Ian Kelk, Yash Nahar, Chandrakant Shendarkar,
Christina Wang, Guilherme Zuanazzi

Harvard Extension School
Data Science Capstone
NASA Project 2
July 10, 2024

---

cloud2cloud

Generating 3D-Mesh Cloud Height Estimations Using Computer Vision

In collaboration with NASA

Table of Contents

Chapter 1- Introduction

1.1. Purpose Statement

Measuring cloud-top heights and their evolution in time is extremely important for NASA. Accurate cloud height data enhances the calibration and validation of radiometric instruments on satellites, ensuring precise measurements of Earth's atmosphere. This in turn helps in understanding the interaction between clouds and radiation, which is needed for studying Earth's energy balance and climate. By accurately measuring cloud heights, NASA can refine radiometric algorithms, improve the quality of satellite data, and enhance the overall accuracy of atmospheric observations, resulting in better climate models and weather predictions.

NASA aims to reduce the cost of future missions but also solve limitations with cloud height measurements from LiDAR installations with limited field of view. This project aims to develop a predictive model to estimate cloud height utilizing high-definition images captured from the Fly's Eye GLM Simulator (FEGS) along with (LiDAR) data.

1.2. Problem & Setting

The Geostationary Lightning Mapper (GLM) is a lightning detection sensor onboard the GOES-16 satellite. To validate the GLM, NASA developed the FEGS. The FEGS is a multi-spectral radiometer system that features 30 radiometers and an HD camera. It is mounted aboard the NASA ER-2 aerial laboratory, a plane that operates at altitudes ranging from 20,000 feet to 70,000 feet, above 99% of the Earth's atmosphere. During a flight campaign in 2017 to validate the recently launched GOES-16, the ER-2 flew a suite of instruments that included FEGS and the Cloud Physics LiDAR (CPL) for measuring the height of clouds. NASA's current objective is to reduce the cost of future aerial meteorological missions; cameras are lower cost, more prevalent, and have much larger 90-degree fields of view.

While LiDAR is effective at accurately measuring the height of cloud-tops, it only provides a single point value corresponding to where the lasers are pointed. This can provide a ground truth for the center of the cloud-top heights, allowing us to train and validate a model to estimate cloud-top height using the HD images from FEGS.

The main problem can be defined as: *How can we use high-definition images from the FEGS combined with LiDAR data to accurately estimate cloud-top heights and create a 3D-mesh cloud projection?*

1.3. Subproblems

1.3.1. Estimate Cloud-Top Height Mesh using object tracking and parallax

- Identify cloud-top features to track across fields of view.
- Adjust for lens distortion and viewing geometry.
- Use parallax and aircraft velocity to calculate cloud-top height.

1.3.2. Predictive Model for Cloud Height Estimation

We aim to utilize high-definition images captured from FEGS along with metadata about the aircraft and Light Detection and Ranging (LiDAR) data for the development of the predictive model for cloud-top height estimation.

By combining high-definition (HD) images from FEGS with their metadata and LiDAR, we can use the wide viewing angles of cameras and the accurate distance measurements from LiDAR to construct a full and accurate three-dimensional picture of cloud structures.

1.3.3. Atmospheric Condition Correlation

Investigate the relationship between the generated point clouds and various atmospheric conditions.

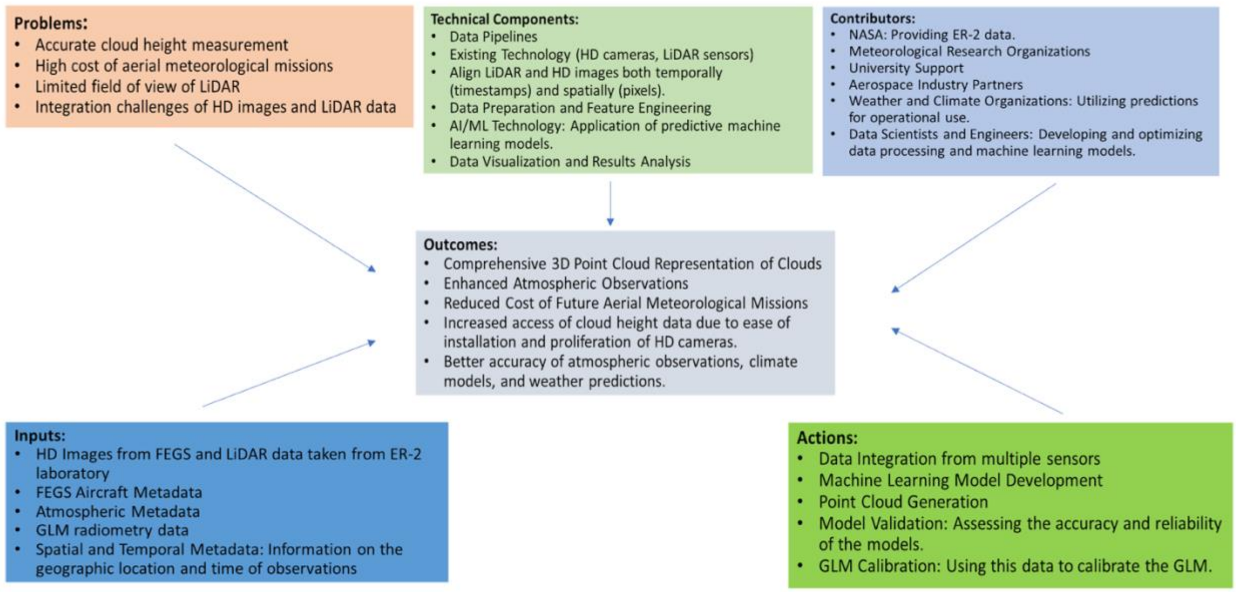
- Develop models to recognize specific weather patterns or phenomena from the point cloud data.
- Create methods to integrate the point cloud data into existing climate models for improved accuracy.

1.3.4. GLM Data Validation Using Point Clouds

Develop methods to utilize the generated point cloud data to validate and improve the GLM measurements.

- Correlate lightning events detected by the GLM with the 3D cloud structures represented in the point clouds.
- Use detailed cloud structure information from the point clouds to analyze and potentially improve the sensitivity of GLM lightning detection.
- Develop techniques to use the point cloud data to help identify and reduce false-positive lightning detections in the GLM data.
- Investigate how different cloud structures and opacities, as represented in the point clouds, affect GLM lightning detection accuracy.

1.4. Theoretical & Conceptual Framework



Note: This project will be valuable to the ultimate goal of improving atmospheric observations, but our Capstone focuses on innovation and improvement of an AI/ML solution for predicting point clouds of cloud tops.

1.5. A Priori Hypothesis

The project posits that by integrating high-definition images with LiDAR data and applying machine learning models, we can generate accurate and comprehensive 3D point cloud representations of atmospheric cloud-tops.

1.6. Variables & Key Concepts

1.6.1. Image Quality Assessment

Independent Variables:

- Image Resolution: Higher-resolution images provide more detailed information, improving the model's ability to identify cloud structures.
- Brightness Level: Adequate brightness is needed so that cloud features are visible and distinguishable.
- Blurriness Level: Lower blurriness levels contribute to clearer images, which enhance the model's accuracy.
- Distance from Ground: This affects the perspective and scale of the captured images, influencing the model's interpretation of cloud height and structure.

Dependent Variable

Image Quality Score: A custom score assessing the overall quality of the image for model consumption.

1.6.2. Machine Learning Model for Point Cloud Generation

Independent Variables:

- LiDAR Point Data: Provides the height of clouds at specific points, serving as ground truth for model training.
- Pixel Intensity: This represents the brightness and color information of each pixel in the HD images, contributing to the detection of cloud boundaries and features.
- Spatial Coordinates: The location of each pixel in the image is essential for aligning LiDAR data with corresponding image points.
- FEGS Aircraft Navigation Metadata: track angle, drift angle, altitude, pitch, yaw, roll, heading, speeds (indicated, true, ground), bank, aircraft vertical velocity, etc.

Dependent Variable:

Point Cloud Accuracy: A metric evaluating the precision and completeness of the generated point cloud in representing the actual cloud structure.

1.6.3. Evaluation of Predictive Models for Cloud Height Estimation

Independent Variables

- Model Architecture:
    - Type of Machine Learning Model: Algorithms such as Convolutional Neural Networks (CNNs) for image data and PointNet for LiDAR data.
    - Layer Configuration: Convolutional layers for image features and point cloud layers for LiDAR data.
    - Training Parameters: Learning rate, batch size, and number of epochs.

Feature Engineering

- Raw Data Features: Pixel intensity, spatial coordinates, and LiDAR point height.
- Derived Features: Spatial features capture the position and distribution of cloud points.
- Data Augmentation: Techniques like rotation, scaling, and noise addition.

Dependent Variables

Performance consistency across different weather conditions and cloud types, and accuracy on unseen data through cross-validation and independent testing.

1.7. Assumptions, Delimitations, & Limitations

This project assumes that:

- **Data Quality**: High-definition images and LiDAR data captured by the NASA ER-2 aerial laboratory will be of high quality and sufficient for training and validating the machine learning models.
- **Timestamp Alignment**: Timestamps for HD images and LiDAR measurements can be accurately synchronized to ensure precise alignment of the datasets
- **Data Source Consistency**: Cloud structures captured during different flights will exhibit enough consistency to allow for effective model training and validation.

- **Field of View Requirements**: The 90-degree field of view for the camera will be sufficient to capture a comprehensive representation of the cloud structures needed for model training and validation.
- **Flight Metadata Accuracy**: Accurate metadata with respect to the ER-2 aircraft carrying the FEGS is made available to us, such as altitude, pitch, yaw, roll, heading, speeds (indicated, true, ground), bank, and aircraft vertical velocity.

The project will focus on data collected from the NASA ER-2 aerial laboratory, specifically the HD images and LiDAR data obtained during specific flight missions. The study will only include the geographical areas that the ER-2 flights covered, which might not include all cloud types and atmospheric conditions on a global scale. Furthermore, the project will utilize data collected over a defined period, limiting the analysis to specific seasonal and weather conditions.

Integrating HD images and LiDAR data presents technical challenges, including aligning data spatially and temporally as well as handling discrepancies in data quality and resolution. Noise is also a relevant factor, which can arise from motion blur, lighting issues, or inherent noise from LiDAR sensors, potentially affecting the data quality and model accuracy. The ability of the machine learning models to generalize to unseen data and different atmospheric conditions may be limited, impacting the robustness of point cloud generation. The availability of computational resources necessary for processing sizable datasets and training computationally intensive machine learning models may also place restrictions on the project. Lastly, variations in environmental conditions, such as cloud density, lighting, and weather patterns, may affect the parallax calculations and, consequently, the accuracy and consistency of the generated point clouds.

1.8. The Importance of the Study

Understanding cloud-top heights and their temporal evolution is vital to comprehending cloud-top processes and their interactions with local meteorology and climate. HD cameras, being cost-effective and easy to deploy, can be readily installed on CubeSats, weather satellites, and space stations, and they are already ubiquitous in many current systems. If HD image data could be used to estimate cloud-top height, this widespread deployment of cameras will provide global monitoring capabilities for cloud-tops and significantly enhance the accuracy of weather predictions by offering more detailed and frequent observations. It would also help to refine the radiometric algorithms that satellites, such as the GOES-16, use to gather data. Ultimately, this research will support better climate models, improved weather predictions—especially extreme weather forecasts—and more cost-effective future aerial meteorological missions.

Chapter 2 – Literature Review

To properly consider the cloud2cloud (c2c) project within the broader field of atmospheric remote sensing and inform a methodological approach, we examine a range of relevant research papers. The selected papers cover important topics related to measuring cloud height, such as satellite imagery, stereoscopic and multi-angle imaging methods, LiDAR applications, and 3D cloud structure reconstruction methods. We both explore current capabilities and limitations in cloud height estimation, find new methodologies for data integration and validation, and identify key challenges that the c2c project must address.

The review is structured thematically, focusing on three areas:

1. **Cloud Height Measurement and Estimation**: Various techniques for estimating cloud heights using instruments such as LiDAR, all-sky cameras, and airborne sensors.
2. **3D Cloud Structure Reconstruction**: Methods for reconstructing three-dimensional cloud structures from two-dimensional imagery.
3. **Calibration and Validation**: Ensuring the accuracy of cloud measurements.

2.1 Cloud Height Measurement and Estimation

2.1.1. Motivation and Challenges

Accurate cloud base and top height measurements are needed in various applications in meteorology, aviation, and climate science. Mussa et al. (2022) and Bedka et al. (2007) both discuss the significance of precise cloud height data for improving the accuracy of atmospheric observations and models. Mussa et al. investigated cloud height and coverage because they affect the analysis of cosmic rays at the Pierre Auger Observatory. Bedka et al. checked the accuracy of cloud-top pressure products made by satellites using LiDAR data from aircraft.

Borisov et al. (2023) and Nied et al. (2023) explore the use of artificial neural networks and convolutional neural networks, respectively, to estimate cloud heights from imagery. Borisov et al. (2023) introduce a novel approach using parallax effects observed with wide-angle cameras and neural networks to estimate cloud base heights, addressing the limitations of expensive and complex equipment in maritime conditions. Nied et al. (2023) focus on detecting high-level clouds above aircraft using CNNs, which is necessary to improve the accuracy of aerosol measurements in remote sensing. They also note that existing calculation methods are too slow to be used in near-real time scenarios, such as where the aircraft must avoid flying through or below such clouds for aerosol measurement. Nied et al. (2023) also notes photo imagery as being superior to some other sensor methods in detecting the presence of clouds. Both papers demonstrate the effectiveness of machine learning techniques in overcoming the limitations of traditional methods and providing cost-effective and adaptable solutions for cloud height estimation.

McGill et al. (2002) developed advanced instruments for high-resolution cloud measurements. The paper discusses the Cloud Physics LiDAR (CPL), designed to overcome the limitations of previous LiDAR systems by using state-of-the-art technology. The paper emphasizes the challenges of high-altitude measurements and the need for high temporal and spatial resolution while maintaining sensitivity. The

CPL's innovative design, featuring a high-repetition-rate laser and photon-counting detectors, addresses these challenges, demonstrating significant improvements in data accuracy and reliability.

2.1.2. Proposed Solutions

Mussa et al. (2022) and Bedka et al. (2007) address the limitations of traditional instruments and methods for providing accurate cloud height measurements. Existing solutions like the rotating-beam ceilometer (RBC) and standard satellite-based instruments often struggle with issues such as sensitivity to noise, limited spatial resolution, and difficulty in differentiating between cloud and aerosol layers. These limitations result in inconsistent and unreliable data, which can significantly impact the accuracy of atmospheric and climatic models.

Mussa et al. (2022) propose a dynamic threshold-based algorithm for cloud detection using elastic multiangle LiDARs at the Pierre Auger Observatory. They use a flexible threshold to find big changes in backscatter that show there are clouds, which overcomes the problems with differential-based methods.

Similarly, Bedka et al. (2007) focus on validating satellite-derived CTP products using high-resolution, aircraft-based CPL data. They leverage the precise and continuous measurements from the CPL as a benchmark for assessing satellite data accuracy. The analysis compares cloud-top height (CTH) products from the GOES-12 Imager and Sounder with CPL measurements obtained during the ATReC field campaign. By evaluating various cloud types and conditions, the paper highlights the strengths and weaknesses of the satellite measurements, providing evidence that GOES-12 Imager data generally aligns better with CPL measurements, particularly for mid-level and low clouds.

Borisov et al. (2023) and Nied et al. (2023) explore the use of advanced machine learning techniques to estimate cloud heights from imagery. Traditional methods often involve expensive and complex equipment like LiDARs and aircraft, which are impractical in unstable conditions, such as maritime environments, or may rely on subjective visual assessments that introduce high variability and uncertainty.

Borisov et al. (2023) address these challenges by using the parallax effect observed with two wide-angle cameras mounted apart to estimate the cloud base height (CBH). By capturing synchronous optical images of the sky from different angles, the parallax phenomenon allows for accurate CBH calculation. The analysis uses the Sail Cloud v.2 system to make observations in the field, neural network-based transformations to fix for camera positioning errors, and the pretrained SuperGlue graph-based neural network (GNN) to match key points in the images. The paper validates its claims by comparing the CBH estimates with ERA-5 reanalysis data, demonstrating consistency, particularly for cirrus and cumulus clouds. The study found that the SuperGlue GNN demonstrated superior performance compared to SIFT, delivering a significantly higher number of matching keypoints.

S92 BORISOV et al.

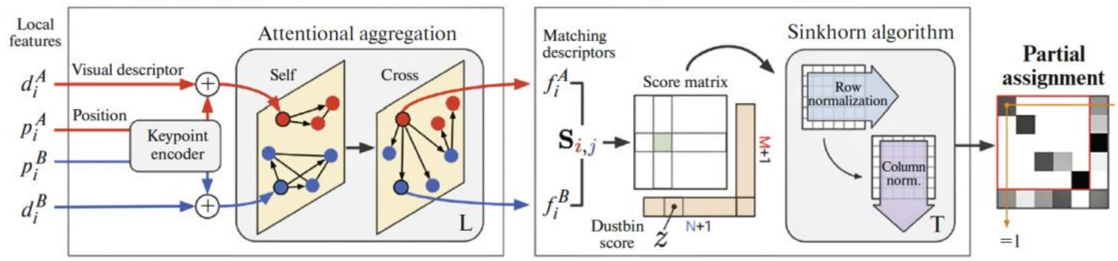


Fig. 8. Original SuperGlue neural network architecture [11].

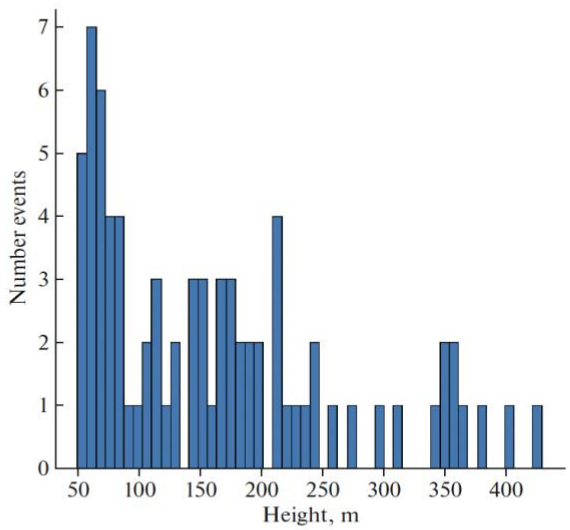


Fig. 9. Cloud base height estimation using SIFT.

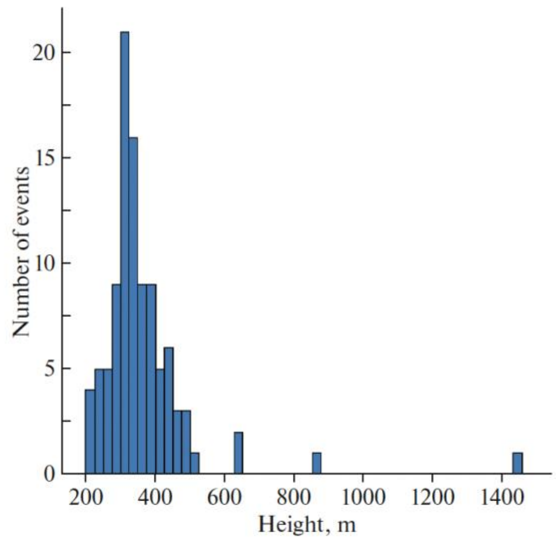


Fig. 12. Cloud base height estimation using SuperGlue.

Nied et al. (2023) leverage convolutional neural networks (CNNs) to detect high-level clouds above aircraft using upward-looking cameras. Existing solutions often rely on satellite-based instruments or additional onboard sensors, which are costly and complex. Instead, they trained CNN models on labeled images from airborne missions to learn cloud detection with high accuracy. As part of the setup, images mostly from NASA's ACTIVATE and CAMP2Ex missions were used to train the models. They were then checked against human-labeled datasets and compared to the SPN-S instrument's cloud mask product. The paper demonstrates the effectiveness of the CNN models, achieving a detection accuracy of 96%, and proves its claims through experiments and validation.

McGill et al. (2002) describe the development of the CPL to address the limitations of previous LiDAR systems, such as multiple scattering issues and limited dynamic range. Existing solutions, like the Cloud

LiDAR System (CLS), suffered from these issues, along with the aging hardware and practical limitations of analog signal acquisition. The idea behind the CPL is to use high-repetition-rate laser and photon-counting detectors to improve eye safety, compactness, reliability, and data accuracy. The setup involved deploying the CPL on a high-altitude ER-2 aircraft during the SAFARI 2000 field campaign, with multiwavelength measurements to provide detailed profiles of clouds and aerosols. The paper proves its claims through field deployment results, comparative analysis with other instruments, and detailed high-resolution profiles.

2.2 3D Cloud Structure Reconstruction

2.2.1. Motivation and Challenges

A key problem identified across multiple papers, including those by Hasler (1981) and Yu et al. (2019), is the accurate measurement and reconstruction of cloud heights and structures. Traditional methods relying on two-dimensional (2-D) satellite observations and infrared-based techniques have significant problems. Hasler (1981) notes that infrared methods suffer from low horizontal resolution and inaccuracies due to various assumptions about cloud emissivity and atmospheric conditions. Yu et al. (2019) emphasize that 2-D data are insufficient for precise radiation transmission calculations and climate modeling, indicating the need for three-dimensional (3-D) data to improve these predictions.

The importance of accurate cloud measurements is universally acknowledged across all the papers. Hasler (1981) highlights the necessity for precise cloud height data to understand and predict severe weather phenomena like hurricanes and thunderstorms. Yu et al. (2019) extend this by emphasizing the role of 3-D cloud structures in enhancing the accuracy of radiation transmission calculations and climate change forecasts.

The challenge of this research lies in the complexity of accurate 3-D reconstruction. Hasler (1981) points to the synchronization issues between satellite images and the geometric irregularities that complicate precise height measurements. Yu et al. (2019) discusses the substantial data requirements and the complex algorithms needed for multi-angle observations, which are essential for defining cloud structures accurately. Hadjitheophanous et al. (2010) talk about how hard it is to do real-time 3-D reconstruction on general-purpose processors. They say that the technique needs to be able to do efficient computations to work in embedded and mobile systems.

Existing solutions often fall short due to their reliance on limited observational angles and computational inefficiencies. Hasler (1981) notes the inaccuracies and low resolution of infrared methods, while Yu et al. (2019) points out the insufficient coverage satellites like CloudSat and CALIPSO provide. Hadjitheophanous et al. (2010) criticize software-based solutions for their high computational and power requirements, which are unsuitable for real-time applications.

2.2.2. Proposed Solutions

Hasler (1981), Yu et al. (2019), and Hadjitheophanous et al. (2010) offer novel approaches to overcoming the limitations of existing methods for 3D cloud structure reconstruction.

The solutions proposed in these studies involve using advanced observational techniques and hardware implementations. Hasler (1981) proposes the use of stereographic observations from geosynchronous satellites. This technique relies on straightforward geometric relationships to provide much higher horizontal resolution and accuracy compared to infrared methods. By capturing stereo images from geosynchronous satellites positioned at different longitudes, the approach leverages the parallax effect to measure cloud heights with greater precision. Yu et al. (2019) suggests using multi-angle, multi-spectral and polarization data from the DPC onboard the GF-5 satellite. The DPC takes pictures of the same target from different angles. This let's cloud structures be reconstructed in 3D space using a ray casting algorithm. This algorithm figures out where cloud voxels are located by comparing where rays from different angles meet. Hadjitheophanous et al. (2010) suggest putting the 3D reconstruction algorithm on FPGA hardware, taking advantage of the algorithm's built-in parallelism. This approach includes a Sobel edge detection unit to reduce the amount of data processed, increase frame rates, and achieve real-time performance.

Each study successfully proves its claims through empirical verification and comparative analysis. Hasler (1981) demonstrates the accuracy of stereo height measurements by comparing them with known altitudes of high-altitude mountain lakes, achieving accuracy within ±0.1–0.2 km near reference points and ±0.5 km for general cloud features. The application of these measurements to meteorological problems, such as severe thunderstorms and hurricanes, further substantiates the efficacy of stereographic observations. Yu et al. (2019) show that their 3D reconstruction method works by showing that it matches up with CALIOP data in terms of vertical profiles and cloud boundaries. They also give accurate measurements of the reconstructed clouds. Hadjitheophanous et al. (2010) achieve real-time performance with frame rates up to 75 fps for 320x240 image pairs, demonstrating the system's efficiency across different parameter settings. The integration of the Sobel edge detector significantly enhances performance by reducing the amount of data processed.

2.3 Calibration and Validation

2.3.1 Motivation and Challenges

Bedka et al. (2007) focus on validating satellite-derived CTP products using aircraft-based CPL data. The problem here is ensuring the accuracy of CTH measurements, which are needed for meteorology, aviation safety, and climate studies. The difficulty lies in the discrepancies introduced by the differences in spatial and temporal resolution between satellite instruments and in-situ measurements, along with the inherent complexity of cloud structures. Existing solutions often fall short due to these complexities.

Vaughan et al. (2010) addresses the calibration of the CALIOP 1064 nm LiDAR channel, which is necessary for reliable atmospheric measurements. The challenge here involves weaker signal levels, noise, and the natural variability in the backscatter color ratio of cirrus clouds. Existing calibration methods based on potentially flawed assumptions can lead to significant errors.

2.3.2 Proposed Solutions

In 2007, Bedka et al. looked for and fixed problems in satellite data by comparing CTH products from the GOES-12 Imager and Sounder with CPL measurements taken during the ATReC field campaign. The study focuses on various cloud types and conditions and includes a brief comparison with MODIS-Aqua CTH retrievals. Their results indicate that the GOES-12 Imager generally agrees better with CPL measurements than the GOES-12 Sounder, particularly for mid-level and low clouds. The study clearly shows the pros and cons of current satellite measurements and backs up its claims with in-depth analysis and real-world examples. Using high-resolution LiDAR measurements provides a more accurate standard, showing that satellite-derived data could be more accurately validated.

Vaughan et al. (2010) calibrated the CALIOP 1064 nm LiDAR channel using cirrus clouds. They employed a measurement-based approach to determine the best estimate of the mean backscatter color ratio for cirrus clouds. They studied how cirrus cloud backscatter coefficients change with wavelength by examining more than 400 hours of LiDAR data from the CPL. As part of this study, they found backscatter color ratios for cirrus clouds and made sure the calibration process worked by comparing the color ratios found by CPL with the assumptions used in the CALIPSO calibration scheme. They accounted for aerosol loading in the normalization region using data from multiple satellites to apply a parameterized correction factor. Their findings showed that the best estimate for the backscatter color ratio of cirrus clouds is 1.01 ± 0.25, aligning with pre-launch assumptions and measurements. This result reassured the validity of the CALIOP 1064 nm calibration algorithm and revealed the need for a large sample size to minimize calibration errors.

2.4 Conclusion – Relevance To This cloud2cloud Project

The literature reviewed provides a comprehensive overview of previously studied approaches relevant to the cloud2cloud project. Mussa et al. (2022) and Bedka et al. (2007) discuss how important it is to get accurate measurements of cloud height and suggest ways to make them more accurate by using advanced instruments such as the flexible threshold-based algorithm for cloud detection and validation techniques using high-resolution LiDAR data. This is similar to our project's plan to combine HD images and LiDAR data to improve estimates of cloud height. Borisov et al. (2023) and Nied et al. (2023) show that ML techniques, especially artificial neural networks and CNNs in particular, are good at estimating cloud heights from images and Nied et al. (2023) notes that HD imagery is superior to detecting cloud presence as opposed to some other sensors used. This supports our plan to use CNN and RNN models in this project and the focus on using camera imagery. The work of McGill et al. (2002) on high-resolution

LiDAR systems provides a foundation for using LiDAR data as a benchmark for validating our predictive models. Hasler (1981), Yu et al. (2019), and Hadjitheophanous et al. (2010) all look into different ways to reconstruct 3D cloud structures while dealing with issues like data synchronization and computational efficiency. Finally, Bedka et al. (2007) and Vaughan et al. (2010) talk about calibration and validation, which stress how important it is that the cloud measurements we receive from NASA are reliable.

Together, these references inform the methodologies we plan to employ in the cloud2cloud project, ensuring a well-rounded approach to cloud height estimation and 3D reconstruction.

Chapter 3 – Methodology

3.1 Infrastructure

We are proposing to do this project on the AWS Platform which will help manage large storage requirement as well as compute power for intensive ETL and model training jobs:

- AWS S3 Buckets for Data Storage (100 GB of storage will cost ~ USD 2 per month)
- AWS Glue for Data Preparation and any potential ETL pipelines (~ USD 4.4 per month)
- AWS Sagemaker for Model Training, Tuning and Validation (1 Sagemaker notebook shared by the group may lead to a cost of ~ USD 350 per month)
- AWS Quicksight for Results Analysis & Visualization (~ USD 24 per month)

These costs are just initial ballpark estimates and may change significantly as we progress on the actual implementation.

SageMaker Instance proposed with compute power to support deep learning for this project -

> ml.p3.2xlarge
> Compute Type: Accelerated Computing Instances
> V CPU: 8
> Memory: 61 GiB
> Clock Speed: 2.3 GHz
> GPU: 1
> Network Performance: N/A
> Storage: EBS only
> GPU Memory: 16

3.2 Data

3.2.1. Data Sourcing and Collection

Following data sources obtained from NASA are being considered for the project:

- The sky videos from the HD Camera onboard FEGS will be the primary data source for training the proposed model.
- Aircraft navigation metadata data: plane speed, pitch, yaw, altitude, etc.
- Light Detection and Ranging (LiDAR) data for validation. This data provides the cloud height at the center of each image frame from the HD camera.

3.2.2. Data Storage

We plan to store all the data in shared AWS S3 buckets with strict access control. This location will serve as the primary source for the entire machine learning pipeline, including model training and validation. The output from various model runs will also be stored in the same shared location in a dedicated output folder with version control for every significant / important run.
Trained models will also be archived in a trained_models folder in the same location.
Key model performance metrics in a delimited text file and key visualization plots from various runs will be logged in this shared location for comparison and future analysis.

3.2.3. Data Cleaning and Preparation

Remaining inside the AWS platform will allow for easy integration between services, so we will use AWS Glue. AWS Glue allows the creation of extract, transform, and load (ETL) data pipelines. The data cleaning and preparation process will involve the following steps:

1) Defining tables for the AWS Glue Catalog
    - Some of the source tables will be video files from the Fly's Eye GLM Simulator (FEGS), LiDAR data, and aircraft navigation metadata.
    - The destination tables will contain processed and cleaned data ready for modeling consumption.
2) Creating transformation jobs. Develop ETL jobs in AWS Glue to transform the raw data into a usable format.
    - Standardize the images for size, resolution, brightness, and blurriness.
    - Remove features from the images that are not related to the clouds (e.g., the aircraft window and dirt particles on the lens or frames where the plane is turning around).
    - Correct the images for noise.
    - Handling missing values.
    - Completeness and accuracy of the metadata.
    - Synchronize all data sources from a time-series perspective (align all data sources temporally and spatially). We would also need to subtract the LiDAR measurement from the aircraft altitude to get the cloud height.
    - Outlier analysis

3) Running the jobs and analyzing results using AWS Glue Studio
    - Run ETL jobs as described
    - Use AWS Glue Studio's visual tools to analyze the results and ensure the ETL process is functioning correctly.
    - Verify the integrity of the cleaned data for subsequent modeling phases.

3.2.4. Data Sampling Methods

- Our camera videos are 30 frames per second while LiDAR is once per second. We can remove most extra image frames and use 1 image frame per second to match with the Lidar data. This will greatly reduce the number of image frames we work with.
- We can start with one flight video (~40 minutes) for initial modeling and then increase sampling of flights across different days, regions, and weather patterns.

3.3 Feature Engineering

We will identify key features from the prepared dataset that could improve the predictive power of the model. It might be worth exploring and potentially beneficial to enrich the data with certain derived composite features that could further enhance the model performance. We will use Amazon Sagemaker Data Wrangler to perform feature engineering which involves features selection and feature enrichment. Some computer vision techniques will be explored as:

i. **Image Stitching:** creating a mosaic image from the multiple image frames by finding matching key points on the images.
ii. **Image Augmentation:** exploring color contrast and saturation (examples in figures A and B).
iii. **Projective Transformations:** calculating trigonometric distances to extract information about the height of other points on the image rather than the center.
iv. **Feature extraction:** using pre-trained neural networks to extract commonly used features such as edges, corners and textures. This can be done also using SIFT or BRIEF algorithms.

Figure A: Original Image from HD Camera

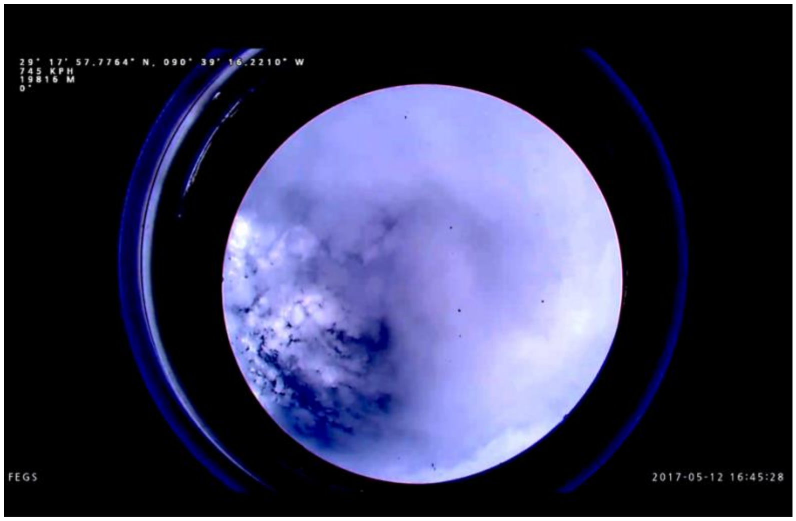


Figure B: Same image from HD Camera with changes in saturation, high contrast and gamma

3.4 Modeling

3.4.1 Proposed Models

1. "Structure From Motion" Traditional Methods for 3D Cloud Reconstruction:

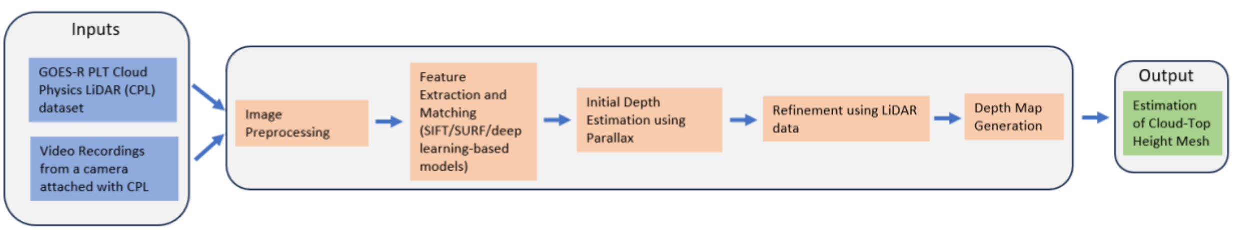


Figure 1: Workflow for 3D Reconstruction using Multiple View Geometry

Output 3D cloud-top shape using n-view reconstruction algorithms from traditional methods of multiple view geometry such as structure from motion (sfm) – a technique to recover 3D structure of a scene by making use of a sequence of 2D images. The general idea is that the images result from two factors: the relative motion between the camera and the object and the object shape. Knowing camera metadata (such as speed, angle, etc.) and time between frames, we can apply geometric calculations to obtain our height estimations.

Pre-requisites for Applying Traditional Multiview Geometry Methods:

a. Object boundaries can be identified within each frame. Match key points between sequential images to find corresponding points.

b. Identify assumptions made or make reasonable estimates such as 1) cloud shape does not change or 2) cloud has minimal movement between frames. A great advantage is that we have the true height data from LIDAR to ensure assumptions are reasonable (in addition to checking correct algorithm applications).
c. Translate plane terminology (pitch, yaw, etc.) into meaningful mathematical equivalents for calculating camera angle, speed, etc.

Hybrid Approach with Traditional Computer Vision and Deep Learning Methods:

i. For any of the given pre-requisite steps above, we could try to augment with other known algorithms or even deep learning methods. Specifically, for pre-requisite a, we can use a feature extraction method (e.g. computer vision algorithms SIFT and SURF; or deep learning-based models like SuperPoint, D2-Net) to detect key points in the images.
ii. We can use these depth calculations using this algorithmic approach as additional features to feed into a deep learning model (see #2 **CNN-RNN-Transposed Convolution architecture**).

Tomasi-Kanade Factorization Algorithm:

We can also use the **factorization algorithm**, in which every image frame in a video is considered a product of two parameters: motion and shape. Assuming the object shape does not change, then if the motion of the camera is constant the image frames would be linear transforms of each other. Thus, the factorization algorithm extracts shape parameters from the images to give the 3D representation.

2. **CNN + RNN type models:** Incorporate sequential image frames and flight metadata per each frame to learn relationships between frames.

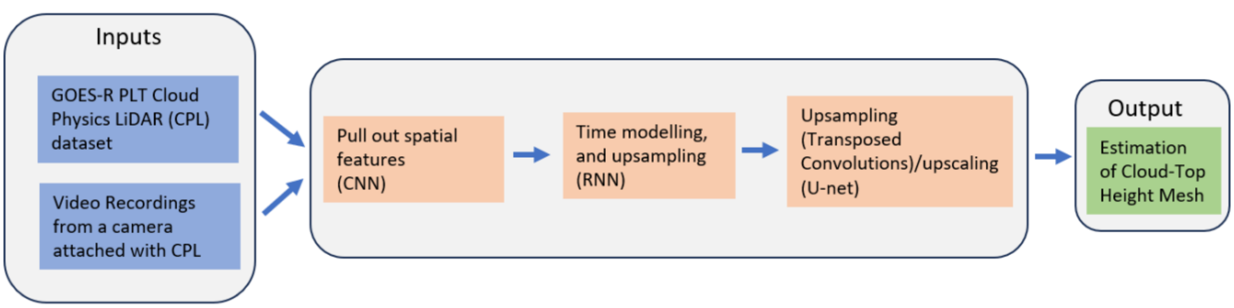


Figure 2: CNN-RNN-Transposed Convolution architecture

a. Neural Networks (CNNs) - CNNs are popular and commonly used models for image classification, regression, object detection, and other computer vision applications. They work well with grid-like data, much like our case, which is a 2D frame sampled every second from the HD video. They use convolution layers to apply convolutional filters to the input images for feature extraction. Pooling layers then reduce the dimensionality of the feature maps, making the computation more efficient and reducing the risk of overfitting. Fully connected layers are used at the end for regression or classification. The addition of convolutional layers to the CNN makes them hierarchical, as the later layers can see the pixels within the receptive fields of prior layers. This could be particularly useful in the problem we are dealing with here, enabling the network to capture both low and high-level details from the images.

b. UNETs - U-Nets are typically used for segmentation rather than regression tasks like predicting a height field/point cloud/mesh. Adapting them for this task might be complicated, and they add a lot of computational overhead. However, they do have an encoder-decoder architecture, which could improve the upsampling step, and could be an option to try if we have the capacity to do so.

c. Given a set of images and LiDAR measurements taken corresponding to the center of each one, we can make a model that uses convolutional neural networks (CNNs) to pull out spatial features, recurrent neural networks (RNNs) to model temporal dependencies, and upsampling techniques to make the height field (point cloud).

3. ViT-Temporal Transformer-U Net Implementation

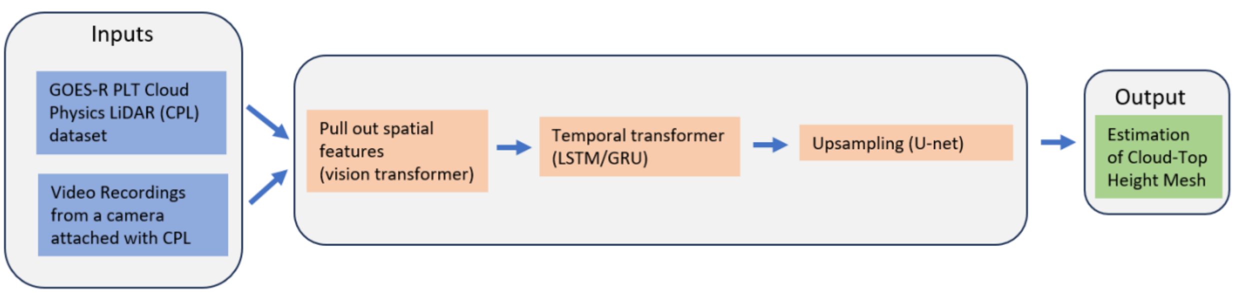


Figure 3: ViT-Temporal Transformer-U Net Implementation

a. Depending on the amount of data and computer resources we have; we could swap the CNN for a vision transformer (ViT) to process each image in the sequence and extract meaningful features. ViTs could potentially capture more complex and longer temporal dependencies within the images using self-attention mechanisms. We could also swap the LSTM or GRU for a temporal transformer.

b. Finally, we could use a U-net for upsampling, as this would preserve long term dependencies in the sequential images. These options could work even better but would require perhaps 10x the data and a lot more compute power. We plan on considering this a stretch goal and will only be attempted once we have the CNN-RNN-Transposed Convolution working.

3.4.2. Model Components

CNN for Feature Extraction:

- Extract spatial features from each image that represent the cloud structure.

Recurrent Neural Network (RNN):

- Capture temporal dependencies between sequential images using the feature maps extracted by CNN. These could be LSTM or GRU layers.

Upsampling Network:

- Convert the RNN output into a 2D n x n height field using upsampling layers that represent the cloud-top.

Custom Loss Function:

- Use the LiDAR measurement at the center of the image to make a custom loss function to train the model.

3.4.3. Model Training

Training Procedure:

- Use the normalized and processed image sequences as input to the chosen model.
- Use a custom loss function to train the model so that it minimizes the error at the center of the LiDAR measurement's predicted height field.

Optimization:

- Use the Adam optimizer with an initial learning rate of 0.001 to train the model.
- Implement a learning rate scheduler to reduce the learning rate if the validation loss plateaus.

Training Loop:

- Split the data into training, test and validation sets.
- Perform forward passes, calculate the loss using the custom loss function, and backpropagate to update the model weights.
- Monitor training and validation loss to prevent overfitting.

3.4.4. Model Tuning

Hyperparameter Tuning (using validation loss):

- Experiment with different learning rates, batch sizes, and the number of epochs.
- Adjust the number of convolutional layers and LSTM layers and their respective sizes.
- Use grid search techniques to explore hyperparameter combinations.

Regularization:

- Apply drop-out layers to the CNN and RNN to prevent overfitting.

- Use L2 regularization to penalize large weights in the network.

Early Stopping:

- Implement early stopping based on validation loss to prevent overfitting and reduce unnecessary training time.

3.4.5. Model Validation & target performance metric

Validation Procedure:

- Use a separate validation set to evaluate the model during training.
- Calculate the loss on the validation set after each epoch to monitor performance.
- Adjust hyperparameters and training strategies based on validation performance.
- Finally, when hyperparameters and architecture appear optimal using the validation set, run it on the test set.

Target Performance Metric:

- A primary performance metric could be the root mean squared error (RMSE) of the height prediction at the center of the height field as compared to the LiDAR measurement.

Cross-Validation:

- Alternatively, we could perform k-fold cross-validation to really test the model, averaging the performance metrics across folds to get a reliable estimate of the model's generalization ability.

Performance Goals:

- Set specific performance targets, such as a minimum MSE.
- Aim for low variance between training and validation losses to know we're not overfitting.

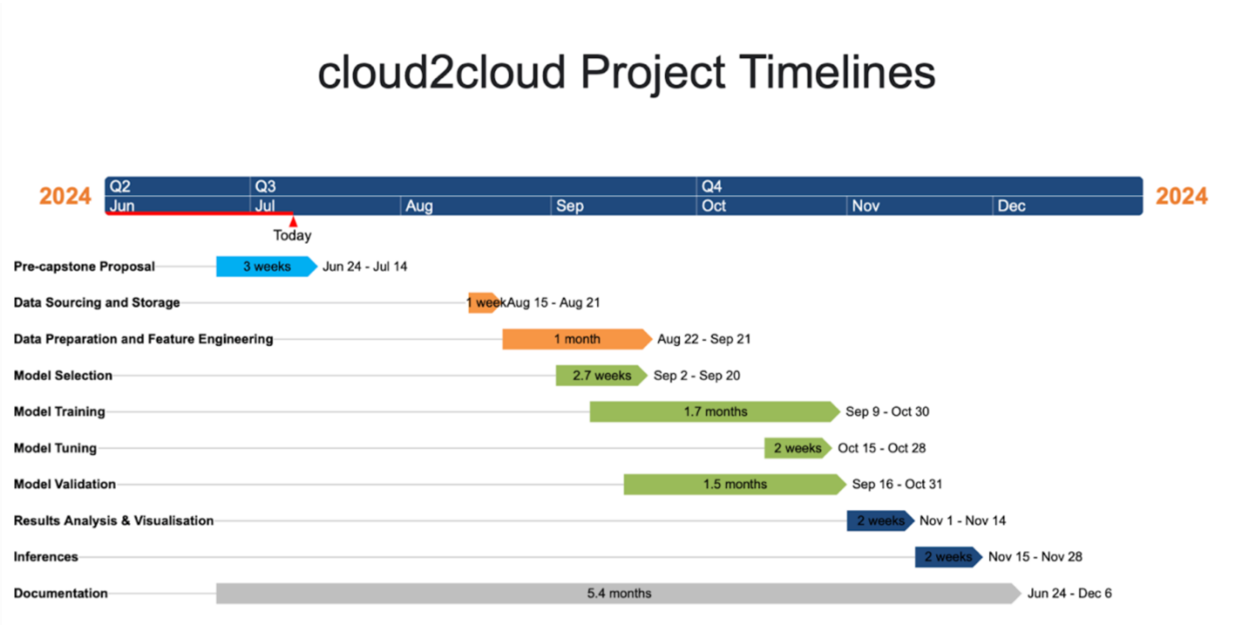
3.5 Ethical, Legal, and Privacy Considerations

We are using two datasets. The first one is GOES-R PLT Cloud Physics LiDAR (CPL) dataset. The dataset consists of data collected during the period April 13, 2017, through May 14, 2017, by Cloud Physics LiDAR instrument flown aboard the NASA ER-2. The data is in HD5 format and publicly available at: https://cmr.earthdata.nasa.gov/search/concepts/C1979112912-GHRC_DAAC.html.
The second dataset has recordings from video cameras fitted on the same aircraft and will be publicly available.

Since NASA had made the first dataset publicly available and NASA is planning to make the second dataset publicly available, there are no ethical, legal and privacy considerations.

Since NASA used LiDAR and HD cameras with their ER-2 aircraft to collect the data, we assume that it is reliable and trustworthy.

3.6 Execution Timeline



Overall project execution plan is to spend significant time for data preparation and infrastructure set-up given the size of the dataset and need for cleaning and time-stamp matching. Next, in model selection we will explore traditional computer vision and algorithmic approaches, deep-learning methods, and hybrid approaches. Finally, we expect to spend time trying to train and further fine tune our most promising methods. During this whole process, we expect to constantly update documentation and milestone visualizations.

References

Bedka, Sarah & Feltz, W. & Schreiner, A. & Holz, Robert. (2007). Satellite-derived cloud top pressure product validation using aircraft-based cloud physics lidar data from the ATReC field campaign. International Journal of Remote Sensing - INT J REMOTE SENS. 28. 2221-2239. 10.1080/01431160500391965.

Borisov, Mikhail & Krinitskiy, Mikhail & Tilinina, N.. (2024). Estimating Cloud Base Height From All-Sky

Imagery Using Artificial Neural Networks. Moscow University Physics Bulletin. 78. S85-S95. 10.3103/S0027134923070020.

Hadjitheophanous, Stavros & Ttofis, Christos & Georghiades, A. & Theocharides, Theo. (2010). Towards hardware stereoscopic 3D reconstruction: A real-time FPGA computation of the disparity map. 1743-1748. 10.1109/DATE.2010.5457096.

Hasler, A.. (1981). Stereographic Observations from Geosynchronous Satellites: An Important New Tool for the Atmospheric Sciences. Bulletin of the American Meteorological Society. 62. 10.1175/1520-0477(1981)062<0194:SOFGSA>2.0.CO;2.

McGill, Matthew & Hlavka, Dennis & Hart, William & Scott, V. & Spinhirne, James & Schmid, Beat. (2002). Cloud Physics Lidar: Instrument Description and Initial Measurement Results. Applied optics. 41. 3725-34. 10.1364/AO.41.003725.

Mussa, R & Pallotta, Juan. (2022). Measuring Cloud Base Height and Cloud Coverage using Elastic Multiangle Lidars at Pierre Auger Observatory. Journal of Physics: Conference Series. 2398. 012018. 10.1088/1742-6596/2398/1/012018.

Nied, Joseph & Jones, Michael & Seaman, Shane & Shingler, Taylor & Hair, Johnathan & Cairns, Brian & Van Gilst, David & Bucholtz, Anthony & Schmidt, Sebastian & Chellappan, Seethala & Zuidema, Paquita & van Diedenhoven, Bastiaan & Sorooshian, Armin & Stamnes, Snorre. (2023). A cloud detection neural network for above-aircraft clouds using airborne cameras. Frontiers in Remote Sensing. 4. 1118745. 10.3389/frsen.2023.1118745.

Vaughan, Mark & Liu, Zhaoyan & McGill, MJ & Hu, Yongxiang & Obland, Michael. (2010). On the spectral dependence of backscatter from cirrus clouds: Assessing CALIOP's 1064 nm calibration assumptions using cloud physics lidar measurements. Journal of Geophysical Research-Atmospheres. 115. 10.1029/2009JD013086.

Yu, Haixiao & Ma, Jinji & Ahmad, Safura & Sun, & Li, & Hong,. (2019). Three-Dimensional Cloud Structure Reconstruction from the Directional Polarimetric Camera. Remote Sensing. 11. 2894. 10.3390/rs11242894.