Data Assimilation of Lightning in WRF 3/4-D VAR Using Observation Operators

Răzvan Ştefănescu[1], Michael I. Navon[1], Max Marchand[2], Henry Fuelberg[2]

[1] *The Florida State University, Department of Scientific Computing,*

Tallahassee, Florida 32306, USA

rstefanescu@fsu.edu, inavon@fsu.edu

[2] *The Florida State University, Department of Meteorology,*

Tallahassee, Florida 32306, USA

maksmarchand@gmail.com, hfuelberg@fsu.edu

Abstract

Compared to other types of satellite-derived data, assimilating lightning data into operational numerical models has received relatively little attention. NASA will launch in 2015 the GOES-R Lightning Mapper (GLM) that will provide continuous, full disc, high resolution total lightning (IC + CG) data. Previous efforts of lightning assimilation mostly have employed nudging. However, we propose to develop a more sophisticated approach involving 3-D VAR and 4-D VAR toward which both NCEP and NRL are moving. The early stages of our research will utilize existing ground-based lightning data that can be assimilated prior to the launch of GLM; later phases will utilize GLM proxy data that will mimic what GLM will detect.

A major difficulty associated with this exercise is the complexity of the observation operator defining the model equivalent of the lightning. It is using Convective Available Potential Energy (CAPE) as a proxy between lightning data and model variables. This operator is highly nonlinear. Marecal and Mahfouf (2003) have shown that nonlinearities can prevent a direct assimilation of rainfall rates in the ECMWF 4D-VAR (using the incremental formulation proposed by Courtier et al. (1994)) from being successful. In our case we also proved that the direct assimilation of lightning into the WRF 3D - VAR schemes is limited due to this incremental approach. Severe threshold limits must be imposed on the innovation vectors in order to obtain an improved analysis.

By adjusting the temperature lapse rate, we directly assimilate WTLN (Worldwide Total

Lightning Network) total lightning data for the 2011 Tuscaloosa, AL tornado outbreak in a domain of 406 × 305 km^2 with a mesh resolution of 1 km in each horizontal direction and 60 vertical levels.

Next we developed a new scheme similar to the one outlined in Marecal and Mahfouf (2002, 2003) i.e., use 1-D VAR to adjust rainfall rate from the moist physics (mass-flux convection scheme and large-scale condensation) closer to an observed value as in Mahfouf et al. (2005). In our case, the 1D-VAR temperature columns retrievals are considered as new observations and are assimilated in the 3D/4D-VAR systems . These approaches are denoted the '1-DVAR + 3-DVAR','1-DVAR + 4-DVAR' approaches (Mahfouf et al. 2005). It minimizes the problem that nonlinearities of the moist convective scheme can introduce discontinuities in the cost function between inner and outer loops of the incremental 3-D/4-D VAR minimization.

Next we present some results obtained with the 1-DVAR +3-DVAR approach. Again, we applied the scheme for the 2011 Tuscalosa, Al tornado outbreak on a horizontal domain of 100 × 100 km^2 with a mesh resolution of 1 km and 60 vertical levels.

The 1-DVAR +4-DVAR approach requires tremendous computational resources at $1km$ resolution. An assimilation window of 6 hours demands 10^{11} double precision variables to be stored in the virtual memory, which will require 1000 Gbs. Even with I/O checkpoints the computer needs to store 10^7 variables which make every 4-DVAR run impossible at $1km$ resolution. One solution will be the split of the assimilation period in small windows with the length of one hour, with the derived 4D-Var problems being solved as new data become available. The release of the parallel version of WRFDA will allow us to test both approaches, the direct lightning assimilation scheme and the 1-DVAR method, in 4-DVAR framework.

1 Introduction

Lightning data constitutes one of the relatively new sources of information considered in the framework of data assimilation. The earliest effort of lightning data assimilation can be attributed to Alexander et al (1999) using assimilating rain-rates derived from satellites and lightning on forecast of 1993 super storm.

Here we investigate effects of lightning on a severe super storm. Prior efforts aimed at assimilating lightning data into numerical models, focused on Newtonian nudging in which vertical profiles of humidity and/or latent heat were altered.

More efforts were carried out by Pessi and Businger (2009) and their group that developed a lightning - convective rainfall rate based on long range lightning data and TRMM PR rainfall data using the MM5 model. Mansell et al (2007) used NLDN and LMA lightning flashing data to control the Kain - Fritsch convective parametrization scheme. Papadopoulos et al (2005) used real time C-G flash rate data from LRLDN network to nudge model generated humidity profiles as o function of lightning intensity.

In as far as lightning data, use was make of NLDN (Orville 2008, Biagi et al 2007) and other long range network Pacific Lightning detection, as limited area VHF network LMA (Rison et al, 1999, Thomas et al 2004). We utilize total lightning data such as WWLLN (see Rodger et al (2008)).

There is considerable interest in discovering of nearly continuous lightning data, could be used to improve convection forecast skill. Our interest will fall on assimilating total lightning data from the Geostationary Lightning Mapper (GLM) and the Geostationary Operational Environmental Satellite "R" (GOES -R) into a tropical cyclone model and a super storm situation characterized by the Tuscaloosa super storm.

The GLM will map total lightning (CC + CG) for a uniform resolution varying between 8-12 km supposed to achieve a detection efficiency of 70% or more.

As we mentioned before we proposed two assimilation technique using CAPE as a proxy between lightning data and model variables in conjunction with the WRFDA-3DVAR system. First we briefly introduce the WRFDA 3-DVAR framework. Next we describe the direct assimilation approach which involves discussion about the observation operator, background error covariance matrix, observation error covariance matrix,etc. Finally we provide information about the 1-DVAR+3-DVAR technique, an approach which is more consistent with the incremental WRFDA 3D-VAR, allowing us to assimilate most of the lightning observations without the necessity of imposing the observation filter used in the direct assimilation approach.

2 WRFDA 3D-VAR System

As a reminder, the 3-DVAR method produces an "optimal" estimate of the true atmospheric state at the analysis time through the iterative solution of a prescribed cost function,

$$J(X_0) = \frac{1}{2}(X_0 - X_0^b)^T \mathbf{B}^{-1}(X_0 - X_0^b) + \frac{1}{2}(H(X_0) - Y_0)^T \mathbf{R}^{-1}(H(X_0) - Y_0).$$

This solution represents the a posteriori maximum likelihood (minimum variance) estimate of the true atmospheric state given two sources of a priori data: the background (previous forecast) X_0^b and observations Y_0. H is the observation operator that converts the initial state into observed equivalents for comparison with the corresponding observations, $\mathbf{R}$ is the error covariance matrix of the observations and matrix $\mathbf{B}$ contains the background error covariances for each atmospheric variable. The control variable are: streamfunction, unbalanced velocity potential, unbalanced temperature, relative humidity and unbalanced surface pressure.

The WRFDA-3-DVAR adopts the incremental VAR formulation that is commonly used in operational systems. The incremental approach is designed to find the analysis increment that minimizes a cost function defined as a function of the analysis increment by using a linearized observation operator.

$$J(v) = \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}(\mathbf{d} - \mathbf{H}'\mathbf{U}\mathbf{v})^T R^{-1}(\mathbf{d} - \mathbf{H}'\mathbf{U}\mathbf{v}),$$

where $\mathbf{d} = Y_0 - H(X_0^b)$ is the innovation vector, $\mathbf{H}'$ is the linearization of the observation operator H, $\mathbf{B} = \mathbf{U}^T\mathbf{U}$ and $X_0 - X_0^b = \mathbf{U}\mathbf{v}$.

3 Direct assimilation of lightning using WRFDA -3-DVAR

For WRFDA to assimilate lightning our choice depends on the horizontal resolution of the WRF model. If we choose a mesh which is cloud resolving, we can calculate flash rates based on ice fluxes. These are the approaches described by Barthe, C et al 2010. Thus we want to exploit the strong linear correlation between the maximum vertical velocity and the total flash rate:

$$f = 5 \cdot 10^{-6}(0.677 w_{max} - 17.286)^{4.55},$$

where f is the total flash rate and w_{max} the maximum vertical velocity.

So we link the maximum vertical velocity to the lightning flash rate and then translate it in temperature lapse rate using CAPE (Convective available potential energy)

$$w_{max} = \sqrt{2 \cdot CAPE},$$

according with parcel theory, so we have to adjust for entrainment.

$$\text{CAPE} = \int_{z_g}^{z_n} g \frac{T_{\text{parcel}} - T_{\text{env}}}{T_{\text{env}}} dz,$$

with T_{parcel} the virtual temperature of the specific air parcel, T_{env} the environment temperature, z_f and z_n the heights of free convection and that of equilibrium (neutral buoyancy).

Thus, our operator has the following form

$$H(X) = 5 \cdot 10^{-6}(0.677\sqrt{2 \cdot CAPE(X)} - 17.286)^{4.55}.$$

The tangent linear and the adjoint models were obtained by using the TAPENADE Automatic Differentiation Engine and were tested with the alpha test described in I.M. Navon et al 1992.

The background error covariance matrix B describes the PDF of forecasts errors. For our experiments we estimate B with the NMC method. We used $12h$ and $24h$ forecasts valid for the same time for a one month long dataset generated by WRF model.

The observation error correlations were assumed to be zero so in consequence the observation error covariance matrix is a diagonal matrix. The assimilation tests were performed using simply the identity matrix for the matrix R.

Next we describe the performances of this scheme for the 2011 Tuscaloosa, AL tornado outbreak on a domain of 406×305 km^2 with a mesh resolution of 1 km and 60 vertical levels. When we assimilated all the available lightning observations the conjugate gradient minimization algorithm failed to find a global minimum due to the moist convective scheme which introduced discontinuities in the cost function between inner and outer loops of the incremental 3-DVAR minimization caused by the lightning nonlinear operator. By taking into account only the observations which have corresponding innovation vectors smaller then a prescribed threshold we imposed an upper limit for the temperature perturbations introduced in the moist convective scheme thus allowing the WRFDA 3-DVAR minimization algorithm to work properly.

When comparing the $6h$ forecast started with an analysis obtained with a 5.5 threshold by assimilating both the available conventional observations and the lightning observations with a $6h$ forecast initialized with a WRFDA 3-DVAR analysis improved only with the conventional observations we get the maximum benefit in term of the temperature root mean square error calculated for the entire domain using the NCEP FNL Operational global analysis as truth. We reduced the temperature root mean square error by 3.1%.

4 1-DVAR+3-DVAR lightning assimilation scheme

To avoid the nonlinearities effects introduced by our lightning operator and to extract the benefit from all the available flash observations we propose a 1-DVAR+3-DVAR technique similar with the one outlined in Marecal and Mahfouf (2002, 2003) to assimilate surface rain - rate retrievals from the SSM/I and TRMM microwave imager. First, the raw measurements are used in 1-DVAR procedure to produce column temperature retrievals. These pseudo -observations are then assimilated into the WRFDA 3-DVAR system.

The 1-DVAR method searches for an optimal estimate of the temperature profile at the analysis

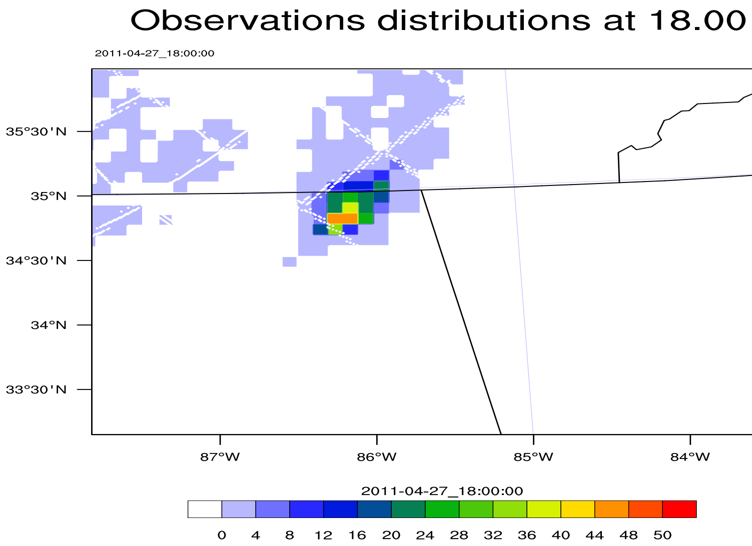


Figure 1: The lightning profile for the 2011 Tuscaloosa, AL tornado outbreak

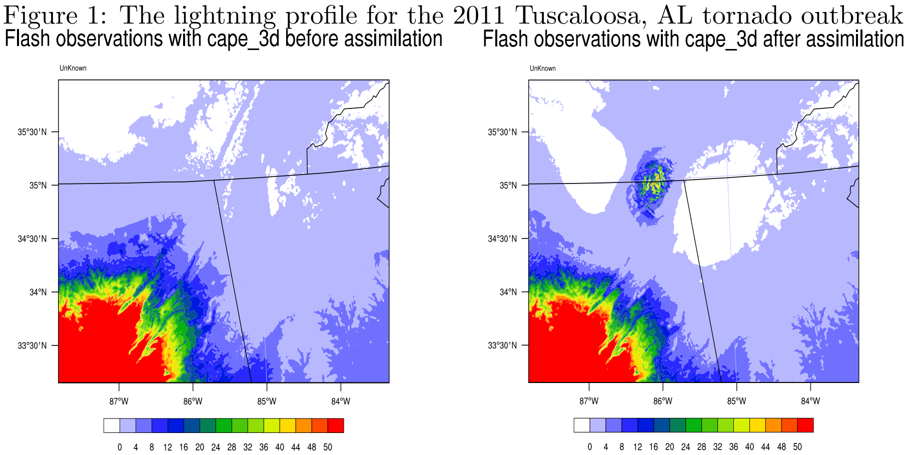


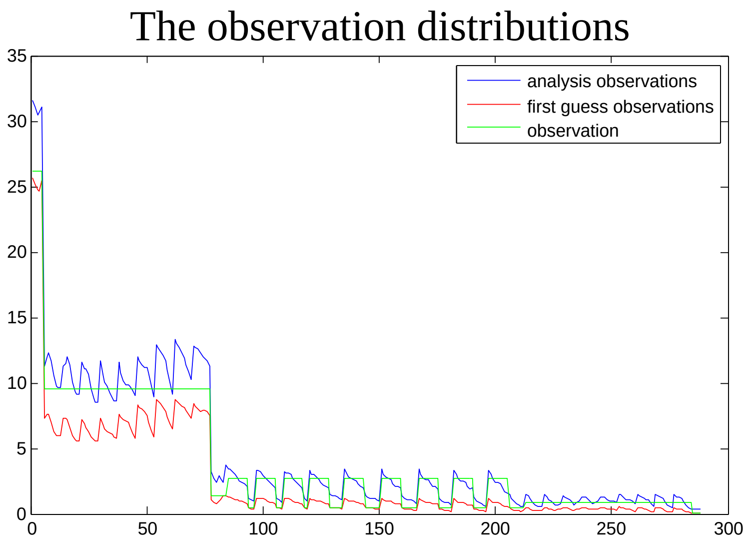
Figure 2: The flash distributions before (left) and after (right) the direct 3D-Var assimilation

time through the iterative solution of the following prescribed cost function,

$$J(X_0) = \frac{1}{2}(X_0 - X_0^b)^T B^{-1}(X_0 - X_0^b) + \frac{1}{2}\left(\frac{H(X) - y_0}{\sigma_0}\right)^2,$$

where H is the lightning observation operator introduced in the previous section and σ_0 is the observation variance. Using the NMC method we construct the background error vertical covariances matrix B for temperature profiles. The minimization is carried out not in the observation space rather in the space determined by the B eigenvectors. This approach is used for preconditioning the analysis equation. For the unconstrained minimization of the cost we used the conjugate gradient algorithm CONMIN proposed by Shanno and Phua (1980).

Our 1-DVAR + 3DVAR approach was tested again for the 2011 Tuscalosa, AL tornado outbreak on a horizontal domain of 100×100 km^2 with a mesh resolution of 1 km and 60 vertical levels. The impact of 1-DVAR is depicted in the next figure where we present a comparison between the simulated observations calculated before (first guess) and after the assimilation (analysis). The assimilation time was $18:00$ and we choose $\sigma_0 = 1$ for all the observations.



The vertical temperature profiles obtained after the 1-DVAR minimization were prepared as AIREP observations and assimilated in WRFDA 3D-Var system. We performed 6 assimilation exercises at $18:00$, $20:00$, $21:00$, $22:00$, $23:00$ and $00:00$. The results are shown in the next figures. Figure 3 shows the average change in the temperature profile after assimilation performed at $18:00$. We may notice an increase in temperature magnitude close to surface which is highly correlated with an increase of CAPE in order to support more flash lightnings.

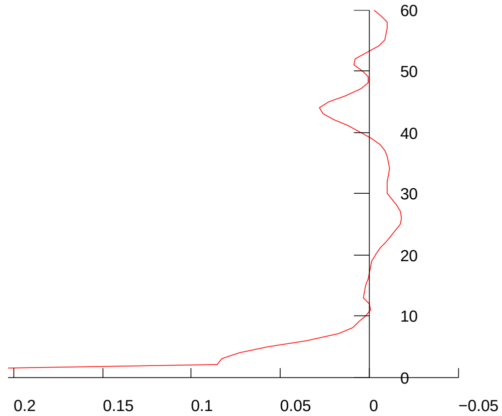


Figure 3: The average change in the temperature profile - 18 : 00

.

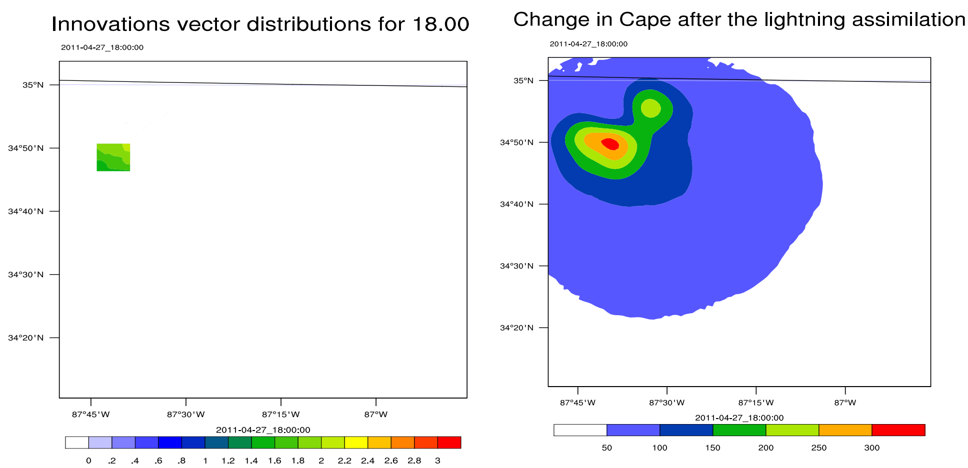


Figure 4: The innovation vectors distributions and change in CAPE after the assimilation - 18 : 00

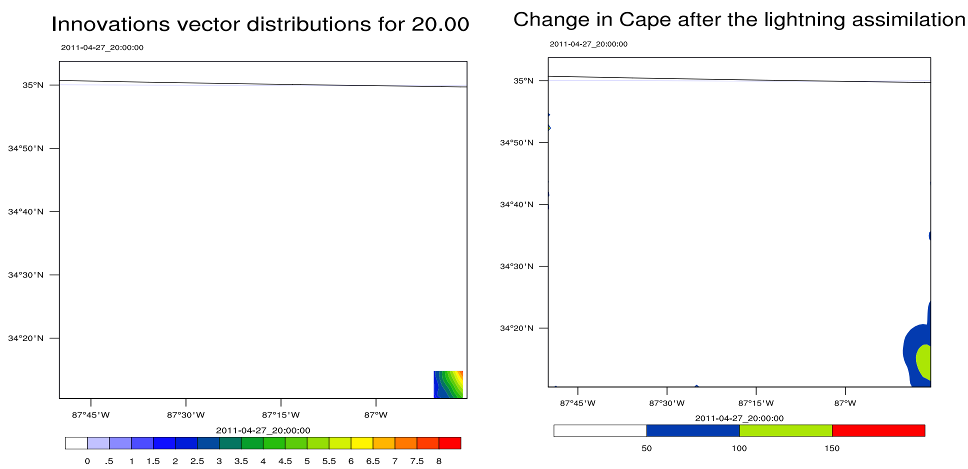


Figure 5: The innovation vectors distributions and change in CAPE after the assimilation - 20 : 00

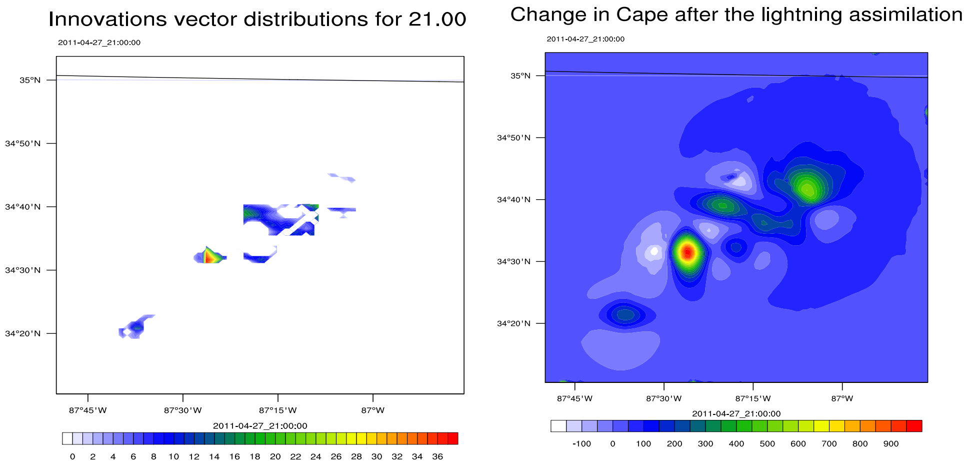


Figure 6: The innovation vectors distributions and change in CAPE after the assimilation - 21 : 00

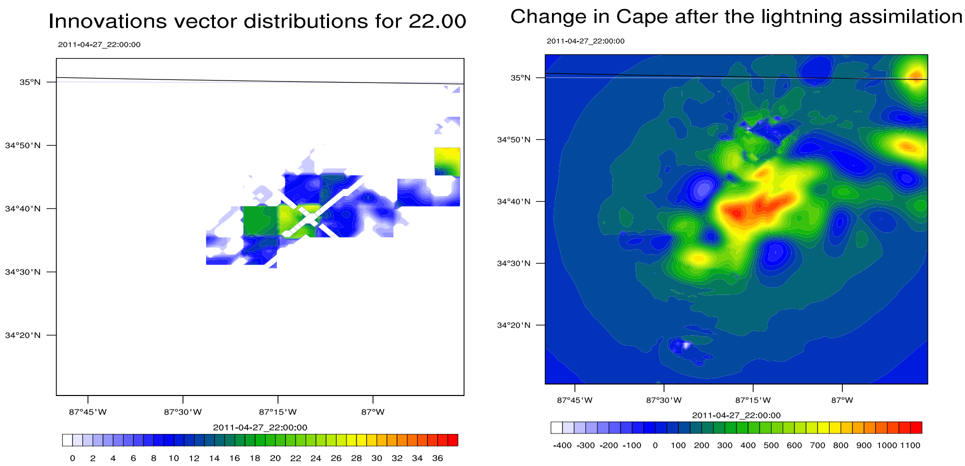


Figure 7: The innovation vectors distributions and change in CAPE after the assimilation - 22 : 00

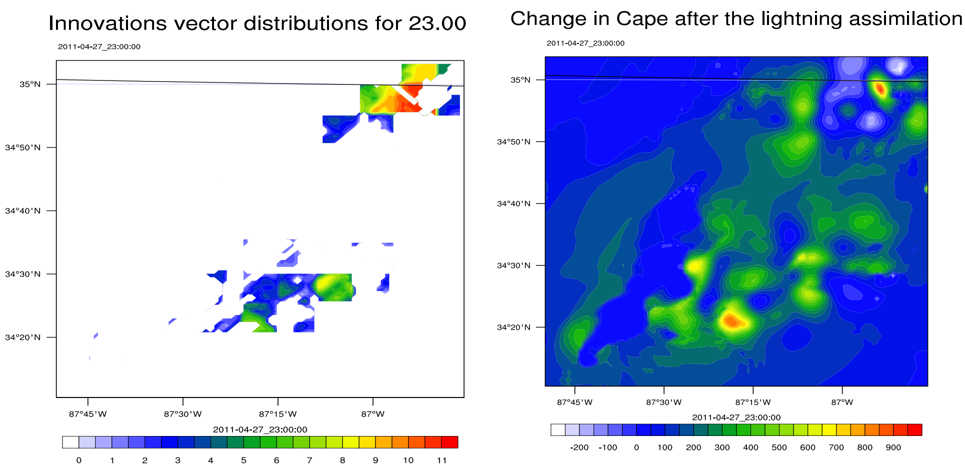


Figure 8: The innovation vectors distributions and change in CAPE after the assimilation - 23 : 00

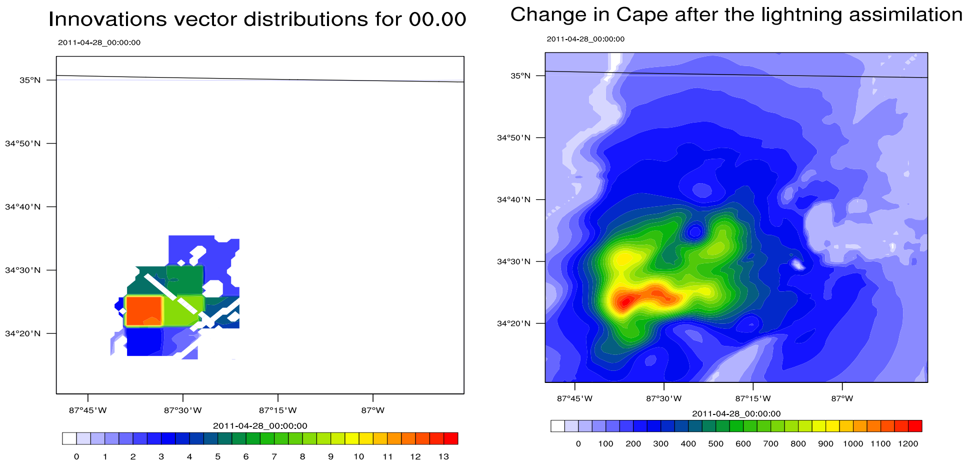


Figure 9: The innovation vectors distributions and change in CAPE after the assimilation - 00 : 00

The $6h$ forecast generated at 18 : 00 with an analysis improved by the lightning observations assimilation has temperature profiles closer to the temperature columns of the truth (NCEP FNL Operational global analysis at 00:00) than the $6h$ forecast initialized with the NCEP FNL analysis at 18 : 00. The impact was measured using the temperature root mean square error. Due to flash lightning assimilation we decreased the error by 16.1%.

5 Conclusions

NASA will launch in 2015 the GOES-R Lightning Mapper (GLM) that will provide continuous, full disc, high resolution total lightning (IC + CG) data. Previous efforts of lightning assimilation employed nudging scheme.

We proposed 2 new data assimilation methods using CAPE as a proxy between lightning data and model variables in conjunction with the WRFDA 3-DVAR system. By adjusting the temperature profiles we assimilate WTLN total lightning data for the 2011 Tuscaloosa, AL super storm for different model domains.

The 1-DVAR + 3-DVAR scheme performed better than the direct assimilation scheme in terms of both the number of assimilated observations and the accuracy of the $6h$ forecast temperature profiles.

The release of the parallel version of WRFDA will allow us to test both approaches, the direct lightning assimilation scheme and the 1-DVAR method, in a full 4-DVAR framework.

Acknowledgments

Professor I.M Navon and Dr. R. Stefanescu acknowledge support of NOOA award: $NA10NES4400008$.

References

[1] Alexander G. David, Weinman James A., Karyampudi V. Mohan, Olson William S., Lee A. C. L. (1999). The Effect of Assimilating Rain Rates Derived from Satellites and Lightning on Forecasts of the 1993 Superstorm. *Mon. Wea. Rev.,* 127 : 1433–1457.

[2] Barthe C., Deierling W., and Barth M.C., (2010). Estimation of total lightning from various storm parameters: A cloud-resolving model study, *J. Geophys. Res.*, 115, D24202, doi:10.1029/2010JD014405.

[3] Biagi C.J., Cummins K.L., Kehoe K.E., Krider E.P., (2007). National Lightning Detection Network (NLDN) performance in southern Arizona, Texas, and Oklahoma in 2003-2004. *Journal of Geophysical Research*, 112, Issue D5.

[4] Courtier P., Thepaut J.-N. and A. Hollingsworth, (1994). A strategy for operational implementation of 4DVAR using an incremental approach. Quart. J. Roy. Meteor. Soc., 120: 1367–1388.

[5] Deierling W., Petersen W.A., Latham J., Ellis S., Christian H.J. (2008). The relationship between lightning activity and ice fluxes in thunderstorms. *Journal of Geophysical Research* 113 : D15210, doi:10.1029/2007JD00970.

[6] Deierling W., Petersen W.A. (2008). Total lightning activity as an indicator of updraft characteristics. *Journal of Geophysical Research*, 113, D16210, doi:10.1029/2007JD009598.

[7] Deierling W., Latham J., Petersen W.A., Ellis S., Christian H.J. (2005). On the relationship of thunderstorm ice hydrometeor characteristics and total lightning measurements. *Atmospheric Research* 76 : 114 – 126.

[8] Fita L., Romero R., Luque A., and Ramis C. (2009). Effects of assimilating precipitation zones derived from satellite and lightning data on numerical simulations of tropical-like Mediterranean storms. *Adv. Geosci.,* 27 : 3297–3319.

[9] Lynn B., Yair Y. (2010). Prediction of lightning flash density with the WRF model. *Adv. Geosci.,* 23 : 11–16.

[10] Mansell E.R., Ziegler C.L., Macgorman D.R. (2007). A Lightning Data Assimilation Technique for Mesoscale Forecast Models. *Mon. Wea. Rev.,* 135 : 1732 – 1748.

[11] Mahfouf J.-F., Bauer P., and Marecal V., (2005). The assimilation of SSM/I and TMI rainfall rates in the ECMWF 4DVAR system. *Quart. J. Roy. Meteor. Soc.*, 131: 437–458.

[12] Marecal V., and Mahfouf J.-F. (2003).Experiments on 4DVAR assimilation of rainfall data using an incremental formulation. *Quart. J. Roy. Meteor. Soc.,* 129 : 3137–3160.

[13] Marecal V., and Mahfouf J.-F. (2002). Four-dimensional variational assimilation of total column water vapor in rainy areas. *Mon. Wea. Rev.,* 130 : 43–58.

[14] Navon I.M., Zou X., Derber J., and Sela J., (1992). Variational Data Assimilation with an Adiabatic Version of the NMC Spectral Model *Mon. Wea. Rev.,* 120 : 1433–1446.

[15] Orville, R.E., (2008). Development of the National Lightning Detection Network. *Bulletin of the American Meteorological Society*, 89 : 180–190.

[16] Papadopoulos A., Serpetzoglou E., Anagnostou E.N. (2009). Evaluating the impact of lightning data assimilation on mesoscale model simulations of a flash flood inducing storm.*Atmospheric Research,* 94 : 715 – 725.

[17] Papadopoulos A., Chronis T.G. and Anagnostou E. N., (2005). Improving Convective Precipitation Forecasting Through Assimilation of Regional Lightning Measurements in a Mesoscale Model, *Mon. Wea. Rev.,* 133 : 1961–1977.

[18] Pessi A.T., Businger S. (2009). The Impact of Lightning Data Assimilation on a Winter Storm Simulation over the North Pacific Ocean. *Mon. Wea. Rev.,* 137 : 3177–3195.

[19] Pessi A.T., Businger S. (2008). Relationships among Lightning, Precipitation, and Hydrometeor Characteristics over the North Pacific Ocean. *Journal of Applied Meteorology and Climatology,* 48 : 833 – 848.

[20] Rison W., Thomas R.J., Krehbiel P.R., Hamlin T., and Harlin J., (1999). A GPS-based three-dimensional lightning mapping system: Initial observations in central New Mexico. *Geophys. Res. Lett.,* 26 : 3573–357.

[21] Rodger C.J., Brundell J.B., Holzworth R.H., and Lay E.H., (2009). Growing Detection Efficiency of the World Wide Lightning Location Network, *Am. Inst. Phys. Conf. Proc.*, Coupling of thunderstorms and lightning discharges to near-Earth space: Proceedings of the Workshop, Corte (France), 23-27 June 2008, 1118, 15-20, DOI:10.1063/1.3137706.

[22] Shanno D.F., Phua K.H., (1980). Remark on algorithm 500. *ACM Trans. on Math. Software,* 6: 618–622.

[23] Thomas R.J., Krehbiel P.R., Rison W., Hunyady S.J., Winn W.P., Hamlin T., and Harlin J., (2004): Accuracy of the Lightning Mapping Array. *J. Geophys. Res.,* 109, D14207, doi:10.1029/2004JD004549