National Aeronautics and Space Administration
Goddard Earth Science Data Information and Services Center (GES DISC)

README Document for Low Cloud Morphology and Deep Convective Cloud Objects, version 001

Last Revised: March 14, 2022

Goddard Earth Sciences Data and Information Services Center (GES DISC)
http://disc.gsfc.nasa.gov
NASA Goddard Space Flight Center
Code 610.2
Greenbelt, MD 20771 USA

Prepared By:

Ashley Heath	**Tianle Yuan**
Name	Name
GES DISC	Room A305 BLDG 33
GSFC Code 610.2	Code 613
12/13/2021	
Date	

Reviewed By:

Suhung Shen	01/13/2022
Reviewer Name	Date
GES DISC	
GSFC Code 610.2	

Goddard Space Flight Center
Greenbelt, Maryland

Revision History

Revision Date	Changes	Author
12/13/2021	Initial version	Ashley Heath
03/14/2022	Final draft revisions	Ashley Heath

Table of Contents

1.0 Introduction

This document provides basic information for using data records produced by the Making Earth System Data Records for Use in Research Environments (MEaSUREs) 2017 Program project titled "A Comprehensive Data Record of Marine Low-level and Deep Convective Cloud Systems Using an Object-Oriented Approach". The major goal of this project is to build a novel and science-based Earth system data record of marine low-level morphology type and deep convective cloud objects and their environmental conditions. The principal investigator responsible for this project is Tianle Yuan.

This release is composed of a beta version for three individual products from the MODerate resolution Imaging Spectrometer (MODIS) on board the Aqua satellite for the year 2017.

1.1 Dataset/Mission Instrument Description

The Low Cloud Morphology and Deep Convective Cloud Objects project aims to create an observational record that spans about two decades of low-cloud mesoscale morphology at a near-global (60°S-60°N) scale. This data collection contains marine low-cloud observations classified into six mesoscale morphology types: stratus, closed cellular convection, disorganized convection, open cellular convection, clustered cumulus convection, and suppressed cumulus convection. These low-cloud observations come primarily from the MODerate resolution Imaging Spectrometer (MODIS) on board the Aqua satellite. Individual MODIS granules are broken into 128x128 pixel scenes and are classified for morphology type. The specifics of this analysis can be found in Yuan et. al 2020.

1.2 Product Table

The Low Cloud Morphology and Deep Convective Cloud Objects project is composed of a beta version for three individual products described in Table 1.

Table 1. Summary of the beta version data products in the Low Cloud Morphology and Deep Convective Cloud Objects project.

Dataset Title	Shortname	DOI
MODIS Aqua L2 chopped blocks (block size: 128 pixels x 128 pixels, daytime) V001	MYD_L2_CB	10.5067/DFDGJR6707D8
MODIS Aqua L2 model predicted low cloud types of chopped blocks in which low cloud dominates	MYD_L2_MPLCT	10.5067/8TDZURGRLN9I

(block size: 128 pixels x 128 pixels, daytime) V001		
MODIS Aqua L3 occurrence frequency of low cloud types monthly mean and annual mean 2x2 degree resolution V001	MYD_L3_OFLCT	10.5067/3FAIC739DQRH

1.3 Data Disclaimer

Questions about the data can be sent to Dr. Yuan (tianle.yuan@nasa.gov). The methodology for creating the data and its usage can be found in Yuan et al. (2020) and Mohrmann et al. (2021).

1.3.1 Data Citation and Acknowledgment

If you use our data in a publication, we hope you will acknowledge the project appropriately. For instance:

> We thank the NCCS Team, esp. the ADAPT team, led by Dr. Carriere and previously Dr. Duffy, for their excellent computing support and the GES DISC team for their efforts in producing the data records.

NASA requests that you include the following acknowledgment in papers published using this dataset:

> "The data used in this study were acquired as part of the mission of NASA's Earth Science Division and archived and distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC)."

The digital object identifier (DOI) can also be used to precisely cite a specific dataset. The DOI for each product is listed in the IdentifierProductDOI global attribute. They are also listed in Table 1 above for reference.

The DOI registered to each data product can simply be pasted into a web browser and will resolve to a landing page that provides a recommended form for citing the data. For example, the recommended form when citing MYD_L3_OFLCT is as follows:

> Yuan, T., NASA/GSFC/GES_DISC (2021), MODIS Aqua L3 occurrence frequency of low cloud types monthly mean and annual mean 2x2 degree resolution, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: [Data Access Date], 10.5067/3FAIC739DQRH

We would appreciate receiving a copy of your publication, which can be sent to the GES DISC Help Desk by email: gsfc-dl-help-disc@mail.nasa.gov

1.3.2 Contact Information

If you have questions or feedback about this dataset and data access, please contact:

GES DISC Help Desk
Code 610.2 NASA/Goddard Space Flight Center Greenbelt, MD 20771
Phone: 301-614-5224 Fax: 301-614-5268
Email: gsfc-dl-help-disc@mail.nasa.gov

1.4 Quality Issues

The cloud type classification may have a dependence on the viewing angle, which is not related to the classification algorithm, but the nature of the observations. Though it is still under investigation, please take note of this potential dependence.

2.0 Data Organization

The Low Cloud Morphology and Deep Convective Cloud Objects is a compilation of data consisting of MODIS observations from the Aqua satellites. The observations are split into granules and for each granule in shape (nx, ny), the whole granule is chopped to small blocks in shape (np_x, np_y), where np_x = 128 and np_y = 128. A model simulates the cloud type for the low-cloud-dominated blocks over ocean. There are six low cloud types: Closed-cellular MCC, Clustered Cumulus, Disorganized MCC, Open-cellular MCC, Solid Stratus, and Suppressed Cumulus.

There are three datasets for each instrument: 1) chopped blocks, 2) model prediction of cloud types for low-cloud-dominated-blocks over the ocean, and 3) aggregated occurrence frequency of low-cloud types.

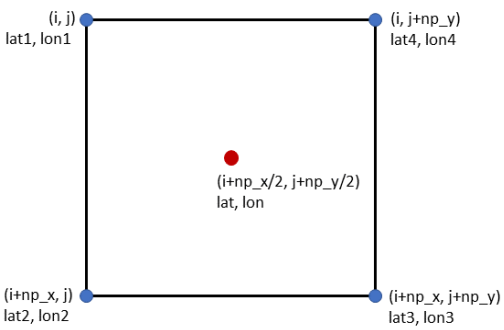
Dataset 1: Chopped block (CB)

This file provides the geolocations of all the chopped blocks in the MODIS granule data. Only daytime granule data are included. Blocks with sensor zenith angle > 45 are excluded.

block_obs: the name of the chopped block

For example, from a block named as
“MYD021KM.A2017001.0005.061_index_0128_index_1024”,
we can find the granule filename as “MYD021KM.A2017001.0005.061”,
index i = 128 and index j = 1024.

There are five latitudes and longitudes for each block shown in the figure below. More details can be found in section 3.1.1.



Dataset 2: Model predictions of cloud types for low-cloud-dominated-blocks over the ocean

This file provides the model predictions of cloud types for low-cloud-dominated blocks over the ocean. The variables include pred_cat, and pred_prob, block names, latitudes and longitudes, high-cloud fraction, and low-cloud fraction of the chopped block. Low-cloud-dominated blocks are defined by the conditions: the ratio of high-cloud fraction and low-cloud fraction is smaller than 0.2, with high-cloud fraction < 0.3 and low-cloud fraction > 0.05.

block_low: the name of the low-cloud-dominated block, from which we can find the location of the chopped block in the granule data.
pred_cat: the predicted cloud type of each block; the value is from 0 to 5
- 0: Closed-cellular MCC
- 1: Clustered Cu
- 2: Disorganized MCC
- 3: Open-cellular MCC
- 4: Solid Stratus
- 5: Suppressed Cu

pred_prob: the prediction probability of cloud type; the value is from 0 to 1.0, the higher the value, the more reliable the prediction of cloud type
hcf: high-cloud fraction of the low-cloud-dominated block valued 0-1
lcf: low-cloud fraction of the low-cloud-dominated block valued 0-1
sensor_zenith: sensor zenith angle at the center of low-cloud-dominated block

Dataset 3: Aggregated occurrence frequency of low-cloud types

This file provides the aggregated occurrence frequency of low-cloud types for both annual mean and monthly means. Variables lat and lon in Dataset 1, and variables lat, lon, and pred_cat in Dataset 2 are used to calculate occurrence frequency of six low cloud types in 2.0°x2.0° grids for 60°S-60°N and 180°W-180°E.

2.1 File Naming Convention

The Low Cloud Morphology and Deep Convective Cloud Objects data files are named in accordance with the following convention:

MODIS_<Instrument>_<ProductLevel>_<Descriptor>.V001.<TemporalResolution>.nc

Table 2.1. Attribute descriptions for the File names

Attribute	**Description**
<Instrument>	"Aqua" for the Aqua satellite
<ProductLevel>	"L2" for Level 2 data
	"L3" for Level 3 data
<Descriptor>	"CB" for Chopped blocks (Dataset 1)
	"MPLCT" for the Model predicted low cloud types (Dataset 2)
	"OFLCT" for the Aggregated occurrence frequency of low cloud types (Dataset 3)
<TemporalResolution>	"YYYYDDMM-YYYYDDMM" for daily
	"YYYYMM-YYYYMM" for monthly

Examples:
The file name for the MODIS Aqua chopped block data for the year 2017 is:
MODIS_Aqua_L2_CB.V001.20170101-20171231.nc

The file name for the MODIS Aqua model predicted low cloud types data for the year 2017 is:
MODIS_Aqua_L2_MPLCT.V001.20170101-20171231.nc

The file name for the MODIS Aqua occurrence frequency of low cloud types data for the year 2017 is:
MODIS_Aqua_L3_OFLCT.V001.201701-201712.nc

2.2 File Format

The Low Cloud Morphology and Deep Convective Cloud Objects files are in NetCDF-4 format. NetCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data.
More information can be found here: https://www.unidata.ucar.edu/software/netcdf/docs/

3.0 Data Contents

3.1 Data Products

The Low Cloud Morphology and Deep Convective Cloud Objects has a total of three data products and is a beta version for testing. It includes three data collections (CB, MPLCT, and OFLCT) from the MODIS Aqua instrument summarized in Table 3.1.

Table 3.1. Low Cloud Morphology and Deep Convective Cloud Objects beta version data products and parameters

Data Product Shortname	**Instrument**	**Spatial Resolution**	**Number of parameters**
MYD_L2_CB	MODIS Aqua	128 pixels x 128 pixels	12, See Table 3.1.1
MYD_L2_MPLCT	MODIS Aqua	128 pixels x 128 pixels	16, See Table 3.1.2
MYD_L3_OFLCT	MODIS Aqua	2.0° x 2.0°	6, See Table 3.1.3

3.1.1 Chopped Block (CB) Parameters

Table 3.1.1 below lists the parameters in the CB products.

- MYD_L2_CB

Table 3.1.1. Parameters in the Chopped Block data products

Short Name	**Long name**	**Unit**
block_obs	block name in the granule file	1
lat	Latitude at the center of the block	degrees_north
lat1	Latitude at the top left corner of the block	degrees_north
lat2	Latitude at the bottom left corner of the block	degrees_north
lat3	Latitude at the bottom right corner of the block	degrees_north
lat4	Latitude at the top right corner of the block	degrees_north
lc_flag	Low cloud flag	1
lon	Longitude at the center of the block	degrees_east
lon1	Longitude at the top left corner of the block	degrees_east
lon2	Longitude at the bottom left corner of the block	degrees_east
lon3	Longitude at the bottom right corner of the block	degrees_east
lon4	Longitude at the top right corner of the block	degrees_east

3.1.2 Model prediction low-cloud type (MPLCT) Parameters

Table 3.1.2 below lists the parameters in the MPLCT products.

- MYD_L2_MPLCT

Table 3.1.2. Parameters in the Model Predicted low-cloud type data products

Short Name	**Long name**	**Unit**
block_low	Low cloud dominated block name in granule file	1
lat	Latitude at the center of the block	degrees_north
lat1	Latitude at the top left corner of the block	degrees_north
lat2	Latitude at the bottom left corner of the block	degrees_north
lat3	Latitude at the bottom right corner of the block	degrees_north
lat4	Latitude at the top right corner of the block	degrees_north
lon	Longitude at the center of the block	degrees_east
lon1	Longitude at the top left corner of the block	degrees_east
lon2	Longitude at the bottom left corner of the block	degrees_east
lon3	Longitude at the bottom right corner of the block	degrees_east
lon4	Longitude at the top right corner of the block	degrees_east
hcf	high cloud fraction of the low cloud dominated block	1
lcf	low cloud fraction of the low cloud dominated block	1
pred_cat	predicted cloud type of the low cloud dominated block	1
pred_prob	probability of predicted cloud type	1
sensor_zenith	sensor zenith angle of the low cloud dominated block	degrees

3.1.3 Occurrence Frequency low-cloud type (OFLCT) Parameters

Table 3.1.3 below lists the parameters in the OFLCT products.

- MYD_L3_OFLCT

Table 3.1.3. Parameters in the Occurrence Frequency low-cloud type data products

Short Name	**Long name**	**Unit**
cloud_type	Low cloud type	1
lat	Latitude	degrees_north
lon	Longitude	degrees_east
time	time	months since 2017-01-01 00:00:00
frequency_annual	annual mean occurrence frequency of low cloud type	1
frequency_monthly	monthly mean occurrence frequency of low cloud type	1

3.2 Data Set Attributes (File Metadata)

In addition to the Scientific Data Set (SDS) arrays containing variables and dimension scales, global metadata is
also stored in the files. Some metadata are required by standard conventions, some are present to meet data provenance requirements and others as a convenience to users of Low Cloud Morphology and Deep Convective Cloud Objects products. A summary of global attributes present in all files is shown in Table 3.2.

Table 3.2. Global attributes in the Low Cloud Morphology and Deep Convective Cloud Objects data files

Global Attribute	**Description**	**Type**
Filename	Name of the data granule	string
ShortName	Short name for the data collection	string
LongName	Long, descriptive name for the data collection	string
VersionID	Data version	string
GranuleID	Name of the data granule (same as Filename)	string
Format	File format of the data (NetCDF)	string
RangeBeginningDate	Start date of the data in the file	string
RangeBeginningTime	Start UTC time of the data	string
RangeEndingDate	End date of the data in the file	string
RangeEndingTime	End UTC time of the data	string
SouthernmostLatitude	Southernmost latitude of global grid of data set	string
NorthernmostLatitude	Northernmost latitude of global grid of data set	string
WesternmostLongitude	Westernmost longitude of global grid of data set	string
EasternmostLongitude	Easternmost longitude of global grid of data set	string
IdentifierProductDOIAuthority	Authority through which DOI can be resolved	string
IdentifierProductDOI	Digital object identifier	string

ProductionDateTime	Date and time the current file was produced	string
ProcessLevel	Level of data processing	string
Conventions	the metadata conventions followed in the file, e.g. CF-1.8	string
Source	Instruments related to the origin of the data product	string
DataSetQuality	Overall assessment of quality of the data	string

4.0 Options for Reading the Data

4.1 Command Line Utilities

4.1.1 ncdump

The ncdump tool can be used as a simple browser for HDF data files, to display the dimension names and sizes; variable names, types, and shapes; attribute names and values; and optionally, the values of data for all variables or selected variables in a netCDF file. The most common use of ncdump is with the –h option, in which only the header information is displayed.

ncdump [-c|-h] [-v ...] [[-b|-f] [c|f]] [-l len] [-n name] [-d n[,n]] filename
Options/Arguments:
[-c] Coordinate variable data and header information
[-h] Header information only, no data
[-v var1[,...]] Data for variable(s) <var1>,... only data
[-f [c|f]] Full annotations for C or Fortran indices in data
[-l len] Line length maximum in data section (default 80)
[-n name] Name for netCDF (default derived from file name)
[-d n[,n]] Approximate floating-point values with less precision filename File name of input netCDF file

For more information about installing and using ncdump can be found from Unidata (https://www.unidata.ucar.edu/software/netcdf)

4.1.2 NCO

The netCDF Operator (NCO) is a powerful command line toolkit developed by the Earth System Science group in University of California, Irvine, which can manipulate and analyzes data stored in NetCDF. More information can be found at http://nco.sourceforge.net.

4.1.3 CDO

Climate Data Operators (CDO) is a collection of command line operators to manipulate and analyze Climate and NWP model Data. CDO is a large tool set for working on climate and NWP model data. NetCDF 3/4, GRIB 1/2 including SZIP (or AEC) and JPEG compression, EXTRA, SERVICE and IEG are supported as IO-formats. Apart from that, CDO can be used to analyze any kind of gridded data not related to climate science. CDO has very small memory requirements and can process files larger than the physical memory. More information can be found at Overview - CDO - Project Management Service (mpg.de).

4.2 Tools/Programming

This section lists some tools, but not limited, that may be used to read, visualize, and process this dataset:

4.2.1 Panoply

Panoply is a data viewer that displays geo-referenced arrays in NetCDF, HDF, and GRIB formats. The first-time user may download the software from NASA Goddard Institute for Space and Studies
(http://www.giss.nasa.gov/tools/panoply/).
Examples to use Panoply for GES DISC archived data can be found in data HowTo:
https://disc.gsfc.nasa.gov/information/howto?keywords=panoply&page=1

4.2.2 GrADS

The Grid Analysis and Display System (GrADS) is an interactive tool developed by the COLA (Center for Ocean-Land-Atmosphere Studies) group at George Mason University, which can read, visualize, and analyze gridded (i.e. Level 3 and Level 4) data files in a number of formats, including NetCDF, HDF, binary, and GRIB, as well as station data in BUFR format (http://cola.gmu.edu/grads/).
Examples of HowTo articles on using GrADS with GES DISC data can be found here:
https://disc.gsfc.nasa.gov/information/howto?keywords=grads&page=1

5.0 GES DISC Data Services

Access to GES DISC data requires a free Earthdata login profile. Please see the following instructions for creating an Earthdata login profile:

https://wiki.earthdata.nasa.gov/display/EL/How+To+Register+For+an+EarthData+Login+Profile

This data is stored online and may be accessed through several methods.

5.1 Direct Data Access

The data can be downloaded or remote accessed through HTTPS service:
MYD_L2_CB:
https://measures.gesdisc.eosdis.nasa.gov/data/CLOUD_MODIS/MYD_L2_CB_001/
MYD_L2_MPLCT:
https://measures.gesdisc.eosdis.nasa.gov/data/CLOUD_MODIS/MYD_L2_MPLCT/
MYD_L3_OFLCT:
https://measures.gesdisc.eosdis.nasa.gov/data/CLOUD_MODIS/MYD_OFLCT_L3/

or through OPeNDAP service:
MYD_L2_CB:
https://measures.gesdisc.eosdis.nasa.gov/opendap/CLOUD_MODIS/MYD_L2_CB_001/
MYD_L2_MPLCT:
https://measures.gesdisc.eosdis.nasa.gov/opendap/CLOUD_MODIS/MYD_L2_MPLCT/
MYD_L3_OFLCT:
https://measures.gesdisc.eosdis.nasa.gov/opendap/CLOUD_MODIS/MYD_OFLCT_L3/

5.3 Help Resources

If you need assistance or wish to report a problem:
Email: gsfc-dl-help-disc@mail.nasa.gov
Voice: 301-614-5224
Fax: 301-614-5268
Address:
Goddard Earth Sciences Data and Information Services Center NASA Goddard Space Flight Center Code 610.2 Greenbelt, MD 20771 USA

5.3.1 How To Articles

The GESDISC web site contains many informative articles under the "How To Section", "FAQ" (frequently asked questions), "News", "Glossary", and "Help" . A sample of these articles includes:

Earthdata Login for Data Access

How to Download Data Files from HTTPS Service with wget

How to Obtain Data in NetCDF Format via OpeNDAP

Quick View Data with Panoply

How to Read Data in HDF-5 or netCDF Format with GrADS

6.0 Acknowledgments

This project is funded by the NASA's MEaSUREs program managed by Dr. Lucia Tsaoussi.

7.0 References

J. Mohrmann, R. Wood, **T. Yuan**, H. Song, R. Eastman, L. Oreopoulos. Identifying meteorological influences on marine low cloud mesoscale morphology using deep learning classifications, Atmos. Chem. And Phys., 21 (12), 9629-9642 https://doi.org/10.5194/acp-2020-1026, 2021.

Yuan, T., Song, H., Wood, R., Mohrmann, J., Meyer, K., Oreopoulos, L., and Platnick, S.: Applying deep learning to NASA MODIS data to create a community record of marine low-cloud mesoscale morphology, Atmos. Meas. Tech., 13, 6989–6997, https://doi.org/10.5194/amt-13-6989-2020, 2020.