NIES-GOSAT2-ALG-20240206-028-00

GOSAT-2 Level 4
Algorithm Theoretical Basis Document

March 2024

National Institute for Environmental Studies
GOSAT-2 Project

SAITO Makoto[1)]
SAEKI Tazu[1)]
MURAKAMI Kazutaka[1)]
NIWA Yosuke[1)]

1) Earth System Division, National Institute for Environmental Studies

Revision History

Version	Revised	Page	Description
00	Mar. 2024	-	-

Table of Contents

1 Introduction

1.1 Objective

This document describes the algorithm and processing method for the Greenhouse gases Observing SATellite-2 (GOSAT-2) Level 4 (L4) carbon dioxide (CO_2) Product and provides an overview of the latest version (01.02). The algorithm is intended to estimate the global surface CO_2 flux based on the GOSAT-2 Thermal And Near infrared Sensor for carbon Observation-Fourier Transform Spectrometer-2 (TANSO-FTS-2) short wavelength infrared (SWIR) Level 2 (L2) Column-averaged Dry-air Mole Fraction Product, as well as the global CO_2 distribution. The system consists of an atmospheric tracer transport model, an inverse analysis scheme, and *a priori* information. This document provides a description and references for each of these components.

1.2 Product revision history

Table 1. Revision history

Version	Date	Author	Description
01.01	4 December 2022	M. Saito	Initial version
01.02	22 March 2024	M. Saito	Second delivery version

2 GOSAT-2 observations

GOSAT-2 is a satellite dedicated to greenhouse gas observations of CO_2 and methane (CH_4). The satellite carries a Fourier transform spectrometer (TANSO-FTS-2) and a push-broom imaging radiometer TANSO Cloud and Aerosol Imager-2 (TANSO-CAI-2). TANSO-FTS-2 measures SWIR sunlight reflected from Earth's surface and thermal infrared (TIR) radiation emitted from the ground and Earth's atmosphere. TANSO-FTS-2 has a high spectral resolution of 0.2 cm^{-1} and operates in five spectral bands: three in the SWIR spectral range (0.75–0.77, 1.56–1.69, and 1.92–2.33 μm; bands 1, 2, and 3, respectively), and two in the TIR spectral range (5.5–8.4 and 8.4–14.3 μm; bands 4 and 5, respectively). Column-averaged dry-air mole fractions of CO_2 and CH_4 (denoted as XCO_2 and XCH_4, respectively) are retrieved using the 1.6 and 2.0 μm bands for CO_2 and the 1.6 and 2.3 μm bands for CH_4. TANSO-FTS-2 spectral data can also resolve carbon monoxide (CO) using the 2.3 μm band, in addition to XCO_2 and XCH_4. Spectral radiance in the two TIR bands is used to obtain information on vertical profiles of atmospheric concentrations of CO_2 and CH_4. TANSO-FTS-2 has an intelligent pointing mechanism that immediately identifies cloud positions in the field of view using an onboard camera and points to a cloud-free location. The camera has a spatial resolution of ~0.1 km with 608 × 1024 pixels over 30 km in the along-track field and 50 km in the cross-track field.

TANSO-CAI-2 has five observation bands for forward viewing at 343, 443, 674, 869, and 1630 nm, and backward viewing at 380, 550, 674, 869, and 1630 nm. It provides data for identifying clouds and aerosol conditions in the cross-track field over a distance of 1000 km.

The instruments on GOSAT-2 have been described in detail by *Suto et al.* (2021).

GOSAT-2 flies in a sun-synchronous orbit at an altitude of 613 km. The equator-crossing local time of the descending node is 13:00 with a repeat cycle of 6 days (inclination angle of 98.0° ± 0.1°). The pointing mechanism for TANSO-FTS-2 covers a range of ±40° in the along-track direction and ±35° in the cross-track direction. The observation interval of TANSO-FTS-2 is 4.024 s, with a nominal turnaround time of 0.65 s required for changing the pointing location, taking an image, identifying cloud locations in the image, and repointing to a cloud-free location. The field of view of TANSO-FTS-2 is 15.8 mrad for all bands, and the instantaneous ground field of view is a circle with diameter 9.6 km.

3 Product design

The GOSAT-2 mission is designed to enhance the space-borne measurements of major greenhouse gases that began with GOSAT observations, and to monitor the impacts of climate change and human activities on the carbon cycle. GOSAT observations have improved the accuracy of single shot measurements of greenhouse gases (*Yokota et al.*, 2009) relative to former satellite missions, providing confidence in the use of XCO_2 and XCH_4 data from space to constrain models that generate global flux estimates using atmospheric inversions (e.g., *Maksyutov et al.*, 2013). However, GOSAT observations provide sparse coverage as a trade-off for the high data quality, resulting in difficulties in quantifying regional fluxes using satellite observations alone. This limited observational coverage, together with the need to minimize the computational cost of atmospheric inversion, means that flux estimates from GOSAT data are resolved on a sub-continental scale (64 regions over the globe) using the National Institute for Environmental Studies (NIES) atmospheric tracer transport model (NIES-TM) and a fixed-lag Kalman smoother with ground-based observations and GOSAT XCO_2 and XCH_4 observations (*Maksyutov et al.*, 2013). This combined application of ground-based and satellite observations to atmospheric inversion allows more accurate inverse estimation of sub-continental fluxes; however, it does not allow a quantitative assessment of the degree to which satellite observations contribute to filling the gaps in greenhouse gas observations for carbon flux estimates at regional and even national scales.

TANSO-FTS-2 measures XCO_2 and XCH_4 over land with better sampling (more than twice the sampling rate) than TANSO-FTS, the main sensor aboard GOSAT, using an intelligent pointing mechanism. In addition, TANSO-FTS-2 has wider pointing angles than those of TANSO-FTS, especially in the along-track direction, allowing wider coverage of observation locations, which contributes to an increase in the available sun glint points over the ocean. Observations by GOSAT-2, which is equipped with enhanced versions of the instruments aboard GOSAT, are expected to facilitate the use of satellite observations in carbon cycle assessments and further improve the spatial resolution of flux estimates to better understand regional sources and sinks. The improvement in satellite observations afforded by GOSAT-2 is illustrated by the fact that using GOSAT-2 observations alone, the GOSAT-2 L4 Product estimates global surface CO_2 and CH_4 fluxes with a higher spatial resolution than those estimated by GOSAT. The atmospheric transport model and inverse scheme used for the GOSAT-2 L4 Product represent an upgraded version of the model system used in the GOSAT mission (NIES-TM with a fixed-lag Kalman smoother). The new model system, which is the Non-hydrostatic Icosahedral Atmospheric Model (NICAM)-based Inverse Simulation for Monitoring CO_2 (NISMON-CO_2), as described by *Niwa et al.* (2021), improves the spatial resolution of flux estimates. The NISMON-CO_2 consists of a NICAM-based transport model (NICAM-TM; *Niwa et al.*, 2011) for forward simulation and an atmospheric inversion using the four-dimensional variational (4D-Var) method (*Niwa et al.*, 2017a, b). The NISMON-CO_2 is operated on an icosahedral grid obtained by five iterations of recursive division (glevel-5, horizontal spatial resolution of ~223 km) and 40 vertical layers with a time step of 20 min. Information on the GOSAT-2 L4 Product plays an important role in assessing the robustness of GOSAT-2 measurements and the contribution of GOSAT-2 observations to the identification of regional sources and sinks.

4 Algorithm description

4.1 Overview

The GOSAT-2 L4 Product consists of global surface CO_2 and CH_4 flux estimates from GOSAT-2 XCO_2 and XCH_4 data, and three-dimensional fields of atmospheric CO_2 and CH_4 concentrations that are simulated using the estimated surface fluxes. The GOSAT-2 L4 product is provided by simulation frameworks that make up the GOSAT-2 L4 computational system.

The global surface CO_2 flux is estimated using NISMON-CO_2 in the context of Bayesian inference. The Bayesian least-squares estimate is obtained by minimizing the cost function as follows:

$$\mathbf{J} = \frac{1}{2}(\mathbf{x} - \mathbf{x}_{\mathrm{pri}})^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_{\mathrm{pri}}) + \frac{1}{2}(\mathbf{Mx} - \mathbf{y})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{Mx} - \mathbf{y}), \tag{1}$$

where $\mathbf{x}$ and $\mathbf{x}_{\mathrm{pri}}$ are vectors for modeled and *a priori* source and sink strengths, respectively; $\mathbf{M}$ is a matrix of a linear forward transport model used to obtain estimates of concentrations at each measurement; $\mathbf{y}$ is a vector of observed concentrations; and $\mathbf{B}$ and $\mathbf{R}$ are error covariance matrices for the *a priori* flux estimates and the misfit of concentrations between observations and model predictions, respectively. The superscript T denotes the transpose operator. In practical operation of NISMON-CO_2, Eq. (1) is replaced with $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_{\mathrm{pri}}$ as follows:

$$\mathbf{J} = \frac{1}{2}\delta\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{M}\delta\mathbf{x} - \mathbf{d})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{M}\delta\mathbf{x} - \mathbf{d}), \tag{2}$$

where $\mathbf{d} = \mathbf{y} - \mathbf{Mx}_{\mathrm{pri}}$.

In performing the inversion, GOSAT-2 XCO_2 data are the primary source of the observations, $\mathbf{y}$, for deducing global surface CO_2 fluxes, $\mathbf{x}$. The *a priori* CO_2 source and sink data, $\mathbf{x}_{\mathrm{pri}}$, as used in the GOSAT-2 L4 CO_2 computational system, consist of six types: monthly fossil fuel CO_2 ($FFCO_2$) emissions; hourly gross primary productivity (GPP); hourly ecosystem respiration (RE); monthly land use change (LUC) emissions; monthly biomass burning (BB) emissions; and monthly ocean–atmosphere (OCN) CO_2 exchanges. In the initial version 01.01, the error covariance matrix for model–observation misfit of concentrations, $\mathbf{R}$, was determined based on the difference between the retrieved and simulated XCO_2 values at each measurement. It was assumed that the values simulated using surface fluxes estimated from measurements by global networks of near-surface atmospheric observations provide relevant variability in XCO_2 over regional and global scales. Under these conditions, as the misfit becomes larger the inversion does not focus on fitting the retrieved values at the expense of a smaller misfit, so the *a posteriori* fluxes can result in behavior similar to those being constrained by the near-surface atmospheric observations. However, this implies that the near-surface atmospheric observations are used primarily as a constraint on the *a priori* CO_2 source and sink data, and there is little use for GOSAT-2 XCO_2 data. Hence, in version 01.02, to evaluate the ability of the GOSAT-2 XCO_2 data to deduce the global surface CO_2 fluxes, $\mathbf{R}$ is represented using a uniform value ($\mathbf{R}$ = 4 ppm) for all retrieved concentrations. For the error covariance matrix $\mathbf{B}$, we apply arbitrary scaling factors to *a priori* strengths. In the GOSAT L4 Product, the global surface CO_2 fluxes are deduced by optimizing only the *a priori* information in terms of net ecosystem exchange and OCN fluxes, with $FFCO_2$ and BB emissions being prescribed with an assigned uncertainty of zero

(*Maksyutov et al.*, 2013). It remains unclear to what extent the spatiotemporal variations of natural fluxes appear over the globe relative to anthropogenic emissions. Therefore, the GOSAT-2 L4 Product optimizes all *a priori* fluxes except for $FFCO_2$ emissions in the estimate of global fluxes; i.e., uncertainty is distributed to five *a priori* fluxes: 30%, 30%, 100%, 100%, and 20% for the GPP, RE, LUC, BB, and OCN fluxes. For the GPP, RE, and OCN fluxes, off-diagonal elements of the error covariance matrix $\mathbf{B}$ are prescribed using a Gaussian function with spatial error correlation lengths of 500, 500, and 1000 km, respectively (*Niwa et al.*, 2017b).

The outputs of the GOSAT-2 L4 CO_2 computational system are an estimate of the monthly averaged global surface CO_2 flux at a spatial resolution of 2.5° and 6-hourly atmospheric CO_2 concentrations on a three-dimensional grid with the same horizontal resolution (2.5°) and 17 pressure levels in the vertical along with a near-surface level. These outputs are provided as the GOSAT-2 L4A Global CO_2 Flux Product and the GOSAT-2 L4B Global CO_2 Distribution Product, respectively.

4.2 Processing outline

The GOSAT-2 L4 computational system was constructed on the NEC SX-Aurora TSUBASA A511-64 supercomputer at NIES, which features a maximum of 256 nodes, each with eight cores; the vector processor has a peak performance of up to 622.8 teraflops. The HPE Apollo2000 scalar computer with a peak performance of 86.0 teraflops at NIES is also used for preprocessing to convert the input information to the model grid data.

The process of deducing the global surface CO_2 flux from GOSAT-2 XCO_2 data begins with the assembly of various input data that are required for operation of the system. The input data are reanalyzed meteorological fields, *a priori* CO_2 sources and sinks, and observations of atmospheric CO_2. In the simulation of atmospheric transport, horizontal winds of the model are nudged toward those of the reanalysis to reproduce past and current atmospheric transport fields. In the operation of the GOSAT-2 L4 computational system, an atmospheric tracer transport simulation is first performed with nudging to generate and archive three-dimensional transport fields (air mass density, air mass flux, vertical diffusion coefficient, water substances, temperature, and cumulus base mass flux). The instantaneous values of these fields are archived every hour for the cumulus base mass flux and every 3 h for other variables, excluding the air mass flux. For the air mass flux, the variables are averaged every 3 h to maintain better consistency with continuity (CWC). The archive data are then used as input for an iterative operation of the atmospheric tracer transport model to deduce the surface fluxes using a four-dimensional variational (4D-Var) method (*Niwa et al.*, 2017a, b; see Section 4.4).

The *a priori* CO_2 source and sink data are prepared for a given analysis period and interpolated onto the model grid of the atmospheric tracer transport model. GOSAT-2 XCO_2 data are used as the atmospheric observational data to drive the GOSAT-2 L4 CO_2 computational system, and a ground-based atmospheric observational dataset is used as ancillary data to prepare an initial field of atmospheric CO_2. In system operation, using the ground-based atmospheric observation data alone, an inversion is first performed to infer monthly *a posteriori* fluxes just before the analysis period and simulate corresponding three-dimensional atmospheric CO_2 variability with a forward simulation using the *a posteriori* fluxes for the initial field data. Then, we calculate the monthly global surface CO_2 fluxes and their three-dimensional variability in the atmosphere over the analysis period. The forward and backward simulations of atmospheric CO_2 are performed for a duration of 2 months before the analysis period and 2 months after.

4.3 Input data

4.3.1 Meteorological reanalysis data

Horizontal winds in the atmospheric tracer transport model are nudged using the Japanese 55-year Reanalysis data (JRA-55; *Kobayashi et al.*, 2015). Reanalysis data are used in our model system for the u- and v-components of wind ("anl mdl ugrd" and "anl mdl vgrd"; m s^{-1}) in the TL319 model grid field with 60 hybrid vertical levels. The JRA-55 horizontal winds are provided on a 6 h time step at 0000, 0600, 1200, and 1800 UTC. As 40 vertical layers are implemented in the atmospheric tracer transport model used in our system, the vertical coordinate system in the JRA-55 horizontal wind data is interpolated to that of the atmospheric tracer transport model; subsequently, the horizontal winds of the model simulation are nudged every 6 h to the wind fields in JRA-55.

4.3.2 *A priori* fluxes

To prescribe $FFCO_2$ emissions in our system, we use the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC version "ODIAC2020 FFCO2 emission dataset"; *Oda et al.*, 2018). ODIAC provides data products for $FFCO_2$ emissions at 1-km and 1° grid resolutions at monthly time step. The 1° data are used in our system. ODIAC data comprise FF emissions from FF combustion, cement production, and gas flaring over land, and international bunker emissions from international aviation and marine bunkers and Antarctic fisheries over the ocean. An aggregation of both sets of emissions is used as $FFCO_2$ emissions in our system. The ODIAC ver. ODIAC2020 covers the period between 1979 and 2019. To extend the period of ODIAC data up to 2020, $FFCO_2$ emissions in 2020 are approximated by scaling the ODIAC emissions in 2019 by a factor of 0.94, which is obtained using the ratio of CO_2 emissions from energy sources to total global CO_2 emissions in 2020 and 2019 (32,078.5 and 34,095.8 million ton CO_2, respectively) reported in the BP Statistical Review of World Energy (*BP*, 2022).

CO_2 flux components associated with the terrestrial biosphere (i.e., GPP, RE, and LUC) are derived from a prognostic biosphere model, the Vegetation Integrative SImulator for Trace gases (VISIT; *Ito*, 2019). The VISIT model comprises three independent modules that simulate carbon exchanges between the atmosphere and biosphere at hourly, daily, and monthly time steps. At present, only the VISIT module with a monthly time step includes the processes for evaluation of the impact of minor carbon flows, such as methane and biogenic volatile organic compound emissions and subsurface carbon exports and disturbances, all of which influence the carbon budget estimates of terrestrial ecosystems such as GPP and RE. We selected the carbon emissions associated with land-use conversion from the components of minor carbon flows to represent LUC emissions with a monthly time step. GPP and RE fluxes were derived from the VISIT module with an hourly time step to represent the immediate response of the terrestrial biosphere to changes in environmental conditions and their impact on atmospheric CO_2 variability. However, the GPP and RE estimates using the module with an hourly time step are impacted fewer times by disturbance processes such as LUC, which is likely to result in biases in estimates of regional carbon budgets. To reduce these biases, the values of GPP and RE with an hourly time step are scaled each month and at each grid cell with those estimated from the module with a monthly time step. Accordingly, the hourly GPP and RE used in our system indirectly consider the effects of minor carbon flows.

The BB CO_2 emissions are provided using a bottom-up approach with a burned area method; i.e., the Global Biomass Burning Emissions Inventory (GBEI version "2022a"; *Shiraishi et al.*, 2021; *Saito et al.*, 2022). The BB emissions in GBEI are estimated by combining the remote sensing products related to fire distribution with aboveground biomass and landcover classification distributions, all of which are derived from satellite observations. GBEI provides data products for CO_2, CH_4, and CO emissions from BB at 1-km and 1° grid resolutions with a monthly time step. The 1° resolution data are used as monthly BB CO_2 emissions in our system.

We use the Japan Meteorological Agency (JMA) carbon dioxide mapping data (JMA Ocean CO_2 Map; *Iida et al.*, 2021) as the *a priori* OCN flux. The data provide information on monthly oceanic pCO_2 and CO_2 uptake at 1° grid resolution. The pCO_2 field is calculated using analytical sea surface temperatures, salinity, and chlorophyll-a concentration data from satellite observations, and the field of CO_2 uptake is calculated from the difference between oceanic and atmospheric pCO_2 and 10-m wind speeds.

Table 2 lists the *a priori* fluxes and their data sources with respective references.

Table 2. List of *a priori* fluxes used in the GOSAT-2 L4 CO_2 computational system.

Prior	Model/Product	Reference	Temporal/Spatial resolution
FF	ODIAC	*Oda et al.* (2018)	monthly/1° × 1°
GPP	VISIT	*Ito* (2019)	hourly/0.5° × 0.5°*
RE	VISIT	*Ito* (2019)	hourly/0.5° × 0.5°*
LUC	VISIT	*Ito* (2019)	monthly/0.5° × 0.5°
BB	GBEI	*Shiraishi et al.* (2021); *Saito et al.* (2022)	monthly/1° × 1°
OCN	JMA Ocean CO_2 Map	*Iida et al.* (2021)	monthly/1° × 1°

* Hourly GPP and RE were originally simulated with a spatial resolution of 0.3125° × 0.3125° then interpolated onto a 0.5° × 0.5° grid. Scaling of hourly values of GPP and RE to monthly values was performed on a 0.5° × 0.5° grid.

4.3.3 Atmospheric observational data

The estimation of global surface CO_2 flux using the inverse scheme requires atmospheric concentration data to infer the spatial distribution of surface fluxes. The primary information on atmospheric observations used in our system is GOSAT-2 XCO_2 data retrieved from the SWIR spectra acquired by TANSO-FTS-2 (*Yoshida and Oshio*, 2022). The initial products of GOSAT-2 XCO_2 data ver. 01.04/07 were released in November 2020/December 2021, and the revised new data ver. 02.00 were released in August 2022. The revision details of GOSAT-2 XCO_2 data from ver. 01.04/07 to ver. 02.00 are given by *Yoshida and Oshio* (2022). A summary of the main revisions is as follows:

- A zero-level offset and an instrument line shape stretch factor for each sub-band are newly incorporated in the state vector.
- Some parameters in the retrieval algorithm, such as the coefficient of empirical noise and post-screening criteria, are revised.
- The Solar Pseudo-Transmittance Spectrum that is used as the solar Fraunhofer line model is updated from version 2015 of the Disk-Integrated Spectrum to version 2016 (*Toon*, 2015). In addition, the TSIS-1 Hybrid Solar Reference Spectrum (*Coddington et al.*, 2021) is used as the solar baseline.

NIES GOSAT-2 Project (2020, 2022) reported comparison results of XCO_2 data between GOSAT-2 and the Total Carbon Column Observing Network (TCCON; *Wunch et al.*, 2011). In this comparison, the GOSAT-2 XCO_2 data ver. 01.04 for the period from 1 March 2019 to 18 May 2020 and those of ver. 02.00 for the period from 1 August 2019 to 31 July 2020 were analyzed over four target areas within radii of ±0.1°, ±1°, ±2°, and ±5° from the TCCON site. As summarized in Table 3, the standard deviations of mean biases (ppm) are explicitly improved in the GOSAT-2 XCO_2 data ver. 02.00 relative to those of ver. 01.04; the standard deviations of the GOSAT-2 XCO_2 data ver. 02.00 over land and ocean are 41%–46% and 47%–75%, respectively, which represent an improvement over ver. 01.04. The mean biases over land are improved in ver. 02.00 by 0.21–0.44 ppm (9%–19%), whereas the mean biases over ocean increased.

Table 3. Comparison of XCO_2 (over land and ocean) between GOSAT-2 and TCCON using *NIES GOSAT-2 Project* (2020, 2022). Distance indicates the target area for comparison, N is the number of comparison data, Bias is the mean bias (ppm), and STDEV is the standard deviation (ppm).

Data	Distance	Land			Ocean		
		N	Bias (ppm)	STDEV (ppm)	N	Bias (ppm)	STDEV (ppm)
Ver. 01.04	±0.1°	532	2.63	3.29	1	2.92	-
	±1°	1981	2.29	3.86	31	0.27	6.85
	±2°	2640	2.34	4.04	92	−0.14	5.79
	±5°	5510	2.14	4.31	733	0.26	4.65
Ver. 02.00	±0.1°	408	2.27	1.94	0	-	-
	±1°	1715	2.08	2.08	52	2.43	1.74
	±2°	2505	1.90	2.21	117	2.35	1.59
	±5°	5397	1.84	2.38	645	2.29	2.46

Initial atmospheric CO_2 concentration field data are required for effective reconstruction of the source and sink distribution from the GOSAT-2 column concentrations in the atmospheric tracer transport model simulation. We prepared the initial field data by preprocessing the GOSAT-2 L4 computational system with *in situ* measurements from the Observation Package (ObsPack) Data Products (Ver. obspack_co2_1_GLOBALVIEWplus_v7.0; https://doi.org/10.25925/20210801; *Schuldt et al.*, 2021). The latest ObsPack product includes 587 atmospheric CO_2 datasets derived from observations made by 66 laboratories in 23 countries. The atmospheric CO_2 observations at 70 sites in the ObsPack product, which are the same as the observation sites used in the GOSAT Level 4 CO_2 Product (*Maksyutov et al.*, 2013), were used for preparation of the initial field data (Table 4).

Table 4. List of the ObsPack observation sites used for construction of the initial field data. Abbreviations: CSIRO, Commonwealth Scientific and Industrial Research Organization; ECCC, Environment and Climate Change Canada; ICOS-ATC, Integrated Carbon Observation System (ICOS) Atmosphere Thematic Centre; JMA, Japan Meteorological Agency; KUP, University of Bern, Physics Institute, Climate and Environmental Physics; NOAA, National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Laboratory and Center for Atmospheric and Oceanic Studies; SAWS, South African Weather Service; TU, Tohoku University.

Site	Name	Country	Lab	Measurement Type
ABT	Abbotsford British Columbia	Canada	ECCC	surface in situ
ALT	Alert Nunavut	Canada	NOAA	surface flask
ALT	Alert Nunavut	Canada	ECCC	surface in situ
AMT	Argyle Maine	United States	NOAA	tower in situ
ASC	Ascension Island	United Kingdom	NOAA	surface flask
BHD	Baring Head Station	New Zealand	NOAA	surface flask
BRA	Bratt's Lake Saskatchewan	Canada	ECCC	surface in situ
BRW	Barrow Atmospheric Baseline Observatory	United States	NOAA	surface flask
CAR	Briggsdale Colorado	United States	NOAA	aircraft pfp
CBA	Cold Bay Alaska	United States	NOAA	surface flask
CDL	Candle Lake Saskatchewan	Canada	ECCC	surface in situ
CGO	Cape Grim Tasmania	Australia	NOAA	surface flask
CHM	Chibougamau Quebec	Canada	ECCC	surface in situ
CIB	Centro de Investigacion de la Baja Atmosfera (CIBA)	Spain	NOAA	surface flask
CMA	Cape May New Jersey	United States	NOAA	aircraft pfp
CPS	Chapais Quebec	Canada	ECCC	surface in situ
CPT	Cape Point	South Africa	SAWS	surface in situ
CYA	Casey Antarctica	Australia	CSIRO	surface flask
EGB	Egbert Ontario	Canada	ECCC	surface in situ
ESP	Estevan Point British Columbia	Canada	NOAA	aircraft pfp
ETL	East Trout Lake Saskatchewan	Canada	NOAA	aircraft pfp
FSD	Fraserdale	Canada	ECCC	surface in situ
HBA	Halley Station Antarctica	United Kingdom	NOAA	surface flask
HDP	Hidden Peak (Snowbird), Utah	United States	NOAA	surface in situ
HIL	Homer Illinois	United States	NOAA	aircraft pfp
HUN	Hegyhatsal	Hungary	NOAA	surface flask
ICE	Storhofdi Vestmannaeyjar	Iceland	NOAA	surface flask
IZO	Izana Tenerife Canary Islands	Spain	NOAA	surface flask
JFJ	Jungfraujoch	Switzerland	KUP	surface in situ
KAS	Kasprowy Wierch, High Tatra	Poland	NOAA	surface in situ
KEY	Key Biscayne Florida	United States	NOAA	surface flask
KUM	Cape Kumukahi Hawaii	United States	NOAA	surface flask
LEF	Park Falls Wisconsin	United States	NOAA	aircraft pfp

Site	Name	Country	Lab	Measurement Type
LLB	Lac La Biche Alberta	Canada	ECCC	surface in situ
LMP	Lampedusa	Italy	NOAA	surface flask
MEX	High Altitude Global Climate Observation Center	Mexico	NOAA	surface flask
MHD	Mace Head County Galway	Ireland	NOAA	surface flask
MID	Sand Island Midway	United States	NOAA	surface flask
MLO	Mauna Loa Hawaii	United States	NOAA	surface in situ
MNM	Minamitorishima	Japan	JMA	surface in situ
MQA	Macquarie Island	Australia	CSIRO	surface flask
NHA	Worcester Massachusetts	United States	NOAA	aircraft pfp
NWR	Niwot Ridge Colorado	United States	NOAA	surface flask
OXK	Ochsenkopf	Germany	NOAA	surface flask
PFA	Poker Flat Alaska	United States	NOAA	aircraft pfp
POC	Pacific Ocean		NOAA	shipboard flask
PSA	Palmer Station Antarctica	United States	NOAA	surface flask
RPB	Ragged Point	Barbados	NOAA	surface flask
RYO	Ryori	Japan	JMA	surface in situ
SEY	Mahe Island	Seychelles	NOAA	surface flask
SGP	Southern Great Plains Oklahoma	United States	NOAA	surface flask
SMO	Tutuila	American Samoa	NOAA	surface flask
SNP	Shenandoah National Park	United States	NOAA	surface in situ
SPO	South Pole Antarctica	United States	NOAA	surface flask
SSL	Schauinsland Baden-Wuerttemberg	Germany	ICOS-ATC	surface in situ
SUM	Summit	Greenland	NOAA	surface flask
SYO	Syowa Station Antarctica	Japan	TU	surface in situ
TAP	Tae-ahn Peninsula	Republic of Korea	NOAA	surface flask
THD	Trinidad Head California	United States	NOAA	surface flask
USH	Ushuaia	Argentina	NOAA	surface flask
UTA	Wendover Utah	United States	NOAA	surface flask
UUM	Ulaan Uul	Mongolia	NOAA	surface flask
WBI	West Branch Iowa	United States	NOAA	tower in situ
WGC	Walnut Grove California	United States	NOAA	tower in situ
WIS	Weizmann Institute of Science at the Arava Institute Ketura	Israel	NOAA	surface flask
WKT	Moody Texas	United States	NOAA	tower in situ
WLG	Mt. Waliguan	People's Republic of China	NOAA	surface flask
WSA	Sable Island Nova Scotia	Canada	ECCC	surface in situ
YON	Yonagunijima	Japan	JMA	surface in situ
ZEP	Ny-Alesund Svalbard	Norway and Sweden	NOAA	surface flask

4.4 Atmospheric simulation and flux estimate

We use the NICAM-TM to simulate the transport of atmospheric CO_2. NICAM uses a quasi-homogeneous distribution of hexagonal or pentagonal grid cells derived from recursive division of an icosahedron to perform global simulations with high spatiotemporal resolution (Tomita and Satoh, 2004). The dynamic core of the model involves the use of nonhydrostatic equations expressed with finite volume methods, which can implement the CWC for tracer transport in the model. The NICAM-TM ensures strict mass conservation to produce realistic simulations of atmospheric tracer transport (*Niwa et al.*, 2011).

The solution of $\delta\mathbf{x}$ that minimizes Eq. (2) is given by the gradient of the objective function, $\mathbf{g} = \partial\mathbf{J}/\partial\delta\mathbf{x}$, so that

$$\mathbf{g} = \mathbf{B}^{-1}\delta\mathbf{x} + \mathbf{M}^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{M}\delta\mathbf{x} - \mathbf{d}). \tag{3}$$

The second term on the right-hand side in Eq. (3) denotes that a vector of model–observation misfit, $\mathbf{M}\delta\mathbf{x} - \mathbf{d}$, is integrated backward in time by an adjoint operator, $\mathbf{M}^{\mathrm{T}}$. The optimized vector of $\delta\mathbf{x}$ is deduced by minimizing the gradient $\mathbf{g}$ using the 4D-Var method with the iterative operation of a forward model and its backward integration that is represented using an adjoint model.

The adjoint calculation requires program codes to step backward in time for integrating sensitivities to source components, as shown in the second term of Eq. (3). The adjoint model in NISMON-CO_2 implements the backward integration by reading in reverse order the meteorological variables that are archived for forward simulations with NICAM-TM. In the model, adjoint codes for vertical diffusion and cumulus convection are written based on a so-called discrete approach, and the expression for advection is given by both the discrete and continuous approaches (*Niwa et al.*, 2017a). In the GOSAT-2 L4 computational system, the continuous approach is used to calculate advection processes.

In the estimates of global surface fluxes using NISMON-CO_2, the POpULar scheme (*Fujii*, 2005), based on a quasi-Newton method, is applied to obtain the $\delta\mathbf{x}$ that minimizes the cost function $\mathbf{J}$ (*Niwa et al.*, 2017b). The POpULar scheme uses the Broyden–Fletcher–Goldfarb–Shanno algorithm (BFGS) to estimate the inverse Hessian of $\mathbf{J}$, which gives the approximate Newton's direction $\mathbf{d}_{\mathrm{k}} = -\mathbf{H}_{\mathrm{k}}\mathbf{g}_{\mathrm{k}}$, where $\mathbf{H}_{\mathrm{k}}$ is the approximated inverse Hessian matrix of $\mathbf{J}$, and $\mathbf{g}_{\mathrm{k}}$ is the gradient shown in Eq. (3) at the *k*-th iteration. The approximate Newton's direction $\mathbf{d}_{\mathrm{k}}$ is then used to find the next point of the vector $\mathbf{x}$ with step size α_k in the direction $\mathbf{x}_{\mathrm{k+1}} = \mathbf{x}_{\mathrm{k}} + \alpha_k\mathbf{d}_{\mathrm{k}}$. Practical algorithms of POpULar for the iterative solution to obtain $\delta\mathbf{x}$ have been described by *Fujii* (2005) and *Niwa et al.* (2017b).

The sensitivity of the remote sensing measurements to the atmosphere is generally not uniform with altitude. Therefore, an accurate representation of vertical atmospheric profiles retrieved from TANSO-FTS-2 measurements in the model simulation results is essential for comparison between GOSAT-2 XCO_2 data and the simulations. We apply the averaging kernel matrix $\mathbf{A}$ with *a priori* information used in the retrieval to the atmospheric inversion of simulated atmospheric concentrations, as follows:

$$x_s = x_a + \mathbf{a}(\mathbf{x}_{\mathrm{s}} - \mathbf{x}_{\mathrm{a}}), \tag{4}$$

where x is the column-averaged dry-air mole fraction, and $\mathbf{a}$ and $\mathbf{x}$ are vectors for the column averaging kernel of the dry-air mole fraction and the vertical profile of atmospheric concentrations, respectively. Subscripts a and s refer to *a priori* and simulated atmospheric concentrations, respectively. Here, the entire depth of the atmosphere is divided into 15 vertical layers in the retrieval of the GOSAT-2 TANSO-FTS-2 SWIR Level 2 Column-averaged Dry-air Mole Fraction.

The *a priori* column-averaged dry-air mole fraction is given by

$$x_a = \mathbf{h}^{\mathrm{T}} \mathbf{x}_{\mathrm{a}}, \tag{5}$$

where $\mathbf{h}$ is a pressure weighting function, which is expressed using a vector of the partial column amount of dry air ω:

$$\mathbf{h}_i = \frac{\omega_{i,j}}{\sum_j \omega_{i,j}}, \tag{6}$$

where i refers to a discrete observation point and j is the vertical level. The total column averaging kernel $\mathbf{a}$ is determined as follows using $\mathbf{h}$ and the averaging kernel matrix $\mathbf{A}$:

$$\mathbf{a} = \mathbf{h}^{\mathrm{T}} \mathbf{A}. \tag{7}$$

We applied the variables $\mathbf{x}_{\mathrm{a}}$, ω, $\mathbf{A}$, pressures in the vertical layers, and surface pressure (hPa) derived from the GOSAT-2 SWIR L2 CO_2 product to Eqs (4)–(7). The pressures and surface pressure were used to adjust the vertical profile of partial column amount given by $\mathbf{x}_{\mathrm{s}}$ to that of $\mathbf{x}_{\mathrm{a}}$.

5 Level 4A and Level 4B Products

The GOSAT-2 L4 CO_2 Product consists of the L4A Global CO_2 Flux Product and the L4B Global CO_2 Distribution Product. The L4A Product is created by gridding the *a posteriori* CO_2 fluxes to a 2.5° latitude/longitude grid on a monthly timescale, and the L4B Product provides global distributions for instantaneous values of atmospheric CO_2 concentrations every 6 h at 17 fixed atmospheric pressure levels (975, 925, 900, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, and 10 hPa) and near the surface on a 2.5° latitude/longitude grid. The L4B Product is produced by performing forward simulation using the atmospheric tracer transport model with the *a posteriori* CO_2 fluxes.

The GOSAT-2 L4 Product is stored in NetCDF format data files (Conventions CF-1.6). In the L4A Product, *a priori* fluxes for six types of source and sink strengths and *a posteriori* fluxes for four types of source and sink strengths are provided on a monthly timescale together with total fluxes (g C m^{-2} day^{-1}). The *a posteriori* fluxes for GPP (flux_apos_gpp), RE (flux_apos_re), and LUC (flux_apos_luc) were merged to a terrestrial biosphere flux as flux_apos_teb = flux_apos_re + flux_apos_luc − flux_apos_gpp. The L4B Product provides three-dimensional atmospheric CO_2 concentration fields (mol mol^{-1}) and surface pressures derived from the atmospheric simulations. For comparison of the GOSAT-2 XCO_2 data, column concentrations of the L4B Product can be estimated using variables stored in the product, and some additional parameters may be obtained from the GOSAT-2 SWIR L2 CO_2 product, as described in Section 4.4.

References

BP (2022), bp Statistical Review of World Energy, 2022.

Coddington, O. M., E. C. Richard, D. Harber, P. Pilewskie, T. N. Woods, K. Chance, X. Liu, and K. Sun (2021), The tsis-1 hybrid solar reference spectrum, *Geophysical Research Letters*, *48*(12), e2020GL091709.

Fujii, Y. (2005), Preconditioned optimizing utility for large-dimensional analyses (POpULar), *Journal of Oceanography*, *61*(1), 167–181.

Iida, Y., Y. Takatani, A. Kojima, and M. Ishii (2021), Global trends of ocean CO_2 sink and ocean acidification: an observation-based reconstruction of surface ocean inorganic carbon variables, *Journal of Oceanography*, *77*(2), 323–358.

Ito, A. (2019), Disequilibrium of terrestrial ecosystem CO_2 budget caused by disturbance-induced emissions and non-CO_2 carbon export flows: a global model assessment, *Earth System Dynamics*, *10*(4), 685–709.

Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi (2015), The JRA-55 reanalysis: General specifications and basic characteristics, *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5–48.

Maksyutov, S., H. Takagi, V. Valsala, M. Saito, T. Oda, T. Saeki, D. A. Belikov, R. Saito, A. Ito, Y. Yoshida, I. Morino, O. Uchino, R. J. Andres, and T. Yokota (2013), Regional CO_2 flux estimates for 2009-2010 based on GOSAT and ground-based CO_2 observations, *Atmospheric Chemistry and Physics*, *13*(18), 9351–9373.

NIES GOSAT-2 Project (2020), Summary of the validation on GOSAT-2 TANSO-FTS-2 SWIR L2 Column-averaged Dry-air Mole Fraction Product.

NIES GOSAT-2 Project (2022), Summary of the validation on GOSAT-2 TANSO-FTS-2 SWIR L2 Column-averaged Dry-air Mole Fraction Product (Ver. 02.00).

Niwa, Y., H. Tomita, M. Satoh, and R. Imasu (2011), A three-dimensional icosahedral grid advection scheme preserving monotonicity and consistency with continuity for atmospheric tracer transport, *Journal of the Meteorological Society of Japan. Ser. II*, *89*(3), 255–268.

Niwa, Y., H. Tomita, M. Satoh, R. Imasu, Y. Sawa, K. Tsuboi, H. Matsueda, T. Machida, M. Sasakawa, B. Belan, and N. Saigusa (2017a), A 4D-Var inversion system based on the icosahedral grid model (NICAM-TM 4D-Var v1.0) – Part 1: Offline forward and adjoint transport models, *Geoscientific Model Development*, *10*(3), 1157–1174.

Niwa, Y., Y. Fujii, Y. Sawa, Y. Iida, A. Ito, M. Satoh, R. Imasu, K. Tsuboi, H. Matsueda, and N. Saigusa (2017b), A 4D-Var inversion system based on the icosahedral grid model (NICAM-TM 4D-Var v1.0) – Part 2: Optimization scheme and identical twin experiment of atmospheric CO_2 inversion, *Geoscientific Model Development*, *10*(6), 2201–2219.

Niwa, Y., Y. Sawa, H. Nara, T. Machida, H. Matsueda, T. Umezawa, A. Ito, S.-I. Nakaoka, H. Tanimoto, and Y. Tohjima (2021), Estimation of fire-induced carbon emissions from Equatorial Asia in 2015 using in situ aircraft and ship observations, *Atmospheric Chemistry and Physics*, *21*(12), 9455–9473

Oda, T., S. Maksyutov, and R. J. Andres (2018), The open-source data inventory for anthropogenic CO_2, version 2016 (odiac2016): a global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions, *Earth System Science Data*, *10*(1), 87–107.

Saito, M., T. Shiraishi, R. Hirata, Y. Niwa, K. Saito, M. Steinbacher, D. Worthy, and T. Matsunaga (2022), Sensitivity of biomass burning emissions estimates to land surface information, *Biogeosciences*, *19*(7), 2059–2078.

Schuldt, K. N., J. Mund, I. T. Luijkx, T. Aalto, J. B. Abshire, K. Aikin, A. Andrews, S. Aoki, F. Apadula, B. Baier, P. Bakwin, J. Bartyzel, G. Bentz, P. Bergamaschi, A. Beyersdorf, T. Biermann, S. C. Biraud, H. Boenisch, D. Bowling, G. Brailsford, G. Chen, H. Chen, L. Chmura, S. Clark, S. Climadat, A. Colomb, R. Commane, S. Conil, A. Cox, P. Cristofanelli, E. Cuevas, R. Curcoll, B. Daube, K. Davis, M. De Mazière, S. De Wekker, J. D. Coletta, M. Delmotte, J. P. DiGangi, E. Dlugokencky, J. W. Elkins, L. Emmenegger, S. Fang, M. L. Fischer, G. Forster, A. Frumau, M. Galkowski, L. V. Gatti, T. Gehrlein, C. Gerbig, F. Gheusi, E. Gloor, V. Gomez-Trueba, D. Goto, T. Griffis, S. Hammer, C. Hanson, L. Haszpra, J. Hatakka, M. Heimann, M. Heliasz, A. Hensen, O. Hermanssen, E. Hintsa, J. Holst, V. Ivakhov, D. Jaffe, W. Joubert, A. Karion, S. R. Kawa, V. Kazan, R. Keeling, P. Keronen, P. Kolari, K. Kominkova, E. Kort, E. Kozlova, P. Krummel, D. Kubistin, C. Labuschagne, D. H. Lam, R. Langenfelds, O. Laurent, T. Laurila, T. Lauvaux, J. Lavric, B. Law, O. S. Lee, J. Lee, I. Lehner, R. Leppert, M. Leuenberger, I. Levin, J. Levula, J. Lin, M. Lindauer, Z. Loh, M. Lopez, T. Machida, I. Mammarella, G. Manca, A. Manning, A. Manning, M. V. Marek, M. Y. Martin, H. Matsueda, K. McKain, H. Meijer, F. Meinhardt, L. Merchant, N. Mihalopoulos, N. Miles, C. E. Miller, J. B. Miller, L. Mitchell, S. Montzka, F. Moore, E. Morgan, J. Morgui, S. Morimoto, B. Munger, D. Munro, C. L. Myhre, M. Mölder, J. Müller-Williams, J. Necki, S. Newman, S. Nichol, Y. Niwa, S. O'Doherty, F. Obersteiner, B. Paplawsky, J. Peischl, O. Peltola, S. Piacentino, J. M. Pichon, S. Piper, C. Plass-Duelmer, M. Ramonet, R. Ramos, E. Reyes-Sanchez, S. Richardson, H. Riris, P. P. Rivas, T. Ryerson, K. Saito, M. Sargent, M. Sasakawa, D. Say, B. Scheeren, T. Schuck, M. Schumacher, T. Seifert, M. K. Sha, P. Shepson, M. Shook, C. D. Sloop, P. Smith, M. Steinbacher, B. Stephens, C. Sweeney, P. Tans, K. Thoning, H. Timas, M. Torn, P. Trisolino, J. Turnbull, K. Tørseth, A. Vermeulen, B. Viner, G. Vitkova, S. Walker, A. Watson, S. Wofsy, J. Worsey, D. Worthy, D. Young, S. Zaehle, A. Zahn, M. Zimnoch, A. G. di Sarra, D. van Dinther, and P. van den Bulk (2021), Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2020; obspack_co2_1_GLOBALVIEWplus_v7.0_2021-08-18, doi:10.25925/20210801.

Shiraishi, T., R. Hirata, and T. Hirano (2021), New inventories of global carbon dioxide emissions through biomass burning in 2001–2020, *Remote Sensing*, *13*(10), doi:10.3390/rs13101914.

Suto, H., F. Kataoka, N. Kikuchi, R. O. Knuteson, A. Butz, M. Haun, H. Buijs, K. Shiomi, H. Imai, and A. Kuze (2021), Thermal and near-infrared sensor for carbon observation Fourier transform spectrometer-2 (TANSO-FTS-2) on the Greenhouse gases Observing SATellite-2 (GOSAT-2) during its first year in orbit, *Atmospheric Measurement Techniques*, *14*(3), 2013–2039.

Tomita, H. and M. Satoh (2004), A new dynamical framework of nonhydrostatic global model using the icosahedral grid, *Fluid Dynamics Research*, *34*(6), 357–400.

Toon, G. C. (2015), Solar Line List for the TCCON 2014 Data Release, doi:10.14291/TCCON.GGG2014.SOLAR.R0/1221658.

Wunch, D., G. C. Toon, J.-F. L. Blavier, R. A. Washenfelder, J. Notholt, B. J. Connor, D. W. T. Griffith, V. Sherlock, and P. O. Wennberg (2011), The total carbon column observing network, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *369*(1943), 2087–2112.

Yokota, T., Y. Yoshida, N. Eguchi, Y. Ota, T. Tanaka, H. Watanabe, and S. Maksyutov (2009), Global concentrations of CO_2 and CH_4 retrieved from GOSAT: first preliminary results, *SOLA*, *5*, 160–

163.
Yoshida, Y. and H. Oshio (2022), GOSAT-2 TANSO-FTS-2 SWIR L2 Retrieval Algorithm Theoretical Basis Document.