

*Report for 2021 DTC Visitor Program*

**Retrospective Experiments with GSI Multiscale EnKF Data  
Assimilation for Convection-Allowing Stand-Alone Regional FV3**

Chong-Chi Tong (cctong@ou.edu)  
Center for Analysis and Prediction of Storms, University of Oklahoma, Norman  
OK 73072

Project Period: 1 April 2021 to 31 March 2022  
On-Site Visit (NCAR Foothills Lab Boulder, CO): 1-29 October to 2021  
DTC Host: Will Mayfield (DTC Data Assimilation Team)

May 2022

## 1. Background and Introduction

Accurate prediction of weather down to the convective storm scales requires initial conditions that accurately represent the atmospheric state at all scales (from the planetary through synoptic, mesoscale to the convective) given the importance of their interactions. Therefore, a well-performing data assimilation (DA) system must accurately analyze flow features at all scales. In this project, a multi-scale DA capability within the GSI-based EnKF system is proposed for the Stand-Alone Regional (SAR) version of the FV3 model that can optimally assimilate both dense convective-scale data such as those of radar, and high-resolution GOES-R observations as well as all other mostly coarser-resolution data. Recently the operational GSI hybrid EnVar system has been adopted to work with SAR FV3 at NCEP, which does not yet have a self-consistent multi-scale EnKF system for FV3. The latter is a prerequisite for an optimal multi-scale hybrid EnVar system because EnKF is needed to provide ensemble perturbations for reliable flow-dependent covariance estimation. For the planned operational use of SAR FV3 for CAM forecasts over CONUS or larger domains, the multi-scale DA issue must be properly addressed.

Two main goals proposed for this project include:

- 1) Develop a GSI-based multi-scale EnKF DA system capable of effectively assimilating all observations, sampling synoptic through convective scales for balanced NWP initial conditions on a continent-sized CAM-resolution (~3 km) grid. Ensemble covariance spatial filtering will be combined with data-type-dependent localization to ensure low-noise balanced analysis increments for synoptic scale (e.g., soundings) through convective-scale (e.g., radar) observations. When assimilating radar observations that sample precipitation regions, narrower localization will be used to retain small-scale details and to allow for close fit to observations while much larger localization will be employed when assimilating coarser-resolution observations. For the latter, a low-pass filter will be applied to the ensemble perturbations to remove noise in covariances at large spatial separations.
- 2) Test the multi-scale DA system coupled with SAR FV3 using retrospective cases, tune and optimize the system configurations, including the filter separation length scale, localization radii, covariance inflation, etc. Sensitivity tests will be conducted to determine the optimal filter separation length scale for the scale selective

filter. Given the sequential nature of EnKF, smoother analysis increments from conventional data will be added to the full 3-km ensemble background before radar data are assimilated using the native-resolution covariances and tight localization.

During the DTC Visitor Program period, the author was able to collect all the data required for the planned retrospective experiments under assistance of the DTC Data Assimilation team. Valuable input toward the work were obtained through the regular weekly meeting with DTC members Drs. Ming Hu and Guoqing Ge., the program host Mr. Will Mayfield, and the visitor (of another project) Miss Ivette Hernandez.

## 2. Methodology and Development of Multiscale EnKF

### 2.1 Formulation

The multiscale EnKF in this study is developed based on the Ensemble Square Root Filter (EnSRF, Whitaker and Hamill 2002) available in the GSI. The key of the multiscale EnKF is to eliminate the impact of unreasonably noisy background error covariance (BEC) on the analyses, due to the fine grid-spacing adopted, when assimilating observations that represent relatively large scale environments (i.e., conventional observations). That is, a modified Kalman gain,  $\tilde{\mathbf{K}}$ , that can reasonably represent the correlations over the large scale will be used. In practice, this is achieved by applying a low-pass filter to the ensemble background to screen out all signals falling below the scale of interest. With this  $\tilde{\mathbf{K}}$ , the final ensemble mean analysis is updated following

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \tilde{\mathbf{K}}[\mathbf{y}^o - H(\bar{\mathbf{x}}^b)],$$

and the ensemble perturbations are updated as

$$\mathbf{x}^{a'} = \mathbf{x}^{b'} - \tilde{\alpha}\tilde{\mathbf{K}}\mathbf{H}\mathbf{x}^{b'},$$

where  $\tilde{\alpha}$  is a reduced factor applied to the Kalman gain constructed as

$$\tilde{\alpha} = \left[ \mathbf{1} + \sqrt{\mathbf{R}(\mathbf{H}\tilde{\mathbf{P}}^b\mathbf{H}^T + \mathbf{R})^{-1}} \right]^{-1},$$

where  $\mathbf{R}$  is the observation error covariance and  $\mathbf{H}\tilde{\mathbf{P}}^b\mathbf{H}^T$  is the background error covariance. The tilde hat of  $\alpha$  and  $\mathbf{P}^b$  indicates that the background error covariance matrix is constructed with a filtered ensemble background.

## 2.2 Lanczos Filtering

As mentioned above, the ensemble background is filtered by a low-pass Lanczos filter before assimilating conventional observations for providing large-scale BECs. The FV3 prognostic variables that are filtered include  $delp$  (pressure difference between adjacent model levels),  $u$ - and  $v$ -component winds,  $T$  (sensible temperature), and  $sphum$  (specific humidity). As a test, the upper-air (including rawinsonde and VAD) and surface observations (e.g., SYNOP, METAR, mesonet, etc.) are assimilated sequentially using a background filtered with filtering length scale ( $l$ ) of 240 and 60 km, respectively.

Figure 1 shows the analysis increments of  $u$  of the first ensemble member at the fifth level above ground from the two analysis passes of conventional data assimilation for the 20 May 2019 MCS, for single-scale (upper row, CNTL) and multi-scale (bottom row, MDA) experiments. The analysis increments from the two experiments have broadly similar patterns in terms of the sign, although the increments of MDA from pass 1 are much smoother and are free of small-scale structures, consistent with expectations (Fig. 1a and d). The increments of MDA are also smoother in most regions (Fig. 1b and e) except for the northern Texas panhandle region that exhibits more structures. This is related to an east-west frontal boundary at that location. With the much smoother ensemble perturbations and ensemble spatial covariance, increments from the first pass appear too smooth across the frontal boundary, so that the second pass is making compensating corrections, creating smaller scale increments near the frontal boundary (Fig. 1e). Based on this observation, we plan to test a reduced filter cutoff scale for the first pass, to better retain flow-dependent structures at scales where the covariance may still be reliable. The total analysis increments from the two passes do appear smoother from MDA (Fig. 1c and f) throughout the domain.

Figure 2 shows the power spectra of analysis increments, from different passes for CNTL and MDA. Regardless of state variables, consistent features of the power distribution are present; increments from MDA show abrupt drops from CNTL at  $\sim 150$  km and  $\sim 40$  km in the first and second pass, respectively while for the third pass the power of increments of radar DA are very similar for short wavelengths. Between 150 and 40 km wavelengths where the power of the first pass increments have dropped in MDA, the increments from the second pass actually have higher power, suggesting that surface observations are correcting errors in these

scales not corrected by sounding/VAD data. Note that the power gains from radar DA are orders of magnitude smaller than those from conventional data, mainly due to the limited regions that radar data exist (only in precipitation regions) and influence (within much smaller localization radii). For this reason, the total increments of the analysis cycle (right column) come primarily from contributions of the first two passes, with the most differences in the power spectra between CNTL and MDA seen below the 40 km wavelength.

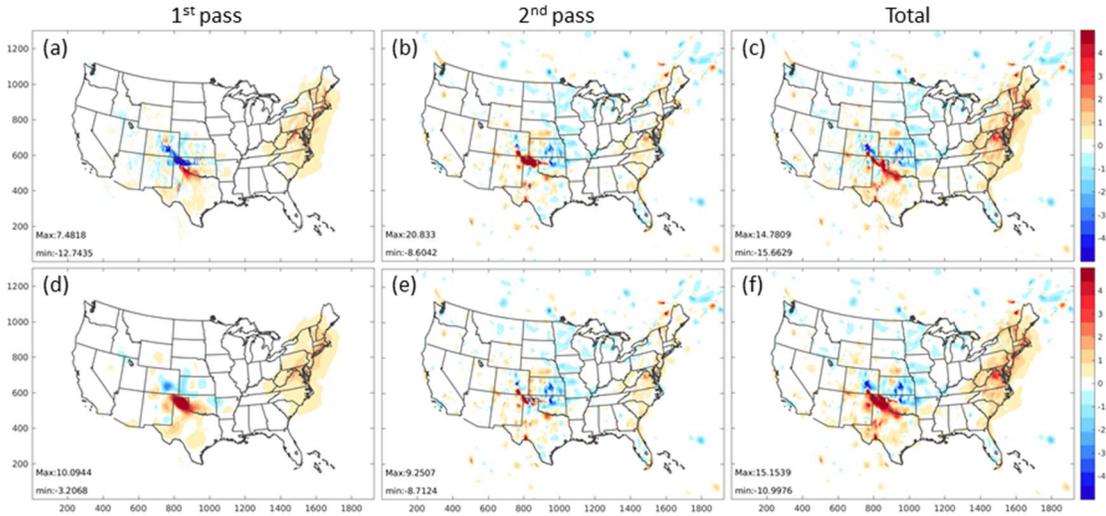


Fig. 1. Analysis increments of  $u$  ( $m s^{-1}$ ) at the 5<sup>th</sup> model level above ground of the first ensemble member from (a) first pass, (b) second pass, and (c) entire cycle for single-scale EnKF analysis (i.e., CNTL). (d), (e), and (f) are as in (a), (b), and (c), but for the multi-scale EnKF analysis. The increment maxima and minima are denoted at the lower left corner of each panel.

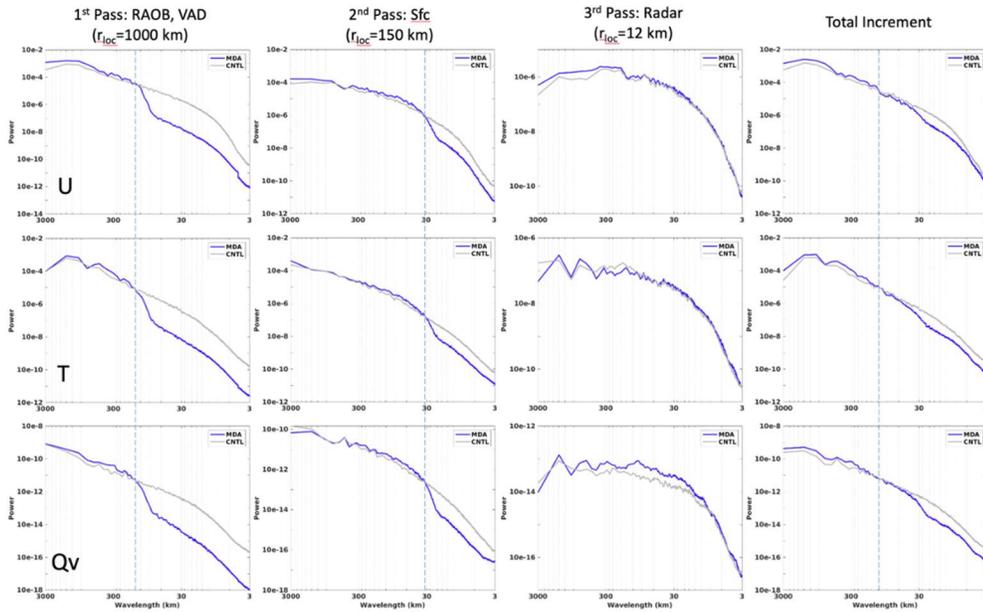


Fig. 2. Power spectra of analysis increments of  $u$  (upper row),  $T$  (middle row), and  $qv$  (bottom row) at the 5<sup>th</sup> model level above ground as a function of wavelength (km) from first pass (1<sup>st</sup> column), second pass (2<sup>nd</sup> column), third pass (3<sup>rd</sup> column) and entire cycle (4<sup>th</sup> column) for CNTL (gray lines) and MDA (blue lines).

### 2.3 Optimization of Filtering Length Scale ( $l$ )

Sensitivity experiments are conducted to determine the optimal degree of filtering for assimilating synoptic and meso-scale observations. The 21 May 2019 Central Plains squall line is studied with one-day continuously cycled DA configurations; sounding and radar VAD profiles (in first pass) and surface data (in second pass) are assimilated in 3-hr intervals with the horizontal and vertical localization radii for the first/second passes being 1000/300 km and 0.4/0.2 log(p). A deterministic forecast is launched from the ensemble mean analysis at the final cycle.

Multiscale DA experiments (MDAs) with different separation scale lengths ( $l$ ) for background filtering are performed: 1) MDA\_120\_60 that uses  $l = 120$  and 60 km for the first and second passes, respectively, and 2) MDA\_30\_30 that uses  $l = 30$  km for both passes. The forecast performances are verified against surface observations and compared with the single-scale EnKF experiment (CNTL). Fig. 3 shows that MDA\_120\_60 outperforms CNTL with smaller forecast errors in most variables throughout the 24-hr forecast period, with some exceptions in  $p_s$  (surface pressure) forecasts. MDA\_30\_30 is generally between MDA\_120\_60 and CNTL and is close to CNTL, suggesting that the longer  $l$  is preferred for assimilating upper-air data.

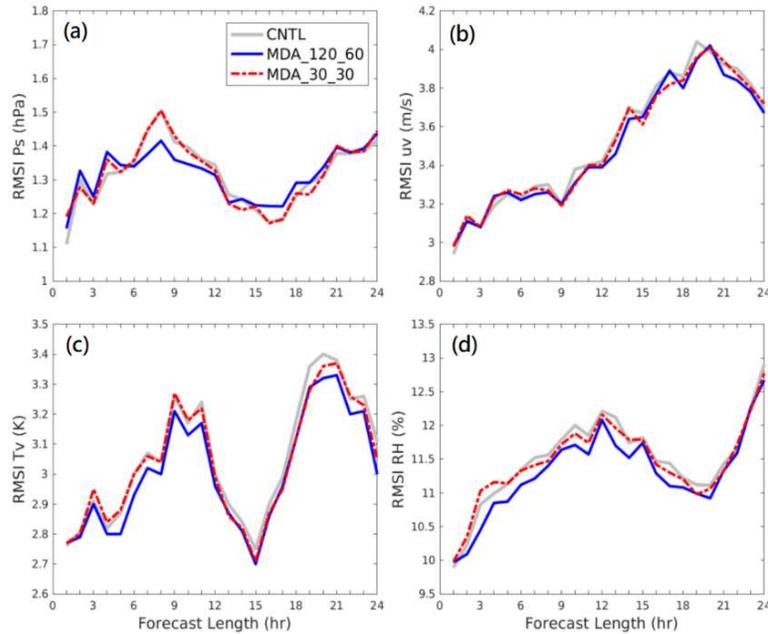


Fig. 3. Hourly root mean square (RMS) innovations of forecast (a)  $p_s$ , (b)  $uv$ , (c)  $T_v$ , and (d)  $RH$  calculated against conventional surface observations for experiments CNTL, MDA\_120\_60, and MDA\_30\_30.

Based on the MDA\_120\_60 configuration, another set of experiments is conducted with additional one-time radar DA in the last DA cycle. Reflectivity and radial velocity data are assimilated as the third pass using updated unfiltered perturbations. The root mean squared innovations (RMSIs) of CNTL and MDA with and without radar DA are shown in Fig. 4. MDA generally outperforms CNTL and the difference is largest in the relative humidity forecast. The benefit of radar DA is limited to the first 6-h, which is expected with a single radar data cycle. Fig. 5 shows the neighborhood ( $\sim 42$  km) ETS for forecast composite reflectivity for different DA experiments. The ETS of MDA experiment is higher than CNTL at most forecast times with exceptions between 9 and 13 h. Larger differences are found for the 25 dBZ threshold.

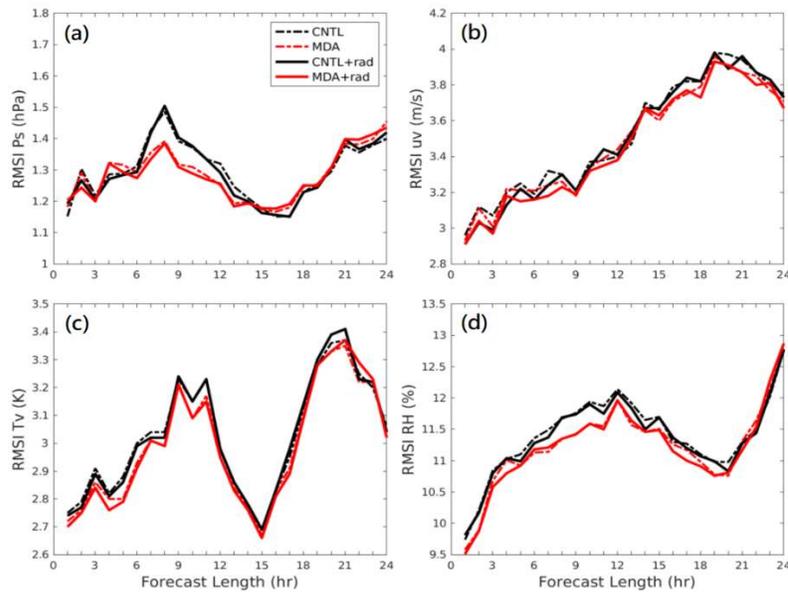


Fig. 4. Same as in Fig.3, but for experiments with and without additional radar DA.

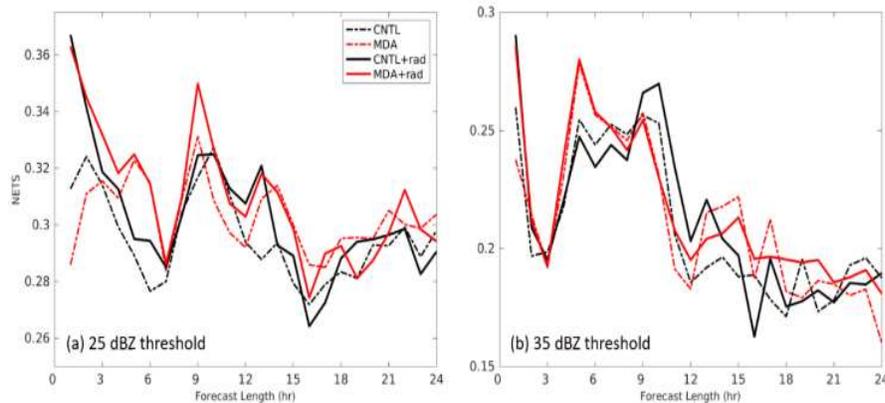


Fig. 5. Neighborhood ETSSs of 24-hr forecast composite reflectivity at (a) 25-dBZ and (b) 35-dBZ thresholds for CNTL and MDA experiments with and without radar DA.

## 2.4 Height-Dependent Filtering

Considering the horizontal variability or horizontal spatial correlation scale of atmosphere states tends to increase with height, multi-scale DA filter length scale as a function of model vertical level (a proxy for height above ground) is introduced for the multiscale EnKF. The height-dependent filtering enables assimilation of all conventional data at a single pass while accounting for correlation at different spatial scales represented by different type of observations at different altitudes. As shown in Fig. 6a, the filter length increases from 20 grid points or  $\sim 60$  km following a hyperbolic tangent function to 120 km at model level 21 ( $\sim 700$  hPa). Its impact on large-scale DA (sounding and VAD) and the subsequent deterministic forecast from the ensemble mean analysis is examined. The most significant benefit of this implementation appears in the 24-h humidity forecast in terms of smaller RMSI against sounding observations (Fig. 6b), as compared to the results using a constant filter length.

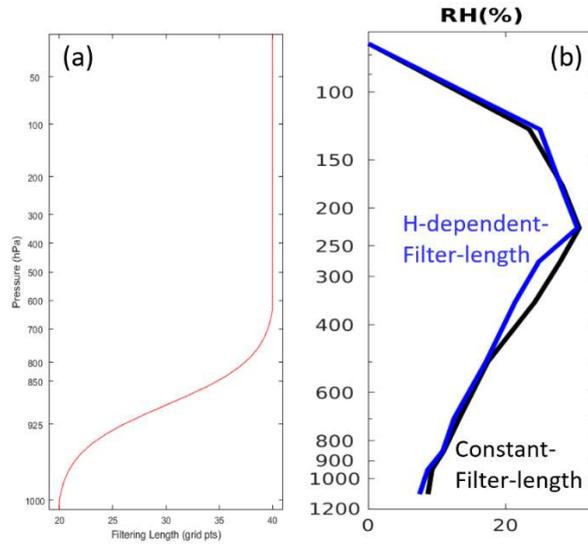


Fig. 6. (a) Height-dependent filter length as a function of model level (domain-averaged pressure at each model level is denoted). (b) RMSIs of 24-hr humidity forecasts at different height levels verified against sounding data.

## 2.5 Post-Inflation for Large-Scale DA

The original relaxation to the prior spread (RTPS) scheme in GSI EnKF restores the spread of analysis perturbations to the spread of background perturbations at all scales, corresponding to the grid spacing. For large-scale DA, such application can reintroduce small-scale features in analysis increments which should be

avoided, given that the key of the proposed multiscale EnKF is to obtain improved low-noise large-scale analyses. A multiscale RTPS procedure is implemented in GSI EnKF enabling restoration of spread at large-scale only. Fig. 7 is provided as a demonstration of this procedure with a single temperature observation EnKF analysis; the analysis increments are effectively inflated in terms of increased values (as compared to the left, the original analysis without inflation) while retaining a smoother pattern (as compared to the middle, original full-scale RTPS).

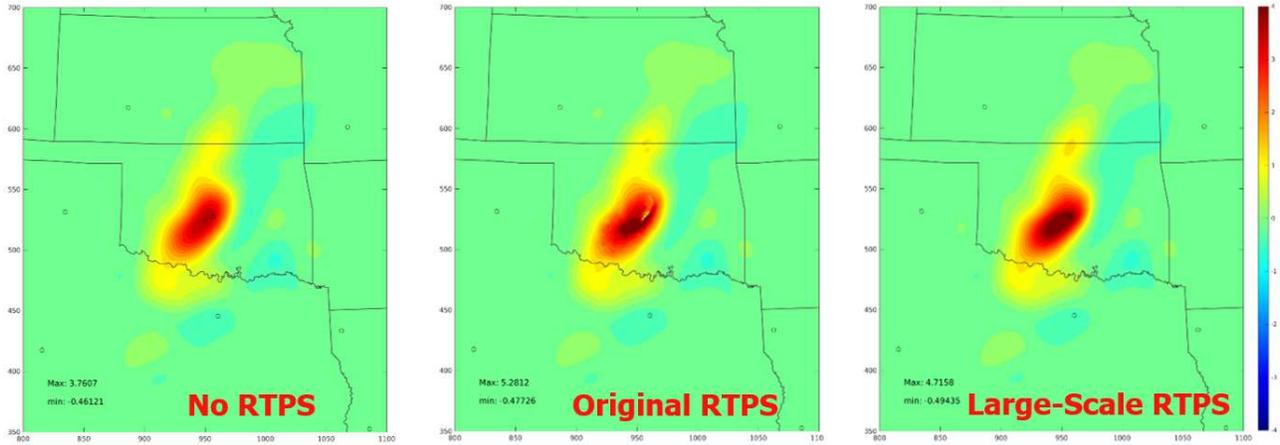


Fig. 7. Analysis increments of single temperature observation (in central Oklahoma) multiscale EnKF with no RTPS (left panel), original RTPS (middle panel), and large-scale RTPS (right panel).

### 3. Real Case Application

#### 3.1 Model and DA Configurations

The SAR FV3 model version used is the NOAA GSL develop branch (of Feb 2020). For the real-case experiments, a CONUS domain that consists of  $1921 \times 1297$  horizontal grid points with spacing of  $\sim 3$ -km is utilized. In the vertical direction, a 64-level hybrid coordinate with model top at  $\sim 0.4$  hPa is used. Model physics applied include Thompson microphysics (Thompson et al. 2008), MYNN PBL (Mellor and Yamada 1982), RRTMG radiation (Iacono et al. 2008), and Noah land surface model (LSM; Yang et al. 2011).

The real case experiments are initialized with the first 40 ensemble member forecasts of GDAS EnKF, following a 3-h spin-up of forecasts. A total of 12-h hourly cycled DA begins at 1200 UTC of the studied days (Fig. 8). For the MDA experiment, the height-dependent background filtering and large-scale RTPS as described in the previous section are used. Conventional and radar reflectivity observations are assimilated

sequentially in two passes. As for localization, a height dependent horizontal localization length, 500-km at surface increasing to 1000-km at 700 hPa, is utilized when assimilating conventional observations; a constant 12-km horizontal localization length is used for radar DA. The direct radar reflectivity DA used in this study for the SAR FV3 can be referred to in Tong et al. 2020.



Fig. 8. Cycled DA configuration for real-case experiments.

The 21 May 2021 Central Plains convective storm case is first tested with assimilation of only conventional data. After all 13 analysis cycles are finished, a 48-h deterministic forecast is launched for each experiment from the ensemble mean analysis of the final cycle. According to the verification against surface observations as shown in Fig. 9, the error of temperature and humidity predictions, in terms of RMSI, is found noticeably smaller in the multiscale EnKF experiment as compared to the CNTL. The outperformance of MDA persists for up to 36 hours for temperature prediction and slightly longer for humidity.

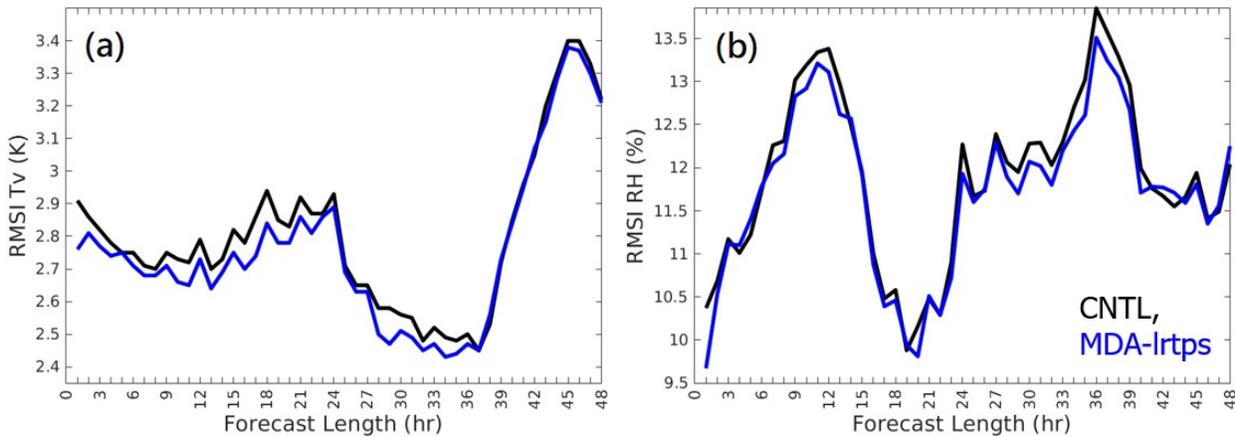


Fig. 9. 48-hour forecast RMSI against surface observations of (a) temperature and (b) humidity for the 21 May 2021 storm experiments with single-scale EnKF, referred to as CNTL (black) and multi-scale EnKF (blue).

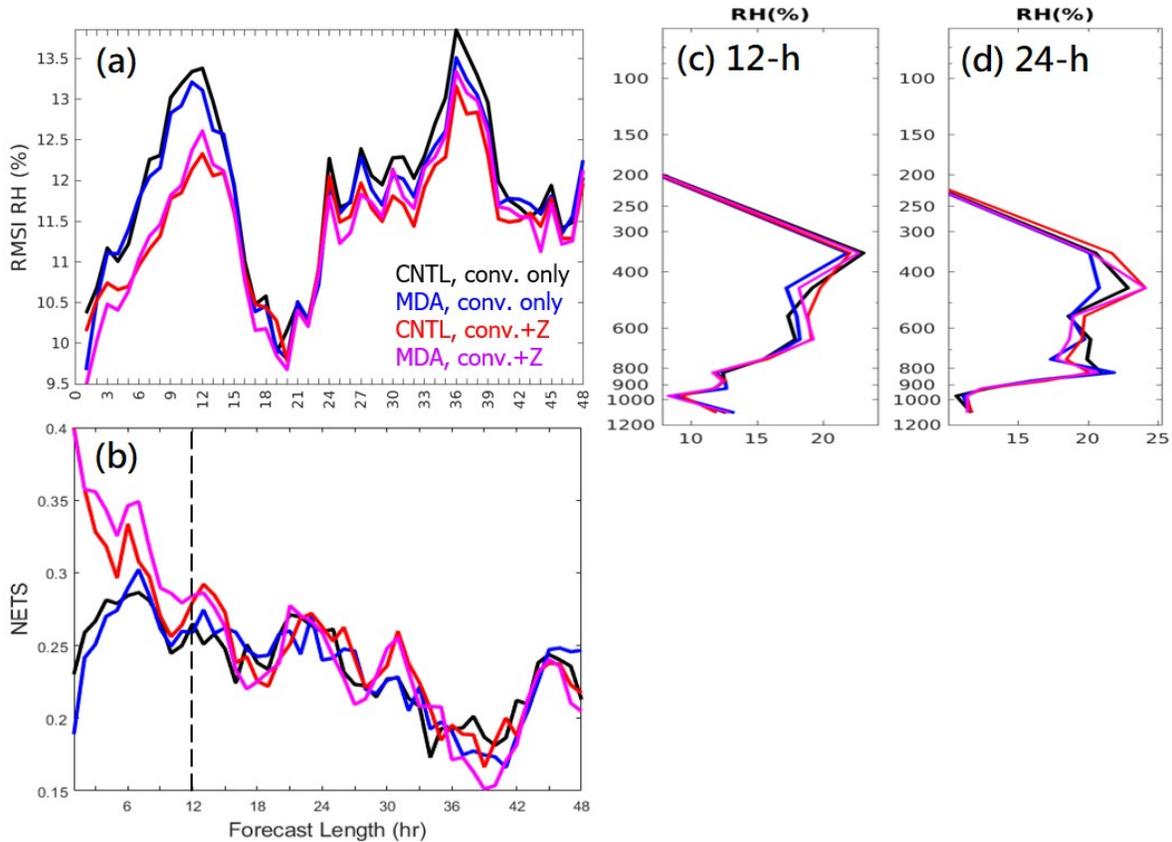


Fig. 10. (a) Same as in Fig. 9b, but with two additional experiments including radar DA (denoted by red and magenta lines for CNTL and MDA, respectively). (b) 42-km neighborhood ETS of 48-hour forecast composite reflectivity at a 25 dBZ threshold. (c) 12-hour and (d) 24-hour forecast RMSI profile against sounding humidity observations.

The additional pass of radar reflectivity assimilation in each cycle shows a benefit mainly in humidity forecasts while the advantage of multiscale EnKF over CNTL is limited to the first 24 hours forecast, with periods of exception (Fig. 10a). The benefit of multiscale EnKF on humidity further contributes to storm prediction in terms of the higher neighborhood equitable threshold score (NETS) of forecast composite reflectivity; however, it is mostly limited to the first 12-h forecasts (Fig. 10b). The positive impact of multiscale EnKF on humidity forecasts also presents in the RMSI verification against sounding observations over most vertical levels of atmosphere (Fig. 10c and d). The improvement of the storm prediction by MDA appears mainly on reducing overforecasting, as denoted by the red blocks in Fig. 11.

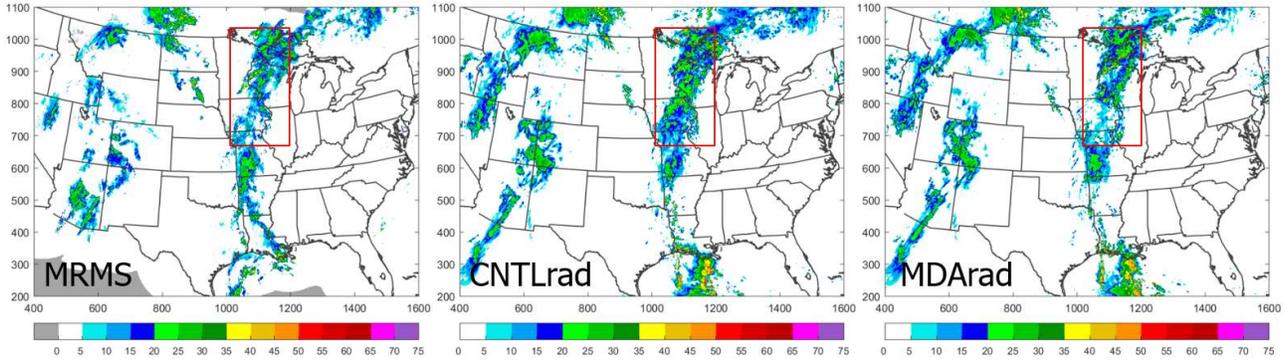


Fig. 11. 12-h forecast composite reflectivity, valid at 1200 UTC 21 May 2021, for CNTL (middle) and MDA (right) experiments with both conventional and radar reflectivity DA, as compared to the MRMS observation (left).

The impact of multiscale EnKF is further investigated by looking into the forecast performance of individual ensemble members. Fig. 12b and d shows the forecast performance of the first ensemble member. As compared to the deterministic forecast from the ensemble mean analysis (Fig. 12a and c), greater positive impact can be found for certain individual members, in terms of forecast errors in significantly smaller quantitative measures for a consistent forecast lead-time, shown for verifications against both surface and sounding humidity observations.

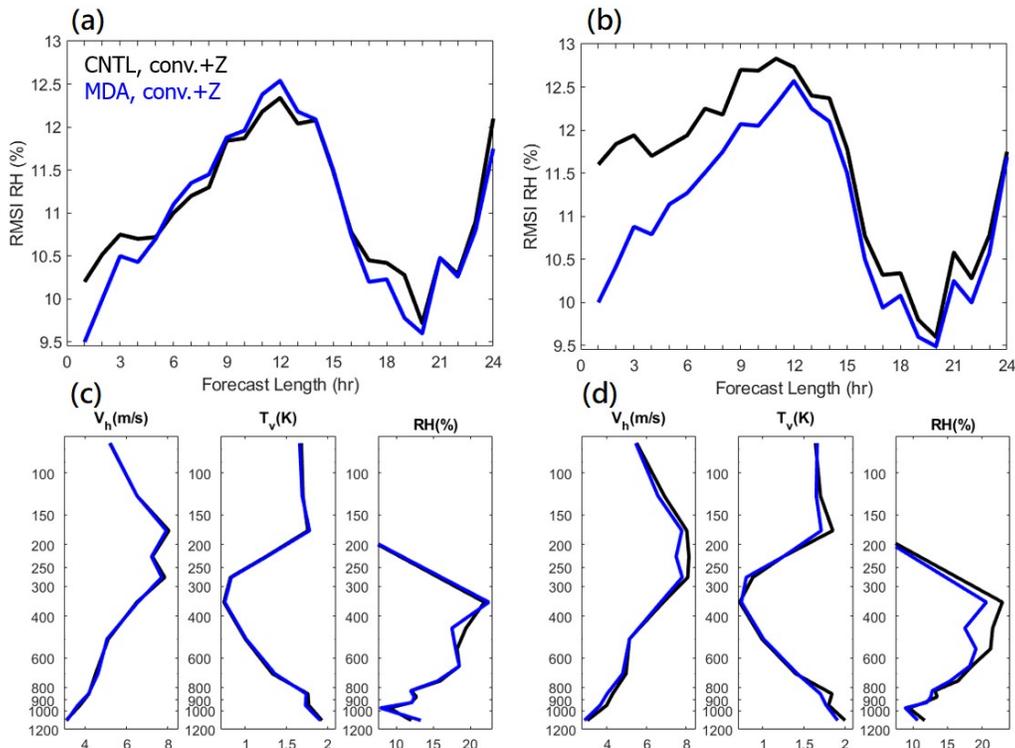


Fig. 12. RMSI of 24-hour deterministic forecasts from (a) ensemble mean analysis and (b) the first ensemble member against surface humidity observations. (c) and (d) are as in (a) and (b), respectively, but for the 24-hour forecast verified against sounding horizontal winds, temperature and humidity observations.

As the greater potential of multiscale EnKF is shown for the individual members given the merit of EnKF on perturbation updates, we further look into the average performance of all ensemble member forecasts. For the 28 May 2021 South Plains Squall Line case, Fig. 13a, b, and c show the ensemble forecast evaluation for 24 hours against conventional observations. The RMSIs shown are averaged from the RMSI calculated for each individual ensemble member, which represents the average performance of ensemble forecasts. According to the plots, significant and time-persistent outperformance of MDA over CNTL is suggested. As for the sounding verification (Fig. 13c), the largest improvement of MDA appears in the humidity forecast. Furthermore, the benefit of MDA is found to be promising in storm prediction as well, in terms of the averaged ETS of hourly precipitation forecasts (Fig. 13 d and e), especially in light rain regions.

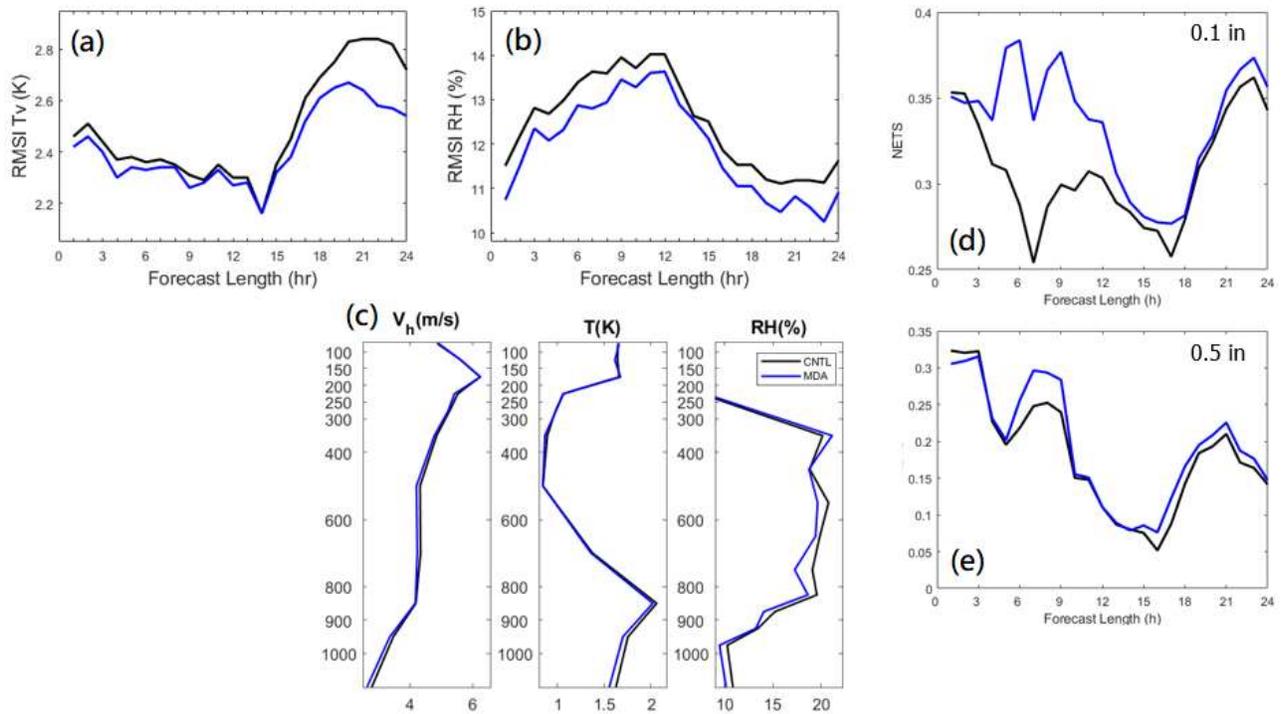


Fig. 13. Ensemble averaged RMSI against surface (a)  $T_v$  and (b) RH observations. (c) Ensemble averaged RMSI of 12-h ensemble forecasts against sounding observations. Ensemble averaged neighborhood ETS of forecast hourly precipitation at a threshold of (d) 0.1 inches and (e) 0.5 inches.

#### 4. Summary

The GSI-based multiscale EnKF (MDA) capabilities are developed for the convection-allowing SAR FV3, aiming at improving the large-scale environment for storm-scale forecasts. The proposed MDA method uses filtered background covariances with long localization lengths for assimilation of conventional observations that sample synoptic to meso-scale perturbations. Sensitivity experiments are performed to determine ideal filtering length scales sufficient to diminish unfavorable noises in analyses. In addition, the height-dependent filtering length is proposed and its impact is examined with one-time upper-air data assimilation; the benefit in subsequent forecasts show for up to 24 hours, in particular for prediction of humidity. The post inflation in the GSI, relaxation to prior spread (RTPS), is optimized accordingly for MDA to restore only the large-scale background perturbations, which avoids reintroducing small-scale noises in analyses.

The MDA is examined with a 12-hour hourly cycled update configuration for real cases and its impact is evaluated. In terms of the deterministic forecasts from the final ensemble mean analysis, consistent improvement of MDA can be found in prediction of most variables for up to 48 hours when only assimilating conventional data; when including radar DA, the benefit of MDA is relatively limited on the storm prediction and humidity forecast, for a shorter lead time. The positive impact of the MDA is found to be greater in the performance of individual ensemble members as well as the ensemble average. It is in our plan to apply the MDA method to more cases for a statistically robust conclusion.

#### References:

- Iacono, M.J., J.S. Delamere, E.J. Mlawer, M.W. Shephard, S.A. Clough, and W.D. Collins, Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.*, **113**, D13103, doi:10.1029/2008JD009944, 2008
- Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure-model for geophysical fluid problems. *Reviews of Geophysics*, **20(4)**, 851–875. <https://doi.org/10.1029/RG020i004p00851>
- Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, <https://doi.org/10.1175/2008MWR2387.1>
- Tong, C.-C., Y. Jung, M. Xue, and C. Liu, 2020: Direct assimilation of radar data within the National Weather

Service operational GSI EnKF and hybrid En3DVar systems for the stand-alone regional FV3 model at a convection-allowing resolution. *Geophys. Res. Lett.*, **47**, e2020GL090179, <https://doi.org/10.1029/2020GL090179>.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1

Yang, Z.-L., G.-Y. Niu, K. E. Mitchell, F. Chen, M. B. Ek, M. Barlage, K. Manning, D. Niyogi, M. Tewari, and Y. Xia, 2011: The Community Noah Land Surface Model with Multi-Parameterization Options (Noah-MP): 2. Evaluation over Global River Basins. *J. Geophys. Res.*, doi:10.1029/2010JD015140