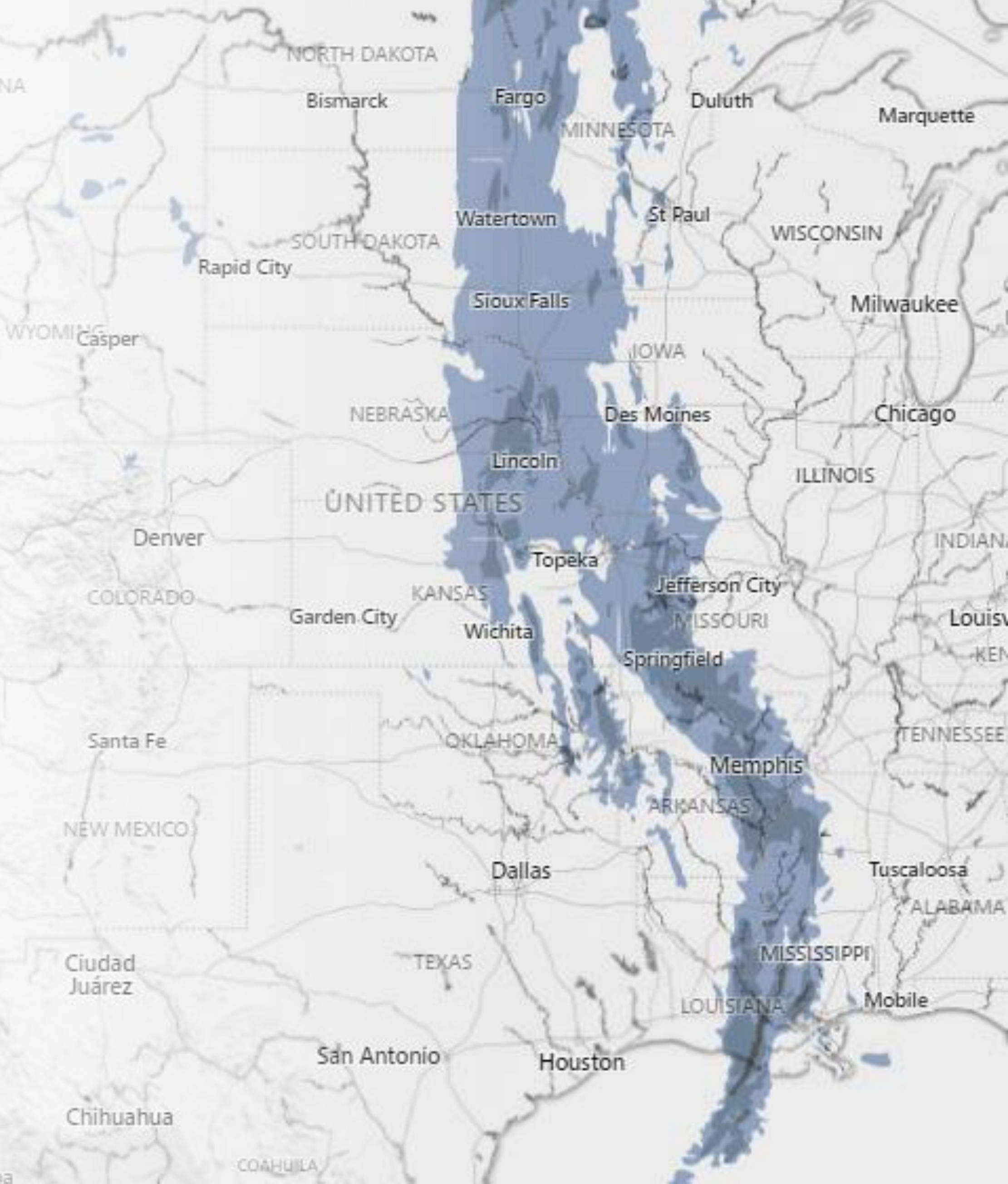
[Microsoft Weather - Live Demo]

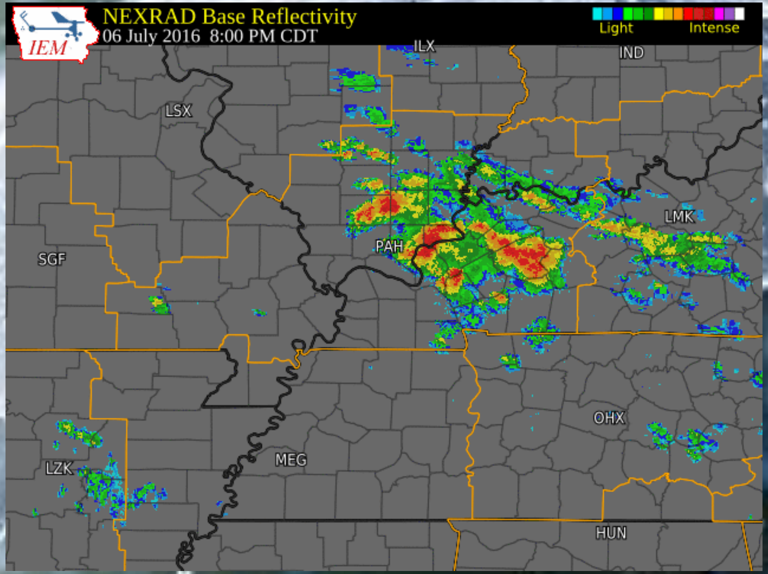
MS-nowcasting: Operational Precipitation Nowcasting with Convolutional LSTMs at Microsoft Weather

Sylwester Klocek
Haiyu Dong
Mathew Dixon
Panashe Kanengoni
Najeeb Kazmi
Pete Luferenko
Zhongjian Lv
Shikhar Sharma
Jonathan Weyn
Siqi Xiang



Motivation

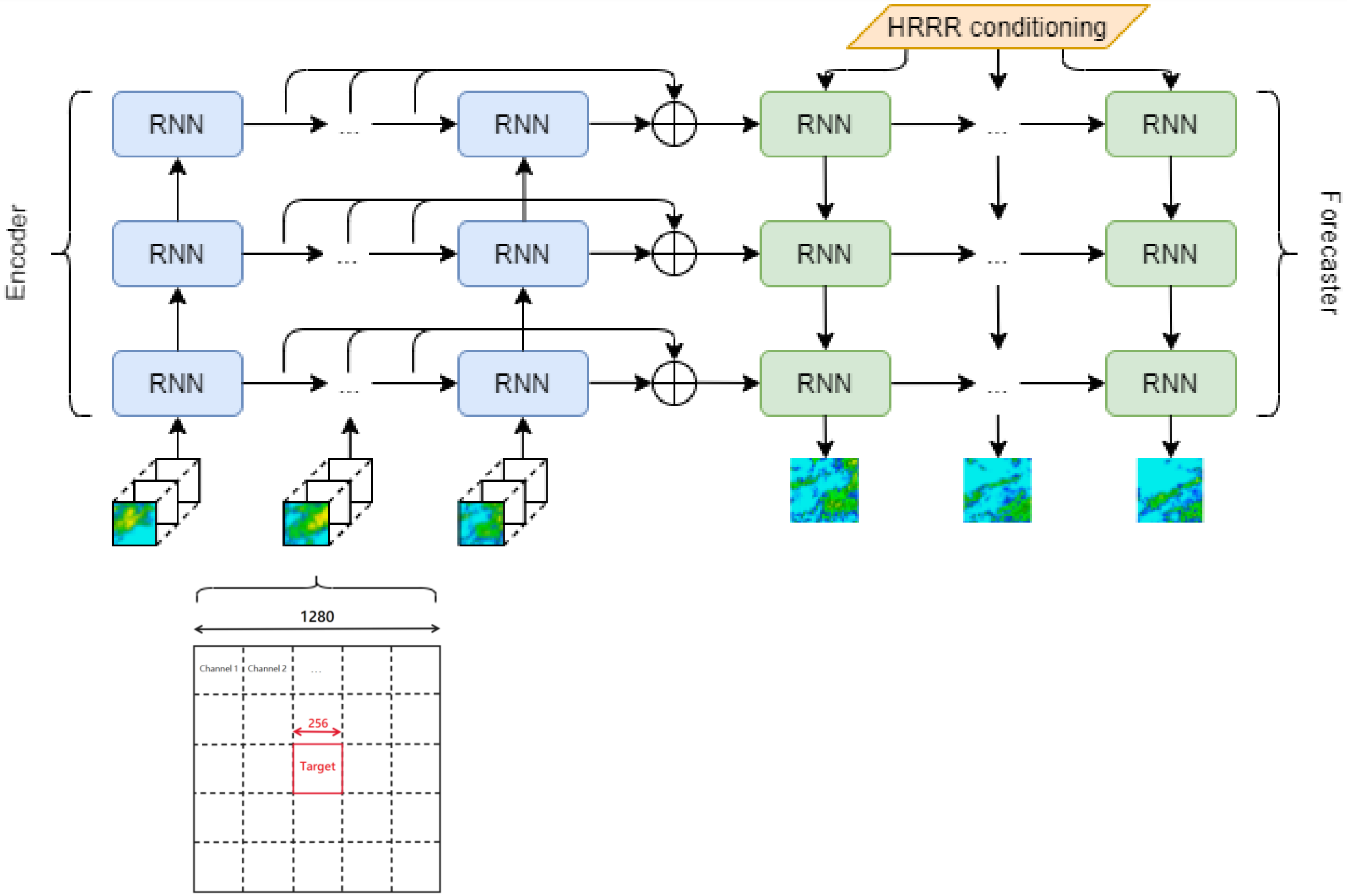
- Accurate short-term precipitation forecasts (nowcasts) are vital for many peoples' daily lives
- Nowcasts can also help in emergency preparedness for extreme weather events
- Extreme precipitation events are expected to occur more frequently in a warming climate, increasing the need for accurate prediction tools
- Traditional physics-based numerical weather prediction models such as the High-Resolution Rapid Refresh (HRRR) suffer from lack of model spin-up at very short lead times and loss of deterministic predictability beyond a few hours
- Despite extensive research work, few operational products based on deep learning exist



Iowa Environmental Mesonet via NWS (Summary of the Flash Flood event of July 6-7 (weather.gov))

Model architecture

- Adapts the Convolutional LSTM architecture introduced by Shi et. al. (NeurIPS 2015) with some improvements
- Divide large input receptive field into 5x5 segments stacked in the channel dimension
- Condition forecaster cells on reflectivity forecasts from HRRR, which helps the model learn formation and dissipation of precipitation
- Use **all** hidden encoder states to produce forecaster hidden states
- Stochastic Weight Averaging helped in generalization and training stability



Metrics

- Adding HRRR conditioning and the large input receptive field (LV) improve the model metrics
- Model with both performs best at all lead times

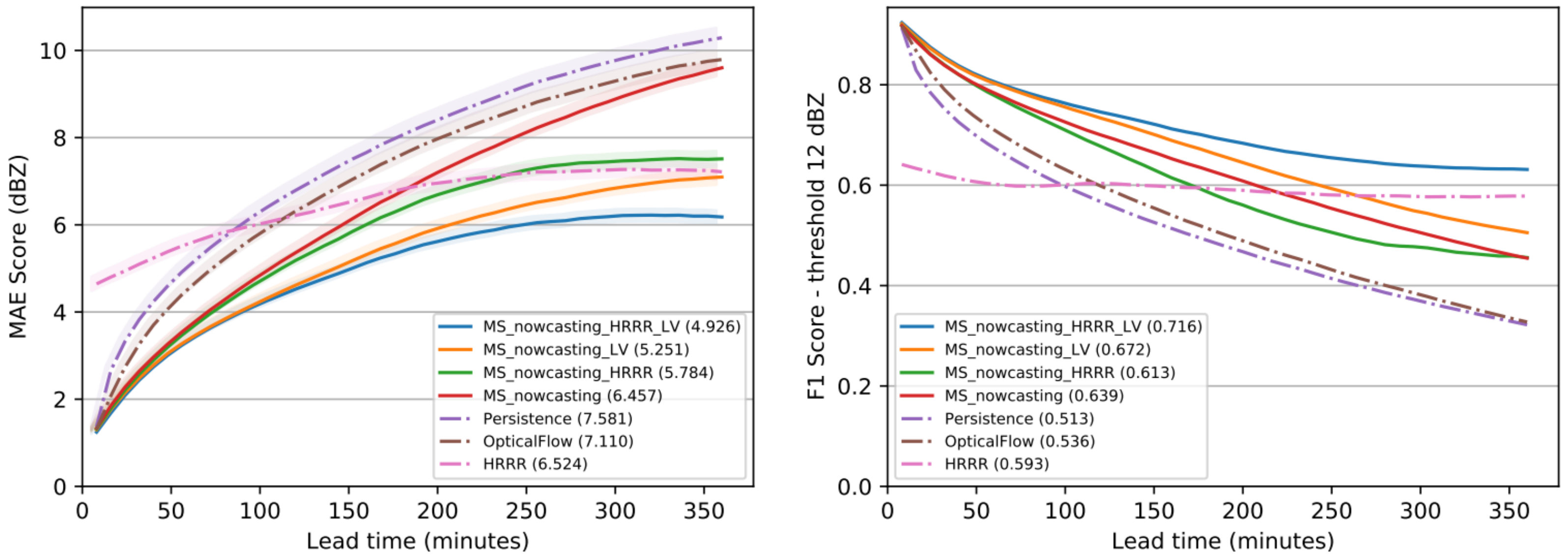
Table 1: Metrics averaged over lead times 0–2 hours

Model	MAE	F1 (12 dBZ)	MS-SSIM	PSNR
Persistence	4.88	0.686	0.5156	21.21
Optical Flow	4.39	0.720	0.5527	21.83
HRRR	5.54	0.610	0.4634	20.01
MS-nowcasting	3.64	0.788	0.6419	23.23
+ HRRR	3.55	0.780	0.6510	23.29
+ LV	3.28	0.808	0.6678	23.86
+ HRRR + LV	**3.23**	**0.813**	**0.6721**	**24.03**

Table 2: Metrics averaged over all lead times 0–6 hours

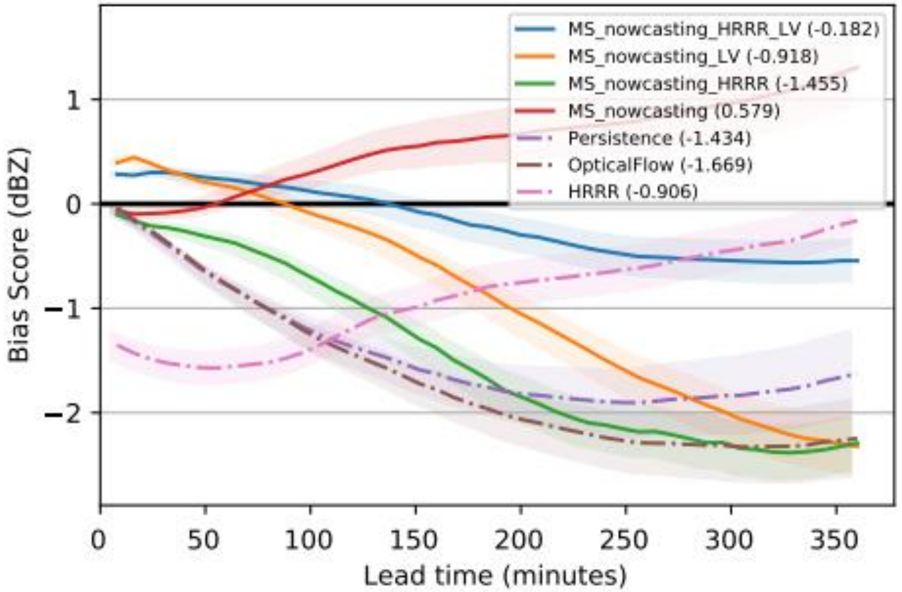
Model	MAE	F1 (12 dBZ)	MS-SSIM	PSNR
Persistence	7.58	0.513	0.3626	16.98
Optical Flow	7.11	0.536	0.3780	17,38
HRRR	6.52	0.593	0.3905	17.72
MS-nowcasting	6.46	0.639	0.4788	18.78
+ HRRR	5.78	0.672	0.4952	19.10
+ LV	5.25	0.613	0.5137	19.68
+ HRRR + LV	**4.92**	**0.716**	**0.5291**	**20.16**

Metrics



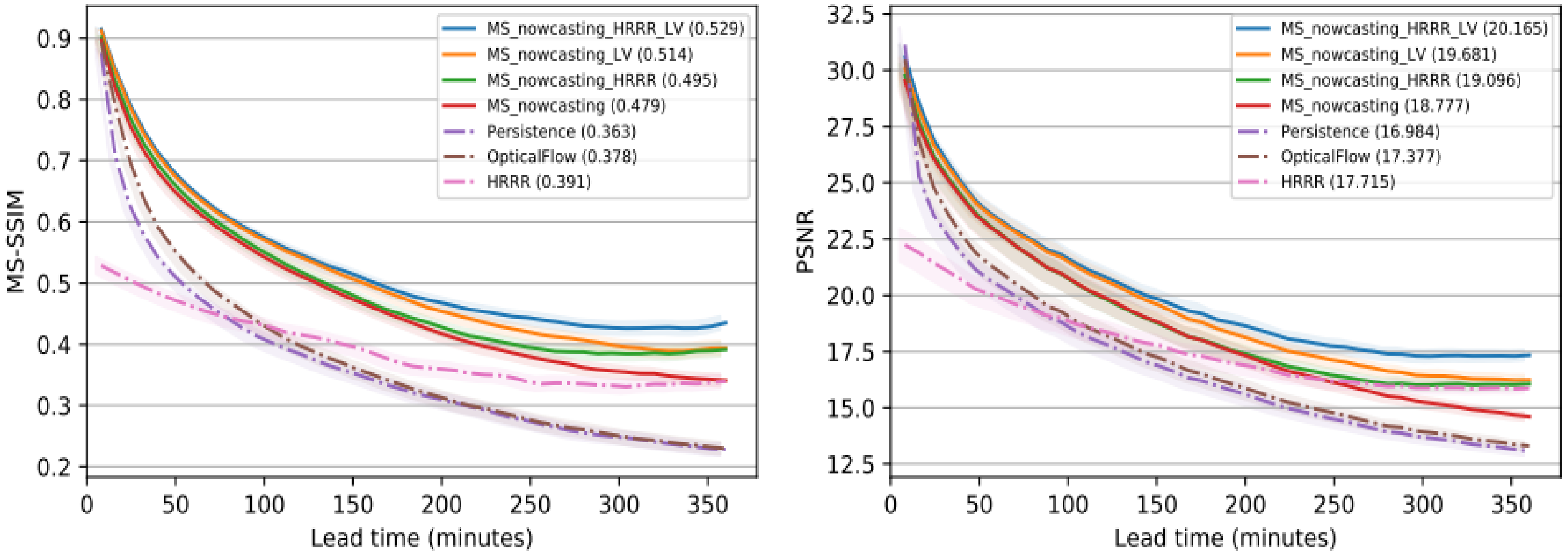
MAE and F1 metrics over lead times for the ablation study. F1 threshold of 12 dBZ corresponds to a rainfall intensity of about 0.2 mm/h. Numbers in parentheses in the plot legends are the average metric value over all lead times.

Metrics



Bias over lead times for the ablation study. The bias metrics for the models show that our baseline model (MS-nowcasting) had an overprediction bias. While MS-nowcasting +HRRR +LV improved other quality metrics, it did lead to an underprediction bias, which, in absolute terms, is still less than the overprediction bias of MS-nowcasting.

Metrics



MS-SSIM and PSNR metrics over lead times for the ablation study. These metrics give a measure of image quality of the predictions relative to ground truth.

Operations

- Our design choices allow the model to run on a single GPU
- Using only 8 GPUs in production, we can produce nowcasts over the entire US in under 2 minutes from **radar availability**
- End-to-end, high-quality nowcast maps are available to consumers within about 6 minutes from **radar observation**

t + 0 min

t + 32 min

t + 64 min

t + 96 min

[Live Demo]