Sensitivity of Intrinsic Error Growth to Large-Scale Uncertainty Structure in a Record-Breaking Summertime Rainfall Event

YUNJI ZHANG[a,b,c]

[a] Center for Advanced Data Assimilation and Predictability Techniques, The Pennsylvania State University, University Park, Pennsylvania
[b] Alliance for Education, Science, Engineering, and Design with Africa, The Pennsylvania State University, University Park, Pennsylvania
[c] Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, Pennsylvania

(Manuscript received 25 October 2022, in final form 25 February 2023, accepted 27 February 2023)

ABSTRACT: It is recognized that the atmosphere's predictability is intrinsically limited by unobservably small uncertainties that are beyond our capability to eliminate. However, there have been discussions in recent years on whether forecast error grows upscale (small-scale error grows faster and transfers to progressively larger scales) or up-amplitude (grows at all scales at the same time) when unobservably small-amplitude initial uncertainties are imposed at the large scales and limit the intrinsic predictability. This study uses large-scale small-amplitude initial uncertainties of two different structures—one idealized, univariate, and isotropic, the other realistic, multivariate, and flow dependent—to examine the error growth characteristics in the intrinsic predictability regime associated with a record-breaking rainfall event that happened on 19–20 July 2021 in China. Results indicate upscale error growth characteristics regardless of the structure of the initial uncertainties: the errors at smaller scales grow fastest first; as the forecasts continue, the wavelengths of the fastest error growth gradually shift toward larger scales with reduced error growth rates. Therefore, error growth from smaller to larger scales was more important than the growth directly at the large scales of the initial errors. These upscale error growth characteristics also depend on the perturbed and examined quantities: if the examined quantity is perturbed, then its errors grow upscale; if there is no initial uncertainty in the examined quantity, then its errors grow at all scales at the same time, although its smaller-scale errors still grow faster for the first several hours, suggesting the existence of the upscale error growth.

SIGNIFICANCE STATEMENT: This study compared the error growth characteristics associated with the atmosphere's intrinsic predictability under two different structures of unobservably small-amplitude, large-scale initial uncertainties: one idealized, univariate, and isotropic, the other realistic, multivariate, and flow dependent. The characteristics of the errors growing upscale rather than up-amplitude regardless of the initial uncertainties' structure are apparent. The large-scale errors do not grow if their initial amplitudes are much bigger than the small-scale errors. This study also examined how the error growth characteristics will change when the quantity that is used to describe the error growth is inconsistent with the quantity that contains uncertainty, suggesting the importance of including multivariate, covariant uncertainties of state variables in atmospheric predictability studies.

KEYWORDS: Mesoscale processes; Ensembles; Mesoscale forecasting; Numerical weather prediction/forecasting

1. Introduction

Intrinsic predictability of the weather, or how long can we predict the weather accurately given almost perfect knowledge of the atmospheric processes and almost perfect estimations of the atmospheric states (Lorenz 1963; Melhauser and Zhang 2012; Zhang et al. 2019), defines the ultimate limit of our day-to-day weather forecasts. Studies using simplified turbulence models suggest the existence of an intrinsic predictability limit depends on the slope of the background energy spectrum (Lorenz 1969; Rotunno and Snyder 2008): for a background spectrum obeying a power law of the form $k^{-5/3}$, like what was observed at the mesoscale (Nastrom and Gage 1985), the theories find limited intrinsic predictability. On the other hand, some recent studies suggest other relationships between the slope of the background energy spectrum and the atmosphere's intrinsic predictability limit (Leung et al. 2019; Lloveras et al. 2022).

Zhang et al. (2007) proposed a three-stage upscale error growth mechanism describing how small-scale small-amplitude uncertainties limit the intrinsic predictability. In this mechanism, initial uncertainties that are unobservably small will grow and saturate at the convective scale with the help of the moist convective processes (first stage), then transfer upscale to the mesoscale and synoptic scales (second stage) and grow at a slower speed at synoptic scales (third stage). The upscale transfer of small-amplitude uncertainties is also observed by many studies using convection-allowing regional models (Selz and Craig 2015; Zhang et al. 2016, 2022, hereafter Z22), convection-parameterized global models (Selz 2019; Selz et al. 2022), or convection-allowing global models (Judt 2018; Zhang et al. 2019). Other studies confirm the critical role of moist convective processes in establishing the mesoscale $k^{-5/3}$ kinetic energy spectrum (Sun et al. 2017; Selz et al. 2019; Fan et al. 2022) and limiting the intrinsic predictability (Hohenegger and Schär 2007a,b; Sun and Zhang 2016, 2020; Baumgart et al. 2019; Leung et al. 2020). Detailed analyses using numerical models of different complexities suggest that small-scale initial errors amplify and

---

Corresponding author: Yunji Zhang, yuz31@psu.edu

DOI: 10.1175/JAS-D-22-0231.1

saturate the convective (cloud) scale with the help of moist convective processes and latent heat release, and these errors grow upscale through gravity wave propagations and geostrophic adjustments (e.g., Selz and Craig 2015; Bierdel et al. 2018; Baumgart et al. 2019; Selz et al. 2022) before entering the third stage of slower and persistent growth associated with baroclinic instability (Selz and Craig 2015; Judt 2018) or near-tropopause nonlinear Rossby wave dynamics (Baumgart et al. 2018, 2019).

On the other hand, there are studies arguing that the small-amplitude uncertainties are growing up-amplitude at all scales simultaneously, and the larger-scale uncertainties are therefore at least as important as—if not more than—smaller-scale uncertainties because of larger background energy at larger scales. Using numerical models of different complexity—including the two-dimensional barotropic vorticity model in Lorenz (1969), the surface quasigeostrophic model in Rotunno and Snyder (2008), the cloud model in Durran and Klemp (1983), and the Advanced Research version of the Weather Research and Forecasting (WRF-ARW; Skamarock et al. 2008) Model—several studies suggest that larger-scale errors cascade downscale, which lead to a more uniform amplification of forecast errors at all scales ("up-amplitude"; Durran and Gingrich 2014; Durran and Weyn 2016; Nielsen and Schumacher 2016; Weyn and Durran 2017; Lloveras et al. 2022). Weyn and Durran (2019) suggest that the impact of larger- versus smaller-scale errors, and therefore the relative role of upscale and up-amplitude mechanism in the overall error growth processes, is sensitive to synoptic-scale forcing. Several studies among them reported a uniform growth of forecast error spectra across all scales, and this type of error spectra does not contain a peak at or close to the small-scale error saturation wavelengths as predicted by the homogeneous turbulence theory in Rotunno and Snyder (2008) but rather flatten out at larger scales (Durran and Gingrich 2014; Weyn and Durran 2017, 2019; Lloveras et al. 2022). Similar temporal evolution of the error kinetic energy spectra is also observed in Judt (2020), different from the findings in Judt (2018), although Rotunno et al. (2023) show that it is an artifact of inappropriate wavenumber spectral analyses.

Yet two recent studies show some interesting results regarding the forecast error's upscale growth. In the global convection-parameterized twin-experiments study of Selz et al. (2022), as they reduce the initial uncertainty from 100% of the current global model analysis uncertainty to 10% of that, the spatial scales of the fastest error growth rate during the first several days change from larger scales to smaller scales, indicating the change from the practical predictability limit to the intrinsic predictability limit. When they further reduce the initial uncertainty to 1% of their original values, the error growth rate during the first day increases. The change in the predictability regime is also accompanied by a change in the error growth characteristics: for the 100% initial uncertainty forecast, errors at larger scales always grow the fastest; however, for the 10% and 1% initial uncertainty forecasts, the fastest error growth scale gradually moves upscale as the forecasts continue. Z22 also reports similar behavior using convection-allowing regional ensembles with small-amplitude initial uncertainties—the fastest error growth occurs at smaller scales at the beginning and gradually moves upscale as the forecasts continue, and this behavior is insensitive to the spatial scale of the initial uncertainties. Furthermore, in the several experiments of Z22 that imposed large-scale initial uncertainties derived from the Global Ensemble Forecast System (GEFS), the forecast errors at larger scales (>200 km) do not grow for the first several hours, and they only start to grow once the magnitude of errors at smaller scales had increased to be comparable to that of larger-scale errors. It is worth noting that the magnitude of these large-scale initial uncertainties in Z22 is also one order of magnitude smaller than the original uncertainties that GEFS represents. Based on Zhang et al. (2019) and Selz et al. (2022), this magnitude of initial uncertainty (10% of current global model analysis uncertainty) already enters the intrinsic predictability regime.

However, these studies showing different error growth mechanisms in the intrinsic predictability regimes should not be directly compared because of the different characteristics of the initial uncertainties that they impose:

1) Although three-dimensional numerical models were used, Durran and Weyn (2016), Weyn and Durran (2017, 2019), and Lloveras et al. (2022) used a function to generate relatively uniform and isotropic initial uncertainties, while Z22, among many other similar previous studies, used flow-dependent initial uncertainties derived from a global model.
2) Durran and Weyn (2016), Weyn and Durran (2017, 2019), and Lloveras et al. (2022) only imposed initial uncertainties on one, thermodynamical variable [Durran and Weyn (2016) perturbed potential temperature and the other three perturbed water vapor mixing ratio], instead of using a set of multivariate, covariant uncertainties, and they primarily examined the temporal evolution of errors in kinetic energy.

To fill these gaps, following the experiments conducted in Z22, this study examines several sets of ensembles imposing large-scale initial uncertainties with unobservably small amplitude and different structures, aiming at the potential impact of different structures of large-scale initial uncertainties on the error growth characteristics in the intrinsic predictability regime. Section 2 introduces the event of interest, the numerical model and its settings, and the ensemble configurations with corresponding methods to generate different structures of initial uncertainties. Section 3 examines the power spectra of the forecast uncertainties, and section 4 examines the temporal evolutions of the forecast uncertainties in the physical space. Section 5 includes conclusions and discussion.

2. Methodology

a. *Overview of the event*

On 19–20 July 2021, devastating rainfall hit Henan Province in central China. Zhengzhou, the province's capital city, recorded 24-h (1200 UTC 19 July–1200 UTC 20 July) accumulated rainfall of over 600 mm (Fig. 1b) with a peak hourly

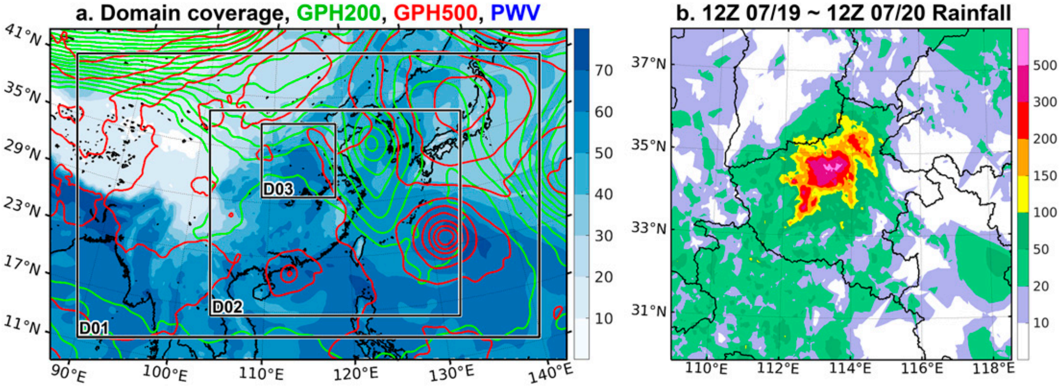

FIG. 1. (a) Model domain configuration (black rectangles), geopotential height (GPH) at 200 hPa (green contours; every 20 gpm) and 500 hPa (red contours; every 20 gpm), and column-integrated total precipitable water vapor (PWV; shaded; in kg m^{-2}) from the European Centre for Medium-Range Weather Forecasts (ECMWF) fifth-generation atmospheric reanalysis (ERA5) valid at 1200 UTC 19 Jul 2021; (b) observed 24-h accumulated rainfall from 1200 UTC 19 Jul to 1200 UTC 20 Jul 2021 within model domain D03 using available surface automatic weather stations from the China Meteorological Administration (CMA).

rainfall of 201 mm during 0800–0900 UTC 20 July. This record-breaking amount of rainfall led to severe flash flooding in the metropolitan region, costing more than 300 lives.

This rainfall event occurred under uncommon synoptic conditions (Fig. 1a). Typhoon Cempaka and Typhoon In-Fa were located to the south and southeast of mainland China, respectively. Together with the western Pacific subtropical high, they transported abundant moisture inland and created a region of total precipitable water vapor exceeding 60 kg m^{-3} when the southeasterlies were stalled by the mountains. While this region was generally under or to the west of a 200-hPa ridge, a persistent midlevel mesoscale convective vortex with a small low pressure center helped the maintenance of the rainstorms. It should also be pointed out that while synoptic conditions and local topography help to build a favorable environment for convection initiation, development, and organization over this region, they do not directly trigger the convection. A more detailed analysis of the controlling synoptic factors and physical processes that led to the record-breaking rainfall can be found in Z22 and references therein.

b. *Configurations of the numerical model*

All the simulations are conducted using the WRF Model with its fully compressible, nonhydrostatic ARW dynamical core, version 4.2 (Skamarock et al. 2019). Three one-way nested domains with 27-, 9-, and 3-km horizontal grid spacings, respectively, are designed. The outer two domains cover the primary synoptic weather systems (Fig. 1a), and the inner 3-km domain covers a 900 km × 900 km region where rainfall occurred (Fig. 1b). Fifty hybrid pressure–terrain-following levels are employed with the uppermost level located at 50 hPa. Physical parameterization schemes include the aerosol-aware Thompson and Eidhammer (2014) microphysics scheme, the Rapid Radiative Transfer Model for Global Circulation Models (RRTMG) longwave and shortwave radiation schemes (Iacono et al. 2008), the revised MM5 Monin–Obukhov surface layer scheme (Jiménez et al. 2012), the thermal diffusion land surface model, the Yonsei University scheme for planetary boundary layer (PBL) processes (Hong et al. 2006), and the modified Tiedtke cumulus parameterization scheme (Zhang et al. 2011) which is only applied in the outermost 27-km domain. This model configuration can reasonably reproduce the spatial distribution of the overall precipitation with a maximum 24-h accumulation (1200 UTC 19 July–1200 UTC 20 July) of 558 mm when initialized with the Global Forecast System (GFS) analysis valid at 0600 UTC 19 July.

c. *Experiment design*

Six 40-member 30-h ensemble forecasts that are initialized at 0600 UTC 19 July and end at 1200 UTC 20 July are performed (the analyses in this study will be based on forecast lead times). Each ensemble's initial conditions (ICs) are centered on the 0600 UTC 19 July GFS analysis. All forecasts share the same lateral boundary conditions (LBCs) generated using the GFS forecast from 0600 UTC 19 July; the impacts of the unperturbed LBCs are minimal to the innermost 3-km domain given the distance between the location of the 3-km domain and the LBCs (Fig. 1a).

Three ensembles employed realistic, multivariate, flow-dependent initial uncertainties. The baseline of all six experiments, "REAL," is the same "LARGE" ensemble in Z22. It used perturbations derived from the GEFS forecasts, with the archived GEFS outputs having a horizontal grid spacing of 0.5° × 0.5° (although GEFS runs at a higher grid spacing). Specifically, the average of the 20 GEFS ICs at 0600 UTC 19 July 2021 was subtracted from these 20 ICs to form 20 perturbations of potential temperature (T), water vapor mixing ratio (Qv), and the two horizontal wind components (U and V). Then, the amplitudes of these perturbations were rescaled by a factor of 0.1 (i.e., making the uncertainties one order of magnitude smaller than current global model accuracies) and added to the 0600 UTC 19 July GFS analysis to form 20 ICs. The other 20 ICs of the REAL ensemble were generated similarly but used the 6-h GEFS forecasts from 0000 UTC 19 July 2021

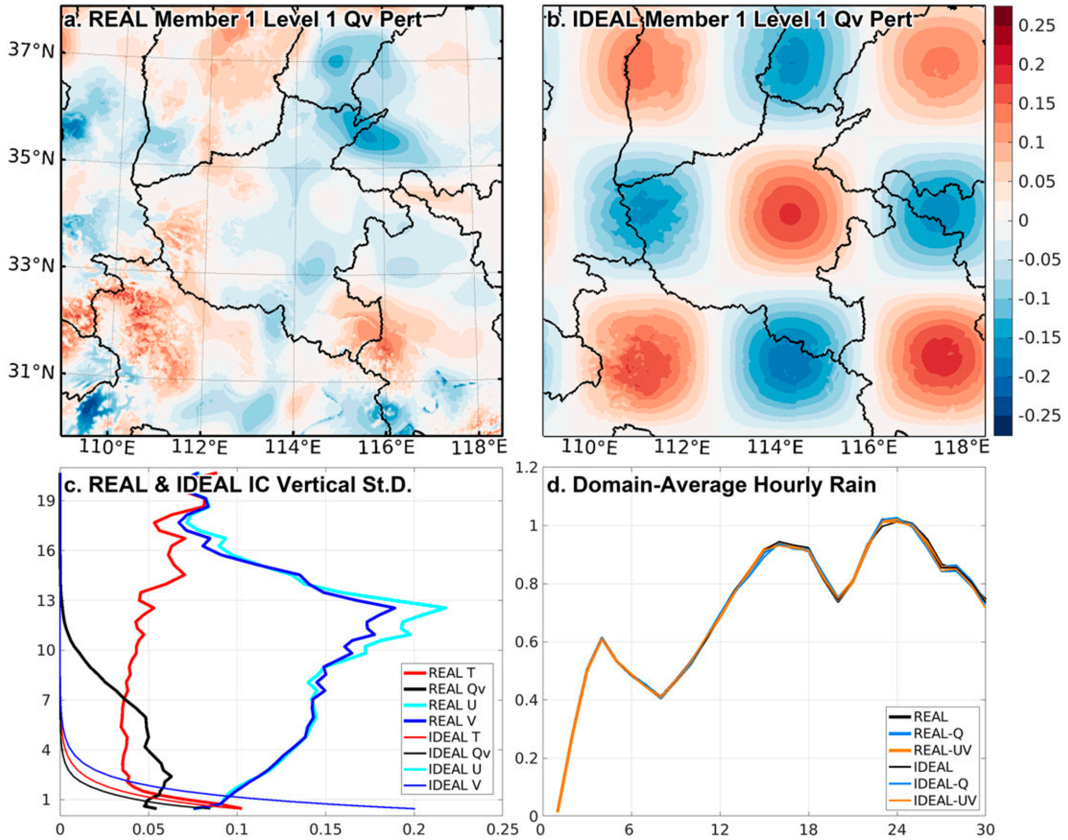

FIG. 2. Qv perturbation (in g kg^{-1}) at the first model level for (a) the first member of REAL and (b) the first member of IDEAL, (c) vertical distribution of level-average standard deviation for the initial uncertainties of T (in K), Qv (in g kg^{-1}), U (in m s^{-1}), and V (in m s^{-1}) from the REAL and IDEAL ensembles, and (d) domain-average ensemble-mean hourly rain (in mm) for the six ensembles.

valid at 0600 UTC 19 July 2021. Since Weyn and Durran (2017, 2019) only added Qv perturbations, the second ensemble forecasts also only added Qv perturbations to the initial conditions (i.e., all ICs have identical T, U, and V, as well as other unperturbed variables) and it will be referred to as "REAL-Q." As a comparison, the third ensemble forecasts added U and V perturbations instead, and it will be referred to as "REAL-UV." Note that perturbations of REAL's ICs are used for REAL-Q's Qv perturbations and REAL-UV's U and V perturbations. Pressure, geopotential, and vertical wind component were not perturbed, because they are calculated using other variables during the WRF initializations and their perturbations are ignored. Because of the small amplitude of these initial perturbations, the additional imbalances introduced by these perturbations compared with a forecast initialized using unperturbed GFS analysis are very small.

The other three ensembles employed idealized, univariate, isotropic initial uncertainties based on the following equation described in Weyn and Durran [2017, 2019; their Eq. (1)]:

$$\delta = ae^{-z/H}\sin\left[2\pi\left(\frac{x}{L}-\phi_x\right)\right]\sin\left[2\pi\left(\frac{y}{L}-\phi_y\right)\right],$$

where δ is the perturbations of T, Qv, U, or V, a is the perturbation amplitude, H is the e-folding height scale, L is the horizontal scale of the perturbations, ϕ_x and ϕ_y are random, member-specific phase offsets, and x, y, and z are the model-grid distances in kilometers. We set H to 1 km and L to 600 km (corresponds to a power spectrum of the perturbations that peaks at 300 km). The differences between two sets of perturbations are primarily determined by their zonal and meridional phase differences, ϕ_x and ϕ_y. The three ensembles with idealized IC perturbations are formed in the same manner as the three ensembles with realistic IC perturbations: IDEAL's ICs contain perturbations of T, Qv, U, and V; IDEAL-Q's ICs only contain Qv perturbations; IDEAL-UV's ICs contain U and V perturbations. The magnitude constant in Eq. (1) for Qv perturbations follows Weyn and Durran (2019) and was set to 1% of the background Qv values (using GFS analysis), and was set to 0.2 K, 0.4 m s^{-1}, and 0.4 m s^{-1} for T, U, and V, respectively. These values are selected to make the peak values of the power spectra of their perturbations (at 300 km) comparable to the power spectra values at the same wavelengths for REAL's IC perturbations.

Figure 2 shows examples of the structure of the IC perturbations of REAL and IDEAL. IDEAL's IC perturbations

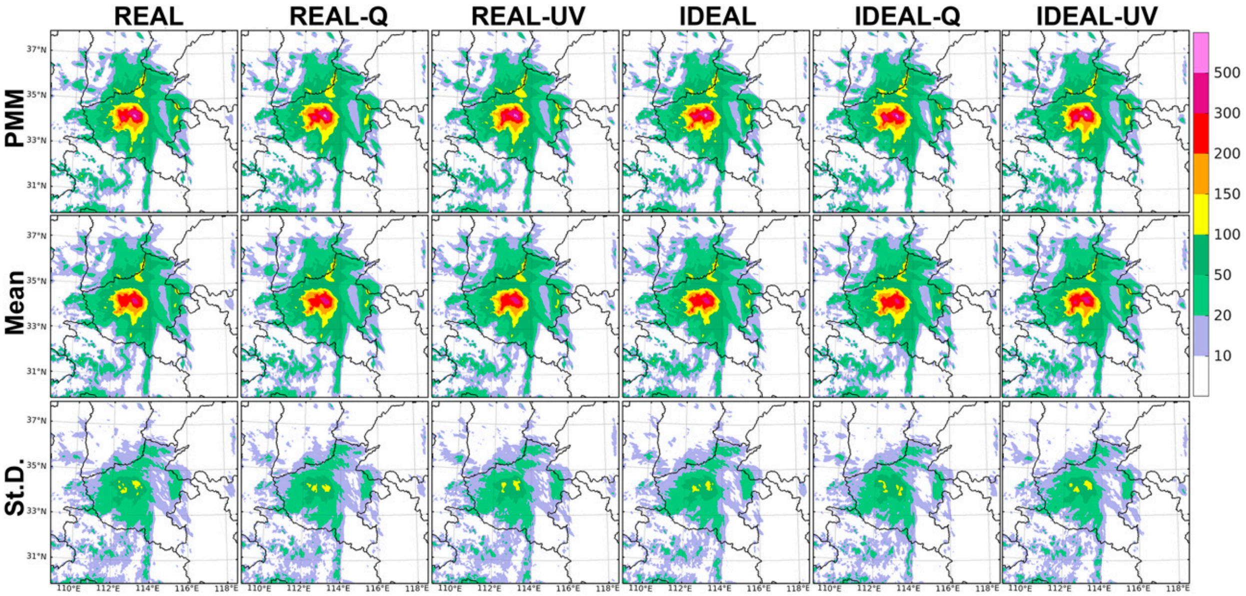

FIG. 3. (top) Probability-matched mean (PMM), (middle) arithmetic mean, and (bottom) standard deviation (St.D.) of 24-h accumulated rainfall (unit: mm) from 1200 UTC 19 Jul to 1200 UTC 20 Jul 2021 for the six ensembles.

have the largest standard deviations occurring at the lowest model level and decreasing exponentially with increasing height for all four variables (Fig. 2c). On the other hand, REAL's IC perturbations show different vertical characteristics for each variable, and Qv does not decrease exponentially going upward (Fig. 2c). Their IC perturbations' horizontal distributions are also vastly different (Figs. 2a,b): compared with the more uniform pattern in IDEAL's IC perturbations (Fig. 2b), REAL's IC perturbations have more horizontal variability (Fig. 2a).

The probability-matched ensemble mean (Ebert 2001), arithmetic mean, and standard deviation of the 24-h accumulated rainfall from 1200 UTC 19 July to 1200 UTC 20 July for the six ensembles is shown in Fig. 3. Apparently, all six ensembles share a common structure for both their rainfall forecasts and associated uncertainties, with only minimal differences in the placement of the maximum values. The arithmetic means show good correspondence with the probability-matched means (PMMs) for 100-mm rainfall, but the magnitudes of the strongest rainfall differ notably between arithmetic means and PMMs. Combined with large ensemble standard deviations that exceed 100 mm in the intense rainfall region, they suggest good consistencies of the intense rainfall region but large uncertainties in predicting the locations of the most extreme rainfall. Previous studies suggest that forecast uncertainties beyond 6–8 h are insensitive to the amplitude (when it is already very small) or the spatial scale of the initial uncertainties (e.g., Durran and Gingrich 2014; Nielsen and Schumacher 2016; Weyn and Durran 2017; Z22); it seems that the forecast uncertainty of rainfall forecast in the deterministic forecast sense is also not sensitive to the structure of the initial uncertainties in this event.

All power spectra and scale decompositions in this study are produced using the 1Dκ method as suggested by Rotunno et al. (2023) on the ensemble member perturbations (subtracting the ensemble mean from each ensemble member at a given forecast lead time) of the 3-km domain. All vertical levels are used when calculating the power spectra to better represent their characteristics over the entire atmosphere.

3. Power spectra of forecast uncertainties

Let us start with examining the power spectra of forecast uncertainties ("error spectra" hereafter) of different variables in the ensembles, represented by the 40-member logarithmic-scale averages of each member's power spectra of its deviations from the ensemble mean for a certain variable. At 0-h forecast lead time (Fig. 4a), Qv error spectra of REAL and REAL-Q show a plateau of error energy at scales larger than ~200 km and error energy decreases when moving toward smaller scales, while IDEAL and IDEAL-Q show a peak of error energy at 300 km and error energy decreases when moving toward either larger or smaller scales (note that the REAL and IDEAL lines are covered by the REAL-Q and IDEAL-Q lines, respectively, because they have identical values), both are consistent with the characteristics of how these initial uncertainties are generated. On the other hand, the error spectra of REAL-UV and IDEAL-UV—which do not include any initial Qv perturbations—are 6–12 orders of magnitude smaller than the other four experiments.

An hour later, while REAL and REAL-Q (IDEAL and IDEAL-Q) have identical initial Qv perturbations, REAL-Q (IDEAL-Q) has smaller Qv error energy than REAL (IDEAL) at scales roughly smaller than 60 (100) km (Fig. 4b). This difference suggests that REAL-Q and IDEAL-Q have slightly slower error growth rate at the small scales during the first hour than REAL and IDEAL. On the other hand, while there are practically no initial Qv perturbations in REAL-UV and IDEAL-UV

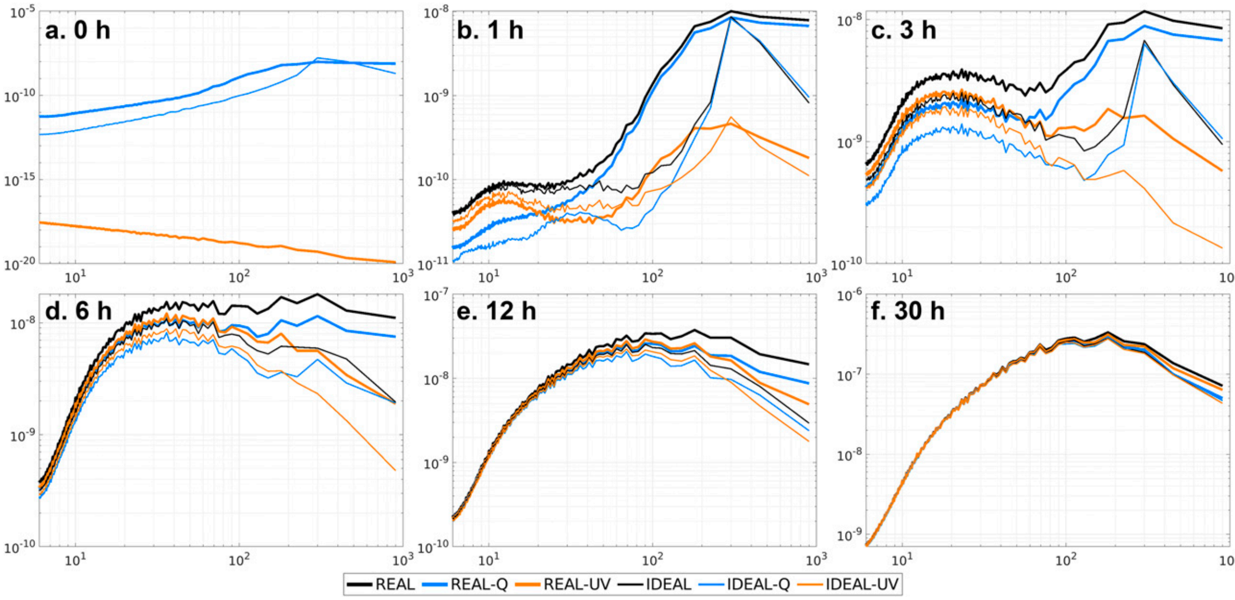

FIG. 4. Qv error spectra of all ensembles at (a) 0-, (b) 1-, (c) 3-, (d) 6-, (e) 12-, and (f) 30-h forecast lead times. The *x* axis is wavelengths (in km).

(Fig. 4a), their errors at the smallest scales grow rapidly during the first hour and even exceed the magnitude of the error energy of REAL-Q and IDEAL-Q at scales smaller than 20 km and are much closer to the magnitude of the error energy of REAL and IDEAL (Fig. 4b). These differences in the error growth rates during the first hour suggests that not considering initial uncertainties of all atmospheric states may lead to underestimation of their error growth rate even when flow-dependent initial uncertainties are applied. We also find that the error spectra of REAL, REAL-Q, IDEAL, and IDEAL-Q at wavelengths > 200 km do not grow for the first hour.

The faster Qv error growth in REAL-UV and IDEAL-UV than in REAL-Q and IDEAL-Q continues. As of 3 h into the forecast, the Qv error energy of REAL-UV (IDEAL-UV) remains greater than that of REAL-Q (IDEAL-Q) at scales up to 50 (80) km (Fig. 4c). In general, the error spectra of all six ensembles at small scales have grown more than one order of magnitude compared with 2 h before, while the error spectra of REAL, REAL-Q, IDEAL, and IDEAL-Q at wavelengths > 200 km still grow slowly when comparing Figs. 4b and 4c. The fast error growth at small scales and the stagnation of error growth at large scales create a shallow "pit" of lower-magnitude error spectra from roughly 20 to 200 km in REAL, REAL-Q, IDEAL, and IDEAL-Q. As the forecast lead time extends, the gaps between Qv error spectra of the six ensembles diminish at smaller scales first (i.e., errors at these scales have saturated) when there are still notable differences at larger scales (Figs. 4d,e), and the differences at larger scales also continue to decrease until the end of the ensemble forecasts—the differences are more than one order of magnitude at forecast lead time of 6 h (Fig. 4d) and reduce to slightly under one order of magnitude at 12 h (Fig. 4e) and less than 2 times at 30 h (the end of the ensemble forecasts; Fig. 4f).

The temporal evolutions of Qv error spectra of IDEAL, IDEAL-Q, and IDEAL-UV notably share many similarities with those of REAL, REAL-Q, and REAL-UV in Fig. 4. Despite very different initial Qv uncertainties (Fig. 4a), Qv error spectra of REAL and IDEAL are very similar for smaller scales up to ~20 km after 1 h of forecast, which is also the situation between REAL-Q and IDEAL-Q, as well as REAL-UV and IDEAL-UV (Fig. 4b). IDEAL-UV has faster Qv error growth at small scales than IDEAL-Q, just like REAL-UV and REAL-Q. IDEAL-UV's smaller-scale Qv errors also grow larger than its larger-scale Qv errors, just like REAL-UV (e.g., Fig. 4c). The similarities suggest that the error growth mechanisms might be insensitive to the two structures of initial large-scale uncertainties that are applied in this study and section 4 will provide examinations of their evolutions in physical space.

To signify the characteristics of the early error growth stage, Fig. 5 shows the 0–12-h temporal evolution of Qv error spectra, as well as the hourly error growth ratio (ratios of error spectra from consecutive hours), of each of the ensembles. The stagnation of error growth during the first several hours previously discussed in Fig. 4 is shown more clearly: errors at scales greater than ~200 km in REAL and REAL-Q grow slowly for the first several hours when errors at scales smaller than 200 km show rapid growth during this period (Figs. 5a1,b1). This behavior is not particular to the magnitude of initial uncertainties of REAL and REAL-Q; similar stagnation occurs at large scales even if we further reduce the amplitude of the initial uncertainties by five orders of magnitudes, although the stagnation lasts for a shorter time as we reduce the amplitude (figure not shown). There is also a clear error saturation at small scales as the forecast extends (the decrease at smallest scales around 6 h is associated with reduced hourly rainfall as shown in Fig. 2d).

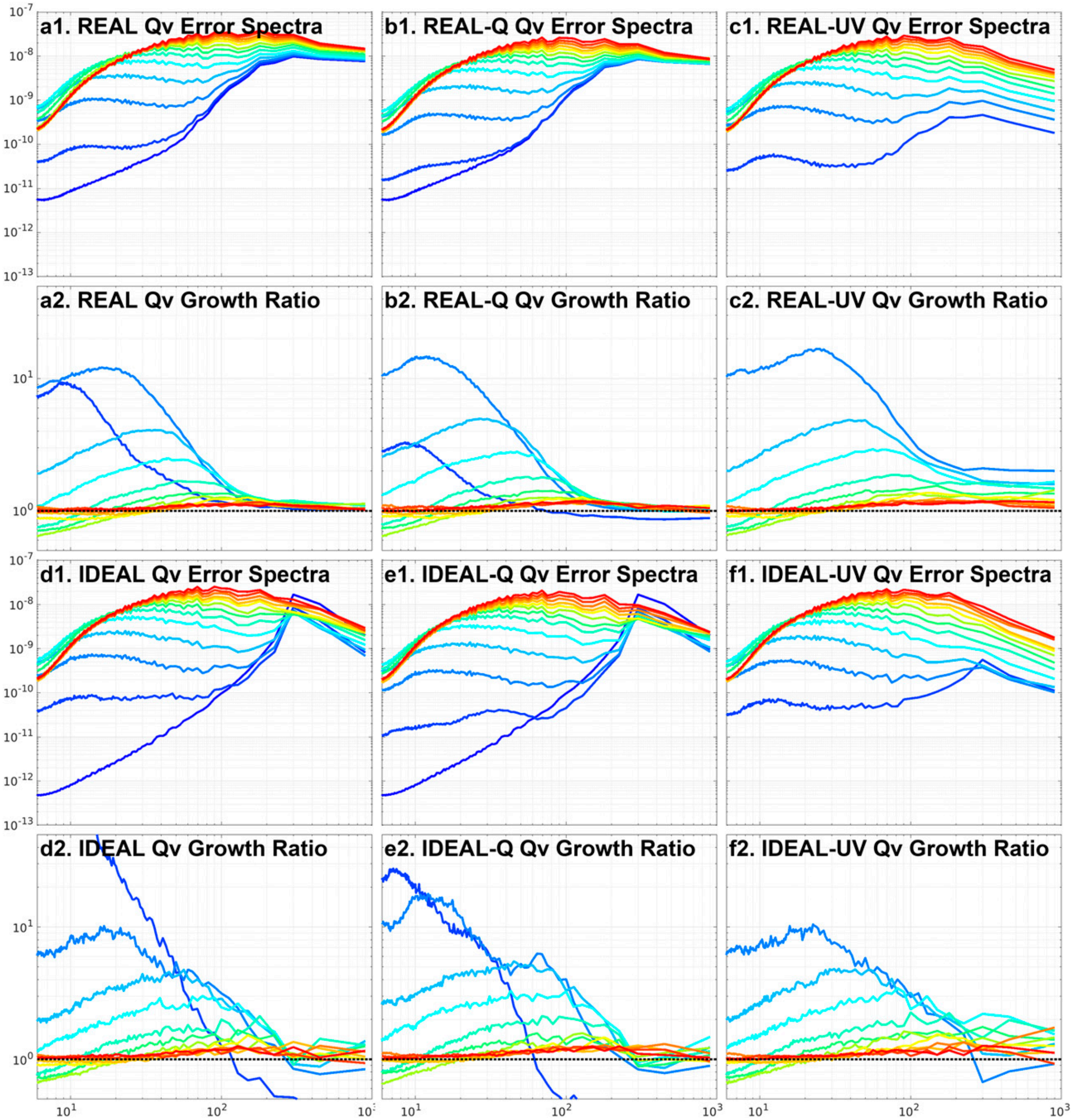

FIG. 5. Qv's 0–12-h (first row),(third row) forecast error spectra and (second row),(fourth row) hourly forecast error spectra growth ratio from the (a) REAL, (b) REAL-Q, (c) REAL-UV, (d) IDEAL, (e) IDEAL-Q, and (f) IDEAL-UV ensembles. Blue colors are earlier hours (shorter forecast lead times), and red colors are later hours (longer forecast lead times). Error spectra start from 0-h forecasts, whereas hourly growth ratios start from 1-h forecasts. The *x* axis is wavelengths (in km).

If we disregard the part of the error spectra at the large scales, the evolution of the error spectra is typical of an upscale error growth behavior: the error spectra peak at a small-scale wavelength that is very close to—but not directly at—the largest error saturation wavelength, and the error growth slows down as the errors saturate progressively larger scales. Unlike Fig. 1a of Rotunno and Snyder (2008), the error spectra peaks in a real-world NWP model forecast never coincide with the largest error saturation scale because of the strong inhomogeneities of small-scale error saturation associated with convective activities (e.g., Judt 2018; Selz 2019; Selz et al. 2022; Rotunno et al. 2023). The slowing down of error growth is even more apparent when looking at the error growth ratios (Figs. 5a2,b2): from earlier to later hours, the peaks of the error growth ratios move from smaller to larger scales with a decreasing magnitude of the peak, meaning that

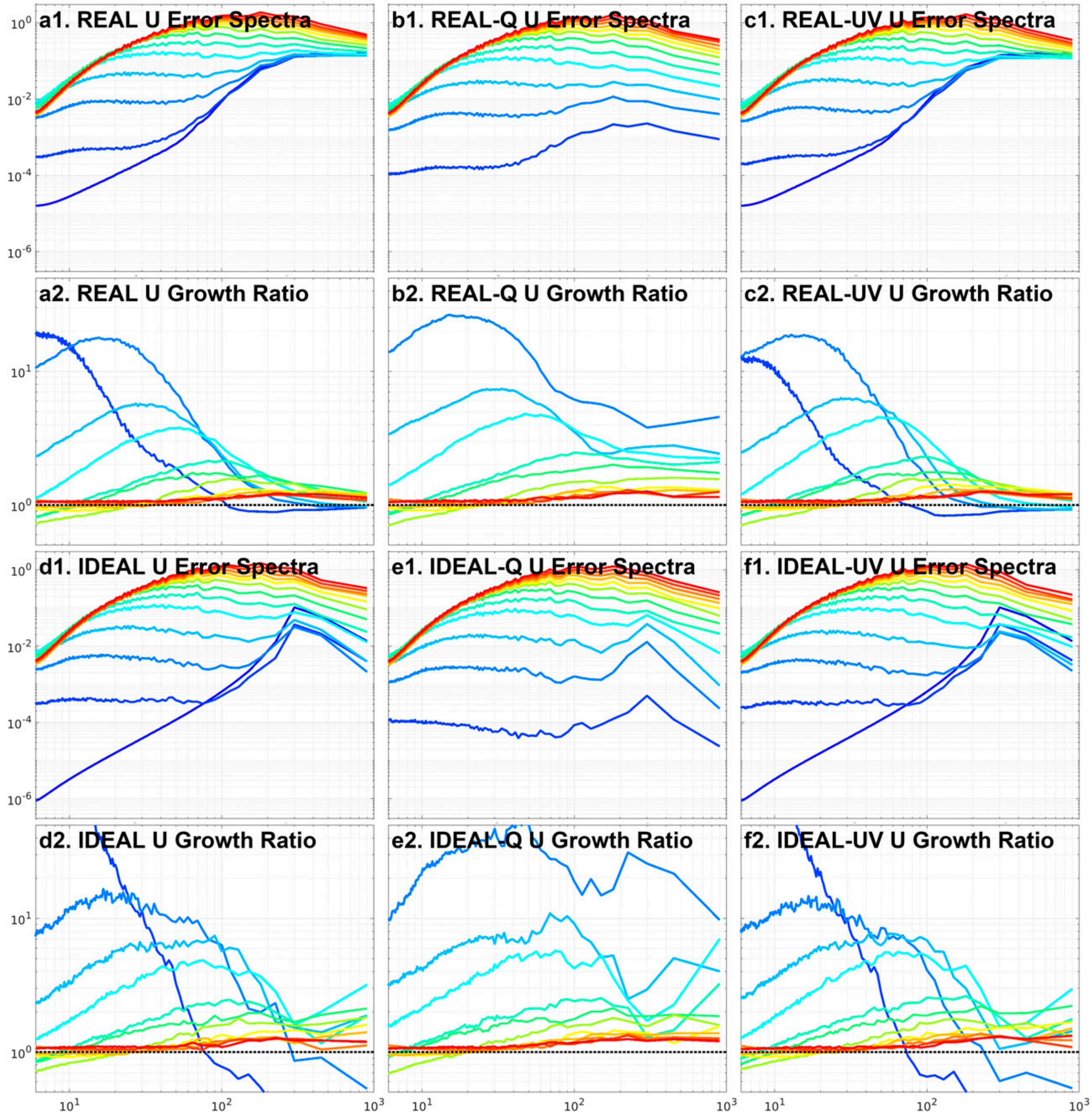

FIG. 6. As in Fig. 5, but for *U* error spectra and their hourly growth ratios.

the scales with the fastest error growth shift from smaller to larger scales with a slowing down of the fastest error growth speed as the forecast lead time extends and progressively larger scales become saturated.

On the other hand, REAL-UV's Qv error spectra—which have no uncertainties initially—show slightly different characteristics (Fig. 5c1) that errors grow simultaneously at all scales, which signifies the up-amplitude error growth characteristics in previous studies. Errors localized in physical space are broad in spectral space (e.g., Lloveras et al. 2022), which could have partly contributed to this feature. However, even from Fig. 5c1, we can still observe that during the first several hours, error growth at scales of ~10–100 km is much faster than at scales greater than ~100 km. This is also apparent in its error growth ratio (Fig. 5c2). Although there is notable error growth at larger scales, the ratios flatten out at scales greater than ~300 km and are much lower than the error growth ratios at smaller scales. Furthermore, the peaks of error growth rates at smaller scales in REAL-UV's Qv error spectra (Fig. 5c2) are similar to those of REAL and REAL-Q (Figs. 5a2,b2) with a shift toward larger scales and decreasing peak magnitudes as the forecasts continue, suggesting that the upscale error growth

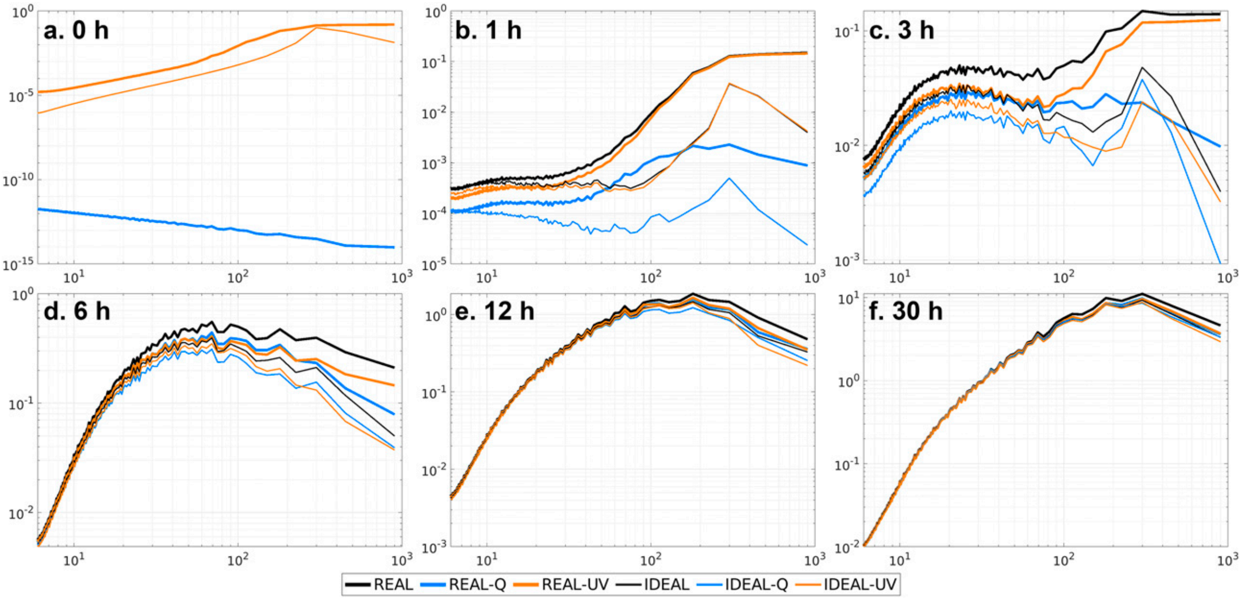

FIG. 7. As in Fig. 4, but for U error spectra.

mechanisms are at least partially responsible for the error growth processes of Qv in this ensemble at these scales.

When idealized instead of flow-dependent uncertainties are imposed, they do not fundamentally change the characteristics of the evolution of the error spectra. There are apparent adjustment processes with error spectra at scales larger than 300 km decreasing for the first several hours (Figs. 5d1–f1), likely due to imbalances introduced by ignoring the covariances between the

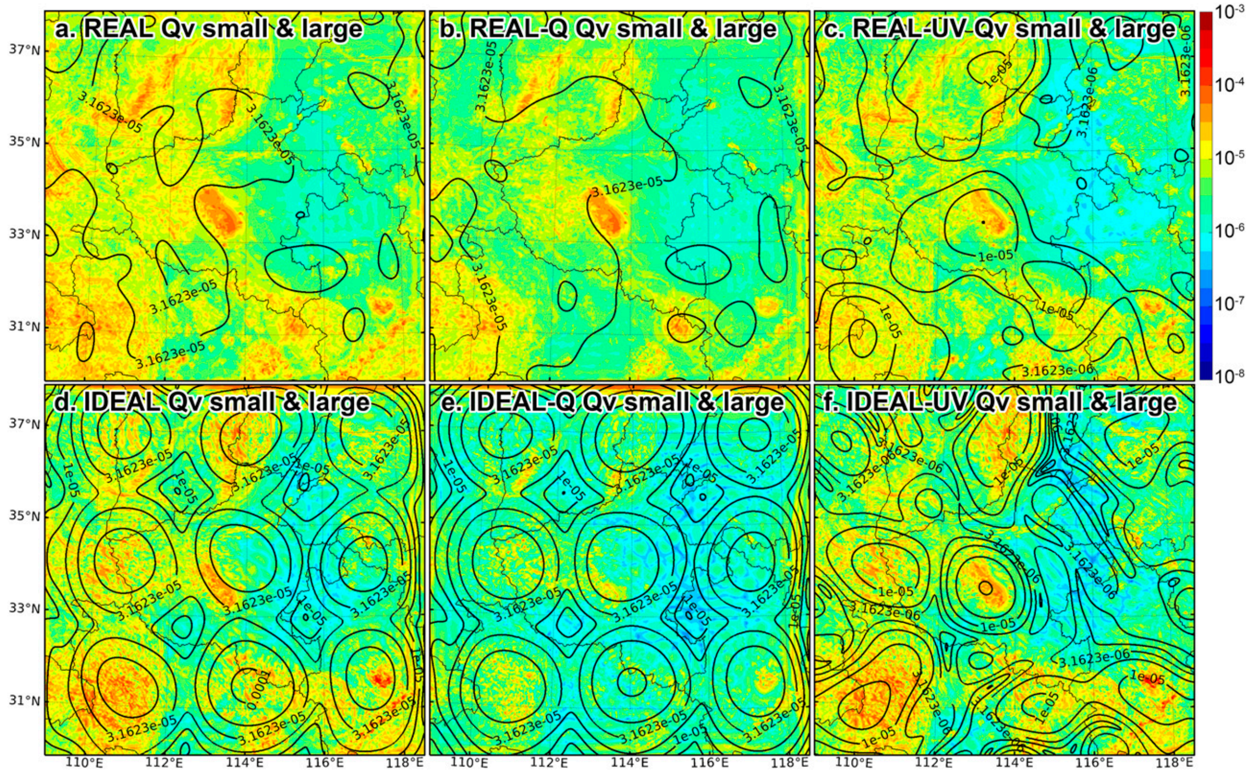

FIG. 8. Small-scale (shaded) and large-scale (contoured every 0.25 on a base-10 logarithmic scale) Qv standard deviation (in kg kg^{-1}) at the first model level for the (a) REAL, (b) REAL-Q, (c) REAL-UV, (d) IDEAL, (e) IDEAL-Q, and (f) IDEAL-UV ensembles at 1-h forecast lead time.

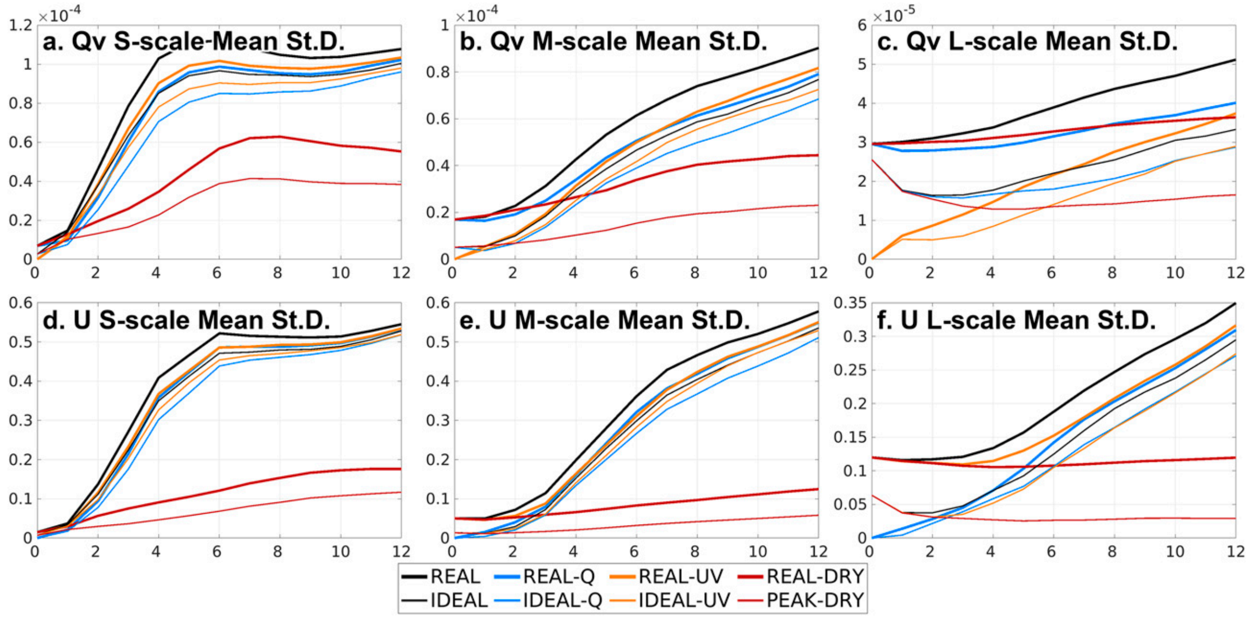

FIG. 9. Temporal evolution (for the first 12 h) of the domain-average standard deviation of (top) Qv (in kg kg^{-1}) and (bottom) U (in m s^{-1}) at the (left) small (<50 km), (center) medium (50–200 km), and (right) large (>200 km) scales.

uncertainties of different variables, and therefore, they are decorrelated. Their error growth ratios, albeit noisier than the experiments with flow-dependent initial uncertainties, still show peaks moving toward larger scales with decreasing peak magnitudes as the forecasts continue (Figs. 5d2–f2) that are similar to the behaviors of the REAL, REAL-Q, and REAL-UV ensembles (Figs. 5a2–c2), suggesting that the leading role of upscale over up-amplitude error growth is insensitive to the structure of imposed large-scale initial uncertainties in this event.

The above characteristics of error spectra and error growth ratios in Figs. 4 and 5 are not just limited to the forecast errors of Qv alone. Figures 6 and 7 show the error spectra and error growth ratios for the U wind component for the six ensembles. Again, we see that errors at scales larger than ~200 km do not grow for the first several hours in REAL and REAL-UV (Figs. 6a1,c1 and 7a–c), consistent with Qv error spectra in REAL and REAL-Q (Figs. 4a–c and 5a1,b1); we see that peaks of error growth ratios shift toward larger scales with decreasing magnitudes of the peaks as the forecast lead time extends, more apparent in REAL and REAL-UV (Figs. 6a2,c2) and less apparent in IDEAL and IDEAL-UV (Figs. 6d2,f2); and we see that when there are no initial uncertainties in U, as in REAL-Q and IDEAL-Q, their U errors grow simultaneously at all scales (Figs. 6b1,e1); however, errors at smaller scales still grow faster for the first several hours (Figs. 6b2,e2).

When comparing the U error spectra of the six ensembles (Fig. 7), they show characteristics that are generally consistent with the Qv error spectra (Fig. 4), but it is REAL-Q and IDEAL-Q that do not have initial perturbations in the examined variable (U), instead of REAL-UV and IDEAL-UV that do not have initial Qv perturbations when Qv is examined. We can see a quick fill-up of error spectra at small scales during the first hour of forecast when large-scale errors show larger discrepancies across the experiments (Fig. 7b). REAL and REAL-UV (as well as IDEAL and IDEAL-UV) have almost identical error spectra at scales larger than 200 km at 1-h lead time with very little error growth at these scales (e.g., Figs. 6a,c,d,f), while REAL-UV (IDEAL-UV) has slightly smaller error spectra at scales smaller than 200 km than REAL (IDEAL), again confirming that incomplete descriptions of uncertainties in terms of perturbed atmospheric state variables may lead to underestimated error growth rates. The smaller-scale U errors become indistinguishable the earliest as the forecasts continue, followed by larger-scale errors becoming progressively closer across the ensembles (Figs. 7c–f).

Based on the analyses of temporal evolutions of error spectra of these six ensembles, it is found that the upscale error growth may have played a dominant role in this event based on the error growth characteristics, and its importance is insensitive to the structure of initial large-scale uncertainties. It also shows that the error growth characteristics of variables that do not have initial uncertainties (i.e., U error spectra of REAL-Q and IDEAL-Q and Qv error spectra of REAL-UV and IDEAL-UV) will show characteristics of up-amplitude error growth, but their error growth ratios reveal that error growth at smaller scales and their upscale transfer still likely dominates their error growth processes. The analyses so far are performed completely in error-spectrum space, and next, we will examine their evolutions in physical space.

4. Horizontal and vertical evolutions of forecast uncertainties at different scales

To complement previous spectrum analyses and provide more insights into how forecast errors grow, this section

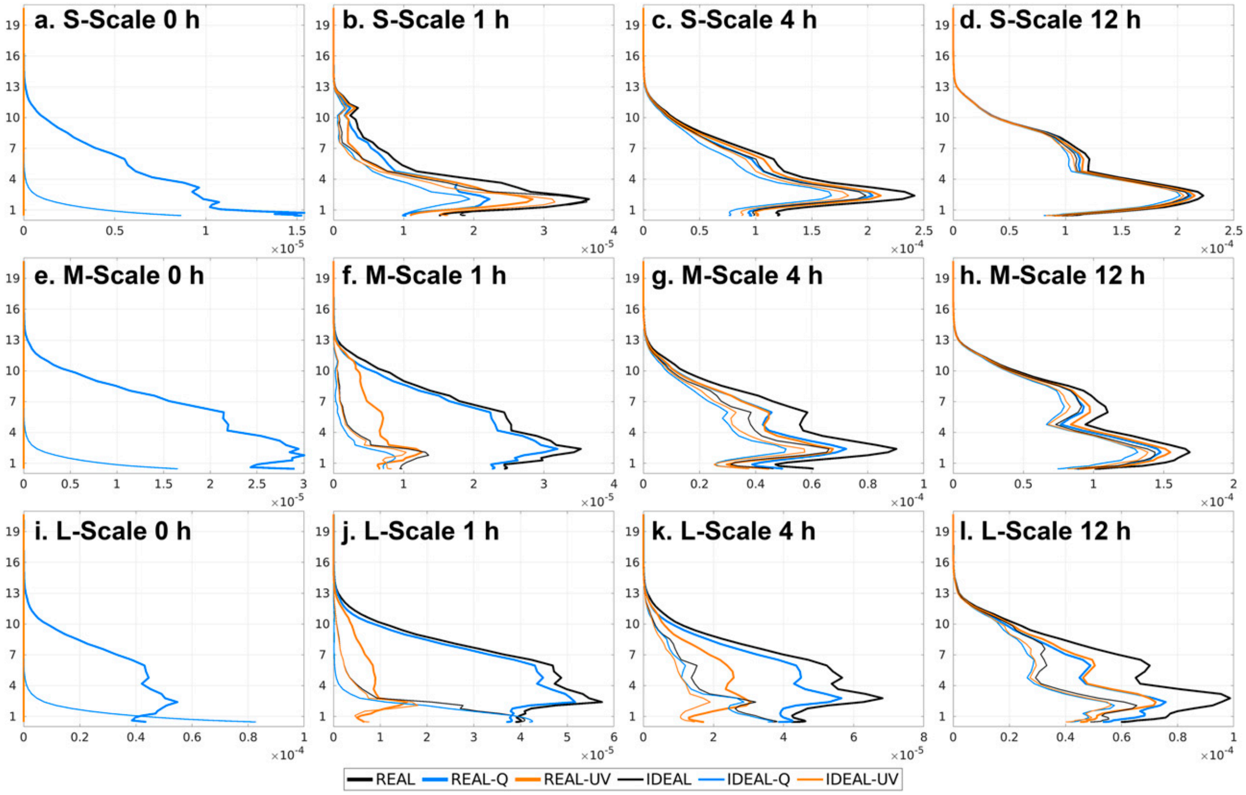

FIG. 10. Vertical distributions of Qv standard deviation (in kg kg^{-1}) at (top) small (<50 km), (middle) medium (50–200 km), and (bottom) large (>200 km) scales at (from left to right) 0-, 1-, 4-, and 12-h forecast lead times, respectively.

examines how ensemble spread at different scales evolves in the physical space. We first calculate the gridwise ensemble spread (in terms of standard deviation) of different state variables, then decompose the ensemble spread into three bands of scales: small scale (<50 km), medium scale (50–200 km), and large scale (>200 km). The threshold of 200 km is chosen based on the characteristics that the error spectra of initial uncertainties of REAL, REAL-Q, and REAL-UV show a rapid decrease when moving toward smaller scales across ~200 km, and those of IDEAL, IDEAL-Q, and IDEAL-UV peak at 300 km; the threshold of 50 km is chosen as an intermediate value between 200 km and the smallest resolvable scale of the ensembles (6 times the model's grid spacing, or 18 km here).

The previous spectrum analyses suggest that error growth at smaller scales during the first several hours is very similar among the six ensembles. This is supported by Fig. 8, which shows the small-scale and large-scale Qv spread at the first model level after one hour of forecast as an example. The small-scale errors at this time are very similar among the six ensembles, although the magnitudes of the errors in REAL-Q and IDEAL-Q (Figs. 8b,e) are smaller than the other four ensembles, which is consistent with their slightly smaller Qv error spectra at this time (Fig. 4b). In fact, the temporal evolution of small-scale errors for all six ensembles are almost parallel to each other during the first several hours of forecast (Fig. 9a). On the other hand, the large-scale errors of REAL-Q (IDEAL-Q) have very similar structures and magnitudes with REAL (IDEAL) and strongly resemble their structures of initial uncertainties (e.g., Fig. 2). Large-scale Qv errors of REAL-UV and IDEAL-UV show completely different structures with overall lower values than the other four ensembles. Yet there are relatively higher values in REAL-UV and IDEAL-UV collocated with larger small-scale Qv errors (Figs. 8c,f), consistent with the error spectra analysis (Figs. 5c,f) that isolated features in the physical space can project to a wide range of wavenumbers in the spectral space.

To quantify the evolutions of errors that originate from the initial large-scale uncertainties and serve as baselines of error evolution at different scales (primarily at the large scale), two additional "fake-dry" experiments were executed by turning off the cumulus parameterization scheme and latent heating associated with the microphysical parameterization scheme of the REAL and IDEAL ensembles. In this way, the small-scale error amplification associated with latent heating is eliminated, and its upscale transfer is also therefore excluded. These two "fake-dry" ensembles are denoted as "REAL-DRY" and "IDEAL-DRY" in Fig. 9. We can see in Fig. 9a that small-scale Qv error growth in REAL-DRY and IDEAL-DRY are much slower compared with the other ensemble. This slow small-scale growth is associated with interactions between hydrometeors and water vapor within the microphysical scheme (that is not turned off). However, since these interactions have no latent heating and therefore no amplification and upscale transfer of the small-scale errors, there are

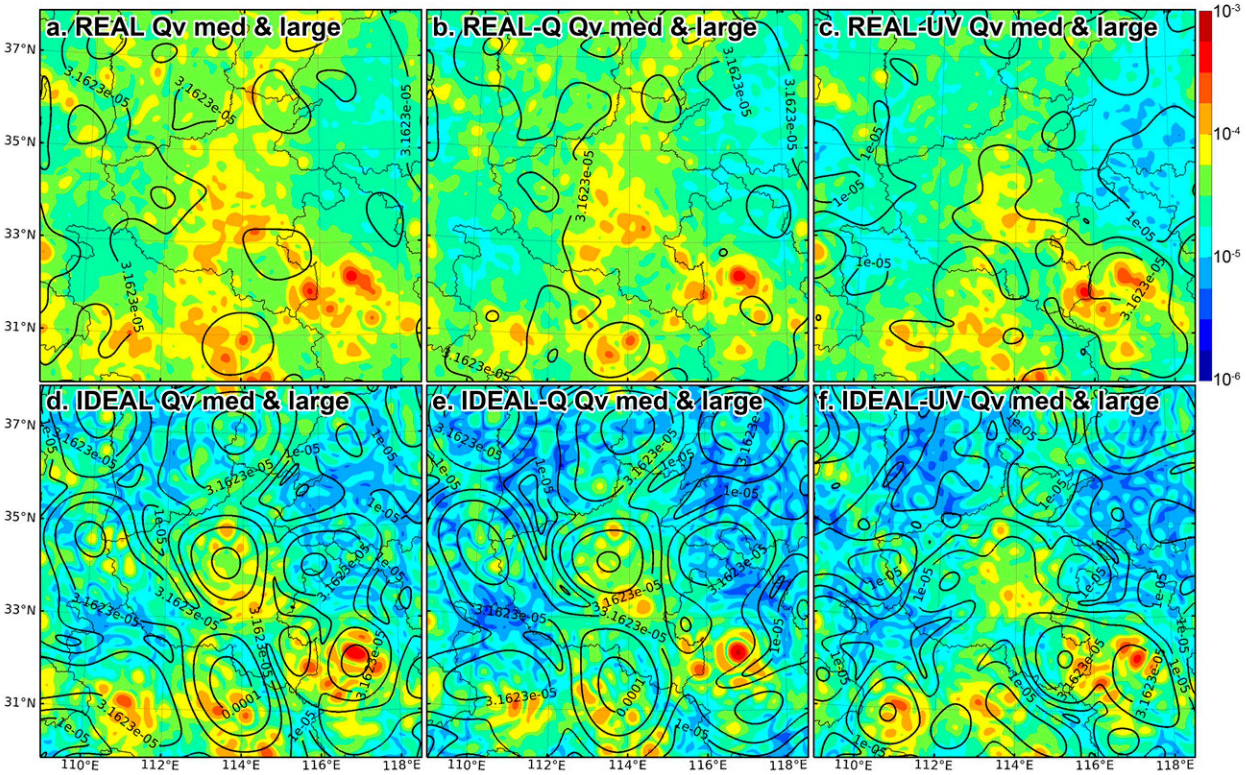

FIG. 11. As in Fig. 8, but for medium-scale (shaded) and large-scale (contoured every 0.25 on a base-10 logarithmic scale) at 4-h forecast lead time.

only limited error growth at the medium scale (Fig. 9b) and almost no growth at all at the large scale (Fig. 9c) in REAL-DRY and IDEAL-DRY. This behavior also confirms the inability of the initial large-scale uncertainties in REAL or IDEAL to amplify.

The similarities and differences of Qv errors at different scales among the six ensembles also occur at other levels (figure not shown), as well as their vertical distributions of level averages (Fig. 10). Despite very different vertical structures and magnitudes for different experiments at different scales in their ICs (Figs. 10a,e,i), all six experiments show very similar vertical structure at the small scales at the 1-h lead time (Fig. 10b). On the other hand, although vertical structures of REAL and REAL-Q (IDEAL and IDEAL-Q) are also very similar at medium and large scales, REAL-UV (IDEAL-UV) shows a distinctly different vertical structure (Figs. 10f,j). These characteristics are consistent with what we have seen in the horizontal distributions of errors at different scales in Fig. 8.

As the forecasts extend, the distributions of errors at small scales remain similar across the six ensembles (Figs. 10c,d). On the other hand, as the small-scale errors saturate after about 3–4 h (e.g., Figs. 4 and 5), they start to transfer to relatively larger scales. At the medium scale, after a 2-h period of no error growth at the beginning of the forecast (Fig. 9b), errors across the six ensembles become very similar at 4-h forecast lead time both horizontally (Fig. 11) and vertically (Fig. 10g). At the same time, large-scale Qv errors grow very slowly as shown by the domain-average large-scale Qv standard deviation and compared with the "fake-dry" baselines (Fig. 9c), except for REAL-UV and IDEAL-UV. This is consistent with the temporal evolutions of their respective error spectra at larger scales that almost do not increase for the first several hours (Figs. 5a,b,d,e). Additionally, the large-scale errors of REAL and REAL-Q (IDEAL and IDEAL-Q) remain relatively similar to each other (Figs. 10k and 11a,b,d,e) as well as how they look like three hours earlier (Figs. 10j and 8a,b,d,e), and the magnitude of the large-scale errors remain largely unchanged if we compare the contours' values in Figs. 8 and 11. These characteristics suggest that the evolution of the large-scale errors until this time is primarily advection rather than amplification.

As the forecast lead time increases, large-scale errors of the six ensembles also start to become more and more alike both vertically (Fig. 10l) and horizontally (Fig. 12), despite their different initial uncertainty structures (Fig. 2) and very different earlier evolutions (Figs. 8 and 11). This suggests that the errors grown from smaller scales to larger scales have concealed the evolution (e.g., primarily advection with little growth as in the "fake-dry" baselines) of errors originated from the large scale. Additionally, domain-average large-scale error growth starts to speed up and deviate from the "fake-dry" baselines (Fig. 9c). Eventually, after 6–12 h, the errors of the six ensembles with different structures of initial uncertainties become qualitatively indistinguishable and impossible to trace back to

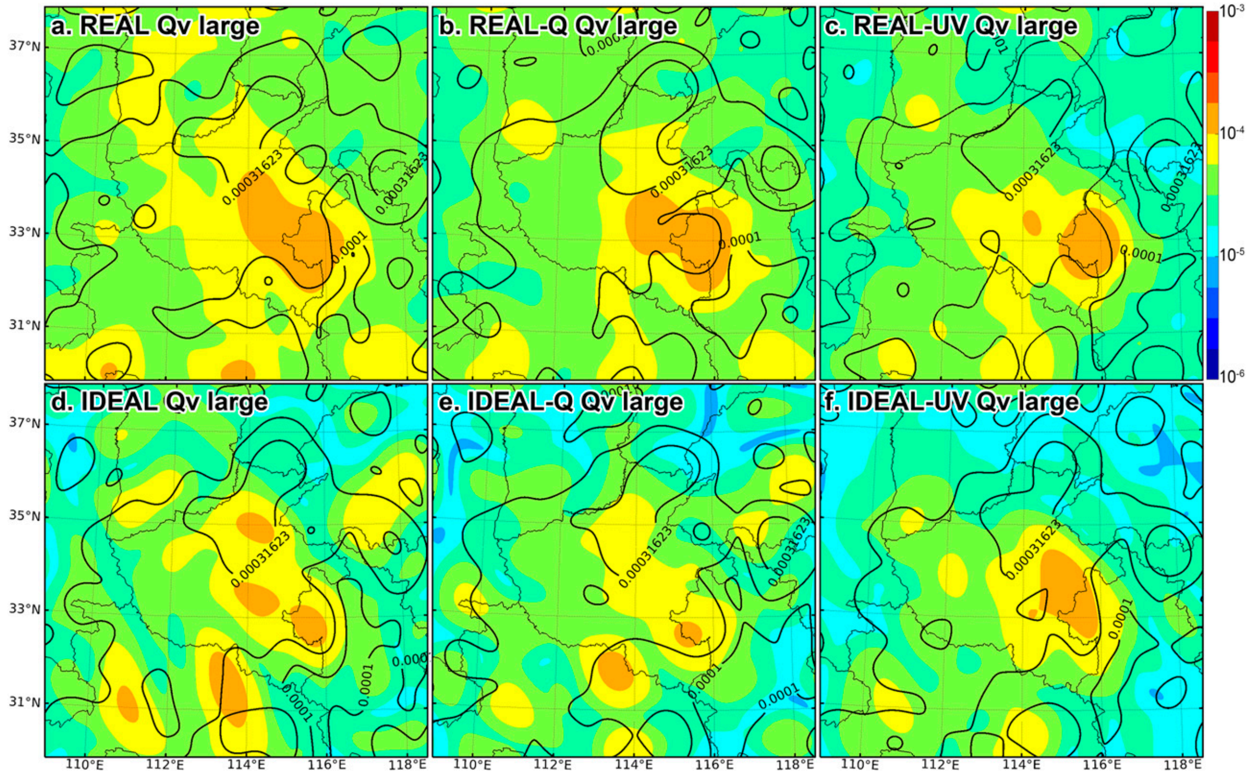

FIG. 12. As in Fig. 8, but for large-scale at 12-h (shaded) and 30-h (contoured every 0.25 on a base-10 logarithmic scale) forecast lead times.

their different initial structures (Fig. 12), and their outcomes at the end of the forecasts in terms of the accumulated rainfall are also qualitatively identical (Fig. 3).

Many previous studies have extensively examined the physical processes that are responsible for the growth and upscale transfer of smaller-scale errors to larger scales (e.g., Selz and Craig 2015; Bierdel et al. 2018; Baumgart et al. 2019; Selz et al. 2022). They found that moist convection (latent heat release) is primarily responsible for the amplification and saturation of errors at small (cloud and convective) scale; gravity wave is responsible for the outward propagation of the small-scale errors; geostrophic adjustment is responsible for transferring the propagated small-scale errors to more geostrophic balanced large-scale errors. Here, we examine the power spectra of horizontal divergence (simply "divergence" hereafter) and vertical vorticity (simple "vorticity" hereafter) of REAL's

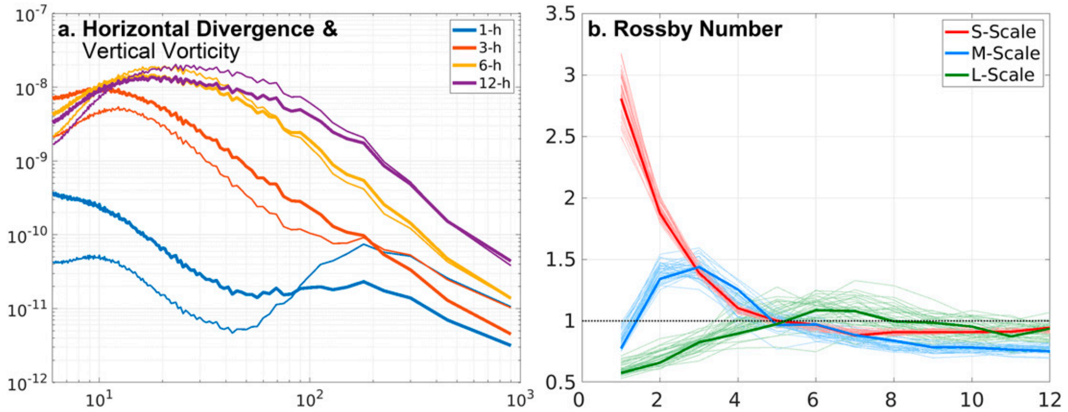

FIG. 13. (a) Power spectra of logarithmic-scale ensemble averages of horizontal divergence (thick lines) and vertical vorticity (thin lines) of REAL's ensemble perturbations at 1-, 3-, 6-, and 12-h lead times and (b) temporal evolutions of Rossby numbers between horizontal divergence and vertical vorticity calculated using REAL's perturbations decomposed into small, medium, and large scales. In (b), thin lines are Rossby numbers of each ensemble member and thick lines are the average of the ensemble members' Rossby numbers.

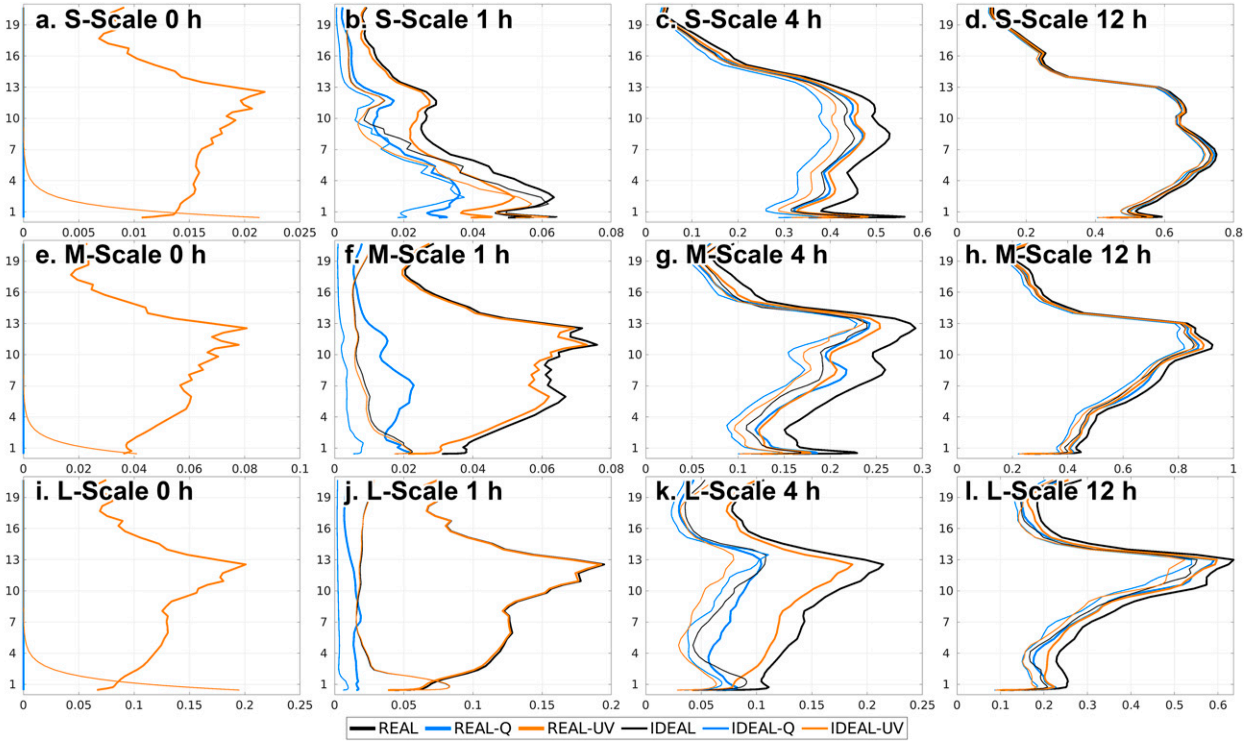

FIG. 14. As in Fig. 9, but for the vertical distribution of *U* standard deviations.

forecast ensemble perturbations (errors) of wind, as well as the ratios between the domain averages of these two quantities [i.e., "Rossby number" defined in Selz and Craig (2015) and Bierdel et al. (2018)]. Divergence and vorticity represent the irrotational and nondivergent wind components, respectively, which are not straightforward to derive using the Helmholtz decomposition on a regional model. These two quantities can reveal the contributions of moist convection and geostrophic adjustment to the overall error growth processes when calculated using the ensemble perturbations of wind.

Figure 13a shows that at the beginning (1-h forecast lead time) divergence of the wind perturbations is greater than the vorticity at smaller scales and vice versa at larger scales, leading to large small-scale Rossby number and small medium- and large-scale Rossby numbers at 1 h in Fig. 13b. However, smaller-scale vorticity quickly catches up with divergence, while at larger scales vorticity virtually does not increase for the first several hours when divergence steadily increases (Fig. 13a), consistent with the stagnation of error growth at these scales. This suggests that the geostrophic adjustment (vorticity) is becoming more important than the convective processes (divergence) at the smaller scales as the forecast extends, leading to decreasing small-scale Rossby number (Fig. 13b); at the same time, medium- and large-scale Rossby numbers increase as their divergence strengthen, although medium-scale Rossby number also starts to decrease after 3 h as vorticity (and hence geostrophic adjustment) strengthen (Fig. 13b). Last, large-scale Rossby number also starts to decrease after 6–7 h as the result of increasing geostrophic adjustment. After 12 h, vorticity is greater than divergence for most wavelengths (Fig. 13a), and the Rossby numbers at all three scale ranges are smaller than 1 (Fig. 13b). The evolutions of divergence, vorticity, and Rossby number are generally consistent with the upscale-growth mechanisms summarized above that convective activity leads the error amplification, followed by the geostrophic adjustment that transfers errors from smaller to larger scales.

The above evolution of Qv errors at different scales also holds for *U*, although there are notable differences. Despite vastly different structures in the initial perturbations (Figs. 14a,e,i), all six ensembles show very similar small-scale *U* errors at the 1-h forecast lead time both vertically (Fig. 14b) and horizontally (Fig. 15). In the meantime, the large-scale errors of REAL and REAL-UV (IDEAL and IDEAL-UV) also remain similar (Figs. 14j and 15a,c,d,f) and different from those of REAL-Q (IDEAL-Q; Figs. 14j and 15b,e). However, the onset of *U* error growth at the large scale is earlier than that of Qv and as early as 2 h after the start of the forecast (Fig. 9f). In fact, at 4-h lead time, in addition to very similar horizontal and vertical structures at the medium scale (Figs. 14g and 16), the large-scale errors of the six ensembles also start to look similar to each other both vertically (Fig. 14k) and horizontally (Fig. 16), even for the experiments that do not contain *U* initial uncertainties, such as REAL-Q (Fig. 16b) compared with REAL (Fig. 16a), and IDEAL-Q (Fig. 16e) compared with IDEAL (Fig. 16d), especially at regions collocated with larger medium-scale errors. The similarity of large-scale *U* errors at 4-h lead time suggests that the upscale error transfer in *U* starts earlier than Qv. This

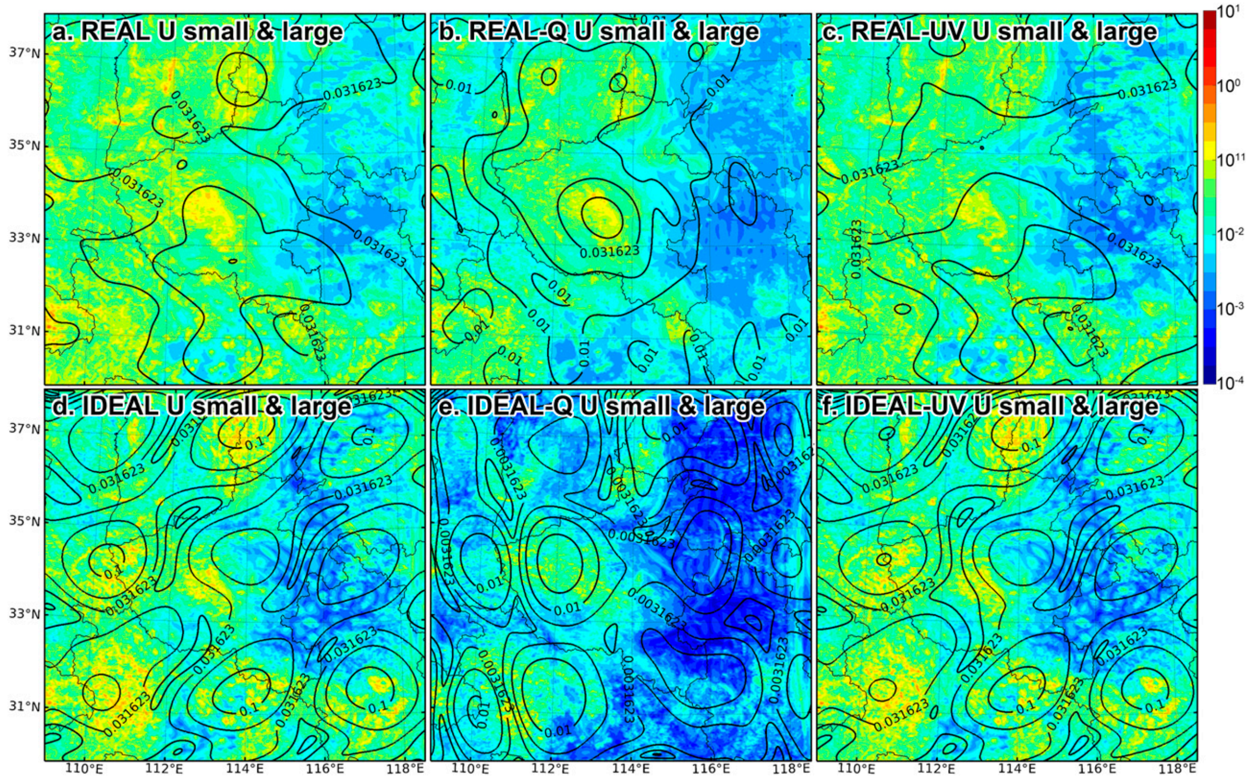

FIG. 15. As in Fig. 8, but for *U*.

earlier onset of upscale error transfer in *U* could be associated with the more direct impact of *U* and *V* in triggering and organizing convection than Qv, and the exact reason deserves further investigation.

Based on the analyses and comparisons of errors at different scales for the six ensembles, it is clear that the error growth processes at the small scales at the beginning of the forecasts are insensitive to the structure of the initial perturbations. The evolutions of the medium- and large-scale errors confirm the upscale transfer of small-scale errors which lead to similar structures of medium- and large-scale errors across the six ensembles. The evolution of the large-scale errors during the first several hours—when small- and medium-scale errors are too small and not comparable to the large-scale errors—are primarily advection instead of amplification, and they are eventually overwhelmed by the errors that grow upscale from smaller scales.

5. Conclusions and discussion

This study examines the impact of the structure of large-scale initial uncertainties with unobservably small amplitude on the intrinsic predictability's error growth characteristics. Six convection-allowing ensembles, each with 40 members, are executed for a record-breaking rainfall event that happened on 19–20 July 2021 in China, and they are initialized with different structures of large-scale small-amplitude initial uncertainties: the REAL ensemble adopted initial perturbations derived from the Global Ensemble Forecast System (GEFS), which has a grid spacing of 0.5° × 0.5°, but multiplied by a factor of 0.1, to represent realistic, multivariate, flow-dependent initial uncertainties; the IDEAL ensemble adopted initial perturbations generated by the same equation that Weyn and Durran (2017, 2019) used to represent idealized, univariate, isotropic initial uncertainties; the REAL-Q and REAL-UV ensembles only include REAL's initial uncertainties in Qv and *U*/*V*, respectively; and the IDEAL-Q and IDEAL-UV ensembles only include IDEAL's initial uncertainties in Qv and *U*/*V*, respectively. It should be noted that these initial uncertainties are much smaller than current global model analysis uncertainties, and this study examines the error growth characteristics in the intrinsic predictability regime.

Through both spectrum-space and physical-space analyses, it is shown that the error growth characteristics in this event are insensitive to the structure of the imposed large-scale small-amplitude initial uncertainties of the ensembles. For all six ensembles, forecast uncertainties always occur at similar locations with similar magnitude/amplitude at relatively smaller scales. For the first several hours, small-scale errors grow at similar locations with similar rates. On the contrary, large-scale errors grow very little, and further examination suggests that it is primarily advection rather than amplification that controls their evolutions during the first several hours. As the forecast extends, small-scale errors approach saturation in all six ensembles and start to transfer upscale; they overwhelm the errors at progressively larger scales and

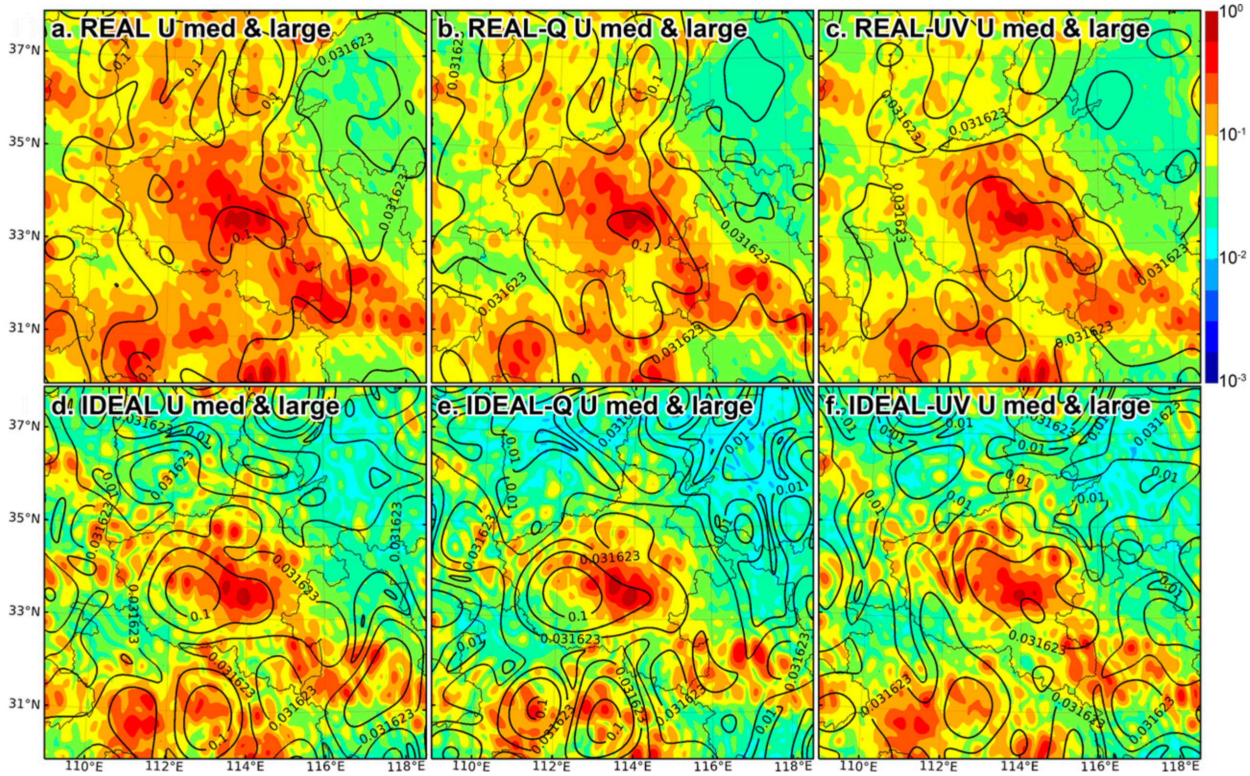

FIG. 16. As in Fig. 11, but for *U*.

make errors at these scales—for which the structures were different at the beginning—become similar. The error spectra also show clear upscale error growth characteristics regardless of the structures of the initial uncertainties: errors grow the fastest at relatively smaller scales; as the forecasts extend and errors at smaller scales saturate, the scales at which the errors grow the fastest gradually shift toward larger scales, accompanied by decreases in the fastest error growth rates.

The results also show that the impact of not including coherent uncertainties or not considering the covariances across the primary atmospheric state variables should not be overlooked. When an incomplete set of uncertainties are considered, it is very likely that the error growth rate—primarily at the smaller scales—will be underestimated. Compared with REAL and IDEAL which include initial uncertainties of *T*, Qv, *U*, and *V*, REAL-UV and IDEAL-UV that include *U* and *V* uncertainties show comparable error growth rates at the beginning, while REAL-Q and IDEAL-Q that include only Qv uncertainties grow notably slower at the smaller scales. This could be because perturbations in *U* and *V* can trigger deep convection more easily than Qv perturbations. *U* and *V* perturbations can also directly affect Qv through advection, while Qv perturbations affect *U* and *V* more indirectly through their impact on convective activities. Furthermore, in situations where the perturbed variable and the examined variable are inconsistent, e.g., where uncertainties in Qv are imposed but the evolution of forecast errors in kinetic energy are investigated, a more up-amplitude evolution of the error spectra could show up while the error spectra of the perturbed variable still show an upscale evolution, and these different behaviors in the error spectra of perturbed versus unperturbed variables could lead to incomplete conclusions. The broad wavenumber projection of localized features in the physical space is also reported by Durran et al. (2013) and Lloveras et al. (2022).

It should be pointed out that the sensitivities (or rather, insensitivities) of error growth characteristics to small-amplitude large-scale initial uncertainties in this study are tested in one case. The amplitudes and the structures of the initial uncertainties applied in this study are likely incapable of fundamentally changing the environmental conditions and hence the regions where convection initiations are more favored. As Nielsen and Schumacher (2016) and Weyn and Durran (2019) pointed out, the sensitivity of forecast error to large- and small-scale initial uncertainties and associated error growth mechanisms is determined by the environmental conditions, primarily synoptic-scale forcing. It is possible that a stronger sensitivity of forecast error to the structure of the small-amplitude large-scale initial uncertainties could emerge under a different situation of synoptic-scale forcing. For example, the large-scale errors are likely to amplify simultaneously with small-scale errors under a strong synoptic-scale forcing scenario with abundant baroclinic instability, and the impact of large-scale initial uncertainties' structures, as well as the relative contributions of upscale and up-amplitude error growth processes, could be different from what was observed in this study. The contributions of upscale error growth from smaller scales and up-amplitude error growth from larger scales can be approximated using "fake-dry" experiments that eliminate small-scale error amplification and upscale growth, similar to the comparisons performed in section 4 of this study. More systematic evaluations of events covering a broader range of synoptic-scale forcing are warranted. Last, specific characteristics of the error growth processes in this study, such as the underestimation of small-scale error growth speed when an incomplete set of initial uncertainties is used, and the different onset times of upscale error transfer for different atmospheric state variables, also need to be further investigated.

Acknowledgments. I would like to thank Y.-Qiang Sun (Rice University), Zhiyong Meng (Peking University), and Xingchao Chen and Da Fan (The Pennsylvania State University) for their helpful discussions, and the three anonymous reviewers for their constructive comments and discussions that substantially improved this manuscript. This study is supported by the National Science Foundation (NSF) Grant AGS-1712290. The numerical simulations were performed on the Stampede 2 supercomputer of the Texas Advanced Computing Center (TACC) through the Extreme Science and Engineering Discovery Environment (XSEDE) and the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) projects supported by the NSF.

Data availability statement. The simulations can be shared upon request.

REFERENCES

Baumgart, M., M. Reimer, V. Wirth, F. Teubler, and S. T. K. Lang, 2018: Potential vorticity dynamics of forecast errors: A quantitative case study. *Mon. Wea. Rev.*, **146**, 1405–1425, https://doi.org/10.1175/MWR-D-17-0196.1.

——, P. Ghinassi, V. Wirth, T. Selz, G. C. Craig, and M. Riemer, 2019: Quantitative view on the processes governing the upscale error growth up to the planetary scale using a stochastic convection scheme. *Mon. Wea. Rev.*, **147**, 1713–1731, https://doi.org/10.1175/MWR-D-18-0292.1.

Bierdel, L., T. Selz, and G. C. Craig, 2018: Theoretical aspects of upscale error growth on the mesoscales: Idealized numerical simulations. *Quart. J. Roy. Meteor. Soc.*, **144**, 682–694, https://doi.org/10.1002/qj.3236.

Durran, D. R., and J. B. Klemp, 1983: A compressible model for the simulation of moist mountain waves. *Mon. Wea. Rev.*, **111**, 2341–2361, https://doi.org/10.1175/1520-0493(1983)111<2341:ACMFTS>2.0.CO;2.

——, 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.

——, P. A. Reinecke, and J. D. Doyle, 2013: Large-scale errors and mesoscale predictability in Pacific Northwest snowstorms. *J. Atmos. Sci.*, **70**, 1470–1487, https://doi.org/10.1175/JAS-D-12-0202.1.

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461–2480, https://doi.org/10.1175/1520-0493(2001)129<2461:AOAPMS>2.0.CO;2.

Fan, D., S. J. Greybush, X. Chen, Y. Lu, F. Zhang, and G. S. Young, 2022: Exploring the role of deep moist convection in the wavenumber spectra of atmospheric kinetic energy and brightness temperature. *J. Atmos. Sci.*, **79**, 2721–2737, https://doi.org/10.1175/JAS-D-21-0285.1.

Hohenegger, C., and C. Schär, 2007a: Atmospheric predictability at synoptic versus cloud resolving scales. *Bull. Amer. Meteor. Soc.*, **88**, 1783–1793, https://doi.org/10.1175/BAMS-88-11-1783.

——, and ——, 2007b: Predictability and error growth dynamics in cloud-resolving models. *J. Atmos. Sci.*, **64**, 4467–4478, https://doi.org/10.1175/2007JAS2143.1.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

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.

Jiménez, P. A., J. Dudhia, J. F. González-Rouco, J. Navarro, J. P. Montávez, and E. García-Bustamante, 2012: A revised scheme for the WRF surface layer formulation. *Mon. Wea. Rev.*, **140**, 898–918, https://doi.org/10.1175/MWR-D-11-00056.1.

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.

——, 2020: Atmospheric predictability of the tropics, middle latitudes, and polar regions explored through global storm-resolving simulations. *J. Atmos. Sci.*, **77**, 257–276, https://doi.org/10.1175/JAS-D-19-0116.1.

Leung, T. Y., M. Leutbecher, S. Reich, and T. G. Shepherd, 2019: Atmospheric predictability: Revisiting the inherent finite-time barrier. *J. Atmos. Sci.*, **76**, 3883–3892, https://doi.org/10.1175/JAS-D-19-0057.1.

——, ——, ——, and ——, 2020: Impact of the mesoscale range on error growth and the limits of atmospheric predictability. *J. Atmos. Sci.*, **77**, 3769–3779, https://doi.org/10.1175/JAS-D-19-0346.1.

Lloveras, D. J., L. H. Tierney, and D. R. Durran, 2022: Mesoscale predictability in moist midlatitude cyclones is not sensitive to the slope of the background kinetic energy spectrum. *J. Atmos. Sci.*, **79**, 119–139, https://doi.org/10.1175/JAS-D-21-0147.1.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

——, 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21A**, 289–307, https://doi.org/10.3402/tellusa.v21i3.10086.

Melhauser, C., and F. Zhang, 2012: Practical and intrinsic predictability of severe and convective weather at the mesoscales. *J. Atmos. Sci.*, **69**, 3350–3371, https://doi.org/10.1175/JAS-D-11-0315.1.

Nastrom, G. D., 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.

Nielsen, E. R., and R. S. Schumacher, 2016: Using convection-allowing ensembles to understand the predictability of an extreme rainfall event. *Mon. Wea. Rev.*, **144**, 3651–3676, https://doi.org/10.1175/MWR-D-16-0083.1.

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.

——, ——, and F. Judt, 2023: Upscale versus "up-amplitude" growth of forecast-error spectra. *J. Atmos. Sci.*, **80**, 63–72, https://doi.org/10.1175/JAS-D-22-0070.1.

Selz, T., 2019: Estimating the intrinsic limit of predictability using a stochastic convection scheme. *J. Atmos. Sci.*, **76**, 757–765, https://doi.org/10.1175/JAS-D-17-0373.1.

——, 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.

——, L. Bierdel, and G. C. Craig, 2019: Estimation of the variability of mesoscale energy spectra with three years of COSMO-DE analyses. *J. Atmos. Sci.*, **76**, 627–637, https://doi.org/10.1175/JAS-D-18-0155.1.

——, M. Reimer, and G. C. Craig, 2022: The transition from practical to intrinsic predictability of midlatitude weather. *J. Atmos. Sci.*, **79**, 2013–2030, https://doi.org/10.1175/JAS-D-21-0271.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

——, and Coauthors, 2019: A description of the Advanced Research WRF Model version 4. NCAR Tech. Note NCAR/TN-556+STR, 145 pp., https://doi.org/10.5065/1dfh-6p97.

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.

——, and ——, 2020: A new theoretical framework for understanding multiscale atmospheric predictability. *J. Atmos. Sci.*, **77**, 2297–2309, https://doi.org/10.1175/JAS-D-19-0271.1.

——, R. Rotunno, and F. Zhang, 2017: Contributions of moist convection and internal gravity waves to building the atmospheric −5/3 kinetic energy spectra. *J. Atmos. Sci.*, **74**, 185–201, https://doi.org/10.1175/JAS-D-16-0097.1.

Thompson, G., and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.*, **71**, 3636–3658, https://doi.org/10.1175/JAS-D-13-0305.1.

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.

——, and ——, 2019: The scale dependence of initial-condition sensitivities in simulations of convective systems over the southeastern United States. *Quart. J. Roy. Meteor. Soc.*, **145**, 57–74, https://doi.org/10.1002/qj.3367.

Zhang, C., Y. Wang, and K. Hamilton, 2011: Improved representation of boundary layer clouds over the southeast Pacific in ARW-WRF using a modified Tiedtke cumulus parameterization scheme. *Mon. Wea. Rev.*, **139**, 3489–3515, https://doi.org/10.1175/MWR-D-10-05091.1.

Zhang, F., 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.

Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2016: Intrinsic predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma at storm scales. *Mon. Wea. Rev.*, **144**, 1273–1298, https://doi.org/10.1175/MWR-D-15-0105.1.

——, H. Yu, M. Zhang, Y. Yang, and Z. Meng, 2022: Uncertainties and error growth in forecasting the record-breaking rainfall in Zhengzhou, Henan on 19–20 July 2021. *Sci. China Earth Sci.*, **65**, 1903–1920, https://doi.org/10.1007/s11430-022-9991-4.