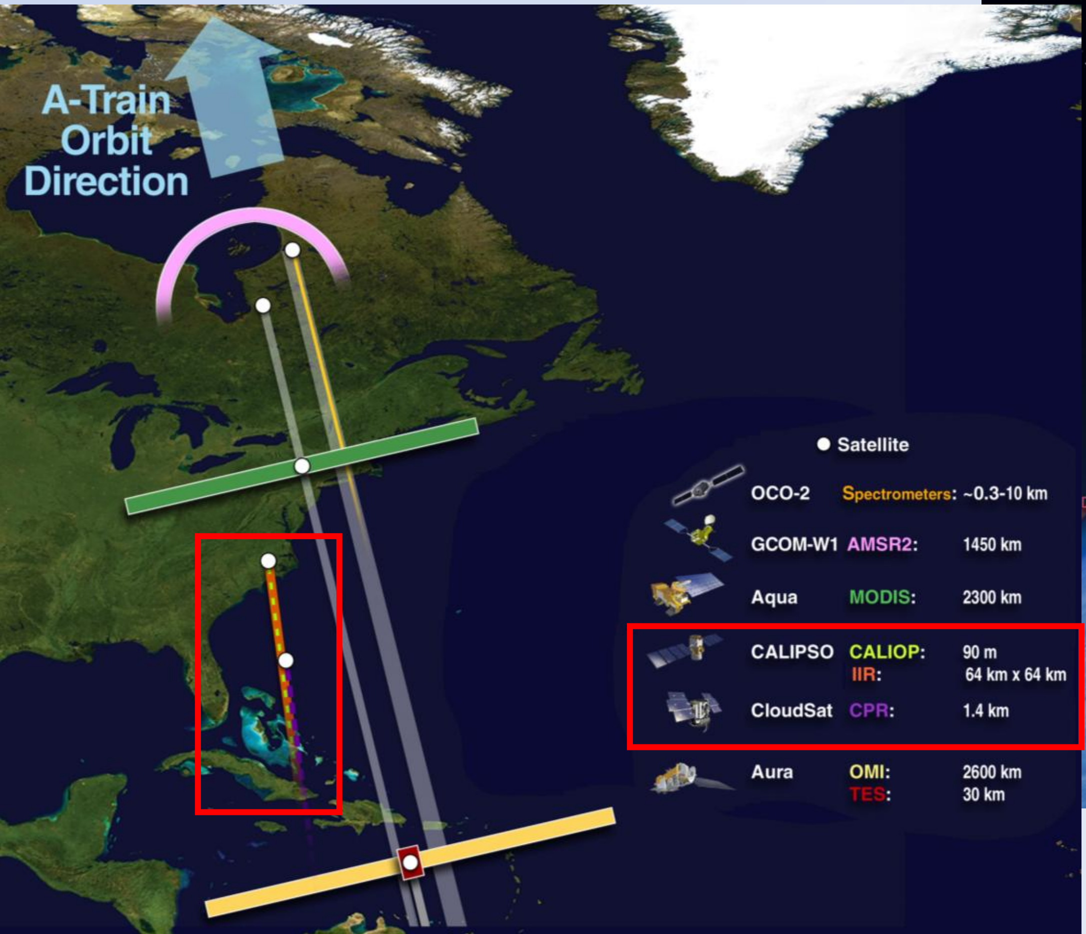
2025 Cloud and Environment

Week 6: AEL satellite module – CloudSat part

ATMOSPHERIC ENVIRONMENT LAB

Peng Jen Chen



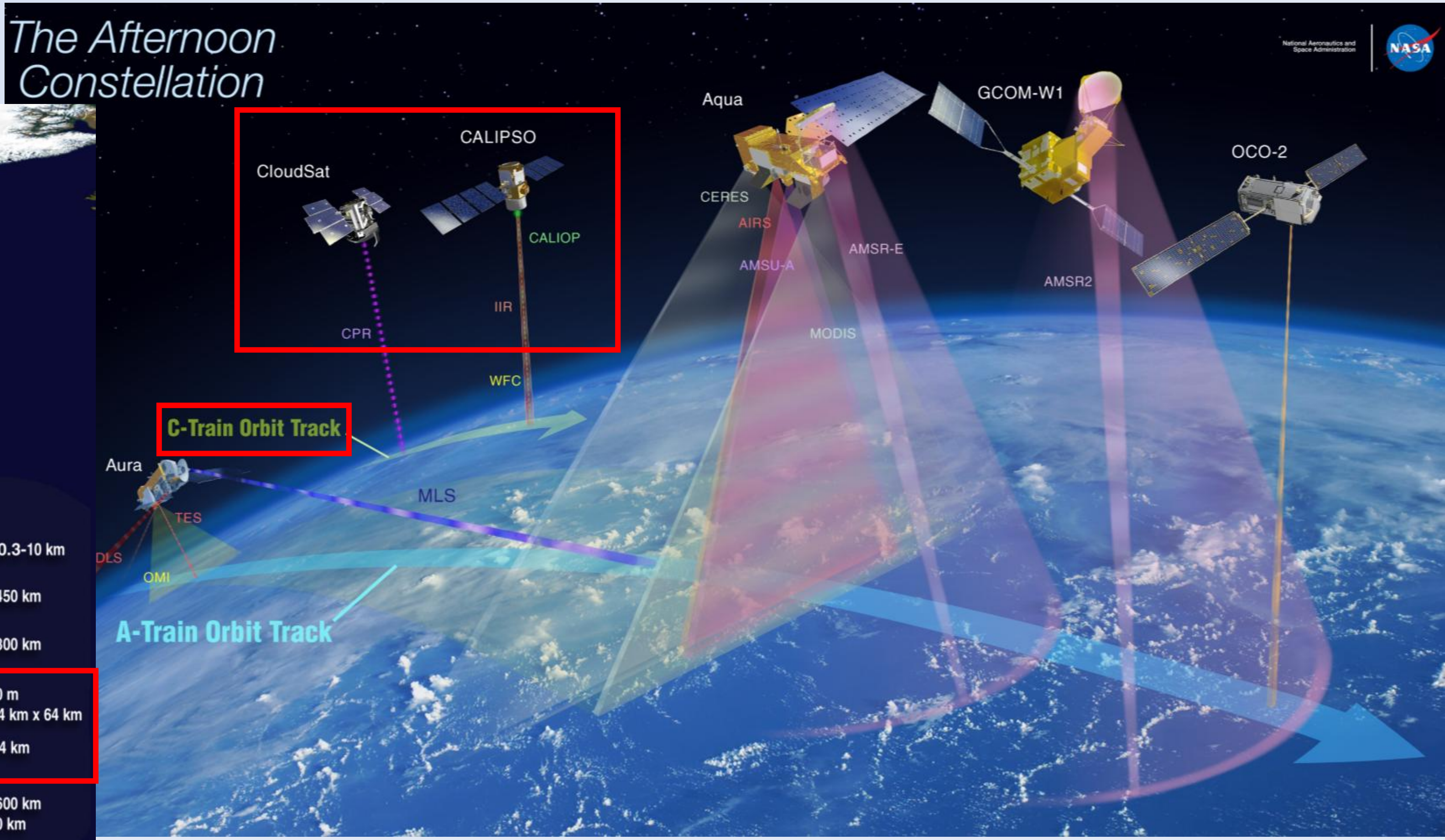
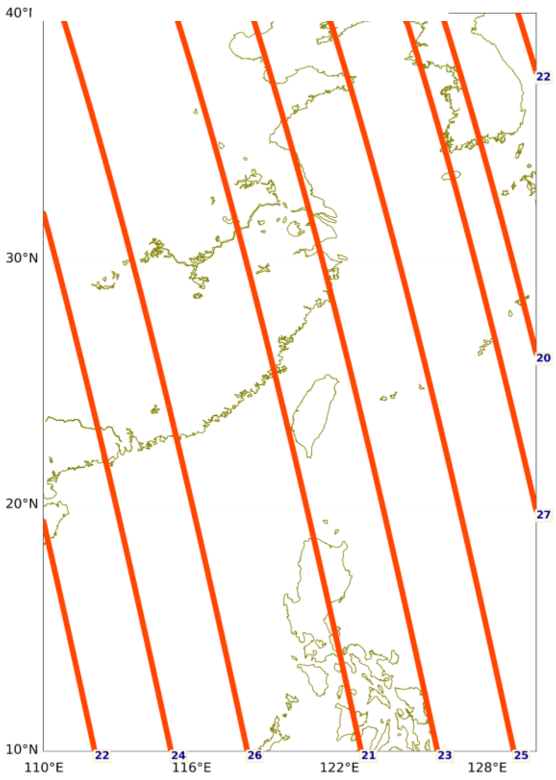


Figure Credit: NASA

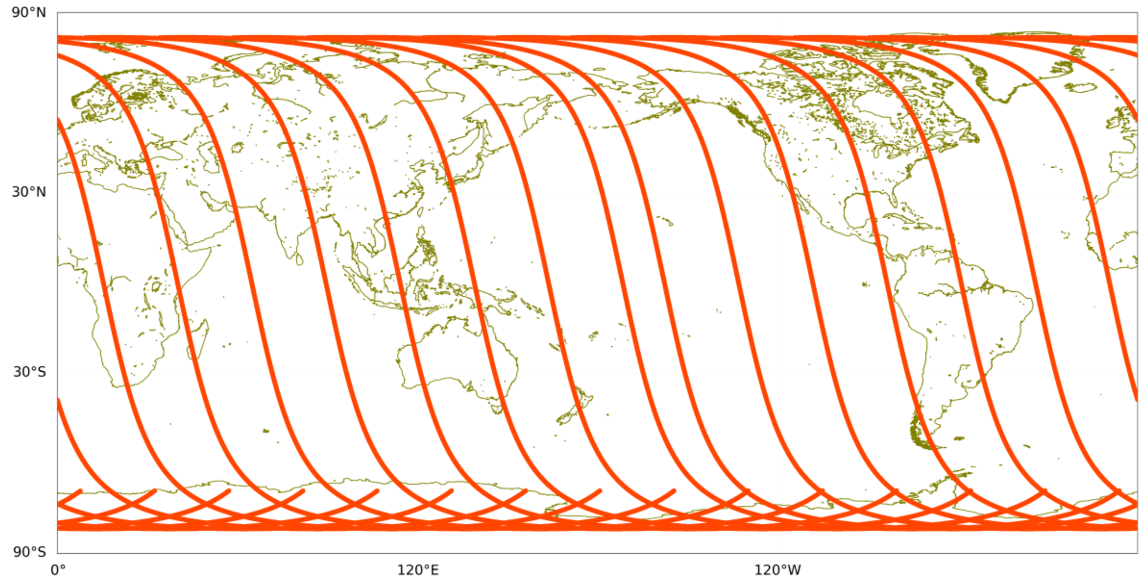
CloudSat & CALIPSO – introduction

- Polar-orbiting satellite; launch at 2006/04/28; CloudSat stop operation at 2023/12/20 (CALIPSO at 8/1)
- Cross the equator at local time ~1:30 pm and ~1:30 am
  - Due to battery malfunction: CloudSat only day-time have observation after 2011
- High resolution vertical profile cross the atmosphere
  - ~15 tracks in one day; ~44,000 swaths for ten years observation
- Weakness:
  - Narrow observation domain; Temporal resolution is limited
- From A-train to C-Train
  - After September 2018
  - CloudSat is 55 seconds behind CALIPSO
- Next gen satellite: EarthCARE
  - Integrated CPR & Lidar instruments

Tracks in one week



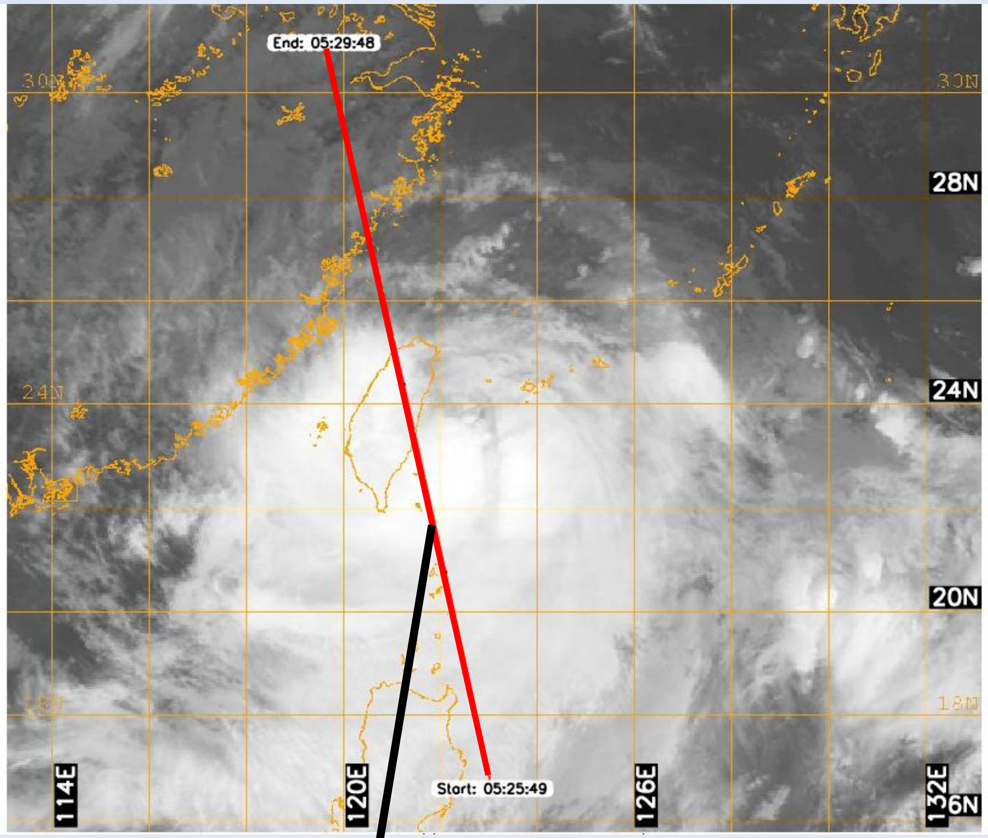
Tracks in one day



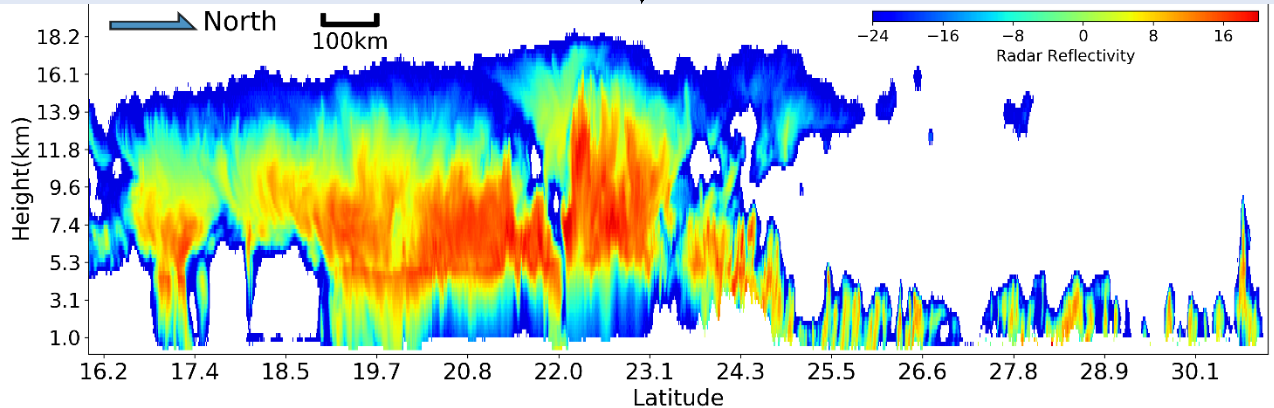
CloudSat & CALIPSO – introduction

- CloudSat
  - Cloud Profiling Radar (CPR) at 94-GHz: **sensitive to larger hydrometeors**
  - **Vertical resolution = 240m; Horizontal resolution: along track = 1.8km; cross track = 1.4km**
- CALIPSO
  - Calipso Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP): sensitive **to aerosols and fine particals**
  - **Vertical resolution = 60m; Horizontal resolution: along track = 333m; cross track = 90m**

Typhoon Morakot 2009/8/7

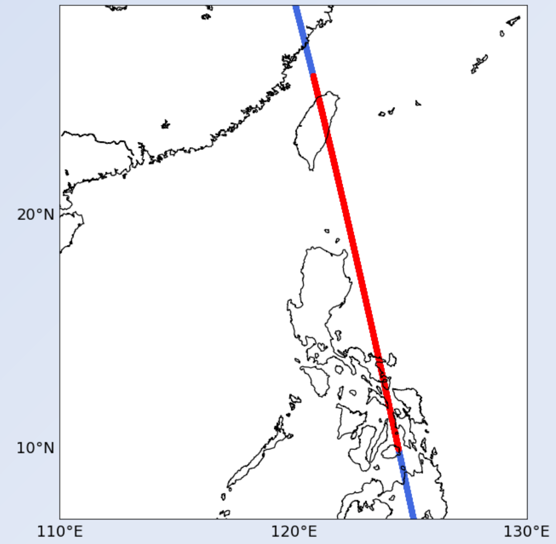


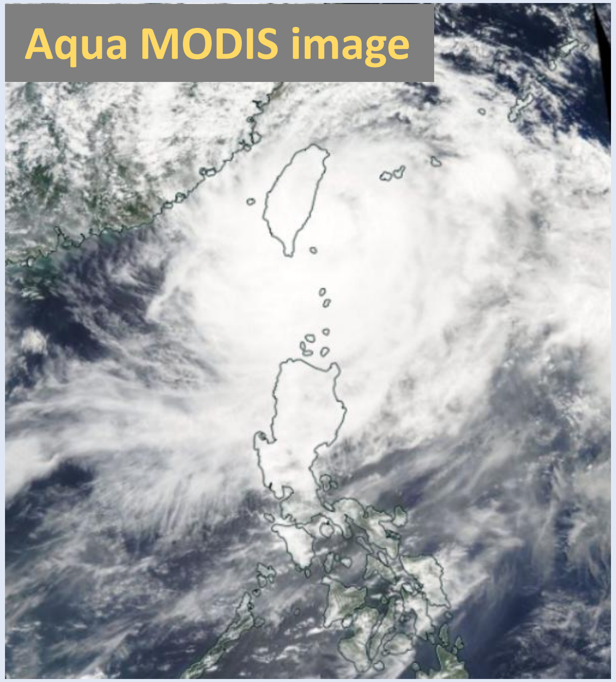
Cloud mask & Reflectivity



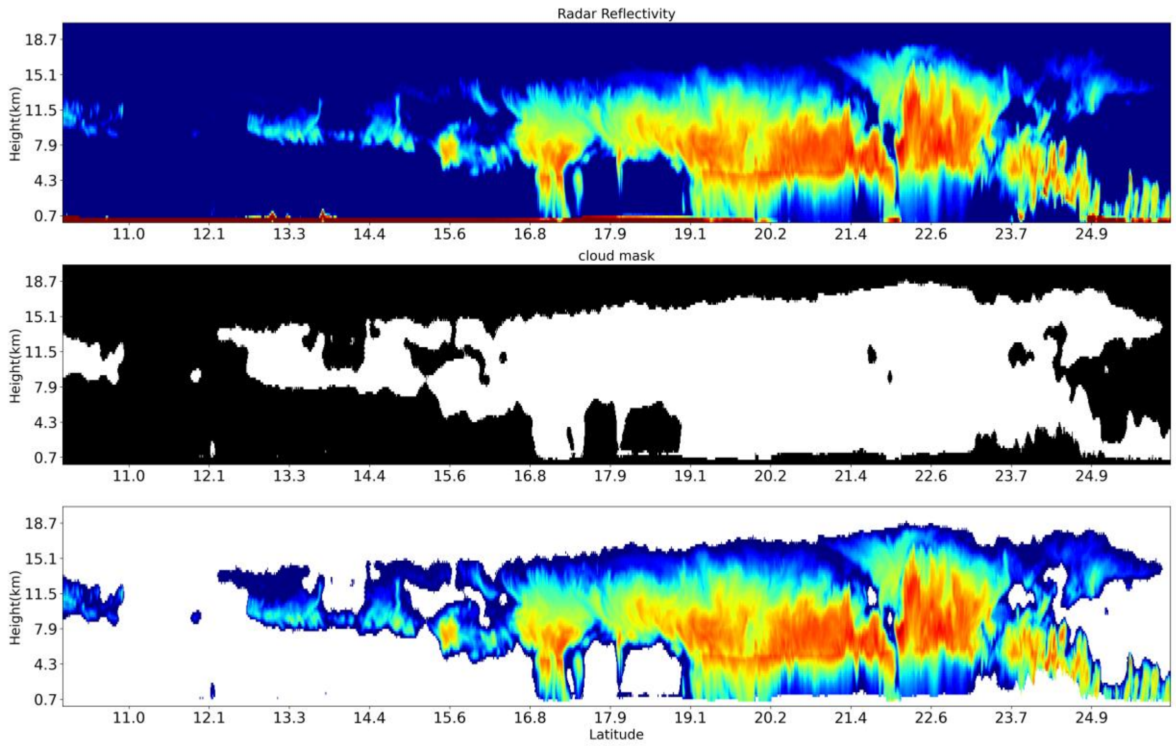
CloudSat– 2B product

- CloudSat 2B product
  - 2B-GEOPROF
    - Product of CPR
  - 2B-GEOPROF-LIDAR
    - Re-grid to CloudSat's resolution

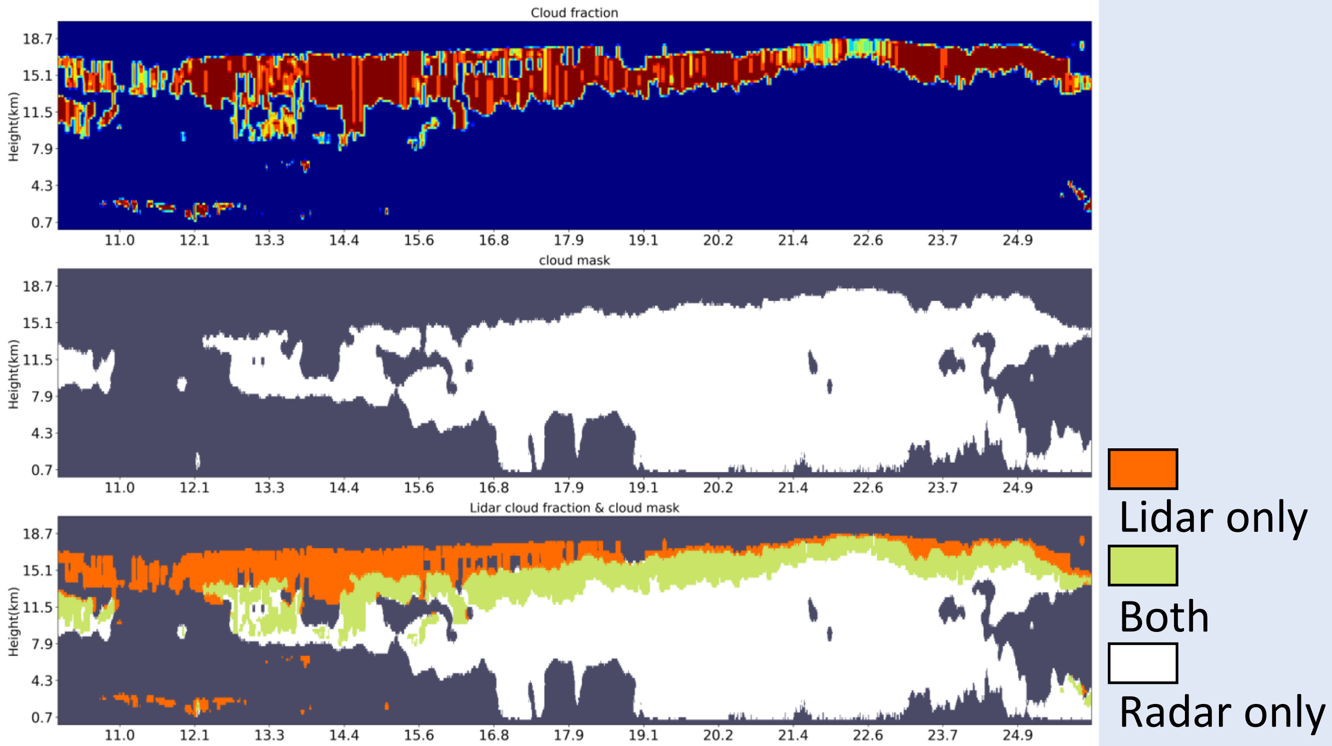




2B-GEOPROF



2B-GEOPROF-LIDAR



Paper – CloudSat clustering

Geophysical Research Letters

RESEARCH LETTER
10.1029/2021GL092733

Convective Cloud Regimes From a Classification of Object-Based CloudSat Observations Over Asian-Australian Monsoon Areas

Peng-Jen Chen[1], Wei-Ting Chen[1], Chien-Ming Wu[1], and Ting-Shuo Yo[1]

[1]Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan

Key Points:

- Five convective cloud regimes are derived from CloudSat observations based on object-based cluster analysis of five physical properties
- Two regimes are highly organized systems over coastal areas and three regimes are the less organized systems mainly over land
- The coastal-intense regime exhibits strong cloud radiative effects and extreme rainfall, clearly coupled with the large-scale circulation

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

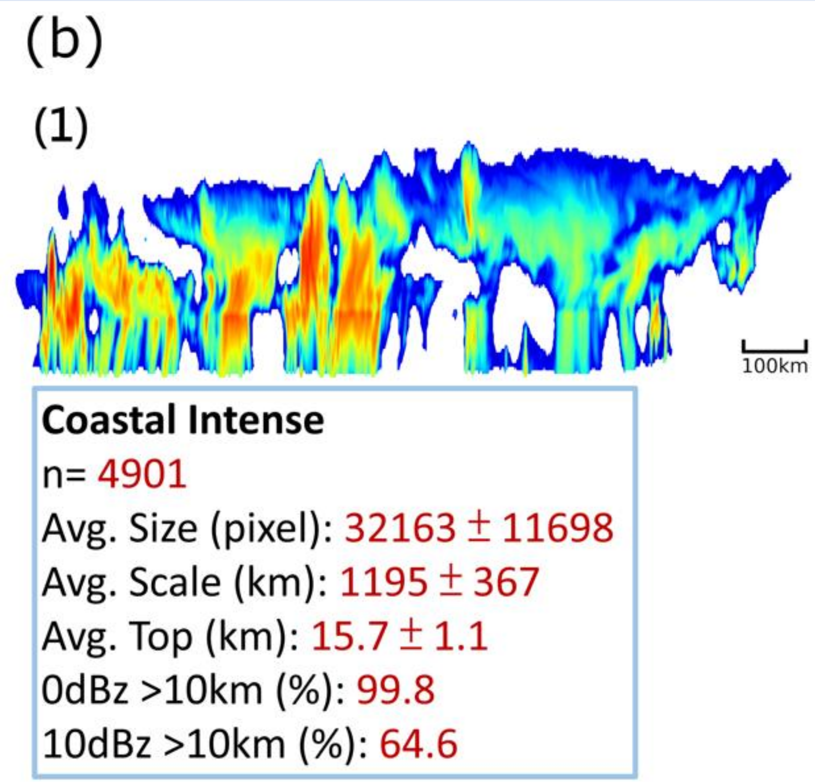
W.-T. Chen,
weitingc@ntu.edu.tw

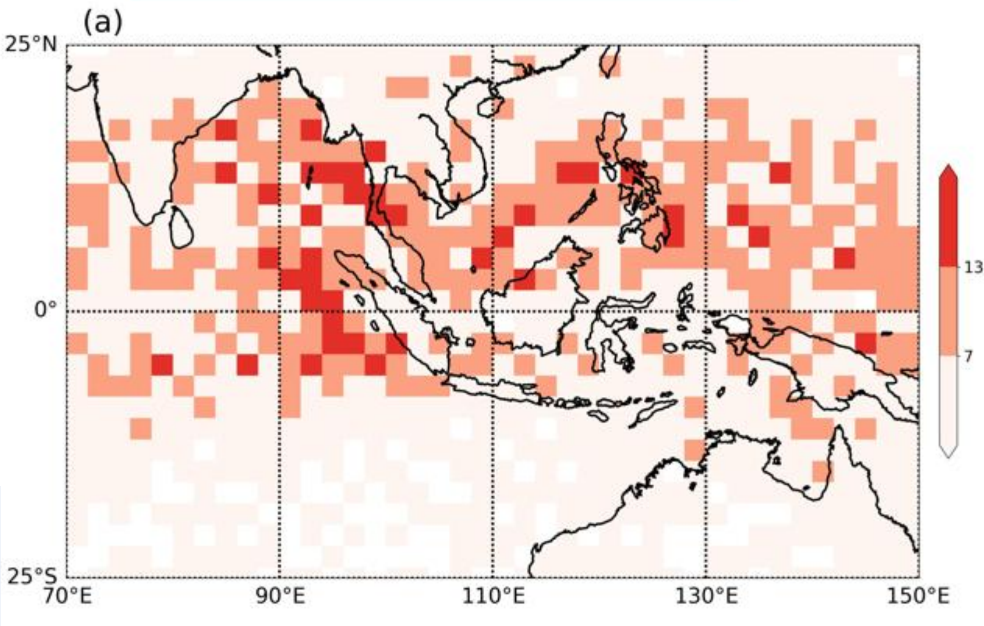
Citation:

Chen, P.-J., Chen, W.-T., Wu, C.-M., & Yo, T.-S. (2021). Convective cloud regimes from a classification of object-based CloudSat observations over Asian-Australian monsoon areas. *Geophysical Research Letters*, *48*, e2021GL092733. https://doi.org/10.1029/2021GL092733

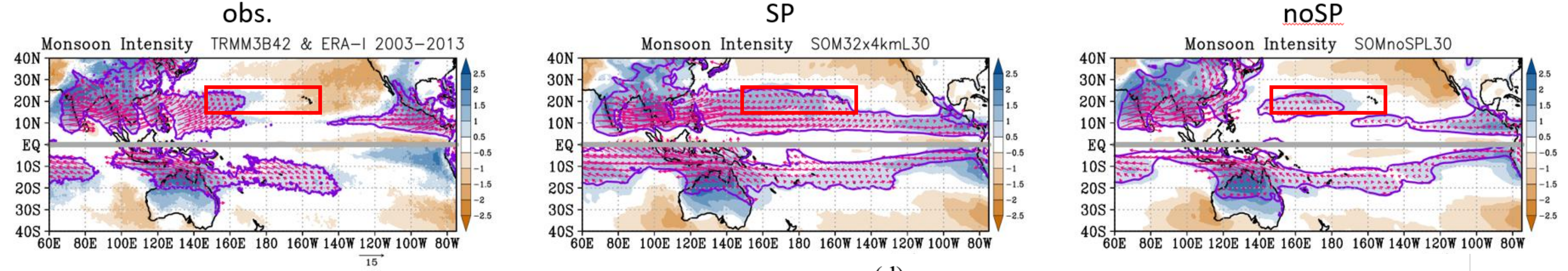
Abstract The present study objectively classified the convective cloud objects detected by the space-borne CloudSat radar over the Asian-Australian monsoon region using the hierarchical agglomerative clustering algorithm. Based on key properties representing the morphological features and convective intensity of the systems, five distinct convective cloud regimes were derived. The unique Coastal-Intense (CI) regime exhibits the most expansive horizontal scales (>1,000 km), high convective strength, the strongest cloud radiative effects, the highest probability of extreme rainfall, and a significant coupling with the sharp onset of the Asian summer monsoon circulation. Second, the Coastal regime illustrates smaller but also highly organized coastal convections, with the strongest convective strength. Less than 10% of the systems in the CI and Coastal regimes overlap with the tropical cyclones. The rest three regimes mark the less organized convection at various life cycle stages mainly over the land areas, with small seasonal variation in their occurrence.

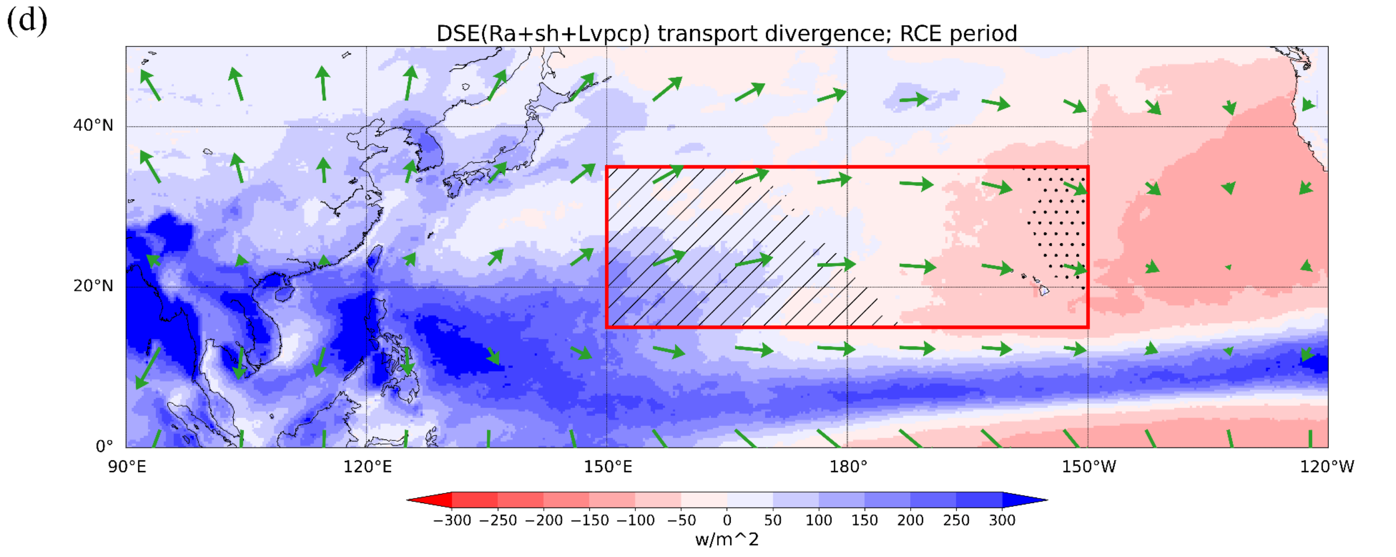
Plain Language Summary One of the key features of the Asian-Australian monsoon is the occurrence of intense rainfall and organized convective systems coupled with the change of large-scale circulation in the summer hemisphere. In this study, the convective systems from the multi-year satellite observations over the Asian-Australian monsoon region were classified objectively into five distinct regimes by a data-driven approach based on systems' physical properties. The most organized Coastal Intense regime, mainly occurring over the coastal area, exhibits the highest probability of extreme rainfall among all regimes, and tightly follows the sharp seasonal switch of the monsoon circulation. The Coastal regime illustrates smaller organized coastal convections with the strongest convective strength. Three of the less organized regimes represent convection at various life stages over the land areas.

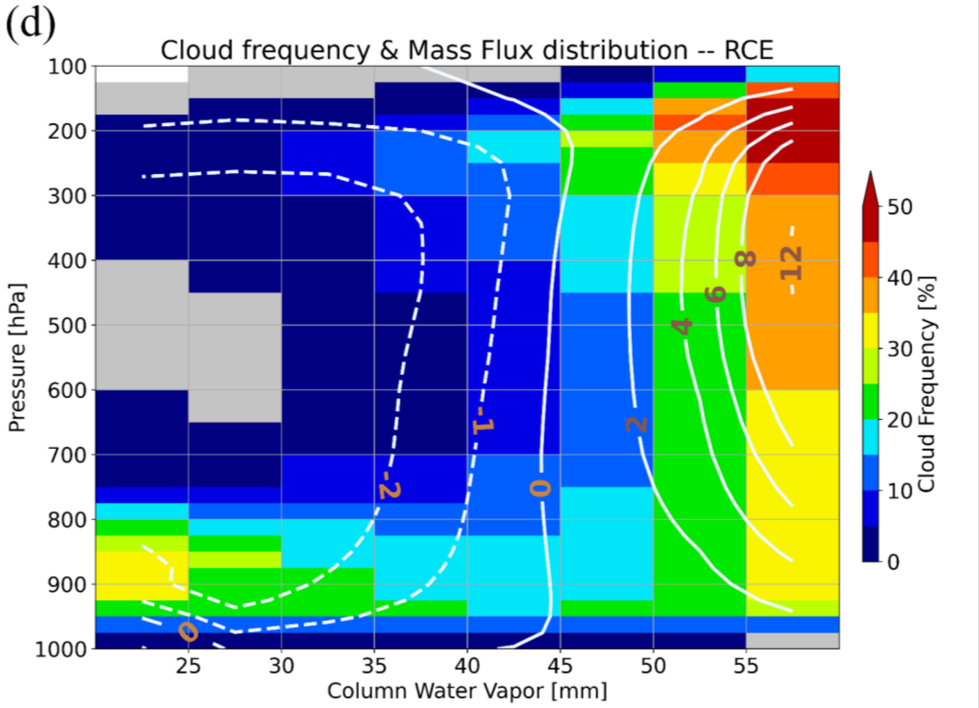




Under reviewing Paper – Observations of Convection Aggregation under Seasonal Radiative Convective Equilibrium Over Northwest Pacific Monsoon Region – **conditional sampling of CloudSat data**

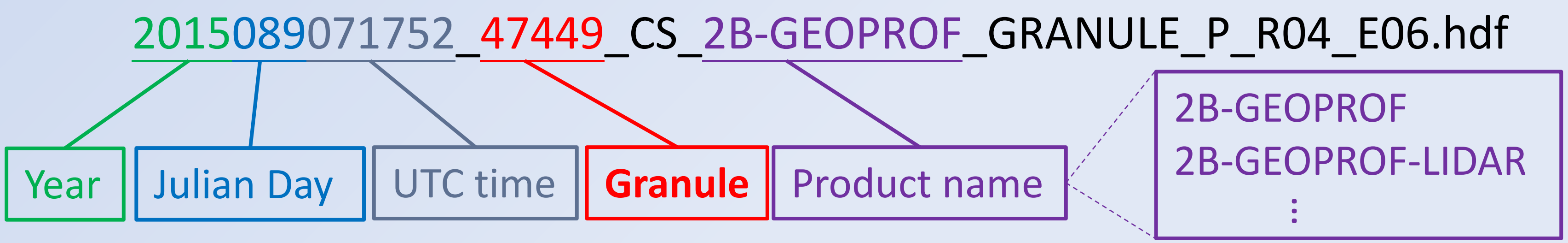






CloudSat – basic product information

File name example:



Geolocation Fields

Field	Type	Dimensions	Units	Valid Range	Missing	Missing Op.	Factor	Offset
Profile_time	FLOAT32	nray	seconds	[0.0, 6000.0]			1.0	0.0
UTC_start	FLOAT32	scalar	seconds	[0.0, 86400.0]			1.0	0.0
TAI_start	FLOAT64	scalar	seconds	[0.0, 600000000.0]			1.0	0.0
Latitude	FLOAT32	nray	degrees	[-90.0, 90.0]			1.0	0.0

Data Fields

Field	Type	Dimensions	Units	Valid Range	Missing	Missing Op.	Factor	Offset
CPR_Cloud_mask	INT8	nbin,nray		[0, 40]	-9	==	1.0	0.0
Gaseous_Attenuation	INT16	nbin,nray	dBZe	[0, 1000]	-9999	==	100.0	0.0
Radar_Reflectivity	INT16	nbin,nray	dBZe	[-4000, 5000]	-8888	==	100.0	0.0

AEL satellite module – CloudSat

In this section, you will learn…

- **Read CloudSat 2B-GEOPROF & 2B-GEOPROF-LIDAR product data**
  0. Read .hdf file by pyhdf module
  1. Read CloudSat 2B product by **AEL satellite module**
  2. Extract sub-domain data array
  3. Plot vertical profile of atmosphere
- **Filtering correct target file(s) over your study domain during specific time period**
  1. Generate file list in the specific time period
  2. Filtering files over the over your study domain
  3. Plot data tracks and pick target file(s)
- **The concept of conditional sampling**
  1. Invalid column and value

AEL satellite module – CloudSat

- **pip install --upgrade ael_satellite_tools (current version 0.0.6)**
- **from ael_satellite_tools.preprocess import CloudSat**
- **data_path = '/data/dadm1/obs/CloudSat' (default setting)**
  - No need to download data
  - data_path/**[product name]/[YYYY]**
- lat = [-10, 60]
- lon = [90, 150]
- cloudsat = CloudSat(work_path=[],lat_range=lat,lon_range=lon)
- cloudsat.product_information()

  CloudSat product on 107

  hdf-GEOPROF; hdf-GEOPROF-LIDAR; hdf-FLXHR-LIDAR

Course demo (**week6*.py & week6*.ipynb**)

- /data/cloud2025/homework_data/
- https://github.com/jerryjerry9/cldenv_2025

AEL satellite module – CloudSat

Read CloudSat 2B-GEOPROF & 2B-GEOPROF-LIDAR product data

0. **Read .hdf file by pyhdf module**

- **from pyhdf.HDF import HDF, HC**
- **from pyhdf import V**
- **from pyhdf.SD import SD, SDC**

SD (Scientific Dataset)

```
### read Scientific Data
hdfFile = SD(geoprof_file[0], SDC.READ)
Var_data = hdfFile.select('Radar_Reflectivity')
var_data = Var_data[:]
hdfFile.end
```

nbin*nray data:

CPR_Cloud_mask,

Radar_Reflectivity, …

Vector data (Vdata)

```
### read Vector data
reading_vdata = HDF(geoprof_file[0], HC.READ).vstart()
vdata_list = reading_vdata.vdatainfo()
target_vdata = 'Latitude'
for ref in vdata_list:
    vdata_name = ref[0]
    if vdata_name == target_vdata:
        vdata_length = ref[3]
        Var_data = reading_vdata.attach(target_vdata)
        var_data = Var_data.read(vdata_length)
        Var_data.detach()
        break
reading_vdata.end()
```

nray or **scalar** data:

Latitude, Longitude, Data_quality,…

AEL satellite module – CloudSat

Read CloudSat 2B-GEOPROF & 2B-GEOPROF-LIDAR product data

1. **Read CloudSat 2B product by AEL satellite module**

- cloudsat.hdf_information()
  - Return variable name of Vdata & SDdata
- cloudsat.cloudsat_var_attribute()
  - Return factor and missing value
- cloudsat.read_ori_vdata() & cloudsat.read_ori_sddata()
  - Read Vdata & SDdata
  - **Variables are already divided by factor**
- Vdata: Longitude, Latitude, SurfaceHeightBin (identical cross different products)
- SDdata:
  - GEOPROF: Radar_Reflectivity, CPR_Cloud_mask , Height (identical cross different products)
  - GEOPROF-LIDAR: CloudFraction, UncertaintyCF

AEL satellite module – CloudSat

Read CloudSat 2B-GEOPROF & 2B-GEOPROF-LIDAR product data

2. Extract sub-domain data array

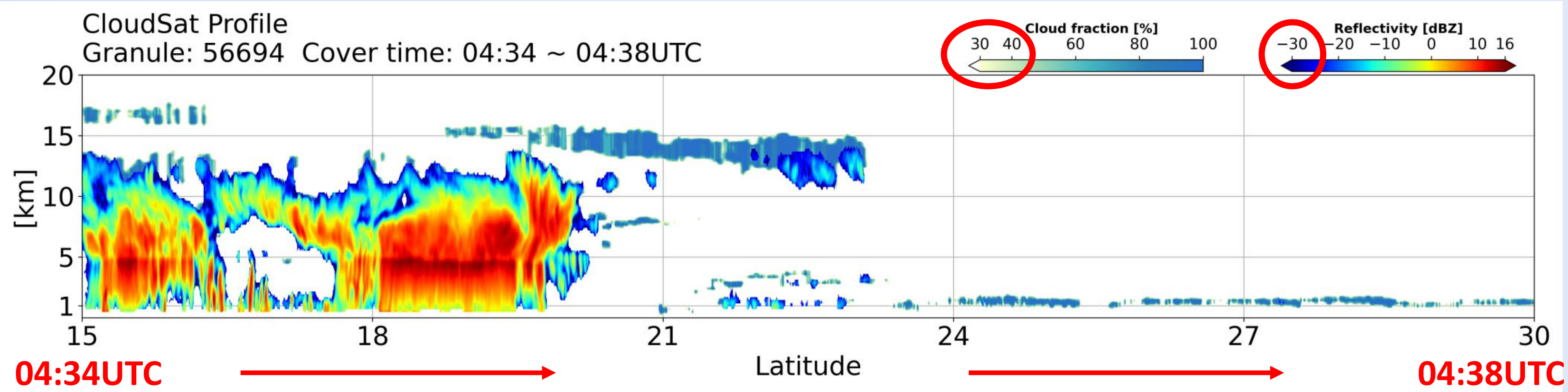
- cloudsat.plot_track_w_rgb()
  - Plot CloudSat track on RGB image as the reference for extracting data array
  - Plot sub-domain track on the original track
- cloudsat.sub_domain_check(**file_list,lon_range,lat_range,mask_output=True**)
  - Check CloudSat track passing the target lon-lat range
  - Return full length masked array of target lon-lat range (1 for target lonlat range; 0 for outside the lonlat range or bad data quality)
- cloudsat.read_vdata() & cloudsat.read_sddata()
  - Read & **extract** Vdata & SDdata
  - Use **extracted_mask** from cloudsat.sub_domain_check()
  - Longitude will shift to 0 ~ 360 (original: -180 ~ 180)

AEL satellite module – CloudSat

Read CloudSat 2B-GEOPROF & 2B-GEOPROF-LIDAR product data

3. Plot vertical profile of atmosphere

- cloudsat.read_geometric_info()
  - Jointed function for reading vdata & extracting file name information
  - Extracted data array: date, granule, start time, end time….
- cloudsat.plot_profile()
  - Plot **Radar Reflectivity** **(filtering by cloud mask)** & **cloud fraction (LIDAR)**
  - Plot vertical profile along **latitude** or **longitude**



AEL satellite module – CloudSat

Filtering correct target file(s) over your study domain during specific time period

1. **Generate file list in the specific time period**

- cloudsat.generate_list()
  - Generate file list via input **product name** and **time period**
- cloudsat.cross_product_match(file_list_1,file_list_2, …)
  - Remove inconsistency files cross different 2B products

2. **Filtering files over the over your study domain**

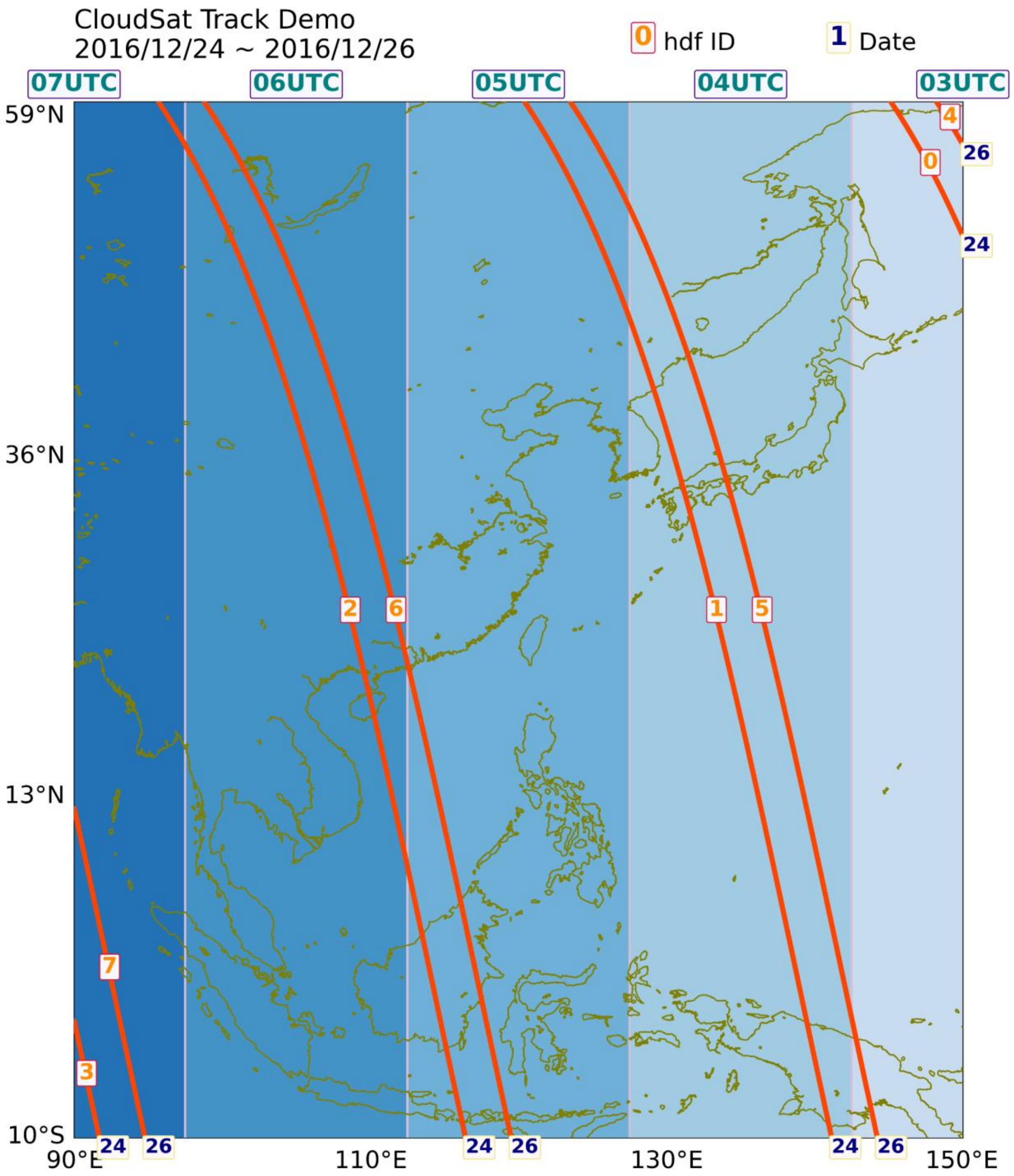
- sub_domain_file_list = cloudsat.sub_domain_check(**file_list**)
  - Use default lon-lat range of the study domain (90E~150E, 10S~60N) if there is no input lon-lat range
  - Check CloudSat track passing the study domain

AEL satellite module – CloudSat

Filtering correct target file(s) over your study domain during specific time period

3. Plot data tracks and pick target file(s)

- cloudsat.plot_track()
  - CloudSat **tracks**
  - Hdf ID: sub_domain_file_list[**0**]
  - Examine period: 2016/12/**24** ~ 2016/12/**26**
  - Approximate **UTC time** when tracks pass through



AEL satellite module – CloudSat

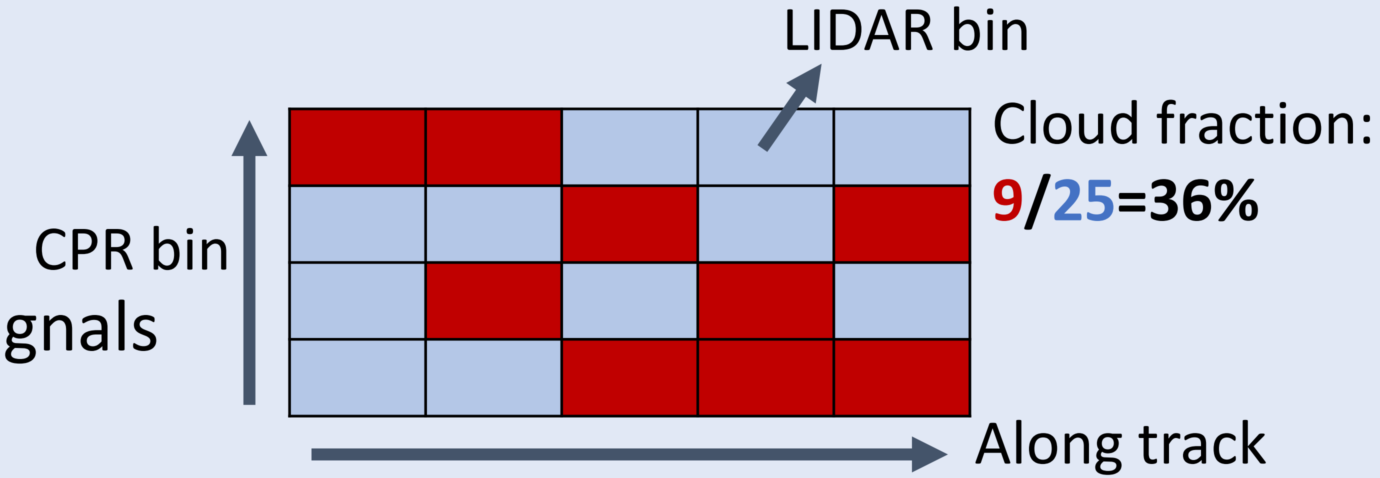
The concept of conditional sampling

1. Invalid column and masked value

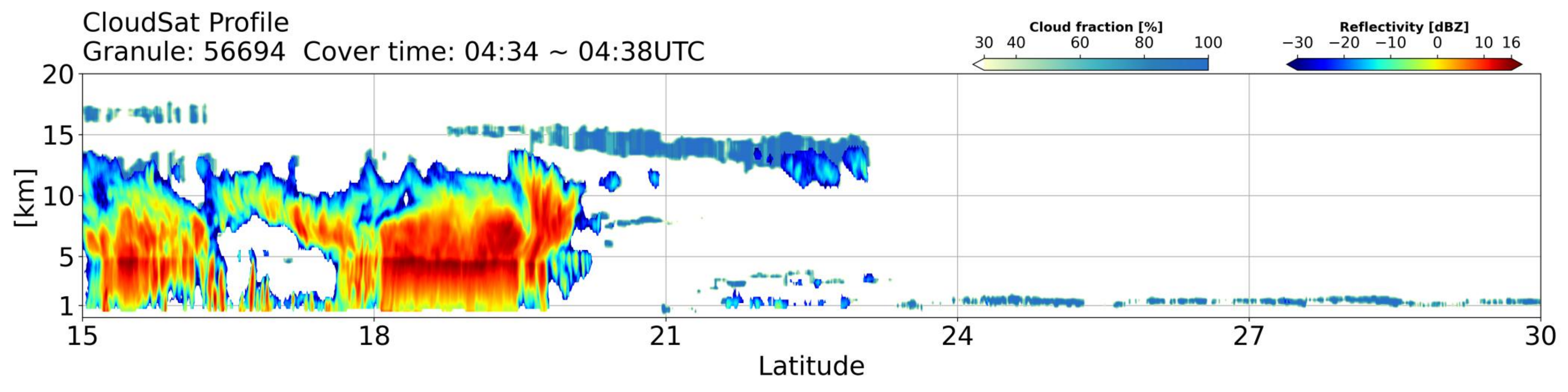
- Data quality (0: good quality column), already processed in read_data function
- Missing value (from cloudsat.cloudsat_var_attribute)
- Self-judgement masked value: **Which value will be your cloud area?**

  2B-GEOPROF
  - CPR_Cloud_mask(**>=20, defined as cloud**, see Appendix)
  - Radar Reflectivity **filtered by CPR_Cloud_mask**

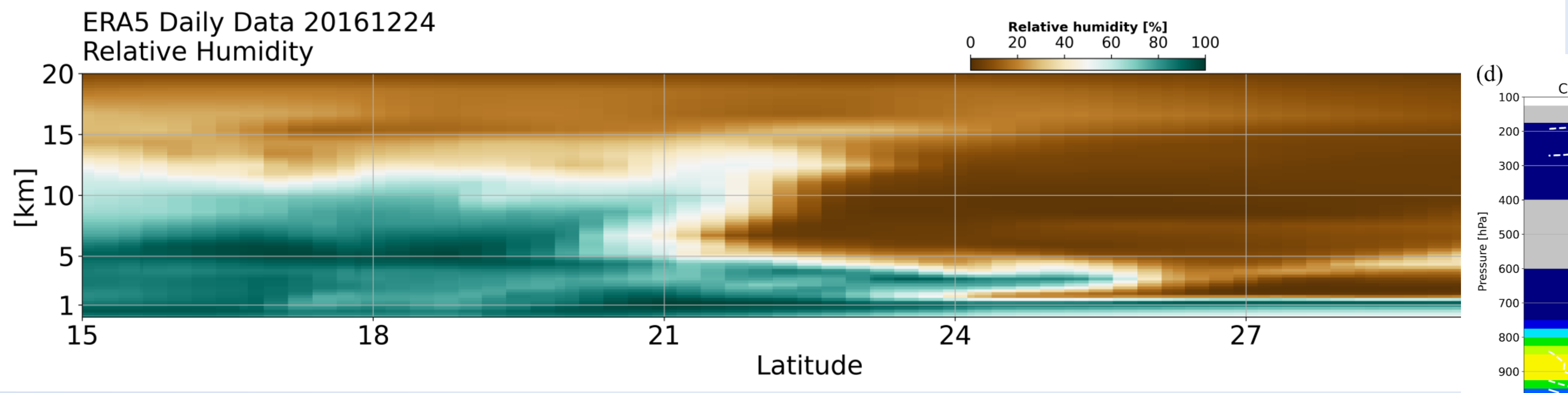
  2B-GEOPROF-LIDAR
  - UncertaintyCF
  - CloudFraction (0~100)
- Overlapping of CPR and LIDAR signals

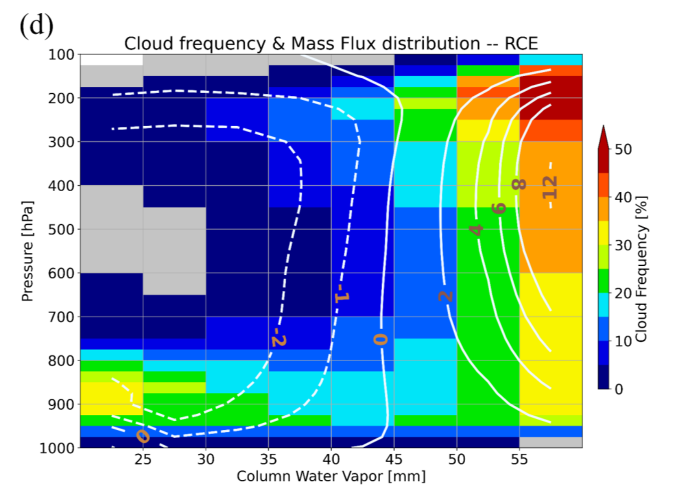


HW6: draw the CloudSat cross section using satellite module



HW7: combine satellite clouds (Himawari+CloudSat) and the environment in the vertical cross section or conditional sampling





AEL satellite module – CloudSat – function list

Filtering data list function

- cloudsat.product_list()
- cloudsat.generate_list()
- cloudsat.cross_product_match()
- cloudsat.sub_domain_check()

Plotting function

- cloudsat.plot_track()
- cloudsat.plot_track_w_rgb()
- cloudsat.plot_profile()

Read data function

- cloudsat.product_information()
- cloudsat.hdf_information()
- cloudsat.sub_domain_check()
- cloudsat.read_ori_vdata()
- cloudsat.read_ori_sddata()
- cloudsat.fit_era5_lon(**lon_list**)
- cloudsat.read_vdata()
- cloudsat.read_sddata()
- cloudsat.read_geometric_info()
- cloudsat.cloudsat_var_attribute()

Online resource

- CloudSat Data Process Center(DPC)
  - https://www.cloudsat.cira.colostate.edu/
- CALIPSO Home page
  - https://www-calipso.larc.nasa.gov/about/
- eoPortal: CloudSat introduction
  - https://www.eoportal.org/satellite-missions/cloudsat
- eoPortal: CALIPSO introduction
  - https://www.eoportal.org/satellite-missions/calipso
- The A-Train & C-Train
  - https://svs.gsfc.nasa.gov/31049/
- EarthCARE
  - https://earth.esa.int/eogateway/missions/earthcare

Reference

- Bacmeister, J. T., & Stephens, G. L. (2011). Spatial statistics of likely convective clouds in CloudSat data. Journal of Geophysical Research, 116(4), 1–21. https://doi.org/10.1029/2010JD014444
- Chen, P.-J., Chen, W.-T., Wu, C.-M., & Yo, T.-S. (2021). Convective cloud regimes from a classification of object-based CloudSat observations over Asian-Australian monsoon areas. Geophysical Research Letters, 48, e2021GL092733. https://doi.org/10.1029/2021GL092733
- Mace, G. G., & Zhang, Q. (2014). The CloudSat radar-lidar geometrical profile product (RL-GeoProf): Updates, improvements, and selected results. *Journal of Geophysical Research: Atmospheres*, *119*(15), 9441–9462. https://doi.org/https://doi.org/10.1002/2013JD021374
- Marchand, R., Mace, G. G., Ackerman, T., & Stephens, G. (2008). Hydrometeor detection using Cloudsat - An earth-orbiting 94-GHz cloud radar. *Journal of Atmospheric and Oceanic Technology*, *25*(4), 519–533. https://doi.org/10.1175/2007JTECHA1006.1
- Riley, E. M., & Mapes, B. E. (2009). Unexpected peak near−15°C in CloudSat echo top climatology. Geophysical Research Letters, 36(9), 1–5. https://doi.org/10.1029/2009GL037558
- Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al. (2002). The cloudsat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, *83*(12), 1771-1790+1742. https://doi.org/10.1175/BAMS-83-12-177
- Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., et al. (2008). CloudSat mission: Performance and early science after the first year of operation. Journal of Geophysical Research, 113(8). D00A18. https://doi.org/10.1029/2008JD009982
- Takahashi, H., Luo, Z. J., & Stephens, G. L. (2017). Level of neutral buoyancy, deep convective outflow, and convective core: New perspectives based on 5 years of CloudSat data. Journal of Geophysical Research: Atmospheres, 122(5), 2958–2969. https://doi.org/10.1002/2016JD025969

END

Appendix

Product	Core Contents
2B-GEOPROF	**Radar reflectivity, cloud mask**
2B-GEOPROF-LIDAR	**Combined CPR+CALIOP cloud mask**
2B-CLDCLASS	Cloud classification
2B-CLDCLASS-LIDAR	Cloud classification supplemented by CALIOP lidar
2B-CWC-RO	Cloud liquid/ice water content and microphysical parameters
2B-CWC-RVOD	Cloud liquid/ice water content and microphysical parameters, constrained by MODIS optical depth
2B-FLXHR-LIDAR	Radiative flux/heating rate profiles supplemented by CALIOP lidar
2C-ICE	Cloud ice water content
2C-PRECIP-COLUMN	Precipitation type and occurrence
2C-RAIN-PROFILE	Rain rate profiles
2C-SNOW-PROFILE	Snow rate profiles
2C-TB94	94 GHz brightness temperature

CPR Cloud mask

Mask Value	Meaning	% False Detections Goal	Estimated % False Detection via CALIPSO comparison
-9	Bad or missing radar data		
5	Significant return power but likely surface clutter		
6-10	Very weak echo (detected using along-track averaging)	< 50 %	44 %
20	Weak echo (detection may be artifact of spatial correlation)	< 16%	5 %
30	Good echo	< 2 %	4.3 %
40	Strong echo	< 0.2 %	0.6 %

Table 1 – Description of CloudSat cloud mask values, false detection rates, and percentage of false detections. The percent of false detection is given by 100 times the number of false detections divided by the total number of detections for the specified cloud mask value.

CPR cloud mask >= 20, defined as cloud from Bacmeister and Stephens (2011), Chen et al., (2021), Riley and Mapes (2009), and Takahashi et al. (2017)

quality check variables

Data_quality

- Dimensions: nray (alone tracks)
- **Flags indicating data quality. If 0, then data is of good quality.**
- Otherwise :

  1 = GPS data not valid

  2 = Temperatures not valid

  3 = Radar telemetry data quality is not normal

  4 = Peak power is not normal

  5 = CPR calibration maneuver.

  6 = Missing frame

  7 = Not used

quality check variables

GEOPROF-LIDAR: UncertaintyCF

- Range: 0 to 100
- Missing value: -9
- Dimensions: nbin,nray (vertical bins*alone tracks)
- The UncertaintyCF is a description of the quality of radar and lidar data. It is recorded per ray and per bin as 1-byte integer.

0 means that neither radar data nor lidar data was found.

1 indicates that only radar data has been found.

2 indicates that only lidar data has been found.

3 indicates that both radar data and lidar data were found.