Project Title: **Investigating the FV3-CAM data assimilation capability representing convective processes and structures of squall lines over the Great Plains**

Graduate Student: **Ivette Hernández Baños**[1]

Advisor: **Dr. Luiz Fernando Sapucci**[2]

Collaborator: **Dr. Daryl Kleist**[3]

1 Abstract

The capability of the GFDL's FV3 dynamical core with advanced physics in a convective-scale has been explored and has shown promising results. Currently, efforts are being made to generate initial conditions that account for the model and the observations contribution on that scale. Into this framework, this proposal aims to investigate the dynamic and thermodynamic structure of convection associated with squall lines generated over the Great Plains represented by an FV3-based convective allowing model coupled with a data assimilation system. Following the installation and compilation of the system, domain configuration, tuning and adjustment of parameters in the data assimilation algorithm; numerical experiments will be conducted and results will be carefully analyzed. We expect to contribute with a preliminary evaluation of the FV3-CAM data assimilation system, highlighting its capability to adequately represent squall lines over the Great Plains and give indications to the development of the future operational convective-scale data assimilation scheme at CPTEC/INPE.

2 Introduction

Continued efforts have made possible the gradually increased forecast skill of the Global Forecast System (GFS) throughout its 39 years of history. Nowadays, GFS is the foundation of the National Centers for Environmental Prediction (NCEP) numerical prediction system and also of many operational centers and private initiatives around the world (EMC et al., 2019). However, Hurricane Sandy (2012) marked a turning point in the GFS evolution history because of the late landfall forecast of the Superstorm (Blake et al., 2013). Thereafter, the Research to Operation Initiative was approved to build the Next Generation Global Prediction System (NGGPS) aiming to improve the day-to-day weather and extended forecasts from larger to convective scales through the unification of the Earth system components (NOAA, 2014). After a thorough evaluation process, the non-hydrostatic finite volume cubed-sphere (FV3) dynamical core developed at the Geophysical Fluid Dynamics Laboratory (GFDL) was selected as the dynamical core of the NGGPS (Ji; (DTG), 2016). On June 12, 2019, NOAA successfully released its new operational global model, the FV3-based GFS model (NOAA, 2019). In addition, an FV3-based stand-alone regional model (SAR FV3) is under development to provide higher resolution forecasts in a unified manner from global to convective-scales (SAR FV3-based convective-allowing model – FV3-CAM) (Harris et al., 2019).

Besides the dynamical core, another important step is the implementation of a common physics package (EMC, 2018). Zhou et al. (2019) used the GFDL's FV3 with an upgraded six-category cloud microphysics scheme (fvGFS) to investigate the grid stretching ability to produce skillful forecast varying horizontal resolutions. Two case studies were carried out: a squall line developed in the Great Plains and Hurricane Harvey. In both, small scale structures of the convective activity were correctly resolved, though an overprediction of the precipitation and radar reflectivity was observed. Moreover, the track of Hurricane Harvey was accurately predicted, but the simulated squall line did not match the observed by 1 degree of difference. The importance of the initial condition was highlighted, which could have led to better results.

Since its origins, numerical weather prediction has strongly relied on the accuracy of the "current" state of the atmosphere given to the model as initial conditions for the integration of the primitive equations. Thus, the generation of accurate initial conditions is another important component of the NGGPS, since only an advanced dynamic core with a more realistic model physics cannot produce by itself a successful numerical weather forecast. The Joint Effort for Data assimilation Integration (JEDI), a collaborative effort to develop a unified data assimilation system, is currently under development which will allow a better representation of the

[1]Graduate Program in Meteorology, National Institute for Space Research, São José dos Campos, São Paulo 12227–010, Brazil

[2]Center for Weather Forecast and Climate Studies, National Institute for Space Research, Cachoeira Paulista, São Paulo 12630–000, Brazil

[3]NOAA/National Centers for Environmental Prediction/Environmental Modeling Center, College Park, MD 20740, United States of America

Earth system through the coupling of all its components across time scales with efficient research and operation transitions (Tremolet; Auligne, 2018; JEDI, 2019). The JEDI is a very broad and challenging project that is gradually maturing. Initially, for the new FV3-based GFS, analyses were enhanced with various modifications while maintaining the data assimilation system and configuration as close as possible to the previously used for the operational GFS (Kleist; Mahajan; Thomas, 2018). At present, too much effort is being made to produce the initial conditions for FV3-CAM systems in the JEDI framework (EMC, 2018).

Data assimilation aims to improve the analyses and, in turn, weather forecasts. However, in the convective scale, forecast errors can grow faster in a short period of time than on larger scales, nonlinearities arises and must be taken into account, circulations at these scales are strongly coupled to thermodynamic and wet microphysics processes, and movements are characterized by strong vertical velocity fluctuations (Gustafsson et al., 2018). Linear balances and constraints of the static background error commonly used for data assimilation on the synoptic scale are no longer appropriate. Therefore, the assumption of Gaussian background and observation errors with zero bias and precisely known covariances are no more recommended (Bastarz, 2017). A convenient approach is to combine two covariance matrices by means of a hybrid method. As much as possible, these elements are being considered in the generation of initial conditions for FV3-CAM systems. An important step of the process is the verification of the analysis and forecasts produced. Aspects directly related to convective processes and structures should be evaluated besides the use of advanced verification statistics. A proper verification can help to determine uncertainties and clarify aspects that require more attention into the system.

As mentioned before, first tests using fvGFS have demonstrated its capability simulating tropical and mid-latitude weather systems in a convective scale. Furthermore, with the current developments in the integration of an FV3-CAM with a data assimilation system, new horizons are being opened, but also some questions arise: Will FV3-CAM be able to better place squall lines over the Great Plains when using its own background for the initial conditions generation? Will FV3-CAM coupled with a data assimilation system adequately represent the initiation, propagation, and dissipation of squall lines over the Great Plains? It is well known that squall lines over the Great Plains are responsible for a large portion of the precipitation that falls in that region as well as severe weather occurrence (*e.g.*, Hane (1986), Klimowski et al. (2003), Hocker and Basara (2008)). Thus, an evaluation of the dynamic and thermodynamic structure of these squall lines when simulated using an FV3-CAM with data assimilation may give us a deeper understanding about the system's ability to adequately represent convection.

Goals

The goal of this proposal is to investigate the FV3-CAM data assimilation capability representing the dynamic and thermodynamic structure of convection generated by squall lines over the Great Plains. It will be accomplished by fulfilling the following specific objectives:

1. Assess the performance of the analysis generated by a convective-scale data assimilation technique coupled to an FV3-CAM system;
2. Investigate the skill of FV3-CAM forecasts representing patterns of the dynamic and thermodynamic structure of convection associated with mid-latitude squall lines, such as the developed over the Great Plains.

3 Methodology

The Gridpoint Statistical Interpolation (GSI) analysis system was initially developed by the NCEP's Environmental Modeling Center (EMC) (Wu; Purser; Parrish, 2002) and operationally implemented in May 2007 coupled to the GFS model. Since 2016, GSI has been using operationally a hybrid four dimensional (4D) ensemble-variational (EnVar) algorithm to generate GFS initial conditions (Kleist; Ide, 2015). The JEDI is expected to support several features currently available in GSI, while also going beyond in other aspects (EMC, 2018). As part of these efforts, data assimilation from global to convective-scale is already ongoing and with this proposal, we intend to contribute by carrying out an evaluation of the FV3-CAM data assimilation capability. A system available for tests will be used for the development of this research. The methodology that will be used is presented by phases as follows:

Phase A– Setup of the system: This phase will include setting up an account on the Developmental Testbed Center (DTC) computing resources, compiling and configuring the FV3-CAM and data assimilation systems, and performing a single observation test. In order to mitigate computational requirements, the Center for Weather Forecast and Climate Studies of the Brazilian National Institute for Space Research (CPTEC/INPE) will provide support in terms of processing and disk space. Thus, this phase will also include to compile and configure these systems on the CPTEC supercomputer. A 3 km grid-length covering the Contiguous United States (CONUS) will be used with lateral boundary conditions generated from an independent external forecast without a global parent domain (Black et al., 2019).

Phase B– Sensitivity study of setup parameters: This study will focus on tuning and adjusting important elements in the data assimilation system to generate a more accurate analysis. So, it will be conditioned to the data assimilation technique implemented at the time of execution of this proposal, which could be a variational, ensemble or a combination of both approaches in a hybrid method. If an FV3-CAM ensemble or hybrid data assimilation is available, this study will be oriented to assess parameters related to perturbation inflation and covariance localization, together with the evaluation and adjustment of stochastic physics techniques already expanded to the FV3 SAR system. Initially, the ensemble could be composed by 8 members and increase this size as computational resources allow it.

Phase C– Numerical experiments execution: Numerical experiments will be conducted to simulate squall lines that occurred in the Great Plains through the spring and early summer of 2019. Prior to the execution of this proposal, an observational study will be conducted to identify the squall line cases. Representative case studies will be analyzed and the more intense will be selected and simulated during this phase. The same model and data assimilation system setup will be used to execute various cases as long as time and computational resources allow. Available conventional and unconventional observations will be assimilated in the experiments.

The first steps of the observational study are already underway. The Forecasting and Tracking the Evolution of Cloud Clusters (Fortracc) algorithm proposed by Vila et al. (2008) and running operationally at CPTEC/INPE, was installed and tests are being executed. Data from the GOES-16 infrared channel 13 and radar reflectivity fields are being used to track and characterize squall line cases.

Phase D– Results evaluation: The dynamic and thermodynamic structure of the simulated squall line will be verified using observations from available sources, such as automatic and conventional stations, METeorological Aerodrome Reports (METARs), upper-air radiosondes and radar data. Model forecasts will be spatially verified by applying the method for object-based evaluation available in the Model Evaluation Tools (MET) verification package (Bullock; Brown; Fowler, 2016). Radar data will be used to verify the model reflectivity as well as precipitation fields.

Phase E– Final report: A technical report summarizing each phase outcomes will be elaborated and submitted to the DTC and a seminar will be presented highlighting the results achieved. In addition, a scientific article will be prepared.

The execution of this proposal has an expected period of twelve months and it is intended to be carried out at DTC and NOAA/Environmental Modeling Center (EMC) under guidance of Dr. Daryl Kleist.

The execution of the work phases detailed above is planned to be executed according to Table 1:

Table 1: Timeline for the project execution

Phases	1	2	3	4	5	6	7	8	9	10	11	12
A	x	x	x	x								
B				x	x	x	x					
C							x	x	x			
D								x	x	x	x	
E											x	x

Computational resource requirements

The estimated necessary disk space is approximately up to 20 TB. However, as mentioned before, in order to decrease this demand, part of the research will be supported by the computational infrastructure of CPTEC/INPE designed for Graduate Student research. 10 TB of disk space will be available to store some of the experiments performed.Dias and Genhe, (2018) suggested that the tropical forecast skill in the first days can improve mid-latitude forecast at lead times beyond day-3.

4 Expected Outcomes

For several years, NOAA (United States) and INPE (Brazil) have had cooperation agreements on Earth observation, which have allowed, data analysis and dissemination, data exchange without charges, support of global and regional numerical weather prediction developments, among several other benefits. Following NOAA recent initiative, INPE is on the first steps of the implementation of a new dynamical core using finite volume. Current efforts are also focusing on regional modeling developments, including convective-scale data assimilation. This proposal forms part of the Doctoral Research of the Graduate Student at INPE and aims to further narrow the cooperation in this area and bring results that can benefit both institutions. Some of which are expected:

- Since the FV3-CAM project is underway, several tests of all its components are necessary to improve it until a robust version is achieved. This proposal expects to contribute with some of the necessary adjustments in the convective-data assimilation system, determine possible shortcomings and indicate potential solutions for an analysis more representative and appropriate for this scale;
- Over the Great Plains, severe thunderstorms in the form of squall lines have a high frequency, so its early and accurate forecast is necessary. With this proposal, we expect to obtain a preliminary insight about the capacity of FV3-CAM with a coupled data assimilation system to predict convective process associated with these squall lines;
- Convective-scale data assimilation is an approach very incipient at CPTEC/INPE. Through the use of a high-resolution state-of-the-art numerical model and data assimilation system we hope to improve our knowledge on this subject, learn to how use this type of systems, and give indications to the development of the future operational convective-scale data assimilation scheme at CPTEC/INPE. In addition, this research will be used as a basis for future studies on squall lines over the Brazilian Amazon.

References

Bastarz, C. F. *Assimilação de dados global híbrida por conjunto-variacional no CPTEC*. 275 p. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2017-07-18 2017.

Black, T. L. et al. A Standalone Limited Area Capability for the Finite-Volume Cubed-Sphere Dynamic Core. In: *Section 5: Development of and studies with regional and convective-scale atmospheric models and ensembles, Astakhova, E. (Report 49) Research Activities in Atmospheric and Oceanic Modelling (WGNE Blue Book 2019)*. Geneva, CH: [s.n.], 2019. p. 3–4. Available from Internet: ⟨http://bluebook.meteoinfo.ru⟩.

Blake, E. S. et al. *Tropical Cyclone Report Hurricane Sandy*. Rep. AL182012, 2013. 157 p.

Bullock, R. G.; Brown, B. G.; Fowler, T. L. Method for Object-Based Diagnostic Evaluation. NCAR Technical Note NCAR/TN-532+STR. p. 84, 2016.

EMC. *Strategic Implementation Plan for evolution of NGGPS to a national Unified Modeling System (First Annual Update)*. [S.l.], 2018. 171 p. Available from Internet: ⟨https://www.weather.gov/sti/stimodeling_nggps_implementation⟩.

EMC, C. P. et al. The Development and Success of NCEP's Global Forecast System. In: *99th American Meteorological Society Annual Meeting*. Phoenix, AZ: [s.n.], 2019. Available from Internet: ⟨https://ams.confex.com/ams/2019Annual/webprogram/⟩.

Gustafsson, N. et al. 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 from Internet: ⟨https://doi.org/10.1002/qj.3179⟩.

Hane, C. E. *Extratropical Squall Lines and Rainbands*. [S.l.]: American Meteorological Society, Boston, MA, 1986. 359–389 p. ISBN 978-1-935704-20-1.

Harris, L. M. et al. Explicit prediction of continental convection in a skillful variable resolution global model. *Journal of Advances in Modeling Earth Systems*, v. 11, p. 1847–1869, 2019. Available from Internet: ⟨https://doi.org/10.1029/2018MS001542⟩.

Hocker, J. E.; Basara, J. B. A 10-year spatial climatology of squall line storms across oklahoma. *Int. J. Climatol.*, v. 28, p. 765–775, 2008.

JEDI, C. T. *JEDI Documentation Release 1*. [S.l.], 2019. Available from Internet: ⟨https://readthedocs.com/projects/jointcenterforsatellitedataassimilation-jedi-docs/downloads/pdf/latest/⟩.

Ji, M.; (DTG), N. D. core T. G. *Dynamical Core Evaluation Test Report for NOAA's Next Generation Global Prediction System (NGGPS)*. [S.l.], 2016. 93 p.

Kleist, D.; Ide, K. An OSSE-Based Evaluation of Hybrid Variational–Ensemble Data Assimilation for the NCEP GFS. Part II: 4DEnVar and Hybrid Variants. *Mon. Wea. Rev.*, v. 143, p. 452–470, 2015.

Kleist, D.; Mahajan, R.; Thomas, C. Data Assimilation in the Next Generation Global Prediction System (NGGPS) Era: Initial Implementation of FV3-based Global Forecast System (GFS). *JCSDA NEWSLETTER*, n. 61, p. 1–9, 2018. Available from Internet: ⟨https://www.jcsda.org/news⟩.

Klimowski, B. A. et al. Severe convective windstorms over the northern high plains of the United States. *Wea. Forecasting*, v. 18, p. 502–519, 2003.

NOAA. *R2O Initiative: Next Generation Global Prediction System (NGGPS) Implementation Plan (v1.0)*. [S.l.], 2014. 51 p. Available from Internet: ⟨http://cmalibrary.cn/xxcp/zt1/201702/P020170220814756089088.pdf⟩.

NOAA. *NOAA upgrades the U.S. global weather forecast model*. 2019. Available from Internet: ⟨https://www.noaa.gov/media-release/noaa-upgrades-us-global-weather-forecast-model⟩.

Tremolet, Y.; Auligne, T. The Joint Effort for Data assimilation Integration (JEDI). In: *Joint Center for Satellite Data Assimilation (JCSDA) NOAA Testbed and Proving Ground Workshop*. Kansas City, MO: [s.n.], 2018. Available from Internet: ⟨https://www.testbeds.noaa.gov/events/2018/workshop/presentations/⟩.

Vila, D. A. et al. Forecast and tracking the evolution of cloud clusters (fortracc) using satellite infrared imagery: Methodology and validation. *Wea. Forecasting*, v. 23, p. 233–245, 2008. Available from Internet: ⟨https://doi.org/10.1175/2007WAF2006121.1⟩.

Wu, W.; Purser, R. J.; Parrish, D. Three-Dimensional Variational Analysis with Spatially Inhomogeneous Covariances. *Mon. Wea. Rev.*, v. 130, p. 2905–2916, 2002.

Zhou, L. et al. Toward Convective-Scale Prediction within the Next Generation Global Prediction System. *Bulletin of the American Meteorological Society*, v. 100, p. 1225–1243, 2019. Available from Internet: ⟨https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018GL081702⟩.