NOAA ROSES Semi-Annual Report

Reporting Period: September 2020 – February 2021 (1st report)

PI: Kyle Hilburn
Co-PI(s): Yoonjin Lee, Milija Zupanski

Project Title: Assimilating GOES-R Latent Heating in FV3 using Machine Learning

Executive Summary (1 paragraph max)

The main objective of this project is to use the high-resolution information from GOES-R Advanced Baseline Imager (ABI) and Geostationary Lightning Mapper (GLM) to improve short-term forecasts of high-impact weather hazards. This will be accomplished through using machine learning (ML) to derive three-dimensional fields of latent heating to spin-up convection in the Rapid Refresh Forecast System. The secondary goal of this project is to provide new data assimilation capabilities for the new generation of FV3 dynamical core models at NOAA, utilizing the Joint Effort for Data Assimilation Integration (JEDI) framework.

Progress toward FY20 Milestones and Relevant Findings (with any Figs)

- The first major accomplishment was production of the *CONUS3* dataset, which brings together GOES-16 ABI, GLM, MRMS, and HRRR data resampled onto the HRRR mass grid. This dataset includes all times-of-year and covers all of CONUS. The dataset provides over 63K images (35x more than *CONUS2*, which was used in *Hilburn et al.* 2021) and the images cover 29x greater area than *CONUS2*.

- The second major accomplishment was integration of the prototype system (Figure 1) for producing 3D latent heating fields, which is now running in near real-time. An example of *GREMLIN* performance for a variety of convective modes, locations, and times-of-day is shown in Figure 2. *GREMLIN* data for a one-week retro period has been sent to GSL and we are working with them on a retrospective simulation test.

- The third major accomplishment is that we have begun model architecture and loss function experiments for *GREMLIN V2*, training deeper models to better resolve synoptic scale precipitation, and adding convective/stratiform classification to *GREMLIN*. We have also been assessing the temporal consistency of *GREMLIN* estimates relative to MRMS, which is driving loss function and architecture decisions.

- Kyle and Yoonjin presented their latest work at a virtual meeting with our NOAA ESRL GSL colleagues on 20-January-2021. This meeting was attended by Curtis Alexander, Steve Weygandt, Amanda Back, Dan Lindsey, Chris Kummerow, Yoonjin Lee, Steve Miller, Anton Kliewer, Milija Zupanski, and Kyle Hilburn. We developed a plan to run HRRR retrospective simulations in Spring 2021.

- Kyle presented results on *GREMLIN* development at the CIRA ML Core Meeting on 16-December-2020. This included examination of alternative architectures, their training properties, and their receptive fields.

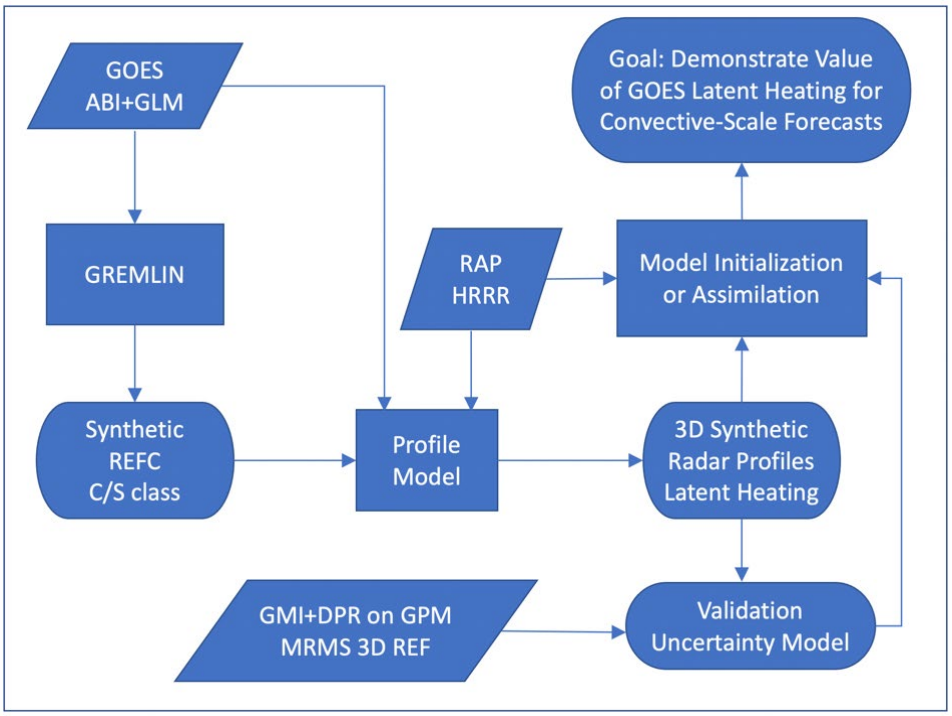

Figure 1. Flowchart showing the data flow and model components in the LH ML system.

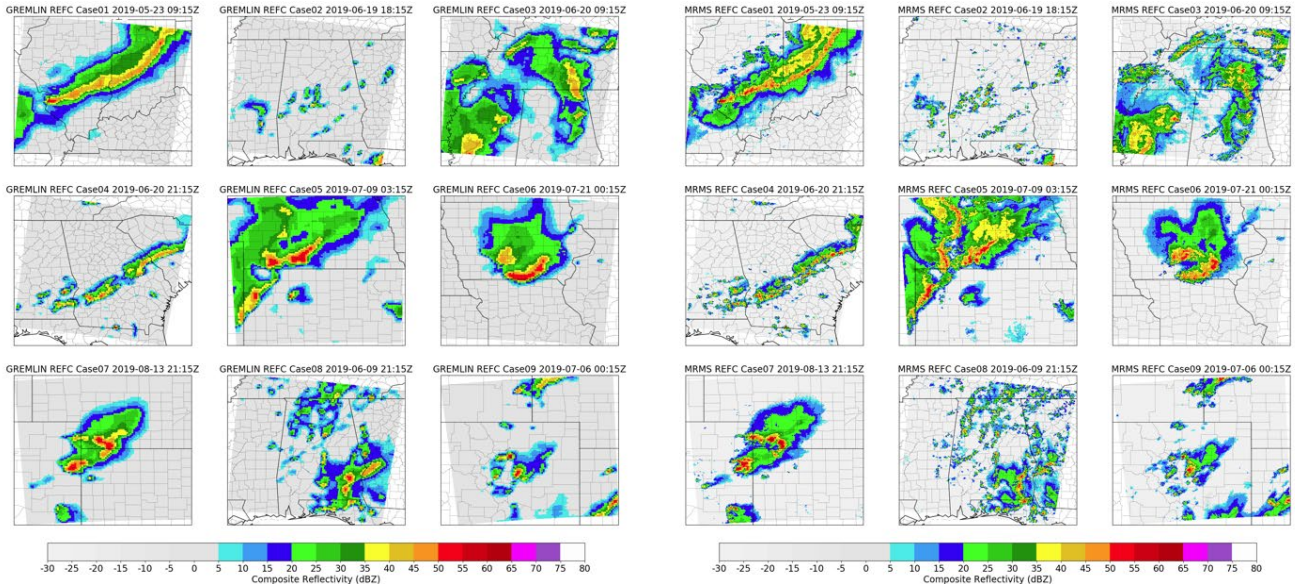

Figure 2. Example of GREMLIN ML predictions (left nine panels) versus MRMS composite reflectivity (right nine panels).

Plans for Next Reporting Period

The three main goals for the next reporting period are: 1) conduct and evaluate results from GOES LH assimilation in HRRR retrospective simulations for warm season convection over CONUS, 2) improving and extending GREMLIN model capabilities using multi-task learning and advanced loss functions, to get the maximum benefit from the additional training data, and 3) extending the model to run on the portion of the G16 and G17 Full Disk domains that have GLM coverage.