**Project Title:** R20 transition of the GOES-R GLM lightning assimilation capability in GSI for use in NCEP GDAS
**DTC Visitor:** Karina Apodaca (Colorado State University/CIRA and NOAA/AOML/Hurricane Research Division-Affiliate)
**DTC Host(s):** Ming Hu and Chunhua Zhou

## 1. Project overview

Great strides are being made to advance the operational weather forecasting enterprise and a vital objective is having the ability to accurately predict high-impact weather. However, this is a highly complex endeavor that requires a holistic approach to improve the multiple components involved in making numerical weather prediction successful. Two key ingredients that can help in this quest are the integration of new observations not previously used in operations and the enhancement of the data assimilation systems themselves to effectively process them. With that vision, the National Oceanic and Atmospheric Administration (NOAA) and the National Aeronautics and Space Administration (NASA) joined forces to create the most advanced fleet of geostationary weather satellites to date; *i.e.* the R-Series of Geostationary Operation Environmental Satellites (GOES, https://www.goes-r.gov/) to provide advanced imagery and measurements of Earth's weather and oceans, space weather, and solar activity. However, the success of this and future observing system missions ultimately depends on the capabilities of data assimilation systems to extract the maximum information from a given set of measurements.

In preparation for the Geostationary Lightning Mapper (GLM) aboard the now called GOES-16 and 17 satellites, several observation operators for the assimilation of lightning information have been developed at Colorado State University/Cooperative Institute for Research in the Atmosphere. Given that lightning activity is strongly tied to severe weather, which often originates in data-sparse regions, it is in these places where remotely sensed observations of lightning can be particularly useful. The GLM instrument is an optical transient detector, capable of seeing changes in luminosity in the optical scene of its field of view from which the frequency, location, and extent of lightning strikes can be estimated. Essentially, measurements of total lightning from the GLM instrument can be regarded as "pictures" or two-dimensional fields that indicate where and when lightning strikes occur (https://www.goes-r.gov/spacesegment/glm.html). Therefore, how can we extract the maximum information possible from these "pictures" of lightning activity for the benefit of operational assimilation and prediction?

To evaluate the benefit of the potential incorporation of the GLM instrument into the operational data stream at NCEP, we enhanced the NCEP operational Gridpoint Statistical Interpolation (GSI) data assimilation system (Parish and Derber, 1992, Kleist et al. 2009) by adding a new forward operator for lightning flash rate capability, as described in Apodaca et al, 2014, and by following a variational data assimilation framework. This forward operator is suitable for the current operational Global Data Assimilation System (GDAS) and accounts for the coarse resolution and simplified cloud microphysics of the Global Forecasting System, in which convection cannot be resolved explicitly; therefore, it is possible to evaluate the impact of lightning observations on the large-scale environment around and prior to storm initiation.

Moreover, in preparation for NOAA's next generation forecast systems, which are based on the finite-volume cubed-sphere dynamical core and operate at convection allowing model (CAM) resolutions, this new lightning assimilation capability was augmented by incorporating a second observation operator for lightning flash rate that extends from the first one. This particular observation operator has already been incorporated in the CSU/Maximum Likelihood Ensemble Filter (Zupanski, 2005) and it has been tested with the WRF-ARW model for tornado and tropical cyclone applications. In addition to updating typical control variables used in assimilation (e.g. temperature, humidity, and wind) it can also impact cloud hydrometeor fields; therefore, improving storm representation and forecasting of lightning activity in the WRF-ARW model (Apodaca and Zupanski, 2018).

In this technical report, an overview of the development within the GSI system is provided. For a more detailed explanation on the theoretical aspects associated with the adaptation of the nonlinear observation operator, as well as linearization, and development of the adjoint model for lightning flash rate, please consult Apodaca and Zupanski (2018).

## 2. Technical aspects of the implementation in GSI

As a starting point, we used surface-based Lightning Detection Network (LDN) data from the World Wide Lightning Location Network (WWLLN) as a GLM-proxy. From these data, the real-earth latitude and longitude and timing of total lightning strikes were extracted in a way similar to what the GLM instrument detects. The processing of observations begins by reading information associated with lightning strikes (real-Earth latitude and longitude and time of occurrence). As required by the GSI system, these data are pre-processed and converted offline into the BUFR format. After reading the BUFR data in GSI, subroutines that transform the lightning strikes into lightning flash rate and perform a time correction are applied. These processes are included in the newly developed "read_goesglm.f90" program.

The calculation of the nonlinear lightning observation operator is performed in a new program named "setuplight.f90". This program eventually computes the tangent linear coefficients at observation locations, as required by GSI. Due to a somewhat more complex definition of the nonlinear observation operator, the interpolation to observation location requires the guess from surrounding grid points, as depicted in Fig.1. Five quadrants are selected for bi-linear interpolation and for finite difference derivation (e.g. central, north, south, east, and west). For instance, in the calculation of the tangent linear of horizontal advection at point i1, geopotential height is calculated using the difference in values at points i3 and i9 in the east-west direction, and i2 and i5 in the south-north direction, similarly for points i2 – i4.

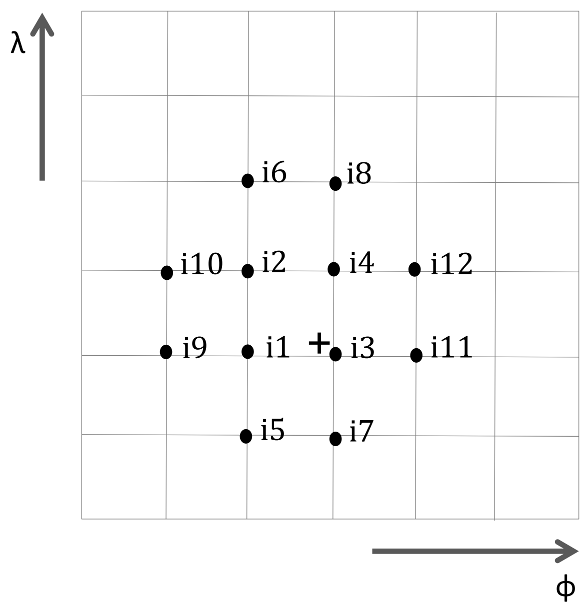


*Figure* 1. Guess grid used for finite difference approximation derivation and for bi-linear interpolation. i1 to i12 are the model grid points surrounding an observation **+.** ϕ and λ are model latitudes and longitudes, respectively.

The tangent linear and adjoint calculations of the lightning flash rate observation operator are done in the new program “intlight.f90.” One additional component of this new lightning assimilation capability is the inclusion of an online bias correction scheme based on optimal parameter estimation. The goal is to correct any skewness on the probability density function statistics of the normalized vectors of departures at observation points (innovations). Skewness can indicate that observed values are considerably larger than a first guess. This online bias correction scheme is described in detail in Apodaca et al. (2014) and it is performed inside “setuplight.f90.” A conjugate gradient-based minimization algorithm (steepest descent) is included in “stplight.f90.” This algorithm calculates the gradient and subsequent correction of all control variables (**x** = *t*, *u*, *v, q) and* (**x** = *t*, *u*, *v, q, qi, qs, qg*), for global and CAM applications, respectively.

The following is a list of all the necessary source code modifications for the incorporation of this lightning assimilation capability in GSI. Bold letters denote newly-developed programs.

1. gsimod.F90
2. gsisub.F90
3. **lightinfo.f90**
4. obs_sensitivity.f90
5. read_obs.F90
6. **m_lightNode.F90**
7. m_obsdiags.F90
8. m_obsNodeTypeManager.F90

9. obsmod.F90
10. stpjo.f90
11. tintrp2a.f90
12. m_obsHeadBundle.F90
13. **intlight.f90**
14. **read_goesglm.f90**
15. setuprhsall.f90
16. **statslight.f90**
17. **setuplight.f90**
18. **stplight.f90**
19. **sumslightbias.f90**
20. Makefile
21. Makefile.dependency
22. Makefile.src

## Additional settings

A new text file "global_lightinfo.txt" containing information associated with lightning observations is included in the /fix directory and this new file needs to be added to a GSI namelist. It is also important to verify that the number of levels in the model-specific "anavinfo" files match the number of model levels from the background. The following line needs to be added to the GSI namelist to allow processing of GOES-16/GLM observations and similarly for GOES-17/GLM:

OBS_INPUT::

```
! dfile      dtype      dplat    dsis      dval    dthin    dsfcalc
  glmbufr    goes_glm   g16      goes_glm  0.0     0        0
```

It is also important to note that in order for the lightning assimilation to function properly with a global forecast system, the “cw” variable, inside fix/global_anavinfo.l64.txt, must be activated. Given that this capability was initially tested by reading background fields from a non-hydrostatic, cloud-resolving model that includes a six-class hydrometeor microphysics option (e.g. WRF-ARW), this lightning assimilation capability can also be tested in “regional” mode. If this is desired, a regional flag in the GSI namelist must be set to true. Additionally, the “ql, qi, qr, qs, qg” variables have to be activated and the “qnr” variable has to be added to the anavinfo_arw_netcdf file located in the /fix directory. Furthermore, with upcoming modifications to this lightning assimilation capability, testing with other systems that are based on a non-hydrostatic mesoscale model dynamical core (e.g. WRF-NMM, HWRF) will be possible.

## 3. Testing and evaluation: Evaluation of the GOES/GLM-proxy lightning assimilation

### 3.1 Processing of lightning flash rate observation

After post processing, the spatial distribution of the lightning strikes is shown in Fig. 2(a) is compared to the lightning flash rate density (# hits/ $km^2$/6-hours) in Fig 2(b). We found that the location of the observations read in by the GSI system (Fig. 2b), coincide with the location of the raw observations shown in Fig. 2(a).

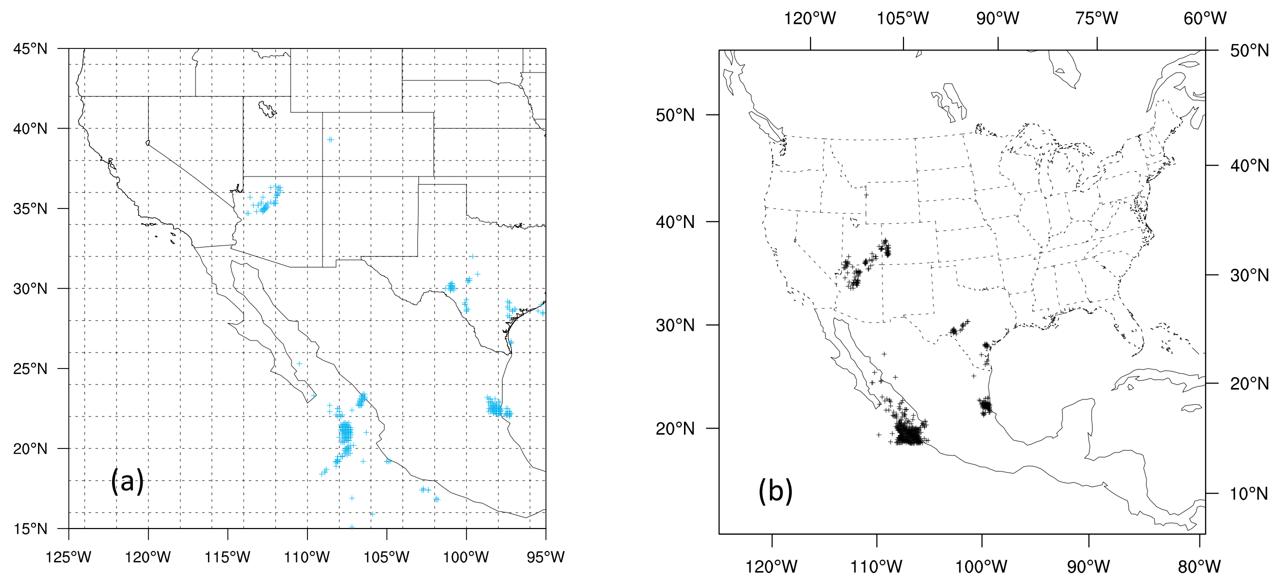


*Figure* 2. Verification of gridded lightning flash rate observations, valid at 1200 UTC 27 August 2013 for (a) Raw WWLLN lightning observations and (b) gridded lightning flash rate used in assimilation (Source: Apodaca and Zupanski, 2018).

### 3.2 Impacts to the analysis

Two experiments were conducted with model background fields, one with the assimilation of lightning (light) and another without (control). A panel of analysis increments, or a difference between both experiments, for the selected set of control variables: temperature, the u and v components of the wind, and specific humidity is shown in Fig. 3. Regions of adjustments to the control variables (**x** = *t*, *u*, *v, q*) by the assimilation of lightning observations are evident and these regions coincide with the areas of a high density of lightning observations shown in Fig. 2(b). The most pronounced increments can be seen over the Nevada-Arizona region and over South-central Texas, in the United States.

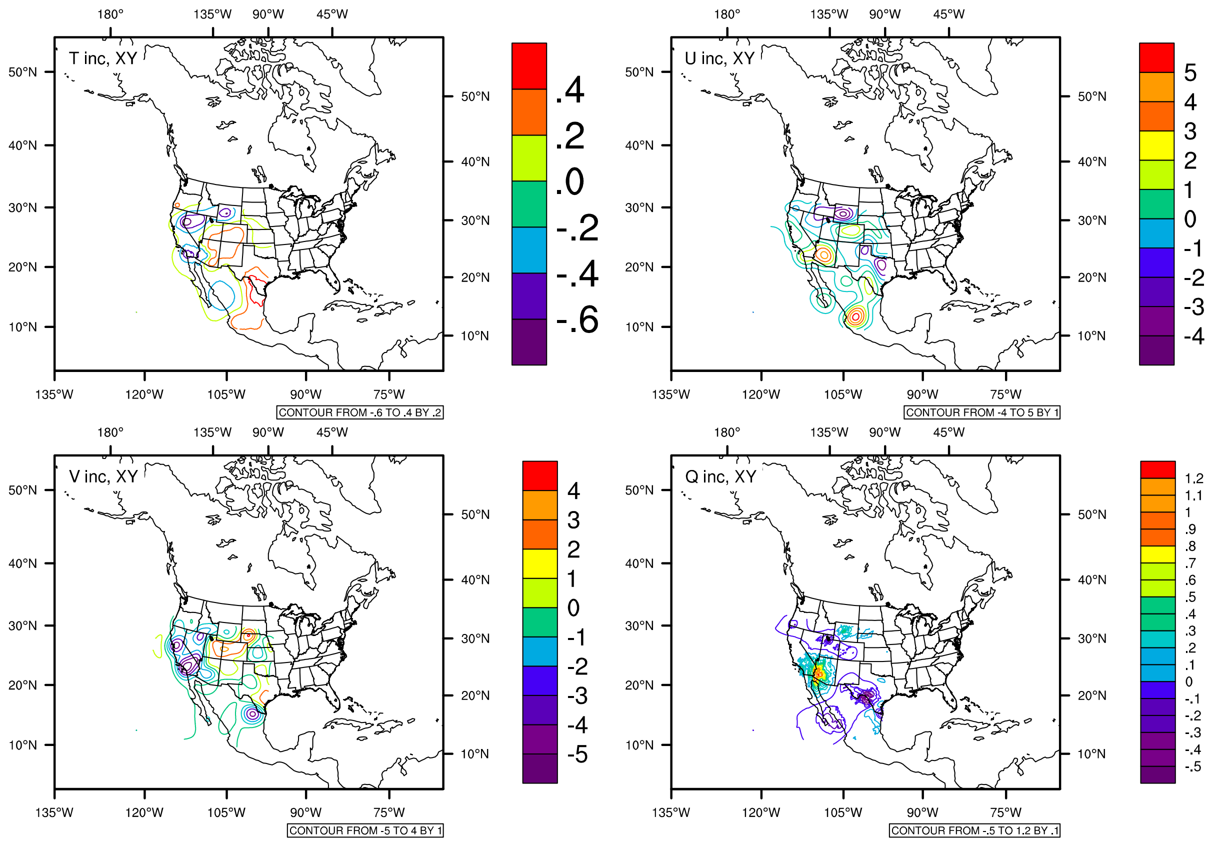


*Figure* 3. Analysis increments of temperature, humidity and wind, between the control and the light experiments, valid at 1200 UTC 27 August 2013 between the control experiment and the lightning data assimilation experiment (Source: Apodaca and Zupanski, 2018).

For verification, observed 24-hour accumulated surface precipitation (Fig. 4) from the National Weather Service, also valid at 1200 UTC 27 August 2013 was used. This graph shows areas of flash flooding produced by the remnants of tropical storm Ivo in the Southwestern United States. The regions of adjustments to the control variables (Fig. 3) and areas with a high density of assimilated lightning observations (Fig. 2(b)), coincide with the location of high precipitation contours shown in Fig. 4, specifically for Nevada, Arizona, Utah, and Southern Texas. It would be worthwhile to investigate if lightning assimilation can have an impact on the forecast of surface precipitation from the GFS model. This work; however, could be performed in subsequent testing and verification of the lightning data assimilation capability in the GDAS.

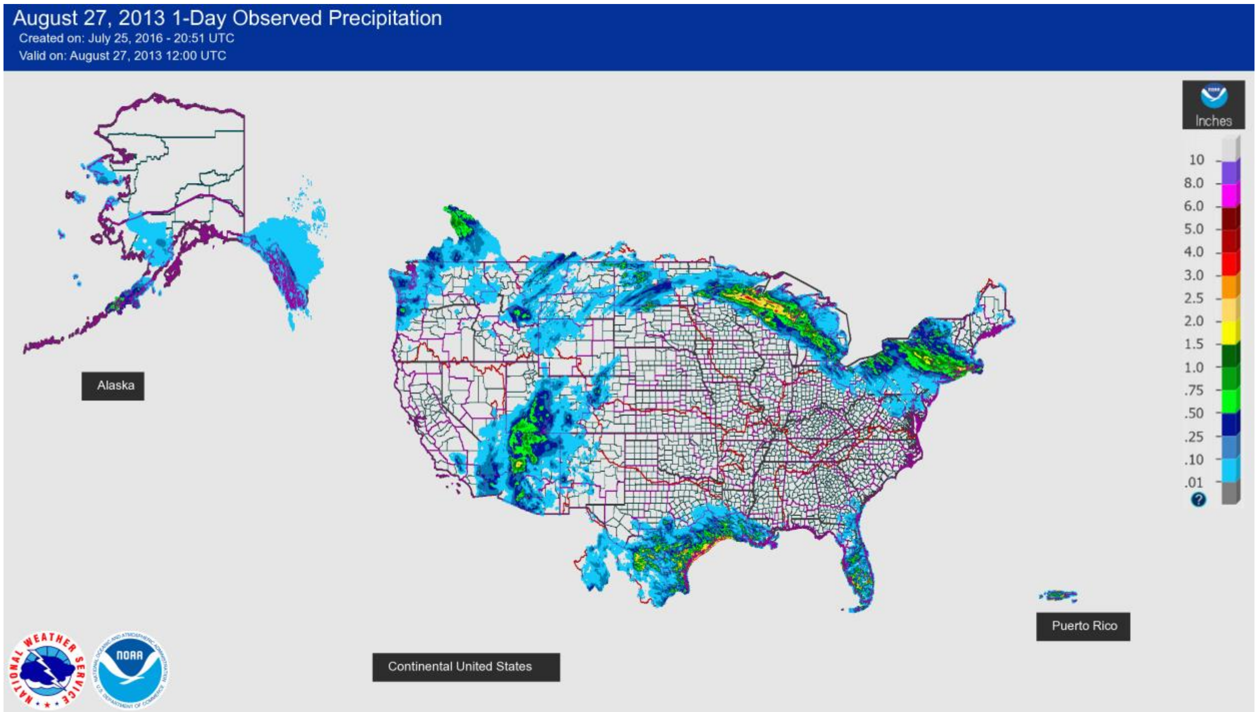


*Figure* 4. 24-hr accumulated precipitation, valid at 1200 UTC 27 August 2013. Note the region of maximum precipitation near the Arizona-Nevada border, which coincided with the region of positive analysis increment in specific humidity shown in Fig. 4. The assimilation of lightning observations has a positive impact in the initial conditions (Source: Apodaca and Zupanski, 2018).

## 4. Summary and Future work

An overview of the development efforts to successfully incorporate this variational lightning flash rate observation operator suitable for the GOES/GLM instrument in GSI was provided. The implementation of this lighting assimilation capability in GSI represents a baseline for the assimilation of lightning information following advanced data assimilation methods, but further work is planned. In preparation for “Next-Gen” operational prediction systems and data assimilation systems, a second observation operator for lightning suitable for non-hydrostatic, cloud-resolving models that permits the inclusion of precipitating and non-precipitating hydrometeors as analysis cloud control variables was incorporated in the GSI system and testing is planned. Additionally, verification of the impacts to the forecast with a global forecast system from NCEP is expected via testing in global parallel cycling experiments. GLM observations are being prepared for conversion into the BUFR format and their true impact in operational assimilation and numerical weather prediction can potentially be evaluated in the near term.

## Acknowledgements

Financial support for this project was provided by the 2017 Developmental Testbed Center Visitor Program Award. I would like to recognize the valuable support from my DTC hosts, administrative staff, and the DTC leadership. I also wish to thank the World Wide Lightning Location Network (http:// wwlln.net), a collaboration among over 50 universities and institutions, for providing the lightning location data used in this project.

## References

Apodaca, K., and M. Zupanski, 2018: Variational and hybrid (EnVar) methodologies to add the capability to assimilate GOES-16/GLM observations into GDAS. *Joint Center for Satellite Data Assimilation Quarterly*, No. 58, Winter 2018, p. 12-20. doi: 0.7289/V5CJ8BR2

Apodaca, K., Zupanski, M., DeMaria, M., Knaff, J. A., and Grasso, L. D.: Development of a hybrid variational-ensemble data assimilation technique for observed lightning tested in a mesoscale model, Nonlin. Processes Geophys., 21, 1027-1041, doi:10.5194/npg-21-1027-2014, 2014

Zupanski, M.: Maximum likelihood ensemble filter: theoretical aspects, Mon. Weather Rev., 133, 1710–1726, 2005.