Inferring Airmass Properties from GOES-R ABI Observations

Kyle Hilburn[1]

[1]Cooperative Institute for Research in the Atmosphere/Colorado State University

November 23, 2022

Abstract

Radiosonde observations are the gold-standard for quantifying vertical profiles of atmospheric state variables. Knowledge of which is critical for quantifying moisture and instability, two main ingredients for severe weather. Unfortunately, radiosondes are very sparse, averaging just one observation per 500 x 500 km area over CONUS, and most locations have only two observations per day. This creates uncertainty in the representation of short wavelength and rapidly evolving synoptic and mesoscale features in numerical weather prediction (NWP) and provides few points of comparison for human forecasters to interpret NWP in making forecasts. To fill this gap in our knowledge of the atmospheric state, human forecasters make use of satellite imagery to estimate airmass properties for incrementing NWP outputs. Data from geostationary satellites have been especially useful because of its high temporal resolution (5-minutes) and high spatial resolution (2 km). While the Advanced Baseline Imager (ABI) was not designed as a sounding sensor, the three water vapor bands and three infrared window bands do provide some sounding capabilities. Satellite data are particularly useful in assessing position and timing errors, the representation of short waves, and humidity. The key question addressed by this work is can the mental process used by human forecasters be translated into a machine learning (ML) algorithm to provide automated and objective estimates of airmass properties from ABI? Experiments with convolutional neural networks (CNNs) show that ML can indeed be used. Related research efforts, such as NOAA Unique Combined Atmospheric Processing System (NUCAPS) has explored use of dense neural networks (DNNs), which are essentially replacing a radiative transfer model with a ML model. However, we find more skill can be achieved by making use of the spatial information captured with CNNs. This more closely mimics the human imagery interpretation process: it is the spatial patterns in the features (as much as the pixel-wise values themselves) that carry the useful information content. We will present our latest results, focusing especially on relative humidity, compare against radiosondes, and discuss whether skill is enough to potentially make a positive impact on NWP analyses.

Inferring Airmass Properties from GOES-R ABI Observations

Kyle Hilburn
Cooperative Institute for Research in the Atmosphere

Introduction

Statement of the Problem

- Vertical profiles of moisture and instability in the troposphere are two main ingredients for severe weather.
- Radiosondes are the gold-standard for measuring vertical profiles.
- Radiosondes are very sparse, averaging just one observation per 500x500 km^2 area over CONUS.
- Moreover, most locations only have two radiosonde observations per day.
- This creates uncertainty for the representation of vertical profiles in numerical weather prediction (NWP) models in particular at short wavelengths and for rapidly evolving features.

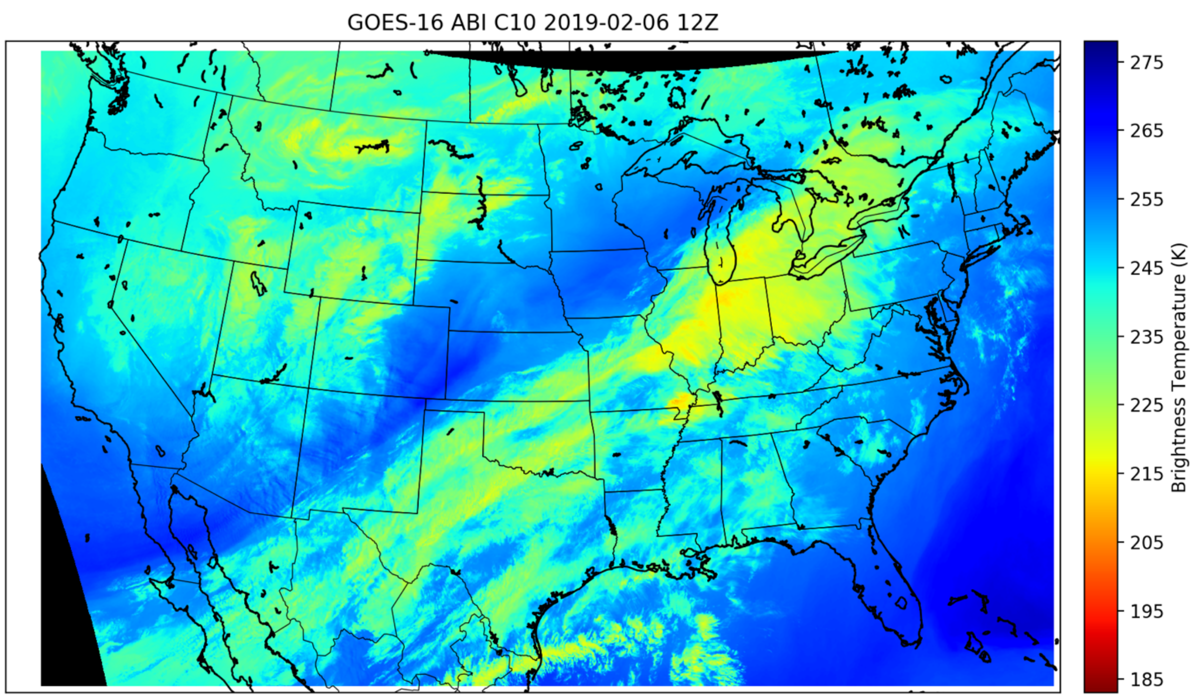
Identification of the Opportunity

- To fill in this gap in our knowledge of the atmospheric state, for decades human forecasters have made use of satellite imagery to estimate airmass properties for incrementing NWP outputs, performing these adjustments mentally.
- Data from the GOES-R Series of geostationary sensors provides very high temporal resolution (5-minutes CONUS-wide) and very high spatial resolution (0.5-2.0 km).
- The Advanced Baseline Imager (ABI) is not a sounder, but multiple water vapor and window bands provide some sounding capability.
- The key question addressed by this work is, can the mental process used by human forecasters be translated into a machine learning (ML) algorithm to provide automated and objective estimates of airmass properties from ABI?

Data and Methodology

GOES-R Data

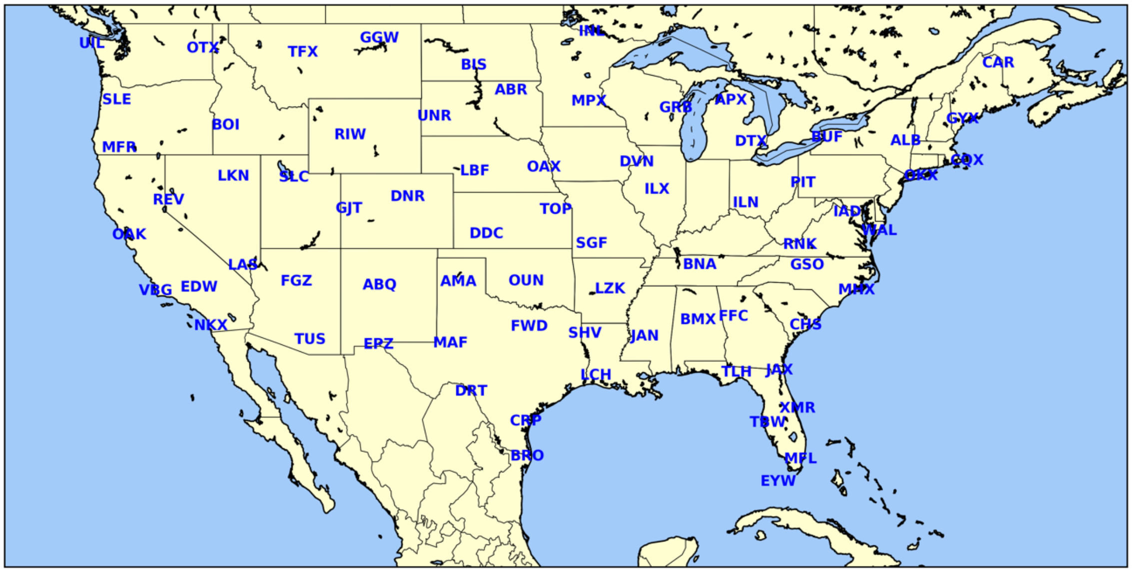
- We are using GOES-16 ABI CONUS sector, focusing on bands:
    - C08 (6.2 µm) upper-level ~334 mb, water vapor band
    - C09 (6.9 µm) mid-level ~442 mb, water vapor band
    - C10 (7.3 µm) lower-level ~618 mb, water vapor band
    - C12 (9.6 µm) ozone band
    - C13 (10.3 µm) clean longwave infrared window
- These bands have a nominal 2x2 km resolution, and we have resampled to the HRRR-CONUS 3x3 km Lambert Conformal Conic grid.
- Figure below shows an example of C10.
- There are slivers of missing data over the northern and southwestern parts of the domain, which represents 2.15% of the grid (black fill below).



Radiosonde Data

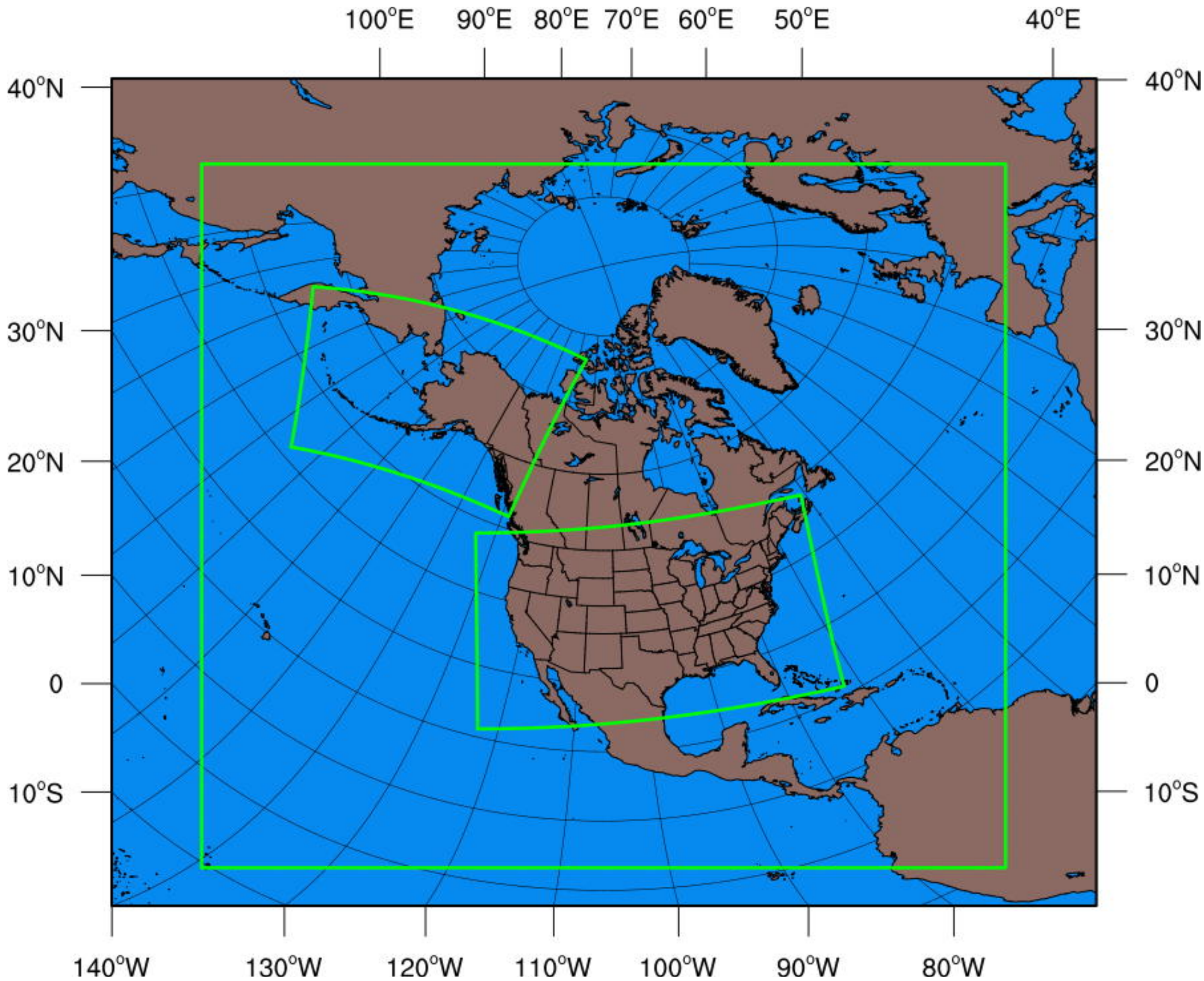
- Using the Integrated Global Radiosonde Archive (IGRA) Version 2.
- Figure below shows 72 stations over CONUS.
- Using the fixed location of launch stations.
- Focusing on levels: 300, 400, and 500 mb.

- Examined 600 mb also, but with 30% fewer observations (using 10 mb threshold) statistics for both NWP and ML vs radiosonde were questionable.
- Validation dataset is 2019, every 12 hours, 52,000 samples.
- Only 1.8% of radiosonde profiles are not at 00Z or 12Z times.



NWP Data

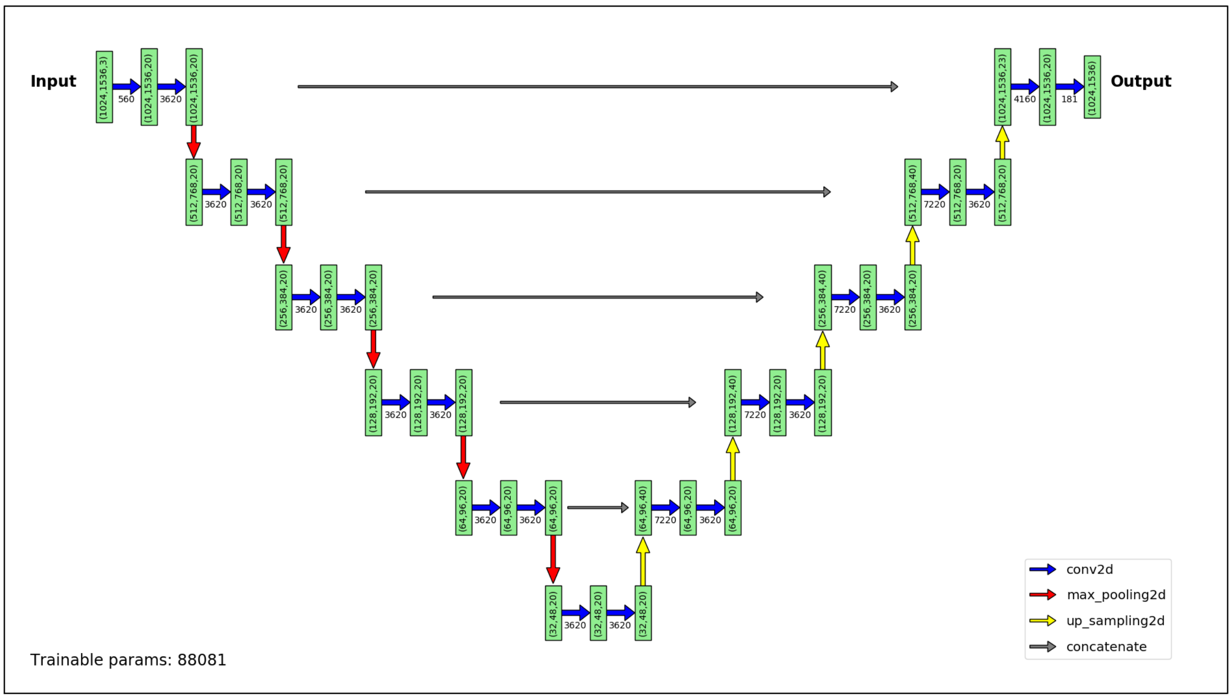
- Radiosondes are too sparse spatially to directly train a convolutional neural network.
- Instead, train a convolutional neural network to map from GOES to NWP and evaluate against radiosondes.
- HRRR CONUS domain shown in figure below.
- Using HRRR F01 fields for results shown here, but analysis with HRRR F00 and GFS Analysis produces very similar results.
- Subset the HRRR grid (1799 x 1059) to 1536 x 1024 to avoid odd parity issues with pooling / upsampling.
- Translational invariance of convolutional neural networks makes adapting the trained model back to the full HRRR grid a trivial operation.
- Training dataset is 2018, every 3 hours, 2920 images.
- Testing dataset is 2019, every 12 hours, 730 images.
- Note that HRRR assimilates GOES-R water vapor bands in clear sky conditions.



Machine Learning Methodology

- Spatial context is important:
  - Fully connected (dense) neural networks that operate on a pixel-wise level had at best R^2 of 0.5 (e.g., 500 mb RH).
  - Convolutional neural networks (CNNs) that can utilize spatial patterns and spatial context are able to achieve R^2 of 0.7 or greater.
- Using a convolutional neural network architecture based on a U-Net.
- Preliminary model has 3 encoding/decoding blocks, C-P blocks, 32 filters/layer, ReLU activation.
- Final architecture has 5 encoding/decoding blocks, C-C-P blocks, 20 filters/layer, ReLU activation, shown in figure below.
- In the figure, the array sizes are given in the green boxes and the number of parameters is given below the blue arrows.
- Going any deeper or wider produced overfitting.

- Model has a total of 88,081 trainable parameters.
- Loss function is the mean-squared-error.
- Evaluate with metric: coefficient of determination (R^2).
- Implemented in Google TensorFlow.
- On one NVIDIA Telsa P100 GPU it took 9 hours for 100 epochs of training.
- The model required 5 GB of memory during training with batch size of 5.
- The dataset required 90 GB memory.
- The saved model in HDF5 is 1.2 MB.



Results and Discussion

ABI Band Selection

- Considered three sets of inputs,
  - GOES heritage water vapor band: C09
  - GOES-R water vapor bands: C08, C09, C10
  - Airmass RGB: C08-C10, C12-C13, C08
- Table below shows the coefficient of determination (R^2) for 500 mb level
  - Shows R^2 for CNN estimate vs HRRR test dataset
  - Using preliminary CNN architecture
  - For three fields:

- Geopotential height (HGT)
- Temperature (TMP)
- Relative humidity (RH)

- The three GOES-R water vapor bands provide much better skill than C09-alone or the Airmass RGB.
- The lower skill for the Airmass RGB is surprising since it contains two of the water vapor bands.

CNN vs HRRR R^2	C09	C08, C09, C10	Airmass RGB
HGT	0.37	0.64	0.35
TMP	0.40	0.65	0.39
RH	0.48	0.69	0.39

ABI Information Content

- Tables below compare coefficient of determination (R^2) for HRRR versus radiosondes and for the final CNN model versus radiosondes.
- Results are not conditioned on cloud/no-cloud, but are for all scenes.
- After training the final CNN architecture, applied linear regression on the training dataset to remove CNN biases relative to radiosondes for 2018 training period; validation is on the 2019 period.
- HRRR does tremendously well for HGT and TMP, and CNN does not achieve that level of skill.
- However, the CNN has better skill for RH at 300 and 400 mb (see tables below), suggesting that GOES can provide value to NWP for RH.

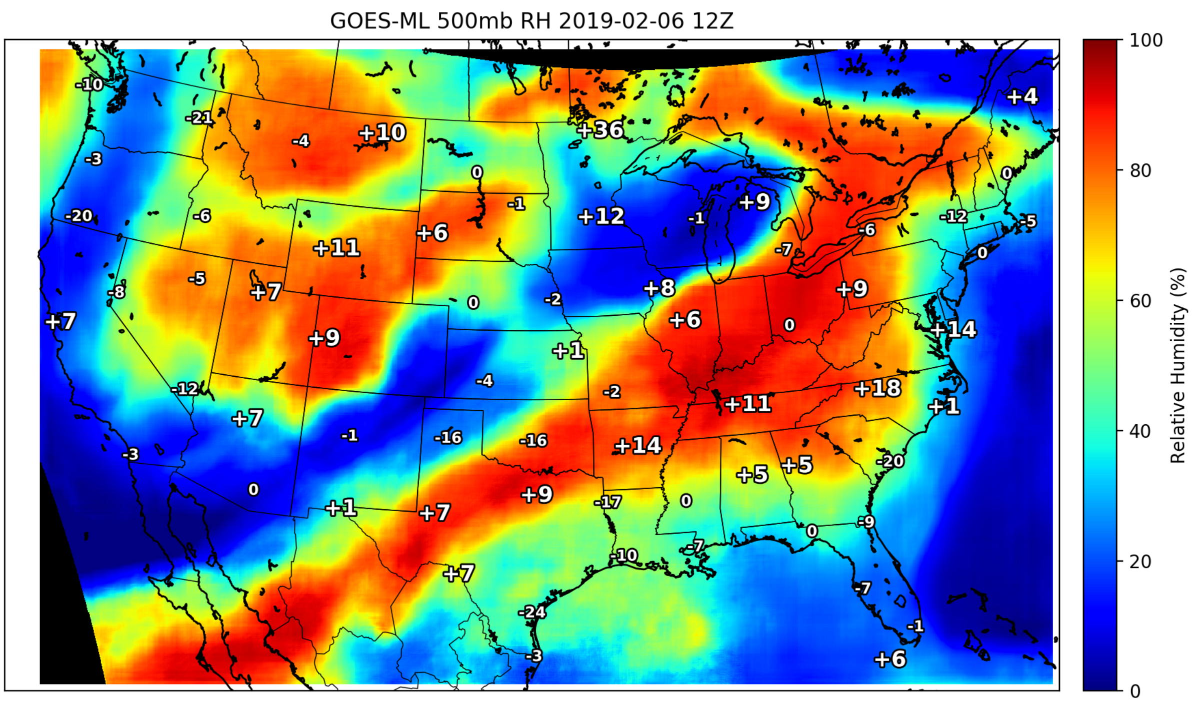
HRRR vs IGRA R^2	HGT	TMP	RH
300 mb	0.99	0.98	0.51
400 mb	0.99	0.99	0.66
500 mb	0.99	0.98	0.73

CNN vs IGRA R^2	HGT	TMP	RH
300 mb	0.88	0.82	0.66
400 mb	0.89	0.83	0.70
500 mb	0.87	0.86	0.68

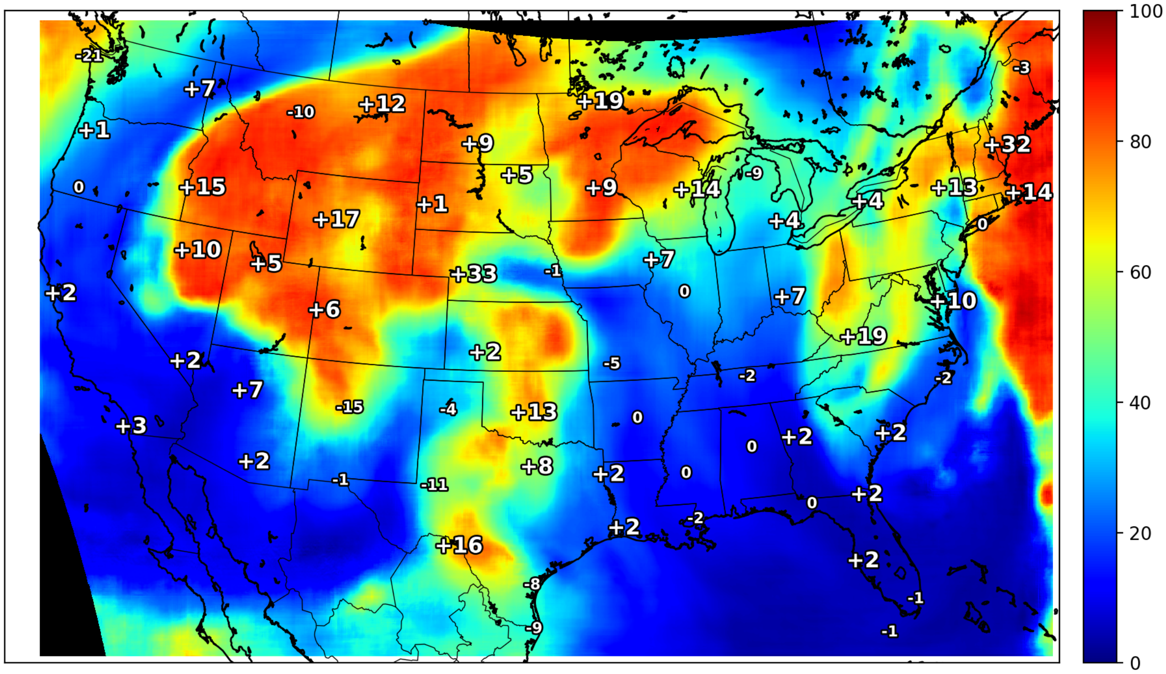
Examples of RH using GOES-ML

- Four figures below show examples of 500 mb RH derived from GOES using ML at different times-of-year (Feb, Apr, Aug, Dec).
- Numbers on figures give the percentage point improvement of GOES-ML over HRRR relative to radiosondes:
    - D = |HRRR-IGRA| - |GOES-IGRA|
    - Where HRRR, IGRA, GOES are relative humidities in %.
    - Positive values indicate locations where GOES has better skill than HRRR.

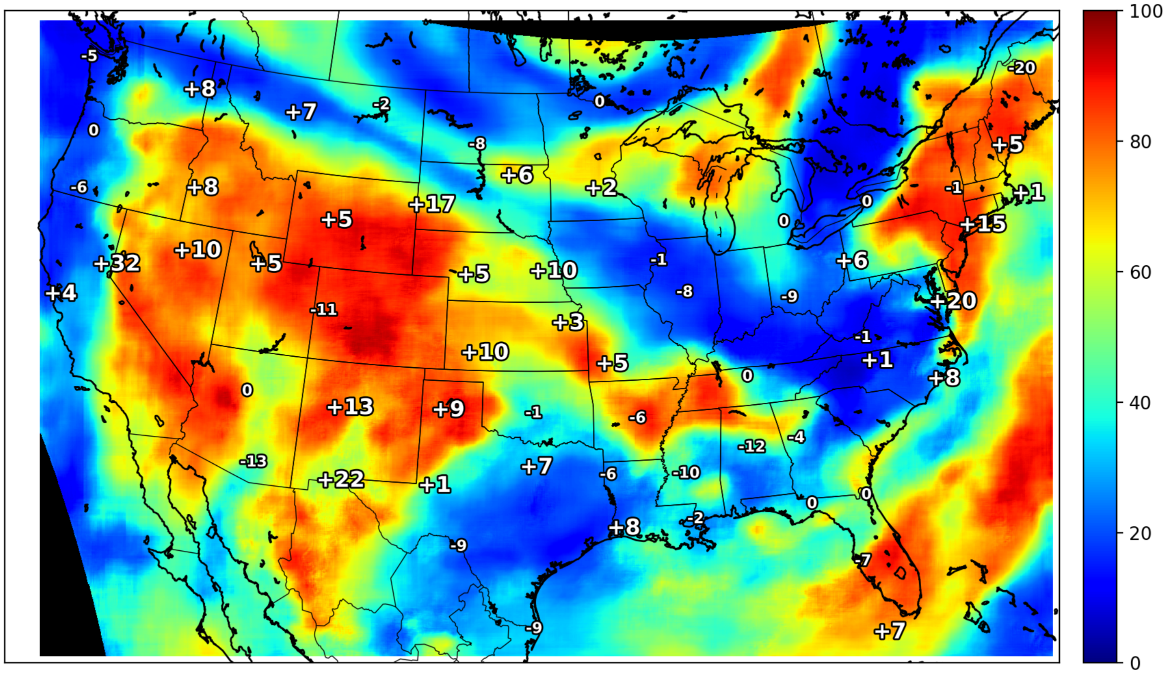
- Improvements are largest or most numerous along gradients or meso-scale features, indicating that GOES provides information that can be used to correct for position/timing errors in NWP.
- It appears the improvements are largest where clouds are present. In that situation, the band weighting functions will peak at cloud top, which is consistent with the good skill at 300-400 mb.

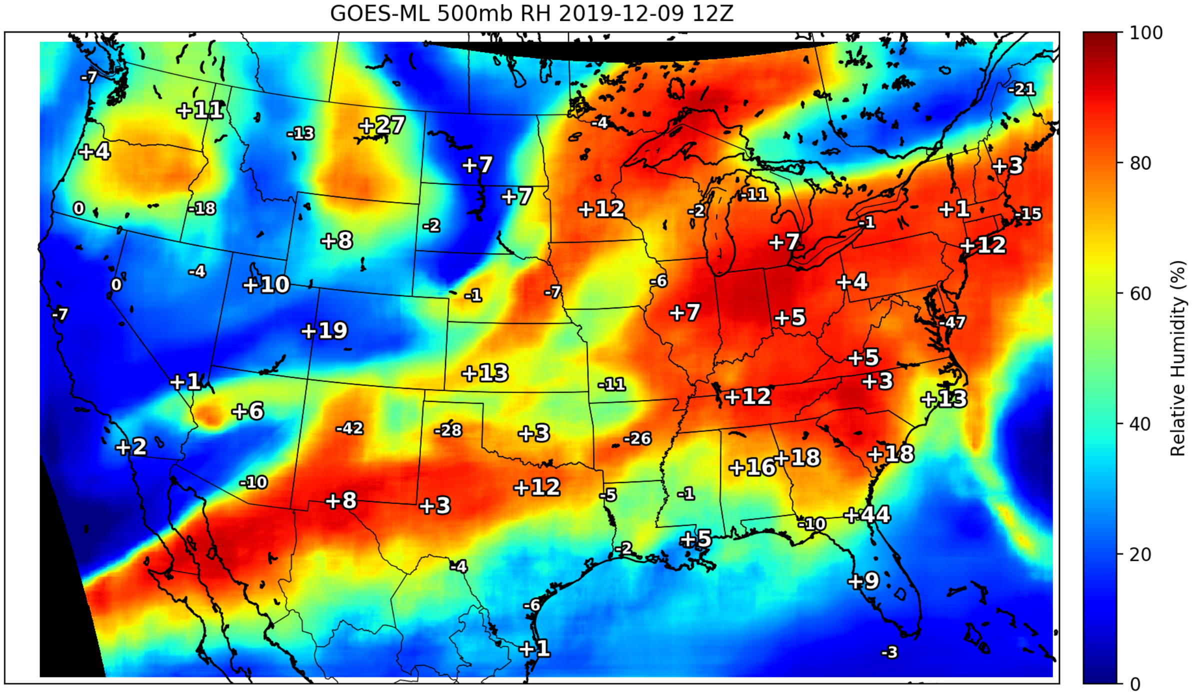


GOES-ML 500mb RH 2019-04-21 12Z



GOES-ML 500mb RH 2019-08-08 00Z





- Combining GOES-ML estimates with HRRR using a fully-connected (dense) network provides the best skill with R^2 of 0.8 or more.
- While that dense network is trained at radiosonde locations, it is applied to all locations with CNN estimates.
- Additional analysis of GOES-ML error characteristics is ongoing.

Summary and Conclusions

Key Findings

- Machine learning is extremely powerful, but not magic; the underlying data must have information content to provide value.
- The three water vapor bands on GOES-R series ABI provide significant additional skill for deducing airmass properties compared to the previous GOES generation.
- GOES water vapor bands provide information content on upper-tropospheric relative humidity that beats NWP.
- Use of a convolutional architecture is essential for extracting this skill from GOES.
- It appears that GOES information content for HGT and TMP is not large enough to improve upon NWP estimates.

- Combining GOES-ML with NWP yields the best estimates of upper-tropospheric RH.
- ML offers a way to make use of GOES-R radiances in cloudy and precipitating scenes.

Acknowledgements

- This work is supported by NOAA GOES-R Program under grant NA19OAR4320073.
- Thank you to NOAA RDHPCS for access to the Fine Grain Architecture System on Hera.
- Thank you to Imme Ebert-Uphoff (CSU) for our discussions that helped shape this research.