

# Insights into Atmospheric Predictability through Global Convection-Permitting Model Simulations<sup>✉</sup>

FALKO JUDT

*Advanced Study Program, National Center for Atmospheric Research, Boulder, Colorado*

(Manuscript received 21 November 2017, in final form 13 February 2018)

## ABSTRACT

Global convection-permitting models enable weather prediction from local to planetary scales and are therefore often expected to transform the weather prediction enterprise. This potential, however, depends on the predictability of the atmosphere, which was explored here through identical twin experiments using the Model for Prediction Across Scales. The simulations were produced on a quasi-uniform 4-km mesh, which allowed the illumination of error growth from convective to global scales. During the first two days, errors grew through moist convection and other mesoscale processes, and the character of the error growth resembled the case of  $k^{-5/3}$  turbulence. Between 2 and 13 days, errors grew with the background baroclinic instability, and the character of the error growth mirrored the case of  $k^{-3}$  turbulence. The existence of an error growth regime with properties similar to  $k^{-5/3}$  turbulence confirmed the radical idea of E. N. Lorenz that the atmosphere has a finite limit of predictability, no matter how small the initial error. The global-mean predictability limit of the troposphere was estimated here to be around 2–3 weeks, which is in agreement with previous work. However, scale-dependent predictability limits differed between the divergent and rotational wind component and between vertical levels, indicating that atmospheric predictability is a more complex problem than that of homogeneous, isotropic turbulence. The practical value of global cloud-resolving models is discussed in light of the various aspects of atmospheric predictability.

## 1. Introduction

Over the last decade, a tremendous increase in computing power has facilitated the advent of global convection-permitting numerical weather prediction (NWP) models (GCPMs). GCPMs are able to simulate Earth's atmosphere with astonishing realism and allow the prediction of weather seamlessly from local to planetary scales (e.g., Satoh et al. 2008; Putman and Suarez 2011; Miyamoto et al. 2013; Skamarock et al. 2014; Heinzeller et al. 2016). Not surprisingly, GCPMs are often expected to revolutionize weather prediction, for example, by predicting high-impact weather up to two weeks ahead (ECMWF 2016). However, many aspects of the atmosphere's predictability are not well understood, especially processes that involve interactions across a wide range of scales. Consequently, it is not clear what forecast problems are potentially tractable and how

GCPMs may be used in practice. This study addresses this issue by exploring the predictability of the atmosphere in the context of GCPM simulations.

Atmospheric flow is extremely complex, which hampers efforts to comprehensively quantify its predictability. To keep the problem manageable, predictability has often been studied in simplified settings, for example, by using idealized numerical experiments (e.g., Lorenz 1969; Métais and Lesieur 1986; Rotunno and Snyder 2008). In fact, most of our knowledge about the predictability of fluid flow is based on idealized flows and theory, provoking questions about the degree to which this knowledge applies to the real atmosphere.

One way to explore atmospheric predictability in a more realistic framework is to employ either global or regional NWP models (e.g., Lorenz 1982; Zhang et al. 2002; Tribbia and Baumhefner 2004; Selz and Craig 2015). Yet both global and regional model studies have suffered from distinct shortcomings. For example, classic global models with grid spacings of >10 km were generally not able to explicitly resolve mesoscale processes such as moist convection—a major disadvantage because convection is the principal process associated with the initial

<sup>✉</sup> Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JAS-D-17-0343.s1>.

Corresponding author: Falko Judt, fjudt@ucar.edu

growth of forecast error (Zhang et al. 2003). Regional models, on the other hand, require lateral boundary conditions, which constrain error growth on synoptic scales. The constrained error growth in turn leads to artificially enhanced predictability estimates (Errico and Baumhefner 1987; Vukicevic and Errico 1990).

GCPMs are excellent tools to study atmospheric predictability because they combine the high resolution of regional models with the unrestricted geometry of global models. With this advantage in mind, the purpose of this study is to better understand the deterministic predictability of scales of motions generally referred to as weather. Specifically, the goals are 1) to illuminate the error growth process from convective to planetary scales, 2) to compare error growth characteristics with those predicted by theory, and 3) to quantify the atmosphere's predictability. Furthermore, this study is intended to provide an update on previous low-resolution global model predictability studies (e.g., Lorenz 1982; Tribbia and Baumhefner 2004; Simmons and Hollingsworth 2002) and complement a recent series of high-resolution predictability studies that employed regional models (Durran and Gingrich 2014; Selz and Craig 2015; Durran and Weyn 2016; Weyn and Durran 2017). Given the study's focus on *intrinsic predictability* (Lorenz 1996), model error and initial condition error with realistic amplitude are not considered. Consequently, the results provide an upper bound on what we can possibly predict. Moreover, this study does not examine the predictability of processes from seasonal to subseasonal time scales nor the predictability of average quantities such as monthly means (Shukla 1981). Finally, questions regarding the ocean's effect on atmospheric predictability cannot be addressed because the present model is not coupled to an ocean model.

This paper is structured as follows: Relevant previous work is discussed in section 2. Section 3 introduces the model and experiment setup, followed by a general description of the simulations in section 4. Sections 5 and 6 discuss error growth and predictability in physical and spectral space, respectively. These two sections also introduce a few novel analytics to the predictability literature, such as a comparison of different error metrics; computations of error doubling times for different error growth regimes; and a quantitative analysis of predictability limits as a function of scale, altitude, and underlying flow dynamics. The paper closes with the summary and conclusions in section 7.

## 2. Previous literature

The first classic predictability study was conducted by Thompson (1957), who explored error growth in a simple

barotropic model and noted that small-scale errors do not necessarily foil the prediction of large-scale motions. Made at a time when NWP was still in its infancy, Thompson's conclusion raised hope for the possibility of accurate long-term predictions of synoptic-scale weather systems. This optimistic view was soon challenged by Lorenz (1963), who employed a highly simplified model of atmospheric convection to show that even the smallest errors eventually lead to the loss of predictability of the entire system. Lorenz (1963) postulated that if the atmosphere behaved like the simple system he studied, accurate long-term weather predictions would not be possible.

Intrigued by the question of how long the weather can be predicted, Lorenz studied error growth and predictability in more fundamental ways. Using a spectral turbulence model, Lorenz (1969) demonstrated that the predictability of homogeneous isotropic turbulence depends on the logarithmic slope of the flow's kinetic energy spectrum. Specifically, flows whose spectral slope is shallower than  $-3$  have a finite intrinsic limit of predictability. In this case, error growth is scale dependent (i.e., errors on progressively smaller scales grow progressively faster). Moreover, the error growth rate is time dependent, slowing monotonically as the error saturates on progressively larger scales. In contrast, flows whose kinetic energy spectrum falls off with  $-3$  or steeper have formally unlimited predictability. In the case of a  $-3$  spectrum, error growth is not scale dependent and the growth rate is constant, meaning that predictability can be extended arbitrarily long by making the initial error sufficiently small. (For a spectrum with a slope steeper than  $-3$ , error growth is scale dependent again, but small-scale errors grow slower than large-scale errors, and therefore, these types of flows also have unlimited predictability.) Over the years, turbulence models of varying degrees of sophistication have confirmed Lorenz's theory, which has become a universally accepted tenet in theoretical meteorology (e.g., Leith and Kraichnan 1972; Métais and Lesieur 1986; Boffetta et al. 1997; Rotunno and Snyder 2008; Durran and Gingrich 2014).

Up to this day, however, it remains unclear how predictability theory—which strictly speaking only applies to homogenous isotropic turbulence—relates to the real atmosphere (e.g., Harlim et al. 2005; Ngan and Eperon 2012). In particular, the question of whether the atmosphere exhibits an intrinsic limit of predictability has not been conclusively answered (Tennant 2009; Palmer et al. 2014). Theory implies that *mesoscale motions* have limited predictability, because the atmospheric kinetic energy spectrum follows a power law close to  $k^{-5/3}$  at the mesoscales ( $k$  is horizontal wavenumber). By the same

argument, *synoptic-scale motions* would have *unlimited predictability*, since the spectrum follows a power law close to  $k^{-3}$  at synoptic scales (Nastrom and Gage 1985).

To better understand how predictability theory relates to the predictability of geophysical flows, scientists have studied error growth in models of intermediate complexity (e.g., Morss et al. 2009; Ngan et al. 2009). These studies generally indicate that certain aspects of predictability theory apply to geophysical flows, whereas other aspects do not. For example, in accord with theory, Morss et al. (2009) demonstrated that quasigeostrophic flow exhibits limited predictability if the slope of the kinetic energy spectrum is shallower than  $-3$ . On the other hand, the error growth behavior in geophysical flows seems to be different from homogenous isotropic turbulence. Specifically, instead of an error cascade from smaller to larger scales (upscale growth), errors in more complex flows tend to grow uniformly at all scales (up-magnitude growth; e.g., Mapes et al. 2008; Ngan et al. 2009; Durrant and Gingrich 2014). Idealized large-domain full-physics simulations, such as those of Waite and Snyder (2013) and Sun and Zhang (2016), have also helped to shed light on the nature of kinetic energy spectra and the relationship between dynamics and predictability.

In the end, the study of atmospheric predictability requires realistic NWP models, especially with regard to quantifying the predictability of the real atmosphere. Most real-world global model studies conducted over the last decades agree that the predictability limit of the atmosphere is about two weeks (e.g., Lorenz 1982; Dalcher and Kalnay 1987; Mapes et al. 2008), although a few studies yielded somewhat longer estimates of up to three weeks (Simmons and Hollingsworth 2002; Buizza and Leutbecher 2015). Several global model studies have also noted that atmospheric error growth does not concur with predictability theory; in particular, predictability does not seem to be limited by upscale growth of initially small-scale errors but rather by the direct excitation and amplification of errors on synoptic scales (Boer 1994; Tribbia and Baumhefner 2004; Ngan and Eperon 2012).

As briefly mentioned in the introduction, the global models used in those studies were limited by low resolution. Specifically, they required the parameterization of convection, and they could not generate the  $k^{-5/3}$  part of the atmospheric kinetic energy spectrum. Hence, important aspects of the atmospheric error growth process were not taken into account.

Indeed, there has been increasing evidence that mesoscale error growth is more in line with the idealized  $k^{-5/3}$  turbulence case. Studies using convection-permitting regional models indicate that errors grow fastest on the

smallest resolved scales, which results in the loss of mesoscale predictability within hours (Zhang et al. 2003, 2007; Selz and Craig 2015; Durrant and Weyn 2016; Weyn and Durrant 2017). The practical consequence of this behavior is the well-known difficulty to forecast convective phenomena, such as tropical cyclones (Sippel and Zhang 2008; Judt et al. 2016) and severe convective storms (e.g., Hawblitzl et al. 2007; Zhang et al. 2015). Because of their restricted domains, however, regional models cannot address error growth from mesoscale to synoptic scales and vice versa or, in other words, error growth across the “kink” that links the  $k^{-5/3}$  and  $k^{-3}$  parts of the spectrum. Furthermore, regional models are almost exclusively employed in case studies focusing on particular mesoscale phenomena, which makes it difficult to generalize the results.

### 3. Methods

#### a. Model and model configuration

The foundation of this study is a set of GCPM simulations produced with the atmospheric component of the Model for Prediction Across Scales (MPAS; Skamarock et al. 2012). MPAS is a global nonhydrostatic NWP model that uses C-grid staggering of the prognostic variables and centroidal Voronoi meshes<sup>1</sup> to discretize the sphere. The model employs a hybrid terrain-following vertical height coordinate (Klemp 2011), which is configured such that horizontal coordinate surfaces are constant height surfaces above approximately 15 km above mean sea level.

This particular study used a quasi-uniform mesh with a mean cell center spacing of 4 km. Specifically, the mesh comprised 36 864 002 cells, most of which were hexagons. Only a few prior studies have employed global models with a comparable or higher horizontal resolution (Miyamoto et al. 2013; Skamarock et al. 2014; Heinzel et al. 2016). The height coordinate was configured with 55 layers, and the model top was at 30 km. Subgrid-scale processes were parameterized with the parameterization schemes listed in Table 1. Of note is the Grell–Freitas convection scheme, a scale-aware cumulus parameterization that enables a smooth transition in the partitioning between parameterized and resolved precipitation (Grell and Freitas 2014; Fowler et al. 2016). On the 4-km mesh, most deep convection is considered resolved, and the scheme produces little parameterized precipitation.

---

<sup>1</sup> Voronoi meshes are unstructured grids that allow for both quasi-uniform discretization of the sphere and local mesh refinement.

TABLE 1. List of physics parameterizations used with MPAS.

Parameterization	Reference
Grell–Freitas convection scheme	Grell and Freitas (2014)
Thompson microphysics scheme	Thompson et al. (2008)
Mellor–Yamada–Nakanishi–Niino (MYNN) boundary layer and surface layer scheme	Nakanishi and Niino (2006, 2009)
Rapid Radiative Trans Model for GCMs (RRTMG) short- and longwave scheme	Iacono et al. (2008)
Noah LSM	Niu et al. (2011), Yang et al. (2011)

### b. Experiment setup

Error growth was investigated by means of identical twin experiments, which is a common approach in predictability studies (e.g., Tribbia and Baumhefner 2004; Zhang et al. 2007; Selz and Craig 2015). This particular experiment comprised three sets of identical twins, that is, a 21-day-long control simulation (CTRL) and three time-staggered perturbed simulations with lengths of 20, 15, and 10 days (Pert-20d, Pert-15d, Pert-10d; Fig. 1). Of course, a larger number of ensemble members would be more desirable, but computational and data storage constraints limited this study to three identical twins. The perturbed runs were staggered in time to investigate the dependence of error growth on the large-scale flow configuration, which evolved during the 20-day experiment period.

As is usual in identical twin experiments, the error was defined as the difference between CTRL, that is, the “truth,” and the perturbed simulations (which were identical to CTRL except for slightly perturbed initial conditions). CTRL was initialized with ERA-Interim fields valid at 0000 UTC 19 October 2012, and a subsequent 24-h spinup period allowed the model to generate initially unresolved scales before the actual 20-day experiment period from 0000 UTC 20 October to 0000 UTC 9 November 2012 (the reason why this period was chosen is given in section 4). The perturbed runs were initialized at 0000 UTC 20 October (Pert-20d), 25 October (Pert-15d), and 30 October 2012 (Pert-10d). Similar to Selz and Craig (2015), the initial conditions of the perturbed runs were created by saving restart files from CTRL and seeding the 3D temperature field in the restart files with small-amplitude Gaussian noise (mean  $\mu = 0$  K, standard deviation  $\sigma = 0.01$  K). This minuscule initial “error” is much smaller than any observational uncertainty and, according to the well-known metaphor, can be thought of as mimicking the effect of butterflies.

### c. Error and predictability metrics

The predictability literature offers a variety of metrics that quantify error, most of which measure the distance between pairs of simulations by computing squared differences. One of these metrics is the difference total energy (DTE; Zhang et al. 2003), which is defined as

$$\text{DTE}(\mathbf{x}, t) = \frac{1}{2} \left( \Delta u^2 + \Delta v^2 + \frac{c_p}{T_r} \Delta T^2 \right). \quad (1)$$

Here,  $\Delta$  indicates a difference between any of the perturbed simulations and CTRL. The variables  $u$ ,  $v$ , and  $T$  have their usual meteorological meanings,  $c_p$  is the heat capacity of dry air at constant pressure ( $1004 \text{ J kg}^{-1} \text{ K}^{-1}$ ), and  $T_r = 287 \text{ K}$  is a reference temperature.

Two other error metrics were used in this study, namely, the difference kinetic energy of the 10-m wind ( $\text{DKE}_{10\text{m}}$ ), and the root-mean-square error of the 500-hPa geopotential height field ( $\text{Z500}_{\text{RMSE}}$ ). The  $\text{DKE}_{10\text{m}}$  was computed analogously to DTE (excluding the temperature term), and  $\text{Z500}_{\text{RMSE}}$  was computed according to

$$\text{Z500}_{\text{RMSE}}(t) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\Delta Z_{500})^2}. \quad (2)$$

The  $\text{Z500}_{\text{RMSE}}$  is a legacy metric that has been frequently used in global model predictability studies (e.g., Lorenz 1982; Simmons and Hollingsworth 2002; Buizza and Leutbecher 2015). One reason for evaluating  $\text{DKE}_{10\text{m}}$  and  $\text{Z500}_{\text{RMSE}}$  in addition to DTE is to test whether predictability depends on a specific metric.

The limit of predictability is usually defined as the forecast time at which the error saturates. For flows that completely decorrelate, such as idealized turbulence, the error saturation limit is twice the variance of the flow itself. For flows with climatological components, such as atmospheric flow, the saturation limit is customarily defined as

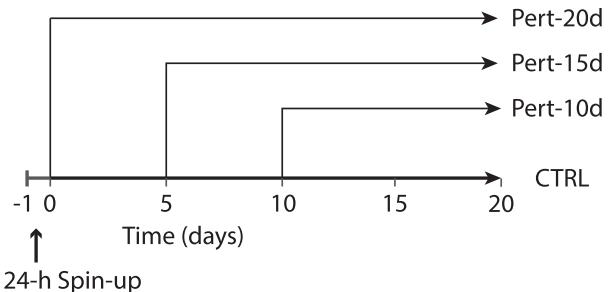


FIG. 1. Schematic of the four simulations that compose the predictability experiment. The time from  $-1$  to  $0$  denotes the 24-h spinup period.

twice the climatological variance. In the latter case, the variance of climatological components is excluded since their prediction is usually not considered skillful.

The climatological variance defining the saturation limit for DTE is given by

$$\text{var}_{uvT}(\mathbf{x}, t) = \text{var}(u) + \text{var}(v) + \frac{c_p}{T_r} \text{var}(T). \quad (3)$$

Here,  $\text{var}(u)$ ,  $\text{var}(v)$ , and  $\text{var}(T)$  are the variances of the zonal and meridional wind components and temperature, computed over the 30-yr period of 1987–2016 using ERA-Interim data. More specifically, 3D variance fields were computed for 0000 UTC of each day in October and November, and the resulting 61 fields were averaged to obtain a single 3D variance field representative of the experiment period. Climatological variances defining the saturation limits for  $DKE_{10m}$  and  $Z500_{\text{RMSE}}$  were computed similarly. The reanalysis data are much coarser resolution (80 km) than the model output and therefore do not account for the variance of smaller-scale motions. Potential implications of this disparity are discussed in section 6.

To explore error growth as a function of spatial scale, error fields are usually decomposed spectrally. Here, error kinetic energy spectra and Z500 error variance spectra were computed from the MPAS output following section 3 of Skamarock et al. (2014). Specifically, the unstructured MPAS-native  $u$ ,  $v$ , and Z500 fields were first interpolated to a regular latitude-longitude grid with ~2-km grid spacing. Then a spherical harmonics transform was applied to the interpolated  $u$ ,  $v$ , and Z500 fields to obtain the background spectra, which, multiplied by two, denote the saturation limit (including climatological components). To obtain the error spectra, the spherical harmonics transform was applied to the difference fields  $\Delta u$ ,  $\Delta v$ , and  $\Delta Z500$ . All resulting 2D wavenumber decompositions were summed over spherical harmonics with the same total spherical wavenumber to produce one-dimensional (1D) spectra and truncated at the minimum resolvable wavelength of 8 km. To illuminate the spectral error growth in relation to the governing dynamics, the spherical harmonics representation of the horizontal wind were decomposed into a divergent and a rotational component.

#### 4. Global weather simulated by the 4-km MPAS

October–November 2012 featured elevated global weather activity (Blunden and Arndt 2013), making this period a compelling case for studying atmospheric predictability. Besides powerful extratropical cyclones, October–November 2012 saw the initiation

of a Madden–Julian oscillation event and the development of several tropical cyclones in the Atlantic, western Pacific, and Indian Oceans.

Based on a brief qualitative analysis, the 4-km MPAS simulated the atmosphere quite realistically. The model captured many observed cloud features, and for the untrained eye, it is at first glance difficult to identify the MPAS simulation in a side-by-side comparison with a satellite image (Fig. 2). Specifically, MPAS seems to do a decent job at simulating tropical convection, which emphasizes that the 4-km convection-permitting configuration is adequate in this regard. The agreement is not perfect, and there are some biases; for example, the cloud distribution appears too extensive over water in the ITCZ and too limited in the Amazon. Notwithstanding these biases, the overall realism is relevant given the study's aim to explore atmospheric predictability with "a model that comes as close to nature as currently possible" (R. Rotunno 2017, personal communication). Note that it was not important to produce an accurate forecast, because the model was assumed perfect and CTRL was treated as the truth. Hence, a formal model verification with observations is not part of this study.

Figure 3 illustrates select weather phenomena in more detail and highlights mesoscale processes that past generations of global models were generally not able to resolve. One example is the cellular wind speed pattern around 50°N, 30°W in the cold air advection sector of a strong extratropical cyclone (highlighted by the box in Fig. 3a but much better demonstrated by the 10-m wind speed animation provided in the online supplemental material). This pattern is likely a manifestation of shallow convection often seen in association with cold air moving over a relatively warmer ocean. Another example is the realistic depiction of tropical cyclones. CTRL and Pert-20d captured the full life cycle of Typhoon Son-Tinh, including cyclogenesis, mature phase (Fig. 3b), and landfall in Vietnam. CTRL and Pert-20d also captured the development of Hurricane Sandy in the Caribbean but failed to reproduce the correct track after Sandy moved into the Bahamas (not shown). Finally, the development of afternoon surface cold pools over the Amazon basin demonstrates that the 4-km MPAS is able to explicitly simulate diurnally driven deep convection (Figs. 3c,d). To assess whether the magnitude of the cold pools and the diurnal temperature range agree quantitatively with observations, a more rigorous model evaluation—which is beyond the scope of this study—would be necessary.

#### 5. Error growth in physical space

Maps and time series present a basic overview of the atmospheric error growth process, including an assessment

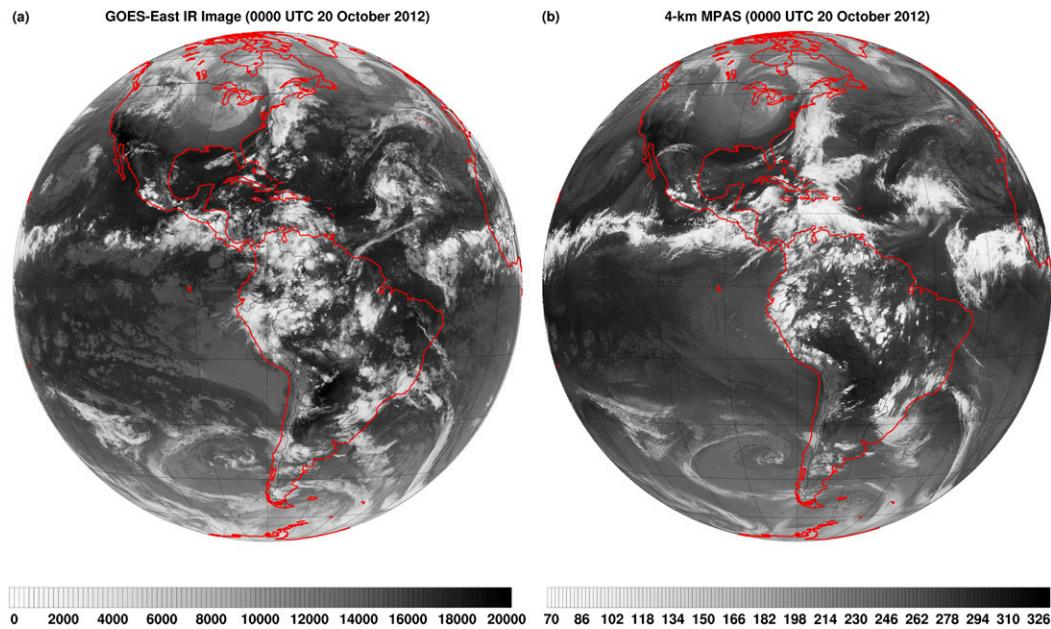


FIG. 2. Comparison of (a) observed and (b) simulated clouds at 0000 UTC 20 Oct 2012; 10- $\mu\text{m}$  channel-scaled radiances from the GOES-East satellite are shown in (a), and outgoing longwave radiation ( $\text{W m}^{-2}$ ) from CTRL is shown in (b).

of the global-mean limit of predictability. Specifically, this section addresses differences in error growth and predictability between the troposphere and stratosphere, the height dependence of error growth in the troposphere, and the impact of different error metrics.

#### *a. Error growth from convective to planetary scales*

The sequence of global DTE maps in Fig. 4 illustrates the tropospheric error growth process in magnitude, scale, and spatial extent over 20 days. During the first 12 h, the initially minuscule error amplified rapidly in convective regions, such as the ITCZ and extratropical fronts (Fig. 4a). Zooming in on the front off the U.S. Atlantic coast in Fig. 5a revealed the quasi-linear shape of the DTE field with embedded cellular maxima, which suggests a close relationship between error growth and precipitating convective bands. During days 1 and 2, the DTE field expanded substantially, mainly because the error spread out beyond the convective zones (Figs. 4b,c). At the same time, the narrow frontal DTE bands became less pronounced and coalesced into a larger-scale feature (Fig. 5b). Sun and Zhang (2016) observed qualitatively similar upscale error growth in idealized simulations of a baroclinic wave. By day 5, the expanding error field had contaminated the entire troposphere, and previously elongated midlatitude mesoscale DTE features that were associated with precipitating frontal zones had expanded into synoptic-scale patches (Fig. 4d). The extratropics experienced considerable error amplification between days 5 and 10, leaving behind a clear DTE minimum in

the tropics (Fig. 4e). Error growth continued beyond day 10, but during the final days, the DTE pattern evolved without a noticeable change in magnitude or scale, indicating that the error growth process had concluded by day 20 (Fig. 4f). The magnitude of DTE had also increased in the tropical belt by this time, but the tropics still represented a DTE minimum.

The error growth process described above concurs with the conceptual error growth model proposed by Zhang et al. (2007): errors initially grow with moist convection, quickly spread through the mesoscales, and eventually contaminate the baroclinic scales. It is not clear, however, what processes cause error growth beyond the mesoscale in the tropics, where no baroclinic instability exists. Error growth processes in the tropics are therefore an excellent topic for future research. An animation visualizing the entire error growth process between 0 and 20 days is available in the supplemental online material. The 12-min time step of the animation highlights the initial significance of rapidly evolving moist convection and exposes the “radiation” of error away from convective systems, which may signify the dispersion of error by gravity waves (Bierdel et al. 2017).

The evolution of stratospheric DTE (Figs. 6 and 7) broadly resembled the tropospheric DTE evolution, although there were some noteworthy differences. Initially, stratospheric DTE was strongly collocated with tropospheric DTE (Figs. 6a,b vs Figs. 4a,b). This suggests that stratospheric errors were induced by tropospheric moist convection, likely through upward

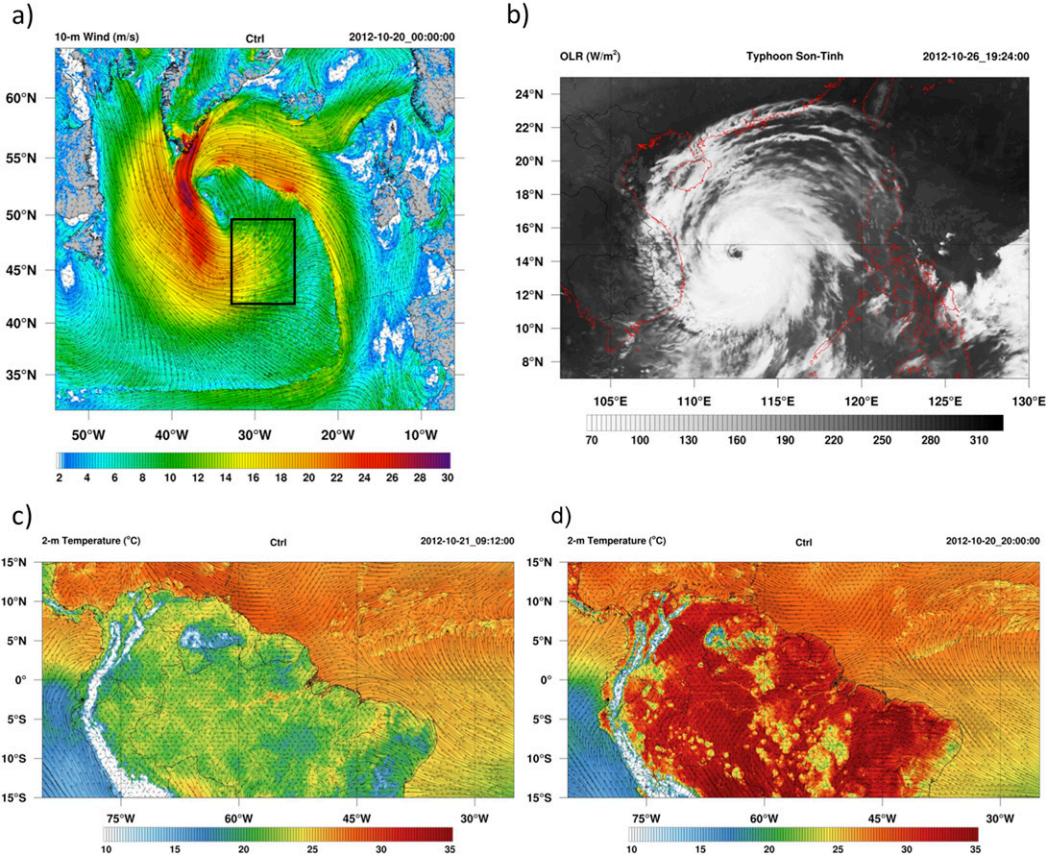


FIG. 3. Examples of meteorological phenomena ranging from synoptic scale to convective scale as simulated by CTRL: (a) extratropical cyclone in the North Atlantic (shading; 10-m wind speed;  $\text{m s}^{-1}$ ), (b) Typhoon Son-Tinh in the South China Sea (shading; outgoing longwave radiation;  $\text{W m}^{-2}$ ), and (c),(d) diurnal minimum and maximum of 2-m temperature (shaded;  $^{\circ}\text{C}$ ) and deep convection over the Amazon basin. Deep convection manifests in the development of cold pools during local afternoon in (d).

propagating gravity waves (Zhang et al. 2007; Ngan and Eperon 2012). The stratospheric DTE field was generally smoother and more diffuse than the tropospheric DTE field, and stratospheric DTE did not exhibit the linear structures resembling precipitating frontal bands (Fig. 7a vs Fig. 5a). Furthermore, errors in the stratosphere seemed to radiate away faster from the convective sources, which is illustrated by the greater areal extent of stratospheric DTE compared to DTE at the same time (Fig. 6b vs Fig. 4b). In contrast with the troposphere, there was also no distinct stratospheric DTE minimum in the tropics on day 5 (Fig. 6e). These qualitative differences between the troposphere and stratosphere indicate that errors grow through distinct physical processes, confirming earlier findings by Ngan and Eperon (2012).

#### b. Evolution of global-mean error

Time series of global, volume-averaged DTE summarize the information discussed above and quantify

error growth in a global-mean sense (Fig. 8). The data are presented in linear (Fig. 8a) and log-linear graphs (Fig. 8b) to better reveal distinct regimes of error growth and highlight the initial growth period when the error magnitude is still small. While much of the discussion in this subsection involves error growth rates, the growth rates themselves are investigated more quantitatively by way of error doubling times in section 5c.

Error growth began with a relatively short initial burst, during which DTE amplified by three to four orders of magnitude (Fig. 8b). Because of the close relationship between error growth and mesoscale processes during this time, in particular moist convection, the early growth phase from 0 to 48 h will be referred to as the *convective-mesoscale phase*. Closer inspection of the time series revealed that the growth rate (i.e., the slope of the lines in Fig. 8b) decreased monotonically during the convective-mesoscale phase, which is characteristic of error growth in  $k^{-5/3}$  turbulence and a hallmark of limited predictability.

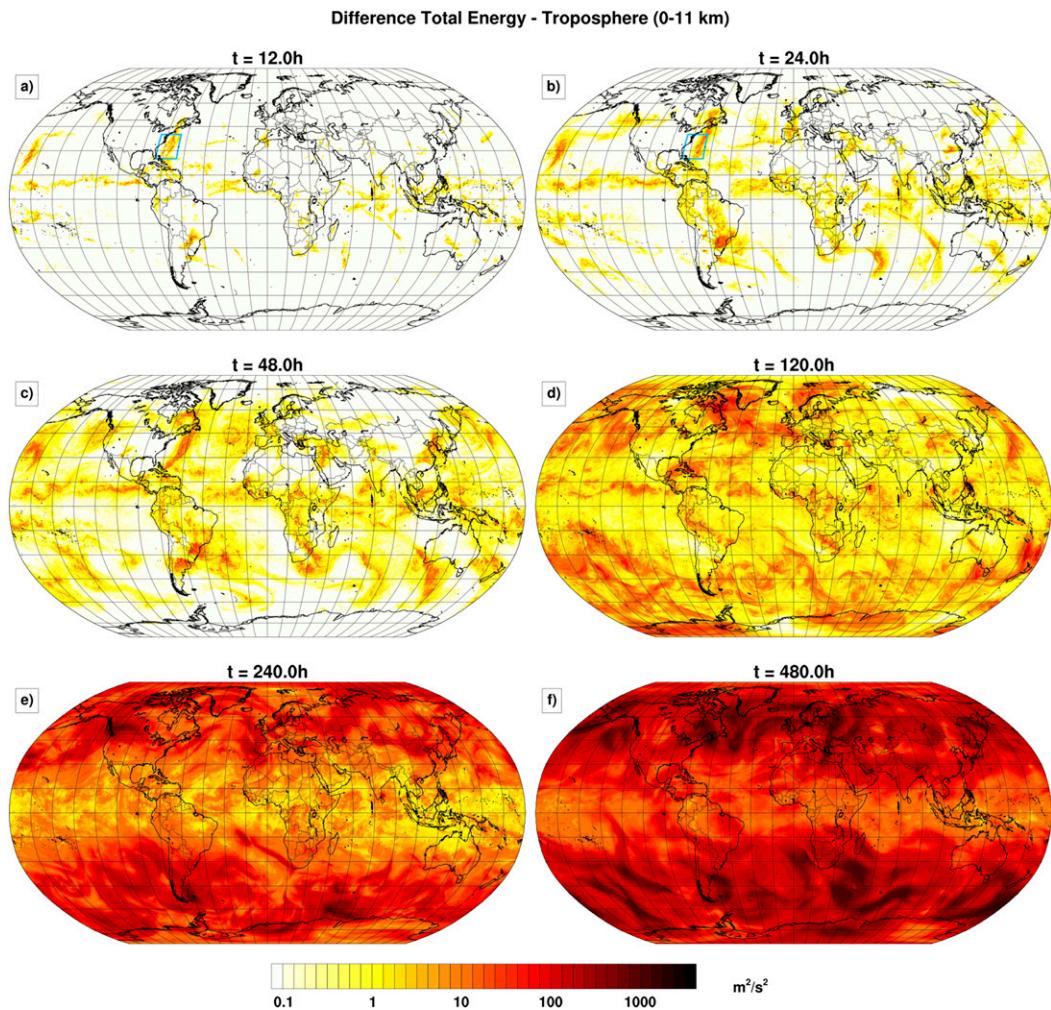


FIG. 4. Sequence of DTE maps illustrating tropospheric error growth over a 20-day period. DTE in this example is based on the twins CTRL/Pert-20d and vertically averaged between 0 and 11 km. Blue rectangles in (a) and (b) outline the zoomed-in area in Fig. 5.

The convective-mesoscale phase was followed by a prolonged phase of quasi-exponential error growth, which seemed to last for about 10–12 days (only Pert-20d was long enough to complete this phase; Fig. 8). The 10–12-day duration and the near-constant growth rate (i.e., near-constant slopes in Fig. 8b) suggest that the error grew with the background baroclinic instability, in agreement with the conceptual model of Zhang et al. (2007) and the much lower-resolution experiments of Tribbia and Baumhefner (2004). Therefore, this period will be referred to as the *baroclinic phase*. Variability between members increased during the baroclinic phase, indicating that the large-scale flow configuration affects the growth rate during this regime.

The tropospheric error growth rate decreased abruptly on day 13, announcing the end of the baroclinic phase (again, only in Pert-20d, in Fig. 8a). Thereafter, DTE

grew unsteadily, reached its saturation limit on day 17 (the predictability limit), attained an overall maximum on day 18, and then decreased. Such fluctuations are typical of errors approaching saturation due to changes in the mean-state kinetic energy (Boffetta and Musacchio 2017). Considering that only one pair of twins exhibited error saturation, one should not take the predictability limit of 17 days too literally, and a vaguer statement like “the tropospheric predictability limit is around 2–3 weeks” seems more appropriate. The latter estimate is consistent with previous studies, especially the more recent work by Ngan and Eperon (2012) and Buizza and Leutbecher (2015).

During the convective-mesoscale phase, the globally averaged *stratospheric* error traced the tropospheric error closely (Fig. 8, orange lines). However, during the quasi-exponential phase, the growth rate of

### Difference Total Energy - Troposphere (0-11 km)

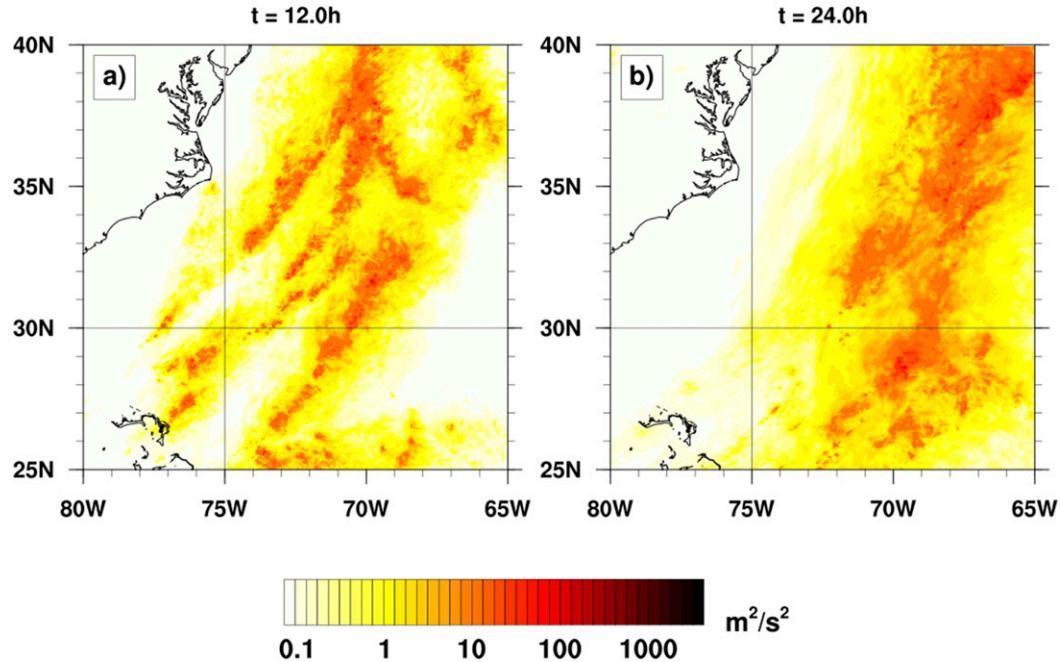


FIG. 5. As in Figs. 4a and 4b, but zoomed in on the midlatitude front over the western North Atlantic.

the stratospheric error was substantially smaller. Consequently, the stratospheric quasi-exponential phase extended beyond 20 days, and the error never saturated. This result implies that the intrinsic predictability limit of stratospheric flow is greater than 20 days, but it is unclear what mechanisms contribute to error growth in the stratosphere during the quasi-exponential phase, given that baroclinic instability plays a lesser role (Ngan and Eperon 2012).

The troposphere and stratosphere evidently differ in error growth and predictability, but it is not apparent whether this is also true for different levels within the troposphere. Error growth turned out to be height dependent, or more specifically, the baroclinic-phase error growth rate increased with height (Fig. 9). However, the climatological variance also increased with height and saturation occurred around day 17 at all levels; therefore, the predictability limit was not a function of height. Given that both  $\text{var}_{uvT}$  and DTE are dominated by the kinetic energy component (85%, not shown), the increase with height is likely because the wind speed increases with height.

The last question to be addressed in this section is whether error growth and predictability depend on the error metric. The answer seems to be no. Aside from differences in their growth rates, both  $Z500_{\text{RMSE}}$  and  $DKE_{10m}$  follow the familiar DTE evolution (Fig. 10). Specifically,

$Z500_{\text{RMSE}}$  and  $DKE_{10m}$  also undergo the convective-mesoscale and baroclinic growth phases, with error saturation on day 17 in Pert-20d. The fact that volume-averaged DTE,  $Z500_{\text{RMSE}}$ , and  $DKE_{10m}$  all saturate at the same time indicates that the troposphere exhibits an unequivocal predictability limit independent of altitude and metric, at least in the simple bulk sense discussed here. This finding somewhat disagrees with Hohenegger and Schär (2007), who noted that meteorological surface variables have shorter predictability than variables in the free troposphere. However, Hohenegger and Schär (2007) used a regional model over complex terrain, which may be the reason for this discrepancy.

#### c. Error doubling times

Error doubling times are another way of quantifying error growth, one that is helpful to estimate the margin for forecast improvement (a doubling time implies that the predictability horizon can be doubled by halving the initial error). Early studies found that the atmosphere's error doubling time is about 5 days, but this number steadily decreased as models became more realistic. In the last two decades or so, doubling times have settled to 1.2–1.7 days (Simmons et al. 1995; Simmons and Hollingsworth 2002; Tribbia and Baumhefner 2004). Here, error doubling times were

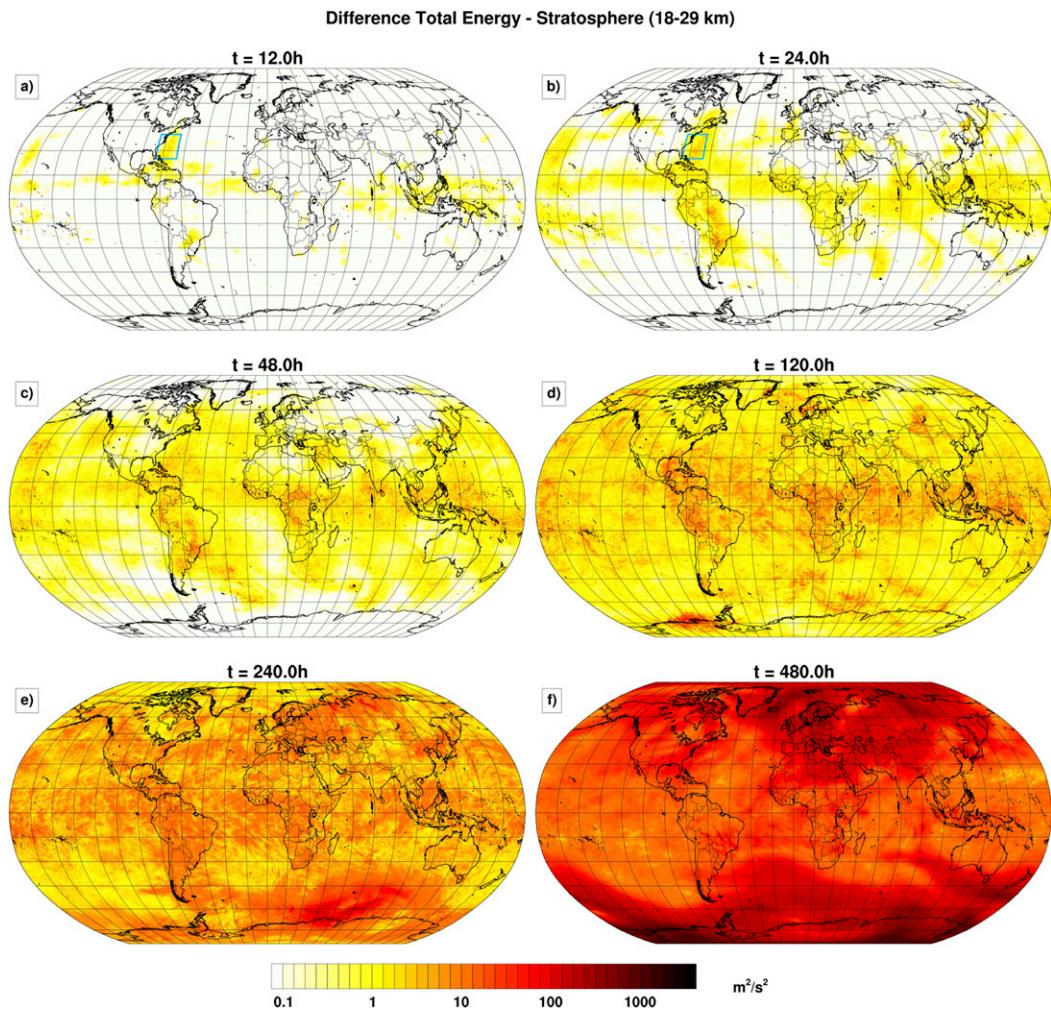


FIG. 6. As in Fig. 4, but for stratospheric DTE vertically averaged between 18 and 29 km.

calculated for tropospheric DTE and  $Z500_{\text{RMSE}}$  according to

$$T_d = (t_2 - t_1) \left\{ \frac{\ln(2)}{\ln \left[ \frac{E(t_2)}{E(t_1)} \right]} \right\}, \quad (4)$$

where  $E$  is either tropospheric DTE or  $Z500_{\text{RMSE}}$  and  $t_2 - t_1 = \Delta t = 12 \text{ h}$ .

DTE doubling times were initially very small ( $<1 \text{ h}$ ) but increased steadily throughout the first two days (Fig. 11, red lines). This increase is a consequence of the decreasing error growth rate during the convective-mesoscale growth phase and typical of flows with limited predictability. To illustrate this point, an initial error doubling time of  $<1 \text{ h}$ , as is the case here, means that decreasing the initial error amplitude by 50% lengthens the global predictability horizon by less than 1 h.

The  $Z500_{\text{RMSE}}$  doubling times differed quite drastically from the DTE doubling times during the first two days and displayed two pronounced diurnal cycles (Fig. 11, turquoise lines). These diurnal cycles, which according to the knowledge of the author have not been reported in the literature, are again evidence that convection plays an important role in the early error growth phase. During the early *baroclinic* phase (roughly between days 2 and 6), both DTE and  $Z500_{\text{RMSE}}$  error doubling times leveled off around 24–36 h before increasing again after day 6. Averaging the doubling times over the core of the baroclinic phase between days 3 and 10 and over the three pairs of twins yielded 39 h (1.6 days) for DTE and 42 h (1.8 days) for  $Z500_{\text{RMSE}}$ . These values are in close agreement with previous studies (Simmons and Hollingsworth 2002; Tribbia and Baumhefner 2004), which indicates a broad consensus: errors growing with the background

### Difference Total Energy - Stratosphere (18-29 km)

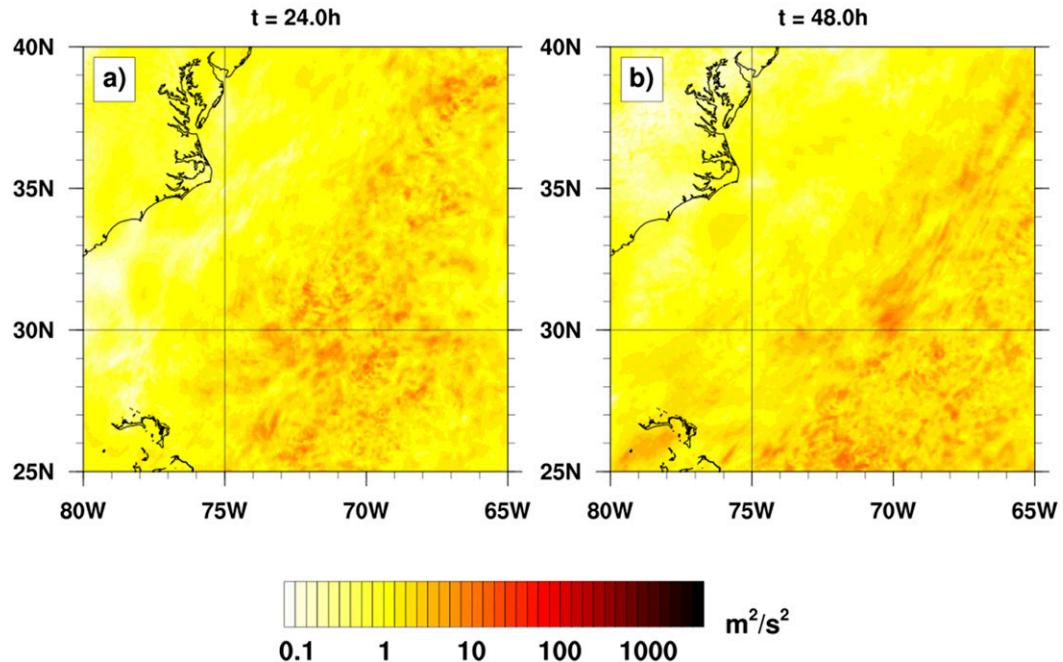


FIG. 7. As in Fig. 5, but for stratospheric DTE vertically averaged between 18 and 29 km.

baroclinic instability double in a little more than 1.5 days. Error doubling times rose more sharply toward the end of the baroclinic phase and fluctuated wildly when the error approached saturation (not shown).

## 6. Error growth in spectral space

The overall goal of this section is to explore the scale dependence of error growth and quantify the atmosphere's scale-dependent predictability limits. Three particular questions are addressed. First, how do error growth and predictability differ between the  $k^{-5/3}$  and  $k^{-3}$  parts of the atmospheric kinetic energy spectrum? Second, how do error growth and predictability differ between the divergent and rotational spectrum? And third, how do error growth and predictability depend on altitude? Additionally, the relationship between atmospheric error growth and predictability theory is discussed. Section 6b advances analytics of previous predictability studies, which were mostly in the form of spectra, by explicitly depicting the predictability limits of the divergent, rotational, and total wind as a function of spatial scale and altitude.

### a. Evolution of error spectra

Figure 12 shows the background and error kinetic energy spectra of the divergent, rotational, and total

wind at three different altitudes. The spectra complement those shown by Weyn and Durrant (2017), who computed similar spectra from a doubly periodic Cartesian domain. The spectra of the total wind are the respective sums of the divergent and rotational spectra and thus dominated by the rotational spectra at larger scales (e.g., Waite and Snyder 2013; Skamarock et al. 2014; Bierdel et al. 2016). Scales smaller than  $6\Delta = 24$  km, to the right of the vertical gray lines, are not fully resolved and will not be considered further.

The background spectra of the rotational and total wind clearly illustrate the transition between the  $k^{-3}$  and  $k^{-5/3}$  segments of the kinetic energy spectrum. In agreement with Skamarock et al. (2014) and Bierdel et al. (2016), the transition occurs at longer wavelengths in the stratosphere than in the troposphere. In the lower stratosphere and upper troposphere (Figs. 12a–f), the divergent spectrum differed markedly from the rotational spectrum. Specifically, the divergent spectrum had a shallower slope and lacked the  $k^{-3}$  segment (Figs. 12a,d). Descending toward the surface, the spectra became generally shallower and the transition between the  $k^{-3}$  and  $k^{-5/3}$  segments became less well defined. In addition, the differences between the divergent and rotational spectra became less obvious. For instance, both the divergent and rotational spectrum of the 10-m wind

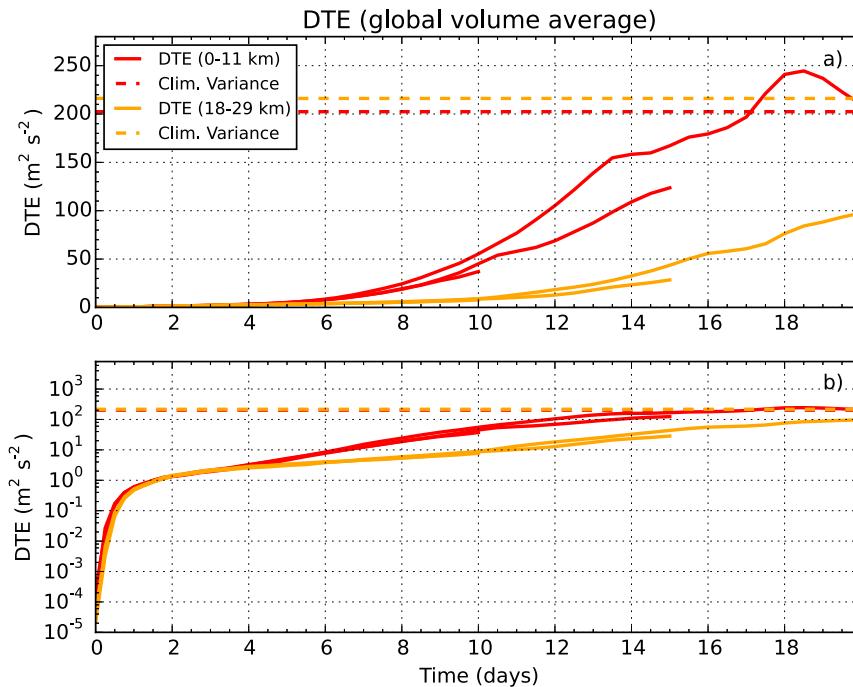


FIG. 8. (a) Time series of volume-averaged DTE from all the three pairs of identical twin experiments (solid lines). Tropospheric DTE is in red, and stratospheric DTE in orange. Dashed lines depict the error saturation limits, which were calculated according to Eq. (3). (b) As in (a), but with a logarithmic y axis.

featured a transition between a steeper and shallower segment (Figs. 12g–i).

The growth of error growth in magnitude and scale is manifested in the progression of error spectra in Fig. 12. The error swept out nearly the entire spectrum within 20 days except for scales of motion larger than about wavenumber 5 (physical scale: 8000 km). According to Boer (1994), the unsaturated large-scale motions represent climatological flow features, such as stationary waves. Thus, the “retained predictability” at scales  $> 8000$  km does not contradict the 17-day predictability limit described in section 5, where predictability was evaluated with respect to climatology. In agreement with many previous studies, the error grew up-magnitude instead of cascading from smaller to larger scales (e.g., Tribbia and Baumhefner 2004; Mapes et al. 2008; Durran and Gingrich 2014).

Superficially, it seems that error growth did not differ substantially between the divergent and rotational component, especially in the troposphere (Figs. 12d–i). However, the shape of the error spectra differed appreciably between the troposphere and the lower stratosphere at 20 km. Specifically, the error spectra in the stratosphere were flatter than the tropospheric error spectra and almost horizontal in the case of the rotational component and total wind. The “flatness” of the

stratospheric error spectra is in agreement with the finding that the stratospheric error spread to larger scales more quickly (section 5a), and the discrepancy between tropospheric and stratospheric error spectra is further evidence that error growth differs between the two layers.

Spectral error growth in the troposphere underwent two distinct phases, consistent with the two phases discussed in section 5. Initially (i.e., during the convective-mesoscale phase), the error peaked at the smallest resolved scales, and the growth rate decreased monotonically (Figs. 12d–i, 13a). Both of these properties are consistent with spectral error growth in the case of  $k^{-5/3}$  turbulence (Fig. 13c), which provides compelling evidence that the atmosphere has indeed a finite limit of predictability. Figures 13a and 13c also differ in certain aspects, because the initial error in Fig. 13a is white noise, whereas the initial error in Fig. 13c is saturated at the smallest scale and zero everywhere else. Durran and Gingrich (2014) demonstrated that the form of the evolving error spectra in the Lorenz turbulence model is a function of the initial error spectrum as well as the slope of the background kinetic energy spectrum.

At the later stages of the convective-mesoscale phase, the initially well-defined peak broadened and shifted toward larger scales (Figs. 12d–f). During the baroclinic

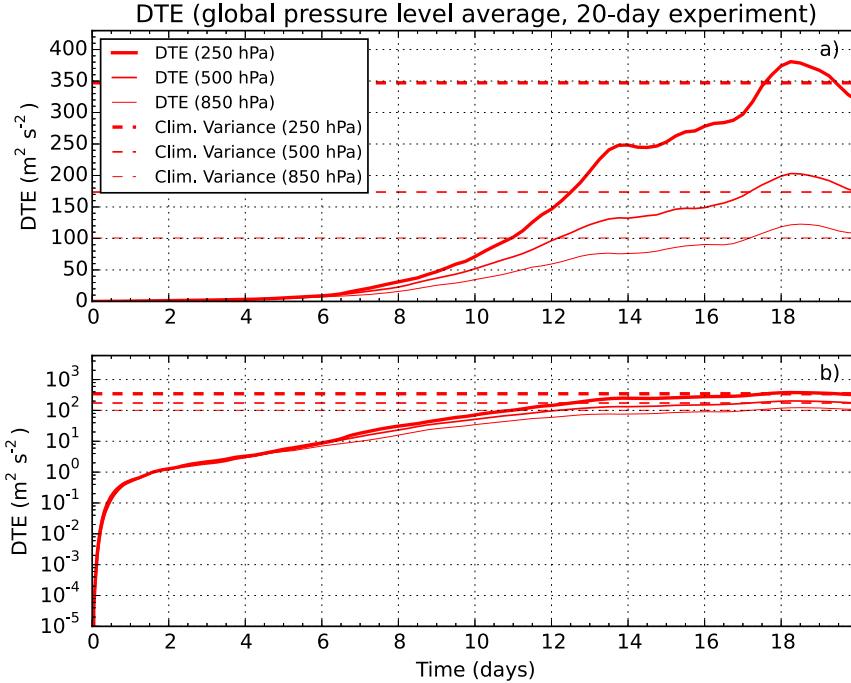


FIG. 9. As in Fig. 8, but for horizontally averaged DTE on pressure levels of 250, 500, and 850 hPa from the twins CTRL/Pert-20d.

phase, the error spectra developed a peak in the baroclinically active band between wavenumbers 10 and 20, in agreement with Tribbia and Baumhefner (2004). This peak was especially pronounced in the rotational wind (and therefore also in the total wind), indicating that errors grew mainly with the balanced rotational flow. During the baroclinic phase, the growth rate was nearly constant (Fig. 13b), and error growth generally mirrored the case of  $k^{-3}$  turbulence (Fig. 13d).

Compared with the error kinetic energy spectra, the evolution of the Z500 error variance spectra was quite different (Fig. 14). The Z500 error variance spectra lacked both the early peak at the smallest resolved scales and the later peak at the baroclinically active scales. Quite remarkably, the Z500 error did not saturate at scales  $< 300$  km, mainly because error growth slowed drastically in the shallower mesoscale part of the spectrum before reaching the saturation limit. The reason for this unexpected behavior is unknown, and it may be due to climatological features associated with topography (Boer 1994). Future research is necessary to shed more light onto this peculiar result.

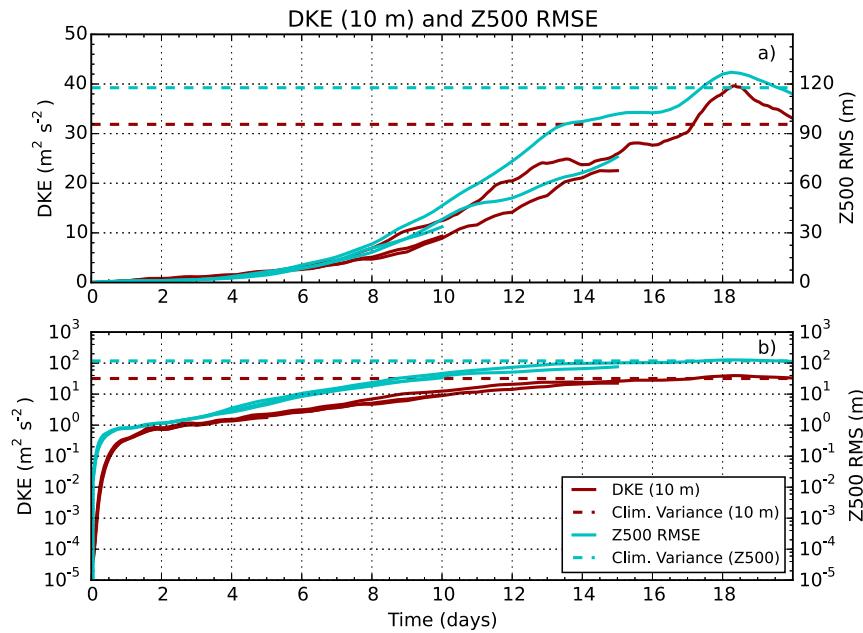
#### b. Scale-dependent predictability limits

Sequences of error spectra such as the ones shown in Fig. 12 illustrate the growth of error as a function of spatial scale, but they are not ideal for quantifying the scale-dependent predictability limits of atmospheric

flow. Here, following Judt et al. (2016), the predictability limit of a given wavenumber was explicitly calculated by determining the forecast time at which the error reaches 95% of the saturation limit. The resulting values are plotted as red dots in Fig. 15 for the divergent, rotational, and total wind components at various altitudes. Analogously, points in orange show forecast times at which the error reaches 60% of its saturation value, a percentage that is often used to define useful prediction skill (Žagar et al. 2017). As before, data points corresponding to scales smaller than  $6\Delta$  will not be considered further.

The patterns traced by the red data points differ substantially between the divergent and rotational wind component and between different vertical levels. Evidently, the predictability of atmospheric flow is much more complex than what could be conveyed by the simple global averages in section 5 or by the spectra in Fig. 12. In particular, the predictability limits of atmospheric motions are not only scale dependent but also affected by the underlying dynamics (divergent vs rotational motions) and altitude. Although the altitude dependence is also apparent when considering a threshold of 60% error saturation, differences between the divergent and rotational motions are far less pronounced (Fig. 15, orange dots).

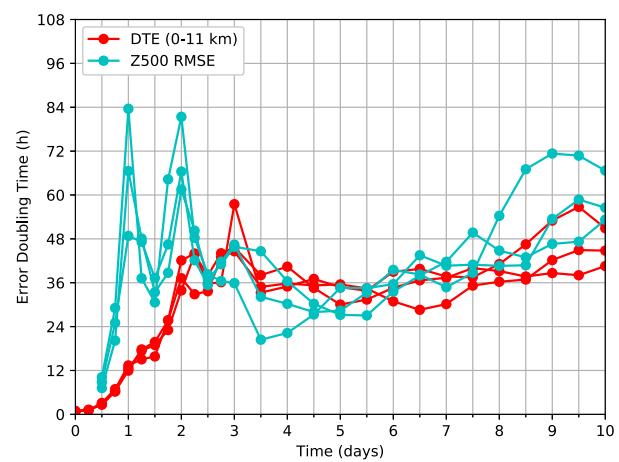
Only rotational motions far above the boundary layer exhibited a classic monotonic relationship between scale

FIG. 10. As in Fig. 8, but for  $DKE_{10m}$  and  $Z500_{RMSE}$ .

of motion and limit of predictability (Figs. 15b,e,h; red dots). For divergent motions in general and rotational motions closer to the surface, the functional relationship between the spatial scale and predictability limit is convex, meaning that smaller mesoscale motions have longer predictability than larger mesoscale motions. Predictability limits are shortest at intermediate scales, around wavenumber 500 for divergent motions in the free atmosphere, and near wavenumber 100 for divergent and rotational motions at and below 850 hPa. (The predictability limits of the total wind in the right column of Fig. 15 are averages of the predictability limits of the divergent and rotational motions, weighted by the contribution of each wind component's kinetic energy to the total kinetic energy.) Although not described with as much detail, Boer (1994) noticed the peculiar inverse relationship between predictability and spatial scale at the mesoscales and attributed it to local topographic forcing. This hypothesis is substantiated by the fact that the signal is stronger for the divergent flow, which is dominated by gravity waves (Waite and Snyder 2013). The effect of Earth's surface on the predictability of atmospheric flow also manifests in the increase of predictability time toward the surface. In fact, regarding the 10-m wind, only a few wavenumbers suffer from error saturation, and many scales of motion retain predictability for at least 20 days (Figs. 15m–o).

Considering a threshold of 60% saturation, the relationship between spatial scale and the time it takes the error to reach this threshold is more in line with the

classic picture (i.e., the smaller the spatial scale, the shorter the time, at least at the 500-hPa level and above). More specifically, there exists a kink that coincides with the transition zone between the  $k^{-5/3}$  and  $k^{-3}$  segments of the kinetic energy spectrum. This kink is further evidence that the errors grow differently depending on the spectral slope of the background spectrum, and in particular, errors in the  $k^{-5/3}$  regime grow upscale faster than errors in the  $k^{-3}$  regime. In the free troposphere, the limit of the useful prediction skill of motions in the  $k^{-5/3}$  regime is consistently  $<5$  days. Closer to the surface, there seems to be a tendency for

FIG. 11. Error doubling times for global volume-averaged DTE and  $Z500_{RMSE}$ , computed from the data shown in Figs. 8 and 10.

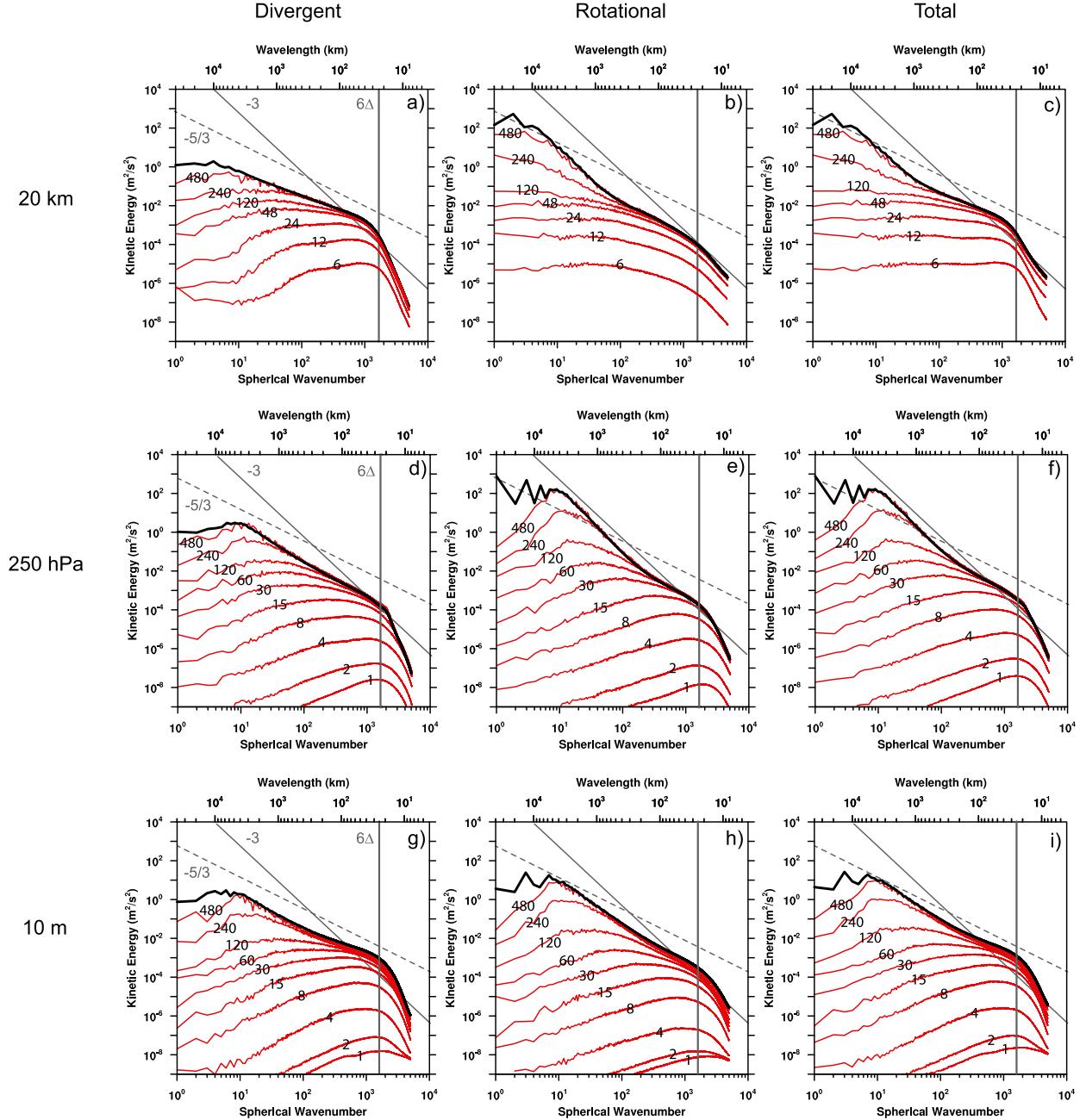


FIG. 12. Background kinetic energy spectra (black) and error kinetic energy spectra (red) of the (left) divergent, (center) rotational, and (right) total wind at (a)–(c) 20 km, (d)–(f) 250 hPa, and (g)–(i) 10 m. Error spectra are plotted in geometric sequences, i.e., in (a)–(c), error spectra are valid at 6 and 12 h and 1, 2, 5, 10, and 20 days; in all other panels, error spectra are valid at 1, 2, 4, 8, and 15 h and 1.25, 2.5, 5, 10, and 20 days. Error spectra are averaged over the number of experiments available at a given time, while the background spectra represent the time-mean ensemble mean multiplied by 2 to denote the saturation limit. For reference, each panel includes graphs with logarithmic slopes of  $-3$  and  $-5/3$  (gray lines). The vertical gray lines denote the effective model resolution (24 km).

developing a convex relationship akin to the scale-dependent predictability limits.

The height dependence of error saturation times in Fig. 15 contradicts the earlier finding that the troposphere's predictability limit is not height dependent (section 5b).

The reason for this contradiction is not obvious, but following similar arguments as in the previous paragraph, it could be related to climatological components of the flow. For example, topographic forcing could hinder the flow from decorrelating completely, especially near the surface.

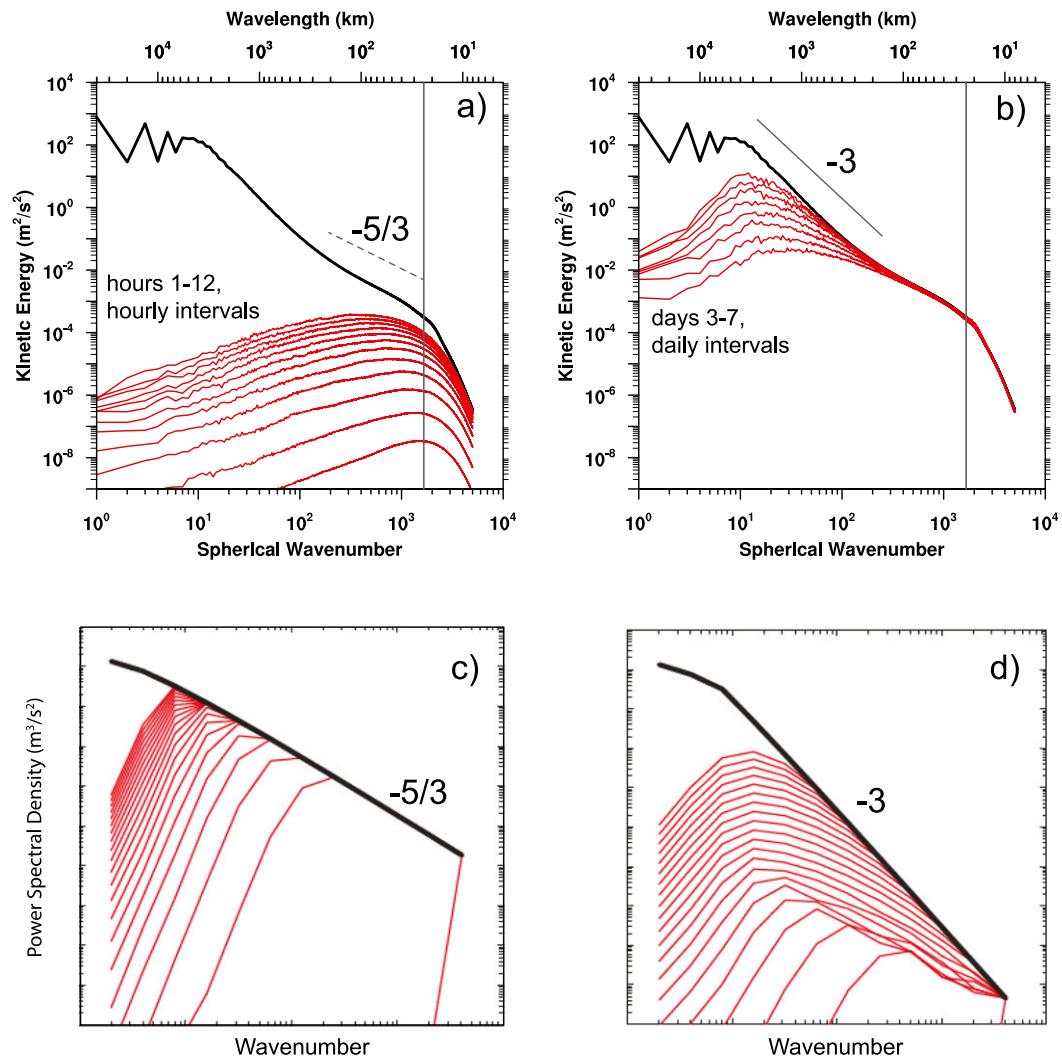


FIG. 13. (a),(b) Spectral error growth in the predictability experiment compared with (c),(d) spectral error growth in idealized  $k^{-5/3}$  and  $k^{-3}$  turbulence. Shown in (a) is the evolution of the 250-hPa error kinetic energy spectra between 1 and 12 h and in (b) between 3 and 7 days; (c) and (d) are adapted from Rotunno and Snyder (2008, their Fig. 1).

Consequently, the error does not saturate, which leads to artificially “enhanced” predictability limits. This hypothesis could be tested with an aquaplanet predictability experiment similar to that of Bretherton and Khairoutdinov (2015).

On the other hand, it is possible that small-scale low-level flow has indeed longer predictability. The climatological variance  $\text{var}_{uvT}$ , which was used to quantify predictability in section 5, was computed from a dataset with approximately 80-km grid spacing. Because of the low resolution, mesoscale components of the flow were therefore not included. This means that the total variance (and thus the saturation limit) may have been estimated too low because of the “missing” variance from the mesoscale components. A potential low bias in the

error saturation limit implies that the predictability limits in section 5 have a potential short bias. Missing variance from mesoscale motions weighs particularly heavy in the lower levels, where the kinetic energy spectrum is shallower (Fig. 12).

## 7. Summary and conclusions

The main goals of this study were 1) to investigate atmospheric error growth with a global convection-permitting NWP model (GCPM), 2) to compare error growth characteristics with those predicted by theory, and 3) to quantify the intrinsic predictability of the atmosphere. Specifically, error growth was explored through a set of identical twin experiments, which were

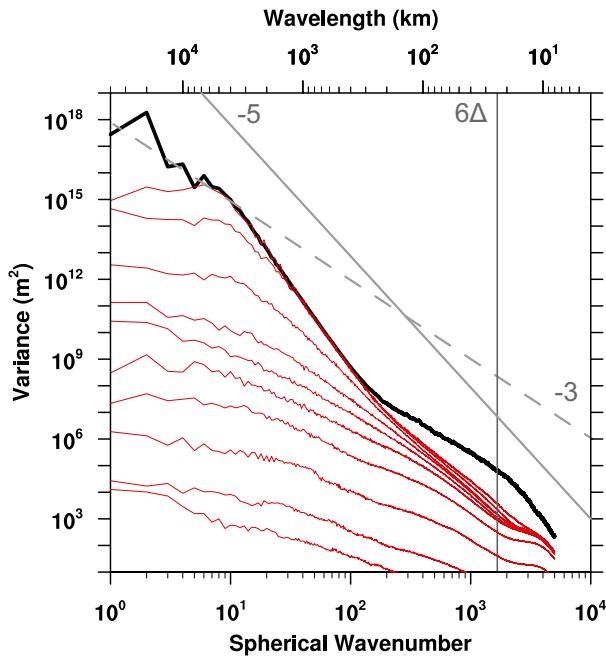


FIG. 14. Background spectrum (black) and error spectra (red) of Z500 variance. The error spectra are valid at 1, 2, 4, 8, and 15 h and 1.25, 2.5, 5, 10, and 20 days. The vertical gray line denotes the effective model resolution (24 km).

produced with the atmospheric component of the Model for Prediction Across Scales (MPAS) on a quasi-uniform 4-km mesh. The focus was on the intrinsic predictability of motions from convective to synoptic scales; hence, the results establish an upper bound on the extent to which we could predict the weather if we had a perfect model and nearly perfect initial conditions.

Errors grew in accordance with the conceptual model proposed by Zhang et al. (2007); that is, initial error growth was tied to moist convection. The convective-scale errors quickly grew in scale, magnitude, and spatial extent while contaminating the mesoscale. Error growth during this convective-mesoscale phase was consistent with predictability theory pertaining to  $k^{-5/3}$  turbulence except that errors did not grow by cascading from smaller to larger scales but grew up in magnitude at all scales. After about 2 days, the errors began to affect the synoptic scales and continued to grow with the background baroclinic instability. During the baroclinic phase, error growth was consistent with predictability theory pertaining to  $k^{-3}$  turbulence.

The fact that initial error growth resembled the  $k^{-5/3}$  turbulence case—the error peaks at the smallest resolved scales and the growth rate decreases monotonically—is evidence that the atmosphere possesses a finite limit of predictability, which, in a global average sense, seems to be between 2 and 3 weeks. This estimate is in line with

previous studies that used much lower-resolution models and remarkably close to the “16–23 days forecast skill horizon” noted by Buizza and Leutbecher (2015). It is also broadly consistent with the 15–30-day limit that was determined by Bretherton and Khairoutdinov (2015) using a global convection-permitting aquaplanet model. This result implies that, even with the prospect of superior future technology, it is most likely not possible to predict the weather beyond one month.

The underlying problem that limits predictability is the rapid error growth during the convective-mesoscale phase. To make this issue more transparent, consider these two hypothetical cases: 1) all scales of motions are correctly specified except for some random noise, and 2) only synoptic-scale motions are correctly specified, while the error on all other scales is saturated from the beginning. Case 1, which from today’s perspective sounds like science fiction, would extend the predictability horizon by maximally two days relative to the more realistic case 2. In the end, the rapid error growth on small scales provokes questions about the practical value of convective-scale data assimilation and future operational GCPM predictions. From a predictability perspective, it seems unlikely that GCPMs can drastically improve deterministic forecasts of mesoscale weather phenomena over the current approach, which is to run lower-resolution global models in conjunction with convection-permitting regional models. Durran and Weyn (2016) and Weyn and Durran (2017), who studied mesoscale convective systems in idealized simulations with even finer grid spacings, arrived at similar conclusions. One area where GCPMs may be useful is to provide tropical cyclone track and intensity forecasts from a single source, but more research is necessary to test this hypothesis.

On the other hand, the above interpretation may be unduly pessimistic in light of the spectral predictability analysis in section 6b, which suggested that mesoscale flow may have longer predictability than commonly thought (especially the flow closer to the surface). It is not yet clear whether this extended predictability is due to climatological features, but the results instill hope that there is potential for longer predictability of processes such as convective initiation, which are rooted in the boundary layer. This potential predictability could be exploited with GCPMs. Furthermore, the notion that GCPMs may not lengthen the predictability horizon of small-scale high-impact weather phenomena does not mean that GCPMs are not valuable. For example, because of explicitly resolving convection, GCPMs may be able to reduce the longstanding biases of current NWP models in the tropics and, more generally, lead to an overall gain in forecast accuracy because of an increased realism in the way the atmosphere is modeled.

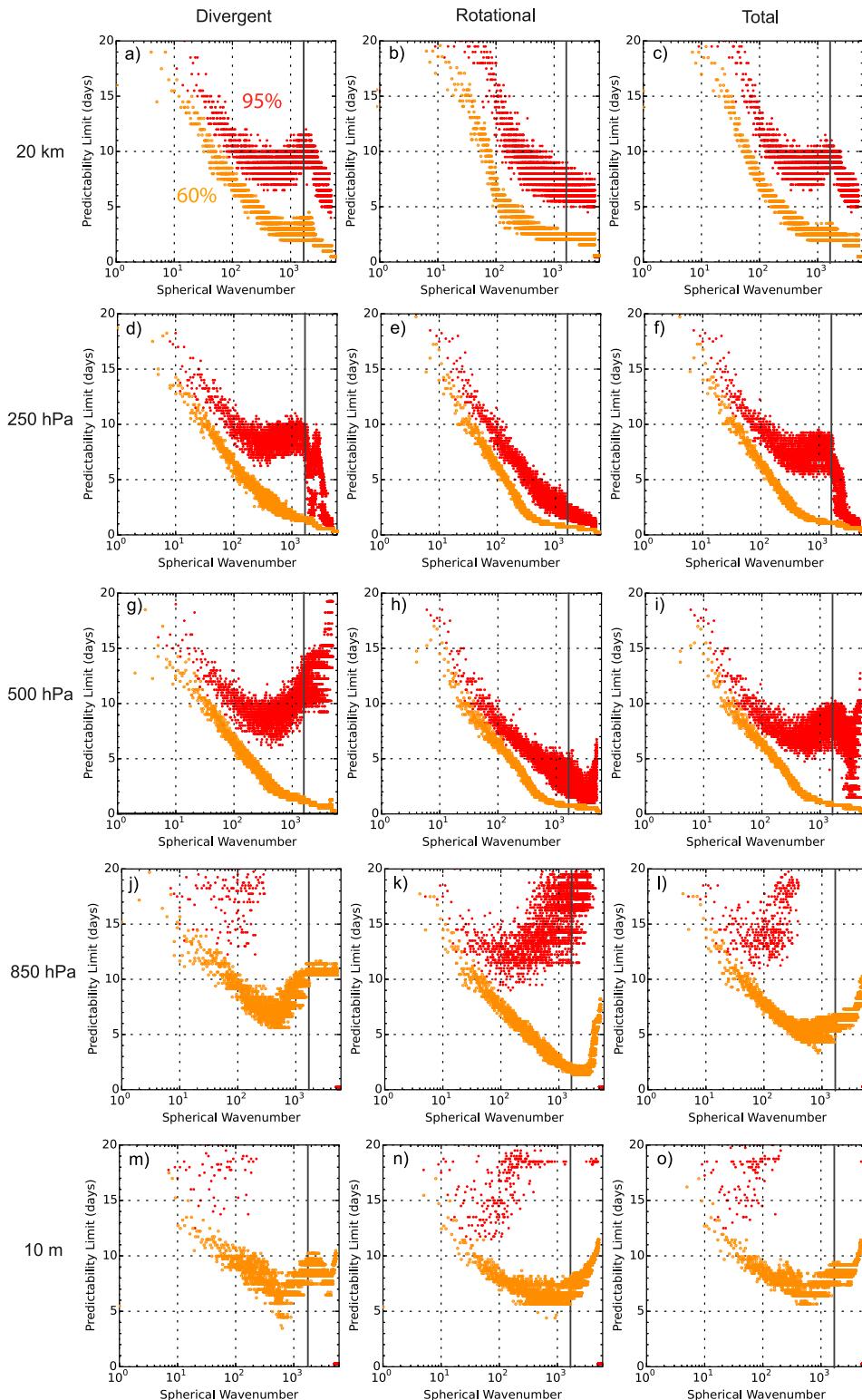


FIG. 15. Red dots show the predictability limits of the divergent, rotational, and total wind as a function of wavenumber at the indicated vertical levels. The predictability limit is defined as the time at which the error reaches 95% saturation. If the error does not reach 95% saturation, no value is plotted. Orange dots show the same as red dots, but for a threshold of 60% error saturation.

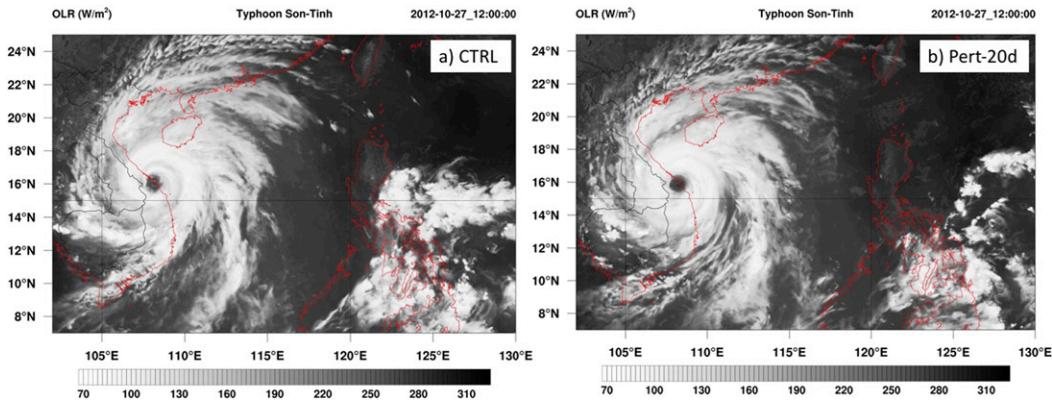


FIG. 16. Typhoon Son-Tinh in (a) CTRL and (b) Pert-20d at simulation day 7.5, as depicted by outgoing longwave radiation ( $\text{W m}^{-2}$ ).

Despite being a considerable undertaking from a computational, data storage, and data analysis standpoint, this study has several limitations. First, it is only a case study with a small number of ensemble members. Second, the possible effect of the ocean on atmospheric predictability could not be addressed. Third, this study relied on classic predictability metrics, (i.e., Eulerian squared differences evaluated at grid points). These metrics are not useful for assessing the predictability of meteorological phenomena that are localized and intermittent (e.g., Mapes et al. 2008; Ngan and Eperon 2012; Potvin et al. 2017). The issue can be illustrated with the following example: Fig. 16 shows two realizations of Typhoon Son-Tinh, one from the control simulation and one of its identical twins, at forecast day 7.5. The typhoons are in almost identical locations and closely resemble each other, which implies high predictability. At the same time, the error spectra in Fig. 12 suggest that at day 5, the mesoscale spectrum is almost saturated and there is considerable error at synoptic scales. The difficulty to reconcile the classic predictability view with the predictability of specific phenomena or “objects” motivates future research into objects-based predictability metrics that can explicitly quantify the predictability of discrete meteorological phenomena.

**Acknowledgments.** This study would not have been possible without Michael Duda (NCAR), who generated the 4-km MPAS mesh and helped tremendously with setting up the model runs. Bill Skamarock (NCAR) assisted in designing the project and determining the computational resources. The study benefitted from science discussions with Rich Rotunno, Chris Snyder, Chris Davis (all NCAR), and many other NCAR employees. Rich Rotunno also provided an informal review of this manuscript, and the comments from three anonymous reviewers helped to further improve the paper. Finally, I would also like to acknowledge high-

performance computing support from Yellowstone (ark:/85065/d7wd3xhc) provided by NCAR’s Computational and Information Systems Laboratory, sponsored by the National Science Foundation. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

## REFERENCES

- Bierdel, L., C. Snyder, S.-H. Park, and W. C. Skamarock, 2016: Accuracy of rotational and divergent kinetic energy spectra diagnosed from flight-track winds. *J. Atmos. Sci.*, **73**, 3273–3286, <https://doi.org/10.1175/JAS-D-16-0040.1>.
- , T. Selz, and G. Craig, 2017: Theoretical aspects of upscale error growth through the mesoscales: An analytical model. *Quart. J. Roy. Meteor. Soc.*, **143**, 3048–3059, <https://doi.org/10.1002/qj.3160>.
- Blunden, J., and D. S. Arndt, 2013: State of the Climate in 2012. *Bull. Amer. Meteor. Soc.*, **94** (8), S1–S258, <https://doi.org/10.1175/2013BAMSSStateoftheClimate.1>.
- Boer, G. J., 1994: Predictability regimes in atmospheric flow. *Mon. Wea. Rev.*, **122**, 2285–2295, [https://doi.org/10.1175/1520-0493\(1994\)122<2285:PRAF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1994)122<2285:PRAF>2.0.CO;2).
- Boffetta, G., and S. Musacchio, 2017: Chaos and predictability of homogeneous-isotropic turbulence. *Phys. Rev. Lett.*, **119**, 054102, <https://doi.org/10.1103/PhysRevLett.119.054102>.
- , A. Celani, A. Crisanti, and A. Vulpiani, 1997: Predictability in two-dimensional decaying turbulence. *Phys. Fluids*, **9**, 724–734, <https://doi.org/10.1063/1.869227>.
- Bretherton, C. S., and M. F. Khairoutdinov, 2015: Convective self-aggregation feedbacks in near-global cloud-resolving simulations of an aquaplanet. *J. Adv. Model. Earth Syst.*, **7**, 1765–1787, <https://doi.org/10.1002/2015MS000499>.
- Buizza, R., and M. Leutbecher, 2015: The forecast skill horizon. *Quart. J. Roy. Meteor. Soc.*, **141**, 3366–3382, <https://doi.org/10.1002/qj.2619>.
- Dalcher, A., and E. Kalnay, 1987: Error growth and predictability in operational ECMWF forecasts. *Tellus*, **39A**, 474–491, <https://doi.org/10.1111/j.1600-0870.1987.tb00322.x>.
- Durran, D. R., and M. Gingrich, 2014: Atmospheric predictability: Why butterflies are not of practical importance. *J. Atmos. Sci.*, **71**, 2476–2488, <https://doi.org/10.1175/JAS-D-14-0007.1>.

- , 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>.
- ECMWF, 2016: ECMWF strategy 2016–2025: The strength of a common goal. European Centre for Medium-Range Weather Forecasts Tech. Rep., 32 pp., [https://www.ecmwf.int/sites/default/files/ECMWF\\_Strategy\\_2016-2025.pdf](https://www.ecmwf.int/sites/default/files/ECMWF_Strategy_2016-2025.pdf).
- Errico, R., and D. Baumhefner, 1987: Predictability experiments using a high-resolution limited-area model. *Mon. Wea. Rev.*, **115**, 488–504, [https://doi.org/10.1175/1520-0493\(1987\)115<0488:PEUAHR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<0488:PEUAHR>2.0.CO;2).
- Fowler, L. D., W. C. Skamarock, G. A. Grell, S. R. Freitas, and M. G. Duda, 2016: Analyzing the Grell–Freitas convection scheme from hydrostatic to nonhydrostatic scales within a global model. *Mon. Wea. Rev.*, **144**, 2285–2306, <https://doi.org/10.1175/MWR-D-15-0311.1>.
- Grell, G. A., and S. R. Freitas, 2014: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, **14**, 5233–5250, <https://doi.org/10.5194/acp-14-5233-2014>.
- 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>.
- Hawblitzel, D. P., F. Zhang, Z. Meng, and C. A. Davis, 2007: Probabilistic evaluation of the dynamics and predictability of the mesoscale convective vortex of 10–13 June 2003. *Mon. Wea. Rev.*, **135**, 1544–1563, <https://doi.org/10.1175/MWR3346.1>.
- Heinzeller, D., M. G. Duda, and H. Kunstmann, 2016: Towards convection-resolving, global atmospheric simulations with the Model for Prediction Across Scales (MPAS) v3.1: An extreme scaling experiment. *Geosci. Model Dev.*, **9**, 77–110, <https://doi.org/10.5194/gmd-9-77-2016>.
- Hohenegger, C., and C. Schär, 2007: Atmospheric predictability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.*, **88**, 1783–1793, <https://doi.org/10.1175/BAMS-88-11-1783>.
- Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, <https://doi.org/10.1029/2008JD009944>.
- Judit, F., S. S. Chen, and J. Berner, 2016: Predictability of tropical cyclone intensity: Scale-dependent forecast error growth in high-resolution stochastic kinetic-energy backscatter ensembles. *Quart. J. Roy. Meteor. Soc.*, **142**, 43–57, <https://doi.org/10.1002/qj.2626>.
- Klemp, J. B., 2011: A terrain-following coordinate with smoothed coordinate surfaces. *Mon. Wea. Rev.*, **139**, 2163–2169, <https://doi.org/10.1175/MWR-D-10-05046.1>.
- Leith, C. E., and R. H. Kraichnan, 1972: Predictability of turbulent flows. *J. Atmos. Sci.*, **29**, 1041–1058, [https://doi.org/10.1175/1520-0469\(1972\)029<1041:PTOF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1972)029<1041:PTOF>2.0.CO;2).
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNONP>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNONP>2.0.CO;2).
- , 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, <https://doi.org/10.3402/tellusa.v21i3.10086>.
- , 1982: Atmospheric predictability experiments with a large numerical model. *Tellus*, **34**, 505–513, <https://doi.org/10.3402/tellusa.v34i6.10836>.
- , 1996: Predictability—A problem partly solved. *Proc. Seminar on Predictability*, Reading, United Kingdom, ECMWF, 1–18.
- Mapes, B., S. Tulich, T. Nasuno, and M. Satoh, 2008: Predictability aspects of global aqua-planet simulations with explicit convection. *J. Meteor. Soc. Japan*, **86**, 175–185, <https://doi.org/10.2151/jmsj.86A.175>.
- Métais, O., and M. Lesieur, 1986: Statistical predictability of decaying turbulence. *J. Atmos. Sci.*, **43**, 857–870, [https://doi.org/10.1175/1520-0469\(1986\)043<0857:SPODT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1986)043<0857:SPODT>2.0.CO;2).
- Miyamoto, Y., Y. Kajikawa, R. Yoshida, T. Yamaura, H. Yashiro, and H. Tomita, 2013: Deep moist atmospheric convection in a subkilometer global simulation. *Geophys. Res. Lett.*, **40**, 4922–4926, <https://doi.org/10.1002/grl.50944>.
- Morss, R. E., C. Snyder, and R. Rotunno, 2009: Spectra, spatial scales, and predictability in a quasigeostrophic model. *J. Atmos. Sci.*, **66**, 3115–3130, <https://doi.org/10.1175/2009JAS3057.1>.
- Nakanishi, M., and H. Niino, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407, <https://doi.org/10.1007/s10546-005-9030-8>.
- , and —, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. *J. Meteor. Soc. Japan*, **87**, 895–912, <https://doi.org/10.2151/jmsj.87.895>.
- Nastrom, D. G., and K. S. Gage, 1985: A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.*, **42**, 950–960, [https://doi.org/10.1175/1520-0469\(1985\)042<0950:ACOAWS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1985)042<0950:ACOAWS>2.0.CO;2).
- Ngan, K., and G. E. Eperon, 2012: Middle atmosphere predictability in a numerical weather prediction model: Revisiting the inverse error cascade. *Quart. J. Roy. Meteor. Soc.*, **138**, 1366–1378, <https://doi.org/10.1002/qj.984>.
- , P. Bartello, and D. N. Straub, 2009: Predictability of rotating stratified turbulence. *J. Atmos. Sci.*, **66**, 1384–1400, <https://doi.org/10.1175/2008JAS2799.1>.
- Niu, G.-Y., and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**, D12109, <https://doi.org/10.1029/2010JD015139>.
- Palmer, T. N., A. Döring, and G. Seregin, 2014: The real butterfly effect. *Nonlinearity*, **27**, R123–R141, <https://doi.org/10.1088/0951-7715/27/9/R123>.
- Potvin, C. K., E. M. Murillo, M. L. Flora, and D. M. Wheatley, 2017: Sensitivity of supercell simulations to initial-condition resolution. *J. Atmos. Sci.*, **74**, 5–26, <https://doi.org/10.1175/JAS-D-16-0098.1>.
- Putman, W. M., and M. Suarez, 2011: Cloud-system resolving simulations with the NASA Goddard Earth Observing System global atmospheric model (GEOS-5). *Geophys. Res. Lett.*, **38**, L16809, <https://doi.org/10.1029/2011GL048438>.
- Rotunno, R., and C. Snyder, 2008: A generalization of Lorenz's model for the predictability of flows with many scales of motion. *J. Atmos. Sci.*, **65**, 1063–1076, <https://doi.org/10.1175/2007JAS2449.1>.
- Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga, 2008: Nonhydrostatic icosahedral atmospheric model (NI-CAM) for global cloud resolving simulations. *J. Comput. Phys.*, **227**, 3486–3514, <https://doi.org/10.1016/j.jcp.2007.02.006>.
- 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>.
- Shukla, J., 1981: Dynamical predictability of monthly means. *J. Atmos. Sci.*, **38**, 2547–2572, [https://doi.org/10.1175/1520-0469\(1981\)038<2547:DPOMM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<2547:DPOMM>2.0.CO;2).

- Simmons, A. J., and A. Hollingsworth, 2002: Some aspects of the improvement in skill of numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **128**, 647–677, <https://doi.org/10.1256/003590002321042135>.
- , R. Mureau, and T. Petroliagis, 1995: Error growth and estimates of predictability from the ECMWF forecasting system. *Quart. J. Roy. Meteor. Soc.*, **121**, 1739–1771, <https://doi.org/10.1002/qj.49712152711>.
- Sippel, J. A., and F. Zhang, 2008: A probabilistic analysis of the dynamics and predictability of tropical cyclogenesis. *J. Atmos. Sci.*, **65**, 3440–3459, <https://doi.org/10.1175/2008JAS2597.1>.
- Skamarock, W. C., J. B. Klemp, M. G. Duda, L. D. Fowler, S.-H. Park, and T. D. Ringler, 2012: A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tessellations and C-grid staggering. *Mon. Wea. Rev.*, **140**, 3090–3105, <https://doi.org/10.1175/MWR-D-11-00215.1>.
- , S.-H. Park, J. B. Klemp, and C. Snyder, 2014: Atmospheric kinetic energy spectra from global high-resolution nonhydrostatic simulations. *J. Atmos. Sci.*, **71**, 4369–4381, <https://doi.org/10.1175/JAS-D-14-0114.1>.
- Sun, Y. Q., 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>.
- Tennant, W., 2009: The butterfly effect—Myth or reality? *Weather*, **64**, 299, <https://doi.org/10.1002/wea.464>.
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, <https://doi.org/10.1175/2008MWR2387.1>.
- 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/10.1111/j.2153-3490.1957.tb01885.x>.
- 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).
- Vukicevic, T., and R. M. Errico, 1990: The influence of artificial and physical factors upon predictability estimates using a complex limited-area model. *Mon. Wea. Rev.*, **118**, 1460–1482, [https://doi.org/10.1175/1520-0493\(1990\)118<1460:TIOAAP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1990)118<1460:TIOAAP>2.0.CO;2).
- Waite, M. L., and C. Snyder, 2013: Mesoscale energy spectra of moist baroclinic waves. *J. Atmos. Sci.*, **70**, 1242–1256, <https://doi.org/10.1175/JAS-D-11-0347.1>.
- Weyn, J. A., and D. R. Durran, 2017: The dependence of the predictability of mesoscale convective systems on the horizontal scale and amplitude of initial errors in idealized simulations. *J. Atmos. Sci.*, **74**, 2191–2210, <https://doi.org/10.1175/JAS-D-17-0006.1>.
- Yang, Z.-L., and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *J. Geophys. Res.*, **116**, D12110, <https://doi.org/10.1029/2010JD015140>.
- Žagar, N., M. Horvat, Ž. 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, 2002: Mesoscale predictability of the surprise snowstorm of 24–25 January 2000. *Mon. Wea. Rev.*, **130**, 1617–1632, [https://doi.org/10.1175/1520-0493\(2002\)130<1617:MPOTSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1617:MPOTSS>2.0.CO;2).
- , —, and —, 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>.
- Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2015: Practical predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma: Sensitivity to synoptic timing and topographical influence. *Mon. Wea. Rev.*, **143**, 2973–2997, <https://doi.org/10.1175/MWR-D-14-00394.1>.