U.S. Department of Commerce

National Oceanic and Atmospheric Administration

National Weather Service

National Centers for Environmental Prediction

5830 University Research Court

College Park, MD 20740-3818

Office Note 515

https://doi.org/10.25923/pytq-6575

Impacts of Doppler Radial Wind Assimilation in the GFS with a Global Observing System Simulation Experiment

Donald E. Lippi, Jacob R. Carley, Daryl T. Kleist

NOAA/NWS/NCEP Environmental Modeling Center

College Park, Maryland

November 2023

E-mail: Donald.E.Lippi@noaa.gov

ABSTRACT

Doppler radial winds have been an underutilized observation in U.S. operational forecast systems. This has typically been owing to limitations in the formulation of the observation operator, the amount of data thinning via super-obbing, the large disparity between observation density and model resolution, or simple exclusion from assimilation in global modeling systems. At the time of this study, there are no global systems that assimilate radial wind observations and, as a result, there have not been any motivation for data sharing agreements of radial wind observations on a global scale. To overcome the lack of global data access, an observing system simulation experiment was used to test radial wind assimilation in the National Centers for Environmental Prediction's Global Forecast System (GFS) version 15. This work is a first attempt to assimilate radial wind observations in the GFS. This effort demonstrates the use and benefits of radial wind assimilation in the GFS, a potentially important observation type in future versions of the GFS and provides evidence to pursue data sharing agreements of radial wind observations.

1. Introduction and Background

More than 800 Doppler radars are registered in the World Meteorological Organization's world radar database (https://wrd.mgm.gov.tr/Home/Wrd). Despite their availability, no radial wind data are assimilated within the National Centers for Environmental Prediction's (NCEP) Global Forecast System (GFS). The assimilation of radial winds has been proven to be beneficial for improving the skill of regional, convection-permitting NWP model forecasts (Gao and Stensrud 2014; Gao et al. 2004; Johnson et al. 2015; Lippi et al. 2019; Xiao et al. 2005), but the value of their assimilation is not well known for global modeling systems (Sun 2005).

Radial winds, which are highly dense and frequent 3D observations, are not typically assimilated in global models like the GFS (e.g., the NCEP GFS; Kleist et al. 2009a) due to the large mismatch between the spatiotemporal resolution of the observations and that of global modeling systems. Until recently, it has been impractical to run global systems at sufficiently high resolution and with a fine enough update frequency to make the best use of radial winds. Therefore, the need to assimilate such observations has not extended beyond those limited area systems for which NWP centers employ high-resolution data assimilation.

Furthermore, assimilation of radial winds has been limited by the lack of a service to collect and distribute the global dataset of radar observations. Unlike most other observing platforms, radial wind observations across the globe are not typically available for assimilation beyond the region from which they originate with the exception of a few regional exchanges. For example, radar information is coordinated among the weather services of European countries via the Operational Program on the Exchange of Weather Radar Information (OPERA; Saltikoff et al. 2019). The U.S. also receives Canadian radars, but ultimately, there is no global exchange of this data.

Year	GFS Version	Dynamic Core	Horizontal Resolution (control/ensemble)	Number of Vertical Levels	DA
2015	GFSv13	Global Spectral Model (GSM)	T1534 (13km)/ T574 (35km)	64	Hybrid 3DEnVar
2016	GFSv14	‘’	‘’	‘’	Hybrid 4DEnVar
2019	GFSv15	Finite-Volume Cubed Sphere (FV3)	C768 (13km)/ C384 (25km)	‘’	‘’

Table 1. Updates made to the operational GFS during 2015–2019.

The operational GFS has recently undergone many upgrades that justify exploring the impacts of assimilating radial wind observations in a global system. Some of the major changes that are discussed are also summarized in Table 1. In 2015, the horizontal resolution was increased to T1534 (~13-km) and T574 (~35-km) for the deterministic GFS and ensemble Kalman Filter members, respectively (NWS 2014). This is a similar grid-spacing to the 12-km North American Mesoscale (NAM) forecast system, which first began assimilating radial winds in 2006. In 2016, the hybrid 3DEnVar data assimilation scheme was upgraded to a hybrid 4DEnVar algorithm; adding a temporal component to the analysis (Kleist and Ide 2015a, 2015b; Lorenc 2003; Wang and Lei 2014; Wu et al. 2017). In 2019, the GFS was

upgraded to version 15 (Kleist et al. 2018; NWS 2019) which included replacing the global spectral model with the non-hydrostatic Finite-Volume Cubed-Sphere (FV3) dynamical core (Harris et al. 2020; Harris and Lin 2013; Lin 2004; Putman and Lin 2007). The GFSv15 upgrade also included the introduction of a more sophisticated microphysics parameterization which expanded the number of prognostic cloud species from one (Zhao and Carr 1997) to five (Chen and Lin 2013; Zhou et al. 2019). During this implementation, the deterministic GFS was maintained at roughly the same resolution but on a cubed sphere grid at C768 (~13-km), with the ensemble resolution upgraded to C384 (~25-km).

These improvements in physics, ensemble resolution, and data assimilation methods merit an evaluation of including higher-resolution sources of observational data. Historically, high spatiotemporal resolution observations in global numerical weather prediction (NWP) have long since been the primary domain of satellite radiances. However, while such observations are dense in time and space, they do not match the capabilities of radar which is able to observe fine scale structures of storms.

Aside from the key impediments to operational global radial wind assimilation such as data density, model resolution, various types of radars (C, S, and X band), and numerous data quality difficulties a major challenge that this work faces is the lack of routinely available Doppler radial wind observations from the global network of radars. Therefore, an observing system simulation experiment (OSSE) methodology (Errico and Privé 2018) is adopted to assess the potential impact of assimilating these observations in a controlled manner. In an OSSE, prospective observations can be simulated from a nature run. For example, using the geographical information in the world radar database, it is possible to create a simulated version of the global radar network to test the potential usefulness of assimilating all radar observations in the GFS.

Keeping in mind the limitations of an OSSE, the goal of this study is to lay the groundwork for assimilating radial wind observations within the GFS by first assessing the impact of assimilating radial wind observations from a network that is accessible today (i.e., the NEXRAD network). This framework is then extended to the potential scenario of having access to all radial wind observations worldwide (i.e., the GLOBAL network). Finally, a purely hypothetical scenario where the worldwide network was designed with highly uniform spatial coverage (i.e., hypothetical network; HYPO for short). This final experiment is designed to provide an upper limit on the impact of assimilating radial winds observations useful for assessing the relative values of the real networks. In each of these experiments, only radial wind observations are assimilated to simplify and isolate the impact of such observations.

In Section 2, the model and data assimilation systems are described along with a description of the OSSE design. The results are presented in Section 3 and conclusions with a short summary and discussion of future work are presented in Section 4.

2. Methods

a. Model and Data Assimilation Configuration

Version 15 of the GFS (GFSv15) was used as the basis for this study, allowing for some modifications to conserve both computational resources and improve model stability. For computational resources, the maximum forecast length was reduced from the operational 384-hour to 168-hour (1-week) forecast. For stability, the model time step was reduced (from

225s to 90s) to overcome imbalances associated with the assimilation of highly dense simulated radial wind observations. The same 90s time step was used for all experiments; the longer time step (225s) was used only in the Nature run. No other modifications were made relative to the operational configuration of GFSv15. The system maintained the standard 6-hourly data assimilation cycle using the hybrid 4DEnVar framework issuing forecasts four times per day at the usual synoptic times of 0000, 0600, 1200, and 1800 UTC respectively. Each experiment was cycled over a 1-week period.

The data assimilation experiments begin 6 hours into the nature run (after the initial spin up period described in Section 2.b.2) and are performed using the Global Data Assimilation System (GDAS) cycling algorithm and the Gridpoint Statistical Interpolation system (GSI; Kleist et al. 2009a; Wu et al. 2002). The GSI is a variational analysis system formulated in model grid space and is used in many operational applications at NCEP, such as the North American Mesoscale Forecast System (Gustafsson et al. 2018; Rogers et al. 2017; Wu et al. 2017), the GFS (Kleist et al. 2009a), the Rapid and High-Resolution Rapid Refresh systems (Benjamin et al. 2016; Dowell et al. 2022; James et al. 2022), and the Real-Time Mesoscale Analysis system (De Pondeca et al. 2011).

The GSI is configured to be consistent with the GFSv15 which uses hybrid 4DEnVar (Kleist and Ide 2015b; Lorenc 2013; Wang et al. 2013; Wu et al. 2017) where the static covariance is implicitly blended with an ensemble covariance through the extended control variable method (Lorenc 2003; Wang 2010). An 80-member ensemble with dual resolution capabilities was used (C768 and C384 for the deterministic and ensemble resolution respectively) and updated from a serial square root filter form of the EnKF (Whitaker and Hamill 2002). The weighting between the static and the flow-dependent ensemble background error covariance (BEC) is set to be 12.5% and 87.5% respectively. The static BEC is identical to that used in the operational GFSv15. The tangent linear normal mode constraint (TLNMC; Kleist et al. 2009b), implemented to help improve the balance of the initial conditions, is enabled during the GSI minimization for each of the data assimilation experiments. All the previously mentioned settings and parameters are consistent with the GSI configuration for GFSv15.

b. Observing System Simulation Experiment (OSSEs)

Because radial wind observations from the existing global network of radars are not currently available for this study, an OSSE methodology was used. Furthermore, the use of an OSSE methodology also allows the possibility to test an idealized global network.

1) OSSE Design

The OSSE was designed to investigate the potential impact of assimilating radial wind observations from three configurations: the U.S. NEXRAD network, a global radar network based upon the known worldwide distribution of radars, and a hypothetical global radar network (Fig. 1).

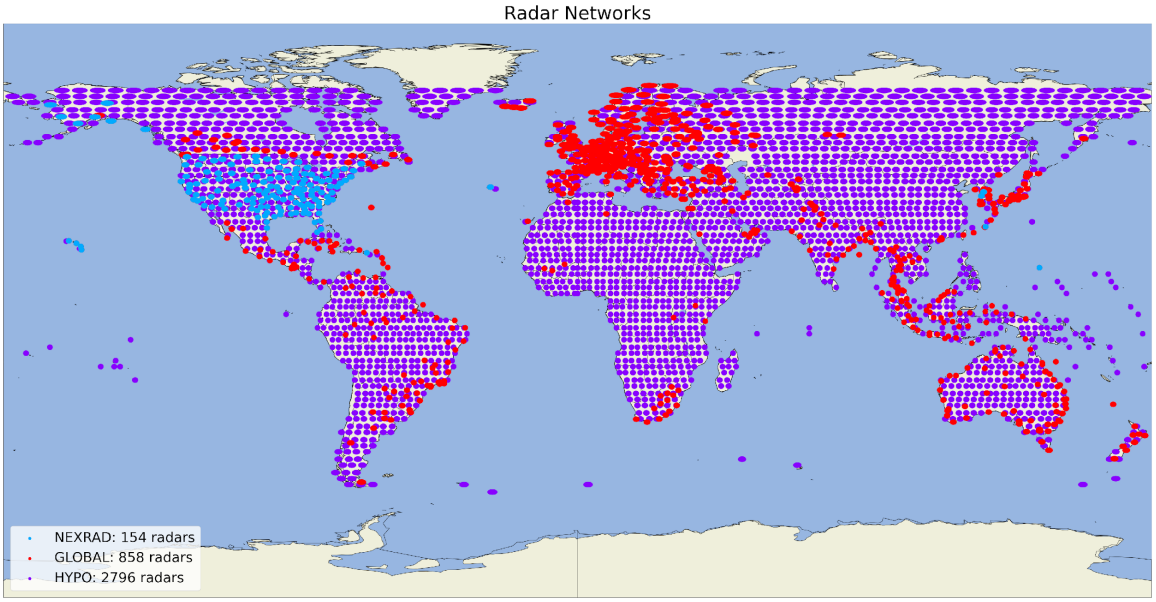

Fig. 1. Radar networks with 100-km radius circles including NEXRAD (blue), GLOBAL (red), and HYPO (purple). Note that the GLOBAL also includes all the NEXRAD radars.

For an OSSE (Errico and Privé 2018), a reference state or nature run is first generated by making a climatologically sound, free-running simulation using a reliable atmospheric model; the nature run is considered as a proxy truth against which subsequent assimilation experiments will be verified (e.g., Tong and Xue 2005; Xue et al. 2001). Simulated, imperfect observations from the observing systems are then generated from the nature run using the radial wind observation operator from the assimilation and adding random Gaussian, unbiased errors (e.g., Gao et al. 2004; Miller and Sun 2003; Xu and Gong 2003). These observations are then used in the data assimilation experiments where the impact of the simulated observations on the analyses and forecasts are assessed.

For simplicity, the same model is used for both the nature run and the experiments. While there are certain limitations to consider while using this identical twin approach, it is not uncommon practice (e.g., Gao et al. 2004; Tong and Xue 2005; Xue et al. 2006; Xue et al. 2001) and can be useful if results are interpreted with the correct perspective. Therefore, one should consider the results from such an idealized design as a “best-case scenario” and a demonstration of potential plausibility.

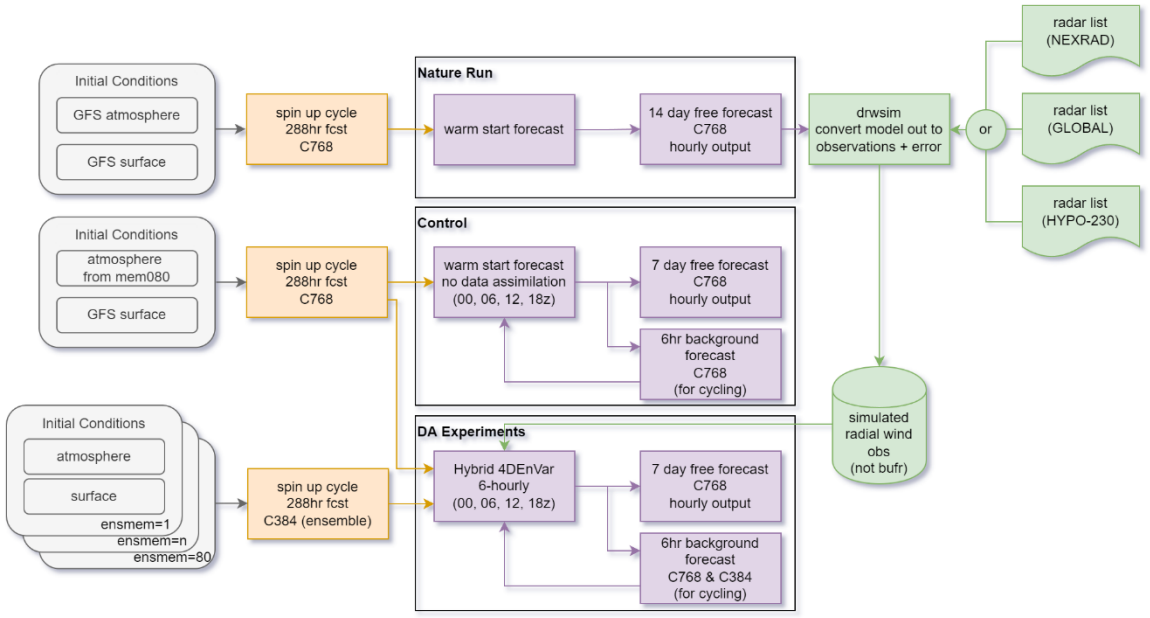

Fig. 2. Diagram describing the OSSE setup and the spin-up of initial conditions for the nature run and experiments. The spin-up cycle is highlighted in orange and represents the initial conditions after running the control and ensemble state for 288h. The forecast and data assimilation steps are shown in purple. The observation simulator and where those observations are used are highlighted in green.

2) Initial Conditions and Model Spin-Up

The model-spin up process is shown in the OSSE flowchart (Fig. 2). The perturbed initial conditions used for the control and all other experiments were generated by replacing the atmospheric state (lower boundary states were not replaced) with that of an arbitrary ensemble member from the operational GFS EnKF. In this case, the 80th member was used. The initial conditions for the nature run used the GFS control member atmospheric state.

Each set of initial conditions were then used to initialize a model free-forecast to determine the point at which the root mean squared difference (RMSD) of temperature, specific humidity, and 500-hPa horizontal wind (Fig. 3) between the two forecasts (the nature run and perturbed initial conditions) saturates (i.e., when the RMSDs no longer continue to grow). This saturation criteria identifies the point at which the model has tracked to its climatology so that both simulations are in a balanced state. The RMSD of 500-hPa temperature, specific humidity, and horizontal wind saturate at approximately 288-hours for both C384 and C768 model resolutions. The 288-hour model state was then used as initial conditions for the OSSE experiments.

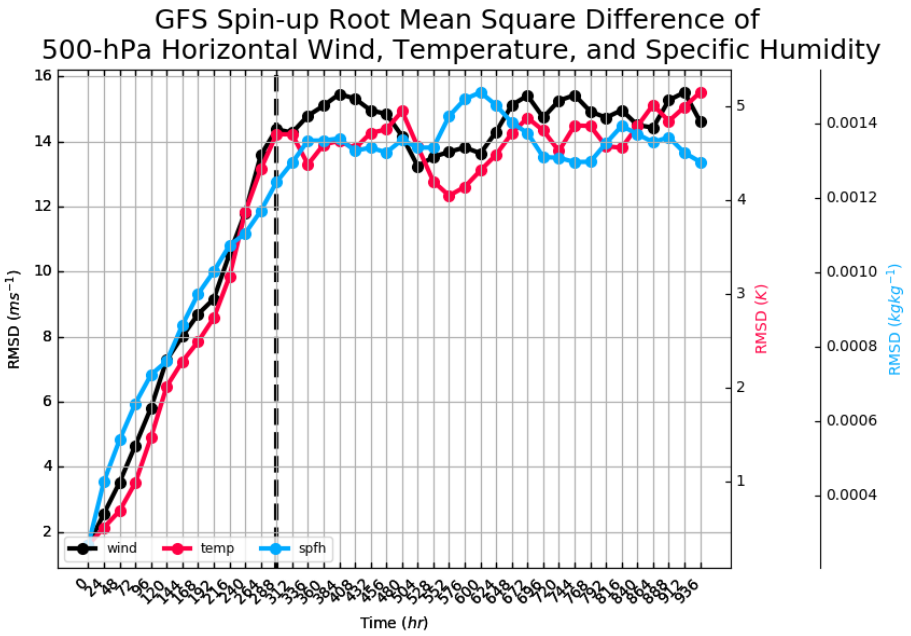

Fig. 3. Root mean square difference (RMSD) for 500-hPa wind (black), temperature (red), and specific humidity (blue) for model resolution C384 for the period starting at 0000 UTC on 11 September 2018. Similar results were found with C768 resolution (not shown). Forecast hour 288 is highlighted by the vertical, black dashed line.

3) The Nature Run

The nature run is a 336-hour, deterministic, free-running simulation (starting after the 288-hour spin up period; Fig. 2) of the GFSv15, absent of any data assimilation, run at C768 (13-km) and 64 vertical levels with hourly output for the 1-week period. Reproducing actual weather events is not the goal in this study as might be the case in some OSSE based studies (e.g., Ge et al. 2012). The initial spin up cycle was initialized with an actual GFS analysis as a convenient approach to get a realistic initial model state.

4) Simulation of Doppler Radial Wind Observations

A Doppler radial wind observation simulator[1] was developed based on the GSI radial wind observation operator, and is represented in Fig. 2 via the green box labeled “drwsim”. The two main inputs are the nature run forecasts and a radar list containing location data for each radar network. For simplicity, all radars are considered to operate similarly to the radars of the NEXRAD network (e.g., S-band, 250-m gates, 1° azimuth; OFCM 2017 p. 2-22 V#99) but also imposed a maximum observation range (R_{max}) of 100-km which is consistent with convention for radial wind pre-processing for assimilation purposes (Alpert and Kumar 2007; Lippi et al. 2019). We chose 100-km under the assumption that all radars operate in a scan mode that uses a higher pulse repetition frequency, which yields a larger Nyquist velocity, yet a lower maximum unambiguous range. Thus, only data within the 100-km range is used to match real-world application of this data, which is meant to avoid the decay of signal quality near the edge of a radial beam.

[1] https://github.com/delippi/drwsim

To obtain realistic simulated radial wind observations, observations were created only in locations where radar reflectivity exceeded a minimum threshold of 10 dBZ. This imposes a requirement that scatterers of sufficient size be present. If using the Marshall Palmer Z-R relationship (Marshall and Palmer 1948)

$$Z = 200R^{8/5} \quad (1)$$

where Z is the reflectivity factor ($mm^6\ m^{-1}$) and R is the rain rate ($mm\ h^{-1}$) and the conversion for Z to obtain the logarithmic reflectivity L_z (dBZ) is,

$$L_z = 10log_{10}(Z), \quad (2)$$

then the relationship between dBZ and rain rate can be estimated as,

$$R = \left(\frac{10^{\frac{L_z}{10}}}{200}\right)^{\frac{5}{8}} \quad (3)$$

A reflectivity value of 10 dBZ approximately corresponds to a light mist with a rain rate of <0.01 in. h^{-1}. Furthermore, studies such as Tong and Xue (2005), Xue et al. (2006), and Liu et al. (2020) have also used a 10 dBZ minimum threshold for detecting regions of precipitation for simulating radar observations. Therefore, in this study, a threshold of 10 dBZ is used as a proxy for determining regions with precipitation.

Radars have different scanning patterns (Volume Coverage Pattern; OFCM 2017) depending on the weather regime. The VCP 212 (OFCM 2017) is the typical radar scan pattern when precipitating regions exist. In our study, we assume all radars are operating in this mode, have Doppler capabilities, and operate at the S-band frequency such that signal attenuation can be ignored. The scan elevation angles of VCP 212 include 0.5, 0.9, 1.3, 1.8, 2.4, 3.1, 4.0, 5.1, 6.4, 8.0, 10.0, 12.5, 15.6, and 19.5 degrees. Clear air techniques for ground-based radar are not considered in this study.

The radial wind observations are simulated using the following equation:

$$V_r(\theta, \alpha) = u\,cos(\theta)\,cos(\alpha) + v\,sin(\theta)\,cos(\alpha) + w\,sin(\alpha) + R(0, 1) \quad (4)$$

where $\mathbf{V_r}$ is the radial wind observation, u and v are the nature run horizontal wind components, w represents the vertical wind, θ is 90° minus the azimuth angle of the radar, and α is the elevation (or tilt) angle of the radar. The formulation here uses the azimuthal directions based on the unit circle rather than the Cardinal Directions (i.e., 0° is East). The last term in Eq. (4), $R(0,\ 1)$, represents random, uncorrelated error drawn from a Gaussian distribution with a mean of 0 m s^{-1} and standard deviation of 1 m s^{-1}, which follows prior OSSE studies (Gao et al. 2004; Liu et al. 2020; Snyder and Zhang 2003; Tong and Xue 2005; Xue et al. 2006). These are errors that may arise from representation errors (coarse model vs. high resolution observations and an imperfect observation operator) and measurement errors. In the case of this OSSE study, the radial wind observations are simulated from the nature run, which is an identical model and thus there are no errors of representation or of the observation operator; however, instrument measurement errors, are included by adding

Gaussian, unbiased errors to the observations. For simplicity, observation errors are assumed uncorrelated.

Radial wind observations were then simulated using the settings found in Table 2. Simulated observations are constrained to the spatiotemporal resolution of the nature run; therefore, each simulated observation is thinned based on a prescribed three-dimensional box. The parameters include: azimuthal width ($\Delta\theta$), the elevation angle width ($\Delta\alpha$), and the gate spacing length (Δr) which describe the width, height, and length of the box respectively. The limiting parameters are the maximum elevation angle (α_{max}) and maximum range (R_{max}) which put a maximum threshold for the elevation angle and maximum distance from the radar, respectively.

Azimuth range (degrees)	**Elevation angle width (degrees)**	**Gate spacing (meters)**	**Min/Max elevation angle (degrees)**	**Max observation range (meters)**	**Ob error mean (m s^{-1})**	**Ob error standard deviation (m s^{-1})**	**Minimum dBz threshold (dBz)**	**VCP**
$\Delta\theta$	$\Delta\propto$	Δr	$\alpha_{max}/\alpha_{min}$	R_{max}	μ	σ	dBz_{min}	
1	0.5	13,000	0.5/20	100,000	0.0	1.0	10	212

Table 2. List of parameters used for simulating the radial observations. Azimuth range, radial range, min/max elevation angle, and max observations range are as described in Lippi et al. (2019) and as used in the GSI. The remaining parameters are specific to the observation simulator.

The horizontal 2D spatial geometry for these observations was specified to be 1° by 13-km. The 13-km radial range was the finest resolution afforded by the nature run. The 1° azimuthal width is consistent with the configuration of the radial wind products of the NEXRAD network.

c. Experimental Design

To test the impact of assimilating Doppler radar radial winds, three DA experiments (see Fig. 2) were designed in addition to a control simulation (see Fig. 2) that does not assimilate any observations (NODA). Each of these is explained below and summarized in Table 3:

1. **NEXRAD**: A data assimilation experiment to assess the impact of assimilating radial wind observations from the only network that is currently accessible to the GFS, the NEXRAD radar network (Fig. 1; blue).
2. **GLOBAL**: A data assimilation experiment to assess the impact of assimilating radial wind observations from the global network of radars as if they were accessible to the GFS, the global network of currently deployed radars (Fig. 1; red). This also includes the NEXRAD network over the U.S.
3. **HYPO**: A data assimilation experiment to assess the impact of assimilating radial wind observations from a hypothetical, optimally designed global network of radars (Fig. 1; purple) based on homogeneous spacing matching that of the average spacing for the NEXRAD network, which is 230-km (Huber and Trapp 2009). This experiment, while highly idealized, is designed to contextualize the results from the

NEXRAD and GLOBAL experiments by providing an estimate for an upper limit on the impact of radial wind assimilation.

Exp	ICs	Res.	Length (hours)	Cycled (hours)	DA	Ensemble	Time step (seconds) (C768/C384)
Nature	GFS	C768	336	N/A	N/A	N/A	225
NODA	Perturbed GFS	C768	168	6	None	N/A	90
NEXRAD	NODA	C768	168	6	Hybrid 4DEnVar	80 (C384)	90/150
GLOBAL	NODA	C768	168	6	Hybrid 4DEnVar	80 (C384)	90/150
HYPO	NODA	C768	168	6	Hybrid 4DEnVar	80 (C384)	90/150

Table 3. List of GFSv15 radial wind OSSE experiments and their configurations.

d. Verification

Verification is performed using the Model Evaluation Tools (MET) verification software (Newman et al. 2022). The impacts of assimilating radial wind observations from the various network configurations are objectively assessed by calculating the fraction skill score (FSS) and frequency bias (FBIAS) of forecast quantitative precipitation as well as the bias corrected root mean square error (BCRMSE) and mean error (bias) of geopotential height, temperature, horizontal wind, and surface pressure forecasts. The verification statistics were computed on a common, 0.125° grid over the following six verification domains: a) Northern Hemisphere (NHM), b) Tropics (TRP), c) Southern Hemisphere (SHM), d) Contiguous U.S. (CONUS), e) Atlantic Ocean (ATL), and f) Europe (EUR) (Fig. 4). All variables and metrics were evaluated via scorecards (not shown) for each verification domain; the major results are summarized into series plots. A bootstrapping technique, using 2000 replications with replacement, was used to test for statistical significance at the 95% confidence interval for series plots.

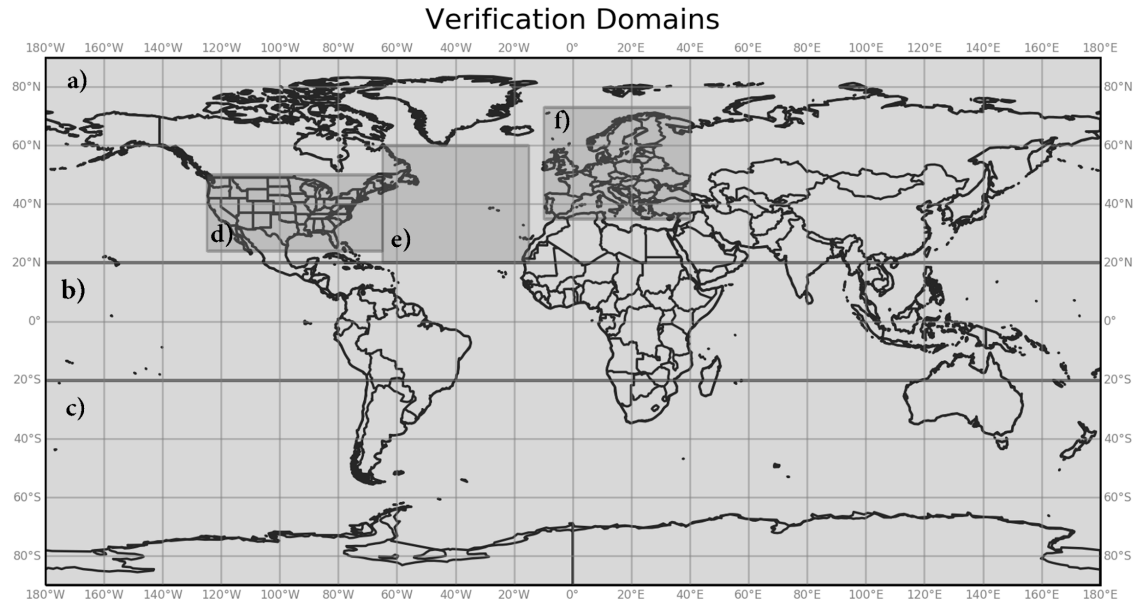

Fig. 4. Verification domains denoted by the gray shaded areas: a) Northern Hemisphere (NHM), b) Tropics (TRP), c) Southern Hemisphere (SHM), d) Continuous U.S. (CONUS), e) Atlantic Ocean (ATL), and f) Europe (EUR).

FBIAS is a categorical metric and is computed as the ratio of "yes" forecasts to "yes" observations. FBIAS greater than 1 indicates more points predicting a given threshold compared to observations (a wet bias). FBIAS less than 1 indicates fewer points predicting a given threshold compared to observations (a dry bias). FBIAS is calculated in the following manner

$$FBIAS = \frac{hits + false\ alarms}{hits + misses} \tag{5}$$

		Observed	
		Yes	No
Forecast	Yes	*hit*	*false alarm*
	No	*miss*	*correct rejection*

Fig. 5. 2x2 contingency table which shows the relationship between events for a dichotomous verification situation (Wilks 2011).

where each of the terms in Eq. (5) are described in Fig. 5. A *hit* corresponds to a scenario where the forecast correctly matches the observations. A *miss* corresponds to a "no" forecast event occurred where there was an observed event. A *false alarm* corresponds to an event that

was forecast but did not occur in reality. A *correct rejection* corresponds to an event that was correctly forecast to not occur.

The FSS is a neighborhood verification approach that relaxes the requirement for forecast and observed events to match exactly at the grid scale. Instead, the fractional coverage of predicted and observed grid-point events above a specified threshold are compared over a range of increasingly large box widths (Roberts and Lean 2008). Thus, FSS reveals how well the forecast resembles the observations at a given spatial scale. The FSS is computed in the following manner.

$$FSS = 1 - \frac{\frac{1}{N}\sum_{N}\left(P_f - P_o\right)^2}{\frac{1}{N}\left[\sum_{N}P_f^2+\sum_{N}P_o^2\right]} = 1 - \frac{FBS}{FBS_{worst}} \qquad (6)$$

where N is the number of grid-points contained within the neighborhood area; P_f is the fractional coverage of forecast events that exceed a pre-determined threshold; and P_o is the fractional coverage of observed events that exceed the threshold. The fractional values (P_f and P_o) are what is computed at the increasingly large box widths. The FSS ranges from 0 to 1 where 0 would be a complete mismatch and 1 would be a perfect match of forecast events to observed events. FSS at a constant box width can also be used to verify the forecast similarly to the Equitable Threat Score (ETS or Gilbert's Skill Score; Wilks 2011).

Bias corrected root mean squared error (BCRMSE) is the standard deviation of the forecast errors that is not accounted for by the bias (i.e, mean error) and is computed in the following manner in the MET verification software by taking the square root of the bias corrected mean squared error (BCMSE)

$$BCRMSE = \sqrt{BCMSE} = \sqrt{RMSE^2 - bias^2} = \overline{(f - t)^2} - \overline{f - t}^2 \qquad (7)$$

where f is the experimental forecast, t is the truth from the nature run and the over bars denote an arithmetic mean.

3. Results

To determine how well each experiment performed in this OSSE, we consider the FSS and FBIAS of forecast precipitation as well as the BCRMSE and bias of 500-hPa geopotential heights. Other fields including wind, temperature, and surface pressure were also verified and generally showed similar results. For simplicity, those results are not shown.

a. FSS and FBIAS of Precipitation

The FSS of precipitation over increasingly longer lead times (Fig. 6) shows that the benefits of assimilating radial winds can persist for several days in this OSSE, especially for the experiments with networks with broader coverage (e.g., ~5 days over the Northern Hemisphere for the HYPO and GLOBAL experiments). Since the impacts are strongest and

clearest at shorter lead times, the day 1 forecast (i.e., average of the 1–24-h forecast) results will be the focus of the remainder of this study.

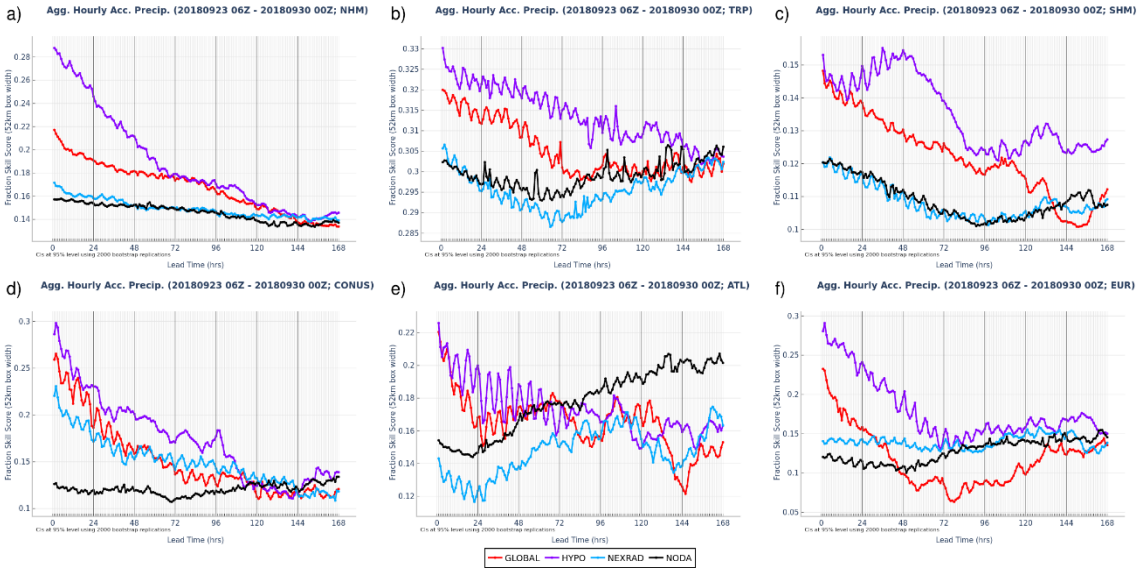

Fig. 6. FSS at 52km box width of the hourly-accumulated precipitation as a function of forecast lead time over the 5 verification domains (Fig. 4). The vertical black lines indicate 24-hour periods.

The FSS difference (relative to NODA) of day 1 aggregated hourly-accumulated precipitation at 52-km box width over each of the verification domains are shown in Fig. 7 and the corresponding FBIAS in Fig. 8. Fig. 7 reveals that the assimilation of only NEXRAD is modestly helpful in the Northern Hemisphere especially over the CONUS and Europe; however, it is not necessarily universally helpful such as over the Atlantic Ocean (Fig. 7e). This suggests that network design and spatial uniformity have downstream impacts – a topic which is explored more completely in Section 3c. Furthermore, the NEXRAD experiment was the only experiment that did *not* produce statistically significant improvements over the Southern Hemisphere and in the Tropics, which is consistent with the fact that there are no NEXRAD radars in those regions nor are there radars upstream from those regions. It also shows that the GLOBAL experiment has significant improvements over the NEXRAD for all domains. Therefore, future implementation of radial wind assimilation in a global system might be met with greater success if the full existing global network of radars is used instead of those just over the U.S. Finally, the GLOBAL often underperforms compared to the HYPO experiment by as much as a 24% difference at short forecast leads over the Northern Hemisphere (Fig. 7a). The HYPO was expected to perform better since the network was designed with greater spatial uniformity and coverage and therefore shows that the existing global network of radars could also be further enhanced providing additional benefit.

The FBIAS of day 1 (1–24-h) precipitation is shown in Fig. 8. In the context of an identical twin OSSE (although a reduced time step was used), it would be expected that there is no bias in the FBIAS; however, this is not the case. This discrepancy is thought to be due to initial imbalances within the assimilation that manifest more prominently in the shorter forecast leads (e.g., day 1 FBIAS scores). Recall that the time step was reduced for the data assimilation experiments due to imbalances in the model. Reducing the timestep did not

remove those initial imbalances but allowed the model to better adjust to those imbalances without causing numerical instability in the model. Over longer forecast leads and larger verification domains the FBIAS is relatively unbiased (not shown). At the shorter forecast leads, the lower FBIAS thresholds generally have good agreement with the truth.

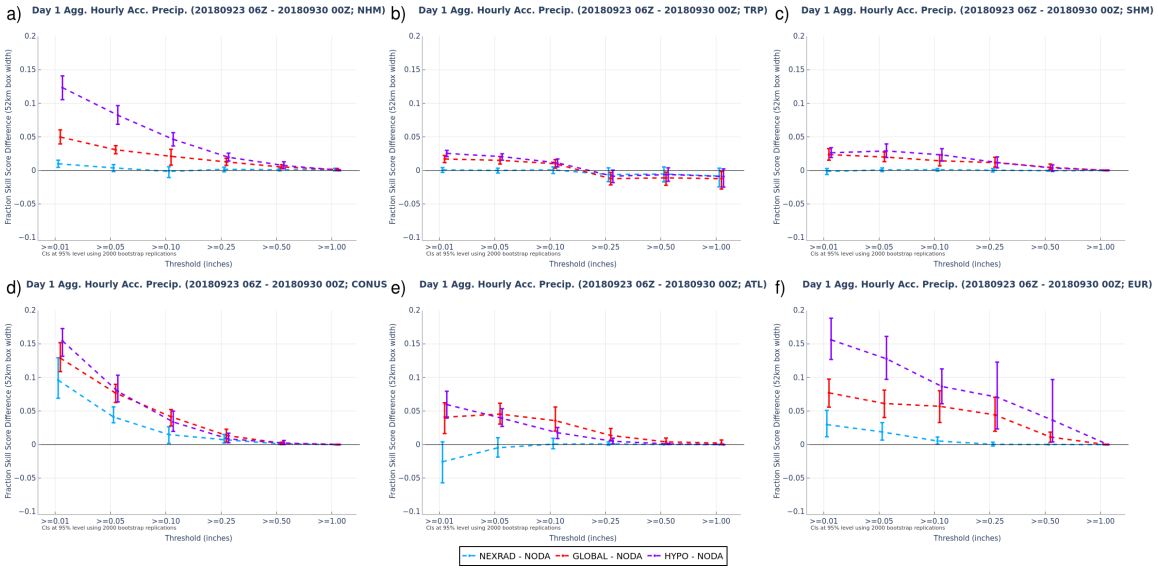

Fig. 7. FSS difference (relative to NODA) of Day 1 aggregated hourly-accumulated precipitation at 52-km box sizes over the 1-week cycled period over the 5 verification domains (Fig. 4). FSS difference scores greater than zero indicate the respective experiment performed better than NODA.

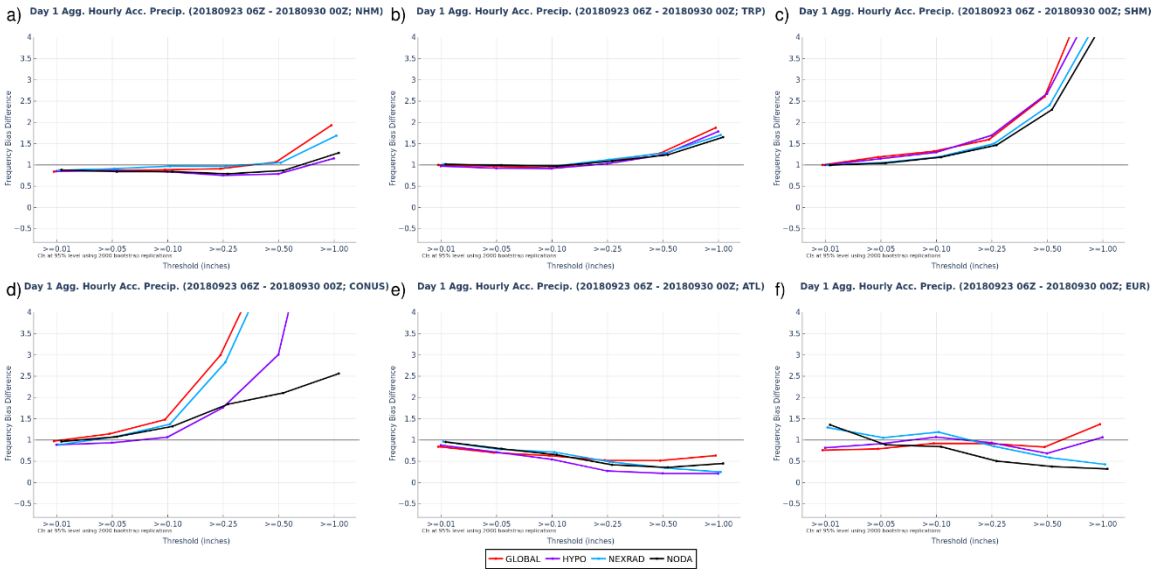

Fig. 8. FBIAS of Day 1 aggregated hourly accumulated precipitation over the 1-week cycled period over the 5 verification domains (Fig. 4).

b. BCRMSE of 500hPa Geopotential Height

To determine the potential impacts of assimilating observations from the different radar networks, the day 1 500-hPa geopotential height BCRMSE (Fig. 9) were compared as a function of cycle time (i.e., forecast initialization time).

Over the Northern Hemisphere (Fig. 9a), the NEXRAD experiment has little change relative to the NODA while the GLOBAL and HYPO experiments have a significant reduction in BCRMSE in 500-hPa geopotential height, especially during the later cycles. The improvements for the GLOBAL and HYPO experiments can be attributed to their larger area of observational coverage relative to NEXRAD. Fig. 10 shows the observation counts (Total, Northern Hemisphere, and Southern Hemisphere) for each of the different experiments as a function of assimilation cycle. On average, the NEXRAD, GLOBAL, and HYPO experiments assimilate roughly 1×10^6, 3.5×10^6, and 15×10^6 observations per assimilation cycle respectively with most of those observations coming from the Northern Hemisphere. To reiterate, it is not only that the GLOBAL and HYPO have more observations, but that those networks have greater spatial coverage.

Over the Tropics (Fig. 9b) there are slight degradations among all experiments (~0–50m). The NEXRAD experiment shows the least amount of change which is consistent with the fact that most NEXRAD radars reside north of the Tropical verification domain and thus should not have much impact on the forecast there. Both GLOBAL and HYPO experiments show larger BCRMSE in the Tropical verification domain than the NODA. This result is currently not well understood, but it is suspected that the limited sampling of the convectively environment of the tropics might be triggering area of convection due to mass convergence/divergence. It would be important to further investigate the impacts of radial wind assimilation in the tropics using real-data experiments.

Over the Southern Hemisphere (Fig. 9c), there is virtually no change for the NEXRAD experiment relative to the control in BCRMSE 500-hPa geopotential height. This is consistent since there are no NEXRAD radar sites in the Southern Hemisphere. There are slight improvements for the GLOBAL and HYPO in the second half of the period; note the observation counts (Fig. 10; dotted lines) for the Southern Hemisphere for each experiment. The relatively smaller impact and lower observation counts in the Southern Hemisphere, relative to what was found in the Northern Hemisphere, is due to the smaller percentage of landmass available for ground-based observing networks (see Fig. 1).

Over the CONUS (Fig. 9d), all experiments have similar performance with small improvements (less than about 50-m height differences) prior to 0600 UTC 26 September 2018 and larger improvements begin after this period (100–150-m height differences). It is expected that each data assimilation experiment performed with similar skill over the CONUS since the GLOBAL and NEXRAD coverage is identical, and the HYPO network was based on the average spacing of the NEXRAD network but are evenly spaced.

Over the Atlantic Ocean domain (Fig. 9e), there is improvement in the earlier cycles among all experiments but the NEXRAD and GLOBAL exhibit an unexpected, anomalous spike in forecast error peaking around 0600 UTC 29 Sept 2018 (as indicated by the vertical line); a brief discussion will follow on this point in section 3c.

Finally, over Europe, all experiments show some level of improvement. During the early cycles there are comparatively smaller differences (0-25 m height differences) between the NODA and each experiment. It takes approximately 5 to 6 assimilation cycles for improvements to appear in the NEXRAD experiment whereas improvements are immediate and large (~150-m height differences) in the GLOBAL and HYPO experiments. The improvement in the NEXRAD experiment is evidence of upstream DA impacting the time lagged downstream forecast, hence why we observe a time lag of approximately 5 to 6 cycles before improvements are detected.

The positive downstream impacts over Europe and the negative impacts over the north Atlantic could be further investigated by running additional, radial wind only experiments over a longer period to give more robust results as the results from this study are based on a limited sample size.

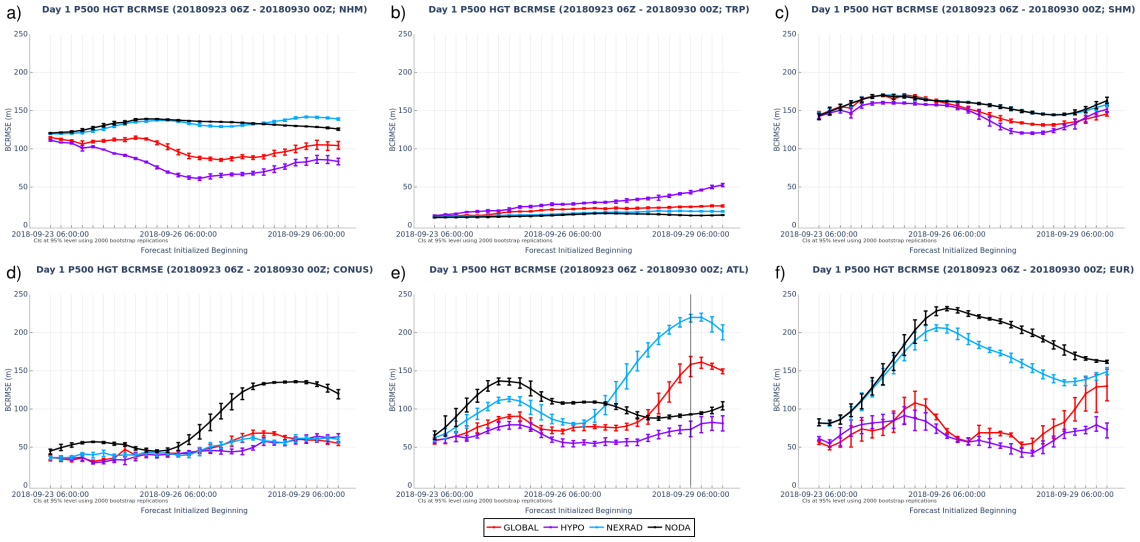

Fig. 9. BCRMSE of day 1 500 hPa geopotential height forecast by forecast initialization time over the 5 verification domains (Fig. 4).

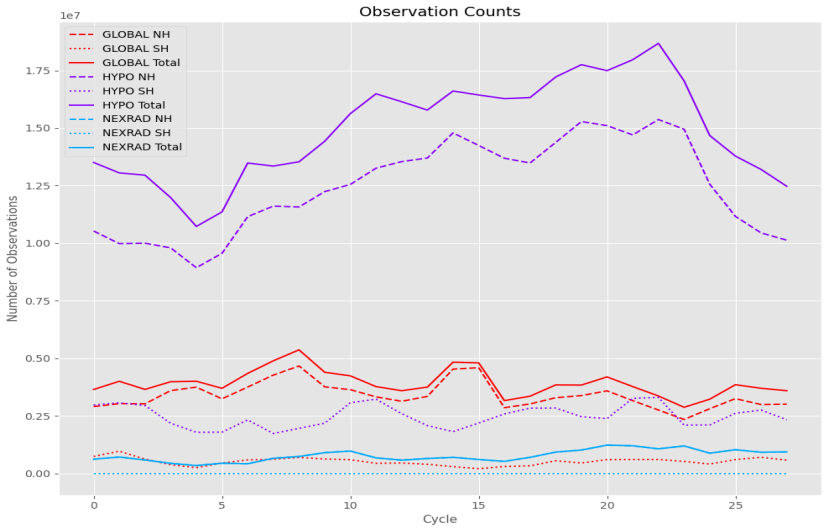

Fig. 10. Observation counts from the NEXRAD (blue), GLOBAL (red), and HYPO (purple) experiments. The counts for each experiment are also given for the Northern Hemisphere (dashed), Southern Hemisphere (dotted), and total (solid) regions.

c. Error Growth Over the Atlantic

The day 1 BCRMSE for the 500-hPa geopotential heights are plotted (Fig. 9e) as a function of forecast initialization time for each of the different experiments over the Atlantic Ocean. The first 18 cycles the NEXRAD experiment, as expected, performed better than the control. By 0600 UTC 29 September 2018, the NEXRAD 500-hPa geopotential height BCRMSE sharply rises to more than 200 m (more than twice that of the control; ~100m). Similarly, in the GLOBAL experiment, the 500-hPa geopotential height BCRMSE also sharply rises to more than 150 m which peaks roughly at the same time to that of the NEXRAD, but the upward trend in error does not occur until 4-5 cycles (24–30hrs) later. The 6-hr geopotential height forecast from the 0600 UTC cycle on 29 September 2018 (the 25th DA cycle; vertical line in Fig. 9e) is shown comparing the NODA (Fig. 11a), NEXRAD (Fig. 11b), GLOBAL (Fig. 11c), and the HYPO (Fig. 11d) experiments all against the NATURE. The NEXRAD and GLOBAL experiments generate a deep upper-level low positioned over the northeast corner of the Atlantic verification domain and noted by the large negative height differences relative to NATURE (blue shading, Fig. 11b, c). Neither the NODA nor the HYPO experiments featured this large error growth.

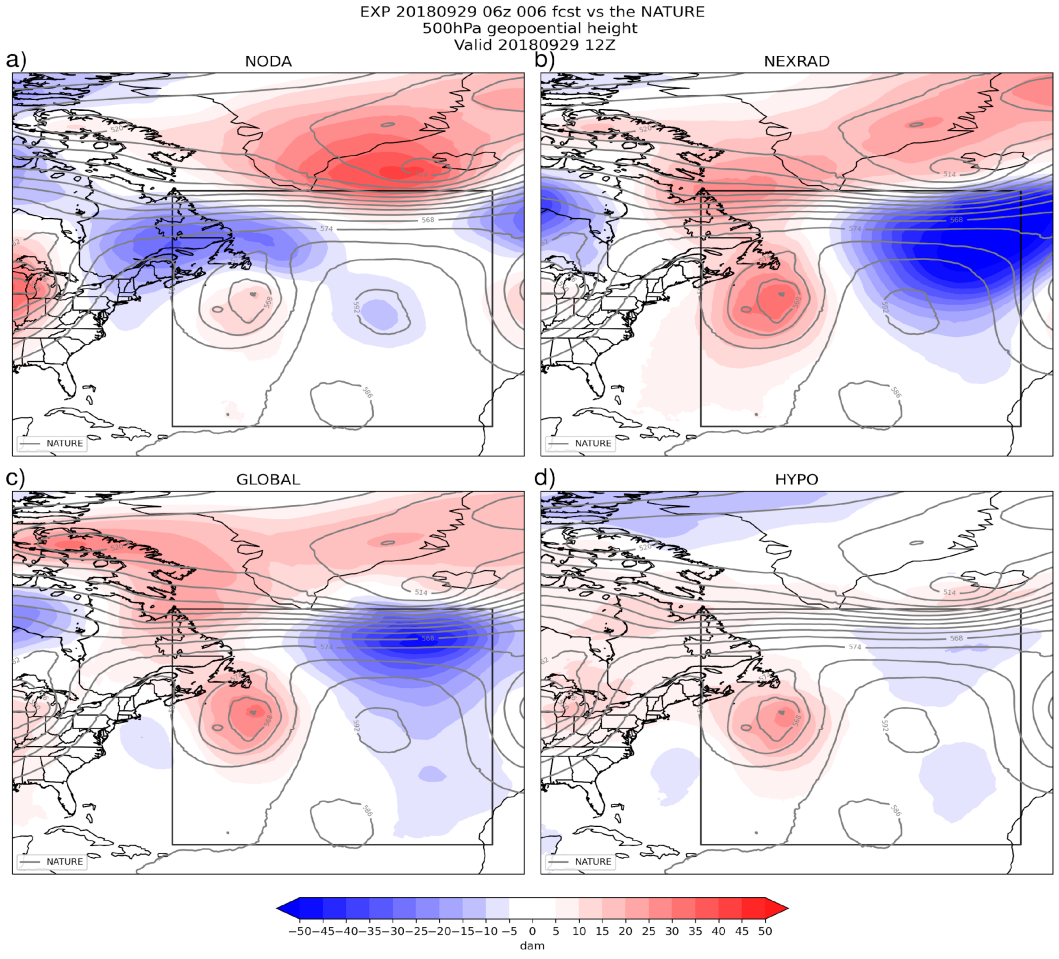

Fig. 11. Comparison of a 6-hr, 500-hPa geopotential height forecast for (a) NODA, (b) NEXRAD, (c) GLOBAL, and (d) HYPO vs. the NATURE (gray contours) from the 25th DA cycle denoted by the vertical line in Fig. 9e. Geopotential height differences are denoted by the red/blue shading where blue indicates where the

respective experiment has lower geopotential heights relative to the NATURE. The Atlantic Ocean verification domain is indicated by the black square.

We hypothesize that the NEXRAD and GLOBAL experiments have this large error growth for the following reasons: 1) the discrete nature of the radar network creates the opportunity for partial sampling to occur, especially along coastlines, and thus the possibility of aliasing of the atmospheric state when these observations are assimilated and 2) they lack the observational coverage that the HYPO network provides (even though HYPO also suffers from the first point).

Aliasing occurs when only a portion of a signal is projected onto a pattern that is not fully observed. This issue of aliasing has also been noted in a similar study (Djalalova et al. 2016) and is analogous to what we have found here. Djalalova et al. (2016) tested the effectiveness of assimilating winds from a network of onshore wind profiling radars (WPRs) located in the Northeastern U.S. In this study, the authors found that assimilating observations from these coastal WPRs generally improved downstream forecasts but had one case that showed clear degradation. Upon further analysis the authors found the degradation was due to aliasing. The similar network configuration and flow regime in our case leads us to hypothesize that we see a similar phenomenon.

Benjamin et al. (2004) also noted that aliasing is a possibility when using any discrete observational network. Therefore, boundaries in the radar (or any observing network) due to natural (i.e., coasts) or cultural (i.e., undeveloped regions with no explicit need for radar observations), can produce areas that are potentially prone to aliasing. What we have found is very similar to that found in Djalalova et al. (2016), having a surface low pressure near the Atlantic Coast of the U.S. with a land confined observing network (of wind). The HYPO network, which also has discrete boundaries, likely avoided the large error over the Atlantic because it has better spatial coverage (e.g., Fig. 1). For this specific case, it is hypothesized that the radars located over Nova Scotia and over Greenland were crucial to the success of the HYPO experiment.

The 10-m wind (black wind barbs) 6-hour forecasts from the NEXRAD (a) and HYPO (b) experiments are compared against the NATURE (c) at the same valid time (Fig. 12). It is at the time shown in which a low initially formed off the mid-Atlantic Coast that led to the large degradation over the Atlantic verification domain. Several variables are shown in each plot and include: the 10-m wind speed (black wind barbs), surface pressure, reflectivity (< 10dBZ), the 100-km radar range rings, and lowest level wind increments (red barbs) with the magnitude also shaded in gray. Superimposed on the NEXRAD and NATURE experiments are the corresponding storm tracks for either scenario (gray line). The HYPO experiment did not exhibit a low forming off the coast at this time.

At the time shown in Fig. 12, the main low in the NATURE is located east of Nova Scotia with a secondary low beginning to develop off the mid-Atlantic coast (as indicated by the gray storm track line; the red dot identifies the center location of the low at the corresponding valid time). The NEXRAD experiment produced low level wind increments of about 10 m s^{-1} in magnitude of offshore flow in Fig. 12a enhancing the cyclonic flow and aiding in the development of this surface low pressure. The low that develops in the

NEXRAD tracks northeastward and northward following the gray line whereas the low in the NATURE tracks more eastward during the same period. Since, in the NATURE, the low tracks out of range of the radars, there are no radars to constrain the currently formed low in the NEXRAD experiment. On the other hand, the HYPO experiment does not generate strong offshore increments, nor does it generate a low off the mid-Atlantic coast. The HYPO is able observe the main low in the NATURE as is traverses Nova Scotia noticeable by the large increments in Fig. 12b.

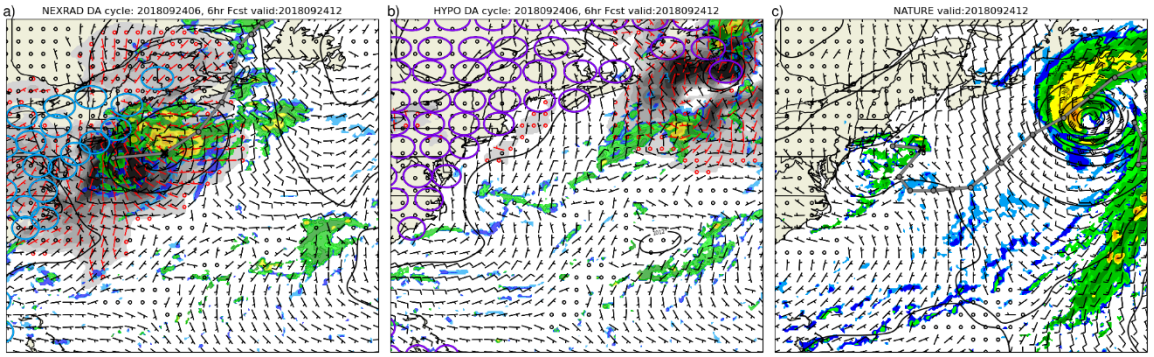

Fig. 12. A comparison of the 6-hour 10m wind forecast (black barbs) for the NEXRAD (a) and HYPO (b) against the NATURE (c) at a time when the error growth begins in the NEXRAD experiment. The 100-km radar rings for the NEXRAD and HYPO networks are both shown. In addition, the DA experiments have wind increments (red wind barbs) with the magnitude of the wind speed increment shaded in gray. Storm tracks (gray lines) are superimposed on the NEXRAD and NATURE experiment noting the location of the storm forming off the mid-Atlantic Coast. The current storm location is noted by a black X.

Ultimately, spatial design of the observing network is important and brings into view the deficiencies in our data assimilation. For example, a larger ensemble could be one element in our data assimilation system that may allow for better utilization of observations of partially sampled features. However, in real-data studies, the full suite of observations will be available and would likely be less susceptible to such issues but is an important issue to document. Furthermore, we suspect that these issues will not be fully mitigated by the use of the full suite of observations, and hence, it was a useful exercise to perform an OSSE experiment testing the assimilation of only radial wind data that might have been obscured by other observations.

4. Summary and Conclusion and Future Work

This work is motivated by several GFS upgrades over recent years that warrant the consideration for using radial winds within the global paradigm. Some of these improvements include upgrades in DA methods, replacing dynamic cores, and increasing model resolution. While the data from the global network of radars currently exists, it is not currently distributed and accessible for use by any individual forecast system. Therefore, an OSSE was used to evaluate the potential impacts of assimilating Doppler radial winds from the global network of radars (as opposed to only the NEXRAD network) in the GFS.

Six verification domains were created, including the Northern Hemisphere, Tropics, Southern Hemisphere, CONUS, Atlantic Ocean, and Europe to test the impacts of assimilating radial winds. The Atlantic Ocean (for all experiments) and European (for only

the NEXRAD experiment) were intended to test the impacts of upstream DA on the downstream forecast. The forecasts were verified, and their relative performance was evaluated using Fraction Skill Score (FSS) and frequency bias (FBIAS) of precipitation as well as the bias corrected root mean squared error (BCRMSE).

Three data assimilation experiments were designed, each using a different radar network. The HYPO experiment used a hypothetical radar network and was used as an upper limit for assessing the real radar networks. Two real radar networks were tested: the NEXRAD which is currently the only network available to the U.S. for operational use and the full GLOBAL network of radars registered within the World Radar Database (https://wrd.mgm.gov.tr/Home/Wrd). A control experiment was also created which did not assimilate any observations (NODA).

The HYPO experiment performed the best and generally set a clear upper limit to the benefits that could be potentially gained from assimilating radial wind observations. A global average of day 1 accumulated precipitation FSS scores showed that the NEXRAD performed about 1% better than NODA, the GLOBAL experiment performed 13% better than NEXRAD, and the HYPO experiment performed 13% better than the GLOBAL experiment. These relative performances will change depending on the verification region. There were clear improvements for the NEXRAD data, especially over the CONUS but also over the European verification domain. While there was not much difference between the NEXRAD and GLOBAL experiments over the CONUS verification domain, there is evidence that upstream DA impacts the downstream forecast. In both cases of the real radar networks, there was downstream degradation of the forecast over the Atlantic Ocean verification domain; however, the severity of this degradation is not expected when using the full suite of observations. Additional testing would be needed to assess the relative magnitude of which radar locations (outside the NEXRAD network) would make the greatest impacts on the downstream forecast over the CONUS. For U.S. landfalling tropical cyclone forecasts, the Caribbean radar network would be likely most useful.

Based on the results of this study, there is support for testing Doppler radar radial wind assimilation in the GFS in real-data experiments. Future, real-data experiments will include testing the assimilation of radial winds using the most recent version of the GFS (i.e., GFSv16; NWS 2021) and using the hourly GFS system workflow (Slivinski et al. 2022). It is hypothesized that the hourly GFS could make better use of observations, especially high spatiotemporal observations like radial winds in global NWP, by updating with observations more frequently such as is done in convective-scale NWP models. It is hypothesized that an hourly-updating global system would be able to better constrain rapidly evolving systems such as hurricanes and convective storms and make better use of high temporal observations including observations such as radial winds.

Acknowledgments.

This work was completed as part of the first author's PhD dissertation requirements at the University of Maryland, College Park and was supported by IMSG, Lynker, and NOAA/EMC. Portions of this research was also supported by Further Additional Supplemental Appropriations for Disaster Relief Requirements Act, 2018, Pub. L. 115-123, div. B, subdiv. 1, Feb. 9, 2018, 132 Stat. 65, 70. The Act allocated the National Oceanic and

Atmospheric Administration (NOAA) $50 million to improve weather forecasting, hurricane intensity forecasting, and flood forecasting and mitigation capabilities, including data assimilation from ocean observing platforms and satellites. Computational resources were provided by the NOAA Research and Development High Performance Computing and Storage Systems and the NOAA/WCOSS (Weather and Climate Operational Supercomputing System). The authors acknowledge the World Meteorological Organization for sharing the meta data for each radar in the World Radar Database (https://wrd.mgm.gov.tr/Home/Wrd). Dr. Alicia Bentley and Dr. Shun Liu are also acknowledged for providing an internal review of an earlier version of this paper.

Data Availability Statement.

Version 15 of the GFS was used along with the modifications already described in the body of this manuscript. Output is too voluminous to host on a public server and is archived behind a firewall on the NOAA High Performance Storage System. For simulating radial wind observations, a radial wind observation simulator was created and is stored in a GitHub repository (https://github.com/delippi/drwsim).

REFERENCES

Alpert, J. C., and V. K. Kumar, 2007: Radial Wind Super-Obs from the WSR-88D Radars in the NCEP Operational Assimilation System. *Monthly Weather Review*, **135,** 1090-1109.

Benjamin, S. G., and Coauthors, 2004: An Hourly Assimilation–Forecast Cycle: The RUC. *Monthly Weather Review*, **132,** 495-518.

Benjamin, S. G., and Coauthors, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, **144,** 1669-1694.

Chen, J.-H., and S.-J. Lin, 2013: Seasonal Predictions of Tropical Cyclones Using a 25-km-Resolution General Circulation Model. *Journal of Climate*, **26,** 380-398.

De Pondeca, M. S. F. V., and Coauthors, 2011: The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: Current Status and Development. *Weather and Forecasting*, **26,** 593-612.

Djalalova, I. V., and Coauthors, 2016: The POWER Experiment: Impact of Assimilation of a Network of Coastal Wind Profiling Radars on Simulating Offshore Winds in and above the Wind Turbine Layer. *Weather and Forecasting*, **31,** 1071-1091.

Dowell, D. C., and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 1: Motivation and System Description. *Weather and Forecasting*.

Errico, R. M., and N. C. Privé, 2018: *Some general and fundamental requirements for designing observing system simulation experiments (OSSEs).* Vol. 8, WMO Rep. WWRP 2018-, 33 pp.

Gao, J., and D. J. Stensrud, 2014: Some Observing System Simulation Experiments with a Hybrid 3DEnVAR System for Storm-Scale Radar Data Assimilation. *Monthly Weather Review*, **142,** 3326-3346.
Gao, J., M. Xue, K. Brewster, and K. K. Droegemeier, 2004: A Three-Dimensional Variational Data Analysis Method with Recursive Filter for Doppler Radars. *Journal of Atmospheric and Oceanic Technology*, **21,** 457-469.
Ge, G., J. Gao, and M. Xue, 2012: Diagnostic Pressure Equation as a Weak Constraint in a Storm-Scale Three-Dimensional Variational Radar Data Assimilation System. *Journal of Atmospheric and Oceanic Technology*, **29,** 1075-1092.
Gustafsson, N., and Coauthors, 2018: Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, **144,** 1218-1256.
Harris, L., X. Chen, L. Zhou, and J.-H. Chen, 2020: The Nonhydrostatic Solver of the GFDL Finite-Volume Cubed-Sphere Dynamical Core. *NOAA Technical Memorandum OAR GFDL*, **2020-003**.
Harris, L. M., and S.-J. Lin, 2013: A Two-Way Nested Global-Regional Dynamical Core on the Cubed-Sphere Grid. *Monthly Weather Review*, **141,** 283-306.
Huber, M., and J. Trapp, 2009: A Review of NEXRAD Level II: Data, Distribution, and Applications. *Journal of Terrestrial Observation*, **1,** 5-15.
James, E. P., and Coauthors, 2022: The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part 2: Forecast Performance. *Weather and Forecasting*.
Johnson, A., X. Wang, J. R. Carley, L. J. Wicker, and C. Karstens, 2015: A Comparison of Multiscale GSI-Based EnKF and 3DVar Data Assimilation Using Radar and Conventional Observations for Midlatitude Convective-Scale Precipitation Forecasts. *Monthly Weather Review*, **143,** 3087-3108.
Kleist, D. T., and K. Ide, 2015a: An OSSE-Based Evaluation of Hybrid Variational–Ensemble Data Assimilation for the NCEP GFS. Part I: System Description and 3D-Hybrid Results. *Monthly Weather Review*, **143,** 433-451.
——, 2015b: An OSSE-Based Evaluation of Hybrid Variational–Ensemble Data Assimilation for the NCEP GFS. Part II: 4DEnVar and Hybrid Variants. *Monthly Weather Review*, **143,** 452-470.
Kleist, D. T., R. Mahajan, and C. Thomas, 2018: Data assimilation in the Next-Generation Global Prediction System (NGGPS) Era: Initial implementation of FV3-based Global Forecast System (GFS). *JCSDA Quarterly*, **61**.
Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009a: Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather and Forecasting*, **24,** 1691-1705.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, R. M. Errico, and R. Yang, 2009b: Improving Incremental Balance in the GSI 3DVAR Analysis System. *Monthly Weather Review*, **137,** 1046-1060.
Lin, S.-J., 2004: A "Vertically Lagrangian" Finite-Volume Dynamical Core for Global Models. *Monthly Weather Review*, **132,** 2293-2307.
Lippi, D. E., J. R. Carley, and D. T. Kleist, 2019: Improvements to the Assimilation of Doppler Radial Winds for Convection-Permitting Forecasts of a Heavy Rain Event. *Monthly Weather Review*, **147,** 3609-3632.
Liu, C., M. Xue, and R. Kong, 2020: Direct Variational Assimilation of Radar Reflectivity and Radial Velocity Data: Issues with Nonlinear Reflectivity Operator and Solutions. *Monthly Weather Review*, **148,** 1483-1502.
Lorenc, A. C., 2003: The potential of the ensemble Kalman filter for NWP—a comparison with 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, **129,** 3183-3203.
Lorenc, A. C., 2013: Recommended nomenclature for EnVar data assimilation methods. *Research Activities in Atmospheric and Oceanic Modeling*, **5**.
Marshall, J. S., and W. M. K. Palmer, 1948: THE DISTRIBUTION OF RAINDROPS WITH SIZE. *Journal of Atmospheric Sciences*, **5,** 165-166.
Miller, L. J., and J. Sun, 2003: Initialization and forecasting of thunderstorms: Specification of radar measurement errors.
Newman, K., and Coauthors, 2022: The MET Version 10.1.2 User's Guide. Developmental Testbed Center.
NWS, 2014: Technical Implementation Notice 14-46 Corrected.
——, 2019: Service Change Notice 19-40.
——, 2021: Service Change Notice 21-20.
OFCM, 2017: Federal Meteorologist Handbook No. 11 WSR-88D Meteorological Observations Part C. OFCM, Ed.
Putman, W. M., and S.-J. Lin, 2007: Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, **227,** 55-78.
Roberts, N. M., and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Monthly Weather Review*, **136,** 78-97.
Rogers, E., and Coauthors, 2017: Upgrades to the NCEP North American Mesoscale (NAM) System. *Research activities in atmospheric and oceanic modelling, CAS/JSC Working Group on Numerical Experimentation. Report No. 47. WCRP Report No.12/2017. WMO, Geneva*.
Saltikoff, E., and Coauthors, 2019: OPERA the Radar Project. *Atmosphere*, **10,** 320.
Slivinski, L. C., D. E. Lippi, J. S. Whitaker, G. Ge, J. R. Carley, C. R. Alexander, and G. P. Compo, 2022: Overlapping Windows in a Global Hourly Data Assimilation System. *Monthly Weather Review*, **150,** 1317-1334.

Snyder, C., and F. Zhang, 2003: Assimilation of Simulated Doppler Radar Observations with an Ensemble Kalman Filter. *Monthly Weather Review*, **131,** 1663-1677.
Sun, J., 2005: Convective-scale assimilation of radar data: progress and challenges. *Quarterly Journal of the Royal Meteorological Society*, **131,** 3439-3463.
Tong, M., and M. Xue, 2005: Ensemble Kalman Filter Assimilation of Doppler Radar Data with a Compressible Nonhydrostatic Model: OSS Experiments. *Monthly Weather Review*, **133,** 1789-1807.
Wang, X., 2010: Incorporating Ensemble Covariance in the Gridpoint Statistical Interpolation Variational Minimization: A Mathematical Framework. *Monthly Weather Review*, **138,** 2990-2995.
Wang, X., and T. Lei, 2014: GSI-Based Four-Dimensional Ensemble–Variational (4DEnsVar) Data Assimilation: Formulation and Single-Resolution Experiments with Real Data for NCEP Global Forecast System. *Monthly Weather Review*, **142,** 3303-3325.
Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-Based Ensemble–Variational Hybrid Data Assimilation for NCEP Global Forecast System: Single-Resolution Experiments. *Monthly Weather Review*, **141,** 4098-4117.
Whitaker, J. S., and T. M. Hamill, 2002: Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, **130,** 1913-1924.
Wilks, D. S., 2011: *Statistical methods in the atmospheric sciences.* Vol. 100, Academic press.
Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Monthly Weather Review*, **130,** 2905-2916.
Wu, W.-S., D. F. Parrish, E. Rogers, and Y. Lin, 2017: Regional Ensemble–Variational Data Assimilation Using Global Ensemble Forecasts. *Weather and Forecasting*, **32,** 83-96.
Xiao, Q., Y.-H. Kuo, J. Sun, W.-C. Lee, E. Lim, Y.-R. Guo, and D. M. Barker, 2005: Assimilation of Doppler Radar Observations with a Regional 3DVAR System: Impact of Doppler Velocities on Forecasts of a Heavy Rainfall Case. *Journal of Applied Meteorology*, **44,** 768-788.
Xu, Q., and J. Gong, 2003: Background error covariance functions for Doppler radial-wind analysis. *Quarterly Journal of the Royal Meteorological Society*, **129,** 1703-1720.
Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE Framework Based on the Ensemble Square Root Kalman Filter for Evaluating the Impact of Data from Radar Networks on Thunderstorm Analysis and Forecasting. *Journal of Atmospheric and Oceanic Technology*, **23,** 46-66.
Xue, M., and Coauthors, 2001: The Advanced Regional Prediction System (ARPS) – A multi-scale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteorology and Atmospheric Physics*, **76,** 143-165.

Zhao, Q., and F. H. Carr, 1997: A Prognostic Cloud Scheme for Operational NWP Models. *Monthly Weather Review*, **125,** 1931-1953.
Zhou, L., S.-J. Lin, J.-H. Chen, L. M. Harris, X. Chen, and S. L. Rees, 2019: Toward Convective-Scale Prediction within the Next Generation Global Prediction System. *Bulletin of the American Meteorological Society*, **100,** 1225-1243.