Document No.	NIES-GOSAT-PO-014
Version	2.0
date of issue	Jun. 2017

ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD) FOR CO_2, CH_4 and H_2O COLUMUN AMOUNTS RETRIEVAL FROM GOSAT TANSO-FTS SWIR

NIES GOSAT Project

Authors: Yukio Yoshida[1)], Nawo Eguchi[1)*], Yoshifumi Ota[1)**], Nobuhiro Kikuchi[2)***], Koji Nobuta[2)], Tadao Aoki[1)], and Tatsuya Yokota[1)]

1) Center for Global Environmental Research, National Institute for Environmental Studies
2) Fujitsu FIP Corporation
* Currently at: Research Institute for Applied Mechanics, Kyushu University
** Currently at: Meteorological Research Institute
*** Currently at: Earth Observation Research Center, Japan Aerospace Exploration Agency

Inspected: A. Kamei

Approved: Tsuneo Matsunaga

Revision history

Version	Revised on	Description	Remarks
1.0	May 2010	－	For SWIR L2 V00.50, V00.80, and V00.90.
1.1	Sep. 27, 2010	Change SWIR L2 Product Version name. Revise sequential numbers of sections, equations, figures, and tables due to add/remove sections. Add section 2.4 Degradation Correction of TANSO-FTS. Revise section 3.1 L1 Quality Evaluation. Revise section 3.7 2-micron Band Scattering Material Determination. Revise section 4.3 Gas Absorption Coefficient. Revise section 6.2 Instrument Line Shape Function. Revise section 6.3 Vertical Grid. Remove section 6.4 Optimum Vertical Grid. Revise section 6.5 A Priori Information of Relative Radiometric Calibration Adjustment Factor. Add section 6.10 A Priori Information of Surface Pressure. Add section 6.11 A Priori Information of Temperature Shift. Add section 6.12 A Priori Information of Wavenumber Dispersion. Revise chapter 7 Atmospheric Column Amount Estimation Algorithm. Revise section 8.1 Channel Selection Algorithm. Revise section 9.2 Quality Check Process. Revise References	For SWIR L2 V01.10, V01.20, and V01.30.
2.0	Jun. 06, 2017	Change SWIR L2 Product Version name. Revise sequential numbers of sections, figures, and tables due to add/remove descriptions. Revise section 2.1 Observation Time Correction Revise section 2.2 Observation Position Correction. Revise section 2.4 Degradation Correction. Revise section 3.1 L1 Quality Evaluation Revise section 3.4 CAI Cloud Determination Revise section 3.6 TIR Cloud Determination. Revise section 3.7 2-micron Band Scattering Material Determination. Add section 3.9 Spectrum Quality Determination. Revise section 3.10 Combined Use of Data Screening Criteria. Revise section 4.5 Aerosol. Revise section 4.7 Solar Irradiance Spectra. Revise section 5.2 LM Method Revise section 6.2 Instrument Line Shape Function. Revise section 6.3 Vertical Grid. Revise section 6.4 A Priori Information of Gas	For SWIR L2 V02.00, V02.10, V02.11, V02.20, V02.21, V02.31, V02.40, V02.50, and V02.60.

		Concentration Revise section 6.5 A Priori Information of Relative Radiometric Calibration Adjustment Factor. Revise section 6.9 A Priori Information of Aerosol. Revise section 6.10 A Priori Information of Surface Pressure. Revise section 6.11 A Priori Information of Temperature Shift. Add section 6.13 A Priori Information of Zero-Level Offset. Revise section 7.1 Retrieval of Carbon Dioxide (CO_2), Methane (CH_4) and Water Vapor (H_2O) Column Amounts. Revise section 8.1 Channel Selection Algorithm. Revise section 9.2 Quality Check Process. Revise References. Add Appendix 2. Add Appendix 3. Correct typos.	

Table of Contents

1. Introduction

1.1 Objectives and Scope of This Document

The objective of this document is to derive and describe the algorithms by which global atmospheric column amounts of carbon dioxide (CO_2), methane (CH_4), and water vapor (H_2O) are retrieved from the Short-Wavelength Infrared (SWIR) spectrum data obtained by the Thermal and Near infrared Sensor for Carbon Observation (TANSO) - Fourier Transform Spectrometer (FTS) onboard the Greenhouse Gases Observing Satellite (GOSAT). The retrieval process accomplishes the creation of Level 2 (L2) Products based on the TANSO-FTS Level 1B (L1B) Products.

The scope of this document is limited to the algorithms for the TANSO-FTS SWIR L2 (V02.00, V02.10, V02.11, V02.20, V02.21, V02.31, V02.40, V02.50, and V02.60; hereafter generically called V02.xx; see Appendix 3 for more detail) Products. The algorithm of Level 1 Products that include TANSO-FTS L1B and TANSO-CAI L1B Products, and those for TANSO-CAI L2 Product as well as the higher level processing, such as Level 3 and Level 4 Products, will be provided in separate documentations.

1.2 Overview of Instruments

The TANSO onboard GOSAT actually consists of two instruments as shown in Table 1.2-1. One is the TANSO-FTS that observes the greenhouse gases; and the other is the TANSO-CAI that senses clouds and aerosols. The TANSO-FTS measures the amounts of greenhouse gases by using their spectrum absorption bands. The effective optical path-length varies according to the presences of clouds and/or aerosols in the field of view (FOV; the diameter of approximately 10.5 km) of the TANSO-FTS, thereby adversely affecting the atmospheric column amount determination. In order to cope with this difficulty, the TANSO-CAI monitors the clouds and aerosols in the TANSO-FTS's FOV. Furthermore, a monitor camera (CAM) is installed on the same optical bench together with the TANSO-FTS, which facilitates the accurate determination of the relative lines of sight (LOS's) of the TANSO-FTS and the TANSO-CAI. *Kuze et al.* [2009, 2012, 2016] gives more detailed information.

Table 1.2-1: Instruments onboard GOSAT and their roles.

Instruments	Active Bands	Purpose
TANSO-FTS	SWIR; Three bands (with two polarization each)	Column amount determination of CO_2, CH_4, and H_2O
	TIR; Single band	Profile determinations of CO_2, CH_4, temperature and H_2O
CAM	Visible; Single band	Determination of relative LOS's of TANSO-FTS and TANSO-CAI
TANSO-CAI	Ultra-violet; Single band Visible; Two bands SWIR; Single band	Measurements of cloud and aerosol distributions

(a) TANSO-FTS

The TANSO-FTS measures two kinds of radiance spectra. One is the SWIR radiance of the sun's light, which has traveled through the earth's atmosphere and is reflected by the earth's surface, clouds and aerosols. The other is the thermal radiance of the earth's atmosphere. From these observed spectra, the greenhouse gas concentrations are retrieved by means of absorption band spectrum distributions and structures. The specifications of the TANSO-FTS instrument are summarized in Table 1.2-2.

Table 1.2-2: Specifications of the TANSO-FTS.

Pointing mirror		2-axis pointing mirror (Completely redundant)			
Scan		Cross track (± 35 deg.), Along track (± 20 deg.)			
FOV		IFOV < 10.5 km, 790 km (Scan width)			
Observation time		4.0, 2.0, 1.1 sec / Interferogram			
Band (Polarization)		1 (P, S)	2 (P, S)	3 (P, S)	4 (N/A)
Bandwidth	μm	0.758 ~ 0.775	1.56 ~ 1.72	1.92 ~ 2.08	5.56 ~ 14.3
	cm^{-1}	12900 ~ 13200	5800 ~ 6400	4800 ~ 5200	700 ~ 1800
Instrument line shape function's half width		< 0.6 cm^{-1}	< 0.27 cm^{-1}	< 0.27 cm^{-1}	< 0.27 cm^{-1}
Signal to noise ratio(1)		> 300	> 300	> 300	> 300
Detector		Si	InGaAs	InGaAs	PC-MCT
On-orbit calibration		Sensitivity: Solar diffuser board, Deep space, Moon Instrument Line Shape function (for Band 2): Diode laser			Sensitivity: Black body, Deep space

Note: (1) Defined for the value when surface albedo= 0.3, solar zenith angle= 30 degrees.

A Fourier Transform Spectrometer (FTS) is adopted as a GOSAT spectrometer to observe the wide spectrum range of light from short-wavelength infrared to thermal infrared. In the FTS, first the incident light is split into transverse and reflective beams. Then the two split-beams are reflected back (by two individual mirrors) to the beam splitter, where an interference phenomenon occurs. Finally a detector measures the intensity of that interference. It is well known that a light phase difference varies as a function of the optical path differences (OPDs) defined by the distances of the beam splitter and the two separate reflection mirrors. Thus, for a given constant OPD, the in-phase lights of an identical wavelength increase in intensity, whereas the out-of-phase lights decrease in intensity. By continuously changing the OPD, it is possible to measure the intensity of interference. The relationship between the OPD and the intensity of interference is defined as an interferogram (TANSO-FTS L1A Product), and its Fourier transformation yields the spectra of the incident light (TANSO-FTS L1B Product). In the FTS, three dichroic filters separate the incident light into four pre-determined observation wavenumber ranges. For this purpose, zinc selenide (ZnSe) has been selected as the material for the beam splitter in order to be effective for both short-wavelength infrared and thermal infrared. However, zinc selenide is also known to have slightly nonlinear polarization characteristics. Namely, incident lights of an identical total intensity can produce different results if its polarizations are different. While direct sunlight and thermal radiance are considered to be non-polarizing, reflected lights from the earth's surface and scattering lights from clouds and aerosols are considered to be polarizing. Therefore, two components (i.e., P- and S-polarization states; refer to Table 1.2-2) of polarization are separately measured, particularly for the three bands of short-wavelength infrared, which is susceptible to scattering.

(b) Pointing Mirrors

The pointing mirrors are designed to facilitate observation of an arbitrary point of the earth's surface, and are mounted on two-orthogonal axes. However, in order to eliminate the sun's direct and stray lights, a hood (also known as a baffle) is attached, which effectively restricts the instrument LOS to an observation point to within ±20 degrees in Along Track (AT) and ±35 degrees in Cross Track (CT).

Another important role of the pointing mirrors is to enable a staring observation of the earth's surface. Typically, an acquisition of one interferogram takes four seconds, during which GOSAT traverses about 28 km. Hence it becomes necessary to activate steering mirrors to stare at an observation point. This functionality is called Image Motion Compensation (IMC).

(c) CAM

Some of the TANSO-FTS input light is diverted into a monitoring camera (CAM), which in turn produces observation images and facilitates the LOS alignment of the TANSO-FTS and the TANSO-CAI. The effective horizontal resolution of the CAM is

about 100 m, and its Instantaneous Field of View (IFOV) is about 23 km x 30 km at the intersection of the CAM's LOS and the earth's surface. This IFOV is three times larger than that of the TANSO-FTS.

The CAM is solidly mounted on the same optical bench as the TASNO-FTS, and hence their relative LOS's are fixed. This means that their relative LOS's are invariant for pre- and post-launch, whereas the TANSO-CAI and the TANSO-FTS are mounted on different optical benches, implying that their pre- and post-launch (relative) LOS's may vary. Furthermore, the relative LOS's may also change as a function of satellite internal temperature. Both of these causes could contribute to an error in the detection of clouds as well as aerosols in the FTS's FOV. Therefore, a precise alignment of images of the TANSO-CAI and the CAM (and eventually the TANSO-FTS and the TANSO-CAI) is indispensable for minimizing this kind of errors.

Since the primary purpose of the CAM is the geometric calibration of the TANSO-FTS by JAXA, the original CAM's image data is not distributed to the public. In this document, it is assumed that the relative LOS's of the TANSO-FTS and the TANSO-CAI during flight are known by means of an in-flight calibration etc.

NOTE: After the TANSO-FTS L1B V200, the decimated CAM's image data are stored in the TANSO-FTS L1B product.

(d) TANSO-CAI

While the primary function of the TANSO-CAI is the detection of clouds and/or aerosols in the FOV of TANSO-FTS, its data is also utilized to estimate the optical properties of aerosol like aerosol optical thickness. The radiometric and geometric calibrated radiance data (TANSO-CAI L1B Product) and the integrated clear-sky confidence level for each CAI pixel (TANSO-CAI L2 Product) are used in the TANSO-FTS SWIR L2 processing.

Table 1.2-3: Specifications of the TANSO-CAI.

Band	**1**	**2**	**3**	**4**
Band widths [μm]	0.370 ~ 0.390	0.664 ~ 0.684	0.860 ~ 0.880	1.555 ~ 1.645
SNR	> 200	> 200	> 200	> 200
Observation swath	1000 km	1000 km	1000 km	750 km
Horizontal resolution	0.5 km	0.5 km	0.5 km	1.5 km

Table 1.2-3 summarizes the specifications of the TANSO-CAI. As is clear from this table, the TANSO-CAI consists of four bands whose center wavelengths are 380, 670, 870 and 1600 nm. The corresponding FOV's and observation swaths are 0.5 km and 1000 km, respectively, for the first three bands, and 1.5 km and 750 km, respectively, for the last band. Therefore, except for the band with 1600 nm at the center

wavelength, the TANSO-CAI covers the TANSO-FTS's scan width of 790 km, and includes about 500 pixels in its FOV of about 10.5 km in diameter.

Cloud detection is done through a combination of tests at each pixel with a pre-determined threshold value [*Ishida and Nakajima*, 2009; *Ishida et al.*, 2011]. Since the individual pixel size of the TANSO-CAI is sufficiently smaller than the TANSO-FTS's FOV, accurate cloud amount determinations are feasible. Furthermore, the optical thickness and the effective radius of cloud particulates for those pixels, which have been judged to contain clouds, are retrieved from the visible and near-infrared radiance observed with the TANSO-CAI (TANSO-CAI L2 Cloud Property Product (Research)) [*Nakajima et al.*, 2008].

For those pixels that have been determined not to contain clouds, the optical thicknesses of specific kinds of aerosols are retrieved from the all observed radiances with TANSO-CAI (TANSO-CAI L2 Aerosol Property Product (Research)). Particularly, 380 nm radiance is quite effective for identifying absorptive aerosols, and detecting aerosols over land. The optical properties of aerosol should be used for TANSO-FTS's data analyses. Furthermore, the assimilation analysis of global aerosol distributions by a three-dimensional aerosol transport model SPRINTARS (Spectral Radiation-Transport Model for Aerosol Species) [*Takemura et al.*, 2000] is planning to utilize the retrieved aerosol optical properties in conjunction with the ground-based observations of AERONET (AErosol Robotic NETwork) [*Holben et al.*, 1998] and SKYNET (Skyradiometer network) [*Takamura et al.*, 2004].

1.3 Overview of TANSO-FTS SWIR L2 Algorithms

The TANSO-FTS SWIR L2 gas concentration retrieval process consists of (1) the pre-process, (2) the data screening process, (3) the retrieval process, and (4) the quality check process. In the pre-process, the identification of a particular observation point and time as well as the necessary spatiotemporal interpolation should be performed. In the data screening process, a suitable event selection should be made as a precursor to the retrieval process. In the retrieval process, the atmospheric column amounts of CO_2, CH_4 and H_2O should be retrieved, followed by a quality check process to evaluate the quality of entire data processes. Only the data that passes the quality check should be distributed to the public as a final product.

These processes are described below.

2. Pre-processing

TANSO-FTS SWIR L2 retrieval processing requires pre-processing, which includes various corrections of TANSO-FTS L1B's radiance spectrum data products.

The observation time of the TANSO-FTS L1B data products is defined as the start of the interferogram acquisition, whereas the preferred time in the TANSO-FTS SWIR L2 retrieval process is the Zero Path Difference (ZPD) passage time. Hence a temporal correction becomes necessary as described in Section 2.1, entitled, Observation Time Correction. The temporal correction is naturally followed by a spatial correction, as explained in Section 2.2, entitled, Observation Position Correction. Then, the necessary meteorological data should be computed through a spatiotemporal interpolation of the Grid Point Values (GPV), as defined in Section 2.3, entitled, Preparation of Meteorological Data. Finally, the degradation correction, the wavenumber correction and the polarization synthesis process, outlined in Sections 2.4, 2.5 and 2.6, respectively, should be performed.

2.1 Observation Time Correction

The interferogram ZPD passage time should be used as the observation time in TANSO-FTS SWIR L2 retrieval processing. Although the timing of ZPD passages for each of the seven TANSO-FTS bands is different, the maximum difference amounts to only 0.02 sec, which corresponds to 0.14 km (= 7 km/sec x 0.02 sec) in the ground track. Hence the ZPD passage time of the Thermal Infrared (TIR) data should be used as the representative time for all TANSO-FTS bands. The reason for this choice lies in that the TIR data applies to both sun-illuminated and sun-shaded areas, whereas the SWIR data applies only to the sun-illuminated area. Thus by using the TIR data, data consistency can be preserved.

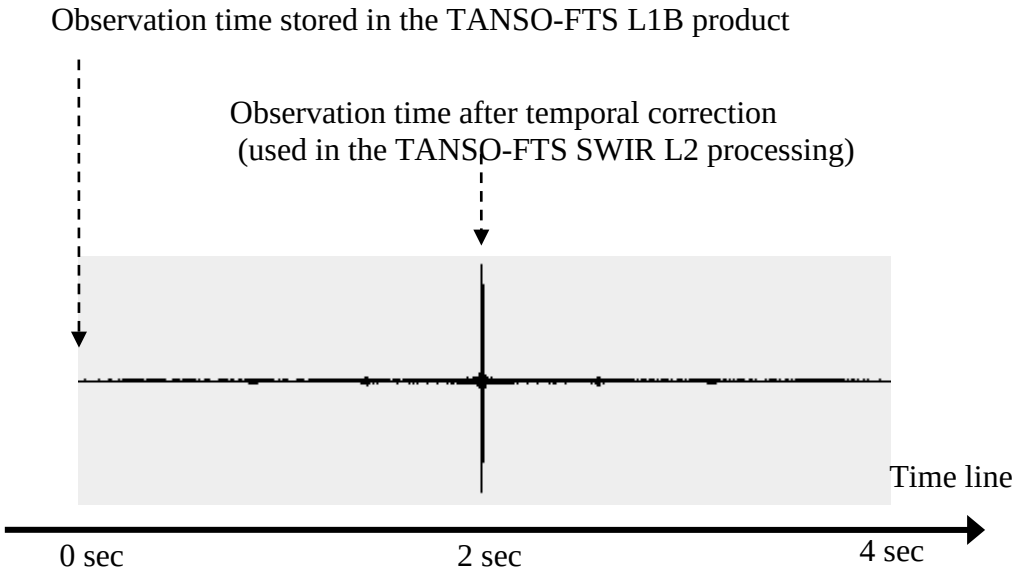

Figure 2.1-1: The conceptual diagram of observation time correction.

The FTS observation time for TANSO-FTS SWIR L2 algorithm is defined by the interferogram acquisition start time(1), which is stored in the TANSO-FTS L1B products, added by the TIR data ZPD passage time(2) (relative time in seconds from the interferogram start time).

Note:

(1) /exposureAttribute/pointAttribute/Time

(2) /exposureAttribute/pointAttribute/RadiometricCorrectionInfo/ZPDPassTime_TIR

Note (V02.50):

Due to the cryocooler shutdown, no TIR data was available for TANSO-FTS SWIR L2 V02.50 period [*Kuze et al.*, 2016]. Therefore, an average of the ZPD passage time for the six SWIR bands is used as the representative time.

2.2 Observation Position Correction

The observation position information stored in the TANSO-FTS L1B product refers to the LOS of the TANSO-FTS's FOV at the start of the interferogram acquisition. It represents the intersection of the TANSO-FTS's LOS with an ellipsoidal surface of the earth (WGS84). However, a position correction in the TANSO-FTS SWIR L2 algorithm should refer to the observation point (latitude, longitude, and height) at the ZPD passage time of the interferogram. A pointing anomaly in the TANSO-FTS pointing mechanism [*Kuze et al.*, 2012] is also corrected after the TANSO-FTS SWIR L2 V02.xx processing.

The spatial correction of observation points is performed in three steps. First, the pointing anomaly is corrected. Second, a position correction due to GOSAT satellite motion should be performed. Third, a correction for the elevation effect (i.e., Ortho-correction) of observation points should also be performed.

Two kinds of observation points should be considered for correction; (i) one is a single observation point corresponding to the center of the FOV referred to as the LOS or simply the bore-sight of FTS and (ii) the other is actually a collection of observation points that lie inside the FOV perimeter, referred to as the FOV-area. The single observation point corresponding to the center of the FOV should typically be utilized when performing data interpolation of a reference database, which has a coarse resolution compared with the size of FTS's FOV. On the other hand, the observation points within the FOV-area should typically be utilized when statistical quantities of a reference database, whose mesh-size is enough smaller than the TANSO-FTS's FOV, within the FTS's FOV are sought. The examples of the former reference database is the GPV data and those of the latter ones are the TANSO-CAI data products and the digital elevation model (DEM) data.

In the spatial correction, the operational state of the onboard Image Motion Compensator (IMC), which tracks the observation points on the earth's surface, should

be taken into consideration. The operational state of the IMC (i.e., whether it is on or off) becomes an important piece of information in the TANSO-FTS SWIR L2 algorithm. Usually the TANSO-FTS is operated with the IMC in the "on" state. The TANSO-FTS operation with the IMC in the "off" state is limited to sunglint-mode observations over the ocean until the end of July 2010. A list, which contains TANSO-FTS L1B scene with the IMC off, is prepared to identify the IMC operational state.

(a) Correction of the pointing anomaly

There is a known pointing anomaly in the TANSO-FTS [*Kuze et al.*, 2012]. The anomaly changes with time and has been quantified using the CAM image. Figure 2.2-1 shows examples of evaluated anomalies in the AT and CT angles as a function of uncorrected AT angle. Two clusters of AT/CT anomalies can be separated by using a threshold value for the AT angle (6.0 degrees for this case). Following correction is applied after the TANSO-FTS SWIR L2 V02.xx processing.

The pointing anomaly is corrected by adding averaged anomaly values to the AT and CT angles (red dashed-lines in Figure 2.2-1). The TANSO-FTS bore-sight vector is calculated from the corrected AT and CT angles. The averaged anomaly values can be obtained via the GOSAT Data Archive Service[1)].

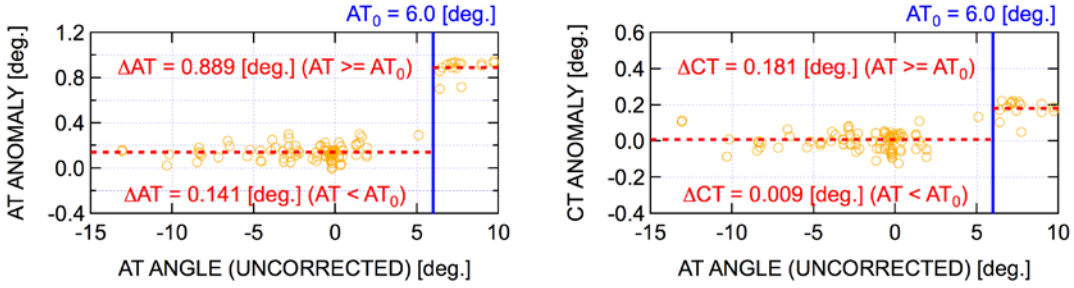

Figure 2.2-1: Examples of evaluated anomalies (specific observation; 2009/07/01 – 2009/11/03). The vertical blue line of AT = 6.0 degrees indicates the threshold value for correction. The horizontal red dotted line represents averaged anomalies.

(b) Spatial Correction of the FTS Bore-sight Vector

If the IMC is on, the observation point does not change during the acquisition of the interferogram. However, the instrument bore-sight vector changes due to the motion of the satellite, as shown in Figure 2.2-2. The orbital position of the satellite can be corrected by means of an extrapolation in a two-body problem that involves the position

1)
https://data2.gosat.nies.go.jp/techInfo/Correction%20Table%20for%20delta%20AT%20&%20delta%20CT%20Angles%20for%20geolocation%20error%20in%20FTS%20L1B.pdf

and velocity of the satellite at the start of the interferogram acquisition stored in the TANSO-FTS L1B data product. In this way, a corrected LOS can be obtained.

If the IMC is off, the observation point shifts due to the motion of the satellite (Figure 2.2-3). However, the bore-sight vector remains unchanged.

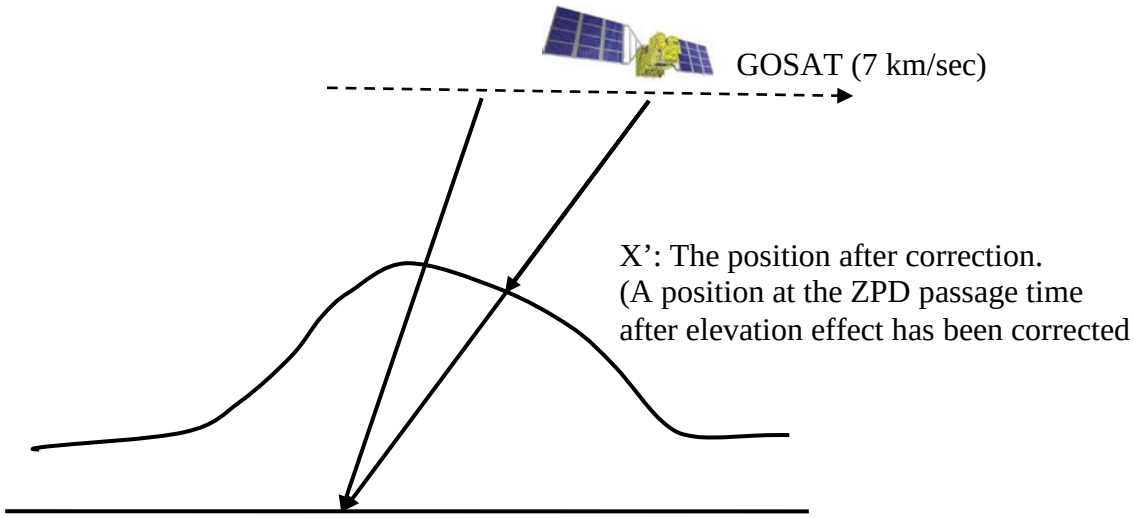

Figure 2.2-2: Overview of observation position correction (when IMC is in operation).

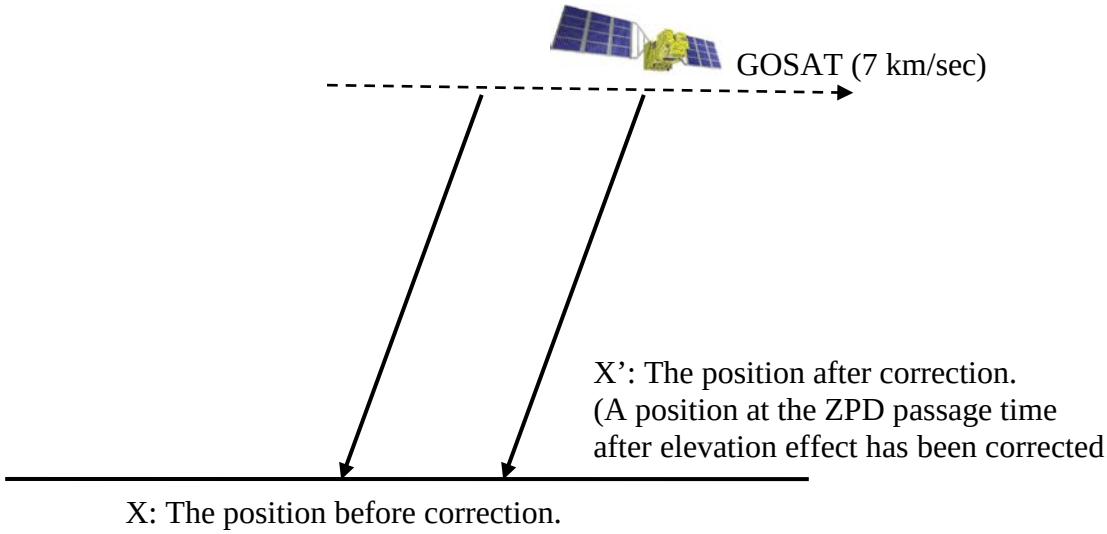

Figure 2.2-3: Overview of observation position correction (when IMC is not in operation).

After determining the LOS of the FTS at the ZPD passage time with respect to the satellite position, the next step is to determine the intersection of the earth's surface and the LOS in terms of longitude, latitude and elevation. For the elevation, the ASTER GDEM (Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model; which has a spatial mesh resolution of 1 arc-second, or equivalently 0.03 km mesh resolution at the equator) and digital elevation model data GTOPO30 (Global 30 Arc-Second elevation Data Set; which has a spatial mesh resolution of 30 arc-seconds, or equivalently 0.9 km mesh resolution at the equator) should be used. The position shift due to Ortho-correction is considered less than several kilometers for a side view observation on a mountainous region having an elevation of 1 km. Of course, for an observation of the ocean, the Ortho-correction is zero since the elevation is zero.

(c) Methodology for Obtaining the FTS's FOV-Area

The FTS's FOV area should be determined in the following manner. First, consider a group of conical vectors, spreading in 7.9 x 10^{-3} steradian, around the FTS's LOS at the ZPD passage time. The intersections of these conical vectors with the earth's surface are defined as the FTS's FOV area. In order to simplify the computation, 36 pyramids approximate this cone. Thus, the FOV-area is approximated by 36 polygons. Ortho-corrections should be applied to all the nodes of those 36 pyramids.

If the IMC is operating, the FOV-area becomes circular or elliptical (Figure 2.2-4 left). If the IMC is inactive or in the "off" state, the FTS's LOS moves during the acquisition period of the interferogram due to the orbital motion of the satellite. Since the orbital speed of GOSAT is about 7 km/s, the LOS moves about 28 km on the ground track during the 4-sec interferogram acquisition period. Thus the FOV-area becomes an oval shape of about 39 km in length and about 11 km in width. The oval FOV-area may be drawn by connecting two ellipses where one is the instantaneous FOV-area at the start of the interferogram and the other at the end of the interferogram. (Refer to the right depiction in Figure 2.2-4)

Figure 2.2-4: Examples of FTS FOV-areas projected on the earth's surface (a 36-polygon). The left depiction is for a computation while the IMC is "on" and the right depiction is for one with the IMC "off". Dark dots represent FOV-area before Ortho-correction, while red dots show after Ortho-correction. The right depiction shows a

sunglint-mode observation over the ocean; hence the FOV-area boundary is identical before and after the Ortho-correction.

2.3 Preparation of Meteorological Data

The TANSO-FTS SWIR L2 algorithms should refer to the meteorological data such as atmospheric pressure, temperature, and water vapor provided in Grid Point Values (GPV) by the Japan Meteorological Agency. The GPV contains the meteorological database of atmospheric pressure at 21 pressure levels ranging from 1000 hPa to 10 hPa (Table 2.3-1), with a horizontal resolution of 0.5 degrees and a temporal resolution of six hours. From this database, an atmospheric vertical profile of the FTS observation points should be constructed through temporal and spatial interpolations. The upper boundary of the GPV pressure of 10 hPa is insufficient to estimate the column amount accurately. It is preferable for the pressure range of the forward model computation (see Chapter 4) to provide the column amount negligibly small outside the computational boundary, but high enough for the local thermo-dynamic equilibrium (LTE) to hold inside the boundary. For this reason, the atmospheric pressure profile is extended with 10 additional levels as shown in Table 2.3-1 (i.e., the upper bound is set at 0.1 hPa).

Described below are the procedures through which to obtain the atmospheric vertical profile and earth's surface data at the FTS observation time (i.e., at the ZPD passage time) and at the observation point (i.e., the FTS bore-sight vector-determined and Ortho-corrected position at the ZPD passage time) along with the gravitational values, which are required by the computation of the column amount of dry air. This procedure starts with spatiotemporal interpolations in the horizontal direction (i.e., within a pressure level plane) of the atmospheric level data, followed by a similar interpolation in the vertical direction to construct earth's surface data. Next, the atmospheric profile is extended to 0.1 hPa pressure level, and finally, gravitational values for the extended 31 atmospheric pressure levels are calculated.

Table 2.3-1: Pressure levels in the atmospheric profile.

GPV pressure levels [hPa]	1000, 975, 950, 925, 900, 850, 800, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10
Augmented pressure levels [hPa]	7, 5, 3, 2, 1, 0.7, 0.5, 0.3, 0.2, 0.1

The physical quantities to be interpolated in the atmospheric level data include geo-potential height, temperature and relative humidity as well as meridional (north-south) and zonal (east-west) winds. In the GPV, the atmospheric pressure values in 21 levels (as shown in Table 2.3-1) are stored. A linear interpolation is the baseline for both spatial and temporal corrections.

To understand this process, let the time, longitude and latitude of a point for interpolation be denoted as (t, λ, φ). Also let the eight points selected from the GPV for interpolating physical quantities for (t, λ, φ) be defined as

$$t_0,\ t_1,\ \lambda_0,\ \lambda_1,\ \varphi_0,\ \varphi_1 \quad (t_0 \le t \le t_1,\ \lambda_0 \le \lambda \le \lambda_1,\ \varphi_0 \le \varphi \le \varphi_1)\,. \tag{2.3-1}$$

Then physical quantities denoted by x's to be interpolated at (t, λ, φ) are given by the linear relationships listed below:

Linear interpolation in time

$$x(t,\lambda_0,\varphi_0) = \left(\frac{t_1 - t}{t_1 - t_0}\right) x(t_0,\lambda_0,\varphi_0) + \left(\frac{t - t_0}{t_1 - t_0}\right) x(t_1,\lambda_0,\varphi_0)\,, \tag{2.3-2}$$

$$x(t,\lambda_0,\varphi_1) = \left(\frac{t_1 - t}{t_1 - t_0}\right) x(t_0,\lambda_0,\varphi_1) + \left(\frac{t - t_0}{t_1 - t_0}\right) x(t_1,\lambda_0,\varphi_1)\,, \tag{2.3-3}$$

$$x(t,\lambda_1,\varphi_0) = \left(\frac{t_1 - t}{t_1 - t_0}\right) x(t_0,\lambda_1,\varphi_0) + \left(\frac{t - t_0}{t_1 - t_0}\right) x(t_1,\lambda_1,\varphi_0)\,, \tag{2.3-4}$$

$$x(t,\lambda_1,\varphi_1) = \left(\frac{t_1 - t}{t_1 - t_0}\right) x(t_0,\lambda_1,\varphi_1) + \left(\frac{t - t_0}{t_1 - t_0}\right) x(t_1,\lambda_1,\varphi_1)\,. \tag{2.3-5}$$

Linear interpolation for meridional (north-south) direction

$$x(t,\lambda_0,\varphi) = \left(\frac{\varphi_1 - \varphi}{\varphi_1 - \varphi_0}\right) x(t,\lambda_0,\varphi_0) + \left(\frac{\varphi - \varphi_0}{\varphi_1 - \varphi_0}\right) x(t,\lambda_0,\varphi_1)\,, \tag{2.3-6}$$

$$x(t,\lambda_1,\varphi) = \left(\frac{\varphi_1 - \varphi}{\varphi_1 - \varphi_0}\right) x(t,\lambda_1,\varphi_0) + \left(\frac{\varphi - \varphi_0}{\varphi_1 - \varphi_0}\right) x(t,\lambda_1,\varphi_1)\,. \tag{2.3-7}$$

Linear interpolation for zonal (east-west) direction

$$x(t,\lambda,\varphi) = \left(\frac{\lambda_1 - \lambda}{\lambda_1 - \lambda_0}\right) x(t,\lambda_0,\varphi) + \left(\frac{\lambda - \lambda_0}{\lambda_1 - \lambda_0}\right) x(t,\lambda_1,\varphi)\,. \tag{2.3-8}$$

As for the earth's surface data, the GPV surface physical data *per se* should not be used. Instead, interpolated data derived from the GPV pressure level data corresponding to the average elevation of the TANSO-FTS's FOV-area, as shown below, should be used. The horizontal resolution of the GPV is about 55 km (0.5 deg), which is greater than the size of the TANSO-FTS's IFOV of about 10.5 km. Hence, it may happen that the elevation of the TANSO-FTS observation point is not accurately represented. To cope with this difficulty, average elevation of the observation FOV-area (obtained from ASTER GDEM/GTOPO30 that has a better horizontal resolution of about 1 km) and the

GPV pressure level data should be used for the interpolation in the vertical direction in order to compute the surface data.

In the derivation of the surface data, an extrapolation should be performed only if the average elevation of the FOV-area is lower than the altitude of the lowest pressure level (1000 hPa) in the GPV. For the remainder of cases, an interpolation should be performed. The physical quantities for interpolation include: pressure, temperature and relative humidity as well as meridional (north-south) and zonal (east-west) winds.

In an event that the average elevation of the FOV-area is lower than the elevation corresponding to 1000 hPa level, the surface pressure should be linearly extrapolated from the geo-potential heights at 1000 hPa and 975 hPa levels. Also, the surface temperature should be extrapolated with an assumption of temperature lapse rate of –5 K/km as elevation increases. For the surface values of relative humidity, meridional (north-south) wind and zonal (east-west) wind, an extrapolation approach should not be employed. Rather, the same values as those at 1000 hPa pressure level should be used without correction. If the average elevation of the FTS FOV-area is higher than the elevation of 1000 hPa level, the physical quantities should be computed using the third order spline interpolation scheme in the vertical direction.

In the GPV database, saturated water vapor pressure is represented by three equations as a function of temperature. If the temperature is 0 degrees C or higher, then a saturated water vapor equation by *Tetens* [1930] is used.

$$E_w = 6.11 \times 10^{\left(7.5t / (237.3+t)\right)}. \qquad (2.3\text{-}9)$$

If the temperature is –15 degrees C or lower, then a saturated water vapor equation with respect to ice by *Alduchov and Eskridge* [1996] is used.

$$E_i = 6.1121\, e^{\left(22.587t / (273.86+t)\right)}. \qquad (2.3\text{-}10)$$

If the temperature is –15 degrees C – 0 degrees C, then a linear interpolation (with respect to a temperature) of the above two expressions is used

$$E_s = \frac{(t+15)E_w - t\,E_i}{15}. \qquad (2.3\text{-}11)$$

where t stands for a temperature in degrees C, and the saturated water vapor pressure is measured in hPa.

Initially, it is necessary to select the atmospheric model that has the closest match with the GPV temperature profile after the spatiotemporal correction to extend the atmospheric profile to 0.1 hPa level. The selection is made from the five available

standard atmospheric models, which include models for tropical, mid-latitude summer, mid-latitude winter, high-latitude summer and high-latitude winter. In this process, one must first relate the temperature of the standard atmospheric model with the GPV atmospheric pressure; then compute the difference in temperature of the GPV profile through the following equation:

$$\sum_{l=l_s}^{21}\left[\frac{T_{GPV}(l)-T_{AFGL}(l)}{T_{GPV}(l)+T_{AFGL}(l)}\right]^2 . \qquad (2.3\text{-}12)$$

where $T_{GPV}(l)$ and $T_{AFGL}(l)$ stand for temperatures at atmospheric pressure l of the GPV and of the standard atmosphere model, respectively. l_s defines the GPV atmospheric pressure level that is closest to the pressure of the observation point. Select the atmospheric model from before-mentioned five standard atmospheric models that minimizes the values of Eq. (2.3-12).

Next, augment the temperature profile up to the atmospheric pressure level of 0.1 hPa by relating the temperature of the selected standard atmospheric model to the extended pressure levels. During this process, add an offset value to the temperature profile of the standard atmospheric model so that the temperatures of the standard atmospheric model and that of the GPV become identical at 10 hPa. If this offset value is too large, the temperature differences in the mesosphere become proportionally overlarge. In order to cope with this difficulty, variable offset values (of a decreasing function of an altitude and zero at 0.1 hPa) should be assigned. Similarly the water vapor profile should be extended up to the atmospheric pressure of 0.1 hPa. Figure 2.3-1 shows an example of the temperature profile and the atmospheric water vapor profile.

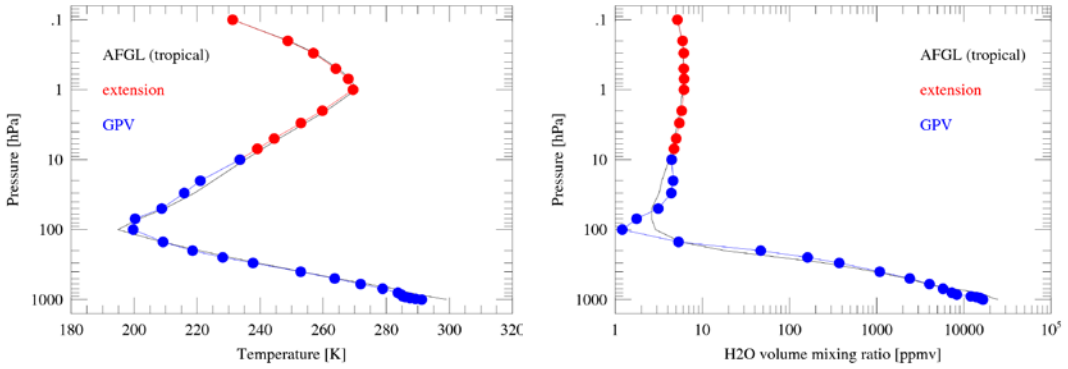

Figure 2.3-1: Example of temperature profile (left) and water vapor profile extended to 0.1 hPa level.

Once the atmospheric temperature and water vapor levels are determined, the next task is to compute the gravity by the gravitational potential through the following equation:

$$U(r,\theta) = -\frac{GM}{r}\left[1 - J_2\left(\frac{a}{r}\right)^2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)\right] - \frac{1}{2}r^2\sin^2\theta\omega^2, \qquad (2.3\text{-}13)$$

where r is the distance from the center of the earth, θ stands for the polar angle (Figure 2.3-2), a represents the equatorial radius of the earth, GM is the gravitational constant, ω is the angular velocity of the earth, and J_2 is the second order coefficient in the spherical harmonic expansion. Also, the directional components of the gravity in r and θ are defined through the following equations, respectively:

$$g_r = -\frac{\partial U(r,\theta)}{\partial r} = -\frac{GM}{r^2}\left[1 - 3J_2\left(\frac{a}{r}\right)^2\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)\right] + r\sin^2\theta\omega^2, \quad (2.3\text{-}14)$$

$$g_\theta = -\frac{1}{r}\frac{\partial U(r,\theta)}{\partial \theta} = \frac{GM}{r^2}3J_2\left(\frac{a}{r}\right)^2\sin\theta\cos\theta + r\sin\theta\cos\theta\omega^2. \qquad (2.3\text{-}15)$$

An equivalent equation may be presented in terms of altitude z and latitude φ, instead of r and θ, of the observation point as:

$$g_z = -\frac{\partial U(z,\varphi)}{\partial z} = g_r\cos(\varphi - \phi) - g_\theta\sin(\varphi - \phi), \qquad (2.3\text{-}16)$$

where $\phi = \pi/2 - \theta$. The looking-down vertical gravity is $g = -g_z$.

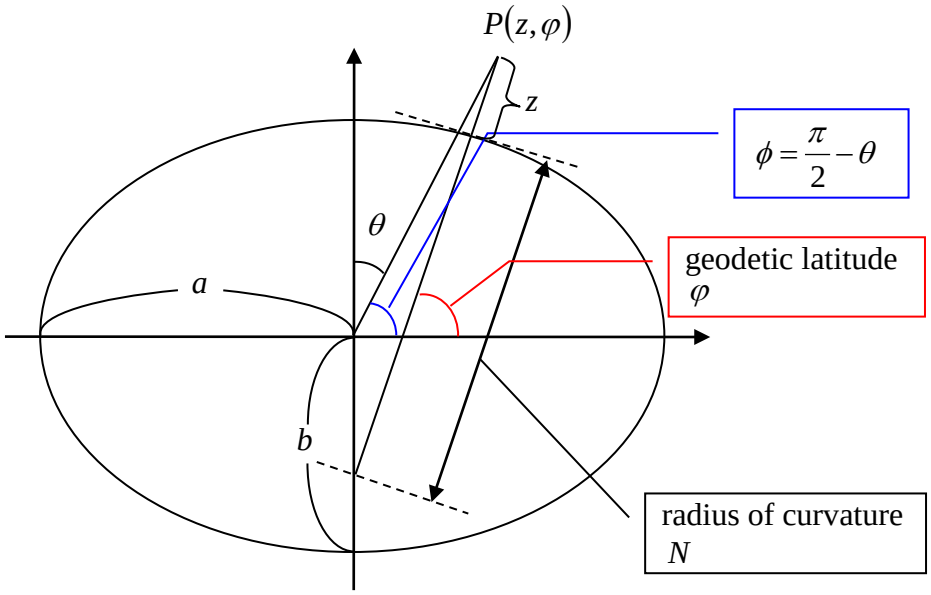

Figure 2.3-2: Ellipsoidal Earth.

In order to compute the gravity in Eq. (2.3-16), it is necessary to know the value for the altitude. It may be computed by assuming that geo-potential is identical with gravitational potential at a given atmospheric pressure. The geo-potential is computed as shown below.

Let the atmospheric pressure be p, and mass density ρ. Then the hydrostatic equilibrium equation is given by

$$-\frac{1}{\rho}\frac{dp}{dz} = -g\,, \tag{2.3-17}$$

and a state equation for moist atmosphere is presented by

$$p = \rho R_d T_v\,, \tag{2.3-18}$$

where R_d stands for a gas constant for dry air. Now the virtual-temperature T_v can be represented by

$$T_v = \frac{1 + x_{H2O}}{1 + (\mu_v/\mu_d)x_{H2O}} T\,, \tag{2.3-19}$$

where T refers to temperature, x_{H2O} denotes the volume mixing ratio of water vapor, μ_v and μ_d, respectively, refer to the mean molecule-weights of moisture and dry air. Eq. (2.3-17) can be re-written using the relationship given by Eq. (2.3-18) as:

$$gdz = -R_d T_v \frac{dp}{p}. \tag{2.3-20}$$

An integration along the vertical direction from $z = z_l$ to z_{l+1} yields

$$U_{l+1} - U_l = -R_d \int_{p_l}^{p_{l+1}} T_v \frac{dp}{p}, \tag{2.3-21}$$

where U_l stands for the geo-potential at atmospheric pressure l. Now, handle T_v inside the integral in Eq. (2.3-21) in terms of an arithmetic average. Then it follows that

$$U_{l+1} - U_l = R_d \frac{T_v(l) + T_v(l+1)}{2} \ln \frac{p_l}{p_{l+1}}. \tag{2.3-22}$$

Thus, once the geo-potential at the reference atmospheric pressure is given, the geo-potentials at all atmospheric pressure levels may be computed through Eq. (2.3-22).

Even when the elevation *h* of the observation point and atmospheric pressure at the surface p_s are given, the computational procedures are different depending upon the situation where p_s is higher or lower than the atmospheric pressure p_1 (= 1000 hPa) at the lowest layer in the GPV.

If p_s is higher than p_1, then start with the geo-potential at atmospheric pressure level $l = 1$, where

$$U_1 = U(h, \varphi) + R_d T_v(1) \ln \frac{p_s}{p_1}, \tag{2.3-23}$$

and use Eq. (2.3-22) to compute the geo-potential along the vertical and upward direction for each atmospheric pressure at $l+1$.

On the other hand, if the surface pressure is in-between k-th and $k+1$-st levels, i.e., $p_k \geq p_s \geq p_{k+1}$, one should start with the geo-potential

$$U_{k+1} = U(h, \varphi) + R_d \frac{T_v(k) + T_v(k+1)}{2} \ln \frac{p_s}{p_{k+1}}, \tag{2.3-24}$$

and use Eq. (2.3-22) to compute the geo-potential along the vertical direction, upward or downward, for each atmospheric pressure.

Once geo-potential is obtained, solve the equation

$$U(z_l, \varphi) = U_l, \tag{2.3-25}$$

for z_l. This means that altitude at each atmospheric pressure level becomes known. Incidentally, Eq. (2.3-25) is easily solvable by means of the Newton-Raphson method.

2.4 Degradation Correction

Degradation of the sensitivity of TANSO-FTS is corrected by utilizing the radiometric degradation model described in *Yoshida et al.* [2012]. This radiometric degradation model is evaluated based on the on-orbit solar calibration data, and is given as a function of time and wavenumber for each SWIR band. The absolute value for radiometric calibration is based on the vicarious calibration campaign results at Railroad Valley [*Kuze et al.*, 2011]. The degradation-corrected spectra of P- and S-polarizations are used for the wavenumber axis correction and the polarization synthesis process. Although the degradation correction factor is a predicted value for a semi-real time processed data, the effect of this prediction error is expected to be small, because the degradation becomes slowly with time in general.

2.5 Wavenumber Axis Correction

The wavenumber axis in the TANSO-FTS L1B data products may not represent the true value due to the self-apodization and temperature dependency of the sampling laser's wavelength. For this reason, it is base-lined in the TANSO-FTS SWIR L2 processing to perform a wavenumber correction as described below.

Let the observed spectra be expressed as

$$S_{P/S}^{obs}(i), i = 1,2,\cdots,N\,, \tag{2.5-1}$$

where $S_P^{obs}(i)$ is P-polarization component, and $S_S^{obs}(i)$ is S-polarization component. Let the wavenumber axis of the observation spectra stored in the TANSO-FTS L1B data product be denoted as

$$v_i^{obs} = v_1^{obs} + (i-1)\Delta v^{obs}, i = 1,2,\cdots,N\,. \tag{2.5-2}$$

It is required that v_1^{obs} is an integer-multiple of Δv^{obs}. Then the true wavenumber axis may be expressed in terms of a correction factor ρ times the observed spectral wavenumber, as

$$v_i = \rho v_i^{obs}, i = 1,2,\cdots,N\,. \tag{2.5-3}$$

The correction factor ρ should be determined in such a way that the cross correlation between the observed and the reference spectra (having the true wavenumber axis) becomes maximum. In the TANSO-FTS L2 algorithm, the reference spectra are defined as the product of solar irradiance and atmospheric transmittance. That is, the reference monochromatic spectrum is given by the following equation:

$$I^{ref}(\nu) = F_0(\nu)\exp\left[-\left(\frac{1}{\cos\theta_0}+\frac{1}{\cos\theta_1}\right)\tau(\nu)\right], \tag{2.5-4}$$

where F_0 is solar irradiance spectra, τ stands for the optical thickness of the atmosphere, θ_0 refers to solar zenith angle and θ_1 represents satellite zenith angle. Thus the convolution integral with the instrument line shape function yields that

$$S_{P/S}^{ref}(i) = \int_{-\infty}^{\infty} ILS_{P/S}(\nu_i - \nu) I^{ref}(\nu) d\nu, \tag{2.5-5}$$

where ILS_P stands for the instrument line shape function for the P-polarization, and ILS_S stands for the S-polarization. The approximation of the above integration in terms of a summation produces the following expression:

$$S_{P/S}^{ref}(i) = \sum_{j=-N_W}^{N_W} ILS_{P/S}(-j\delta\nu) I^{ref}(\nu_i + j\delta\nu)\delta\nu. \tag{2.5-6}$$

In the above equation, $\delta\nu$ is the integration step size of the convolution integral and typically takes a value of 0.01 cm^{-1}. An integer N_W is determined so that the convolution integral half width $N_W\delta\nu$ becomes about 20 cm^{-1}.

The aforementioned cross correlation function is defined by

$$C(\rho) = \sum_{i=1}^{N}\left(S_P^{obs}(i)S_P^{ref}(i) + S_S^{obs}(i)S_S^{ref}(i)\right), \tag{2.5-7}$$

where ρ is to be determined to maximize $C(\rho)$, as explained earlier. In the TANSO-FTS SWIR L2 algorithm, the Golden Section Search method is used to determine the value of ρ.

In order to obtain the value of ρ through an iterative method, it becomes necessary to facilitate a computation of reference monochromatic spectra I^{ref} at any arbitrary wavenumber. To accomplish this computation in high-speed, it is convenient if a table for an optical thickness τ has been built through a separate pre-computation. The wavenumber axes for such a table consist of wavenumbers for a static earth-atmospheric coordinate system, and are expressed as:

$$\nu_i^{earth} = \nu_1^{earth} + (i-1)\Delta\nu^{earth}, i = 1,2,\cdots,N^{ref}. \tag{2.5-8}$$

Use is made of such a table in order to compute the reference monochromatic spectra through Eq. (2.5-4).

$$I_i^{ref}, i = 1,2,\cdots, N^{ref}. \tag{2.5-9}$$

The wavenumbers of the spectra observed by the satellite v_i^{sat} would Doppler-shift in accordance with the following equation:

$$v_i^{sat} = \left(1 + \frac{v_{dop}}{c}\right) v_i^{earth}, i = 1,2,\cdots, N^{ref}, \tag{2.5-10}$$

where v_{dop} represents the relative velocity between the earth's atmosphere and the satellite instrument along the instrument's direction of sight, and is positive if both are approaching each other.

The monochromatic spectrum appearing in Eq. (2.5-6) is computed through (a) obtaining the spectra I_i^{ref} through Eq. (2.5-10), and (b) carrying out a four-point Lagrange-interpolation. Incidentally, the optical thickness table should be built for an atmospheric profile of US standard where the carbon dioxide concentration is increased to a level of 380 ppm.

2.6 Polarization Synthesis Process

The polarization synthesis stands for the process of converting the observed spectra of P- and S-polarizations into estimated values of the total radiance spectra of the light entering the sensor. Consequently, it is viable for the forward model to use a scalar radiative transfer code. Generally speaking, the polarization of atmospheric radiation is expressed in terms of Stokes parameters $(I \quad Q \quad U \quad V)^T$ with respect to a local meridian plane as a reference. On the other hand, it is quite natural that the instrument's sensitivity of polarization is expressed in terms of the reflection plane of a pointing mirror (i.e., the plane defined by the incident and reflected light vectors, as the polarization reference plane). The set of Stokes parameters expressed in terms of the polarization sensitivity reference plane is denoted as $(I' \quad Q' \quad U' \quad V')^T$. These two sets of Stokes parameter are related to each other through a rotation matrix $\mathbf{L}(\theta_{PM})$ to be discussed later. θ_{PM} is an angle between the local meridian plane and the reflection plane of the pointing mirror. The explanation of the polarization synthesis process is made in three steps: First a polarization sensitivity model is defined. Then, a methodology of computing θ_{PM} is presented, and finally, an algorithm for polarization synthesis is detailed.

Let the parallel and perpendicular polarized components of the incident light with respect to the polarization reference plane be denoted as $I_{\parallel}$ and $I_{\perp}$, respectively. Then I and Q are given as

$$I = I_{\parallel} + I_{\perp}, \tag{2.6-1}$$

$$Q = I_{\parallel} - I_{\perp}. \tag{2.6-2}$$

Instead of traditional Stokes parameters, let the polarization parameters for the incident light into the sensor be expressed in terms of $\begin{pmatrix} I'_{\parallel} & I'_{\perp} & U' & V' \end{pmatrix}^T$ components. Then the radiance detected by the P- and S-polarization detectors are given, respectively, by

$$S_P = a'_P I'_{\parallel} + b'_P I'_{\perp} + c'_P U', \tag{2.6-3}$$

$$S_S = a'_S I'_{\parallel} + b'_S I'_{\perp} + c'_S U'. \tag{2.6-4}$$

In atmospheric radiative transfer, V' is ignored since it is sufficiently small in comparison with other Stokes parameters. When the optical system in the FTS is perfectly aligned and the polarization beam splitter is functioning perfectly, the coefficients $a'_{P/S}$, $b'_{P/S}$ and $c'_{P/S}$ in Eqs. (2.6-3) and (2.6-4) can be presented as functions of the wavenumber ν, AT angle θ_{AT} and CT angle θ_{CT}, as:

$$a'_P(\nu, \theta_{AT}, \theta_{CT}) = W_P(\nu) R_{\parallel}(\nu, \theta_{AT}) \cos^2 \theta_{CT}, \tag{2.6-5}$$

$$b'_P(\nu, \theta_{AT}, \theta_{CT}) = W_P(\nu) R_{\perp}(\nu, \theta_{AT}) \sin^2 \theta_{CT}, \tag{2.6-6}$$

$$c'_P(\nu, \theta_{AT}, \theta_{CT}) = -\frac{1}{2} W_P(\nu) \sqrt{R_{\parallel}(\nu, \theta_{AT}) R_{\perp}(\nu, \theta_{AT})} \sin 2\theta_{CT} \cos \delta(\nu, \theta_{AT}), \tag{2.6-7}$$

$$a'_S(\nu, \theta_{AT}, \theta_{CT}) = W_S(\nu) R_{\parallel}(\nu, \theta_{AT}) \sin^2 \theta_{CT}, \tag{2.6-8}$$

$$b'_S(\nu, \theta_{AT}, \theta_{CT}) = W_S(\nu) R_{\perp}(\nu, \theta_{AT}) \cos^2 \theta_{CT}, \tag{2.6-9}$$

$$c'_S(\nu, \theta_{AT}, \theta_{CT}) = \frac{1}{2} W_S(\nu) \sqrt{R_{\parallel}(\nu, \theta_{AT}) R_{\perp}(\nu, \theta_{AT})} \sin 2\theta_{CT} \cos \delta(\nu, \theta_{AT}), \tag{2.6-10}$$

where $R_{\parallel}$ and $R_{\perp}$ denote the reflectivities of the pointing mirror in the parallel and perpendicular directions, respectively. Let the electric field vectors of pre- and post-reflections of the pointing mirror be expressed as $\begin{pmatrix} E^t_{\parallel} & E^t_{\perp} \end{pmatrix}^T$ and $\begin{pmatrix} E^{tt}_{\parallel} & E^{tt}_{\perp} \end{pmatrix}^T$, respectively. Then they are related as

$$\begin{pmatrix} E''_{\parallel} \\ E''_{\perp} \end{pmatrix} = \begin{pmatrix} \sqrt{R_{\parallel}} e^{i\delta_{\parallel}} & 0 \\ 0 & \sqrt{R_{\perp}} e^{i\delta_{\perp}} \end{pmatrix} \begin{pmatrix} E'_{\parallel} \\ E'_{\perp} \end{pmatrix}, \tag{2.6-11}$$

where δ denotes the phase shift caused by the protective film of a pointing mirror and is defined as

$$\delta = \delta_{\parallel} - \delta_{\perp}. \tag{2.6-12}$$

The values for $R_{\parallel}$, $R_{\perp}$ and δ are dependent upon the AT angle. Since their values have not been measured for the flight articles, similar sample devices manufactured together with the flight articles should be used as a substitute.

Coefficients W_P and W_S depends upon the transmittance of the band pass filter, efficiency of the optical system, sensitivity of detectors etc., and should be determined by means of data analyses of the pre-launch integration sphere calibration data. If the integration sphere with known radiance L is measured with zero AT and CT angles, then the following equations hold:

$$S_P(v) = \frac{L}{2} W_P(v) R_{\parallel}(v, \theta_{AT} = 0), \tag{2.6-13}$$

$$S_S(v) = \frac{L}{2} W_S(v) R_{\perp}(v, \theta_{AT} = 0). \tag{2.6-14}$$

Thus, W_P and W_S can readily be obtained from the above equations.

Let the Stokes parameters having a local meridian plane as a polarization reference plane be denoted as $\mathbf{I} = (I \quad Q \quad U \quad V)^T$, and the Stokes parameters having the reflection plane of a pointing mirror as a polarization reference plane be denoted as $\mathbf{I}' = (I' \quad Q' \quad U' \quad V')^T$. Then $\mathbf{I}$ and $\mathbf{I}'$ are related as:

$$\mathbf{I}' = \mathbf{L}(\theta_{PM})\mathbf{I}, \tag{2.6-15}$$

where the rotation matrix $\mathbf{L}$ for the Stokes parameters is given by

$$\mathbf{L}(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta & 0 \\ 0 & -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.6-16}$$

Now, the derivation method of θ_{PM} is described below.

The vector $\mathbf{n}_{PM}$ perpendicular to the plane of pointing mirror is expressed by Eq. (2.6-17) in the sensor coordinate system XYZ.

$$\mathbf{n}_{PM} = \begin{pmatrix} -\sin(\pi/4 + \theta_{AT}) \\ -\sin\theta_{CT} \cos(\pi/4 + \theta_{AT}) \\ \cos\theta_{CT} \cos(\pi/4 + \theta_{AT}) \end{pmatrix}, \tag{2.6-17}$$

where θ_{AT} is called as a mechanical AT angle of the pointing mirror. It should be noted that the value of θ_{AT} is a half (1/2) of the AT angle value in the TANSO-FTS L1B data products. Since the alignment errors of the optical bench are deemed to be sufficiently small, the senor coordinate system and the satellite coordinate system are identical. The X-axis lies in the direction of the satellite velocity vector, and the positive Z-axis is pointed in the nadir direction. The incident light into the pointing mirror is reflected in

the negative X -axis direction. As shown later, it is necessary to express the directional vector of the X -axis in the geo-centric coordinate system. The transformation from the satellite coordinate system into the geo-centric coordinate system may be accomplished through the quaternion elements $(q_0 \quad q_1 \quad q_2 \quad q_3)$ [1)] provided in the TANSO-FTS L1B data products. The transformation matrix $\mathbf{A}$ is given as shown below:

$$\mathbf{A} = \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_0^2 & 2q_1q_2 - 2q_3q_0 & 2q_1q_3 + 2q_2q_0 \\ 2q_1q_2 + 2q_3q_0 & -q_1^2 + q_2^2 - q_3^2 + q_0^2 & 2q_2q_3 - 2q_1q_0 \\ 2q_1q_3 - 2q_2q_0 & 2q_2q_3 + 2q_1q_0 & -q_1^2 - q_2^2 + q_3^2 + q_0^2 \end{pmatrix}. \quad \text{(2.6-18)}$$

With this transformation matrix, the unit vector $\mathbf{n}_X$ of the X -axis of the satellite coordinate system can be derived as shown below:

$$\mathbf{n}_X = \mathbf{A}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_0^2 \\ 2q_1q_2 + 2q_3q_0 \\ 2q_1q_3 - 2q_2q_0 \end{pmatrix}. \quad \text{(2.6-19)}$$

Consider a coordinate system depicted in Figure 2.6-1, in which the origin O of the coordinate system is the position of the satellite and the vertical direction at the observation point is ON. Let the direction of the incident light be denoted by OP_1. Then the local meridian plane to be used as the polarization reference plane in the atmospheric radiative transfer is defined by ONP_1. If it is assumed that the incident light is reflected in the OP_2 direction by the pointing mirror, then the reflection plane of the pointing mirror can be defined by $\mathrm{OP}_1\mathrm{P}_2$.

If the azimuth angle ϕ_1 of the incident light is greater than the azimuth angle ϕ_2 of the reflected light as shown in Figure 2.6-1, the local meridian plane coincides with the reflection plane of the pointing mirror by rotating the former by $-\sigma$ as seen from the incident light propagation direction. In this case, $\theta_{PM} = -\sigma$.

If the azimuth angle ϕ_1 of the incident light is smaller than the azimuth angle ϕ_2 of the reflected light as shown in Figure 2.6-2, the local meridian plane coincides with the reflection plane of the pointing mirror by rotating the former by $+\sigma$. That is, $\theta_{PM} = \sigma$.

According to the formulae for a spherical triangle, σ in Figures 2.6-1 and 2.6-2 can be obtained from

$$\cos\sigma = \frac{\cos\theta_2 - \cos\theta_1\cos\theta}{\sin\theta_1\sin\theta}, \quad \text{(2.6-20)}$$

1) /exposureAttribute/pointAttribute/satellite/satAtt

where θ is the angle between the incident light and the reflected light, and is given by

$$\cos\theta = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2). \qquad (2.6\text{-}21)$$

In this equation, the angles associated with the incident and reflected lights are given in terms of their polar angles θ_1 and θ_2, respectively.

(θ_1, ϕ_1) and (θ_2, ϕ_2) can be derived as shown in the sequel. Let $\mathbf{x}_{sat}$ and $\mathbf{x}_{obs}$ be the position vectors of the satellite and the observation point, respectively. Then the directional vector $\mathbf{n}_{inc}$ of the incident light is given by

$$\mathbf{n}_{inc} = \frac{\mathbf{x}_{sat} - \mathbf{x}_{obs}}{|\mathbf{x}_{sat} - \mathbf{x}_{obs}|}, \qquad (2.6\text{-}22)$$

The value for $\mathbf{x}_{sat}$ in the geo-centric coordinate system[1)] can be obtained from the TANSO-FTS L1B data products. Whereas $\mathbf{x}_{sat}$ in the TANSO-FTS L1B refers to the satellite position vector at the start of the observation, $\mathbf{x}_{sat}$ for the TANSO-FTS SWIR L2 algorithm is the same vector at the middle of the observation (ZPD pass time). For this reason, correct $\mathbf{x}_{sat}$ should be computed through an appropriate interpolation. Also

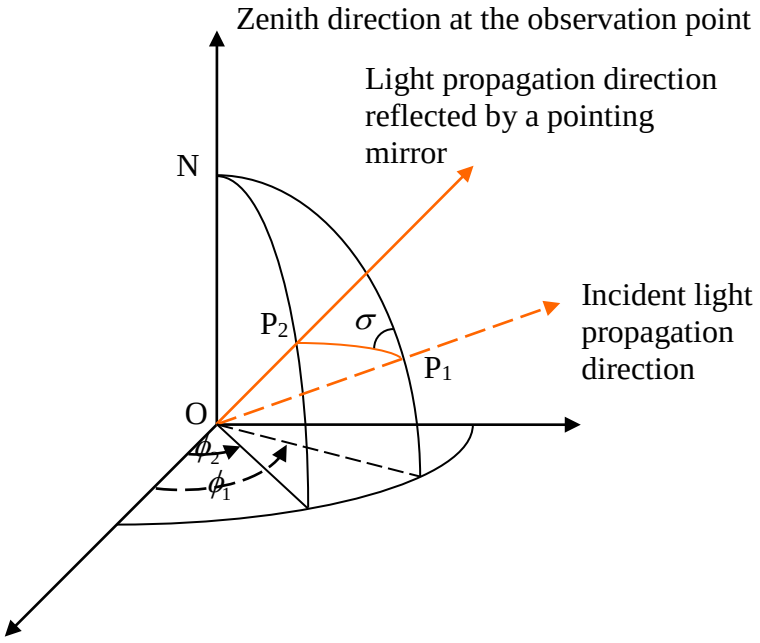

Figure 2.6-1: The local meridian plane (ONP_1) and the reflection plane of the pointing mirror (OP_1P_2). In the case of $0 < \phi_1 - \phi_2 < \pi$.

1) /exposureAttribute/pointAttribute/satellite/satPos

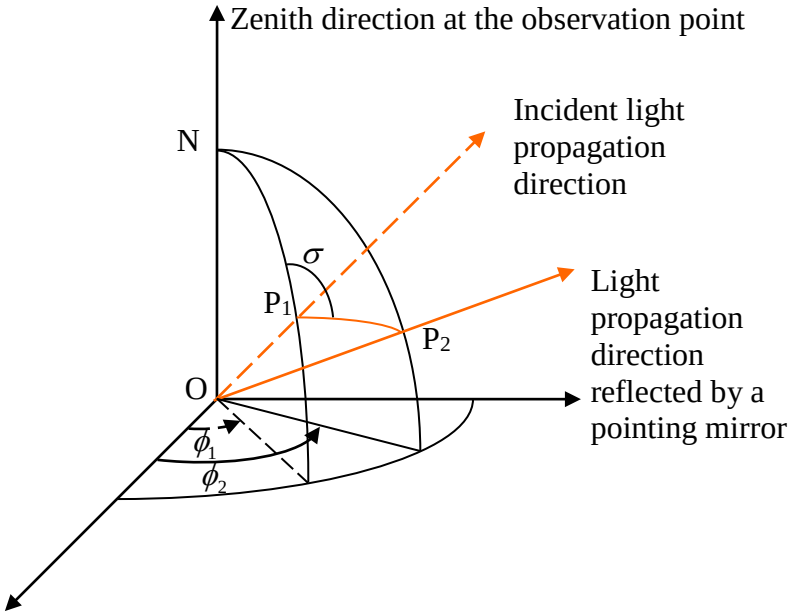

Figure 2.6-2: The local meridian plane (ONP_1) and the reflection plane of the pointing mirror ($\mathrm{OP}_1\mathrm{P}_2$). In the case of $0 < \phi_2 - \phi_1 < \pi$.

the $\mathbf{x}_{obs}$ vector can readily be computed based on the longitude and latitude information of the observation point. In this way, numerical values of $\mathbf{n}_{inc}$ can be computed from Eq. (2.6-22). Also the $\mathbf{n}_{inc}$ vector is expressed as

$$\mathbf{n}_{inc} = \begin{pmatrix} \sin\theta_1 \cos\phi_1 \\ \sin\theta_1 \sin\phi_1 \\ \cos\theta_1 \end{pmatrix}, \tag{2.6-23}$$

when the coordinate system transformation is conducted so that the vertical axis of the observation point coincides with the polar axis. From Eqs. (2.6-22) and (2.6-23), numerical values for (θ_1, ϕ_1) can be obtained. In order to obtain numerical values of (θ_2, ϕ_2), follow the similar process after performing a rotation around the instrument's $-X$ axis in the geo-centric coordinate system.

If $(I'_{\parallel}, I'_{\perp}, U')$ is used in place of $(I_{\parallel}, I_{\perp}, U)$, Eqs. (2.6-3) and (2.6-4) can be rewritten as follow:

$$S_P = a_P I_{\parallel} + b_P I_{\perp} + c_P U \,, \tag{2.6-24}$$

$$S_S = a_S I_{\parallel} + b_S I_{\perp} + c_S U \,, \tag{2.6-25}$$

where the coefficients $a_{P/S}$, $b_{P/S}$, and $c_{P/S}$ are defined below:

$$a_{P/S} = \frac{1}{2}\left(a'_{P/S} + b'_{P/S}\right) + \frac{1}{2}\left(a'_{P/S} - b'_{P/S}\right)\cos 2\theta_{PM} - c'_{P/S}\sin 2\theta_{PM}\,, \qquad (2.6\text{-}26)$$

$$b_{P/S} = \frac{1}{2}\left(a'_{P/S} + b'_{P/S}\right) - \frac{1}{2}\left(a'_{P/S} - b'_{P/S}\right)\cos 2\theta_{PM} + c'_{P/S}\sin 2\theta_{PM}\,, \qquad (2.6\text{-}27)$$

$$c_{P/S} = \frac{1}{2}\left(a'_{P/S} - b'_{P/S}\right)\sin 2\theta_{PM} + c'_{P/S}\cos 2\theta_{PM}\,. \qquad (2.6\text{-}28)$$

Eqs. (2.6-24) and (2.6-25) together with Eqs. (2.6-26), (2.6-27) and (2.6-28) present instrument response to the atmospheric radiation having arbitrary polarization characteristics. Investigated in the sequel is the algorithm that represents the total radiance $I = I_{\parallel} + I_{\perp}$ incident into the instrument, based on the instrument responses S_P and S_S. Generally speaking, there are three unknowns (i.e., (I, Q, U) or $(I_{\parallel}, I_{\perp}, U)$, with only two known values (i.e., S_P and S_S). Hence, it is not feasible to determine uniquely the polarization parameters of the incident light or the total radiance. However, as a special case, the total radiance may be determined uniquely when $\theta_{CT} = 0$ or $R_{\parallel} = R_{\perp}$. When these special conditions are not satisfied, the total radiance is derived in an approximate sense as shown below.

Let the angle between the polarization reference plane and the major axis of a vibrational ellipse be denoted as χ as shown in Figure 2.6-3. With a shorthand notation of $P = \cos 2\beta$, the polarization parameters are expressed as shown below:

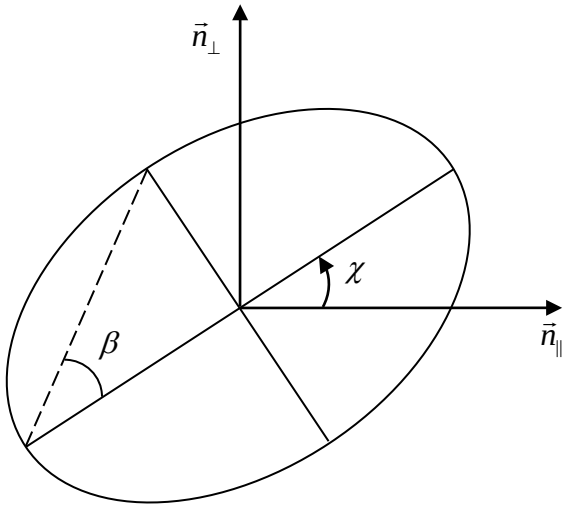

Figure 2.6-3: Vibrational ellipse.

$$I_{\parallel} = \frac{1}{2}\left(1 + P\cos 2\chi\right)I\,, \qquad (2.6\text{-}29)$$

$$I_{\perp} = \frac{1}{2}\left(1 - P\cos 2\chi\right)I\,, \qquad (2.6\text{-}30)$$

$$U = P\sin 2\chi I\,. \qquad (2.6\text{-}31)$$

The substitution of these equations into Eqs. (2.6-24) and (2.6-25) yields:

$$S_P = \frac{a_P + b_P}{2} I + \left(\frac{a_P - b_P}{2} \cos 2\chi + c_P \sin 2\chi \right) PI \,, \tag{2.6-32}$$

$$S_S = \frac{a_S + b_S}{2} I + \left(\frac{a_S - b_S}{2} \cos 2\chi + c_S \sin 2\chi \right) PI \,. \tag{2.6-33}$$

The solution of the above equations with respect to I is

$$I = d_P S_P + d_S S_S \,, \tag{2.6-34}$$

where the coefficients are given by the following expressions:

$$d_P = \left(\frac{a_S - b_S}{2} \cos 2\chi + c_S \sin 2\chi \right) \Big/ D \,, \tag{2.6-35}$$

$$d_S = -\left(\frac{a_P - b_P}{2} \cos 2\chi + c_P \sin 2\chi \right) \Big/ D \,, \tag{2.6-36}$$

$$\begin{aligned} D = & \left(\frac{a_P + b_P}{2} \right) \left(\frac{a_S - b_S}{2} \cos 2\chi + c_S \sin 2\chi \right) \\ & - \left(\frac{a_S + b_S}{2} \right) \left(\frac{a_P - b_P}{2} \cos 2\chi + c_P \sin 2\chi \right) \end{aligned} \,. \tag{2.6-37}$$

Thus the true total radiance values can be obtained if the value of χ is correctly given in Eqs. (2.6-34), (2.6-35), (2.6-36) and (2.6-37). For instance, if the observation point happens to be a specular reflection point over the ocean, then $\chi = 0$ because of symmetry. However, this is a special case, and the value of χ is not readily discernible. In order to cope with this situation, the values of χ may be estimated by means of a single-scattering approximation. This method provides a value for χ, which is determined only by the geometric shape of the reflector independently from the scattering matrix of the aerosols and/or clouds and those vertical distribution, or from the reflectance matrix of earth's surface. This is shown in the discussions below.

Let the scattering matrix $\mathbf{F}$ be defined as

$$\mathbf{F}(\Theta) = \begin{pmatrix} a_1(\Theta) & b_1(\Theta) & 0 & 0 \\ b_1(\Theta) & a_2(\Theta) & 0 & 0 \\ 0 & 0 & a_3(\Theta) & b_2(\Theta) \\ 0 & 0 & -b_2(\Theta) & a_4(\Theta) \end{pmatrix}, \tag{2.6-38}$$

where Θ stands for a scattering angle. Use is made of a rotation matrix **L** defined in terms of Stokes parameters. Then the phase matrix **Z** is given by

$$\mathbf{Z} = \mathbf{L}(\pi - \sigma_2)\mathbf{F}(\Theta)\mathbf{L}(-\sigma_1). \tag{2.6-39}$$

In this equation, σ_1 and σ_2 are related to the rotation angle of the polarization reference plane and can be derived in a similar way as the derivation of θ_{PM}. The details of the derivation are not presented here, but can be found in the reference by *Hovenier and van der Mee* [1983].

Let the solar irradiance vector be denoted as $\mathbf{F}_0 = (F_0 \quad 0 \quad 0 \quad 0)^T$. Then the source term of the radiative transfer equation is proportional to the phase matrix **Z** multiplied by the solar irradiance vector as shown in the following equation:

$$\mathbf{Z}\mathbf{F}_0 = F_0 \begin{pmatrix} a_1(\Theta) \\ b_1(\Theta)C_2 \\ b_1(\Theta)S_2 \\ 0 \end{pmatrix}, \tag{2.6-40}$$

where C_2 and S_2 are parameters to be determined only by the rotational angle of the polarization reference plane. From the above equation, it is clear that the path-radiance component ratio is $C_2 : S_2$. Consider Q and U of the surface-reflected components. They are both zero for a Lambertian surface, and $Q:U = C_2 : S_2$ for an ocean surface since the reflection matrix of an ocean surface can be presented in a similar way as Eq. (2.6-40). Therefore the addition of the path-radiance and surface-reflected components produces $Q:U = C_2 : S_2$.

Based on the above investigation, χ obtained by means of a single-scattering approximation is given by

$$\tan 2\chi = \frac{S_2}{C_2}. \tag{2.6-41}$$

Eq. (2.6-34) holds for a monochromatic radiation. Hence the left-hand side of corresponding equation for the spectra of the TANSO-FTS should be given in terms of a convolution integral of radiance $I(v)$ convolved with an instrument line shape function $ILS(v)$ where

$$ILS(v) = \frac{1}{2}\left[ILS_P(v) + ILS_S(v)\right]. \tag{2.6-42}$$

The terms S_P and S_S in the right hand side of the above equation represent the P- and S-polarization components, respectively. If the instrument line shape functions for the P- and S-polarizations are different, the polarization synthesize equation (i.e., Eq. (2.6-34)) yields an error proportional to the product of the two components (i.e., the difference between ILS_P and ILS_S and the degree of a linear polarization). However, the pre-launch simulation verified that this error was in the order of 1/100 of the observation noise. Hence it is considered negligible in TANSO-FTS SWIR L2 processing.

3. Data Screening

Of the 9000 or so pieces of scans collected by the TANSO-FTS SWIR per day, only the high quality data should be processed during the TANSO-FTS SWIR L2 retrieval. The means by which high quality data and low quality data are sifted is called "data screening". Criteria of data screening are often condition-based. For example, one criterion for high quality data is that the data was collected while the pointing mirror was in a stable tracking mode. Another criterion is a cloud-free condition determined by means of TANSO-CAI L2 cloud flag product.

Data screening should be done through the selection and use of individual quality criteria and/or their use in combination. Described in the following sections is a description of data screening through the use of single criterion, followed by a contrasting summary discussion of the comprehensive data quality evaluation by means of a combination of criteria.

3.1 L1 Quality Evaluation

Multiple data-quality flags, primarily relevant to the satellite and onboard instrument operations, are defined and stored in the TANSO-FTS L1B data product. The TANSO-FTS L1B data product should be evaluated and filtered by referring to those L1 data-quality flags.

Listed below are the thirteen data-quality flags and the associated evaluation criteria. The criteria are meant to define low quality. Thus the TANSO-FTS SWIR data should be judged to be of low quality and is subject to being filtered out if it cannot satisfy at least one of the following criteria.

Following items are stored under the "/exposureAttribute/pointAttribute/".

(1) geometricInfo/CenterLat is flagged as invalid (−8192)

(2) geometricInfo/CenterLon is flagged as invalid (−8192)

(3) RadiometricCorrectionInfo/ELU_StableFlag is set as unstable (0)

(4) RadiometricCorrectionInfo/pointingStableFlag is not set as stable, namely (0)

(5) pointingMirrorStableFlag is set as invalid (0), or as undeterminable data quality (2)

(6) RadiometricCorrectionInfo/AT_Ang_Error_Flag is set as (1), meaning the error is greater than ±0.1 degrees

(7) RadiometricCorrectionInfo/CT_Ang_Error_Flag is set as (1), meaning the error is greater than ±0.1 degrees

(8) satellite/satelliteAttitudeStabilityStatusFlag is set as (1) meaning satellite attitude control is underway, (2) meaning satellite control is possibly underway, or (3) undeterminable on satellite attitude control operations

(9) satellite/yawSteeringFlag is set to indicate that yaw-steering has not been conducted (0)

(10) RadiometricCorrectionInfo/CT_obsPoints is set as (0) meaning that other than electric calibration being conducted, or other than (10) for a target observation mode, day (SPOD)
(11) missingFlag is set as (1) meaning that partially missing data, or (2) meaning that totally missing data of at least one of the six SWIR bands
(12) RadiometricCorrectionInfo/ZPD_SaturationFlag_SWIR is set as other than (0) meaning that at least one of six SWIR bands being saturated
(13) RadiometricCorrectionInfo/spikeNoiseFlag_SWIR is set as other than (0 (until the TANSO-FTS SWIR L2 V02.21) or 0, 3, 4, and 5 (since V02.31)) meaning that at least one of the six SWIR bands has a spike noise

3.2 Solar Zenith Angle Determination

If the solar zenith angle is greater than or equal to 70 degrees, the observations by the TANSO-FTS should be judged to yield low quality data. The reasoning behind this determination lies in that an assumption of a plane-parallel atmosphere made in the radiative transfer calculation. This approximation approach yields about 1% error in an optical path if the solar zenith angle is 72 degrees. In order to not exceed an error of 1%, the threshold value has been set at 70 degrees.

3.3 Ground Surface Roughness Determination

A ground surface roughness determination should be performed to select those FTS observation data points that have a lesser degree of ground surface roughness. In the retrieval process, the *a priori* value of the atmospheric pressure at ground level is determined based on the average elevation of the FOV-area and/or the GPV's pressure values. Hence, a large deviation in the elevation of an observation point caused by a pointing error of the onboard IMC might affect the retrieval data quality.

During the evaluation of an average elevation of the FOV-area, one should consider that the effect of a small vibration of IMC, thereby resulting pointing direction error around the stable pointing, effectively enlarges the FOV-area as depicted in Figure 3.3-1 (left). The theoretical FOV-area of the TANSO-FTS is represented by the intersection of a cone of 7.9×10^{-3} steradian and the earth's surface. In order to incorporate the effect of a small vibration of the IMC, and thereby integrating the effective pointing directions of the FTS, an enlarged cone of $(7.9 + 1.0) \times 10^{-3}$ steradian should be used, and the average elevation data should be obtained using the digital elevation model ASTER GDEM/GTOPO30.

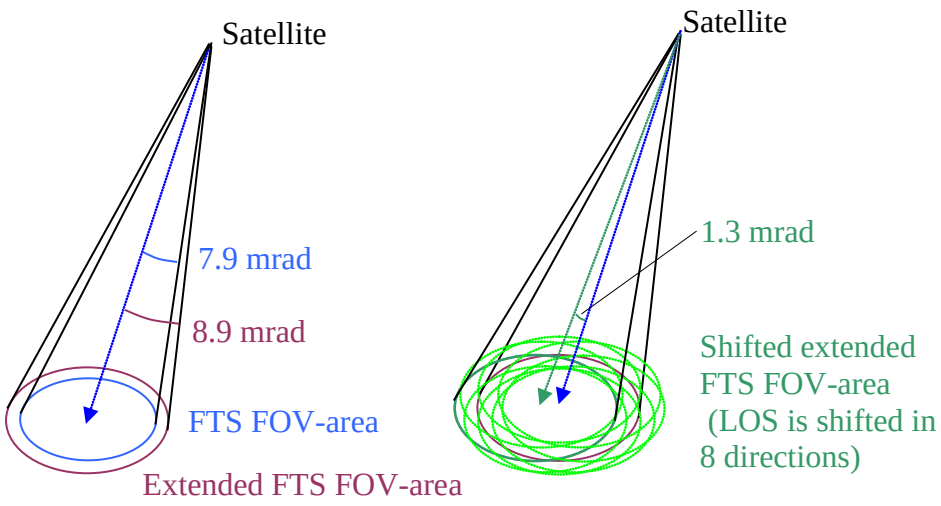

Figure 3.3-1: The conceptual diagram of FTS FOV-area and extended FOV-area to be used for the Ground Surface Roughness Determination.

Furthermore, the effect of an error in the stable IMC-based pointing vector should be taken into consideration as described below. Namely, this error effect is incorporated by considering eight FTS LOS's, which are shifted by 1.3×10^{-3} radian around the center LOS, as shown in Figure 3.3-1 (right). This small shift is projected to be at about a 1 km shift on the earth's surface near a sub-satellite point. Then an average elevation should be obtained in each of the extended FOV-areas.

In the ground surface roughness determination, the observation data that satisfies following criteria should pass the quality test.

(1) The difference between the average elevation of the center FOV-area and the average elevation of each of the eight extended FOV-areas should be 25 m or less.

3.4 CAI Cloud Determination

In this process, the existence of clouds in the observation path is determined by means of the TANSO-CAI L2 cloud flag product. Clouds along the observation path of the TANSO-FTS as well as a sun ray's path through the atmosphere (hereafter abbreviated as "observation and sun ray paths"; see Figure 3.4-1) will cause an error in the determination of the atmospheric column amount. For this reason, it is necessary to filter observation data as a function of the existence of clouds along the observation and sun ray paths.

As shown in Figure 3.4-2, the observation and sun ray paths consist of two hypothetical cylinders along the sun's ray path and the observation path. The tropopause altitude is computed based on the temperature gradient obtained from the GPV's temperature vertical profile. Namely, the atmospheric temperature-decrease gradient is

2 K or less per 1 km within a range of the geo-potential height at 500 hPa to 50 hPa. If the above gradient condition continues 1 km or more in the higher altitudes, then the lowest altitude that satisfies such a condition is defined as the tropopause altitude. In an event that the above gradient condition is not satisfied, the 50 hPa geo-potential altitude is defined as the tropopause altitude.

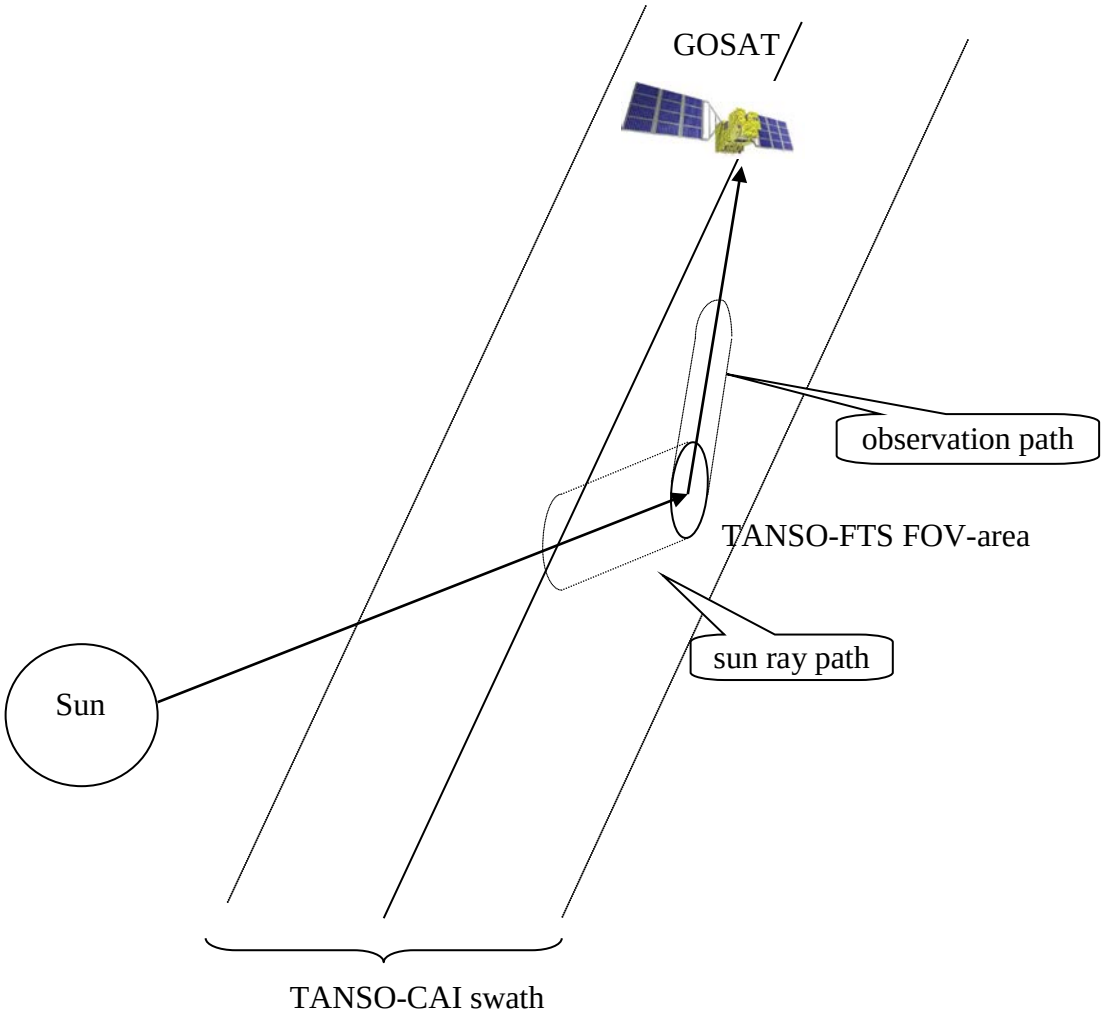

Figure 3.4-1: The conceptual diagram for the observation and sun ray paths.

Three projected regions are considered: the sun-side projected region, the satellite-side projected region and the FOV-area. All of these have internal and external regions, thereby creating a total of six regions altogether. Each internal region is extended by 1 x 10^{-3} steradian of the theoretical FTS FOV-area; each external region has an internal boundary overlapped with the corresponding internal region and an external boundary, which extended by 5 x 10^{-3} steradian of the theoretical FTS FOV-area. Even if those regions are partially overlapped, each region is treated as an independent entity when the existence of clouds is being considered. Each region is approximated by 36

polygons. As shown in Figure 3.4-3, those regions are partially overlapped. Such regions are classified to a specific region according to the priority as shown below;

(AREA 1) internal FTS FOV-area
(AREA 2) external FTS FOV-area
(AREA 3) internal sun-projected region
(AREA 4) external sun-projected region
(AREA 5) internal satellite-projected region
(AREA 6) external satellite-projected region

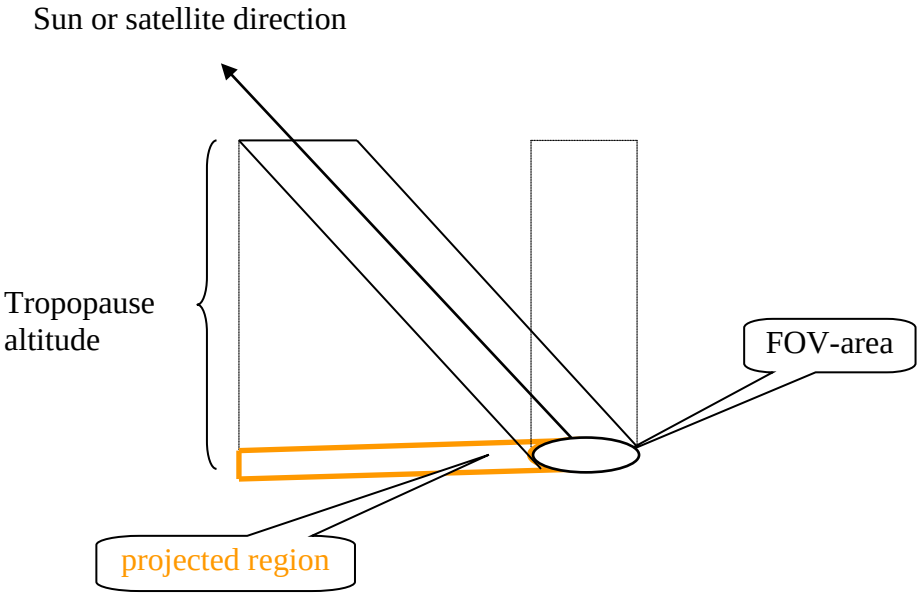

Figure 3.4-2: The conceptual diagram for the projected region.

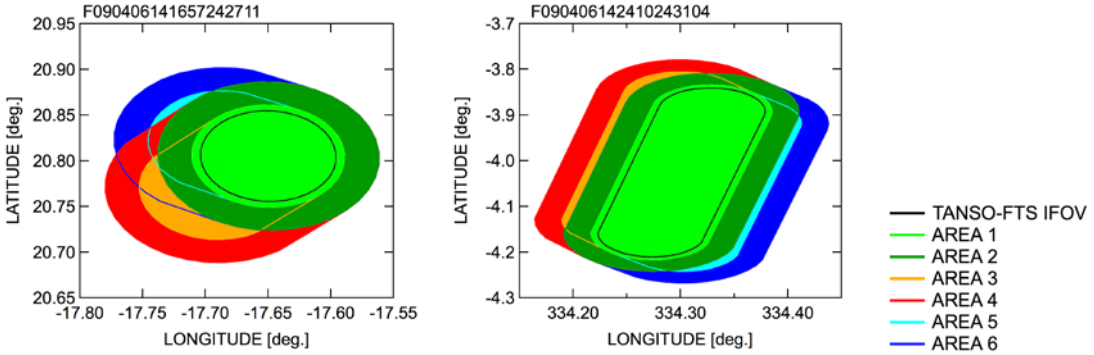

Figure 3.4-3: Definition of projected region. The figure shows internal and external TANSO-FTS FOV-areas, the internal and external sun-projected regions, and the internal and external satellite-projected regions.

The cloud coverage for each region is computed based on the integrated clear-sky confidence level for each pixel of the TANSO-CAI. In the TANSO-FTS SWIR L2

processing, the integrated clear-sky confidence level, represented by a floating value between 0.0 and 1.0, is categorized as shown below:

(1) 0.00 ~ 0.15 for confident cloudy
(2) 0.15 ~ 0.33 for probably cloudy
(3) 0.33 ~ 1.00 for confident clear

Any TANSO-CAI pixel that has a value between 0.00 ~ 0.33 is treated as cloudy-pixel. Then a cloud coverage ratio is computed in terms of the ratio of the number of cloudy-pixels in the region of interest and the total number of the TANSO-CAI's effective pixels in the same region.

Until the TANSO-FTS SWIR L2 V02.11, it is determined to be a cloud free if the TANSO-FTS data has a cloud coverage ratio that satisfies all of the criteria listed below. While only the criteria (1) has been used since V02.20.

(1) 0% in the TANSO-FTS internal FOV-area
(2) 10% or less in the sun-projected region and
(3) 8% or less in the satellite-projected region

3.5 CAI Coherent Test

The average radiance and its standard deviation of the TANSO-CAI in the extended FOV region of the TANSO-FTS are utilized for another cloud detection method. The TANSO-CAI L2 cloud flag product explained earlier may not be sensitive enough to detect the existence of clouds of a sub-pixel size in ocean observations. In order to cope with this difficulty, cloud-free data is selected by paying special attention to the standard deviation of the TANSO-CAI radiance since it will increase its value in a distributed cloud condition. The application of this determination approach should be limited to ocean observations since the standard deviation of the TANSO-CAI radiance is known to change significantly due to ground albedo variations in addition to cloud coverage.

Figure 3.5-1 shows the standard deviation versus the average radiance of TANSO-CAI Band 1 within the TANSO-FTS's FOV, where "CAI cloud determination test" described in Section 3.4 indicated as cloud-free for the TANSO-FTS observation over an ocean surface from April 23 through April 25, 2009. As shown in this figure, a recurrent segregation line has been defined by assigning colors for cloud-free and cloudy sky conditions, respectively, thereby permitting visual recognition of the conditions. This approach is currently utilized in the TANSO-CAI Band 1 only in the TANSO-FTS SWIR L2 algorithm. The parameters associated with the recurrent segregation line were determined only for the month of April 2009, and need further study of possible seasonal variations. Also, it should be noted that these parameters must be updated every time the TANSO-CAI instrument's radiance is re-calibrated.

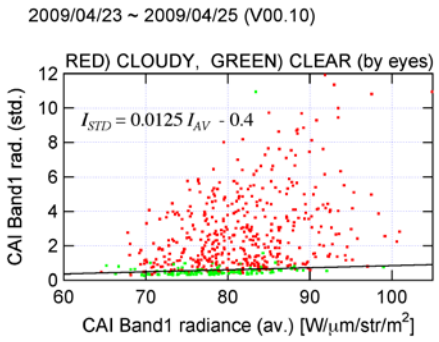

Figure 3.5-1: The relationship between the average and standard deviation of TANSO-CAI Band 1 radiance within the TANSO-FTS FOV. The assigned colors indicate eye-balled sky condition in the TANSO-FTS's FOV: red for cloudy and green for clear-sky.

3.6 TIR Cloud Determination

TIR-based cloud determination is not conducted in the TANSO-FTS SWIR L2 V02.xx processing.

3.7 2-micron Band Scattering Material Determination

The TANSO-FTS Band 3 contains many strong water vapor absorption bands. If there is no scattering by clouds and/or aerosols, almost no radiance at the strong water vapor absorption band region is observed by the TANSO-FTS due to the saturation of the water vapor absorption. In general, because the water vapor mixing ratio decreases with height, reflected radiance due to scattering may not be completely absorbed by water vapor and reaches the TANSO-FTS, which depends on the partial column amount of water vapor above the scattering. Therefore, if non-zero radiance is observed at the strong water vapor absorption band, it is the evidence of the existence of some scattering within the TANSO-FTS IFOV. By utilizing this characteristic, a cloud-free observation is identified. The 2-micron band scattering material determination method explained below has the benefit of allowing the detection of the optically thin cirrus clouds and elevated aerosols, neither of which can be achieved through the CAI cloud determination test (Section 3.4).

Eleven water vapor absorptive channels are preliminary selected within a wavenumber range from 5150 cm^{-1} to 5200 cm^{-1}. An average value of the radiance that normalized with the noise level is computed. If the average normalized radiance value is greater than unity, it can be determined that cirrus clouds and/or aerosol are/is present at extremely high altitudes.

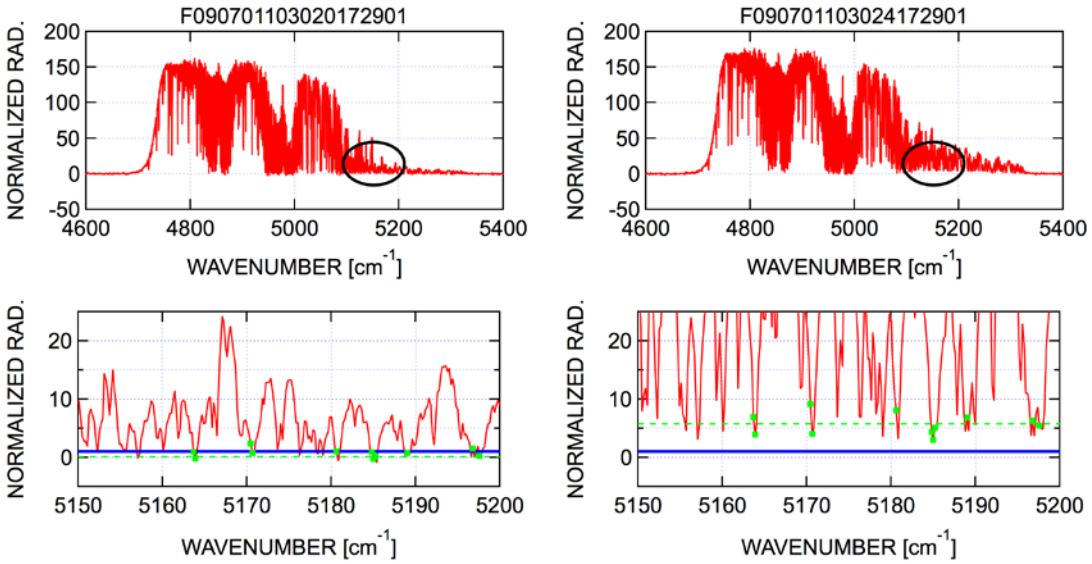

Figure 3.7-1: Examples of the TANSO-FTS Band 3 radiance of P-polarization components normalized with its noise level. The left figure shows a clear-sky observation, while the right figure depicts a cloudy sky observation. The green points are the selected channels, the green line is their average radiance value, and the blue line indicates the threshold value (1.0).

3.8 Elevated Aerosol Determination

It could happen that the atmospheric column amount is determined to be less than what it should be due to the effects of the path radiance components having short optical path lengths, particularly when aerosol is scattered in extremely high altitudes. The elevated aerosol determination sub-process resolves this problem by means of aerosol extinction coefficients computed through the aerosol transport model SPRINTARS.

The aerosol extinction coefficients computed by SPRINTARS on a global scale are with about 2.8 degrees horizontal resolution and 20 layers (σ-level) in the vertical direction with a time resolution of 6 hours. The aerosol extinction coefficients have also been computed at almost both edge wavelengths of each band of the TANSO-FTS SWIR, namely at the wavelengths of 0.74, 0.78, 1.49, 1.81, 1.85 and 2.17 μm.

First, a similar spatiotemporal correction of observation data as explained in Section 2.3 should be performed by a linear interpolation in the horizontal direction and along the time axis. Then, the optical thickness in the TANSO–FTS Band-2 should be computed via a vertical integration of the extinction coefficients. The extinction coefficient in Band-2, denoted as $\sigma_{band\,2,\,i}$ is given by

$$\sigma_{band\,2,\,i} = \frac{\sigma_{1.49,\,i} + \sigma_{1.81,\,i}}{2}. \qquad (3.8\text{-}1)$$

Namely, it is represented by an average of extinction coefficients $\sigma_{1.49,\,i}$ and $\sigma_{1.81,\,i}$ at the both ends of the band. The second subscript, i, represents the layer in the vertical direction (where $i = 1, 2, \cdots, 20$). Next, compute the two optical thicknesses $\tau_{band\,2\,all}$; the total optical thickness from the earth's surface to the top of the atmosphere; and $\tau_{band\,2\,part}$; the optical thickness from 600 hPa pressure level (about 4 km in altitude) to the top of the atmosphere:

$$\tau_{band\,2\,all} = \sum_{i=1}^{20} \sigma_{band\,2,\,i}\ , \tag{3.8-2}$$

$$\tau_{band\,2\,part} = \tau_{band\,2,\,ista} + \sum_{i=ista+1}^{20} \sigma_{band\,2,\,i}\ , \tag{3.8-3}$$

$$\tau_{band\,2,\,ista} = \left(\frac{\log(600) - \log(p_s\ \sigma_{ista+1})}{\log(\sigma_{ista}) - \log(\sigma_{ista+1})} \right) \sigma_{band\,2,\,ista}\ , \tag{3.8-4}$$

where the suffix *ista* stands for the layer containing the 600 hPa level, σ_i means the i-th component along the σ-axis (vertical direction), $\tau_{band\,2,\,ista}$ means the optical thickness integrated over the extinction coefficients from the 600 hPa pressure level, p_s represents the surface pressure derived from the GPV database and average elevation of the FOV-area. Note that σ with two suffices denotes an extinction coefficient.

Based on the above computation, if $\tau_{band\,2\,all}$ is greater than a threshold value of 0.05 and if $\tau_{band\,2\,part}$ is 20 % or more of $\tau_{band\,2\,all}$, then the existence of aerosol in extremely high altitudes can be concluded.

3.9 Spectrum Quality Determination

The TANSO-FTS L1B contains three types of data-quality flags for measured interferograms (i.e., missing, saturation, and spike noise), but there is no flag for spectra. Because outlier spectra with no bad L1B data-quality flag are often seen, spectrum-based quality check is newly added.

Averages and standard deviations of the real (R_{AVG} and R_{STD}) and imaginary (I_{AVG} and I_{STD}) part of the out-of-band spectra are calculated. Here, low-wavenumber side and high-wavenumber side are treated separately (wavenumber range is shown in Table 3.9-1). If the out-of-band statistics for both sides satisfy all of the criteria listed below (threshold values T are shown in Table 3.9-2), then it is determined to be a good spectrum.

(TEST 1) $|\log_{10}(R_{STD} / I_{STD})| <= T_{UL1}$
(TEST 2) $\log_{10}(R_{AVG}) <= T_{UL2}$
(TEST 3) $T_{LL3} <= \log_{10}(R_{STD}) <= T_{UL3}$
(TEST 4) $|I_{AVG}| <= T_{UL4}$

Table 3.9-1: Wavenumber range used for the spectrum quality determination.

	low-wavenumber side	high-wavenumber side
Band 1	12450 - 12700 cm^{-1}	13450 - 13650 cm^{-1}
Band 2	5350 - 5500 cm^{-1}	6700 - 6850 cm^{-1}
Band 3	4450 - 4600 cm^{-1}	5450 - 5650 cm^{-1}

Table 3.9-2: Threshold values for the spectrum quality determination. Upper and lower values are threshold values for low- and high-wavenumber sides, respectively. These values are evaluated by using the TANSO-FTS L1B V141.141.

		Band 1P	Band 1S	Band 2P	Band 2S	Band 3P	Band 3S
T_{UL1} (Gain-H, M)		0.055 0.060	0.055 0.060	0.078 0.077	0.076 0.079	0.080 0.062	0.073 0.063
T_{UL2}	(Gain-H)	-4.72 -4.90	-4.88 -4.90	-5.26 -5.36	-5.28 -5.32	-5.42 -5.42	-5.42 -5.16
	(Gain-M)	-5.42 -5.36	-5.36 -5.34	-5.68 -5.60	-5.60 -5.62	-5.90 -5.98	-5.92 -6.00
T_{LL3}	(Gain-H)	-5.22 -5.18	-4.90 -4.86	-5.52 -5.48	-5.42 -5.36	-5.42 -5.38	-5.26 -5.22
	(Gain-M)	-5.42 -5.44	-5.20 -5.20	-5.64 -5.62	-5.62 -5.60	-5.62 -5.62	-5.58 -5.56
T_{UL3}	(Gain-H)	-4.78 -4.76	-4.60 -4.58	-5.00 -4.98	-4.92 -4.88	-4.90 -4.90	-4.78 -4.68
	(Gain-M)	-5.06 -5.06	-4.94 -4.92	-5.32 -5.26	-5.22 -5.20	-5.32 -5.30	-5.16 -5.14
T_{UL4}	(Gain-H)	3.95e-7 4.25e-7	6.75e-7 5.00e-7	4.25e-7 8.65e-7	4.85e-7 4.65e-7	4.75e-7 5.00e-7	6.50e-7 5.85e-7
	(Gain-M)	1.55e-7 1.90e-7	2.05e-7 1.70e-7	2.15e-7 2.60e-7	2.50e-7 2.50e-7	2.30e-7 2.75e-7	2.70e-7 2.20e-7

3.10 Combined Use of Data Screening Criteria

A comprehensive determination of the suitability of the data for the TANSO-FTS SWIR L2 algorithm can be done through the combined use of the individual data screening determinations discussed in the previous Sections.

Only data that has passed all seven data screening criteria (i.e., (1) L1 Quality Evaluation (Section 3.1), (2) Solar Zenith Angle Determination (Section 3.2), (3) Ground Surface Roughness Determination (Section 3.3), (4) CAI Cloud Determination (Section 3.4), (5) CAI Coherent Determination (Section 3.5), (6) 2-micron Based Scattering Material Determination (Section 3.7), and (7) Spectrum Quality Determination (Section 3.9)) should be judged suitable until the TANSO-FTS SWIR L2 V02.11 Retrieval Processing. After V02.20, the Ground Surface Roughness Determination (Section 3.3) has been removed from the data screening criteria.

4. Forward Model

4.1 Radiative Transfer Equation

A general expression of the radiative transfer for a plane-parallel atmosphere is given by the following equation if thermal radiation can be ignored:

$$\mu \frac{dI(\tau,\mu,\phi)}{d\tau} = -I(\tau,\mu,\phi) + \frac{\omega}{4\pi}\int_{-1}^{1} d\mu' \int_{0}^{2\pi} d\phi' P(\tau,\mu,\phi;\mu',\phi') I(\tau,\mu',\phi') + \frac{\omega}{4\pi} P(\tau,\mu,\phi;\mu_0,\phi_0) F_0 \exp(-\tau/\mu_0) \quad , (4.1\text{-}1)$$

where I is a diffused component of the radiance, τ stands for the optical thickness measured from the top of the atmosphere, μ denotes directional cosine, and ϕ represents an azimuth angle. Also, ω is the single-scattering albedo, and P is the scattering phase function, F_0 stands for the solar irradiance, μ_0 and ϕ_0 denote directional cosine and azimuth angles of the sun. The radiance must satisfy the boundary conditions presented below:

$$I(0,\mu,\phi) = 0 \text{ for } 0 \le \mu \le 1, \quad (4.1\text{-}2)$$

$$I(\tau_0,\mu,\phi) = \frac{1}{\pi}\int_{0}^{1} d\mu' \int_{0}^{2\pi} d\phi' \mu' A(\mu,\phi;\mu',\phi') I(\tau_0,\mu',\phi') + \frac{\mu_0}{\pi} A(\mu,\phi;\mu_0,\phi_0) F_0 \exp(-\tau_0/\mu_0) \quad \text{for } -1 \le \mu \le 0 , (4.1\text{-}3)$$

where τ_0 is the total optical thickness of the atmosphere, and A stands for the bi-directional reflectance.

In the TANSO-FTS SWIR L2 algorithm, the discrete ordinance method [*Siewert*, 2000] is used in order to solve Eq. (4.1-1). In order to allow an accurate computation of the radiance even for a small number of streams, the TMS method [*Nakajima and Tanaka*, 1988] is used. In the current version of the TANSO-FTS SWIR L2 algorithm, the number of streams is set at eight for a hemisphere.

A Lambert surface is assumed for the land region of the earth. Then the bi-directional reflectance is determined as being equal to the non-directional surface albedo. For the reflectivity of an ocean region of the earth, the ocean reflection matrix [*Deuze et al.*, 1989] is used. The distribution function for wave slope of the facets refers to the model presented by *Cox and Munk* [1954].

4.2 High Speed Computation of Radiative Transfer

In the spectrum computations of a forward model, it becomes necessary to calculate monochromatic radiances for several tens of thousands of points within the applicable wavenumber range. For this reason, a radiative transfer calculation that

includes multiple scatterings becomes an extremely time consuming process. However, in the TANSO-FTS SWIR L2 processing, this vast computational load is significantly reduced by means of a modified / improved algorithm of Duan's high-speed computational approach for the radiative transfer equation [*Duan et al*, 2005]. In this Section, an overview of Duan's approach is provided together with the new, specialized improvements developed and applied to the TANSO-FTS SWIR L2 processing.

(a) Overview of the highspeed computation of radiative transfer by *Duan et al.* [2005]

The key idea for the computational improvements discussed below lies in grouping the atmospheric parameters into several categories based on the redundancy or similarity to one another. For instance, the optical thickness, denoted by k, due to gaseous absorption is a good candidate for parameter grouping (refer to Figure 4.2-1).

For the scattering atmosphere, the typical altitudes of scattering in the atmosphere may be another beneficiary of parameter grouping. Let the scattering component be denoted as k_{sca} as shown in Figure 4.2-1. Then the typical altitude of scattering is defined as that altitude where the optical thickness measured from the top of the atmosphere becomes $k'_{sca} = \min(1, 1/2\,k_{sca})$. Hence the ratio $k'/k \equiv \xi$ can be organized in grouping parameters. Figure 4.2-2 depicts k and ξ values corresponding to each wavenumber in the O_2 A-band.

It is possible to separate the observed radiance I into the single scattering component I^{ss} and multiple scattering component I^{ms}.

Namely

$$I = F_0\left(I^{ss} + I^{ms}\right). \qquad (4.2\text{-}1)$$

I^{ss} can be calculated exactly without approximation for each wavenumber. On the other hand, I^{ms} can be parameterized as follows:

$$I^{ms}(k, \xi) = e^{g(k) - \beta(k)[\ln \xi - \ln \xi_0(k)]}, \qquad (4.2\text{-}2)$$

where $\xi_0(k)$ is the average value of ξ at the optical thickness value k. For a high speed computation of a radiative transfer, first (a) perform multiple scattering computations for an appropriately chosen reference wavenumber; then (b) construct tables for $g(k)$, $\beta(k)$ and $\xi_0(k)$; and finally (c) compute the values for k and k' and the multiple scattering components for all the wavenumbers of interest through table-interpolations.

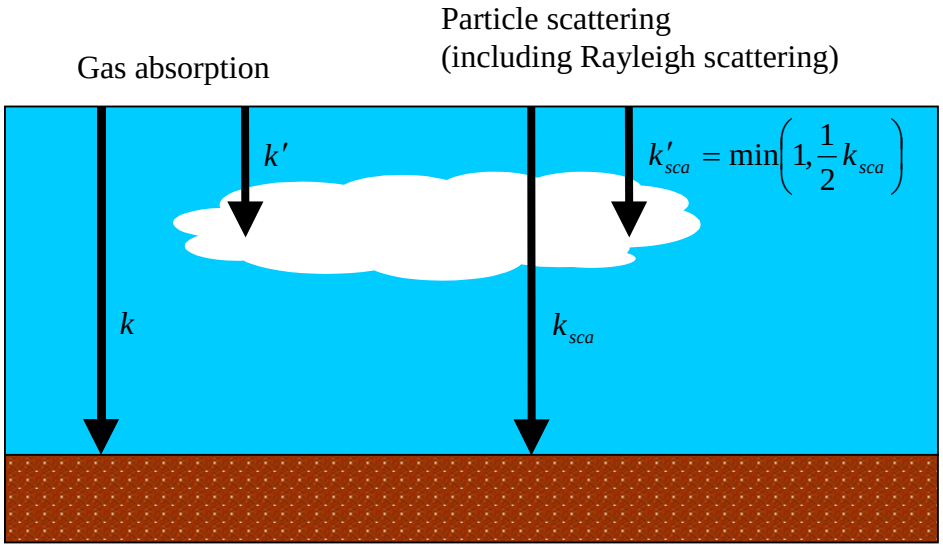

Figure 4.2-1: Grouping parameters.

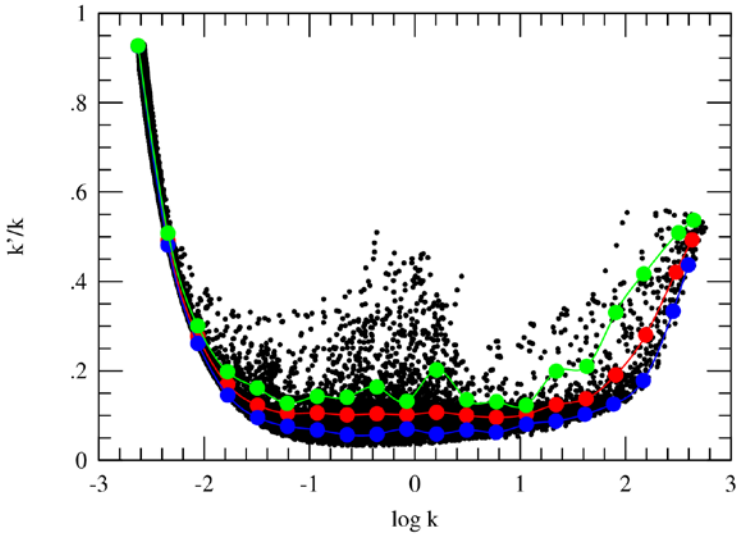

Figure 4.2-2: Grouping parameters in O_2 A-band.

(b) Dependency of Cirrus-cloud and Aerosol Scattering Properties on Wavenumbers

Duan's high speed computational method [*Duan et al.* 2005] briefly explained above assumes that the scattering properties of cirrus clouds, aerosols or the surface reflectivity does not vary within the wavenumber bands of interest. Hence Duan's method without any modifications may not be directly applicable to the TANSO-FTS SWIR L2 processing. In order to cope with this difficulty, the dependency of cirrus cloud and aerosol scattering properties on wavenumbers within a band is addressed in the following manner. First, assume that the optical properties of cirrus clouds and aerosols within a band are invariable with properties at the lower bound. With this

assumption in mind, the interpolation tables for $g(k)$, $\beta(k)$ and $\xi_0(k)$ should be constructed. Second, a set of similar tables should be constructed with the assumption that the scattering properties of cirrus clouds and aerosols are invariable with the properties at the upper bound. Thus, two sets of tables are completed for cirrus clouds and aerosols corresponding to the upper and lower bound property assumptions. Third, the radiance for an arbitrary wavenumber should then be computed by linearly interpolating the values of these two tables.

(c) Dependency of Surface Reflectivity on Wavenumbers

The dependency of surface reflectivity on wavenumbers can be handled as shown in the following discussion. First, assume that the ground surface is a Lambert surface with reflectivity α. Let the radiance at the upper bound of the atmosphere be I_0 when $\alpha = 0$. Then, according to *Liou* [2002], the radiance I_α at the top of the atmosphere is given by the following expression, where an arbitrary reflectivity α at the ground surface is assumed.

$$I_\alpha = I_0 + (I_m - I_0)\frac{1-\bar{r}\alpha_m}{\alpha_m}\frac{\alpha}{1-\bar{r}\alpha}. \qquad (4.2\text{-}3)$$

In this equation, α_m is the average reflectivity of the ground surface, and I_m denotes the corresponding satellite-observed radiance. Also $\bar{r}$ represents the spherical albedo. Once the multiple scattering I_0 and I_m tables are created by means of Eq. (4.2-2), the multiple scattering I_α components can be computed through Eq. (4.2-3). During this computation, spherical albedo values will become necessary, which can be handled conveniently by creating a table prior to the computation, in a similar way as in the generation of I_0 and I_m tables.

Through the technique briefly explained above, the number of reference points in the tables amounts to only about 300, where the effects of the wavenumber dependency of cirrus clouds and aerosols as well as the ground surface albedo are fully taken into considerations. Thus a dramatic 100-fold increase in computational speed can be accomplished for the voluminous computations of Eq. (4.2-2) with respect to the several tens-of-thousands of wavenumber points in a multiple scattering.

(d) Computation of Jacobian

From Eq. (4.2-2), it is possible to derive rapidly the Jacobian associated with the optical thickness due to gas absorption at the top of the atmosphere. First, divide the atmosphere into as many as *n* layers. Then, with the optical thickness τ_l of the l-th layer from the upper most layer, it holds that

$$k = \sum_{l=1}^{n} \tau_l \,. \tag{4.2-4}$$

Since

$$\frac{dk}{d\tau_l} = 1, \tag{4.2-5}$$

it follows that

$$\frac{dI^{ms}}{d\tau_l} = \frac{I^{ms}}{k}\left[\frac{dg}{d\ln k} - \beta\left(\frac{1}{\xi}\frac{dk'}{d\tau_l} - 1 - \frac{d\ln\xi_0}{d\ln k}\right)\right]. \tag{4.2-6}$$

In this derivation, the dependency of β on k is ignored.

If it is assumed that the optical thickness due to gas absorption at l_{sca} is k', then such quantity can analytically be computed as

$$k' = \sum_{l=1}^{l_{sca}-1} \tau_l + s_{sca}\tau_{l_{sca}} \,. \tag{4.2-7}$$

Since $0 \le s_{sca} \le 1$, it follows that

$$\frac{dk'}{d\tau_l} = \begin{cases} 1 & 1 \le l \le l_{sca} - 1 \\ s_{sca} & l = l_{sca} \\ 0 & l_{sca} + 1 \le l \le n \end{cases} \,. \tag{4.2-8}$$

For the tables of g and ξ_0 expressed as functions of $\ln k$, compute the derivative values together with the interpolated values. Then, from Eqs. (4.2-6) and (4.2-8), the Jacobians can be computed.

Also the Jacobian for the radiance at the top of the atmosphere with respect to the surface albedo can readily be obtained by differentiating Eq. (4.2-3) with respect to α.

$$\frac{\partial I_\alpha}{\partial \alpha} = \left(I_m - I_0\right)\frac{1 - \bar{r}\alpha_m}{\alpha_m}\frac{1}{\left(1 - \bar{r}\alpha\right)^2} \,. \tag{4.2-9}$$

4.3 Gas Absorption Coefficient

In the TANSO-FTS SWIR L2 processing, the optical thickness due to gas absorption is computed by interpolating the gas absorption coefficient in given pressure

and temperature from the pre-calculated Look-Up-Table (LUT). LUT was prepared based on *Tran et al.* [2006] and *Tran and Hartmann* [2008] for oxygen including the line-mixing and the collision-induced absorption effect. For carbon dioxide, the line-mixing effect is considered according to *Lamouroux et al.* [2010], and the HITRAN 2008 database [*Rothman et al.*, 2009] with Voigt line shape is used for methane and water vapor. LUT has 70 pressure grids from 0.06 hPa to 1040 hPa and 10 temperature grids for each pressure. The full-width of the temperature grid is 90 K (i.e., 10 K intervals) and its median temperature values for each pressure grid is set to that of the U.S. standard atmospheric model. The continuous absorption of water vapor is calculated by the MT CKD model by *Clough et al.* [2005].

4.4 Cirrus Cloud

In the TANSO-FTS SWIR L2 processing, the gaseous column amounts of carbon dioxide (CO_2), methane (CH_4), and water vapor (H_2O) should be estimated with a compensation of the effect of cirrus clouds if the clouds are optically thin within the FTS's FOV. However, at the time of this writing, the TANSO-FTS SWIR L2 algorithm assumes cloud-free condition (see Sections 3.7 and 3.10), and the following part in this section is not used in the actual data processing.

In the forward model, scattering properties of cirrus cloud are characterized by four parameters, which include: optical thickness, effective particulate size, cloud-top pressure and base pressure. The cirrus cloud is modeled as a logarithmical normal distribution of spherical ice particles

$$n(r) = \frac{N}{\sqrt{2\pi}\sigma r}\exp\left[-\frac{\left(\ln r - \ln r_0\right)^2}{2\sigma^2}\right], \tag{4.4-1}$$

where r is the radius of the particle, N refers to the total number density and σ is the standard deviation constant. The relationship between the mode radius r_0 and the effective radius r_e is given by

$$r_e = r_0 \exp\left(5\sigma^2/2\right). \tag{4.4-2}$$

The standard deviation is; $\sigma = 0.35$.

The volume extinction coefficient, volume scattering coefficient and scattering phase function of cirrus clouds should be computed by means of the Mie theory with the complex refractive index of ice by *Warren* [1984]. One should compute the scattering properties of cirrus clouds at the maximum, minimum and middle values of the wavenumber ranges of the TANSO-FTS SWIR L2 processing. Then the volume extinction coefficient $C^{ext}(v)$ for wavenumber v is linearly interpolated based on the wavenumber range $v_1 \le v \le v_2$ as

$$C^{ext}(\nu) = C_1^{ext} + \frac{C_2^{ext} - C_1^{ext}}{\nu_2 - \nu_1}(\nu - \nu_1), \qquad (4.4\text{-}3)$$

where C_1^{ext} and C_2^{ext} are the volume extinction coefficients corresponding to minimum and maximum wavenumbers, respectively, used in the TANSO-FTS L2 process. The volume scattering coefficients should be computed in a similar way. However, the scattering phase function must be evaluated at the center value of the wavenumber range used in the retrieval process.

A cirrus cloud is assumed that it has no vertical structure, and that it is uniform between the cloud-top pressure and base pressure.

4.5 Aerosol

The characteristics of aerosols are known to be diverse. For this reason, GOSAT carries the onboard instrument TANSO-CAI in order to measure aerosol characteristics. Also, utilize the simulation results by the aerosol transport model SPRINTARS [*Takemura et al.*, 2000] for such areas where TANSO-CAI aerosol information is not available (e.g., an ocean sunglint region).

In the TANSO-FTS SWIR L2 process, the types and size distributions of aerosols are provided as externally input data for each TANSO-FTS observation point. The scattering properties of aerosols (optical thickness, single-scattering albedo, and scattering phase function) should be evaluated (a) first, at the maximum and the minimum values of the wavenumbers used in the retrieval process and then (b) the scattering properties at an arbitrary wavenumber of interest should be linearly interpolated from the values of two points obtained in step (a) above.

In the TANSO-FTS SWIR L2 process, two types of aerosol (fine-mode and coarse-mode) can be handled. Layer-averaged aerosol scattering properties should be calculated as a sum and/or weighted mean of the scattering properties of fine- and coarse-mode aerosols.

4.6 Rayleigh Scattering

The Rayleigh scattering cross section for a single air molecule is given by the following equation [*Bucholz*, 1995].

$$\sigma(\lambda) = \frac{24\pi^3\left(n_s^2 - 1\right)^2}{\lambda^4 N_s^2\left(n_s^2 + 2\right)^2}\left(\frac{6 + 3\rho_n}{6 - 7\rho_n}\right), \qquad (4.6\text{-}1)$$

where λ is a wavelength, n_s refers to the refractive index of standard air, N_s stands for the number density of standard air (2.54743 x 10^{19} cm^{-3}) and ρ_n denotes depolarization factor. The “standard air” is dry air, which includes carbon dioxide (CO_2) in the amount of 0.03% in the pressure of 1013.25 hPa and the temperature of 15 degrees C.

A set of fitting formulas for the refractive index of standard air is given as a function of wavelength, as shown below [*Bucholtz*, 1995; *Peck and Reeder*, 1972].

$$(n_s - 1)\times 10^8 = \frac{5791814}{238.0185 - (1/\lambda)^2} + \frac{167929}{57.362 - (1/\lambda)^2} \text{ for } \lambda > 0.23\mu\text{m}, \quad (4.6\text{-}2)$$

$$(n_s - 1)\times 10^8 = 8060.51 + \frac{2480990}{132.274 - (1/\lambda)^2} + \frac{17455.7}{39.32957 - (1/\lambda)^2} \text{ for } \lambda \le 0.23\mu\text{m}. \quad (4.6\text{-}3)$$

Although *Bucholtz* [1995] showed the depolarization factor in Table 1 of his paper, the entity values are not a smooth function of wavelengths. For this reason, it should be computed from the following fitting formula:

$$\rho_n = 1.007482\times 10^{-2} + \frac{7.990914\times 10^{-1}}{47.48717 - (1/\lambda)^2}. \quad (4.6\text{-}4)$$

Incidentally the data source of the Table 1 of Bucholtz' paper is originated in Bates [*Bates*, 1984]. The unit of wavelengths λ in Eqs. (4.6-2), (4.6-3) and (4.6-4) is μm.

4.7 Solar Irradiance Spectra

In the TANSO-FTS SWIR L2 processing, solar irradiance database provided by Dr. G.C. Toon is used. This database was based on balloon FTS spectra from 30+km altitude in addition to ground-based Kitt Peak and TCCON spectra.

5. Inverse Model

5.1 MAP Analysis Method

According to the MAP (Maximum *A Posteriori*) retrieval method [*Rodgers*, 2000], the optimal estimate of state vector $\mathbf{x}$ is the value of $\mathbf{x}$ that minimizes the cost function presented below:

$$J(\mathbf{x}) = [\mathbf{y} - \mathbf{F}(\mathbf{x})]^T \mathbf{S}_\varepsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x})] + (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a), \tag{5.1-1}$$

where $\mathbf{y}$ is the observed spectra, $\mathbf{F}(\mathbf{x})$ stands for the forward model, $\mathbf{x}_a$ refers to *a priori* state vector, $\mathbf{S}_a$ and $\mathbf{S}_\varepsilon$ denote covariance matrices of $\mathbf{x}_a$ and observation errors, respectively.

A Gauss-Newton method is used to obtain the solution in an iterative manner,

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \left(\mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}_i + \mathbf{S}_a^{-1}\right)^{-1} \left[\mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} (\mathbf{y} - \mathbf{F}_i) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a)\right], \tag{5.1-2}$$

where $\mathbf{F}_i = \mathbf{F}(\mathbf{x}_i)$, and $\mathbf{K}_i$ is the Jacobian matrix calculated with respect to $\mathbf{x} = \mathbf{x}_i$.

$$\mathbf{K}_i = \left.\frac{\partial \mathbf{F}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_i}. \tag{5.1-3}$$

The solution $\mathbf{x}$ must satisfy the following constrains:

$$\mathbf{x}_{\min} \le \mathbf{x} \le \mathbf{x}_{\max}. \tag{5.1-4}$$

5.2 LM Method

The Gauss-Newton iterative process tends to become unstable if the *a priori* value $\mathbf{x}_0$ is significantly different from the true value, or the forward model $\mathbf{F}(\mathbf{x})$ happens to be highly non-linear. To cope with this difficulty, the LM (Levenberg-Marquardt) method may be used in order to stabilize the iterative process and to obtain a stable solution.

$$\left(\mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} \mathbf{K}_i + \mathbf{S}_a^{-1} + \lambda \mathbf{D}^2\right)(\mathbf{x}_{i+1} - \mathbf{x}_i) = \mathbf{K}_i^T \mathbf{S}_\varepsilon^{-1} (\mathbf{y} - \mathbf{F}_i) - \mathbf{S}_a^{-1} (\mathbf{x}_i - \mathbf{x}_a), \tag{5.2-1}$$

where $\mathbf{D}$ is a diagonal matrix introduced to stabilize the iteration process. The value of the parameter λ is adjusted in each step of the iteration so that the values of the cost function decrease steadily.

Next, carry out a Cholesky decomposition process for the observation error and *a priori* covariance matrices as shown below:

$$\mathbf{S}_{\varepsilon} = \mathbf{T}_{\varepsilon}^{T}\mathbf{T}_{\varepsilon}, \tag{5.2-2}$$

$$\mathbf{S}_{a}^{-1} = \mathbf{T}_{ainv}^{T}\mathbf{T}_{ainv}, \tag{5.2-3}$$

Now, define matrix $\mathbf{A}$ as

$$\mathbf{A} = \begin{pmatrix} \mathbf{T}_{\varepsilon}^{-T}\mathbf{K}_{i} \\ \mathbf{T}_{ainv} \end{pmatrix}, \tag{5.2-4}$$

and vectors $\mathbf{b}$ and $\Delta\mathbf{x}$ as

$$\mathbf{b} = \mathbf{K}_{i}^{T}\mathbf{S}_{\varepsilon}^{-1}\left(\mathbf{y} - \mathbf{F}_{i}\right) - \mathbf{S}_{a}^{-1}\left(\mathbf{x}_{i} - \mathbf{x}_{a}\right), \tag{5.2-5}$$

$$\Delta\mathbf{x} = \mathbf{x}_{i+1} - \mathbf{x}_{i}. \tag{5.2-6}$$

Then Eq. (5.2-1) is written in a compact form as

$$\left(\mathbf{A}^{T}\mathbf{A} + \lambda\mathbf{D}^{2}\right)\Delta\mathbf{x} = \mathbf{b}. \tag{5.2-7}$$

Now, set the diagonal matrix $\mathbf{D}$ as

$$\mathbf{D}^{2} = diag\left(\mathbf{A}^{T}\mathbf{A}\right). \tag{5.2-8}$$

Then Eq. (5.2-7) becomes

$$\left[\left(\mathbf{A}\mathbf{D}^{-1}\right)^{T}\left(\mathbf{A}\mathbf{D}^{-1}\right) + \lambda\mathbf{I}\right]\mathbf{D}\Delta\mathbf{x} = \mathbf{D}^{-1}\mathbf{b}. \tag{5.2-9}$$

Now perform a singular-value decomposition of matrix $\mathbf{A}\mathbf{D}^{-1}$ as

$$\mathbf{A}\mathbf{D}^{-1} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{T}. \tag{5.2-10}$$

Then Eq. (5.2-9) becomes

$$\left(\mathbf{\Lambda}^{2} + \lambda\mathbf{I}\right)\mathbf{V}^{T}\mathbf{D}\Delta\mathbf{x} = \mathbf{V}^{T}\mathbf{D}^{-1}\mathbf{b}. \tag{5.2-11}$$

Define the norm $\left|\mathbf{D}\Delta\mathbf{x}\right|$ of the variation of the solutions. Then it is readily derived as

$$\left|\mathbf{D}\Delta\mathbf{x}\right| = \left|\left(\mathbf{\Lambda}^{2} + \lambda\mathbf{I}\right)^{-1}\mathbf{V}^{T}\mathbf{D}^{-1}\mathbf{b}\right|. \tag{5.2-12}$$

Note:

$$\begin{aligned}\left|\mathbf{D}\Delta\mathbf{x}\right|^2 &= \left(\mathbf{D}\Delta\mathbf{x}\right)^T\mathbf{D}\Delta\mathbf{x} = \left(\mathbf{D}\Delta\mathbf{x}\right)^T\mathbf{V}\mathbf{V}^T\mathbf{D}\Delta\mathbf{x} \\ &= \left(\mathbf{V}^T\mathbf{D}\Delta\mathbf{x}\right)^T\mathbf{V}^T\mathbf{D}\Delta\mathbf{x} = \left|\mathbf{V}^T\mathbf{D}\Delta\mathbf{x}\right|^2\end{aligned} \tag{5.2-13}$$

where $\mathbf{V}\mathbf{V}^T = \mathbf{I}$ .....The end of Note.

Define δ as the radius of confidence region

$$\left|\mathbf{D}\Delta\mathbf{x}\right| \le \delta, \tag{5.2-14}$$

and the parameter λ should be determined to satisfy the above equation.

Next, the degree of non-linearity may be defined in terms of a parameter *r* as

$$r = \frac{\text{Actual variation of the cost function}}{\text{Maximum variation that allows a linear approximation}}, \tag{5.2-15}$$

where the denominator in the above equation is given by

$$\delta J\left(\Delta\mathbf{x}\right) = -\Delta\mathbf{x}^T\left[\left(\mathbf{T}_\varepsilon^{-T}\mathbf{K}_i\right)^T\left(\mathbf{T}_\varepsilon^{-T}\mathbf{K}_i\right) + \mathbf{S}_a^{-1} + 2\lambda\mathbf{D}^2\right]\Delta\mathbf{x}. \tag{5.2-16}$$

The radius of the confidence region may be changed depending upon the degree of nonlinearity as shown below.

(1) If $\left|r-1\right| < 0.25$:
Double the value of δ.

(2) If $\left|r-1\right| \ge 0.25$:

(2.a) If $r > 0$,
$\delta \to \max\left(0.5, 0.5/\left|r-1\right|\right)\times\delta$.

(2.b) If $r \le 0$,
This means that the successive difference did not decrease. Hence abort $\mathbf{x}_{i+1}$ and repeat the process with a new delta as $\delta \to 0.5\delta$.

Now Eq. (5.2-1) may be rewritten as

$$\begin{aligned}\left(\mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}\mathbf{K}_i + \mathbf{S}_a^{-1} + \lambda\mathbf{D}^2\right)\left(\mathbf{x}_{i+1} - \mathbf{x}_a\right) = \mathbf{K}_i^T\mathbf{S}_\varepsilon^{-1}\left\{\mathbf{y} - \mathbf{F}_i + \mathbf{K}_i\left(\mathbf{x}_i - \mathbf{x}_a\right)\right\} \\ + \lambda\mathbf{D}^2\left(\mathbf{x}_i - \mathbf{x}_a\right)\end{aligned}. \tag{5.2-17}$$

It is noted that solving Eq. (5.2-17) for $\mathbf{x}_{i+1} - \mathbf{x}_a$ is equivalent to seeking a least square solution for the following equation.

$$\begin{bmatrix} \mathbf{T}_{\varepsilon}^{-T}\mathbf{K}_i \\ \mathbf{T}_{ainv} \\ \sqrt{\lambda}\mathbf{D} \end{bmatrix}\left(\mathbf{x}_{i+1}-\mathbf{x}_a\right)=\begin{bmatrix} \mathbf{T}_{\varepsilon}^{-T}\left\{\mathbf{y}-\mathbf{F}_i+\mathbf{K}_i\left(\mathbf{x}_i-\mathbf{x}_a\right)\right\} \\ \mathbf{0} \\ \sqrt{\lambda}\mathbf{D}\left(\mathbf{x}_i-\mathbf{x}_a\right) \end{bmatrix}. \tag{5.2-18}$$

In the TANSO-FTS SWIR L2 algorithm, Eq. (5.2-18) is solved using a subroutine DGELSS of LAPACK.

Note: Pre-multiply the following amount to sides of Eq. (5.2-18)

$$\begin{bmatrix} \mathbf{T}_{\varepsilon}^{-T}\mathbf{K}_i \\ \mathbf{T}_{ainv} \\ \sqrt{\lambda}\mathbf{D} \end{bmatrix}^T. \tag{5.2-19}$$

Then Eq. (5.2-17) is obtained…….....The end of Note.

Continue the iterative improvements of the solutions until both of the following two conditions are satisfied:

(1) The successive difference becomes saturated; and

(2) A visible change ceases to occur in the successive solutions.

For the first condition, define a normalized χ^2 as

$$\chi_i^2 = J\left(\mathbf{x}_i\right)/m. \tag{5.2-20}$$

Then the saturation condition is defined as

$$\left|\chi_{i+1}^2-\chi_i^2\right| < f_{tol}, \tag{5.2-21}$$

and this condition must be satisfied, where *m* stands for the number of wavenumber channels.

For the second condition, normalize the difference of successive solutions with error covariance matrix for solution $\mathbf{S}$,

$$d_i^2 = \left(\mathbf{x}_{i+1}-\mathbf{x}_i\right)^T\mathbf{S}^{-1}\left(\mathbf{x}_{i+1}-\mathbf{x}_i\right). \tag{5.2-22}$$

The following equation dictates the condition for the iteration termination.

$$\frac{\left|d_i^{\,2}\right|}{n} < x_{tol}\,, \tag{5.2-23}$$

where n stands for the number of unknown parameters. In the TANSO-FTS SWIR L2 algorithm, $f_{tol} = 10^{-3}$ and $x_{tol} = 10^{-2}$.

Finally, the algorithm for minimizing the cost function given by Eq. (5.1-1), under the condition that the retrieval variable $\mathbf{x}$ is constrained as $\mathbf{x}_{\min} \le \mathbf{x} \le \mathbf{x}_{\max}$, will be explained. First, the retrieval variables $\mathbf{x}$'s should be sorted into two categories **(**i.e., $\mathbf{x}_E$ and $\mathbf{x}_S$, where $\mathbf{x}_E$ is the group when $\mathbf{x}$ satisfies the condition with the equality sign, and $\mathbf{x}_S$ is the group when $\mathbf{x}$ satisfies the condition with an inequality sign). Under this setup, an iterative process should be performed as explained below.

(1) First, assume that all the variables belong to $\mathbf{x}_S$.

(2) Compute $\Delta\mathbf{x} = \mathbf{x}_{i+1} - \mathbf{x}_i$.

(3) Consider all retrieval variables x_j that belong to $\mathbf{x}_S$ and satisfy the condition $x_j + \Delta x_j < x_{j,\min}$ (i.e., $\Delta x_j < 0$). Obtain a maximum value of α_1 that satisfies $x_j + \alpha_1 \Delta x_j \ge x_{j,\min}$, where $\alpha_1 (< 1)$ is a positive value.

(4) Consider all retrieval variables x_j that belongs to $\mathbf{x}_S$ and satisfies the condition $x_j + \Delta x_j > x_{j,\max}$ (i.e., $\Delta x_j > 0$). Obtain a maximum value of α_2 that satisfies $x_j + \alpha_2 \Delta x_j \le x_{j,\max}$, where $\alpha_2 (< 1)$ is a positive value.

(5) Define $\alpha = \min(\alpha_1, \alpha_2)$.

(6) Move the retrieval variable to $\mathbf{x}_E$ group corresponding to α.

(7) If $\alpha = 0$, then this means that $\mathbf{x}_E$ was included in $\mathbf{x}_S$. In this case, the column corresponding to the particular $\mathbf{x}$ that gives this α in a Jacobian matrix be set to 0. Then return to step (2).

(8) Update the retrieval variable that belongs to $\mathbf{x}_S$ as

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha\Delta\mathbf{x}. \tag{5.2-24}$$

6. Input Data and Pre-processing

6.1 Wavenumber Axis and Setting a Range for Wavenumbers

The observed spectra within the selected wavenumber ranges are utilized in the retrieval processing. The selected wavenumber range is called as a retrieval sub-band (or simply sub-band). Note that the retrieval sub-band does not correspond to the observational band of the TANSO-FTS.

From the wavenumber axes of observed spectra, select a particular wavenumber with which to conduct a monochromatic radiative transfer calculation in a forward model (Chapter 4). First, let the wavenumbers, which lie within the range of retrieval wavenumbers after the correction (Section 2.5), be defined as

$$v_i = v_1 + (i-1)\Delta v, i = 1,2,\cdots,N\ . \tag{6.1-1}$$

The forward spectrum F_i can be computed through a convolution integral of monochromatic radiance I and an instrument line shape function ILS as shown below:

$$F_i = \sum_{j=-N_w}^{N_w} ILS(-j\delta v) I(v_i + j\delta v)\delta v\ , \tag{6.1-2}$$

where the increment of the integration should be pre-selected as $\delta v_0 = 0.01$ cm^{-1}. However, the actual increment δv used should be the one that makes $\Delta v/\delta v$ an integer, and is closest to δv_0. Mathematically this may be expressed as

$$\delta v = \Delta v / n_s\ , \tag{6.1-3}$$

where an integer n_s is defined as

$$n_s = f(\Delta v/\delta v_0), \tag{6.1-4}$$

where $f(x)$ is a function that rounds off a pertinent variable x.

The convolution integral should be performed in the wavenumber range of $v_i - w \le v \le v_i + w$, where a pre-set known number of $w_0 = 20$ cm^{-1} is used. The integer value of N_w in Eq. (6.1-2) is given by

$$N_w = f(w_0/\delta v). \tag{6.1-5}$$

The actual wavenumber range (of one side) is

$$w = N_w \delta v\ . \tag{6.1-6}$$

From the above discussion, it follows that the wavenumber axis v_i^{lbl} of a monochromatic radiance is given by the following,

$$v_i^{lbl} = v_1^{lbl} + (i-1)\delta v, i = 1,2,\cdots, N^{lbl} . \quad (6.1\text{-}7)$$

The initiation and the termination wavenumbers are given, respectively, as

$$v_1^{lbl} = v_1 - w, \quad (6.1\text{-}8)$$

and

$$v_{N^{lbl}}^{lbl} = v_N + w . \quad (6.1\text{-}9)$$

Next, consider a correction of wavenumber axis deviation due to Doppler shift. A light, whose wavenumber is v_i^{lbl} when it is observed by a stationary coordinate on FTS, has a wavenumber v_i^{cal} when it is observed in a stationary earth's atmosphere coordinate

$$v_i^{cal} = \left(1 - \frac{v_{dop}}{c}\right) v_i^{lbl} , \quad (6.1\text{-}10)$$

where v_{dop} refers to the LOS direction component of the relative velocities between the earth's atmosphere and GOSAT, and takes a positive value if the two elements are approaching each other. Thus the monochromatic radiative transfer must be conducted at the wavenumber v_i^{cal} obtained through Eq. (6.1-10).

Next it is necessary to investigate the Doppler shifts of solar irradiance spectra. Let the relative velocity of the earth's atmosphere at the observation point and the sun be denoted as v_{sun}, where it takes a positive value when two elements are approaching each other. In this case, let the solar irradiance, whose wavenumber is v_i^{sun}, be observed as a light having a wavenumber v_i^{cal}. Then these two wavenumbers are related as:

$$v_i^{sun} = \left(1 - \frac{v_{sun}}{c}\right) v_i^{cal} . \quad (6.1\text{-}11)$$

Thus, if the monochromatic radiative transfer calculation is conducted at the wavenumber axis of v_i^{cal}, then the corresponding solar spectrum should be evaluated at the wavenumber of v_i^{sun}.

In the computation of v_{dop}, earth's rotation and the orbital motion of GOSAT satellite must be taken into consideration. Similarly, in the computation of v_{sun}, the effects of earth's rotation and revolution must be taken care of.

6.2 Instrument Line Shape Function

A set of Instrument Line Shape Functions has been provided by JAXA where such functions are based on models constructed from actual observations. In the TANSO-FTS SWIR L2 processing, an instrument line shape function, which has been evaluated at the both edges of the retrieval wavenumbers range, is interpolated linearly to each target wavenumber. The line shape functions provided by JAXA consist of tables in the wavenumber interval of 0.01 cm^{-1}. In order to conduct a convolution integral, a third order spline-interpolation should be performed where the wavenumber points are separated by an interval of $\delta\nu$ as shown in Eq. (6.1-3).

During the long-term operation of the TANSO-FTS, a stable zero-path difference (ZPD) position has been changed several times. And this causes a small change in the spectral resolution (i.e., the line shape function) of the TANSO-FTS [*Kuze et al.*, 2016]. The relationship between the observation period and the line shape functions used in the retrieval processing is summarized in Appendix 3.

6.3 Vertical Grid

The profiles for atmospheric pressure, temperature and water vapor are provided in terms of an extended GPV in the 31 levels of atmospheric vertical pressure (Section 2.3). However, in the TANSO-FTS SWIR L2 process, the GPV's vertical grid should not be used as is; rather, it is necessary to re-define a vertical grid for retrieval.

In the TANSO-FTS SWIR L2 algorithm, three vertical grids (i.e., main layer, sub-layer and RT-layer) are defined as explained below.

(a) Main-Layer

The vertical profile of the mixing ratio of carbon dioxide (CO_2), methane (CH_4) and water vapor (H_2O) are defined by the average mixing ratios in the main layer (Figure 6.3-1 left). In the event of simultaneous estimation of column amounts of carbon dioxide (CO_2), methane (CH_4) and water vapor (H_2O), the vertical grid which has a constant pressure difference should be used as the main-layer. In the TANSO-FTS SWIR L2 processing, the main-layer consists of 15 layers regardless of the surface pressure.

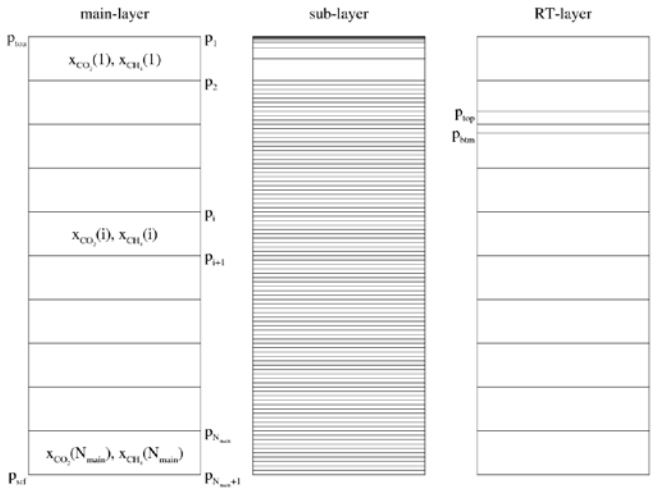

Figure 6.3-1: Vertical Grids.

(b) Sub-Layer

One caveat is that, if the atmospheric temperature and pressure are treated constant in each of a small number of layers such as the main-layer, then it is possible that the optical thickness computation results in a large error.

In order to cope with this difficulty, in the TANSO-FTS SWIR L2 algorithms, the main layer is further divided into sub-layers, and the optical thickness due to gas absorption is computed in each sub-layer. Let the number of the main-layers be N_{main}, and that of sub-layer N_{sub}. Then N_{sub} is determined as an integer value satisfying the following equation.

$$N_{main} N_{sub} \leq 200. \qquad (6.3\text{-}1)$$

That is, each main layer is sub-divided into as many as N_{sub} layers which have a constant pressure difference. However, the top layer should be divided into sub-layers of logarithmically equal layers of atmospheric pressure (Figure 6.3-1 Center). In the TANSO-FTS SWIR L2 processing, N_{sub} is 12.

The optical thickness due to gas absorption for any particular main layer should be equal to the cumulative optical thicknesses of the sub-layers in that main layer. In order to compute the optical thicknesses of sub-layers, it is necessary to compute the average values of atmospheric pressure and temperature as well as the column amount. For this purpose, first compute the temperature profile through the interpolated values of the pressure at the sub-layer's boundaries (Figure 6.3-2); then compute the arithmetic average temperature from the boundary values. The interpolation for both temperature and pressure utilizes the CIP method [*Xiao et al.*, 1996] after obtaining logarithmic values of both. The pressure value for each sub-layer should use an arithmetic average of pressure values at its boundaries in the same manner as the computation of

temperature. The computational methodology for the column amount will be explained in sub-sections (d) and (e).

(c) RT-Layer

If cirrus clouds are detected during observations, then the cloud top and base pressure should be added to the main-layer as boundaries. Whenever a boundary is added, each of the pre-existed layers is divided into two layers. Consequently the optical thickness due to gas absorption is also divided into two parts, where the original optical thickness value is prorated with the pressure difference of the upper and lower boundaries.

The layers thus created are called as RT-layers (Figure 6.3-1 Right). The radiative transfer calculation should be conducted for the RT-layers.

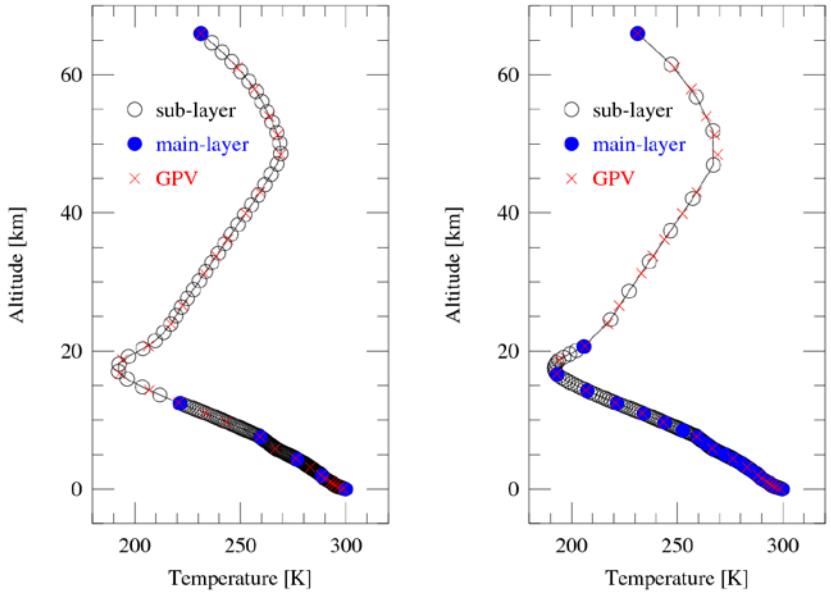

Figure 6.3-2: Interpolation of a temperature profile into the sub-layer.
Left: $N_{main} = 5, N_{sub} = 40$, Right: $N_{main} = 20, N_{sub} = 10$.

(d) Column Amount of Dry Air

Estimating the total column amount, among other physical quantities of interest, is the ultimate goal. Also, the retrieval grid is different for each observation point according to the surface pressure. Therefore, it is important to utilize such a computational approach that would assure the correct total column amount regardless of the particular selection of the vertical grid or the different column amounts in the selected vertical grid. One computational approach that satisfies the above requirements is: (a) first, the column amounts of dry air in the pressure layers of the GPV is calculated under the hydrostatic equilibrium assumption, and (b) then, remap the column amounts to each sub-layer. This approach is briefly explained below.

Let ρ stand for the mass density of air, ρ_d be the mass density of dry air and ρ_v refer to the mass density of water vapor. Then, ρ is a sum of ρ_d and ρ_v, and hence the following relationship holds:

$$\rho = \rho_d + \rho_v = u(\mu_d n_d + \mu_v n_v) = n_d u(\mu_d + \mu_v x_{H2O}), \tag{6.3-2}$$

where n_d and n_v are, respectively, molecular densities of dry air and water vapor; μ_d and μ_v denote average molecular weights of dry air and water vapor, respectively; u stands for atomic mass unit; and $x_{H2O} = n_v / n_d$ represents the volume mixing ratio (or dry air mole fraction) of water vapor. Then, Eq. (2.3-17), the so-called "hydrostatic equilibrium equation", can now be rewritten using Eq. (6.3-2) as:

$$n_d dz = \frac{\rho dz}{u(\mu_d + \mu_v x_{H2O})} = -\frac{dp}{gu(\mu_d + \mu_v x_{H2O})}. \tag{6.3-3}$$

Let the altitude and pressure at the top of the j-th layer be denoted, respectively, as z_j and p_j; let same physical quantities at the bottom of the layer be z_{j+1} and p_{j+1}, respectively. Then the partial column amount $w_d(j)$ of the dry air in the j-th layer is given by the following equation:

$$w_d(j) = \int_{z_{j+1}}^{z_j} n_d dz = \int_{p_j}^{p_{j+1}} \frac{dp}{gu(\mu_d + \mu_v x_{H2O})}. \tag{6.3-4}$$

The gravity and the mixing ratio of water vapor are found in the pressure levels in the GPV. Hence, the average values of these physical quantities in the j-th layer should be expressed as

$$g_{j+1/2} = \frac{g_j + g_{j+1}}{2}, \tag{6.3-5}$$

$$x_{H2O}(j+1/2) = \frac{x_{H2O}(j) + x_{H2O}(j+1)}{2}. \tag{6.3-6}$$

From those equations, the partial column amount $w_d(j)$ of the dry air in the j-th layer is derived as

$$w_d(j) = \frac{p_{j+1} - p_j}{g_{j+1/2} u[\mu_d + \mu_v x_{H2O}(j+1/2)]}. \tag{6.3-7}$$

Now, define the accumulated amount $W_d(p)$ of dry air from the top of the atmosphere to an arbitrary pressure level p as

$$W_d(p) = \int_{z_1}^{z(p)} n_d(z')dz' . \quad (6.3\text{-}8)$$

Then it follows that $W_d(p)$ is computed, at the pressure of the GPV, as

$$W_d(p_1) = 0 , \quad (6.3\text{-}9)$$
$$W_d(p_{j+1}) = W_d(p_j) + w_d(j), j = 1,2,\cdots,30 . \quad (6.3\text{-}10)$$

Then the partial column amount of dry air $w_d(i)$ in the i-th sub-layer can be computed from the $W_d(p)$ as

$$w_d(i) = W_d(p_{i+1}) - W_d(p_i) . \quad (6.3\text{-}11)$$

The value of $W_d(p_i)$ for pressure p_i in the sub-layer can be computed by means of linear interpolation from the value of $W_d(p_j)$. If the pressure $p_i > 1000$ hPa, then it should be linearly extrapolated. Thus no matter how the sub-layer may be defined, it holds that

$$\sum_{i=1}^{N_{main}N_{sub}} w_d(i) = W_d(p_s) . \quad (6.3\text{-}12)$$

From the above, it is assured that the total column amount of dry air remains invariant.

(e) Remapping Gas Mixing Ratio

In the TANSO-FTS SWIR L2 algorithm, it becomes necessary to transform the gas-mixing ratio between different pressure conditions. For example, the partial column amount in a sub-layer must be computed based on the gas-mixing ratio provided by the GPV's pressure levels. Likewise, it is necessary to compute an average gas-mixing ratio in a main layer. Similar to the dry air case, it is preferable to utilize a gas-mixing ratio transformation method that allows the total column amount to remain invariant regardless of any vertical layer setup.

Let the gas-mixing ratio at pressure level p_j^a be denoted as $x_j^a (1 \le j \le N)$. Discussed in the following is the methodology for computing the partial column amount w_i in the i-th layer of a main-layer or a sub-layer.

$$w_i = \int_{z_{i+1}}^{z_i} x(z) n_d(z) dz . \quad (6.3\text{-}13)$$

From Eq. (6.3-8), it follows that $n_d\,dz = dW_d$. It is assumed that the mixing ratio is expressed incrementally as a linear function of W_d. For brevity, let $a_j = W_d(p_j^a)$ and $b_i = W_d(p_i)$. Then it follows that

$$\begin{aligned} w_i &= \int_{b_i}^{b_{i+1}} x(W_d)\,dW_d \\ &= \sum_{j=1}^{N-1} \int_{c_j}^{c_{j+1}} \left[x_j^a + \frac{x_{j+1}^a - x_j^a}{a_{j+1} - a_j}(W_d - a_j) \right] dW_d + x_1^a \Delta_0 + x_N^a \Delta_N , \\ &= \sum_{j=1}^{N-1} \left[(1 - A_j) x_j^a + A_j x_{j+1}^a \right] \Delta_j + x_1^a \Delta_0 + x_N^a \Delta_N \end{aligned} \tag{6.3-14}$$

where

$$c_0 = b_i , \tag{6.3-15}$$

$$c_1 = \min(a_1, b_{i+1}), \tag{6.3-16}$$

$$c_N = \max(a_N, b_i), \tag{6.3-17}$$

$$c_{N+1} = b_{i+1}, \tag{6.3-18}$$

$$c_j = \max(a_j, b_i), \tag{6.3-19}$$

$$c_{j+1} = \min(a_{j+1}, b_{i+1}), \tag{6.3-20}$$

$$\Delta_j = \max(c_{j+1} - c_j, 0), \tag{6.3-21}$$

$$A_j = \frac{\frac{c_j + c_{j+1}}{2} - a_j}{a_{j+1} - a_j} . \tag{6.3-22}$$

In the above, it was set that $x = x_1^a$ if $p < p_1^a$, and $x = x_N^a$ (by extrapolation) if $p > p_N^a$.

Define the average value of gas-mixing ratio in the i-th layer as $x_i = w_i / w_d(i)$. Also, let $\mathbf{x}$ and $\mathbf{x}^a$ be defined as, $\mathbf{x} = (x_1, x_2, \cdots)^T$ and $\mathbf{x}^a = (x_1^a, x_2^a, \cdots)^T$, respectively. Then Eq. (6.3-14) is expressed in a compact form as

$$\mathbf{x} = \mathbf{W}^* \mathbf{x}^a . \tag{6.3-23}$$

The matrix $\mathbf{W}^*$ can be used when remapping *a priori* values of gas mixing ratios and the corresponding variance and covariance matrices given by the GPV in 21 pressure levels to the main-layer (Figure 6.3-3, see Section 6.4).

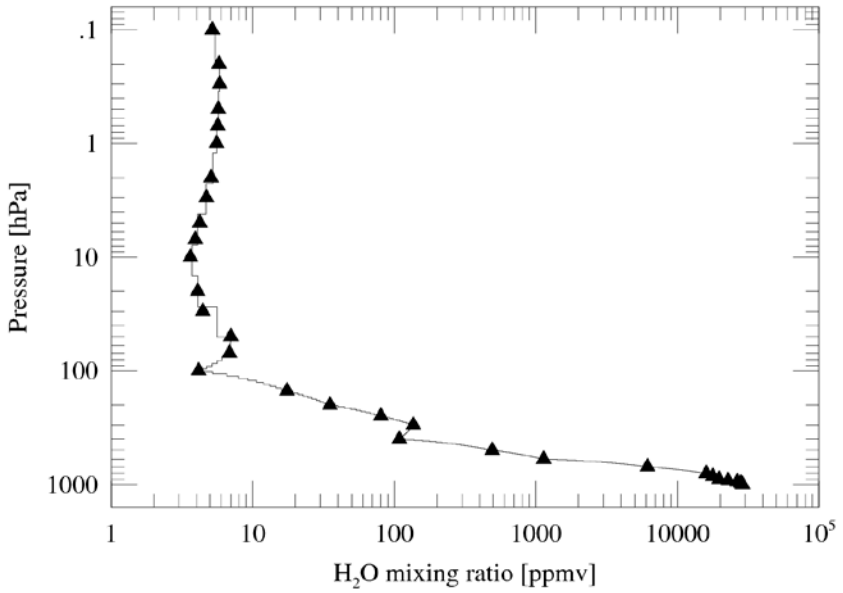

Figure 6.3-3: Remap of Gas Profile.

6.4 *A Priori* Information of Gas Concentration

Here, we describe the *a priori* and the variance and covariance matrices of CO_2, CH_4, and H_2O.

(a) Carbon Dioxide (CO_2)

A priori Information

Use the simulated data by the NIES Transport Model [*Maksyutov et al.*, 2008] as the *a priori* information for the observation point by the TANSO-FTS SWIR. From the standard grid resolution of this model (i.e., 0.5 degrees and a time interval of 3 hours) appropriate interpolations should be performed for the particular time and location of an observation point.

Variance and Covariance Matrices

Compute the variance and covariance values at 12-month for each horizontal grid point of the model from the simulated data for the year 2008 by the NIES Transport Model and the ground-based observation data GLOBALVIEW-2008 (1979 through 2007) [*Eguchi et al.*, 2010]. Use the variance and covariance data of the nearest location to the observation point. Such data should be linearly interpolated to fit to the observation time.

(b) Methane (CH_4)

A priori Information

Similar to the case of carbon dioxide, use the simulated data by the NIES Transport Model for the *a priori* information at the TANSO-FTS SWIR observation point.

Variance and Covariance Matrices

Compute the variance and covariance values at 12-month for each horizontal grid point of the model from the simulated data for the year 2008 by the NIES Transport Model, the ground-based observation data GLOBALVIEW-2008 (1979 through 2007) and the flight data observed by the HALOE satellite (1994 through 2005; data exists for the stratosphere only) [*Eguchi et al.*, 2010]. Use the variance and covariance data of the nearest location to the observation point. Such data should be linearly interpolated to fit to the observation time.

(c) Water Vapor

***A priori* Information**

The *a priori* values may be obtained from the GPV and temporally and spatially interpolated for the particular observation point.

Variance and Covariance Matrices

Based on the GPV data for the year 2008, pre-compute the average values of the relative humidity for each month and associated variance and covariance matrices. Then, for the retrieval process, use the monthly variance and covariance matrices of the GPV's horizontal grid nearest to the observation point.

(d) Remapping in the Vertical Direction

The *a priori* variance and covariance matrices are pre-computed for carbon dioxide, methane and water vapor, interpolated at 21 pressure levels of the GPV. In order to remap the *a priori* values and the variance and covariance matrices into the main layer, use is made of the transformation matrix $\mathbf{W}^*$ presented in Eq. (6.3-23).

Namely, let the *a priori* value be $\mathbf{x}_a^0$ and the covariance matrix $\mathbf{S}_a^0$. Then their corresponding expressions in the main-layer are remapped as

$$\mathbf{x}_a = \mathbf{W}^* \mathbf{x}_a^0 , \tag{6.4-1}$$

$$\mathbf{S}_a = \mathbf{W}^* \mathbf{S}_a^0 \mathbf{W}^{*T} . \tag{6.4-2}$$

(e) Update of the NIES Transport model

The *a priori* values and the variance and covariance matrices of CO_2 and CH_4 are updated according to the update of the NIES Transport Model. The stratospheric part of the updated NIES Transport Model is nudged to climatological values using three-dimensional CO_2 climatology that was adjusting to observed age of air for CO_2, and to the long-term satellite observations from HALOE for CH_4. Also, the variance and covariance matrices are updated to give more realistic stratospheric values [*Saeki et al.*, 2013].

6.5 *A Priori* Information of Relative Radiometric Calibration Adjustment Factor

In the TANSO-FTS SWIR L2 V02.xx algorithm, the relative radiometric calibration adjustment factor is not considered.

6.6 *A Priori* Information of Ground Surface Albedo

Use the average albedo data over 16 days, which was obtained by the MODIS (Moderate Resolution Imaging Spectrometer) instrument flown on the Aqua and Terra satellites over the period of five years from 2000 through 2004 (Black-sky albedo of MOD43B3). All the data with a one arc-minute angular resolution of MODIS Band 2 (0.858 μm, 841 ~ 876 nm), MODIS Band 6 (1.64 μm, 1628 ~ 1652 nm), and MODIS Band 7 (2.13 μm, 2105 ~ 2155 nm) should be used after interpolating for the observation time and observation point of the TANSO-FTS SWIR. Partly because the Black-Sky albedo is applicable to about 90% of the cases and partly because it is preferable to simplify the computational process, the White-Sky albedo is not used for the computation of *a priori* values. The standard deviation values for surface albedo is set at 1.0.

6.7 *A Priori* Information of Ocean Winds

For retrievals over the ocean, the ocean wind speed, instead of albedo, is selected as an estimating physical parameter of interest. The ocean wind speed values given by the GPV are used as the *a priori* values after performing interpolations for the particular observation time and observation point of the TANSO-FTS SWIR. Based on the GPV data for the year 2008, pre-compute the average values of the wind speed for each month and associated variance. Then, for the retrieval process, use the monthly variance of the GPV's horizontal grid nearest to the observation point.

6.8 *A Priori* Information of a Cirrus Cloud

Currently it is assumed that a cirrus cloud does not come into play in the TANSO-FTS SWIR L2 processing (Also refer to Section 3.7 and Section 3.10).

6.9 *A Priori* Information of Aerosol

The *a priori* information of aerosols can be obtained from the simulation results by the aerosol transport model SPRINTARS [*Takemura et al.*, 2000]. In SPRINTARS, the mass mixing ratio of aerosol over 20 vertical layers are computed, with a horizontal resolution of 2.8 degrees and a temporal resolution of 6 hours, for carbon-, sulfate-, soil dust- and sea salt-based aerosol types. A logarithm of the mass mixing ratios of fine-mode (carbon- and sulfate-based aerosol) and coarse-mode (soil dust- and sea salt-based aerosol) aerosols simulated by SPRINTARS are used as the *a priori* values after performing interpolations for the particular observation time and observation point of the TANSO-FTS SWIR (see section 2.3). The number of aerosol layers are six (i.e., two

types x six layers = 12 parameters for aerosols). A diagonal variance-covariance matrix with its diagonal value of $(0.5)^2$ is considered.

6.10 *A Priori* Information of Surface Pressure

The surface pressure values calculated from the geo-potential height of the GPV and the elevation data (ASTER GDEM/GTOPO30) are used as the *a priori* values after performing interpolations for the particular observation time and observation point of the TANSO-FTS SWIR (see section 2.3). The standard deviation value for surface pressure is set at 5 hPa.

6.11 *A Priori* Information of Temperature Shift

Let consider a constant bias for the temperature profile of the GPV and is simultaneously retrieved with gas concentrations. The *a priori* value and the standard deviation should be set at 0.0 and 5 K, respectively.

6.12 *A Priori* Information of Wavenumber Dispersion Correction Factor

Although the wavenumber dispersion is corrected during the pre-processing (Eq. (2.5-3)), it should be re-estimated in the retrieval processing. A small dispersion $\Delta\rho$ to the correction factor ρ is retrieved simultaneously with the gas concentration.

$$v_i = \rho(1+\Delta\rho)v_i^{obs}, i = 1,2,\cdots,N \qquad (6.12\text{-}1)$$

Consider the wavenumber dispersion correction factors for each retrieval sub-band, and the *a priori* value and the standard deviation should be set at 0.0 and 10^{-5}, respectively.

6.13 *A Priori* Information of Zero-Level Offset

A problem with the analog circuitry and a separate problem with the analog-to-digital converter (ADC) of the TANSO-FTS Band 1 channel cause a non-linear response of the signal chains [*Suto et al.*, 2011; *Kuze et al.*, 2012]. To reduce the impact of the non-linear response on the retrieval results, a wavenumber-independent offset of TANSO-FTS Band 1 (hereafter referred to as zero-level offset) is retrieved simultaneously [*Butz et al.*, 2011]. The *a priori* value and the standard deviation should be set at 0.0 and 10^{-8} W/cm^2/str/cm^{-1}, respectively.

7. Atmospheric Column Amount Estimation Algorithm

7.1 Retrieval of Carbon Dioxide (CO_2), Methane (CH_4) and Water Vapor (H_2O) Column Amounts

The total column amounts of carbon dioxide, methane and water vapor, retrieved from the radiance spectra in the ranges of 12950 ~ 13200 cm^{-1} (O_2 sub-band), 5900 ~ 6150 cm^{-1} (CH_4 sub-band), 6180 ~ 6380 cm^{-1} (CO_2 sub-band), and 4800 ~ 4900 cm^{-1} (SCO_2 sub-band), should be computed by multiplying the products of the average gas mixing ratio estimated for each layer by the column amount of dry air. Also computed are the column-averaged dry air mole fractions (XCO_2, XCH_4 and XH_2O), which are defined as the total column amount of carbon dioxide and methane divided by the total column amount of dry air.

Together with the layer-averaged gas mixing ratio of CO_2 CH_4 and H_2O, simultaneously retrieve the aerosol profile, surface pressure, temperature shift, wavenumber dispersion correction factor, zero-level offset and surface albedo for land cases or ocean wind speed for ocean cases.

Table 7.1-1: Elements of state vector for carbon dioxide and methane column amounts retrieval.

Range of retrieval wavenumbers	O_2 sub-band (12950 ~ 13200 cm^{-1}) CO_2 sub-band (6180 ~ 6380 cm^{-1}) CH_4 sub-band (5900 ~ 6150 cm^{-1}) SCO_2 sub-band (4800 ~ 4900 cm^{-1})
Estimating parameter of interest	Layer-averaged CO_2 gas-mixing ratio (Standard Product) Layer-averaged CH_4 gas-mixing ratio (Standard Product) Layer-averaged H_2O gas-mixing ratio (Standard Product)
Other parameters to be retrieved simultaneously	Aerosol mass mixing ratio profile (2 types x 6 layers = 12 parameters) Surface pressure Temperature shift Wavenumber dispersion correction factor Zero-level offset for TANSO-FTS Band 1 Ground surface albedo (When observing ground) Ocean wind speed (When observing ocean)

(a) Gas-Mixing Ratios of Carbon Dioxide, Methane, and Water Vapor

The estimating parameter of interest is the average gas-mixing ratio for each layer of the main-layer explained in Section 6.3. The *a priori* values and the variance and covariance matrices should be prepared as explained in Section 6.4.

(b) Aerosol

The logarithm of mass mixing ratios of fine- and coarse-mode aerosols should be included among parameters to be estimated simultaneously. The *a priori* values and the variance and covariance matrices should be prepared as explained in Section 6.9.

(c) Surface Pressure

The surface pressure should be included among parameters to be estimated simultaneously. The *a priori* values and the variance should be prepared as explained in Section 6.10.

(d) Temperature Shift

The temperature shift should be included among parameters to be estimated simultaneously. The *a priori* values and the variance should be prepared as explained in Section 6.11.

(e) Wavenumber Dispersion Correction Factor

The wavenumber dispersion correction factor should be included among parameters to be estimated simultaneously. The *a priori* values and the variance should be prepared as explained in Section 6.12.

(f) Zero-level offset for TANSO-FTS Band 1

The zero-level offset for TANSO-FTS Band 1 should be included among parameters to be estimated simultaneously. The *a priori* values and the variance should be prepared as explained in Section 6.13.

(g) Ground Surface Albedo or Ocean Wind Speed

When the TANSO-FTS observation point falls on land, the dependency of ground surface albedo on wavenumbers should be plotted in a piece-wise manner with connected straight lines of a resolution of 250 cm^{-1} for O_2 sub-band, 25 cm^{-1} for CO_2 and CH_4 sub-bands, and 100 cm^{-1} for SCO_2 sub-band. Then the local peak values of albedo in the piece-wisely connected straight lines should be included among parameters to be estimated simultaneously. The *a priori* value and the variance are to be set based on MODIS data explained in Section 6.6.

When the TANSO-FTS observation point falls on the ocean surface, the ocean wind speed should be included among the parameters to be estimated simultaneously. The *a priori* values and the variance should be prepared as explained in Section 6.7.

(h) Scaling of O_2 Absorption Cross Section

Retrieved surface pressure tends to show positive bias even for the low aerosol concentration case. This remnant bias was removed by adjusting the absorption cross

section of O_2 [*Butz et al.*, 2011; *Crisp et al.*, 2012], although its root cause is unclear. The scaling factor of 1.01 was evaluated according to the method by *Butz et al.* [2011].

(i) Empirical Noise

The residual spectrum $\boldsymbol{\varepsilon}$ between the measured and simulated spectra comprises the measurement error $\boldsymbol{\varepsilon}_{meas}$, the forward model error $\Delta\mathbf{f}$, and the model parameter error $\mathbf{K}_b\left(\mathbf{b}-\hat{\mathbf{b}}\right)$.

$$\boldsymbol{\varepsilon} \equiv \mathbf{y} - \mathbf{F}(\mathbf{x},\mathbf{b}) = \boldsymbol{\varepsilon}_{meas} + \Delta\mathbf{f} + \mathbf{K}_b\left(\mathbf{b}-\hat{\mathbf{b}}\right) \tag{7.1-1}$$

Here, $\mathbf{K}_b$ is the Jacobian matrix for a model parameter $\mathbf{b}$, and $\hat{\mathbf{b}}$ is the true values of $\mathbf{b}$.

In the TANSO-FTS SWIR L2 V01.xx retrievals, we neglected the contribution of the forward model error and the model parameter error and assumed a diagonal matrix $\mathbf{S}_\varepsilon$ (Eq. (5.1-1)) with diagonal elements $\left(\boldsymbol{\varepsilon}_{meas}\right)^2$. However, in this case, the chi-squared cost function (Eq. (5.1-1)) of the retrieved state increases as the SNR increased, indicating that the contribution of the forward model error and/or the model parameter error became large with SNR. Because it was hard to evaluate the forward model error and the model parameter error adequately, the contributions of these error components were empirically modeled as a function of SNR.

$$\boldsymbol{\varepsilon}_{empirical} = \boldsymbol{\varepsilon}_{meas}\sqrt{a_0 + a_1 SNR + a_2 SNR^2} \tag{7.1-2}$$

Here, coefficients a_0, a_1, and a_2 of each spectral window were preliminary evaluated from the relationships between SNR and the mean-squared values of the residual spectra. A diagonal matrix $\mathbf{S}_\varepsilon$ with its diagonal value of $\left(\boldsymbol{\varepsilon}_{empirical}\right)^2$ is used in the SWIR L2 V02.xx retrieval.

8. Channel Selection

8.1 Channel Selection Algorithm

No channel selection is applied in the TANSO-FTS SWIR L2 V02.xx processing.

9. A Posteriori Processing

9.1 Internal Error Evaluation

In the MAP analysis method, the error covariance is given as

$$\mathbf{S} = \left(\mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1}. \tag{9.1-1}$$

The error consists of a smoothing error

$$\mathbf{S}_S = (\mathbf{A} - \mathbf{I})\mathbf{S}_a (\mathbf{A} - \mathbf{I})^T, \tag{9.1-2}$$

and retrieval noise

$$\mathbf{S}_M = \mathbf{G}\mathbf{S}_\varepsilon \mathbf{G}^T, \tag{9.1-3}$$

where $\mathbf{A}$ stands for the averaging kernel

$$\mathbf{A} = \mathbf{G}\mathbf{K}. \tag{9.1-4}$$

The gain matrix $\mathbf{G}$ is given by

$$\begin{aligned} \mathbf{G} &= \left(\mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \mathbf{K} + \mathbf{S}_a^{-1}\right)^{-1} \mathbf{K}^T \mathbf{S}_\varepsilon^{-1} \\ &= \mathbf{S}_a \mathbf{K}^T \left(\mathbf{K}\mathbf{S}_a\mathbf{K}^T + \mathbf{S}_\varepsilon\right)^{-1} \end{aligned}. \tag{9.1-5}$$

For the estimation of other parameters other than a column amount, the diagonal elements of Eq. (9.1-1) should be the variances of retrieval errors.

The retrieval error of the column amount should be evaluated as shown below. Let the state vector $\mathbf{x}$ be divided into two categories (i.e., one category for the gas concentration and denoted as $\mathbf{x}_{gas}$, and the rest as $\mathbf{x}_{nongas}$).

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_{gas} \\ \mathbf{x}_{nongas} \end{pmatrix}. \tag{9.1-6}$$

Similarly the Jacobian and *a priori* covariance matrices are presented in terms of partition matrices, respectively, as

$$\mathbf{K} = \left(\mathbf{K}_{gas} \quad \mathbf{K}_{nongas}\right), \tag{9.1-7}$$

$$\mathbf{S}_a = \begin{pmatrix} \mathbf{S}_{a,gas} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_{a,nongas} \end{pmatrix}. \tag{9.1-8}$$

Then it becomes clear that the retrieval error matrix consists of a smoothing error matrix,

$$\mathbf{S}_{S,gas} = \left(\mathbf{A}_{gas} - \mathbf{I}\right)\mathbf{S}_{a,gas}\left(\mathbf{A}_{gas} - \mathbf{I}\right)^T, \tag{9.1-9}$$

retrieval noise

$$\mathbf{S}_{M,gas} = \mathbf{G}_{gas}\mathbf{S}_{\varepsilon}\mathbf{G}_{gas}^T, \tag{9.1-10}$$

and interference error

$$\mathbf{S}_{I,gas} = \mathbf{A}_{nongas}\mathbf{S}_{a,nongas}\mathbf{A}_{nongas}^T, \tag{9.1-11}$$

where $\mathbf{G}_{gas}$, $\mathbf{A}_{gas}$ and $\mathbf{A}_{nongas}$ are defined as shown below:

$$\mathbf{G}_{gas} = \mathbf{S}_{a,gas}\mathbf{K}_{gas}^T\left(\mathbf{S}_{\varepsilon} + \mathbf{K}_{gas}\mathbf{S}_{a,gas}\mathbf{K}_{gas}^T + \mathbf{K}_{nongas}\mathbf{S}_{a,nongas}\mathbf{K}_{nongas}^T\right)^{-1}, \tag{9.1-12}$$

$$\mathbf{A}_{gas} = \mathbf{G}_{gas}\mathbf{K}_{gas}, \tag{9.1-13}$$

$$\mathbf{A}_{nongas} = \mathbf{G}_{gas}\mathbf{K}_{nongas}. \tag{9.1-14}$$

If the column amount z is given in terms of a product of weighting vector $\mathbf{g}$ and the estimation parameter vector $\mathbf{x}_{gas}$

$$z = \mathbf{g}^T\mathbf{x}_{gas}, \tag{9.1-15}$$

then the variance σ^2 of the retrieval error is

$$\sigma^2 = \sigma_S^2 + \sigma_M^2 + \sigma_I^2, \tag{9.1-16}$$

where

$$\sigma_S^2 = \mathbf{g}^T\mathbf{S}_{S,gas}\mathbf{g}, \tag{9.1-17}$$

$$\sigma_M^2 = \mathbf{g}^T\mathbf{S}_{M,gas}\mathbf{g}, \tag{9.1-18}$$

$$\sigma_I^2 = \mathbf{g}^T\mathbf{S}_{I,gas}\mathbf{g}. \tag{9.1-19}$$

Since it is rather cumbersome to compute $\mathbf{G}_{gas}$ directly from Eq. (9.1-12), first $\mathbf{G}$ is computed from Eq. (9.1-5) and $\mathbf{G}_{gas}$ is extracted from the partition matrix.

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_{gas} \\ \mathbf{G}_{nongas} \end{pmatrix}. \tag{9.1-20}$$

Similarly $\mathbf{A}_{gas}$ and $\mathbf{A}_{nongas}$ can be obtained from the partition matrices of

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{gas} & \mathbf{A}_{nongas} \\ \mathbf{A}_{NG} & \mathbf{A}_{NN} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{gas}\mathbf{K}_{gas} & \mathbf{G}_{gas}\mathbf{K}_{nongas} \\ \mathbf{G}_{nongas}\mathbf{K}_{gas} & \mathbf{G}_{nongas}\mathbf{K}_{nongas} \end{pmatrix}. \tag{9.1-21}$$

9.2 Quality Check Process

For quality control purposes, first, evaluate and select the TANSO-FTS SWIR L2 data that satisfies the retrieval data quality defined by the internal error evaluation process outlined below. Then, only the data selected through the quality check process should be opened to the public as TANSO-FTS SWIR L2 Products.

The following eight criteria should be conducted. Only the data selected through the following quality check process should be archived as TANSO-FTS SWIR L2 Products (V02.xx), where the same threshold values for carbon dioxide and methane should be used for both evaluations.

(1) Select data that satisfies 70 or greater in SNR (Polarized Synthesis Lights) of O_2 sub-band, where SNR is defined as the ratio of the maximum radiance within the wavenumber range to the average value of noise level.

(2) Select data that satisfies the Degree of Freedom for Signal (DFS) being 1 or greater, where DFSs for CO_2, CH_4 and H_2O are defined as a trace (the sum of diagonal elements of a matrix)

$$\mathrm{DFS} = \mathrm{tr}(\mathrm{A}_{\mathrm{gas}}). \tag{9.2-1}$$

(3) Select data that satisfies the Mean Squared of the Residual spectrum (MSR) being less than or equal to 1.2, 1.2, 1.3, 1.4 for O_2, CO_2, CH_4, and SCO_2 sub-bands, respectively.

$$MSR = \frac{\left[\mathbf{y} - \mathbf{F}(\mathbf{x})\right]^T \mathbf{S}_{\varepsilon}^{-1}\left[\mathbf{y} - \mathbf{F}(\mathbf{x})\right]}{m} \tag{9.2-2}$$

(4) Select data that has the aerosol optical thickness (at wavelength of 1.6 μm) of less than or equal to 0.1, which has been estimated together with the gas concentration.

(5) Select data that has the so-called blended albedo [*Wunch et al.*, 2011] less than or equal to 1.0.
(6) Select data that has the deviation of the retrieved surface pressure from its *a priori* value less than or equal to 20 hPa.
(7) Select data that has the ocean wind speed larger than 0.1 m/sec and less than 20.0 m/sec, which has been estimated together with the gas concentration for ocean case.
(8) Select data that has been obtained through the FTS's FOV of land fraction of greater than or equal to 60% or ocean region only. (The observation data is removed if the corresponding FOV-land-fraction is greater than 0% and less than 60%.)

Acknowledgement

Authors thank to Isamu Morino (NIES), Osamu Uchino (NIES), Tomoaki Tanaka (NIES), Nobuyuki Uemura (Fujitsu FIP corporation), Tadayoshi Murakami (Adecco Ltd.), and Hiroshi Otake (Jet Propulsion Laboratory) for their cooperation to complete this document (Version 1.0).

NOTE: Affiliation as of May, 2010.

References

Alduchov, O. A., and R. E. Eskridge (1996), *Complex quality control of upper-air parameters at mandatory and significant levels for the CARDS dataset*, NTISPB-97-132286, 125pp.

Bates, D. R. (1984), Rayleigh scattering by air, *Planet. Space Sci.*, **32**, 785-790.

Bucholtz, A. (1995), Rayleigh-scattering calculations for the terrestrial atmosphere, *Appl. Opt.*, **34**, 2765-2773.

Butz, A., et al. (2011), Toward accurate CO_2 and CH_4 observations from GOSAT, *Geophys. Res. Lett.*, **38**, L14812, doi10.1029/2011GL047888.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P.D. Brown (2005), Atmospheric radiative transfer modeling: a summary of the AER codes, *J. Quant. Spectrosc. Radiat. Transfer*, **91**, 233-244.

Cox, C., and W. H. Munk (1954), Measurement of the roughness of the sea surface from photographs of the Sun's glitter, *J. Opt. Soc. Am.*, **44**, 838-850.

Crisp, D., et al. (2012), The ACOS CO_2 retrieval algorithm - Part II: Global X_{CO2} data characterization, *Atmos. Meas. Tech.*, **5**, 687-707, doi:10.5194/amt-5-687-2012.

Deuze, J. L., M. Herman, and R. Santer (1989), Fourier series expansion of the transfer equation in the atmosphere-ocean system, *J. Quant. Spectrosc. Radiat. Transfer*, **41**, 483-494.

Duan, M., Q. Min, and J. Li (2005), A fast radiative transfer model for simulating high-resolution absorption bands, *J. Geophys. Res.*, **110**, D15201, doi:10.1029/2004JD005590.

Eguchi, N., R. Saito, T. Saeki, D. Belikov and S. Maksyutov (2010), A priori covariance estimation for CO_2 and CH_4 retrievals, *J. Geophys. Res. Atmosphere*, in press.

GLOVALVIEW-CH4 (2008), Cooperative Atmospheric Data Integration Project - Methane. CD-ROM, NOAA ESRL, Boulder, Colorado [Also available on

Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/ch4/GLOBALVIEW].

GLOVALVIEW-CO2 (2008), Cooperative Atmospheric Data Integration Project - Carbon Dioxide. CD-ROM, NOAA ESRL, Boulder, Colorado [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW].

Holben, B.N., T.F. Eck, I. Slutsker, D. Tanré, J.P. Buis, A. Setzer, E. Vermote, J.A. Reagan, Y.J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, and A. Smirnov (1998), AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, *Remote Sens. Environ.*, **66**, 1-16.

Hovenier, J. W., and V. M. van der Mee (1983), Fundamental relationships relevant to the transfer of polarized light in a scattering atmosphere, *Astron. Astrophys.*, **128**, 1-16.

Ishida, H., and T. Y. Nakajima (2009), Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager, *J. Geophys. Res.*, **114**, D07206, doi:10.1029/2008JD010710.

Ishida, H., T. Y. Nakajima, T. Yokota, N. Kikuchi, and H. Watanabe (2011), Investigation of GOSAT TANSO-CAI cloud screening ability through an intersatellite comparison, *J. Appl. Meteor. Climatol.*, **50**, 1571-1586.

Kuze, A., H. Suto, M. Nakajima, and T. Hamazaki (2009), Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, *Appl. Opt.*, **48**, 6716-6733.

Kuze, A., et al. (2011), Vicarious calibration of the GOSAT sensors using the railroad valley desert playa, *IEEE T. Geosci. Remote*, **49**, 1781-1795, doi:10.1109/TGRS.2010.2089527.

Kuze, A., H. Suto, K. Shiomi, T. Urabe, M. Nakajima, J. Yoshida, T. Kawashima, Y. Yamamoto, F. Kataoka, and H. Buijs (2012), Level 1 algorithms for TANSO on GOSAT: processing and on-orbit calibrations, *Atmos. Meas. Tech.*, **5**, 2447-2467, doi:10.5194/amt-5-2447-2012.

Kuze, A., H. Suto, K. Shiomi, S. Kawakami, M. Tanaka, Y. Ueda, A. Deguchi, J. Yoshida, Y. Yamamoto, F. Kataoka, T.E. Taylor, and H.L. Buijs (2016),

Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space, *Atmos. Meas. Tech.*, **9**, 2445-2461, doi:10.5194/amt-9-2445-2016.

Lamouroux, J., H. Tran, A.L. Laraia, R.R. Gamache, L.S. Rothman, I.E. Gordon, and J.-M. Hartmann (2010), Updated database plus software for line-mixing in CO_2 infrared spectra and their test using laboratory spectra in the 1.5-2.3 μm region, *J. Quant. Spectrosc. Raidat. Transfer*, **111**, 2321-2331.

Liou, K. N. (2002), *An Introduction to Atmospheric Radiation: Second Edition*, Academic Press, New York.

Maksyutov, S., P. K. Patra, R. Onishi, T. Saeki, and T. Nakazawa (2008), NIES/FRCGC global atmospheric tracer transport model: Description, validation, and surface sources and sinks inversion, *J. Earth Simulator*, **9**, 3-18.

Nakajima, T., and M. Tanaka (1988), Algorithms for radiative intensity calculations in moderately thick atmosphere using truncation approximation, *J. Quant. Spectrosc. Radiat. Transfer*, **40**, 51-69.

Nakajima, T., T. Y. Nakajima, A. Higurashi, I. Sano, T. Takamura, H. Ishida and N. Schutgens (2008), A study of Aerosol and Cloud Information Retrievals from CAI Imager on Board GOSAT Satellite, *J. Remote Sens. Soc. Japan*, **28**, 178-189 (in Japanese with English abstract and figure captions).

Peck, E. R., and K. Reeder (1972), Dispersion of air, *J. Opt. Soc. Am.*, **62**, 958-962.

Rodgers, C. D. (2000) *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Sci., Singapore.

Rothman, L. S., et al. (2009), The HITRAN 2008 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer*, **110**, 533-572.

Saeki, T., R. Saito, D. Belikov, and S. Maksyutov (2013), Global high-resolution simulations of CO_2 and CH_4 using a NIES transport model to produce a priori concentrations for use in satellite data retrievals, *Geosci. Model Dev.*, **6**, 81-100, doi:10.5194/gmd-6-81-2013.

Siewert, C. E. (2000), A concise and accurate solution to Chandrasekhar's basic problem in radiative transfer, *J. Quant. Spectrosc. Radiat. Transfer*, **64**, 109-130.

Suto, H., A. Kuze, K. Shiomi, and M. Nakajima (2011), Updated Level-1 processing after two-years operation of TANSO-FTS, *Proc. SPIE*, **8154**, 81541A, doi:10.1117/12.893405.

Takamura, T., T. Nakajima, and SKYNET community group (2004), Overview of SKYNET and its Activities, *Optica Pura Y Aplicada*, **37(3)**, 3303-3308.

Takemura, T., H. Okamoto, Y. Maruyama, A. Numaguti, A. Higurashi, and T. Nakajima (2000), Global three-dimensional simulation of aerosol optical thickness distribution of various origins, *J. Geophys. Res.*, **105**, 17853-17873.

Tetens, O. (1930), Ubereinige meteorologische Begriffe, *Z. Geophys.*, **6**, 297-309.

Tran, H., C. Boulet, and J.-M. Hartmann (2006), Line mixing and collision-induced absorption by oxygen in the A band: Laboratory measurements, model, and tools for atmospheric spectra computations, *J. Geophys. Res.*, **111**, D15210, doi:10.1029/2005JD006869.

Tran, H. and J.-M. Hartmann (2008), An improved O_2 A band absorption model and its consequences for retrievals of photon paths and surface pressures, *J. Geophys. Res.*, **113**, D18104, doi:10.1029/2008JD010011.

Warren, S. G. (1984), Optical constants of ice from the ultraviolet to the microwave, *Appl. Opt.*, **23**, 1206-1225.

Wunch, D., et al. (2011), A method for evaluating bias in global measurements of CO_2 total columns from space, *Atmos. Chem. Phys.*, **11**, 12317-12337, doi:10.5182/acp-11-12317-2011.

Xiao, F., T. Yabe, and T. Ito (1996), Constructing oscillation preventing scheme for advection equation by rational function, *Comput. Phys. Commun.*, **60**, 1-12.

Yoshida, Y., N. Kikuchi, and T. Yokota (2012), On-orbit radiometric calibration of SWIR bands of TANSO-FTS onboard GOSAT, *Atmos. Meas. Tech.*, **5**, 2515-2523, doi:10.5194/amt-5-2515-2012.

Yoshida , Y., N. Kikuchi, I. Morino, O. Uchino, S. Oshchepkov, A. Bril, T. Saeki, N. Schutgens, G.C. Toon, D. Wunch, C.M. Roehl, P.O. Wennberg, D.W.T. Griffith, N.M. Deutscher, T. Warneke, J. Notholt, J. Robinson, V. Sherlock, B. Connor, M. Rettinger, R. Sussmann, P. Ahonen, P. Heikkinen, E. Kyrö, J. Mendonca, K. Strong, F. Hase, S. Dohe, and T. Yokota (2013), Improvement of the retrieval algorithm for GOSAT SWIR XCO_2 and XCH_4 and their validation using TCCON data, *Atmos. Meas. Tech.*, **6**, 1533-1547, doi:10.5194/amt-6-1533-2013.

Appendix 1. ACRONYMS

ADC	Analog-to-Digital Converter
AERONET	AErosol Robotic NETwork
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
ASTER GDEM	ASTER Global Digital Elevation Model
AT	Along Track
BS	Beam Splitter
CAM	monitor CAMera
CT	Cross Track
DEM	Digital Elevation Model
DFS	Degree of Freedom for Signals
EORC	Earth Observation Research Center
FOV	Field Of View
GOSAT	Greenhouse gases Observing SATellite
GOSAT DHF	GOSAT Data Handling Facility
GPV	Grid Point Value
GTOPO30	Global 30 Arc-Second elevation Data Set
HALOE	Halogen Occultation Experiment
HITRAN	High-Resolution Transmission Molecular Absorption Database
IFOV	Instantaneous Field Of View
ILS	Instrumental Line Shape
IMC	Image Motion Compensation
JAXA	Japan Aerospace Exploration Agency
L	Level
LBLRTM	Line-By-Line Radiative Transfer Model
LM	Levenberg-Marquardt
LOS	Line Of Sight
LTE	Local Thermodynamic Equilibrium
LUT	Look Up Table
MAP	Maximum A Posteriori
MODIS	Moderate Resolution Imaging Spectroradiometer
MOE	Ministry of the Environment
MSR	Mean Squared of the Residual spectrum
NIES	National Institute for Environmental Studies
OPD	Optical Path Difference
TANSO	Thermal And Near-infrared Sensor for carbon Observation
TANSO-CAI	TANSO-Cloud and Aerosol Imager
TANSO-FTS	TANSO-Fourier Transform Spectrometer

SKYNET	SKYradiometer NETwork
SIM	Spectral Irradiance Monitor
SNR	Signal to Noise Ratio
SORCE	SOlar Radiation & Climate Experiment
SPRINTARS	Spectral Radiation-Transport Model for Aerosol Species
SWIR	Short-Wavelength Infrared
TIR	Thermal Infrared
ZPD	Zero Path Difference

Appendix 2. Differences between SWIR L2 V01.xx and V02.00

As for the description of the SWIR L2 V01.xx, please refer to "Algorithm theoretical basis document (ATBD) for CO_2 and CH_4 column amounts retrieval from GOSAT TANSO-FTS SWIR (version 1.1)".

Differences of the retrieval results are shown in *Yoshida et al.* [2013].

Related Sections	Item	SWIR L2 V01.xx	SWIR L2 V02.00
2.2	Pointing mirror anomaly	Not corrected	Corrected
2.2	IMC operation status	Based on the simplified judgment	Based on the IMC "off" state list
2.2	DEM	GTOPO30	ASTER GDEM / GTOPO30
2.4	Degradation correction	Based on the degradation model provided by JAXA	Based on the degradation model described in *Yoshida et al.* [2012]
3.6	TIR Cloud Determination	Based on the simplified judgment	Not applied
3.7	2-micron band scattering material determination	Select 10 water vapor absorptive channels for each scan	Use fixed 11 water vapor absorptive channels
3.9	Spectrum quality determination	Not applied	Applied
3.10	Combined use of data screening criteria (revised item only)		Add spectrum quality determination
4.5, 6.9	Aerosol	Assume 2-km thickness layer from the surface with single type aerosol	Handle aerosol profile for two type (fine- and coarse-mode) aerosols
4.7	Solar irradiance spectra	Dr. R.L. Kurucz data	Dr. G.C. Toon data
6.2	Instrument line shape function	Apply single instrument line shape function at the center of sub-band for each sub-band	Linearly interpolated to each target wavenumber

6.4	A priori and variance and covariance matrix		Updated
6.5	Relative radiometric calibration adjustment factor	Retrieved	Not retrieved
6.10	Surface pressure	Standard deviation for a priori value is 20 hPa	Standard deviation for a priori value is 5 hPa
6.11	Temperature shift	Standard deviation for a priori value is 20 K	Standard deviation for a priori value is 5 K
6.13	Zero level offset	Not retrieved	Retrieved
7.1	Wavenumber range (revised item only)		Add 4800 ~ 4900 cm^{-1}
7.1	Simultaneously retrieved parameter (revised item only)	AOD Relative radiometric adjustment factor (ocean)	Aerosol profile Zero-level offset
7.1	Scaling of O_2 absorption cross section	Not applied	Applied
7.1	Empirical noise	Not applied	Applied
8.1	Channel selection	Applied	Not applied
9.2	Quality check items (revised item only)	SNR >= 100 χ^2 <= 5 MSR <= 3 AOD <= 0.5 Relative radiometric adjustment factor FOV-land-fraction (0% or 100%)	SNR >= 70 MSR <= threshold values AOD <= 0.1 blended albedo <= 1 surface pressure ocean wind speed FOV-land-fraction (0% or >= 60%)

Appendix 3. Differences among the SWIR L2 V02.xx

Minor versions of the SWIR L2 V02.xx products vary with the update of the TANSO-FTS L1B product and some major anomalies of the satellite system possibly affecting the TANSO-FTS L1B quality.

The details of the TANSO-FTS L1B updates and major anomalies of the satellite system are shown in *Kuze et al.* [2016].

Instrumental Line Shape Function data is available via the following web page.

https://data2.gosat.nies.go.jp/doc/document.html#TechnicalInformation

FTS L1B	SWIR L2	Observation Period	Instrumental Line Shape	Ponting Mirror	Combined use of data screening criteria
V141.141	V02.00	June 3, 2009 ~ July 31, 2010	ILSF2011	Primary	
V150.150	V02.10	August 1, 2010 ~ June 21, 2012	ILSF2011	Primary	
V150.151	V02.11	June 3, 2009 ~ April 30, 2013	ILSF2011	Primary	
V160.160	V02.20	January 1, 2010 ~ June 30, 2013	ILSF2011	Primary	Ground surface roughness determination CAI cloud determination
V161.160	V02.21	April 23, 2009 ~ May 24, 2014	ILSF2011	Primary	
< Solar-paddle-rotation anomaly : May 24, 2014 >					
V161.160	V02.31	June 16, 2014 ~ December 14, 2014	ILSF2011-800	Primary	L1 quality evaluation
< Switch to the secondary pointing mechanism : January, 2015 >					
V161.161	V02.40	February 1, 2015 ~ August 2, 2015	ILSF2011-800	Secondary	
< Cryocooler shutdown : August 2, 2015 >					
V161.161	V02.50	August 2, 2015 ~ September 13, 2015	ILSF2011-600	Secondary	
< Cryocooler restart : September 14, 2015 >					
V161.161	V02.60	September 15, 2015 ~ February 28, 2017	ILSF2011-1100	Secondary	