Potentials in Improving Predictability of Multiscale Tropical Weather Systems Evaluated through Ensemble Assimilation of Simulated Satellite-Based Observations

YUE YING AND FUQING ZHANG

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

(Manuscript received 19 August 2017, in final form 31 January 2018)

ABSTRACT

As a follow-up of our recent paper on the practical and intrinsic predictability of multiscale tropical weather and equatorial waves, this study explores the potentials in improving the analysis and prediction of these weather systems through assimilating simulated satellite-based observations with a regional ensemble Kalman filter (EnKF). The observing networks investigated include the retrieved temperature and humidity profiles from the Advanced TIROS Operational Vertical Sounder (ATOVS) and global positioning system radio occultation (GPSRO), the atmospheric motion vectors (AMVs), infrared brightness temperature from *Meteosat-7* (*Met7*-Tb), and retrieved surface wind speed from the Cyclone Global Navigation Satellite System (CYGNSS). It is found that assimilating simulated ATOVS thermodynamic profiles and AMV winds improves the accuracy of wind, temperature, humidity, and hydrometeors for scales larger than 200 km. The skillful forecast lead times can be extended by as much as 4 days for scales larger than 1000 km. Assimilation of *Met7*-Tb further improves the analysis of cloud hydrometeors even at scales smaller than 200 km. Assimilating CYGNSS surface winds further improves the low-level wind and temperature. Meanwhile, the impact from assimilating the current-generation GPSRO data with better vertical resolution and accuracy is comparable to assimilating half of the current ATOVS profiles, while a hypothetical 25-times increase in the number of GPSRO profiles can potentially exceed the impact from assimilating the current network of retrieved ATOVS profiles. Our study not only shows great promises in further improving practical predictability of multiscale equatorial systems but also provides guidance in the evaluation and design of current and future spaceborne observations for tropical weather.

1. Introduction

The tropical weather systems are multiscale in nature. The Madden–Julian oscillation (MJO; Madden and Julian 1971, 1972) dominates the planetary-scale intraseasonal variability. The synoptic and mesoscales feature a variety of waves that interact with moist convection, also known as convectively coupled equatorial waves (CCEWs; Wheeler and Kiladis 1999; Kiladis et al. 2009), which include the equatorial Rossby (ER) waves (Kiladis et al. 2009), Kelvin waves (Dunkerton and Crum 1995), mixed-Rossby–gravity (MRG) waves (Dickinson and Molinari 2002), and inertia–gravity (IG) waves (Haertel and Kiladis 2004). Skillful modeling of these multiscale tropical weather systems is an important and challenging task for global numerical weather prediction. Lorenz (1963) first discovered that the atmospheric flow can be intrinsically unpredictable because of its chaotic nature even with near-perfect model and initial/boundary condition. Moist convective processes are identified as responsible for the intrinsic limit of predictability for moist baroclinic waves (Zhang et al. 2003, 2007; Sun and Zhang 2016), mesoscale convective systems (Bei and Zhang 2007; Melhauser and Zhang 2012; Selz and Craig 2015; Y. Zhang et al. 2016), and tropical cyclones (Zhang and Sippel 2009; Tao and Zhang 2015). CCEWs are also susceptible to the moist convective processes that will limit their predictability. Despite such intrinsic limits in predictability, there is still room for improvement in the current prediction skill by assimilating better observations and improving the models. For example, Reynolds et al. (1994) found that the then-operational global prediction system had deficiencies in the tropics that lead to large error growth, although the internal error growth rate is actually slower in the tropics than in midlatitudes.

Corresponding author: Professor Fuqing Zhang, fzhang@psu.edu

DOI: 10.1175/JAS-D-17-0245.1

To model the multiscale tropical weather systems, a high spatial resolution is required to resolve finer-scale wave components and their interaction with larger-scale systems. Recent developments in cloud-permitting model simulation of the tropical weather systems, such as Wang et al. (2015), who conducted a convective-permitting simulation of the October 2011 MJO event observed during the Dynamics of the MJO (DYNAMO; Zhang et al. 2013) field campaign, start to show better agreement with observation than previous coarse-resolution global models. Zhang et al. (2017) demonstrated that initiation of this MJO event may result from a global circumnavigating signal and thus can be potentially predictable at the planetary scale several weeks in advance. Most recently, Ying and Zhang (2017; hereafter referred to as YZ17) systematically assessed the intrinsic and practical predictability of multiscale tropical weather systems within the moist phase of the same MJO event. They estimated that the large-scale ER, Kelvin, MRG, and IG waves have a practical predictability as much as 8 days, while the small-scale IG waves have a much more limited predictability that is less than l day. The intrinsic predictability of large-scale CCEW modes may be achievable beyond 2 weeks, but the limit is likely still less than 3 days for small scales.

Because of the lack of in situ soundings with high resolution and good coverage, the tropics heavily rely on satellite remote sensing to provide observations for forecast systems. As a follow-up study of YZ17, we seek to systematically evaluate the impact of assimilating current/future satellite observing networks to the predictability of each CCEW mode and at different scales. To the best of our knowledge, this study is the first satellite data assimilation experiment focusing on multiscale tropical weather systems beyond tropical cyclones. The MJO active phase simulation from YZ17 is used as a test case, and an idealized observing system simulation experiment (OSSE) is conducted under the assumption of perfect dynamic model and perfect observation error models used in data assimilation. For observing networks with different spatial density, several studies suggest that a shorter localization distance is required for denser observations to achieve better analyses (Dong et al. 2011; Kirchgessner et al. 2014; Periáñez et al. 2014). In this study, the localization for each observing network is manually tuned to allow a relatively fair comparison of their impact.

This paper is organized as follows. Section 2 describes the dynamic model, test case, and data assimilation method. Section 3 provides details of the satellite observing networks under investigation. The sections that follow will discuss results regarding scale- and variable-dependent predictability (section 4), the relative contribution from each observing network (section 5), and the specific improvement found for each CCEW (section 6). Section 7 will summarize the findings from this study.

2. Experimental design

a. The verifying truth simulation

A similar configuration of the Weather Research and Forecasting (WRF) Model is adapted from YZ17 to conduct a monthlong simulation from 12 October to 12 November 2011. The model domain contains 777 × 444 horizontal grid points with 9-km spacing and has 45 vertical levels with 9 in the lowest 1 km and a model top at 20 hPa. Cloud microphysics processes are represented by the WRF double-moment 6-class (WDM6) scheme (Lim and Hong 2010). No cumulus parameterization is used, and organized convections are explicitly represented by the 9-km model grid. Shortwave and longwave radiative processes are simulated using the new Goddard (Chou and Suarez 1999) and RRTM for GCMs (RRTMG; Iacono et al. 2008) schemes, respectively. Surface processes are simulated using the unified Noah (Chen and Dudhia 2001) with variable surface skin temperature (Zeng and Beljaars 2005). Subgrid-scale turbulent mixing is treated with the Yonsei University (YSU) boundary layer scheme (Hong et al. 2006). The initial conditions (IC) and lateral boundary conditions (LBC) are specified by the ERA-Interim data (Dee et al. 2011). The sea surface temperature (SST) for the lower boundary conditions is updated every 6 h according to the National Centers for Environmental Prediction (NCEP) Final Operational Global Analysis (FNL) data.

Figure 1 shows the computation domain that covers the precipitation associated with the active MJO phase. The simulation successfully captures the slow eastward propagation of the active phase as well as most of the CCEW modes. Figure 2 shows the Wheeler and Kiladis (1999) space–time spectra for surface precipitation and 850-hPa zonal wind. For precipitation, the dominant signals are the MJO, Kelvin, and westward-propagating IG (WIG) waves. The WIG waves contain a hierarchy of spatial and temporal scales. For spatial scales larger than 2000 km, there is the "2-day wave," while the diurnal and semidiurnal waves have wavelengths slightly shorter than 2000 km. The Kelvin wave has a precipitation signal at a shorter wavelength than the wind signal. It is also evident that the propagation speed of the convectively coupled Kelvin wave is slower than the dry Kelvin waves. For ER and MRG waves, the precipitation signal is relatively weaker than their wind signal. The precipitation signal for the ER wave also has a shorter wavelength than the wind signal, similar to the Kelvin waves. The

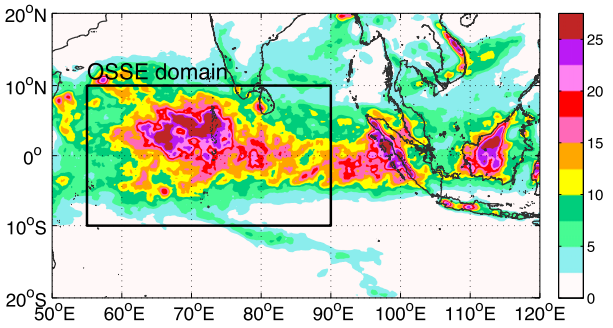

FIG. 1. Horizontal map of 15-day averaged daily precipitation (mm day^{-1}) shown in the computational domain with ERA-Interim as initial and boundary condition. The OSSE domain is shown in black. The color scale has a value of 25 at the top of the purple band and a value of 0 at the bottom, and has an interval of 5 mm day^{-1}.

signals related to eastward-propagating IG waves are not as strong as their westward-propagating counterparts.

Because of the limitation of computation resources, the OSSE is conducted in a smaller domain with 378 × 222 grid points (10°S–10°N, 55°–90°E) during the 17-day period from 16 October to 2 November (phases 1–3 of the MJO). The location of this OSSE domain is shown in Fig. 1. The monthlong large-domain simulation provides IC and LBC for the smaller OSSE domain. The OSSE domain has the same 45 vertical levels and 9-km grid spacing as the larger domain.

To extract a certain CCEW mode, a space–time bandpass filter is applied to the monthlong large-domain model simulation. Table 1 lists the filter parameters for each CCEW. Figure 3 shows the filtered surface precipitation and 850-hPa zonal wind during the 17-day test period and within the OSSE domain. Hereafter, the time from 16 October to 2 November will be labeled as days 1 to 18. During the test period, the MJO precipitation signal slowly propagates across the entire OSSE domain from west to east. The low-level westerly wind anomaly is trailing the precipitation peak. The ER wave features two strong westerly wind signals associated with the cyclonic Rossby gyres within the test domain and replaced by easterlies (anticyclonic gyre) later into the test period. The Kelvin wave has three successive phases propagating across the test domain during the 17 days; each one progresses more to the east because of modulation from the MJO. MRG waves have a strong wind signal but very weak precipitation signal during the test period. WIG waves consist of a wide range of spatial and temporal scales. Large-scale (>1000 km) WIG signals have a life cycle of about 2 days and propagate across the whole test domain, while intermediate-scale (500–2000 km) WIG signals are more transient in space and time. Both types of WIG waves appear to be modulated by the MJO envelope as well.

b. *Data assimilation method: EnKF*

The Pennsylvania State University (PSU) WRF EnKF (Meng and Zhang 2008) is employed as the data assimilation method. The EnKF utilizes an ensemble of model realizations to estimate flow-dependent error covariance, which is an approximation of the extended Kalman filter (Evensen 1994). The ensemble perturbations

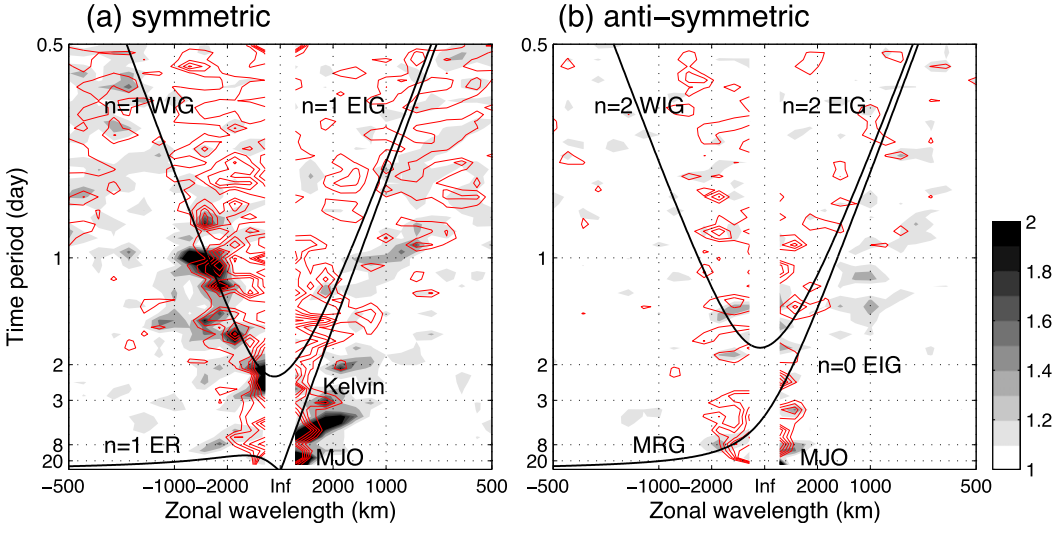

FIG. 2. Wheeler–Kiladis space–time spectra for the truth simulation: (a) symmetric and (b) asymmetric component about the equator. Signal strength from 1.1 to 2 is shown in gray shadings for precipitation and red contours for 850-hPa zonal wind. Dispersion relations for dry waves are shown with equivalent depth of 25 m. Horizontal (vertical) axis is zonal wavenumber (time frequency) but labeled with zonal wavelength (time period).

TABLE 1. Space–time filtering parameters for each tropical weather system (CCEW type).

CCEW type	Spatial scale l (km)	Time scale τ (days)	Propagation direction	Symmetry about the equator
MJO	>2000	>20	Eastward	Both
ER	>1000	8–20	Westward	Symmetric
Kelvin	>1000	3–10	Eastward	Symmetric
MRG	>1000	4–10	Westward	Antisymmetric
WIG (L)	>1000	1–3	Westward	Symmetric
WIG (M)	500–2000	1/4–1	Westward	Both

around the mean are updated with the square root algorithm from Whitaker and Hamill (2002). Covariance relaxation to prior perturbation (Zhang et al. 2004) with $\alpha = 0.8$ inflates the ensemble variance to prevent filter divergence. Covariance localization is applied with the Gaspari and Cohn (1999) tapering function to remedy sampling noises in error covariance.

The OSSE is conducted under a perfect-model assumption; that is, the forecast model and LBC are the same from the truth simulation. An ensemble of 60 members is created one day before the data assimilation cycle begins. The initial ensemble perturbations are randomly sampled from the climatological error covariance matrix created by the National Meteorological

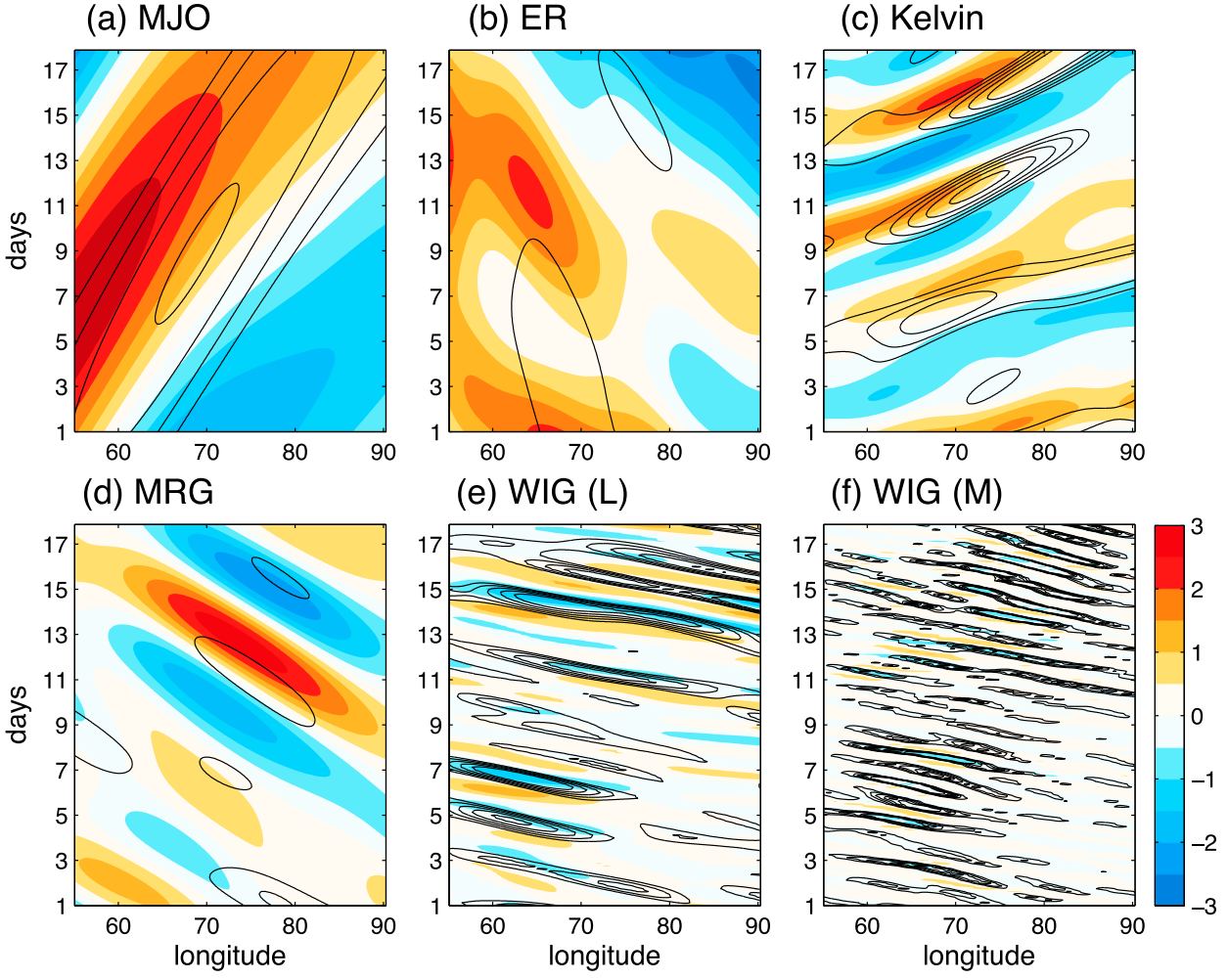

FIG. 3. Longitude–time Hovmöller diagrams of precipitation (black contours from 2 to 10 mm day^{-1} every 2 mm day^{-1}) and 850-hPa zonal wind (color shadings; m s^{-1}) averaged over 0°–5°N from the truth simulation and filtered for each CCEW: (a) MJO, (b) ER, (c) Kelvin, (d) MRG, (e) large-scale WIG, and (f) intermediate-scale WIG waves.

TABLE 2. Observed variables, their associated uncertainty (observation error), spatial resolution, and approximate data count within a 3-h window of each tested observing network.

Observing network	Observed variable	Observation error	Horizontal spacing	Vertical spacing–coverage	Observation count in 3-h window
ATOVS retrievals	T	~2 K	90 km	1 km; surface to 0.1 hPa	~35 000
	Q	~30%		2 km; surface to 200 hPa	~11 000
GPSRO retrievals (COSMIC-2)	T	~1.7 K	Irregular (~600 km)	~100 m in low troposphere	~1660
	Q	~10%		1 km in stratosphere	~1350
AMV wind	u, v	4.5 m s^{-1}	Irregular (~100 km)	Irregular spacing mostly covers ~200 hPa	~1700
Met7 IR Tb (channel 3)	T_b	3 K	9 km	Column averaged with weighting function	~80 000
CYGNSS retrievals	Wind speed	2 m s^{-1}	45 km	At ocean surface	~3200

Center (NMC) method (Parrish and Derber 1992). The ensemble mean is created by adding another perturbation to the truth. This perturbation is drawn from the same error covariance, but we ensured that its magnitude is close to the ensemble spread. The initial ensemble is run forward for one day to develop flow-dependent error covariance.

Synthetic observations are created by randomly perturbing the observed variable simulated from the truth with prescribed observation error. The perfect forward operator is used, and the correct observation errors matching the uncertainties in synthetic observations are specified in EnKF. The prescribed observation errors are unbiased and uncorrelated in this OSSE study. Table 2 lists the prescribed observation error and resolution for each observing network, which will be discussed in more details in the next section. Table 3 provides a succinct list of experiments conducted in this paper. No data assimilation (NoDA) is a free ensemble run from the perturbed initial ensemble. Control (CNTL) is a benchmark case assimilating Advanced TIROS Operational Vertical Sounder (ATOVS) temperature (T) and specific humidity (Q) profiles and atmospheric motion vector (AMV) winds. The observations are assimilated every 3 h during the 17-day test period, which results in 136 data assimilation cycles in total. We will show that the EnKF spinup period appears to be shorter than 4 days; therefore, some of the sensitivity experiments are only conducted for the first 9 days. The localization cutoff distance [i.e., the radius of influence (ROI)] is specified separately for each observing network in consideration of the respective observation density. A larger ROI is used for sparser networks, and some manual tuning with a few cycles are performed to reach a reasonable ROI. A sensitivity experiment that doubles the horizontal ROI in CNTL confirmed that the resulting change in analysis error is minor. As pointed out by Zhen and Zhang (2014) and Lei and Whitaker (2017), the analysis sensitivity to localization is minor when approaching the best-performing ROI. Thus, we will not further discuss how the exact choice of ROI changes analysis accuracy in this paper.

3. Observing networks

a. *Synthetic ATOVS retrieved temperature and humidity profiles*

The ATOVS is a collection of polar-orbiting satellite instruments that are designed to sample the atmospheric T and Q at different vertical levels. These instruments currently include the High Resolution Infrared Radiation Sounder (HIRS), the Advanced Microwave Sounding Unit-A (AMSU-A) and Microwave Humidity Sounder (MHS) on board *NOAA-19*, *MetOp-A*, and *MetOp-B* satellites. Operational centers routinely produce atmospheric sounding products from the ATOVS instruments using sophisticated retrieval methods (Reale 2001), typically including cloud detection, radiative transfer, and other physical process–based algorithms. The T and Q profiles can be retrieved for both clear and cloudy atmospheric conditions. Although direct assimilation of ATOVS radiances may yield more realistic results, current regional-scale data assimilation systems (including the one employed here) cannot effectively and simultaneously assimilate all the ATOVS radiances, especially for cloud-affected microwave-channel radiances that are extremely sensitive to parameterizations in model microphysical processes and related representations of scattering properties. Therefore, we assimilate synthetic retrieved T and Q profiles with regular spatiotemporal resolutions to evaluate the impact from these ATOVS radiances (similar to English et al. 2000), which also serves a baseline for other observing systems.

TABLE 3. Description of assimilated observing networks and their benefits (improvement in prediction skills in variables, scales, and weather systems) found in each experiment.

Experiment name	Assimilated observing networks	Benefits in analysis/prediction skill
NoDA	None	Rely on LBC to constrain solution in domain Large-scale CCEWs (MJO, Kelvin, MRG, and ER) have some predictability in wave phase but with large amplitude errors
CNTL	ATOVS profiles + AMV wind	Significantly improve winds, T, Q, and hydrometeors at large to intermediate scales Extend the practical predictability limit by ~4 days for large scales Further reduce the amplitude error for large-scale CCEW and improve intermediate-scale WIG waves
+*Met7*	CNTL + *Met7*-Tb	Further improve ice, snow, and graupel mixing ratio near cloud top for smaller scales
+CYGNSS	CNTL + CYGNSS wind speed	Further improve large-scale low-level u, v, and T
NoProfile	AMV wind + *Met7*-Tb + CYGNSS wind speed	AMV + *Met7* + CYGNSS can still improve u, v, T, and Q although less effectively than ATOVS profiles
+ATOVS	NoProfile (AMV + *Met7* + CYGNSS) + ATOVS profiles	ATOVS profiles are essential in improving u, v, T, and Q thanks to their fine spatial resolution and coverage
+ATOVScoarse	NoProfile + ATOVS profiles reduced to 6-hourly and 180-km grid spacing	Although worse than the full-resolution profiles, there are still persistent improvements at large to intermediate scales Reduced resolution degrades smaller-scale WIG
+ATOVSthinV	NoProfile + ATOVS profiles thinned in the vertical	Reduced information content in the vertical does not degrade analysis accuracy as much as reduction in horizontal and temporal resolution
+ATOVSerror1.5	NoProfile + ATOVS profiles with errors inflated to 150%	
+GPSRO	NoProfile + GPSRO profiles at COSMIC-2 resolution	Approaching but less than improvement from low-resolution ATOVS because of sporadic sampling
+GPSROdense	NoProfile + GPSRO profiles at CubeSat resolution (25 times more data than COSMIC-2)	More improvement than ATOVS thanks to finer vertical resolution and more accurate measurements
+GPSRON +GPSRODenseN	Same as +GPSRO and +GPSROdense but assimilating refractivity (N) profiles	Nonlinear observation operator requires shorter localization distance Similar sensitivity to observation density

The spatial and temporal resolution of ATOVS profiles depend on both the availability of satellite instruments and the level of quality control during retrieval. The current profiles have a horizontal spacing of about 40 km. In the vertical, there are 40 layers from surface up to 0.1 hPa for T profiles and 17 layers from surface up to 200 hPa for Q profiles. Miyoshi et al. (2013) demonstrated how nonorthogonal observation operators reduce the available information from observations. Because of the overlapping of weighting functions of radiances used in the retrieval process, the number of independent pieces of information from real ATOVS profiles is likely to be less that what is assumed in this study, the impact of which will be examined with sensitivity experiments through reducing the information content in either the horizontal and/or vertical directions and in time. The horizontal spacing is set to a coarser 90 km to account for missing data due to gaps between swaths and/or quality control processes. The current three-satellite constellation could sample the whole tropics 6 times every day. This observation interval can be shortened with an increase in the number of participating satellites in the constellation. Since 2015, with the launch of the Global Precipitation Measurement (GPM) mission (Hou et al. 2014), it has been estimated that there are enough satellites to achieve an average revisiting time of 1.5 h. Therefore, a 3-h cycling period is reasonable for ATOVS in this study. Figure 4a shows the vertical profile of prescribed ATOVS observation errors in solid lines. Li et al. (2000) validated the ATOVS retrievals with respect to radiosonde observations and found the accuracy of retrieved T is about 2 K and dewpoint temperature is 3–6 K (about 30% error in Q). The horizontal distribution of ATOVS profiles is shown as black dots in Fig. 4b. The horizontal localization radius (ROI) is set to 400 km. The vertical ROI is

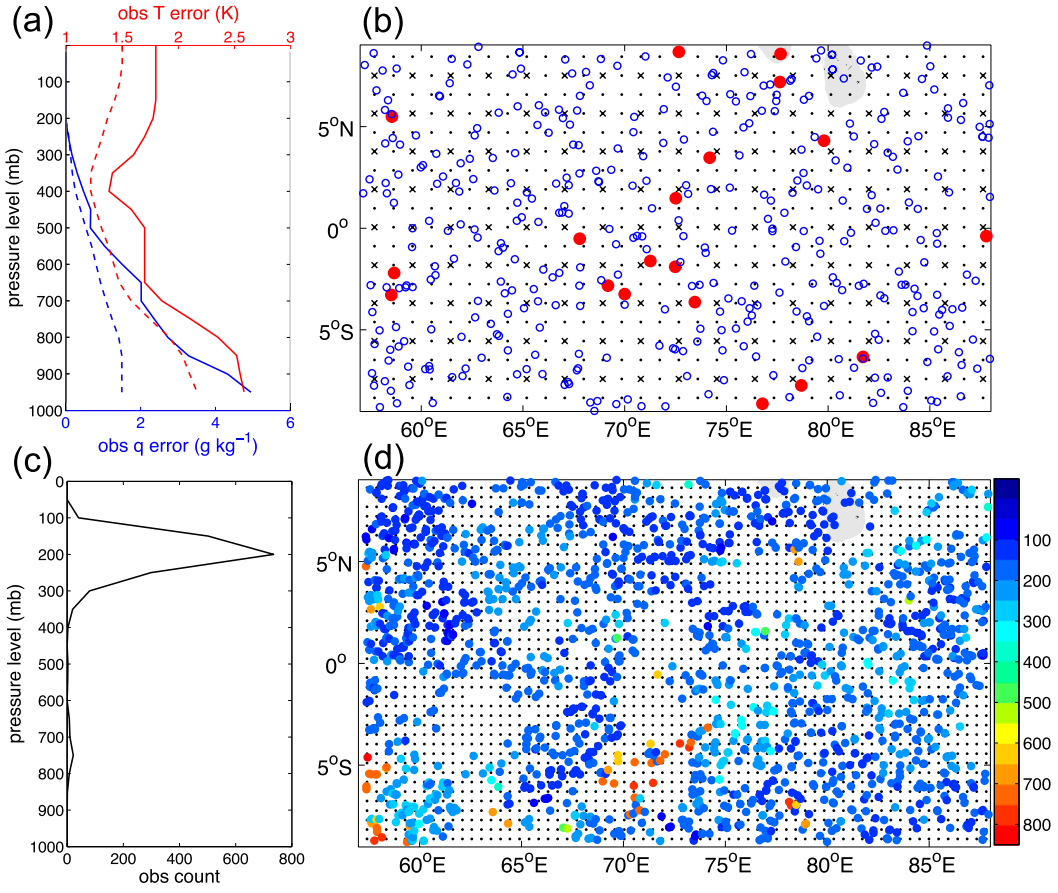

FIG. 4. (a) Vertical observation error profiles from ATOVS (solid) and GPSRO (dashed) for T (red) and Q (blue); (b) horizontal map of observation location during a 3-h window for ATOVS (solid black dots), ATOVS with reduced resolution (black cross), GPSRO at COSMIC-2 resolution (red dots), and GPSRO at CubeSat resolution (blue circles); (c) AMV observation count (per 50-hPa vertical layer) during a 3-h window; (d) horizontal map of AMV observation location color coded with observation pressure height (hPa); black dots show location of CYGNSS wind observations.

set to 5 levels, which correspond to ~500 m near the surface and increases to ~3000 m near the tropopause.

b. *Synthetic GPSRO observations*

The global positioning system radio occultation (GPSRO) samples the atmosphere with limb radio signals transmitted by the GPS satellites and received by low-Earth-orbiting (LEO) satellites. Atmospheric T and Q can be retrieved from the refraction of GPS signals. Since radio signals are not influenced by particulates, the retrieval can be performed for both clear and cloudy weather conditions. The limb sounding geometry provides a higher vertical resolution than the passive nadir sounders (Kursinski et al. 1997; Yunck et al. 2009). The achievable vertical spacing is ~100 m for the low troposphere and reduces to 1 km into the stratosphere for both temperature and humidity. The horizontal density of retrieved profiles depends on the number of LEO satellites in operation. For the second Constellation Observation System for Meteorology, Ionosphere and Climate (COSMIC-2) mission (Cook et al. 2011), 12 LEO satellites will be launched that provide more than 8000 profiles globally within a 3-h window. The retrieval products also have better accuracy than ATOVS. Wang et al. (2013) validated the GPSRO profiles with global radiosondes and found that T error is 1.72 K and Q error is 0.67 g kg^{-1}. Figure 4a shows the GPSRO error profiles as dashed lines in comparison to ATOVS error profiles.

In this study, synthetic GPSRO profile location is randomly specified. At COSMIC-2 density, there are about 20 profiles in the OSSE domain. A much higher horizontal density may be achieved with the future launch of the CubeSat constellation (U-class spacecraft; Mannucci et al. 2010). Assuming a 300-satellite CubeSat constellation, there will be 25 times more profiles in the domain. Figure 4b compares the horizontal distribution of

GPSRO at COSMIC-2 and CubeSat density to the ATOVS profile grid. The horizontal ROI is set to 1600 km for GPSRO profiles at COSMIC-2 density, while a shorter 400 km is used for denser CubeSat GPSRO profiles. Vertical ROI is set to five levels, the same as for the ATOVS profiles.

To compare the assimilation of retrieved T and Q profiles with the direct assimilation of refractivity (N) data, synthetic N profiles are generated at the same location of the T and Q profiles according to the following equation (Smith and Weintraub 1953):

$$N = 77.6\frac{p}{T} + 3.73 \times 10^5 \frac{e}{T^2}, \quad (1)$$

where p is pressure, T is temperature, and e is water vapor pressure. According to Kursinski et al. (1997), the observation errors for N is ~0.2% above 500 hPa and increases to ~1% at lower levels. Since the observation operator is nonlinear, a shorter localization distance is specified. The horizontal ROI is 600 km for coarser network and 200 km for denser network. Vertical ROI is set to 3 levels.

c. *Synthetic AMV wind*

The AMV is derived from a sequence of geostationary satellite images by tracking features such as cloud edges and water vapor gradients (Nieman et al. 1997). A cross-correlation tracking algorithm locates sharp gradients in raw satellite images, and their displacement vectors are retrieved as wind observations. Heights are assigned to each AMV according to its observed brightness temperature. Operational centers now produce routine AMV wind products from the Geostationary Operational Environmental Satellite (GOES) images. NCEP GFS assimilates AMVs in real time every 6 h with relatively strict quality control processes. CIMSS prepares hourly AMV products with more detailed coverage over tropical regions (Velden et al. 1997). A "rapid scan" mode can be switched on to achieve even higher temporal frequency (every 15 min). Velden and Bedka (2009) performed a careful comparison of large volumes of AMV data with collocated rawinsonde wind profiles and estimated that AMV observation errors are about 5 m s^{-1} and the height assignment is the dominant error source (up to 70%).

To generate synthetic AMV observations, instead of simulating the cloud detection and tracking algorithms, the real AMV observation locations from the Global Telecommunication System (GTS) dataset are directly used to interpolate wind from the truth simulation. Such treatment does not guarantee the synthetic AMVs to be located at the simulated clouds, especially for cloud features at smaller scales. The simulated AMV observations at real-data observed locations may sometimes be located in the clear-air regions of the simulation experiments. At smaller scales, such biased coverage may potentially introduce flow dependence in analysis performance; thus, the results obtained from the OSSE may vary for real-case scenarios. However, for larger scales, the results should be less influenced since simulated cloud clusters agree well with observations. More important is that such a procedure yields synthetic AMVs that match the spatial resolution and distribution of the current real-world data. Figure 4c shows the number of AMVs at each vertical layer. Most of the AMVs are located at upper levels (~200 hPa), and only a small fraction is at low levels (~750 hPa). At upper levels, the number of AMVs is comparable to ATOVS profiles (~800). Figure 4d shows the horizontal distribution of AMVs. The horizontal ROI is set to 400 km. The vertical coverage is limited to where the cloud or vapor gradient is located; therefore, a larger vertical ROI of 15 model levels is used.

d. *Synthetic* Met7 *infrared brightness temperature*

Among the first-generation geostationary satellites, *Meteorological Satellite 7* (*Meteosat-7*, hereafter *Met7*) has coverage to the Indian Ocean. It provides full-disk brightness temperature (Tb) every 30 min from two infrared channels and a visible channel. The horizontal observation spacing is 5 km. *Met7* was replaced by the second-generation *Meteorological Satellite 8* (*Meteosat-8*) in 2017, which has more channels and increased resolution (similar to *GOES-16*, which achieves a spatial grid spacing of 2 km every 15 min). Although clear-air satellite Tb is routinely assimilated in operational centers, the direct assimilation of cloud-affected Tb is still quite challenging. Recent development in methods that remedy the non-Gaussian observation errors makes the assimilation of all-sky Tb much more effective (Geer and Bauer 2011; Tavolato and Isaksen 2015; F. Zhang et al. 2016; Minamide and Zhang 2017). The Tb from infrared channels are easier to assimilate than those from microwave channels thanks to less sensitivity to scattering processes that are not well represented in current microphysics schemes and radiative transfer models. In this study, the water vapor band Tb (channel 3) is assimilated directly using the Community Radiative Transfer Model (CRTM) as forward operator. The thermal infrared window (channel 2) is preserved for verification purposes. The adaptive observation error inflation (AOEI; Minamide and Zhang 2017) is applied to reduce the negative effect from clear/cloudy-air representativeness errors. Horizontal ROI is set to 30 km. The vertical height of a Tb observation is specified according to its cloudy/clear-air condition. A cloudy Tb is assigned to 250 hPa and a clear-air Tb to 400 hPa (peaks of their

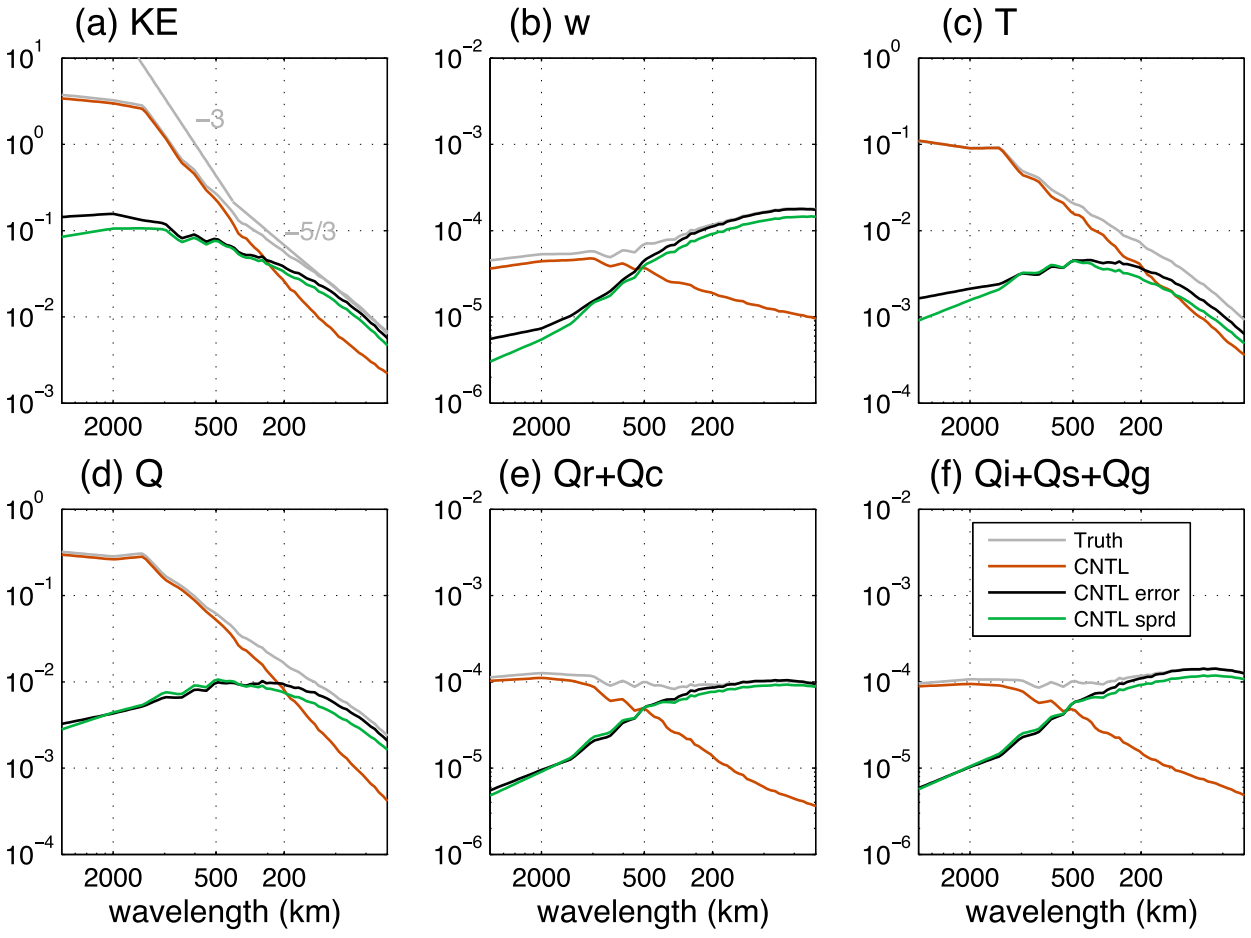

FIG. 5. Temporally and vertically averaged spectra for (a) kinetic energy ($m^2 s^{-2}$), (b) vertical motion ($m^2 s^{-2}$), (c) temperature (K^2), (d) specific humidity ($g^2 kg^{-2}$), (e) rain + cloud water mixing ratio ($g^2 kg^{-2}$), and (f) ice + snow + graupel mixing ratios ($g^2 kg^{-2}$). Gray lines show spectra of the truth signal; red lines the spectra of the CNTL prior mean; black lines show spectra of CNTL error (difference between CNTL prior mean and the truth); and green lines show spectra of CNTL ensemble variance.

weighting function). The vertical ROI is set to 25 levels around the Tb observation at its specified height so that the localization function is similar to its weighting function.

e. Synthetic CYGNSS surface wind speed

The Cyclone Global Navigation Satellite System (CYGNSS) is a constellation of eight microsatellites that receive both direct and reflected signals from GPS satellites (Ruf et al. 2016). The direct signals pinpoint the location of a CYGNSS observation, and the reflected signals carriy information of the roughness of ocean surface, from which wind speed is retrieved. The relatively low-inclination orbits of the microsatellites are designed to provide excellent coverage for the tropics. The observation footprint is 25 km wide with a mean and median revisit time of 7.2 and 2.8 h, respectively. The observation error is about $2 m s^{-1}$ or 10% for a wide dynamic range of wind speed. The synthetic CYGNSS observation is generated at 45-km grid spacing every 3 h. Instead of locating the observations along the realistic orbits that are horizontally inhomogeneous, a slightly coarser uniform grid is used for simplicity. The CYGNSS observation grid is shown as black dots in Fig. 4d. The horizontal ROI is set to 200 km for the CYGNSS observation, and their vertical ROI is 15 model levels from surface up to ~700 hPa.

4. Scale- and variable-dependent improvements in prediction skill

In this section, the benchmark CNTL experiment is compared to NoDA to evaluate the impact from assimilating synthetic ATOVS and AMV observations. First, to assess the EnKF performance in CNTL, the two-dimensional spectra of error in prior ensemble mean are compared to the ensemble spread in Fig. 5. For

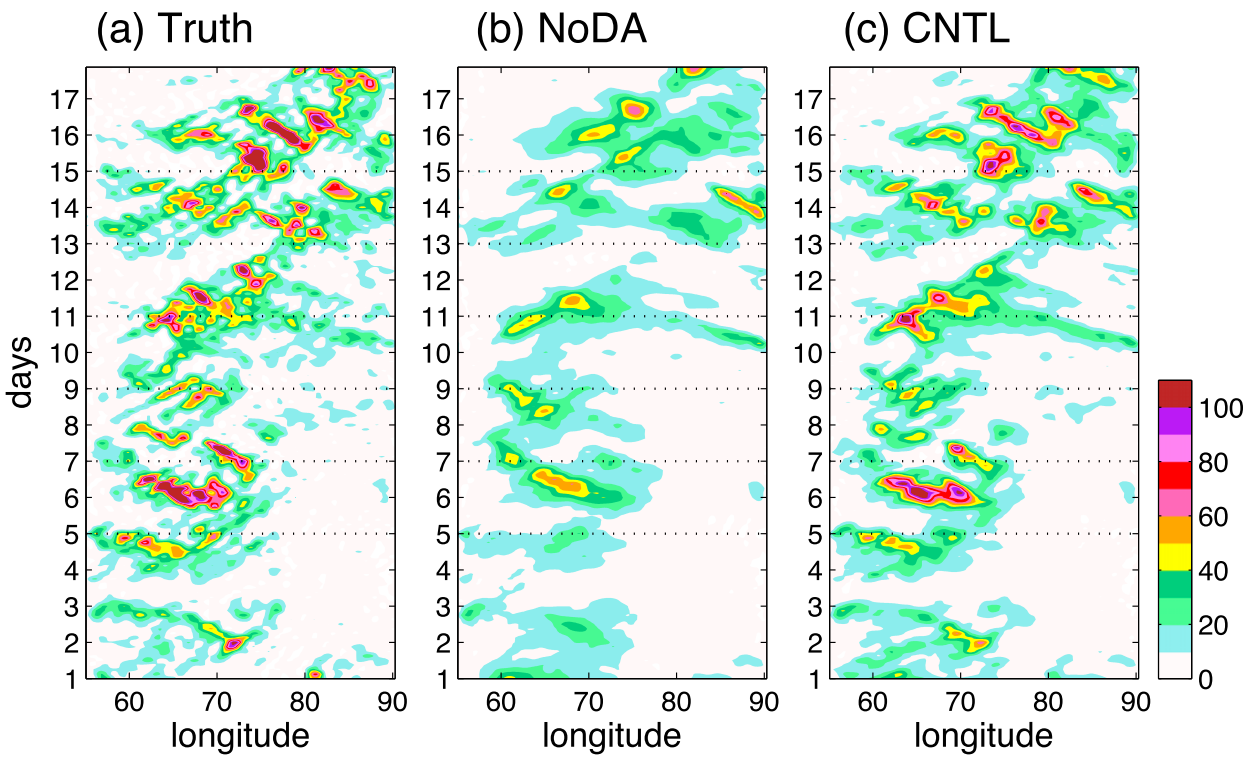

FIG. 6. Longitude–time Hovmöller diagrams of precipitation (mm day^{-1}) averaged over 0°–5°N latitudes from (a) truth simulation, (b) NoDA ensemble mean, and (c) CNTL posterior ensemble mean. The precipitation fields are filtered for $l > 200$ km.

most variables and scales, the ensemble spread well represents the prior error, and no filter divergence occurred. Figure 5 also shows the spectra from the CNTL prior ensemble mean and truth signals as references. The two reference spectra match at the large scales; however, the ensemble-mean spectral power are lower than the truth at smaller scales. This drop in spectral power is due to the smoothing of small-scale features that are dislocated among the members, which to a certain extent reflects the more limited predictability for small scales, where error grows and displacing these features at a higher rate than the data assimilation cycles can constrain. YZ17 characterized the predictability limits by comparing the ensemble spread spectra to the ensemble mean reference spectra. Going from large to small scales, at a given time, the ensemble spread reaches and exceeds the signal level from the reference (signal-to-noise ratio smaller than 1), indicating loss of predictability. Notice that hydrometeors and vertical motion (w) reach predictability limits at larger scales than temperature (T), specific humidity (Q), and winds. At even smaller scales, when signal-to-noise ratio is dropped significantly below 1, the error spectra match the truth signal (error saturation). Scales smaller than 200 km appear to have saturated hydrometeor and w errors and wind, T, and Q errors approaching their predictability limits at 3 h in the prior ensemble. Therefore, the following diagnostics will filter out these smaller scales for a clearer demonstration of observation impact. For the sake of simplicity, "NoDA" will refer to NoDA ensemble mean and "CNTL" to CNTL posterior ensemble mean.

Figure 6 compares the longitude–time plots of surface precipitation from NoDA and CNTL to the truth. The NoDA mean precipitation signal is much weaker than the truth, although it still captures the precipitation events around days 6, 11, and 14–17. On the other hand, the CNTL analysis precipitation better agrees with the truth than NoDA. The amplitude, location, timing, and propagation of the precipitation signals are much improved by assimilating the observations. Snapshots of low-level horizontal wind and relative humidity fields from NoDA and CNTL just prior to the precipitation event at day 6 are compared to truth in Fig. 7. The truth low-level flow forms a convergence line and an anticyclonic shear around the precipitation, and a cyclonic gyre is located northeast to the precipitation. Although NoDA has the perfect

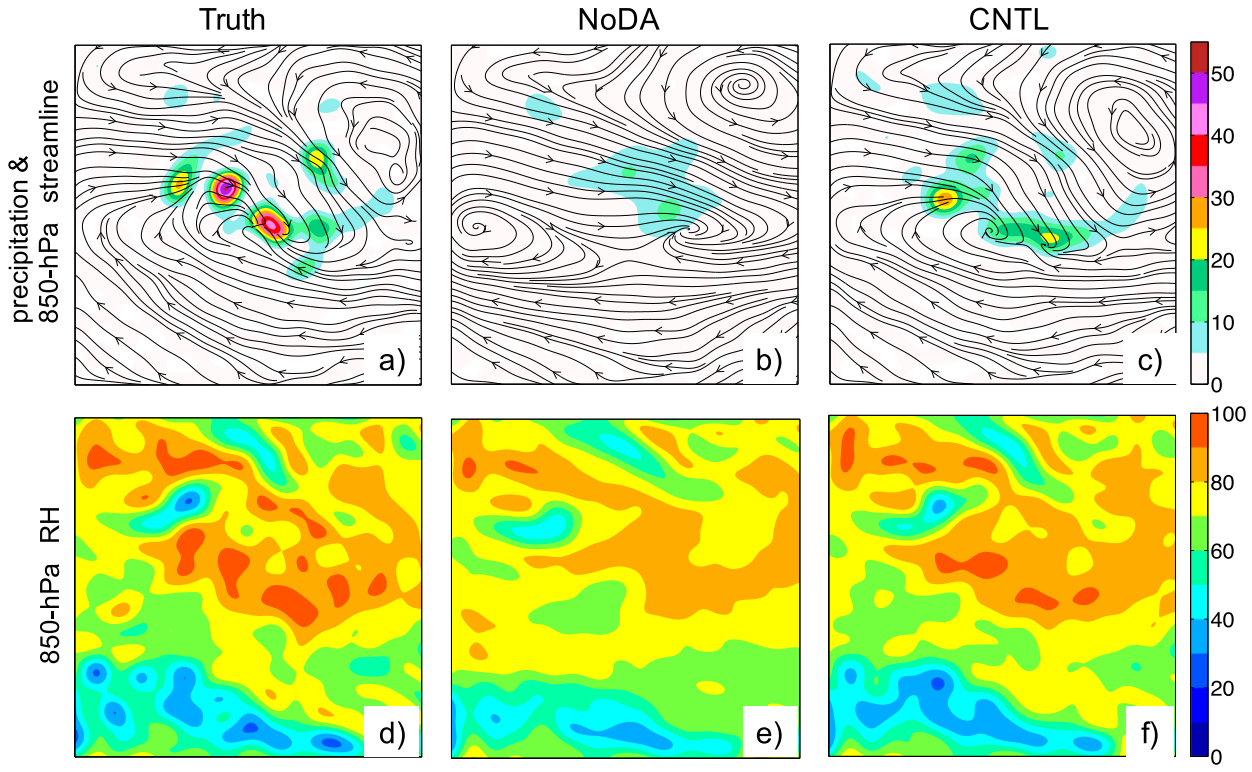

FIG. 7. (a)–(c) Snapshots of 850-hPa streamlines and 3-h accumulated precipitation (shading; mm) and (d)–(f) 850-hPa relative humidity (%) after 5.5 days of cycling (valid at 1200 UTC 21 Oct). Results are compared for (a),(d) truth simulation, (b),(e) NoDA ensemble mean, and (c),(f) CNTL posterior ensemble mean. All fields are filtered for $l > 200$ km. The left half of the OSSE domain is shown (10°S–10°N, 55°–75°E).

lateral boundary conditions by the current OSSE design, the convergence line and precipitation peaks are almost missing by the ensemble simulation without data assimilation. EnKF assimilation of satellite observations in CNTL recovers the flow pattern, although the location of precipitation is still slightly but noticeably different from the truth. Similarly, CNTL has a better representation of atmospheric moisture near the line of precipitation than NoDA as shown from the relative humidity fields. Note that CNTL has a weaker precipitation signal because of the averaging among members that have precipitation at different locations. Although small-scale signals ($l < 200$ km) are filtered out, this effect is still present.

Figure 8 shows time series of the domain-averaged root-mean-square errors (RMSEs) with respect to the truth from NoDA, CNTL, and the forecasts from CNTL. Zonal wind (u) and meridional wind (v) errors are combined and shown as root-mean difference kinetic energy (RM-DKE). Difference kinetic energy (Zhang et al. 2002) is defined as $(u'^2 + v'^2)/2$, where u' and v' are the errors in u and v, respectively. The forecasts are used to evaluate how long in time the prediction skill can be retained. The posterior ensemble mean is lack of small-scale signals, which is unrealistic and not necessarily in physical balance. Therefore, an ensemble forecast is required. Schwartz et al. (2014) suggested that sub-ensembles of 20–30 members will have comparable forecast skills with a full 50-member ensemble. In this study, a small 10-member ensemble forecast is performed, and forecast errors are calculated from the mean of ensemble forecasts. The directly observed variables (u, v, T, and Q) all show evident error reduction in CNTL from NoDA. The hydrometeor mixing ratios and w are also improved slightly from the better flow pattern and thermodynamic structures. The horizontal wind, T, and Q forecast errors grow linearly for the first day and more flow dependent afterward as errors approach the NoDA level. Hydrometeor and w error growth is flow dependent almost right from the beginning. The apparent error peaks at days 6, 11, and 14–17 are associated with the precipitation episodes.

Figure 9 shows vertical error profiles from NoDA, CNTL, and the forecasts, and Fig. 10 shows their spectra.

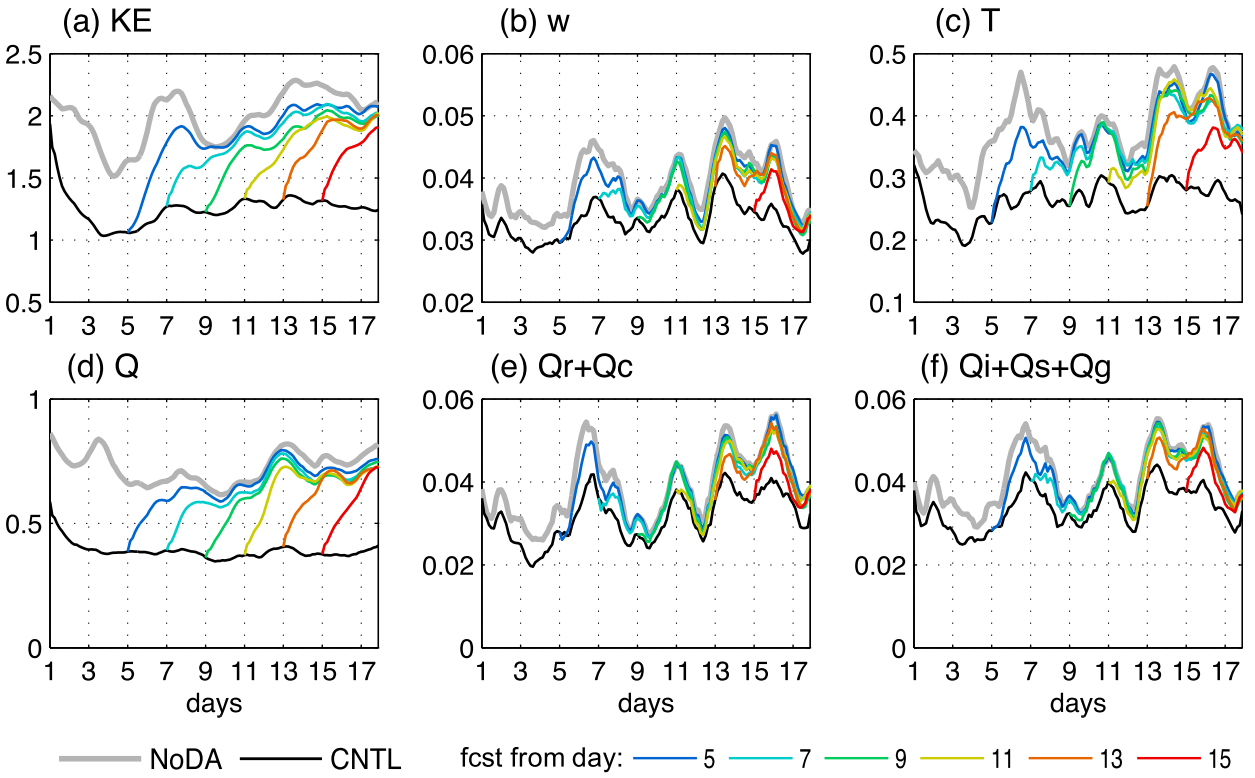

FIG. 8. Times series of domain-averaged (a) RM-DKE ($m\,s^{-1}$) and RMSEs of (b) vertical motion ($m\,s^{-1}$), (c) temperature (K), (d) specific humidity ($g\,kg^{-1}$), (e) rain + cloud water mixing ratio ($g\,kg^{-1}$), and (f) ice + snow + graupel mixing ratio ($g\,kg^{-1}$). The errors are also averaged over the vertical column and filtered for $l > 200$ km. Black lines show error from CNTL posterior ensemble mean and gray lines from NoDA ensemble mean. Colored lines show error evolution from ensemble forecasts initiated at days 5, 7, 9, 11, 13, and 15. All ensemble means are calculated from 10 members.

The forecast errors at certain lead times are averaged over 10 forecast runs during the test period (five forecasts starting from 0000 UTC on days 5, 7, . . . , 13, and five forecasts starting 1200 UTC). Errors peak at different vertical levels for each variable. Horizontal wind, w, T, and solid hydrometeor (ice, snow, and graupel) errors have peaks at upper levels that associate with cloud tops. Deep convective clouds are likely responsible for the liquid hydrometeor (rain and cloud) errors at midlevel, while shallow low-level clouds likely induce errors in horizontal wind, T, and Q around 900 hPa. Their spectral error distribution confirms that improvement from NoDA to CNTL is mostly at scales larger than 200 km. The forecast errors saturate faster for smaller scales. For 200–500-km scales, error saturates after 2 days; for scales larger than 1000 km, errors have not reached saturation after 4 days. The hydrometeors and w have faster small-scale error growth than other variables. This scale- and variable-dependent forecast error growth is consistent with the findings from Bei and Zhang (2014) and YZ17.

5. Relative impact from different observing networks

The +*Met7* experiment assimilates *Met7* channel-3 Tb in addition to CNTL. Figs. 11a–d shows a snapshot of its resulting channel-2 Tb as an independent verification in comparison to other experiments. In CNTL, the assimilation of ATOVS and AMV recovers the large-scale precipitating cloud cluster from NoDA already. However, its small-scale details of convective clouds are still erroneous comparing to the truth. The +*Met7* result has a better agreement with the truth for the location and strength of cloud clusters (low Tb) than CNTL. The horizontal distribution of ice, snow, and graupel, to which infrared Tb is most sensitive, are shown in Figs. 11e–h. Although not perfect comparing to the truth, the +*Met7* hydrometeors are improved over CNTL.

The +CYGNSS experiment assimilates CYGNSS wind speed in addition to CNTL, and another experiment NoProfile assimilates both *Met7*-Tb and CYGNSS

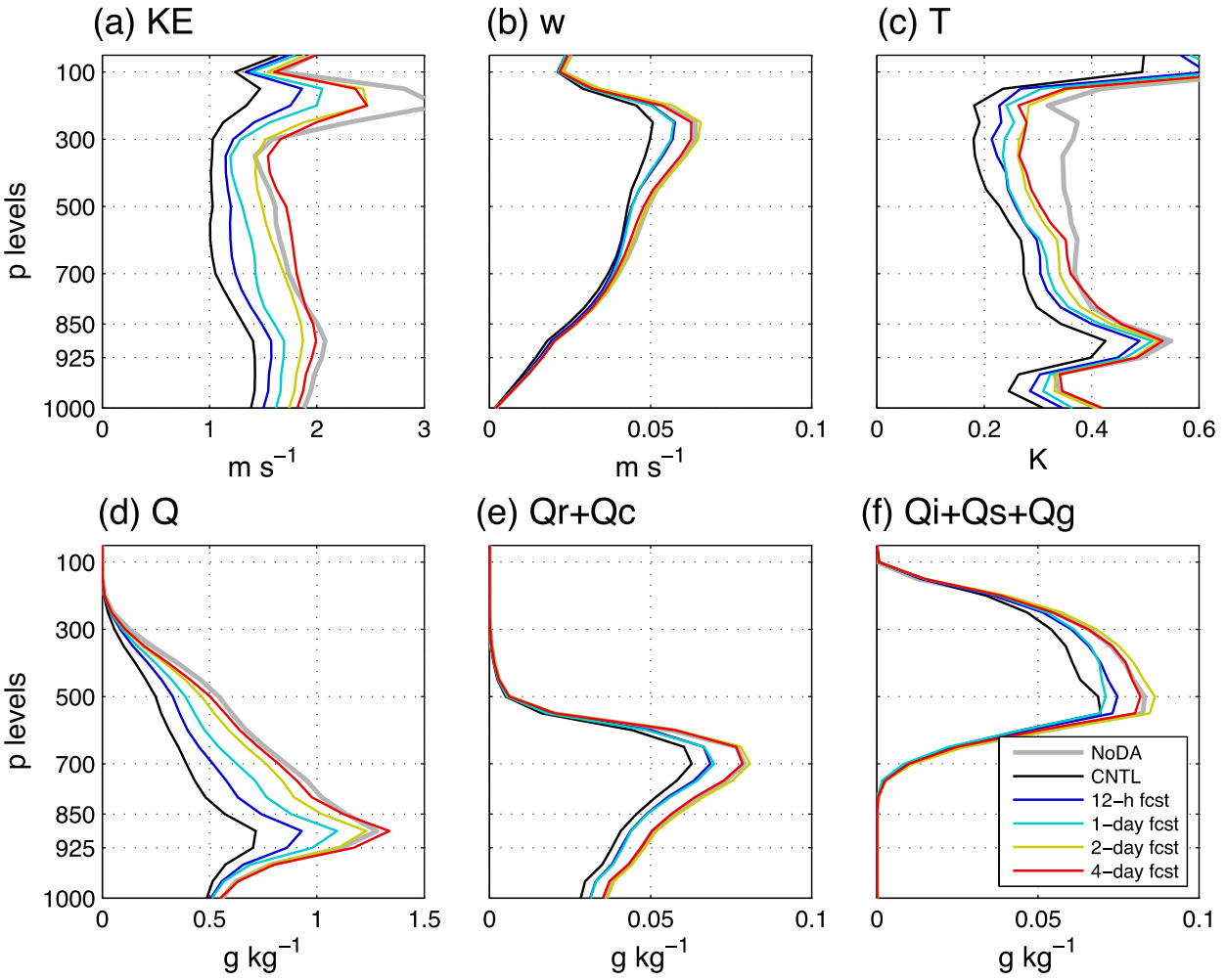

FIG. 9. Vertical profiles of (a) RM-DKE ($m\,s^{-1}$) and RMSEs of (b) vertical motion ($m\,s^{-1}$), (c) temperature (K), (d) specific humidity ($g\,kg^{-1}$), (e) rain + cloud water mixing ratio ($g\,kg^{-1}$), and (f) ice + snow + graupel mixing ratio ($g\,kg^{-1}$). The errors are averaged over days 5–18 and filtered for $l > 200$ km. Black lines show error from CNTL posterior ensemble mean and gray lines from NoDA ensemble mean. Colored lines show error at forecast lead time of 12 h and 1, 2, and 4 days. All ensemble means are calculated from 10 members.

in addition to CNTL but excludes ATOVS profiles. Figure 12 plots the error time series, vertical profiles, and spectra from these experiments for comparison. The +*Met7* can effectively reduce errors in ice, snow, and graupel mixing ratios (Fig. 12d) but not so much for rain and cloud (not shown). Horizontal wind, *T*, and *Q* are also slightly improved by assimilating *Met7*-Tb. Since *Met7*-Tb has a much higher horizontal resolution than ATOVS profiles, it can potentially improve the predictability for scales smaller than 200 km. The summed error spectra of solid hydrometeors (Fig. 12l) show that +*Met7* indeed reduces error at small scales. However, in this study, the 3-h cycling period is too long to constrain errors at these scales before their saturation. Potentially, assimilating *Met7*-Tb at hourly intervals and with a higher-resolution model will better utilize its positive impact. The single-channel Tb assimilated is also not providing sufficient vertical profile information. Comparing NoProfile to CNTL, it is evident that ATOVS profiles constrain better the horizontal wind, *T*, and *Q*. The ATOVS profiles combine the information from multiple satellite Tb images that are sensitive to different vertical heights, effectively having better vertical resolution than a single-channel Tb. The future development in all-sky infrared and microwave Tb assimilation will allow a fairer comparison between direct Tb assimilation and assimilating retrieved profiles. In addition to AMV wind, CYGNSS observations provide extra wind information for the lowest levels, reducing horizontal wind, *T*, and *Q* errors from surface up to 850 hPa. This improvement is mostly found for larger scales because of its moderate resolution. In NoProfile,

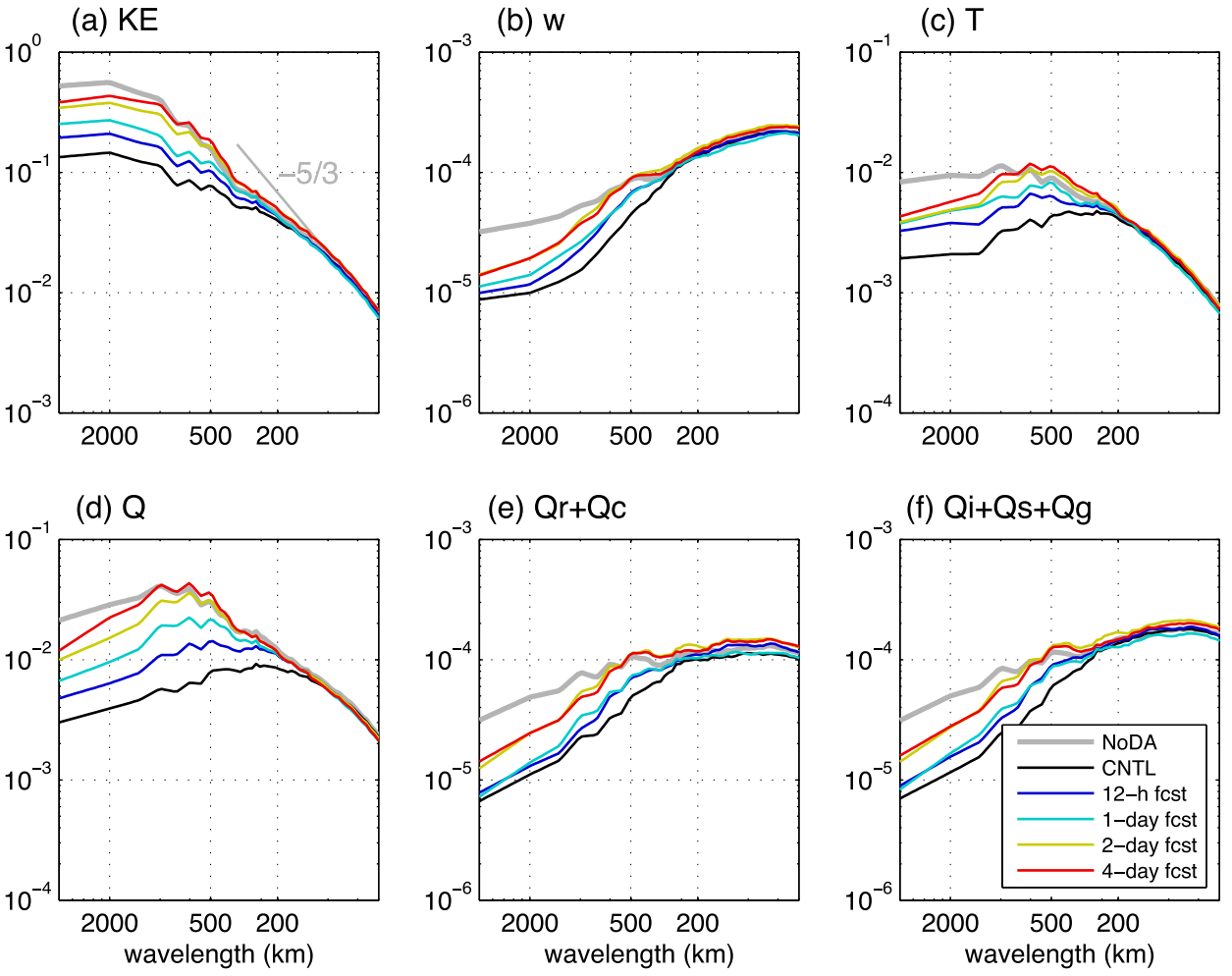

FIG. 10. As in Fig. 9, but showing the temporally and vertically averaged error energy spectra. The spectral errors are in variance units ($m^2 s^{-2}$ for DKE and K^2 for temperature). The spectrum is shown with respect to global horizontal wavenumber, $k = (k_x^2 + k_y^2)^{1/2}$, but labeled with its corresponding wavelength ($l = k^{-1}$).

the denial of ATOVS profiles causes a larger error in midlevel horizontal winds; however, their errors at upper and low levels remain small with the assimilation of AMV and CYGNSS (Fig. 12e).

To further test the sensitivity to resolution of the retrieval T and Q profiles from either ATOVS or GPSRO, several additional experiments are conducted assimilating retrieved profiles in addition to NoProfile. The +ATOVS assimilates the ATOVS profiles at their current resolution, while +ATOVScoarse assimilates them with temporal and horizontal resolution reduced by half. The +GPSRO assimilates GPSRO profiles at COSMIC-2 resolution, while +GPSROdense assimilates a denser version from the hypothetical populated CubeSat constellation. Note that the localization distance is increased for coarse networks according to the increase in observation intervals. Figure 13 compares their resulting error time series, vertical profiles, and spectra. When reduced-resolution ATOVS profiles are assimilated, their impact is also reduced for all variables and scales. The COSMIC-2 GPSRO profiles bring an impact that is approaching but from time to time less than the reduced-resolution ATOVS profiles. The random location of these profiles is a disadvantage over the uniformly spaced ATOVS profiles in continuous sampling of convective systems of interest. This issue is likely resolved with an increase in observation density. As shown in +GPSROdense, the CubeSat constellation with 25 times more data will yield a retrieval product dense enough to surpass the performance of the current network of ATOVS profiles.

In +ATOVS, the ATOVS profiles are assumed to have ~40 independent pieces of information for T in the vertical, which likely overestimates the information content and may yield an analysis that is too optimistic.

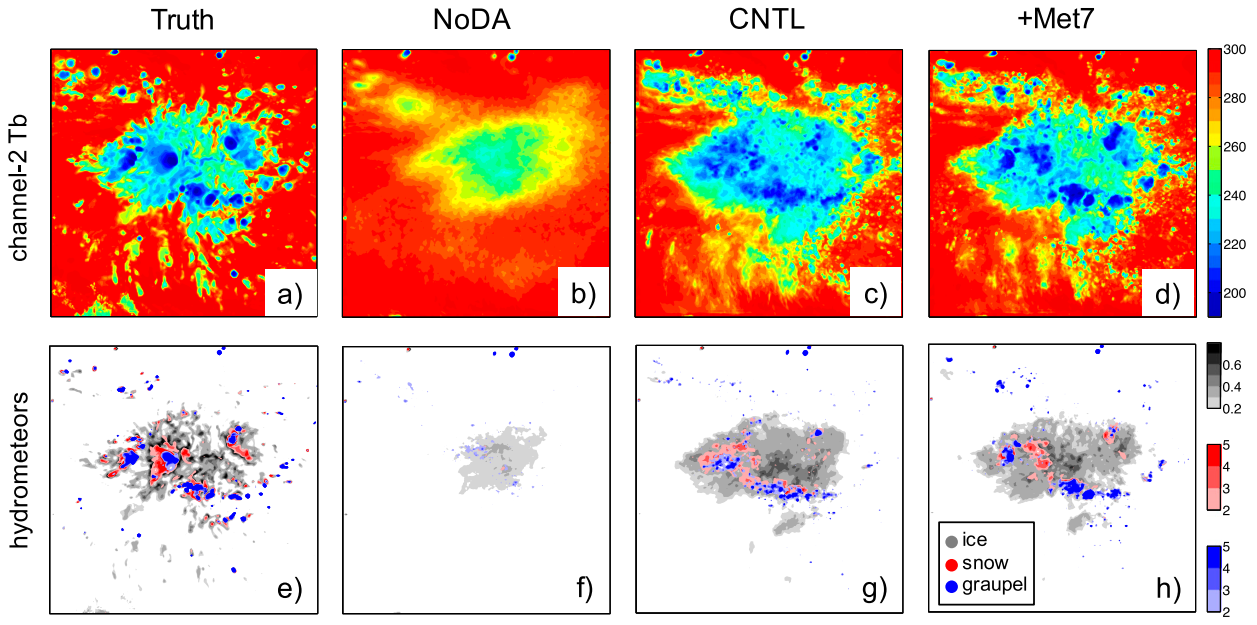

FIG. 11. (a)–(d) Snapshots of channel-2 Tb (K) and (e)–(h) column-integrated ice (gray), snow (red), and graupel (blue) mixing ratios ($g\,kg^{-1}$) after 5.5 days of cycling (valid at 1200 UTC 21 Oct) from (a),(e) truth simulation, (b),(f) NoDA ensemble mean, (c),(g) CNTL posterior ensemble mean, and (d),(h) +*Met7* posterior ensemble mean.

To at least partially address this concern, two additional experiments are conducted. The +ATOVSthinV assimilates ATOVS profiles with vertical levels thinned to 10 levels for T and 4 levels for Q, and +ATOVSerror1.5 assimilates ATOVS profiles with vertical resolution unchanged but errors for both T and Q inflated to 150%. Note that for +ATOVSthinV, the vertical ROI is enlarged to 20 levels as the observations are thinned in the vertical. Figure 14 shows that reduction in vertical resolution/accuracy does not considerably degrade the analysis accuracy. The increase in analysis error in these experiments is much smaller compared to +ATOVScoarse with reduced horizontal and temporal resolutions. This also indicates that the vertical information content estimated to contain ~10 independent pieces of information is likely able to capture the key vertical modes in the tropics.

Figure 15 compares the direct assimilation of GPSRO N profiles (+GPSRON and +GPSROdenseN) with assimilation of retrieved T and Q profiles (+GPSRO and +GPSROdense). The overall analysis error is similar for both assimilation strategies. The nonlinear observation operator results in a shorter decorrelation length scale, which requires the localization distance to be shorter to achieve similar performance. The nonlinearity also causes longer filter spinup period for Q (Fig. 15c) and causes the analysis for T and winds to be less accurate than those from assimilating retrieved profiles (Figs. 15a and 15b). Since our idealized experiment assumes no errors in the retrieval process and error correlation and bias are not considered, the performance of assimilating retrieval profiles is an optimistic expectation compared to real-data cases.

6. Observation impact on CCEWs

The same space–time bandpass filter from section 2 is applied to posterior ensemble mean to evaluate the observation impact for each CCEW mode. The posterior mean fields lack the spatial and temporal coverage to directly filter for large-scale waves; thus, the month-long large-domain truth simulation is used to fill in the gaps before applying the bandpass filter. Figure 16 plots the longitude–time diagrams of the filtered 850-hPa zonal wind from NoDA and CNTL and compares them to the truth, and Table 4 quantifies these errors. Two error sources can lead to a large RMSE: the waves being out of phase with the truth and/or the amplitude of the waves being wrong. The pattern correlation quantifies the contribution from phase errors; if the waves are perfectly in phase with the truth (correlation equals 1), the RMSE is solely due to amplitude errors. Compared to the truth, NoDA does capture most of the large-scale low-frequency wave phases. For the MJO, its correlation is up to 0.98, indicating that using a perfect LBC in the current OSSEs plays a dominant role in modulating the model solution of MJO, which agrees with the sensitivity experiment results from Zhang et al. (2017). The ER, Kelvin, MRG and large-scale WIG waves also have

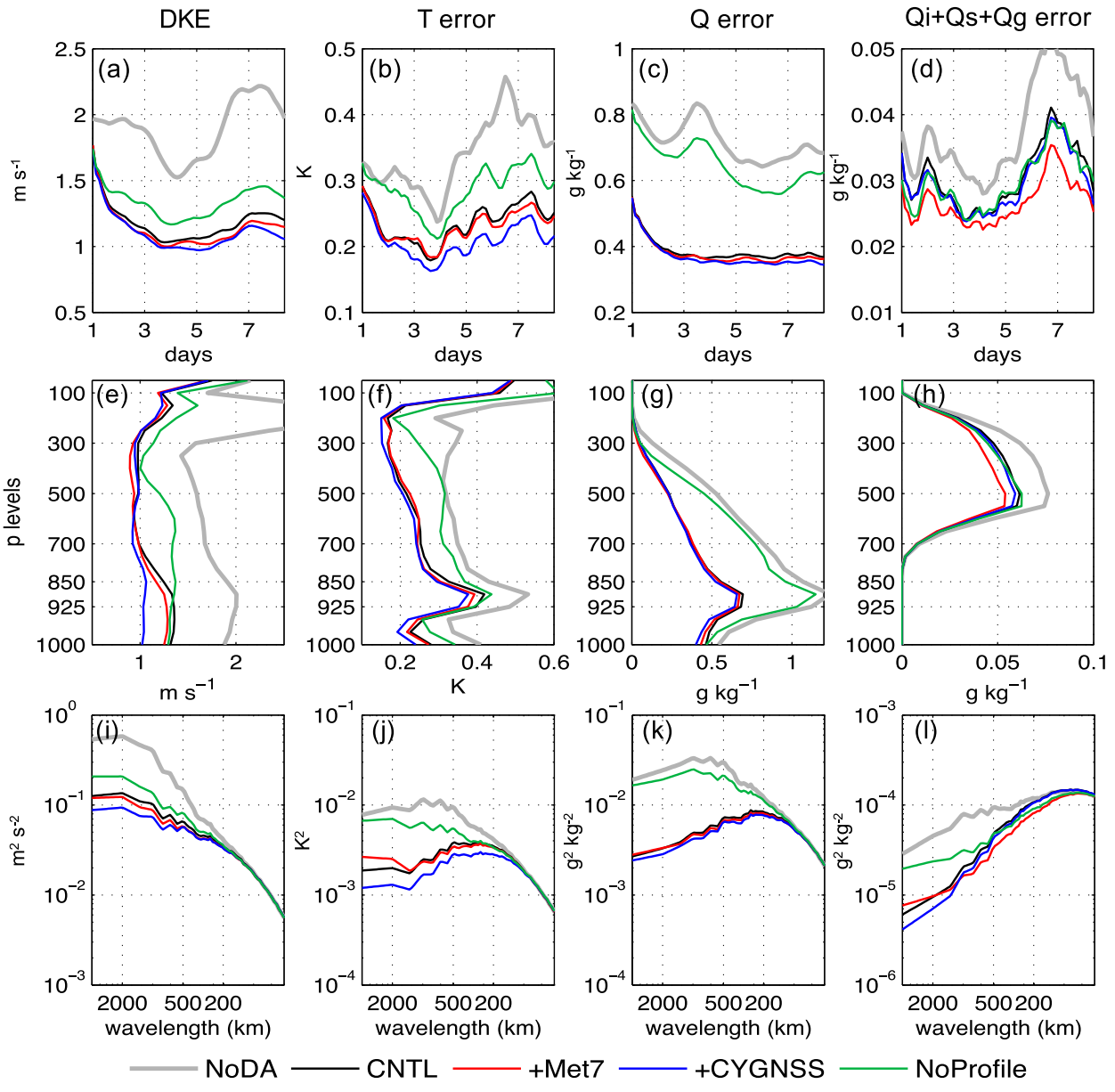

FIG. 12. (a)–(d) Time series, (e)–(h) vertical profiles, and (i)–(l) spectra of domain-averaged analysis errors in (a),(e),(i) kinetic energy, (b),(f),(j) temperature, (c),(g),(k) specific humidity, and (d),(h),(l) summed error from ice, snow, and graupel mixing ratios. Errors are shown as RMSE (RM-DKE for kinetic energy) except that the spectra show error variance. The time series are vertically averaged and filtered for $l > 200$ km; vertical profiles are averaged over time and filtered for $l > 200$ km; spectra are averaged temporally and vertically. Results from NoDA, CNTL, +*Met7*, +CYGNSS, and NoPofile are compared.

reasonable forecasted phases in NoDA, but errors grow as wave phases propagate away from LBC and further downstream. The ER and MRG westerly signals at day 7 start from the middle of the domain, which is not captured by NoDA. The large-scale WIG signals at the east side of the domain are much weaker in NoDA than the truth, and for intermediate-scale WIG waves, the signals are almost wiped out in the east half of the domain.

By assimilating observations in CNTL, the phase and amplitude of these wave modes are significantly improved. Figure 17 shows the error reduction from NoDA to CNTL as measured by root-mean difference total energy (RM-DTE; Melhauser and Zhang 2012). The MJO amplitude error is further reduced in CNTL for all variables. For ER, Kelvin, MRG, and large-scale WIG wave, the phase correlation is improved from ~0.9

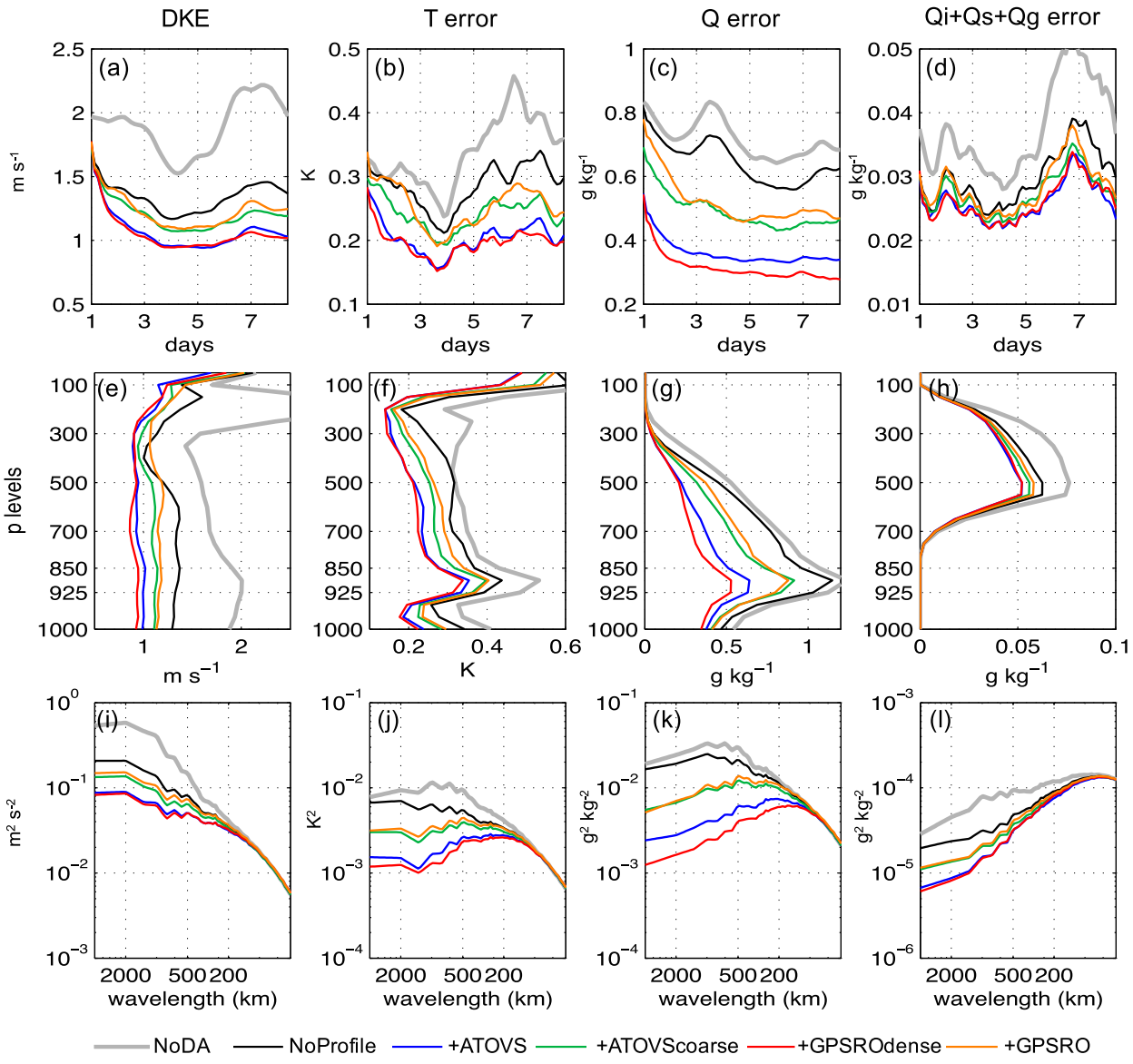

FIG. 13. As in Fig. 12, but showing results from NoDA, NoProfiles, +ATOVS, +ATOVScoarse, +GPSRO, and +GPSROdense.

to almost perfect (Table 4). As for the more challenging intermediate-scale WIG waves, the phase correlation is also much improved for all variables (e.g., from 0.51 to 0.83 for zonal wind). The precipitation has larger phase errors than other variables because of its more limited predictability. Table 5 shows the relative impacts on WIG waves from assimilating different observations. Since these sensitivity experiments are only run for 9 days, they are not diagnosed for lower-frequency large-scale waves. Assimilating AMV, *Met7*-Tb, and CYGNSS (NoProfile) significantly improves both large-scale and intermediate-scale WIG waves. The addition of ATOVS profiles further reduces the errors in all variables. The additional merit from higher-resolution profiles appears to be more important for better analysis and prediction of intermediate-scale WIG waves.

7. Concluding remarks

In this study, a perfect-model OSSE is performed to evaluate the potential impact from assimilating satellite observations on the practical predictability of tropical multiscale weather systems. Following YZ17, a regional WRF Model is configured to simulate the CCEWs during a 17-day period within the active phase of the October 2011 MJO event. The simulation captures the

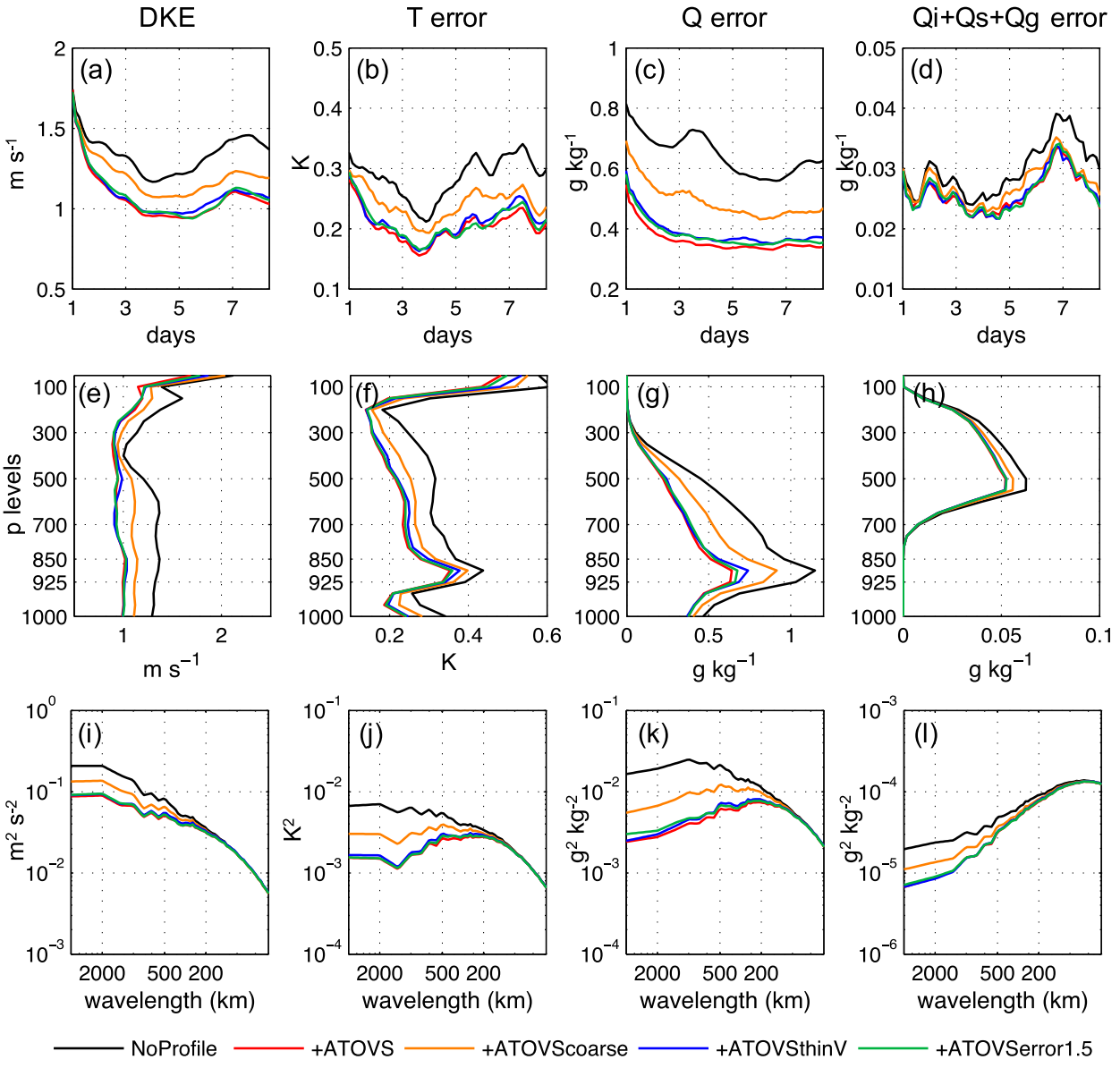

FIG. 14. As in Fig. 12, but showing results from NoProfiles, +ATOVS, +ATOVScoarse, +ATOVSthinV, and +ATOVSerror1.5.

strong precipitation signals associated with the MJO, Kelvin, and IG waves. The ER and MRG waves have relatively weak precipitation but well-defined zonal wind signals. Most wave components are at scales larger than 1000 km except for IG waves that have a smaller-scale subdiurnal component. The OSSE assumes that a perfect forecast model is used, and observation errors are also modeled perfectly during data assimilation.

Table 3 summarizes the benefits of assimilating each observing network every 3 h using EnKF. According to Zhang et al. (2017), the global circumnavigating signals play an important role in MJO initiation inside their domain. Therefore, the perfect LBC alone is able to provide the correct phase of most of the large-scale waves (MJO, ER, Kelvin, MRG), and their errors are dominated by amplitude errors. The smaller-scale high-frequency waves (WIG), on the other hand, are less well constrained by LBC, and large phase errors occur without data assimilation. The CNTL experiment tests assimilation of ATOVS T and Q profiles and AMV wind observations for the 17-day period and showed an improvement in winds, T, Q, and hydrometeors for scales larger than 200 km. For large-scale ($l > 1000$ km) CCEWs, their predictability limit can be extended by about 4 days after assimilation. However, the small-scale ($l < 500$ km) waves have more limited

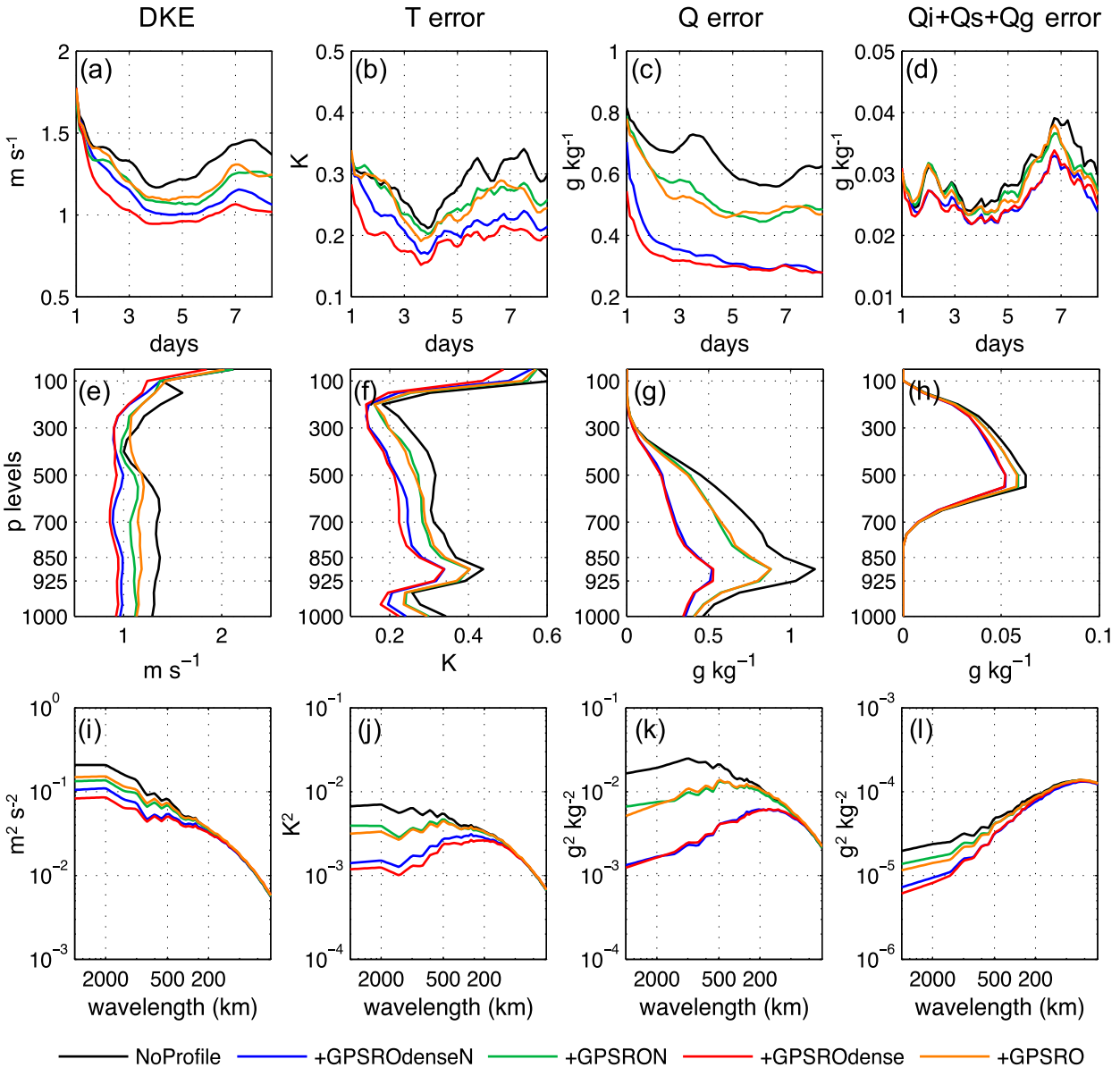

FIG. 15. As in Fig. 12, but showing results from NoProfiles, +GPSROdenseN, +GPSRON, +GPSRODense, and +GPSRO.

predictability with the extra prediction skill lasting for less than 1 day. The EnKF assimilation further improved amplitude for the large-scale waves and reduced phase errors of intermediate-scale WIG waves as well. Assimilating CYGNSS wind speed retrievals reduces errors in large-scale low-level horizontal wind and *T*. It provides complementary wind information to the surface and low levels where not many AMVs are available.

The *Met7* infrared Tb has higher horizontal resolution and thus further improves the model variables at smaller scales. The improvement is found mostly in the ice, snow, and graupel mixing ratios near cloud top. However, the 3-h cycling is probably not frequent enough to maintain the extra prediction skill at small scales. Assimilating Tb from only one infrared channel does not constrain model dynamic and thermodynamic variables as efficiently as the ATOVS profiles. The later has the advantage of combining multiple satellite images that have information at different vertical levels. The uniform horizontal distribution of ATOVS profiles yields a more persistent observation impact than the sporadic sampling of sparser GPSRO profiles. At COSMIC-2 resolution, GPSRO profiles have impacts approaching

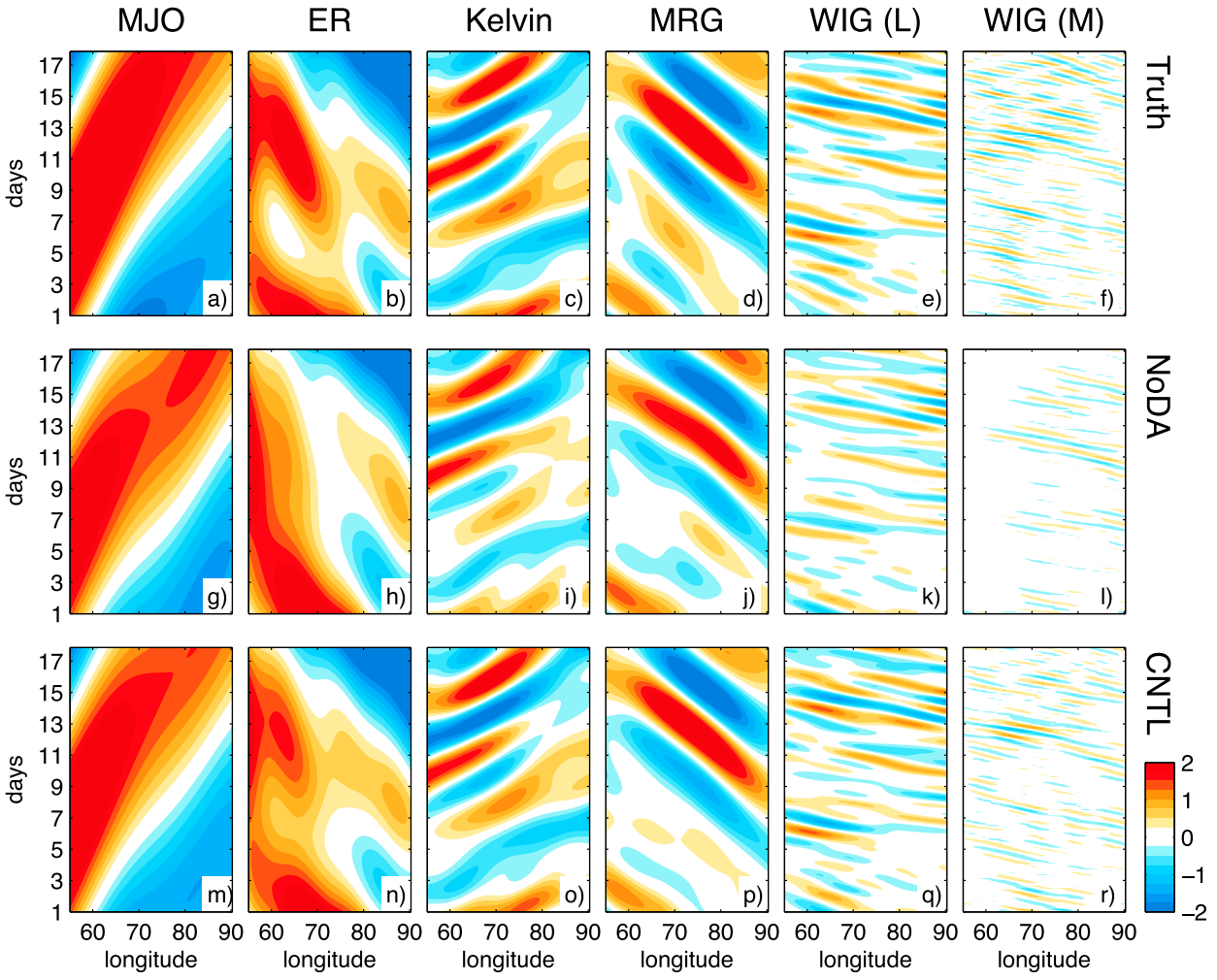

FIG. 16. Longitude–time Hovmöller diagrams of 850-hPa zonal wind (color shading; m s^{-1}) averaged over 0°–5°N from (a)–(f) truth simulation, (g)–(l) NoDA ensemble mean, and (m)–(r) CNTL posterior ensemble mean. (left to right) The fields are filtered for each CCEW.

half of those from ATOVS profiles. Although the number of profiles is likely still 10 times fewer than the reduced-resolution ATOVS, their higher vertical resolution and more accurate retrievals of both T and Q bring more impact per profile. The better horizontal resolution from the ATOVS retrieval profiles and better vertical resolution from the GPSRO profiles can potentially complement each other. With expected future improvement in CubeSat resolution, the GPSRO profiles may have an impact comparable to or exceeding that from the current ATOVS profiles. The high-resolution profiles can potentially improve smaller-scale weather systems, such as the subdiurnal WIG waves coupled with transient convective processes.

While these results encourage the future development and use of satellite-based observations for the analysis and prediction of tropical weather, the real-data application faces more challenges. There may be nontrivial uncertainties in model dynamics, physics schemes, and forward operator not accounted by the present study. The exact retrieval algorithms are not used in this study to generate ATOVS and GPSRO profiles; thus, the simulated errors for these synthetic profiles do not account for complicated error sources from the retrieval processes. Thus, the idealized experiment results should be interpreted with these caveats in mind. The assimilation of these satellite observations in real-data cases needs to be tested in future studies. In this study, a limited-area domain is used and a perfect LBC is specified, and uncertainties from the LBC and forecast model itself are not accounted for. Therefore, the NoDA experiment is not intended to account for all realistic sources of forecast errors that can impact practical predictability. Since we only focused on the

TABLE 4. RMSEs with respect to the truth simulation for u (m s^{-1}), T (K), Q (g kg^{-1}), and precipitation (mm day^{-1}) filtered for CCEW types described in Table 1. The errors are averaged over latitude (0°–5°N) and time (17-day period). Pattern correlations with the truth are shown in parentheses. Results are compared between NoDA ensemble mean and CNTL posterior ensemble mean.

CCEW type	Experiment	u	T	Q	Precipitation
MJO	NoDA	0.39 (0.98)	0.063 (0.98)	0.072 (0.99)	1.97 (0.99)
	CNTL	0.17 (1.00)	0.025 (1.00)	0.020 (1.00)	0.53 (1.00)
ER	NoDA	0.50 (0.90)	0.064 (0.90)	0.073 (0.94)	1.18 (0.88)
	CNTL	0.19 (0.99)	0.033 (0.97)	0.027 (0.99)	0.61 (0.98)
Kelvin	NoDA	0.40 (0.92)	0.089 (0.92)	0.078 (0.87)	2.37 (0.83)
	CNTL	0.13 (0.99)	0.025 (0.99)	0.027 (0.98)	0.69 (0.99)
MRG	NoDA	0.33 (0.91)	0.029 (0.85)	0.088 (0.90)	1.00 (0.77)
	CNTL	0.15 (0.98)	0.012 (0.98)	0.021 (0.99)	0.41 (0.97)
WIG (L)	NoDA	0.38 (0.82)	0.067 (0.90)	0.078 (0.80)	2.83 (0.79)
	CNTL	0.14 (0.98)	0.025 (0.99)	0.032 (0.97)	1.20 (0.97)
WIG (M)	NoDA	0.38 (0.51)	0.097 (0.69)	0.060 (0.54)	4.88 (0.35)
	CNTL	0.25 (0.83)	0.061 (0.89)	0.047 (0.77)	2.83 (0.84)

impact of assimilating data in a limited-area domain for a selected period, one should be cautious in generalizing our findings. Although the 9-km horizontal grid spacing used is comparable to the current operational global models, it will be interesting to further test the sensitivity of EnKF performance in satellite data assimilation to the horizontal resolution, especially for smaller-scale tropical weather phenomena and CCEWs. In this study, we

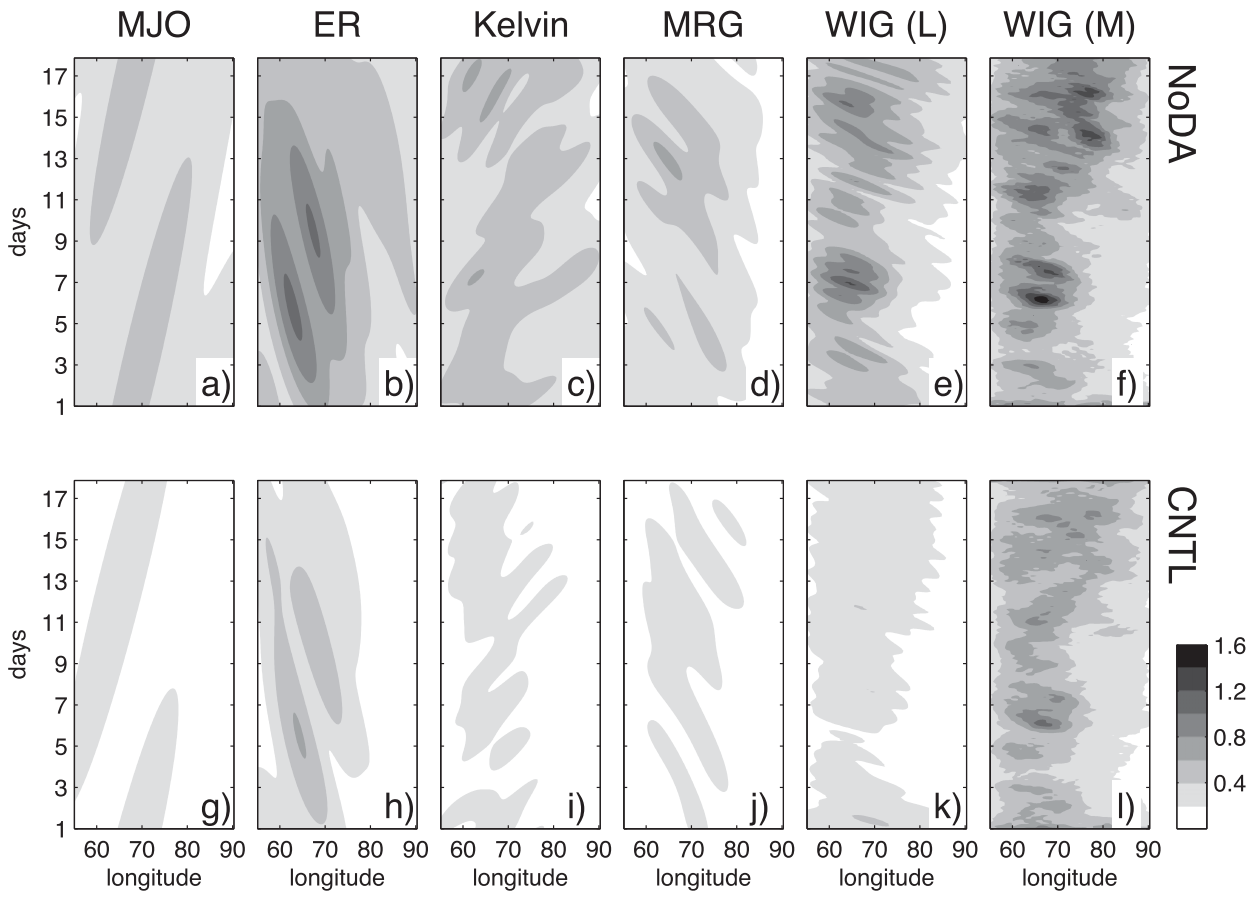

FIG. 17. Longitude–time Hovmöller diagrams of RM-DTE with respect to the truth (m s^{-1}) averaged over 0°–5°N and vertical levels for (a)–(f) NoDA ensemble mean and (g)–(l) CNTL posterior ensemble mean. (left to right) A space–time filter is applied to u, y, and T for each CCEW before calculating the RM-DTE.

TABLE 5. As in Table 4, but comparing results from NoDA, NoProfile, +ATOVScoarse, and +ATOVS posterior ensemble mean from the first 10 days and for WIG waves.

CCEW type	Experiment	u	T	Q	Precipitation
WIG (L)	NoDA	0.37 (0.74)	0.060 (0.90)	0.072 (0.80)	2.92 (0.62)
	NoProfile	0.19 (0.94)	0.043 (0.95)	0.062 (0.86)	1.61 (0.92)
	+ATOVScoarse	0.14 (0.97)	0.031 (0.97)	0.045 (0.93)	1.06 (0.96)
	+ATOVS	0.13 (0.97)	0.025 (0.98)	0.032 (0.96)	0.90 (0.97)
WIG (M)	NoDA	0.32 (0.59)	0.086 (0.73)	0.056 (0.54)	4.30 (0.32)
	NoProfile	0.26 (0.74)	0.076 (0.80)	0.050 (0.66)	3.04 (0.74)
	+ATOVScoarse	0.24 (0.80)	0.069 (0.84)	0.048 (0.71)	2.56 (0.82)
	+ATOVS	0.23 (0.82)	0.059 (0.88)	0.046 (0.74)	2.10 (0.88)

highlighted the different performance from ATOVS and GPSRO profiles due to their spatial coverage and resolution. A similar OSSE can be useful for assessing the cost effectiveness of other future space-based sensors. We also believe that a more systematic direct assimilation of satellite Tb is needed for substituting the assimilation of retrieved profiles. The synthetic observations that are available evenly in space and time allow better quantification of multiscale analysis errors. One caveat of this assumption is that results are rather optimistic in terms of the information content and availability of the current satellite observing networks. On the other hand, although methods like the AOEI make the direct assimilation of Tb possible, we believe the filter performance is still suboptimal, and a better methodology in dealing with nonlinearity in forward operator is still much desired.

Acknowledgments. This research was supported by NSF Grant AGS-1305798, NASA Grants NNX16AD84G and NNX12AJ79G, and Office of Naval Research Grant N000140910526. Yue Ying was also sponsored by the China Scholarship Council (CSC). Computing is performed at the Texas Advanced Computing Center (TACC) and the National Center for Atmospheric Research (NCAR). All experiment data from this study can be freely accessed upon request at the data archive facility of TACC. The authors also appreciate the insightful comments from two anonymous reviewers on an earlier version of this paper.

REFERENCES

Bei, N., and F. Zhang, 2007: Impacts of initial condition errors on mesoscale predictability of heavy precipitation along the mei-yu front of China. *Quart. J. Roy. Meteor. Soc.*, **133**, 83–99, https://doi.org/10.1002/qj.20.

——, and ——, 2014: Mesoscale predictability of moist baroclinic waves: Variable and scale-dependent error growth. *Adv. Atmos. Sci.*, **31**, 995–1008, https://doi.org/10.1007/s00376-014-3191-7.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Chou, M. D., and M. J. Suarez, 1999: A solar radiation parameterization for atmospheric studies. NASA Tech. Rep. Series on Global Modeling and Data Assimilation 104606, Vol. 15, 40 pp.

Cook, K. L. B., P. Wilczynski, C. J. Fong, N. L. Yen, and G. S. Chang, 2011: The Constellation Observing System for Meteorology Ionosphere and Climate follow-on mission. *Proc. Aerospace Conf.*, Big Sky, MT, IEEE, 1–7, https://doi.org/10.1109/AERO.2011.5747291.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Dickinson, M., and J. Molinari, 2002: Mixed Rossby–gravity waves and western Pacific tropical cyclogenesis. Part I: Synoptic evolution. *J. Atmos. Sci.*, **59**, 2183–2196, https://doi.org/10.1175/1520-0469(2002)059<2183:MRGWAW>2.0.CO;2.

Dong, J., M. Xue, and K. Droegemeier, 2011: The analysis and impact of simulated high-resolution surface observations in addition to radar data for convective storms with an ensemble Kalman filter. *Meteor. Atmos. Phys.*, **112**, 41–61, https://doi.org/10.1007/s00703-011-0130-3.

Dunkerton, T. J., and F. X. Crum, 1995: Eastward propagating ~2- to 15-day equatorial convection and its relation to the tropical intraseasonal oscillation. *J. Geophys. Res.*, **100**, 25 781–25 790, https://doi.org/10.1029/95JD02678.

English, S. J., R. J. Renshaw, P. C. Dibben, A. J. Smith, P. J. Rayer, C. Poulsen, F. W. Saunders, and J. R. Eyre, 2000: A comparison of the impact of TOVS arid ATOVS satellite sounding data on the accuracy of numerical weather forecasts. *Quart. J. Roy. Meteor. Soc.*, **126**, 2911–2931, https://doi.org/10.1002/qj.49712656915.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, https://doi.org/10.1029/94JC00572.

Gaspari, G., and S. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, https://doi.org/10.1002/qj.49712555417.

Geer, A. J., and P. Bauer, 2011: Observation errors in all-sky data assimilation. *Quart. J. Roy. Meteor. Soc.*, **137**, 2024–2037, https://doi.org/10.1002/qj.830.

Haertel, P. T., and G. N. Kiladis, 2004: Dynamics of 2-day equatorial waves. *J. Atmos. Sci.*, **61**, 2707–2721, https://doi.org/10.1175/JAS3352.1.

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.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1.

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.

Kiladis, G. N., M. C. Wheeler, P. T. Haertel, K. H. Straub, and P. E. Roundy, 2009: Convectively coupled equatorial waves. *Rev. Geophys.*, **47**, RG2003, https://doi.org/10.1029/2008RG000266.

Kirchgessner, P., L. Nerger, and A. Bunse-Gerstner, 2014: On the choice of an optimal localization radius in ensemble Kalman filter methods. *Mon. Wea. Rev.*, **142**, 2165–2175, https://doi.org/10.1175/MWR-D-13-00246.1.

Kursinski, E. R., G. Hajj, J. Schofield, R. Linfield, and K. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the global positioning system. *J. Geophys. Res.*, **102**, 23 429–23 465, https://doi.org/10.1029/97JD01569.

Lei, L., and J. S. Whitaker, 2017: Evaluating the trade-offs between ensemble size and ensemble resolution in an ensemble-variational data assimilation system. *J. Adv. Model. Earth Syst.*, **9**, 781–789, https://doi.org/10.1002/2016MS000864.

Li, J., W. W. Wolf, W. P. Menzel, W. Zhang, H. Huang, and T. H. Achtor, 2000: Global soundings of the atmosphere from ATOVS measurements: The algorithm and validation. *J. Appl. Meteor.*, **39**, 1248–1268, https://doi.org/10.1175/1520-0450(2000)039<1248:GSOTAF>2.0.CO;2.

Lim, K.-S. S., and S.-Y. Hong, 2010: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models. *Mon. Wea. Rev.*, **138**, 1587–1612, https://doi.org/10.1175/2009MWR2968.1.

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.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708, https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

——, and ——, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. *J. Atmos. Sci.*, **29**, 1109–1123, https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2.

Mannucci, A. J., J. Dickson, C. Duncan, and K. Hurst, 2010: GNSS geospace constellation (GGC): A Cubesat space weather mission concept. California Institute of Technology Jet Propulsion Laboratory Rep., 5 pp.

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.

Meng, Z., and F. Zhang, 2008: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.*, **136**, 522–540, https://doi.org/10.1175/2007MWR2106.1.

Minamide, M., and F. Zhang, 2017: Adaptive observation error inflation for assimilating all-sky satellite radiance. *Mon. Wea. Rev.*, **145**, 1063–1081, https://doi.org/10.1175/MWR-D-16-0257.1.

Miyoshi, T., E. Kalnay, and H. Li, 2013: Estimating and including observation-error correlations in data assimilation. *Inverse Probl. Sci. Eng.*, **21**, 387–398, https://doi.org/10.1080/17415977.2012.712527.

Nieman, S. J., W. P. Menzel, C. M. Hayden, D. Gray, S. Wanzong, C. S. Velden, and J. Daniels, 1997: Fully automated cloud-drift winds in NESDIS operations. *Bull. Amer. Meteor. Soc.*, **78**, 1121–1134, https://doi.org/10.1175/1520-0477(1997)078<1121:FACDWI>2.0.CO;2.

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Periáñez, A., H. Reich, and R. Potthast, 2014: Optimal localization for ensemble Kalman filter systems. *J. Meteor. Soc. Japan*, **92**, 585–597, https://doi.org/10.2151/jmsj.2014-605.

Reale, A. L., 2001: NOAA operational sounding products from Advanced-TOVS polar orbiting environmental satellites. NOAA Tech. Rep. NESDIS 102, 59 pp.

Reynolds, C. A., P. J. Webster, and E. Kalnay, 1994: Random error growth in NMC's global forecasts. *Mon. Wea. Rev.*, **122**, 1281–1305, https://doi.org/10.1175/1520-0493(1994)122<1281:REGING>2.0.CO;2.

Ruf, C. S., and Coauthors, 2016: New ocean winds satellite mission to probe hurricanes and tropical convection. *Bull. Amer. Meteor. Soc.*, **97**, 385–395, https://doi.org/10.1175/BAMS-D-14-00218.1.

Schwartz, C. S., G. S. Romine, K. R. Smith, and M. L. Weisman, 2014: Characterizing and optimizing precipitation forecasts from a convection-permitting ensemble initialized by a mesoscale ensemble Kalman filter. *Wea. Forecasting*, **29**, 1295–1318, https://doi.org/10.1175/WAF-D-13-00145.1.

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.

Smith, E. K., and S. Weintraub, 1953: The constants in the equation for atmospheric refractivity index at radio frequencies. *J. Res. Natl. Bur. Stand. (U.S.)*, **50**, 39–41, https://doi.org/10.6028/jres.050.006.

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.

Tao, D., and F. Zhang, 2015: Effects of vertical wind shear on the predictability of tropical cyclones: Practical versus intrinsic limit. *J. Adv. Model. Earth Syst.*, **7**, 1534–1553, https://doi.org/10.1002/2015MS000474.

Tavolato, C., and L. Isaksen, 2015: On the use of a Huber norm for observation quality control in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **141**, 1514–1527, https://doi.org/10.1002/qj.2440.

Velden, C. S., and K. M. Bedka, 2009: Identifying uncertainty in determining satellite-derived atmospheric motion vector height attribution. *J. Appl. Meteor. Climatol.*, **48**, 450–463, https://doi.org/10.1175/2008JAMC1957.1.

——, C. Hayden, S. Nieman, W. Menzel, S. Wanzong, and J. Goerss, 1997: Upper-tropospheric winds derived from geostationary satellite water vapor observations. *Bull. Amer. Meteor. Soc.*, **78**, 173–195, https://doi.org/10.1175/1520-0477(1997)078<0173:UTWDFG>2.0.CO;2.

Wang, B. R., X. Y. Liu, and J. K. Wang, 2013: Assessment of COSMIC radio occultation retrieval product using global radiosonde data. *Atmos. Meas. Tech.*, **6**, 1073–1083, https://doi.org/10.5194/amt-6-1073-2013.

Wang, S., A. H. Sobel, F. Zhang, Y. Q. Sun, Y. Yue, and L. Zhou, 2015: Regional simulation of the October and November MJO events observed during the CINDY/DYNAMO field campaign at gray zone resolution. *J. Climate*, **28**, 2097–2119, https://doi.org/10.1175/JCLI-D-14-00294.1.

Wheeler, M. C., and G. N. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *J. Atmos. Sci.*, **56**, 374–399, https://doi.org/10.1175/1520-0469(1999)056<0374:CCEWAO>2.0.CO;2.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Ying, Y., and F. Zhang, 2017: Practical and intrinsic predictability of multiscale weather and convectively coupled equatorial waves during the active phase of an MJO. *J. Atmos. Sci.*, **74**, 3771–3785, https://doi.org/10.1175/JAS-D-17-0157.1.

Yunck, T. P., E. J. Fetzer, A. M. Mannucci, C. O. Ao, W. Irion, B. D. Wilson, and G. J. Manipon, 2009: Use of radio occultation to evaluate atmospheric temperature data from spaceborne infrared sensors. *Terr. Atmos. Ocean. Sci.*, **20**, 71–85, https://doi.org/10.3319/TAO.2007.12.08.01(F3C).

Zeng, X., and A. Beljaars, 2005: A prognostic scheme of sea surface skin temperature for modeling and data assimilation. *Geophys. Res. Lett.*, **32**, L14605, https://doi.org/10.1029/2005GL023030.

Zhang, C., J. Gottshalck, E. D. Maloney, M. W. Moncrieff, F. Vitart, D. E. Waliser, B. Wang, and M. C. Wheeler, 2013: Cracking the MJO nut. *Geophys. Res. Lett.*, **40**, 1223–1230, https://doi.org/10.1002/grl.50244.

Zhang, F., and J. A. Sippel, 2009: Effects of moist convection on hurricane predictability. *J. Atmos. Sci.*, **66**, 1944–1961, https://doi.org/10.1175/2009JAS2824.1.

——, C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the ''surprise'' 24–25 January 2000 snowstorm. *Mon. Wea. Rev.*, **130**, 1617–1632, 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.

——, ——, and J. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253, https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>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.

——, M. Minamide, and E. E. Clothiaux, 2016: Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophys. Res. Lett.*, **43**, 2954–2963, https://doi.org/10.1002/2016GL068468.

——, S. Talaphdar, and S. Wang, 2017: The role of global circumnavigating mode in the MJO initiation and propagation. *J. Geophys. Res. Atmos.*, **122**, 5837–5856, https://doi.org/10.1002/2016JD025665.

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.

Zhen, Y., and F. Zhang, 2014: A probabilistic approach to adaptive covariance localization for serial ensemble square root filters. *Mon. Wea. Rev.*, **142**, 4499–4518, https://doi.org/10.1175/MWR-D-13-00390.1.