USING CNNs TO PRODUCE QUANTITATIVE PRECIPITATION ESTIMATES

Micheal Simpson, Lin Tang
University of Oklahoma; Cooperative Institute for Mesoscale Meteorological Studies; National Severe Storms Laboratory, USA

Jian Zhang, Kenneth Howard
National Severe Storms Laboratory, USA

Pao-Liang Chang, Yu-Shuang Tang, Pin-Fang Lin
Central Weather Bureau, Taiwan

Background

Precipitation is an important meteorological variable which is necessary to determine water-related hazards including flash-flooding. While improvements have been made since the advent of utilizing radar for quantitative precipitation estimates (QPEs), there are still many assumptions and small-scale phenomena that are not explicitly incorporated into QPE calculations.

With the recent improvements in computer capabilities and machine-learning technology, leveraging deep learning techniques (such as convolutional neural networks, or CNNs) are a feasible and ideal option to sift through very large datasets that have been gathered by the Taiwan Central Weather Bureau and the Multi-Radar Multi-Sensor (MRMS) groups.

This work will present the capabilities of CNNs to learn on single-radar data as well as mosaicked products to produce accurate QPEs.

Machine Learning Model Outline

The following is the general outline of the deep-learning models implemented for both the single-radar Taiwan data as well as the ConUS mosaicked data.

Input Layer
Layers.Conv2D(64), activation = 'linear'
Layers.LeakyReLU(alpha = 0.4)
Layers.Conv2D(128), activation = 'linear'
Layers.LeakyReLU(alpha=0.4)
Layers.MaxPooling2D((2,2))

Layers.Flatten

Layers.Dense(512)
Layers.GaussianNoise(0.1)
Layers.LeakyReLU(alpha = 0.4)
Layers.Dropout(0.5)

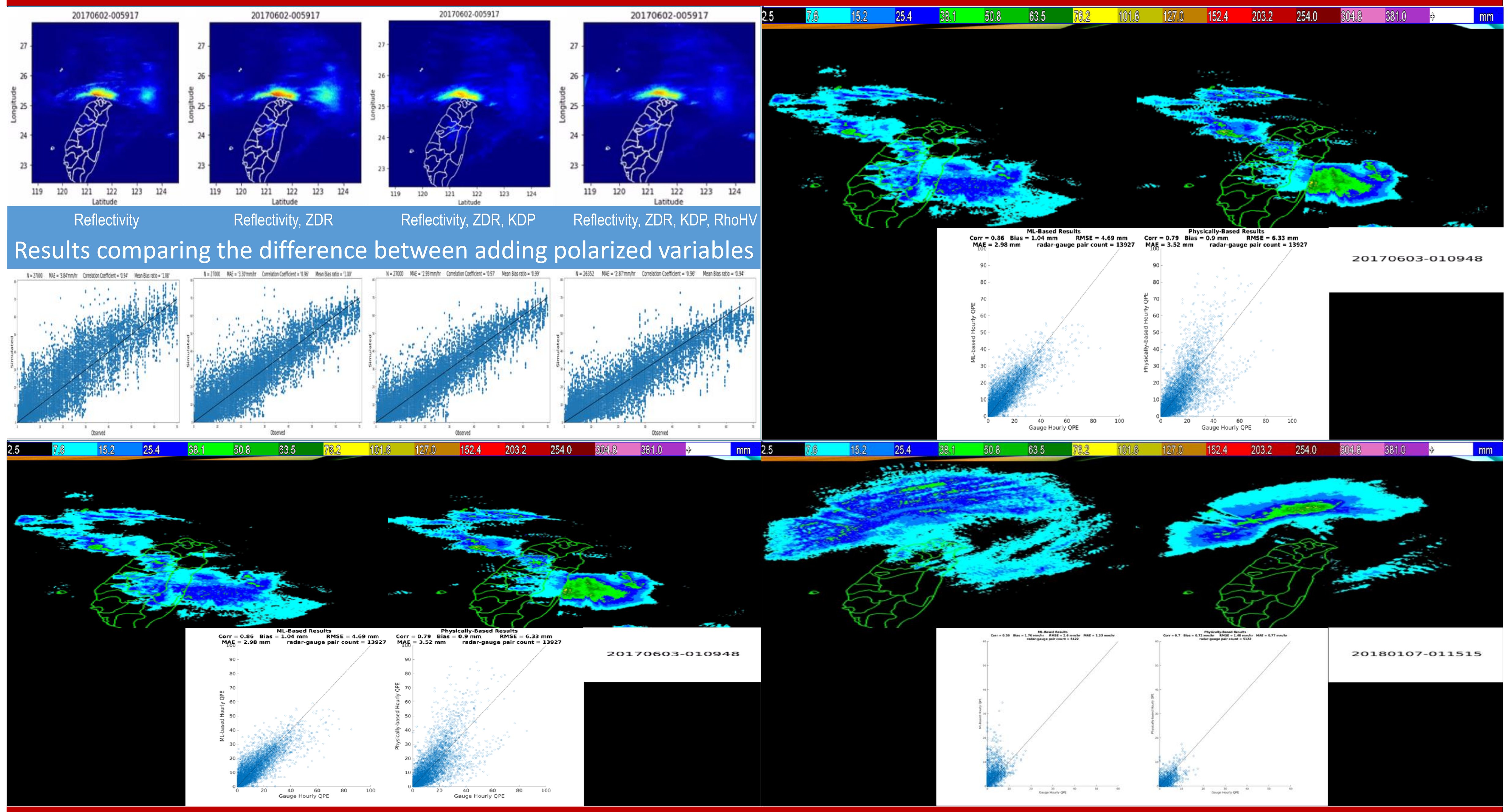
Layers.Dense(32)
Layers.LeakyReLU(alpha = 0.4)
Layers.Dropout(0.5)

Layers.Dense(1)
Layers.LeakyReLU(alpha = 0.4)

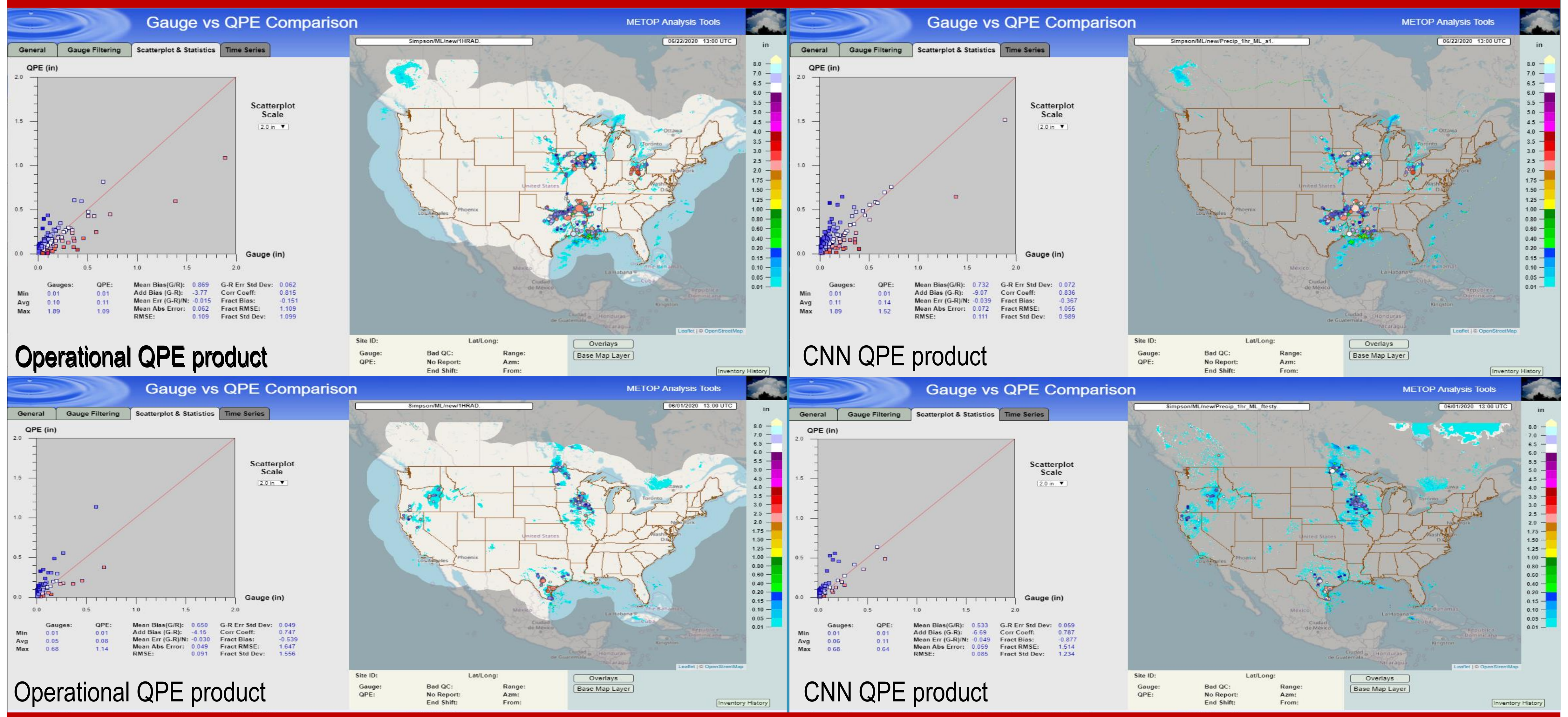
Optimizer = adam
Batch Size = 1,000
Number of Samples (CWB) ~ 17,000,000 (about 3 weeks)
Number of Samples (ConUS) ~ 4,000,000,000 (about 1 year)
Batch Steps = Number of Samples / Batch Size
Epoch Size = 50

Time to train on single-radar CWB data: 5-minutes
Time to train on mosaicked ConUS data: 1-hour

Single-Radar CWB Results



Mosaicked ConUS Results



Conclusions

The CNN model performs well for the single-radar (CWB) data, as well as for mosaicked ConUS results.

The single-radar showed, in general, better improvement to the operational QPE algorithm compared to the ConUS results.

Many more variables from MRMS may be added to improve performance, in addition to static variables such as latitude, longitude, etc.

Other machine-learning models will be implemented, including SVM, Random Forest, LSTM, ConvLSTM, Unet, etc.

Quality Control Methods

1) **Data extraction:** A 5x5 grid was centered over each individual rain gauge which served as the 'ground truth' for the current study.

2) **Domain setup:** MRMS values of -99000 indicate no data, whereas values of -99003 indicate outside of the radar (or mosaicked radar) domain. These values were excluded from the analyses.

3) **Binning data:** Because there are much more instances of zero precipitation compared to large values of rainfall, the data was arbitrarily binned based on the following ranges (mm/hr):

   0, 0-1, 1-3, 3-6, 6-25, 15-25, 25-40, 40-55, 55-120

4) **Precipitation mask:** Although the gauges utilized for ground-truth go through a rigorous quality control scheme, some anomalies still exist. Therefore, a precipitation mask of operational values that closely matched ground-truth were utilized.

5) **Normalization:** The data were normalized before being introduced to the model instead of at every batch for each epoch.

6) **Remove autocorrelation:** The data were randomly shuffled to ensure the same day was not included in any training, validation, or simulation dataset.

Dates

CWB data range: 18 days spanning throughout 2014, 2015, 2017, and 2018.

ConUS data range: 20190501 - 20200630

Input Variables

CWB Input Variables: Single radar (RCWF) data from 9 different elevation tilts:

Reflectivity
ZDR
KDP
RhoHV

ConUS Input Variables: Mosaicked data based on the MRMS product suite:

Reflectivity (SHSR)
Base Reflectivity (RALA)
Composite Reflectivity (CREF)
Reflectivity at 0°C
Reflectivity Height (SHSRH)
RAP temperature
Vertically Integrated Liquid Water Content (VIL)
Terrain (Elevation above Sea Level, m)