

## CORRESPONDENCE

 Reply to “Comments on ‘What Is the Predictability Limit of Midlatitude Weather?’”

Y. QIANG SUN

*Program in Atmospheric and Oceanic Sciences, Princeton University, and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey*

FUQING ZHANG<sup>a</sup>

*Department of Meteorology and Atmospheric Science, and Center for Advanced Data Assimilation and Predictability Techniques, The Pennsylvania State University, University Park, Pennsylvania*

LINUS MAGNUSSON AND ROBERTO BUIZZA

*European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*

JAN-HUEY CHEN

*NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey*

KERRY EMANUEL

*Lorenz Center, Massachusetts Institute of Technology, Cambridge, Massachusetts*

(Manuscript received 22 November 2019, in final form 14 January 2020)

## ABSTRACT

In their comment, Žagar and Szunyogh raised concerns about a recent study by Zhang et al. that examined the predictability limit of midlatitude weather using two up-to-date global models. Zhang et al. showed that deterministic weather forecast may, at best, be extended by 5 days, assuming we could achieve minimal initial-condition uncertainty (e.g., 10% of current operational value) with a nearly perfect model. Žagar and Szunyogh questioned the methodology and the experiments of Zhang et al. Specifically, Žagar and Szunyogh raised issues regarding the effects of model error on the growth of the forecast uncertainty. They also suggested that estimates of the predictability limit could be obtained using a simple parametric model. This reply clarifies the misunderstandings in Žagar and Szunyogh and demonstrates that experiments conducted by Zhang et al. are reasonable. In our view, the model error concern in Žagar and Szunyogh does not apply to the intrinsic predictability limit, which is the key focus of Zhang et al. and the simple parametric model described in Žagar and Szunyogh does not serve the purpose of Zhang et al.

---

**1. Introduction**

Predictability is a fundamental concept for numerical weather prediction (NWP). As the numerical prediction proceeds, dynamical instabilities and chaotic nonlinear interactions cause the forecast uncertainty to increase until the differences between individual ensemble forecasting members are statistically indistinguishable from random draws from the numerical model’s climate. After that, NWP provides no information that could not be readily obtained from a previously generated model

---

 Denotes content that is immediately available upon publication as open access.

---

<sup>a</sup> Deceased.

---

*Corresponding author:* Y. Qiang Sun, ys5@princeton.edu

climate, which means the predictability is lost under this NWP system. The time interval between the initial time of the integration and the time at which predictability is lost is usually called the *practical predictability limit* for an NWP system (see, e.g., Buizza and Leutbecher 2015). If the model and/or initial-condition errors can be reduced, then this practical predictability limit can be extended. Yet Lorenz (1969, hereafter L69) shows that there is likely an intrinsic finite predictability limit for the atmosphere that cannot be further extended due to very rapid error growth at small scales combined with nonlinear interaction between different scales (so-called butterfly effect). This *intrinsic predictability limit* for our atmosphere is supported and verified by follow-up studies (e.g., Tribbia and Baumhefner 2004; Froude et al. 2013; Selz and Craig 2015; Judt 2018) and now widely accepted.

Zhang et al. (2019, hereafter Z19) studied the predictability limit of midlatitude weather using global convection-permitting models, with a focus on the *intrinsic predictability limit* for day-to-day synoptic weather system. In other words, Z19 tried to address the long-lasting intriguing question since L69: how far our NWP skill could be further extended? The methodology used in Z19 is similar to identical-model twin experiments, where the same numerical model is integrated from nearly identical initial states until predictability is lost. Under a perfect-model assumption, this kind of experiment will provide an estimate for the *intrinsic predictability limit*. This is a common approach in predictability studies (e.g., Tribbia and Baumhefner 2004; Zhang et al. 2003, 2007; Mapes et al. 2008; Judt 2018). Given the assumption of a perfect model, the accuracy of this estimate largely depends on how realistically the model can simulate reality. There is growing evidence showing that moist physics plays a critical role in the upscale error growth (e.g., Zhang et al. 2003; Selz and Craig 2015; Sun and Zhang 2016). Therefore, the ability of a model to resolve small-scale moist physics is critical for examining the growth of the errors and providing an accurate estimate for the intrinsic finite range of the predictability. In addressing this problem, Z19 used the most up-to-date (at that time) high-resolution full-physics global models: the 9-km version of the operational model at ECMWF, and a uniform 3-km experimental finite-volume cubed-sphere model developed at GFDL during the next-generation global prediction system (NGGPS) project in the United States. These models can give us state-of-the-art, intrinsic predictability estimates.

Two types of perturbation experiments were conducted in Z19. One ensemble was started from the first 10 of the 21 available members of the operational

ensemble of 4DVar analyses (EDA) in ECMWF, whereas the other was started from the same analyses after a rescaling of the analysis perturbations by a factor of 10. The former ensemble is representative of our current day initial-condition errors, while the latter was aimed to address the intrinsic predictability limit. Both ECMWF and U.S. global convection-permitting models in Z19 show very consistent results, implying that our current day deterministic weather forecast may be further extended by up to 5 days.

In their comment to Z19, Žagar and Szunyogh (2020, hereafter ZS) raised several concerns about Z19's methodology and conclusions. They first questioned the perturbations used in Z19, arguing EDA-only type perturbations and the rescaling of these perturbations "do not provide realistic simulations of either the current-day operational, or the future ideal evolution of the forecast errors." Another major issue they raised is related to the impacts of model error on the growth of forecast uncertainty. They also suggested the use of a parametric model for estimates of the extension of the practical predictability.

We would like to thank ZS for providing the chance to further clarify our findings in Z19, as many of the concerns in ZS are due to their misunderstanding of the aims of Z19 and the approach adopted there. In our reply here, section 2 shows that the use of EDA perturbations in Z19 is reasonable and consistent. In section 3 we will then show that the model error argument and the suggested parametric model in ZS are invalid. A summary is given in section 4.

## 2. Representation of the analysis and forecast uncertainties

### a. EDA-only perturbations

ZS pointed out that the operational ECMWF ensemble prediction system (ENS) needs to supplement the EDA initial-condition perturbations with singular vector initial-condition perturbations and a parameterization of the effect of model errors to realistically simulate the evolution of the forecast uncertainty (Isaksen et al. 2010; Leutbecher et al. 2017; Haiden et al. 2018). Therefore, ZS suspected that including only a subset of 10 of the EDA perturbations and not including the singular vector perturbations and the parameterization of the effect of model errors, may have hampered the ensemble in Z19 from fully capturing the forecast uncertainty for an extended forecast time.

We want to first note that this argument of ZS applies more to the practical predictability scenario (EDA case

in Z19). The study of Z19, however, focuses on the intrinsic predictability scenario (EDA0.1 case). Motivated by L69, the key question Z19 aims to address is “With a nearly perfect model, to what limit does the predictability horizon of NWP could approach as the initial error approaches zero?” In this case, the use of EDA-only perturbations guarantees that no model error is introduced into the integration, which ensures that the initial growth of the forecast uncertainty in EDA0.1 is largely due to the butterfly effect idea Lorenz originally proposed. On the contrary, both the singular vector perturbations and the model error schemes are developed in part to address deficiencies of the model, which is contrary to the perfect model assumption in the EDA0.1 experiment. If they were adopted, it would be hard for us to know whether the initial growth of the spread in the EDA0.1 scenario is due to these model error schemes or due to the upscale error propagation we focused on in Z19. Hence, not using singular vectors and model error schemes are justified in experiments targeting intrinsic predictability.

By using a subset of EDA perturbations, Z19 not only wants to “observe the process of forecast uncertainty growth without the influence of the unique dynamics of singular vectors” as mentioned in ZS but also tries to be more consistent with the intrinsic predictability scenario (the main focus of Z19). While we acknowledge that using only 10 EDA members is not the best way of estimating the practical predictability, we show that the use of 10 members is also reasonable for the study of practical predictability in Z19. As we do not have the computer resources to rerun the experiments, we here explore the THORPEX International Grand Global Ensemble (TIGGE) dataset to better represent forecast uncertainty growth in the operational centers (Bougeault et al. 2010). The TIGGE dataset consists of ensemble forecast data from 10 global NWP centers, starting from October 2006, which has been collected in real time and made available for scientific research. ECMWF in TIGGE dataset consists of one control forecast member and all 50 perturbed forecast members, generated using EDA and singular vector perturbations to simulate initial uncertainties (Buizza et al. 2008), also including stochastic parameterizations to simulate the effect of model errors (Palmer et al. 2009).

Figure 1 shows the evolution of the ensemble variance of 500-hPa winds in the 50 perturbed ensemble members in the TIGGE dataset, calculated using the same method as in Fig. 3 of Z19 for the same summer case. We find that forecast uncertainty in the TIGGE dataset does grow faster in the first few days, which is expected due to the nature of singular vector perturbations

and added model errors. We also see that ensemble variance using only the first 10 of the 50 available TIGGE members produce almost identical results as that using all the 50 members, implying that an ensemble of 10 members is valid for the metric considered in Z19.

It is worth pointing out that, although slower initially, the growth rate of ensemble variance in Z19 shows little difference from that in the 50 member TIGGE (even slightly larger due to less saturation effect) after 2 days. As the nonlinear processes become more important, and the linearly computed singular vectors lose their optimality, the difference between Z19 and TIGGE also becomes smaller and smaller. In summary, we agree that singular vector perturbations could bring faster initial error growth for the EDA case, but we would argue that this has a small impact on the estimation of the practical predictability limit, provided that the estimated limit is beyond this short forecast range, as reported by Z19. Concerning the inclusion of model error schemes in Z19’s ensembles of the practical predictability scenario, we would also argue that including stochastic parameterizations would not have changed fundamentally results of Z19 based on Fig. 1. Results from ECMWF (see, e.g., Palmer et al. 2009) have indeed shown that the inclusion of model error schemes has a small impact, on average and globally, by ~5%–10%.

### b. Rescaling

To investigate the intrinsic predictability limit of midlatitude weather systems, Z19 rescaled the EDA perturbations by a constant factor of 10 to reduce the magnitude of the perturbations and achieved nearly perfect initial conditions. Given that EDA-type perturbations have larger amplitudes in the tropics and at larger scales, the rescaling will lead to, in the absolute sense, more reduction at larger scales and in the tropical region. ZS then suspected that “reduced initial-condition uncertainties in the tropics and at the large scales must have played a major role in the increased predictability in the midlatitudes in the experiments of Z19.” This is speculative. Rescaling the perturbations is one way to investigate the sensitivity to a reduction of the initial-condition errors to a minimum, and thus to assess the intrinsic predictability limit. It is hard to separate the downscale propagation of the large-scale errors and the local error growth in the small scales due to a very short time scale for both processes. In fact, for the intrinsic predictability limit, as long as the amplitude of the perturbation is sufficiently small, the scale dependence of the minute perturbations is not so important. Sun and Zhang (2016) showed that minute perturbations at large scale show a very similar evolution path to the noise type

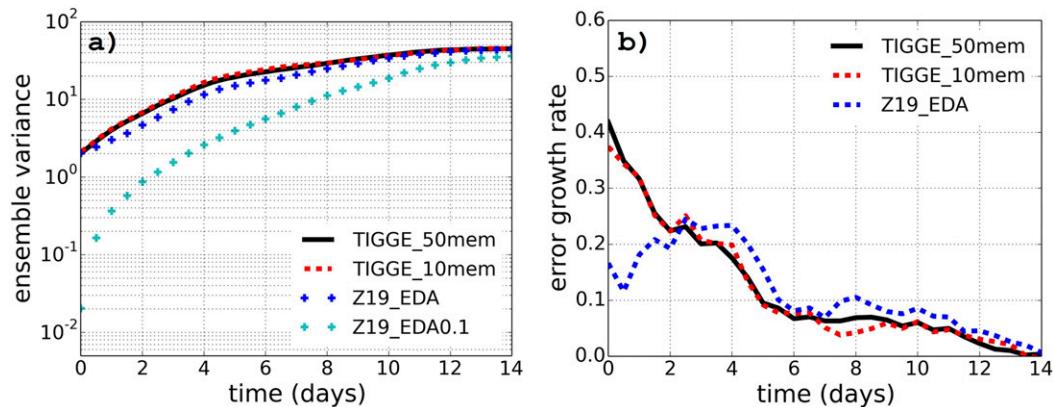


FIG. 1. (a) Evolution of forecast uncertainty growth in terms of ensemble variance of 500-hPa wind energy averaged over the NH midlatitudes (30°–60°N) from TIGGE dataset (black line: all 50 perturbed members; red line: the first 10 members in TIGGE data) and Z19 for June 2016 case. (b) Error growth rate  $[\ln(E_{t+\Delta t}/E_t)]$ , where  $\Delta t = 12$  h calculated using data in (a).

perturbation experiment, which is also seen in Durran and Weyn (2016).

### c. $\beta$ term

ZS discussed also one diagnostic model used in Z19 to study error growth, which is reported here:

$$\frac{d\varepsilon(t)}{dt} = [\alpha\varepsilon(t) + \beta][1 - \varepsilon(t)], \quad (1)$$

where  $\varepsilon(t)$  is the normalized error.<sup>1</sup> One difference between Z19 and some previous studies (e.g., Dalcher and Kalnay 1987; Magnusson and Källén 2013; Žagar et al. 2017) is that  $\varepsilon(t)$  is the ensemble spread for the 10 ensemble members in Z19, which does not contain model error at all. Hence the growth of  $\varepsilon(t)$  in Z19 has zero contribution from model error;  $\beta$  is the “super-exponential” term that represents the intrinsic error growth even if the model is perfect and the initial-condition error  $\varepsilon(0)$  is 0. If  $\beta = 0$ , then Eq. (1) describes a simple exponential error growth, in which our forecast skill could be improved without any limit if we keep reducing the initial-condition error  $\varepsilon(0)$ . We agree with ZS in that similar superexponential error growth can be

found in coarse-resolution full-physics models (Harlim et al. 2005). Otherwise, there would be no intrinsic predictability limit for the weather systems in these coarse-resolution models. Nonetheless, as pointed in Z19 and the introduction, moist physics might be the key to this intrinsic predictability limit. For example, Sun and Zhang (2016) showed that the Gaussian white noise perturbations added to the dry model will decay for the first 36 h, compared to their rapid growth in the moist model. Therefore, to get an accurate estimate for the intrinsic predictability limit, we believe that the global convection-permitting model, which at least partially resolves the moist convection process, is the best tool to use. As stated in Buizza (2010) and Bengtsson et al. (2008), although these simple error-growth models are useful tools to investigate forecast error growth, they have difficulties in describing the error growth at short forecast ranges, and thus should be used with care.

## 3. Model versus initial-condition improvements and the parametric model

### a. Model error

While the *practical predictability limit* for an NWP system is greatly affected by model errors and model bias, it is common to adopt the perfect model assumption when estimating the *intrinsic predictability limit*. As pointed in the introduction and Z19, the accuracy of the estimation of the intrinsic predictability limit does depend strongly on how well the model captures physical processes in the real world. This is also the reason that Z19 chose two of the most up-to-date global models available at that time. Consistency between these two very different models strengthens our confidence in the

<sup>1</sup> We normalize  $\varepsilon(t)$  as  $\varepsilon(t) = (E/E_{\max})$ , where  $E$  and  $E_{\max}$  are defined in Eq. (2);  $\beta$  here is also normalized by  $E_{\max}$ . Z19 did not present the value of  $E_{\max}$ . However, it can be seen from the top panels of Figs. 3 and 4 in Z19 that the errors in both EDA and EDA0.1 experiments approach the same saturation value at the end of the integration. The value of  $E_{\max}$  is calculated as the average of  $E$  in EDA and EDA0.1 for the last day (day 20). The value of  $E_{\max}$  is larger in the winter than that of the summer, on the order of  $\sim 100 \text{ m}^2 \text{ s}^{-2}$  (Figs. 3a,b of Z19). Also, the same parameter is used to fit both EDA and EDA0.1 experiments in Z19.

results reported in Z19. Moreover, some systematic model bias does not necessarily change the intrinsic predictability of the system. For example, a trivial dislocation of small mountains will certainly generate a model bias, but it shall not change the intrinsic predictability of the flow. Nonetheless, there is still a possibility that the scenario and estimation for the intrinsic limit would be different once the grid spacing of the global model drops below about 1 km and moist convection becomes fully resolved. New experiments with more advanced models should be conducted to answer this question.

Z19 stated that achieving the “up to 5 days” potential gain “requires coordinated efforts by the entire community to design *better numerical weather models* [emphasis added], to improve observations, and to make better use of observations with advanced data assimilation and computing techniques.” Although the requirement of reducing model error is implied by this statement and the perfect model assumption in Z19, Z19 did not explicitly state it. We welcome the opportunity to do it here to avoid misunderstandings as in ZS. Reducing initial-condition error alone will surely not achieve this potential “up to 5 days” gain reported in Z19.

The best way to evaluate the impact of model error is running experiments using different models under the same initial and boundary conditions and then compare the model performance (Magnusson et al. 2019). This kind of experiment is usually conducted for continuous operational model development (Haiden et al. 2018). Yet, as model development is generally done piece by piece and takes years, sometimes a decade, to see significant improvements in the weather forecast skill of an NWP system, very few systematic results are published.

*b. Parametric model in ZS*

We agree with ZS that parametric analytical models for the average growth of the forecast uncertainty are useful. Equation (1), which is used in Z19, also serves this purpose. Yet one needs to be cautious not to overinterpret the results of these parametric models, as the parameters are fitted and could be affected by many factors.

For convenience, the model used in Žagar et al. (2017) is listed here [their Eq. (12)]:

$$\frac{dE}{dt} = s(E_{\max} - E)(E - E_{\min}), \tag{2}$$

where  $E$  is the forecast uncertainty,  $E_{\max}$  and  $E_{\min}$  are the maximal and minimal values of the model function according to Žagar et al. (2017) and  $E \in [E_{\min}, E_{\max}]$ . It can be shown that Eq. (2) is just a variant of Eq. (1).

They are equivalent mathematically. Simply let  $\alpha = sE_{\max}$  and  $\gamma = -sE_{\max}E_{\min}$ , then Eq. (2) becomes

$$\frac{dE}{dt} = \alpha \left(1 - \frac{E}{E_{\max}}\right) \left(E + \frac{\gamma}{\alpha}\right) = (\alpha E + \gamma) \left(1 - \frac{E}{E_{\max}}\right). \tag{3}$$

This is similar to Eq. (1), except that Eq. (1) is normalized by  $E_{\max}$ .

Whereas similar mathematical models are used in ZS and Z19, the application of the models and the interpretation are very different. Given that Žagar et al. (2017) and ZS applied this model to the operational dataset, model error plays a role in their study. In the experiment of Z19, as discussed above, the growth of the forecast uncertainty has no contribution from model error; therefore,  $\beta$  is mainly interpreted as the intrinsic upscale error propagation from small scales.

ZS applied the model to different horizontal scales (each different horizontal wavenumber). Figure 1 of ZS also compared their fitted results of June 2018 and May 2015, claiming that more improvement is achieved at subsynoptic and even smaller scales. However, this is not a fair comparison. The forecast skill of operational models has a large variation for different months, showing a strong seasonal cycle (Haiden et al. 2018). The seasonal cycle signal is much larger than the improvement observed in the past few years. The differences between results for June 2018 and May 2015 shown in Fig. 1 of ZS are likely due to different large-scale flow patterns and different forcing for these two periods. A valid comparison would be running the updated model again under May 2015 initial conditions to show the impact of model improvement. Alternatively, the old model in 2015 could be used to run the 2018 initial conditions for a better understanding of the impact of reduced initial conditions.

It is also questionable trying to use this parametric model to predict the effect of reducing initial-condition errors as shown in Fig. 2 of ZS. ZS processes each horizontal scale separately. However, given the strong interactions between different horizontal scales, it is almost meaningless to examine the effect of reducing the initial-condition error for a certain horizontal scale only (as in Fig. 2 of ZS) without knowing the information of the initial-condition errors at other scales. For example, if we reduce the initial-condition error only at  $k = 35$  while keeping the initial-condition errors at other horizontal scales the same, then it is most likely we will not see any improvement in the forecast skill for this same horizontal scale ( $k = 35$ ). The strong cross-scale interaction will very quickly fill the initial small gap

at  $k = 35$ , making the reduction at this certain scale negligible. This example also indicates that the error growth rate at a certain horizontal scale (e.g.,  $k = 35$ ) is affected by the errors at this scale and other horizontal scales. Therefore, reducing the initial-condition errors will surely change the error growth rate and the associated parameters in the parametric model. Estimates made based on fixed parameters as in [ZS](#) then become less convincing.

A similar yet more meaningful problem would be, How much gain shall we expect for each scale if the initial-condition errors are reduced by a certain percentage (e.g., 75%) at all scales? This would be close to the example given in Fig. 2 of [ZS](#).<sup>2</sup> However, it is very difficult to provide an answer to this kind of question using a simple parametric equation due to strong cross-scale interactions. We will briefly illustrate this next. The calculation in [ZS](#) used fixed parameters for their parametric model. Hence the evolution of the errors in their model only depends on the amplitude of the errors. The additional gain calculated in [ZS](#) is therefore simply the time needed for the reduced errors to grow back to their original values. Assume we have these estimates for different horizontal scales as calculated in [ZS](#), defined as  $t_1, t_2, \dots, t_n$  for wavenumber  $k_1, k_2, \dots, k_n$ . Initially, errors at all the scales are reduced (e.g., by 75%), after a certain time  $t_m$  (e.g., the median number of  $t_1, t_2, \dots, t_n$ ), then errors at some scales (with estimated time that is less than  $t_m$ ) will grow larger than their original values according to the fixed parametric model, whereas errors at other scales would still be smaller than their original values. The distribution of the errors would then be very different from the original error distribution unless  $t_1 = t_2 = \dots = t_n$ . The error growth rate afterward would soon be very different due to cross-scale interaction, which correspondingly shall change the parameters of this parametric model. Hence, the estimation made based on fixed parameters becomes invalid. In fact, the best way to answer this question is by using numerical models as in [Z19](#). [Z19](#) also provided qualitatively answers to this kind of question, implying that we shall see more gain at relatively larger scales, consistent with a longer intrinsic predictability limit at larger scales. After all, one should never expect a deterministic model to forecast the genesis of a tornado (with horizontal scales of hundreds of meters) at a specific location 1 day ahead.

In summary, we disagree with the conclusions drawn in [ZS](#) using the parametric model due to strong interactions between different scales. It is also difficult to relate them to the results of [Z19](#).

#### 4. Summary

[Z19](#) investigated the predictability limit of midlatitude weather systems, with a focus on the intrinsic predictability limit of a deterministic forecast. Using the 9-km model currently running operationally at ECMWF and the 3-km experimental model developed in GFDL, [Z19](#) concluded that “assuming the current-generation state-of-the-science NWP models could capture the most essential physical processes in the real world, we can further improve the forecast accuracy of day-to-day weather events by up to 5 days.” Achieving this additional potential gain requires continued coordinated efforts by the entire community to design more accurate and better NWP models (reducing model error), improve and enhance the observing techniques and networks, and make better use of observations with advanced data assimilation (reducing initial-condition error).

[ZS](#) criticizes that [Z19](#) used only a subset of 10 of the EDA perturbations and did not include the singular vector perturbations and parameterizations of model errors. We think that the [ZS](#)'s comments apply more to practical predictability than to intrinsic predictability, which was the focus of [Z19](#). For the intrinsic predictability limit (EDA0.1 experiment in [Z19](#)), adding singular vector perturbations and stochastic schemes will complicate our interpretation of the initial growth of the errors and is not really relevant in addressing our central question. Using a perfect model assumption and running the global convection-permitting model is, in our opinion, the best way to get an accurate estimate for the intrinsic predictability limit. We further note that using a perfect model assumption in the intrinsic predictability experiments also means that the “up to 5 days” potential gain reported in [Z19](#) for midlatitude synoptic weather system requires a reduction of both the model error and the initial-condition errors. For the practical predictability (EDA experiment in [Z19](#)), we agree with [ZS](#) that the choice of perturbation methods and the model error have an impact. Indeed, ECMWF keeps using singular vectors to generate their ensemble's initial perturbations to improve the ensemble reliability in the short forecast range, as they documented in several papers. Nonetheless, through comparison with results using the TIGGE dataset for the same event, we argue that the use of 10 EDA members in [Z19](#) is reasonable and we show that different

<sup>2</sup>Figure 2 of [ZS](#) shows their results for each horizontal scale when the initial-condition errors are reduced by 75% and 99%.

perturbation methods and model errors mainly influence the forecast uncertainty in the first 2 days. The authors also disagree with ZS about the suggestion of using the parametric model in ZS to predict the effects of reducing the initial-condition errors due to strong interactions between different scales.

*Acknowledgments.* The ECMWF services are gratefully acknowledged for the access to the ENS data. The authors thank Baoqiang Xiang and Kun Gao for providing useful comments on an earlier version. Y. Qiang Sun is funded under Award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce.

## REFERENCES

- Bengtsson, L. K., L. Magnusson, and E. Kallen, 2008: Independent estimations of the asymptotic variability in an ensemble forecast system. *Mon. Wea. Rev.*, **136**, 4105–4112, <https://doi.org/10.1175/2008MWR2526.1>.
- Bougeault, P., and Coauthors, 2010: The THORPEX Interactive Grand Global Ensemble. *Bull. Amer. Meteor. Soc.*, **91**, 1059–1072, <https://doi.org/10.1175/2010BAMS2853.1>.
- Buizza, R., 2010: Horizontal resolution impact on short- and long-range forecast error. *Quart. J. Roy. Meteor. Soc.*, **136**, 1020–1035, <https://doi.org/10.1002/QJ.613>.
- , and M. Leutbecher, 2015: The forecast skill horizon. *Quart. J. Roy. Meteor. Soc.*, **141**, 3366–3382, <https://doi.org/10.1002/QJ.2619>.
- , —, and L. Isaksen, 2008: Potential use of an ensemble of analyses in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **134**, 2051–2066, <https://doi.org/10.1002/QJ.346>.
- Dalcher, A., and E. Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. *Tellus*, **39A**, 474–491, <https://doi.org/10.3402/tellusa.v39i5.11774>.
- Durran, D. R., and J. A. Weyn, 2016: Thunderstorms do not get butterflies. *Bull. Amer. Meteor. Soc.*, **97**, 237–243, <https://doi.org/10.1175/BAMS-D-15-00070.1>.
- Froude, L. S. R., L. Bengtsson, and K. I. Hodges, 2013: Atmospheric predictability revisited. *Tellus*, **65A**, 19022, <https://doi.org/10.3402/tellusa.v65i0.19022>.
- Haiden, T., M. Janousek, J. Bidlot, R. Buizza, L. Ferranti, F. Prates, and F. Vitart, 2018: Evaluation of ECMWF forecasts, including the 2018 upgrade. ECMWF Tech. Memo. 831, 54 pp., <https://www.ecmwf.int/node/18746>.
- Harlim, J., M. Oczkowski, J. A. Yorke, E. Kalnay, and B. R. Hunt, 2005: Convex error growth patterns in a global weather model. *Phys. Rev. Lett.*, **94**, 228501, <https://doi.org/10.1103/PHYSREVLETT.94.228501>.
- Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher, and L. Raynaud, 2010: Ensemble of data assimilations at ECMWF. ECMWF Tech. Memo. 636, 48 pp., <https://www.ecmwf.int/node/10125>.
- Judt, F., 2018: Insights into atmospheric predictability through global convection-permitting model simulations. *J. Atmos. Sci.*, **75**, 1477–1497, <https://doi.org/10.1175/JAS-D-17-0343.1>.
- Leutbecher, M., and Coauthors, 2017: Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quart. J. Roy. Meteor. Soc.*, **143**, 2315–2339, <https://doi.org/10.1002/qj.3094>.
- Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, <https://doi.org/10.3402/tellusa.v21i3.10086>.
- Magnusson, L., and E. Källén, 2013: Factors influencing skill improvements in the ECMWF forecasting system. *Mon. Wea. Rev.*, **141**, 3142–3153, <https://doi.org/10.1175/MWR-D-12-00318.1>.
- , J. Chen, S. Lin, L. Zhou, and X. Chen, 2019: Dependence on initial conditions vs. model formulations for medium-range forecast error variations. *Quart. J. Roy. Meteor. Soc.*, **145**, 2085–2100, <https://doi.org/10.1002/qj.3545>.
- Mapes, B., S. Tulch, T. Nasuno, and M. Satoh, 2008: Predictability aspects of global aqua-planet simulations with explicit convection. *J. Meteor. Soc. Japan*, **86A**, 175–185, <https://doi.org/10.2151/jmsj.86A.175>.
- Palmer, T. N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. J. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic parametrization and model uncertainty. ECMWF Tech. Memo. 598, 42 pp., <http://www.ecmwf.int/sites/default/files/elibrary/2009/11577-stochastic-parametrization-and-model-uncertainty.pdf>.
- Selz, T., and G. C. Craig, 2015: Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. *Mon. Wea. Rev.*, **143**, 813–827, <https://doi.org/10.1175/MWR-D-14-00140.1>.
- Sun, Y., and F. Zhang, 2016: Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. *J. Atmos. Sci.*, **73**, 1419–1438, <https://doi.org/10.1175/JAS-D-15-0142.1>.
- Tribbia, J. J., and D. P. Baumhefner, 2004: Scale interactions and atmospheric predictability: An updated perspective. *Mon. Wea. Rev.*, **132**, 703–713, [https://doi.org/10.1175/1520-0493\(2004\)132<0703:SIAAPA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0703:SIAAPA>2.0.CO;2).
- Žagar, N., and I. Szunyogh, 2020: Comments on “What is the predictability limit of midlatitude weather?” *J. Atmos. Sci.*, **77**, 781–785, <https://doi.org/10.1175/JAS-D-19-0166.1>.
- , M. Horvat, Z. Zaplotnik, and L. Magnusson, 2017: Scale-dependent estimates of the growth of forecast uncertainties in a global prediction system. *Tellus*, **69A**, 1287492, <https://doi.org/10.1080/16000870.2017.1287492>.
- Zhang, F., C. Snyder, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, **60**, 1173–1185, [https://doi.org/10.1175/1520-0469\(2003\)060<1173:EOMCOM>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<1173:EOMCOM>2.0.CO;2).
- , N. Bei, R. Rotunno, C. Snyder, and C. C. Epifanio, 2007: Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. *J. Atmos. Sci.*, **64**, 3579–3594, <https://doi.org/10.1175/JAS4028.1>.
- , Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077–1091, <https://doi.org/10.1175/JAS-D-18-0269.1>.