MRMS Machine Learning QPE

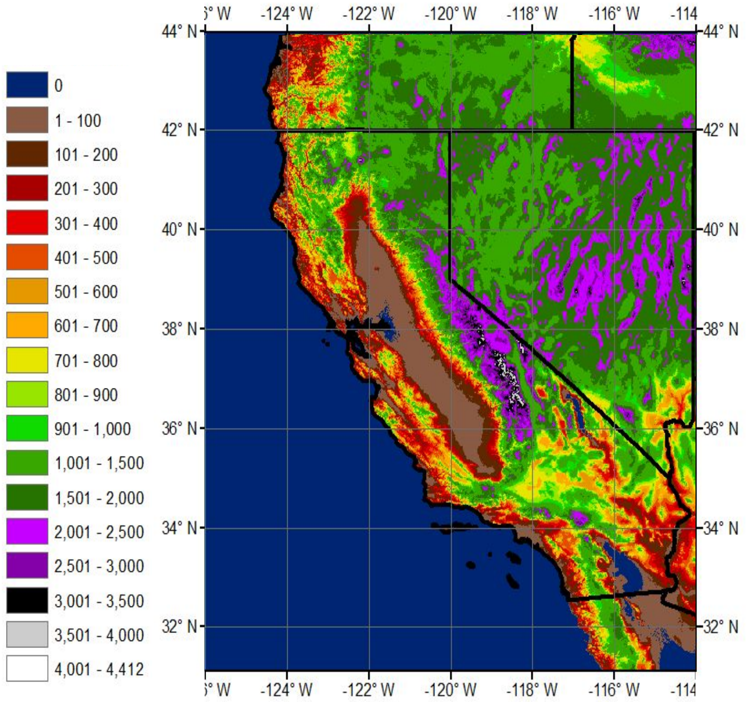
Andrew Osborne[1,2], Jian Zhang[2], Race Clark[2], Steve Cocks[1,2], Micheal Simpson, Ken Howard[2]

[1] *Cooperative Institute for Severe and High Impact Weather Research and Operations, Norman, OK*

[2] *NOAA/OAR National Severe Storms Laboratory, Norman, OK*

FFaIR Seminar Series

June 6, 2023



Introduction

- Looking to improve MRMS QPE in complex terrain using ML based approach

❑ The high terrain inhibits radar coverage in these areas

❑ Orographic enhancement of precipitation is difficult to accurately capture with current rain rate relationships

❑ Goal is to develop an ML QPE product available every 2 minutes to augment radar-based QPE in the trouble areas of the West

- Initial ML study:

❑ Convolutional Neural Network (CNN) model with 13 radar variables shows improvement over radar-based QPE in mountainous terrain for selected cases and year-long study

- Recent ML development:

❑ Long term statistical analysis of CNN with additional model and terrain input variables (29 input variables) for the western CONUS domain

❑ Long term statistical analysis of CNN applied to Hawaii domain

❑ Closer look at individual cases and areas of interest to help interpret and explain the CNN model output

CNN Model Setup

- Convolution Neural Network
- Input data: 5 x 5 grids of radar, model, terrain variables
- Truth datasets:

  -Gauge precipitation value at center of 5 x 5 grid (Hourly QC'ed from HADS dataset)

  -Hourly MRMS Multi-Sensor QPE precip value at center of 5 x 5 grid
- 2 min input variables are accumulated over an hour to compare

to the hourly gauge precipitation value

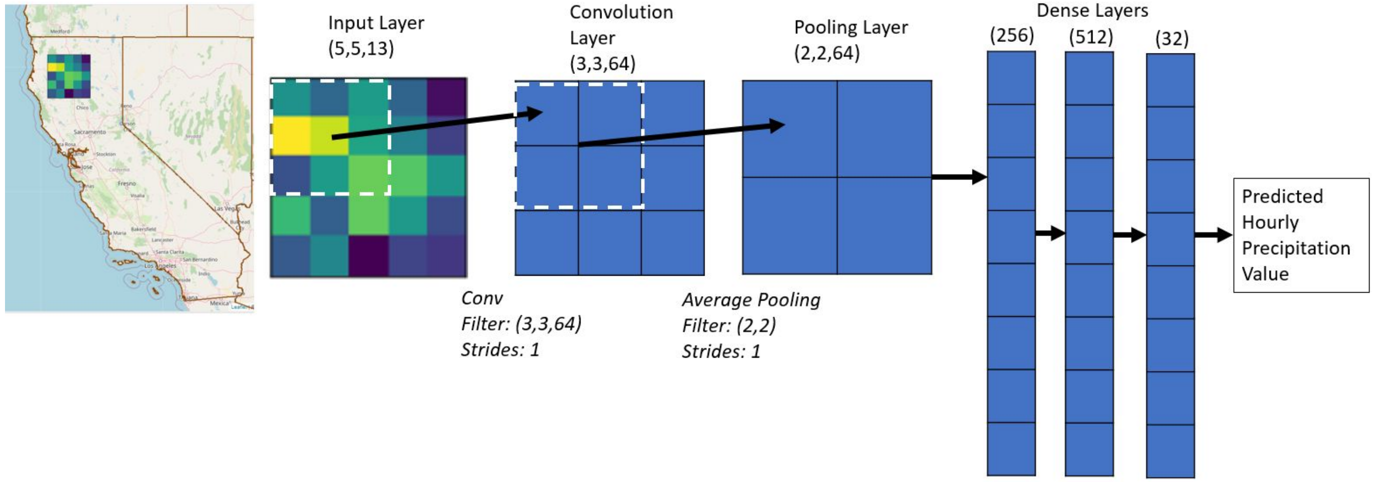
- Preprocessing including binning by precip amount, rotating the input fields, and MinMax scaler applied
- 2 Regional domains: 1 covering CA, 1 covering HI

Precipitation Size Bins

0-1 mm	10-12 mm	22-24 mm
1-2 mm	12-14 mm	24-26 mm
2-4 mm	14-16 mm	26-28 mm
4-6 mm	16-18 mm	28-30 mm
6-8 mm	18-20 mm	>30 mm
8-10 mm	20-22 mm	

CA Dataset Breakdown:

Training	Validation	Simulation
Feb. 2019 Dec. 2019 Dec. 2020	Feb. 2020	6 Mar. 2019 27 Nov. 2019 17 Jan. 2020 13 Mar. 2020 **All Precip Days 2021**



HI Dataset breakdown:

-Training: *12 months*

-Validation: *202003*

-Simulation: *20181009,20210228,20210204,20210314,20210309,20190626,* ***Nov. 2021 – Oct. 2022***

Input Variables

Radar Variables	Terrain Variables	Model Variables
Seamless Hybrid Scan Reflectivity (SHSR)	*Orographic forcing factor*	*Precipitable Water*
Vertically Integrated Liquid (VIL)	*Mean U wind 850-700 mb*	*Surface Temperature*
Radar Quality Index (RQI)	*Mean V wind 850-700 mb*	*Dewpoint Temperature*
Composite Reflectivity (CREF)	*Latitude*	*Wet bulb Temperature*
Reflectivity at Lowest Altitude (RALA)	*Longitude*	*0C Height*
Bright Band Bottom (BB_BOTTOM)	*Terrain Height*	*Prob of Warm Rain*
Bright Band Top (BB_TOP)	*PRISM*	*Precipitation Efficiency*
Reflectivity at 0C		*1 hour QPF*
Reflectivity at -5C		
Reflectivity at -10C		
Reflectivity at -15C		
Reflectivity at -20C		
Precipitation Type		
Seamless Hybrid Scan Height (SHSRH)		

-All 29 input vars used for CONUS domain year-long study

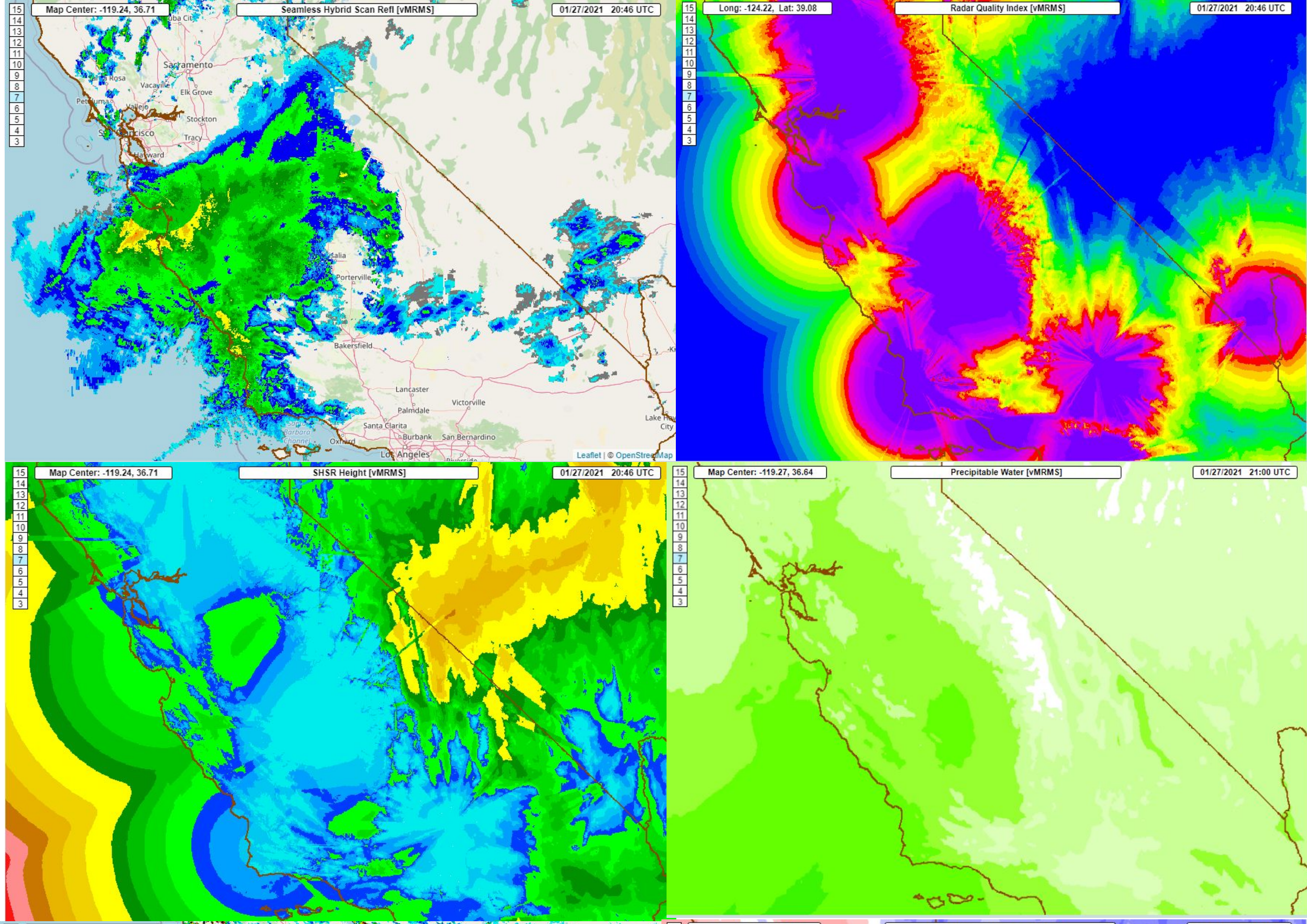
-19 radar and terrain variables used for HI long-term run

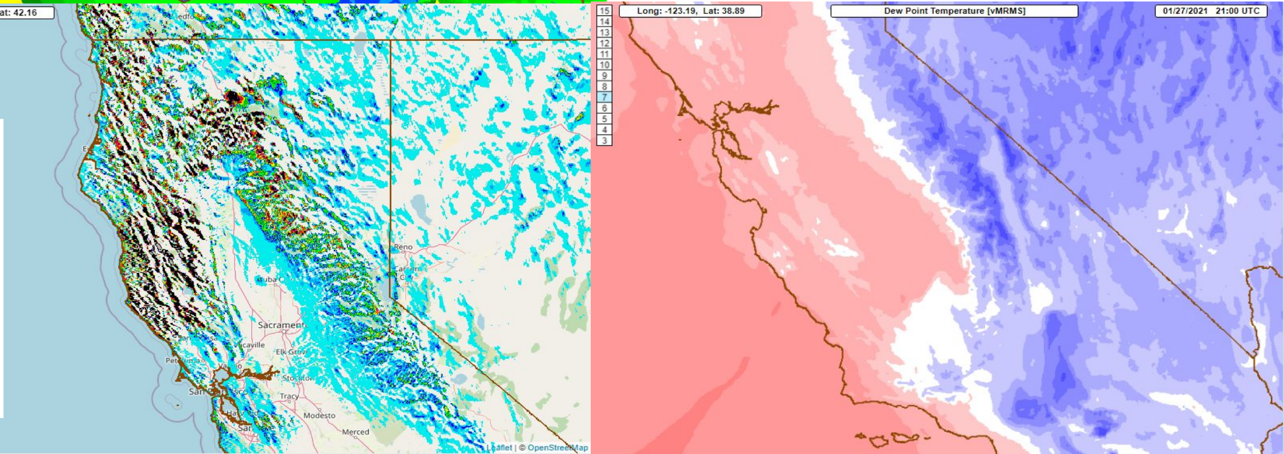
Orographic forcing factor:

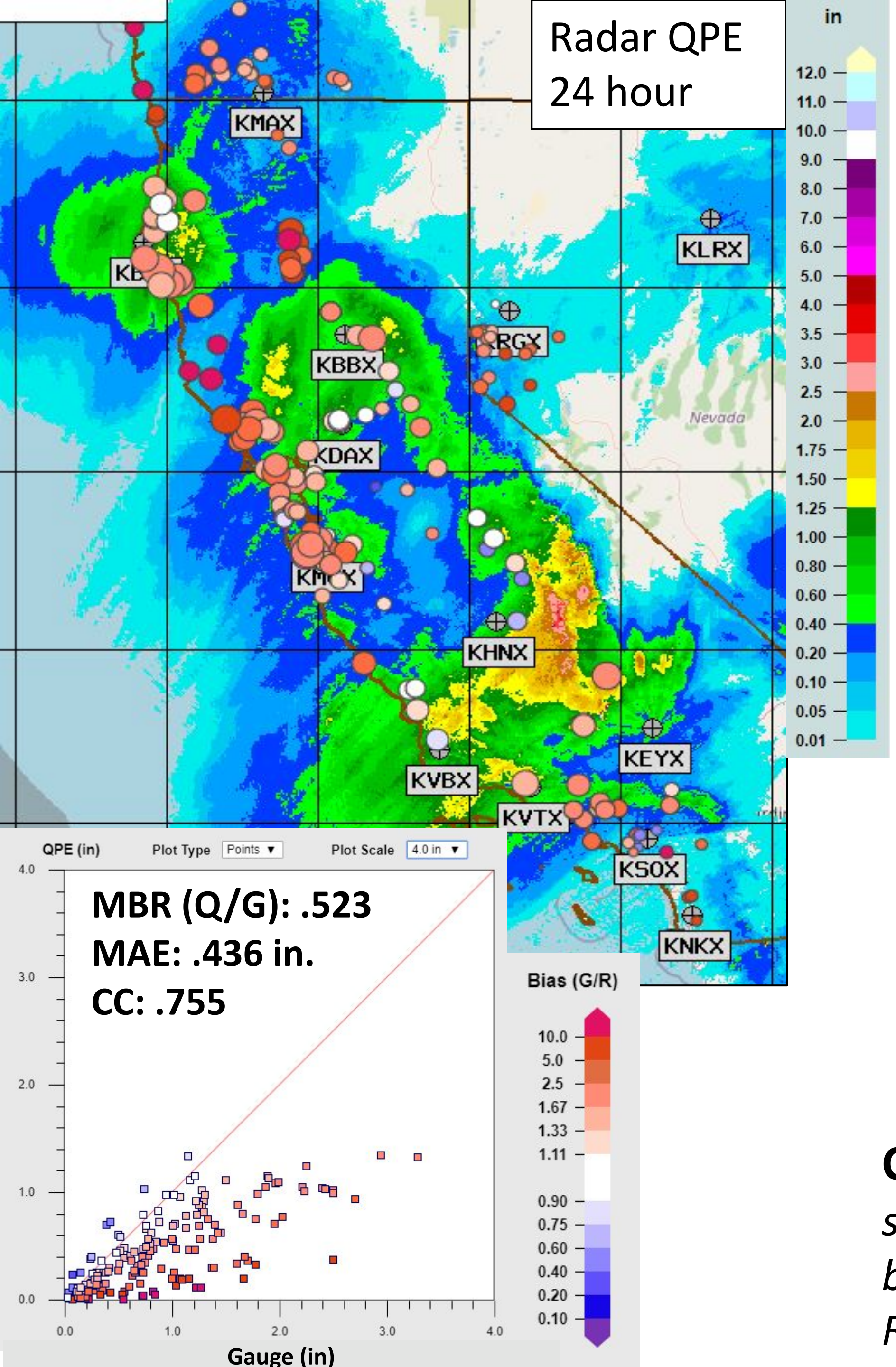
$$f_{orog} = \text{1 hour QPF} * \left(\bar{u} * \frac{dh}{dx} + \bar{v} * \frac{dh}{dy} \right)$$

where $\bar{u}$, $\bar{v}$=850-700 mb mean wind

and h is terrain height



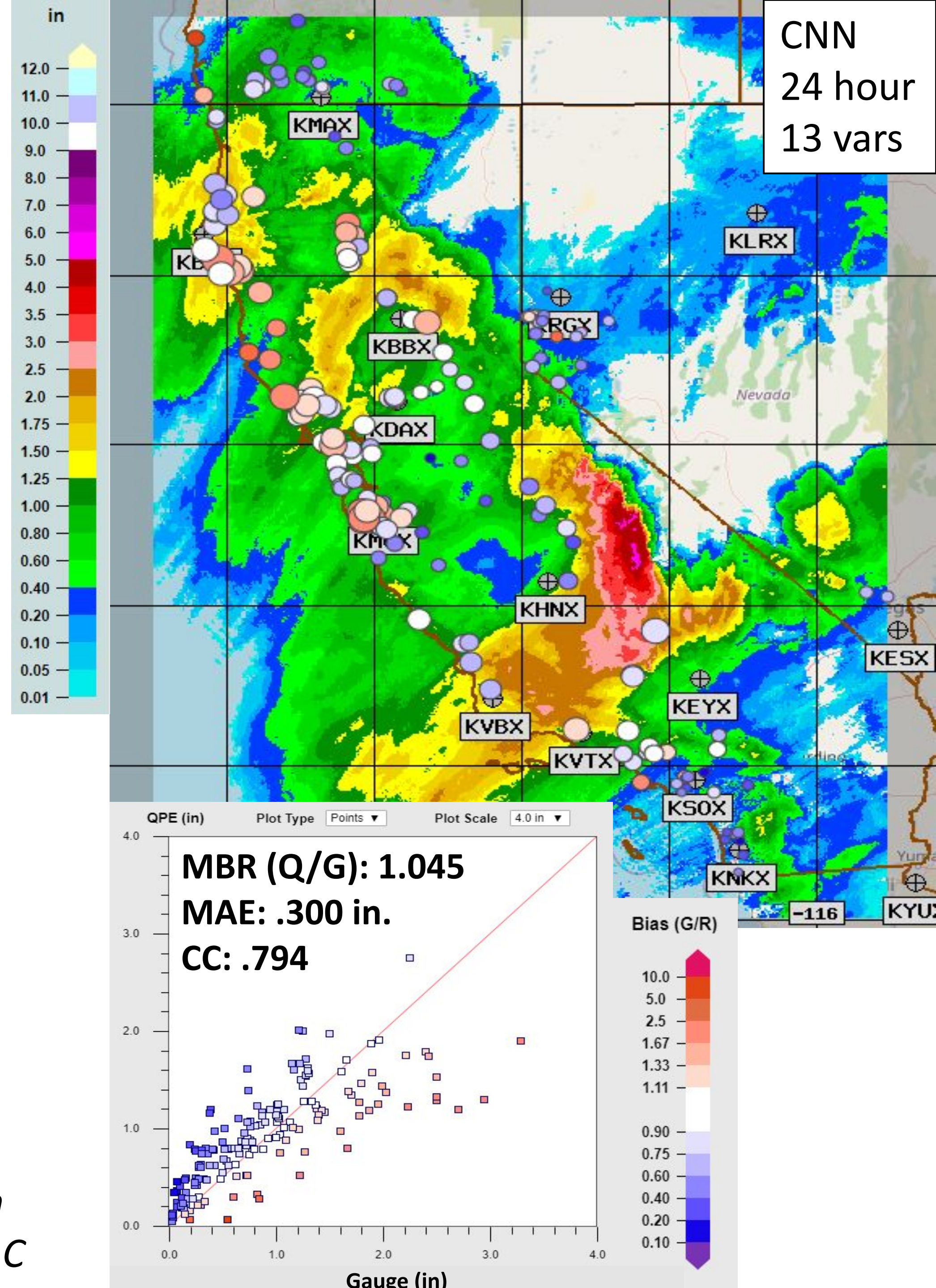




20190306
1500 UTC

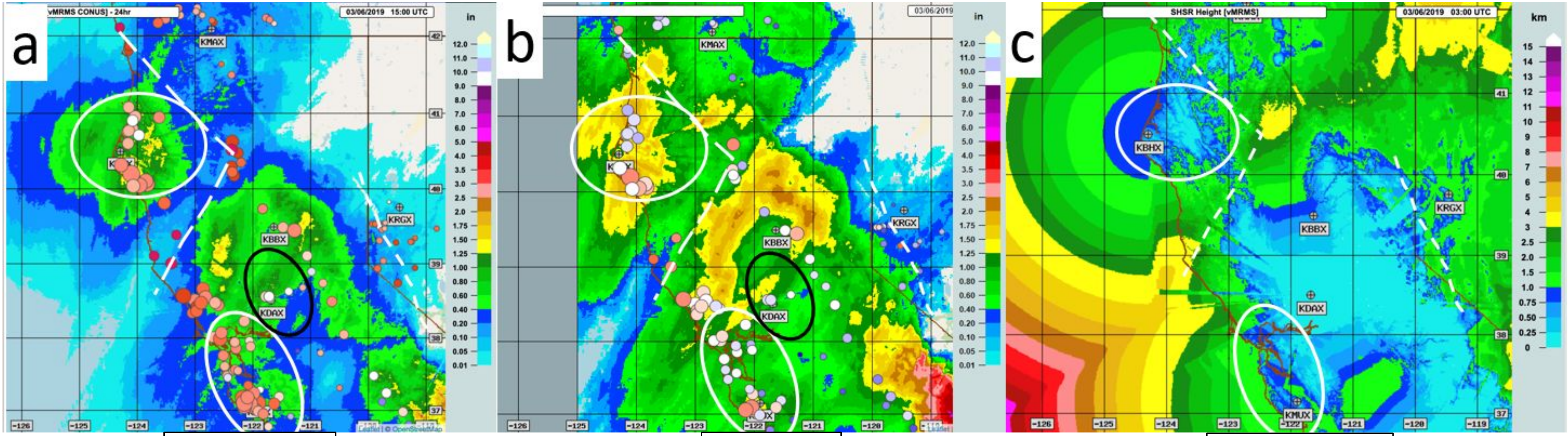
-Initial CNN version shows improvement over radar-based QPE

-CoCoRAHS Gauges used for verification



CNN v1.0 Input variables:

shsr, vil, rqi, shsrh, cref, rala, bright band bottom/top, shsrh Reflectivity at 0,-5,-10,-15,-20 C



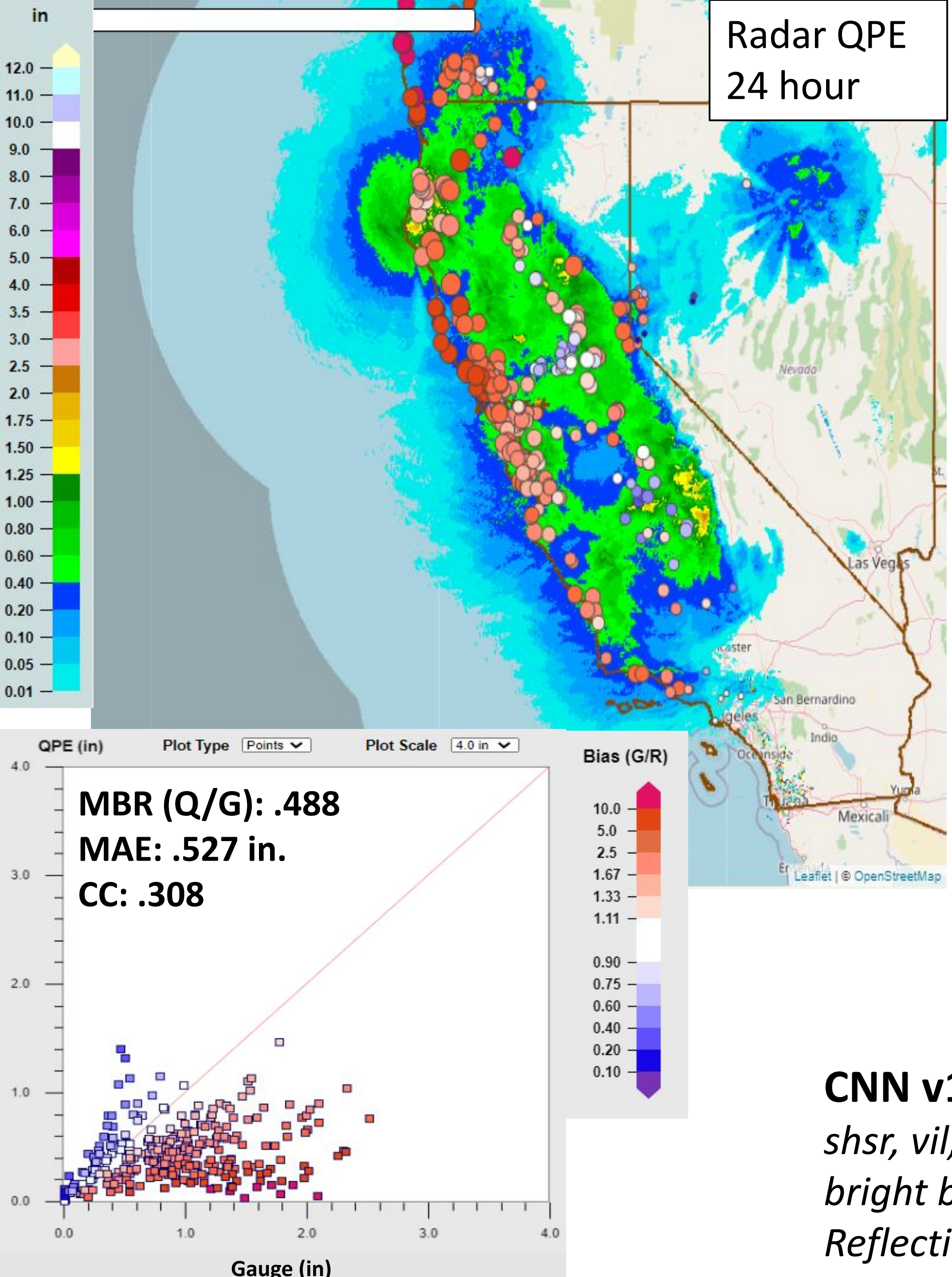
Radar QPE

CNN QPE

SHSR Height

-Zoomed in look at 20190306 1500 UTC
24 hr accumulations Northern CA

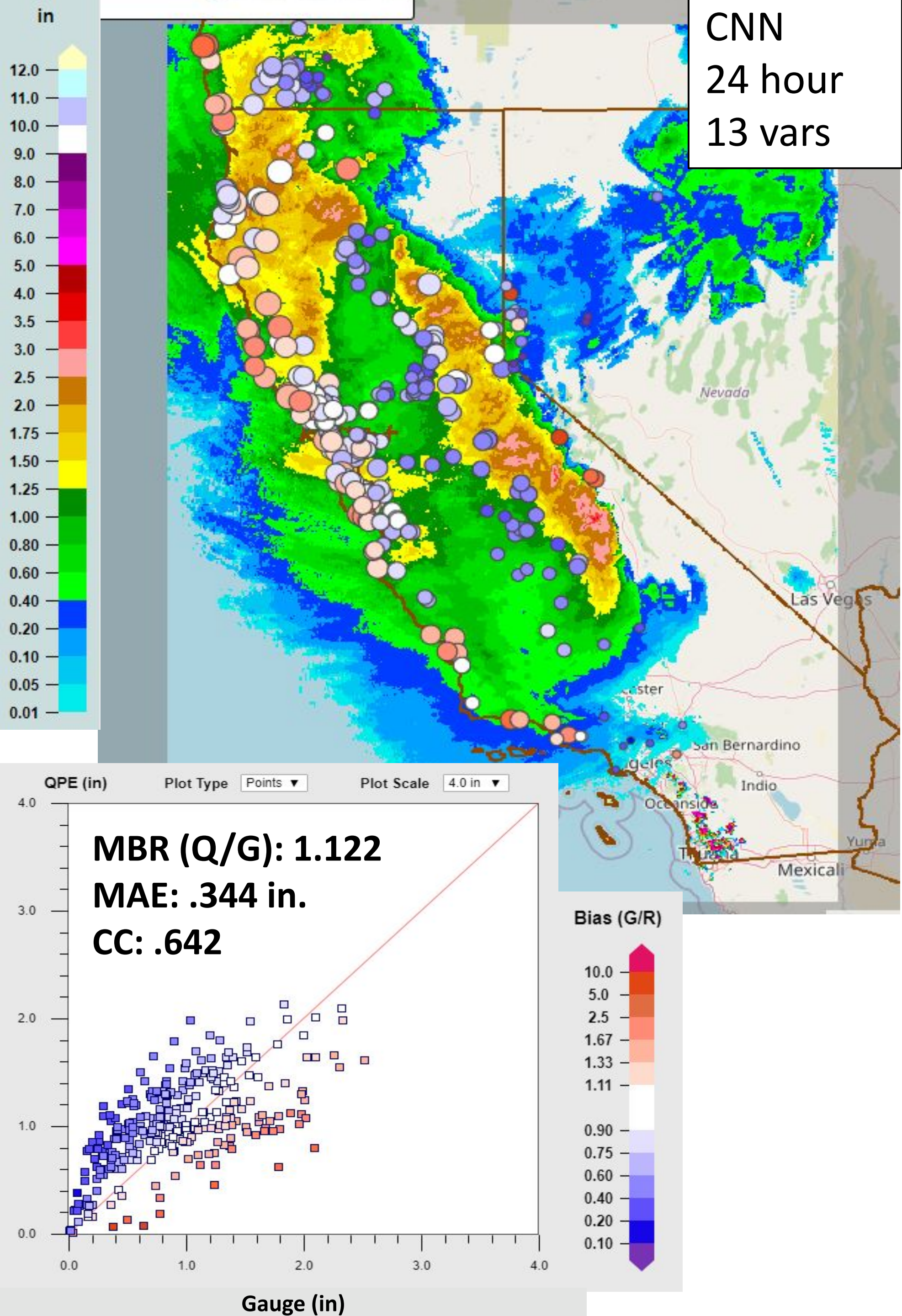
-ML QPE fills in precipitation in areas of radar coverage gaps (high radar beam height) and also improves precip estimates close to radar



20191127
1500 UTC

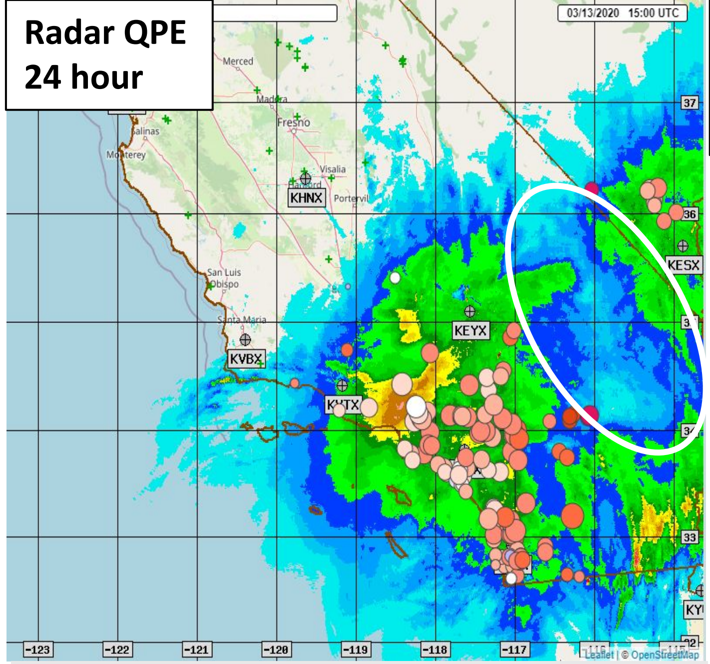
-Initial CNN version shows improvement over radar-based QPE

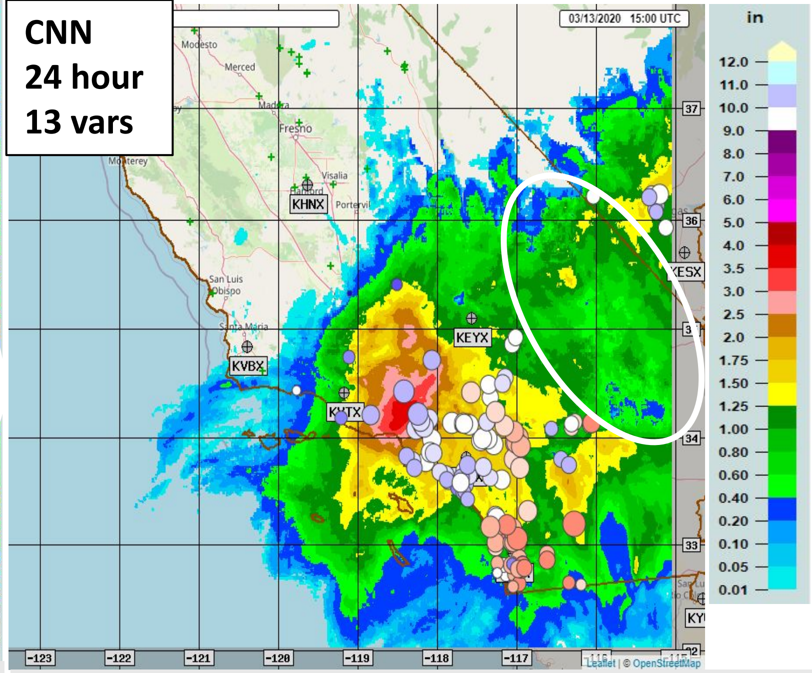
-CoCoRAHS Gauges used for verification

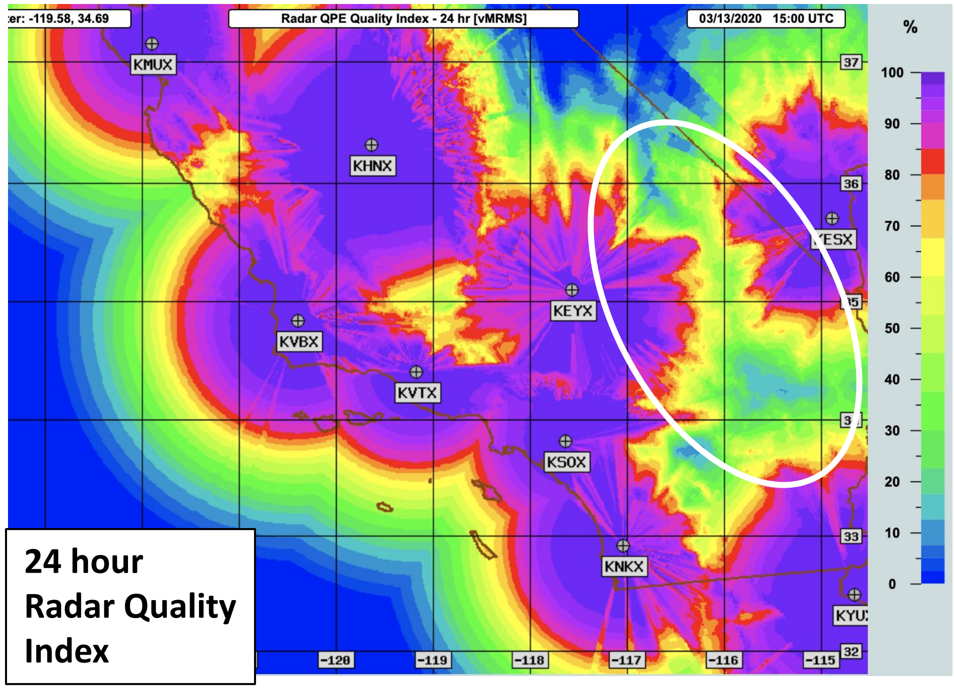


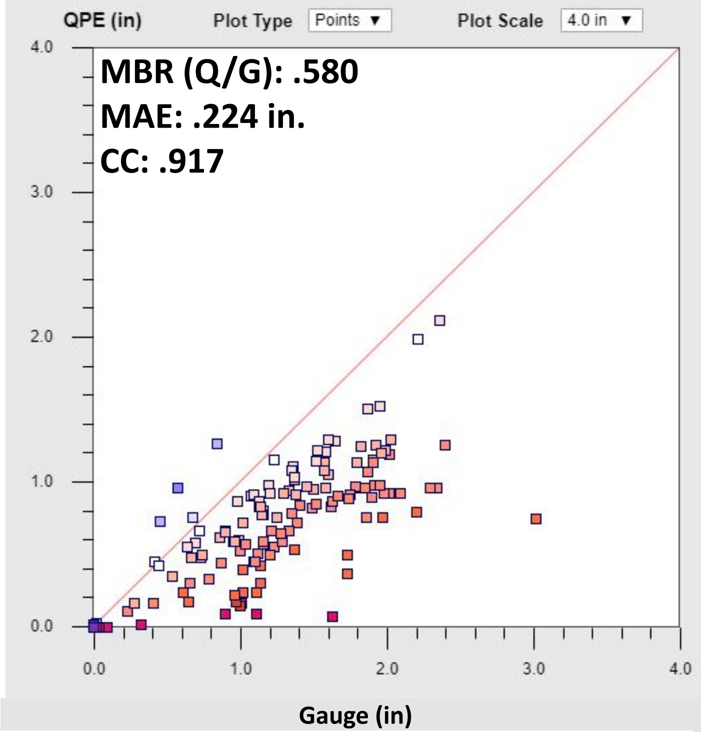
CNN v1.0 Input variables:

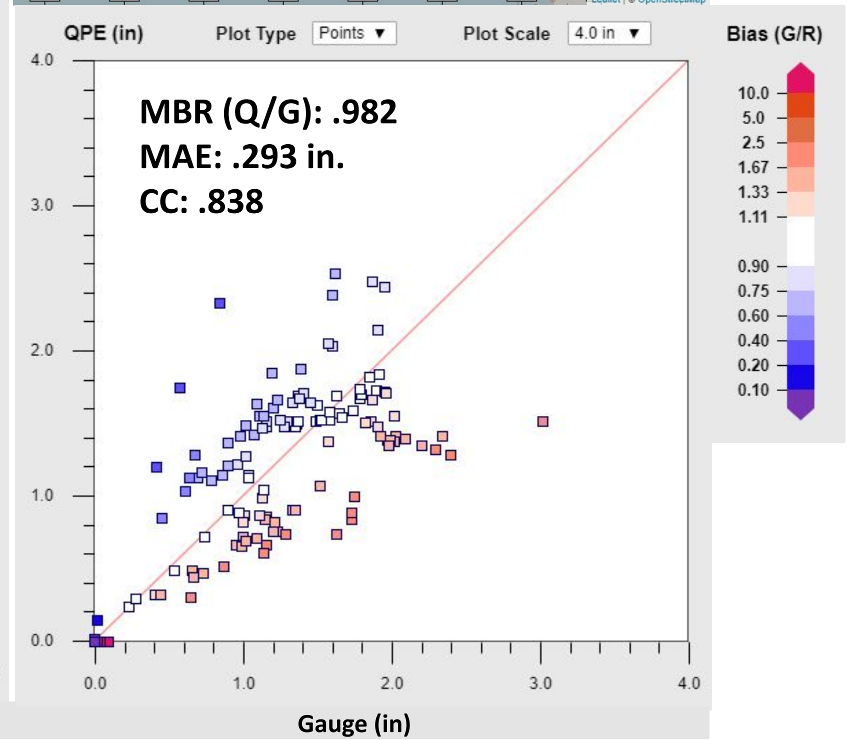
shsr, vil, rqi, shsrh, cref, rala, bright band bottom/top, shsrh Reflectivity at 0,-5,-10,-15,-20 C











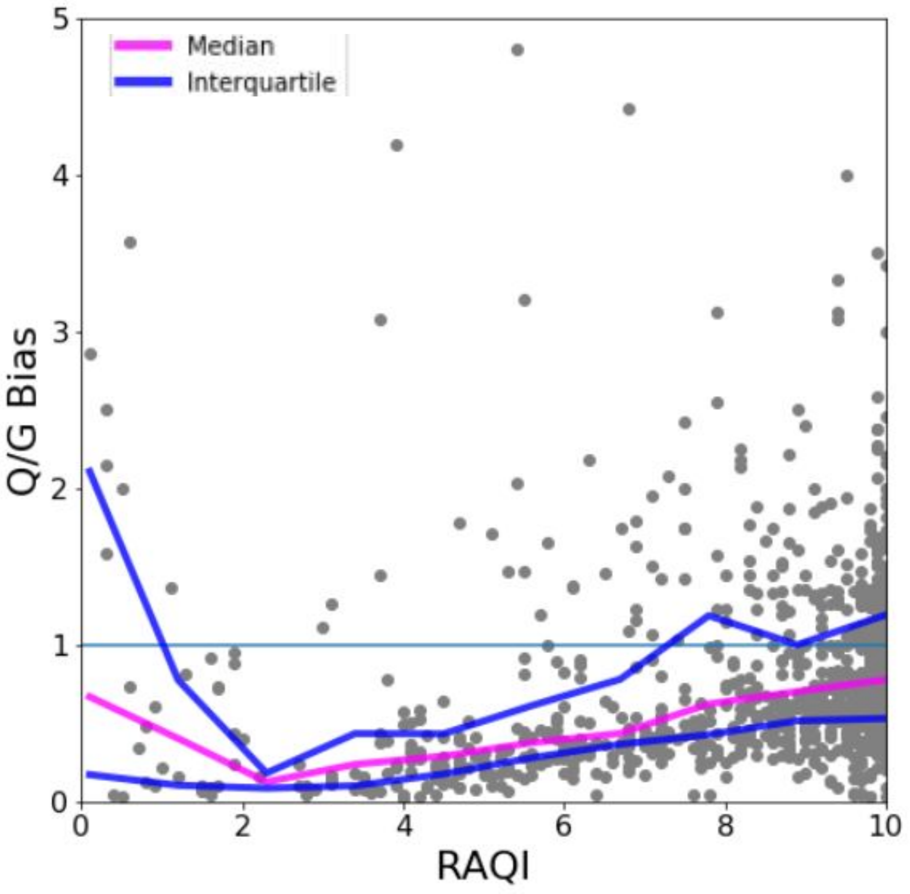
20200313 1500 UTC

-This case is a good example of how the CNN is able to fill in precipitation areas associated with radar coverage gaps

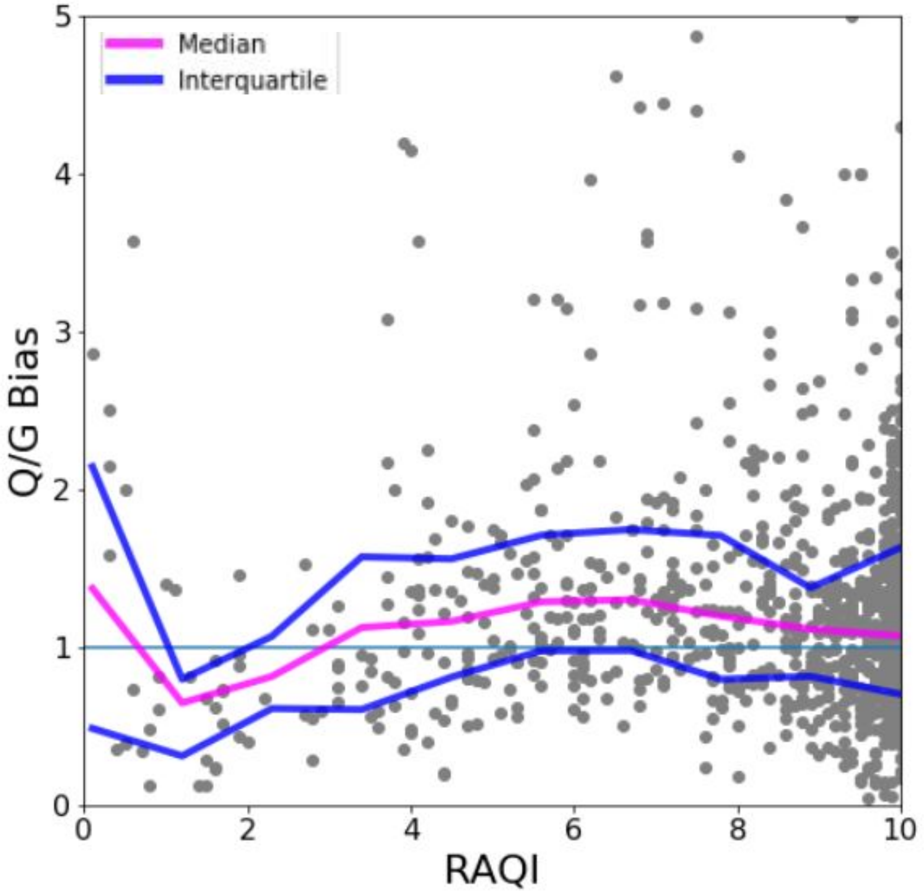
QPE Bias vs. Radar Quality

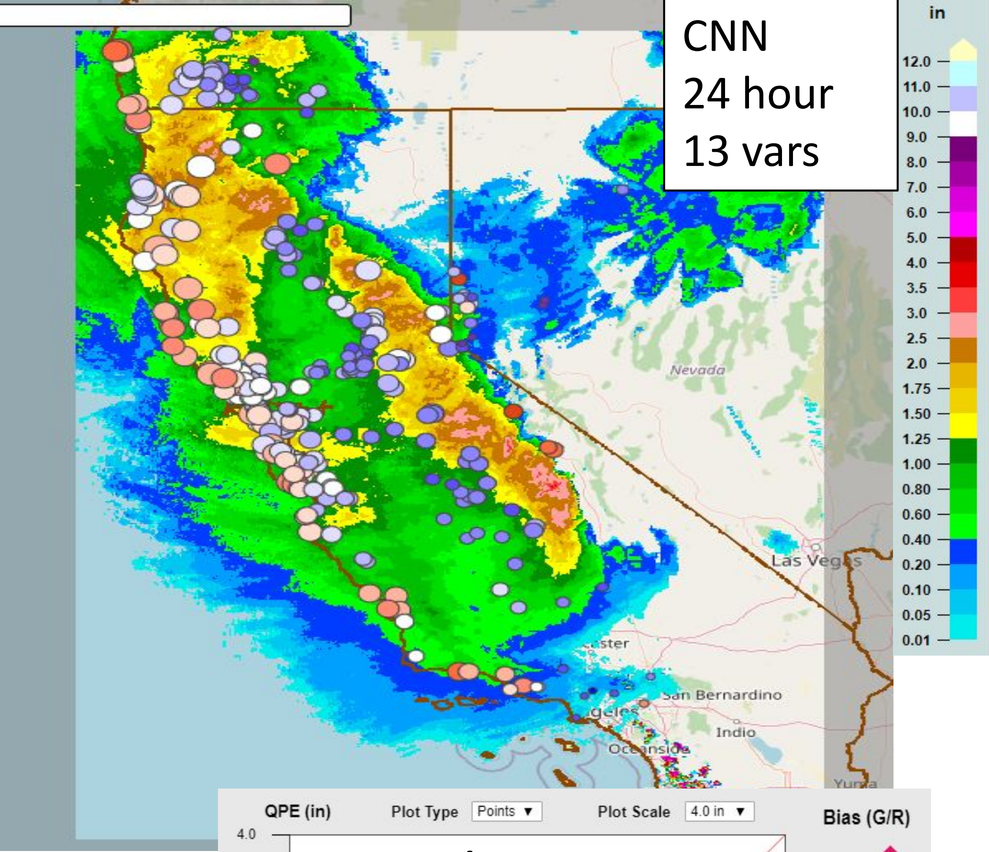
-For several cases studied in the western US, QPE is increased and underestimation bias is reduced in areas of poor radar coverage

Radar QPE



CNN Model





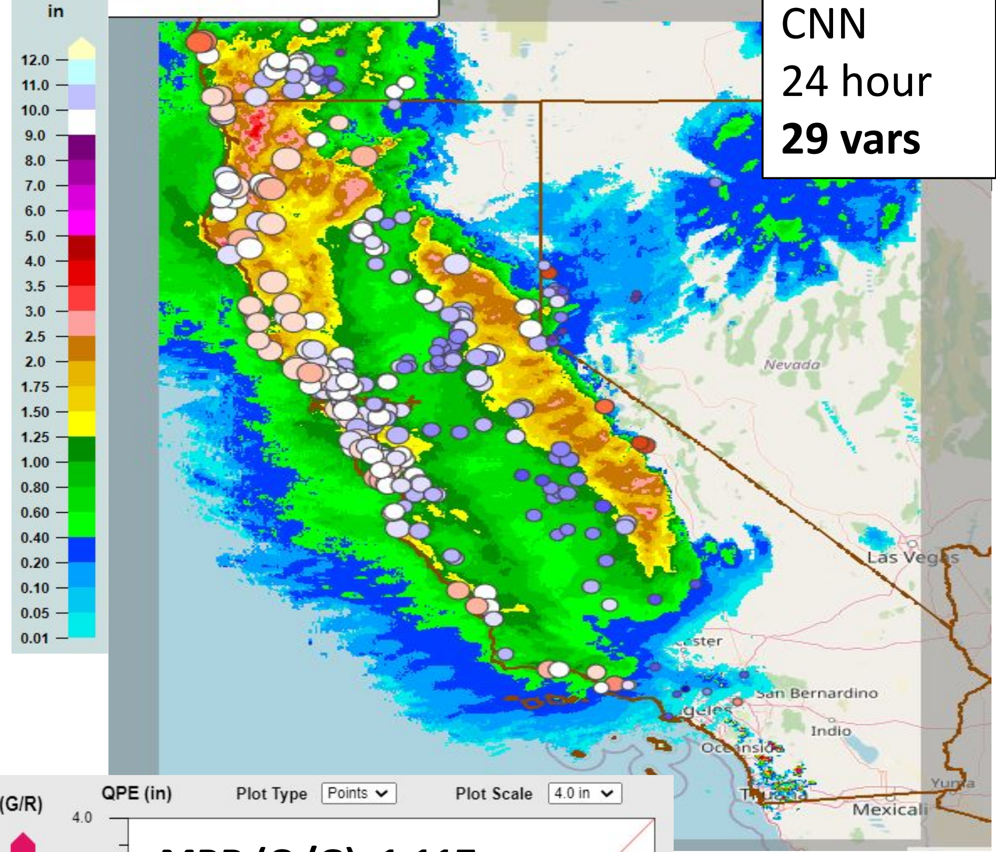
CNN
24 hour
13 vars

20191127
1500 UTC

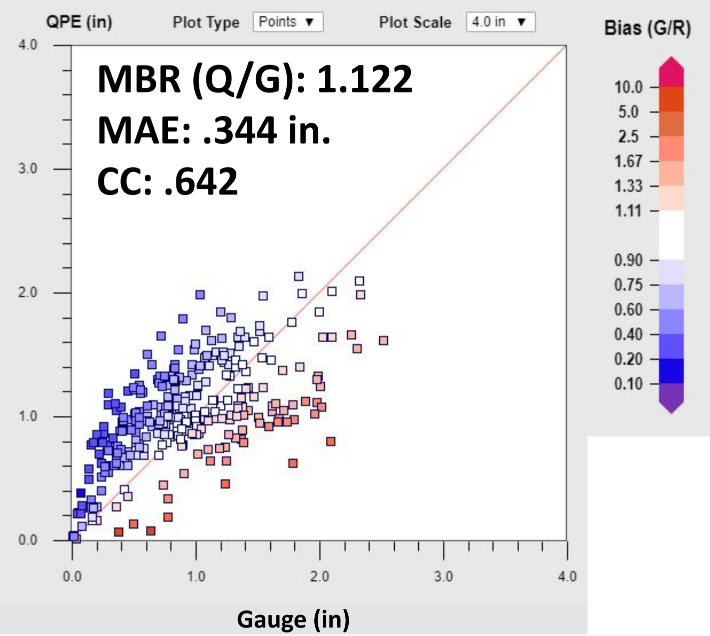
-Initial look at CNN model with additional variables for this case showed clear improvements

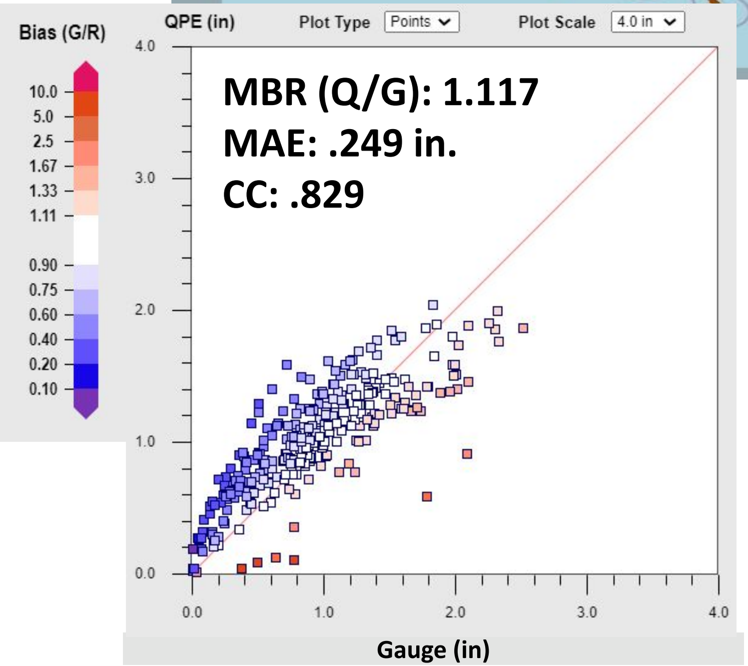
-Model variables likely the key to the improvement based on permutation testing results

CNN
24 hour
29 vars



CNN input:
13 radar
variables





CNN input:
13 radar variables
+
16 model and terrain-related variables

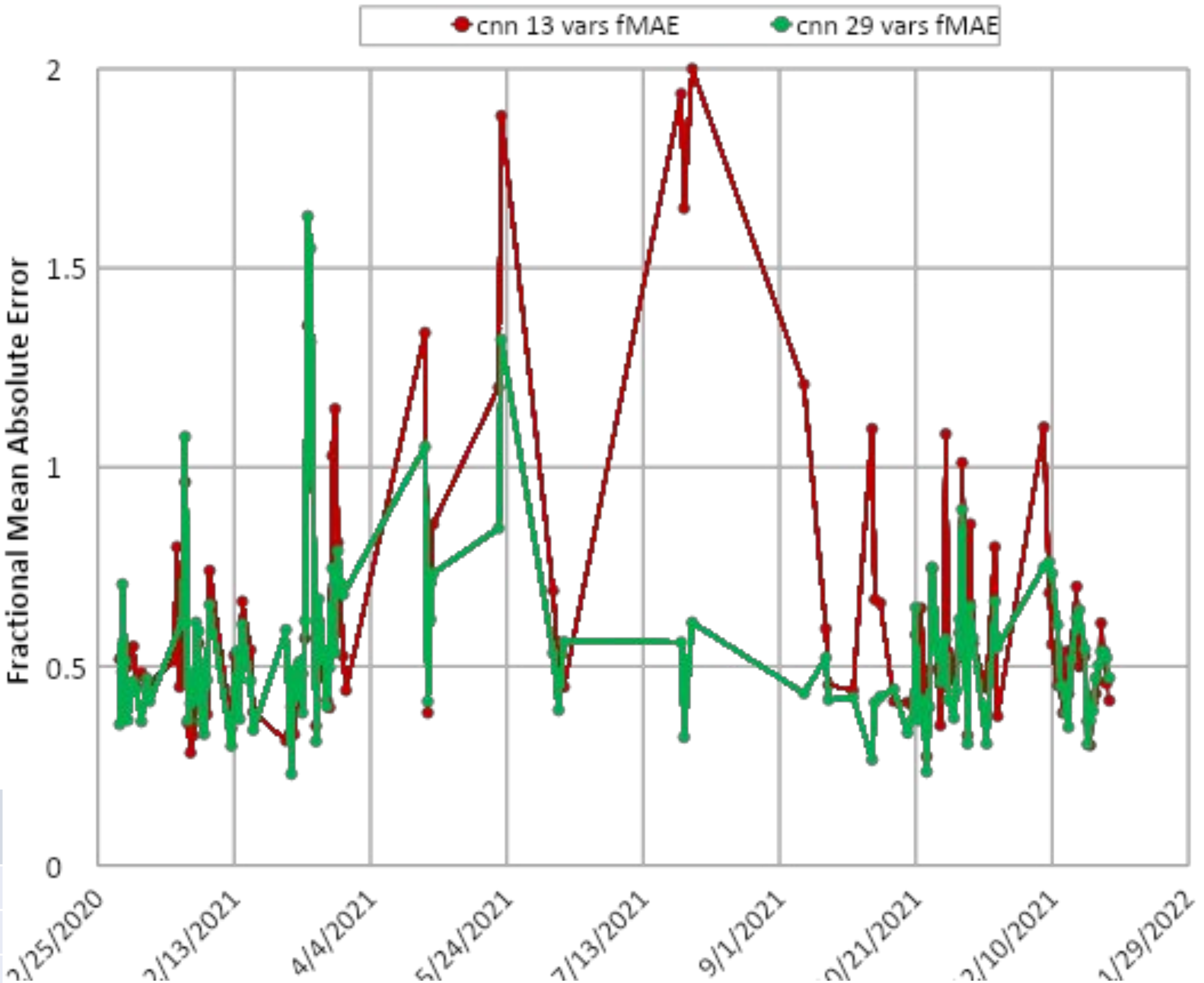
29 vars vs. 13 vars
Full Year 2021

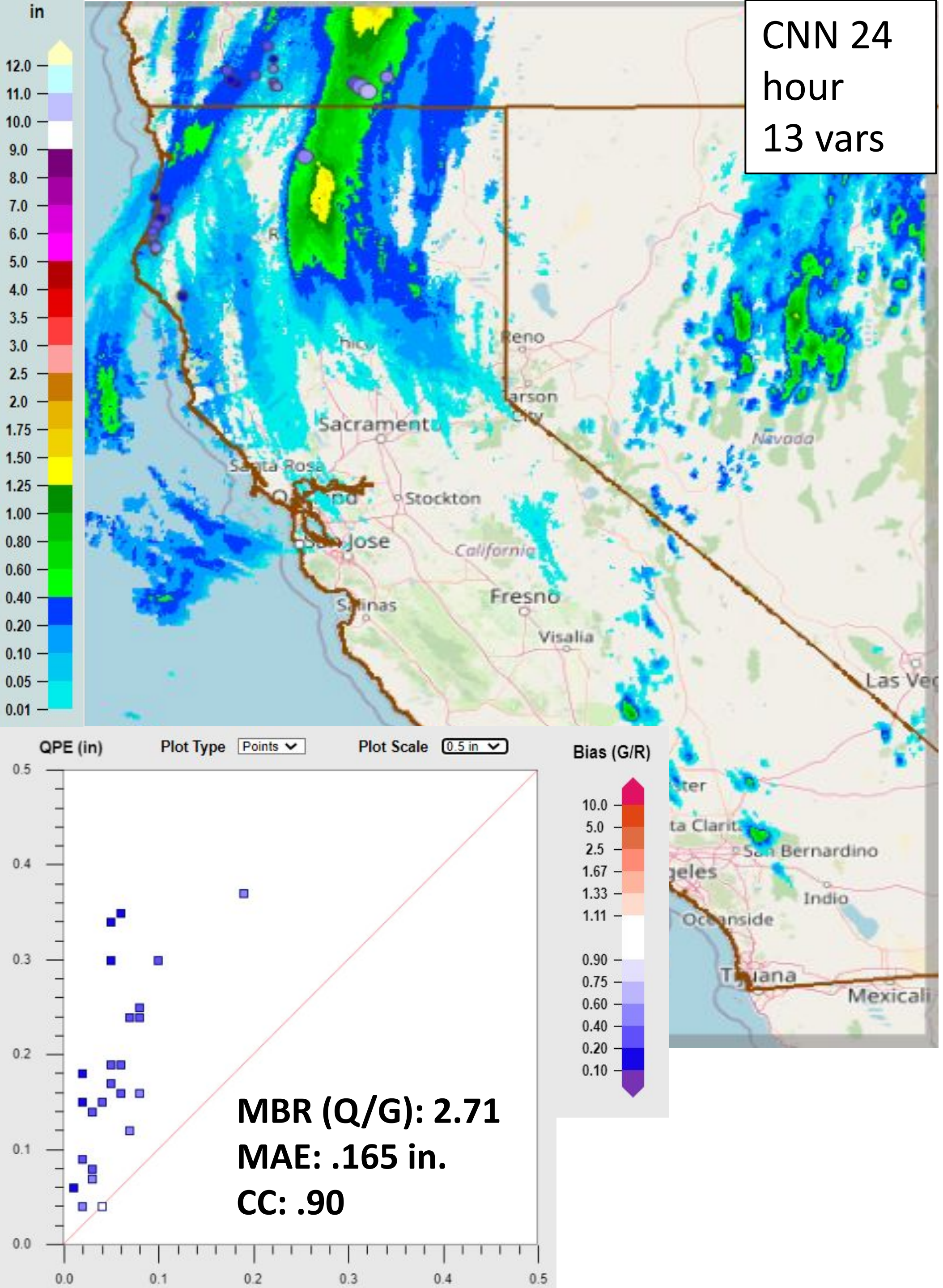
-Decreased fMAE from 29 variable CNN for summer months

-General slight improvement seen throughout other parts of the year

-Overall improved long term stats from the 29 vars CNN run compared to 13 vars CNN

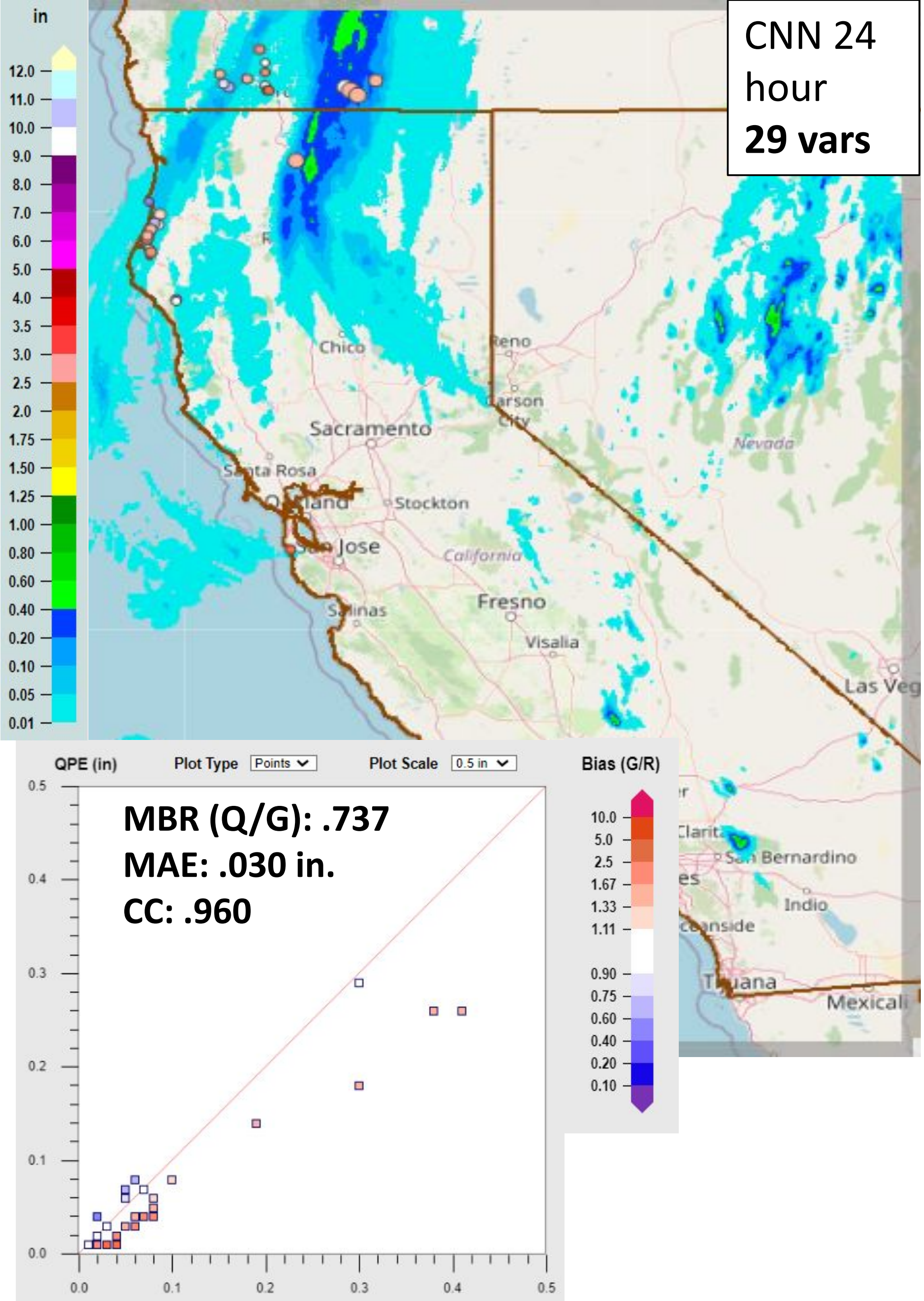
	Q/G Bias	CC	fMAE
CNN (13 vars)	1.126	0.656	0.621 in
CNN (29 vars)	1.089	0.735	0.535 in
Q3EVAP	.604	.575	.592 in

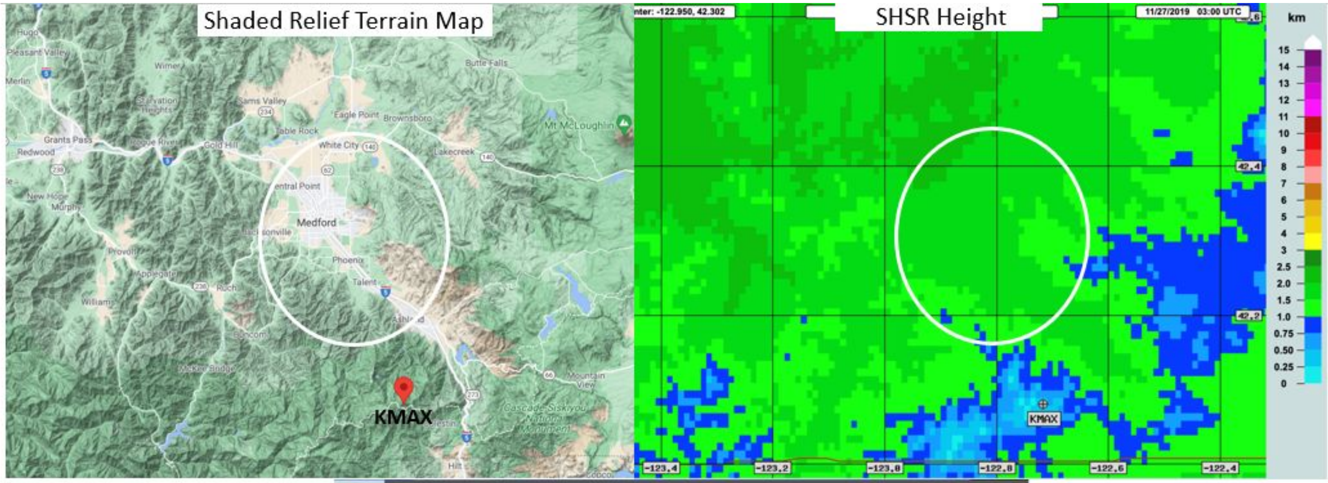


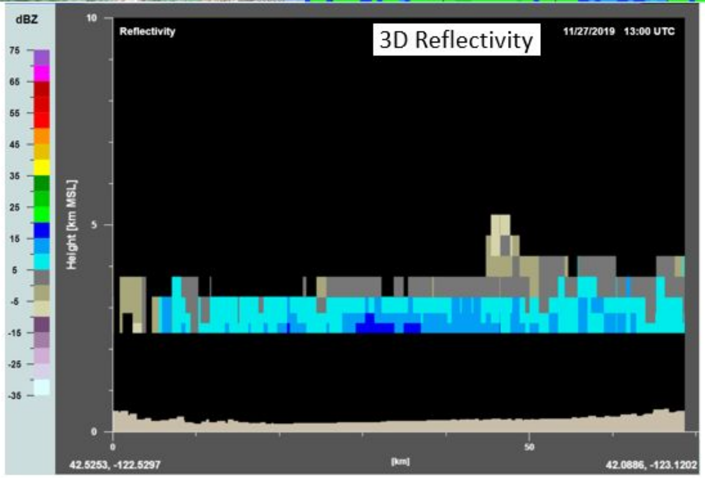


20210728
1500 UTC

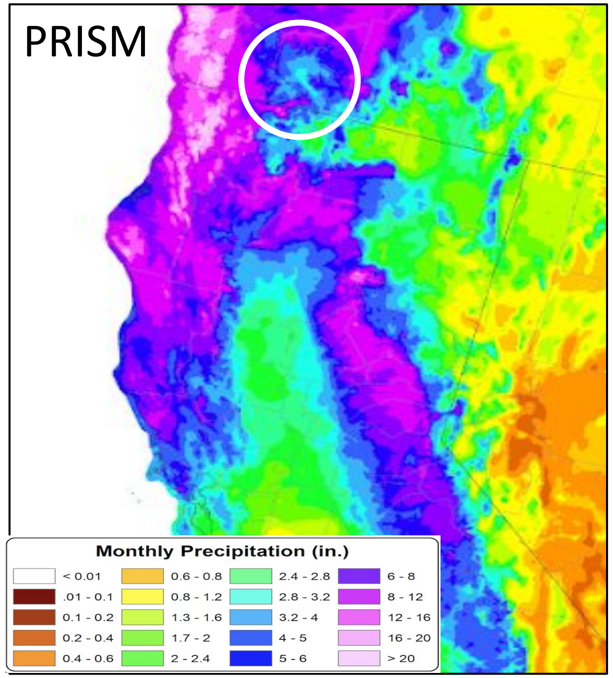
-Additional input variables improve the overestimate bias for this warm season case







Localized Evaporation Effects

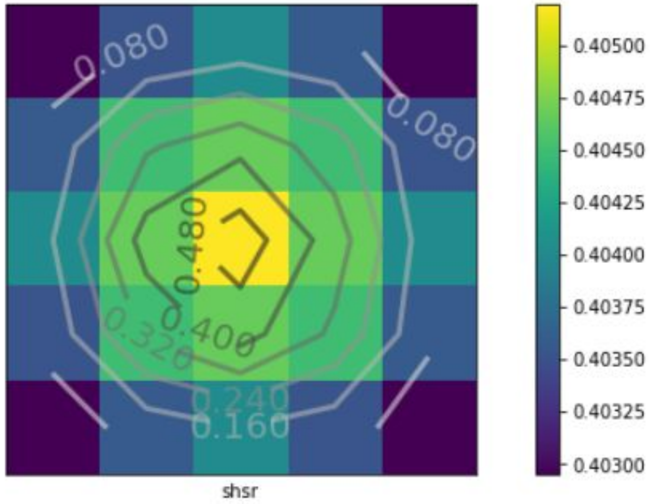


-Notable overestimation bias over Rouge Valley area of southern Oregon, which PRISM shows is climatologically drier than surrounding areas

-The radar beam is well above the ground in this area and the CNN model with only 13 input radar variables likely is not capturing the local evaporative effects in the lowest layer

-Adding some model variables with moisture information likely helps the CNN model account for the evaporative effects and reduce the overestimate bias for this case

Model Interpretation



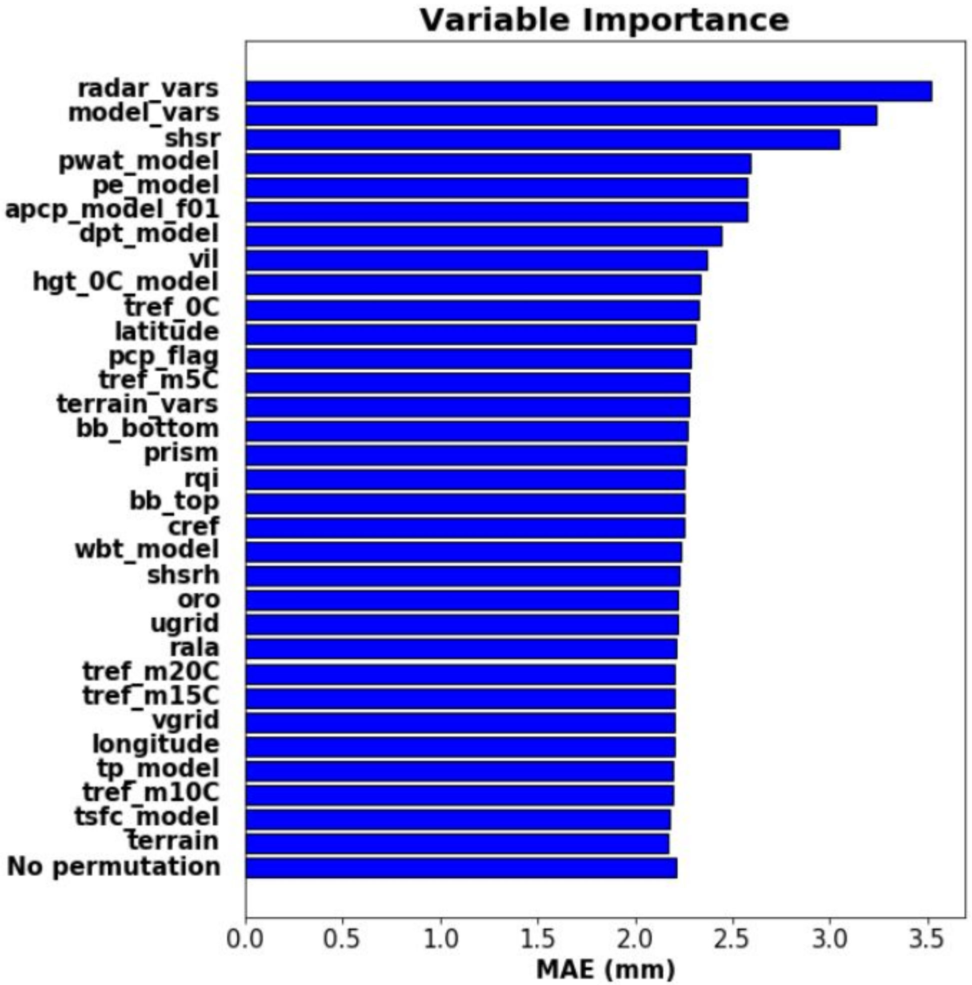
Input*gradient plot for SHSR

-SHSR variable most important to prediction

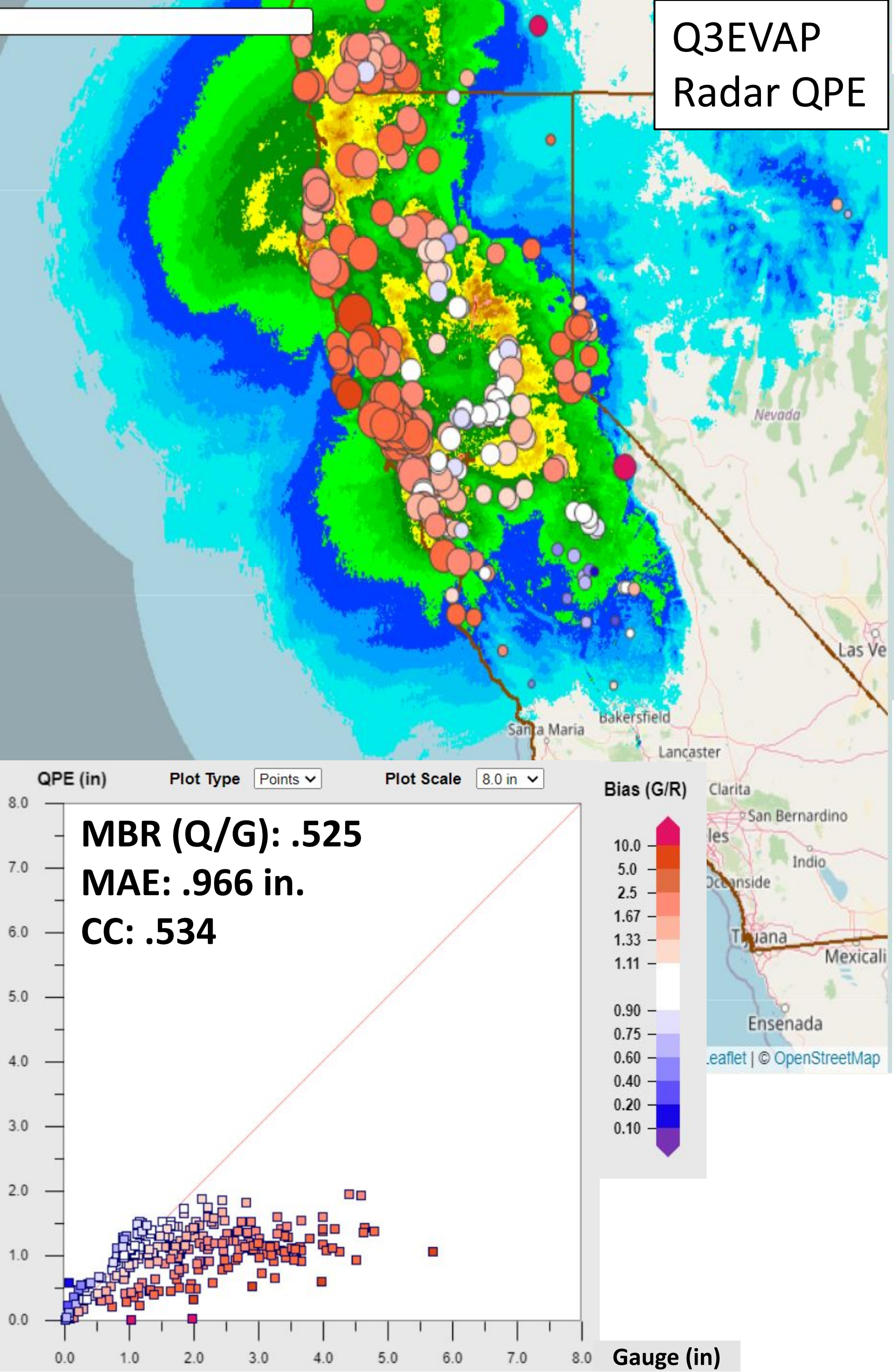
-PWAT model and 1 hour QPF also show strong importance

-For the 3 different categories of input variables:

- *Radar variables most important (due to SHSR)*
- *Model variables also show high importance (spread among a few different moisture related vars)*
- *Terrain variables show low importance*



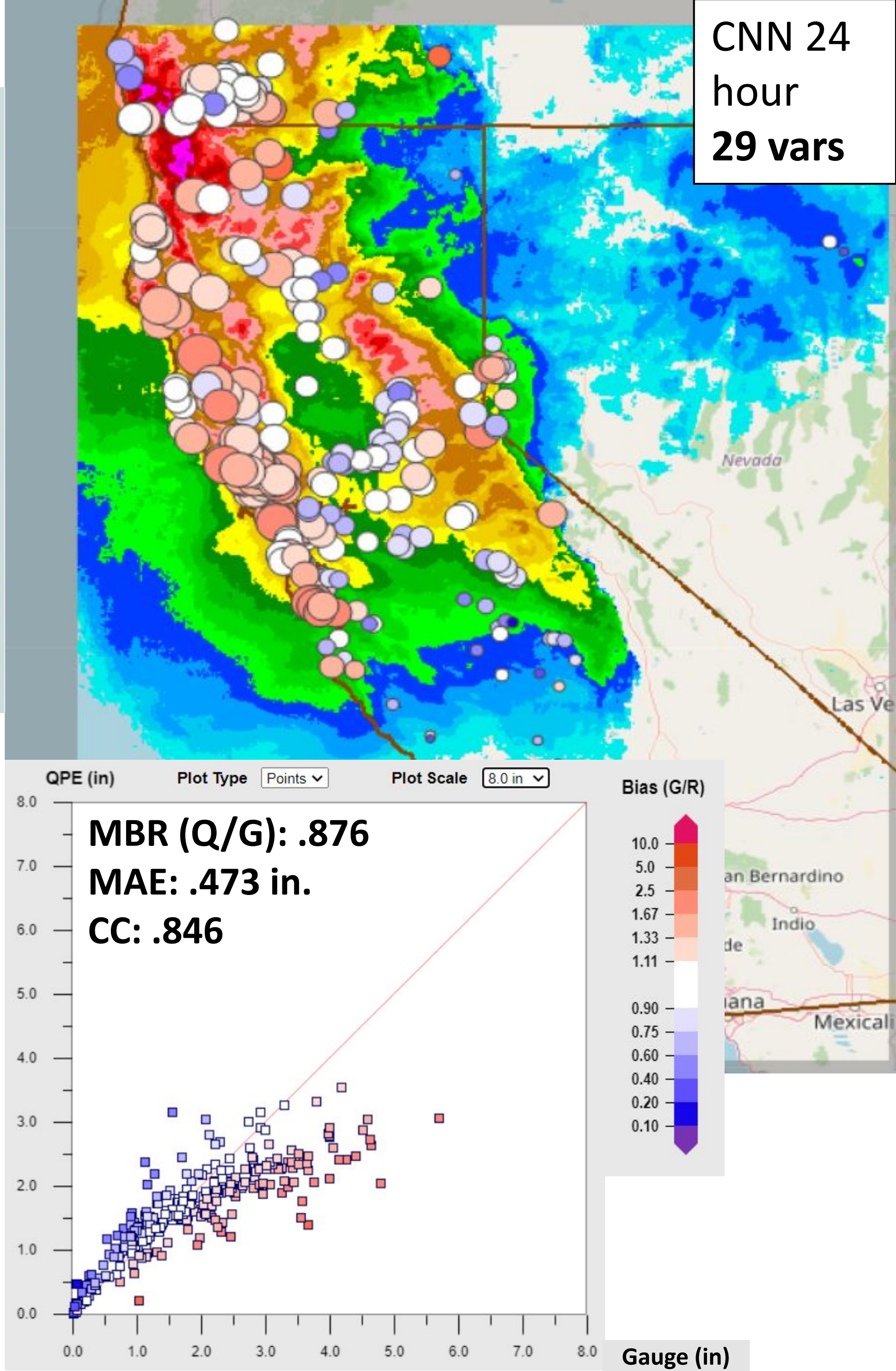
Permutation testing run on validation dataset



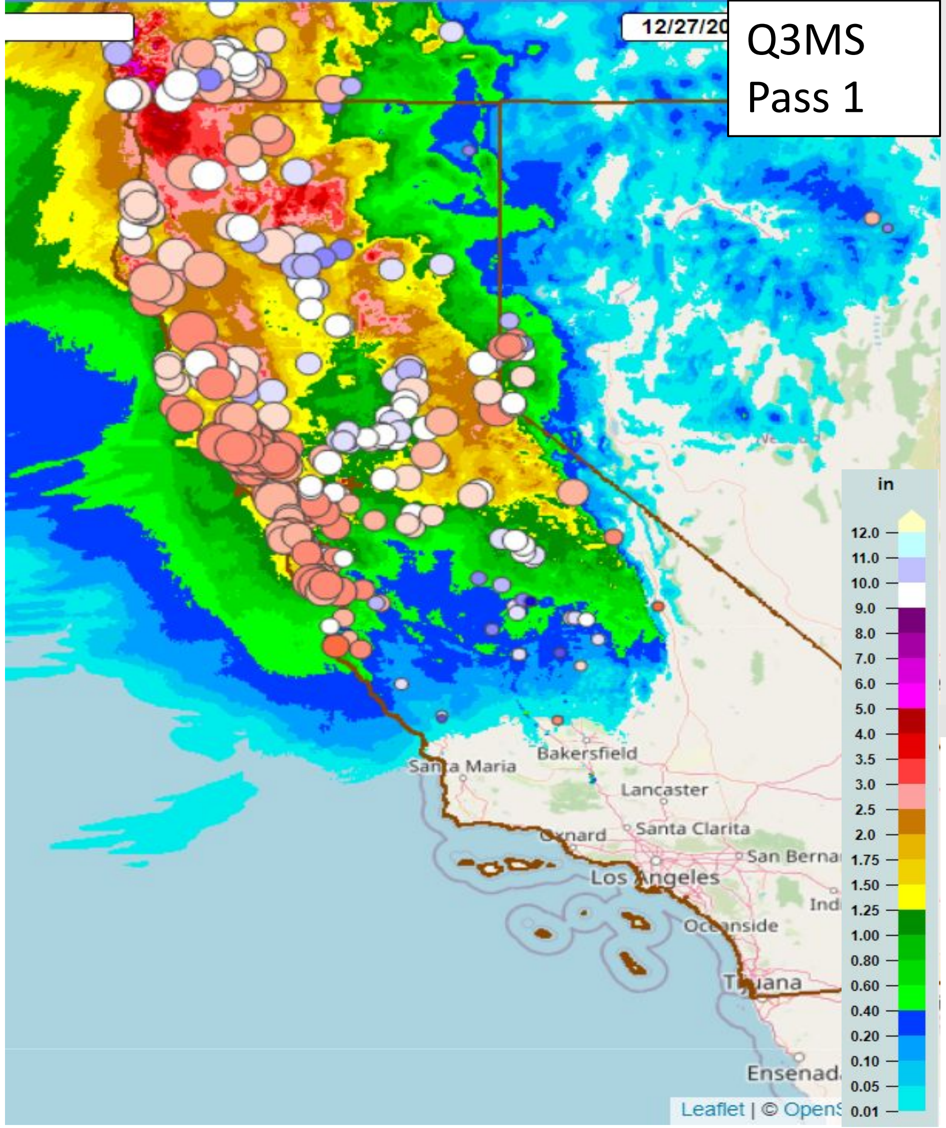
20221227
1500 UTC

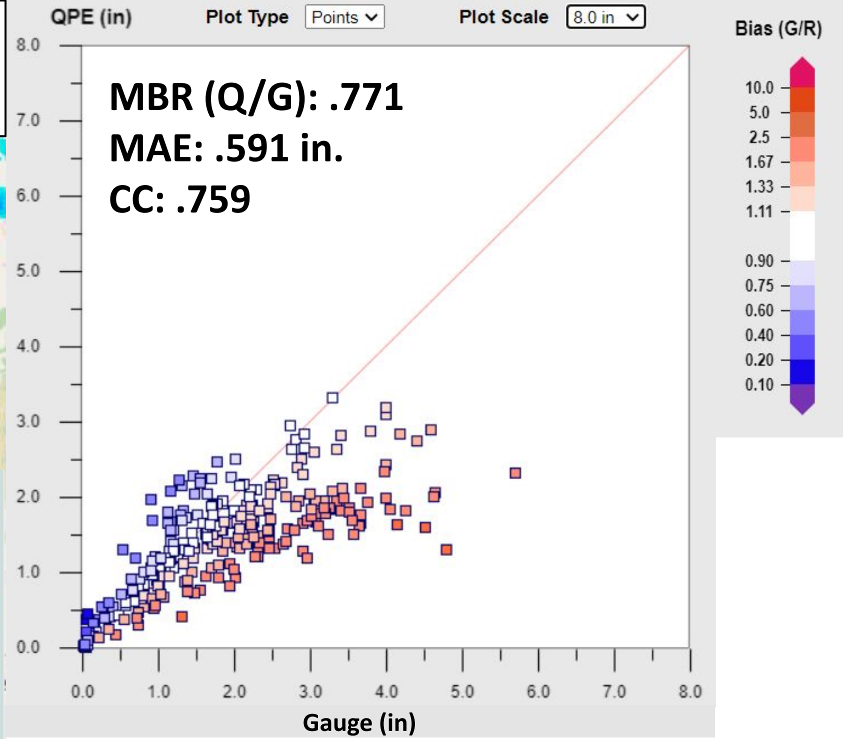
-Current version of the model with 29 input variables tested on recent set of ARs impacting California late Dec 2022 – Jan 2023

-Performance as good or better than Pass 1 of our multi-sensor product for some of the days in this period

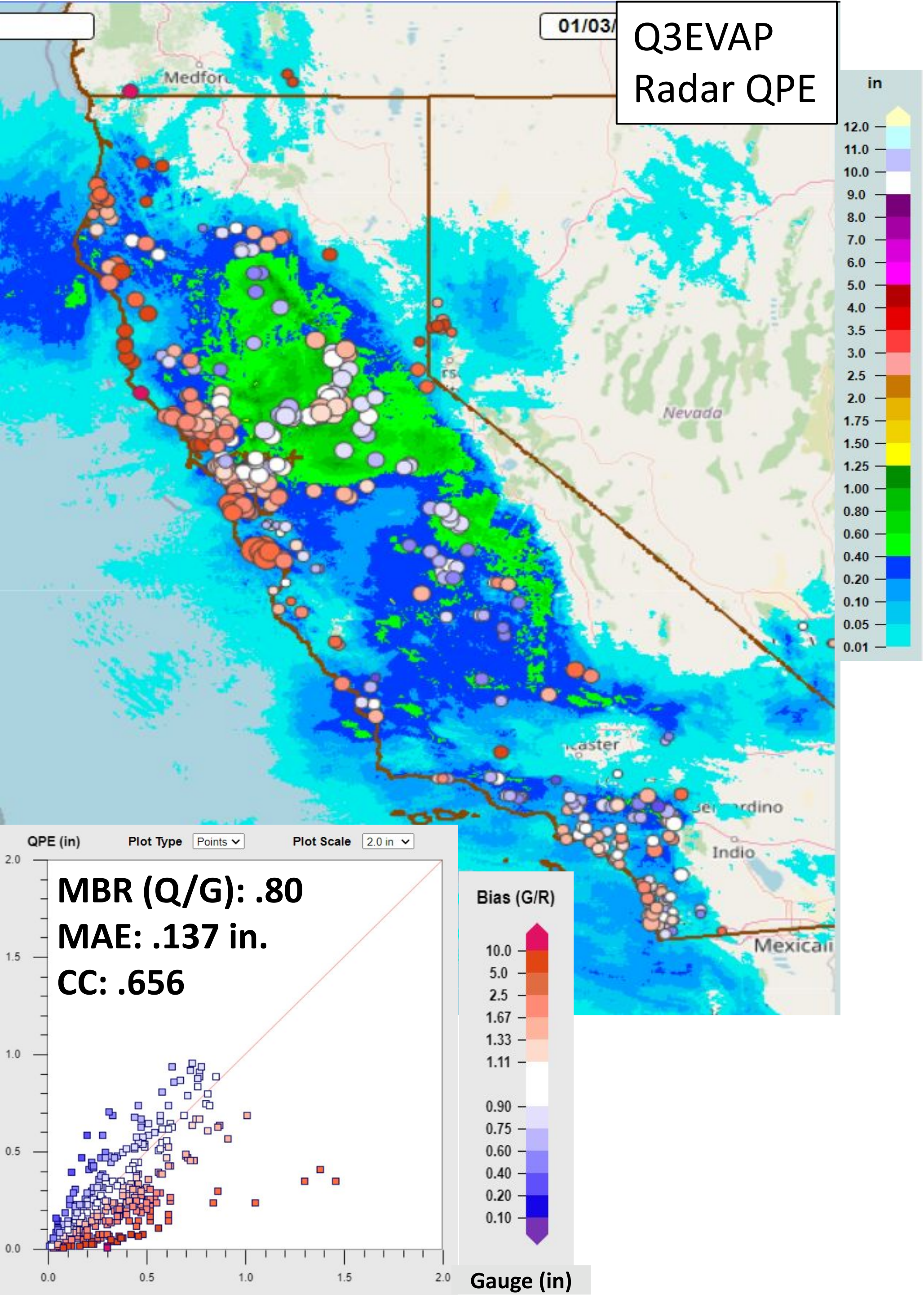


Q3MS
Pass 1





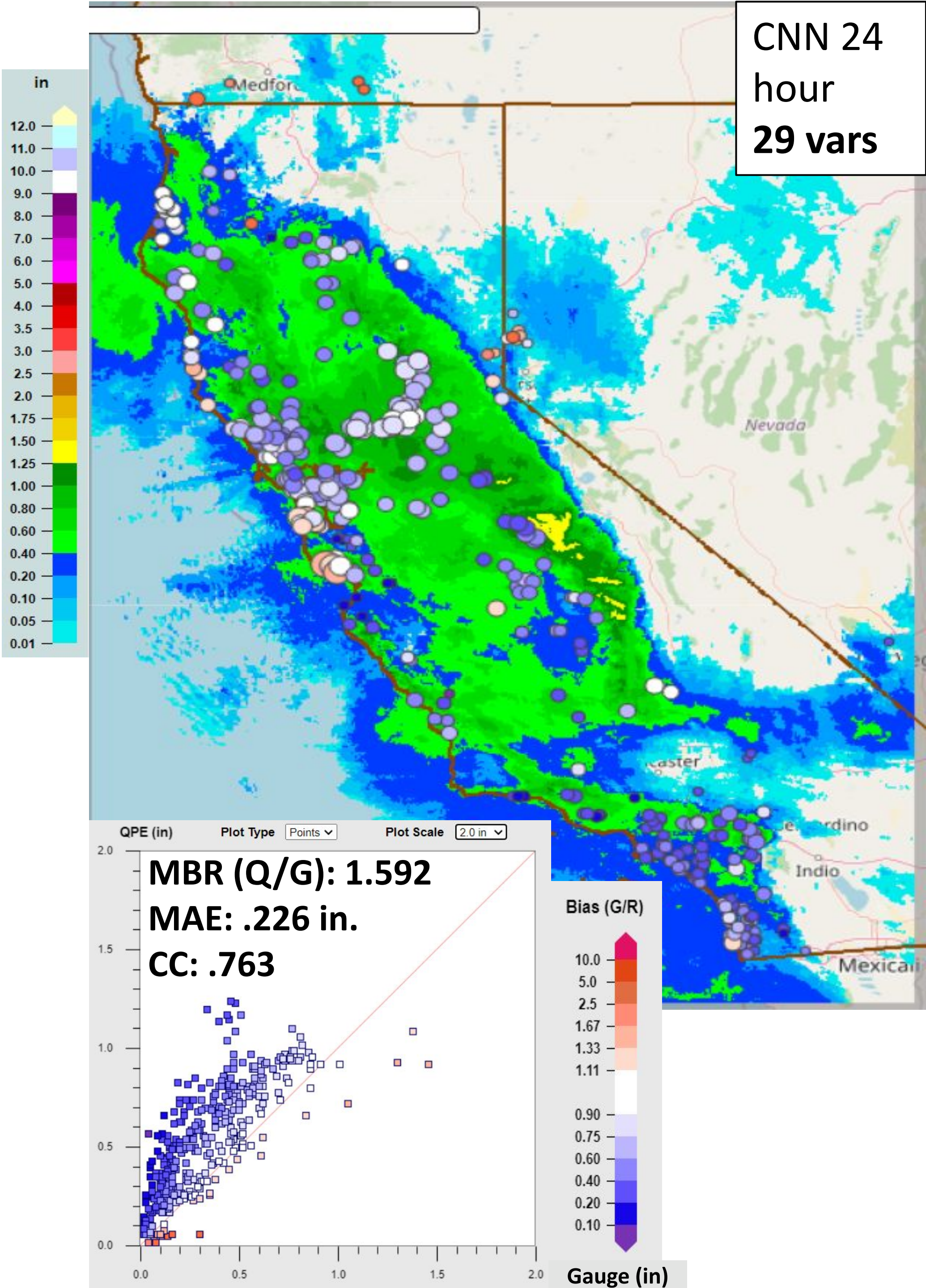
20221227
1500 UTC



20230103
1500 UTC

-Seeing some overestimate bias for lighter events

-Working to mitigate this through use of different models for heavy and lighter precipitation days

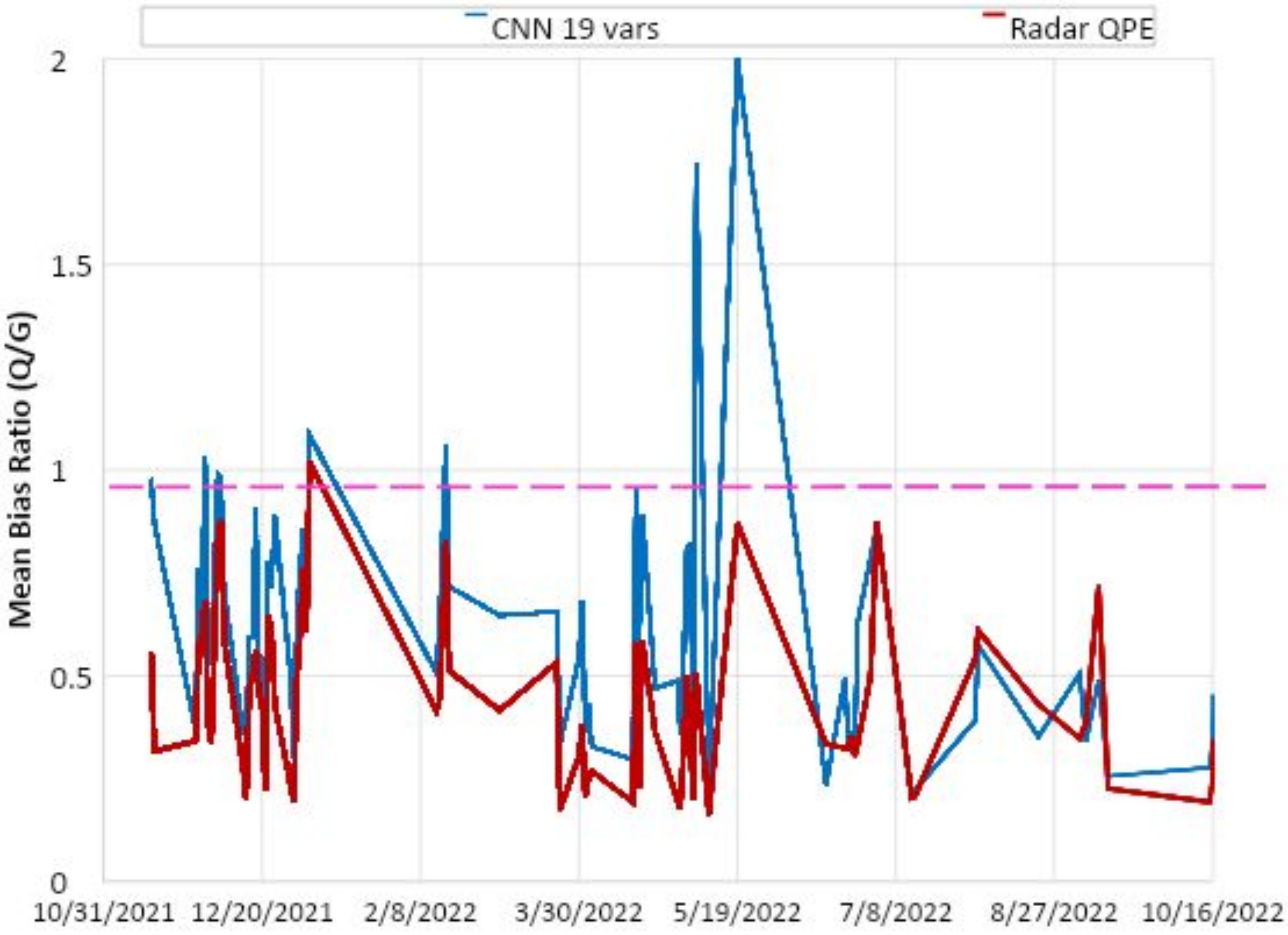


HI Long-Term Statistics

-Overall decreased underestimate bias from the CNN compared to radar-based QPE over the long- term period

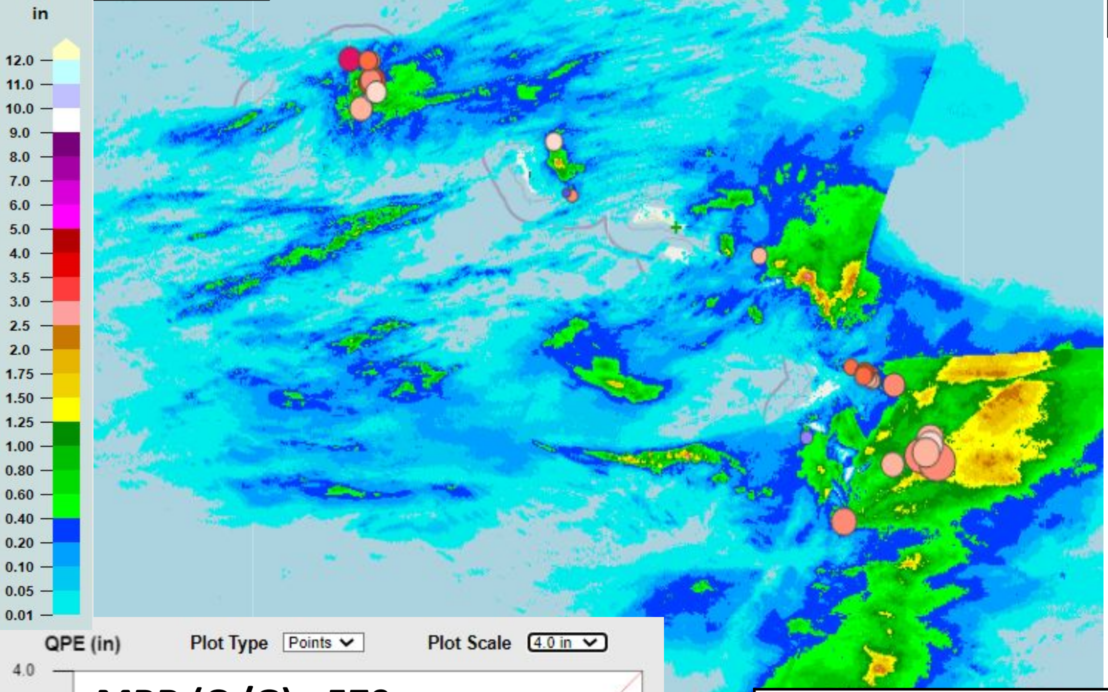
-Slight decrease in fMAE from CNN compared to radar QPE

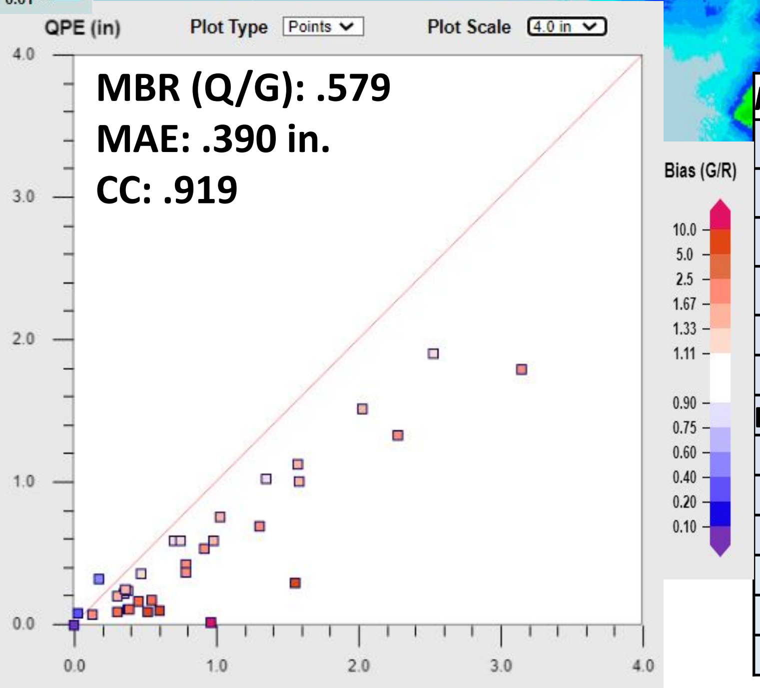
	Q/G Bias	CC	fMAE
Q3EVAP	0.44	0.599	0.641
CNN	0.61	0.597	0.639



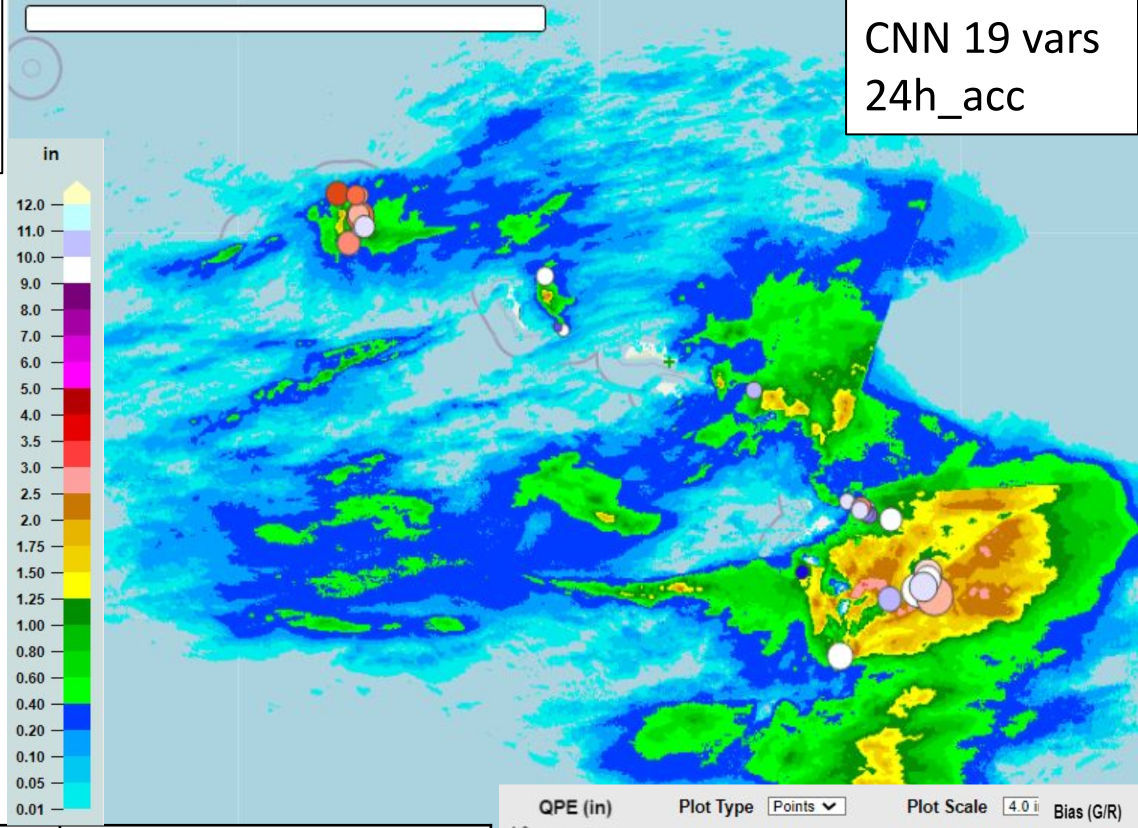
20220417
1700 UTC

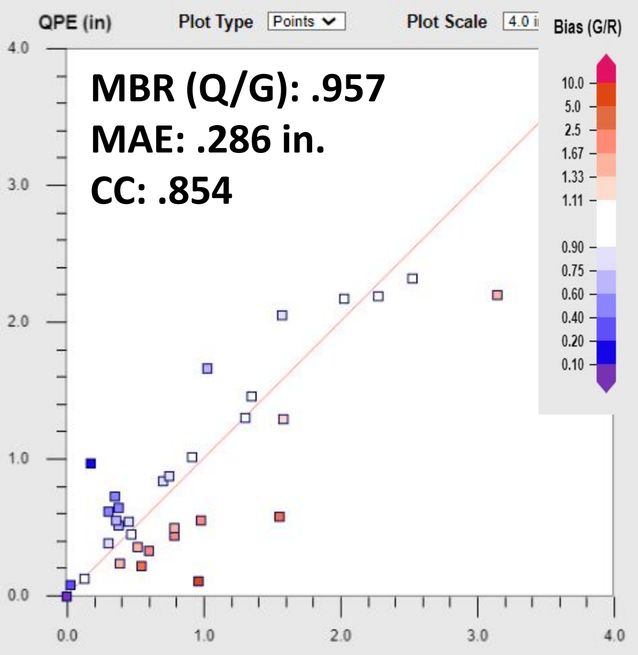
Radar QPE
24h_acc



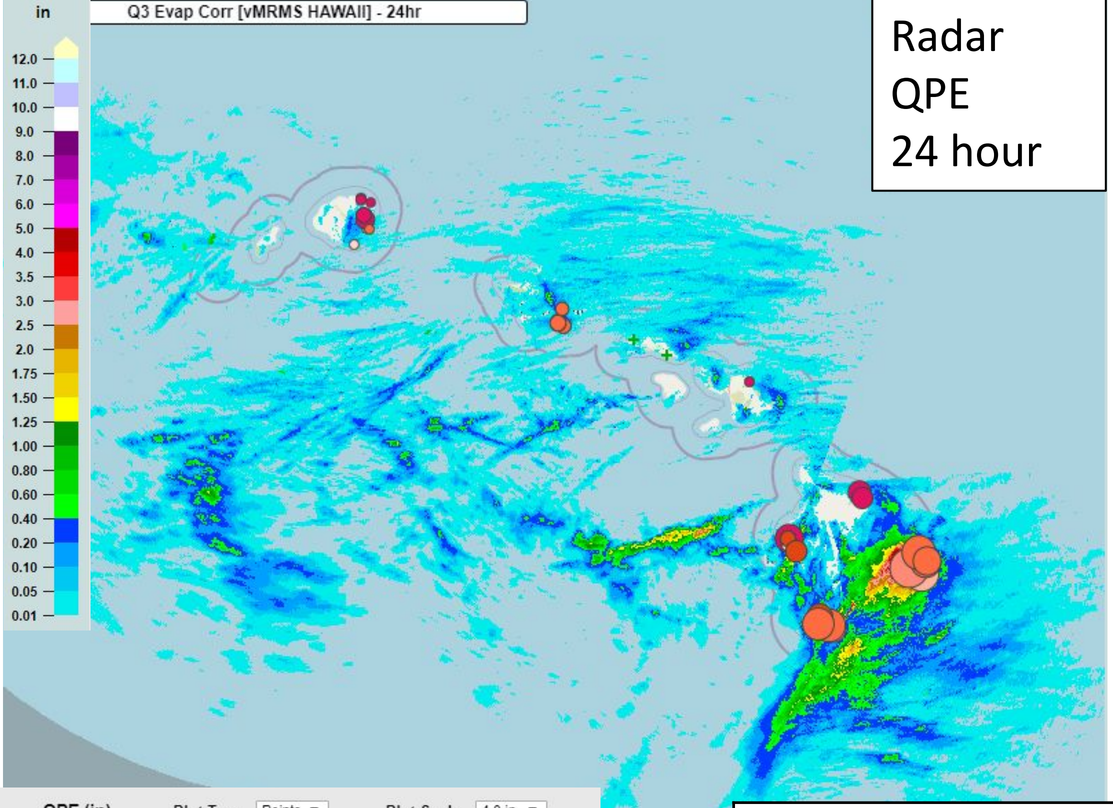


CNN 19 vars
24h_acc

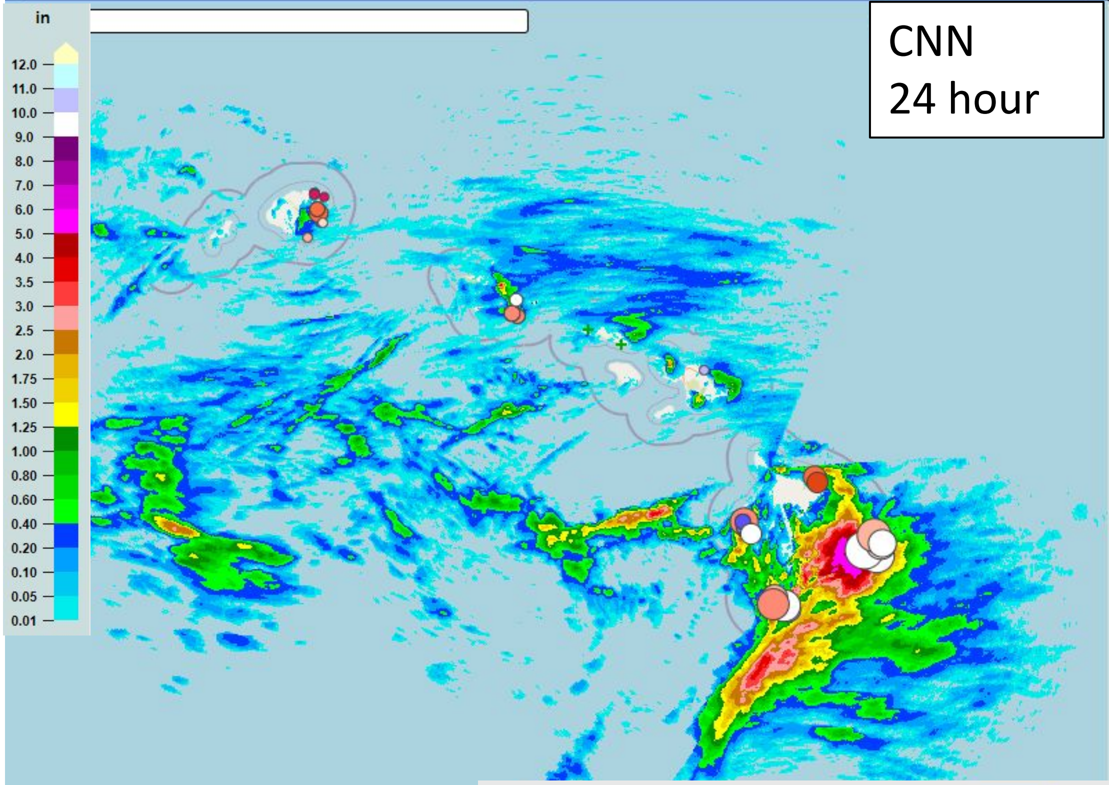


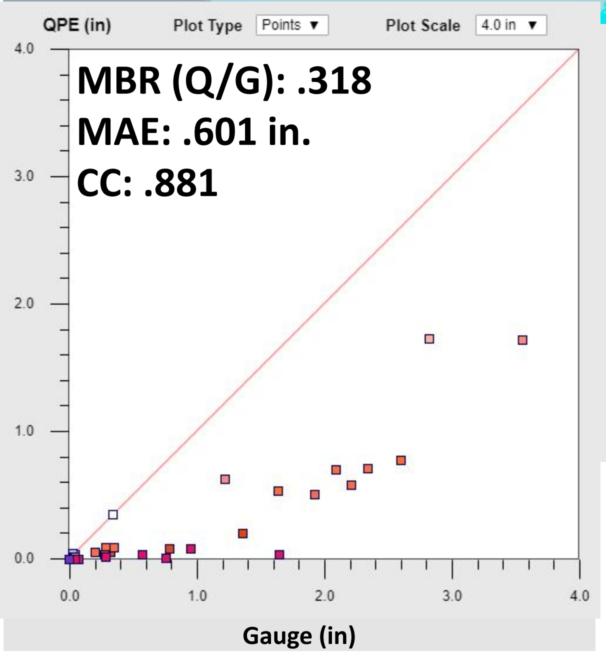


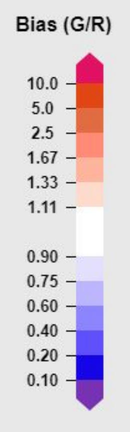
Radar Variables	*Terrain Variables*
Seamless Hybrid Scan Reflectivity (SHSR)	Orographic forcing factor
Vertically Integrated Liquid (VIL)	Mean U wind 850-700 mb
Radar Quality Index (RQI)	Mean V wind 850-700 mb
Composite Reflectivity (CREF)	Latitude
Reflectivity at Lowest Altitude (RALA)	Longitude
Bright Band Bottom (BB_BOTTOM)	Terrain Height
Bright Band Top (BB_TOP)	
Reflectivity at 0C	
Reflectivity at -5C	
Reflectivity at -10C	
Reflectivity at -15C	
Reflectivity at -20C	
SHSR Height (SHSRH)	



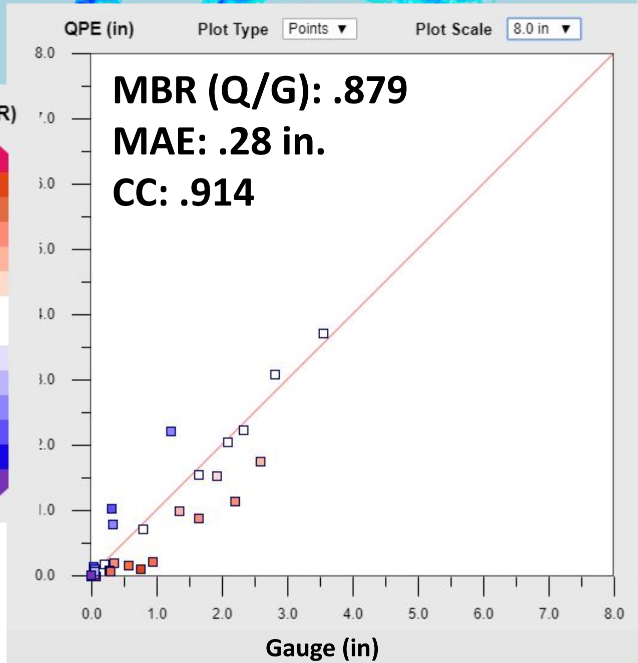
20181009
1700 UTC







Radar Variables	*Terrain Variables*
Seamless Hybrid Scan Reflectivity (SHSR)	Orographic forcing factor
Vertically Integrated Liquid (VIL)	Mean U wind 850-700 mb
Radar Quality Index (RQI)	Mean V wind 850-700 mb
Composite Reflectivity (CREF)	Latitude
Reflectivity at Lowest Altitude (RALA)	Longitude
Bright Band Bottom (BB_BOTTOM)	Terrain Height
Bright Band Top (BB_TOP)	
Reflectivity at 0C	
Reflectivity at -5C	
Reflectivity at -10C	
Reflectivity at -15C	
Reflectivity at -20C	
SHSR Height (SHSRH)	



Summary & Upcoming Work

- CNN is generally improving on the underestimate bias from the radar-based QPE in the western CONUS

  -Model moisture variables important for the further improvement seen relative to the original version using only the 13 radar variables

  -Looking into ways to improve the consistency of the CNN model performance across heavy and lighter precipitation events

- Implementing the optimized version of the CNN model in real-time on the VMRMS system to track to the performance on a daily basis
- Further optimize HI model using lessons learned from western CONUS model
- Expand to full western CONUS domain