2020 DTC Visitor Program Final Report:

Test and evaluation of CCPP physics suites and data assimilation capabilities to improve the Rapid Refresh Forecast System for convection forecasts

DTC Host: DTC Data Assimilation team

Ivette Hernández Baños
Graduate Program in Meteorology, National Institute for Space Research, São José dos Campos, São Paulo, Brazil (ivette.banos@inpe.br)
Ph.D. Advisor: Luiz Fernando Sapucci

1 Background and introduction

In order to simplify the operational suite of the National Centers for Environmental Prediction (NCEP) into a single system capable of representing different spatial and temporal scales as well as the various Earth system components, the National Oceanic and Atmospheric Administration (NOAA) is transitioning toward the Unified Forecast System (UFS), based on the non-hydrostatic finite volume cubed-sphere (FV3) dynamical core (EMC, 2018). The UFS is a community-based system that enables contributions from the research community to become available for 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 United States 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 extensively tested. The first version of the SRW application was recently released including the FV3 Limited Area Model (FV3-LAM) 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; Wu, Purser and Parrish (2002)) has been added as the analysis component of the SRW application to improve initial conditions for the FV3-LAM 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 is conducted in order to provide developers an insight to 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 are conducted testing various aspects of the RRFS along with numerous subjective and objective forecast verifications to evaluate each experiment.

1.1 Objectives

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

1. Assess different CCPP physics suites;
2. Assess data assimilation algorithms and configurations;
3. Examine different cold start initial condition and cycling configurations;
4. 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 radio occultation (GNSS-RO) observations;
5. Evaluate RRFS forecasts for the initiation and evolution of convection.

2 Methodology

2.1 Case study and experiments configuration

A line of storms developed over northeastern Oklahoma ahead of a southward moving cold front during the afternoon of 4 May 2020. Between 19Z and 20Z on 4 May, high values of Mixed Layer Convective Available Potential Energy (MLCAPE) and effective bulk shear were observed over northeastern Oklahoma. At 20Z, the first convective cells were observed on the Multi-Sensor Multi-Radar (MRMS) composite reflectivity observations (see the black circle in Fig. 1B.), and at around 22Z (Fig. 1D.) a line of storms extending across central Oklahoma was observed along the pre-frontal wind shift. The system evolved while slowly moving southeastward causing several instances of large hail and high wind, mostly over northeastern and south-central Oklahoma, southeastern Kansas, southwestern Missouri, and northwestern Arkansas. Small convective cells initiated and developed over Texas between 23Z on 4 May (Fig. 1E.) to 01Z on 5 May (Fig. 1G.) along a dry line extending over western Texas.

For the simulation of this case, a domain is configured covering a 460×460 grid cells centered on Fort Smith, Arkansas (-95.35°W; 35.28°N) with a 3 km grid-spacing and 64 vertical layers. All simulations run hourly cycles with 18-hour forecasts starting at 00Z on 4 May 2020 through 06Z on 5 May 2020. Initial conditions (ICs) and lateral boundary conditions (LBCs) are from the 3-km High-Resolution Rapid Refresh (HRRR) model. Hourly Rapid Refresh (RAP) observations are utilized for the analysis. Upper-air, surface, radar radial velocity and satellite AMV wind, as well as precipitable water are called the ConTroL OBServations dataset (CTLOBS) included in the experiments with data assimilation. BUFR (Binary Universal Form for data Representation) observation files from the Global Data Assimilation System (GDAS) are used to assimilate bending angles derived from GNSS-RO observations.

For the ensemble component of the hybrid analysis, the 9-h forecast from the GDAS Ensemble Kalman Filter (EnKF; Whitaker et al. (2008)) 80-member ensemble are used. However, since they are available only 4 times per day, the same 9-h forecast GDAS ensembles are reused for

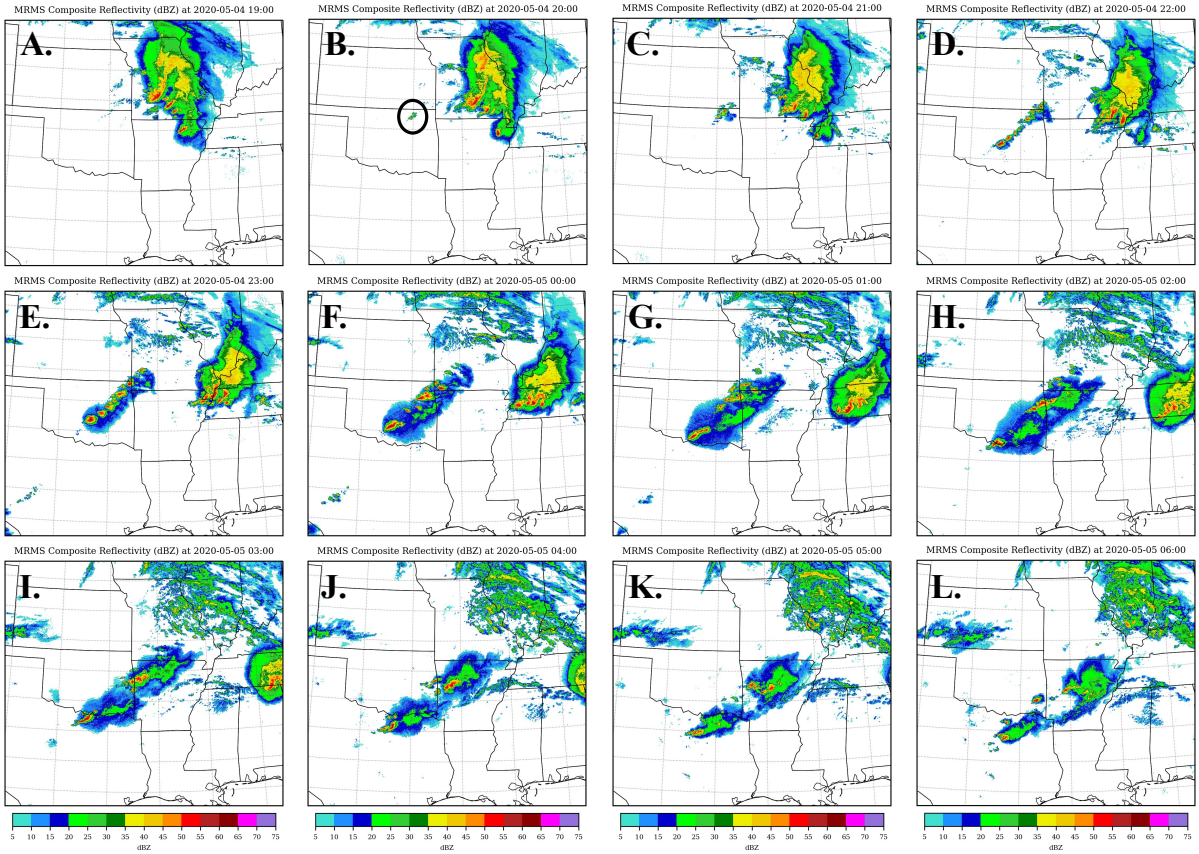

Figure 1: Hourly Multi-Radar Multi-Sensor (MRMS) composite reflectivity from 19Z on 4 May 2020 through 06Z on 5 May 2020 (panels A through L).

the 2 hours before and 3 hours after its valid hour.

2.2 Workflow

The RRFS workflow used is based on the UFS SRW application v1.0.0 (UFS Development Team, 2021) community workflow. Figure 2 shows a diagram of the RRFS workflow used in this study. It includes the tasks to configure the domain by creating the corresponding grid, orography, and surface climatology files, also known as fixed files (brown box); obtain and create the ICs and LBCs from an external model (yellow and green boxes); execute the FV3-LAM (purple oval) and finally, post-process the model outputs (turquoise box). The community workflow was modified to include the analysis tasks and different cycling strategies. The gray box in Fig. 2 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 would be generated from an external model and the analysis would use an external model forecast as background combined with observations valid in the time window around the initialization. Meanwhile, the warm start indicates that the FV3-LAM 1-h forecast from the previous cycle (dashed purple rectangle) would be used as background combined with observations 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. 2 by the dashed lines. Arrows in the diagram

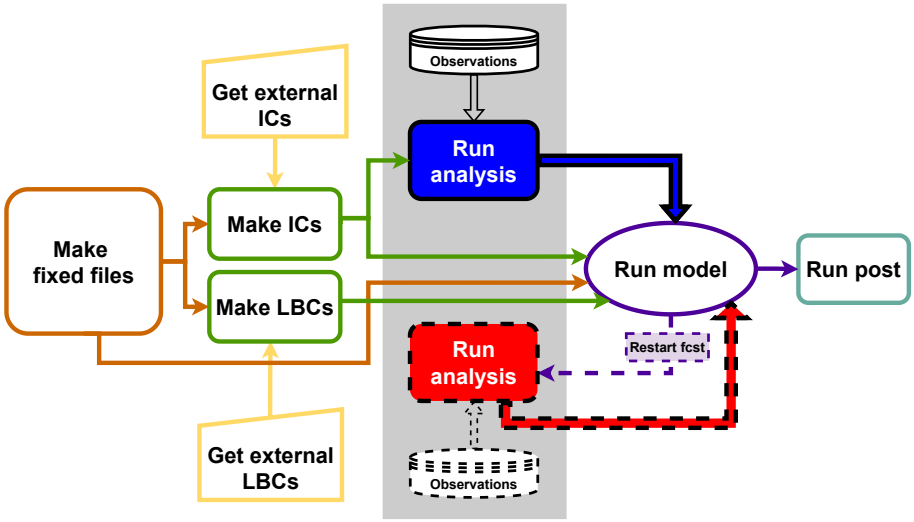

Figure 2: Diagram including the tasks in the current RRFS workflow as well as the tasks added to create the analysis using GSI.

indicate the general pipeline of the RRFS system, where the results of one task are the inputs to the next.

2.3 Cycling configurations

In this project, five cycling strategies are tested. First, the cycling is configured using cold starts at each initialization time. Hourly HRRR analysis results are used to initialize the FV3-LAM, which run 18-h forecasts with hourly outputs. For this cycling configuration, the released SRW Application workflow is used without running any analysis tasks. Experiments using this cycling are referred to as HRRRDA. Second, a baseline cycling configuration is setup building off of HRRRDA. Cold start is used every 12 hours and warm start is instead used between 01Z and 11Z. The model is configured to save the first hour forecast (restart files), which is used to initialize the model in the next cycle. No data assimilation is performed in any cycle. Experiments with this configuration are called NoDA.

Next, a cycling strategy with data assimilation is configured. A GSI analysis is added at each cycle to provide initial conditions for the model forecasts as indicated in Fig. 3A., but the cold and warm start strategy is maintained as in the previous configuration. As indicated in Fig. 2, the analysis task uses the HRRR analysis as background to cold start the model in the cycles at 00Z and 12Z on 4 May 2020 and 00Z on 5 May 2020. In the hours in between, the 1-h forecast from previous cycle is used as background in the analysis task to warm start the model. Observations in CTLOBS are assimilated every hour as indicated in Fig. 3A. In addition, a strategy similar to the RAP partial cycle structure is configured following Hu et al. (2017) (see Fig. 3B.). For continuous cycles (second row in Fig. 3B.), GFS ICs are used to cold start the FV3-LAM in the first cycle and warm starts are used thereafter. Thus, after the first cycle, restart files are used every hour as background to create the analyses combined with CTLOBS. However, twice a day (at 06Z and 18Z), the restart files (atmosphere and surface) are the FV3-LAM 1-h forecast from the 6-h partial cycles executed in parallel to the continuous cycling every 12 hours. These

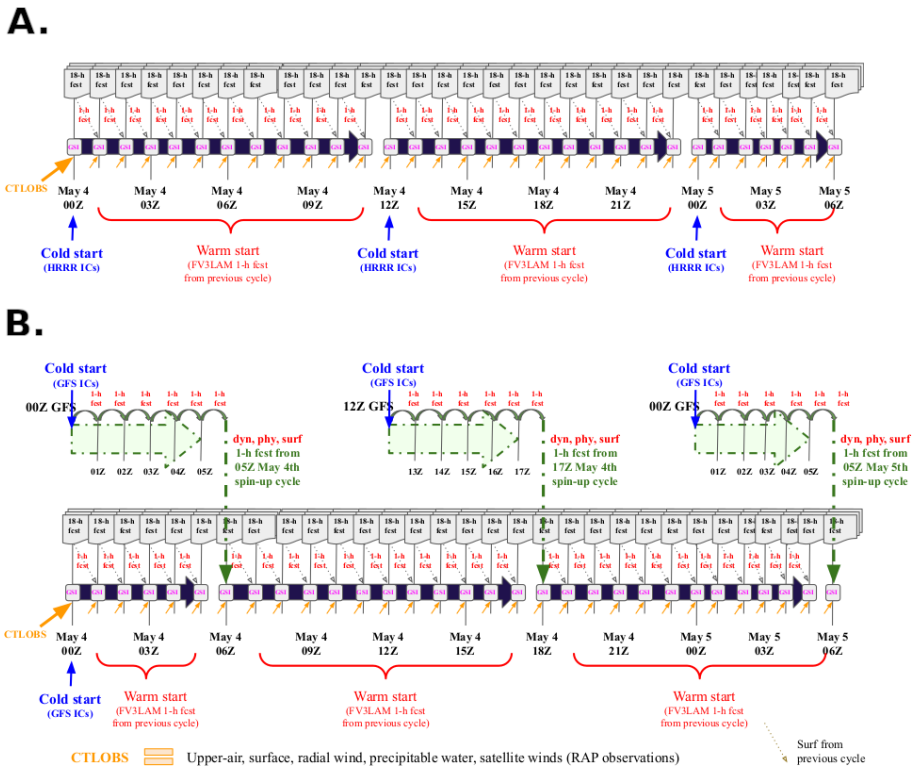

Figure 3: Cycling configuration with hourly data assimilation (A.) and with data assimilation using GFS atmospheric and surface conditions following the RAP partial cycle structure (B.)

parallel partial cycles are cold started using GFS ICs (at 00Z and 12Z) while HRRR forecasts are used as LBCs. The experiment using this cycling configuration is called GFSics+surf. Lastly, a cycling configuration is set up identically to the GFSics+surf (Fig. 3B.) mentioned, but keeping the FV3-LAM surface fully cycled in the continuous cycles. In this case, FV3-LAM surface conditions from previous cycle 1-h forecast are used in the 06Z and 18Z continuous cycles. The experiment using this cycling configuration is called GFSics+CONTsurf.

2.4 Common Community Physics Package (CCPP) suites

CCPP suites (Bernardet et al., 2020) and GSI configurations are tested and evaluated along with cycling configurations and observation impact experiments. The latest CCPP released, CCPP v5.0[1], 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 are tested along with two suites from CCPP v4.0, the GSDsar (FV3_GSD_SAR) developed at GSL based on RAP and HRRR physical schemes and the GFSreg suite (FV3_GFS_2017_gfdlmp_regional)—also based on GFS v15 physics. Table 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 is used.

[1] ⟨https://dtcenter.ucar.edu/GMTB/v5.0.0/sci_doc/index.html⟩

Table 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	uGWP	uGWP	uGWP	uGWP
Ocean	NSST	NSST	NSST	NSST
Ozone	NRL 2015	NRL 2015	NRL 2015	NRL 2015
Water Vapor	NRL 2015	NRL 2015	NRL 2015	NRL 2015

2.5 Forecast verification

After executing the experiments, the Model Evaluation Tools (MET) version 9.0 (Jensen et al., 2020) is used for forecast verification against observations. Specifically, the PB2NC, Point-Stat, Grid-Stat, Stat-Analysis, Method for Object-Based Diagnostic Evaluation (MODE), and MODE-analysis tools are used in this study. Upper-air (ADPUPA) and surface (ADPSFC) RAP observations are used to verify the corresponding forecasts. Precipitation forecasts are verified against hourly Stage IV precipitation observations (Lin; Mitchell, 2005) and hourly MRMS composite reflectivity mosaics (optimal method) observations (Zhang et al., 2016) are used to verify the predicted composite reflectivity. In order to quantitatively identify the experiment configuration that yielded better forecasts, the median of maximum interest (MMI (F+O)) (Davis et al., 2009) is analyzed. This metric 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 be close to zero.

3 Summary of results

3.1 HRRRDA and NoDA experiments

Figure 4 shows 2- and 6-h forecast of composite reflectivity in HRRRDA and NoDA experiments using RRFSv1 and GFSv15 CCPP suites from the 19Z cycle on 4 May 2020, with overlapping MRMS observations valid at 21Z on 4 May 2020 and 01Z on 5 May 2020, respectively. The two CCPP suites represent the squall line differently. Experiments using RRFSv1 (Fig. 4A., E., C., and G.) show stronger cells and smaller coverage while using GFSv15 (Fig. 4B., F., D., and H.) gives weaker and smoother cells with larger convection coverage and more spurious cells. HRRRDA experiments (Fig. 4A., E., B., and F.) capture the convective initiation over northeastern Oklahoma and the convection between northeastern Arkansas and western Tennessee, however the extent and intensity of convective cells is overestimated. NoDA experiments (Fig. 4C., G., D., and H.) predict the main features of the squall line, but the convection initiation over northeastern Oklahoma is misplaced in the 2-h forecast with RRFSv1 (Fig. 4C.)

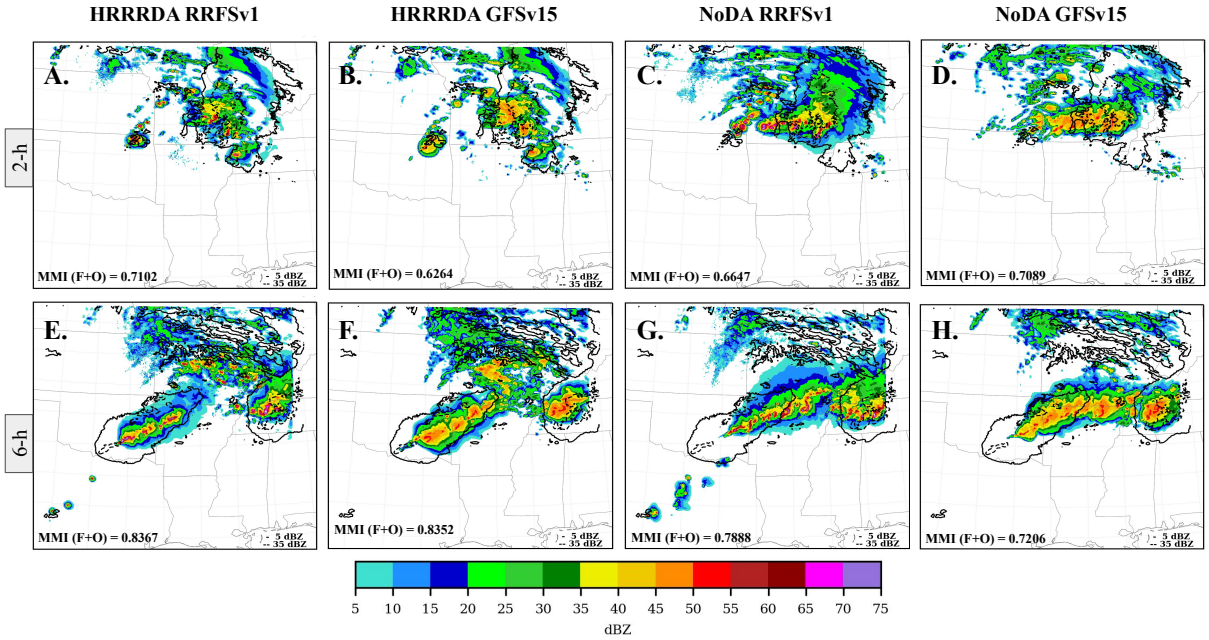

Figure 4: 2- (upper panels) and 6-h (lower panels) composite reflectivity forecasts from experiments using the HRRRDA cycling configuration with RRFSv1 (A. and E.) and GFSv15 (B., and F.) physics suites, and using NoDA cycling configuration with RRFSv1 (C., and G.) and GFSv15 (D., and H.) physics suites, initialized at 19Z on 4 May 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.

and the convective cells over Texas does not appear in the 6-h forecast with GFSv15 (Fig. 4H.). Overall, the extent of the squall line to the south central Oklahoma is not well captured in any of the experiments and the strongest convective cells in the squall line are located ahead of the observations. The MMI (F+O) indicates that the HRRRDA cycling configuration with the RRFSv1 suite produced better storms in both 2- and 6-h forecasts (Fig. 4A. and E.).

The NoDA experiment using GSDsar (not shown) shows a good representation of the convection over eastern Missouri and Illinois for the 2-h forecast as well as over western Tennessee and Kentucky for the 6-h forecast, but the squall line is missing in both forecast lead times. The GFSreg suite (not shown) predicts a better structure in the squall line, but the coverage is overestimated in Texas and northern Arkansas for both forecast lengths. Therefore, hereinafter only experiments using RRFSv1 are presented and discussed.

3.2 3DVar and hybrid 3DEnVar data assimilation

The three dimensional variational (3DVar) and hybrid 3D ensemble variational (3DEnVar) algorithms in GSI are tested using values 3 (G3) and 1 (G1) of the analysis to background grid ratio, which produce analysis grids of 9 and 3 km, respectively. The 3 km analysis grid ratio allows for more detail in the analysis increments, and therefore, only results using G1 are presented in this report. In GSI, a parameter (beta_s) is introduced via namelist to specify the relative weight given to the static background error covariance (BEC). This parameter can have values between

0 and 1, where 1 − beta_s indicates the weight given to the ensemble BEC (EnBEC). Different weights of the EnBEC are tested starting with 100% or pure ensemble, 90%, 75%, 50%, and 0% or 3DVar. In these experiments, the 9-h forecast GDAS ensembles are used with a 3 km ensemble grid ratio (ratio between the analysis and ensemble members grid), the same as the analysis grid ratio.

The specific humidity and temperature analysis increments using 100%, 75%, and 0% (3DVar) EnBEC weights with RRFSv1 CCPP suite are shown in Fig. 5 for the 21Z cycle on 4 May 2020. More flow dependent features can be clearly observed as the weight of EnBEC increases. Analysis increments using 3DVar (Fig. 5C. and F.) are smoother than ones using 100% EnBEC (Fig. 5A. and D.), which show detailed flow-dependent contours. Otherwise, analysis increments in the 100% and 75% (Fig. 5B. and E.) experiments seem to better adjust the temperature and humidity over northern Arkansas and Oklahoma.

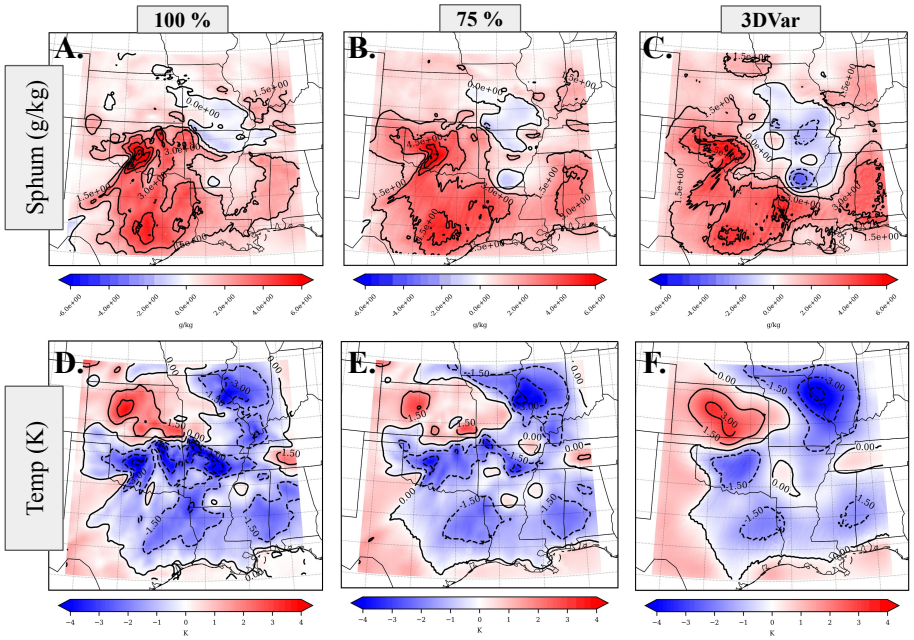

Figure 5: Analysis increment for specific humidity (g/kg) (A., B., and C.) and temperature (K) (D., E., and F.) at the first level above the surface (model level 63) for 21Z on 4 May 2020, using the RRFSv1 CCPP suite and 100% EnBEC (A. and D.), 75% EnBEC (B. and E), and 3DVar (C. and F.).

The 2- and 6-h composite reflectivity forecasts for the same experiments as in Fig. 5 are presented in Fig. 6. The erroneous convection in the 2-h forecast over southwestern Missouri in 3DVar (Fig. 6C.) is reduced as the weight of the EnBEC increases. Less spurious convection and a better representation of the convection over eastern Missouri are achieved in the 2-h forecast with 100% EnBEC (Fig. 6A.), while a degradation is observed over western Missouri and central Tennessee. With 75% of the EnBEC (Fig. 6E.), the simulated convection over eastern Missouri in the 6-h forecast better matches the observations than with 100% EnBEC (Fig. 6D.) or 3DVar (Fig. 6F.), and the evolution of the convective system in this area seems to be better represented with less convection in the domain overall. The MMI (F+O) results for these three

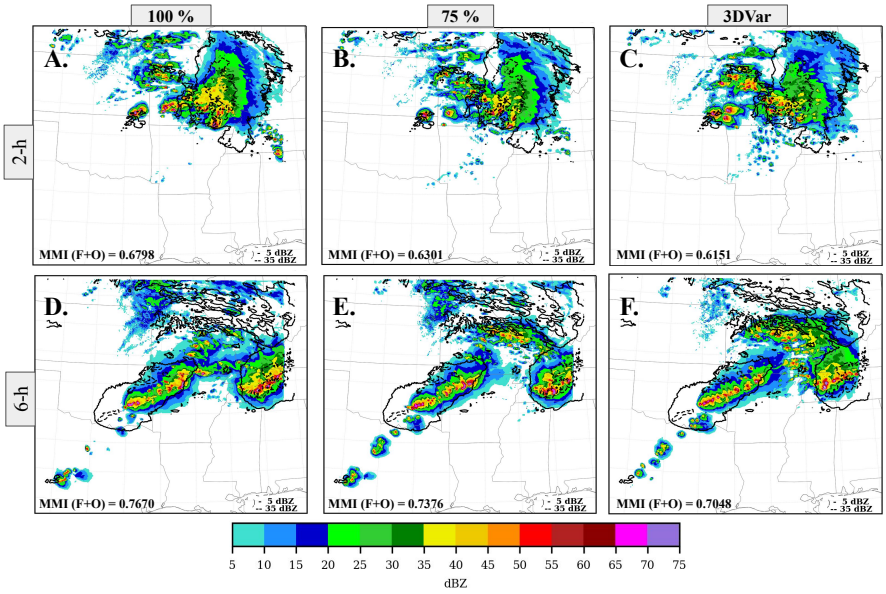

Figure 6: 2- (upper panels) and 6-h (lower panels) composite reflectivity forecasts initialized at 19Z on May 4 2020 from the experiments using the RRFSv1 physics suite with 100% of the EnBEC (A. and D.), 75% of the EnBEC B. and E.), and 3DVar (C. and F.) with G1 and CTLOBS data assimilation. Solid and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

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 (Fig. 6A. and D.). Although the 100% EnBEC experiment shows better results in Fig. 6 for the storm forecast in the 2- forecast, experiment with 75% of the EnBEC shows reasonable good results for the 2-h forecast and the best performance in the 6-h forecast, therefore, subsequent experiments were conducted using 75% EnBEC in order to leverage the influence of the static background—similar to operational configurations—when evaluating forecasts in which other GSI parameters are tested. Hereinafter, this experiment using 75% EnBEC, G1, and CTLOBS is called the control (CTL) experiment.

3.3 Supersaturation removal, PBL Pseudo-observations, and spatial localization in GSI hybrid analysis

Since an overestimation of convection with strong cells is observed in many of the experiments, a GSI parameter is tested that can limit supersaturation in the background specific humidity (CIMSS, 2014). In addition, an experiment is conducted to evaluate the impact of adding PBL pseudo-observations innovations based on surface temperature and moisture observations. The PBL pseudo-observations are used in the operational RAP to give a more realistic analysis for the levels inside the PBL (Benjamin et al., 2016). Another important aspect in hybrid EnVar algorithms is the spatial localization of the ensemble-based covariances (Buehner, 2005). Two experiments are conducted reading the horizontal and vertical localization scales for each level of the analysis grid from an external file. Tested vertical localization scales are of 1 level in the first 10 model vertical levels, and 3 levels (VLOC_1_3) and 2 levels (VLOC_1_2) in higher levels. A neutral impact is found when changing the vertical localization from 3 to 2 levels above level 10, and thus, only results using 3 levels are shown.

Figure 7 presents results of these three experiments (Fig. 7B., C., D., F., G., and H.) and the CTL experiment (Fig. 7A. and E.), where the vertical localization is set to 3 for all levels as in RAP. When supersaturation is removed in the analyses, slightly more spurious convection and intense individual cells are developed in both forecast hours shown, but a better evolution of the squall line is observed in the 6-h forecast with a MMI (F+O) value of 0.8048, the best among these experiments (Fig. 7B. and F.). The forecasts with the PBL pseudo-observations overestimate convection in southern Missouri, western Tennessee and Kentucky, and Texas in either forecast length. The convective initiation over northeastern Oklahoma and southeastern Kansas is also overproduced in the 2-h forecast (Fig. 7C.), but the 6-h forecast had better coverage of the squall line and the convective system over western Tennessee and Kentucky with 0.8032 of MMI (F+O) for this hour (Fig. 7 G.). Reducing the vertical localization to 1 level in the first 10 model vertical levels slightly improved the extent and intensity of convection over northeastern Oklahoma, eastern Missouri, and Illinois in the 2-h forecast (Fig. 7 D.). However, a degradation of the forecast quality was observed in the 6-h forecast with a decrease in the MMI (F+O) value to 0.7080 (Fig. 7 H.). Upper-air and surface verification show, in general, less skillful forecast with this configuration (not shown).

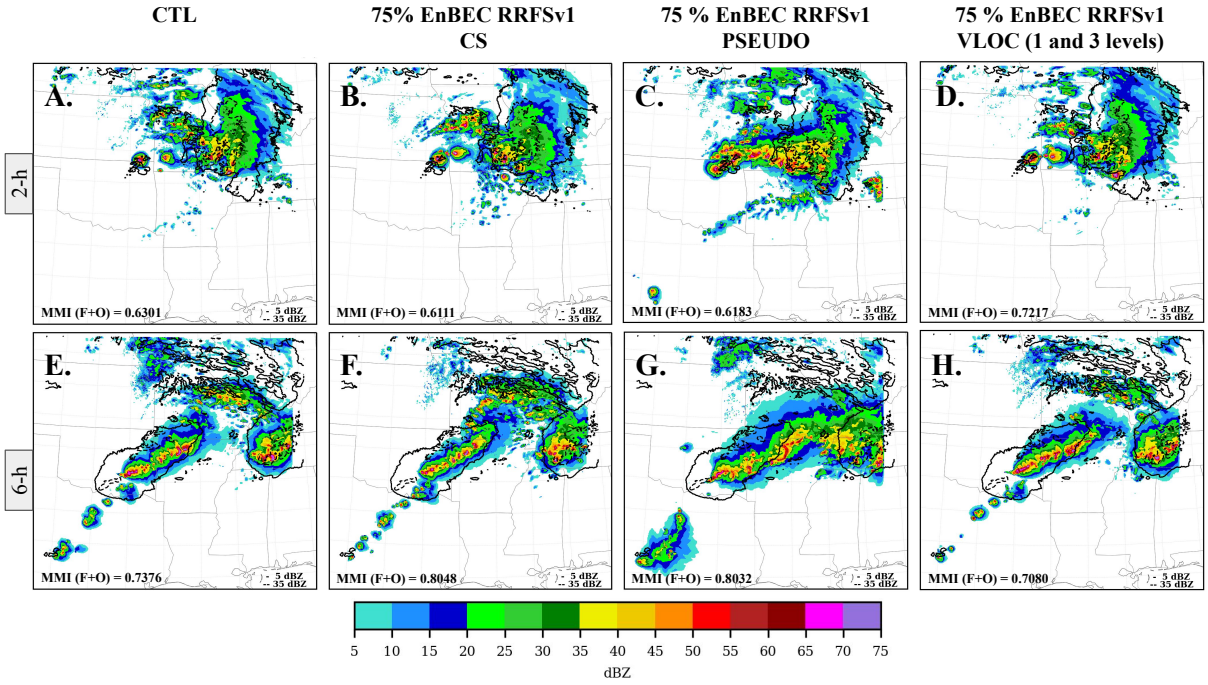

Figure 7: As in Fig. 4, but for experiments using the RRFSv1 physics suite with 75% EnBEC, G1, and CTLOBS data assimilation (A. and E.), with supersaturation removal activated (B. and F.), with the PBL pseudo-observation activated (C. and G.), and with vertical localization of 1 level in the first 10 model vertical levels and 3 levels in higher levels (D. and H.).

3.4 Observation System Experiments

Observation System Experiments (OSEs) are conducted in order to assess the impact of each observation type in CTLOBS as well as of adding GNSS-RO bending angles. Table 2 lists OSEs with the observations used in each experiment. In these experiments, the RRFSv1 suite with

Table 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+RW	T	T	T	**F**	**F**	**F**
UP+SRF+RW+SAT	T	T	T	T	**F**	**F**
CTL	T	T	T	T	T	**F**
CTL+BND	T	T	T	T	T	T

G1 grid and 75% EnBEC are used. Here, the experiment using RRFSv1 suite with 75% of the EnBEC and assimilating all data included in CTLOBS is called CTL.

The 2- and 6-h composite reflectivity forecasts from the UP, UP+SRF, CTL, and CTL+BND experiments are presented in Fig. 8. The positive impact of assimilating surface data can be clearly seen in the UP+SRF experiment (Fig. 8B. and F.), with better convective initiation over northeastern Oklahoma in the 2-h forecasts and the evolution of the squall line in 6-h forecast compared to UP. Without using surface data, the convective initiation is missed in the 2-h forecast over northeastern Oklahoma, as well as over central Texas in the 6-h forecast with UP (Fig. 8A. and E.). Adding radar radial velocity and satellite derived winds, and precipitable water slightly degrade the 2-h forecast quality (Fig. 8C.) when compared with UP+SRF, with MMI (F+O) values of 0.6301 versus 0.6385, respectively. Furthermore, although in CTL the convection over western Tennessee and Kentucky is better simulated in the 6-h forecast (Fig. 8G.), an overall degradation is observed with 0.7306 of MMI (F+O) compared with 0.8326 of MMI (F+O) in UP+SRF (Fig. 8F.). This degradation is mitigated when assimilating bending angle along with CTLOBS (Fig. 8D. and H.). In CTL+BND the spurious convection is not only decreased in both forecast lengths, but a slightly better simulation of the convection is also produced over eastern Missouri in the 2-h forecast and over western Tennessee in the 6-h forecast. However, the coverage and intensity of convection over northeastern Arkansas is overestimated in the 2-h forecast (Fig. 8D.). 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.

3.5 Cycling with GFS initial conditions and spin-up cycles

As described in Sect. 2.3, different cycling configurations are tested in order to use the GFS analyses as cold start ICs. Figure 9 shows the 2- and 6-h composite reflectivity forecasts from the experiments using the GFSics+CONTsurf and GFSics+surf cycling configuration with the RRFSv1 CCPP suite and hybrid data assimilation using 75% EnBEC, G1, and CTLOBS, comparing with the similar forecasts from the CTL experiment. In the 2-h forecast, MMI results show that CTL (Fig. 9A.) is slightly better than GFSics+CONTsurf (Fig. 9B.). Although the coverage of the convection over eastern Missouri, western Illinois, western Tennessee, and western Kentucky is better represented in GFSics+CONTsurf and the spurious convection developed over southwestern Missouri and southern Arkansas in the CTL experiment is greatly diminished

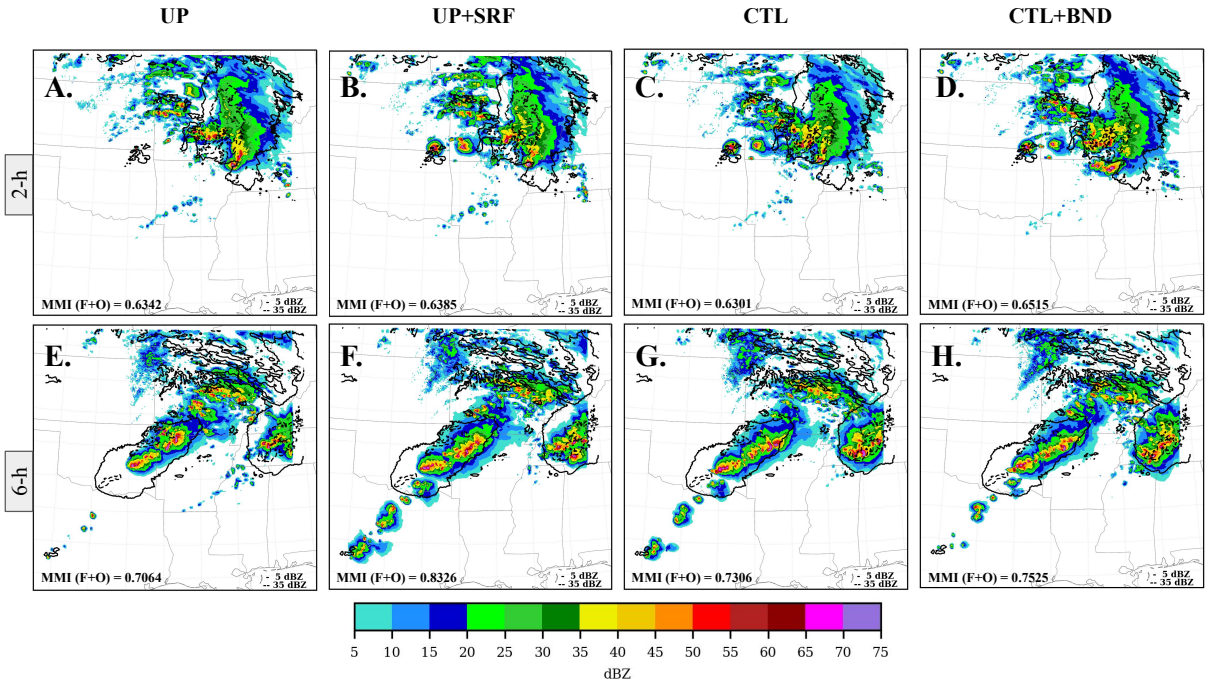

Figure 8: As in Fig. 9, but for experiments UP (A. and E.), UP+SRF (B. and F.), CTL (C. and G.), and CTL+BND (D. and H.).

as well (Fig. 9B.). Using GFS ICs with GFSics+surf cycling configuration give an overall better simulation of the convection compared to CTL and GFSics+CONTsurf results (Fig. 9C.). In the 6-h forecast, results using the GFSics+surf cycling configuration (Fig. 9F.) also indicate a better representation of the convection associated to the squall line and the system over western Tennessee and western Kentucky, reducing the intensity and extent of the area of intense convection and matching the observed composite reflectivity slightly better. MMI (F+O) values confirm that this last experiment has better quality with an increase from 0.6301 in CTL to 0.6501 in the 2-h forecast and from 0.7411 in GFSics+CONTsurf to 0.7771 in the 6-h forecast. In the 6-h forecast, GFSics+CONTsurf shows a slightly better MMI value than CTL.

3.6 Precipitation verification

To further evaluate the experiments conducted in this study, the FV3-LAM 1-h accumulated precipitation forecasts are also analyzed (figure not shown). The 2- and 6-h forecasts of 1-h accumulated precipitation showed the worst performance for values >1.0 inches in all of the experiments assessed. The performance is degraded as the precipitation threshold increases, indicating the difficulty in predicting rarer events. Lighter precipitation forecasts show better skill in all of the experiments and forecast hours. However, it was still far from the desired quality. Among the experiments presented, the ones using 100% EnBEC and the GFSics+surf cycling configuration show the best scores in the 2-h forecast. Also, the experiment using 100% EnBEC had slightly better skill in the 6-h forecast than the other experiments presented.

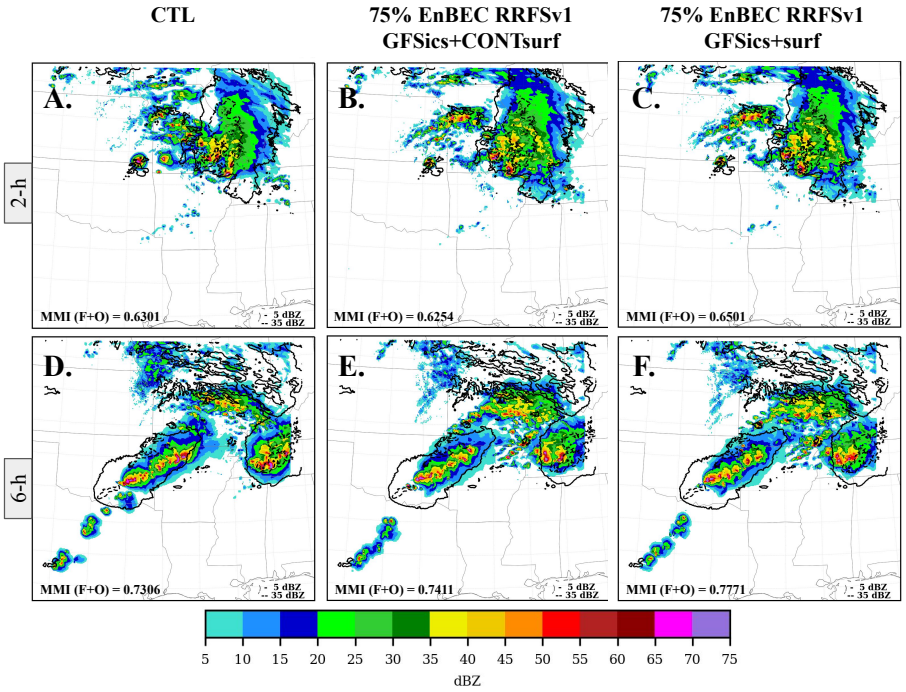

Figure 9: As in Fig. 7, but for experiments using the RRFSv1 physics suite with 75% EnBEC, G1, and CTLOBS data assimilation with (A. and E.), with supersaturation removal activated (B. and F.), with the PBL pseudo-observation activated (C. and G.), and with vertical localization of 1 level in the first 10 model vertical levels and 3 levels in higher levels (D. and H.).

4 Conclusions and future work

In this visitor project, the current RRFS capability to simulate convection is investigated through a case study of a squall line that occurred over Oklahoma during the afternoon of 4 May 2020. Various parameters and options are tested and evaluated in search of the best configuration to produce more realistic convection forecasts. The main findings of the study are listed below:

a) HRRRDA experiments show that FV3-LAM with the RRFSv1 physics suite is able to give reasonably good simulation of the convective initiation and evolution of the squall line. However, as the squall line evolved, the strongest cells developed ahead of what was observed;

b) Different CCPP suites give different characteristics in storm forecasts: the RRFSv1 has more strong, individual cells; the GFSv15 depicts weaker and smoother cells with larger coverage; the GSDsar physics suite misses the squall line initiation and evolution when no observations are assimilated; GFSreg captures the squall line structure well, but over-estimates the convection over Texas, Missouri and northern Arkansas. In general, the RRFSv1 produces the best simulation of the squall line and convective systems in this case study.

c) Data assimilation makes the analysis match the observations better. It can also improve the convection forecasts. Pure ensemble data assimilation produces better 2-h forecast for the storms, but 75% EnBEC produces good forecasts in both the 2- and 6-h forecast

lengths;

d) Supersaturation clipping in GSI can improve the specific humidity and temperature fields in the analyses. It produces more spurious convection and relatively more intense individual cells in the 2-h forecasts, but more skillful forecasts are produced at longer forecast lengths;

e) Convection was greatly over-predicted when using PBL pseudo-observations from surface 2-m temperature and 2-m moisture observations based on RAP configurations, indicating the need for more tuning in this function;

f) Reducing the vertical localization from 3 to 1 level in the lowest 10 levels of the analysis grid slightly improves the 2-h forecast of composite reflectivity, but degrades the 6-h forecast. Neutral impact is found when changing the vertical localization from 3 to 2-level in the model vertical levels above level 10. In general, a negative impact is found when using this configuration;

g) The assimilation of surface observations plays a crucial role in the convection forecasts because it improves convective initiation. Overall, a more complete observation dataset for assimilation produces more promising results;

h) Using GFS ICs with a more appropriate cycling configuration produces more skillful 2- and 6-h forecasts. It helps to mitigate the over-predicted storms seen in other experiments and represents the convective initiation, squall line structure, and convection patterns slightly better. The GFS ICs after spin-up cycles seems promising for RRFS applications;

i) FV3-LAM hourly accumulated precipitation forecasts for different thresholds indicate that heavy precipitation is more difficult to predict than light precipitation. The longer forecast lengths have a slightly better quality. The experiment using 100% EnBEC shows the best 1-h accumulated precipitation forecast quality in the 2- and 6-h forecast among the experiments evaluated, followed by the experiment with the GFSics+surf cycling configuration;

j) Most of the configurations tested are able to capture the main characteristics of the major convective systems during the execution period. However, the convection associated with the squall line tends to be overestimated in intensity and underestimated in its extent. MRMS observations show a stratiform region, often ahead of the convective cells, which is not well captured in the forecasts. This indicates that further testing and evaluation are warranted in addition to the options tested here.

In the future, the RRFS is intended to cover a similar domain as RAP, extending into part of northern South America. This region is also covered by the domain of the Hurricane Analysis and Forecast System (HAFS) (Dong et al., 2020), which is the UFS application for predicting hurricanes. The next part of this project will focus on investigating the RRFS's ability to represent convection in 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⟩.

Benjamin, S. G. et al. 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⟩.

Bernardet, L. et al. Engaging the community in the development of physics for nwp models. In: *EGU General Assembly Conference Abstracts*. [s.n.], 2020. p. 22093. Available at: ⟨https://ui.adsabs.harvard.edu/abs/2020EGUGA..2222093B⟩.

Buehner, M. Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quart. J. Roy. Meteor. Soc.*, v. 131, p. 1013–1043, 2005. Available at: ⟨https://doi.org/10.1256/qj.04.15⟩.

CIMSS. *CIMSS Cooperative Agreement Annual Report*. [S.l.], 2014. 302 p. Available at: ⟨https://cimss.ssec.wisc.edu/reports/CIMSS-CA-Report_2014_Final.pdf⟩.

Davis, C. A. et al. 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⟩.

Dong, J. et al. 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⟩.

EMC. *Strategic Implementation Plan for evolution of NGGPS to a national Unified Modeling System (First Annual Update)*. [S.l.], 2018. 171 p. Available at: ⟨https://www.weather.gov/media/sti/nggps/UFS\%20SIP\%20FY19-21_20181129.pdf⟩.

Garstang, M. et al. 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⟩.

Hu, M. et al. 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⟩.

Jensen, T. et al. *The Model Evaluation Tools v9.0 (METv9.0) User's Guide*. [S.l.], 2020. Available at: ⟨https://dtcenter.org/sites/default/les/community-code/met/docs/user-guide/MET_Users_Guide_v9.0.pdf⟩.

Lin, Y.; Mitchell, K. E. The ncep stage ii/iv hourly precipitation analyses: Development and applications. In: *19th Conf. on Hydrology*. Amer. Meteor. Soc., 2005. Available at: ⟨http://ams.confex.com/ams/pdfpapers/83847.pdf⟩.

UFS. *About UFS Applications*. 2019. Available at: ⟨http://ufs-dev.rap.ucar.edu/index.html\#/science/aboutapps⟩.

UFS Development Team. *Unified Forecast System (UFS) Short-Range Weather (SRW) Application v.1.0.0*. Zenodo, 2021. Available at: ⟨https://doi.org/10.5281/zenodo.4534994⟩.

Whitaker, J. et al. 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⟩.

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⟩.

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⟩.

Zhang, J. et al. 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⟩.