Validation of GOSAT SWIR XCO$_2$ and XCH$_4$ Retrieved by PPDF-S Method and Comparison with Full Physics Method

Chisa Iwasaki$^1$, Ryoichi Imasu$^1$, Andrey Bril$^2$, Tatsuya Yokota$^1$, Yukio Yoshida$^1$, Isamu Morino$^1$, Sergey Oshchepkov$^2$, Debra Wunch$^1$, David W. T. Griffith$^3$, Nicholas M. Deutscher$^1, 6$, Rigel Kivi$^7$, Dave Pollard$^8$, Coleen M. Roehl$^1$, Voltaire A. Velazco$^5$, Ralf Sussmann$^{10}$, Thorsten Warneke$^6$, and Justus Notholt$^6$

$^1$Atmosphere and Ocean Research Institute, The University of Tokyo, Kashiwa, Japan
$^2$Institute of Physics of National Academy of Sciences of Belarus, Minsk, Belarus
$^3$National Institute for Environmental Studies, Tsukuba, Japan
$^4$Department of Physics, University of Toronto, Toronto, Canada
$^5$Centre for Atmospheric Chemistry, University of Wollongong, New South Wales, Australia
$^6$Institute of Environmental Physics, University of Bremen, Bremen, Germany
$^7$Arctic Research Centre, Finnish Meteorological Institute (FMI), Sodankylä, Finland
$^8$National Institute of Water and Atmospheric Research, Lauder, New Zealand
$^9$California Institute for Technology, Pasadena, CA, USA
$^{10}$IMK-IFU, Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany

Abstract

Column-averaged concentrations of carbon dioxide (XCO$_2$) and methane (XCH$_4$) were retrieved from spectra observed by the Greenhouse gases Observing SATellite (GOSAT) using the so-called Photon path length Probability Density Function-Statistical Method (PPDF-S) retrieval method, which explains cloud/aerosol effects in terms of light path modification. The PPDF-S method, as well as the standard products for General Users (GU) of XCO$_2$ and XCH$_4$ retrieved using the full physics (FP) method, were validated through comparison with Total Carbon Column Observing Network (TCCON) data. Results show that bias and its standard deviation of XCO$_2$ over the land are 0.73 and 1.83 ppm for the PPDF-S data, and −0.32 and 2.16 ppm for GU products. For XCH$_4$, they are 1.4 and 14.1 ppb, and −1.9 and 12.5 ppb, respectively. Although the magnitude relations between XCO$_2$ and XCH$_4$ retrieved by the PPDF-S and GU products are identical over the land, they differ over the ocean. This fact emphasizes the importance of additional validation data over the ocean. Results also show that 68% of FP data that were screened out through an Aerosol Optical Thickness (AOT) test passed all screening tests for the PPDF-S method, implying the applicability of the PPDF-S method to denser aerosol conditions.

(Citation: Iwasaki, C., R. Imasu, A. Bril, T. Yokota, Y. Yoshida, I. Morino, S. Oshchepkov, D. Wunch, D. W. T. Griffith, N. M. Deutscher, R. Kivi, D. Pollard, C. M. Roehl, V. A. Velazco, R. Sussmann, T. Warneke, and J. Notholt, 2017: Validation of GOSAT SWIR XCO$_2$ and XCH$_4$ retrieved by PPDF-S method and comparison with Full Physics method. SOLA, 13, 168–173, doi: 10.2151/sola.2017-031.)

1. Introduction

Because of human activity, atmospheric concentrations of carbon dioxide (CO$_2$) and methane (CH$_4$) have increased since the Industrial Revolution. To estimate sources and sinks of CO$_2$ and CH$_4$, the Greenhouse gases Observing SATellite (GOSAT) was launched in 2009. It measures backscattered sunlight by the surface or atmosphere of the earth in Short-Wavelength InfraRed (SWIR) with near-global coverage. Column-averaged concentrations of carbon dioxide (XCO$_2$) and methane (XCH$_4$) can be retrieved from the measured spectra.

Satellite-retrieved XCO$_2$ and XCH$_4$ tend to show some biases, mainly attributable to light scattering by clouds and aerosols. In the Full Physics (FP) retrieval method that is used widely for gas retrieval, the multiple light scattering processes are calculated explicitly by solving radiative transfer equations accounting for optical properties of aerosols under a clear sky condition (Morino et al. 2011; Yoshida et al. 2011, 2013).

A more simplified estimation method, a photon path length probability density function (PPDF)-based method, has been proposed (e.g., Bril et al. 2007; Oshchepkov et al. 2008, 2009, 2013). The method can represent the effects of non-negligible light path modification without exact radiative transfer simulations because the light path modification is evaluated using PPDF statistically. In the new version of the PPDF-based method, the PPDF-simultaneous (PPDF-S) method, gas concentrations are retrieved simultaneously with PPDF parameters and its retrieval uncertainty for GOSAT was evaluated (Oshchepkov et al. 2013). However, because evaluation was made only for XCO$_2$ and because the data period was limited, further validation is necessary.

The objective of this study is to clarify the uncertainty and global characteristics of PPDF-S retrieval results for both XCO$_2$ and XCH$_4$ for an extended data period through comparison of them with ground-based Total Carbon Column Observing Network (TCCON) data that are generally used as reference for satellite measurements (Wunch et al. 2011a; Blumenstock et al. 2014; Wunch et al. 2015). For comparison, we also analyzed data of XCO$_2$ and XCH$_4$ retrieved using the FP method and released by the National Institute for Environmental Studies (NIES).

2. GOSAT retrieval method and data used for this study

2.1 PPDF-S method

The PPDF data reflect statistical characteristics of the photon path length. Using two equations, the radiance, as affected by scattering and absorption in a homogeneous medium $I(k)$, can be represented as the radiance observed at an absorption-free wavelength that is close to the absorption band $I_0(k = 0)$ and PPDF.
\[ I(k) = I_s T(k) = I_s \int_0^\infty \exp(-kL)P(L) dL \]
\[ \int_0^\infty P(L) dL = 1, \]

where, \( T \) is transmittance, \( L \) the photon path length, \( k \) the volume absorption coefficient, and \( P(L) \) denotes the PPDF with respect to the photon geometrical path (Bennartz and Preusker 2006).

According to Bril et al. (2007), the equations above are applicable for retrieving greenhouse gas concentration from satellite observational data. Herein, PPDF is explained using PPDF parameters accounting for single light scattering by a cloud/aerosol layer as well as multiple light scattering/reflection between the layer and the ground surface. Oshchepkov et al. (2009) pointed out that the PPDF can be represented by four PPDF parameters: the altitude of the cloud/aerosol layer \( h \), the relative reflection of the cloud/aerosol layer \( \alpha \), which represents the relative fraction of the photons scattered singly to the detector; \( \rho \), which characterizes the relative mean path length between the layer and ground surface; and an adjustment parameter \( \gamma \) accounting for multiple scattering/reflection.

Oshchepkov et al. (2013) investigated the uncertainty of XCO retrieved using the PPDF-S method (hereinafter called PPDF-S data) for June 2009 through July 2010. The present study evaluated both XCO and XCH\(_4\) retrieved using the PPDF-S method from the spectral radiance products labeled as V161.160 for June 2009 through May 2014. Hereinafter, we use labeling names, Level 1B (L1B) and Level 2 (L2), respectively for spectral radiance and XCO and XCH\(_4\).

### 2.2 NIES standard products

NIES releases the L2 data as the standard products for General Users (i.e., GU products) retrieved using FP method, which is widely used for gas retrieval (Yoshida et al. 2011, 2013). Morino et al. (2011) reported the uncertainty of GU products derived from April 2009 through October 2009 by comparing them with TCCON data. Yoshida et al. (2013) validated GU products in the same way for a longer period of June 2009 through July 2010, but the version of L1B data used for retrieval is a mix of V141.141, V150.150, and V150.151. We used a later version of GU products than those used by Yoshida et al. (2013), which were retrieved from L1B in V161.160 for June 2009 through May 2014. They are labeled as V02.21 of the L2 data. We also used the L2 data released only for Research Announcement principal investigators (RA products). Differences between GU products and RA products are just criteria for data screening tests applied to control the data quality. The GU products are selected from RA products using seven screening tests with thresholds of four parameters (e.g., aerosol optical thickness), which means that the screening criteria for GU products are stricter than those for RA products (Yoshida et al. 2011, 2013). The numbers of observational scenes used in this study are 494,259 and 878,762, respectively, for GU and RA products.

### 2.3 TCCON data and validation methodology

The TCCON is a global network of ground-based Fourier Transform Spectrometers (FTS) measuring the solar spectrum in the near-infrared almost continuously under a clear sky condition, to retrieve column abundances of CO\(_2\), CO, CH\(_4\), N\(_2\)O and other molecules (Wunch et al. 2011a). Although the observational principles of TCCON and GOSAT are similar, TCCON measures the direct solar spectrum with higher spectral resolution and better calibrated. Accordingly, TCCON data have been used to validate GOSAT data because they are more reliable than those obtained from satellite measurements.

For this study, we used TCCON data from 11 sites presented in Table 1. An open source software package (GGG) is used to calculate total column abundances of the gases from measured spectra. TCCON data analyzed using the GGG2009 and GGG2012 version of retrieval algorithm were used respectively by Morino et al. (2011) and by Yoshida et al. (2013). For this study, we used data analyzed using the newest version of the algorithm: GGG2014.

To validate GOSAT data using TCCON data, GOSAT data within a 4 × 4 degree grid box centered at each TCCON site and TCCON data measured within ± 30 min of GOSAT overpass time were selected. To represent the GOSAT data uncertainty, we define the three parameters presented below.

- \( \text{diff.} \): (GOSAT data) – (TCCON data)
- \( \bar{B} \): mean value of \( \text{diff.} \)
- \( \sigma \): standard deviation of \( \text{diff.} \)

| Site            | Latitude | Longitude | Dataset reference                                                                 | TCCON software version and release number |
|-----------------|----------|-----------|-----------------------------------------------------------------------------------|------------------------------------------|
| Bialystok       | 53.23 N  | 23.03 E   | Deuser et al., 2015. doi:10.14291/tccon.ggg2014.bialystok01.R1/1183984.            | GGG2014.R1                               |
| Bremen          | 53.10 N  | 8.85 E    | Notholt et al., 2014. doi:10.14291/tccon.ggg2014.bremen01.R0/1149275.              | GGG2014.R0                               |
| Garmisch        | 47.48 N  | 11.06 E   | Sussmann et al., 2014. doi:10.14291/tccon.ggg2014.garmisch01.R0/1149299.           | GGG2014.R0                               |
| Orléans         | 47.97 N  | 2.11 E    | Warneke et al., 2014. doi:10.14291/tccon.ggg2014.orleans01.R0/1149276.             | GGG2014.R0                               |
| Lamont          | 36.60 N  | 97.49 W   | Wennberg et al., 2016. doi:10.14291/tccon.ggg2014.lamont01.R1/1255070.            | GGG2014.R1                               |
| Park Falls       | 67.37 N  | 90.27 W   | Wennberg et al., 2014. doi:10.14291/tccon.ggg2014.parkfalls01.R0/1149161.         | GGG2014.R0                               |
| Sodankylä       | 67.37 N  | 26.63 E   | Kivi et al., 2014. doi:10.14291/tccon.ggg2014.sodankyla01.R0/1149280.             | GGG2014.R0                               |
| Tsukuba125HR    | 36.05 N  | 140.12 E  | Morino et al., 2016. doi:10.14291/tccon.ggg2014.tsukubao2.R1/1241486.             | GGG2014.R1                               |
| Darwin          | 12.43 S  | 130.89 E  | Griffith et al., 2014a. doi:10.14291/tccon.ggg2014.darwin01.R0/1149290.           | GGG2014.R0                               |
| Lauder125HR     | 45.05 S  | 169.68 E  | Sherlock et al., 2014. doi:10.14291/tccon.ggg2014.laudero2.R0/1149298.            | GGG2014.R0                               |
| Wollongong      | 34.41 S  | 150.88 E  | Griffith et al., 2014b. doi:10.14291/tccon.ggg2014.wollongong01.R0/1149291.      | GGG2014.R0                               |
Generally, when column-averaged gas concentrations observed using instruments of different types are compared, differences in a priori profiles assumed in analyses and the column averaging kernels, which represent the sensitivity profiles, should be considered. Although Wunch et al. (2011b) reported that the bias in XCO$_2$ attributed to the differences in a priori and averaging kernel is about 0.6 ppm for GOSAT data at a TCCON site (Lamont, CA, USA), we do not consider the differences in this study so as to make our validation method to be same as that of an earlier study by Yoshida et al. (2013).

3. Results

3.1 Validation results of global data

Figures 1a and 1b respectively depict global maps of XCO$_2$ and XCH$_4$ of PPDF-S data and GU products. All global maps in this paper describe the mean value in a grid of 2.5 × 2.5 degree box for the whole data period of June 2009 − May 2014. The most distinctive difference over land between PPDF-S data and GU products is found over the middle of Africa, where biomass burning often occurs and where many aerosols exist. In this region, the PPDF-S data have lower XCO$_2$ and XCH$_4$ than GU products, but there are no TCCON data available for validation.

To investigate the temporal variation of biases, time series of monthly biases of PPDF-S data and GU products and their standard deviations averaged over the 11 TCCON sites are presented in Fig. 2. The biases are not uniform during the operational period. Moreover, the temporal features in the variabilities are not the same for XCO$_2$ and XCH$_4$. The bias of XCO$_2$ of GU products has a larger negative bias in 2009 and 2010, which then decreased to zero around 2011 until 2014. However, the XCH$_4$ of PPDF-S data has a positive bias in 2009 and 2010. In later analyses, we shall present validation results separately for 2009–2010 and 2011–2014.

Tables 2 and 3 present the validation results obtained using TCCON data observed at 11 sites from June 2009 through May 2014. Table 2 shows the difference in uncertainty of PPDF-S data and GU products. Results for 2009–2010 (data period (1)) and for 2011–2014 (data period (2)) are also presented separately in the

Fig. 1. (a) Global maps of the mean value of XCO$_2$ (left) and XCH$_4$ (right) in a grid of 2.5 × 2.5 degree box for the whole data period of June 2009−May 2014 for PPDF-S data: (b) as in (a), but for GU products.

Fig. 2. Time series of monthly biases of PPDF-S data (red line) and GU products (blue line) and its deviations: the left panel is for XCO$_2$; the right one is for XCH$_4$.
bias are similar between PPDF-S data and GU products, but the absolute value of $B$ is extremely large for PPDF-S data in Darwin. Furthermore, $XCH_4$ of PPDF-S data is positive, whereas that of GU products is negative.

The slope of the regression line. Results show that the orders of the bias are similar between collocated GOSAT and TCCON data; and $\sigma$ is the standard deviation (in the same units as $B$), $r$ is the correlation coefficient between collocated GOSAT and TCCON data; $a$ is the slope of the regression line. Results show that the orders of the bias are similar between PPDF-S data and GU products, but $B$ of PPDF-S data is positive, whereas that of GU products is negative. Furthermore, $XCO_2$ of both PPDF-S data and GU products have larger absolute values of $B$ and smaller $\sigma$ in period (1) than those in period (2), as depicted in Table 2. Actually, $XCH_4$ of both PPDF-S data and GU products have larger $\sigma$ in (2) than those in (1), but the absolute value of $B$ is extremely large for PPDF-S data in (1) and for GU products in (2). Among the 11 TCCON site data, $B$ of $XCO_2$ for PPDF-S data at Darwin is negative. This point will be discussed in Section 4.

To illustrate the spatial distribution of the difference between PPDF-S data and GU products, we present a global map of the difference between the two products in Fig. 3 by subtracting PPDF-S data from GU products. Values are averages over the whole data period in a grid of 2.5° × 2.5° degree box.

Table 2. Validation results evaluated using TCCON data observed at 11 sites during June 2009–May 2014 for common cases between PPDF-S data and GU products: $N$, number of data; $B$, bias; $\sigma$, standard deviation.

| Site         | $N$ | $B$ [ppm] | $\sigma$ [ppm] | $r$ | $a$ | $B$ [ppb] | $\sigma$ [ppb] | $r$ | $a$ |
|--------------|-----|-----------|----------------|-----|-----|-----------|----------------|-----|-----|
| Bialystok    | 58  | 1.00      | 1.63           | 0.26 | 1.96 | 4.1       | 10.6           | 1.7  | 10.5 |
| Bremen       | 36  | 2.05      | 1.56           | 0.57 | 1.50 | -2.5      | 14.1           | -1.3 | 9.1  |
| Garmisch     | 94  | 1.39      | 1.77           | 0.01 | 2.19 | 4.7       | 14.0           | 3.3  | 15.4 |
| Orléans      | 199 | 0.81      | 1.40           | -0.54| 1.82 | -3.6      | 13.1           | -3.4 | 10.5 |
| Lamont       | 598 | 0.42      | 1.63           | -1.42| 1.68 | 5.5       | 12.4           | -4.4 | 11.5 |
| Park Falls   | 173 | 0.83      | 1.89           | -0.07| 1.91 | 4.0       | 13.1           | 3.7  | 10.0 |
| Sodankylä   | 404 | 1.68      | 1.50           | 0.59 | 1.65 | 3.8       | 9.8            | 1.2  | 9.3  |
| Tsukuba      | 277 | 2.05      | 2.05           | 1.66 | 2.14 | -1.3      | 14.4           | 3.7  | 11.8 |
| Darwin       | 150 | -0.52     | 1.38           | -0.73| 1.64 | -14.5     | 8.7            | -4.5 | 7.6  |
| Lauder       | 132 | 0.34      | 1.25           | -0.35| 1.18 | 8.8       | 11.0           | -1.7 | 8.2  |
| Wollongong   | 352 | 0.47      | 1.96           | -0.25| 2.51 | 1.0       | 15.0           | -5.3 | 16.0 |
| Total        | 2118| 0.73      | 1.83           | -0.32| 2.16 | 1.4       | 14.1           | -1.9 | 12.5 |

Table 3. Averaged validation results obtained using TCCON data observed at 11 sites for all cases of PPDF-S data and GU products for the whole data period: $N$, $B$ and $\sigma$ are the same as in Table 2; $r$ is the correlation coefficient between collocated GOSAT and TCCON data; $a$ is the slope of the regression line.

| Site         | $N$ | $B$ [ppm] | $\sigma$ [ppm] | $r$ | $a$ | $B$ [ppb] | $\sigma$ [ppb] | $r$ | $a$ |
|--------------|-----|-----------|----------------|-----|-----|-----------|----------------|-----|-----|
| Bialystok    | 58  | 0.47      | 2.09           | 0.89 | 0.96 | 0.66      | 15.41          | 0.87 | 0.95 |
| Bremen       | 36  | -0.29     | 2.33           | 0.88 | 0.99 | -2.24     | 13.14          | 0.91 | 0.99 |

Fig. 3. Differences in $XCO_2$ (left) and $XCH_4$ (right) between PPDF-S data and GU products that are calculated using subtracting PPDF-S data from the GU products. Values are averages over the whole data period in a grid of 2.5° × 2.5° degree box.
optical parameters of aerosols (Yoshida et al. 2011).

To investigate the uncertainty of GOSAT data with large AOT, we used AOT estimated in the FP method as a measure of the aerosol concentration, and compared data removed from RA products through the AOT test with the PPDF-S data for the same cases. Table 4 presents results for 2009–2010, 2011–2014, and 2009–2014. The results for 2009–2014 showed that 144 observational scenes (≈68%) out of the 211 scenes passed all the PPDF-S screening tests. When the 144 cases are validated with the TCCON data, both $\sigma$ and $\mu$ of XCO$_2$ of the PPDF-S data is smaller than that of the removed RA products. Contrary to the results of XCO$_2$, $\sigma$ of XCH$_4$ of the PPDF-S is larger than that of the RA products.

### 4. Discussion

From the validation results shown in Tables 2 and 3, the magnitudes of $\sigma$ and $\mu$ were found to be almost identical between PPDF-S data and GU products. Therefore, it can be inferred that PPDF-S data have some potential for use in source–sink inversion analyses of gases as well as GU products after appropriate bias correction, as far as $\sigma$ does not change so much after bias correction. Results also show that the characteristics of the biases have changed since 2010, as depicted in Table 2 and Table 3. This change might result from changes in the spectral radiance characteristics. This point should be investigated further in future studies.

As shown in Table 2, $\mu$ of XCO$_2$ of PPDF-S data is negative only at Darwin, which is the only site where collocated GOSAT data were observed over the ocean. This negativity implies that XCO$_2$ observed over the ocean and analyzed by the PPDF-S method has generally negative $\mu$. Furthermore, as portrayed in Fig. 3, large differences exist in both XCO$_2$ and XCH$_4$ between PPDF-S data and GU products over the ocean. The difference of XCH$_4$ can be as great as 90 ppb. Saitoh et al. (2012) compared the XCH$_4$ of GU product with aircraft data and reported that the difference was ≈8 ppb with the standard deviation of 10 ppb. XCH$_4$ of PPDF-S data might have larger biases compared with GU products over the ocean. It is important to validate the GOSAT data over the ocean to ensure the accuracy and applicability of the PPDF-S method to denser aerosol conditions. To ascertain which method is superior under high aerosol conditions, simulation studies or validation using aerosol measurements are necessary for future researches.

### 5. Summary and concluding remarks

XCO$_2$ and XCH$_4$ retrieved using the PPDF-S method and GU products have been validated using TCCON data at 11 sites. The uncertainties of PPDF-S data and GU products, defined as $\sigma$ and $\mu$, are almost equal for both XCO$_2$ and XCH$_4$. Actually, PPDF-S data have positive $\mu$ in both XCO$_2$ and XCH$_4$. By contrast, GU products show negative $\mu$. The biases also show temporal variation during 2009–2014, with changing characteristics of the biases since 2010.

We found large differences in both XCO$_2$ and XCH$_4$ between PPDF-S data and GU products over the ocean. It is necessary to increase the number of validation data to assess the uncertainty of GOSAT data over the ocean. At the same time, PPDF-S data have smaller values of both XCO$_2$ and XCH$_4$ than those of GU products, particularly over the middle of Africa, where many aerosols might exist because of biomass burning.

GU products were selected from RA products using some screening tests. $\sigma$ and $\mu$ of XCO$_2$ of PPDF-S data are smaller than those of the data removed from RA products through the AOT test for the comparable observational scenes. From results of this analysis, we inferred that PPDF-S and FP methods exhibit different performance under high-AOT conditions implying the applicability of the PPDF-S method to denser aerosol conditions.

To ascertain which method is superior under high aerosol conditions, simulation studies or validation using aerosol measurements are necessary for future researches.

### Acknowledgments

This study was conducted under the framework of joint research between NIES and Atmosphere and Ocean Research Institute (AORI) of The University of Tokyo. The investigation was partially supported by the Belarusian Republican Foundation for Fundamental, Research Agreement No. F15SB-023, and the Bilateral Program of the Japan Society for the Promotion of Science (JSPS). TCCON data were obtained from the TCCON Data Archive (doi:10.14291/tccn.archive/1348407), hosted by the Carbon Dioxide Information Analysis Center (CDIAC) at Oak Ridge National Laboratory, Oak Ridge, Tennessee, US. ACOS/OCO-2 absorption coefficient (ABSCO) tables were provided by the Jet Propulsion Laboratory (JPL) and were used for radiative transfer calculation. I received generous support from Sachiko Hayashida of Nara Women’s University. Nicholas Deutscher is supported by an Australian Research Council Discovery Early Career Researcher Award, DE140100178.

Edited by: S. Morimoto

### References

Aben, I., O. Hasekamp, and W. Hartmann, 2007: Uncertainties in the space-based measurements of CO$_2$ columns due to scattering in the Earth’s atmosphere. J. Quant. Spectrosc. Radiat. Transfer, 104, 450–459, doi:10.1016/j.jqsrt.2006.09.013.

Bennartz, R., and R. Preusker, 2006: Representation of the photon pathlength distribution in a cloudy atmosphere using finite elements. J. Quant. Spectrosc. Radiat. Transfer, 98, 202–219, doi:
