

Cloud masking without thermal infrared bands: a neural network approach based on MODIS

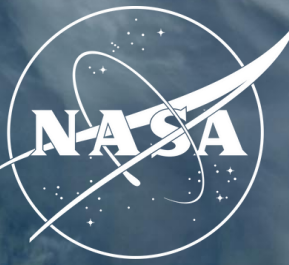
Andrew M. Sayer^{1,2}, Ian T. Carroll^{1,2}, Taha Al-Nufaili³, Xingyan Li¹, Jianwu Wang¹, P. Jeremy Werdell²,

1: University of Maryland, Baltimore County, MD, USA; 2: NASA Goddard Space Flight Center, MD, USA;

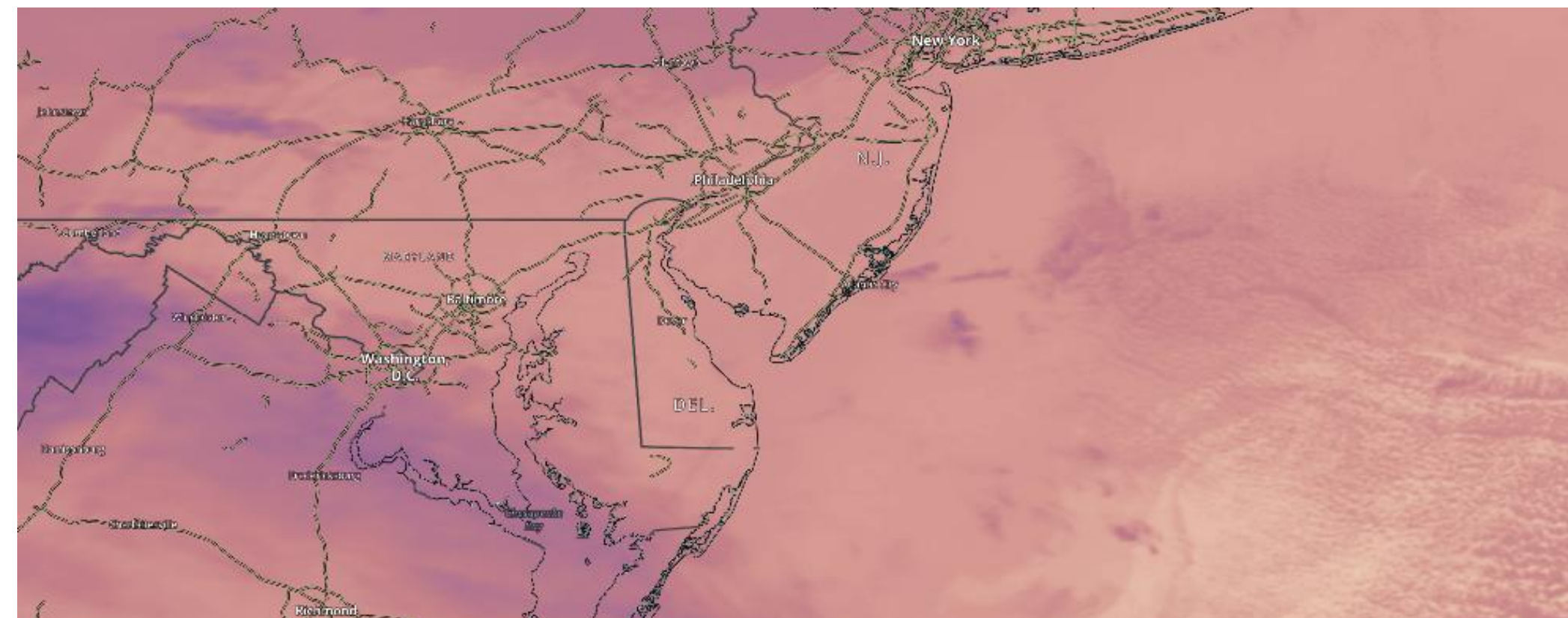
3: formerly with Science Systems & Applications, Inc, Lanham, MD, USA

andrew.sayer@nasa.gov

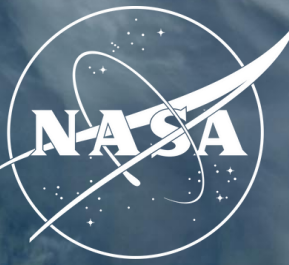
Cloud masking is the first step in most satellite retrieval processing



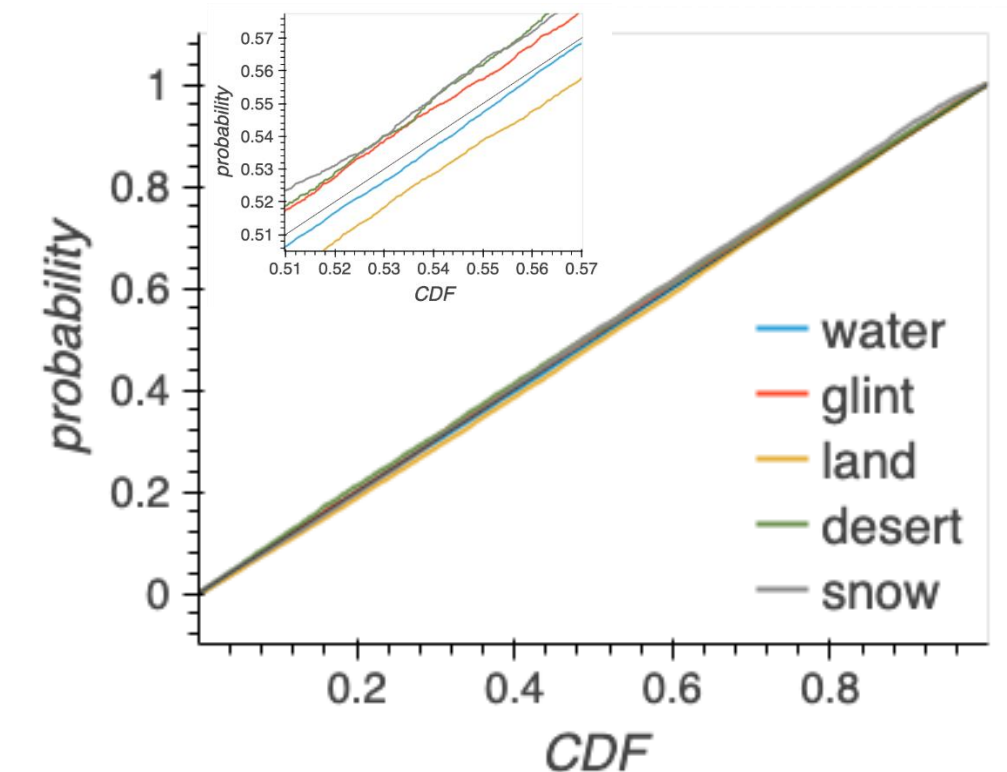
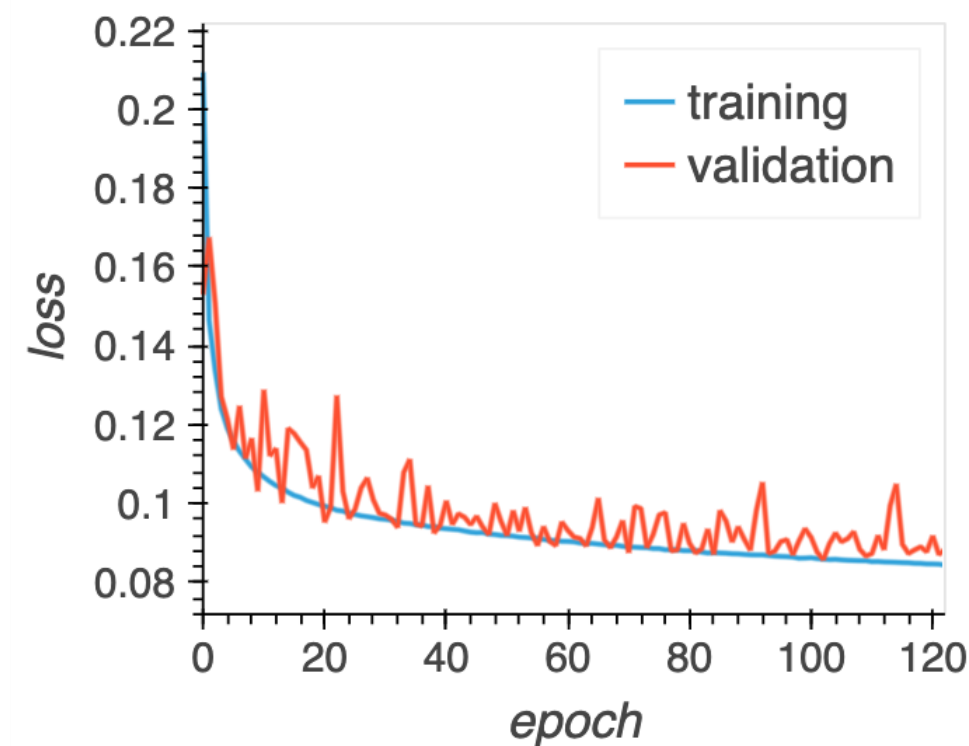
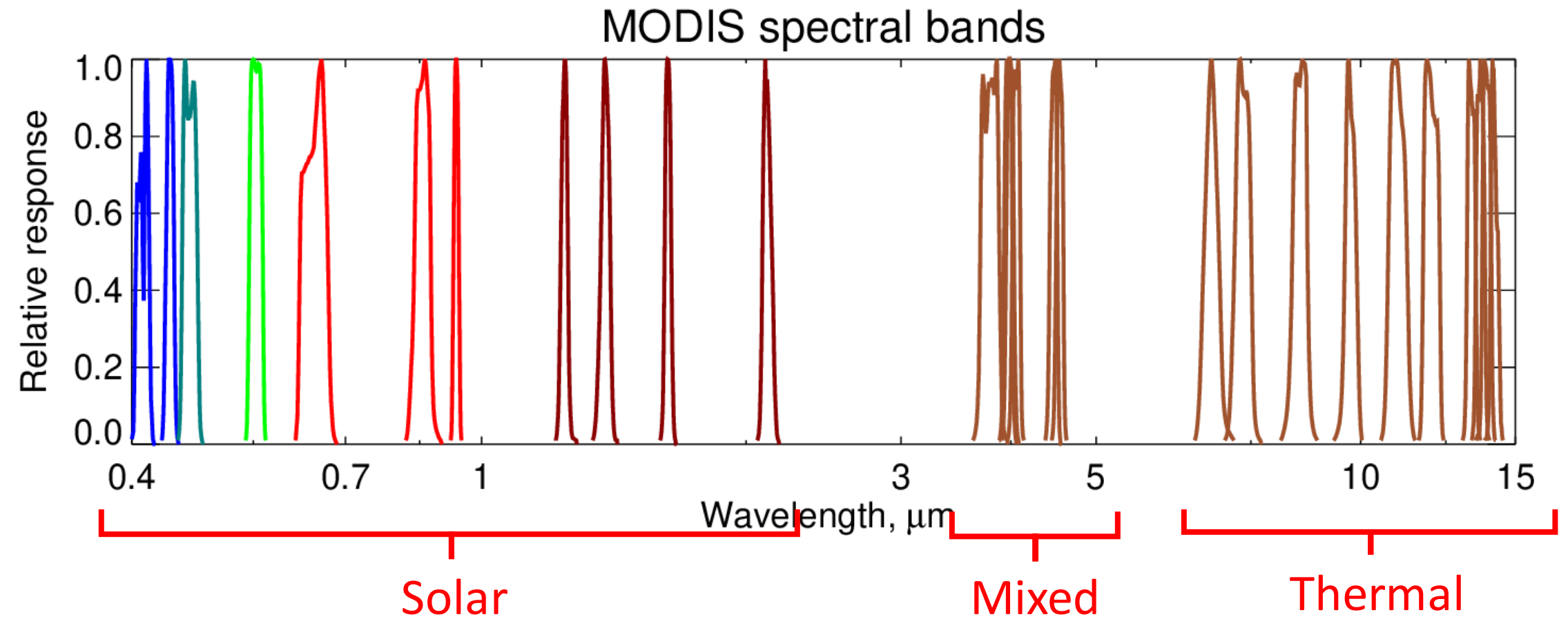
- Satellite cloud masking algorithms (e.g. MODIS) often use both solar and thermal infrared bands
 - Clouds are (generally) bright, white, and cold
 - ... but many satellite imagers lack thermal bands (e.g. PACE OCI)
- How much skill in the cloud mask could be obtained if we didn't have the infrared measurements?
 - Try and reproduce the MODIS cloud mask without using thermal bands



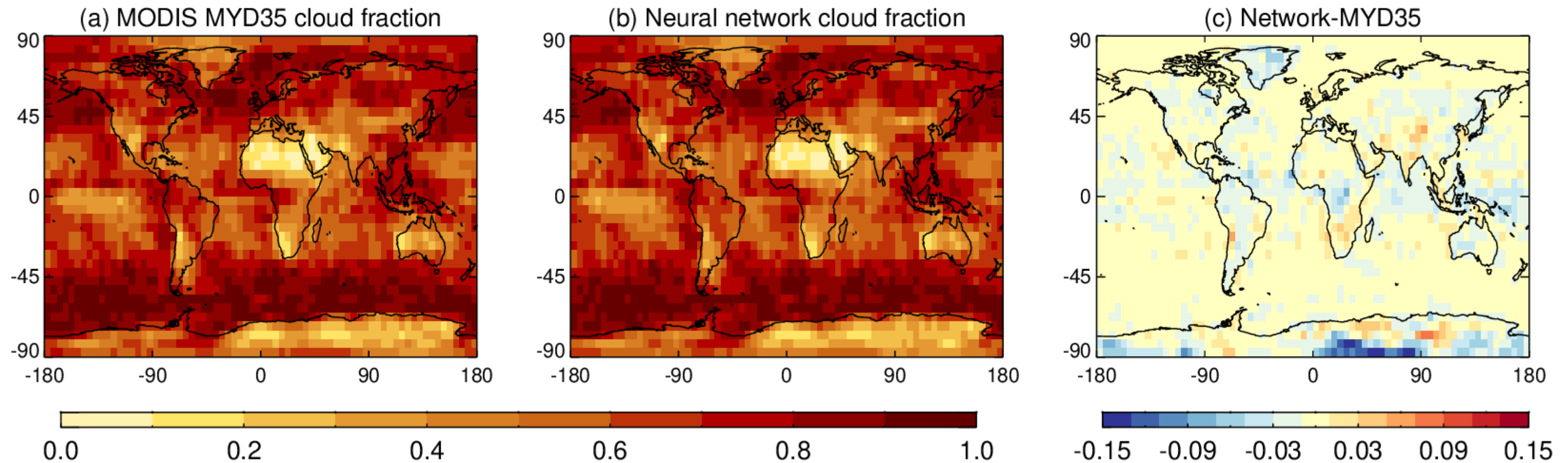
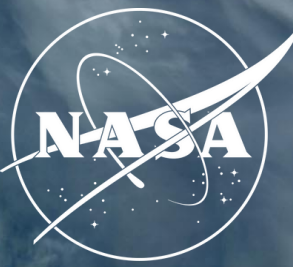
Train a network based on the full MODIS cloud mask, but only provided with reflective solar bands



- Train a neural network to the target of the full MODIS (solar plus thermal) cloud mask
- Inputs
 - 0.1% random sample of MODIS Aqua pixels from 1st and 15th of each month in 2015
 - $n = 5257292$, random train/val/test split 90%/8%/2%
 - Features:
 - 11 MODIS solar bands
 - 3 solar/view angles
 - Reflectance standard deviation in 3 bands
 - Surface type classification
- Output:
 - Binary cloudy/not cloudy mask
- Network:
 - Multi-layer perceptron with 2^9 [9, 8, 7, 7, 7, 5] nodes and ReLU activation
 - Categorical cross-entropy loss computed from batch normalized logits
 - Adam optimizer, learning rate of $3e-4$, 2^9 batch size
 - Tensorflow 2.14.1



Most of the skill in the MODIS cloud mask can be reproduced without the thermal bands



- Further evaluate on unseen MODIS data
- Solar-only neural network agrees with full MODIS cloud mask about 96% of time
- So, likely applicable to similar imagers without thermal bands (domain adaptation)

Neural network

MODIS

	Clear	Cloudy
Clear	32.1%	2.1%
Cloudy	1.5%	64.2%