

• Original Paper •

Predictability of Ensemble Forecasting Estimated Using the Kullback–Leibler Divergence in the Lorenz Model

Ruiqiang DING^{*1,2}, Baojia LIU³, Bin GU⁴, Jianping LI⁵, and Xuan LI^{1,2}

¹State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China

²College of Earth Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

³Institute of Space Weather, School of Math and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China

⁴College of Physics and Optoelectronic Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China

⁵College of Global Change and Earth System Sciences, Beijing Normal University, Beijing 100875, China

(Received 26 February 2019; revised 18 April 2019; accepted 30 April 2019)

ABSTRACT

A new method to quantify the predictability limit of ensemble forecasting is presented using the Kullback–Leibler (KL) divergence (also called the relative entropy), which provides a measure of the difference between the probability distributions of ensemble forecasts and local reference (true) states. The KL divergence is applicable to a non-normal distribution of ensemble forecasts, which is a substantial improvement over the previous method using the ensemble spread. An example from the three-variable Lorenz model illustrates the effectiveness of the KL divergence, which can effectively quantify the predictability limit of ensemble forecasting. On this basis, the KL divergence is used to investigate the dependence of the predictability limit of ensemble forecasting on the initial states and the magnitude of initial errors. The local predictability limit of ensemble forecasting varies considerably with the initial states, as well as with the magnitude of initial errors. Further research is needed to examine the real-world applications of the KL divergence in measuring the predictability of ensemble weather forecasts.

Key words: predictability, ensemble forecasting, Kullback–Leibler divergence

Citation: Ding, R. Q., B. J. Liu, B. Gu, J. P. Li, and X. Li, 2019: Predictability of ensemble forecasting estimated using the Kullback–Leibler divergence in the Lorenz model. *Adv. Atmos. Sci.*, **36**(8), 837–846, <https://doi.org/10.1007/s00376-019-9034-9>.

Article Highlights:

- A new method is introduced to quantify the predictability limit of ensemble forecasting using the KL divergence.
- The KL divergence is applicable to a non-normal distribution of ensemble forecasts, thereby overcoming the limitations of ensemble spread.

1. Introduction

The atmosphere is a chaotic system in which small errors in its initial state can lead to large forecast errors (Thompson, 1957; Lorenz, 1963, 1965; Chou, 1989; Li and Chou, 1997; Bengtsson and Hodges, 2006). We can never observe every detail of the atmosphere's initial state, either in terms of spatial coverage or accuracy of measurements, so the initial conditions from which every forecast starts are inevitably

slightly inaccurate. Small errors in the initial state will be amplified, so there is always a limit to how far ahead we can predict weather events (Lorenz, 1969, 1996; Dalcher and Kalnay, 1987; Li and Ding, 2011). Considering that weather predictions are inherently uncertain, the concept of ensemble forecasting was proposed to provide probabilistic forecasts of the future state of the atmosphere (Epstein, 1969; Leith, 1974). The basic idea of ensemble forecasting is to produce not just one single forecast but an ensemble of many forecasts starting from slightly different initial conditions.

In contrast to a single forecast, the ensemble mean of forecasts acts as a nonlinear filter that reduces forecast er-

* Corresponding author: Ruiqiang DING
Email: drq@mail.iap.ac.cn

ror (Toth and Kalnay, 1993). In general, the ensemble mean of forecasts will, on average, have a smaller error than the error of any of the single forecasts making up the ensemble (Leith, 1974; Murphy, 1988). Most importantly, the spread between the ensemble members (also called the forecast variance), which is an estimate of the standard deviation of ensemble members with respect to the ensemble mean, provides key information on the degree of confidence in the predictions under the assumption that the outputs of the ensemble members follow a normal distribution (Barker, 1991; Buizza, 1997; Palmer et al., 1998; Zhu et al., 2002). A large (small) ensemble spread indicates more (less) uncertainty in the prediction in general. In view of its advantages, ensemble forecasting is commonly performed at most of the major operational weather prediction centers worldwide, including the National Centers for Environmental Prediction (Toth and Kalnay, 1993, 1997; Wei et al., 2006, 2008), the European Centre for Medium-Range Weather Forecasts (Molteni et al., 1996; Buizza, 1997), and the Canadian Meteorological Centre (Houtekamer et al., 1996).

Ensemble forecasting aims to provide an approximate description of the probability distribution of possible future states of the atmosphere. The probability information is typically derived by using a finite number of ensemble members. Assuming that the forecast probability distribution is normal or unimodal, the width of the distribution from forecast to forecast can be measured by the ensemble spread or variance. However, the forecast probability distribution is not always unimodal and can sometimes be bimodal or even multimodal. In this case, the ensemble spread may fail to reflect the ensemble mean skill or predictability of ensemble forecasting. As pointed out by Whitaker and Loughe (1998), even for a perfect ensemble the correlation between the ensemble spread and skill may be very low. In addition, the ensemble spread has limited utility as a predictor of ensemble mean skill (Houtekamer, 1993; Kumar et al., 2000; Gritmit and Mass, 2002; Tang et al., 2008a).

Given that ensembles provide flow-dependent probabilistic forecasts of the future state of the atmosphere, it is more appropriate to investigate the predictability of ensemble forecasting from the standpoint of the flow-dependent probability distribution of ensemble forecasts instead of the ensemble spread. In the present study, in relation to the forecast probability distribution, we introduce the Kullback–Leibler (KL) divergence (also called the relative entropy) to measure the predictability limit of ensemble forecasting. The KL divergence is a measure of how one probability distribution diverges from a second, expected probability distribution (Kullback and Leibler, 1951), thereby enabling an estimate of the difference between the probability distributions of ensemble forecasts and local reference (true) states. By investigating the evolution of the KL divergence with time, we can quantitatively estimate the predictability limit of ensemble forecasting. In contrast to the ensemble spread, the KL divergence not only provides a quantitative measure of the predictability limit of ensemble forecasting but is applicable to a non-normal distribution of ensemble forecasts, thereby over-

coming the limitations of the ensemble spread and providing an effective way to investigate the predictability of ensemble forecasting.

Note that information theory measures, such as the KL divergence or relative entropy, have been used in previous studies to measure the skill of ensemble forecasts (Stephenson and Dolas-Reyes, 2000; Roulston and Smith, 2002; Del-Sole, 2004, 2005; Tang et al., 2005, 2008b). However, in these studies the entropy of ensemble forecasts was used as a measure or predictor of forecast skill, rather than a measure of the predictability limit. In this paper, we present a wider role of information theory in quantifying the predictability limit of ensemble forecasting, which can provide useful information on the time at which ensemble forecasts become meaningless.

The remainder of this paper is organized as follows. Section 2 provides a definition of the KL divergence and presents a method to compute the KL divergence for ensemble forecasting. Section 3 tests the validation and usefulness of the KL divergence in measuring the predictability of ensemble forecasting by applying it to a simple system—the three-variable Lorenz model. Section 4 summarizes the major results of this work and discusses possible limitations and future research.

2. Methods

2.1. KL divergence

The KL divergence measures the difference between two probability distributions P and Q (Kullback and Leibler, 1951). For discrete probability distributions P and Q , the KL divergence from Q to P is defined as

$$D_{\text{KL}}(P\|Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}, \quad (1)$$

where “ $\|$ ” denotes “relative to”, and Eq. (1) is equivalent to

$$D_{\text{KL}}(P\|Q) = - \sum_i P(i) \log \frac{Q(i)}{P(i)}. \quad (2)$$

For distributions P and Q of a continuous random variable x , the KL divergence is defined as

$$D_{\text{KL}}(P\|Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx, \quad (3)$$

where p and q represent the probability densities of P and Q . The KL divergence is always non-negative, with $D_{\text{KL}}(P\|Q)$ zero if and only if $P = Q$.

2.2. Local attractor radius

Let \mathbf{x}_i be a specific state on a compact attractor Ω , then the local attractor radius (LAR, R_L) with respect to the state \mathbf{x}_i is defined by Li et al. (2018) as

$$R_L(\mathbf{x}_i) = \sqrt{E(\|\mathbf{x}_i - \mathbf{x}\|^2)}, \quad \mathbf{x}_i, \mathbf{x} \in \Omega, \quad (4)$$

where the norm $\| \cdot \|$ represents the L^2 -norm and E denotes the expectation. The LAR measures the root-mean-square distance between one specific state \mathbf{x}_i and all other states on an attractor. In terms of the LAR, the local attractor with respect to the state \mathbf{x}_i can be defined as a subset of all states on the attractor whose distance to the state \mathbf{x}_i is less than the LAR. Li et al. (2018) showed that the LAR can be used as an objective metric to quantify the local predictability limit of forecast models. In the present study, the LAR is used to define the local attractor with respect to a specific reference state and to construct the probability distributions of local reference (true) states.

An example from the three-variable Lorenz system is given to illustrate the spatial structure of the LAR over the Lorenz attractor. The three-variable Lorenz system is

$$\begin{cases} \frac{dX}{dt} = -\sigma X + \sigma Y \\ \frac{dY}{dt} = rX - Y - XZ \\ \frac{dZ}{dt} = XY - bZ \end{cases}, \quad (5)$$

where $\sigma = 10$, $r = 28$, and $b = 8/3$, for which the system exhibits chaotic behavior (Lorenz, 1963). Figure 1 shows a projection of the LAR over the Lorenz attractor in the x - y plane. Obviously, the LAR varies widely over the attractor, with a minimum value of the LAR of ~ 15 and the maximum value exceeding 35. The LAR is not randomly distributed but exhibits a distinct organization in phase space, consistent with the results of Li et al. (2018). The LAR is antisymmetric with respect to the x - or y -axis, with minimum values at the intersection of the two wings and maximum values at the outermost rims. As the LAR varies over the attractor, the local attractor with respect to a specific state also changes with the state.

2.3. Calculation of the KL divergence in ensemble forecasting

The definition of the KL divergence in Eq. (1) aims to quantify the difference between two probability distributions, P and Q . To compute the KL divergence in ensemble forecasting, it is necessary to estimate the probability distribution of local reference (true) states (hereafter P) and the probability distribution of ensemble forecasts (hereafter Q). For a specific reference state \mathbf{x}_i , we first calculate the LAR of the state \mathbf{x}_i . Then, we can obtain the subset of all states on the attractor whose distance to the reference state is less than the LAR. Finally, the probability distribution P of local reference (true) states can be obtained based on the subset of the states on the local attractor.

When N random perturbations are added to or subtracted from the reference states, N different results of ensemble forecasts can be generated from the prediction model. Based on N ensemble forecasts, the probability distribution Q of ensemble forecasts can then be obtained. Once both P and Q are obtained, we can directly compute the KL divergence. As the reference state and ensemble forecasts change with the

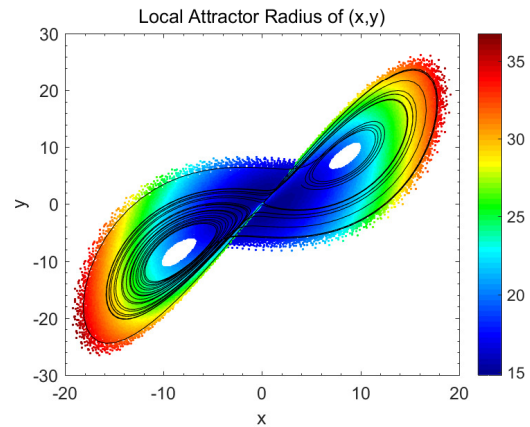


Fig. 1. Projection of the LAR over the Lorenz attractor in the x - y plane.

forecast time, the KL divergence will vary with the forecast time. By examining the evolution of the KL divergence with the forecast time, we can quantitatively estimate the predictability limit of ensemble forecasting.

2.4. Nonlinear local Lyapunov exponent method

The nonlinear local Lyapunov exponent (NLLE), which is a nonlinear extension of the existing linear finite-time or local Lyapunov exponents (Yoden and Nomura, 1993; Boffetta et al., 1998; Ziehmman et al., 2000), measures the mean growth rate of the initial errors of nonlinear dynamical systems without having to linearize the nonlinear equations of motion (Ding and Li, 2007, Ding et al., 2008a; Li and Ding, 2011). The NLLE and its derivative (i.e., the mean relative growth of the initial error) have been widely applied to quantitatively determine the limit of dynamic predictability of weather or climate variables (Ding et al., 2008b, 2010, 2011, 2015), exhibiting superior performance to the existing linear finite-time or local Lyapunov exponents. A brief description of the NLLE method is given in Appendix A.

Note that the NLLE method is defined based on nonlinear error dynamics, while the KL divergence is defined based on probability and information theory. Some differences exist between both methods. For example, the NLLE method uses the root-mean-square error as the measure of error, and therefore depends on the dimension of variables. In contrast, the KL divergence uses the difference between two probability distributions as the measure of uncertainty, and therefore does not depend on the dimension of variables. This may be one advantage of the KL divergence relative to the NLLE method. Nevertheless, although the NLLE method (the KL divergence) is used to determine the predictability limit by exploring the evolution of initial errors (the evolution of forecast probability distributions), considering that the predictability limit is an intrinsic property of a given dynamical system that does not depend on specific methods (Lorenz, 1969; Mu et al., 2017), the predictability limit of ensemble forecasting derived from the KL divergence and from error evolution should be consistent (see Fig. 2). Therefore, we

compare the predictability limits of ensemble forecasting derived from the KL divergence and NLE. Their consistency would support the effectiveness of the KL divergence in measuring the predictability of ensemble forecasting.

3. Results

Taking the three-variable Lorenz model as an example, we examine the evolution of the KL divergence with forecast time t for ensemble forecasting. Starting from a randomly chosen initial state \mathbf{x}_{01} $(-5.76, -0.29, 30.5)$ on the Lorenz attractor, we first integrate the Lorenz model to obtain the long-term model states as the reference states. The local attractor with respect to each reference state can be determined

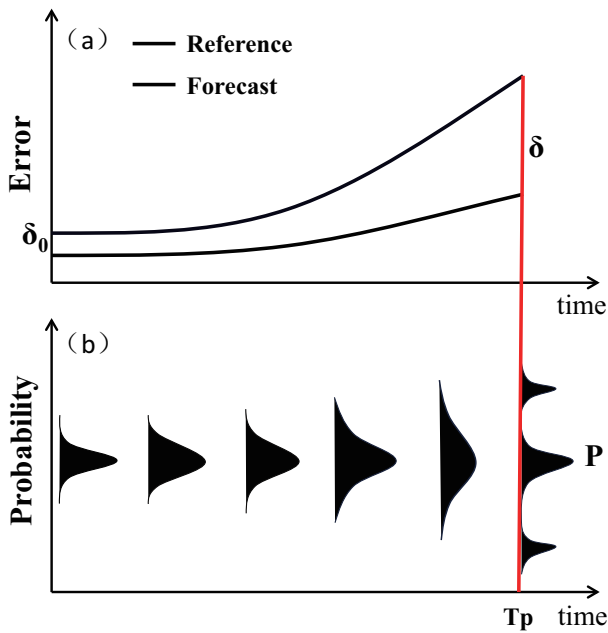


Fig. 2. Schematic illustration of the consistency of the predictability limit (denoted as T_p) derived from (a) error and (b) probability evolutions.

from the LAR, and then the probability distribution P of local reference (true) states can be obtained based on the subset of the states on the local attractor. Note that the local attractor and its probability distribution P depend on the reference states that vary with integration time. To obtain the probability distribution Q of ensemble forecasts, we superpose $N = 10^5$ initial perturbations with the same amplitude, $\varepsilon = 10^{-3}$, and random directions in phase space onto the initial state \mathbf{x}_{01} to generate slightly different initial states. Then, ensemble predictions are made starting from these different initial states. For each forecast time, the probability distribution Q of ensemble forecasts can be obtained based on ensemble members. The KL divergence is calculated based on Eq. (1) for discrete probability distributions P and Q .

Figure 3a shows the variation in the KL divergence as a function of time t for the initial state \mathbf{x}_{01} $(-5.76, -0.29, 30.5)$. The KL divergence shows a nonuniform growth process with time. At time $t = 7$, the KL divergence reaches a maximum value, implying that the probability distribution Q of ensemble forecasts deviates most from the probability distribution P of local reference (true) states. At this time, the forecast distribution yields unreliable probabilistic forecasts, and the ensemble prediction can be considered meaningless. If the time at which the KL divergence reaches its maximum value is specified as the local predictability limit, the predictability limit of ensemble forecasting starting from \mathbf{x}_{01} with $\varepsilon = 10^{-3}$ would be $T_p \approx 7$. For another initial state \mathbf{x}_{02} $(10.3, 0.92, 16.7)$, the KL divergence shows a similar zigzag growth process before it reaches the maximum value at around $t = 11$ (Fig. 4a). According to the definition, the local predictability limit of ensemble forecasting starting from \mathbf{x}_{02} with $\varepsilon = 10^{-3}$ would be $T_p \approx 11$, greater than the predictability limit at \mathbf{x}_{01} .

To understand why the KL divergence varies with time, we examine the evolution of the probability distributions P (local reference states) and Q (ensemble forecasts) with time for \mathbf{x}_{01} (Fig. 5). Both probability distributions P and Q change with time. At the beginning of ensemble forecasting, Q is concentrated in the center of P . Gradually, the range of

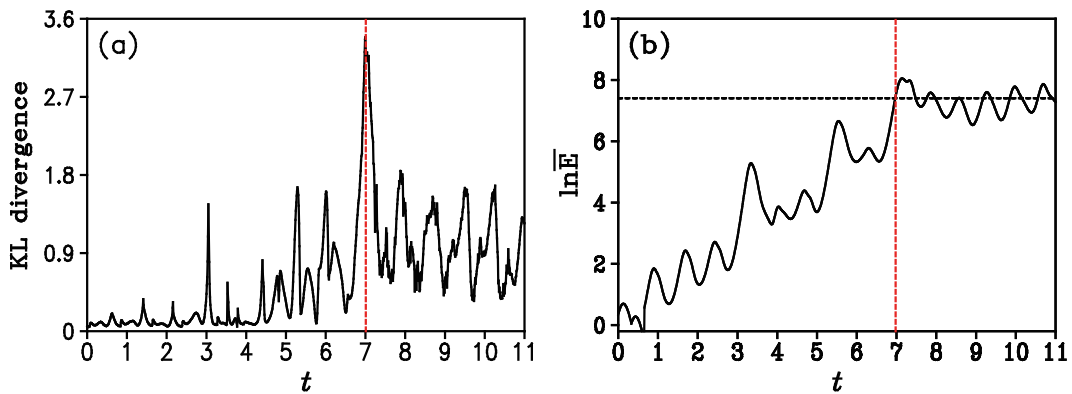


Fig. 3. For the initial state on the Lorenz attractor \mathbf{x}_{01} $(-5.76, -0.29, 30.5)$, we show (a) the KL divergence and (b) the mean error growth obtained using the NLE method with $\varepsilon = 10^{-3}$ as a function of time t . In (a), the time at which the KL divergence reaches its maximum value is indicated by the red dashed line. In (b), the average value of the nonlinear stochastic fluctuation states of the mean error is indicated by the black dashed line, and the time at which the error growth enters the nonlinear stochastic fluctuation states is indicated by the red dashed line.

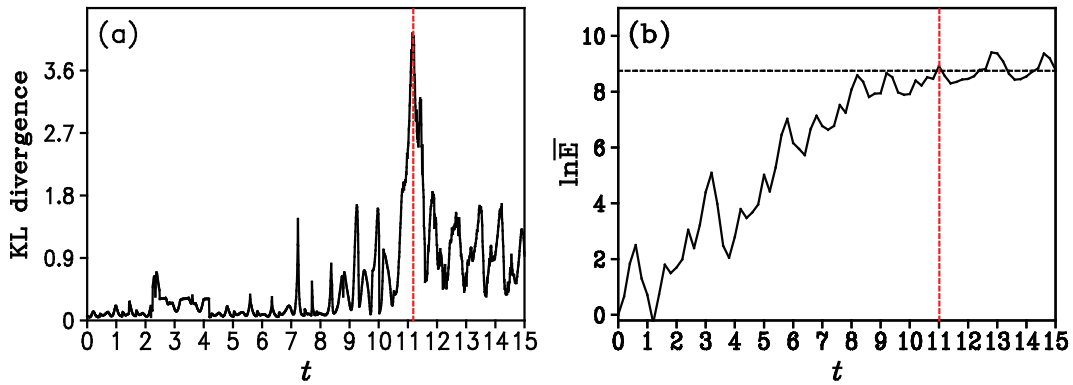


Fig. 4. As in Fig. 3 but for the other initial state on the Lorenz attractor, \mathbf{x}_{02} (10.3, 0.92, 16.7).

Q becomes wider as ensemble perturbations tend to diverge over time. Correspondingly, Q begins to diverge from P and the KL divergence gradually becomes larger. When $t = 7$, the difference between P and Q is significant. As a result, the KL divergence reaches its maximum value at that moment. Afterwards, Q gradually converges to the distribution of the entire attractor as ensemble perturbations expand to the entire attractor, and instead P falls within Q . Correspondingly, the KL divergence drops from the peak and then enters the nonlinear stochastic fluctuation phase. In addition, we note in Fig. 5 that ensemble forecasts do not follow a normal distribution for each forecast time, implying that the ensemble spread may not be appropriate to provide an accurate measure of the predictability of ensemble forecasting. In contrast, the application of the KL divergence in this paper excludes the influence of the type of probability distribution, and therefore ensures the accuracy of estimates of the predictability of ensemble forecasting.

We now examine the evolution of the local attractor (green points in Fig. 6) and ensemble forecast states (red points in Fig. 6) starting from \mathbf{x}_{01} over the entire Lorenz attractor. At the beginning of the ensemble forecast, all ensemble forecast states fall within the local attractor. As the prediction time increases, ensemble forecast states begin to fall outside the local attractor, and gradually expand to the entire attractor. When $t = 7$, almost all ensemble forecast states fall completely outside the local attractor, and the prediction subsequently becomes meaningless as the KL divergence reaches its maximum value (Li et al., 2018). Therefore, it is reasonable to use the maximum value of the KL divergence to measure the predictability of ensemble forecasting.

The predictability of ensemble forecasting derived from the probability and error evolutions should be consistent. It is interesting to compare the local predictability limit obtained using the KL divergence and NLE. Figure 3b shows the ensemble mean error growth over 10^5 initial random perturbations obtained using the NLE method for the initial state \mathbf{x}_{01} and $\varepsilon = 10^{-3}$. The mean error initially shows an oscillating growth, and finally stops increasing and enters the nonlinear stochastic oscillation regime with a constant average value. Once the error growth enters the nonlinear stochastic oscillation regime, almost all predictability is lost and the prediction becomes meaningless. Following the work of Ding

et al. (2008b), we determine the local predictability limit as the time at which the mean error reaches the average value of the nonlinear stochastic fluctuation states. Then, we find that the predictability limit at \mathbf{x}_{01} with $\varepsilon = 10^{-3}$ calculated using the NLE method is $T_p \approx 7$, which is consistent with the predictability limit derived from the KL divergence. Similarly, the predictability limit at \mathbf{x}_{02} with $\varepsilon = 10^{-3}$ calculated using the NLE method is $T_p \approx 11$ (Fig. 4b), which is also consistent with the predictability limit derived from the KL divergence. The consistency across methods lends support to the effectiveness of the KL divergence in measuring the predictability of ensemble forecasting.

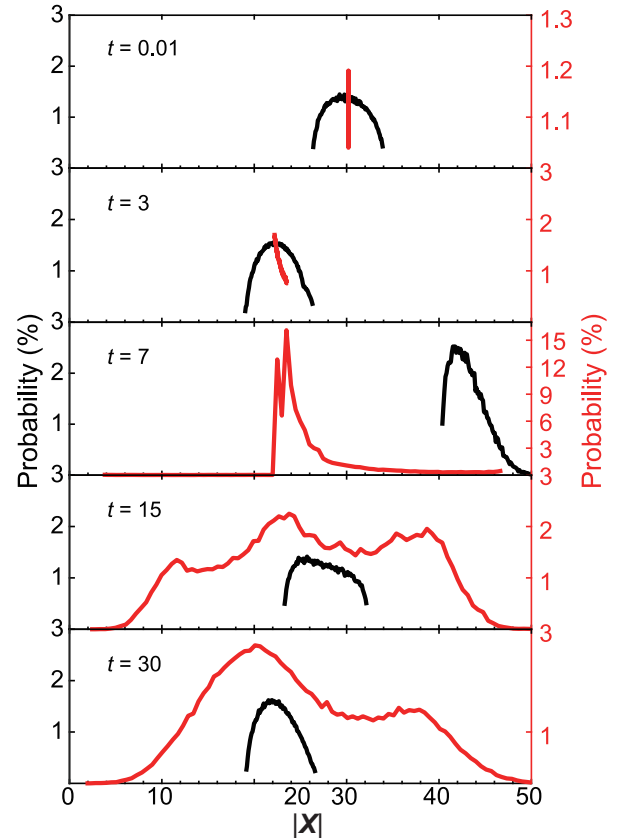


Fig. 5. Evolution of the probability distributions P (local reference states; black line; left axis) and Q (ensemble forecasts; red line; right axis) with time for \mathbf{x}_{01} (-5.76, -0.29, 30.5).

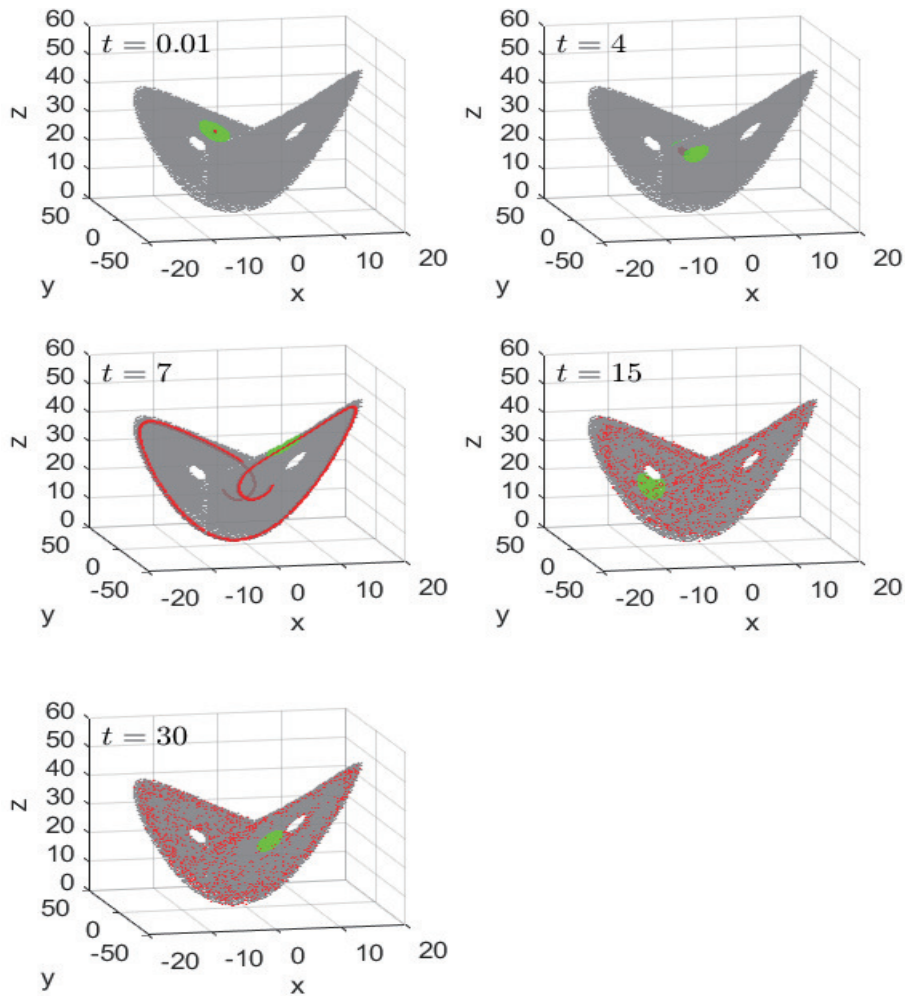


Fig. 6. Evolution of the local attractor (green points) and ensemble forecast states (red points) starting from \mathbf{x}_{01} $(-5.76, -0.29, 30.5)$ over the entire Lorenz attractor (gray points).

Figure 7 shows the variations in the local predictability limit of ensemble forecasting as a function of initial states \mathbf{x}_i with $\varepsilon = 10^{-3}$ for a typical trajectory on the Lorenz attractor. The local predictability limit of ensemble forecasting varies widely with initial state on the Lorenz attractor. For 600 initial states, we find that a minimum value of the local predictability limit is ~ 3.6 , while the maximum value is ~ 16 . Local predictability limits obtained using the KL divergence and NLE closely resemble each other, with a correlation coefficient of 0.92 (significant at the 99.9% confidence level). These results indicate that the predictability of ensemble forecasting depends on the initial states of the ensemble forecast.

We now consider the structure of predictability in phase space by investigating the three-dimensional distribution of the local predictability limit derived from the KL divergence (Fig. 8). The local predictability limit has a distinct organization in phase space. On the whole, the inner and outer rims of each wing of the Lorenz attractor have a relatively high local predictability limit, while the regions between the inner and outer rims of each wing have a relatively low predictability limit, consistent with the distribution of the local predictability limit derived from the NLE method (Huai et al., 2017).

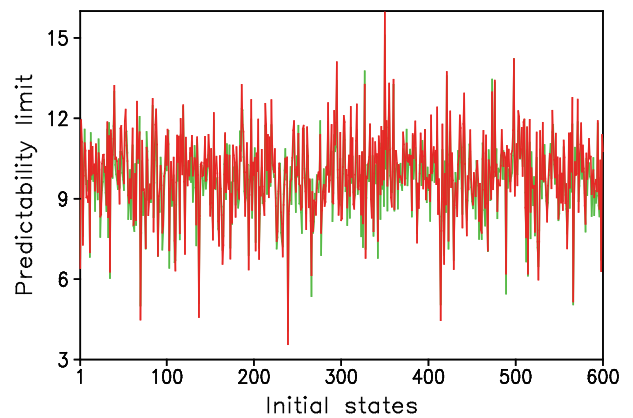


Fig. 7. Variations in the local predictability limit obtained using the KL divergence (red line) and NLE (green line) as a function of initial states \mathbf{x}_i ($i = 1, 2, \dots, 600$) with $\varepsilon = 10^{-3}$ for a typical trajectory on the Lorenz attractor.

Huai et al. (2017) pointed out that this structure of the local predictability limit in phase space may be related to the local dynamics of the Lorenz attractor that affects the length of time that each point remains on the current wing, and this pe-

riod of time is important in determining the local predictability limit of each point. This underlying structure allows the identification of regions in phase space of high and low predictability and may be helpful in estimating the predictability for each point.

The predictability of ensemble forecasting depends on the initial states as well as the magnitude of initial errors. For the analysis presented above, the magnitude ε of initial errors is fixed as 10^{-3} . We next examine the dependence of the predictability of ensemble forecasting on the magnitude of initial errors. Figure 9 shows the local predictability limits derived from the KL divergence as a function of the magnitude of initial errors for the initial state \mathbf{x}_{03} (6.03, 9.71, 16.5). As a comparison, local predictability limits derived from the NLLE method as a function of the magnitude of initial errors are also shown in Fig. 9. Local predictability limits derived from the KL divergence and NLLE decrease approximately linearly as the logarithm of the magnitude of initial errors is increased. For a specific initial error, the local predictability limit derived from the KL divergence is very close to the limit derived from the NLLE method. Similar results are obtained for initial states \mathbf{x}_{01} and \mathbf{x}_{02} (not shown), indicating that the predictability of ensemble forecasting is sensitive to the magnitude of initial errors.

Let us now consider an important question concerning the influence of the number of ensemble members on the predictability estimation of ensemble forecasting. Given that the KL divergence is obtained by computing the difference between the probability distributions of local reference states and ensemble forecasts, an accurate estimate of the KL divergence depends on an accurate estimate of the probability distributions of local reference states and ensemble forecasts. However, a sufficiently large number of ensemble members is required to accurately estimate the probability distribution of ensemble forecasts. It is likely that the method using the KL divergence to estimate the predictability of ensemble

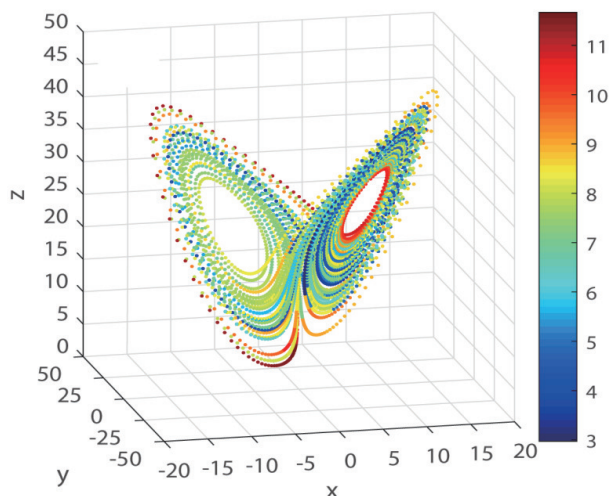


Fig. 8. Three-dimensional distribution of the local predictability limit of 5000 states on the Lorenz attractor derived from the KL divergence.

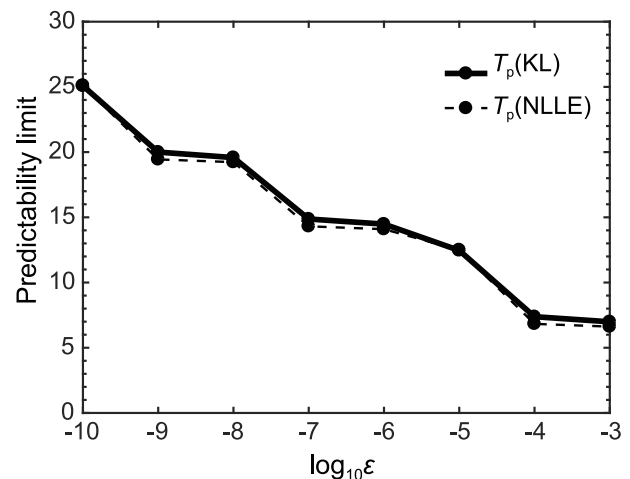


Fig. 9. Local predictability limits derived from the KL divergence (solid line with dots) and NLLE (dashed line with dots) as a function of the magnitude of initial error ε for the initial state \mathbf{x}_{03} (6.03, 9.71, 16.5).

forecasting would give worse results for ensemble predictions using operational weather forecasting models, in which the number of ensemble members is usually restricted due to limitations in computing resources.

The number of ensemble members used in the present study is $N = 10^5$. We examine the dependence of the estimated predictability limit of ensemble forecasting on the number of ensemble members. Figure 10 shows the estimated local predictability limit of ensemble forecasting starting from \mathbf{x}_{01} as a function of the number of ensemble members. The number of ensemble members decreases from 10^5 to 200; the latter is close to the number of ensemble members used in current operational weather forecasting. The local predictability limit is initially almost constant, followed by a gradual decrease with decreasing number of ensemble members. This result might be expected because the estimation of the probability distributions of local reference states and ensemble forecasts would have larger uncertainties for a smaller number of ensemble members. The estimated predictability limit obtained using 200 ensemble members is $T_p \approx 6.5$, which is slightly lower than the limit obtained using 10^5 ensemble members. Similar results were obtained for other initial states (not shown). These results suggest that, although a smaller number of ensemble members tends to underestimate the predictability limit to some extent, such an underestimate is relatively slight. Consequently, it may be feasible to use a relatively small number of ensemble members to estimate the predictability of ensemble forecasting in the Lorenz model.

Note that the above analyses are based solely on a simple toy model: the three-variable Lorenz model. For complex weather forecasting models, the situation may be different when we try to estimate the probability distribution in a higher-dimensional space using a small number of ensemble forecasts. In this case, the probability distribution of ensemble forecasts may be poorly estimated, possibly leading to a large error in the estimation of predictability. This may be a

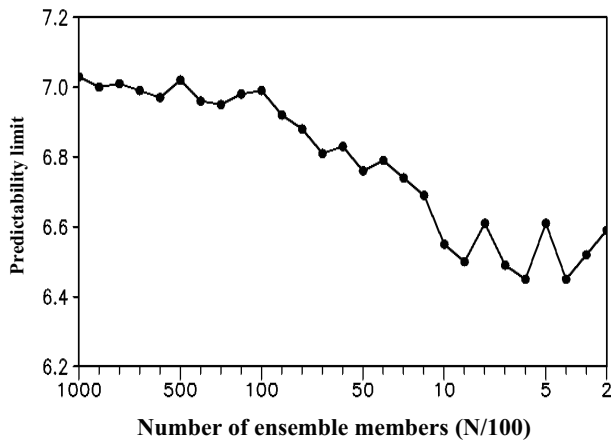


Fig. 10. Estimated local predictability limit of ensemble forecasting starting from x_{01} $(-5.76, -0.29, 30.5)$ over the Lorenz attractor as a function of the number of ensemble members.

limitation of the KL divergence. Hopefully, with increased computing resources available, the number of ensemble members can be further increased in real-world numerical weather models. Further research is required to examine the application of the KL divergence in real-world ensemble weather forecasts and to assess the influence of ensemble size on estimates of predictability.

We then consider another important question regarding the influence of model errors on the accurate estimation of the KL divergence. This study simply uses the Lorenz model without model error. Given the existence of model error, the probability distribution of true states P is generally unknown. If the forecast states are used instead of true states, the KL divergence and hence the estimated predictability limit would possibly include an error. For the Lorenz attractor, the local attractor with respect to a given state is not sensitive to the state itself (see Fig. 1), and the local attractor and its probability distribution of nearby states are similar. Consequently, a small error in the Lorenz model would produce a relatively small initial error in the KL divergence. In real-world ensemble weather forecasts, although models are imperfect, a large amount of observed atmospheric data is available. We can use observations to estimate the probability distribution of true states; this remains a topic for future research.

4. Conclusions

We have presented a new method using the KL divergence to measure the predictability of ensemble forecasting. The KL divergence allows us to estimate the difference between the probability distributions of ensemble forecasts and local reference (true) states. By investigating the evolution of the KL divergence with time, the local predictability limit of ensemble forecasting may be quantitatively determined. The KL divergence is applicable to a non-normal distribution of ensemble forecasts. This represents an improvement over the ensemble spread, which is only applicable under the assumption that the ensemble members follow a normal distribution.

Using the KL divergence, we have performed a quantitative analysis of the predictability of ensemble forecasting in the Lorenz model. The local predictability limit derived from the KL divergence is clearly consistent with that derived from error evolution, lending support to the effectiveness of the KL divergence in measuring the predictability of ensemble forecasting.

In addition, we have investigated the sensitivity of the predictability of ensemble forecasting to the initial states and the magnitude of initial errors. We found that the predictability of ensemble forecasting depends on the initial states as well as on the magnitude of initial errors. The local predictability limit of ensemble forecasting varies considerably with time, but the predictability variability shows organization in phase space. The predictability of ensemble forecasting is also sensitive to the magnitude of initial errors. The local predictability limit decreases approximately linearly as the logarithm of the magnitude of initial errors is increased.

Our study presents a preliminary application of the KL divergence in measuring the predictability of ensemble forecasting in a relatively simple system. For more complex ensemble weather or climate forecasts, there will be higher dimensionality and more complicated models. This implies that there would exist some uncertainties in estimating the KL divergence for operational weather or climate forecasts, which poses a challenge to the accurate estimation of the KL divergence. It would be interesting to extend the current investigation to more realistic ensemble weather forecasts, which we intend to examine in future research. In addition, this study simply used random perturbations as ensemble perturbations. Up to now, various schemes have been developed to generate the initial perturbations in ensemble forecasts, such as the bred vector method (Toth and Kalnay, 1993, 1997), the singular vector method (Molteni et al., 1996; Buizza, 1997), and the ensemble transform Kalman filter (Bishop et al., 2001; Wang and Bishop, 2003). These schemes have been shown to improve operational forecasts compared with random perturbations. It is worthwhile examining from the standpoint of the KL divergence the predictability of these ensemble forecasts using perturbations generated by such schemes.

Acknowledgements. This work was jointly supported by the National Key Research and Development Program of China (Grant No. 2018YFC1506402), the National Program on Global Change and Air–Sea Interaction (Grant Nos. GASI-IPOVAI-03 and GASI-IPOVAI-06), and the National Key Technology Research and Development Program of the Ministry of Science and Technology of China (Grant No. 2015BAC03B07).

APPENDIX A

Introduction to the NLE method

Consider a general n -dimensional nonlinear dynamical system whose evolution is governed by

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}), \quad (\text{A1})$$

where $\mathbf{x} = [x_1(t), x_2(t), \dots, x_n(t)]^T$ is the state vector at time t , the superscript T is the transpose, and \mathbf{F} represents the dynamics. The evolution of a small error $\boldsymbol{\delta} = [\delta_1(t), \delta_2(t), \dots, \delta_n(t)]^T$, superimposed on a state \mathbf{x} is governed by the following nonlinear equation:

$$\frac{d}{dt}\boldsymbol{\delta} = \mathbf{J}(\mathbf{x})\boldsymbol{\delta} + \mathbf{G}(\mathbf{x}, \boldsymbol{\delta}), \quad (\text{A2})$$

where $\mathbf{J}(\mathbf{x})\boldsymbol{\delta}$ are the tangent linear terms and $\mathbf{G}(\mathbf{x}, \boldsymbol{\delta})$ are the high-order nonlinear terms of the error $\boldsymbol{\delta}$. Without a linear approximation, the solutions of Eq. (A2) can be obtained by numerical integration along the reference solution \mathbf{x} from $t = t_0$ to $t_0 + \tau$:

$$\boldsymbol{\delta}_1 = \boldsymbol{\eta}(\mathbf{x}_0, \boldsymbol{\delta}_0, \tau)\boldsymbol{\delta}_0, \quad (\text{A3})$$

where $\boldsymbol{\delta}_1 = \boldsymbol{\delta}(t_0 + \tau)$, $\mathbf{x}_0 = \mathbf{x}(t_0)$, $\boldsymbol{\delta}_0 = \boldsymbol{\delta}(t_0)$, and $\boldsymbol{\eta}(\mathbf{x}_0, \boldsymbol{\delta}_0, \tau)$ is the nonlinear propagator. The NLLE is then defined as

$$\lambda(\mathbf{x}_0, \boldsymbol{\delta}_0, \tau) = \frac{1}{\tau} \ln \frac{\|\boldsymbol{\delta}_1\|}{\|\boldsymbol{\delta}_0\|}, \quad (\text{A4})$$

where $\lambda(\mathbf{x}_0, \boldsymbol{\delta}_0, \tau)$ depends in general on the initial state \mathbf{x}_0 in phase space, the initial error $\boldsymbol{\delta}_0$, and time τ . The NLLE differs from existing local or finite-time Lyapunov exponents defined from linear error dynamics, which depend solely on the initial state \mathbf{x}_0 and time τ , and not on the initial error $\boldsymbol{\delta}_0$. Assuming that all initial perturbations with amplitude ε and random directions are on an n -dimensional spherical surface centered at an initial point \mathbf{x}_0 , then we have

$$\boldsymbol{\delta}_0^T \boldsymbol{\delta}_0 = \varepsilon^2. \quad (\text{A5})$$

The local ensemble mean of the NLLE over a large number of random initial perturbations is given by

$$\bar{\lambda}(\mathbf{x}_0, \tau) = \langle \lambda(\mathbf{x}_0, \boldsymbol{\delta}_0, \tau) \rangle_N, \quad (\text{A6})$$

where $\langle \rangle_N$ denotes the local ensemble average of samples of large enough size N ($N \rightarrow \infty$). Here, $\bar{\lambda}(\mathbf{x}_0, \tau)$ characterizes the average growth rate of random perturbations superimposed on \mathbf{x}_0 within a finite time τ . For a fixed time τ , $\bar{\lambda}(\mathbf{x}_0, \tau)$ depends on \mathbf{x}_0 and reflects the local error growth dynamics of the attractor. The mean local relative growth of the initial error can be obtained by

$$\bar{E}(\mathbf{x}_0, \tau) = e^{[\bar{\lambda}(\mathbf{x}_0, \tau)\tau]}. \quad (\text{A7})$$

For a given initial state \mathbf{x}_0 , $\bar{E}(\mathbf{x}_0, \tau)$ initially increases with time τ and finally reaches a state of nonlinear stochastic fluctuation, which means that error growth reaches saturation with a constant average value. At that moment, almost all information on the initial state is lost and the prediction becomes meaningless. If the local predictability limit is defined as the time at which the error reaches the average value of the nonlinear stochastic fluctuation states, the predictability limit of the system at \mathbf{x}_0 can be quantitatively determined.

REFERENCES

- Barker, T. W., 1991: The relationship between spread and forecast error in extended-range forecasts. *J. Climate*, **4**, 733–742, [https://doi.org/10.1175/1520-0442\(1991\)004<0733:TRBSAF>2.0.CO;2](https://doi.org/10.1175/1520-0442(1991)004<0733:TRBSAF>2.0.CO;2).
- Bengtsson, L., and K. I. Hodges, 2006: A note on atmospheric predictability. *Tellus A*, **58**, 154–157, <https://doi.org/10.1111/j.1600-0870.2006.00156.x>.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436, [https://doi.org/10.1175/1520-0493\(2001\)129<0420:ASWTET>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2).
- Boffetta, G., P. Giuliani, G. Paladin, and A. Vulpiani, 1998: An extension of the Lyapunov analysis for the predictability problem. *J. Atmos. Sci.*, **55**, 3409–3416, [https://doi.org/10.1175/1520-0469\(1998\)055<3409:AEOTLA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<3409:AEOTLA>2.0.CO;2).
- Buizza, R., 1997: Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, **125**, 99–119, [https://doi.org/10.1175/1520-0493\(1997\)125<0099:PFSEOP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0099:PFSEOP>2.0.CO;2).
- Chou, J. F., 1989: Predictability of the atmosphere. *Adv. Atmos. Sci.*, **6**, 335–346, <https://doi.org/10.1007/BF02661539>.
- Dalcher, A., and E. Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. *Tellus A*, **39**, 474–491, <https://doi.org/10.3402/tellusa.v39i5.11774>.
- DelSole, T., 2004: Predictability and information theory. Part I: Measures of predictability. *J. Atmos. Sci.*, **61**, 2425–2440, [https://doi.org/10.1175/1520-0469\(2004\)061<2425:PAITPI>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<2425:PAITPI>2.0.CO;2).
- DelSole, T., 2005: Predictability and information theory. Part II: Imperfect forecasts. *J. Atmos. Sci.*, **62**, 3368–3381, <https://doi.org/10.1175/JAS3522.1>.
- Ding, R. Q., and J. P. Li, 2007: Nonlinear finite-time Lyapunov exponent and predictability. *Physics Letters A*, **364**, 396–400, <https://doi.org/10.1016/j.physleta.2006.11.094>.
- Ding, R. Q., J. P. Li, and K. J. Ha, 2008a: Nonlinear local Lyapunov exponent and quantification of local predictability. *Chin. Phys. Lett.*, **25**, 1919–1922, <https://doi.org/10.1088/0256-307X/25/5/109>.
- Ding, R. Q., J. P. Li, and K. J. Ha, 2008b: Trends and interdecadal changes of weather predictability during 1950s–1990s. *J. Geophys. Res.*, **113**, D24112, <https://doi.org/10.1029/2008JD010404>.
- Ding, R. Q., J. P. Li, and K. H. Seo, 2010: Predictability of the madden-Julian oscillation estimated using observational data. *Mon. Wea. Rev.*, **138**, 1004–1013, <https://doi.org/10.1175/2009MWR3082.1>.
- Ding, R. Q., J. P. Li, and K. H. Seo, 2011: Estimate of the predictability of boreal summer and winter intraseasonal oscillations from observations. *Mon. Wea. Rev.*, **139**, 2421–2438, <https://doi.org/10.1175/2011MWR3571.1>.
- Ding, R. Q., J. P. Li, F. Zheng, J. Feng, and D. Q. Liu, 2015: Estimating the limit of decadal-scale climate predictability using observational data. *Climate Dyn.*, **46**, 1563–1580, <https://doi.org/10.1007/s00382-015-2662-6>.
- Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus A*, **21**, 739–759, <https://doi.org/10.1111/j.2153-3490.1969.tb00483.x>.
- Grimit, E. P., and C. F. Mass, 2002: Initial results of a mesoscale short-range ensemble forecasting system over the Pacific Northwest. *Wea. Forecasting*, **17**, 192–205, [https://doi.org/10.1175/1520-0434\(2002\)017<0192:IROAMS>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<0192:IROAMS>2.0.CO;2).
- Houtekamer, P. L., 1993: Global and local skill forecasts. *Mon. Wea. Rev.*, **121**, 1834–1846, [https://doi.org/10.1175/1520-0493\(1993\)121<1834:GALSF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<1834:GALSF>2.0.CO;2).
- Houtekamer, P. L., L. Lefaire, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble

- prediction. *Mon. Wea. Rev.*, **124**, 1225–1242, [https://doi.org/10.1175/1520-0493\(1996\)124<1225:ASSATE>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2).
- Huai, X. W., J. P. Li, R. Q. Ding, J. Feng, and D. Q. Liu, 2017: Quantifying local predictability of the Lorenz system using the nonlinear local Lyapunov exponent. *Atmos. Oceanic Sci. Lett.*, **10**, 372–378, <https://doi.org/10.1080/16742834.2017.1346457>.
- Kullback, S., and R. A. Leibler, 1951: On information and sufficiency. *The Annals of Mathematical Statistics*, **22**, 79–86, <https://doi.org/10.1214/aoms/1177729694>.
- Kumar, A., A. G. Barnston, P. T. Peng, M. P. Hoerling, and L. Goddard, 2000: Changes in the spread of the variability of the seasonal mean atmospheric states associated with ENSO. *J. Climate*, **13**, 3139–3151, [https://doi.org/10.1175/1520-0442\(2000\)013<3139:CITSOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<3139:CITSOT>2.0.CO;2).
- 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**, 215–220, <https://doi.org/10.1007/BF02878381>.
- Li, J. P., and R. Q. Ding, 2011: Temporal-spatial distribution of atmospheric predictability limit by local dynamical analogs. *Mon. Wea. Rev.*, **139**, 3265–3283, <https://doi.org/10.1175/MWR-D-10-05020.1>.
- 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**, 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**, 321–333, <https://doi.org/10.1111/j.2153-3490.1965.tb01424.x>.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646, [https://doi.org/10.1175/1520-0469\(1969\)26<636:APARBN>2.0.CO;2](https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2).
- Lorenz, E. N., 1996: Predictability: A problem partly solved. *Proc. ECMWF Seminar on Predictability*, Vol. I, Reading, United Kingdom, ECMWF, 1–18.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliaigis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, <https://doi.org/10.1002/qj.49712252905>.
- Mu, M., W. S. Duan, and Y. M. Tang, 2017: The predictability of atmospheric and oceanic motions: Retrospect and prospects. *Science China Earth Sciences*, **60**, 2001–2012, <https://doi.org/10.1007/s11430-016-9101-x>.
- Murphy, J. M., 1988: The impact of ensemble forecasts on predictability. *Quart. J. Roy. Meteor. Soc.*, **114**, 463–493, <https://doi.org/10.1002/qj.49711448010>.
- Palmer, T. N., R. Gelaro, J. Barkmeijer, and R. Buizza, 1998: Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.*, **55**, 633–653, [https://doi.org/10.1175/1520-0469\(1998\)055<0633:SVMAAO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<0633:SVMAAO>2.0.CO;2).
- Roulston, M. S., and L. A. Smith, 2002: Evaluating probabilistic forecasts using information theory. *Mon. Wea. Rev.*, **130**, 1653–1660, [https://doi.org/10.1175/1520-0493\(2002\)130<1653:EPFUIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1653:EPFUIT>2.0.CO;2).
- Stephenson, D. B., and F. J. Dolan-Reyes, 2000: Statistical methods for interpreting Monte Carlo ensemble forecasts. *Tellus A*, **52**, 300–322, <https://doi.org/10.3402/tellusa.v52i3.12267>.
- Tang, Y. M., R. Kleeman, and A. M. Moore, 2005: Reliability of ENSO dynamical predictions. *J. Atmos. Sci.*, **62**, 1770–1791, <https://doi.org/10.1175/JAS3445.1>.
- Tang, Y. M., H. Lin, and A. M. Moore, 2008a: Measuring the potential predictability of ensemble climate predictions. *J. Geophys. Res.*, **113**, D04108, <https://doi.org/10.1029/2007JD008804>.
- Tang, Y. M., R. Kleeman, and A. M. Moore, 2008b: Comparison of information-based measures of forecast uncertainty in ensemble ENSO prediction. *J. Climate*, **21**, 230–247, <https://doi.org/10.1175/2007JCLI1719.1>.
- Thompson, P. D., 1957: Uncertainty of initial state as a factor in the predictability of large scale atmospheric flow patterns. *Tellus*, **9**, 275–295, <https://doi.org/10.3402/tellusa.v9i3.9111>.
- 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, Y. J. Zhu, C. H. Bishop, and X. G. Wang, 2006: Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus A*, **58**, 28–44, <https://doi.org/10.1111/j.1600-0870.2006.00159.x>.
- 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>.
- Whitaker, J. S., and A. F. Lough, 1998: The relationship between ensemble spread and ensemble mean skill. *Mon. Wea. Rev.*, **126**, 3292–3302, [https://doi.org/10.1175/1520-0493\(1998\)126<3292:TRBESA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<3292:TRBESA>2.0.CO;2).
- Yoden, S., and M. Nomura, 1993: Finite-time Lyapunov stability analysis and its application to atmospheric predictability. *J. Atmos. Sci.*, **50**, 1531–1543, [https://doi.org/10.1175/1520-0469\(1993\)050<1531:FTLSAA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<1531:FTLSAA>2.0.CO;2).
- Zhu, Y. J., Z. Toth, R. Wobus, D. Richardson, and K. Mylne, 2002: The economic value of ensemble-based weather forecasts. *Bull. Amer. Meteor. Soc.*, **83**, 73–83, [https://doi.org/10.1175/1520-0477\(2002\)083<0073:TEVOEB>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0073:TEVOEB>2.3.CO;2).
- Ziehmann, C., L. A. Smith, and J. Kurths, 2000: Localized Lyapunov exponents and the prediction of predictability. *Physics Letters A*, **271**, 237–251, [https://doi.org/10.1016/S0375-9601\(00\)00336-4](https://doi.org/10.1016/S0375-9601(00)00336-4).