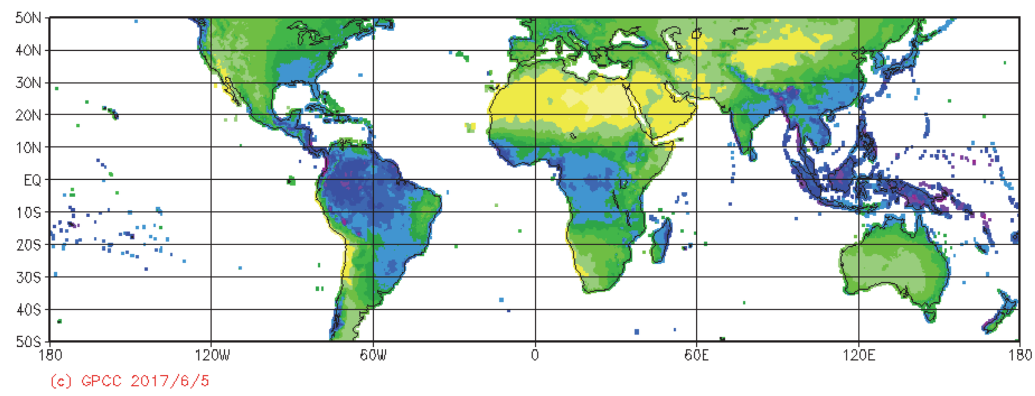
# Wide-area precipitation estimation from satellite imagery

Paul M. Aoki

## Motivation

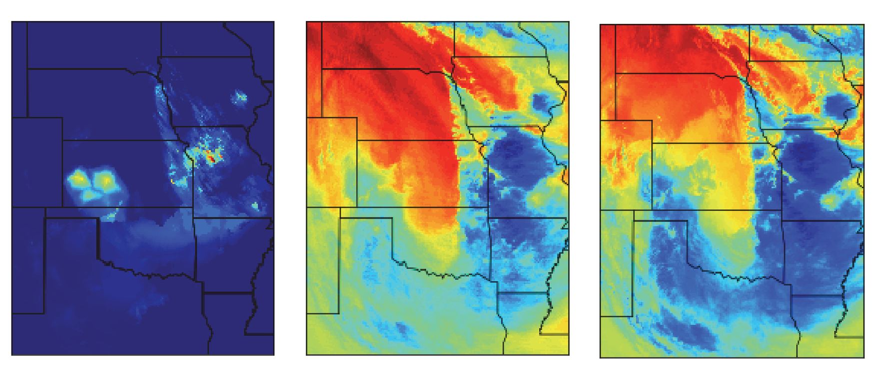
Global precipitation models are crucial in areas ranging from climate change research to wireless network planning in developing countries (my own interest). Raster data from geostationary satellites is still the only whole-globe, near-real-time, multi-decadal observational data record we have, but it is not a direct measurement of rainfall presence (*detection*) or quantity (*estimation*).



*WMO GPCC annual rainfall estimate, 2016*

## Problem statement

Models for rainfall detection and estimation from infrared satellite imagery can be constructed using supervised learning. Rain/no-rain (R/NR) labels and rainfall values are not available globally, but quantitative precipitation estimate (QPE) data is available for select regions based on weather radar networks.



*NWS QPE vs. GOES IR band and WV band (U.S. Great Plains, 2013-02-21 1600Z) Cold (blue) imagery at right corresponds to intense rainfall at left.*

The 2017 state-of-the-art (SOTA) is based on predicting rainfall at the center pixel of every 15x15 patch (1.2° x 1.2°) from each hourly image of the study region and I retain this approach here.

## Data

**Datasets**

"Ground truth": U.S. NWS QPE grid (hourly, 0.4°)
Comparable operational system: UCI PERSIANN-CCS (hourly, 0.4°)
Images: U.S. NOAA GOES infrared imagery (hourly, 0.4°)

**Study region**

U.S. Great Plains region, 30-45°N, 105-90°W (6742,450,450,2)
948 M (30,30,1) patches = 6.8 TB @ 0.4°
238 M (15,15,1) patches = 214 GB @ 0.8°
Train/validation set: Winter/Summer 2012, 4.8% rain pixels
Test set: Winter/Summer 2013, 6.3% rain pixels

## Method

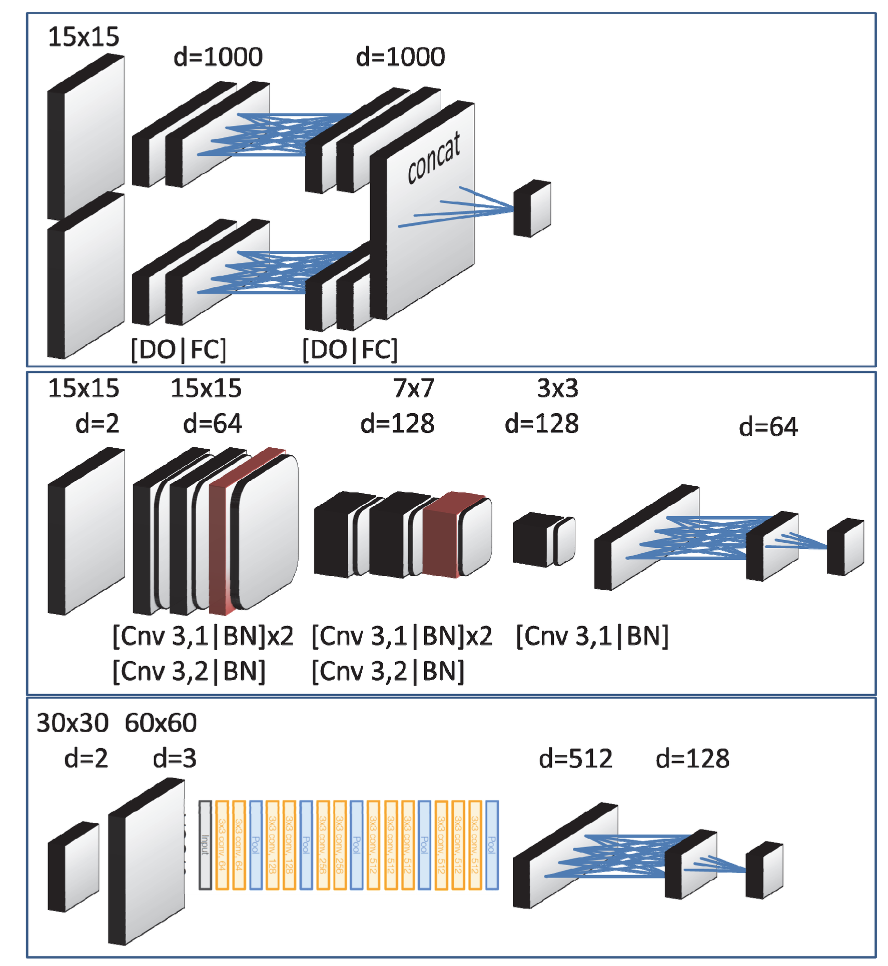
**Implementations** in Keras/TensorFlow:

- Keras "generator" for patches (≈ data augmentation pipeline)
- Several models, including the ones compared here:

***SOTA (SDAE):***
*3-layer fully-connected model w/ SDAE unsupervised pre-training*
- *1 band model*
- *2-band ensemble (concat)*

***Simple CNN:***
*Modified "Simplicity" all-CNN model (+ dense top layers)*

***Transfer CNN:***
*Pre-trained VGG16 model (+ dense top layers)*




**Intuition/motivation:** The SOTA approach does not use any convolutional layers, and so may miss out on any translation-invariant features that affect probability of rain. It is also relatively shallow, and increased depth could improve hierarchical representation.

**Detection measures** : I apply standard measures – binary (sigmoid) cross-entropy for loss, binary accuracy and *critical score index* (CSI) for evaluation metrics.

$$CSI = \frac{TP}{TP+FP+FN} \qquad \text{(c.f. } F_1 = \frac{2TP}{2TP+FP+FN}\text{)}$$

(CSI is more common in meteorology; $F_1$ is not used.)

**Estimation measures** : I apply mean squared error (MSE) as both loss and evaluation metric.

**Class imbalance**: The class distribution is unbalanced; to improve CSI, balancing by oversampling of positive class and by class weighting have been examined.

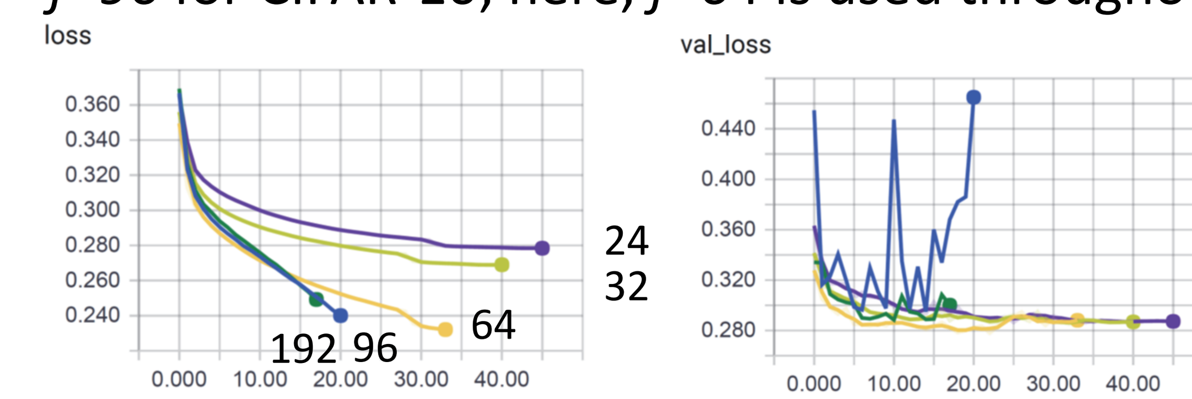

**References (see report)**

1. Detection SOTA. Tao et al., *J. Hydrometeorology*, to appear.
2. Estimation SOTA: Tao et al., IEEE CEC 2016.
3. NOAA GOES. GOES N Series Data Book (Rev. D), 2009.
4. NWS QPE. Lin & Mitchell, Proc. 2005 AMS Hydrology Conf.
5. SDAE. Vincent et al., *JMLR* 2010.
6. Simplicity. Springenberg et al., ICLR 2015. https://arxiv.org/abs/1412.6806
7. VGG16. Simonyan & Zisserman, ICLR 2015. https://arxiv.org/abs/1409.1556


## Selected findings

**Tuning Simple CNN**

In addition to the usual tuning and small-n overfitting tests, I tuned filter depth *f* (24, 32, 64, 96, 192). The Simplicity work used *f*=96 for CIFAR-10; here, *f*=64 is used throughout.

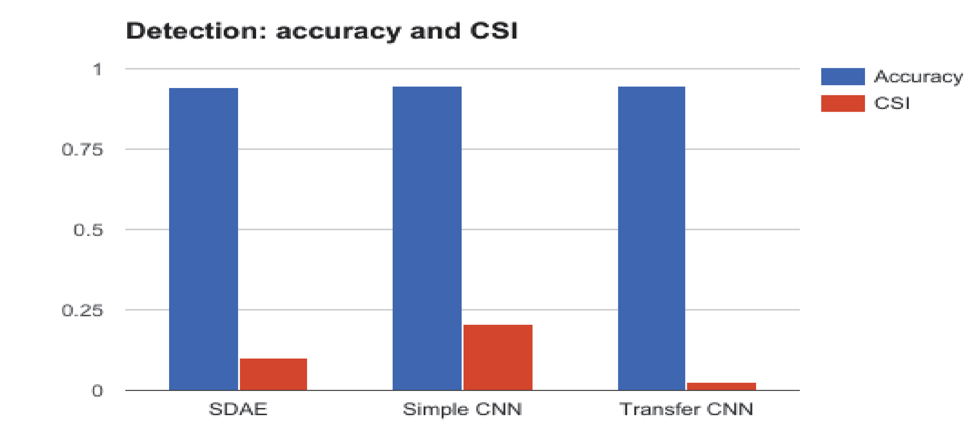



*Filter depths > 64 overfit on this dataset*

**Detection comparison**

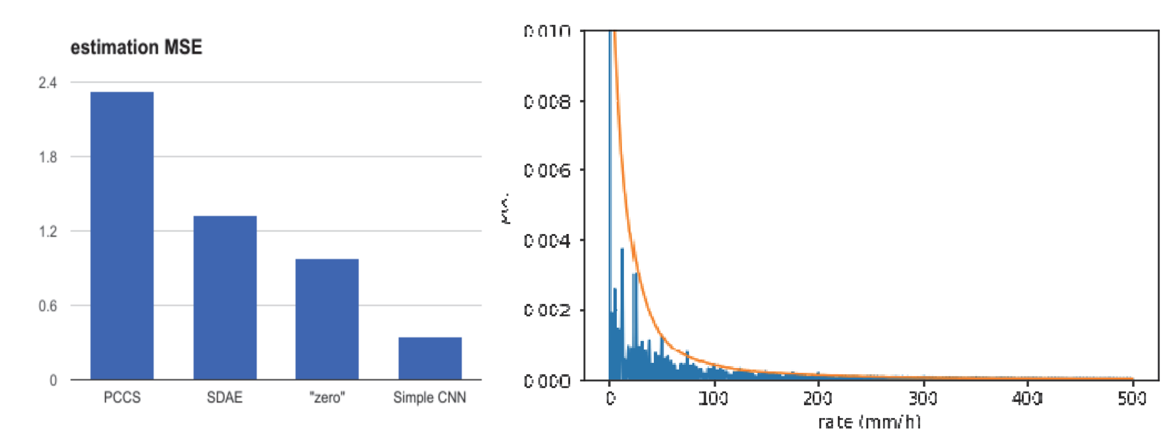
It is impossible to determine Tao et al.'s detection accuracy with certainty, but combining statistics from two of their papers gives a best guess of 94.7%. (the test set majority-class frequency).

*All models achieved majority-class validation and test set accuracy w/o class weighting.*




Tao et al. report a test set CSI of 0.306. So far, my accuracies are substantially lower (e.g., 89.3% SOTA, 93.7% Simple CNN) with class weighting/balancing that achieves comparable CSI.

**Estimation comparison**




*Simple CNN achieves lower MSE (left) while retaining a reasonable frequency distribution (right)*

SOTA attained an MSE of 1.32, far better than the operational PCCS system but larger than that of the counterfactual model that always estimates zero rainfall. Simple CNN appears to do better than either but the frequency distributions still needs to be checked against historical priors in a rigorous way.

## Conclusions / directions

While replication is not complete, it's clear that the SOTA results are plausible. Thus far, it appears that SOTA estimation results can be improved upon; SOTA detection results remain difficult to improve past the majority-class frequency, and additional layers may be needed before large architectural differences can be seen.