National Institute of Information and Communications Technology

JEM/SMILES

Level 2 Research Data (L2r) Product Guide

for Version 3.0.0

STRUCTURE OF SMILES L2r PRODUCTS

MARCH 20, 2014

Table of Contents

1 LEVEL 2 RESEARCH PRODUCTS OVERVIEW

SMILES NICT Level 2 (a.k.a. L2-research) products are distributed as HDF-EOS files with one file per product (e.g., species) and per day. Details of file naming convention are described below.

(1) FILE NAME

The file name is defined as:

SMILES_L2r_{*product_name*}-{l1b_type}{l1b_version}_{*band_name*}_{l2r_*version_name*}_{*date*}.he5

(2) COMBINATION OF PRODUCT_NAME AND BAND_NAME

Combinations of product name and band name are as follows.

No.	Product_name	Band_name
1	O3 (18 - 100 km) H37Cl (18 - 90 km) BrO (20 - 80 km) HOCl (20 - 80 km) Temperature (15 - 100 km)	A
2	O3(18 - 100 km) H35Cl (18 - 90 km) HOCl (20 - 80 km) Temperature (15 - 100 km)	B
3	ClO (20 - 90 km) BrO (20 - 90 km) HO2 (18 - 100 km) HNO3 (20 - 90 km)	C

Other potential products are: Band A: HNO3, CH3CN; Band B: HO2. Also, wind velocity and O3 isotopomers are retrieved from each Band. For details on these products, please contact the SMILES NICT team.

Email: Yasuko Kasai (ykasai@nict.go.jp)

(3) LEVEL1B_TYPE

l1b_type is represented as follows.

std/rev:

std: Level 1B Standard

rev: Level 1B Revision

(4) LEVEL1B_Version

l1b_version is represented as follows.

xxx:

xxx: Level 1B Version

(5) LEVEL2r_VERSION_NAME

l2r_version_name is represented as follows.

x.x.x:

x.x.x: Level2r Algorithm version

(6) DATE

date refers to UTC times and is represented as follows.

yyyymmdd: (ex. 20091009)

yyyy: Observation year

mm: Observation month

dd: Observation day

2 PRODUCT FORMAT

(1) STRUCTURE OF HDF5-EOS DATA FILES

The format structure of the HDF5*-EOS data file is shown below.

No.	Field	Attributes
1	***File Attribute***	**File Level Attributes**: ・ Instrument name ・ Processing level ・ Version ・ Observation day ・ Band name ・ Scan number ・ L1B file name ・ AOS unit number
2	***Geolocation Field***	**Geolocation Field Attributes:** ・ Observation point ・ Time ・ Altitude ・ Solar zenith angle ・ Ascending/descending flag
3	***Data Field***	**Data Field Attributes:** ・ Data value ・ Error estimation ・ Status

*HTTP://WWW.HDFGROUP.ORG/

(2) STRUCTURE OF STANDARD SMILES L2 PRODUCTS

● Standard processing data (HDF5-EOS)

Structure of standard processing data is as follows.

<File Attributes>

No.	HDF-EOS5 Name	Explanation	Dimension	Data type	Byte
1	***L1BID***	L1B file name	(n times)	long	8*n times
2	***Instrument Name***	Instrument name (SMILES)	-	char	6
3	***Process Level***	Processing level (L2r)	-	char	3
4	***Start UTC***	Start time in this file (yyyy-mm-ddT00:00:00)	-	char	19
5	***End UTC***	End time in this file (yyyy-mm-ddT23:59:59)	-	char	19
6	***Granule Month***	Month (mm)	-	long	8
7	***Granule Day***	Day (dd)	-	long	8
8	***Granule Day of Year***	Granule day of year	-	long	8
9	***Granule Year***	Year (yyyy)	-	long	8
10	***PGE Version***	Processing version (X.X.X)	-	char	5
11	***Start Scan***	First scan number in this file	-	long	8
12	***End Scan***	Last scan number in this file	-	long	8
13	***Band Name***	Band name	-	char	1
14	***AOS Unit Num.***	Unit number of the AOS	(n times)	long	1*n times
			Total		101 + 9*n times

<Swath Attributes>

No.	HDF-EOS5 Name	Explanation	Dimension	Data type	Unit	Byte
1	***Altitude***	Calculation altitude	(n levels)	double	km	8*n levels
2	***Vertical Coordinate***	Vertical coordinate system name	-	char		8
			Total			8 + 8*n levels

<Geolocation/Data fields Attributes>: Following information is added to each field item.

No.	HDF-EOS5 Name	Explanation	Dimension	Data type	Byte
1	***Missing Value***	Missing value	-	double	8
2	***Title***	Filed name	-	char	7
3	***Units***	Unit	-	char	5
4	***Unique Field Definition***	Field definition	-	char	15
5	***_FillValue***	(T.B.D)	-	double	8
			Total		43

<Geolocation Fields>

No.	HDF-EOS5 Name	Explanation	Dimension	Data type	Unit	Byte
1	***Time***	Observation time (Total seconds since 1/1/1958)	(n times)	double	*1	8*n times
2	***TimeUTC***	Observation time (UTC) yyyy-mm-dd hh:mm:ss.sss	(n times)	char	-	19*n times
3	***Altitude***	Representative altitude	(n level)	double	km	8*n level
4	***Latitude***	Observation latitude	(n times)	double	degrees	8*n times
5	***Longitude***	Observation longitude	(n times)	double	degrees	8*n times
6	***Solar Zenith Angle***	Solar zenith angle	(n times)	double	degrees	8*n times
7	***Local Time***	Local time (hh:mm:ss)	(n times)	char	-	8*n times
8	***AscendingDescending***	Ascending/descending flag	(n times)	char	-	1*n times
9	***Altitude_tp***	Altitude grid for auxiliary temperature and pressure data	(n level_tp)	double	km	8*n level_tp
			Total	60*n times + 8*n level + 8*n level_tp		

*1: seconds since 1/1/1958

<Data Fields>

No.	HDF-EOS5 Name	Explanation	Dimension	Data type	Unit	Byte
1	***L2 Value***	Value	(n level, n times)	double	vmr	8*n times *n level
2	***L2 Precision***	Calculation error	(n level, n times)	double	vmr	8*n times *n level
3	***Vertical Resolution***	Vertical resolution	(n level, n times)	double	km	8*n times *n level
4	***Measurement Response***	Measurement response	(n level, n times)	double	-	8*n times *n level
5	***Pressure_tp***	Auxiliary pressure data	(n level_tp, n times)	double	hPa	8*n times *n level_tp
6	***FOV Interference***	Flag of field-of-view interference on the L1b data	(n times)	long	-	8*n times
7	***Status***	Status information	(n times)	long	-	8*n times
8	***Temperature_tp***	Auxiliary temperature data	(n level_tp, n times)	double	K	8*n times *n level_tp
9	***cloud_flag***	Cloud flag information	(n times, n cloud flags) n cloud flags = 64	long	-	8*n times* 64
10	***chi2***	Chi2	(n times)	double	-	8*n times
		Total	32*n times*n level + 16*n times*n level_tp + 24*n times			

<StructMetadata>

No.	HDF-EOS5 Name	Explanation	Dimension	Byte
1	***StructMetadata.0***	Matrix information of swath data	1	32000

< coremetadata >

No.	HDF-EOS5 Name	Explanation	Dimension	Byte
1	***coremetadata.0***	HDF-EOS information	1	About 6960

3 JEM/SMILES L2r data guide for release version 3.0.0

3.1 SMILES L2r version 3.0.0

Compared to the previous version (2.1.5), this version 3.0.0 of NICT data processing has several modifications in the retrieval procedures. The concept of utilizing definitive spectral regions (windows) for each species has been kept, however, now the subsequence of the data processing (such as using the tangent heights retrieved from the O_3 spectra in retrieving HCl) is not applied. This is because we found that the tangent height retrieval from O_3 spectra is very sensitive to the accuracy of non-linear gain calibration and using the retrieved tangent heights for other species' retrieval introduces unwilling systematic errors. See the paper by Kasai et al. (2013) for detail.

Updated from version 2.1.5 are as follows:

- Version of the L1b data was updated to **008** from 007. The major impact of using L1b-008 is the availability of SMILES star-tracker tangent height information improved by the auxiliary data from the MAXI instrument (Ochiai et al., 2012). This MAXI + Star tracker tangent height information is used in the 3.0.0 processing instead of using the attitude information from ISS.
- AOS response function was changed to a triple-Gaussian model based on the improved analysis of on-orbit comb measurements by Mizobuchi et al. (2013).
- Deviation on the frequency for each AOS channel was improved by applying a spline interpolation instead of 7-order polynomial correction derived by the SMILES instrumental team.
- A priori temperature profile is changed as same as the ones used in the JAXA/ISAS L2 processing. The difference between the one used in NICT 2.1.5 processing is found only above 40 km. Hydrostatic equilibrium is assumed when retrieving the temperature from Ozone spectra.
- With respect to the temperature profile, the important point is that, when retrieving the species except Ozone, the temperature profile is fixed to the a priori values instead of using the one retrieved from Ozone line. This is in order to avoid the systematic error propagation originated from the correlation between Ozone abundance and temperature (around 50 – 60 km).

- References for the L1b calibration:

  Ochiai et al. (2012): Calibration of superconducting submillimeter-wave limb-emission sounder (SMILES) on the ISS, *Proc. SPIE*, **8527**.

  Mizobuchi et al. (2012): In-orbit measurement of the AOS (acousto-optical spectrometer) response using frequency comb signals, *IEEE J. Sel. Top. Appl.*, **5**, pp. 977 – 983.

3.2 Overview of the retrieval processing (L2r version-3)

In the previous 2.1.5 processing, we applied "sequential" retrieval method. That is, before the actual retrieval, we first went through a pre-processing procedure, which includes a cloud flagging and an initial estimation of the LOS elevation angle offset resulting from uncertainties in the measured instrument attitude. Then, we extracted the limited spectral range around the 625.371 GHz O_3 line (from band A or B, whichever is present) and retrieve the LOS elevation angle correction **for each tangential point** (along with an initial retrieval of $O_3(z)$ and $T(z)$) as described by Baron et al. (2011).

In the current 3.0.0 processing, we keep the pre-processing procedure and the concept of utilizing the limited spectral range for each species. However, we do not "pass" the retrieved results (tangent heights, $O_3(z)$, $T(z)$, etc) between different retrieval steps in order to avoid systematic error propagation.

3.2.1 Pre-processing (cloud flagging)

Cloud flagging and elevation angle offset estimation is done in an iterative process. Cloud flags are derived for each observation from the comparison of the observed signal with the signal simulated assuming clear-sky conditions, where the flag is set when the signal difference exceeds a certain threshold.

Mean signals from selected spectral regions in each band, the cloud flagging windows (CF-A, B, C), are used. The CFs have been chosen in regions with minimum gaseous absorption as far as possible avoiding contamination from absorption lines, particularly in lower stratosphere and troposphere spectra, in order to minimize the effect of variable trace gas concentration and maximize the cloud sensitivity. Hence, major contributors to absorption in the CF are dry air and water vapor continua and (far-)wing contributions of further away strong (557GHz H_2O, 625.4GHz O_3) and weaker, but closer absorption lines (e.g., band A and B HCl lines, band C ClO). We analyzed the impact of the different species. For each SMILES band we derived a set of species whose absorption contribution in the CF is significant (see Table 1) and which therefore are considered when simulating the clear-sky comparison spectra.

Profile information of the absorption species to consider as well as of temperature are taken from best available estimates. That is, temperature water vapor, and ozone profiles are taken from GEOS5 reanalysis data close in time and location to the actual observation (for details of GEOS5 data selection see Baron et al. (2011)). Other species are taken from climatology data

Table 1: Cloud flagging setups

Band	A	B	C
CF position [GHz]	624.5464 – 624.6764	26.1691 – 626.2991	649.5552 – 649.6852
Trace gases	H_2O, O_3, HCl-37, CH_3CN, HNO_3	H_2O, O_3, HCl-35, N_2O	H_2O, O_3, ClO

The clear-sky spectra simulation is done on the exact tangent altitude grid and with the antenna pattern provided by the L1b data. However, attitude data is known to be not exact enough for a brightness temperature (Tb) difference based cloud flagging, particular in the middle to lower stratosphere, where the Tb change with altitude is very steep (order of 10K/km), and where therefore small deviations in the tangent altitudes between observation and simulation can lead to large differences of the observed and simulated signal. This can easily cause false cloud flaggings in this region. In order to avoid this, and elevation angle offset per scan is estimated. Since the relation between altitude and Tb is not linear, elevation angle estimate and cloud flagging are iterated until convergence. The procedure is as follows

1. Simulate clear-sky spectra on fine altitude grid and average over cloud flagging window
2. Antenna pattern convolution on current estimate of tangent altitude grid.
3. Least square estimate of elevation angle offset (assuming step-wise linear Tb-TanH relation) from non-cloudy flagged observations in scan only.
4. Repeat 2.-3. until convergence, i.e., offset correction <6e3 degrees (necessary since Tb-TanH relation is non-linear).
5. For each tangent altitude in scan compare observed and simulated signal and flag cloudy observations (for flagging details see further below).
6. Compare new flagging profile with previous one. Repeat 2.-6. until no changes in cloud flag profile.

The flagging decision itself is based on the fact that clouds significantly modify the signal, for sufficiently thick clouds by several to tens of Kelvin. For lower stratosphere and upper troposphere tangent altitudes, where the atmospheric opacity is low, clouds increase the observed signal. For lower tangent altitudes, clouds lead to Tb depression compared to the clear-sky case. The exact location of the transition region is variable and depends on the actual opacity of the atmosphere, which is mostly determined by the water vapor content. Hence, we do not attempt to estimate it, but base the cloud flagging decision on the absolute Tb difference except for low tangent altitudes, where only Tb depressions physically make sense. Explicitly, the cloud flagging criteria are:

- for TanH > 10km flag as cloud if |Tb_obs–Tb_sim| > X, where X is the maximum of
    - 5K
    - 3 times the standard deviation of Tb_obs–Tb_sim over the whole scan (this criterion shall avoid excessive cloud flagging in case elevation offset estimate is not good enough yet and in case of bad observational data)
- for TanH < 10km flag as cloud if Tb_obs–Tb_sim > 5K (i.e., only Tb depressions >5K)

Cloud flagging and elevation offset estimation is done separately for the two observed channels. Cloud flags are kept per individual band throughout (i.e., the two observed

bands might exhibit deviating cloud flags). Experience showed that band B elevation estimates are less good than the other two, hence in a last step the band B offset is reset to the band A or C value (whichever is present as second band). For simultaneous observation of band A and C, each band keeps its own offset estimate.
The elevation offset estimates from the pre-processing are used in the following sequential retrievals as initial guess for the lineshape based elevation retrieval therein.

Cloud flags are contained in each of the product hdf-files. They are delivered for the respective product band. In the hdf files they are stored as two-dimensional array of size (number of scans per day) x (64), where 64 is the number of observations per scan (if the actual number deviates, the remainder of the array is filled with dummy values). That is, the cloud flags correspond to individual observations and to the observational tangent altitude grid, not to the altitude grid of the retrieved product!

Values in the cloud flag fields mean 0: clear
1: cloudy

3.2.2 Spectral range configurations

Spectral range used in each species' retrieval is kept as same as those listed in Baron et al. (2013).

3.3 Some notes on the L2r-hdf contents

SMILES L1b data contains attitude information from several sources, particularly from the ISS itself and from the Star Tracker (STT) onboard SMILES. The SMILES/ISS attitude data on L2r processing are based on the "STT" origin information of L1b with an improvement from MAXI data. Latitude and longitude of L2r hdf data are derived as position of the tangent point of the 30th (from bottom tangent height) spectrum of each scan (i.e., the observation in the middle of the scan). These configurations cause some differences to L2-operational data in latitude and longitude information for the same scan. This discrepancy will be solved in a future L2r product release.
The "FOV Interference" field is prepared with the intention of providing quality flag information on L1b data. For the current version of the L2r-hdf, this FOV Interference field intertwines quality flag information from the level-2 procedure. This confusing situation will be improved in the next version of the L2r product.

For version 3.0.0, FOV Interference means:

- 0: Normal L1b data and successful level-2 process
- 1: Normal L1b data, although the level-2 process had some problems.
- 2 - 15: L1b data contains warning flag on the data quality.
- 16 - 31: Warning in the event that the antenna line of sight (LOS) has an object in it.
- 32 - 63: Cold-sky observation has an object within its LOS.
- >= 64: Star tracker LOS has an object in it.

If the flag number is larger than 1000, this means that the temperature a priori data is from the one used for 2.1.5 processing. Such scan is better to be discarded from the statistical analyses.

We recommend using the ***measurement response*** and ***chi2*** as measures of the product quality and use those for the data selection of the L2r v3.0.0 product.
Please contact to the SMILES L2r team for further instruction.

3.4 Miscellaneous

Relevant reference papers for data processing:

- Baron, P., et al. (2011): The Level 2 research product algorithms for the Superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES), *Atmos. Meas. Tech.* , **4**, pp. 2105 – 2124.

- Baron, P., et al. (2008): AMATERASU: Model for Atmospheric TeraHertz Radiation Analysis and Simulation, *Journal of the National Institute of Information and Communications Technology*, **55(1)**, pp. 109 – 121. http://www.nict.go.jp/publication/shuppan/kihou-journal/journal-vol55no1/07-04.pdf

- Sato, T. O., et al. (2012): Strato-mesospheric ClO observations by SMILES: error analysis and diurnal variation), *Atmos. Meas. Tech.*, **5**, pp. 2809 – 2825.

- Kasai, Y., et al. (2013): Validation of stratospheric and mesospheric ozone observed by SMILES from International Space Station, *Atmos. Meas. Tech.*, **6**, pp. 2311 – 2338.

- Sagawa, H., et al. (2013): Comparison of SMILES ClO profiles with satellite, balloon-borne and ground-based measurements, *Atmos. Meas. Tech.*, **6**, pp. 3325 – 3347.

- Baron, P., et al. (2013): Observation of horizontal winds in the middle-atmosphere between 30° S and 55° N during the northern winter 2009–2010, *Atmos. Chem. Phys.*, **13**, pp. 6049 – 6064.

- Sato, T. O., et al. (2013): Vertical profile of δ18OOO from middle stratosphere to lower mesosphere derived by retrieval algorithm developed for SMILES spectra, *Atmos. Meas. Tech.*, in press.

For other scientific results using SMILES NICT data, please check our publication list on http://smiles.nict.go.jp/pub/publication/publications_all.html