Machine Learning for Retrieving Cloud Optical Thickness from Observed Reflectance: 3D Effects

CyberTraining 2020 Team 5
Kallista Angeloff¹, Kirana Bergstrom², Tianhao Le³, Chengtao Xu⁴
RA: Chamara Rajapakse⁵, Kevin Zheng⁵, Mentor: Zhibo Zhang⁵

¹Department of Atmospheric Sciences, University of Washington
²Department of Mathematical and Statistical Sciences, University of Colorado Denver
³Division of Geological and Planetary Sciences, California Institute of Technology
⁴Department of Electrical Engineering, Embry-Riddle Aeronautical University, Daytona Beach
⁵Department of Physics, University of Maryland, Baltimore County

http://cybertraining.umbc.edu/
Cloud optical thickness (COT)

• Vertical optical depth

• Determines how much solar radiation reaches the planet’s surface vs. how much is reflected or scattered

• Difficult to measure directly – radar, lidar, pyranometers

• Calculated (“retrieved”) from satellite observations of reflectance
Retrieval algorithm and assumptions

\[ \tau_c' = (1-g)\tau_c = \frac{4K(\mu)K(\mu_0)}{3[R_\infty(\mu, \mu_0, \phi) - R(\tau_c; \mu, \mu_0, \phi) - 2q' - \frac{4A_g}{3(1-A_g)}].} \]

• Underlying assumptions: plane-parallel (2D) clouds, each pixel independent

• Actual clouds: 3D effects – horizontal transport of radiation
Synthetic data

• 1D fractal cloud profiles – “lines of cloud”
Synthetic data

• Three connected parameters:

  • Liquid water path (LWP), a measure of the weight of liquid water between two points in the atmosphere – fractal-like random variation

  • Cloud drop effective radius (CER), a measure of the size distribution of water drops within a cloud – fixed; randomly assigned to each profile

  • Cloud optical thickness (COT) – calculated from other two
3D Radiative Transfer Model (SHDOM)

• Spherical Harmonic Discrete Ordinate Method (SHDOM)

• Inputs:
  • COT, CER, LWP
  • Solar zenith angle (SZA) – angle of the sunlight striking the cloud; 60°

• Outputs:
  • Reflectance
3D Radiative Transfer Model (SHDOM)

• “Step cloud” check
How to get from reflectance to COT?

• Historically an inverse problem

• Bispectral method – still used by MODIS, for example
  • One visible wavelength, one absorbing wavelength

• Simplification of machine learning: pattern recognition (statistical inference) rather than inversion
Deep neural network (DNN)
DNN structure

- Current DNN structure based on DNN-2r [Okamura et al 2017]
- Optimizer: Adam, Loss = mean_square_error
- Batch = 4, epochs = 10
Results from deep neural network (DNN)
Results from deep neural network (DNN)

- Although MSE converged to a relatively low value, overall COT predictions were inaccurate (unphysical)

- Accuracy decreased with increasing spatial position

- Most likely explanation: overfitting
Convolutional neural network (CNN) spatial slicing

CyberTraining: Big Data + High-Performance Computing + Atmospheric Sciences
Convolutional neural network (CNN) spatial slicing

CyberTraining: Big Data + High-Performance Computing + Atmospheric Sciences
Convolutional neural network (CNN) spatial slicing

CyberTraining: Big Data + High-Performance Computing + Atmospheric Sciences
CNN structure

- Current CNN structure based on DNN-4w [Okamura et al]
- Parallelization possible over spatial slices

```
Reflectances 12 x 3
\rightarrow Convolutional 6, 100
\rightarrow Convolutional 1, 4
\rightarrow Dropout
\rightarrow Fully Connected 8
```

CyberTraining: Big Data + High-Performance Computing + Atmospheric Sciences
Results from CNN

- Loss: mean squared logarithmic error
- Dropout rate: 0.5
- Batch size: 1024
- Epochs: 30
Results from CNN

- Loss: mean squared error
- Dropout rate: 0.5
- Batch size: 1024
- Epochs: 30
Discussion of results

• The CNN was more successful than the DNN: why?
  • Not a one-to-one comparison – full data vs. spatial slicing
  • Nature of the data more closely resembles an image problem

• Edge (boundary) cases inherently more difficult

• More successful at predicting local rather than global COT
  • Some outliers and unphysical predictions – pre- or post-processing
Future directions

• Expanded dataset – tuning
  • More data to better cover parameter range
  • More parameter tuning and structural changes

• Emulating a multi-scale model
  • Using a DNN-based global method to inform the local CNN

• 2D data – testing on satellite data, e.g. MODIS
  • Structure of spatially sliced CNN lends itself to parallelization