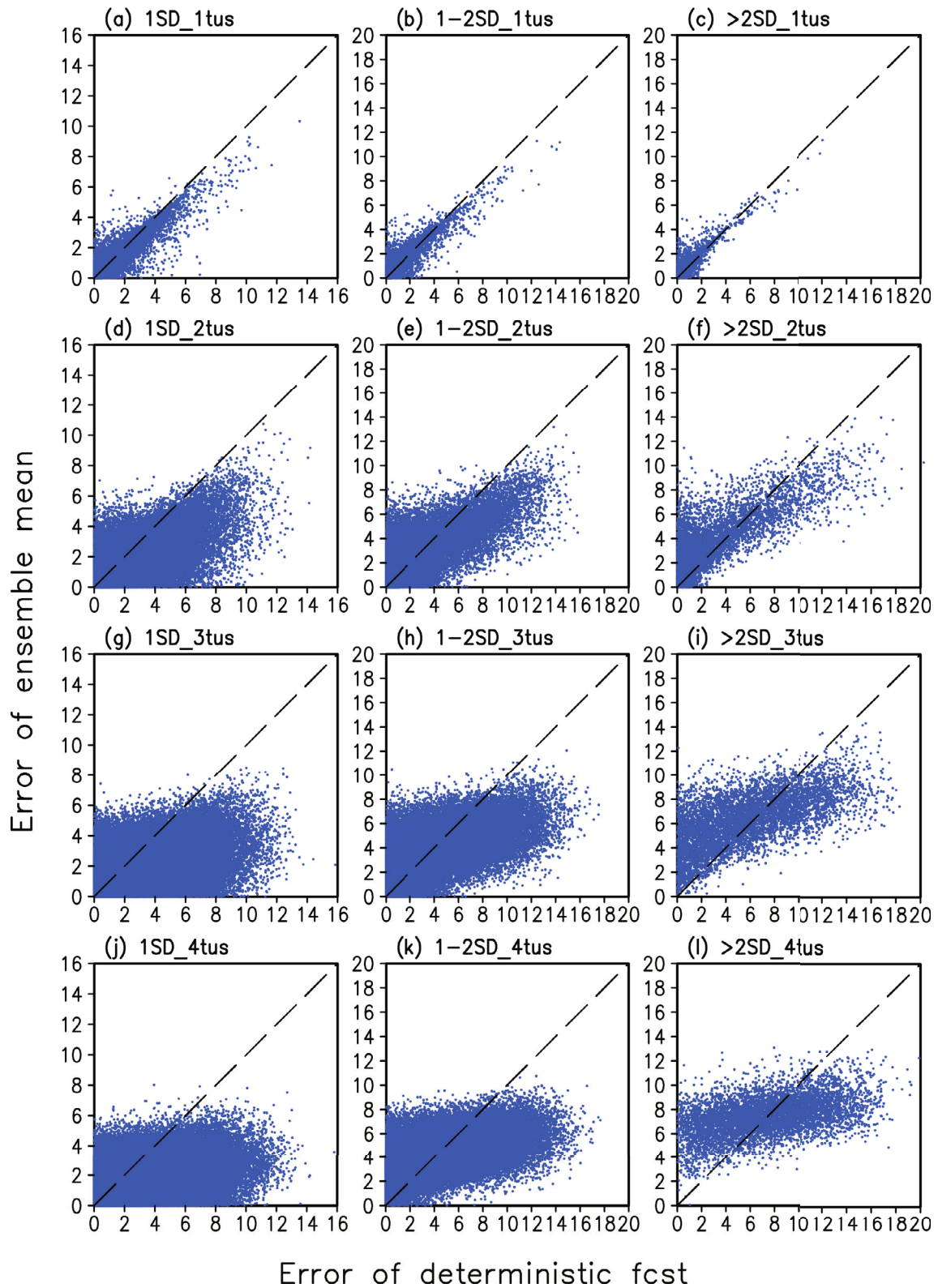


Fig. 6. RMS error of the ensemble mean and the deterministic forecasts for neutral events within one SD at time (a) 1 tu, (d) 2 tu, (g) 3 tu and (j) 4 tu. The second and third columns are same as the first, but for weak extreme states within 1–2 SD and strong extreme states out of 3 SD, respectively.

event at the same lead time, as shown in Table 1, which is essentially related to the distribution of LAR on an attractor. At long lead time (4 tu), the ratios between the average ensemble

mean and deterministic forecast errors are 0.54 (1.69 vs 3.11), 0.87 (4.19 vs 4.79) and 0.99 (7.23 vs 7.33) for neutral, weak and strong extreme events, respectively, which are within the

Table 1. Average RMS errors of ensemble mean (E.EM) and deterministic (E.Det) forecasts and their ratio for different categories of events (neutral, < 1 SD; weak extreme, 1–2 SD; strong extreme, > 2 SD) at different lead times.

	< 1 SD	1–2 SD	> 2 SD
	E.EM/E.Det Ratio		
1 tu	0.47/0.48 0.98	0.47/0.46 1.02	0.59/0.50 1.18
2 tu	1.24/1.61 0.77	1.89/1.89 1.0	3.43/2.89 1.19
3 tu	1.62/2.66 0.61	3.51/3.67 0.96	6.23/5.72 1.09
4 tu	1.69/3.11 0.54	4.19/4.79 0.87	7.23/7.33 0.99

range of the expected ratio in Fig. 5. At shorter lead times, the errors of deterministic and ensemble mean forecasts become closer for neutral and weak extreme events, but the ensemble mean performs much worse (about a 20% error increase at 1 and 2 tu) for strong extreme events. For more extreme events at a given lead time, the ensemble mean forecasts are less likely to have small RMS errors, especially for longer lead times (see Figs. 6c, f, i and l).

5. Discussion and conclusions

In this study we investigate the quantitative relationship between the forecast errors of deterministic and ensemble mean forecasts using the Lorenz96 model as an example. Instead of evaluating the results from a large number of forecast samples as most studies do, the skills of deterministic and ensemble mean forecasts are compared by using two statistics defined on attractors, namely the global and local attractor radii (GAR and LAR, respectively). GAR and LAR quantitatively describe the average distances among states on the same attractors, which are found closely related to forecast errors. The sample mean saturated errors of deterministic and ensemble mean forecasts with a perfect model can be approximately estimated by the GAR and attractor radius, respectively, and their ratio equals $\sqrt{2}$. Moreover, the expected ratio between deterministic and ensemble mean forecast errors in individual cases can be quantified by the LAR-related statistics. The results indicate that the superiority of the ensemble mean over deterministic forecasts significantly reduces from predicting neutral to strong extreme events.

GAR and LAR can be applied to practical weather and climate predictions. Since GAR and LAR are independent of specific forecast models, but derived from the attractor of observed states, they can provide objective and accurate criteria for quantifying the predictability of sample mean forecasts and individual cases in operations, respectively. The deviations of GAR and LAR between observed and practical model states may indicate the level of model deficiencies and give guidance on the development of model performance.

The relative performance of deterministic and ensemble

mean forecasts revealed by GAR and LAR will not change for practical weather and climate forecasts with model errors. However, GAR and LAR calculated based on the observed states may introduce bias when used to estimate the expected errors of deterministic and ensemble mean forecasts in imperfect prediction models. It may be more appropriate to use the other two statistics on attractors introduced by Li et al. (2018)—namely, the global and local average distances (GAD and LAD, respectively), which are similar to GAR and LAR, respectively, but estimate the average distance of states on two different attractors. The application of GAD and LAD to practical deterministic and ensemble mean forecasts will be further studied in the future.

Since the occurrence of neutral events carries large probability, the ensemble mean can still provide a valuable reference for most of the time. However, the filtering effect of the ensemble mean algorithm results in its inherent disadvantage for predicting extreme events, which cannot be easily overcome. In operations, each ensemble forecast usually has not only the amplitude but also the positional errors when predicting specific flow patterns, e.g., a trough. Therefore, the ensemble mean may have stronger smoothing effects than our theoretical results in a simple model, and thus becomes more incapable of capturing extreme flow features. To identify extreme weather, the model performance of deterministic forecasts needs further improvement toward a higher spatial resolution and more accurate model physics and parameterization. Additionally, more efficient post-processing methods for ensemble forecast members need to be developed to extract more accurate probability forecast information.

Acknowledgements. The authors acknowledge funding from the National Natural Science Foundation of China (Grant Nos. 41375110 and 41522502).

APPENDIX

This appendix shows the processes to prove Theorem 1. $R_{L,i}$ and $R_{L,j}$ are the local attractor radii of the compact attractor \mathcal{A} at state \mathbf{x}_i and \mathbf{x}_j , respectively. \mathbf{x}_E and R_E are the mean state and attractor radii of \mathcal{A} . Based on Eq. (1), the expression of $R_{L,i}$ can be derived as follows:

$$\begin{aligned}
 R_{L,i}^2 &= E(\|\mathbf{x}_i - \mathbf{x}\|^2), \quad \mathbf{x}_i, \mathbf{x} \in \mathcal{A}, \\
 &= E(\mathbf{x}^2 - 2\mathbf{x}\mathbf{x}_i + \mathbf{x}_i^2), \\
 &= E(\mathbf{x}^2) - 2\mathbf{x}_i E(\mathbf{x}) + \mathbf{x}_i^2, \\
 &= \mathbf{x}_i^2 - 2\mathbf{x}_E \mathbf{x}_i + (\mathbf{x}_E^2 + R_E^2), \\
 &= (\mathbf{x}_i - \mathbf{x}_E)^2 + R_E^2.
 \end{aligned}$$

$R_{L,i}$ reaches the minimal value R_E , i.e., the attractor radius, when $\mathbf{x}_i = \mathbf{x}_E$; and if $d_i > d_j$, $R_{L,i} > R_{L,j}$, where d_i and d_j denote the RMS distances of \mathbf{x}_i and \mathbf{x}_j from the mean state \mathbf{x}_E .

REFERENCES

- Basnarkov, L., and L. Kocarev, 2012: Forecast improvement in Lorenz 96 system. *Nonlinear Processes in Geophysics*, **19**, 569–575, <https://doi.org/10.5194/npg-19-569-2012>.
- Buizza, R., M. Milleer, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, <https://doi.org/10.1002/qj.49712556006>.
- Corazza, M., and Coauthors, 2003: Use of the breeding technique to estimate the structure of the analysis “errors of the day”. *Nonlinear Processes in Geophysics*, **10**(3), 233–243, <https://doi.org/10.5194/npg-10-233-2003>.
- Duan, W. S., and Z. H. Huo, 2016: An approach to generating mutually independent initial perturbations for ensemble forecasts: Orthogonal conditional nonlinear optimal perturbations. *J. Atmos. Sci.*, **73**, 997–1014, <https://doi.org/10.1175/JAS-D-15-0138.1>.
- Eckmann, J.-P., and D. Ruelle, 1985: Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, **57**, 617–656, <https://doi.org/10.1103/RevModPhys.57.617>.
- Farmer, J. D., E. Ott, and J. A. Yorke, 1983: The dimension of chaotic attractors. *Physica D*, **7**, 153–180, [https://doi.org/10.1016/0167-2789\(83\)90125-2](https://doi.org/10.1016/0167-2789(83)90125-2).
- Feng, J., R. Q. Ding, D. Q. Liu, and J. P. Li, 2014: The application of nonlinear local Lyapunov vectors to ensemble predictions in Lorenz systems. *J. Atmos. Sci.*, **71**, 3554–3567, <https://doi.org/10.1175/JAS-D-13-0270.1>.
- Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.*, **123**, 2181–2196, [https://doi.org/10.1175/1520-0493\(1995\)123<2181:MFEP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<2181:MFEP>2.0.CO;2).
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 230 pp.
- Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418, [https://doi.org/10.1175/1520-0493\(1974\)102<0409:TSOMCF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2).
- Li, J. P., and J. F. Chou, 1997: Existence of the atmosphere attractor. *Science in China Series D: Earth Sciences*, **40**(2), 215–220, <https://doi.org/10.1007/BF02878381>.
- Li, J. P., and R. Q. Ding, 2015: Seasonal and interannual weather prediction. *Encyclopedia of Atmospheric Sciences*, 2nd ed., North et al., Eds., Academic Press and Elsevier, 303–312.
- Li, J. P., J. Feng, and R. Q. Ding, 2018: Attractor radius and global attractor radius and their application to the quantification of predictability limits. *Climate Dyn.*, **51**(5–6), 2359–2374, <https://doi.org/10.1007/s00382-017-4017-y>.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2).
- Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**(3), 321–333, <https://doi.org/10.1111/j.2153-3490.1965.tb01424.x>.
- Lorenz, E. N., 1996: Predictability: A problem partly solved. *Proceedings of the Seminar on Predictability*, Vol. I, Reading, UK, ECMWF Seminar, 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, [https://doi.org/10.1175/1520-0469\(1998\)055<0399:OSFSWO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2).
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, <https://doi.org/10.1002/qj.49712252905>.
- Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A*, **56**, 415–428, <https://doi.org/10.3402/tellusa.v56i5.14462>.
- Szunyogh, I., and Z. Toth, 2002: The effect of increased horizontal resolution on the NCEP global ensemble mean forecasts. *Mon. Wea. Rev.*, **130**, 1125–1143, [https://doi.org/10.1175/1520-0493\(2002\)130<1125:TEOHR>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1125:TEOHR>2.0.CO;2).
- Thompson, P. D., 1957: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*, **9**(3), 275–295, <https://doi.org/10.1111/j.2153-3490.1957.tb01885.x>.
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, [https://doi.org/10.1175/1520-0477\(1993\)074<2317:EFANTG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2).
- Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319, [https://doi.org/10.1175/1520-0493\(1997\)125<3297:EFANAT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2).
- Wang, X. G., and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, **60**, 1140–1158, [https://doi.org/10.1175/1520-0469\(2003\)060<1140:ACOBAE>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<1140:ACOBAE>2.0.CO;2).
- Wei, M. Z., Z. Toth, R. Wobus, and Y. J. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus A*, **60**, 62–79, <https://doi.org/10.1111/j.1600-0870.2007.00273.x>.
- Zheng, F., J. Zhu, H. Wang, and R. H. Zhang, 2009: Ensemble hindcasts of ENSO events over the past 120 years using a large number of ensembles. *Adv. Atmos. Sci.*, **26**(2), 359–372, <https://doi.org/10.1007/s00376-009-0359-7>.
- Ziehmann, C., L. A. Smith, and J. Kurths, 2000: Localized Lyapunov exponents and the prediction of predictability. *Physics Letters A*, **271**(4), 237–251, [https://doi.org/10.1016/S0375-9601\(00\)00336-4](https://doi.org/10.1016/S0375-9601(00)00336-4).
- Zou, C. Z., X. J. Zhou, and P. C. Yang, 1985: The statistical structure of Lorenz strange attractors. *Adv. Atmos. Sci.*, **2**(2), 215–224, <https://doi.org/10.1007/BF03179753>.