ML/AI Applications to the Atmospheric Science Data and Simulations
(Demonstration and Vision)

- ❑ Part 1: Applications to Remote Sensing Data and Vison
- ❑ Part 2: Applications to Scientific Models and Vison
- ❑ Part 3: Applications to Integration/Automation and Vision

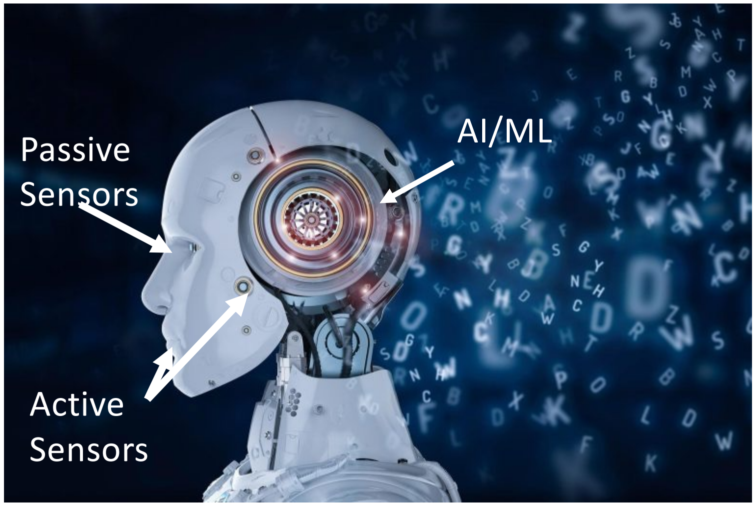
Shan Zeng, Ali Omar, Yongxiang Hu, Charles Trepte, Mark Vaughn, Kuanman Xu et al.
NASA LaRC/CAI
2/13/2024

Motivation

❑ Passive and Active Remote Sensing

Remote sensing is the acquiring of information from a distance. (From NASA website)

=> **Digital sensors to mimic human five facial senses to interact with the environment and gather information**



❑ Artificial Intelligence

Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind (From IBM Website)

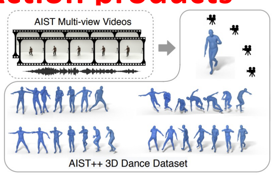
=> A digital brain to mimic different human cognitive approaches to solve different problems

❑ Remote Sensors + Artificial Intelligence + Atmosphere Science

The **Intelligent Science Assistant** can help us conduct atmospheric science research. It simplifies complex issues and enhances our capabilities to solve scientific problems. The ultimate goal is to realize **integration and automation** in the atmospheric science research, enabling us to gain a comprehensive understanding of the Earth-Atmospheric system, pursue higher objectives, and deepen our exploration

Part I: ML Application to Remote Sensing Data

- **can produce operational products + beyond operational products**
  - **Operational Products**
    - Pattern recognitions
    - Microphysical-optical properties retrievals
  - **Innovative products – new classes/new proprieties**
  - **Action products**



- Are clouds/aerosols blooming, growing, disappearing, merging, or breaking?
- What is the speed of the action and how long will this action last?

- **Integrated products**

➔ **Learning among sensors**

- Use passive sensing signal/products to produce active sensing signal/products

(Gong et al., 2023: toward physics-informed Neural Networks for 3D multi-layer Cloud Mask Reconstruction)

- Use low resolution or multispectral signal to produce high resolution or hyperspectral signal

(Liu et al., 2021: research on super-resolution reconstruction of remote sensing images: a comprehensive review)

- Use high noise signal to produce low noise signal

(hu et al. 2021: A Novel Lidar Signal Denoising Method Based on Convolutional Autoencoding Deep Learning Neural Network)

➔ **Learning among sensors and simulations**

By

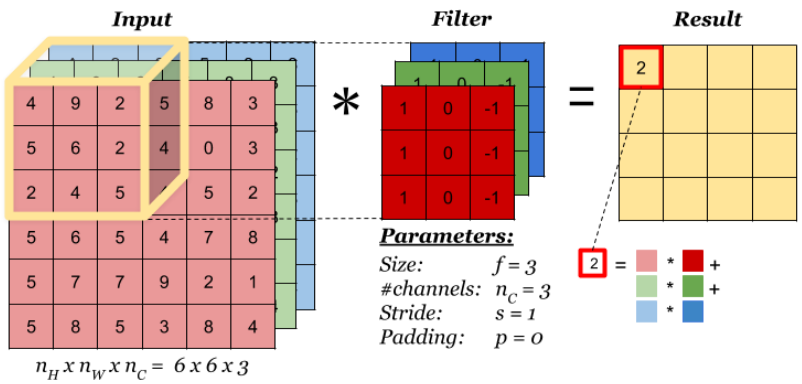
- **Improving products quality with multiple sources**
- **Filling the gaps of observations**

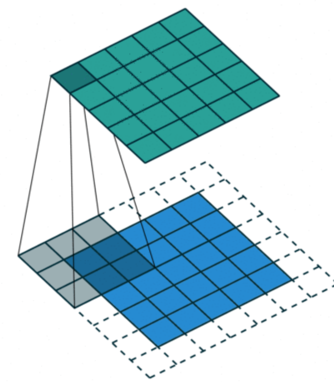
Reconstruct high quality 4D/5D Earth-Atmosphere data systems

❑ can produce and improve all operational products

<u>The Advantages of Convolutional Neural Network (CNN) Architectures? -apply to solve spatial problems</u>

- channel information (absorption, scattering, polarization)
- texture information (relationship with neighborhood pixels/angular/channel)

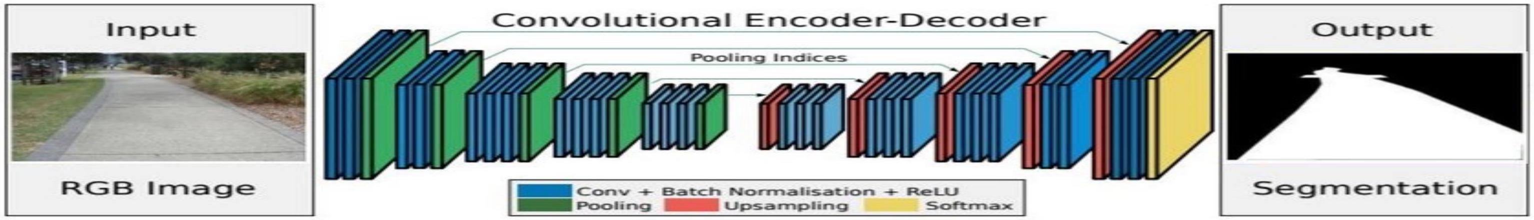




<u>Benefits of using ML algorithm</u>

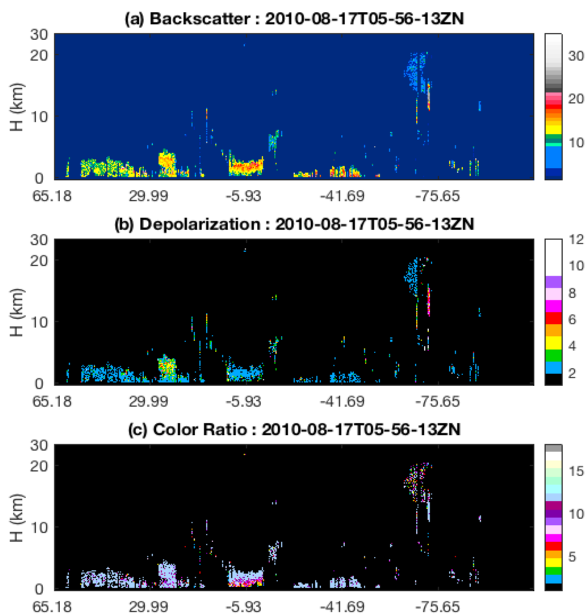
- Allow us to use naïve resolutions signal, and consider both the signal intensity and shape, to do pattern recognitions or properties retrievals at bin-to-bin level (creating profile to profile LUT)

Part I: ML Application to Remote Sensing Data



CNN model

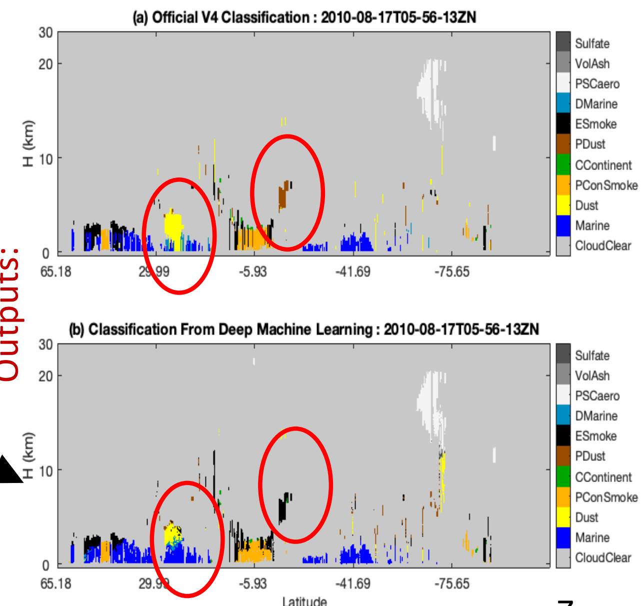
Inputs



Thresholds Tree

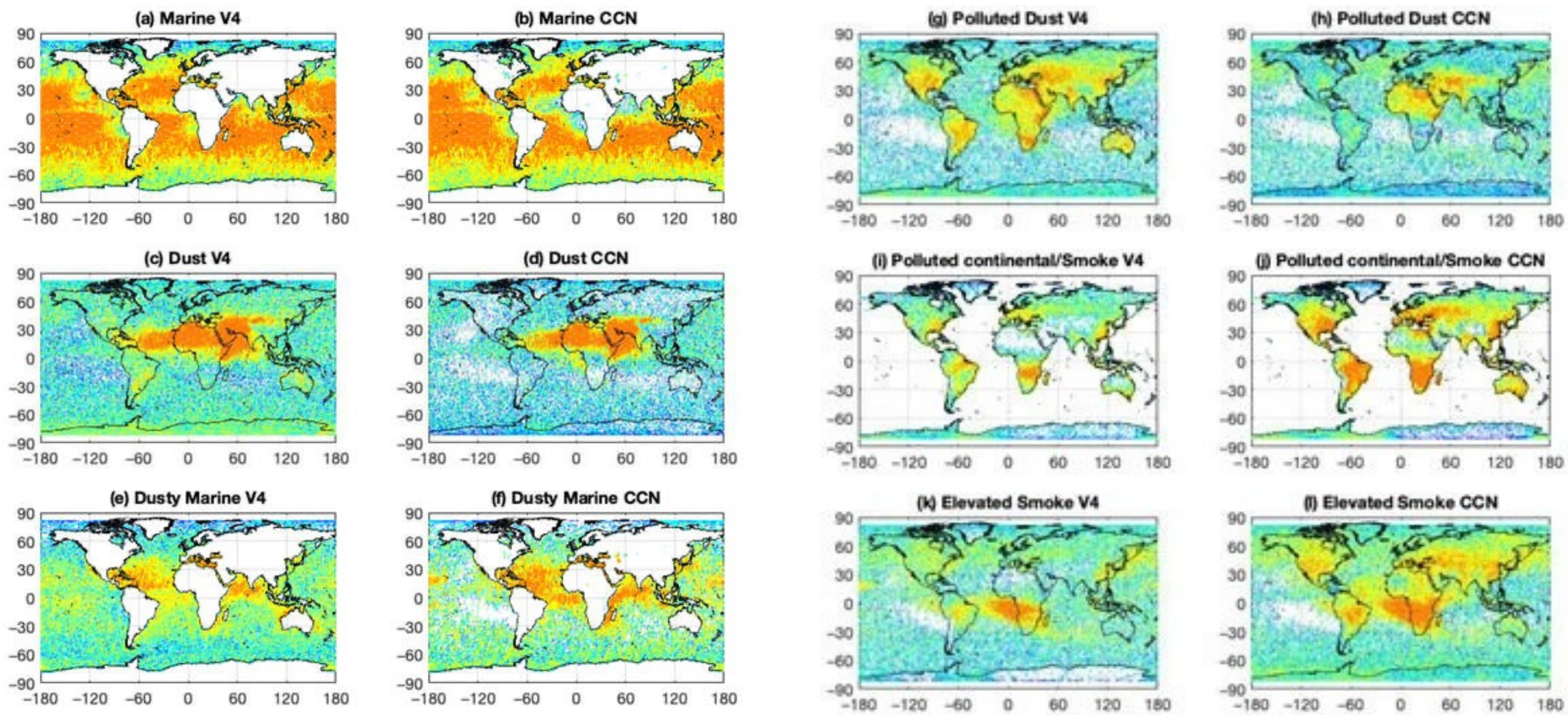
Segnet

Outputs:



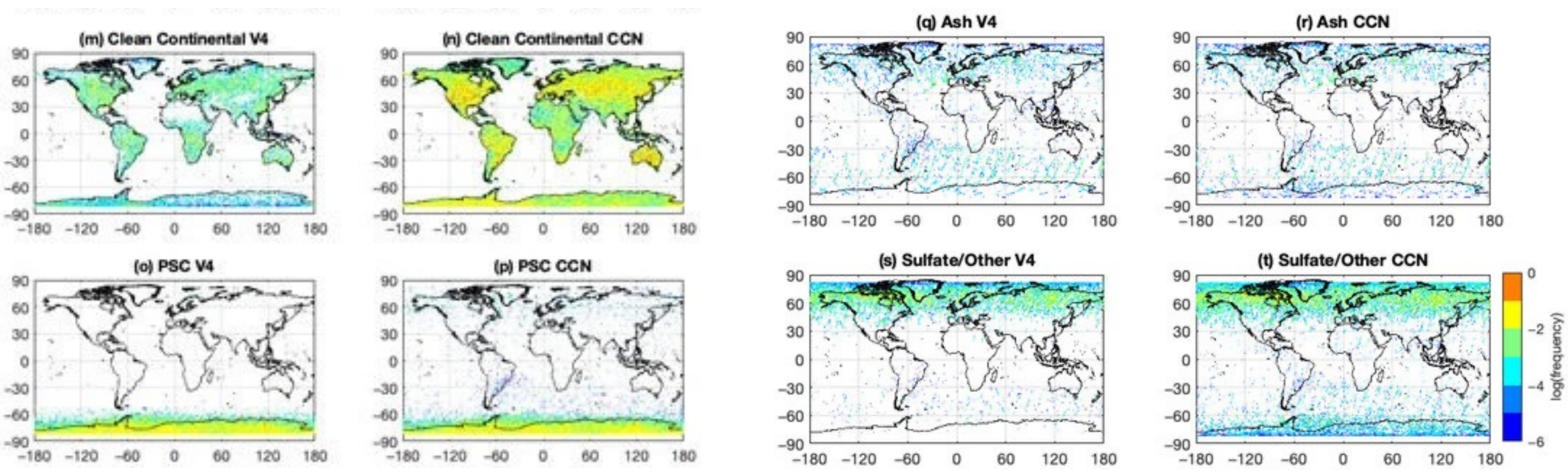
Zeng et al. 2020

Part I: ML Application to Remote Sensing Data



Zeng et al. 2020

Part I: ML Application to Remote Sensing Data

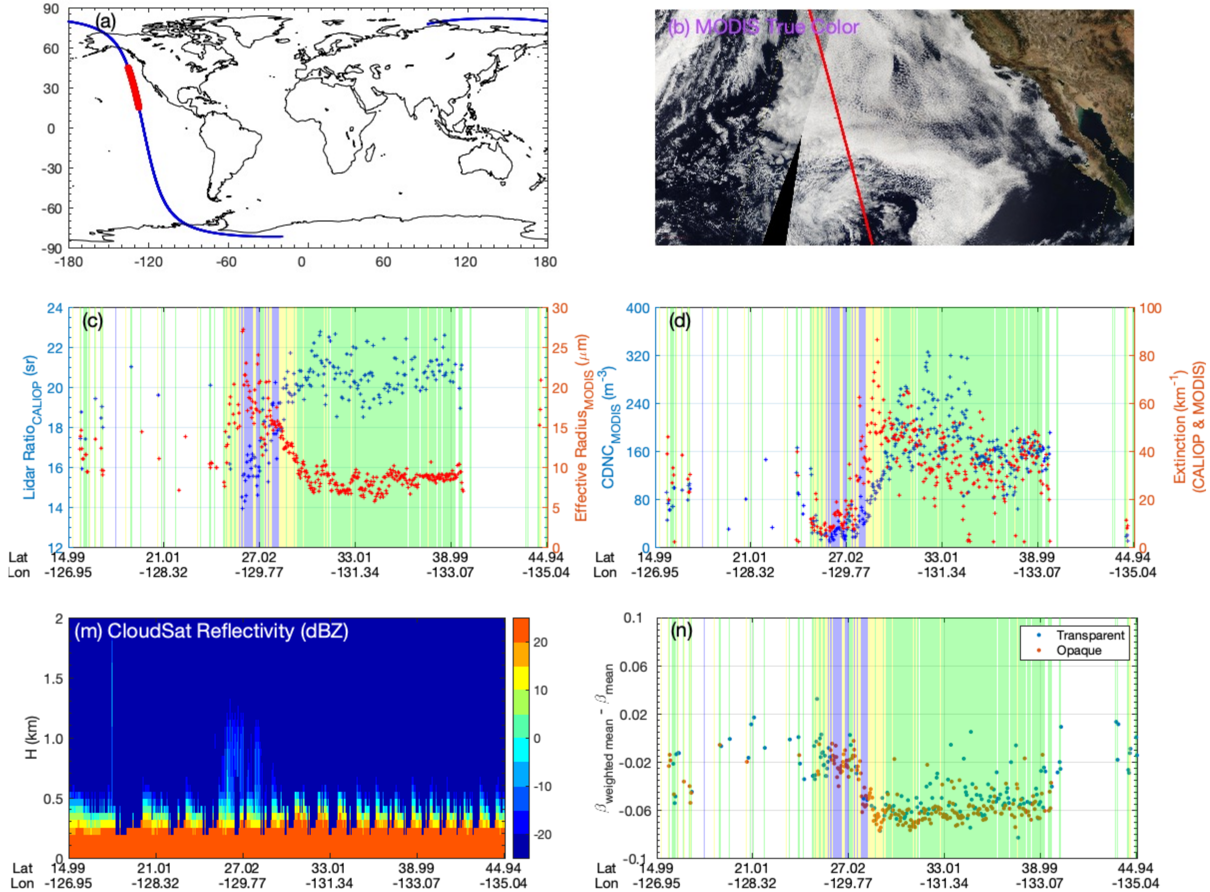


- ❑ Transition area -> mixing of two
- ❑ Add in model product indicator/passive sensor measurements
- knowing the history/dynamic background of aerosols
- lidar ratio
- ❑ Maybe start from smoke?

Zeng et al. 2020

Part I: ML Application to Remote Sensing Data

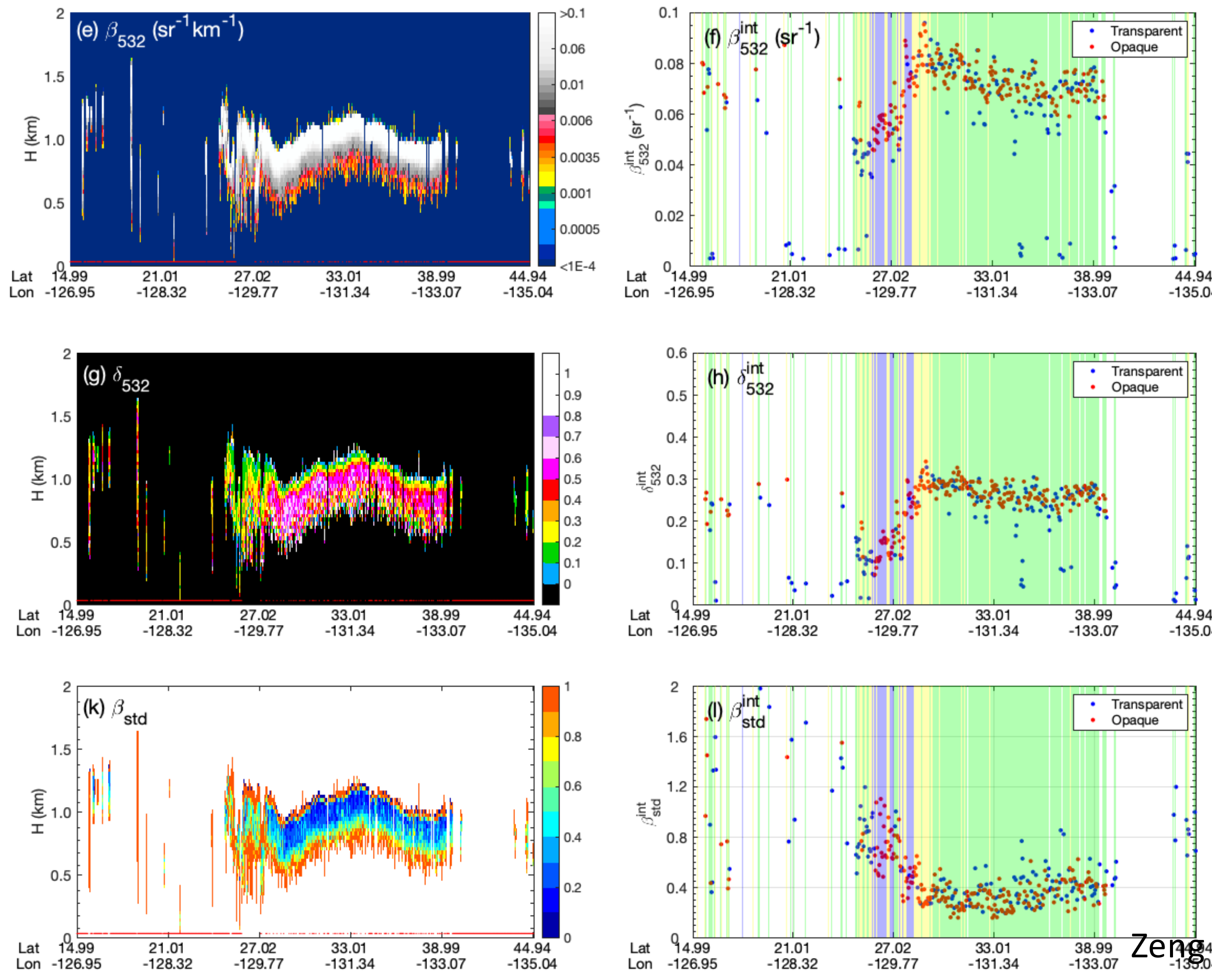
- can produce beyond operational products – new class, thanks to its diverse ML architectures;



Zeng et al. In preparation

Part I: ML Application to Remote Sensing Data

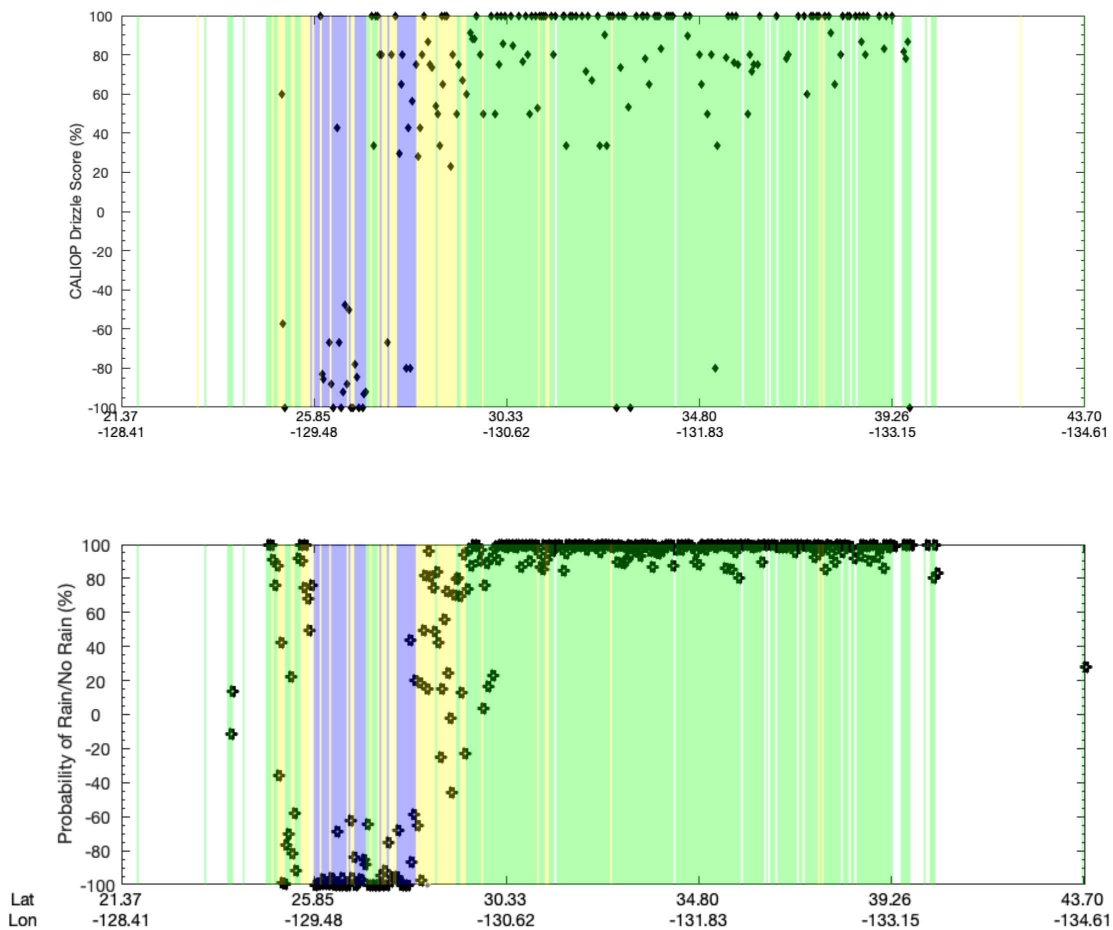
- ❑ **can produce beyond operational products – new class, thanks to its diverse ML architectures;**



Zeng et al. In preparation

Part I: ML Application to Remote Sensing Data

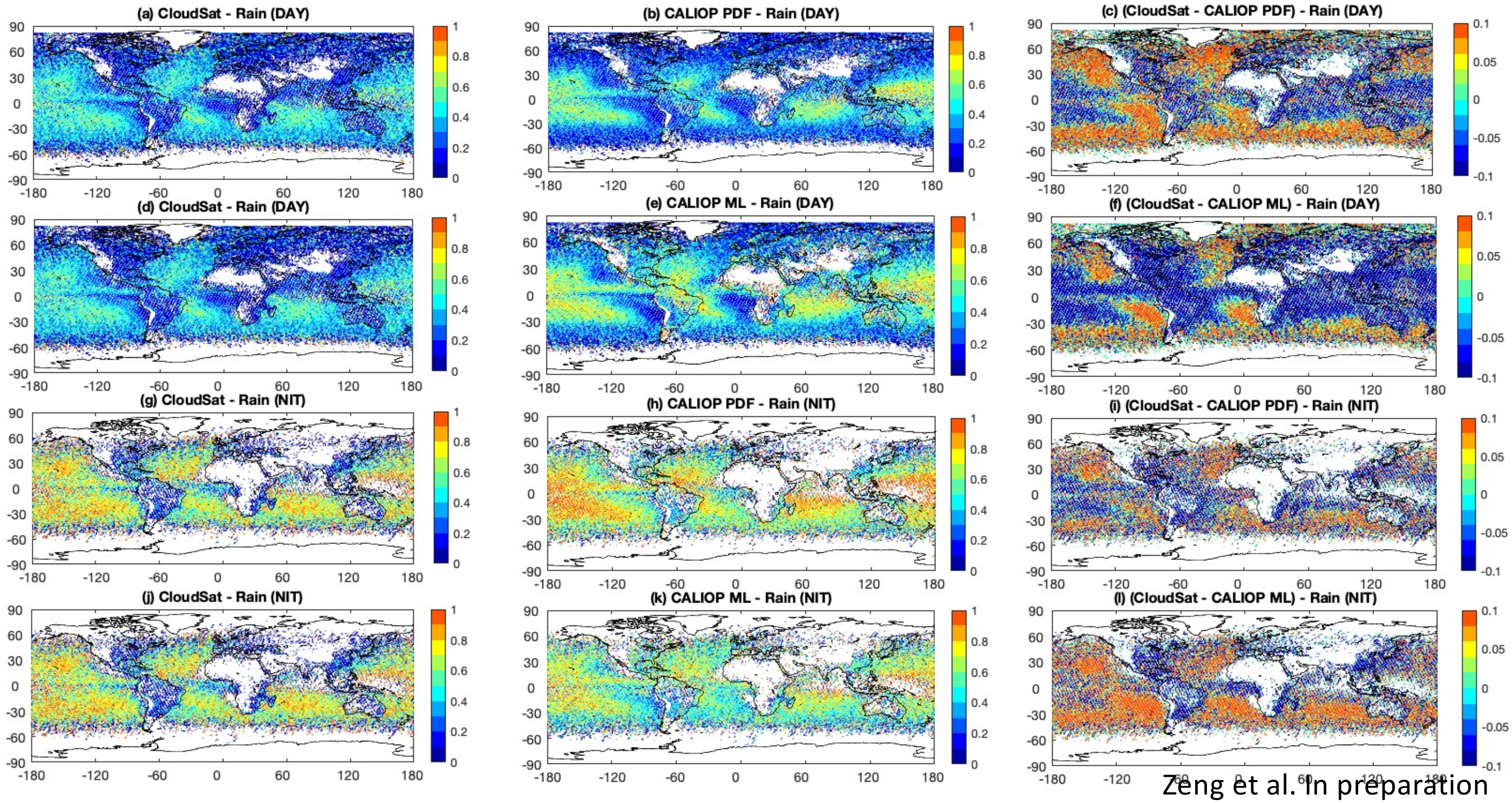
- can produce beyond operational products – new class, thanks to its diverse ML architectures:



Zeng et al. In preparation

Part I: ML Application to Remote Sensing Data

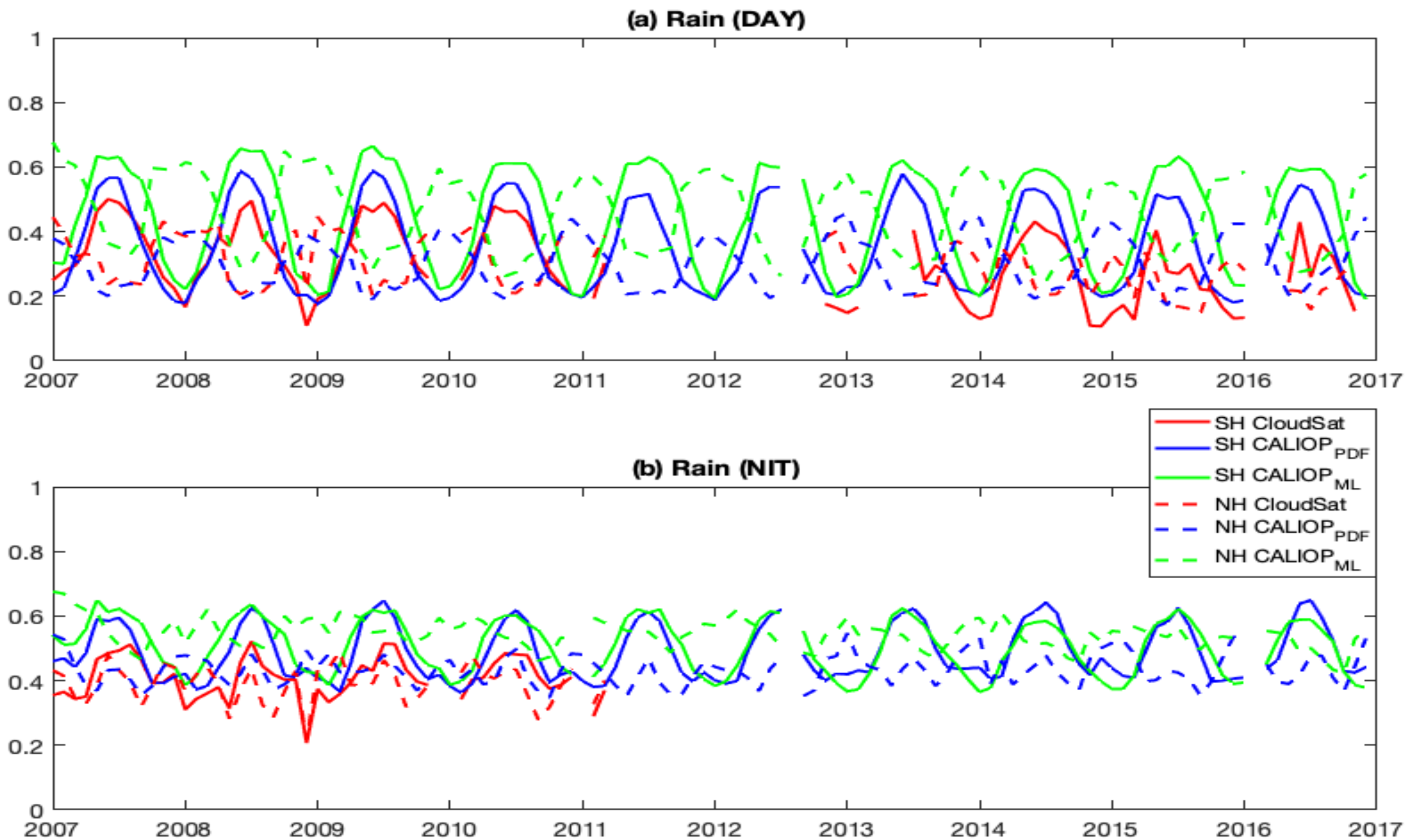
- can produce beyond operational products – new class, thanks to its diverse ML architectures;



Zeng et al. In preparation

Part I: ML Application to Remote Sensing Data

- ❑ **can produce beyond operational products – new class, thanks to its diverse ML architectures;**

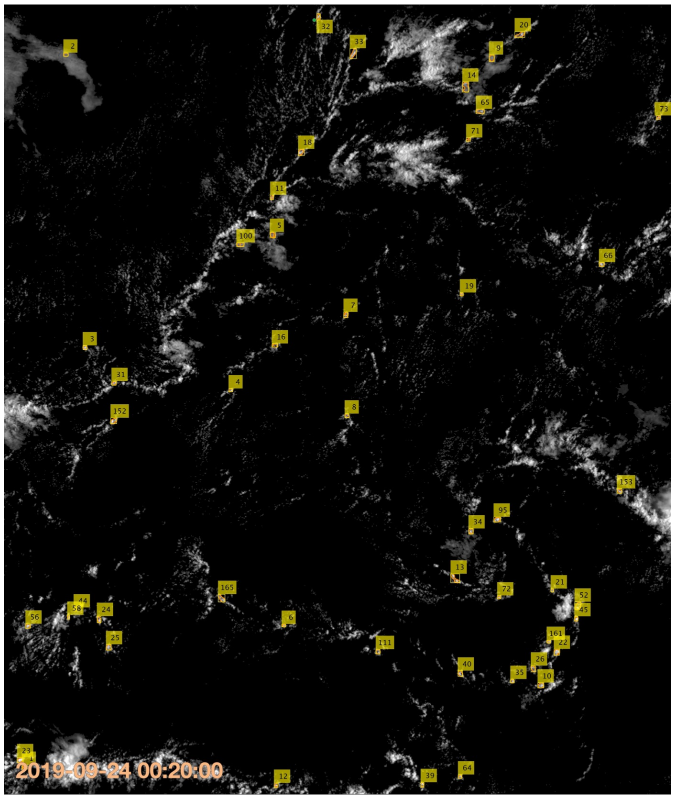


Zeng et al. in preparation

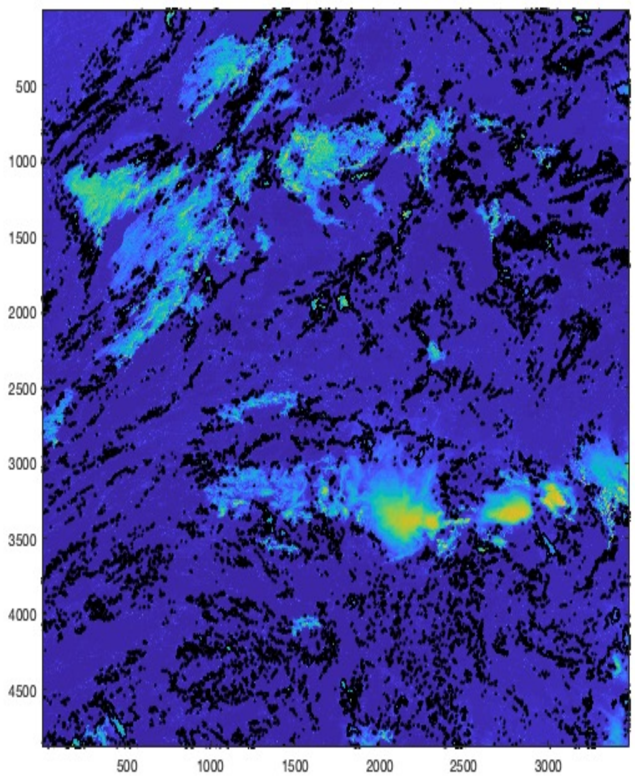
Part I: ML Application to Remote Sensing Data

- **can produce beyond operational products – new class, thanks to its diverse ML architectures;**

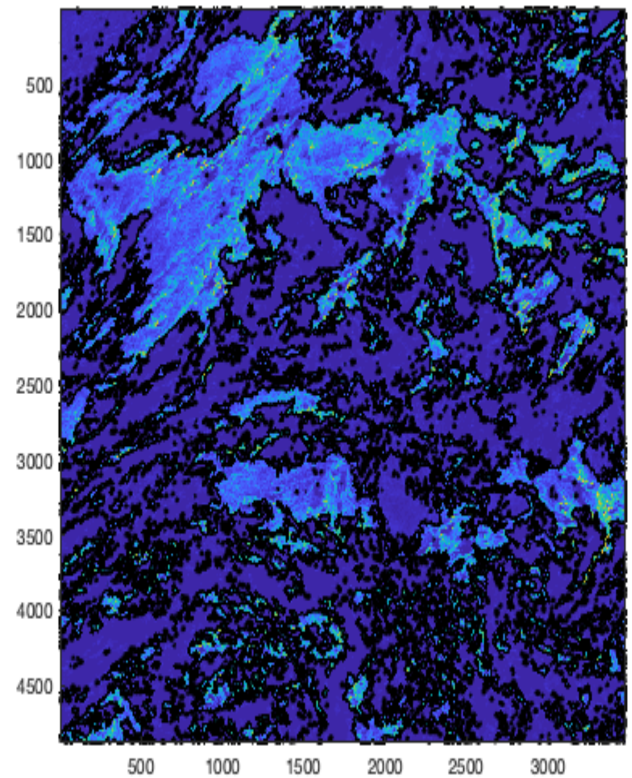
Cloud system identification /Rossby wave clouds



No filter



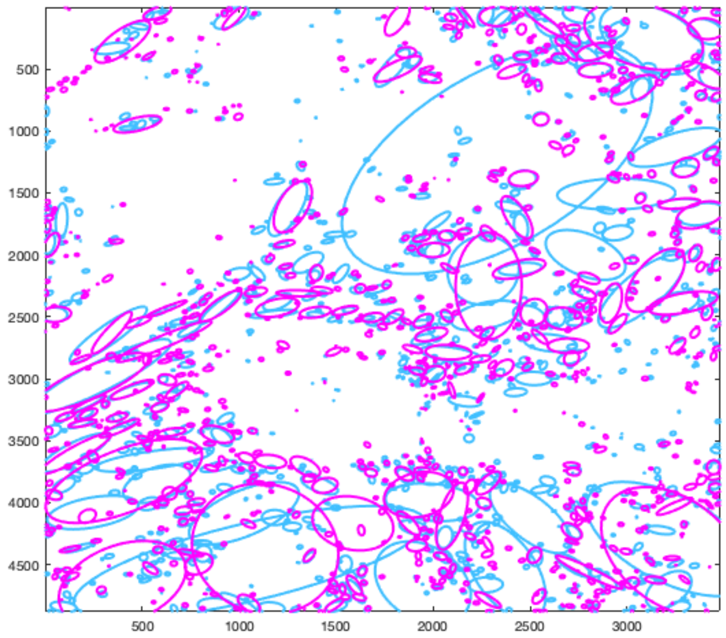
20X20filter

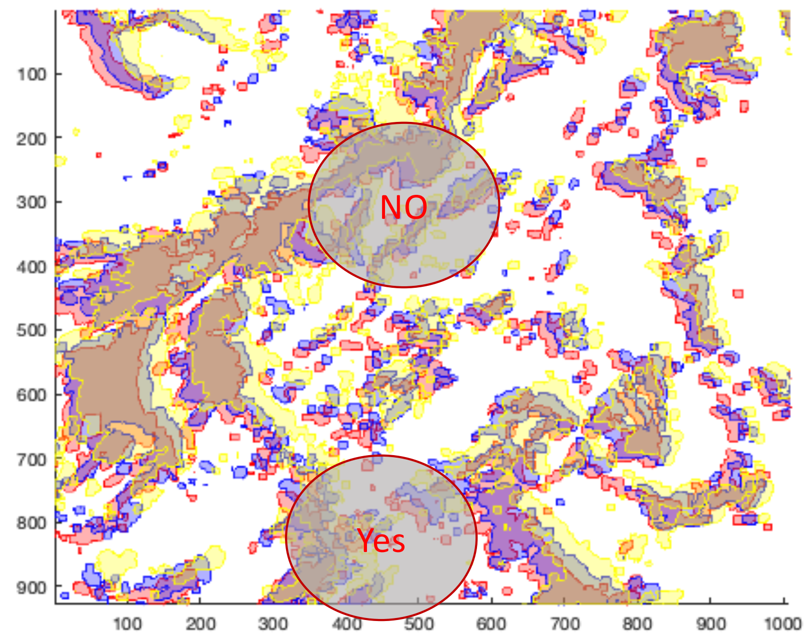


Part I: ML Application to Remote Sensing Data

- **can produce beyond operational products – new class, thanks to its diverse ML architectures;**

Cloud system identification /Rossby wave clouds





Summary and Perspectives of Part I

	Traditional Algorithm	ML
Diverse cognitive approaches	No	Yes
Fast speed (Real-time products)	Hard	Easy
Integration ability (Innovative products)	Hard	Easy

improve training outputs (ground truth):

- ❑ more in-situ, airborne measurements
- ❑ citizen science data
- ❑ simulations
- ❑ collocated other sensor measurements
- ❑ traditional operational products
- ❑ unsupervised learning products (Zeng et al. 2019)
- ❑ manual labels

Improve the training inputs:

- ❑ **collocated** diverse source measurements
- ❑ **accurate** measurements from better designed sensors

Improve/explore ML models:

- ❑ more complicated and diverse
- ❑ standardized model with diverse cognitive approaches for integration
- ❑ manage trained NN for integrations?

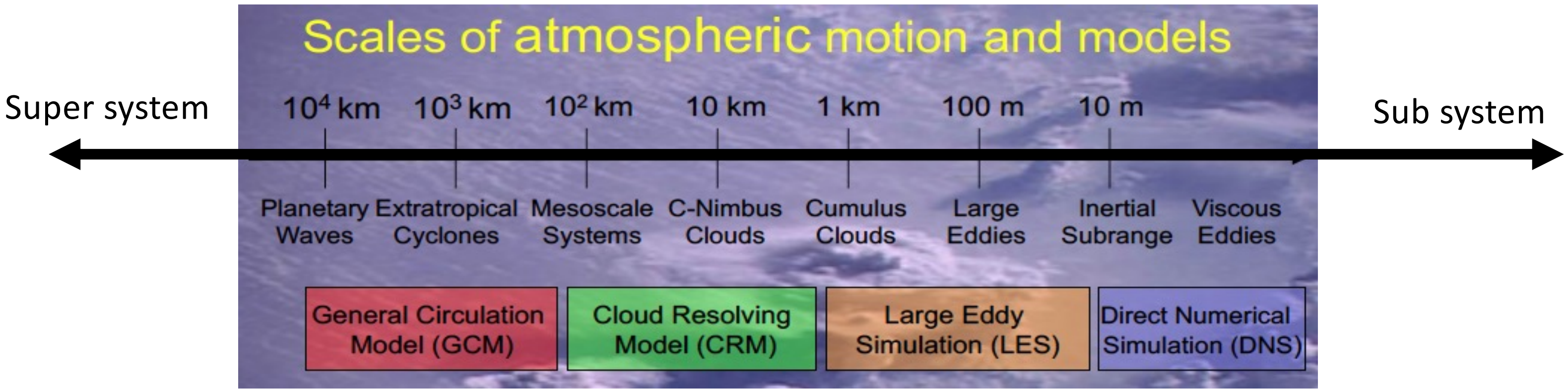
improve prediction outputs:

- ❑ Innovative Products: 4D Earth-atmosphere system
- ❑ Being able to bridge with industry products

Part II: ML Applications to Model (Cloud Parameterization – observation-based relationship)

ML can help to improve atmospheric models/simulations

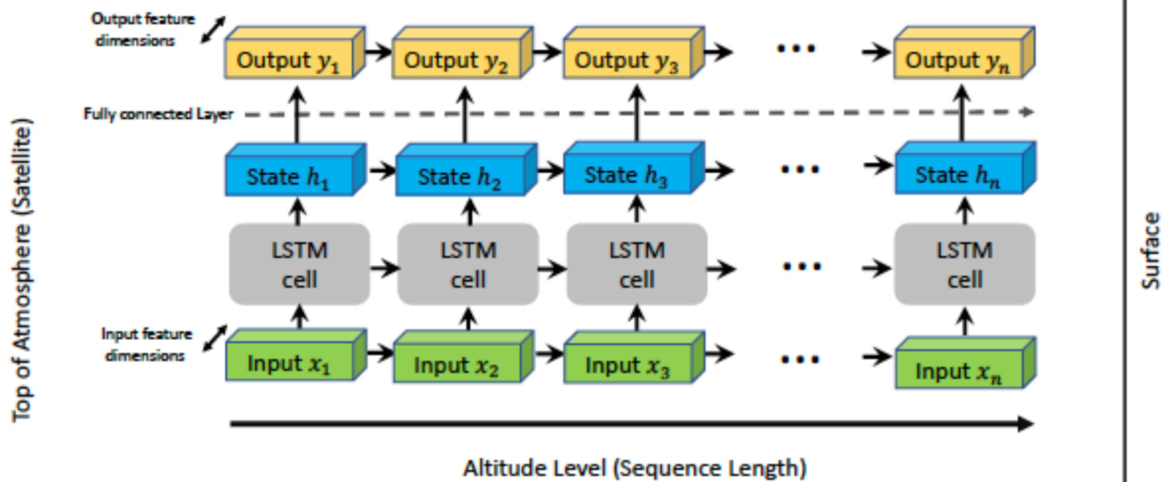
- We are able to obtain the **more accurate** atmospheric dynamic relationships with **fast** processing speed
- We are able to reconstruct and predict a **virtual Earth-Atmosphere system** with **fast** processing speed

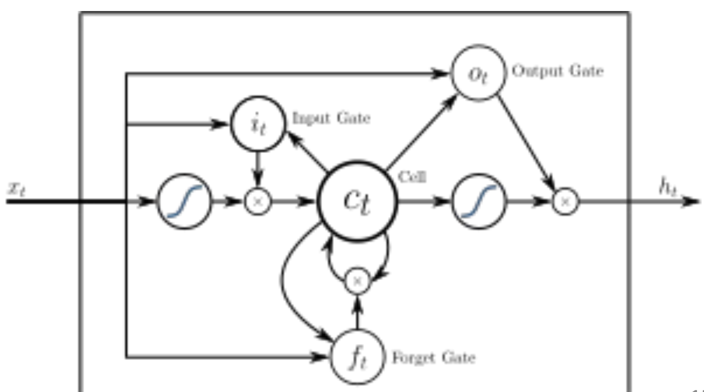


- ❑ Global climate models (GCMs) models are unable to resolve many important sub-grid or super scale physical processes

=> we use parameterization to reconstruct sub-grid cloud dynamical relationships, however traditional methods are too simple and fail to capture the complex nature of the associated sub-grid dynamic processes
=> we hope to improve cloud parameterizations as well as the cloud radiative feedback by using ML methods

Part II: ML Applications to Model (Cloud Parameterization-model)

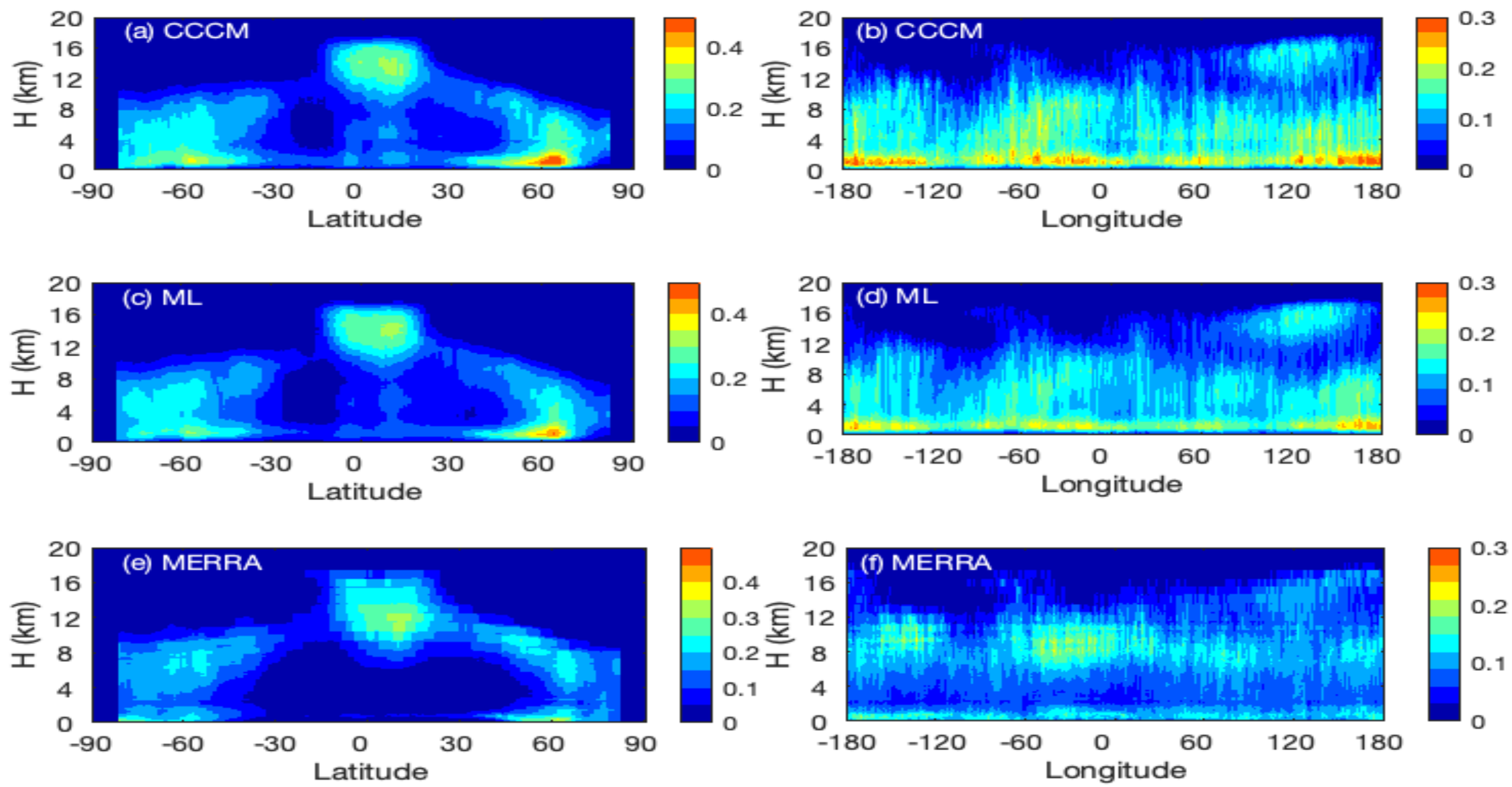




Inputs (MERRA)	Outputs (CCCM)
Wind Profiles	CCCM (CALIPSO-CloudSat-CERES-MODIS) Volumetric Cloud Fraction Profiles
RH/T/P Profiles	
Latent/sensitivity Heat Flux	
Surface T, RH, P	

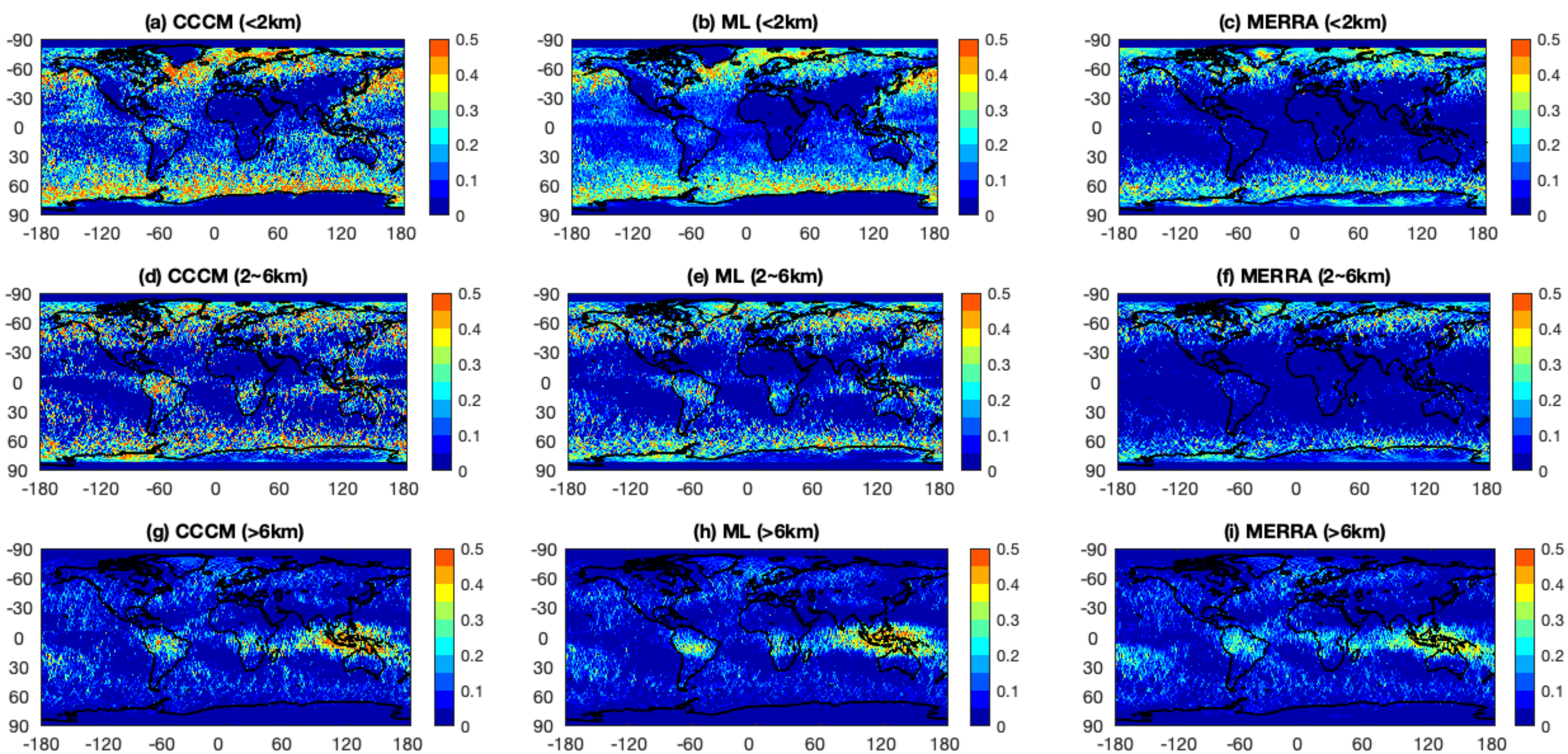
Training data	(2008) 90%
Validation data	(2008) 10%
Test data	(2009)

Part II: ML Applications to Model (Cloud Parameterization-Predictions)



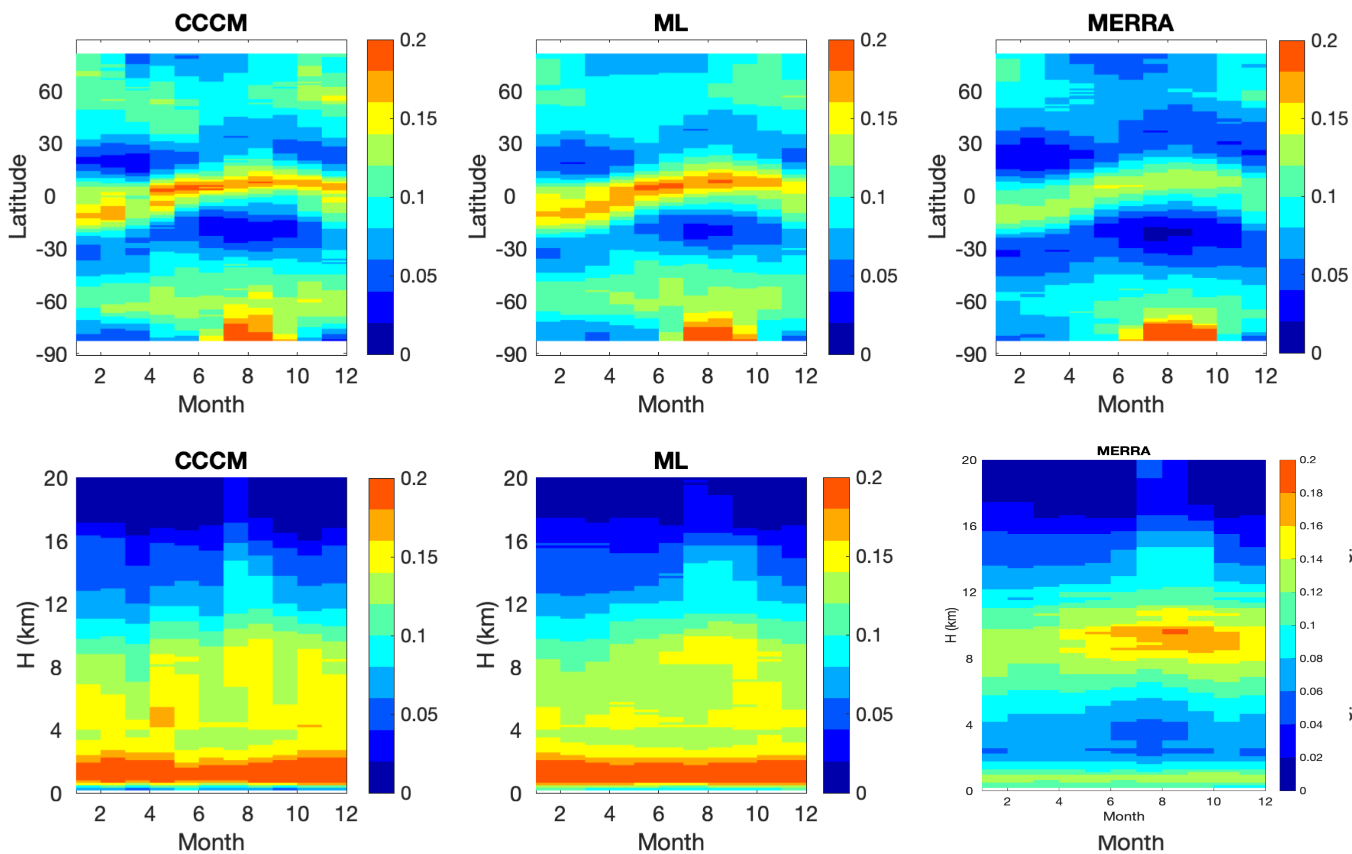
Zeng et al., in preparation

Part II: ML Applications to Model (Cloud Parameterization-Predictions)



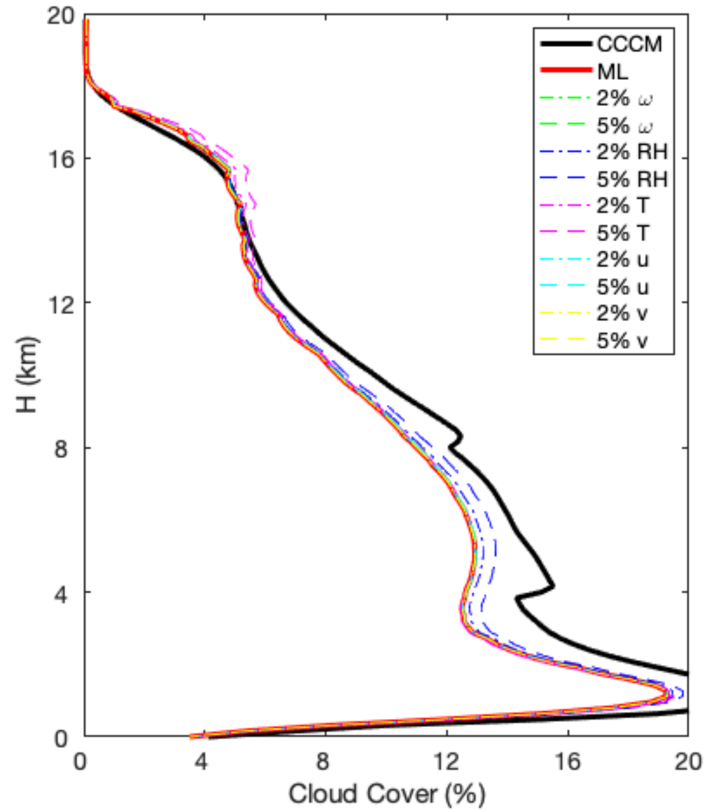
Zeng et al., in preparation

Part II: ML Applications to Model (Cloud Parameterization-Predictions)



Zeng et al., in preparation

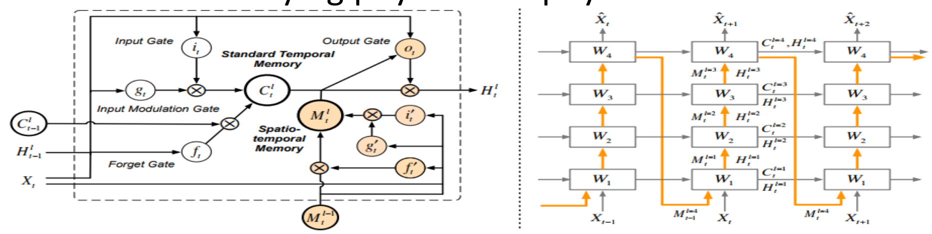
Part II: ML Applications to Model (Cloud Parameterization-Predictions)



Zeng et al., in preparation

Summary and Perspectives of Part II

- ML can learn the cloud dynamic relationships within the sub-grid system and summarize them using a “neural network”, which serves as a surrogate theory. In the future, the key task will be to develop customized ML models that more closely aligns with the underlying physics and physical constraints.



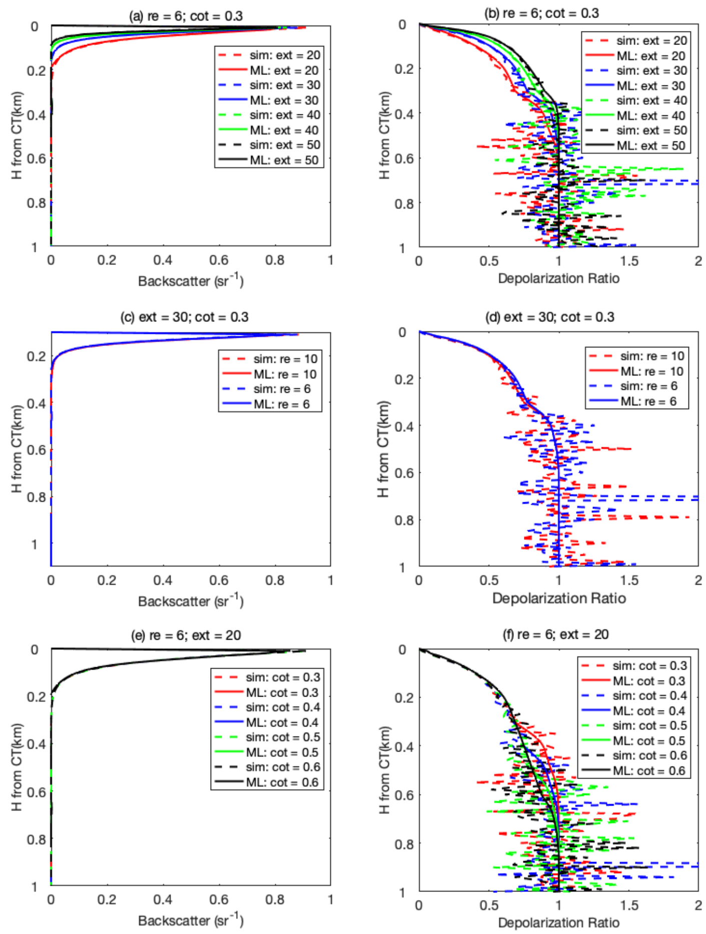
  - 5 axis – conv-rnn
  - Adding drizzle gate, ice cloud gate, seasonal variation gate etc.

- => reconstruct the aerosols dynamic relationships (i.e. fire)
  => finish reconstructing the clouds dynamic relationships with cloud microphysics
  => through integration of different sub-systems, ML provides an innovative way to reconstruct “metaverse” of entire atmosphere and climate system -> improve traditional models and help deeper understanding of cloud/aerosol physical processes as well as better predictions of climate changes.

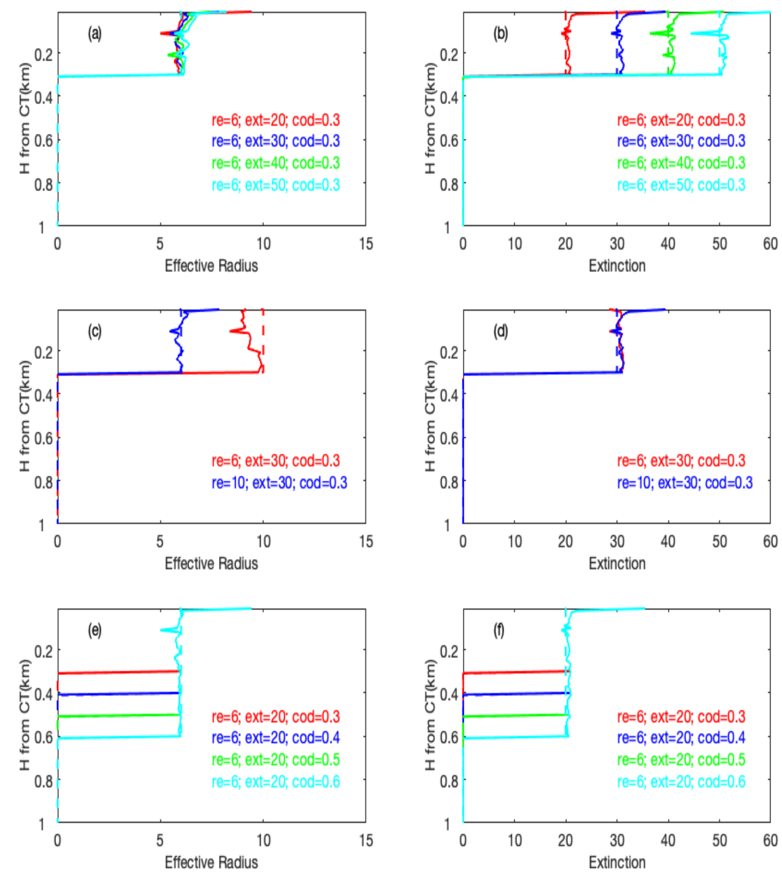
- Through ML, we can continue to reconstruct satellite measurements by integrating atmospheric dynamic models and radiative transfer model (Sensor Emulator);

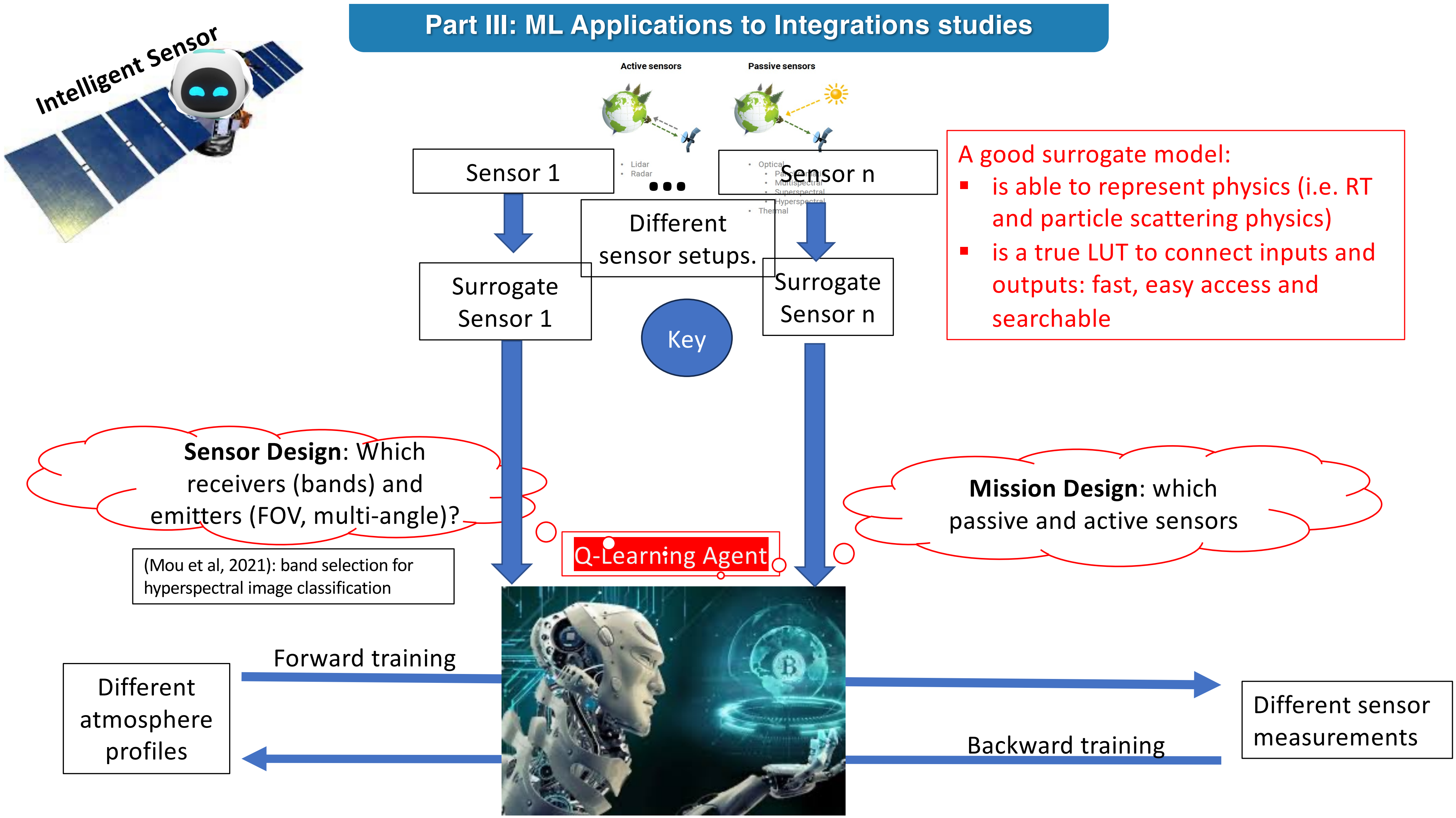
Part II: ML Applications to Model
(fast lidar MCRT– model-based relationships)

Forward training results

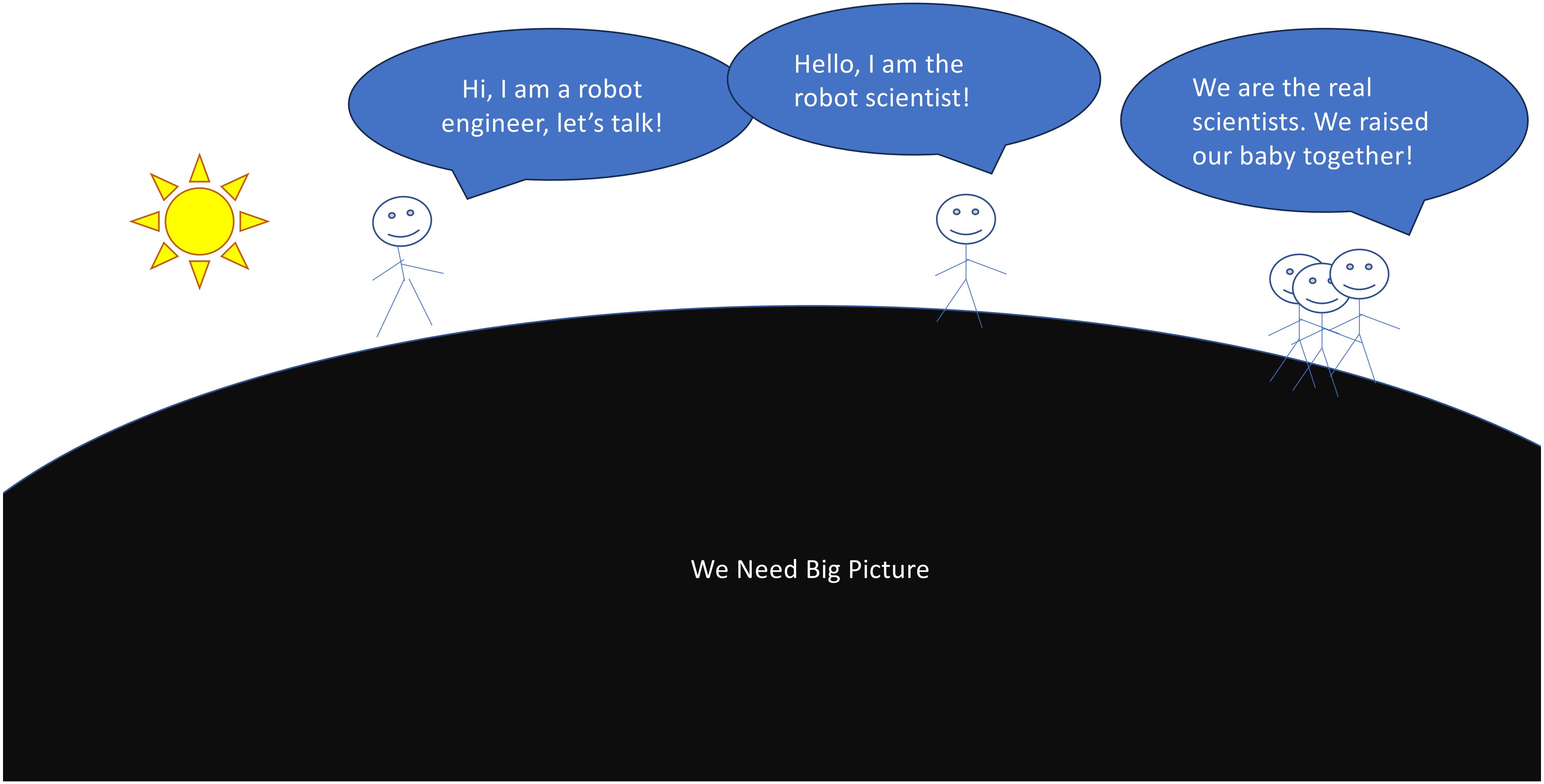


Backward training results




Part III: ML Applications to Integrations studies
Intelligent Sensor
Active sensors
Passive sensors
Lidar
Radar
Optical
Multispectral
Superspectral
Hyperspectral
Thermal
Sensor 1
Sensor n
Different sensor setups.
Surrogate Sensor 1
Surrogate Sensor n
Key
A good surrogate model:
is able to represent physics (i.e. RT and particle scattering physics)
is a true LUT to connect inputs and outputs: fast, easy access and searchable
Sensor Design: Which receivers (bands) and emitters (FOV, multi-angle)?
(Mou et al, 2021): band selection for hyperspectral image classification
Q-Learning Agent
Mission Design: which passive and active sensors
Forward training
Backward training
Different atmosphere profiles
Different sensor measurements

Thank You


Hi, I am a robot engineer, let's talk!
Hello, I am the robot scientist!
We are the real scientists. We raised our baby together!
We Need Big Picture