Machine Learning Emulation across the Earth System

David John Gagne
National Center for Atmospheric Research
Boulder, CO USA

January 22, 2021

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

The NCAR Analytics and Integrative Machine Learning Group

David John Gagne
Group Head, ML Scientist

John Schreck
ML Scientist

Gabrielle Gantos
Associate Scientist II

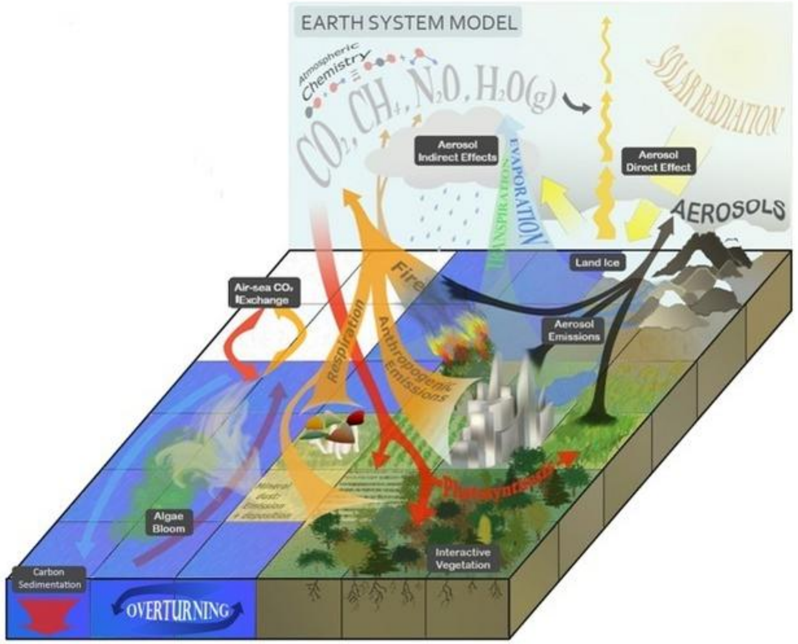
Charlie Becker
Associate Scientist II

Keely Lawrence
Student Assistant

Gunther Wallach
Student Assistant

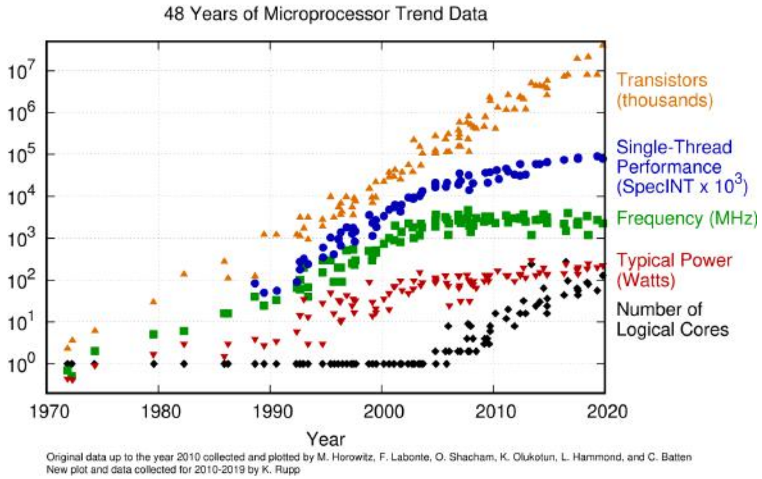
Dallas Foster
Ph.D. Student Intern

The Case for Machine Learning Emulation

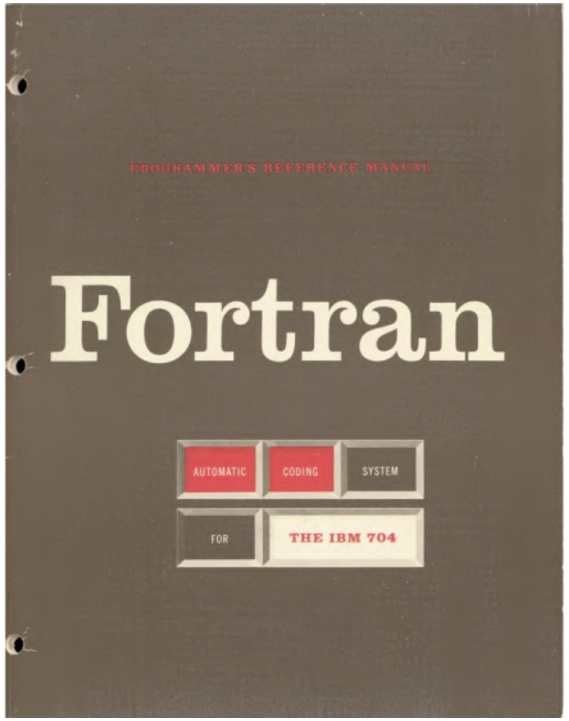


Source: Heavens et al. 2013

Earth system models continue to increase in complexity of processes and scale



Computer performance increases are slowing down



Source: Wikipedia

Lots of legacy software that is hard to manually optimize and interface

Verdict: Business as usual in Earth System Modeling is becoming infeasible

The Case for Machine Learning Emulation

Earth Scientist Strengths

- Designing rigorous modeling experiments
- Generating lots of model output
- Analyzing results for physical consistency and correctness

Earth Scientist Weaknesses

- Writing clean, accessible, optimized, high-performance software
- Writing software for specialized hardware setups
  - Distributed parallel computing
  - Specialized processors (GPUs, TPUs)

Enter Machine Learning

- Can build model from large amounts of specialized data
- Best for optimizing processes that are hard to describe formally and/or are expensive to perform otherwise
- Many accessible libraries to train and evaluate machine learning models on a variety of hardware
- Earth science community can take advantage of software investments from other fields

Result: the NCAR AIML group is working on multiple machine learning emulation projects spanning the Earth System.

Presentation Goals

- Discuss ML emulation implementations for microphysics, atmospheric chemistry, and cloud particle imaging
- Identify ongoing issues with broader deployment of ML emulation and potential solutions

Machine Learning the Warm Rain Process

Collaborators

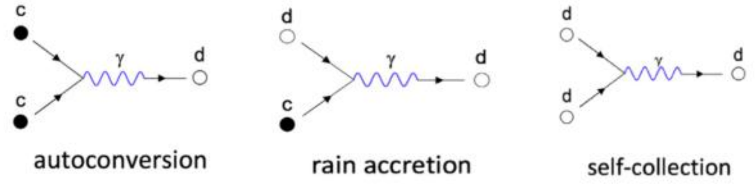
A. Gettelman, D. J. Gagne, C.-C. Chen, M. W. Christensen, Z. J. Lebo, H. Morrison, and G. Gantos

Available online at https://www.essoar.org/doi/abs/10.1002/essoar.10503868.1

Microphysics Emulation: Motivation

Precipitation formation is a critical uncertainty for weather and climate models.

Different sizes of drops interact to evolve from small cloud drops to large precipitation drops.



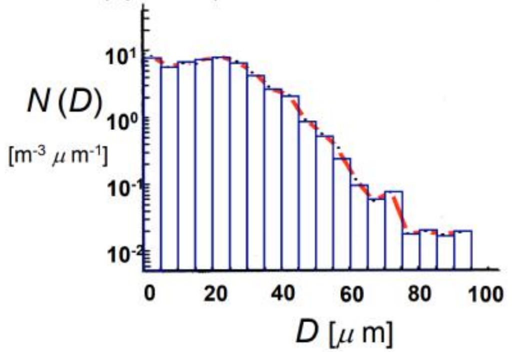
Detailed bin codes are too expensive for large scale models, so empirical functions are used instead.

Can a machine learning approach provide a more accurate emulation of precipitation formation processes without a significant increase in computation?

Goal: Put a detailed bin process into a global general circulation model and emulate it using ML techniques.

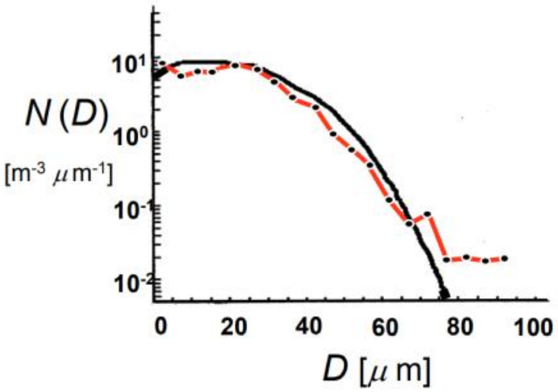
Bin Scheme (Tel Aviv University (TAU) in CAM6): Divide particle sizes into bins and calculate evolution of each bin separately.

Bin-resolving: (spectral) $N(D) = \sum_{i=1}^{I} N_i$

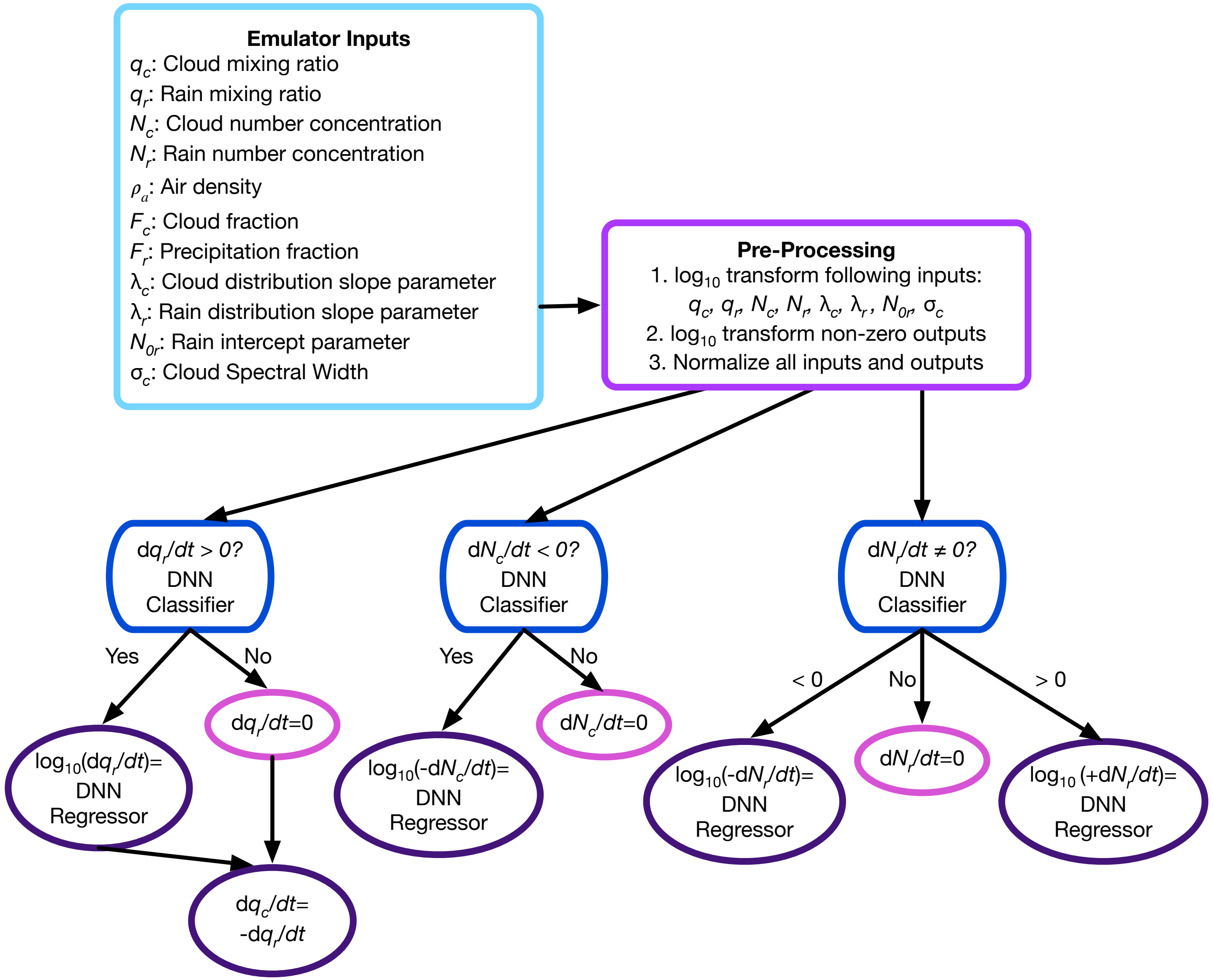


Bulk (MG2 in CAM6): Calculate warm rain formation processes with a semi-empirical particle size distribution (PSD) based on exponential fit to LES bin microphysics runs.

Bulk: $N(D) = N_0 D^{\alpha} e^{-\lambda D}$



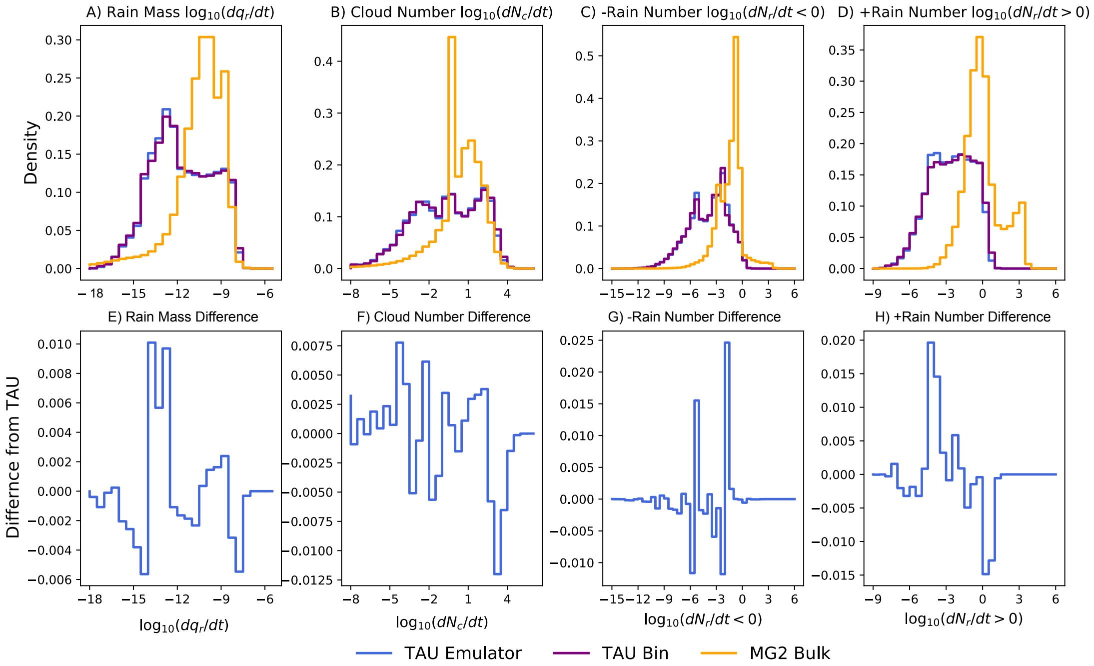
Microphysics Emulator Procedure

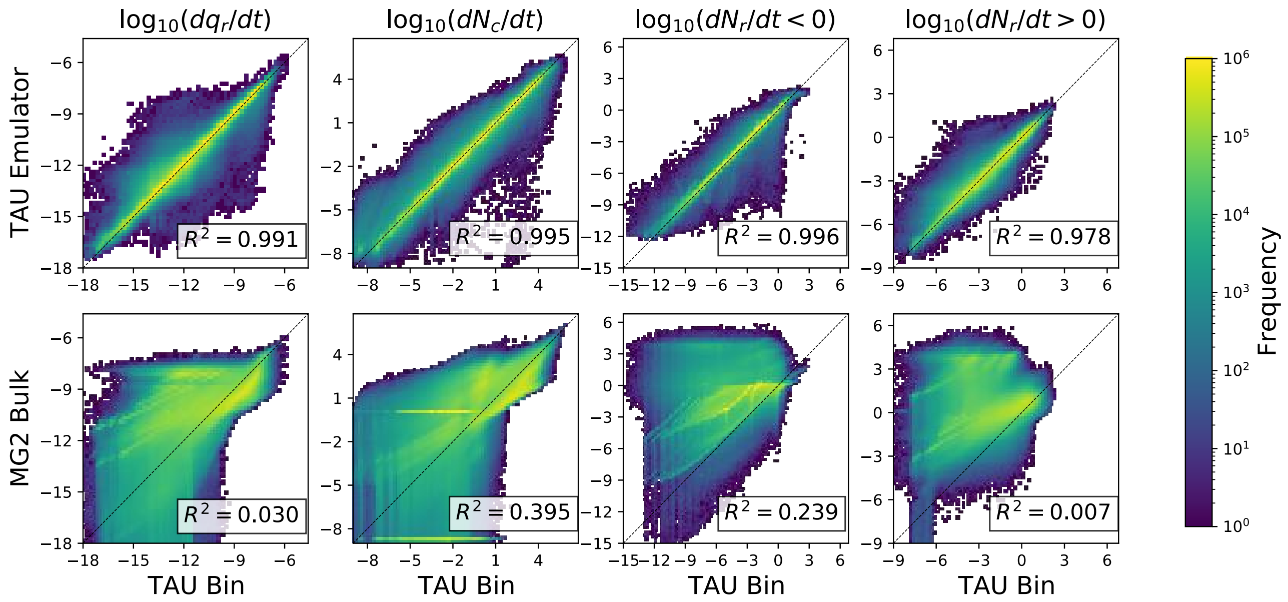


NCAR | UCAR

Contact: dgagne@ucar.edu, @DJGagneDos

Regressor Results





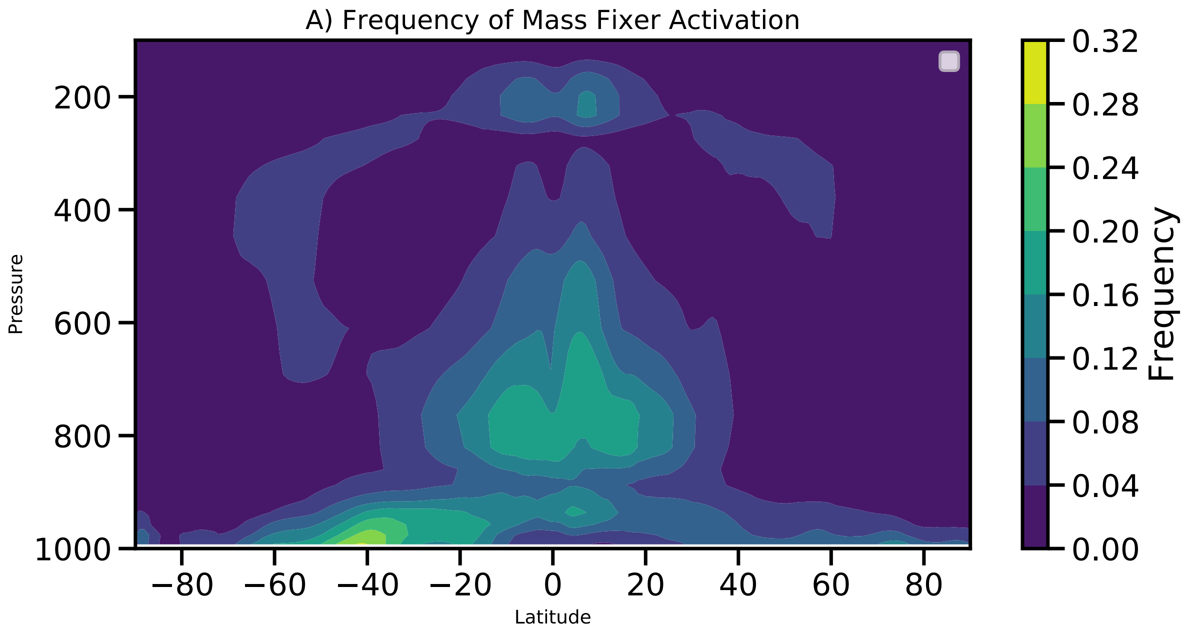
Simulations

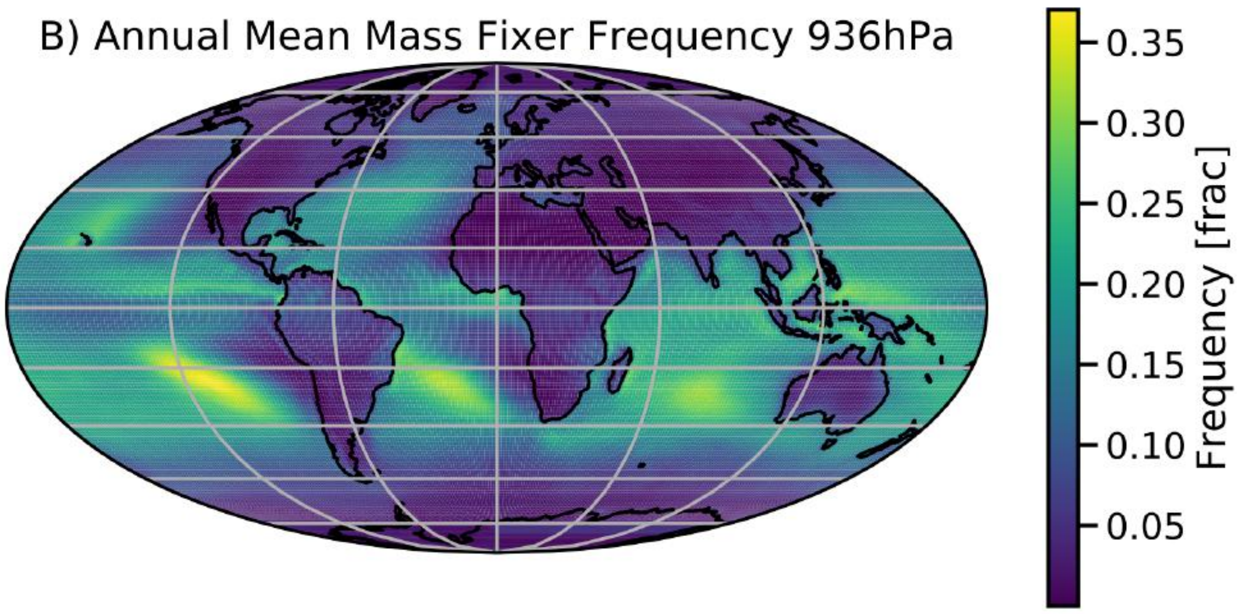
- CAM6: Control
- TAU or TAU-bin: Stochastic Collection Kernel
- TAU-ML: Machine learning Emulator for TAU code (runs neural net inference in Fortran)

- For each, global 0.9°x1.25° simulation, 9 years, 1st year high frequency instantaneous output
  - Base (2000 Climatology)
  - Pre-Industrial (1850) aerosols. (For aerosol cloud interactions)
  - SST+4K (For Cloud Feedbacks)

```fortran
type Dense
   integer :: input_size
   integer :: output_size
   real(kind=8), allocatable :: weights(:, :)
   real(kind=8), allocatable :: bias(:)
   character(len=10) :: activation
end type Dense
```

Mass Fixer for ML Emulator

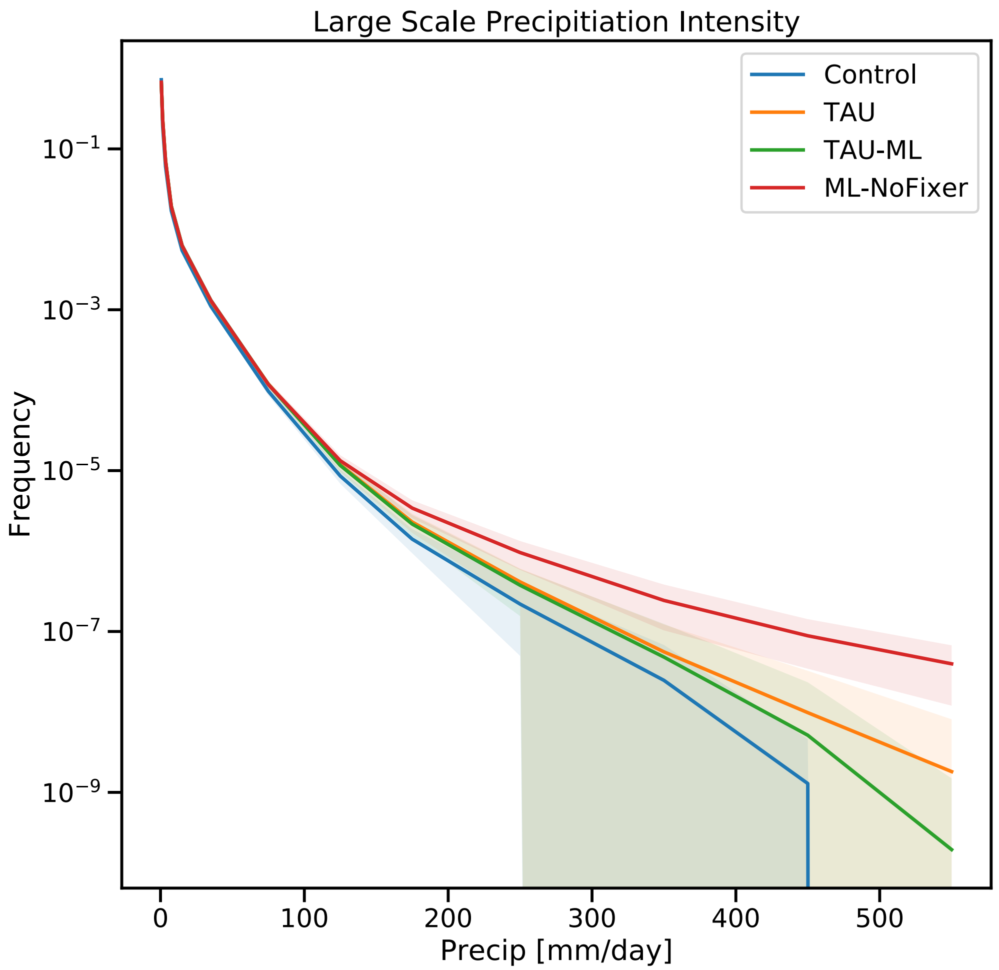


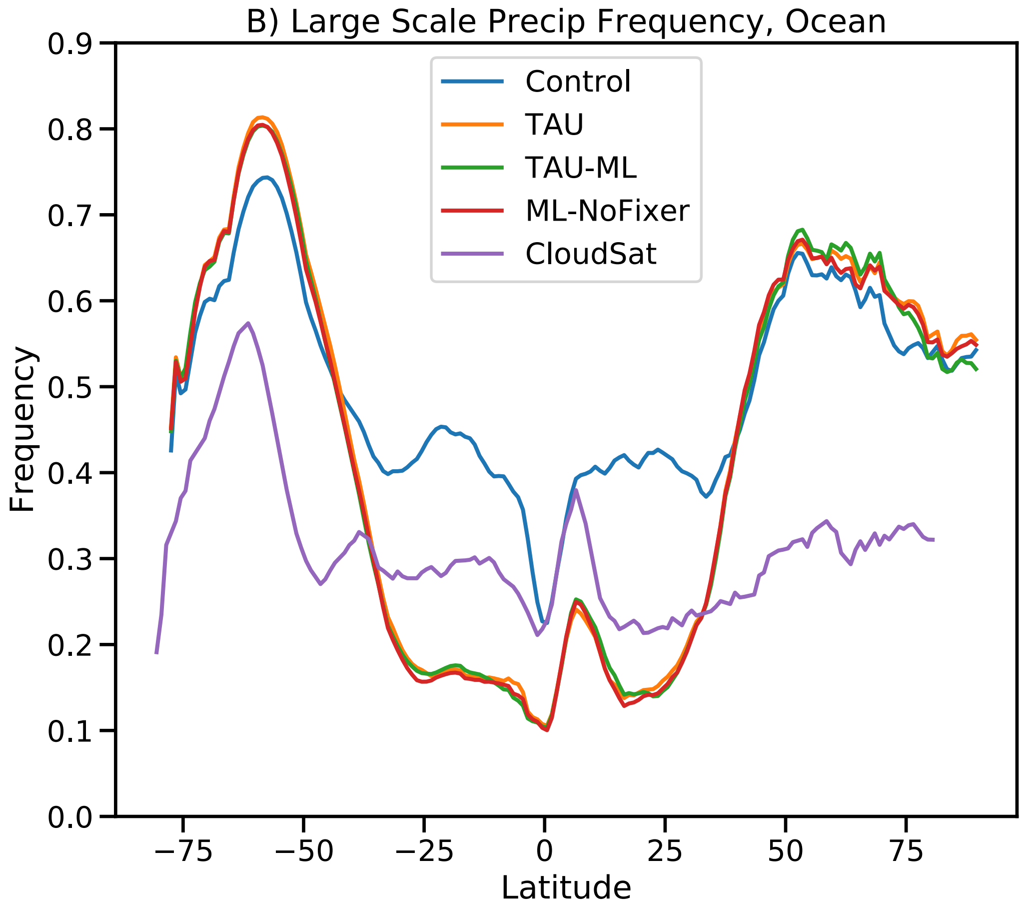


How often does mass fixer kick in and where?

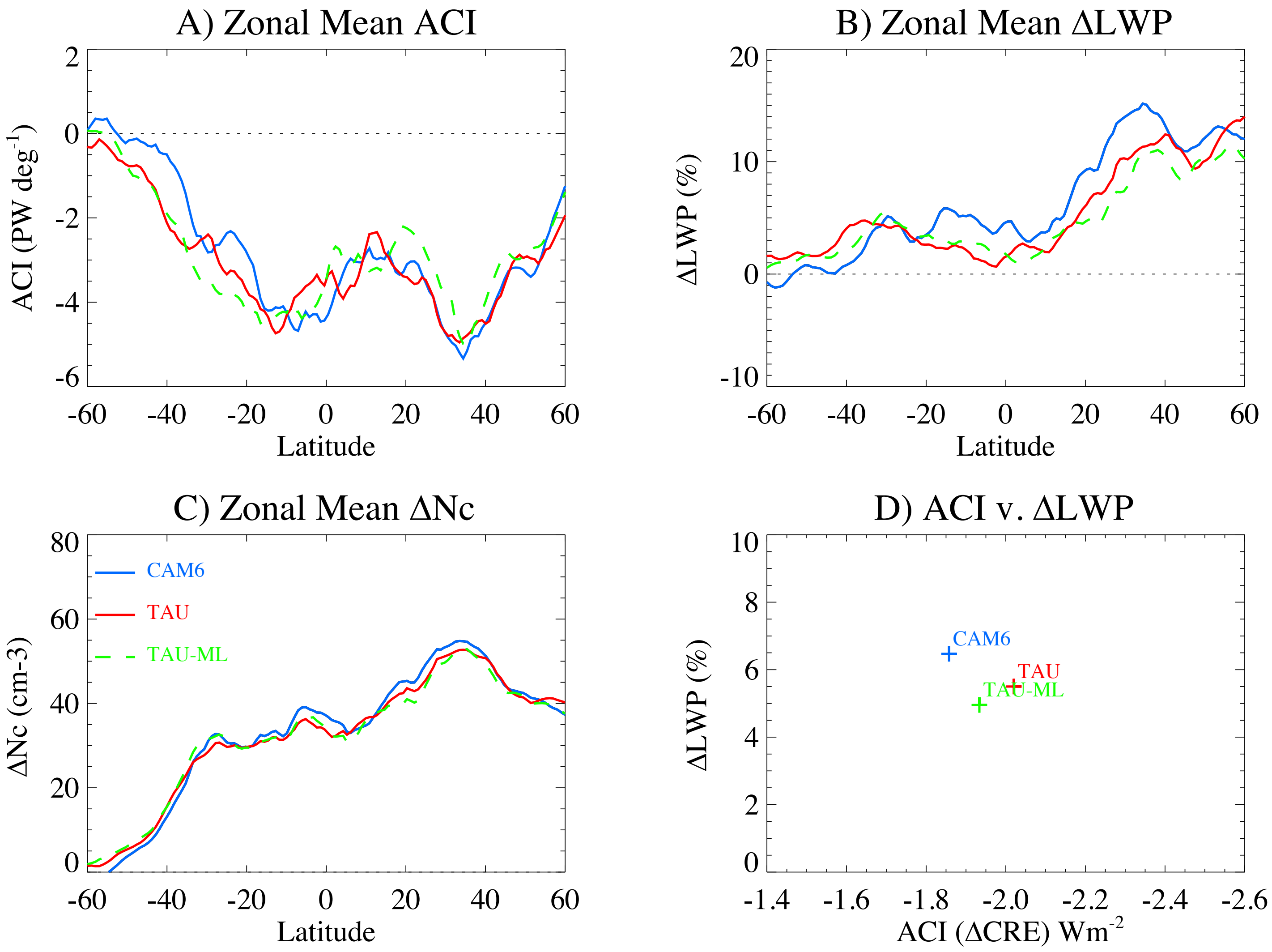
- Low altitudes and tropical high altitudes (cirrus)
- Low altitude (below is 936hPa), mostly in sub-tropical strato-cumulus regions, edge of stratus regions. Mostly SH.
- Also a tropical peak at 800hPa

Precipitation Feedbacks





Cloud Feedbacks



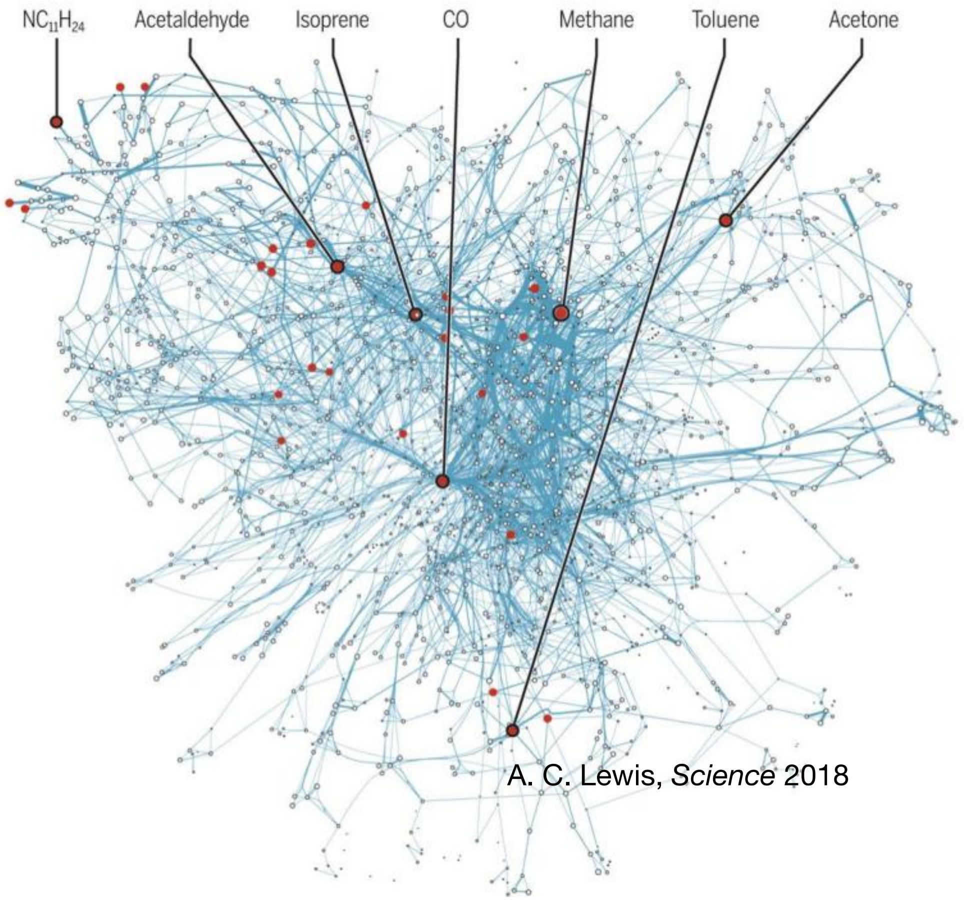
- ACI are similar between control and TAU code.
- Slightly lower LWP change, but forcing is similar, a bit higher in S. Hemisphere.
- Emulator reproduces TAU results.

Machine Learning Emulation of the GECKO-A Chemistry Model

David John Gagne, Charlie Becker, John Schreck, Keely Lawrence, Siyuan Wang, Alma Hodzic

GECKO-A: Generator of Explicit Chemistry and Kinetics of Organics in the Atmosphere

- Natural and anthropogenic sources emit a large number of volatile organic compounds (VOCs)
- VOCs photochemical oxidation the atmosphere leads to hundreds of thousands of volatile products that can condense to form organic aerosols
- **Organic aerosols have significant direct and indirect radiation effect.**



A. C. Lewis, *Science* 2018

GECKO-A

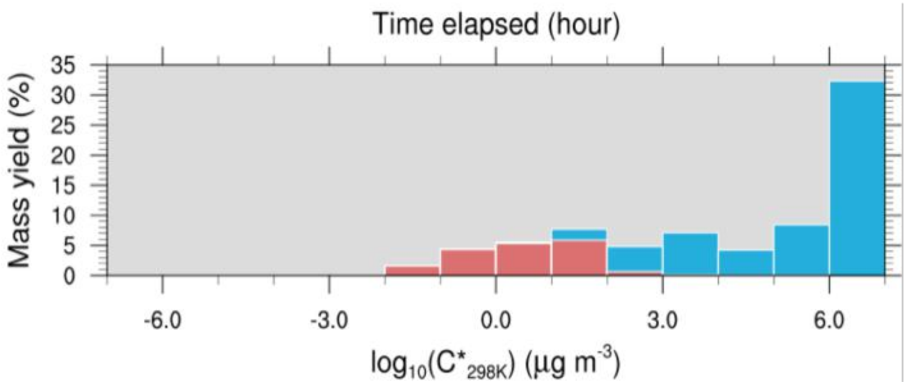
- Explicit - We tell GECKO-A how atoms, bonds, and functional groups in molecules/radicals behave, then GECKO-A will predict what reactions it may undergo.
- Complicated - Oxidation of some compounds involve nearly 400,000 compounds and over 2,000,000 reactions!

GOAL

- **Emulate GECKO-A with ML for a variety of chemical compounds as no 3-dimensional models can afford to run GECKO-A in the foreseeable future.**

Benchmarking Dataset

- Available species (generated as separate simulations)
    - **α-Pinene**
    - **n-Dodecane**
    - **Toluene**
- **Features**
    - Chemical Precursor
    - Gaseous phase mass
    - Aerosol phase mass
    - **Environmental Features**
        - Temperature
        - Solar Zenith Angle
        - Pre-existing aerosols
        - Ozone
        - NOx
        - Hydroxide

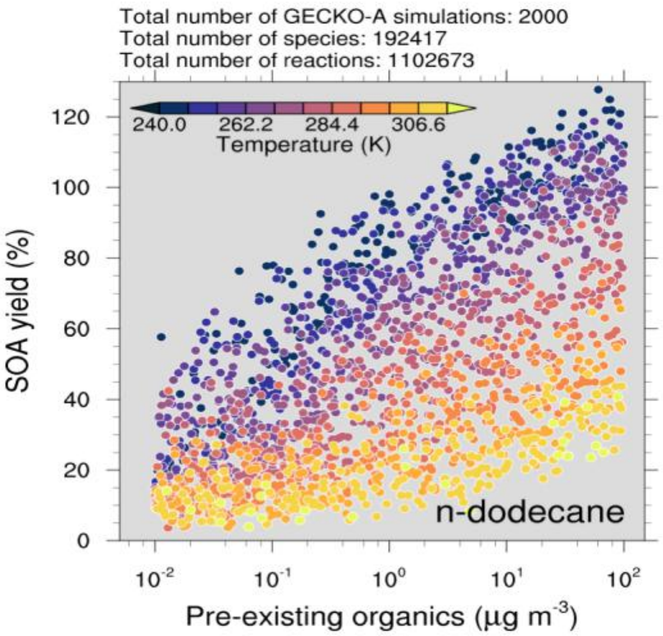


Each feature binned by volatility (14 bins each)

Latin hypercube sampled and held static per experiment

Total data for each species

- 35 Features
- 2000 Experiments
- 1440 time steps
    - Output every 5-minutes (total of 5-days)
- Total of 2,880,000 samples



Initial Benchmarking

Start simple, see what we can learn.

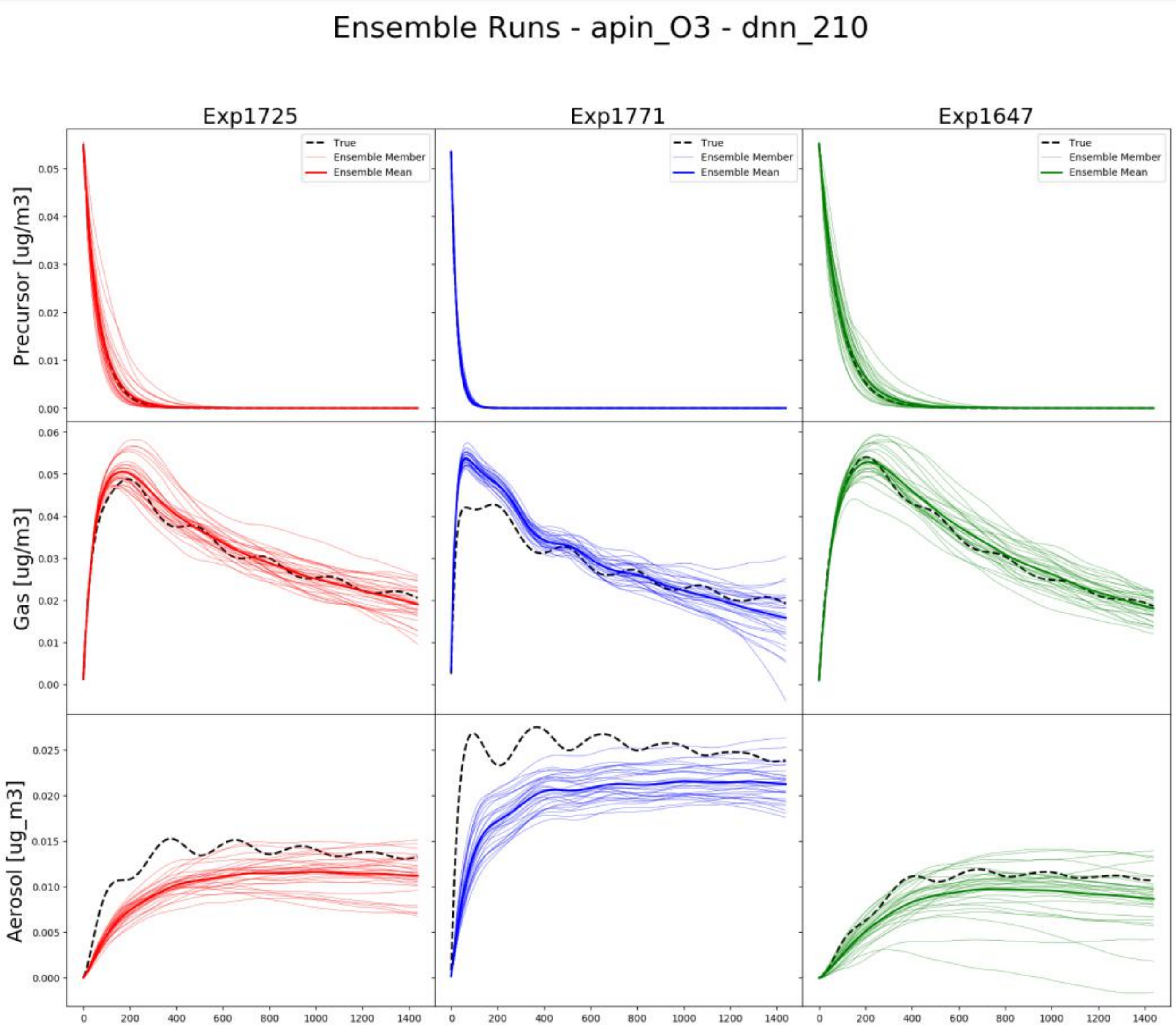
- **Features**
    - Aerosol phase mass
    - Chemical Precursor
    - Gaseous phase mass
    - **Environmental Features**
        - Temperature
        - Solar Zenith Angle
        - Pre-existing aerosols
        - Ozone
        - NOx
        - Hydroxide

$X_{(t)}$

Single layer, 500 neuron dense network

$Y_{(t+1)}$

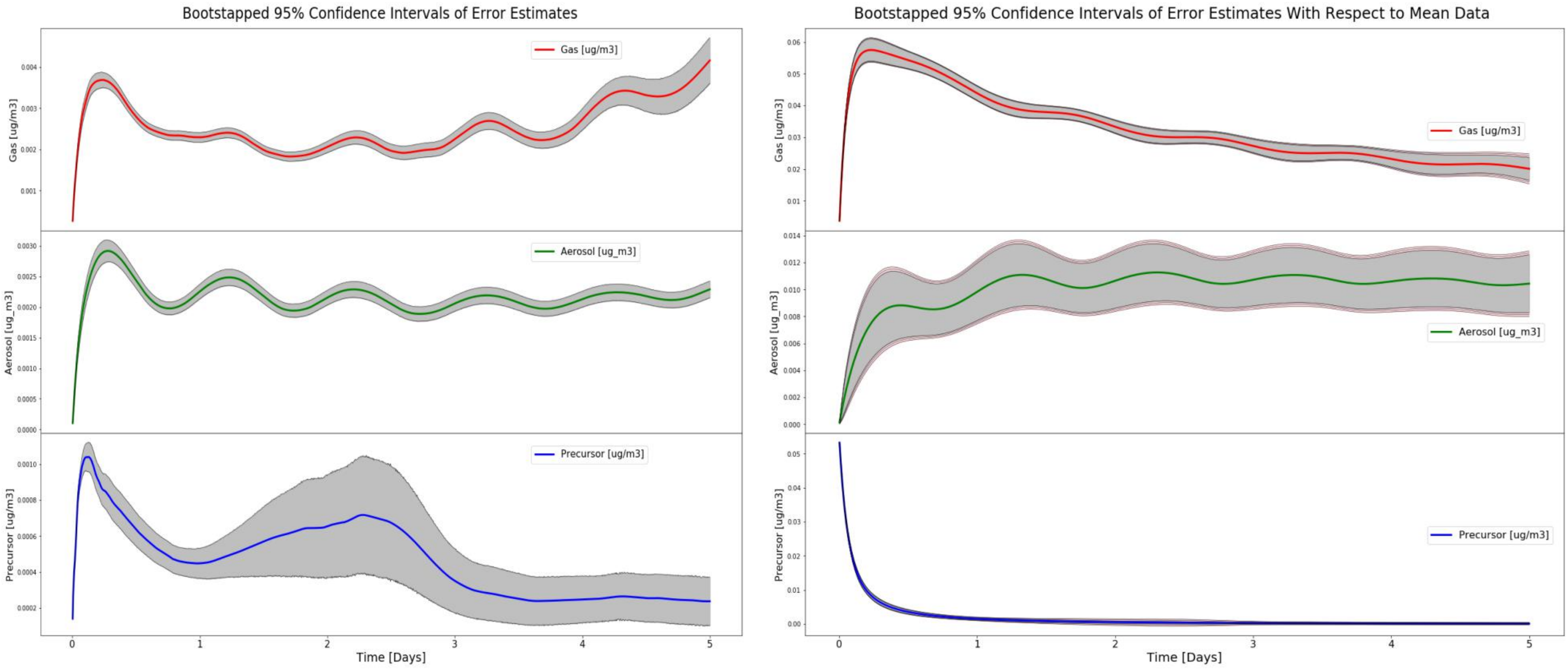
Initial Results: 30 Member Ensemble



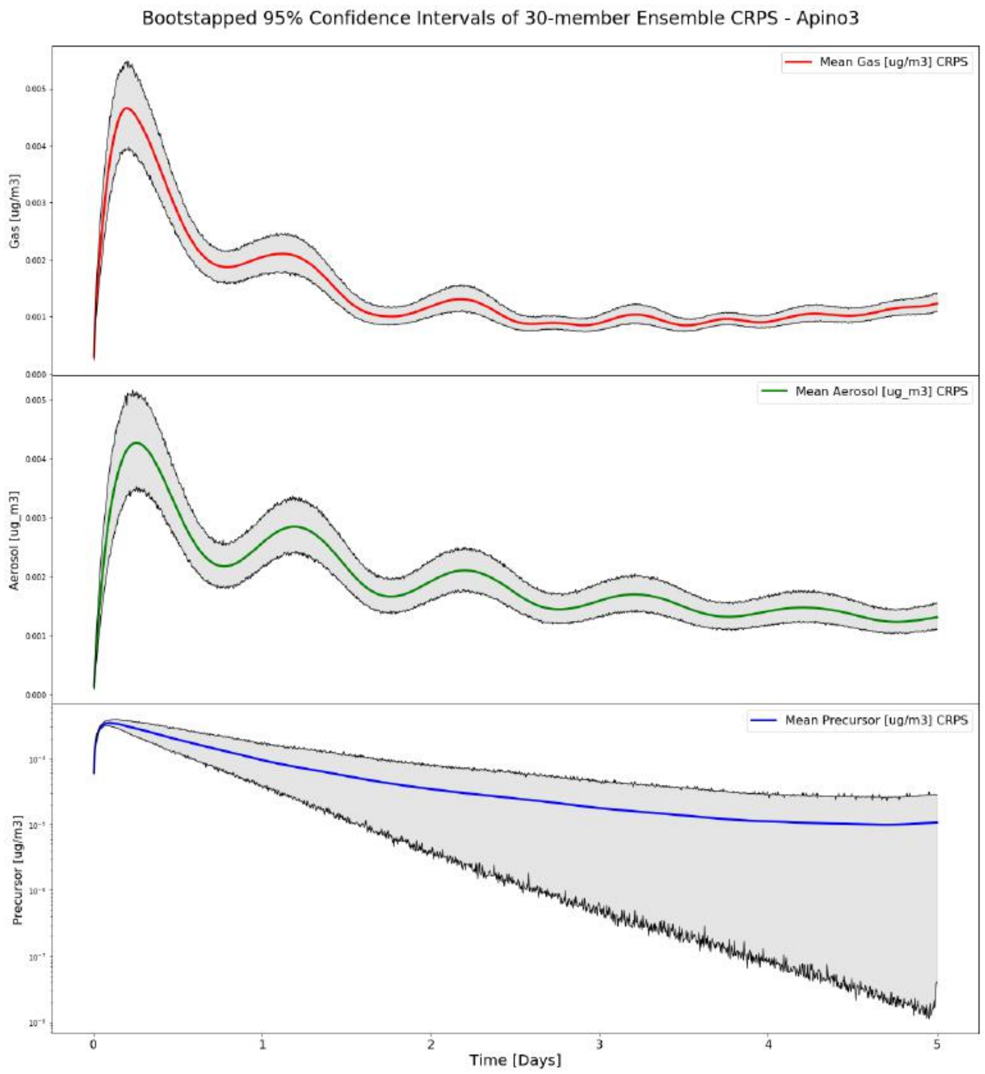
- Generally easy to learn x(t) → y(t+1)
- However, maintaining stability when walking a prediction forward is challenging!
- Extensive hyperparameter tuning
  - Light L1 and L2 regularization
  - Trained for many epochs (1500)

	Precursor	Gas Phase	Particle Phase
R^2	0.54	0.940	0.917
MAE	0.00082	0.0024	0.0027
HD	0.0019	0.0018	0.0501

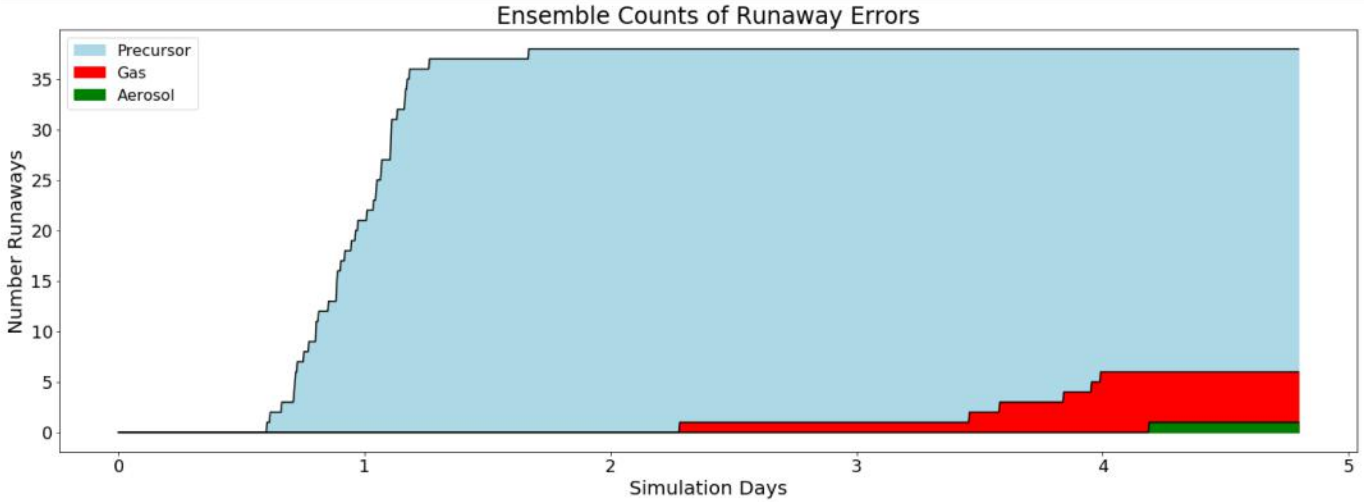
Verification (α-pinene, 200 experiments)



Ensembled Metrics - Continuous Ranked Probability Score

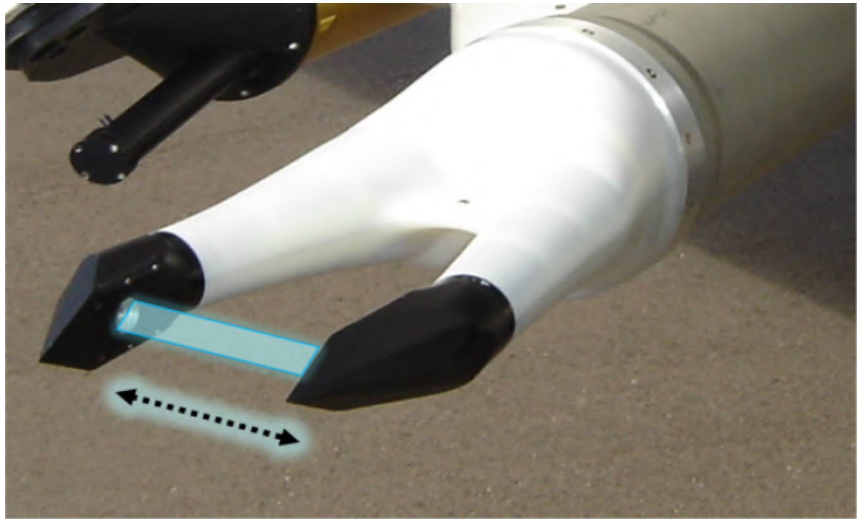


- 30 ensemble members of 200 experiments
  - 99.3% Experiments remained stable
  - *CRPS only calculated for stable experiments*
    - *Runaway threshold +/- 1.0 (raw value)*

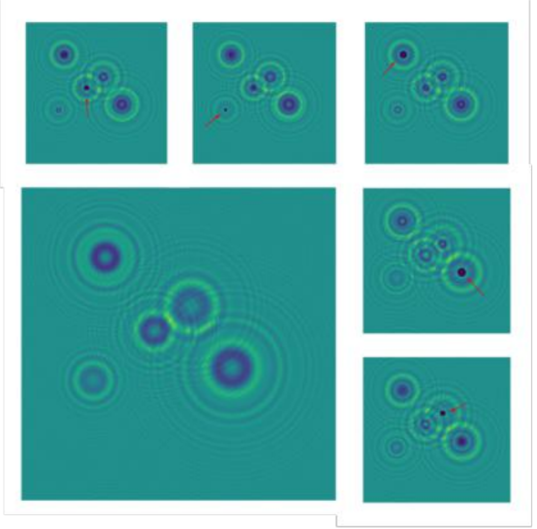


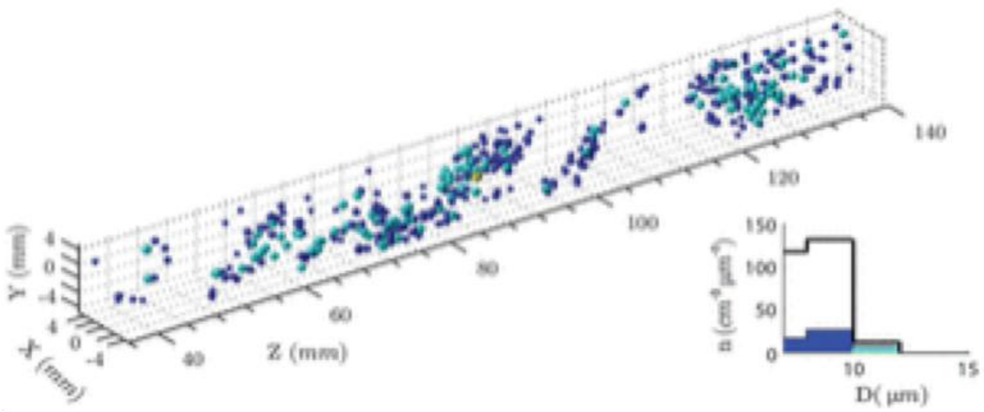
What is HOLODEC?

- Holographic Detector for Clouds (HOLODEC)
- Airborne instrument that measures the microstructure of natural clouds
- Capable of measuring liquid droplets and ice crystals
- Simultaneously measures all particles in the volume between the arms (13 cm^3) in a single picture
- Allows retrieval of the size, shape, and relative 3D position using digital inline holography



Beals et al. 2015





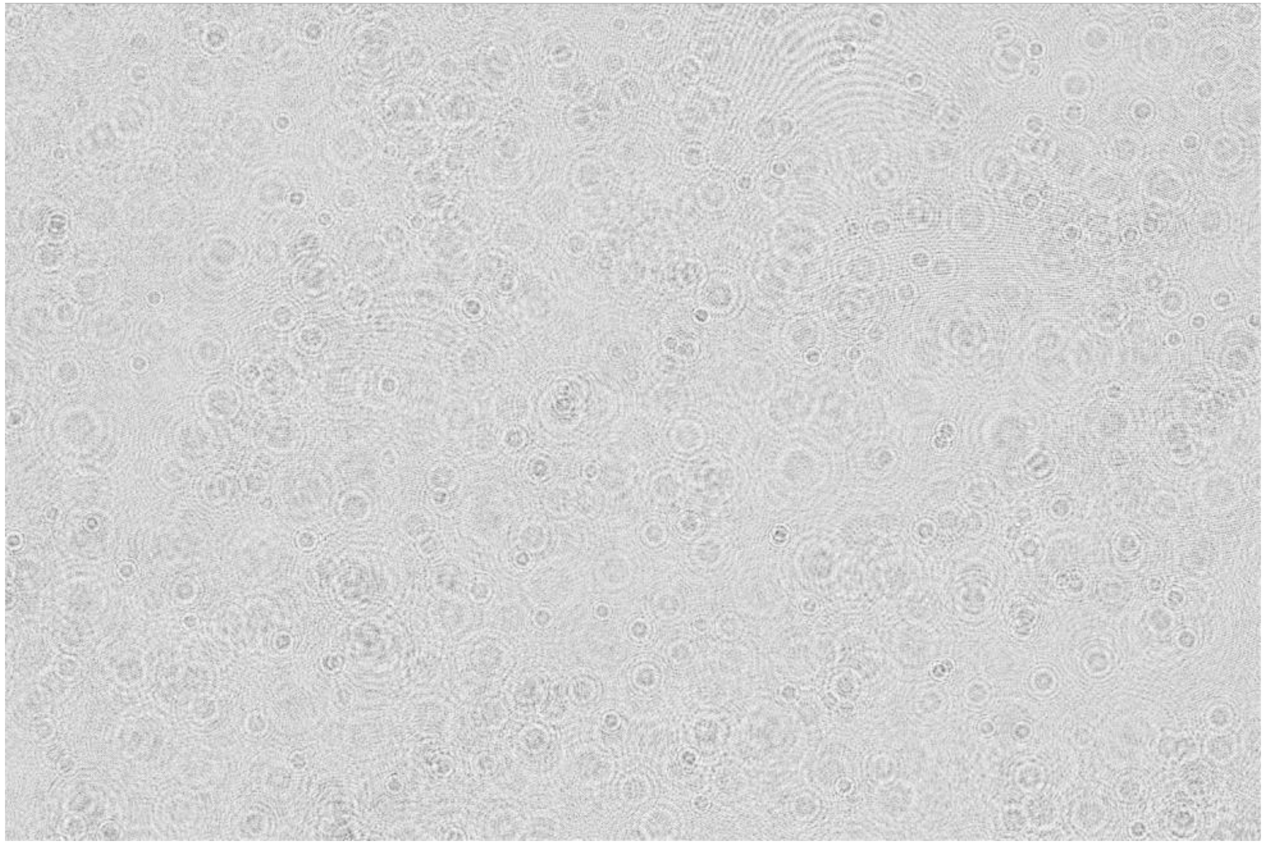
Machine Learning Processing of the HOLODEC Cloud Particle Imager

Collaborators

Matt Hayman, Aaron Bansemer, John Schreck, Gabrielle Gantos, Gunther Wallach, David John Gagne

Instrument Challenges

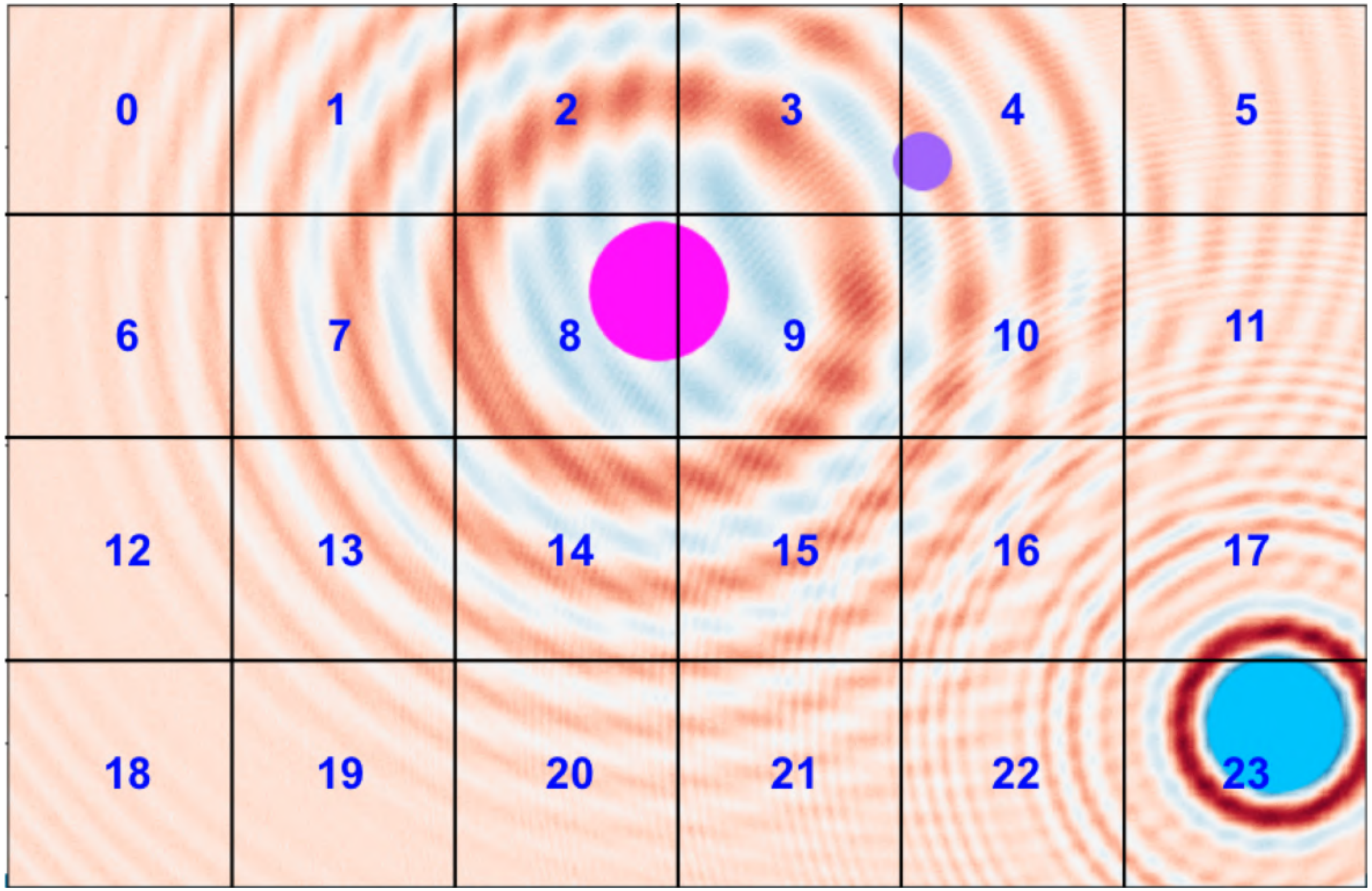
- A single hologram may contain 1000+ particles
- Traditional refocusing is performed 1000 times for each image which searching for particles
- Computationally expensive and labor intensive, up to 2 million core hours per project
- Processing is primary bottleneck in improving probe performance



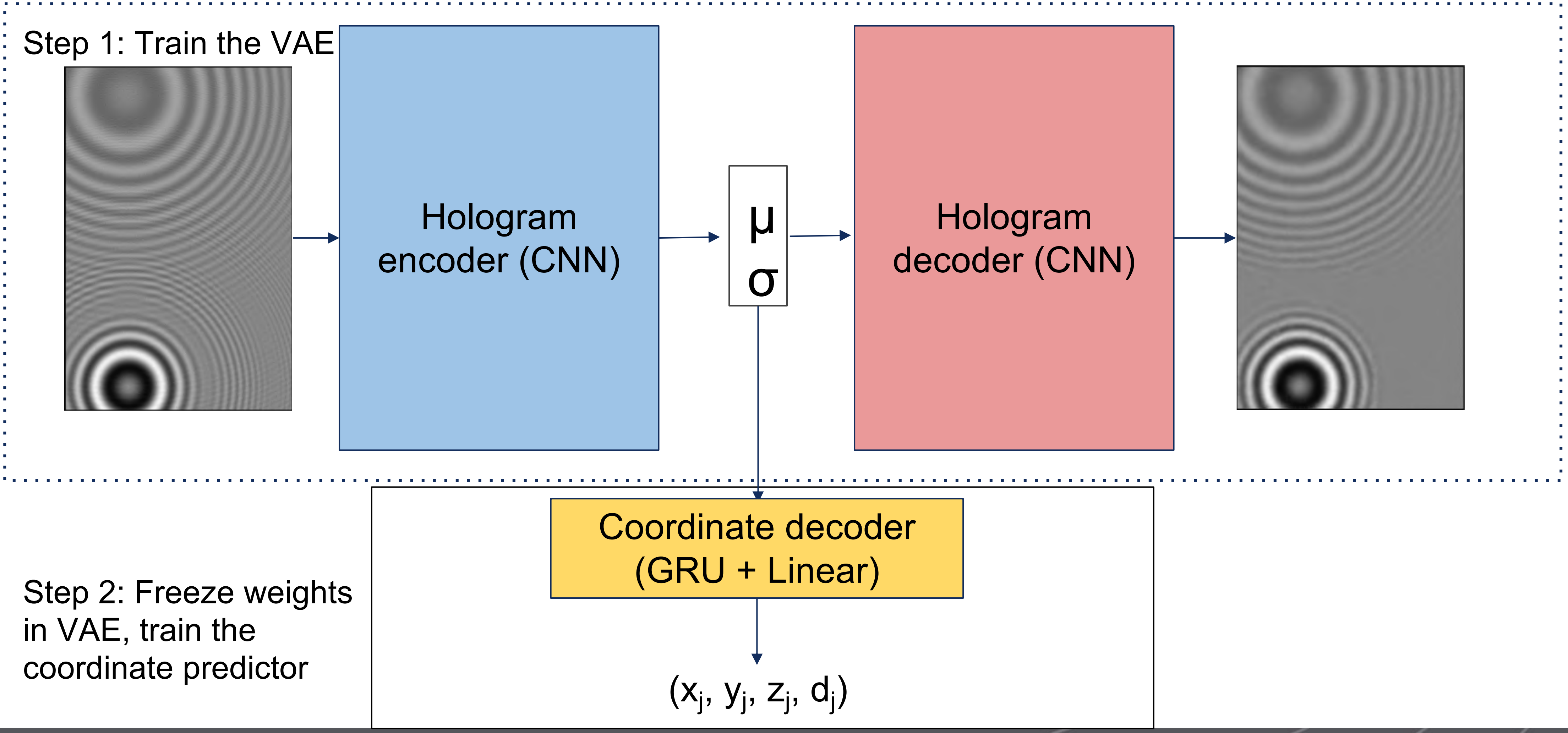
Hologram in Pacific cumulus, 2015

Discretizing hologram coordinates

- Discretize (x, y) each directions into n bins
- Each grid cell associated with a token label
- Each particle is associated with one token
- A sequence of particles as a series of tokens:
  - (START, 23, 8, 4, STOP)
  - Post-localization, a better prediction of (x, y) can be made

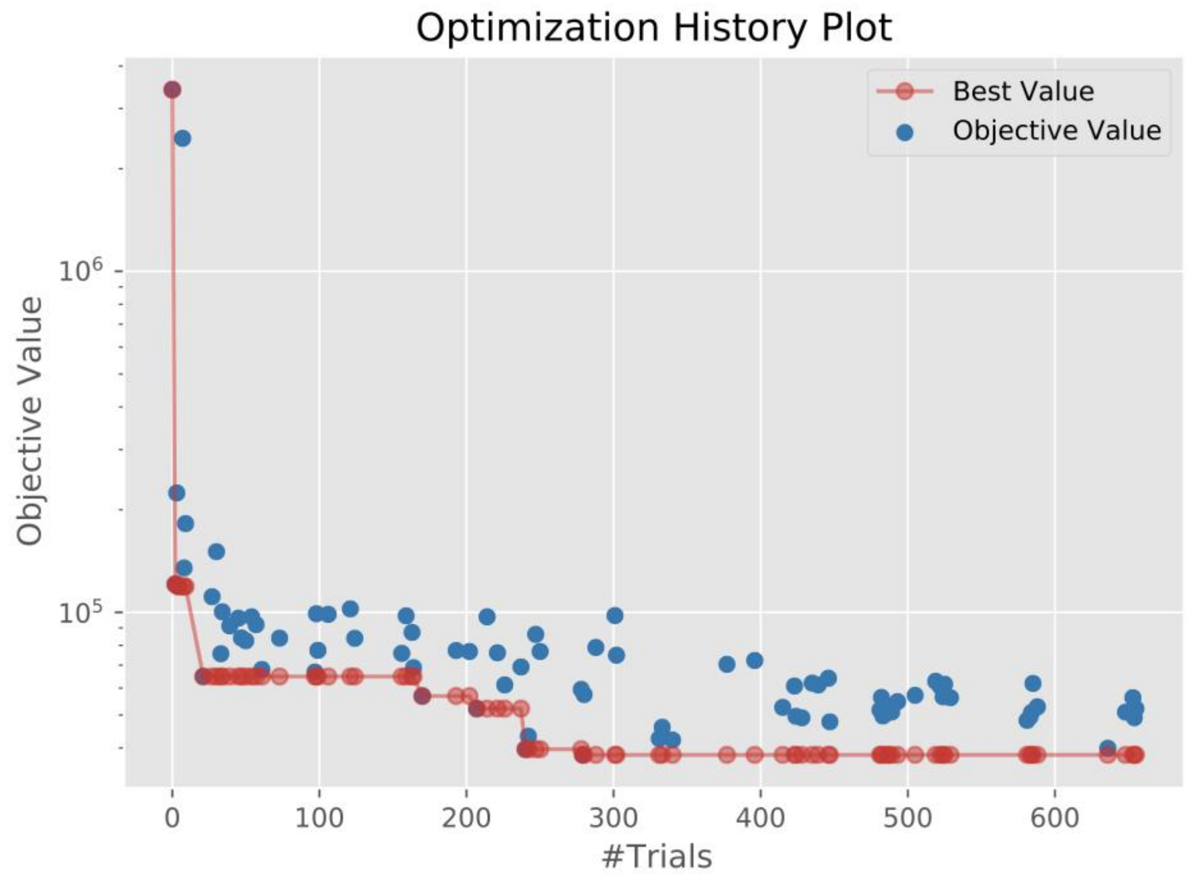


Model architecture:
(1) Variational autoencoder + (2) regression



Results: VAE training and validation

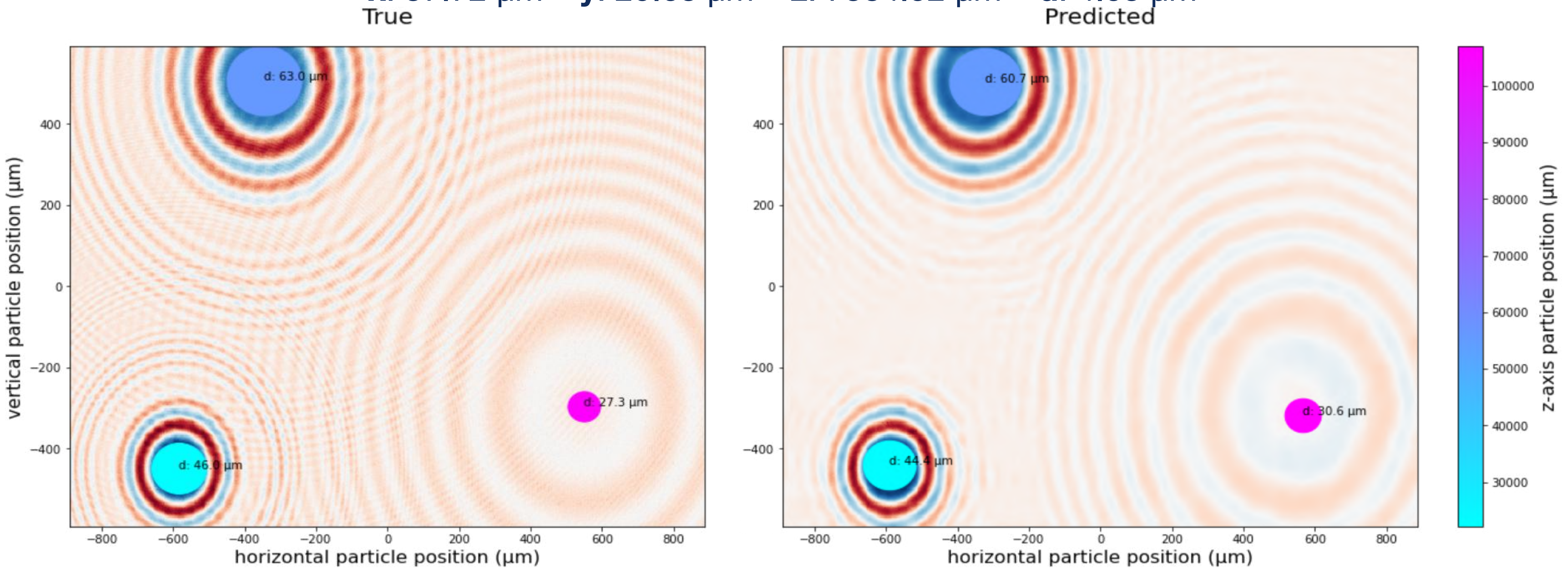
- Hyperparameter search (Optuna):
  - Number of filters in each layer
  - Latent dimension
  - L2 weight decay
  - Learning rate
  - Loss weights on MSE, KLD.
- Validation objective = mse + (fixed weight) * kld
- Using physically-constrained with self-attention (SA)
- The first 100 trails used randomly sampled hyperparameters (within specified ranges)
- After ~650 trails, still finding better models but objective improvement flat since 250 trials.
- Optimization used 500 GPU core-hours, 5,000 CPU core-hours



Results: Particle decoder validation (3-particle, 100 (x,y) bins)

Validation mean-average-error (MAE):

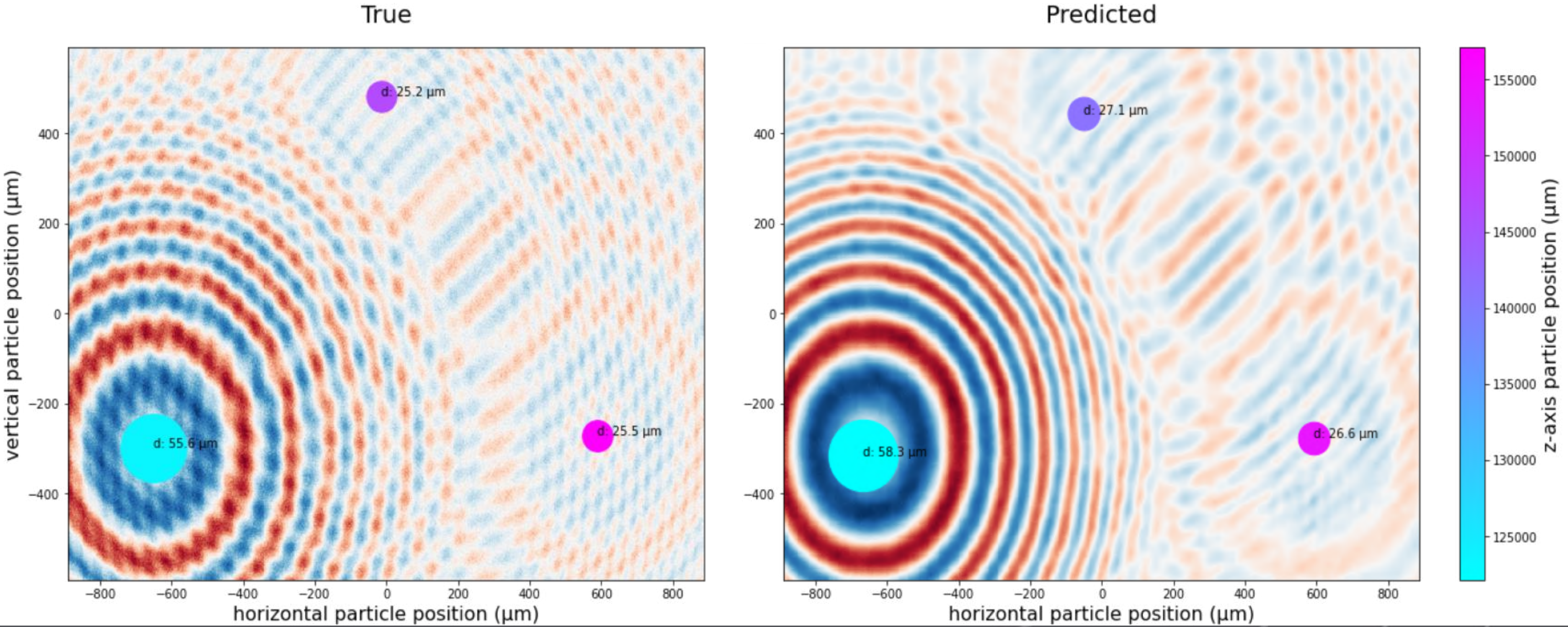
x: 37.72 μm **y**: 29.63 μm **z**: 7384.52 μm **d**: 4.03 μm



Results: Particle decoder validation (3-particle, 100 (x,y) bins)

Validation mean-average-error (MAE):

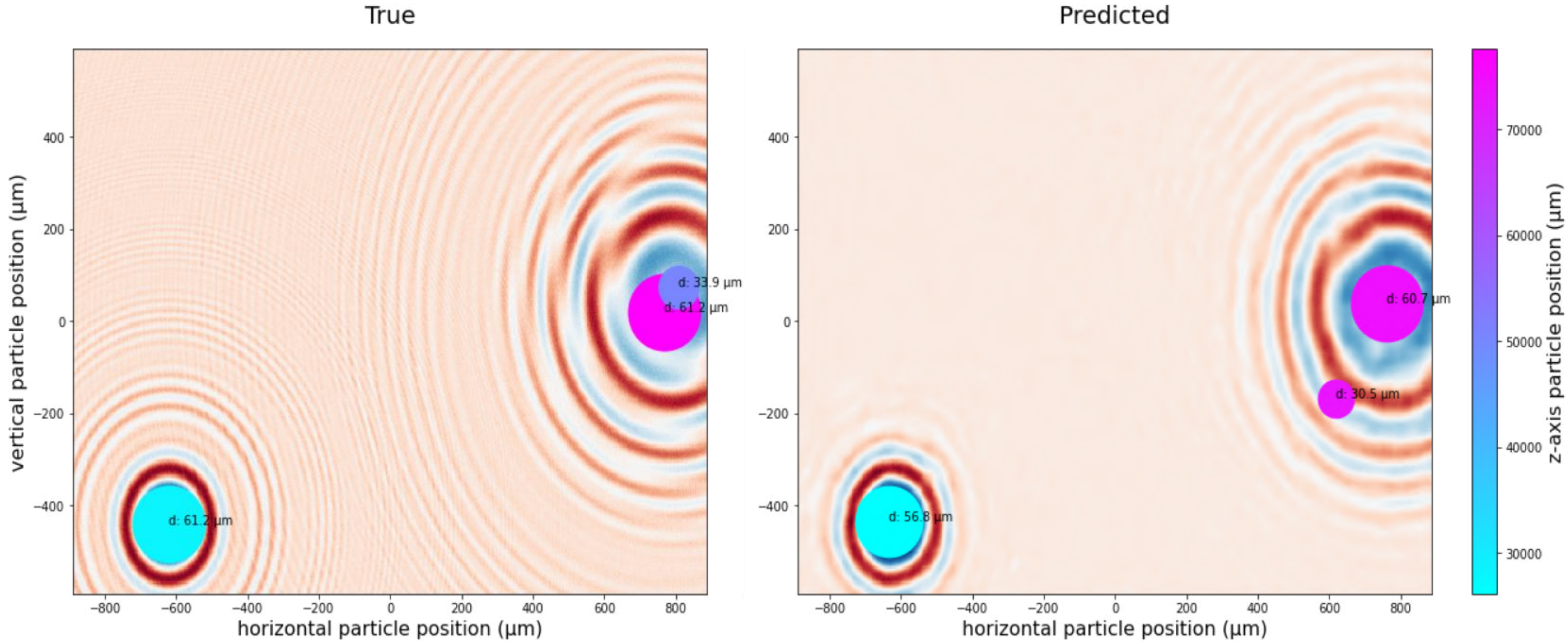
x: 17.75 µm **y**: 20.19 µm **z**: 3272.19 µm **d**: 1.89 µm



Results: Particle decoder validation (3-particle, 100 (x,y) bins)

Validation mean-average-error (MAE):

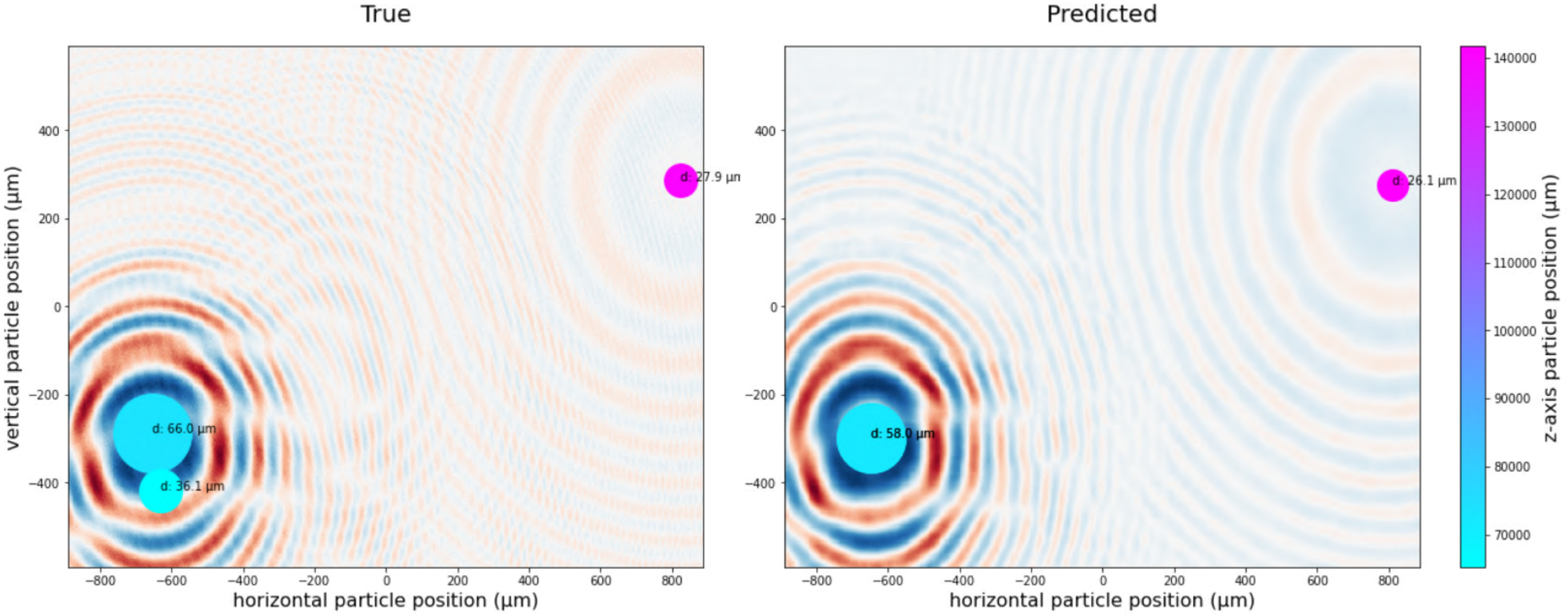
x: 69.33 µm **y**: 88.84 µm **z**: 9344.72 µm **d**: 2.78 µm



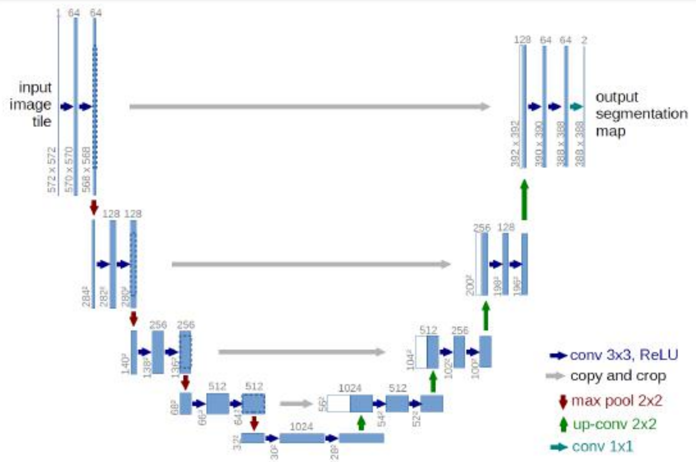
Results: Particle decoder validation (3-particle, 100 (x,y) bins)

Validation mean-average-error (MAE):

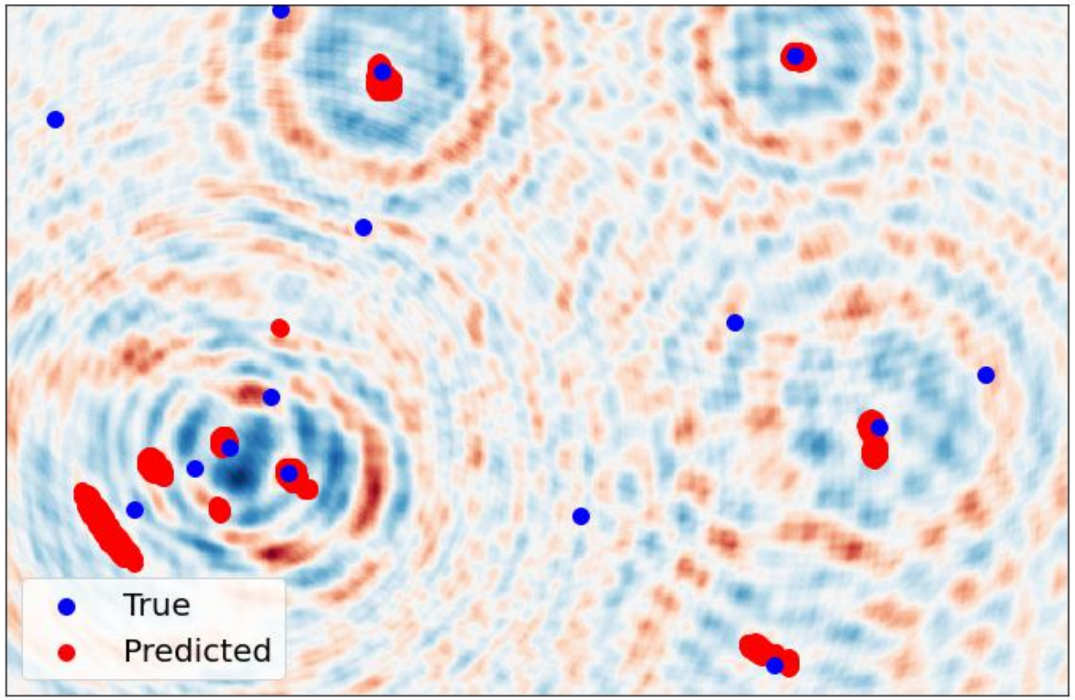
x: 12.19 μm **y**: 46.73 μm **z**: 3189.28 μm **d**: 10.52 μm



Alternative Approach: U-Net



Source: Ronneberger et al. 2015, https://arxiv.org/abs/1505.04597



- U-Net: image to image convolutional neural network with skip connections at each resolution
- Goal: translate hologram to pixel-wise particle location
- Initial results promising but needs further optimization

Infrastructure Costs of ML Emulation

Source: Wikipedia

ML Emulation is like a nuclear plant. Provides lots of power but requires lots of infrastructure to do so.

ML emulation can save significant computational costs but requires its own infrastructure

- Data generation
- ML to Fortran interfaces
- Organization and standards for storing ML model collection
- Specialized hardware
- Physical constraints/fail safes
- Monitoring
- Code and data maintenance/retraining

Is emulation worth the effort?

Summary

- Machine learning emulation is a potentially viable path forward to incorporating complex Earth system process models into ESMs without blowing the computation budget
- Promising initial results for
  - Microphysics
  - Atmospheric Chemistry
  - Cloud particle emulation
- Next steps: integrate into more modeling systems and estimate true infrastructure costs

Contact Me
Email: dgagne@ucar.edu
Twitter: @DJGagneDos