TEST AND EVALUATION OF CCPP PHYSICS SUITES AND DATA ASSIMILATION ALGORITHMS TO IMPROVE THE RAPID REFRESH FORECAST SYSTEM FOR CONVECTION FORECASTS

Ivette Hernández Baños

PhD Student in Meteorology
National Institute for Space Research
São José dos Campos, SP, Brazil
ivette.banos@inpe.br
Ph.D. Advisor: Luiz Fernando Sapucci

Final project report

Developmental Testbed Center (DTC) Visitor Program

DTC Host: Ming Hu, Guoqing Ge, and Will Mayfield

March
2021

ABSTRACT

The Unified Forecast System application for regional and convective scales, or Rapid Refresh Forecast System (RRFS), is under development and aims to replace the operational suite of regional and convective scale modeling systems in the next upgrade. In order to achieve skillful forecasts comparable to current operational modeling systems, such as NAM, RAP, and HRRR, each component needs to be exhaustively tested and best configured. The current development version of RRFS includes a FV3 Limited Area Model with a Common Community Physics Package (CCPP), Unified Post-Processing system, and data assimilation capability using the Gridpoint Statistical Interpolation (GSI) analysis system, providing a suitable research framework to assess its ability to represent convection.

In this study, various physics suites and data assimilation algorithms were assessed to improve the RRFS forecasts of a squall line over Oklahoma on May 4th, 2020. Numerical experiments were conducted running hourly cycles from May 4th 00z to May 5th 06z with 18-h forecasts launched at each cycle. Forecast verification was performed using the Model Evaluation Tools.

Four CCPP physics suites were tested: two Global Forecast System (GFS)-based physics, a suite developed at NOAA's Global Systems Laboratory, and a suite based on RAP/HRRR physics. Various analysis algorithms in GSI were evaluated, such as the three dimensional (3D) variational versus hybrid 3D Ensemble Variational (3DEnVar) data assimilation, different analysis grid ratios, supersaturation removal, and various weights of ensemble background error covariance in the hybrid analysis. Observation impact experiments were conducted and the HRRR and GFS as cold start initial conditions were also evaluated. Results show that the FV3LAM model is able to represent convection close to the observed systems. This study indicates the current RRFS has great potential for convective scale forecasts but more testing and evaluation are needed.

42 # LIST OF ABBREVIATIONS

ACARS	–	Aircraft Communications Addressing and Reporting System
AIREP	–	Aircraft Weather Report
AMDAR	–	Aircraft Meteorological Data Relay
ATLAS	–	Autonomous Temperature Line Acquisition System
MDCRS	–	Meteorological Data Collection and Reporting Service
METAR	–	Surface Weather Observations and Reports (translated from French)
43 MESONET	–	Mesoscale Network
PIBAL	–	Pilot Balloon observation
PIREP	–	Pilot Report
RASS	–	Radio Acoustic Sounding System
SODAR	–	Sonic Detection And Ranging
TAMDAR	–	Tropospheric Airborne Meteorological Data Reporting
WSR88D	–	Weather Surveillance Radar-1988 Doppler

CONTENTS

1 Introduction

More powerful computer resources, optimization of numerical codes, improvements on initial conditions and model physics, and new and better use of existing observational systems have allowed for more reliable and extended forecasts of the Global Forecast System (GFS) since its first version in 1980 (WHITE et al., 2018). Now, the GFS is the foundation of the National Centers for Environmental Prediction (NCEP)'s numerical prediction systems and also of many centers around the world, such as the Center for Weather Forecast and Climatic Studies of the National Institute for Space Research, in São Paulo, Brazil (CPTEC-INPE).

In order to simplify the NCEP's operational suite into a single system capable of representing different spatial and temporal scales as well as all components of the Earth's system, the National Oceanic and Atmospheric Administration (NOAA) is transitioning toward a Unified Forecast System (UFS), based on the non-hydrostatic finite volume cubed-sphere (FV3) dynamical core (SIP, 2018). Following NOAA's initiative, the CPTEC-INPE is on the first steps of the selection of a new dynamical core and also envisions unifying its modeling suite.

The UFS is a community-based system that enables contributions from the research community to operational applications. It encompasses medium- and short-range weather, hurricane, seasonal to sub-seasonal, air quality, coastal, marine and cryosphere as well as space weather applications (UFS, 2019). The future U.S. operational mesoscale modeling system, the Rapid Refresh Forecast System (RRFS), is currently under development and is built upon the UFS Short-Range Weather (SRW) Application (ALEXANDER; CARLEY, 2020). RRFS aims to replace the current suite of operational regional models in the next upgrade, but in order to achieve comparable forecast skill for operational applications, each component needs to be exhaustively tested.

The first version of the SRW application was recently released including the FV3 Limited Area Model (FV3LAM) with pre-processing utilities, the Common Community Physics Package (CCPP), the Unified Post Processor (UPP), and a workflow to run the system (WOLFF; BECK, 2020). Although a data assimilation capability was not yet included in the public release, the Gridpoint Statistical Interpolation (GSI) has been added as the analysis component of the SRW application to improve initial conditions for the FV3LAM in development of the RRFS at NOAA's Global Systems Laboratory (GSL). This provides a suitable research framework with the necessary components to explore current RRFS capabilities. In this study, an extensive testing and evaluation of various CCPP suites and data assimilation configurations was conducted in order to provide developers an insight on the current capabilities of RRFS in predicting convection. This work is focused on the improvement of convection forecasts of a squall line that occurred over Oklahoma on 4 May 2020. Many numerical experiments were conducted testing various aspects of the RRFS. Numerous subjective and objective forecasts verification were performed for evaluating each experiment.

Objectives

This visitor project aims to investigate the RRFS capability to represent convection by fulfilling the following specific objectives:

a) Assess different CCPP physics suites;

b) Assess different data assimilation algorithms and configurations;

c) Evaluate the impact of different data types, such as: upper-air, surface, radar radial velocity, satellite derived winds (AMV- Atmospheric Motion Vectors), and Global Navigation Satellite System (GNSS) radio occultation observations;

d) Evaluate different hybrid 3D EnVar configurations, including vertical dependent ones;

e) Evaluate new hybrid vertical coordinate coefficients from GSL RRFS system;

f) Examine different cold start initial conditions and cycling configurations;

g) Evaluate RRFS forecasts for convection initiation and evolution.

Section 2 presents a brief description of the tasks carried out during the project execution. The materials and methods employed are addressed in Section 3 and the numerical experiments conducted with the analyses of results are presented in Section 4. The summary and future work can be found in Section 5.

2 Project execution

This visitor project forms part of graduate student research at CPTEC/INPE and began in March 2020, with a planned 12 month execution at the NCAR Foothills Laboratory in Boulder, CO. During the first two weeks it was possible to visit the laboratory, but due to mandatory regulations because of the COVID-19 outbreak, the rest of the project was executed virtually. The project was hosted by Ming Hu, Guoqing Ge, and Will Mayfield—the Developmental Testbed Center (DTC) data assimilation team, together with Louisa Nance and other researchers at the DTC. DTC hosts were very helpful, flexible, and reachable, making it possible to complete the project on time and achieve the proposed objectives.

The work during the first two months of this visit was conducted on the National Center for Atmospheric Research (NCAR) High Performance Computing (HPC) system, Cheyenne. During this time, the UFS Medium-Range Weather Application v1.1's Graduate Student Test (DELUCA; JASKOFOR, 2020) was completed and a bibliographic revision of the UFS, FV3 model and its components, and CCPP documentation was begun. In addition, initial tests were carried out using the available NCEP Environmental Modeling Center (EMC) community workflow with and without GSI for analyzing the analysis increments and assimilated observations. In May, the studied case was selected; the initial and lateral boundary conditions and observations were staged; the fixed files were generated; and short runs were performed. At the end of this month, access to the NOAA HPC Orion was obtained. During June and July, the available developments on the SRW application and latest version of GSI were debugged and successfully installed and compiled on Orion. Tests with different CCPP suites were initiated. In August, experiments with the selected CCPP suites were run and Model Evaluation Tools (MET) verification software package was used to calculate statistics.

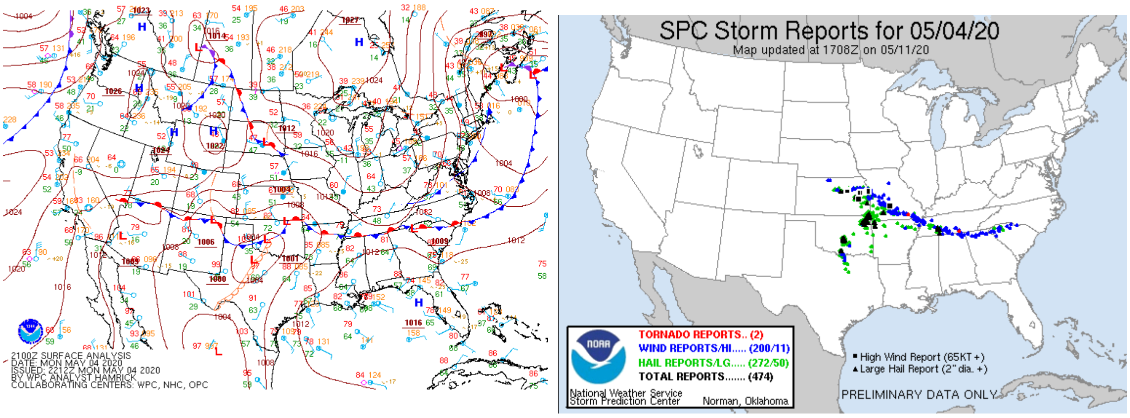
In September and November, data assimilation experiments were run for exploring different configurations of GSI, such as the supersaturation clipping option, three dimensional variational (3D-Var) versus hybrid 3D ensemble variational data analysis (hybrid 3D-EnVar) with different ensemble background error covariance weights. MET matched pairs were carefully analyzed and an observation quality issue in verification was explored. An article was written for the Autumn 2020 edition of the DTC Newsletter including preliminary results obtained up to that time. In December, other data assimilation configurations in the system were tested and more MET verification tools were applied. New experiments were run during January and February, comparing results with previous experiments using other CCPP physics suites. A seminar was prepared and presented on March 4th 2021 covering the most important results obtained.

3 Materials and methods

A description of the case studied, the domain, and data used are presented in Sections 3.1 through 3.3. RRFS workflow and cycling configurations are addressed in Sections 3.4 and 3.5, respectively, and the forecast verification methodology is described in Section 3.6.

3.1 Case study

A line of storms developed over northeastern Oklahoma ahead of a southward moving cold front during the afternoon of 4 May 2020. Embedded supercell and locally severe storms caused several instances of large hail and high wind, mostly over northeast and south-central Oklahoma, southeast Kansas, southwest Missouri, and northwest Arkansas (Fig. 3.1 (b)). The Hydro-meteorological Prediction Center (HPC)'s surface analysis at 18Z May 4th, 2020 (Fig. 3.1 (a)) showed a surface low pressure across western Oklahoma with a dry line extended over western Texas, favoring an environment with low-level convergence, high temperatures and humidity over these areas. A trough in the upper troposphere moving eastward and deepening into the troposphere supported the upper-level divergence with rising movements over Oklahoma and Kansas. A surface warm front over northeast Oklahoma and southeast Kansas favored warm advection to portions of northeast Oklahoma and southeastern Kansas.



(a) (b)

Figure 3.1 - HPC surface analysis at 18Z (a) and the Storm Prediction Center's storm reports for 4 May 2020 (b).

Between 19Z and 20Z, high values of MLCAPE (Mixed Layer Convective Available Potential Energy) and effective bulk shear were observed over northeastern Oklahoma. At 20Z May 4th, the first convective cells were observed on the Multi-Sensor Multi-Radar (MSMR) composite reflectivity observations (see the black circle in Fig. 3.2 B.) and around 22Z (Fig. 3.2 D.), a line of storms extending across central Oklahoma was observed along the pre-frontal wind shift. The system evolved while slowly moving southeastward. Weak winds in the lower levels but strong winds aloft and high thermodynamic instability supported the elevated convection. A supercell developed over far southeast Missouri at 00Z May 5th producing very large hail and locally damaging winds (Fig. 3.1 (b)). Small convective cells initiated and developed over Texas between 23Z May 4th (Fig. 3.2 E.) to 01Z May 5th (Fig. 3.2 G.) along the dry line. Clusters of severe storms developed across south-central Oklahoma and north-central Texas in the intersection between the cold front and the dry line with hail occurrences and associated strong wind.

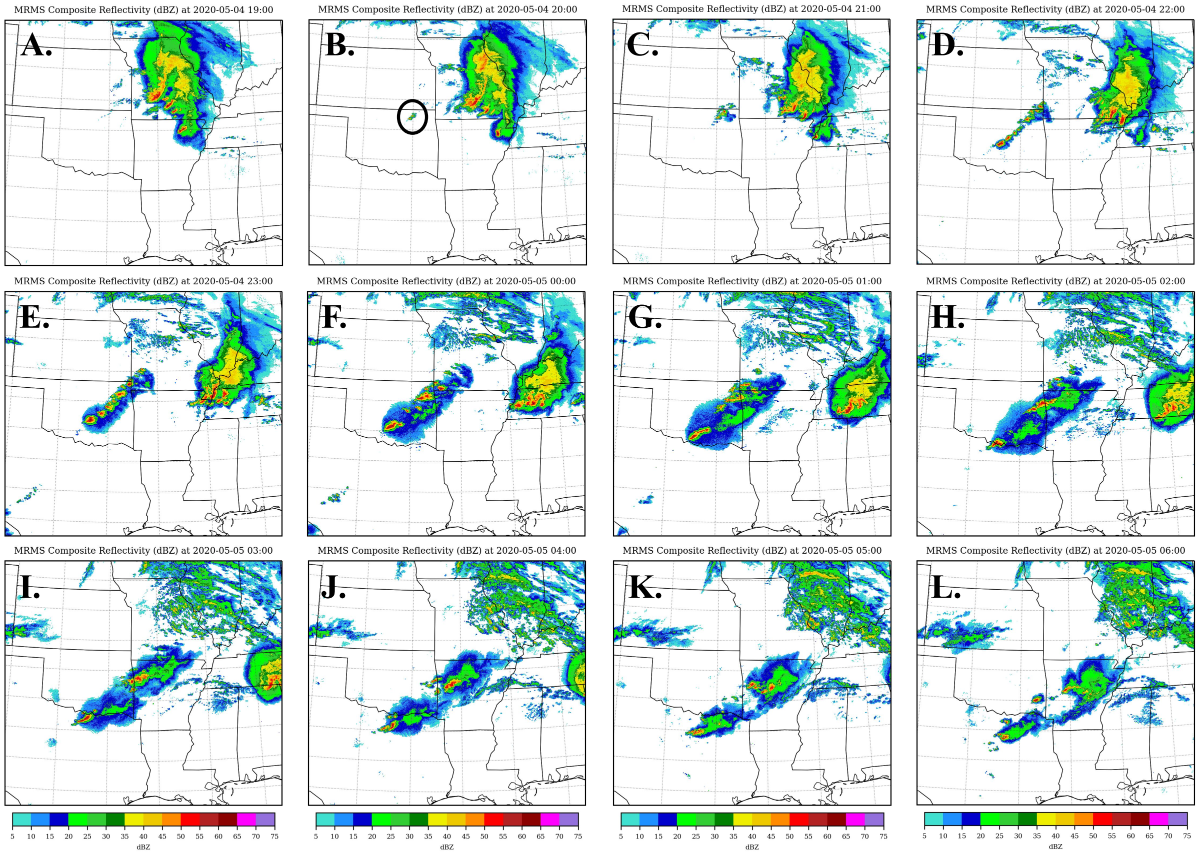


Figure 3.2 - Hourly Multi-Sensor Multi-Radar (MSMR) composite reflectivity since 19Z May 4th, 2020 through 06Z May 5th, 2020 (panels A through L).

3.2 Domain

For the simulation of this case, a domain was configured covering the area of interest, mainly Oklahoma, Kansas, Missouri, Arkansas, and much of Texas. The domain has a 460x460 grid centered on Fort Smith, Arkansas (-95.35°W; 35.28°N) with 3 km of horizontal resolution and 64 vertical levels. The Extended Schmidt Gnomonic method developed by Purser et al. (2020) and implemented in the SRW application v1.0.0 (UFS Development Team, 2021) was used for grid generation. Fig. 3.3 shows the domain coverage and orography created using the pre-processing utilities, where the black star indicates the central point.

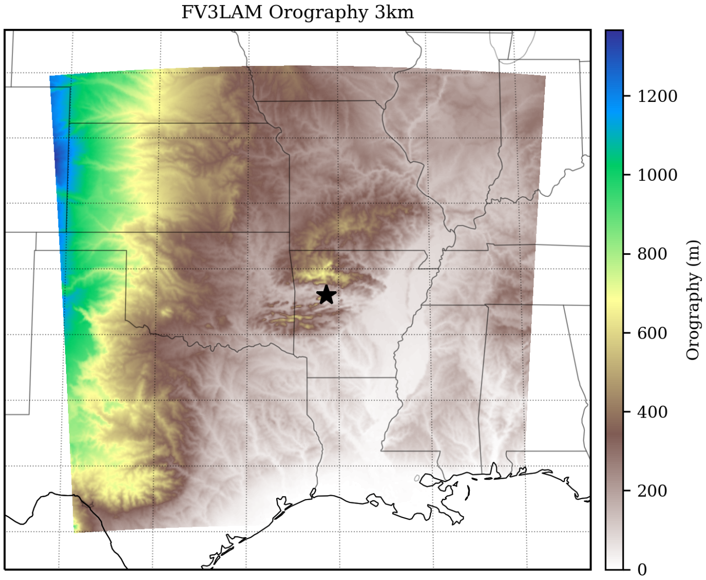


Figure 3.3 - Domain with the orography created for the numerical experiments in this study. The black star indicates the central latitude and longitude on Forth Smith, Arkansas.

3.3 Data

All simulations ran hourly cycles with 18-hour forecasts starting at 00Z on May 4th through 06Z May 5th. Initial conditions (ICs) and lateral boundary conditions (LBCs) were from the 3-km High-Resolution Rapid Refresh (HRRR) model.

Hourly Rapid Refresh (RAP) observations were utilized for the analysis. These observations are generated at NCEP for the hourly updated data assimilation component in RAP and typically include all available data from 30 minutes before to 15 minutes after the analysis hour (HU et al., 2017). Historic data were obtained from NOAA's High Performance Storage System (HPSS) archives. Upper-air, surface, radar radial velocity and satellite AMV wind, as well as precipitable water are called the ConTroL OBServations dataset (CTLOBS) in the experiments. Table 3.1 provides detailed information on the observation types included in CTLOBS as well as their sources, where ps stands for station (surface) pressure; t for virtual temperature and/or sensible (dry bulb) temperature; q for specific humidity; and uv for u- and v-components of wind.

Table 3.1 - Sources of assimilated observations.

Observation Type	**Description**	**Observation variable**				
		ps	t	q	pw	uv
ADPUPA	Rawindsonde	✓	✓	✓		✓
RASSDA	RASS virtual temperature		✓			
AIRCFT	AIREP and PIREP aircraft		✓			✓
AIRCFT	AMDAR aircraft		✓			✓
ADPUPA	Dropsonde		✓	✓		
AIRCAR	MDCRS ACARS aircraft		✓	✓		✓
AIRCFT	TAMDAR aircraft		✓	✓		✓
AIRCFT	Canadian AMDAR aircraft		✓	✓		✓
GPSIPW	GPS Integrated Precipitable Water				✓	
SFCSHP	Ship, Buoy, C-MAN, and Tide Gauge reports	✓	✓	✓		✓
ADPSFC	SYNOPTIC and METAR	✓	✓	✓		✓
SFCSHP	Splash-level Dropsonde over ocean	✓	✓	✓		
ADPSFC	METAR	✓	✓	✓		✓
MESONET	Surface MESONET	✓	✓	✓		
SATWND	Cloud drifts, cloud top, and deep layer from different satellite imagery					✓
ADPUPA	PIBAL					✓
PROFLR	Wind Profiler					✓
VADWND	Vertical Azimuth Display from WSR88D radars					✓
PROFLR	Multi-Agency Profiler and SODAR					✓
PROFLR	Wind Profiler from PIBAL					✓
SFCSHP	ATLAS Buoy					✓
WDSATR	Scatterometer winds over ocean					✓

In addition, observation BUFR (Binary Universal Form for data Representation) files from the Global Data Assimilation System (GDAS) were used to assimilate bending angles derived from Global Navigation Satellite System Radio Occultation (GNSS-RO) observations. GDAS is the operational global data assimilation system used to create the initial conditions for GFS, and its observation files are available daily at synoptic hours. GDAS observations files include all valid data from 3 hours before to 3 hours after the analysis time (KLEIST et al., 2009). In order to use GDAS observation data in the hourly cycles configured in this study (see Section 3.5), the same observation file was reused for each of the 3 hours before and 3 hours after its analysis time, and two files were used at overlapping times (Fig. 3.4). For example, the file for synoptic hour 12Z was used to run the cycles 09Z through 15Z. However, since observations before 09Z are included in the 06Z GDAS observation files, both 06Z and 12Z GDAS observation files were used to run

GSI at 09Z as the purple boxes indicate in Fig. 3.4.

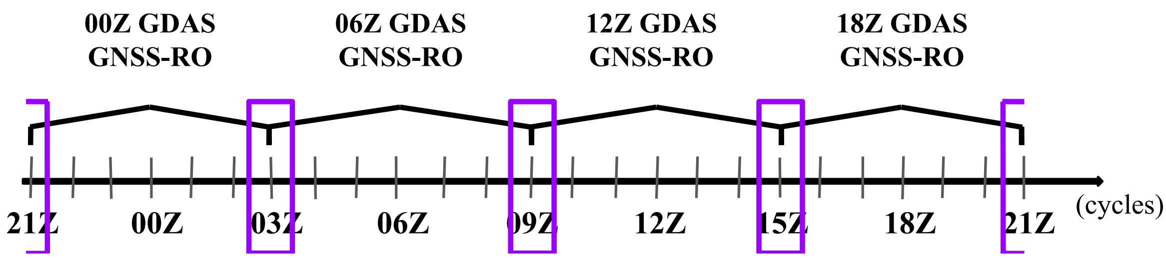


Figure 3.4 - Strategy used to assimilate GNSS-RO bending angles from GDAS observations every hour.

Available Low-Earth Orbit (LEO) satellites during May 4th and 5th, 2020 include the second mission of the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC-2), the European Meteorological Operational (MetOp) A, B, and C satellites, the Korea Multi-Purpose Satellite-5 (KOMPSat5), the Spanish mission PAZ, and the German Tandemx and TerraSARX satellites. COSMIC-2 and MetOp-C were not among the satellites that were currently assimilated in the GSI version used in this study. Because GNSS-RO observations are not well distributed in the limited domain, these satellites were therefore added to the code in order to increase the number of available bending angle observations during the execution period. Quality control procedures were maintained similar to those applied on COSMIC-1 and MetOp-A and -B observations, respectively. Bending angles from all LEO satellites were assimilated up to 50 km.

For the ensemble component of the hybrid analysis, 9-h forecast from GDAS Ensemble Kalman Filter (EnKF; Whitaker et al. (2008)) 80-member ensemble were used. These data are also available on HPSS archives. However, since they are available only 4 times per day, the same 9-h forecast GDAS ensembles were reused for the 3 hours before and 3 hours after its valid hour as indicated in Fig. 3.5. For example, the 9-h forecast GDAS ensembles initialized at 00Z (valid at 09Z) were used for the cycles from 07Z to 12Z. Similarly, the 9-h forecast GDAS ensembles initialized at 06Z (valid at 15Z) were used for the cycles from 13Z to 18Z. This follows the strategy adopted in the RAP hybrid analysis (HU et al., 2017).

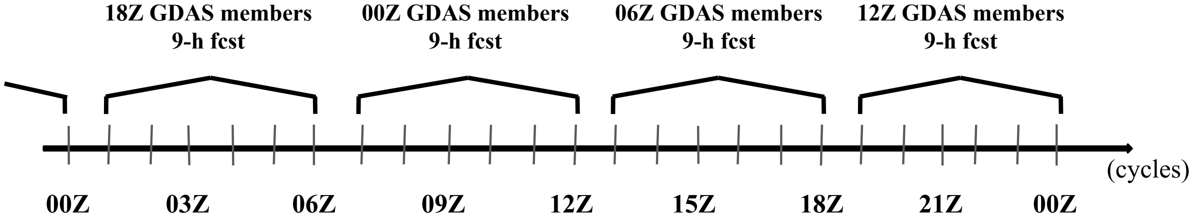


Figure 3.5 - Strategy used to reuse 9-h forecast GDAS ensembles in GSI every hour.

3.4 Workflow

The RRFS workflow is based on the UFS SRW application v1.0.0 (UFS Development Team, 2021) community workflow. A Rocoto workflow management system (`https://github.com/christopherwharrop/rocoto/wiki/Documentation`) is used to schedule the tasks that will be executed and manage the cycling configuration with other crucial information needed to run the system. The UFS SRW application code can be publicly found at `https://github.com/ufs-community/ufs-srweather-app`. The released workflow includes the tasks to configure the domain by creating the corresponding grid, orography, and climatological or fixed information; obtain and create the initial (0-h forecast) and lateral boundary conditions (1-, 2-, ..., Nth-h forecasts) from an external model (*e.g.,* GFS, RAP, and HRRR); execute the limited area model; and finally, post-process the model outputs. In this study, the workflow was modified to include the analysis tasks and different cycling strategies. Fig. 3.6 shows the inclusion of the tasks used to create the analysis for cold start (blue box) or warm start (red box) cycles. The cold start indicates that the initial conditions will be generated from an external model and the analysis use external model forecast as background combined with observations valid within the time window around initialization time. Meanwhile, the warm start will use the 1-h forecast from the previous cycle as background for the analysis to provide the initial condition for the model forecast. At each warm start initialization time, the analysis task and related data flow are highlighted in Fig. 3.6 by the dashed lines.

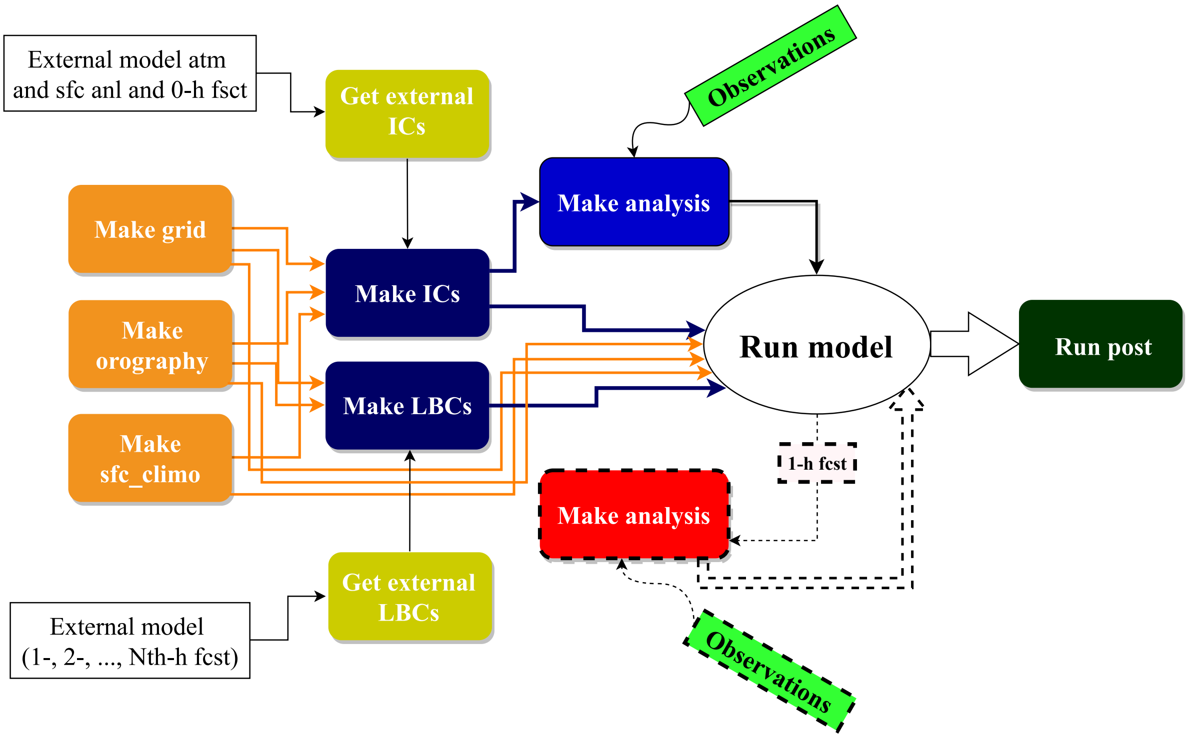


Figure 3.6 - Diagram including the tasks on current RRFS workflow and the tasks added to create the analysis using GSI.

3.5 Cycling configurations

In this project, five cycling strategies were tested. First, a cycling was configured using cold starts at each initialization time as indicated by the blue arrows in Fig. 3.7. Hourly HRRR analysis results were used to initialize the FV3LAM, which ran 18-h forecasts with hourly outputs. For this cycling configuration, the original UFS SRW application workflow was used without running any analysis tasks. Experiments using this cycling were called HRRRDA. HRRR analyses can be considered the best ICs for mesoscale models at present. Thus, this configuration sought to show how well the FV3LAM is capable of simulating this case of squall line using best currently ICs.

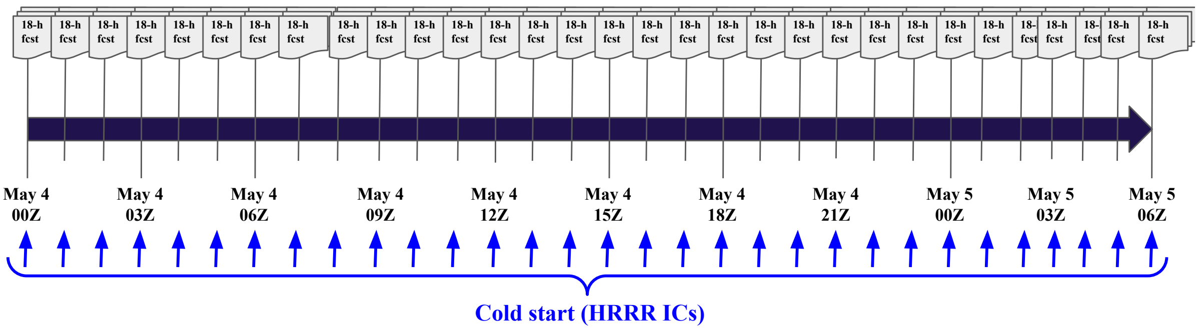


Figure 3.7 - HRRRDA cycling configuration.

Second, a baseline cycling was configured based on HRRRDA. As in HRRRDA, 18-hour forecasts were performed in each cycle, but a cold start was used only every 12 hours. Then, a warm start was used between 01Z and 11Z as specified by the red braces in Fig. 3.8. The model was configured to save the first hour forecast (restart files), which was used to initialize the model in the next cycle. The schematic in Fig. 3.8 shows the start type and forecast hours specified in this cycling configuration. No data analysis performed in each cycle. Experiments with this configuration were called NoDA.

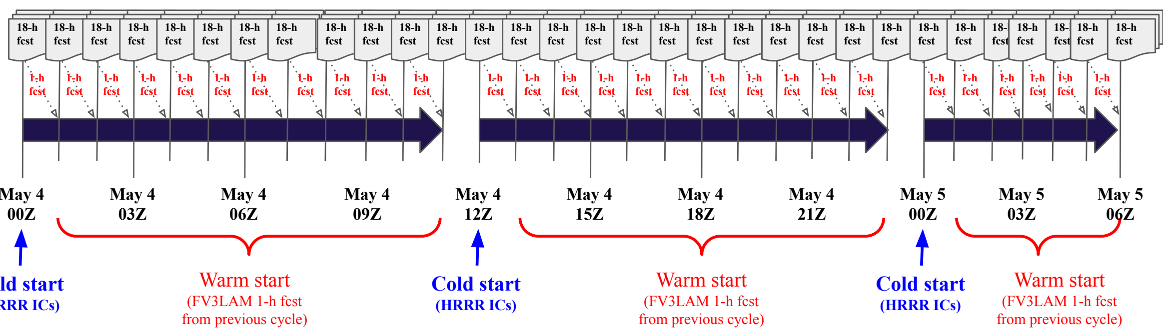


Figure 3.8 - NoDA cycling configuration without data assimilation.

Next, cycling with data assimilation was configured. In this case, the GSI analysis was added at each cycle to provide initial conditions for the model forecasts as indicated in Fig. 3.9, but the cold and warm start strategy was maintained as in the previous configuration. Every 12 hours, HRRR analysis results were used to create the initial conditions for cold start. As indicated in Fig. 3.6, the analysis task uses external model (HRRR analysis) as background to cold start the model in the cycles at 00Z and 12Z May 4th and 00Z May 5th. In the hours in between, the 1-h forecast from previous cycle was used as background in analysis task to warm start the model. Observations in CTLOBS (see Section 3.3) were assimilated every hour as indicated in Fig. 3.9.

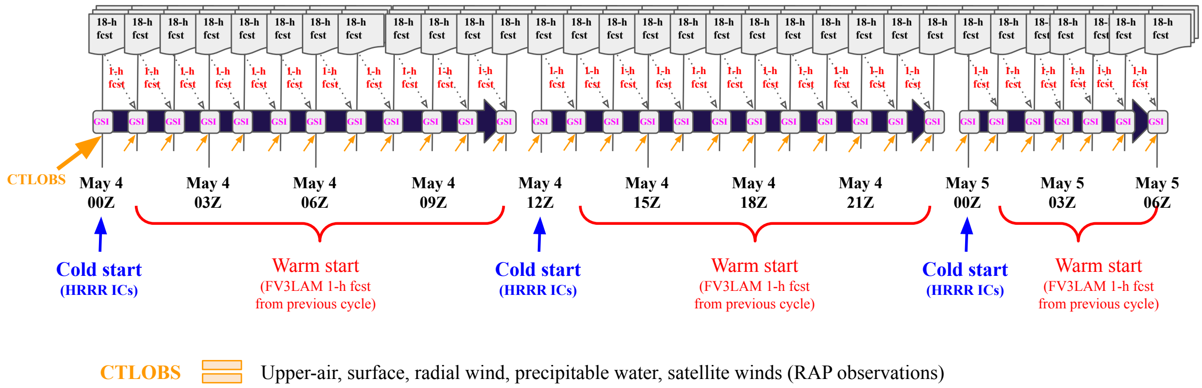


Figure 3.9 - Cycling configuration with hourly Data Assimilation.

In addition, a cycling similar to RAP partial cycles structure was configured (Fig. 3.10) following Hu et al. (2017). For continue cycles, GFS ICs were used to cold start the FV3LAM in the first cycle and warm start was used thereafter. Thus, after the first cycle, restart files were used every hour as background to create the analyses combined with CTLOBS. However, twice a day (at 06Z and 18Z), the restart files used were the 1-h forecast from partial cycles that were executed in parallel to the continuous cycling, as shown in Fig. 3.10 in green. Three parallel cycling (first row in Fig. 3.10) were run every 12 hours with cold start at the first cycle and warm start during the subsequent 5 cycles, executing 1 hour forecast every cycle. The 1-h forecasts from the fifth cycle in the parallel cycling were used to update the atmospheric (dynamical and physical variables in FV3LAM) and surface initial conditions for the continuous cycles at 06Z and 18Z May 4th and 06Z May 5th, as indicated by the green dashed arrows. This was basically to update large scale conditions using more balanced fields in the ICs generation. No digital filter was used. HRRR LBCs were used in the parallel partial cycles. The experiment using this cycling configuration was called GFSics+surf.

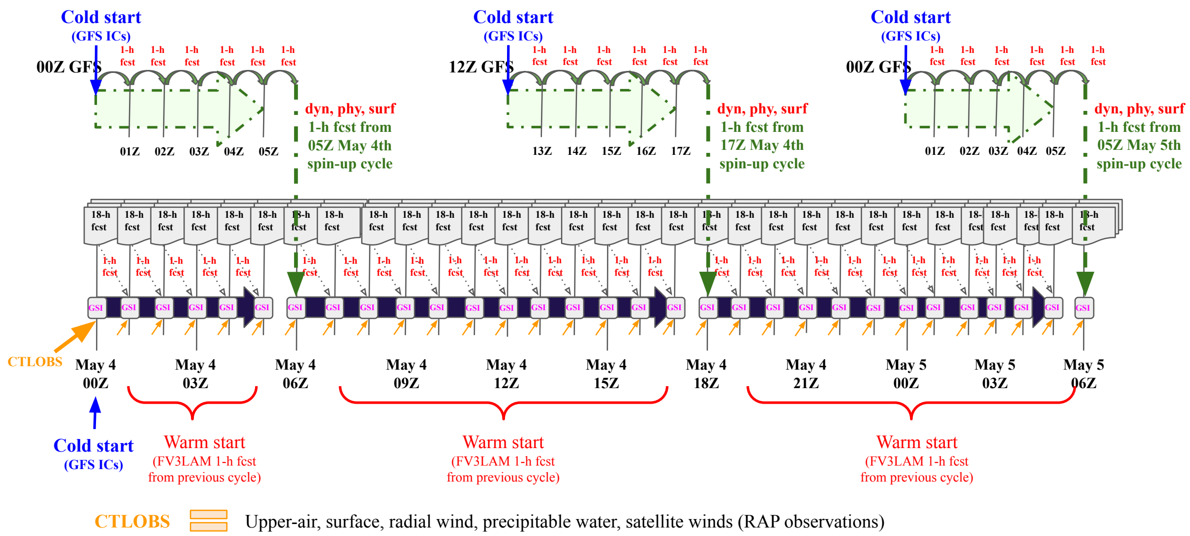


Figure 3.10 - Cycling configuration using GFSFV3 atmospheric and surface conditions starting parallel partial cycles, following RAP partial cycle structure.

Lastly, a cycling configuration was set up identically to the GFSics+surf mentioned above, but keeping the surface fully cycled in the continuous cycles (second row in Fig. 3.11). This is indicated in Fig. 3.11 by the brown dotted arrows in the continuous cycles connecting the cycles initiated at 05Z and 17Z with the cycles 06Z and 18Z, respectively. FV3LAM surface conditions (surf) from previous cycle 1-h forecast are used in the 06Z and 18Z continuous cycles. The green dashed arrows indicate the atmospheric conditions coming from the 1-h forecast from previous cycle in the parallel partial cycle (first row of Fig. 3.11). The experiment using this cycling configuration was called GFSics+CONTsurf.

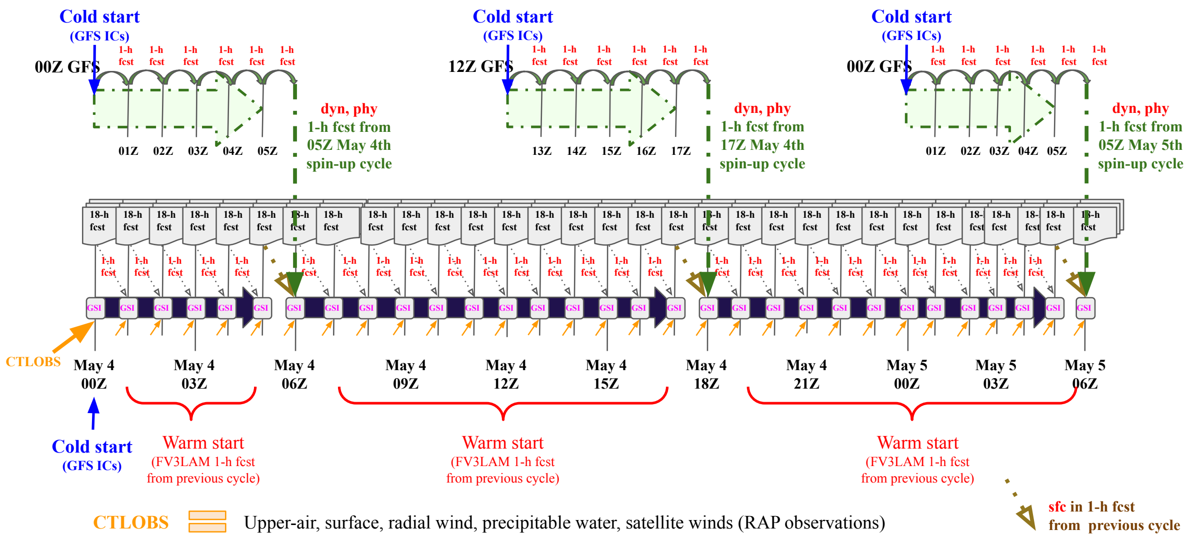


Figure 3.11 - Cycling configuration using GFS atmospheric conditions from parallel partial cycles but fully cycling the surface in the continuous cycle, following RAP partial cycle structure.

3.6 Forecast verification

After the experiments' execution, the Model Evaluation Tools (MET) version 9.0 (JENSEN et al., 2020) was used for forecast verification against observations. MET was developed at the DTC and has been widely used by the NWP community. It includes several tools from re-formatting, re-gridding, masking the input observations to the computation of traditional metrics, wavelet, neighbors and object-based verification as well as analysis tools to process the outputs. Specifically, the PB2NC, Point-Stat, Grid-Stat, Stat-analysis, Method for Object-Based Diagnostic Evaluation (MODE), and MODE-analysis tools were used in this study. Upper-air (ADPUPA) and surface (ADPSFC) RAP observations were used to verify the forecasts.

First, the PB2NC tool was used to convert RAP observations from prepBUFR (prepared BUFR) files to netcdf, the format used by MET tools. For upper-air observations, the time window was from 1 hour and 30 minutes before to 1 hour and 30 minutes after the verification time and it was narrowed to 15 minutes before to 15 minutes after the verification time for surface observations. This was because upper-air data, such as soundings, are available twice a day while surface data are available with a higher frequency. Information on the observations quality provided in the prepBURF files (WMO, 2002) is also taken into account in the PB2NC tool. The maximum quality mark in PB2NC indicates only observations with this quality flag or below will be used in the verification. Here, the maximum quality mark for the observations used was 2.

Next, the Point-Stat tool was used to compute the mean, root mean square error (RMSE), bias, and standard deviation between the point observations and forecasts. Various methods to interpolate model values to the observation locations were tested. Fig. 3.12 shows the bias results for 6-h forecast of 2-m temperature at each cycle when using bilinear and nearest neighbors interpolation methods. Since the difference between the bias calculated with these two methods is very small, the bilinear method was selected for the rest of the calculations.

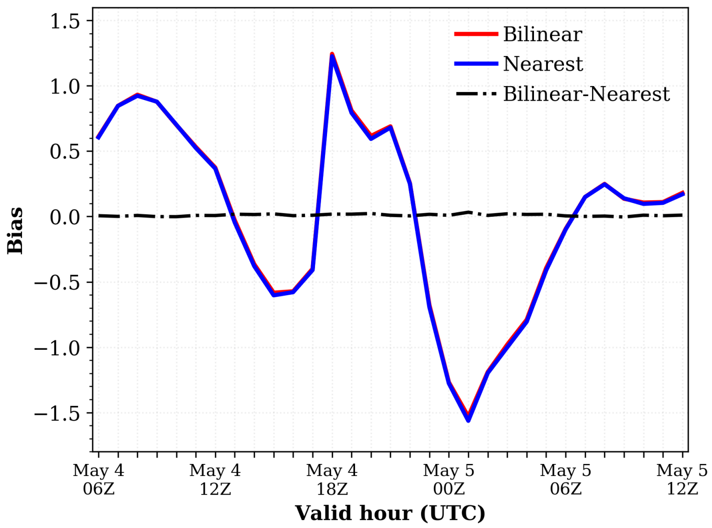


Figure 3.12 - Bias for 6-h forecast of 2-m temperature against METAR and SYNOPTIC reports using bilinear and nearest neighbors interpolation methods.

The calculation of statistics was based on the matched pairs. The analysis of matched pairs found that some stations had observation values diverging significantly from surrounding stations. Thus, a quality control was applied to flag those stations and filter them from the verification. A quality control similar to the gross check in GSI was applied using the observation errors provided in the GSI observation error table (HU et al., 2018). The 2-m temperature and 2-m dew point temperature use the METAR observation errors. The temperature and wind vertical profiles use the observation errors from the radiosonde observations.

Fig. 3.13 shows the spatial distribution of the matched pairs errors (Forecast minus Observation) for 2-m dew point temperature in Kelvin at valid hour 17Z on May 4th, 2020 (0-h forecast), before (Fig. 3.13 A.) and after (Fig. 3.13 C.) applying the quality control. The purple point in Fig. 3.13 B. indicates that the very divergent station was correctly identified. The colorbar in Fig 3.13 A. indicates that the errors before implementing quality control ranged from -60 to 60 K. After the quality control application, this range decreased to -8 to 8 K (see the colorbar in Fig. 3.13 C.). Results in this figure are from the experiment using the GFSreg physics suite and 3DVar data assimilation (see Section 4.2.1).

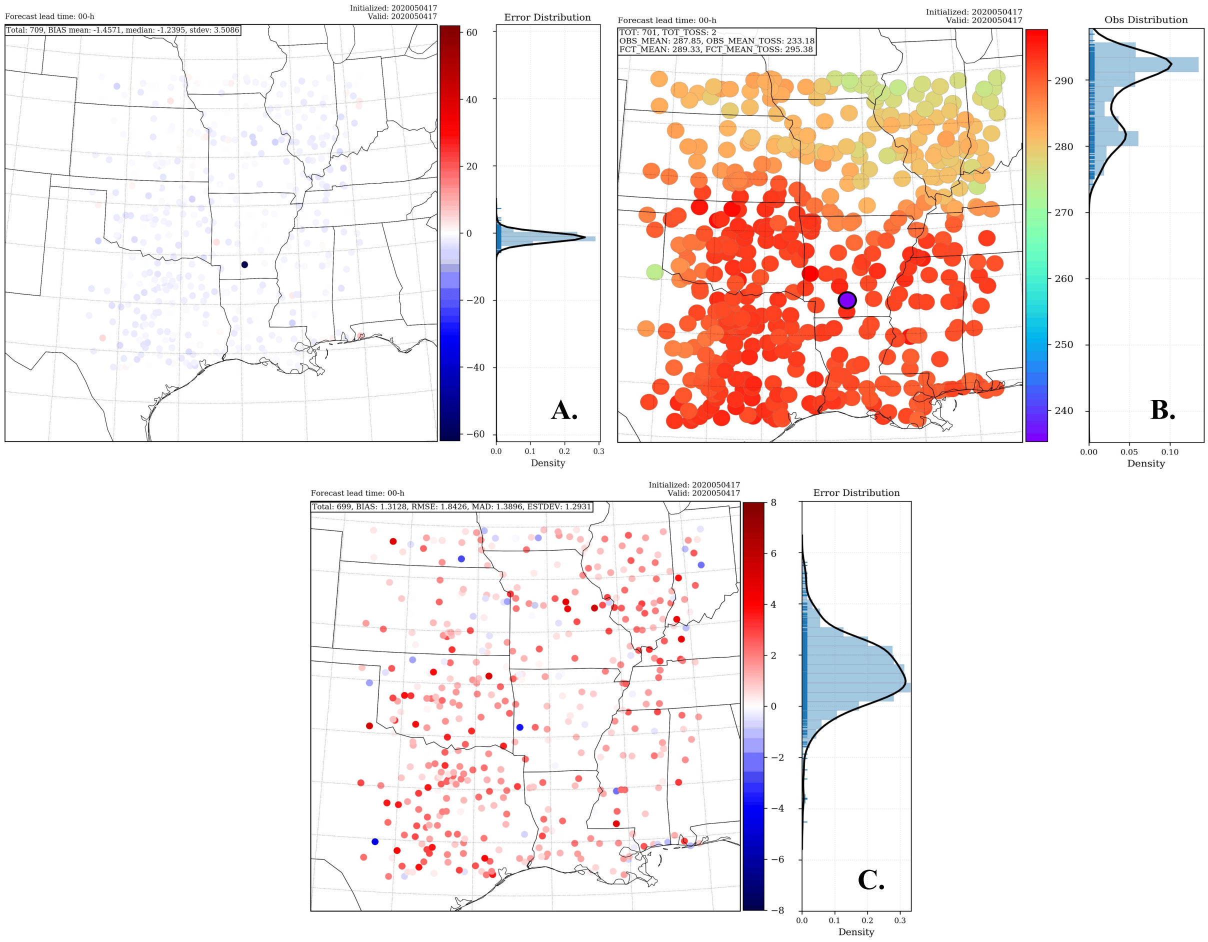


Figure 3.13 - Matched pairs for 2-m dew point temperature (K) forecasts against SYNOPTIC and METAR stations valid at 17Z May 4th, 2020 (0-h forecast), before (A) and after (C) applying quality control. Panel B. shows the spatial distribution of METAR observations at this valid hour.

The Grid-Stat and MODE tools were used to verify the precipitation forecasts against hourly Stage IV precipitation observations for different thresholds: >0.01 in., >0.025 in., >0.05 in., >0.1 in., >0.25 in., >0.5 in., and >1.0 in. Stage IV precipitation data are from the combination of gauges and radar reflectivity mosaics over the Contiguous United States (CONUS) on a 4 km polar-stereographic (1121x881) grid (DU, 2011). These data were downloaded from the NCAR Earth Observing Laboratory data server at https://data.eol.ucar.edu/cgi-bin/codiac/fgr_form/id=21.093.

MODE was also used to verify the predicted composite reflectivity against hourly MSMR composite reflectivity mosaic (optimal method) observations for the thresholds: >5.0 dBZ, >20.0 dBZ, and >35.0 dBZ. MRMS integrates data from several sources into multiple products[1]. Specifically, the MSMR composite reflectivity mosaic product results from the integration of U.S. WSR-88D and TDWRs (Terminal Doppler Weather Radars)

[1] https://www.nssl.noaa.gov/projects/mrms/

radars and the Canadian radar network into a mosaic covering the CONUS and southern Canada, with 1 km and 2 min of spatial and temporal resolution, respectively (ZHANG et al., 2016). These data are available on the Iowa Environmental Mesonet archives at https://mesonet.agron.iastate.edu/archive/.

4 Numerical experiments and results

In this section, CCPP suites and GSI configurations are tested and evaluated as well as cycling configurations and hybrid coefficients for the vertical levels.

4.1 Common Community Physics Package (CCPP) suites

The CCPP contains a set of physical schemes and a framework that facilitates the interaction between the physics and a numerical model (Bernardet et al., 2020). The latest CCPP released, CCPP v5.0 (`https://dtcenter.ucar.edu/GMTB/v5.0.0/sci_doc/index.html`), includes the GFSv15 physics suite (GFS_v15p2) used for GFS v15 operation and the suite targeted for RRFS, RRFSv1 (RRFS_v1alpha). RRFSv1 is based on convection allowed physical schemes implemented in HRRR. These two were tested along with two suites from CCPP v4.0, the GSDsar (FV3_GSD_SAR) developed at the GSL based on RAP and HRRR physical schemes and the GFSreg suite (FV3_GFS_2017_-gfdlmp_regional)—also based on GFS v15 physics. Table 4.1 summarizes the CCPP suites tested in this study with the set of physical schemes included in each suite. The default configuration for each CCPP suite was used.

Table 4.1 - Common Community Physics Package (CCPP) suites tested.

Physical process	**Physics suites**			
	RRFSv1	**GFSv15**	**GSDsar**	**GFSreg**
Deep Cu	*off*	GFS sa-SAS for deepcnv	*off*	GFS sa-SAS for deepcnv
Shallow Cu	MYNN-EDMF	GFS sa-MF for shalcnv	MYNN-EDMF	GFS sa-MF for shalcnv
Microphysics	Thompson	GFDL	Thompson	GFDL
PBL/TURB	MYNN-EDMF	Hybrid EDMF	MYNN-EDMF	Hybrid EDMF
Radiation	RRTMG and SGSCLOUD	RRTMG	RRTMG	RRTMG
Surface Layer	GFS	GFS	GFS	GFS
Land	Noah-MP	Noah	RUC	Noah
Gravity Wave Grag (GWD)	uGWD	uGWD	uGWD	uGWD
Ocean	NSST	NSST	NSST	NSST
Ozone	NRL 2015	NRL 2015	NRL 2015	NRL 2015
Water Vapor	NRL 2015	NRL 2015	NRL 2015	NRL 2015

4.1.1 HRRRDA experiments

The first set of experiments evaluated the FV3LAM capability using RRFSv1 and GFSv15 CCPP physics suites for convection predictions. The cycling type HRRRDA, which has no data assimilation, was used for the experiments with HRRR analysis as initial condition. Using HRRR analysis as initial condition in each cycle gives us a chance to evaluate the capability of the model with each physical suite for storm forecast when the best initial condition is used. Fig. 4.1 shows 2- and 6-h forecast of composite reflectivity in HRRRDA experiments from 19Z cycle on May 4th 2020, with overlapping MRMS observations valid at 21Z on May 4th and 01Z on May 5th, respectively. The two CCPP suites represent the squall line differently. The RRFSv1 experiment showed stronger cells and smaller coverage while the GFSv15 gave weaker and smoother cells with larger convection coverage and more spurious cells. Both experiments captured the convective initiation over northeast Oklahoma and the convection between northeast Arkansas and west Tennessee, however the extent and intensity of convective cells was overestimated. HRRRDA with GFSv15 simulated the evolution of the squall line from northeast Oklahoma to southeast Oklahoma and northwest Arkansas slightly better, but the extent of the squall line to the south central Oklahoma was not well captured in any of the experiments and the strongest convective cells in the squall line are located ahead of the observations. Also in the 6-h forecast, the convection systems over Missouri, Illinois, and western Tennessee and Kentucky was underestimated in both experiments.

In order to quantitatively identify the suite that yields better forecasts, the median of maximum interest (MMI (F+O)) (DAVIS et al., 2009) was also analyzed. This metric results from the median between the maximum interest from each observed object with all predicted objects (MIF) and the maximum interest from each predicted object with all observed objects (MIO). It takes into account all attributes used in the total interest calculation, summarizing them into a single value. The forecast with the best quality or in greater agreement with the observations will give MMI (F+O) values close to one. Otherwise, the value will close to zero. This metric is calculated in MODE for composite reflectivity higher than 5 dBZ and results are shown in the lower left corner of the panels in Fig. 4.1. The MMI (F+O) indicates that the HRRRDA cycling configuration with the RRFSv1 suite produced better storms for both 2- and 6-h forecast. This case study shows FV3LAM with the targeted RRFS physics suite has great potential for storms forecast, but it also suggest many aspects of the suite still need to be improved in current developments of RRFS.

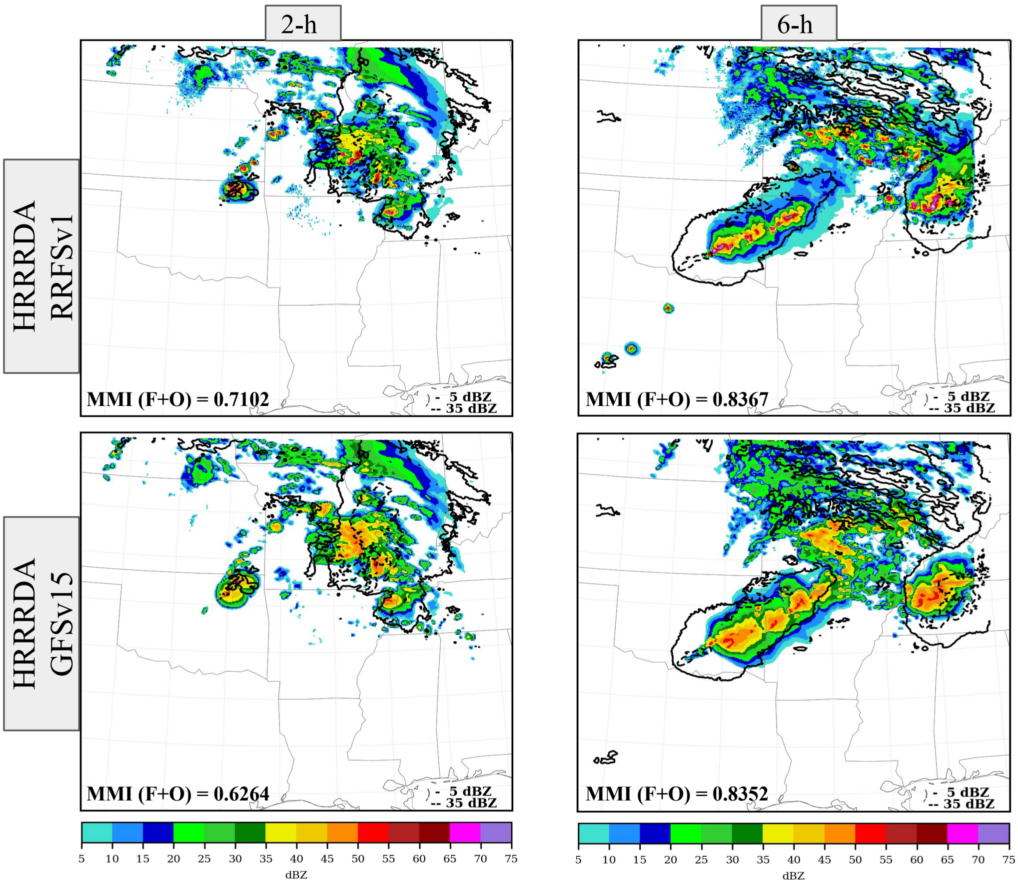


Figure 4.1 - 2-h (left panels) and 6-h (right panels) composite reflectivity forecasts from the experiments using the HRRRDA cycling configuration with RRFSv1 (upper panels) and GFSv15 (lower panels) physics suites, initialized at 19Z May 4th 2020. Solid and dashed black lines are the 5 and 35 dBZ MRMS observation contours, respectively. MMI (F+O) results are shown in the lower left corner of each panel.

4.1.2 NoDA experiments

Experiments using the NoDA cycling configuration and the four CCPP suites described in Table 4.1 were conducted to further study the impact of different CCPP physics suites and to be used as baselines for evaluating the impact from data assimilation. As in Fig. 4.1, Fig. 4.2 presents the 2- and 6-h composite reflectivity forecasts from the 19Z cycle on May 4th, 2020 with overlapping observations, but from NoDA experiments using RRFSv1 and GFSv15 suites. Both NoDA experiments predicted main features of the squall line and even without data assimilation, but the convection initiation over northeast Oklahoma was misplaced in the 2-h forecast with RRFSv1 and the convective cells over Texas were missed in the 6-h forecast with GFSv15. Compared with GFSv15, the experiment with the RRFSv1 gave a better simulation of the convection over eastern Missouri and Illinois in the 2-h forecast as well as the convective initiation over Texas at 01Z on May 5th (6-h forecast). MMI (F+O) indicates that the NoDA GFSv15 experiment is better in the 2-h forecast with a value of 0.7089, but NoDA RRFSv1 is superior at 6-h forecast with an

MMI (F+O) value of 0.7888, while for GFSv15 the MMI (F+O) is 0.7206.

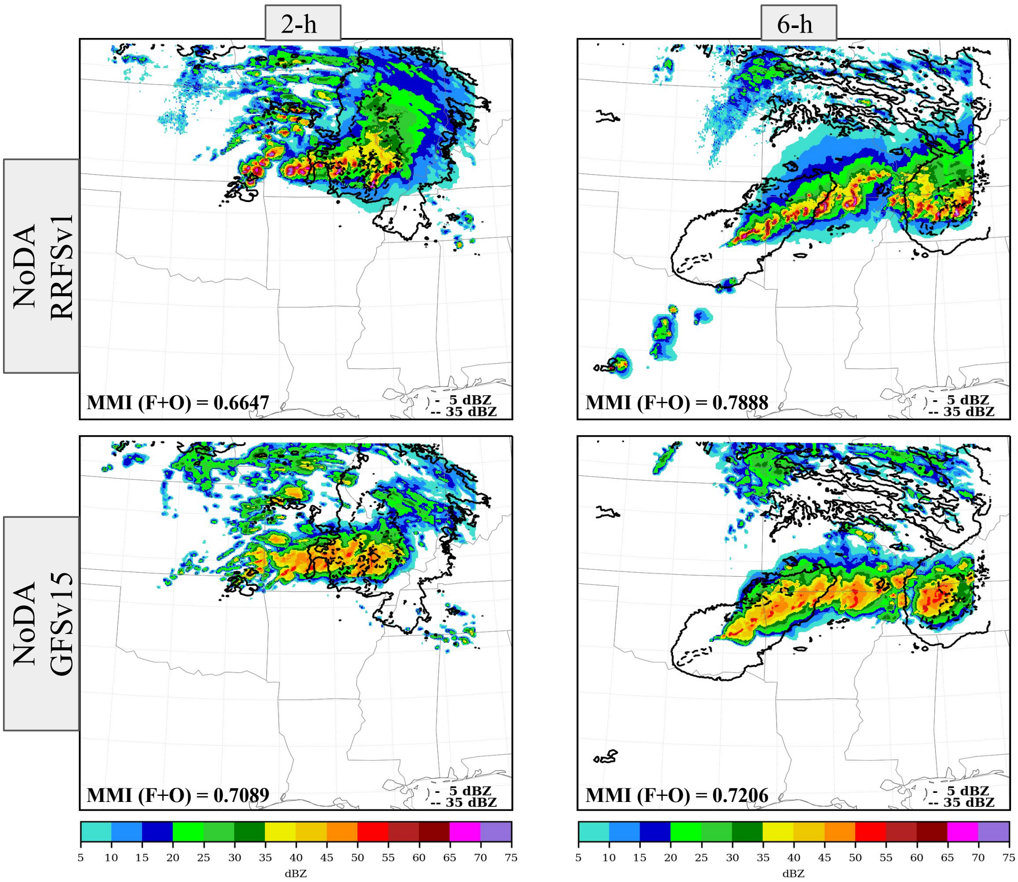


Figure 4.2 - As in Fig. 4.1, but for the NoDA experiments using RRFSv1 (upper panels) and GFSv15 (lower panels) physics suites.

NoDA results using GSDsar and GFSreg CCPP suites are presented in Fig. 4.3. A good representation of the convection over eastern Missouri and Illinois can be observed in NoDA GSDsar for 2-h forecast as well as over western Tennessee and Kentucky for 6-h forecast, but the squall line was missed in both forecasts lead times. The GFSreg suite predicted a better structure of the squall line, but the coverage was overestimated over Texas in both forecast lengths as well as over northern Arkansas, where it is difficult to differentiate the convection associated to the squall line initiated from northeast Oklahoma or evolution of the convective systems occurring between eastern Missouri and western Tennessee and Kentucky. MMI (F+O) corroborates these results with a value of 0.6894 in the 2-h forecast using GSDsar and 0.7418 using the GFSreg suite.

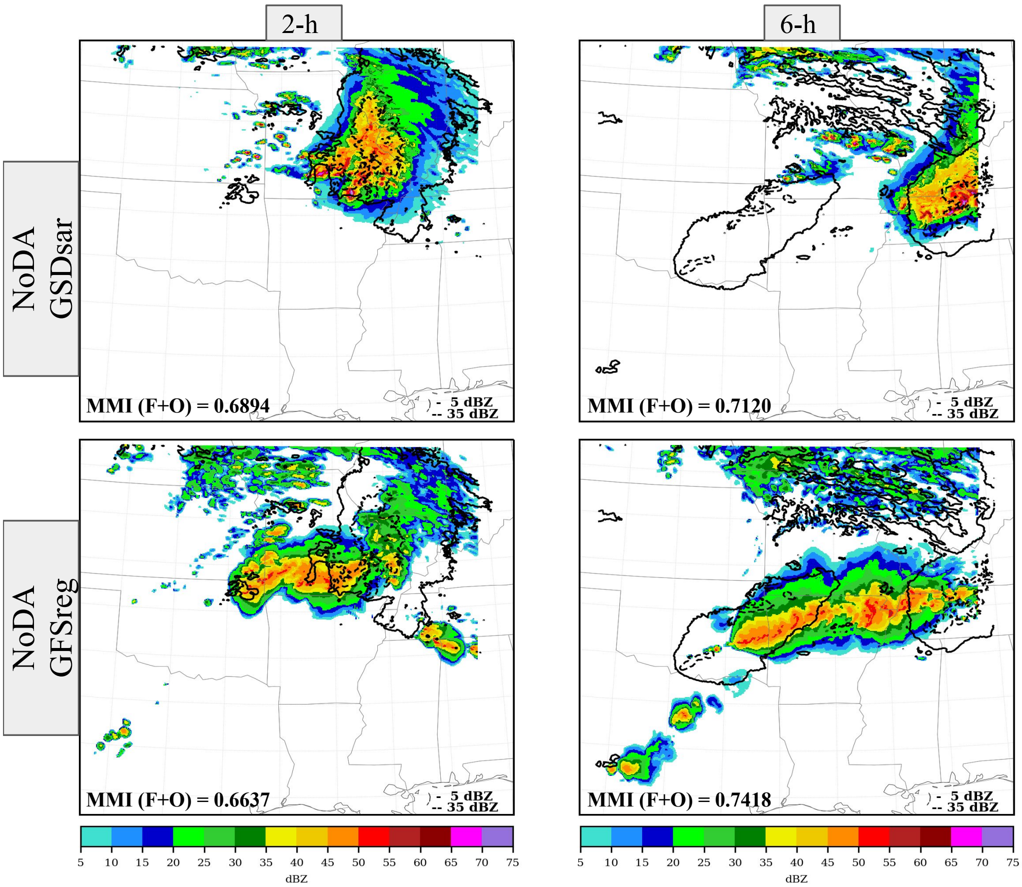


Figure 4.3 - As in Fig. 4.2, but for the NoDA experiments using the GSDsar (upper panels) and GFSreg (lower panels) physics suites.

4.2 GSI configurations

GSI is the analysis system used operationally at NCEP for the global and regional models. It was initially developed by the NCEP/EMC (WU et al., 2002) and was implemented as analysis component in the operational GFS with GFS model in May 2007 (KLEIST et al., 2009) and in the operational RAP with WRF model in May 2012 (BENJAMIN et al., 2016). Over the years, many functionalities have been added and it has shown to be a robust system, capable of performing a skillful analysis on synoptic and mesoscale scales. In order to explore the capability of GSI working with FV3LAM to generate the analysis that produces better forecasts, several GSI parameters were tested with the hourly data assimilation cycling (Fig. 3.9) and results are presented bellow.

4.2.1 The ration of analysis grid to background grid

The ratio between the analysis and background grids provides the resolution in which the analysis is generated. In this study, values 3 (G3) and 1 (G1) were tested which produce analyses grid of 9 and 3 km, respectively. Since 3 km is also the model horizontal resolution, it allows for more details in the analysis increments. Results confirm that hypothesis. Here, only results using 3 km as the analysis grid from the 3DVar analysis are presented. Fig. 4.4 shows the analysis increments for specific humidity and temperature at the first level above surface. The background seems to be drier over Oklahoma and Texas, wetter over Arkansas and southern Missouri, and warmer all over Missouri. The analysis reduced those bias by adjusting the humidity and temperature toward observations in these areas.

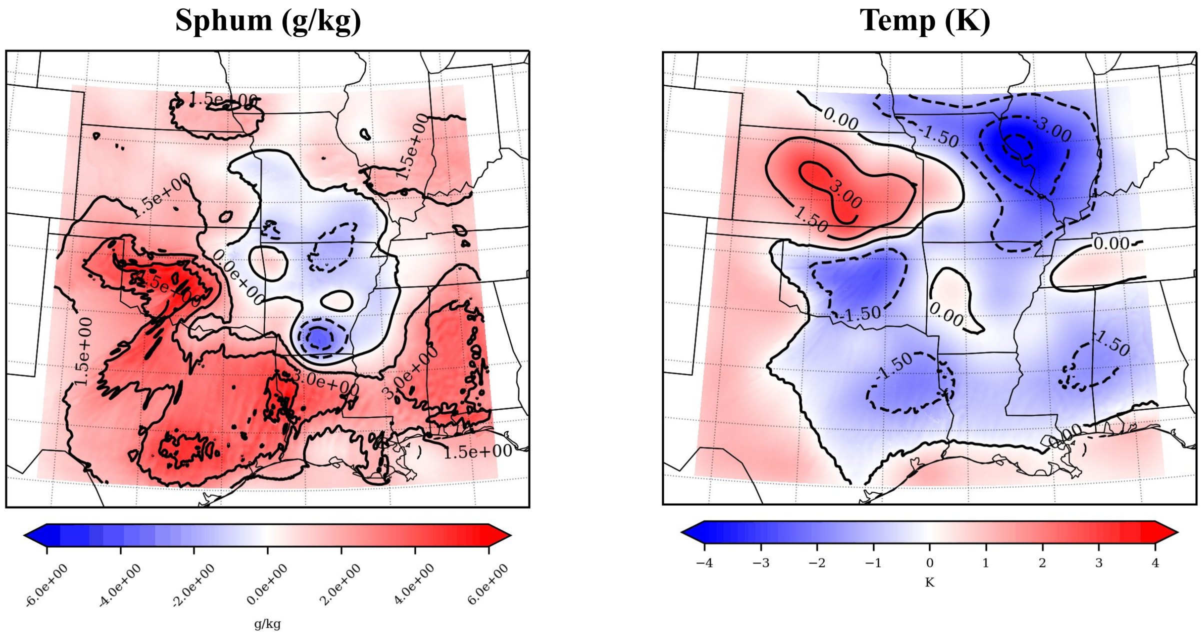


Figure 4.4 - Analysis increment for the specific humidity (g/kg) (left) and temperature (K) (right) at the first level above surface (model level 63) for GSI 3DVar analysis at 21Z on May 4th, 2020.

The mean RMS and bias of the background and analysis against to METAR observations are presented in Fig. 4.5. These are for the specific humidity (upper panels) and temperature (lower panels) in each analysis during the execution period. Lower RMS and bias values of the analysis (OmA–red line) are shown in all the cycles for the temperature and in most of the cycles for the specific humidity, indicating that the analyses are closer to the observations.

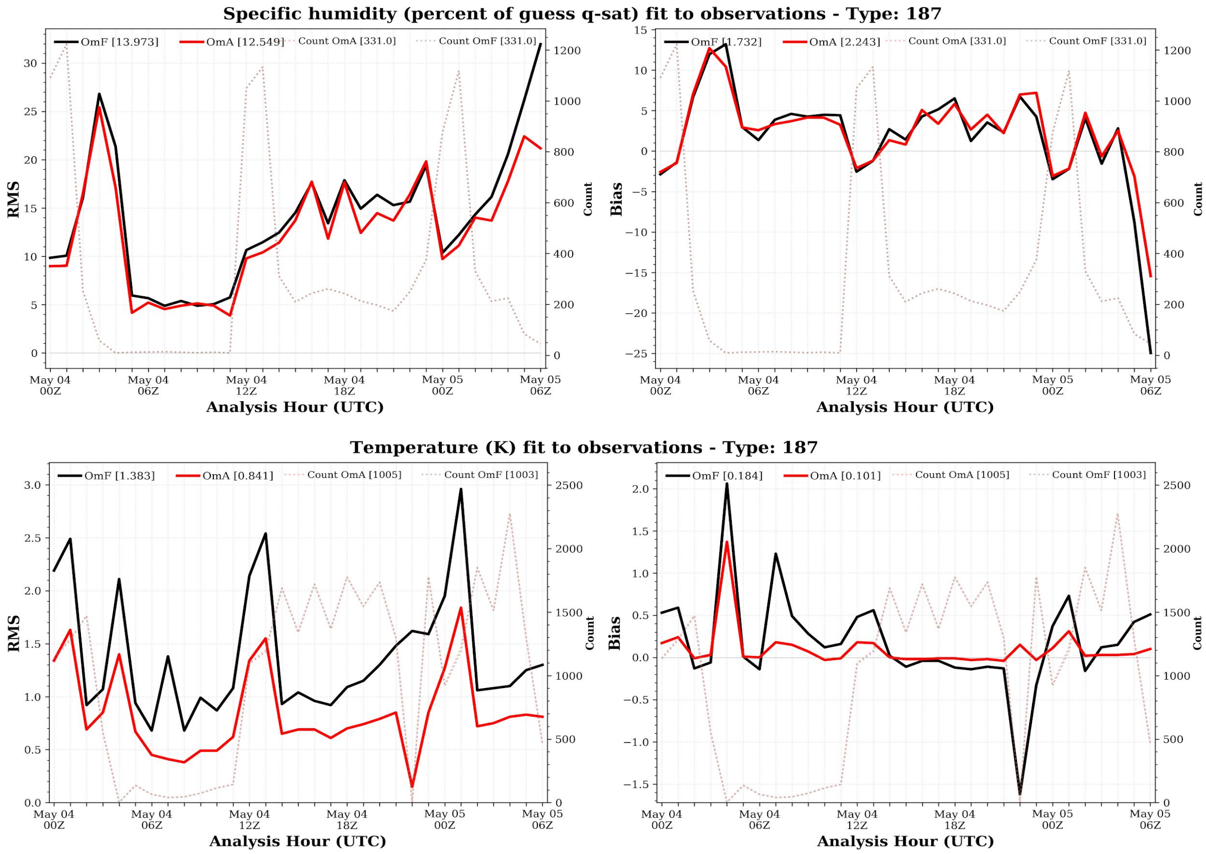


Figure 4.5 - Mean RMS and bias of the background (OmF) and analysis (OmA) against METAR observations (type: 187) for analysis in cycles from 00Z May 4th to 06Z May 5th, 2020. Upper panels are results for the specific humidity (g/kg) and the lower panels present temperature (K) results.

The 2- and 6-h composite reflectivity forecasts initialized at 19Z May 4th, 2020 from the experiment using the RRFSv1 suite with 3DVar analysis over G1 grid with CTLOBS observations are shown in the lower panels of Fig. 4.6. Same results from the NoDA experiment using RRFSv1 (previously presented) are shown in the upper panels. GSI analysis improves the convection initiation and evolution by reducing the spurious convection over Arkansas. Forecast with data assimilation also distinguishes the reflectivity associated with the squall line and the convection over western Tennessee and Kentucky better. However, two spurious cells were developed over southwest Missouri at 2-h forecast and some spurious cells were over Texas at 6-h forecast. MMI (F+O) results still indicate better skill without 3DVar data assimilation, which will be addressed in the next experiments using hybrid GSI analysis.

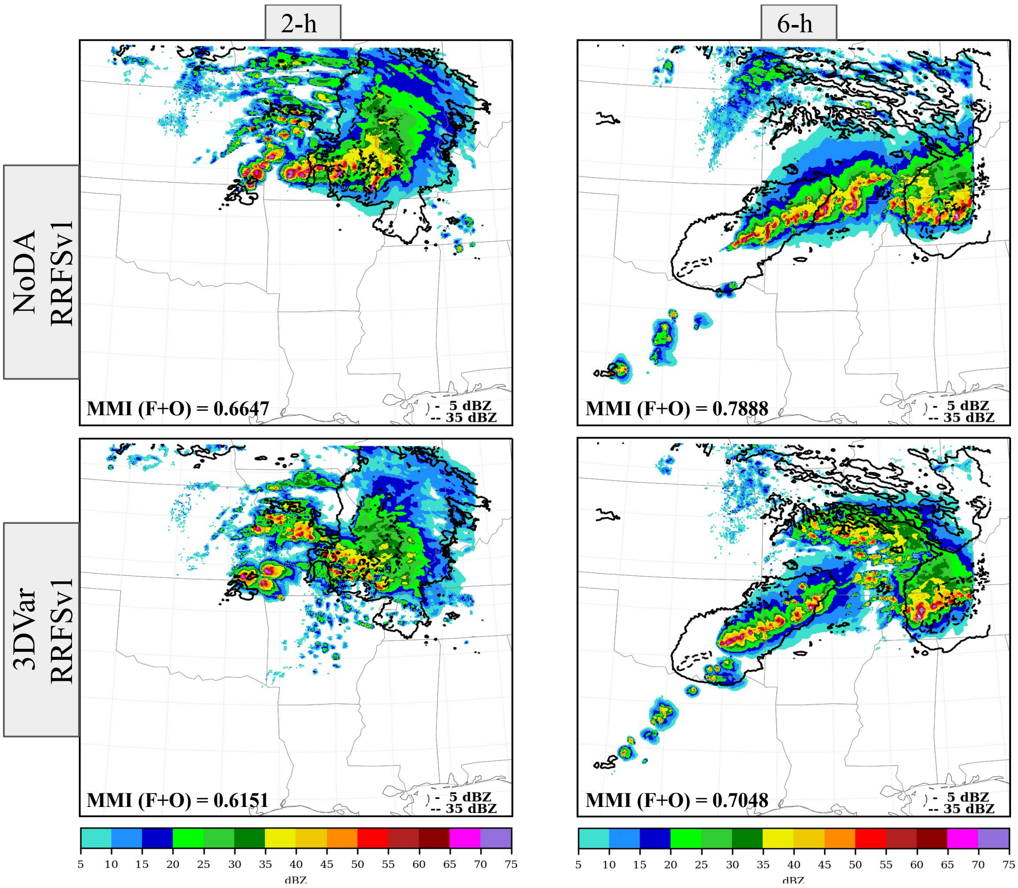


Figure 4.6 - As in Fig. 4.3, but for the experiments using the RRFSv1 suite with 3DVar, G1, and CTLOBS data assimilation (lower panels) and without it (upper panels).

Same as Fig. 4.6, but using the GFSv15 suite are shown in the lower panels of Fig. 4.7. Using 3Dvar with GFSv15 improves the storm forecast over NoDA in general. The convection associated to the squall line evolution was better represented in the 6-h forecast when using 3DVar data assimilation, which may have contributed to improve the MMI (F+O) results reaching a value of 0.7692. But it gives a smaller coverage of the convection, particularly over eastern Missouri and western Illinois, Kentucky, and Tennessee. The convection in those areas were underestimated in both forecast lengths. In addition, the convection initiation over Texas was delayed, which may be related to a bad representation of the dry line in the forecasts.

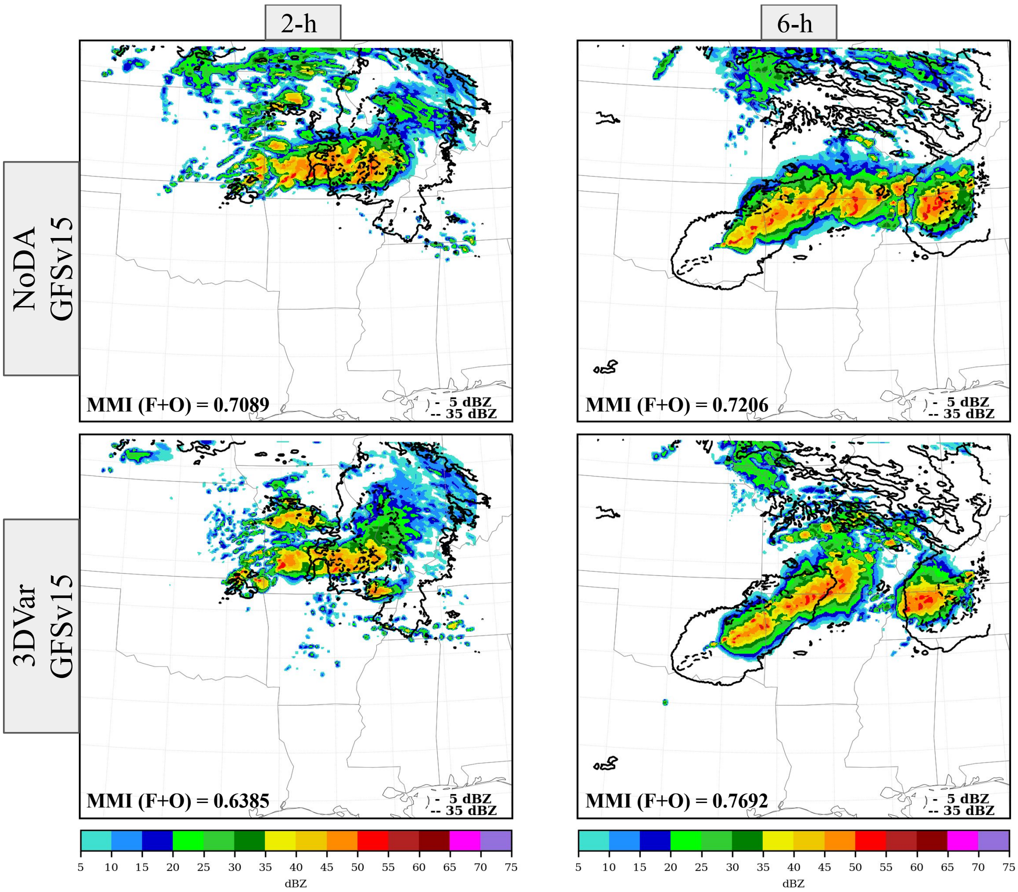


Figure 4.7 - As in Fig. 4.6, but for the experiment using the GFSv15 suite with 3DVar, G1, and CTLOBS data assimilation in the lower panels.

Fig. 4.8 presents results using 3DVar with G1 and CTLOBS data assimilation and the GSDsar and GFSreg physics suites. These experiments were run with PBL pseudo-observation option on in GSI analysis, which were different from previous experiments that had this option off. A bug in the calculation of the planetary boundary layer (PBL) height was found in the GSI code when used with FV3LAM background. Because of wrong PBL height, the surface temperature and moisture observations could impact atmosphere up to 500 mb that leading to stronger cells and wider coverage in the forecast initialized from those analyses results as shown in this figure. This was fixed by GSL and the new code was used to conduct the experiments in Section 4.2.4. All other experiments presented here were performed with this option off. Even though the results in Fig. 4.7 have this error embedded, they are presented here to show how an error in the data assimilation can affect the convection forecasts.

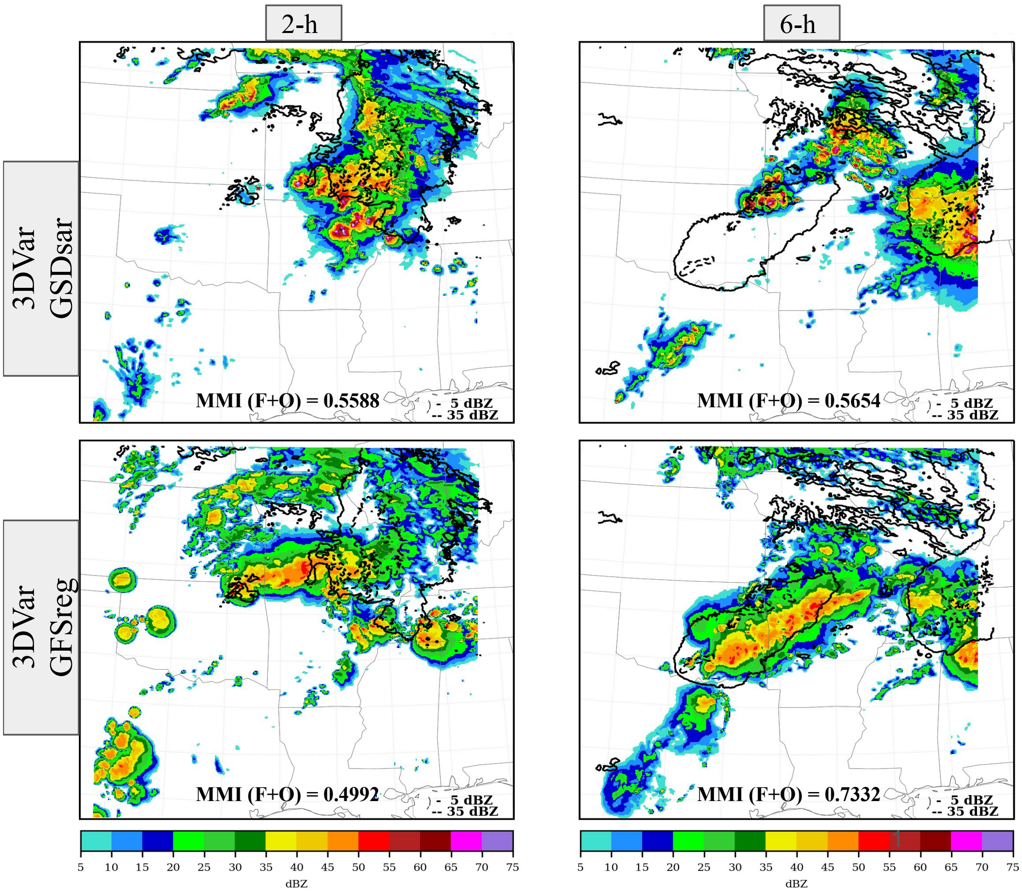


Figure 4.8 - As in Fig. 4.7, but for experiments using the GSDsar (upper panels) and GFSreg (lower panels) suites with a bug version 3DVar, G1, and CTLOBS data assimilation.

4.2.2 Hybrid data assimilation

For convective scales analysis, a more appropriate approach is to run hybrid EnVar analysis, which combines the static background error covariance (BEC) with an ensemble BEC (EnBEC), due to nonlinearities and flow dependence of the storms. The structures of convective scale phenomena are dependent on the actual state of large-scale forcing and also to strong vertical velocity fluctuations present in convective systems (GUSTAFSSON et al., 2018). The hybrid EnVar analysis provides greater flexibility and introduces an alternative for the modeling of non-Gaussian errors (LORENC, 2014). This method has parameters to control the relative weight that is given to each covariance matrix. In GSI, a parameter (beta_s) is introduced via namelist to specify the relative weight given to the static BEC. This parameter can have values up to 1 and 1-beta_s indicates the weight given to the EnBEC.

Different weights of the EnBEC were tested starting with 100% or pure ensemble (0% static), 90% (10% static), 75% (25% static), 50% (50% static), and 0% (100% static). The specific humidity and temperature analysis increments using 100%, 75%, and 0% (3DVar)

EnBEC weights with RRFSv1 CCPP suite are shown in Fig. 4.9 for 21Z cycle on May 4th, 2020. More flow dependence features can be clearly observed as the weight of EnBEC increases. Analysis increments using 3DVar are smoother than ones using 100% EnBEC, which show detailed flow-dependent contours. Besides, analysis increments in 100% and 75% experiments seemed to better adjust the temperature and humidity over northern Arkansas and Oklahoma.

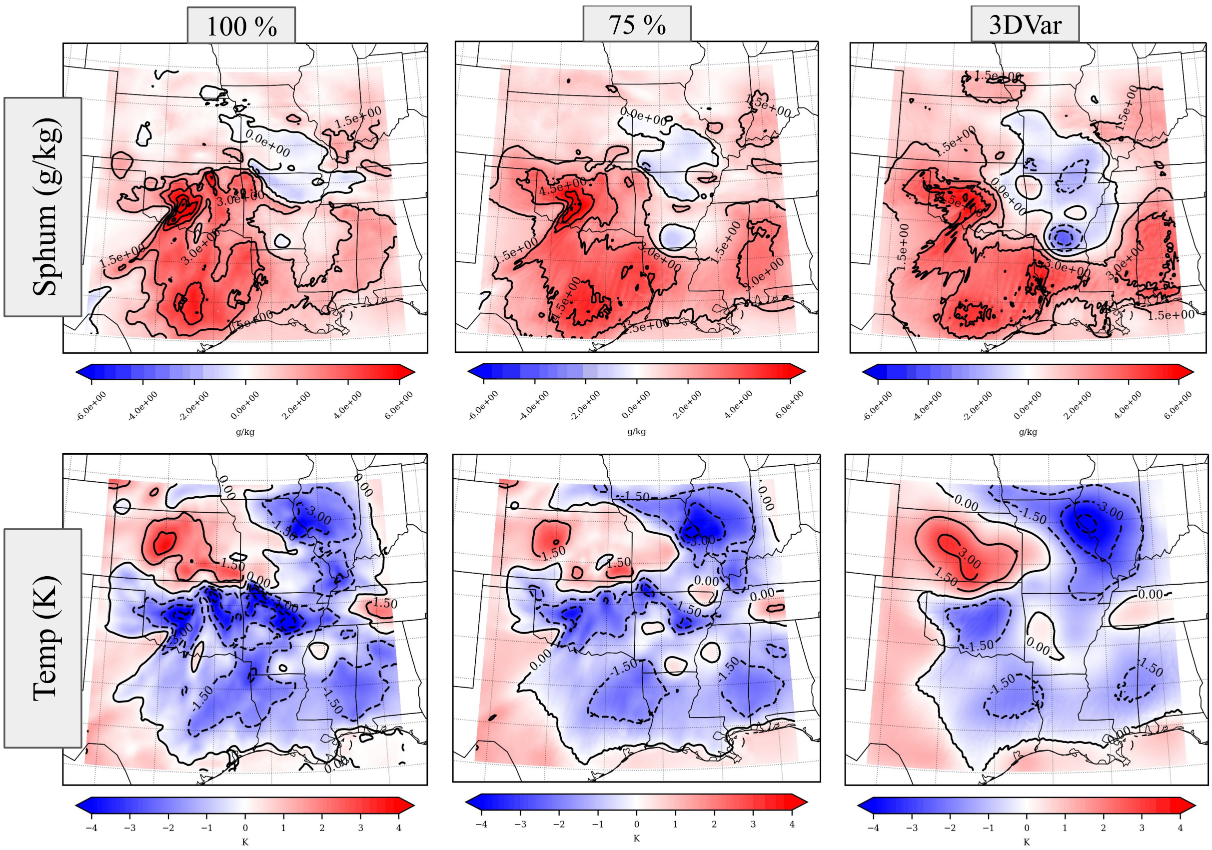


Figure 4.9 - Analysis increment for the specific humidity (g/kg) (upper panels) and temperature (K) (lower panels) at the first level above surface (model level 63) for 21Z May 4th, 2020, using RRFSv1 CCPP suite and 100% EnBEC (left column), 75% EnBEC (middle column), and 3DVar (right column).

The 2- and 6-h composite reflectivity forecasts for the same experiments in Fig. 4.9 are presented in Fig. 4.10. The erroneous convection over southwest Missouri in 3DVar was reduced as the weight of the EnBEC increases. Less spurious convection and a better representation of the convection over eastern Missouri were achieved in 2-h forecast with 100%, although a degradation is observed over western Missouri and central Tennessee. With 75% of the EnBEC, the simulated convection over eastern Missouri in the 6-h forecast matched better with the observations than ones with 100% EnBEC or 3DVar and the evolution of the convective system in this area seemed to be better represented with overall less convection in the domain. The MMI (F+O) results for these three experiments show that the highest values were achieved with 100% of the EnBEC, in both the 2-

(0.6798) and 6-h (0.7670) forecast.

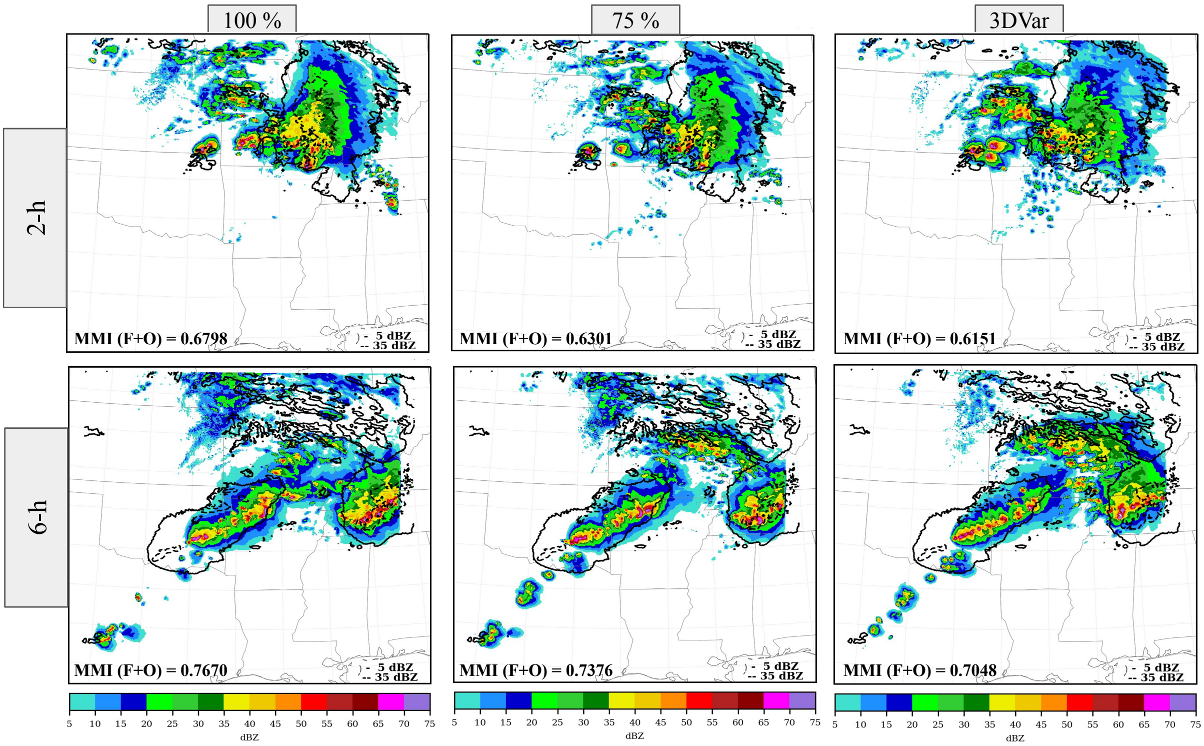


Figure 4.10 - 2- (upper panels) and 6-h (lower panels) composite reflectivity forecasts initialized at 19Z May 4th 2020 from the experiments using the RRFSv1 physics suite with 100% of the EnBEC (left panels), 75% of the EnBEC (middle panels), and 3DVar (right panels) with G1 and CTLOBS data assimilation. Solid and dashed black lines are the 5 and 35 dBZ MRMS observations contours, respectively.

Although the 100% experiment showed the best results in Fig. 4.10 for the storm forecast, following experiments were conducted using 75% in order to leverage the influence of the static background similar to operational options and evaluate the forecasts when other GSI parameters are tested.

Fig. 4.11 includes previous results using 3DVar data assimilation with the GFSv15 suite (upper panels) along with a new experiment using 75% of the EnBEC with the GFSv15 suite (lower panels). The forecast with 75% of the EnBEC produced a better structure of the squall line and slightly better positioning of the convection over western Tennessee in the 6-h forecast as well as over southwest Missouri in the 2-h forecast. The MMI (F+O) of 0.8328 for 6-h forecast is as high as in HRRRDA GFSv15, although the skill in the 2-h forecast is slightly lower than one using 3DVar (lower MMI (F+O)).

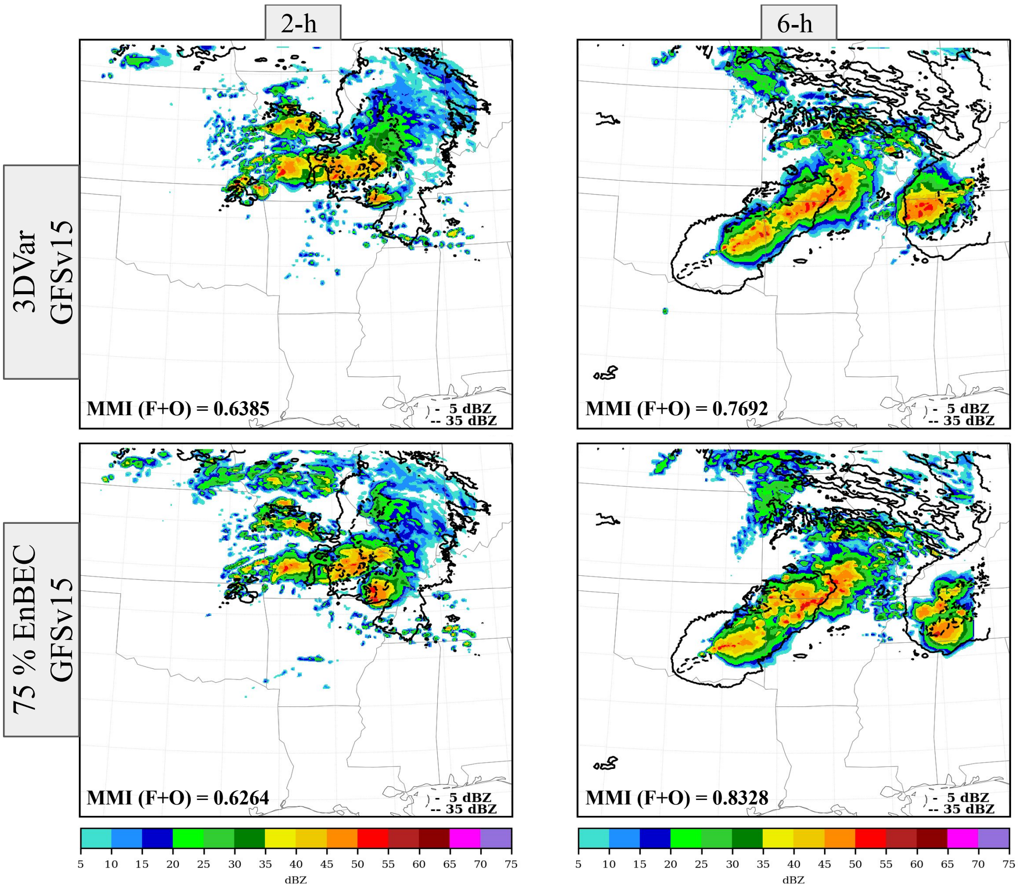


Figure 4.11 - As in Fig. 4.8, but for experiments using the GFSv15 suite with 3DVar (upper panels) and 75% of the EnBEC (lower panels) with G1, and CTLOBS data assimilation.

4.2.3 Supersaturation removal

Since an overestimation of the convection with strong cells were observed in many of the experiments, a GSI parameter that can limit the supersaturation in the background specific humidity was tested. This function can be activated via setting GSI namelist option, the clip_supersaturation, to true. It is designed to restrict the background specific humidity to the minimum between the background specific humidity and the saturation specific humidity calculated using the background fields, in each outer loop during GSI analysis (CIMSS, 2014). Fig. 4.12 shows the difference of the specific humidity (g/kg) between the analyses for 21Z cycle on May 4th, 2020 without and with this function activated at various levels. These experiments are using the RRFSv1 suite with 75% of EnBEC hybrid analysis. The experiment with supersaturation removal shows that supersaturation was removed in the analyses mostly in areas where the convection was overestimated.

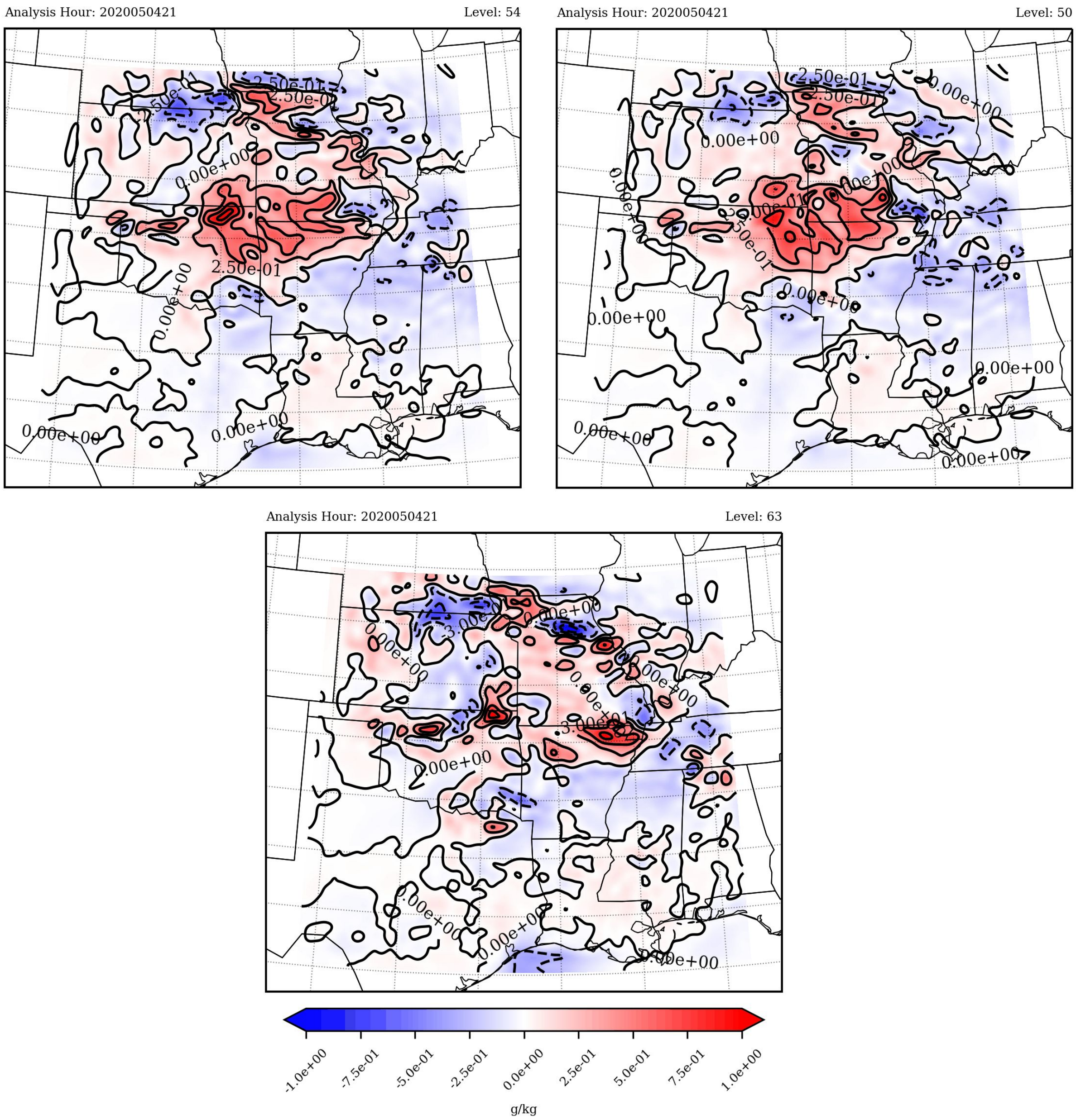


Figure 4.12 - Difference of the specific humidity (g/kg) for 21Z cycle on May 4th, 2020 between analyses without and with the supersaturation clipping activated (NoCS - CS), at model levels 54 (upper left panel), 50 (upper right panel), and 63 (lower panel).

However, the 2- and 6-h composite reflectivity forecasts in Fig. 4.13 indicate neutral to slightly negative impact when using this function in the analysis. The results with this option on are shown in the lower panels and the results without activating the function are upper panels, which are from the previous experiment using RRFSv1 suite with 75% of the EnBEC, and hereinafter is called CTL. When the supersaturation is removed in the analyses, more spurious convection and relatively more intense individual cells were developed in both forecast hours, although a better evolution of the squall line was observed in the 6-h forecast. MMI (F+O) results indicate that activating this parameter can give a better forecast skill for storms in 6-h forecast, with a value of 0.8048.

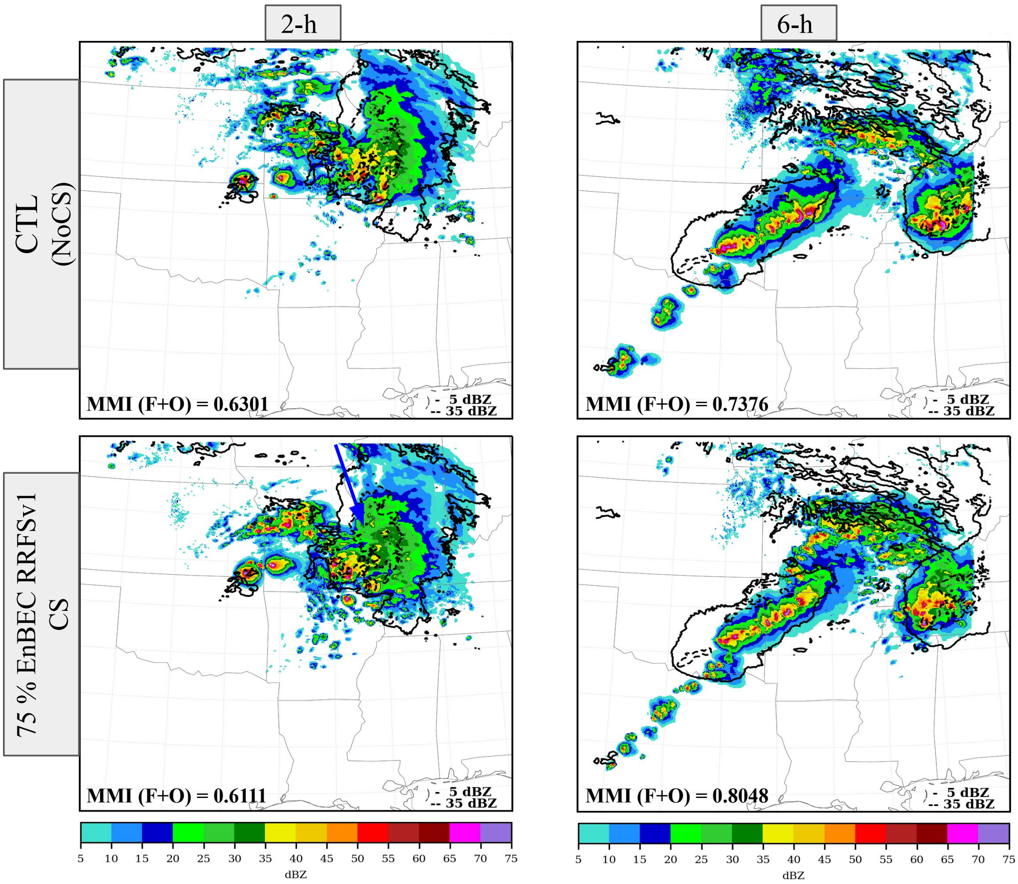


Figure 4.13 - As in Fig. 4.11, but for the experiments using the RRFSv1 physics suite with 75% of the EnBE, G1, and CTLOBS data assimilation with (lower panels) and without (upper panels) the supersaturation removal activated.

4.2.4 The impact of adding PBL Pseudo observations

After the bug in GSI PBL pseudo-observations was fixed, an experiment was conducted to evaluate the impact of adding those PBL pseudo-observations based on the surface temperature and moisture observations. Firstly, this function identifies the PBL height using the model vertical levels and background. Then it computes the 2-m temperature and 2-m moisture observation innovations (OmB). The calculated innovations are treated as new observations located every 20 hPa vertically up to the level corresponding to 75% of the PBL height (BENJAMIN et al., 2016). The PBL pseudo observations are used in RAP/HRRR operation to give more realistic analysis for the levels inside the PBL. They can be activated through options in the GSI namelist.

Storm forecast using the PBL pseudo-observations with 75% of the EnBEC and the RRFSv1 suite are presented in the lowers panels of Fig. 4.14. Results without this option are shown in the upper panels (CTL). The forecast with the PBL pseudo-observations clear add more storms in its forecast than ones without. It overestimated the convection in southern Missouri, western Tennessee and Kentucky, and Texas in both forecast lengths. The convection initiation over northeast Oklahoma and southeast Kansas was also over-produced in the 2-h forecast and the squall line is not well represented in the 6-h forecast after adding PBL pseudo-observations. Since the 6-h forecast has better coverage of the squall line and the convective system over western Tennessee and Kentucky, the MMI (F+O) for this hour is 0.8032, showing better quality than those in the experiment without PBL pseudo-observations. Clearly, more tuning and testing are needed for applying this techniques in RRFS system.

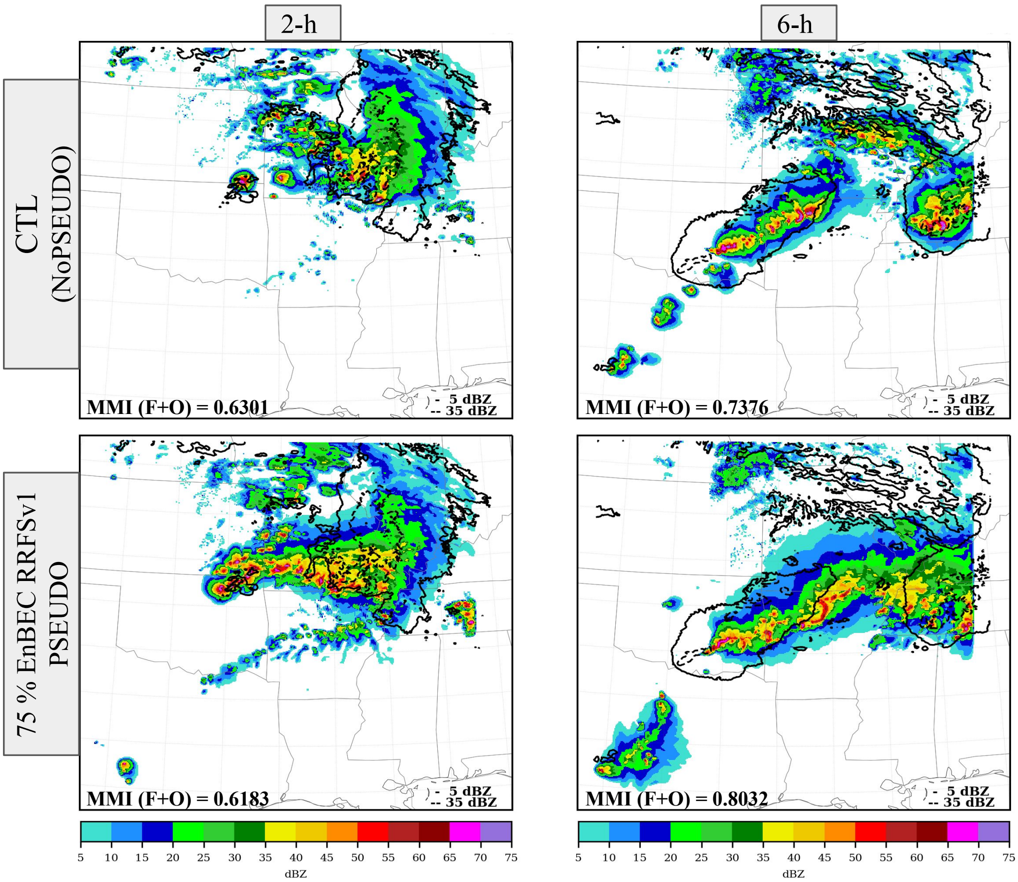


Figure 4.14 - As in Fig. 4.13, but for the experiment using the RRFSv1 physics suite with 75% of the EnBE, G1, and CTLOBS data assimilation with (lower panels) and without (upper panels) the PBL pseudo-observation added.

4.2.5 Observation System Experiments

Observation System Experiments (OSEs) were conducted in order to assess the impact of each observation type in CTLOBS as well as GPSRO bending angles from GNSS-RO. Table 4.2 list OSEs with the observations used in the that experiment. The surface dataset includes observations from SFCSHP, ADPSFC, and MESONET; radar radial wind velocity from radar level-II observations; satellite-derived wind from SATWND;

precipitable water (Pw) from GPSIPW; GNSS-RO bending angles from the LEO satellites listed in Section 3.3; and upper-air observations include all types in Table 3.1 but surface, radial velocity, satellite wind, precipitable water, or bending angles. In these experiments, the RRFSv1 suite with G1 grid and 75% of the EnBEC were used. CTL is the same experiment used in the previous sections.

Table 4.2 - List of Observation System Experiments conducted with the observation set included.

Experiments name	**Observations used**					
	Upper-air	Surface	Radial wind	Satellite wind	Pw	GNSS-RO
UP	T	**F**	**F**	**F**	**F**	**F**
UP+SRF	T	T	**F**	**F**	**F**	**F**
UP+SRF+SAT	T	T	T	**F**	**F**	**F**
UP+SRF+SAT+RW	T	T	T	T	**F**	**F**
CTL	T	T	T	T	T	**F**
CTL+BND	T	T	T	T	T	T

The 2- and 6-h composite reflectivity forecasts from the experiments UP and UP+SRF are presented in Fig. 4.15. The positive impact of assimilating surface data can be clearly seen in experiment UP+SRF with a better simulation of the convective initiation over northeast Oklahoma in the 2-h forecasts and the evolution of the squall line for 6-h forecast. Without using surface data, the initiation of convective was missed in the 2-h forecast over northeast Oklahoma as well as over central Texas in the 6-h forecast in UP. The convection over southwest Missouri was greatly improved in experiment UP=SURF also. However, experiment UP+SFR increased he extent and intensity of spurious convection in the experiment UP. It overestimated the cell over southwest Missouri in the 2-h forecast and over-predicted the convection over Texas in 6-h forecast. The MMI (F+O) values for these experiments show better performance in UF+SRF with a value of 0.8326 for 6-h forecast.

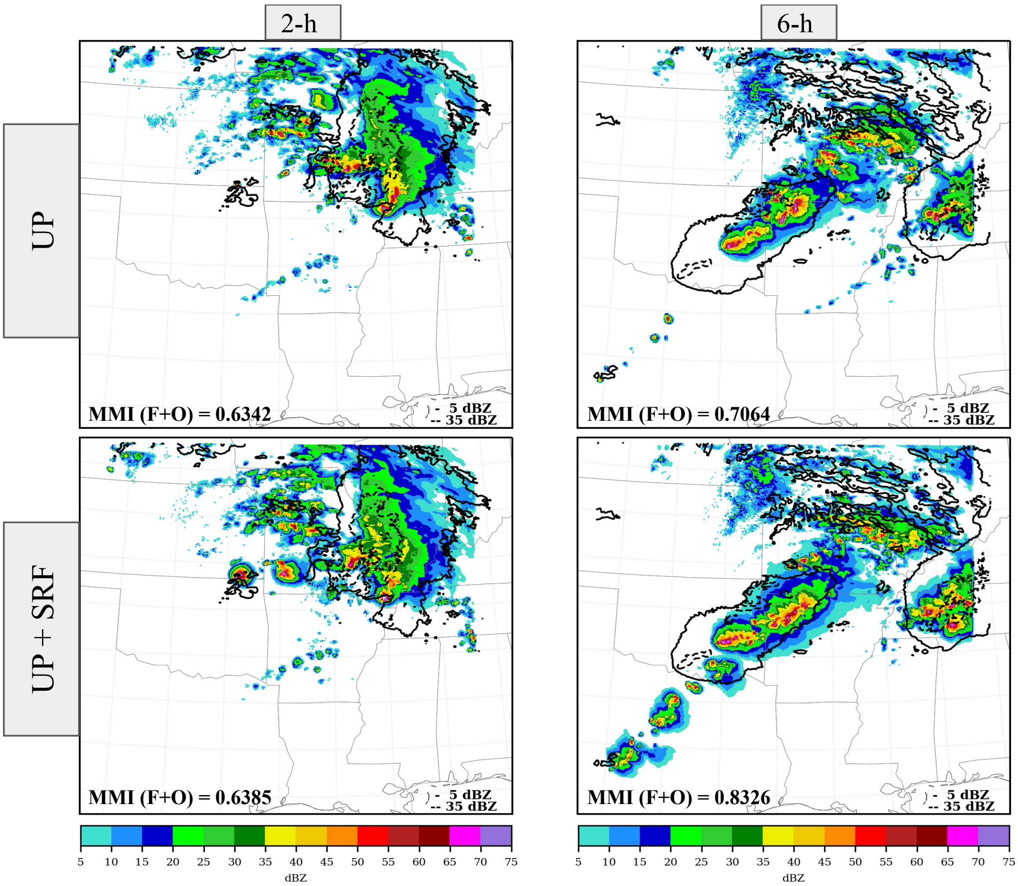


Figure 4.15 - As in Fig. 4.14, but for experiment UP (upper panels) and experiment UP+SRF (lower panels).

The mean RMS and bias of the background and analysis against GNSS-RO bending angles observations in the experiment adding GNSS-RO bending angles are presented in Fig. 4.16. Note that some cycles have no bending angles observation in the domain. However, when there are GNSS-RO data available, the analyses matched better to the observations with lower bias and a 20% reduction of the RMS values in most analyses, approximately.

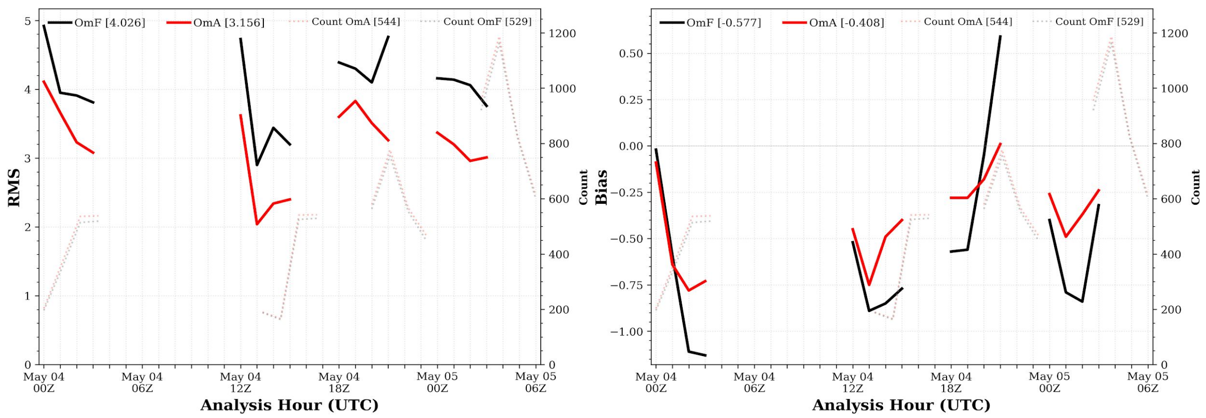


Figure 4.16 - Mean RMS (left) and bias (right) of the background (OmF) and analysis (OmA) against GNSS-RO bending angles observations, in experiment CLTOBS+BND during the cycling period from 00Z May 4th to 06Z May 5th, 2020.

An example of the spatial distribution of the assimilated bending angles profiles is shown in Fig. 4.17 for the analyses at 21Z on May 4th and at 01Z on May 5th. The colors indicate the satellite identification number (id) and time of the observations. For the analysis at 01Z May 5th, a total of 1187 observations were assimilated from COSMIC2 (id: 752), MetOp-B, and -C satellites (id: 4 and 5, respectively). For the analysis at 21Z May 4th, less observations were available. A total of 466 observations were assimilated from four of the six COSMIC2 satellites (id: 751, 752, 753, and 755), including one profile (from id: 752) near the convection initiation over northeast Oklahoma and another one between western Tennessee and northern Alabama near the convection occurring over western Tennessee.

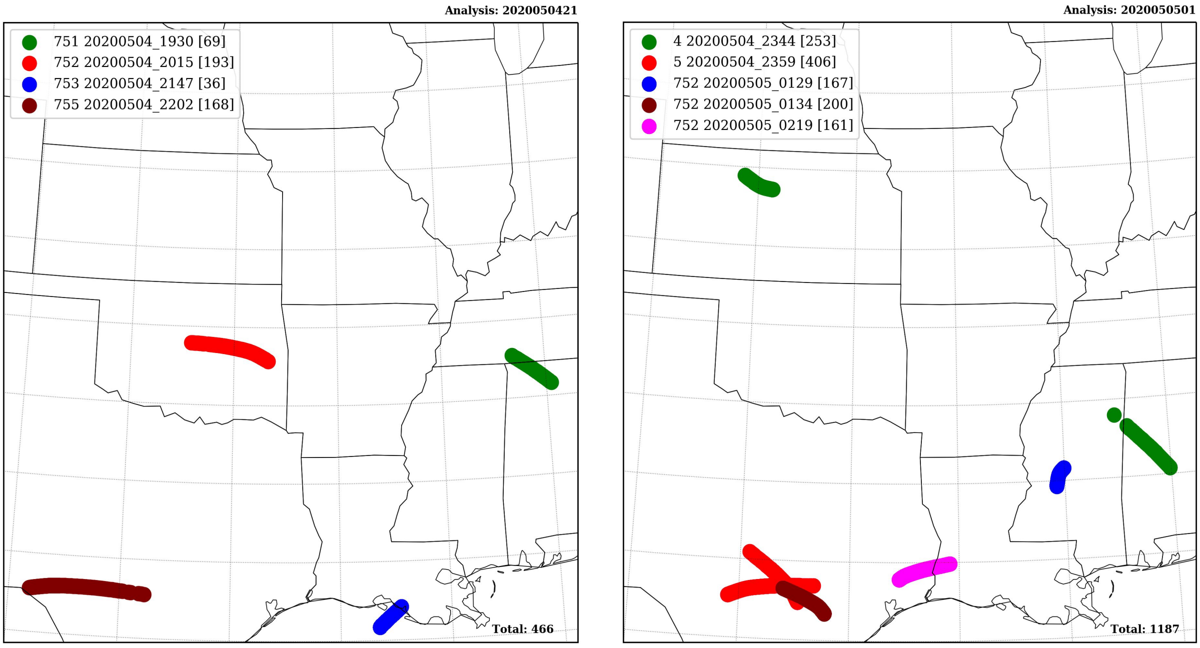


Figure 4.17 - Spatial distribution of assimilated bending angle observations in the analyses at 21Z on May 4th, 2020 (left) and at 01Z on May 5th, 2020 (right). Colored circles indicate the satellite that provided the data.

Fig. 4.18 shows the 2- and 6-h composite reflectivity forecasts initialized at 19Z May 4th from the experiments CTL+BND (lower panels) and CTL (upper panels). Assimilating the bending angles not only decreased the spurious convection in both forecast lengths, but also produced slightly better simulation of the convection over eastern Missouri in 2-h forecast and over western Tennessee at 6-h forecast. Nevertheless, the coverage and intensity of convection over northeast Arkansas was overestimated in the 2-h forecast. Results from MMI (F+O) indicate better results in CTL+BND over CTL, with an increase from 0.6301 to 0.6515 in the 2-h forecast and from 0.7306 to 0.7525 in the 6-h forecast.

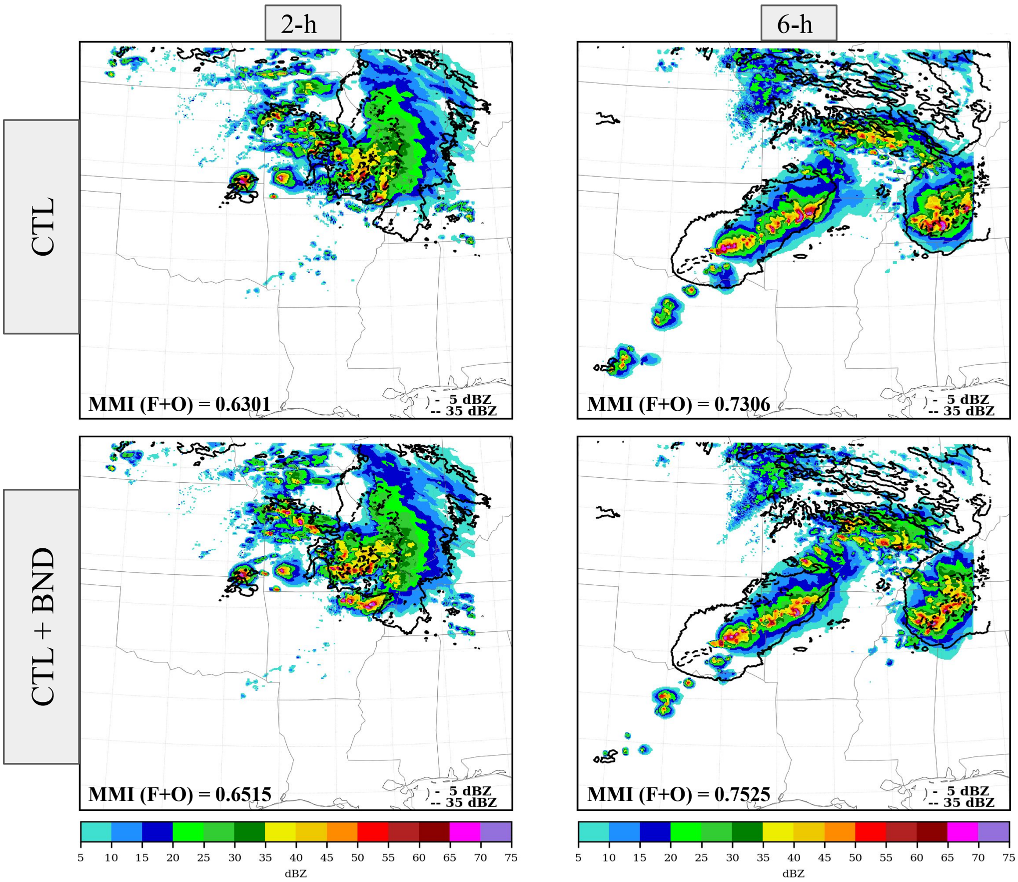


Figure 4.18 - As in Fig. 4.15, but for CTL (upper panels) and CTL+BND (lower panels) experiments.

4.3 Cycling with GFS initial conditions and spin-up cycles

When using HRRR analysis as ICs for the cold start cycle as all previous experiments, spin-up period is not needed because the HRRR is a operational mesoscale model with the same horizontal resolution and cycling intervals as in the experiments performed in this study. However, as described in Section 3.5, different cycling configurations are needed in order to use GFS analysis/forecast as cold start ICs. Therefore, an experiment was conducted using the GFSics+CONTsurf cycling configuration with the RRFSv1 CCPP suite and hybrid data assimilation using 75% of the EnBEC, G1, and CTLOBS. The GFSics+CONTsurf cycling configuration cycles surface fields continually in the full cycles. Fig. 4.19 shows the 2- and 6-h composite reflectivity forecasts from this new experiment comparing with the same forecast from the CTL experiment. The 2-h forecast was clearly improved when using GFS ICs with an GFSics+CONTsurf cycling configuration. The convection over eastern Missouri, western Illinois, western Tennessee, and western Kentucky was better represented and the extent of the convection over northeast Oklahoma was improved to better match the observations. The spurious convection developed over southwest Missouri and southern Arkansas in the experiment CTL were greatly diminished in new experiment (left panels in Fig. 4.19). Using GFS ICs with GFSics+CONTsurf cycling also gave better simulation of the squall line structure and decreased intensity of the convection in the 6-h forecast. However, the convection associated with the dry line in northeast Texas and convection over Missouri and Illinois were overestimated at 6-h forecast when using GFS ICs. Yet, MMI (F+O) results suggest that CTL was slightly better in the 2-h forecast, but GFSics+CONTsurf surpassed CTL forecasts quality in 6-h forecast length.

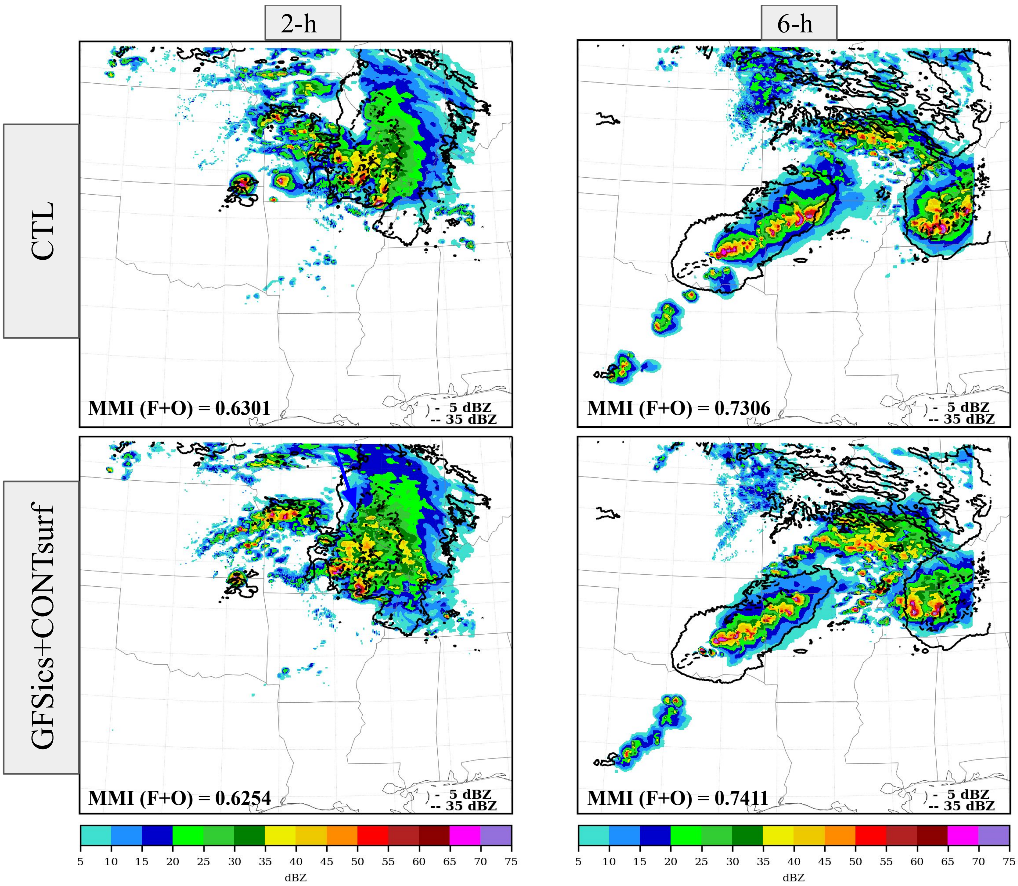


Figure 4.19 - As in Fig. 4.18, but for CTL (upper panels) and GFSics+CONTsurf (lower panels) experiments.

The GFSics+surf cycling configuration uses the GFS surface field at the cold start spin-up cycles and those surface fields will be used in the full cycles also. Results using the GFSics+surf cycling configuration with RRFSv1 physics suite and hybrid analysis with 75% of the EnBEC, G1, and the CTLOBS are shown in the lower panels of Fig. 4.20. In order to compare results with and without the surface fully cycled, the upper panels show results from the GFSics+CONTsurf experiment. Because of the short length of the full cycles in experiment execution period, it is expected to obtain a neutral impact from continue cycling surface fields. However, results in Fig. 4.20 indicate a slightly better representation of the squall line in the 6-h forecast with a reduction of the cells developed, and also less strong convective cells over western Tennessee in both forecast hours. The overestimation of the convection over Texas and areas of Missouri and northern Arkansas was still produced. MMI (F+O) values confirm that this new experiment had better quality with an increase from 0.6254 to 0.6501 in the 2-h forecast and from 0.7411 to 0.7771 in the 6-h forecast.

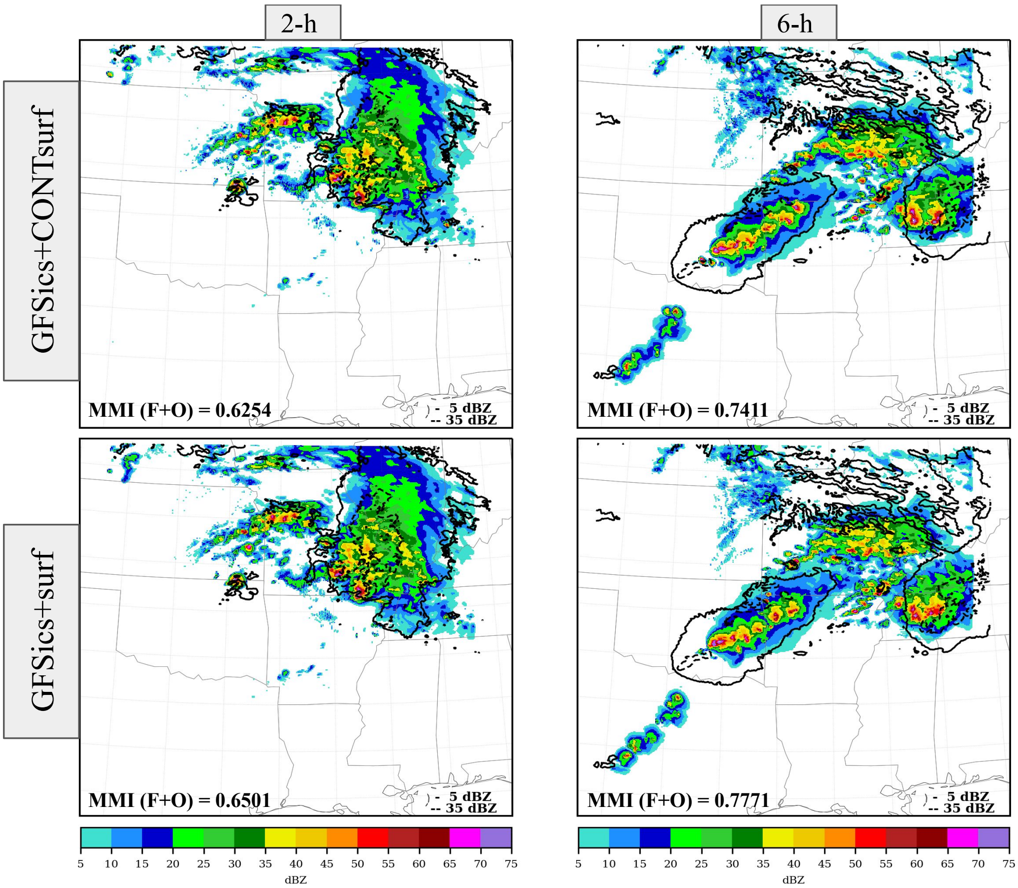


Figure 4.20 - 2- and 6-h composite reflectivity forecasts from GFSics+surf and CTL experiments, initialized at 19Z May 4th 2020. Solid and dashed black lines are the 5 and 35 dBZ MRMS observations contours, respectively.

4.4 Regional levels

It is well known that UFS FV3-based models use a lagrangian vertical coordinate, which is generated using hybrid coefficients. The coefficients used to generate the 64 vertical levels in the SRW application are available in the fixed files used to configure and run the system. In this study, the default 64 vertical levels were used to run all previous experiments, as mentioned in Section 3.2. Moreover, other hybrid coefficients were took into account and tested. These new coefficients were provided by Dr. Chunhua Zhou from GSL. Fig. 4.21 shows the two sets of the ak and bk hybrid coefficients used. The solid orange line corresponds to the default ak and bk values, which generate 64 vertical levels with model top and bottom heights similar to those observed in global models. That is, the first level is located at around 20 m and the last level at around 56 km. New coefficients are represented by the solid black line and also generate 64 vertical levels. However, the model top and bottom heights are similar to those observed in regional models, with the bottom at around 8 m and top around 32 km.

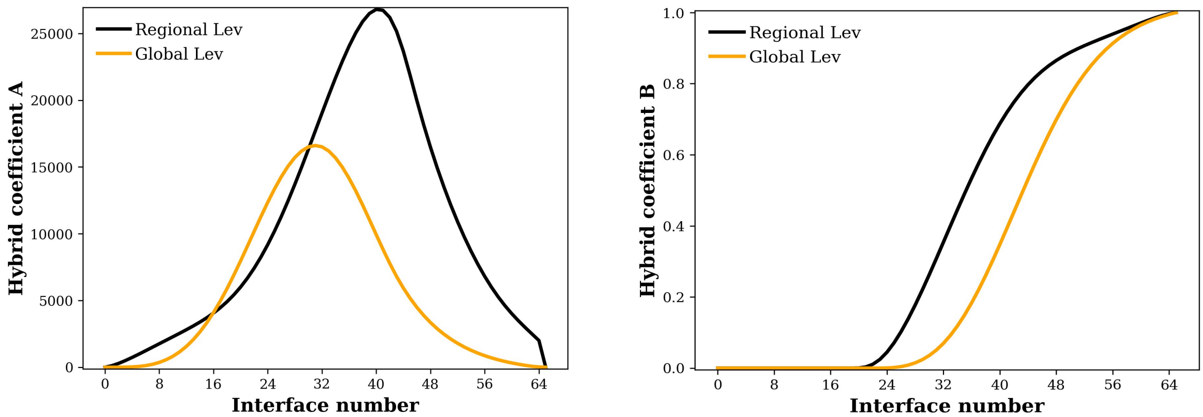


Figure 4.21 - Hybrid ak (left hand figure) and bk (right hand figure) coefficients used to generate the vertical levels in FV3LAM.

An experiment called NLEV was conducted using the new coefficients along with the RRFSv1 physics suite and hybrid analysis with 75% of the EnBEC, G1, and CTLOBS. Neutral impact was found in the convection forecasts, but the RMSE and bias verification for 2-h forecast of temperature, specific humidity, and wind vertical profiles showed improvements in the lower levels during the afternoon (Fig. 4.22). During the night, the impact was neutral as can be observed in the upper panels of Fig. 4.22.

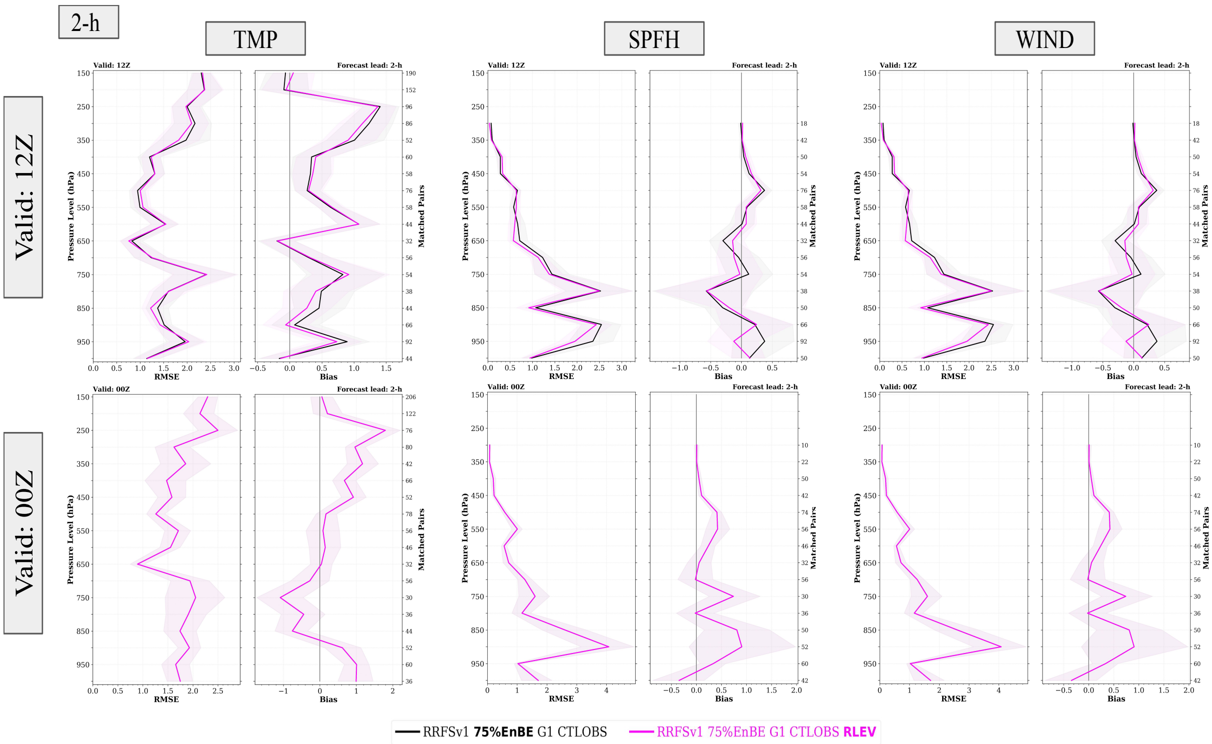


Figure 4.22 - RMSE and bias verification of temperature (TMP), specific humidity (SPHM), and wind (WIND) vertical profiles against RAOB, PIBAL, and DROPS reports in the experiment NLEV for 2-h forecast lead time at valid hours 00Z (lower panels) and 12Z (upper panels). Statistics were computed for a range of 50hPa in the vertical levels, thus, the RMSE and bias with its correspondent bootstrap confidence interval for 950 hPa were calculated for the vertical levels between 950 to 901 hPa (including both).

5 Summary and future work

In this visitor project, the current RRFS capability to represent convection was investigated for a case study of a squall line the occurred over Oklahoma during the afternoon of 4 May, 2020. Various parameters and options were tested and evaluated seeking to find the best configuration that produced more realistic convection forecasts. Bellow are listed the main findings in this study.

a) FV3LAM with HRRR ICs showed the ability of FV3LAM to represent the convection initiation and evolution of the squall line with better results using the RRFSv1 physics suite. However, as the squall line evolved, the strongest cells were developed ahead of what was observed;

b) Different CCPP suites gave a different representation of the convection, with RRFSv1 representing more strong and individual cells and GFSv15 depicting weaker and smoother cells with larger coverage. The GSDsar physics suite missed the squall line initiation and evolution when no observations were assimilated. The GFSreg captured well the squall line structure, but overestimated the convection over Texas, Missouri and northern Arkansas;

c) With data assimilation, the analysis matched better the observations improving the convection forecasts. Pure ensemble data assimilation produced better results in the short-term forecasts, but 75% of the ensemble BEC produced good forecasts in the short-term and larger forecast lengths. More tests may be conducted using 100% of the EnBEC;

d) The assimilation of surface observations was crucial for the convection forecasts with a better representation of the convection initiation. Overall, a more complete observations dataset produced more promising results;

e) The convection was greatly overestimated when using pseudo-innovations from 2-m temperature and 2-m dew point (specific humidity?) observations through 75% of the PBL height. This indicates the need for more tuning in this function;

f) GFS ICs with an appropriate cycling configuration produced skillful short-term forecast by mitigating overprediction and better representing the convection initiation, squall line structure, and convection patterns in other parts of the domain. This configuration seems promising for RRFS;

g) Vertical levels with the first level closer to the surface are able to better represent the temperature, specific humidity, and wind in the lower levels of the troposphere;

h) Most of the configurations tested were able to capture the main convective systems during the execution period, however, the convection associated to the squall line was overestimated in intensity and underestimated in its extent. MRMS observations showed a more stratiform region ahead of the convective cells which was not well captured in the forecasts. This may indicate that despite the several options tested here, more testing and evaluation of these and other options are needed.

In the future, RRFS is intended to cover a similar domain as RAP, extending into part of northern South America. Besides, this region is also covered by the Hurricane Analysis and Forecast System (HAFS) (DONG et al., 2020), which the UFS application for hurricanes. Therefore, the next part of this project will focus on the investigation of RRFS's ability to represent convection on the tropical region, specifically the convection associated to Amazon Coastal Squall Lines (GARSTANG et al., 1994).

REFERENCES

ALEXANDER, C.; CARLEY, J. Short-Range Weather in operations. **Bulletin of the UFS Community**, p. 9, 2020. Available at: <https://www.ufscommunity.org/wp-content/uploads/2021/03/Bulletin_UFS_Community_Winter_2020_Full.pdf>. 1

BENJAMIN, S. G.; WEYGANDT, S. S.; BROWN, J. M.; HU, M.; ALEXANDER, C. R.; SMIRNOVA, T. G.; OLSON, J. B.; JAMES, E. P.; DOWELL, D. C.; GRELL, G. A.; LIN, H.; PECKHAM, S. E.; SMITH, T. L.; MONINGER, W. R.; KENYON, J. S.; MANIKIN, G. S. A north american hourly assimilation and model forecast cycle: The rapid refresh. **Monthly Weather Review**, American Meteorological Society, Boston MA, USA, v. 144, n. 4, p. 1669 – 1694, 2016. Available at: <https://journals.ametsoc.org/view/journals/mwre/144/4/mwr-d-15-0242.1.xml>. 21, 31

Bernardet, L.; Firl, G.; Heinzeller, D.; Carson, L.; Sun, X.; Pan, L.; Zhang, M. Engaging the community in the development of physics for nwp models. In: **EGU General Assembly Conference Abstracts**. 2020. p. 22093. Available at: <https://ui.adsabs.harvard.edu/abs/2020EGUGA..2222093B>. 17

CIMSS. **CIMSS Cooperative Agreement Annual Report**. 2014. 302 p. Available at: <https://cimss.ssec.wisc.edu/reports/CIMSS-CA-Report_2014_Final.pdf>. 29

DAVIS, C. A.; BROWN, B. G.; BULLOCK, R.; HALLEY-GOTWAY, J. The method for object-based diagnostic evaluation (mode) applied to numerical forecasts from the 2005 nssl/spc spring program. **Weather and Forecasting**, American Meteorological Society, Boston MA, USA, v. 24, n. 5, p. 1252 – 1267, 2009. Available at: <https://journals.ametsoc.org/view/journals/wefo/24/5/2009waf2222241_1.xml>. 18

DELUCA, C.; JASKOFOR, S. Graduate Student Tests: Evaluat ing t he Usabilit y of UFS. In: **Unified Forecast System (UFS) Users' Workshop. July 27–29**. 2020. Available at: <https://dtcenter.org/sites/default/files/events/2020/1-deluca-cecelia.pdf>. 3

DONG, J.; LIU, B.; ZHANG, Z.; WANG, W.; MEHRA, A.; HAZELTON, A.; WINTERBOTTOM, H.; ZHU, L.; WU, K.; ZHANG, C.; TALLAPRAGADA, V.; ZHANG, X.; GOPALAKRISHNAN, S.; MARKS, F. The Evaluation of Real-Time Hurricane Analysis and Forecast System (HAFS) Stand-Alone Regional (SAR) Model Performance for the 2019 Atlantic Hurricane Season. **Atmosphere**, v. 11, 2020. Available at: <https://doi.org/10.3390/atmos11060617>. 43

DU, J. CIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data. Version 1.0. **UCAR/NCAR - Earth Observing Laboratory.**, 2011. Available at: <https://doi.org/10.5065/D6PG1QDD>. 15

GARSTANG, M.; MASSIE, H. L.; HALVERSON, J.; GRECO, S.; SCALA, J. Amazon coastal squall lines. Part I: Structure and kinematics. **Monthly Weather Review**, v. 122, p. 608–622, 1994. Available at: <https://doi.org/10.1175/1520-0493(1994)122<0608:ACSLPI>2.0.CO;2>. 43

GUSTAFSSON, N.; JANJI, T.; SCHRAFF, C.; LEUENBERGER, D.; WEISSMAN, M.; REICH, H.; BROUSSEAU, P.; MONTMERLE, T.; WATTRELOT, E.; BUČÁNEK, A.; MILE, M.; HAMDI, R.; LINDSKOG, M.; BARKMEIJER, J.; DAHLBOM, M.; MACPHERSON, B.; BALLARD, S.; INVERARITY, G.; CARLEY, J.; ALEXANDER, C.; DOWELL, D.; LIU, S.; IKUTA, Y.; FUJITA, T. Survey of data assimilation methods for convective-scale numerical weather prediction at operational centres. **Quarterly Journal of the Royal Meteorological Society**, v. 144, p. 1218–1256, 2018. Available at: <https://doi.org/10.1002/qj.3179>. 26

HU, M.; BENJAMIN, S. G.; LADWIG, T. T.; DOWELL, D. C.; WEYGANDT, S. S.; ALEXANDER, C. R.; WHITAKER, J. S. Gsi three-dimensional ensemble–variational hybrid data assimilation using a global ensemble for the regional rapid refresh model. **Monthly Weather Review**, American Meteorological Society, Boston MA, USA, v. 145, n. 10, p. 4205 – 4225, 2017. Available at: <https://journals.ametsoc.org/view/journals/mwre/145/10/mwr-d-16-0418.1.xml>. 6, 8, 11

HU, M.; GE, G.; ZHOU, C.; STARK, D.; SHAO, H.; NEWMAN, K.; BECK, J.; ZHANG, X. **Grid-point Statistical Interpolation (GSI) User's Guide Version 3.7.** 2018. Available at: <http://www.dtcenter.org/com-GSI/users/docs/index.php>. 14

JENSEN, T.; BROWN, B.; BULLOCK, R.; FOWLER, T.; GOTWAY, J. H.; NEWMAN, K. **The Model Evaluation Tools v9.0 (METv9.0) User's Guide.** 2020. Available at: <https://dtcenter.org/sites/default/les/community-code/met/docs/user-guide/MET_Users_Guide_v9.0.pdf>. 13

KLEIST, D. T.; PARRISH, D. F.; DERBER, J. C.; TREADON, R.; WU, W.-S.; LORD, S. Introduction of the gsi into the ncep global data assimilation system. **Weather and Forecasting**, American Meteorological Society, Boston MA, USA, v. 24, n. 6, p. 1691 – 1705, 2009. Available at: <https://journals.ametsoc.org/view/journals/wefo/24/6/2009waf2222201_1.xml>. 7, 21

LORENC, A. C. Advances in data assimilation techniques and their relevance to satellite data assimilation DA methods for NWP. In: **ECMWF Seminar on Use of Satellite Observations in NWP, 8–12 September**. 2014. 26

PURSER, R. J.; JOVIC, D.; KETEFIAN, G.; BLACK, T.; BECK, J.; DONG, J.; CARLEY, J. The Extended Schmidt Gnomonic Grid for Regional Applications. In: **Unified Forecast System (UFS) Users' Workshop. July 27-29**. 2020. Available at: <https://dtcenter.org/sites/default/files/events/2020/2-purser-james.pdf>. 5

SIP. **Strategic Implementation Plan for evolution of NGGPS to a national Unified Modeling System (First Annual Update)**. 2018. 171 p. Available at: <https://www.weather.gov/media/sti/nggps/UFS%20SIP%20FY19-21_20181129.pdf>. 1

UFS. **About UFS Applications**. 2019. Available at: <http://ufs-dev.rap.ucar.edu/index.html#/science/aboutapps>. 1

UFS Development Team. **Unified Forecast System (UFS) Short-Range Weather (SRW) Application**. Zenodo, mar. 2021. Available at: <https://doi.org/10.5281/zenodo.4534994>. 5, 9

WHITAKER, J.; HAMILL, T.; WEI, X.; SONG, Y.; TOTH, Z. Ensemble data assimilation with the NCEP Global Forecast System. **Monthly Weather Review**, v. 136, p. 463–482, 2008. Available at: <https://doi.org/10.1175/2007MWR2018.1>. 8

WHITE, G.; YANG, F.; TALLAPRAGDA, V. The Development and Success of NCEP's Global Forecast System. **National Oceanic and Atmospheric Administration: Silver Spring, MD, USA**, 2018. Available at: <https://ams.confex.com/ams/2019Annual/webprogram/Manuscript/Paper350196/technotegfsh.pdf>. 1

WMO. **Guide to WMO Table Driven Code Forms: FM 94 BUFR and FM 95 CREX**. 2002. Available at: <https://www.wmo.int/pages/prog/www/WDM/Guides/BUFRCREXGuide-English.html>. 13

WOLFF, J.; BECK, J. The UFS Short-Range Weather App. **Bulletin of the UFS Community**, p. 9, 2020. Available at: <https://www.ufscommunity.org/wp-content/uploads/2021/03/Bulletin_UFS_Community_Winter_2020_Full.pdf>. 1

WU, W.; PURSER, R. J.; PARRISH, D. Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. **Monthly Weather Review**, v. 130, p.

2905–2916, 2002. Available at: <https://doi.org/10.1175/1520-0493(2002)130<2905:TDVAWS>2.0.CO;2>. 21

ZHANG, J.; HOWARD, K.; LANGSTON, C.; KANEY, B.; QI, Y.; TANG, L.; GRAMS, H.; WANG, Y.; COCKS, S.; MARTINAITIS, S.; ARTHUR, A.; COOPER, K.; BROGDEN, J.; KITZMILLER, D. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities. **Bulletin of the American Meteorological Society**, v. 97(4), p. 621–638, 2016. Available at: <https://doi.org/10.1175/BAMS-D-14-00174.1>. 16