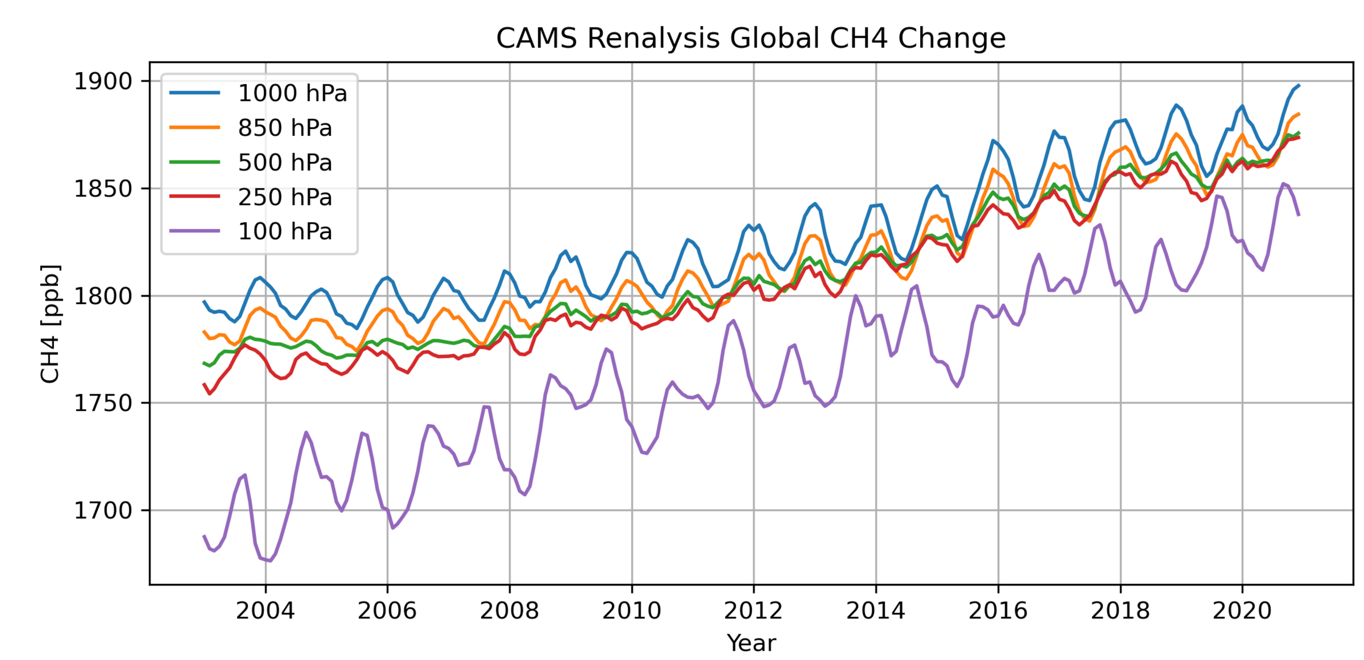
Estimating Tropospheric Methane from Cross-track Infrared Sounder (CrIS) Spectra using a Machine Learning Method

Likun Wang[1], Lihang Zhou [2], Satya Kalluri [2]

1. Earth System Science Interdisciplinary Center /Univ. of Maryland , College Park, MD Email : wlikun@umd.edu ; 2. NOAA Office of Low Earth Orbit Observations, Greenbelt, MD;

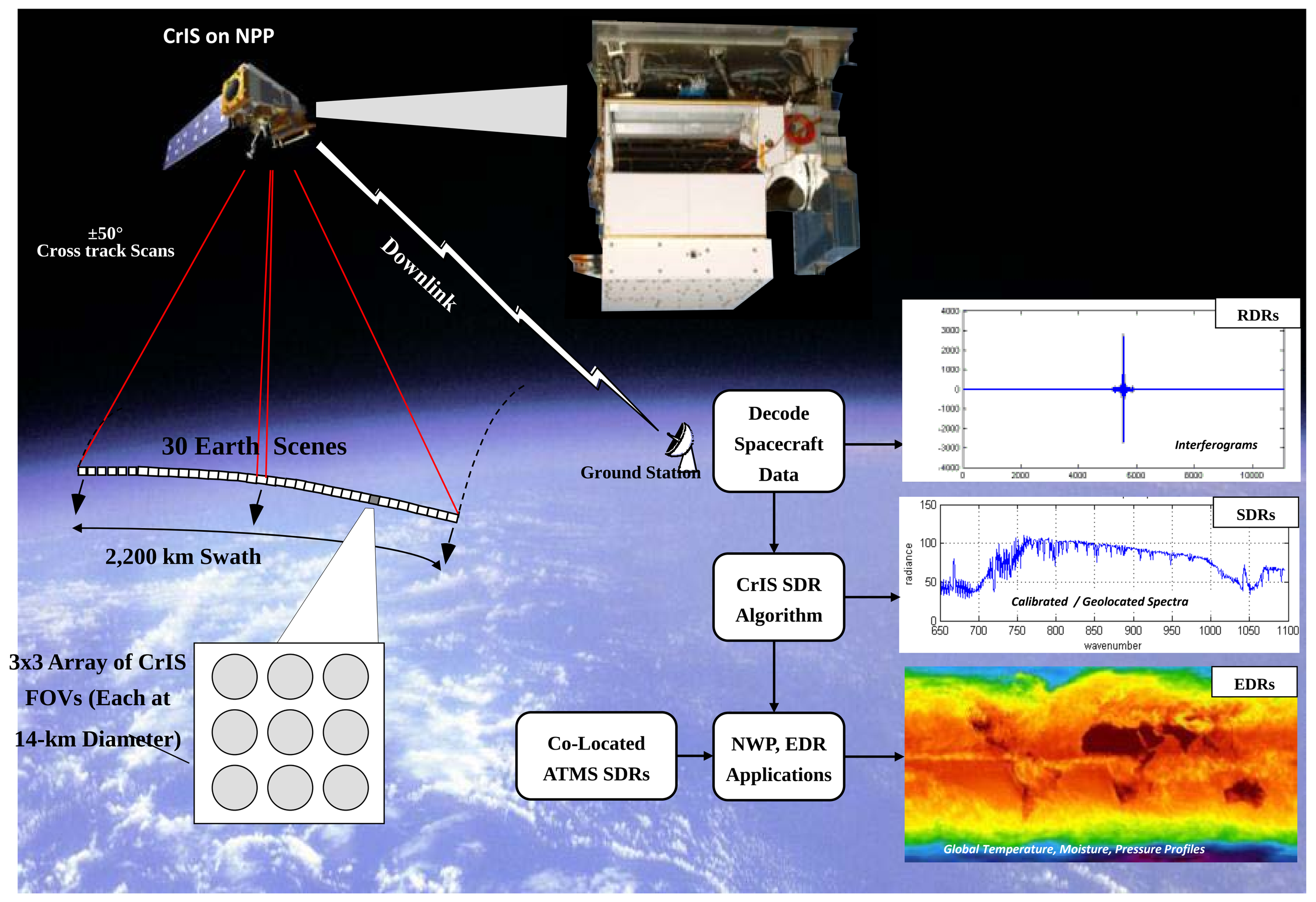
Abstract: The Cross-track Infrared Sounder (CrIS) on the Joint Polar Satellite System (JPSS) satellites is a Fourier transform spectrometer, providing sounding (temperature and humidity) and trace gas the of the atmosphere with 2221 spectral channels along three infrared bands. This poster presents a machine learning (ML) method to retrieve the methane in the middle troposphere from CrIS spectra. Different from traditional physical retrieval methods, the main idea of the ML-based approach is to use a neural network (NN) to approximate the complex inverse function that maps the target trace gas concentration to the CrIS radiances measured based on the training datasets. The authors utilize the Thermodynamic Initial Guess Retrieval (TIGR) dataset coupling with a three-dimensional chemical-transport model outputs (Global Greenhouse Gas Reanalysis) to build the training datasets through radiative transfer model calculations. The training dataset covers a large range not only of the concentration of the target trace gas but also of the auxiliary parameters on the state of the atmosphere. A deep residual neural network (ResNet) is trained to determine model parameters (weights for each node). The preliminary results are encouraging and indicate that the AI-based method has the ability to retrieve the tropospheric methane from CrIS data. In proposed future work, we will use the real CrIS spectral as inputs to estimate the tropospheric methene. These newly developed products will be compared with the existing sounding products from physical methods as well as those directly measured from ground and aircraft.

Importance of Tropospheric Methane



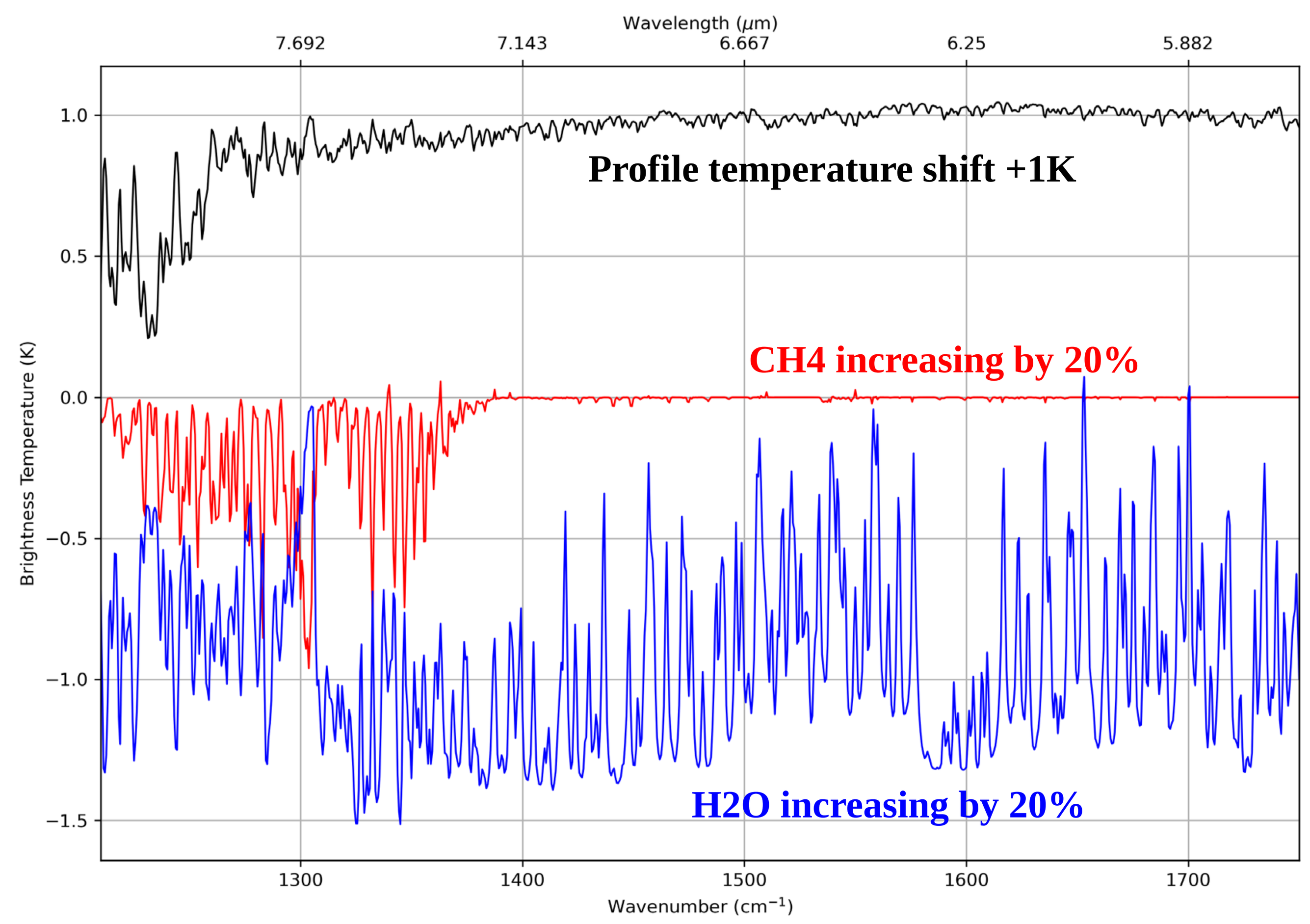
As a potent greenhouse gas, methane (CH_4) is the second-largest contributor to climate warming after carbon dioxide (CO_2) . The concentration of methane in the atmosphere has been continuously increasing over the past 30 years. Methane can come from both natural sources and human activities, and these sources contribute significantly to its presence in the atmosphere. Therefore, globally monitoring methane in the atmosphere is important.

Cross-Track Infrared Sounder Operational Concept



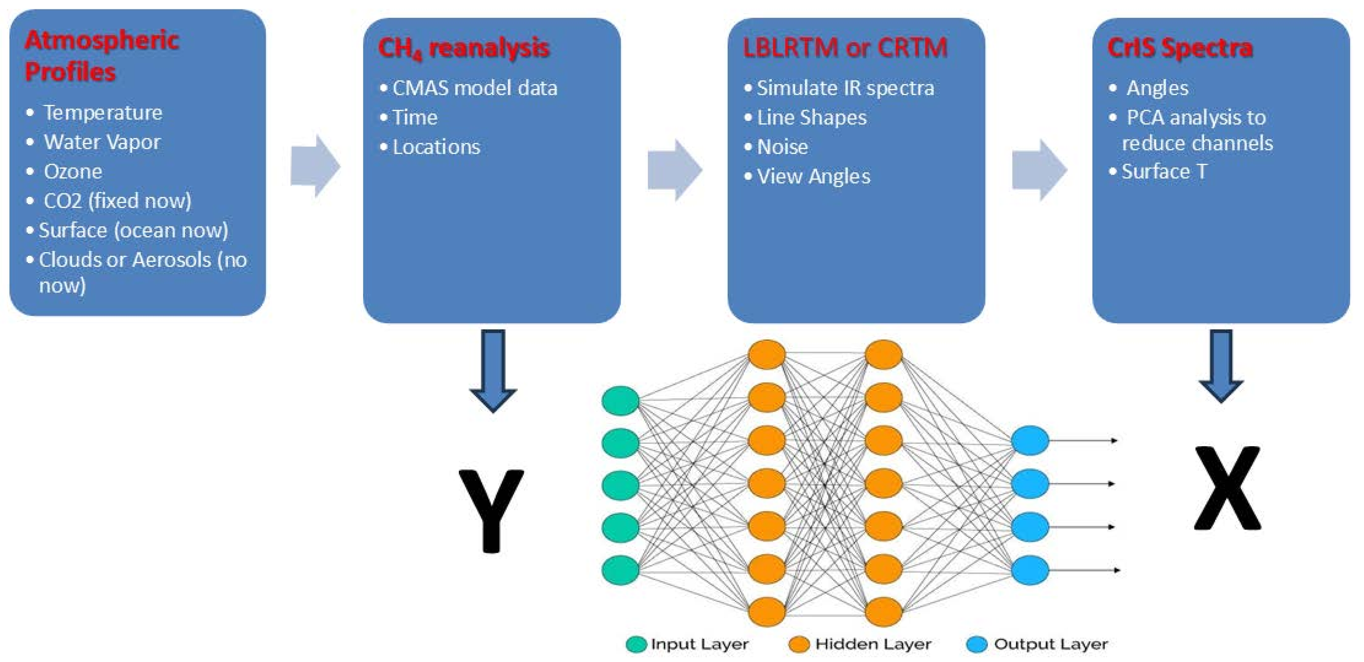
The Cross-track Infrared Sounder (CrIS) on the Joint Polar Satellite System (JPSS) satellites is a Fourier transform spectrometer, providing sounding (temperature and humidity) and trace gas the of the atmosphere with 2221 spectral channels along three infrared bands. CrIS radiance spectrum (without apodization) covers three IR bands from 650 to 1095 cm^{-1}, 1210 to 1750 cm^{-1}, and 2155 to 2550 cm^{-1} with spectral resolutions of 0.625 cm^{-1}. CrIS scans a 2200-kilometers swath width, plus or minus 50 degrees, with 30 Earth scene views. Each view consists of nine fields of view (FOVs) with 14-kilometer diameter spots in a 3x3 array.

Sensitivity of CrIS spectrum to Methane

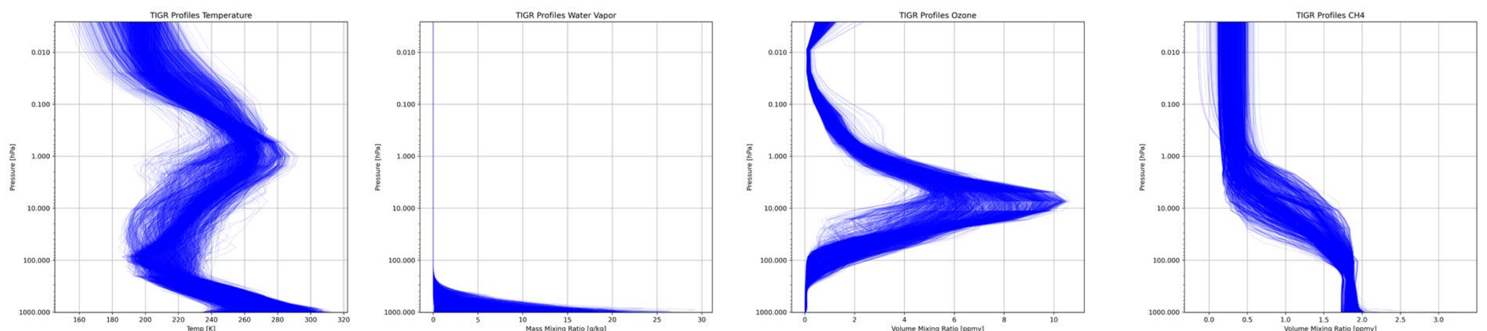


As shown in the above figure, the CrIS spectrum at the range of 1210-1400 cm^{-1} is sensitive to methane and thus can be used to retrieve methane in the atmosphere. The traditional approach to estimate methane from CrIS – so called physical methods - involve interpreting CrIS spectral radiance measurements using radiative transfer principles and optimal estimation methods by choosing several sensitive channels. However, methane absorption bands usually overlap with water vapor and other gases, requiring careful spectral separation (Figure 1). The interference of clouds and complex surface properties make the retrieval even more challenging. By contrast, a machine learning method by training a deep neural network has a capability to 1) capture the nonlinear relationship between methane and other environmental factors and 2) detect complex, subtle patterns in CrIS data that physical methods may miss.

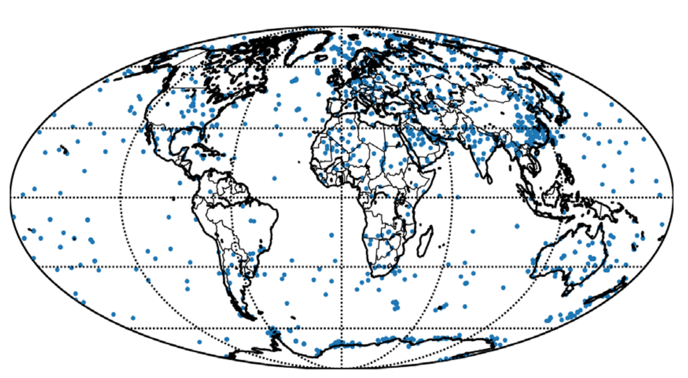
Training Datasets: Using Simulated Datasets



The key step is to build training datasets that encompass a wide range, not only in the concentration of methane but also in the auxiliary parameters describing the state of the atmosphere (atmospheric temperature and humidity profiles). We utilize the Thermodynamic Initial Guess Retrieval (TIGR) dataset coupling with a three-dimensional chemical-transport model outputs (Global Greenhouse Gas Reanalysis) to build the training datasets through radiative transfer model calculations as shown. The above figures shows the detailed step.

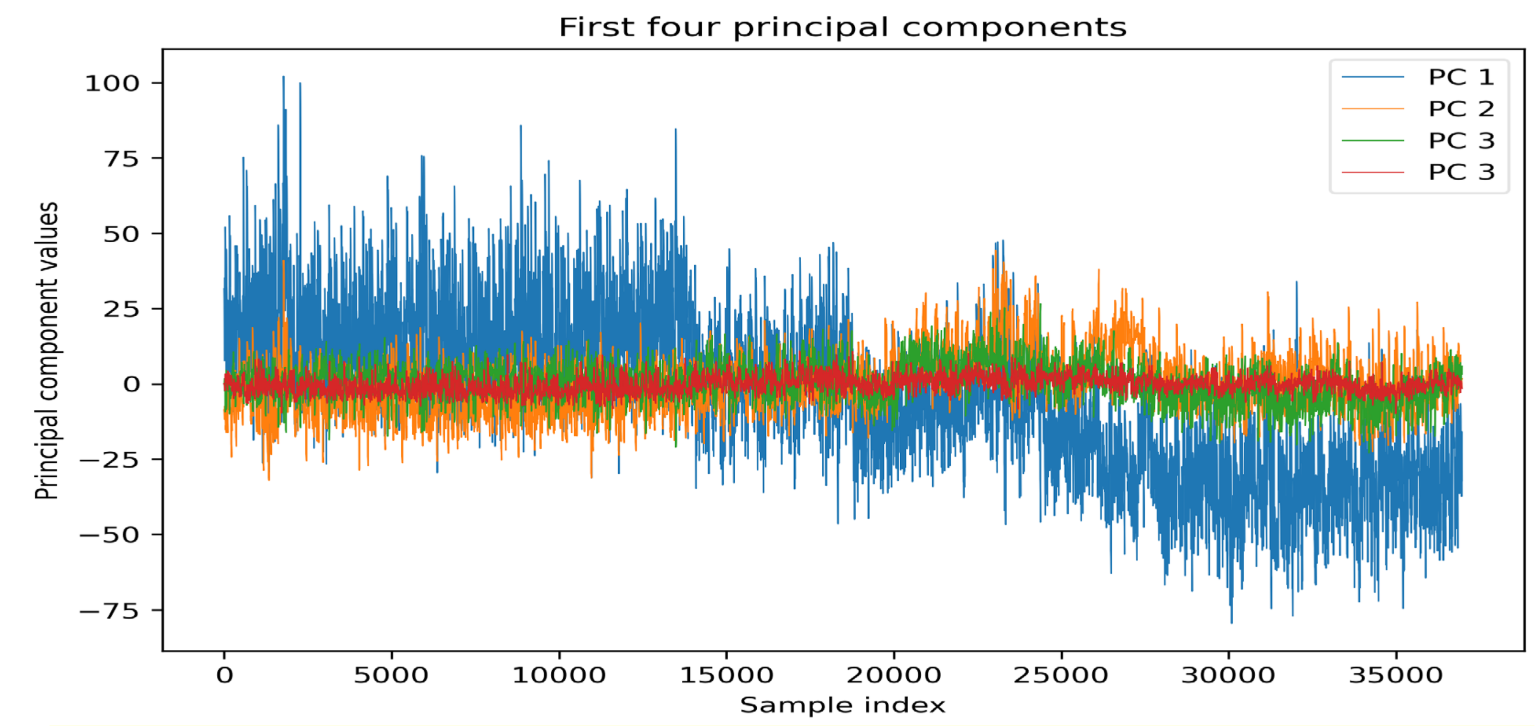


Temperature | Water Vapor | Ozone | CH_4 from CAMS



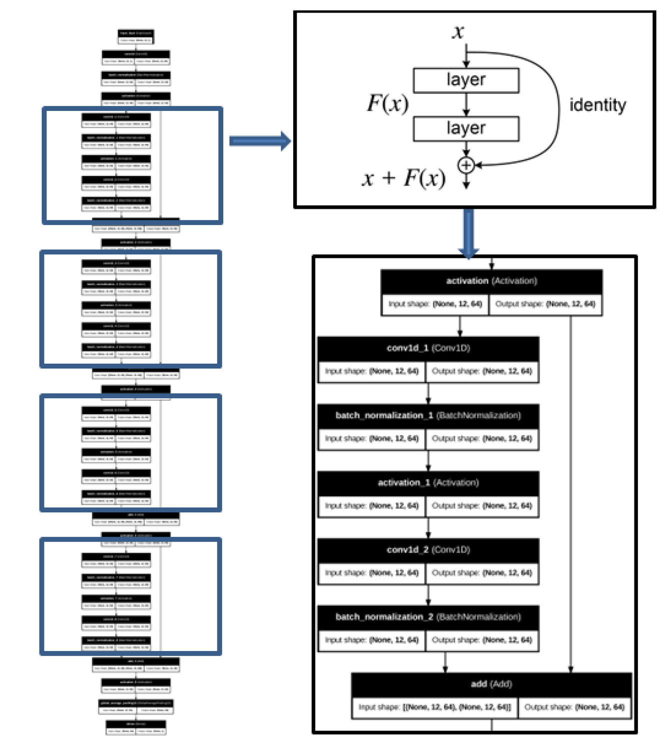
- Approximately 2,311 atmospheric profiles
- A wide range of atmospheric conditions from different climate zones around the world
- Benchmark for Radiative Transfer Models and retrieval algorithms
- Methane profiles from CAMS reanalysis interpolated at different locations and time (2000-2020)

PCA Analysis for CrIS Spectra



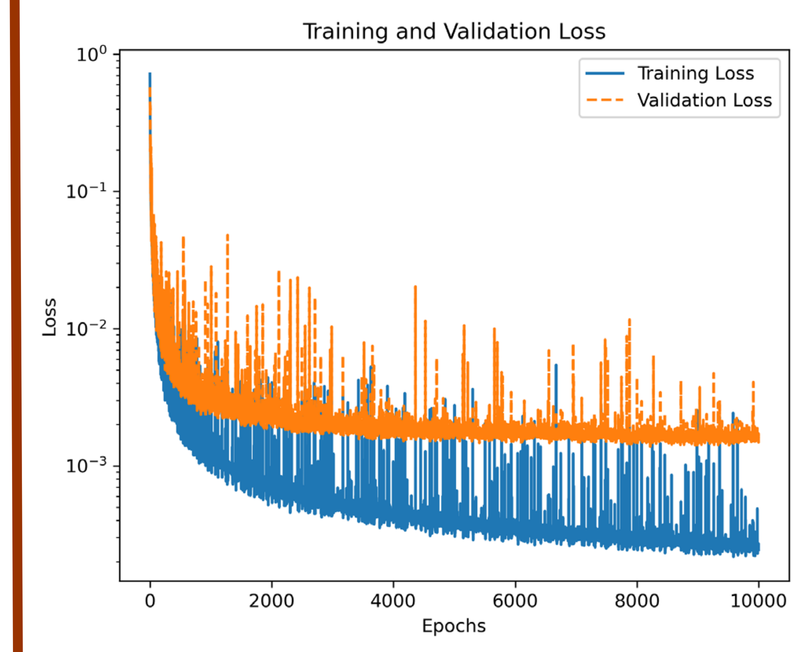
Different from physical methods that rely on limited sensitive channels, the ML method keeps the information from all CrIS channels through the Principal Component Analysis (PCA). Therefore, we use top 12 PC components as feature inputs.

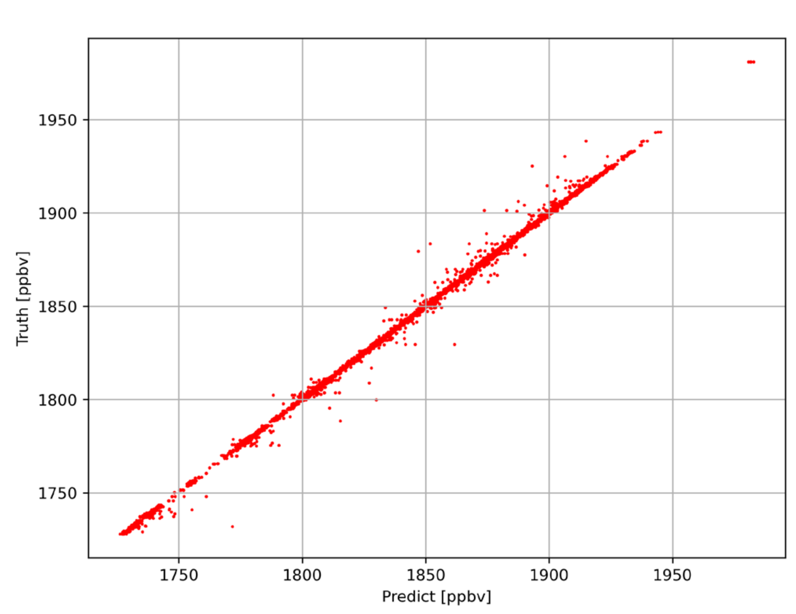
Model Description



- Model Architecture
  - **1D Convolutional Layers :** filters moving along 1D data:
    - 64 filters per CNN layer
  - **Residual Blocks:** Feature skip connections for efficient learning.
    - 4 Residual Blocks and 34 layers
  - Total params: 302,789 (1.16 MB)
- Key Features of ResNet
  - Residual Connections: help information flow, solving the vanishing gradient problem.
  - Deeper Networks with Fewer Issues
  - Efficient Feature Extraction: Learns complex patterns and relationships

Model Training and Preliminary Results





The left figure shows the model's loss over the course of training epochs, in which the blue and orange lines represent the training and validation losses. Both training and validation losses decrease over epochs, indicating that the model is learning and improving. The right one shows the validation results based on the remaining 20% dataset. Statistically, the mean squared error (MSE) is 3.141 ppbv and the Mean Absolute Error (MAE) is 0.733 ppbv between predicated and true values.

Conclusion and Future Work: This paper presents a ML method to retrieve the methane in the middle troposphere from CrIS spectra. Different from traditional physical retrieval methods, the main idea of the ML-based approach is to use a neural network (NN) to approximate the complex inverse function that maps the target trace gas concentration to the CrIS radiances measured based on the training datasets. The preliminary results are encouraging and indicate that the AI-based method has the ability to retrieve the tropospheric methane from CrIS data. In proposed future work, we will use the real CrIS spectral as inputs to estimate the tropospheric methene. These newly developed products will be compared with the existing sounding products from physical methods as well as those directly measured from ground and aircraft.