CloudSat Project

A NASA Earth System Science Pathfinder Mission

CloudSat Level 3 RMCP Gridded Data Product
(Reflectivity - Cloud Mask - Cloud Class - Precipitation)

Process Description and Interface Control Document

Algorithm Version: V0002
Date: 6 October 2020

Questions concerning the document and proposed changes shall be addressed to

John M. Haynes
john.haynes@colostate.edu
(970) 491-8535

Document Revision History

Date	Revision	Description	Section(s) Affected
20 September 2019	0	Initial release	All
6 October 2020	1	Revised algorithm ver. V0002	2,3,4,5,7,8,A

Contents

1 Introduction

This document provides an overview of the CloudSat Level 3 RMCP (Reflectivity - Cloud Mask - Cloud Class - Precipitation) gridded data product. There are two broad categories of Level 3 products.

- Level 3-Full products contain histograms of cloud, precipitation, and reflectivity variables on vertical levels and in columns.
- Level 3-Simplified products are wholly formed using data available in the Level 3-Full products, and contain variables and classifications in a more user-friendly format. These products contain means of radar reflectivity, and cloud/precipitation variables broken into simpler categorizations than the Level 3-Full products.

It is recommended that most users start with the Level 3-Simplified product. Level 3-Simplified products are intended to answer common inquiries about occurrence of clouds of various types (or classes, as they are referred to throughout this document), as well as precipitation occurrence observed by CloudSat, on both vertical levels and in columns. The Level 3-Full products are, in some sense, *expert level* products. They will be most useful to advanced users who find that the Level 3-Simplified products are aggregated in ways that do not meet their particular needs.

2 Product Resolution, Format, and Naming

2.1 Horizontal resolution

Products are produced at at three fixed horizontal resolutions: $2.5^\circ \times 2.5^\circ$, $5^\circ \times 5^\circ$, and $10^\circ \times 10^\circ$ latitude/longitude. The variables *lat* and *lon* (dimensions of the same names) are of type float and contain the latitudes and longitude of the product, respectively, in degrees.

Figure 1 shows the number of unique day and night visits of CloudSat to a horizontally-oriented grid box of each of these sizes (and $1^\circ \times 1^\circ$, for reference) during a 30 day period; Figure 2 shows the zonal average of this number, as well as the average number of associated profiles that are averaged together per grid box. This number of unique visits generally increases from low latitudes to high latitudes since the orbits converge near the poles. At $1^\circ \times 1^\circ$, some grid boxes are never visited. To avoid this situation and to remove many (but not all) orbital artifacts, the finest resolution Level 3 product that is made available is $2.5^\circ \times 2.5^\circ$.

For a given application, it is important to choose a horizontal and temporal resolution that is appropriate for the problem; a discussion of statistical significance can be found Section 2.4. Some examples of cloud fraction from RMCP products plotted at varying horizontal (and temporal) resolutions are shown in Appendix A for reference.

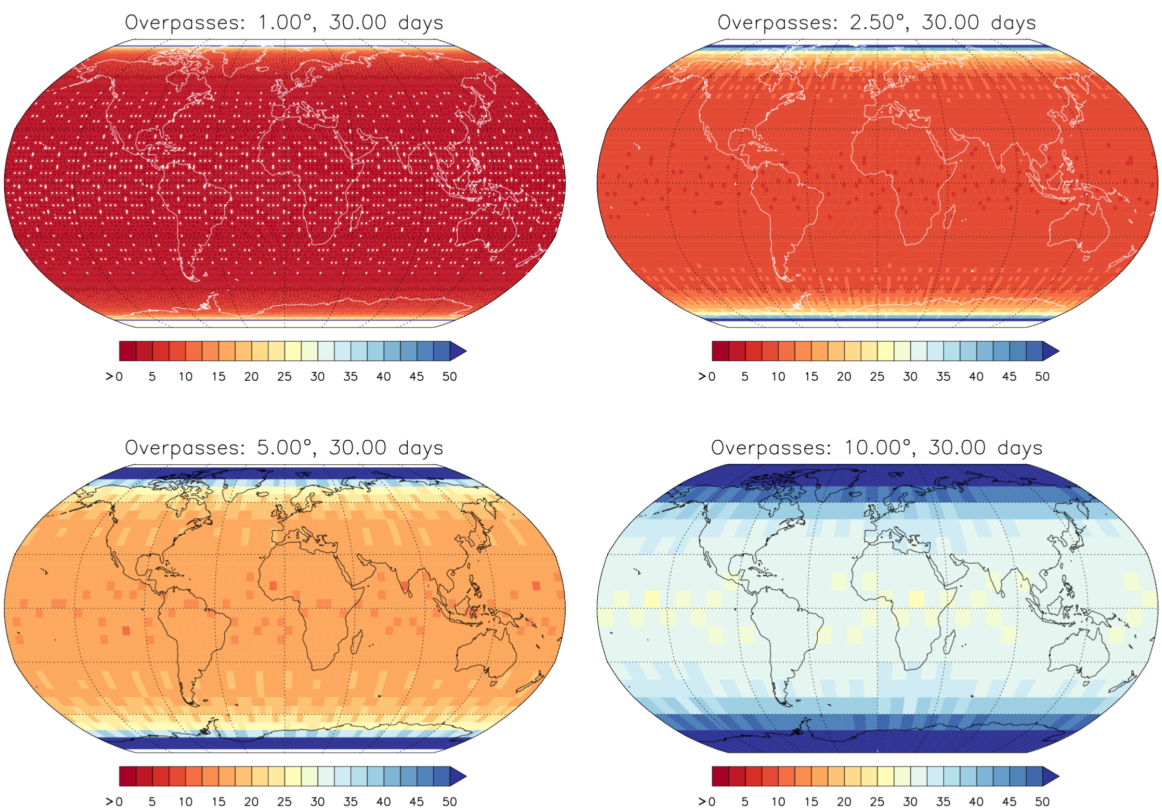

Figure 1: Number of unique visits of CloudSat to a horizontally-oriented grid box of size 1°×1°, 2.5°×2.5°, 5°×5°, and 10°×10° during a 30 day period. Assumes both day and night radar operation.

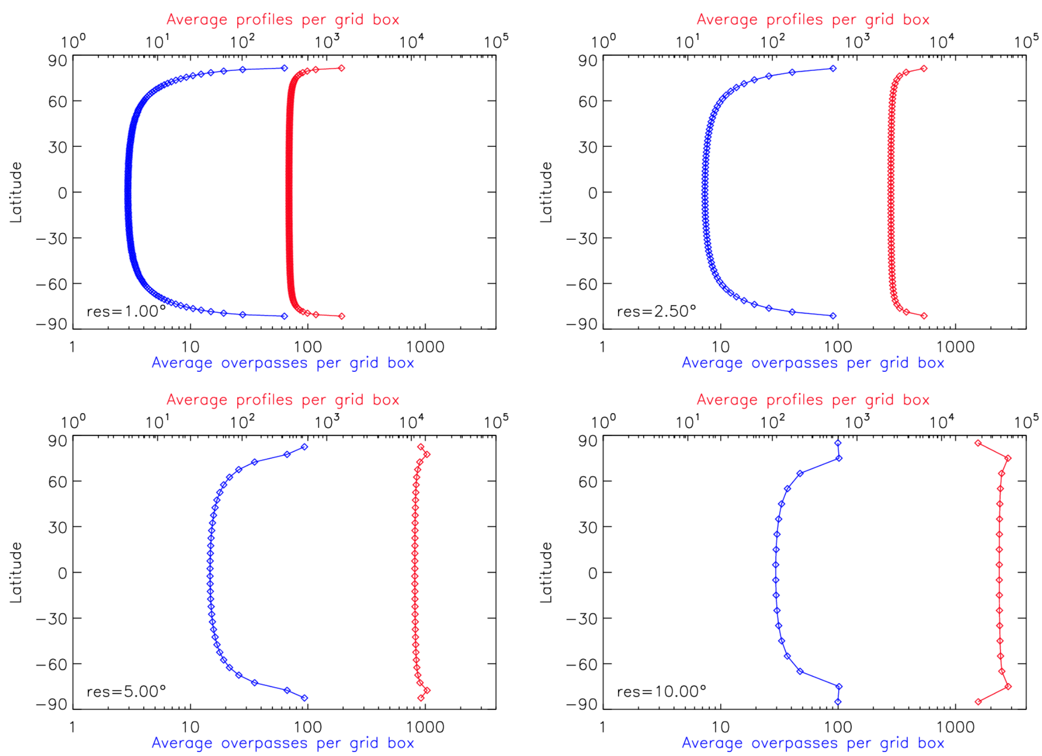

Figure 2: Zonal average number of unique visits of CloudSat to a horizontally-oriented grid box of size $1^\circ \times 1^\circ$, $2.5^\circ \times 2.5^\circ$, $5^\circ \times 5^\circ$, and $10^\circ \times 10^\circ$ during a 30 day period (blue, bottom axis), and the average number of profiles per unique grid box visit (red, top axis). Assumes both day and night radar operation. In the bottom two panels, the change in slope of the curves at the highest latitude points is due to these grid boxes including some area outside the observable limits of the satellite.

2.2 Vertical resolution

The Level 3 products retain the full vertical resolution of the radar, which is 240 m (as oversampled and reported in the 2B-GEOPROF products). Variables are binned for each CloudSat profile based on the height of the bin above mean sea level.

The variable *height* (dimension of the same name) is of type float and contains the height values used in the product, which spans the range of -480 to 18000 m above mean sea level in 240 m increments (77 levels).

There are two broad types of variables in the Level products: "level variables" and "column variables".

Level variables are specified for each *lon*, *lat*, and *height*.

Column variables are specified for each *lon* and *lat*. They are independent of *height*.

2.3 Temporal resolution

The products are regularly produced for three time periods, subject to Level 2 data availability:

- month
- season (a three month period, such as June-July-August)
- year

See section 6.3 for more information about how missing input data may preclude the production of Level 3 products for a given time period. See section 2.4 for more information on statistical significance.

2.4 Statistical significance

Users are encouraged to think carefully about the statistical significance of the products that they are using. For a product of a given horizontal resolution, the statistical significance of results derived from the product increases with the number of months that are aggregated together; and similarly, for a given timescale, the statistical significance increases as horizontal resolution becomes more coarse (but the area of applicability becomes larger). Figure 2, for example, shows that when using a $2.5°\times2.5°$ product, tropical grid boxes are, on average, overpassed less than 10 times in a given 30 day period. It would therefore be inadvisable to use a monthly $2.5°\times2.5°$ product on its own (i.e. without aggregating together with other months) to make a statistically significant statement about the distribution of clouds at a tropical location.

It is challenging to quantify the time period over which one must aggregate products of a given horizontal resolution to achieve statistical significance. It is possible to make a crude estimate of this, however, by assuming that a cloud variable of interest is normally distributed and randomly sampled. The validity of the first assumption depends on the variable itself; the validity of the second assumption requires that the variable of interest be independent of time of day (since the time of sampling is determined by the CloudSat orbital characteristics, and is not random). Since clouds often exhibit diurnal variation, this assumption may be violated in certain regimes, introducing considerable uncertainty into the calculations that follow.

With these caveats noted, if each unique grid box overpass is considered an independent observation of the variable of interest, then the number of observation days that must be aggregated together to achieve an acceptable margin of error (MoE) at the 95% confidence level can be calculated. Some such estimates are shown in Figure 3 at the three available product resolutions, for MoE of 10, 20, and 30%. It is apparent from the top figure, for example, that at an equatorial location one must aggregate the $2.5°\times2.5°$ product for approximately three months to achieve a 20% margin of error; for the $10°\times10°$ product, this can be achieved in under one month.

2.5 Splitting of Level 3-Full $2.5°\times2.5°$ files by latitude band

The Level 3-Full $2.5°\times2.5°$ files would be extremely large if distributed as a single file, so the data is divided into three latitude bands. Data for each band is written to a separate file. These bands are given by:

- Southern band (-90 to $-30°$ latitude)
- Tropical band (-30 to $30°$ latitude)
- Northern band (30 to $90°$ latitude)

The filename (Section 2.6) and global attributes (Section 7) specify the latitude band contained in a Level 3-Full $2.5°\times2.5°$ file.

2.6 File naming convention

Files are named according to the following convention:

`ZZZZZZ_CS_3X-PROD_[AA]x[BB]_R[CC]_V[DDDD]_U[FFF]_L[GG].nc`

ZZZZZZ Indicates time period represented, in one of the following formats:

YYYY-MM month MM of year YYYY

YYYY-MMM season (three month period) of year YYYY; note that the year corresponds to the start of the season, so 2008-DJF represents December 2008 through February 2009

YYYY year YYYY

YYY1-M1-YYY2-M2 month M1 of year YYY1 through month M2 of year YYY2

In the above notation, years are referenced as the four digit year (e.g. 2007). Months are represented by their two digit representation (e.g. 01 for January, 12 for December). Seasons are

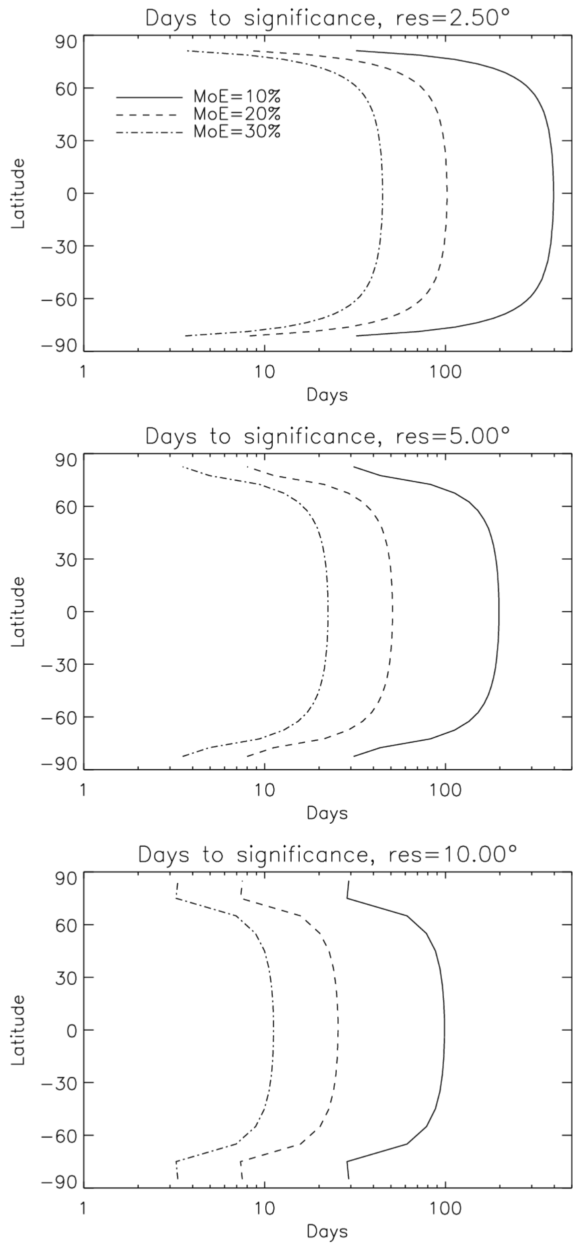

Figure 3: Number of days over which one must aggregate data at a resolution of 2.5°×2.5°, 5°×5°, and 10°×10° to obtain the given margin of error (MoE) with 95% confidence, under the assumptions outlined in the above text.

represented by either DJF (December-January-February), MAM (March-April-May), JJA (June-July-August), or SON (September-October-November).

X ‘F‘ for level 3-Full, or ‘S‘ for level 3-Simplified

PROD product name (in this case, RMCP)

AA longitude resolution in degrees

BB latitude resolution in degrees

CC CloudSat product revision, e.g. 05 for R05 products

DDDD Level 3 algorithm version, starting with 0001 for first production run

FFF A run number, starting with 001. Each run corresponds to a fixed set of input Level 2 granules. See section 6 for a discussion of algorithm version and run number.

GG This is used only with Level 3-Full files at 2.5°×2.5° resolution, and indicates the latitude band contained in the file (see Section 2.5). ’SO’ indicates southern, ’TR’ indicates tropical, and ’NO’ indicates northern band.

2.7 File format

Level 3 products are distributed in netCDF-4 format and are compatible with the CF version 1.6 metadata convention (http://cfconventions.org).

A very capable viewer for the these products is NASA’s Panoply application:
https://www.giss.nasa.gov/tools/panoply

3 Level 3 Product Setup

The Level 3-Full and 3-Simplified products contain data segregated into the following categories:

- **reflectivity** classification (**Z**)

  Classification based on radar reflectivity, with gaseous attenuation not removed (2B-GEOPROF variable *Radar_reflectivity*)

- **cloud mask** classification (**C**)

  Classification based on the presence of cloud according to the 2B-GEOPROF cloud mask (2B-GEOPROF variable *CPR_Cloud_mask*)

- **cloud class** classification (**D**)

  Classification based on the cloud class (type) (2B-CLDCLASS variable *cloud_scenario*)

- **precipitation state** classification (**P**)

Classification based on the presence of precipitation at the surface (2C-PRECIP-COLUMN variable *Precip_flag*)

- **DO-op state** classification (**E**)

  Classification based on Daylight-only Operations mode. See description in the text below.

Level 3 data is created by looping over all available 2B-GEOPROF (Marchand et al. 2008) granules within a given time range, then looping over all vertically-oriented range bins. Whether a granule falls within a given time range is determined by the time of the first profile in the granule. For every 2B-GEOPROF granule that is located, an associated downstream 2B-CLDCLASS (Sassen and Wang 2008) and 2C-PRECIP-COLUMN (Haynes et al. 2009) granule may also be present; if so, information from these granules is also read and processed. This allows any radar bin to be classified into a state for cloud mask (C), cloud class (D), precipitation presence at the surface (P), and DO-Op state (E) as described below. Since precipitation is only classified at the surface, all vertical bins within the same radar profile will be assigned the same precipitation state. The possible values of C, D, P, and E are shown in Tables 1, 2, 3, and 4 respectively.

Table 1: Cloud mask classification (C) based on *CPR_Cloud_mask* from 2B-GEOPROF

Value of C	CPR_Cloud_mask	Meaning
C0	0	clear
C1	>0 and <20	cloud unlikely or clutter
C2	20	cloud possible, but weak echo
C3	30	cloud probable
C4	40	cloud very likely
C5	<0 or >40	unknown or missing

Table 2: Cloud class (D) based on *cloud_class* from 2B-CLDCLASS

Value of D	cloud_class bit values	Meaning
D0	0000	clear
D1	0001	cirrus
D2	0010	altostratus
D3	0011	altocumulus
D4	0100	stratus
D5	0101	stratocumulus
D6	0110	cumulus
D7	0111	nimbostratus
D8	1000	deep convection
D9	other	unknown or missing

Note that while 2B-GEOPROF granules are required for Level 3 processing, the associated 2C-PRECIP-COLUMN and 2B-CLDCLASS granules are not. (These granules may be missing for a variety of reasons, such as missing upstream inputs, algorithm errors, delays in processing, etc.) Therefore there will be instances when the cloud mask can be characterized, but the precipitation state or cloud class cannot. In these instances C or D is set to the ’unknown or missing’ value.

Three variables in the Level 3 products specify which Level 2 granules were used as inputs. The dimension of these variables is num_granule:

Table 3: Precipitation state (P) at the surface based on *Precip_flag* from 2C-PRECIP-COLUMN

Value of P	Precip_flag	Meaning
P0	0	no precipitation
P1	1	rain possible
P2	2	rain probable / drizzle
P3	3	rain certain
P4	4	snow possible
P5	5	snow certain
P6	6	mix possible
P7	7	mix certain
P8	<0 or >7	unknown or missing

- *Granule_2B_GEOPROF* is of type int and contains a list of component 2B-GEOPROF granule numbers
- *Granule_uses_precip_flag* is of type short and is set to 1 if the corresponding 2C-PRECIP-COLUMN granule was used in the analysis (otherwise it is set to 0)
- *Granule_uses_cloudclass_flag* is of type short and is set to 1 if the corresponding 2B-CLDCLASS granule was used in the analysis (otherwise it is set to 0)

Users of these products should be aware that CloudSat operated during both daytime and nighttime from the start of the mission in 2006 through 2011 April 17, but experienced a battery anomaly at that time that resulted in a long period when the radar was shutdown. When operations resumed on 2011 October 28, CloudSat operated in a Daylight-only Operations mode (DO-Op) to preserve charge in the faulty battery. DO-Op operations continue to the present date. Comparisons between data from before and after this period may introduce day/night biases.

To account for this, radar bins are also classified into a DO-Op state (E) as described in Table 4. This indicates whether the observation was made in the pre or post DO-op period. It also reveals, for pre DO-Op observations, whether the observation *would* have been made, had it occurred in the post DO-Op period (in other words whether the radar would have been sending and receiving pulses in this portion of the orbit). To simulate DO-Op data collection, the latitudes of the first and last observable profile in a granule were collected from a random subset of CloudSat data from 2013 to 2018, as a function of day of year, and a curve was fit to these data (Figure 4). This curve allows the simulation of DO-Op data collection patterns in the pre DO-Op period.

Table 4: DO-Op state (E)

Value of E	Meaning
0	Profile collected in pre DO-Op period; would not have been observed in the post DO-Op period
1	Profile collected in pre DO-Op period; would have been observed in the post DO-Op period
2	Profile was collected in post DO-Op period

More information about significant dates and events in the CloudSat mission that may affect interpretation of Level 3 products can be found on the CloudSat Data Processing Center webpage:

http://www.cloudsat.cira.colostate.edu/data-products/information/epochs

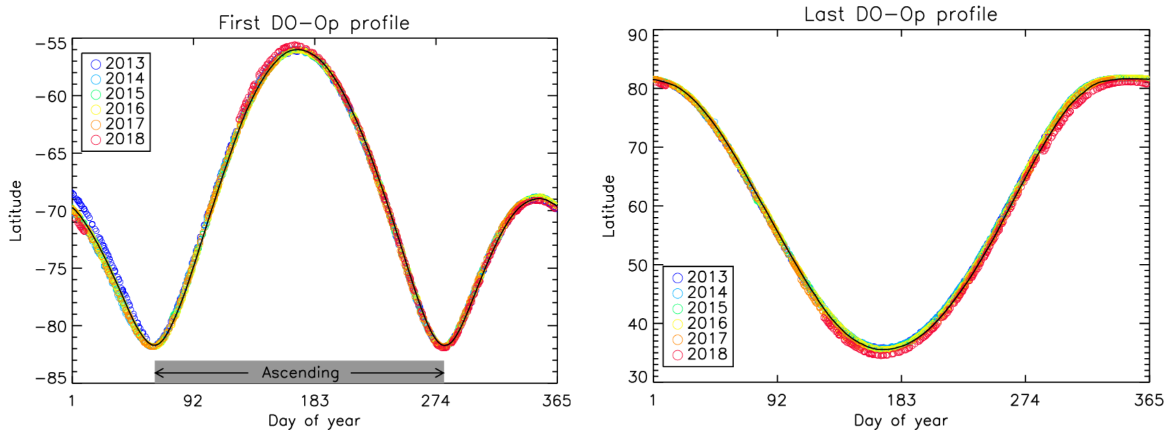

Figure 4: Latitude of first and last DO-Op observed profile from a random set of CloudSat granules from 2013 through 2018 (symbols). Curve fit to the data shown in solid black line. The curve applies to the descending portion of the orbit except for the time period denoted by the gray box labeled "Ascending".

4 Level 3-Full

Level 3-Full products contain raw counts of events occurring as a function of latitude, longitude, height, reflectivity, cloud mask, cloud class, and precipitation type at the surface. An event is defined as an observation from one radar bin within one vertical profile.

Level 3-Full does not further process these raw counts. As noted in the introduction, such processing is instead output in the Level 3-Simplified products.

4.1 Dimensions used by Level 3-Full

4.1.1 `cmask`, `cclass`, and `precip` dimensions

The values associated with the `cmask`, `cclass`, and `precip` dimensions used in Level 3-Full are given by Tables 1, 2, and 3, respectively.

4.1.2 `refl` dimension

There are 39 values associated with the `refl` dimension which are described in Table 5. A range of reflectivity values are retained, including atmospheric returns (reflectivities less than about 25 dBZ) and larger returns associated with the earth's surface. No correction is made for gaseous attenuation. The `refl` dimension labels provided in the product files are a simplification of the actual values shown in Table 5.

Table 5: Meaning of `refl` dimension

Value	Range (dBZ)	Value	Range (dBZ)	Value	Range (dBZ)
0	-36 to -34	13	-10 to -8	26	16 to 18
1	-34 to -32	14	-8 to -6	27	18 to 20
2	-32 to -30	15	-6 to -4	28	20 to 22
3	-30 to -28	16	-4 to -2	29	22 to 24
4	-28 to -26	17	-2 to 0	30	24 to 26
5	-26 to -24	18	0 to 2	31	26 to 34
6	-24 to -22	19	2 to 4	32	34 to 42
7	-22 to -20	20	4 to 6	33	42 to 50
8	-20 to -18	21	6 to 8	34	50 to 58
9	-18 to -16	22	8 to 10	35	58 to 64
10	-16 to -14	23	10 to 12	36	≥ 64
11	-14 to -12	24	12 to 14	37	< -36
12	-12 to -10	25	14 to 16	38	missing

4.1.3 `ccol` dimension

The `ccol` dimension is associated with a test of whether the 2B-GEOPROF *CPR_cloud_mask* indicates there is cloud anywhere within a given radar profile. It has three possible values which are described in Table 6.

Table 6: Meaning of `ccol` dimension

Value	Meaning	Value of C in Table 1	Value of CPR_Cloud_mask
0	No cloud in column	$C2$, $C3$, and $C4$ not in column	>=20 and <=40 not in column
1	Cloud somewhere in column	$C2$, $C3$, or $C4$ in column	>=20 or <=40 in column
2	Column cloud not determined	$C5$ in column	<0 or >40 in column

4.1.4 `cclass_col` dimension

The `cclass_col` dimension is associated with a test of whether the class of cloud provided by 2B-CLDCLASS indicates there is cloud of a given class anywhere within a given radar profile. It has nine possible values which are described in Table 7.

4.1.5 `doop` dimension

The `doop` dimension indicates whether an observation was taken in the pre or post DO-Op period, and in the former case, whether the observation would have been made had it occurred in the post DO-Op period. See Table 4.

Table 7: Meaning of `cclass_col` dimension

Value	Type of cloud in column	Value of D in Table 2
0	No cloud in column	None of $D1$, $D2$, ..., $D8$
1	Cirrus	$D1$
2	Altostratus	$D2$
3	Altocumulus	$D3$
4	Stratus	$D4$
5	Stratocumulus	$D5$
6	Cumulus	$D6$
7	Nimbostratus	$D7$
8	Deep convection	$D8$

4.2 Level 3-Full variables on levels

4.2.1 *Level_count* variable

The variable *Level_count* is dimensioned (`lat` x `lon` x `height` x `cmask` x `cclass` x `precip` x `doop`) and is of type int. It contains counts of event occurrence as a function of the reflectivity, cloud mask, cloud class, and precipitation classifications.

4.3 Level 3-Full variables in columns

4.3.1 *Column_count* variable

The variable *Column_count* is dimensioned (`lat` x `lon` x `ccol` x `precip` x `doop`) and is of type int. It contains counts of event occurrence anywhere in a column as a function of the column cloud and precipitation classifications.

4.3.2 *Column_class_count* variable

The variable *Column_count* is dimensioned (`lat` x `lon` x `cclass_col` x `precip` x `doop`) and is of type int. It contains counts of event occurrence anywhere in a column as a function of cloud class in the column and precipitation classifications.

4.3.3 *Column_count_total* variable

The variable *Column_count_total* is dimensioned (`lat` x `lon` x `precip` x `doop`) and is of type int. It contains counts of event occurrence anywhere in a column as a function of precipitation classifications.

This variable is intended to be used in conjunction with *Column_class_count*. Since multiple cloud classes can occur in one radar profile at the same time, this variable provides a total count of the number of times a particular column was visited during the analysis.

5 Level 3-Simplified

Level 3-Simplified undergoes more processing than Level 3-Full so-as to produce a more user-friendly product, with some choices (such as what cloud mask values constitute "cloudy") made for the user.

5.1 Simplified cloud mask, cloud class, and precipitation dimensions for Level 3-Simplified

For ease of use, three new dimensions are created for the Level 3 products that contain combinations of the C, D, and P classifications.

- `cmask_s`: simplified cloud mask dimension
- `cclass_s`: simplified cloud class dimension
- `precip_s`: simplified precipitation classification dimension
- `doop_s`: simplified DO-Op classification dimension

5.1.1 `cmask_s` dimension

There are two values associated with the `cmask_s` dimension which are described in Table 8.

Table 8: Meaning of `cmask_s` dimension

Value	Meaning	Value of C in Table 1	Value of CPR_Cloud_mask
0	All cases (Clear+Cloudy)	$C0$, $C1$, $C2$, $C3$, $C4$	>=0 and <= 40
1	Cloud present	$C2$, $C3$, $C4$	>=20 and <=40

5.1.2 `cclass_s` dimension

There are nine values associated with the `cclass_s` dimension which are described in Table 9 (these are a subset of the cloud class state values described in Table 1).

5.1.3 `precip_s` dimension

There are seven values associated with the `precip_s` dimension which are described in Table 10 (these are a subset of the precipitation state values described in Table 3).

5.1.4 `doop_s` dimension

There are two values of the simplified DO-Op dimension which are described in Table 11.

Table 9: Meaning of `cclass_s` dimension

Value	Meaning	Value of D in Table 2
0	All cases (Clear+Any cloud type)	$D1$, $D2$, ..., $D8$
1	Cirrus	$D1$
2	Altostratus	$D2$
3	Altocumulus	$D3$
4	Stratus	$D4$
5	Stratocumulus	$D5$
6	Cumulus	$D6$
7	Nimbostratus	$D7$
8	Deep convection	$D8$

Table 10: Meaning of `precip_s` dimension

Value	Meaning	Value of P in Table 3	Value of Precip_flag
0	All cases	$P0$, $P1$, ..., $P7$	0, 1, ..., 7
1	Precip=no	$P0$	0
2	Precip=yes	$P2$, $P3$, $P5$, $P7$	2, 3, 5, 7
3	Drizzle	$P2$	2
4	Rain+drizzle	$P2$, $P3$	2, 3
5	Snow	$P5$	5
6	Mix	$P7$	7

Table 11: Meaning of `doop_s` dimension

Value	Meaning	Value of E in Table 4
0	All cases	$E0$, $E1$, $E2$
1	Profile either was, or would have been, observed in DO-Op	$E1$, $E2$

5.2 Level 3-Simplified variables on levels

5.2.1 *Counts_on_levels* variable

The variable *Counts_on_levels* is dimensioned (`lat` x `lon` x `height` x `cmask_s` x `cclass_s` x `precip_s` x `doop_s`) and is of type int. It contains counts of event occurrence as a function of the simplified cloud mask, cloud class, and precipitation classifications.

5.2.2 *Occurrence_on_levels* variable

The variable *Occurrence_on_levels* is dimensioned (`lat` x `lon` x `height` x `cmask_s` x `cclass_s` x `precip_s` x `doop_s`) and is of type float. It contains the frequency of occurrence of events as a function of the simplified cloud mask, cloud class, and precipitation classifications.

It can be used to determine cloud frequency of occurrence, the frequency of occurrence of a specific cloud class, and the frequency of occurrence of precipitation at a given, latitude, longitude, and height.

For example, at a location with (`lat`, `lon`, `height`, `doop_s`) indices given by (i, j, k, e):

(1) The overall frequency of occurrence of cloud is given by *Occurrence_on_levels* (i,j,k,1,0,0,e)

(2) The frequency of occurrence of cumulus is given by *Occurrence_on_levels* (i,j,k,0,6,0,e)

(3) The frequency of occurrence of nimbostratus when snow is also occurring at the surface is given by *Occurrence_on_levels* (i,j,k,0,7,5,e)

Note that the overall frequency of occurrence of cloud on a given level should be determined as in example 1 (using the `cmask_s` dimension) if the class of cloud is not of interest. Although it is possible to calculate an overall cloud frequency of occurrence using the `cclass_s` dimension by summing over all possible cloud classes, this may not produce an accurate result! i.e.

$$\begin{aligned} &Occurrence\_on\_levels(i,j,k,1,m,n,e) \\ \neq &\sum_{t=1}^{8} Occurrence\_on\_levels(i,j,k,0,t,n,e). \end{aligned} \qquad (1)$$

The reason for the discrepancy is that cloud frequency of occurrence determined using the `cmask_s` dimension is valid regardless of whether 2B-CLDCLASS inputs were available.

5.2.3 ***Reflectivity_on_levels*** variable

The variable *Reflectivity_on_levels* is dimensioned (`lat` x `lon` x `height` x `cmask_s` x `cclass_s` x `precip_s` x `doop_s`) and is of type float. It contains the mean radar reflectivity as a function of the simplified cloud mask, cloud class, and precipitation classifications.

The mean radar reflectivity is calculated from the Level 3-Full reflectivity bins using

$$Z_{mean} = 10\, log_{10} \left(\sum_{i=1}^{36} C_i\, 10^{Ze_i/10} + C_{37}\, 10^{Ze_{36}/10} + C_{38}\, 10^{Ze_0/10} \right), \qquad (2)$$

where C_i is the count of radar reflectivity values in the reflectivity bin i, with midpoint value Ze_i. Reflectivity bins are defined in Table 5.

For example, at a location with (`lat`, `lon`, `height`, `doop_s`) indices given by (i, j, k, e):

(1) The mean radar reflectivity of cloud is given by *Reflectivity_on_levels* (i,j,k,1,0,0,e)

(2) The mean radar reflectivity of cirrus is given by *Reflectivity_on_levels* (i,j,k,0,1,0,e)

(3) The mean radar reflectivity of cirrus when underlying clouds are producing surface precipitation is given by *Reflectivity_on_levels* (i,j,k,0,1,2,e)

5.3 Level 3-Simplified variables in columns

5.3.1 ***Counts_in_column*** variable

The variable *Counts_in_column* is dimensioned (`lat` x `lon` x `cmask_s` x `precip_s` x `doop_s`) and is of type int. It contains counts of event occurrence anywhere in a column as a function of the simplified cloud mask and precipitation classifications.

5.3.2 *Occurrence_in_column* variable

The variable *Occurrence_in_column* is dimensioned (`lat` x `lon` x `cmask_s` x `precip_s` x `doop_s`) and is of type float. It contains the frequency of occurrence of events as a function of the simplified cloud mask and precipitation classifications.

It can be used to determine cloud frequency of occurrence anywhere in the column and the frequency of occurrence of precipitation at a given latitude and longitude, without regard to where the event occurred vertically in the column.

For example, at a location with (`lat`, `lon`, `doop_s`) indices given by (i, j, e):

(1) The overall frequency of occurrence of cloud anywhere in the column is given by *Occurrence_in_column* (i,j,1,0,e)

(2) The frequency of occurrence of precipitation is given by *Occurrence_in_column* (i,j,0,2,e), and the frequency of occurrence of snow is given by *Occurrence_in_column* (i,j,0,5,e)

It is noteworthy that the precipitation frequency of occurrence given by *Occurrence_in_column* (i,j,0,a,e) (for any value of a) is equal to that provided by *Occurrence_on_levels* (i,j,k,0,0,a,e), so long as k is large enough that the associated level is above all possible topography. This is because, as previously noted, precipitation is only determined for the surface and is therefore "mirrored" by all values within a column.

5.3.3 *Counts_in_column_by_class* variable

The variable *Counts_in_column_by_class* is dimensioned (`lat` x `lon` x `cclass_s` x `precip_s` x `doop_s`) and is of type int. It contains counts of event occurrence anywhere in a column as a function of the simplified cloud class and precipitation classifications.

5.3.4 *Counts_in_column_total* variable

The variable *Counts_in_column_total* is dimensioned (`lat` x `lon` x `precip_s` x `doop_s`) and is of type int. It contains counts of event occurrence anywhere in a column as a function of the precipitation classification.

The reason this variable is necessary is that multiple cloud classes can occur in one radar profile at the same time.

5.3.5 *Occurrence_in_column_by_class* variable

The variable *Occurrence_in_column_by_class* is dimensioned (`lat` x `lon` x `cmask_s` x `precip_s` x `doop_s`) and is of type float. It contains the frequency of occurrence of events as a function of the simplified cloud class and precipitation classifications.

It can be used to determine cloud frequency of occurrence by class anywhere in the column and the frequency of occurrence of precipitation associated with that cloud class at a given latitude and longitude.

For example, at a location with (`lat`, `lon`, `doop_s`) indices given by (i, j, e):

(1) The frequency of occurrence of cumulus anywhere in the column is given by *Occurrence_in_column_by_class* (i,j,6,0,e)

(2) The frequency of occurrence of cumulus occurring simultaneously with surface precipitation is given by *Occurrence_in_column_by_class* (i,j,6,2,e)

6 Product Versioning and Production

Level 3 products are assigned both an algorithm version and run number.

From the end-user perspective, it will be generally desirable to obtain Level 3 products with the highest available algorithm version (e.g. V0001), as this will correspond to the most recently updated version of the algorithm.

Having found the latest algorithm version, users should obtain Level 3 products with the highest available run number (e.g. U001) within that algorithm version, as this will contain results incorporating the most recently available Level 2 data.

6.1 Algorithm version

The algorithm version, as specified in the file metadata and filename (e.g. V0001) corresponds to exactly one unique combination of all of the following:

- Level 3 product codebase
- External settings associated with said codebase

The above items will be archived for each algorithm version. If any of the above are changed in a way that alters the Level 3 product that is produced from a given set of Level 2 input granules, the algorithm version number will be incremented.

Exception: Products produced using a specific algorithm version at a given time and space resolutions may undergo simple aggregations to produce products at new time and space resolutions, and be considered the same algorithm version.

Note: The CloudSat product revision, e.g. R05, is included in the file metadata and filename for reference. A change in this revision number will require incrementing either the algorithm version or the run number.

6.2 Run number

For a given algorithm version, the run number (e.g. U001) corresponds to a particular set of Level 2 input files. The 2B-GEOPROF granule numbers used, and whether corresponding 2C-PRECIP-COLUMN and 2B-CLDCLASS files were used, is recorded in every Level 3 file. The run number will be incremented in situations where Level 3 products are to be updated (because new Level 2 granules become available, or are recalled, etc.) but the algorithm version is unchanged.

6.3 Requirements on input data availability and uniformity

For a given time period, Level 3 products are only produced if Level 2 2B-GEOPROF input data are available. Product production does not depend on the availability of 2B-CLDCLASS or 2C-PRECIP-COLUMN input data. The presence or absence of such data, however, is contained in the Level 3 variables described in section 3.

To reduce skewing of statistics by missing input data, there is also a requirement that the 2B-GEOPROF input granules are distributed in a sufficiently uniform way throughout the time period. Specifically, the time period is divided into a number of equally sized segments. It is required that a certain fraction of the potential input granules are available in each of these segments. If this fractional availability criterion is not met in any of the segments, no Level 3 product is produced for that time period.

Global attribute *minimum_data_segments* contains the required number of segments, and *minimum_data_fraction* contains the required fractional availability of input data required per segment. Currently, the former is set to three (four for the year-long time period), and the latter is set to 0.65.

7 Global Attributes

The following global attributes are found in all Level 3 files.

Conventions is currently 'CF-1.6' and defines the convention followed by metadata

description contains a short description of the file contents; it includes information on whether the file is Level 3-Full or 3-Simplified

time_period describes the time period of the file in words, such as 'January 2008 through December 2008'

resolution_lon is the longitude resolution of the product

resolution_lat is the latitude resolution of the product

version is the product version, with elements R[CC]_V[DDDD]_U[FFF] meaning the same as they do in the filename specification (Section 2.6)

geoprof_version contains the information about the version of 2B-GEOPROF files used in production, for example '2B-GEOPROF.P1_R05'

precip_column_version contains the information about the version of 2C-PRECIP-COLUMN files used in production, for example '2C-PRECIP-COLUMN.P1_R05'

cldclass_version contains the information about the version of 2B-CLDCLASS files used in production, for example '2B-CLDCLASS.P1_R05'

minimum_data_fraction contains information about data uniformity requirements; see Section 6.3

minimum_data_segments contains information about data uniformity requirements; see Section 6.3

created is the creation date and time of the file

latitude_band indicates the latitude band for Level 3-Full $2.5° \times 2.5°$ files, as described in Section 2.5: 'SO' indicates southern, 'TR' indicates tropical, and 'NO' indicates northern band. Set to 'All' bands for all other files.

8 Changelog

Changes in algorithm version V0002:

– Classification of bins by Daylight-only Operations mode (DO-Op)
– Addition of $2.5° \times 2.5°$ product

References

Haynes, J. M., T. S. L'Ecuyer, G. L. Stephens, S. D. Miller, C. Mitrescu, N. B. Wood, and S. Tanelli, 2009: Rainfall retrieval over the ocean with spaceborne W-band radar. J. Geophys. Res., 114, D00A22, https://doi.org/10.1029/2008JD009973.

Marchand, R., J. Haynes, G. G. Mace, T. Ackerman, and G. Stephens, 2009: A comparison of simulated cloud radar output from the multiscale modeling framework global climate model with CloudSat cloud radar observations. J. Geophys. Res., 114, D00A20, https://doi.org/10.1029/2008JD009790.

Sassen, K., and Z. Wang, 2008: Classifying clouds around the globe with the CloudSat radar: 1-year of results. Geophysical Research Letters, 35, https://doi.org/10.1029/2007GL032591.

A Appendix: Horizontal and Temporal Resolution

As noted in Section 2.4, users are cautioned to consider the statistical significance of aggregations of products of various horizontal and temporal resolutions. To aid in visualization, cloud fraction has been extracted from a sample of Level 3-Simplified products and is plotted here for all available timescales and horizontal resolutions, on two different map projections (Figures 5 and 6).

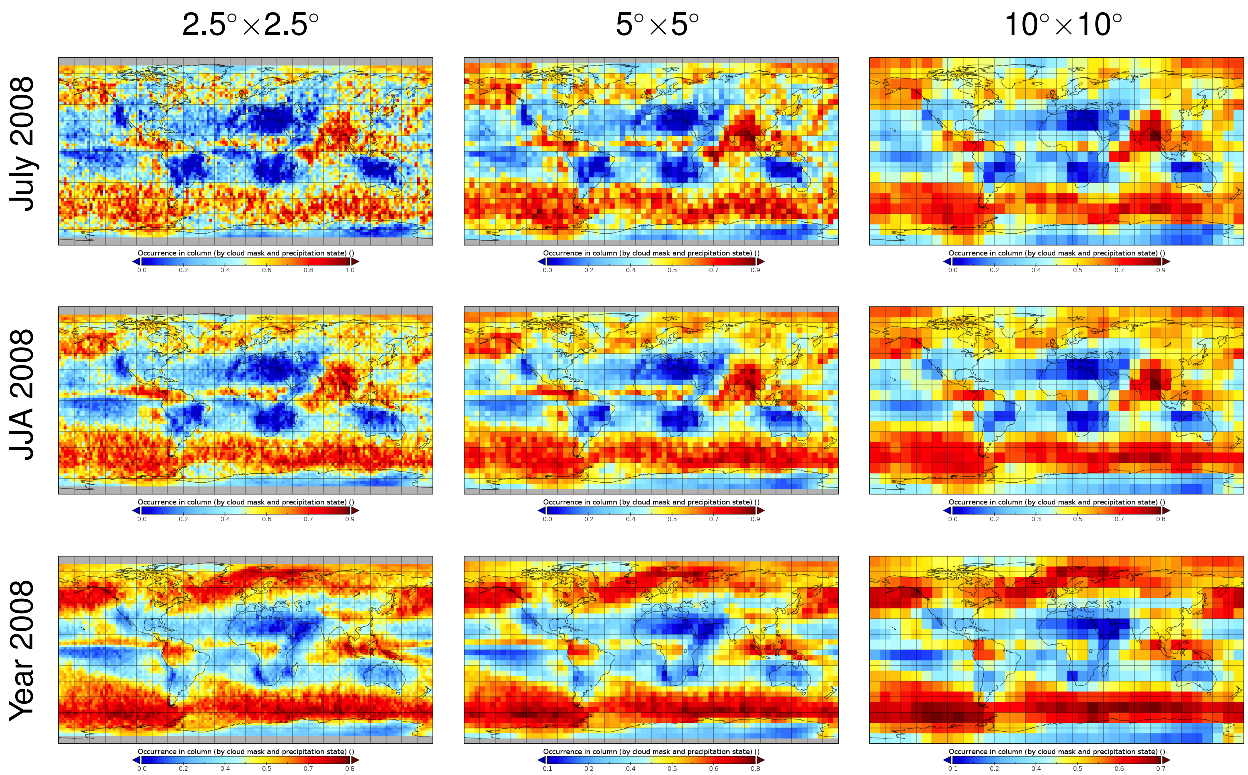

Figure 5: Total cloud fraction at a variety of horizontal and temporal resolutions. Columns from left to right are 2.5°×2.5°, 5°×5°, and 10°×10°. Rows from top to bottom are July 2008, JJA 2008, and year 2008. Data from algorithm version V0002 run U001.

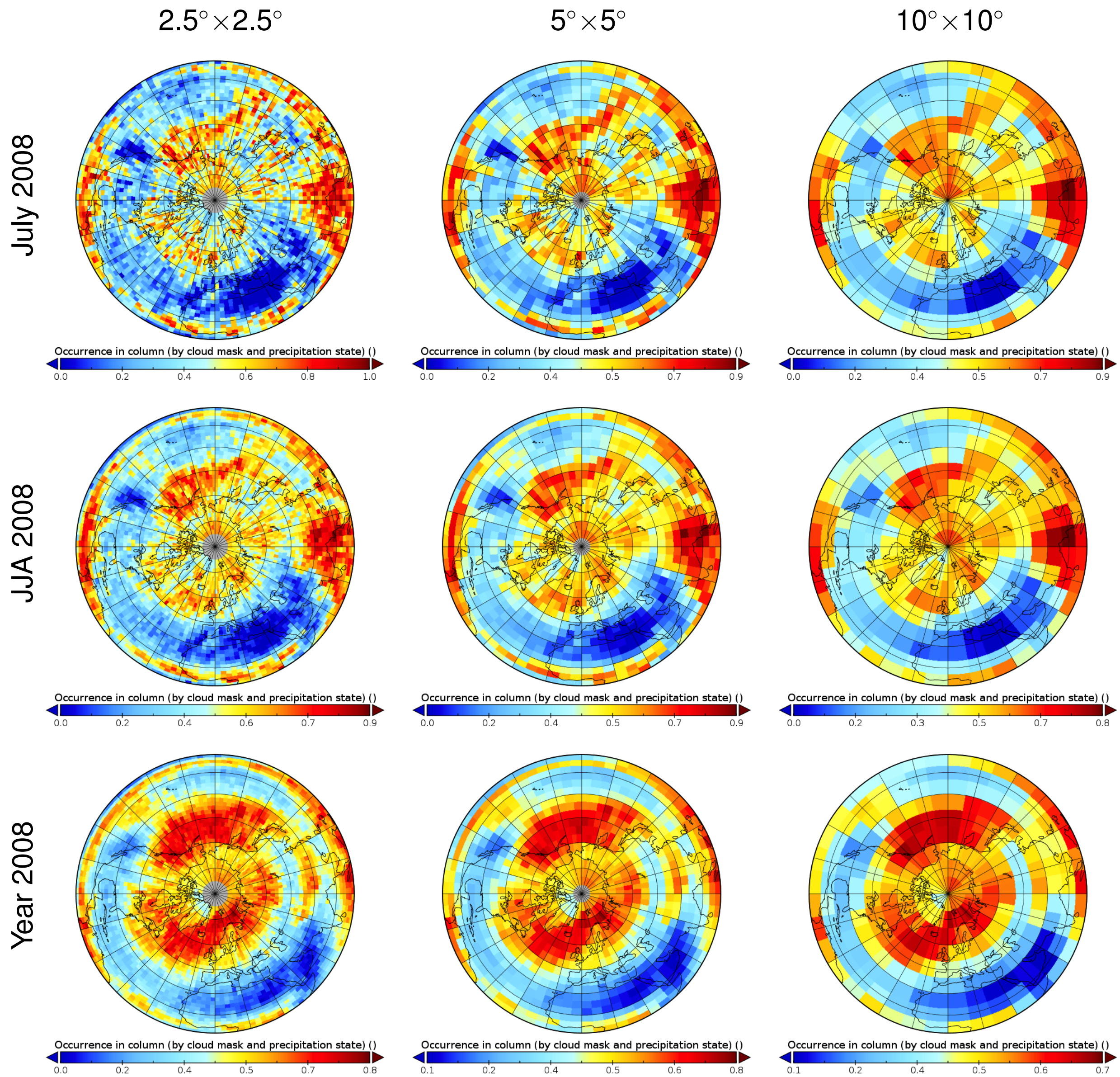

Figure 6: As in Figure 6, but with a map projection centered on the geographic North Pole.