



## Article

# Comparison of XH<sub>2</sub>O Retrieved from GOSAT Short-Wavelength Infrared Spectra with Observations from the TCCON Network

Eric Dupuy <sup>1,\*</sup>, Isamu Morino <sup>1</sup>, Nicholas M. Deutscher <sup>2,3</sup>, Yukio Yoshida <sup>1</sup>, Osamu Uchino <sup>1</sup>, Brian J. Connor <sup>4</sup>, Martine De Mazière <sup>5</sup>, David W. T. Griffith <sup>2</sup>, Frank Hase <sup>6</sup>, Pauli Heikkinen <sup>7</sup>, Patrick W. Hillyard <sup>8,9</sup>, Laura T. Iraci <sup>8</sup>, Shuji Kawakami <sup>10</sup>, Rigel Kivi <sup>7</sup>, Tsuneo Matsunaga <sup>1</sup>, Justus Notholt <sup>3</sup>, Christof Petri <sup>3</sup>, James R. Podolske <sup>8</sup>, David F. Pollard <sup>11</sup>, Markus Rettinger <sup>12</sup>, Coleen M. Roehl <sup>13</sup>, Vanessa Sherlock <sup>11</sup>, Ralf Sussmann <sup>12</sup>, Geoffrey C. Toon <sup>14</sup>, Voltaire A. Velazco <sup>2</sup>, Thorsten Warneke <sup>3</sup>, Paul O. Wennberg <sup>13</sup>, Debra Wunch <sup>13,†</sup> and Tatsuya Yokota <sup>1</sup>

<sup>1</sup> National Institute for Environmental Studies (NIES), 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan; morino@nies.go.jp (I.M.); yoshida.yukio@nies.go.jp (Y.Y.); uchino.osamu@nies.go.jp (O.U.); matsunag@nies.go.jp (T.M.); yoko@nies.go.jp (T.Y.)

<sup>2</sup> Center for Atmospheric Chemistry, School of Chemistry, Northfields Ave, University of Wollongong, NSW 2522, Australia; ndeutsch@uow.edu.au (N.M.D.); griffith@uow.edu.au (D.W.T.G.); voltaire@uow.edu.au (V.A.V.)

<sup>3</sup> Institute of Environmental Physics, University of Bremen, Otto-Hahn-Allee 1, Bremen 28359, Germany; jnotholt@iup.physik.uni-bremen.de (J.N.); christof\_p@iup.physik.uni-bremen.de (C.P.); warneke@iup.physik.uni-bremen.de (T.W.)

<sup>4</sup> BC Consulting Limited, 6 Fairway Dr., Alexandra 9320, New Zealand; bcconsulting@extra.co.nz

<sup>5</sup> Institut d'Aéronomie Spatiale de Belgique (BIRA-IASB), 3 Avenue Circulaire, Brussels B-1180, Belgium; Martine.DeMaziere@bira-iasb.oma.be

<sup>6</sup> Karlsruhe Institute of Technology, IMK-ASF, Hermann-von-Helmholtz-Platz 1, Leopoldshafen 76344, Germany; frank.hase@kit.edu

<sup>7</sup> FMI-Arctic Research Center, Tähteläntie 62, Sodankylä FIN-99600, Finland; Pauli.Heikkinen@fmi.fi (P.H.); Rigel.Kivi@fmi.fi (R.K.)

<sup>8</sup> NASA Ames Research Center, Atmospheric Science Branch, Mail Stop 245-5, Moffett Field, CA 94035, USA; patrick.hillyard@nasa.gov (P.W.H.); Laura.T.Iraci@nasa.gov (L.T.I.); James.R.Podolske@nasa.gov (J.R.P.)

<sup>9</sup> Bay Area Environmental Research Institute, 625 2nd St., Suite 209, Petaluma, CA 94952, USA

<sup>10</sup> Earth Observation Research Center (EORC), Japan Aerospace Exploration Agency (JAXA), 2-1-1 Sengen, Tsukuba-city, Ibaraki 305-8505, Japan; kawakami.shuji@jaxa.jp

<sup>11</sup> National Institute of Water and Atmospheric Research (NIWA), Private bag 50061, Omakau 9352, New Zealand; Dave.Pollard@niwa.co.nz (D.F.P.); vj.sherlock@gmail.com (V.S.)

<sup>12</sup> Karlsruhe Institute of Technology, IMK-IFU, Kreuzeckbahnstr. 19, Garmisch-Partenkirchen 82467, Germany; markus.rettinger@kit.edu (M.R.); ralf.sussmann@kit.edu (R.S.)

<sup>13</sup> California Institute of Technology, MC 131-24, 1200 E. California Blvd., Pasadena, CA 91125, USA; coleen@gps.caltech.edu (C.M.R.); wennberg@gps.caltech.edu (P.O.W.); dwunch@caltech.edu (D.W.)

<sup>14</sup> Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA; geoffrey.c.toon@jpl.nasa.gov

\* Correspondence: dupuy.eric-albert@nies.go.jp; Tel.: +81-29-850-2751

† Current address: Department of Physics, University of Toronto, 60 St. George Street, Toronto, ON M5S 1A7, Canada; dwunch@atmos.physics.utoronto.ca.

Academic Editors: Alfredo R. Huete and Prasad S. Thenkabail

Received: 10 January 2016; Accepted: 4 May 2016; Published: 17 May 2016

**Abstract:** Understanding the atmospheric distribution of water (H<sub>2</sub>O) is crucial for global warming studies and climate change mitigation. In this context, reliable satellite data are extremely valuable for their global and continuous coverage, once their quality has been assessed. Short-wavelength infrared spectra are acquired by the Thermal And Near-infrared Sensor for carbon Observation-Fourier Transform Spectrometer (TANSO-FTS) aboard the Greenhouse gases Observing Satellite (GOSAT).

From these, column-averaged dry-air mole fractions of carbon dioxide, methane and water vapor ( $\text{XH}_2\text{O}$ ) have been retrieved at the National Institute for Environmental Studies (NIES, Japan) and are available as a Level 2 research product. We compare the NIES  $\text{XH}_2\text{O}$  data, Version 02.21, with retrievals from the ground-based Total Carbon Column Observing Network (TCCON, Version GGG2014). The datasets are in good overall agreement, with GOSAT data showing a slight global low bias of  $-3.1\% \pm 24.0\%$ , good consistency over different locations (station bias of  $-1.53\% \pm 10.35\%$ ) and reasonable correlation with TCCON ( $R = 0.89$ ). We identified two potential sources of discrepancy between the NIES and TCCON retrievals over land. While the TCCON  $\text{XH}_2\text{O}$  amounts can reach 6000–7000 ppm when the atmospheric water content is high, the correlated NIES values do not exceed 5500 ppm. This could be due to a dry bias of TANSO-FTS in situations of high humidity and aerosol content. We also determined that the GOSAT-TCCON differences directly depend on the altitude difference between the TANSO-FTS footprint and the TCCON site. Further analysis will account for these biases, but the NIES V02.21  $\text{XH}_2\text{O}$  product, after public release, can already be useful for water cycle studies.

**Keywords:** GOSAT;  $\text{H}_2\text{O}$ ; SWIR; validation

## 1. Introduction

Water ( $\text{H}_2\text{O}$ ) is among the most abundant and ubiquitous species in the Earth's atmosphere. It is the only minor atmospheric constituent present in all three states of matter: as a liquid, a solid and a gas. Due to its numerous absorption lines in infrared wavelengths, it is also the most important natural greenhouse gas. Water vapor has a very short atmospheric lifetime of approximately nine days [1], which is far shorter than any other major greenhouse gas. Because of these characteristics,  $\text{H}_2\text{O}$  strongly influences atmospheric properties and processes such as: the energy and radiation balance (absorption of solar radiation, Earth's thermal emission, non-radiative transport, albedo from cloud cover), atmospheric chemistry (precursor of OH, ice nucleation and heterogeneous chemistry in clouds), atmospheric dynamics, weather and climate (e.g., [2–4]). For example, clouds remain the largest source of uncertainty in climate models, and there are strong feedback mechanisms through which  $\text{H}_2\text{O}$  amplifies climate change [5,6]. Therefore, characterizing the hydrological cycle and quantifying that feedback is a considerable undertaking [7]; the relationship between water vapor and climate change is still not fully understood.

The atmospheric distribution of water vapor is highly variable at spatial and temporal scales relevant to weather and climate (e.g., [8]). This imposes strong observational constraints on water vapor measurements; sensor intercomparisons must take the spatio-temporal mismatch into account [9,10]. No single, standard instrument is capable of measuring  $\text{H}_2\text{O}$  with, at the same time, high accuracy, good geographic and temporal coverage and good vertical sampling. Fortunately, most of it resides in the troposphere, with  $\sim 60\%$  contained in the boundary layer up to 850 hPa and  $\sim 90\%$  below 500 hPa [11]. This enables the use of many different techniques and instruments to study water vapor, not only by remote sensing, but also via *in situ* acquisition of data. Recently, a detailed survey of the available observation techniques was conducted, including a list of existing sensors and networks and a discussion of the existing intercomparison efforts [12]. While ground-based instruments or *in situ* sensors can provide very high quality data, they do not give access to the global scale and must be complemented by satellite instruments. On the other hand, satellite-borne instruments can provide near-global coverage with some restrictions (e.g., daytime-only or land-only measurements), but few are capable of providing vertically-resolved or long-term information. Satellite sensors measuring water vapor operate in different viewing geometries, limb-sounding, occultation and nadir-viewing, and different spectral domains, ultraviolet (UV), visible, near- or short-wavelength infrared (NIR/SWIR), thermal infrared (TIR) and microwave [13].

Accurate quantification of the atmospheric water vapor content is a challenge, and available sensors all have their strengths and weaknesses. Furthermore, existing validation studies have shown that the observed differences are study-dependent and that no single instrument or technique can yet provide continuous data of sufficiently good quality for accurate climate model predictions [14,15]. When new data become available, it is thus of foremost importance to evaluate their quality and limitations in order to use them efficiently in long-term trend evaluation (jointly with pre-existing datasets) and climate modeling.

The latest satellite mission able to provide global measurements of water vapor from space is the Japanese Greenhouse gases Observing Satellite (GOSAT), launched in January 2009 [16]. GOSAT is a joint project of the Ministry of the Environment (MOE), the National Institute for Environmental Studies (NIES) and the Japan Aerospace Exploration Agency (JAXA). Aside from the main target greenhouse gases, carbon dioxide ( $\text{CO}_2$ ) and methane ( $\text{CH}_4$ ), column-averaged dry-air mole fractions of water vapor ( $\text{XH}_2\text{O}$ ) can be retrieved from the SWIR measurements of the Thermal And Near-infrared Sensor for carbon Observation-Fourier Transform Spectrometer (TANSO-FTS) using the NIES retrieval algorithm [17]. In this work, we present our initial comparison of the NIES  $\text{XH}_2\text{O}$  retrievals Version 02.21 (V02.21) with data from the ground-based FT spectrometers of the Total Carbon Column Observing Network (TCCON) [18]. The GOSAT mission and the TCCON are presented in Sections 2 and 3, respectively. The methodology is described in Section 4, the analysis results in Section 5 and the conclusions in Section 6.

## 2. The GOSAT Mission, Instrumentation and L2 Data

GOSAT is the first satellite mission entirely dedicated to monitoring atmospheric  $\text{CO}_2$  and  $\text{CH}_4$ , with the purpose of estimating their emissions, absorptions and fluxes on a subcontinental scale (several thousand square kilometers). Launched on 23 January 2009, GOSAT is in a Sun-synchronous, 98°-inclination orbit at the altitude of 666 km, with an Equator crossing (at the descending node) occurring at 12:48, local time. These orbital parameters yield a revisit time of three days. The payload was described in detail by Kuze *et al.* [16]. It consists of two instruments: the main instrument, TANSO-FTS, and the TANSO-Cloud and Aerosol Imager (TANSO-CAI).

### 2.1. The GOSAT Payload

TANSO-FTS observes both the solar light reflected at the Earth's surface during daytime and the atmospheric thermal emission continuously during day- and night-time. Spectra are acquired in four spectral bands: three located in the SWIR region near 0.76, 1.6 and 2  $\mu\text{m}$  (Bands 1, 2 and 3, respectively) and a broad TIR band between 5.5 and 14.3  $\mu\text{m}$  (Band 4). The spectral sampling is  $0.2\text{ cm}^{-1}$  for all bands, and the spectral resolution is approximately  $0.37\text{ cm}^{-1}$  for Band 1 and  $0.26\text{ cm}^{-1}$  for Bands 2, 3 and 4. The nominal integration time is 4 s. The Instantaneous Field-of-View (IFOV) of TANSO-FTS is 15.8 mrad, which corresponds to a sea-level nadir circular footprint of  $\sim 10.5\text{ km}$  in diameter. The built-in pointing mechanism allows for off-nadir observations up to  $\pm 20^\circ$  and  $\pm 35^\circ$  in the along-track and cross-track directions, respectively [16]. This is very important for SWIR observations, since the reflection properties of the Earth's surface are drastically different over land and over ocean. Observations over land are obtained in the nadir or near-nadir direction. Over ocean, however, usable measurements are obtained in "sunglint mode" by pointing off-nadir at the area of specular reflection for the incident sunlight. Note that TANSO-FTS can also point at a specific location on the globe ("target mode") for validation or scientific purposes. The NIES is responsible for retrieving the  $\text{CO}_2$  and  $\text{CH}_4$  column-averaged dry-air mole fractions (Level 2 (L2) products  $\text{XCO}_2$  and  $\text{XCH}_4$ ) from the SWIR spectra, for validating the retrieved  $\text{XCO}_2$  and  $\text{XCH}_4$ , and for estimating global carbon fluxes (Level 4 products) from the SWIR L2 data [17].

The nominal objective of the TANSO-CAI is to characterize the cloud distribution and aerosol properties in the field-of-view of TANSO-FTS in order to account for the effect of cloud and aerosol particles in the retrieval algorithm. Indeed, scattering by clouds and aerosols has a major impact on

SWIR remote sensing retrievals (e.g., [19]). TANSO-CAI is a near-UV to near-IR push-broom imager with four bands centered at 0.38, 0.674, 0.87 and 1.6  $\mu\text{m}$  (Bands 1 to 4, respectively), with spatial resolutions at nadir of 0.5 km for Bands 1–3 and 1.5 km for Band 4. The images acquired are currently used to determine the cloud coverage within wide areas, including a series of TANSO-FTS consecutive fields-of-view, in order to categorize the IFOV as cloud-free or contaminated. This screening is critical for the L2 retrieval algorithm, which assumes cloud-free conditions. Application of a strict cloud screening leads to the rejection of about 90% of the TANSO-FTS spectra, so that 10% of the observations can be used as input for the processing code.

## 2.2. The SWIR L2 Data at NIES

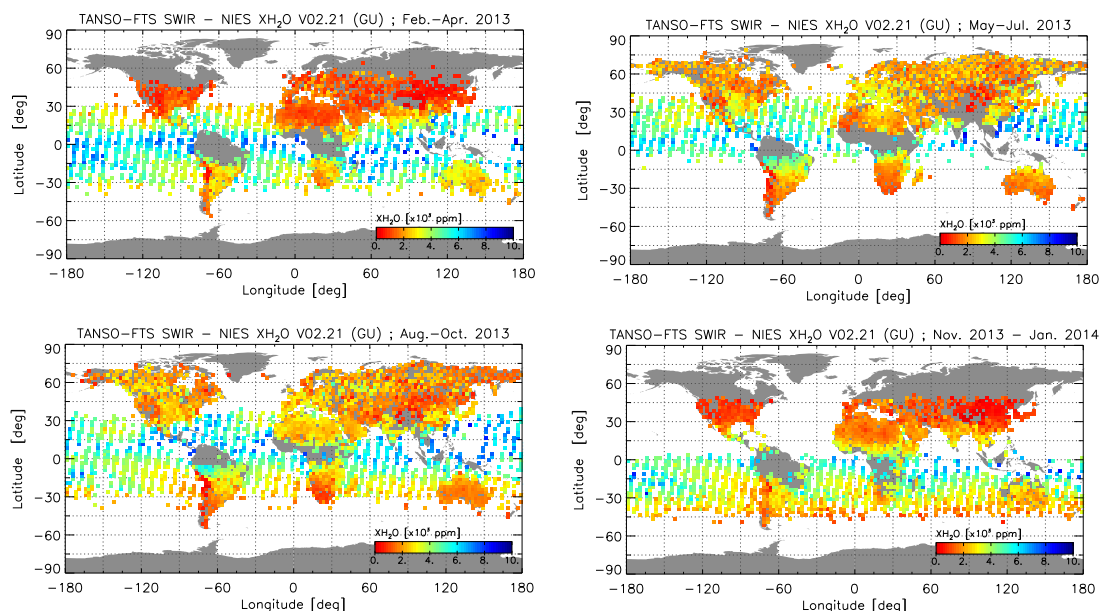
The details of the retrieval algorithm and its successive improvements have been described extensively by Yoshida *et al.* [17,20]. The first step in the algorithm is the rejection of cloud-contaminated measurements using the TANSO-CAI cloud-detection algorithm of Ishida and Nakajima [21]. To further reject cloudy observations, two more pre-screening filters are applied: a CAI “spatial coherence” test to check for the presence of sub-pixel-sized clouds in the CAI images, and a “2  $\mu\text{m}$ -scattering” test using data from TANSO-FTS Band 3 to detect high-altitude cirrus clouds not seen by TANSO-CAI [20]. Restriction on the Solar Zenith Angle (SZA) and the test of the quality of the calibrated spectra (Level 1B) are also performed at this stage. This pre-processing ensures that only data of suitable quality that are acquired in nominal, cloud-free conditions are used for the retrievals.

The NIES retrieval algorithm is based on the optimal estimation method [22] and is used to deliver simultaneously all operational (*i.e.*, validated and publicly released) and research (not validated) products. All target quantities are retrieved jointly; therefore, there are no separate retrievals for water vapor, carbon dioxide or methane. Four spectral regions (“sub-bands”) from Bands 1, 2 and 3 of TANSO-FTS are fitted simultaneously: the oxygen “O<sub>2</sub>-A” sub-band (12,950–13,200  $\text{cm}^{-1}$ ); the “weak CO<sub>2</sub>” sub-band containing a weak spectral absorption of carbon dioxide (6180–6380  $\text{cm}^{-1}$ ); the methane sub-band (5900–6150  $\text{cm}^{-1}$ ); and the “strong CO<sub>2</sub>” sub-band, containing a stronger CO<sub>2</sub> absorption feature (4800–4900  $\text{cm}^{-1}$ ). The target gases are molecular oxygen (O<sub>2</sub>), CO<sub>2</sub>, CH<sub>4</sub> and H<sub>2</sub>O; no other interfering species are considered. Firstly, partial columns are derived for CO<sub>2</sub>, CH<sub>4</sub> and H<sub>2</sub>O over 15 vertical layers, together with aerosol parameters and the surface pressure [17,20]. The partial columns are then integrated to obtain vertical column densities, which are later converted to column-averaged dry-air mole fractions (XCO<sub>2</sub>, XCH<sub>4</sub> and XH<sub>2</sub>O) using the surface pressure values simultaneously retrieved from the O<sub>2</sub>-A sub-band spectra. For land measurements, the surface albedo, and for ocean measurements, the surface wind speed are also retrieved. To simulate water vapor absorption in the forward model (radiative transfer), the spectroscopic parameters are taken from the High-resolution Transmission database, Edition 2008 (HITRAN 2008 [23]), and a Voigt profile is used for the line shape. Scattering is also accounted for in the forward model; the retrieved O<sub>2</sub> columns are not used to correct the CO<sub>2</sub>, CH<sub>4</sub> and H<sub>2</sub>O retrievals for scattering, but only as an *a posteriori* screening parameter to exclude the high-scattering data. Finally, concerning the aerosols, the *a priori* information is taken from the Spectral Radiation-Transport Model for Aerosol Species (SPRINTARS) Version 3.84 [24]. The Aerosol Optical Depth (AOD) is retrieved over six vertical layers and for two broad types of aerosols (12 values): fine mode (carbonaceous and sulphate aerosol) and coarse mode (soil dust, sea salt) [17].

After the retrieval, additional quality checks are performed (post-processing screening). Details of the pre- and post-processing filters are given in Table A1 (Appendix A). The retrieved data are made available by NIES through the GOSAT User Interface Gateway (GUIG) [25]. Depending on how strict the post-processing filtering is, data are released under different labels: “RA” (Research Announcement) with limited filtering, provided to researchers from the GOSAT research announcements and “GU” (General User) with stricter filtering for the general users, after public release of the data. Three major versions (V00.xx, V01.xx and V02.xx) of the NIES retrievals have already been released for the main TANSO-FTS data products, XCO<sub>2</sub> and XCH<sub>4</sub>. Versions 01.xx and 02.xx of the XCO<sub>2</sub> and XCH<sub>4</sub> products

have been extensively validated [17,20,26–28]. Independent algorithms developed separately are also used routinely to retrieve profile or column information from the TANSO-FTS spectra for CO<sub>2</sub>, CH<sub>4</sub> and, in some cases, H<sub>2</sub>O and deuterated water (HDO) [29–32]. However, the NIES XH<sub>2</sub>O data are still considered a research product. The objective of this work was a preliminary quality assessment of the NIES XH<sub>2</sub>O retrievals (Version 02.21) as a pre-requisite to their public release. We therefore applied GU-level screening filters (Table A1, Appendix A) to the data prior to our analyses.

Figure 1 shows global maps of the NIES XH<sub>2</sub>O V02.21 data from one year of observations (from February 2013–January 2014). The data were averaged within latitude/longitude bins of  $2.5^\circ \times 2.5^\circ$  and grouped into three-month periods corresponding approximately to seasons, thus representative of distinct atmospheric situations. Such maps show the difficulty of analyzing H<sub>2</sub>O data due to the large amplitude of its variations (roughly from 0 to 10,000 ppm or 1% in volume) and the spatial inhomogeneities of its distribution. This figure also illustrates the characteristics of GOSAT observations: dense coverage in the mid-latitudes over land; narrow swaths over water corresponding to sunglint retrievals for consecutive orbits; and a limited amount of usable data in the tropics because of the persistent cloud coverage in these regions, notably over the Amazon basin and equatorial Africa. Due to the Sun-synchronous nature of GOSAT's orbit, the accessible latitude domain is locked to the Sun position; thus, it varies with season.



**Figure 1.** Three-month averages of Short-Wavelength Infrared (SWIR) measurements of column-averaged dry-air mole fractions of water vapor (XH<sub>2</sub>O) from the Thermal And Near-infrared Sensor for carbon Observation-Fourier Transform Spectrometer (TANSO-FTS) (National Institute for Environmental Studies (NIES) V02.21 retrievals) representative of specific seasonal patterns, for data acquired between February 2013 and January 2014. From left to right and top to bottom: around the spring equinox (February–April), the summer solstice (May–July), the fall equinox (August–October) and the winter solstice (November–January). Because the orbit is Sun-synchronous, the band in which sunglint (specular reflection) ocean measurements are possible is located at different latitudes depending on the period. Data are averaged for each period and binned using a latitude/longitude mesh of  $2.5^\circ \times 2.5^\circ$ .

### 3. The Total Carbon Column Observing Network

The TCCON is a network of ground-based, high-spectral-resolution FTSs that record solar absorption spectra in the NIR spectral region. From these spectra, total column abundances of CO<sub>2</sub>, CH<sub>4</sub> and other gases are retrieved with high accuracy and precision [18,33]. The main purposes of the



TCCON are to provide reliable, long-term measurements of greenhouse gases and other atmospheric constituents for use in carbon cycle studies, to provide a reference for satellite measurements and, finally, to act as a transfer standard between *in situ* and space-borne measurements [33]. The locations of the operational, previous and potential TCCON sites are shown in Figure 2. The latest version (GGG2014) of the TCCON data product was used in this work. When data from a particular station were used in the nominal comparison as defined in Section 4 (or in the “extended” comparisons reported in the Supplementary Material), we included its geographic coordinates and periods of operation in Table 1, together with a bibliographic reference for the corresponding dataset.



**Figure 2.** Ground-based stations of the Total Carbon Column Observing Network (TCCON). From the TCCON data archive homepage [34]. Credits for the underlying image: Blue Marble: Next Generation, produced by Reto Stockli, NASA Earth Observatory (NASA Goddard Space Flight Center).

**Table 1.** List of the TCCON stations included in the nominal (this paper) or extended (Supplementary Material) comparisons. Their geographic location, status of operations and reference for the dataset are given. Sites are arranged by decreasing latitude from north to south. The data description and site information can be found on the TCCON website [35].

Site/Dataset	Latitude (°)	Longitude (°)	Altitude (km)	Start/End of Operations		Reference
Sodankylä	67.37	26.63	0.188	6 February 2009	Ongoing	Kivi <i>et al.</i> [36]
Bialystok	53.23	23.03	0.180	1 March 2009	Ongoing	Deutscher <i>et al.</i> [37]
Bremen	53.10	8.85	0.030	6 January 2005	Ongoing	Notholt <i>et al.</i> [38]
Karlsruhe	49.10	8.44	0.119	19 April 2010	Ongoing	Hase <i>et al.</i> [39]
Orléans	47.97	2.11	0.130	29 August 2009	Ongoing	Warneke <i>et al.</i> [40]
Garmisch	47.48	11.06	0.743	16 July 2007	Ongoing	Sussmann <i>et al.</i> [41]
Park Falls	45.95	−90.27	0.442	26 May 2004	Ongoing	Wennberg <i>et al.</i> [42]
Indianapolis <sup>†</sup>	39.86	−86.00	0.270	23 August 2012	1 December 2012	Iraci <i>et al.</i> [43]
Lamont	36.60	−97.49	0.320	6 July 2008	Ongoing	Wennberg <i>et al.</i> [44]
Tsukuba	36.05	140.12	0.031	4 August 2011	Ongoing	Morino <i>et al.</i> [45]
Edwards <sup>†</sup>	34.96	−117.88	0.700	20 July 2013	Ongoing	Iraci <i>et al.</i> [46]
JPL <sup>†</sup>	34.20	−118.18	0.390	8 December 2011	31 March 2013	Wennberg <i>et al.</i> [47]
Pasadena	34.14	−118.13	0.230	20 September 2012	Ongoing	Wennberg <i>et al.</i> [48]
Saga	33.24	130.29	0.007	28 July 2011	Ongoing	Shiomi <i>et al.</i> [49]
Darwin*	−12.42	130.89	0.030	28 August 2005	June 2015	Griffith <i>et al.</i> [50]
	−12.46	130.93	0.037	June 2015	Ongoing	
La Réunion**	−20.90	55.49	0.087	6 October 2011	Ongoing	De Mazière <i>et al.</i> [51]
Wollongong	−34.41	150.88	0.030	26 June 2008	Ongoing	Griffith <i>et al.</i> [52]
Lauder	−45.04	169.68	0.370	2 February 2010	Ongoing	Sherlock <i>et al.</i> [53]

<sup>†</sup> The Indianapolis (*cf.* the Supplementary Material), Edwards and JPL datasets were acquired with the same instrument; for the explanation, please see the site-specific notes [54]; \* the instrument in Darwin was recently relocated; \*\* for La Réunion, coincidences with TANSO-FTS ocean scans only and for relaxed geographic criteria of  $\pm 2^\circ$  in latitude and longitude (*cf.* the Supplementary Material).

TCCON retrievals are performed using the non-linear least-squares fitting algorithm Gas Fit (GFIT, “spectral fitting and line-by-line retrieval algorithm” [18,55]), which scales an *a priori* profile to obtain a synthetic spectrum representing the best fit to the measured spectrum. The scaled profile is integrated to compute total column abundances, which are then divided by the total column amount of dry air to obtain the final product: the column-averaged dry air mole fractions of the target species ( $X_{\text{gas}}$ ). While integrated quantities are more frequently expressed as total column or, for water vapor, precipitable water vapor,  $X_{\text{gas}}$  is a useful quantity as it is independent from surface pressure, hence from small-scale temporal and spatial variations and from local topography (e.g., [18]). In the TCCON retrieval scheme, the total column amount of dry air is calculated as the ratio of the retrieved total column of  $\text{O}_2$  to an assumed  $\text{O}_2$  dry-air mole fraction equal to 0.2095 [56]. For water vapor, the TCCON *a priori* profiles are taken from the National Centers for Environmental Predictions/National Center for Atmospheric Research (NCEP/NCAR) reanalysis data, which are tied to the radiosonde network.

Considerable efforts have been made to minimize all known error sources, both in the ancillary data used by the GFIT algorithm (surface temperature and pressure, *a priori* information for the target gases, etc.) and in the measurement and retrieval protocols. The TCCON data (except for hydrogen fluoride (HF) and HDO) have been calibrated through extensive comparisons with aircraft [57,58] and radiosonde data [33]. They are tied to the World Meteorological Organization (WMO) reference scale and show excellent consistency between the different sites. The achieved uncertainties (defined as the root-sum-square of the precision and accuracy values) for the GGG2014 data are of the order of 0.2%–0.3% for  $\text{XCO}_2$  and  $\text{XCH}_4$  and 1%–2% for  $\text{XH}_2\text{O}$  [56].

#### 4. Comparison Methodology

##### 4.1. Datasets Used in This Study

In the present study, we use the NIES L2 SWIR  $\text{XH}_2\text{O}$  research product Version 02.21, based on the TANSO-FTS Level 1B data provided by JAXA (L1B, Version 161.160). It is the most recent reprocessing of all TANSO-FTS data acquired from the start of the operational observation period (early June 2009) to May 2014. Unfortunately, a critical malfunction occurred on one of the two solar paddles on 25 May 2014, and the on-board systems experienced a sudden shutdown. Observations resumed in mid-June 2014 after recovery, but in a reduced-power operating mode to accommodate for the loss of the solar paddle. This led to changes in the scanning sequence and in the nominal properties of the interferograms [59]. Subsequently, a small, but non-negligible degradation of the data quality was noticed. Therefore, although data acquired after June 2014 have also been processed (NIES SWIR data Versions 02.31 and 02.40), we decided to restrict the comparison to V02.21 data acquired while the spacecraft was still operating nominally. As mentioned previously, we applied the nominal filters for GU (Table A1, Appendix A) to the dataset, since the official GU-level  $\text{XH}_2\text{O}$  product has not yet been released. Such screening might not be perfectly adapted to water vapor, whose atmospheric variability is much larger than that of the main target gases, but its optimization (currently in progress) is beyond the scope of this paper.

For the TCCON data, we use the latest version (GGG2014 processing) of the  $\text{XH}_2\text{O}$  product. After a nominal retention period, these data are public and can be freely downloaded from the TCCON data archive [34], hosted by the Carbon Dioxide Information Analysis Center (CDIAC). Uncertainties for  $\text{XH}_2\text{O}$  processed with the GGG2014 code are estimated at ~1.3% or below for SZA values of up to  $85^\circ$  [56], and the released data only include the measurements with SZAs of up to  $82^\circ$ . The TCCON  $\text{XH}_2\text{O}$  data are calibrated using *in situ* aircraft and radiosonde data. The consistency between the different sites is high enough to allow for the determination of a single calibration factor (relative to the WMO scale) for each target species. This scaling is then applied to the retrieved columns prior to data release. Data are also filtered using several criteria, notably the cloudiness of the measurements and the quality of the spectral fits. A final quality check is performed at each site for known issues that impact the quality of the data, but do not raise any flags during routine automated processing.

Data still considered suspect (e.g., due to instrument misalignment or detector saturation) are then further removed (site-specific problems are reported by the instrument teams and can be consulted online [60]). A detailed description of the GGG2014 data version (available from the CDIAC) is given by Wunch *et al.* [56].

#### 4.2. Matching GOSAT and TCCON Measurements

The high variability of the H<sub>2</sub>O distribution, even at very small spatial scales and short time intervals, makes it difficult to choose optimal collocation criteria for validation (e.g., [9,10]). In addition, the nature of GOSAT's orbit and the scanning pattern of TANSO-FTS [16] limit the influence of the temporal criterion on the number of coincidences. Depending on the location of the TCCON site, there might indeed not be any close spatial coincidences with TANSO-FTS "standard" footprints; on the other hand, if there are collocated footprints, then the frequency of the TCCON measurements ensures that there will be TCCON soundings temporally close to a GOSAT overpass. Furthermore, the existence of a target mode allows for close spatial and temporal matches at most TCCON sites, which are among the most frequently-scheduled targets for GOSAT. In Section 5.3, we investigate this issue and evaluate our choice of coincidence criteria.

For this study, we choose simple geophysical collocation criteria. Nominally, we select GOSAT scans acquired within  $\pm 1^\circ$  in latitude and in longitude of the TCCON sites (distances of up to  $\sim 130$  km between the GOSAT footprint and the ground-based instrument) and all TCCON measurements acquired within 30 min of each GOSAT overpass (before or after). These are comparable to, or stricter than, those used in previous GOSAT validation studies [17,26–30]. This provides a sufficiently large sample of matched measurements at most TCCON sites, which is important, because the single-scan measurement noise is the dominant source of error for the GOSAT SWIR observations [29]. For the same reason, we exclude sites with too few coincidences (less than 10) from the nominal comparison. Note that setting a higher threshold reduces the number of sites, but does not significantly impact the comparison results.

There are no coincidences for TANSO-FTS ocean scans using the nominal criteria. These can occur if the spatial coincidence criterion is relaxed to  $\pm 2^\circ$ , but there are too few TCCON sites involved and too few coincidences at each site to draw meaningful statistical conclusions. Therefore, these results will not be discussed here. Nevertheless, they exhibit a different behavior and provide a point of comparison with the results over land. For this reason, we decided to include them in the Supplementary Material, together with results obtained with different coincidence criteria.

#### 4.3. Calculation Steps

The possibility of using TANSO-FTS SWIR spectra to obtain useful information on water vapor and its main isotopologue, HDO, was demonstrated by Frankenberg *et al.* [29] and Boesch *et al.* [30]. Since HDO is not retrieved with the NIES V02.21 algorithm, it was not possible to perform direct comparison with other GOSAT retrievals. Alternately, the NIES XH<sub>2</sub>O data are still a research product that will undergo further improvements. For these reasons, we simplify the comparison methodology. We also investigate the impact of different issues (geographic and temporal proximity, altitude difference, filtering thresholds, retrieval parameters) on the comparison results, in order to assess the validity of our choice. The main steps of the calculation are listed below:

1. For reasons given in the previous section (GOSAT's scanning pattern, revisit time, orbital speed), the time criterion has a limited impact on the number of coincidences. For a specific TCCON site, there are, at best, only a few "standard" footprints (*i.e.*, distinct from target-mode observations) in close geographic proximity to the ground-based instrument. On the other hand, if the geographic coincidence condition is fulfilled, the frequency of the TCCON measurements ensures that a sufficient number of observations are in close temporal coincidence with a given TANSO-FTS scan. These temporally-matched TCCON observations are averaged and the resulting value is



- compared to the single coincident GOSAT scan. This is done in order to minimize the impact of the short-scale variations of the H<sub>2</sub>O distribution on the results of the comparison. Thus, TCCON data are generally counted multiple times, while TANSO-FTS scans are included only once in the calculations. This calculation method was previously used by Yoshida *et al.* [17] and Morino *et al.* [26] for the validation of NIES XCO<sub>2</sub> and XCH<sub>4</sub> retrievals.
2. We use the ground-based TCCON data as the reference for the calculations. The absolute bias for one pair (GOSAT *vs.* the mean of matched TCCON, “single-scan bias”) is thus defined as  $\delta_{\text{XH}_2\text{O}} [\text{ppm}] = \text{XH}_2\text{O}_{\text{GOSAT}} - \overline{\text{XH}_2\text{O}}_{\text{TCCON}}$  and the relative bias as the absolute bias ratioed to the TCCON XH<sub>2</sub>O values:  $\delta_{\text{XH}_2\text{O}} [\%] = 100 \times \delta_{\text{XH}_2\text{O}} [\text{ppm}] / \overline{\text{XH}_2\text{O}}_{\text{TCCON}}$ . Here,  $\overline{\text{XH}_2\text{O}}_{\text{TCCON}}$  represents the arithmetic mean of XH<sub>2</sub>O from all TCCON soundings coincident with one GOSAT scan. We then compute the corresponding global bias (absolute or relative, “ensemble bias”) as the arithmetic mean of the single-scan biases (absolute or relative) with its associated standard deviation. In order to evaluate the overall consistency of the results for different TCCON sites, the average and standard deviation of the station mean biases are also calculated (“station bias”). The linear least-squares fitting parameters and the correlation coefficient (R) are determined for the ensemble set and for each TCCON dataset.
  3. In this study, we directly compare the output of the NIES SWIR V02.21 and TCCON GGG2014 algorithms: no smoothing was applied to either dataset. Rigorously, comparison data should be smoothed following the approach of Rodgers and Connor [61], to account for the differences of instrumentation and observation geometries. The formalism of Rodgers [22], especially, provides and uses *ad hoc* mathematical tools—averaging kernels and *a priori* information—to perform this smoothing (e.g., [27,28,62]). However, smoothing the observed data might unduly constrain the comparison results towards the *a priori* information rather than towards the measured data, if the information content is low. Note that Inoue *et al.* [27] compared the NIES SWIR XCO<sub>2</sub> retrievals to aircraft data with and without applying GOSAT averaging kernels to the higher-resolution aircraft data and did not find a significant difference for XCO<sub>2</sub>.
  4. Discrepancies between the mean altitude within a TANSO-FTS footprint and the elevation of the TCCON sites potentially have a significant impact on the results. This is particularly true for water vapor, whose column abundance is largely dominated by its lower-tropospheric amount. Here, we also assess the impact of the GOSAT/TCCON altitude differences on the XH<sub>2</sub>O bias, but for simplicity reasons, we do not apply any altitude compensation to the GOSAT or TCCON columns prior to the bias calculations.

## 5. Results and Discussion

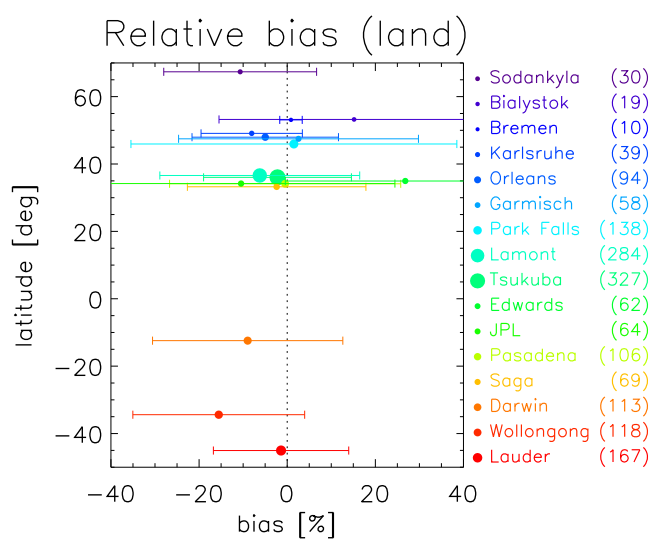
Since the seasonal variations of water vapor and the atmospheric conditions can differ significantly from one geographic location to another, it is difficult to compare absolute values of the XH<sub>2</sub>O differences (in ppm) from different TCCON sites. Thus, we decided to discuss the results primarily in terms of relative differences.

### 5.1. Statistical Comparison

In this section, we present the results of the statistical comparison between the TANSO-FTS and TCCON retrievals. These were obtained with the methodology described in Section 4: GOSAT observations within  $\pm 1^\circ$  in latitude and longitude of the TCCON sites were compared with the average of TCCON observations acquired within  $\pm 30$  min of the corresponding GOSAT overpass. The bias and standard deviation estimates are given in Table 2, and the relative bias is plotted as a function of the TCCON station latitude in Figure 3. The parameters of the linear least-squares fitting and the correlation coefficients are given in Table 3. Results are sorted in order of decreasing latitude from Sodankylä (67.37° N) to Lauder (45.04° S).

**Table 2.** Results of the comparison between TANSO-FTS scans acquired within  $\pm 1^\circ$  in latitude and in longitude of the TCCON sites and the average of TCCON measurements within  $\pm 30$  min of the corresponding GOSAT overpasses. The number of matched scans is given. The absolute and relative values of the mean bias and standard deviation (SD) are indicated for each station. The ensemble and site-by-site results are also given.

TCCON Dataset	# of Scans	Bias $\pm$ SD (ppm)	Bias $\pm$ SD (%)
Sodankylä	30	$-309.3 \pm 555.6$	$-10.67 \pm 17.34$
Bialystok	19	$204.7 \pm 521.3$	$15.18 \pm 30.65$
Bremen	10	$20.9 \pm 40.9$	$0.86 \pm 2.55$
Karlsruhe	39	$-213.6 \pm 307.2$	$-8.06 \pm 11.50$
Orléans	94	$-170.6 \pm 439.4$	$-4.98 \pm 16.61$
Garmisch	58	$-84.9 \pm 446.5$	$2.57 \pm 27.22$
Park Falls	138	$-91.8 \pm 797.6$	$1.52 \pm 36.98$
Lamont	284	$-229.1 \pm 611.9$	$-6.22 \pm 22.68$
Tsukuba	327	$-71.4 \pm 249.8$	$-2.21 \pm 16.77$
Edwards	62	$380.7 \pm 410.4$	$26.79 \pm 27.23$
JPL	64	$-268.0 \pm 636.8$	$-10.47 \pm 34.91$
Pasadena	106	$2.8 \pm 575.4$	$-0.47 \pm 26.23$
Saga	69	$-76.6 \pm 417.3$	$-2.40 \pm 20.25$
Darwin	113	$-413.1 \pm 808.2$	$-8.97 \pm 21.59$
Wollongong	118	$-406.7 \pm 495.7$	$-15.53 \pm 19.50$
Lauder	167	$-41.8 \pm 232.5$	$-1.40 \pm 15.36$
Ensemble bias	1698	$-138.8 \pm 542.9$	$-3.09 \pm 24.04$
Station bias	16	$-110.5 \pm 208.3$	$-1.53 \pm 10.35$



**Figure 3.** Mean relative bias (filled circles) and associated standard deviation (“error bars” representing  $\pm \sigma$ ) as a function of the latitude of the TCCON sites, for coincidence criteria of  $\pm 30$  min and  $\pm 1^\circ$  in latitude and longitude. The dataset names and corresponding number of coincidences are shown on the right-hand side, color-coded from purple to red in order of decreasing latitude from the northernmost site (Sodankylä,  $67.4^\circ$  N) to the southernmost station (Lauder,  $45.0^\circ$  S). The size of the symbols is proportional to the number of coincidences at each site.

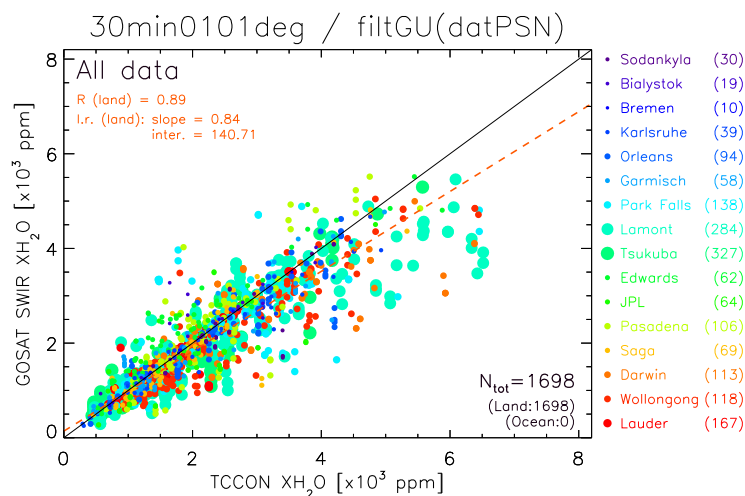
**Table 3.** Linear regression parameters (slope and intercept) and correlation coefficient (R) for TANSO-FTS scans acquired over land within  $\pm 1^\circ$  in latitude and in longitude of the TCCON sites and the average of TCCON measurements within  $\pm 30$  min of the corresponding GOSAT overpasses.

TCCON Dataset	# of Scans	Slope (ppm/ppm)	Intercept (ppm)	R
Sodankylä	30	0.64	558	0.76
Bialystok	19	0.81	553	0.84
Bremen	10	1.03	−39	1.00
Karlsruhe	39	0.90	32	0.98
Orléans	94	0.88	96	0.93
Garmisch	58	0.82	238	0.93
Park Falls	138	0.72	339	0.71
Lamont	284	0.81	141	0.90
Tsukuba	327	0.87	109	0.95
Edwards	62	1.08	165	0.94
JPL	64	0.80	18	0.72
Pasadena	106	1.03	−63	0.87
Saga	69	0.96	−5	0.93
Darwin	113	0.57	878	0.77
Wollongong	118	0.85	−60	0.93
Lauder	167	0.91	87	0.93
All coincidences	1698	0.84	141	0.89
Station-to-station	16	0.77	393	0.86

Overall, the TANSO-FTS measurements compare quite well to the TCCON data, with a negative ensemble bias of  $-3.09\%$ . The associated standard deviation is  $24.04\%$  (Table 2). This value not only includes the combined random errors of both datasets, but also contains a measure of the variability and inhomogeneities of the atmospheric water vapor distribution, which are very large. The comparison of the single-site values with a climatological knowledge of  $H_2O$  variability at each location might later yield a precision estimate for the GOSAT SWIR dataset, but it is beyond the scope of this paper. Site by site, biases range from  $-15.53\%$  (Wollongong) to  $26.79\%$  (largest difference, for Edwards). The best agreement ( $-0.47\% \pm 26.23\%$ ) is found at Pasadena, although the scatter is quite high. The standard deviations range from  $\sim 12\%$  at the Karlsruhe site to  $\sim 37\%$  at Park Falls, except for Bremen, where the standard deviation is very low (but the number of coincidences quite small). While these biases cover a rather broad range, they actually remain within  $\pm 3\%$  for seven of the 16 datasets, with no apparent systematic latitude bias (Figure 3). This shows good overall consistency between the GOSAT and TCCON datasets, further reflected in the reasonable values of the station bias and associated standard deviation of  $-1.53\% \pm 10.35\%$ .

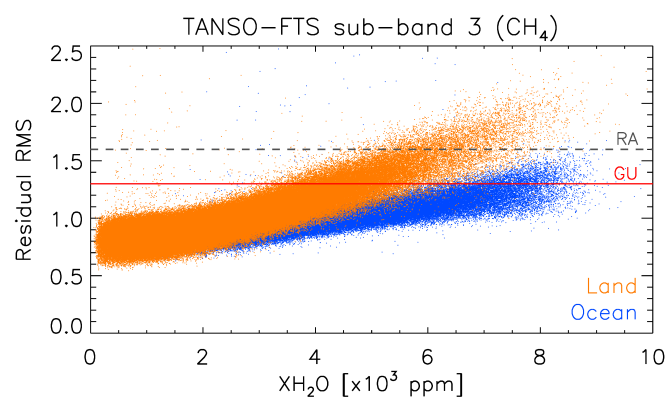
To further examine potential systematic biases, we analyze the  $XH_2O$  scatter diagram for the ensemble comparison and for each TCCON site. Aside from the overall low bias previously noted, the scatter diagram for the ensemble set (Figure 4) reveals another systematic effect. While the coincident TANSO-FTS and TCCON mole fractions are in good agreement up to  $\sim 4500$  ppm, the TANSO-FTS values corresponding to larger TCCON amounts ( $5000$ – $6500$  ppm) are more scattered and consistently smaller (within  $4500$ – $5500$  ppm). This discrepancy increases for larger values of the TCCON  $XH_2O$ . This might be due to the difference in the observation geometries of TANSO-FTS and of the TCCON instruments. The ground-based TCCON FTSs perform direct solar absorption measurements; thus, they are virtually unaffected by atmospheric scattering and measure the full water vapor column. Conversely, TANSO-FTS observes reflected sunlight and likely exhibits a high sensitivity to cloud and aerosol scattering. In situations of high humidity, represented by the largest  $XH_2O$  values of the TCCON data, the increased presence of clouds and aerosols induces more scattering, thereby shortening the solar radiation path length and reducing the possibility of sounding the lowermost

tropospheric layers (where most of the water vapor is located). This, in turn, would explain the underestimated  $\text{XH}_2\text{O}$  values, as well as the worse spectral residuals.



**Figure 4.** Scatter plot of the GOSAT TANSO-FTS  $\text{XH}_2\text{O}$  and coincident TCCON soundings (criteria of  $\pm 30$  min and  $\pm 1^\circ$  in latitude/longitude). For these criteria, there are no coincident TANSO-FTS ocean scans. The caption and color-coding are identical to those of Figure 3.

Such an effect is visible in Figure 5, where the spectral residuals in the methane sub-band (Sub-band 3,  $1.67 \mu\text{m}$ ) have been plotted separately for the complete V02.21 dataset (not only the coincident scans) after applying all pre- and post-processing filters, except the RMS screening. The additional fact that TANSO-FTS ocean measurements, coincident with TCCON sites primarily in the Southern Hemisphere (Darwin, La Réunion), seem relatively unaffected (residuals in blue, Figure 5; scatter diagrams available as part of the Supplementary Material), tends to reinforce this hypothesis, since aerosol amounts are generally larger over land and in the Northern Hemisphere.



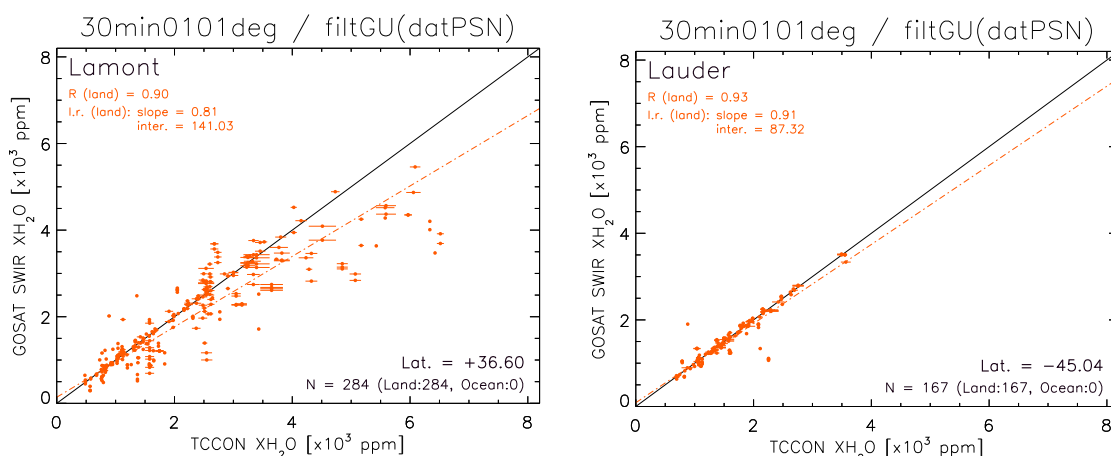
**Figure 5.** Root-mean-square residuals (retrieved spectrum *minus* simulation) in the methane sub-band as a function of the retrieved  $\text{XH}_2\text{O}$ , after applying all filters corresponding to General User (GU) screening, except the RMS filters (Table A1, Appendix A). All successful retrievals for the V02.21 dataset are included. Each dot represents a successful retrieval, in orange and blue for land and ocean measurements, respectively. The filtering thresholds are indicated for the Research Announcement (RA) (dashed grey line) and GU (solid red line) screening levels.

The linear regression (Table 3) reflects the overall low (“dry”) bias of GOSAT, increasing with increasing humidity (slope lower than unity), but also shows a slight wet bias of TANSO-FTS in



low- $\text{XH}_2\text{O}$  cases (the associated intercept is positive). The slope and intercept for the ensemble comparison are 0.84 ppm/ppm and +141 ppm, respectively. Site by site, the properties of the linear regression curves are similar to the ensemble bias, with slopes noticeably lower than one, but positive intercepts. For Bremen and Pasadena, an opposite behavior is noted: slopes larger than unity and negative intercept (Table 3). The exception is Edwards, where a consistent wet bias of TANSO-FTS can be seen for all coincidences (a tentative explanation is given in Section 5.3). Similar general features—overall dry bias with respect to ground-based instruments, reasonably small biases associated with large standard deviations, significant biases in high-humidity cases or cloudy conditions even if the measurements pass the cloud filters—have already been noted for other satellite instruments when compared to ground-based measurements (e.g., [15,63–65]).

We illustrate this further with the examples of Lamont and Lauder. For Lamont (Figure 6, left panel), GOSAT and TCCON are in good agreement up to  $\text{XH}_2\text{O}$  values of  $\sim 4000$  ppm. Above this value, the scatter increases, but the TANSO-FTS mole fractions remain within 2800–5000 ppm, while the TCCON values become significantly larger (4000–6500 ppm). On the contrary, the scatter diagram for Lauder shows very good consistency, most likely due to a low abundance and small variations of  $\text{H}_2\text{O}$  throughout the comparison period, with a maximum value below 4000 ppm (Figure 6, right panel).



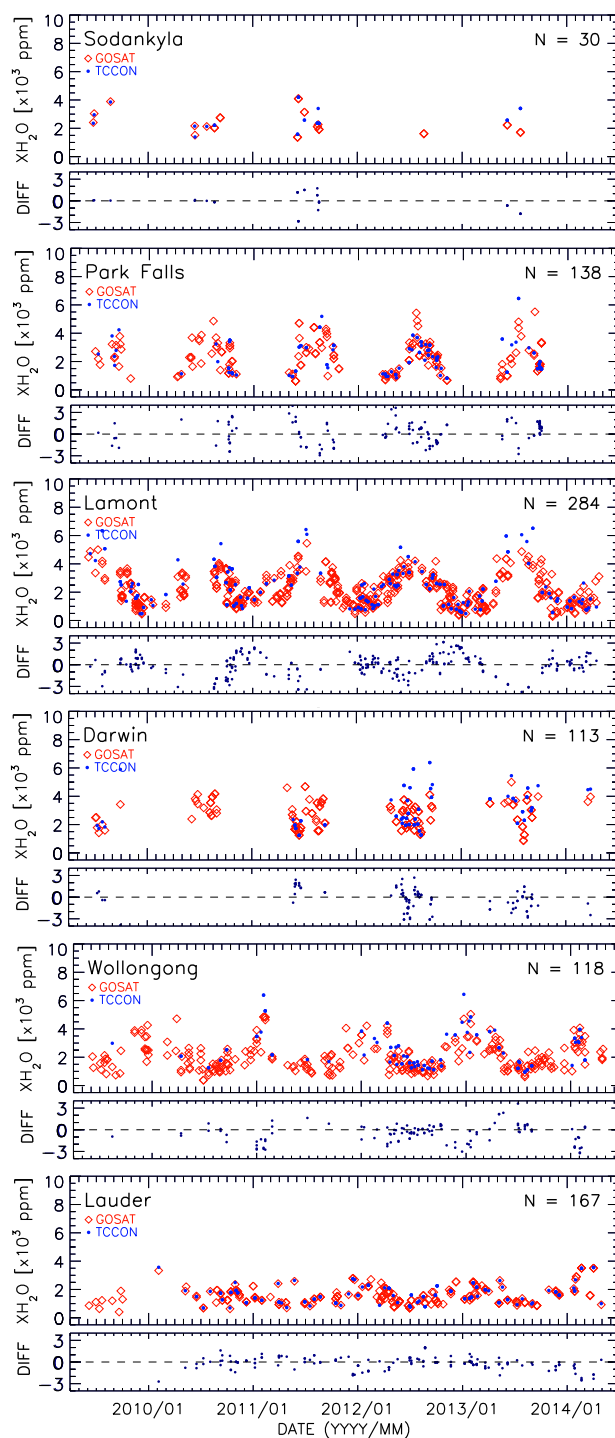
**Figure 6.** Scatter plot of the GOSAT TANSO-FTS  $\text{XH}_2\text{O}$  and coincident TCCON soundings (criteria of  $\pm 30$  min and  $\pm 1^\circ$  in latitude/longitude) at the Lamont (left) and Lauder (right) TCCON sites. All coincidences were found for TANSO-FTS land scans.

## 5.2. Temporal Evolution: Time Series for Selected Stations

To check the consistency of the TANSO-FTS SWIR and TCCON retrievals, not only in terms of absolute amounts, but also in terms of temporal (and latitudinal) variations, it is useful to analyze the time series of the retrieved  $\text{XH}_2\text{O}$  over the whole comparison period (3 June 2009–30 April 2014), separately for each TCCON location. In Figure 7, we present such time series at six selected TCCON sites. For each site, the TANSO-FTS and TCCON  $\text{XH}_2\text{O}$  are shown in the upper panel and the absolute differences (GOSAT-TCCON in ppm) in the lower panel.

The TANSO-FTS data used for these time series are spatially collocated with the TCCON observations, but contrary to the nominal statistical analysis, we did not apply any temporal-matching criterion. Conversely, the TCCON data are the same as those used in the nominal analysis. The six sites are arranged from top to bottom according to their latitude: Sodankylä represents the “northern high latitudes” ( $60^\circ \text{N}$ – $70^\circ \text{N}$ ), Lamont and Park Falls the “northern subtropical and mid-latitudes” ( $30^\circ \text{N}$ – $40^\circ \text{N}$  and  $45^\circ \text{N}$ – $55^\circ \text{N}$ , respectively), Darwin the “Southern Tropics” ( $15^\circ \text{S}$ –Equator) and Wollongong and Lauder the “southern sub-tropical and mid-latitudes” ( $45^\circ \text{S}$ – $30^\circ \text{S}$ ). Sodankylä and the Southern Hemisphere sites are the only TCCON sites, at these latitudes, for which we found coincidences using the nominal criteria. Park Falls and Lamont are shown for the northern

mid-latitudes because they present the longest record of TCCON observations in their respective latitude band, from May 2004 at Park Falls and July 2008 at Lamont.



**Figure 7.** Time series of  $\text{XH}_2\text{O}$  at six TCCON sites for collocated TANSO-FTS data ( $\pm 1^\circ$  latitude/longitude, no time constraint) and for the average of TCCON measurements acquired within  $\pm 30$  min of a GOSAT overpass. TCCON sites are ordered from top to bottom by decreasing latitude. For each site, the top panel shows the  $\text{XH}_2\text{O}$  time series of GOSAT (red diamonds) and TCCON (blue circles). Bottom panel: absolute differences (GOSAT–TCCON) for spatially- and temporally-coincident pairs.

These time series show that the ground-based instruments and TANSO-FTS are well able to trace the natural variations of atmospheric water vapor. Firstly,  $\text{H}_2\text{O}$  is primarily found in the tropical and sub-tropical troposphere (see also Figure 1). Its abundance decreases with increasing latitude. For instance, for Northern Hemisphere sites, the TCCON instruments register yearly (summer) maxima of  $\sim 4200$  ppm only at Sodankylä, but  $\sim 6500$  ppm over Lamont; in the Southern Hemisphere, the maxima are  $\sim 6400$  ppm and  $\sim 3600$  ppm for Darwin and Lauder, respectively.

Secondly,  $\text{XH}_2\text{O}$  shows a clear seasonal variation even at latitudes with generally low  $\text{H}_2\text{O}$  (Sodankylä, Lauder). This is especially verified in the mid-latitudes and tropical region, where the  $\text{XH}_2\text{O}$  values are representative of a rather dry atmosphere (less than 400–500 ppm) in the wintertime or dry season, but increase significantly in the summer or tropical wet season. This strong seasonal cycle is qualitatively well traced by both the ground-based FTSs and TANSO-FTS, although the TANSO-FTS  $\text{XH}_2\text{O}$  values are biased significantly lower for higher TCCON values, as described earlier. The expected half-year phase difference between the Northern and Southern Hemisphere sites is also clearly seen. At Lauder, the time series confirms the conclusions drawn from the scatter diagram: the  $\text{H}_2\text{O}$  abundance remains low throughout the year, with mean values of  $\sim 1600$  ppm and limited seasonal variations.

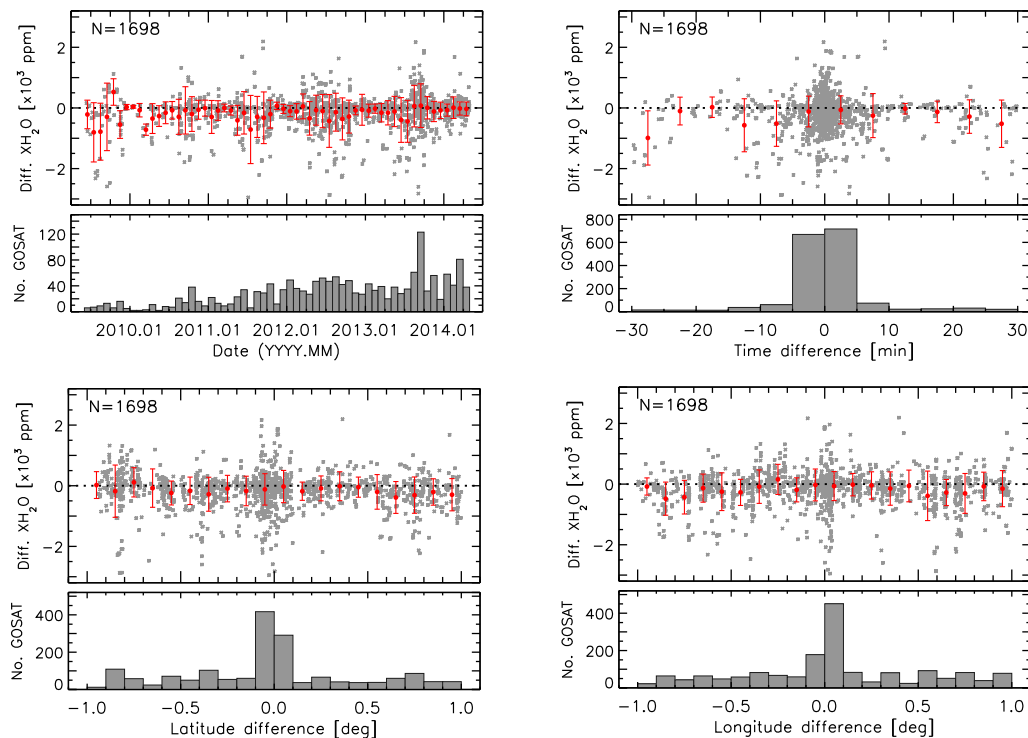
Finally, some characteristics of the TANSO-FTS observations are also illustrated in Figure 7: the gaps in the temporal coverage and the density of GOSAT points in each panel are directly related to the screening of the TANSO-FTS  $\text{XH}_2\text{O}$  data. There are no data points in winter for the northern high latitudes, as shown by the time series at Sodankylä (Figure 7, top panel). Since the screening condition limits the SZA values to below  $70^\circ$  (Table A1, Appendix A), all of the winter high-latitude data are rejected during the pre-screening. Gaps in the temporal coverage are also found at Darwin (fourth panel from the top). This is because of the density of the cloud coverage in the Tropics throughout the year, but especially during the wet season (November–April).

### 5.3. Impact of the Comparison Characteristics on the Single-Scan Differences

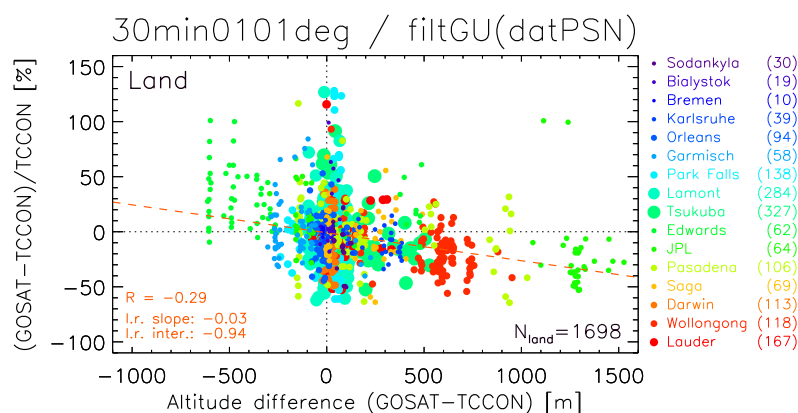
In addition to the statistical comparison, we explore potential issues that are important to understand the comparison results: temporal and spatial mismatches between GOSAT scans and TCCON measurements (Figures 8 and 9) and the impact of geophysical or retrieval parameters on the  $\text{XH}_2\text{O}$  differences (Figure 10). The evolution of the single-scan absolute differences and the corresponding histograms of the number of TANSO-FTS scans, with respect to the measurement date and to the geolocation characteristics, is presented in Figure 8. The relative differences as a function of the altitude difference between TANSO-FTS scans and TCCON sites are shown separately in Figure 9. Lastly, the absolute differences and the corresponding histograms, relative to selected geophysical and retrieval parameters, are plotted in Figure 10. The results shown in Figures 8–10 were obtained using the nominal coincidence criteria.

To assess the pertinence of our choice of geophysical coincidence criteria and to evaluate the consistency of the TANSO-FTS observations over time, we plot the differences as a function of the date of the measurements (time series) and the time, latitude and longitude differences between the TANSO-FTS scans and the TCCON locations (Figure 8). Mean differences are calculated within each histogram bin (red symbols with “error bars” representing  $\pm \sigma$ ). The ensemble time series (Figure 8, top left panel) shows a seasonal variation of the monthly mean differences with an amplitude of  $\sim 500$ – $1000$  ppm, the largest negative values during the Northern Hemisphere summer and near-zero or slightly positive values during northern winter. This is a composite effect of the seasonal variations observed at each site, previously illustrated by the single-site  $\text{XH}_2\text{O}$  time series (Figure 7). The large negative values are likely explained by the increasingly low bias of TANSO-FTS  $\text{XH}_2\text{O}$  for larger TCCON mole fractions. This could appear in the comparisons as a seasonal or latitudinal bias, since it will essentially impact the results in the mid-latitudes in summer, when atmospheric water vapor is the most abundant. Indeed, while the largest  $\text{XH}_2\text{O}$  discrepancies should be found in the Tropics (largest atmospheric amounts of water vapor), most of the TANSO-FTS tropical data were already filtered out by the pre-processing cloud screening. The histogram associated with the time series (Figure 8, top left

panel) shows another interesting feature: the monthly number of coincident GOSAT scans appears to be increasing with time. This is directly due to the increasing number of operational TCCON stations over the duration of the GOSAT mission and, therefore, of the quantity of data available for comparison.

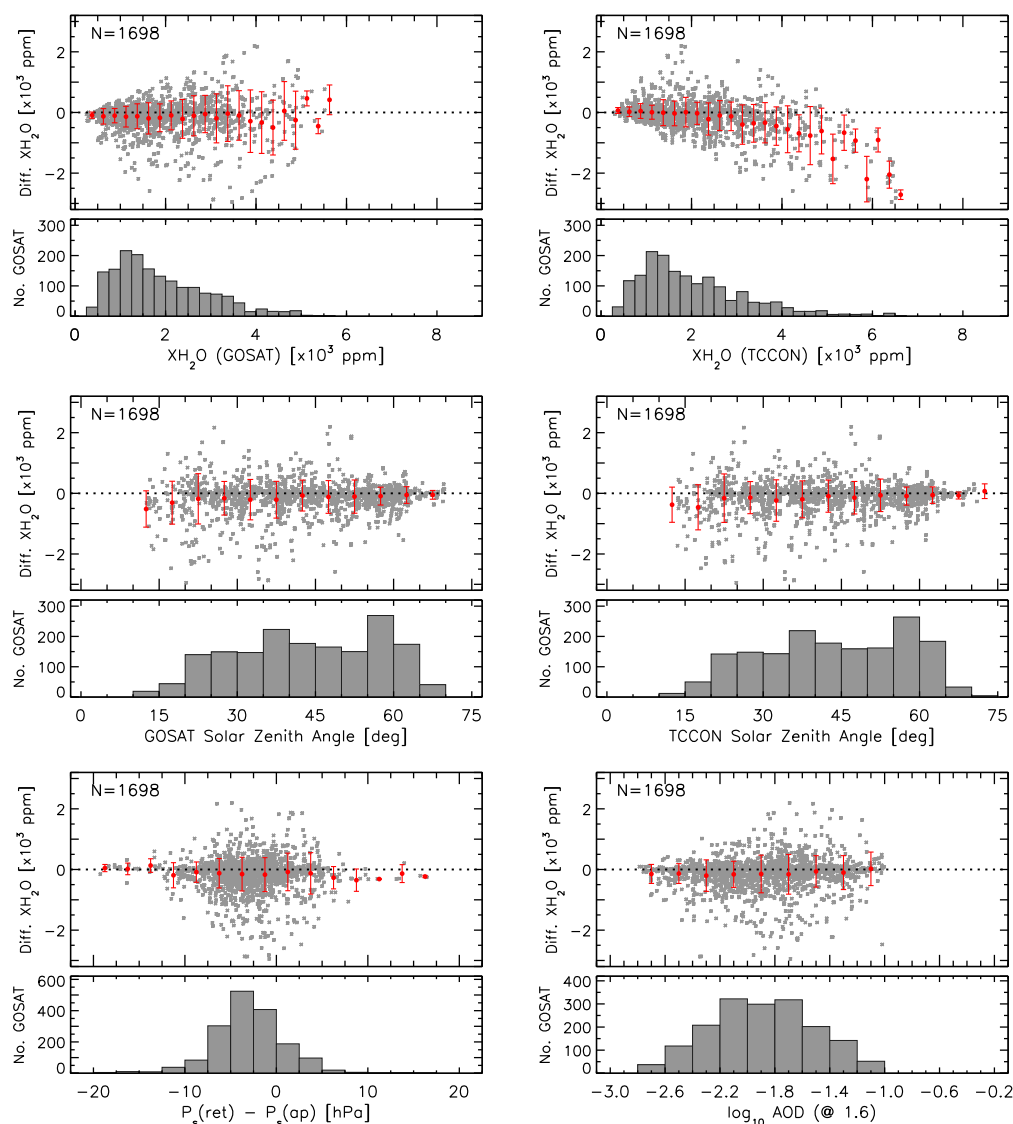


**Figure 8.** Evolution of the XH<sub>2</sub>O absolute differences (GOSAT–TCCON) for the nominal coincidence criteria ( $\pm 1^\circ$  latitude/longitude and  $\pm 30$  min) as a function of the measurement date (time series, top left panel) and of the collocation characteristics: time, latitude and longitude differences (top right, bottom left and bottom right panels, respectively). The corresponding histograms of the number of TANSO-FTS scans are plotted below each panel. The grey dots represent the single-scan differences; the red symbols with “error bars” show the average value and associated standard deviation within each histogram bin.



**Figure 9.** Relative differences (GOSAT–TCCON)/TCCON as a function of the difference, in meters, between the retrieved altitude of the GOSAT footprints and the altitude of the TCCON sites, for GOSAT land scans only. The caption and color-coding are identical to those of Figure 3.





**Figure 10.** Evolution of the  $\text{XH}_2\text{O}$  absolute differences (GOSAT–TCCON) for the nominal coincidence criteria ( $\pm 1^\circ$  latitude/longitude and  $\pm 30$  min), as a function of geophysical and retrieval parameters: the TANSO-FTS and TCCON  $\text{XH}_2\text{O}$  (top row), the solar zenith angle values for GOSAT and TCCON (middle row), the difference between the retrieved and the *a priori* values for the surface pressure (bottom left) and the aerosol optical depth at  $1.6 \mu\text{m}$  retrieved from the TANSO-FTS spectra (bottom right). The corresponding histograms of the number of TANSO-FTS scans are plotted below each panel. The grey dots represent the single-scan differences; the red symbols with “error bars” show the average value and associated standard deviation within each histogram bin.

Finally, there is no apparent systematic bias of the mean  $\text{XH}_2\text{O}$  differences directly related to the collocation parameters (top right, bottom left and bottom right panels of Figure 8). It is interesting to note that the corresponding histograms show a clear peak around the zero-difference values. The small time differences are due to the high frequency of TCCON measurements at each site, which allows for very close temporal matches between TCCON measurements and GOSAT overpasses. The large number of spatially close coincidences illustrates the critical importance of target-mode observations for GOSAT validation studies. Without the target mode, there would indeed be too few coincidences at most TCCON sites. For example, target-mode observations around Park Falls, Lamont and Tsukuba respectively account for 67%, 99% and 100% of the coincidences found at each site using the nominal

criteria ( $\pm 1^\circ$  in latitude/longitude,  $\pm 30$  min in time). The exceptions are Sodankylä (3%), Darwin (8%), the JPL (30%) and Lauder (36%), which are sufficiently close to the standard GOSAT scanning pattern footprints to be observed routinely without requiring target-mode observations.

We also separately check for potential biases relative to the altitude difference between the TANSO-FTS footprints and the elevation of the TCCON instruments (Figure 9). Here, it is necessary to examine relative values, because the range of variation of the absolute  $\text{XH}_2\text{O}$  differences is too large. Due to the vertical distribution of atmospheric water vapor, if the retrieved altitude of a scan is higher than that of the coincident TCCON site (positive altitude difference), the retrieved mole fraction should be lower. Conversely, if the mean altitude within the TANSO-FTS footprint is lower than that of the TCCON instrument, the differences should be positive because of the additional water vapor measured by TANSO-FTS. Such a pattern is visible in Figure 9, with a global tendency for the  $\text{XH}_2\text{O}$  differences to decrease with increasing altitude difference. Specifically, there is a clear negative correlation ( $R = -0.29$ ), and the slope of the linear fitting curve is negative ( $-0.03\%/m$ ) with an intercept close to zero ( $-0.94\%$ ). This is consistent for nearly all individual sites, for which we note a negative trend similar to the ensemble result. It can be illustrated by the comparisons at Edwards, a high altitude site in the Mojave Desert not far inland from Los Angeles in California, and at Wollongong, a seaside city south of Sidney in Australia, which is also close to a nearby inland mountain range. The comparison for Edwards (altitude 700 m, dry weather) yields the largest positive mean bias ( $+26.79\%$ ; Table 2) with single-scan biases almost exclusively positive. This is because there are no spatially close coincidences for Edwards (the shortest distance between coincident measurements is  $\sim 80$  km), and all TANSO-FTS footprints are located in the Los Angeles basin oceanward of the TCCON site, therefore mostly at significantly lower altitudes. Conversely, for Wollongong (altitude 30 m, seaside location), the single-scan biases are mostly negative, and the mean bias is the largest negative result of all of the sites considered here ( $-15.53\%$ ). This is due to the fact that the GOSAT land footprints are exclusively inland at higher altitudes than the TCCON station, which is located on the coastal plain.

To determine whether geophysical conditions or retrieval characteristics have any impact on the TANSO-FTS SWIR retrievals, we analyze the GOSAT–TCCON differences with respect to the physical variables or retrieval parameters of the NIES processing: the retrieved TANSO-FTS and TCCON  $\text{H}_2\text{O}$  mole fractions, the SZA for GOSAT and TCCON, the difference between the retrieved surface pressure and its prior for TANSO-FTS, as well as the retrieved AOD at  $1.6 \mu\text{m}$ . The latter two parameters are used during the NIES processing as post-screening filters [17,62]. The aerosol optical depth filter is particularly stringent for GU-level data, since all data with an AOD value larger than 0.1 are filtered out (Table A1). The low bias of the TANSO-FTS data for large  $\text{XH}_2\text{O}$  values is immediately apparent in the upper panels of Figure 10, with mean differences almost constant and close to zero for all GOSAT  $\text{XH}_2\text{O}$ , but becoming increasingly negative for increasingly larger values of TCCON  $\text{XH}_2\text{O}$ . The corresponding linear correlation coefficient, which should be zero in the absence of a systematic bias (as is the case for GOSAT  $\text{XH}_2\text{O}$ ), is quite large ( $R = -0.48$ ). There also seems to be a slight bias with respect to the SZA values for both datasets, with differences becoming larger and negative for SZAs smaller than  $25^\circ$  and a corresponding correlation of  $\sim 0.17$  for both GOSAT and TCCON. Finally, the comparison results do not seem to show any dependence with respect to the surface pressure difference and the retrieved AOD, with mean differences close to zero and no visible trend (lower panels of Figure 10).

## 6. Conclusions

We conducted initial analyses of the water vapor column-averaged dry-air mole fraction ( $\text{XH}_2\text{O}$ ) research data product, Version 02.21 (V02.21), retrieved with the Level 2 (L2) processing algorithm of the National Institute for Environmental Studies (NIES) from the Short-Wavelength Infrared (SWIR) spectra of the Thermal And Near-infrared Sensor for carbon Observation-Fourier Transform Spectrometer (TANSO-FTS) on board the Greenhouse gases Observing Satellite (GOSAT). We used a

simple methodology to compare the NIES L2 retrievals with the latest processing version (GGG2014) of coincident data from the ground-based Total Carbon Column Observing Network (TCCON) and checked the robustness of the comparison results.

We found a good agreement between GOSAT and TCCON, with an ensemble low bias of the TANSO-FTS  $\text{XH}_2\text{O}$  data of  $-3.1\%$  and single-station mean biases within  $\pm 3\%$  for seven of the 16 TCCON datasets. The related standard deviation is a measure, not only of the combined precision of TANSO-FTS and the TCCON instruments, but also of the natural  $\text{H}_2\text{O}$  variability. We found quite large values ranging from  $11\%$ – $37\%$  with an ensemble value of  $\sim 24\%$ . Furthermore, there was good consistency between different TCCON sites, with a station bias and standard deviation of  $-1.53\% \pm 10.35\%$ . This seems to indicate the absence of significant regional-scale biases in the NIES dataset. Complementary analyses also confirmed the general absence of systematic artifacts due to the comparison characteristics or retrievals parameters.

However, we were able to identify two factors that largely contribute to the observed biases, globally or for individual TCCON sites. We found a negative bias of the NIES  $\text{XH}_2\text{O}$  retrievals relative to the TCCON data, becoming increasingly larger for higher water vapor abundances. For TCCON  $\text{XH}_2\text{O}$  data values larger than  $4000$  ppm and up to  $\sim 6500$  ppm, the coincident NIES V02.21 data seem to reach a peak value of  $\sim 5000$ – $5500$  ppm, and the discrepancy between the datasets increases with increasing TCCON mole fractions. When analyzing the global dataset, this translates into latitudinal and seasonal biases, because the atmospheric water content is largest at mid-latitudes and in the tropical region during summer. This significant dry bias seems to characterize most satellite observations of water vapor. It has been noted in previous studies and shows the impact of cloud coverage and atmospheric aerosols on satellite-borne measurements (especially in the short-wavelength infrared region), amplified for TANSO-FTS by the observation geometry (nadir sounding of reflected sunlight). We also identified an altitude dependence of the GOSAT–TCCON differences. There is a significant negative correlation ( $R = -0.29$ ) between the  $\text{XH}_2\text{O}$  relative differences and the altitude difference between the TANSO-FTS footprints and the TCCON sites. This is expected from the characteristics of the vertical distribution of atmospheric water vapor, with the largest amounts at the lowermost altitudes and an exponential decrease with height. These two factors combined potentially account for most of the  $\text{XH}_2\text{O}$  bias observed between the TANSO-FTS and TCCON datasets.

Additional studies using a refined methodology will be undertaken. Thorough characterization of the data screening and optimization of the filtering parameters for  $\text{XH}_2\text{O}$  are ongoing at NIES. Our results show that the NIES retrieval algorithm (V02.21) is already successfully retrieving  $\text{XH}_2\text{O}$  from the TANSO-FTS spectra. After their public release, these data will be available to general users for scientific studies of the water cycle.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2072-4292/8/5/414/s1](http://www.mdpi.com/2072-4292/8/5/414/s1): Figures S1–S22 showing the ensemble scatter diagrams, the sensitivity of the  $\text{XH}_2\text{O}$  differences to the collocation characteristics and their altitude dependence, for varying coincidence criteria, gathered in one supplementary PDF document; corresponding statistical results presented in tables gathered in one supplementary Microsoft Excel workbook.

**Acknowledgments:** This paper is essential for the NIES GOSAT Project Office to ensure the qualification of the GOSAT TANSO-FTS SWIR  $\text{XH}_2\text{O}$  data as a standard product.

Nicholas Deutscher is supported by an Australian Research Council fellowship, DE140100178. TCCON data were obtained from the TCCON Data Archive, hosted by the Carbon Dioxide Information Analysis Center (CDIAC): <http://tcon.ornl.gov>. The TCCON Network is supported by NASA through a grant to the California Institute of Technology: NNX14AI60G. Park Falls, Lamont and JPL are funded by NASA grants NNX14AI60G, NNX11AG01G, NAG5-12247, NNG05-GD07G and the NASA Orbiting Carbon Observatory Program. We are grateful to the United States of America's Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) program for technical support in Lamont and to Jeff Ayers for technical support in Park Falls. Darwin and Wollongong are funded by NASA grants NAG5-12247 and NNG05-GD07G and the Australian Research Council grants DP140101552, DP110103118, DP0879468, LE0668470 and LP0562346. We are grateful to the DOE ARM program for technical support in Darwin. Bremen, Bialystok and Orléans are funded by the European Union's projects InGOS (Integrated non- $\text{CO}_2$  Greenhouse gas Observing System) and ICOS-INWIRE (Integrated Carbon Observing System-Improved sensors, Network and Interoperability for GMES (Global Monitoring for Environment and Security)) and by the Senate of Bremen. From 2004–2011, the Lauder TCCON program was

funded by the New Zealand Foundation of Research, Science and Technology contracts CO1X0204, CO1X0703 and CO1X0406. Since 2011, the program has been funded by New Zealand's National Institute of Water and Atmospheric Research (NIWA)'s Atmosphere Research Programme 3 (2011/2013 Statement of Corporate Intent). Garmisch water vapor retrievals have been supported by the Bavarian Ministry of the Environment and Consumer Protection. Indianapolis and Edwards measurements are supported by the NASA Earth Science Division.

**Author Contributions:** Eric Dupuy performed all of the NIES/TCCON comparisons for XH<sub>2</sub>O presented in this paper within the NIES GOSAT Project, produced the figures and tables and wrote the entire paper. Nicholas Deutscher conducted the counterpart XH<sub>2</sub>O comparisons among the TCCON partners; he collaborated closely throughout this analysis and the conception of the paper. Isamu Morino leads the GOSAT/GOSAT-2 Validation Team and took extensive part in the data analysis and the writing and review of this paper. Yukio Yoshida is responsible for the Level 2 processing of the GOSAT SWIR data at NIES; he was instrumental in the interpretation of the validation results from the point of view of retrieval biases. Osamu Uchino is the GOSAT Validation Manager and oversees the validation for all GOSAT data products at NIES, including XH<sub>2</sub>O Version V02.21. Tatsuya Yokota and Tsuneo Matsunaga are the NIES Project Leaders for GOSAT and GOSAT-2, respectively, and coordinate all aspects of the GOSAT and GOSAT-2 Projects within NIES. Other authors are members of the TCCON Steering Committee and principal investigators or co-investigators of TCCON stations; they contributed to the acquisition, processing and delivery of the TCCON data products for the sites included in this study. All authors provided feedback on the data analysis and the writing of this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Filtering criteria (pre- and post-processing) used for the TANSO-FTS products at NIES (Version 02.21). Some criteria have different thresholds for, or are specific to, different release levels: the “Research Announcement” (RA) and “General Users” (GU) data.

Variable Tested	Code	Level	Rejection Condition	Explanation
PRE-SCREENING ITEMS				
L1B quality flag <sup>†</sup>	11q	RA, GU	11q = 1	Quality of the calibrated spectra
CAI radiances flag <sup>†</sup>	cai	GU	cai = 1	CAI coherence test
2 $\mu$ m-scattering flag <sup>†</sup>	mu2	RA, GU	mu2 = 1	High-altitude scattering
Solar zenith angle flag <sup>†</sup>	sza	RA, GU	sza = 1	Reject SZA larger than 70°
Land fraction	ldf	RA, GU	0% < ldf < 60%	“Too mixed” (< 60% land)
Signal-to-noise ratio	o2s	RA, GU	o2s < 70.	Minimum SNR in the O <sub>2</sub> -A band
POST-SCREENING ITEMS				
Number of iterations	itr	RA, GU	itr ≥ 20	Convergence not reached
Residuals Sub-band 1	rb1	RA	rb1 ≥ 1.4	RMS Band 1 (O <sub>2</sub> -A) *
		GU	rb1 ≥ 1.2	
Residuals Sub-band 2	rb2	RA	rb2 ≥ 1.5	RMS Band 2 (wCO <sub>2</sub> ) *
		GU	rb2 ≥ 1.2	
Residuals Sub-band 3	rb3	RA	rb3 ≥ 1.6	RMS Band 2 (CH <sub>4</sub> ) *
		GU	rb3 ≥ 1.3	
Residuals Sub-band 4	rb4	RA	rb4 ≥ 1.6	RMS Band 3 (SCO <sub>2</sub> ) *
		GU	rb4 ≥ 1.4	
Degrees of freedom	dfs	RA	dfs < 0.8	Minimum column information content
		GU	dfs < 1.	
Aerosol optical thickness	aot	RA	aot > 0.5	Estimated value at 1.6 $\mu$ m
		GU	aot > 0.1	
Blended albedo **	bla	RA, GU	bla ≥ 1.	Land scans only (if ldf ≥ 60.)
Surface wind speed	wsp	RA, GU	wsp ≤ 0.1 m.s <sup>-1</sup>	Ocean scans only (if ldf = 0.)
			or wsp ≥ 20. m.s <sup>-1</sup>	
Retrieved surface pressure	dsp	RA, GU	dsp > 20. hPa	abs (retrieved – a priori)
DATE SELECTION	dtm	RA, GU	Earlier than 03 June 2003	Operational period only

<sup>†</sup> Flags: tests with binary result values. If the test is passed, the flag value is set to 0, else it is set to 1. \* RMS: root-mean-square of the residuals. The sub-bands are labeled after the major absorption features in the TANSO-FTS spectral bands at these positions: Sub-band 1 (Band 1) for O<sub>2</sub>; Sub-bands 2 and 3 (both parts of Band 2) for a weak CO<sub>2</sub> and a CH<sub>4</sub> absorption features, respectively; Sub-band 4 (Band 3) for a stronger CO<sub>2</sub> absorption. \*\* Blended albedo: defined by Wunch *et al.* [18] and empirically determined as bla = 2.4 × albedo (O<sub>2</sub>-A band) – 1.13 × albedo (SCO<sub>2</sub> band).



## References

1. Trenberth, K.E. Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change. *Clim. Chang.* **1998**, *39*, 667–694.
2. Harries, J.E. Atmospheric radiation and atmospheric humidity. *Quart. J. R. Meteor. Soc.* **1997**, *123*, 2173–2186.
3. Trenberth, K.E.; Fasullo, J.; Smith, L. Trends and variability in column-integrated atmospheric water vapor. *Clim. Dyn.* **2005**, *24*, 741–758.
4. Worden, J.; Noone, D.; Bowman, K.; the Tropospheric Emission Spectrometer science team and data contributors. Importance of rain evaporation and continental convection in the tropical water cycle. *Nature* **2007**, *445*, 528–532.
5. Held, I.M.; Soden, B.J. Robust Responses of the Hydrological Cycle to Global Warming. *J. Clim.* **2006**, *19*, 5686–5699.
6. Flato, G.; Marotzke, J.; Braconnot, P.; Chou, S.C.; Collins, W.; Cox, P.; Driouech, F.; Emori, S.; Eyring, V.; Forest, C.; et al. Evaluation of Climate Models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013; pp. 741–866.
7. Stocker, T.F.; Qin, D.; Plattner, G.K.; Alexander, L.V.; Allen, S.K.; Bindoff, N.L.; Bréon, F.M.; Church, J.A.; Cubasch, U.; Emori, S.; et al. Technical Summary. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Stocker, T.F., Qin, D., Plattner, G.K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., Eds.; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013; pp. 33–115.
8. Vogelmann, H.; Sussmann, R.; Trickl, T.; Reichert, A. Spatiotemporal variability of water vapor investigated using lidar and FTIR vertical soundings above the Zugspitze. *Atmos. Chem. Phys.* **2015**, *15*, 3135–3148.
9. Sussmann, R.; Borsdorff, T.; Rettinger, M.; Camy-Peyret, C.; Demoulin, P.; Duchatelet, P.; Mahieu, E.; Servais, C. Technical Note: Harmonized retrieval of column-integrated atmospheric water vapor from the FTIR network—first examples for long-term records and station trends. *Atmos. Chem. Phys.* **2009**, *9*, 8987–8999.
10. Vogelmann, H.; Sussmann, R.; Trickl, T.; Borsdorff, T. Intercomparison of atmospheric water vapor soundings from the differential absorption lidar (DIAL) and the solar FTIR system on Mt. Zugspitze. *Atmos. Meas. Tech.* **2011**, *4*, 835–841.
11. Ross, R.J.; Elliott, W.P. Tropospheric Water Vapor Climatology and Trends over North America: 1973–1993. *J. Clim.* **1996**, *9*, 3561–3574.
12. Kämpfer, N. (Ed.) *Monitoring Atmospheric Water Vapour*; ISSI Scientific Report Series 10; Springer: New York, NY, USA, 2013.
13. Urban, J. Satellite sensors measuring atmospheric water vapour. In *Monitoring Atmospheric Water Vapour*; ISSI Scientific Report Series 10; Kämpfer, N., Ed.; Springer: New York, NY, USA, 2013; pp. 175–214.
14. Buehler, S.A.; Östman, S.; Melsheimer, C.; Holl, G.; Eliasson, S.; John, V.O.; Blumenstock, T.; Hase, F.; Elgered, G.; Raffalski, U.; et al. A multi-instrument comparison of integrated water vapour measurements at a high latitude site. *Atmos. Chem. Phys.* **2012**, *12*, 10925–10943.
15. Van Malderen, R.; Brenot, H.; Pottiaux, E.; Beirle, S.; Hermans, C.; De Mazière, M.; Wagner, T.; De Backer, H.; Bruyninx, C. A multi-site intercomparison of integrated water vapour observations for climate change analysis. *Atmos. Meas. Tech.* **2014**, *7*, 2487–2512.
16. Kuze, A.; Suto, H.; Shiomi, K.; Nakajima, M.; Hamazaki, T. On-orbit performance and Level 1 data processing of TANSO-FTS and CAI on GOSAT. *Proc. SPIE* **2009**, *7474*, doi:10.1117/12.830152.
17. Yoshida, Y.; Kikuchi, N.; Morino, I.; Uchino, O.; Oshchepkov, S.; Bril, A.; Saeki, T.; Schutgens, N.; Toon, G.C.; Wunch, D.; et al. Improvement of the retrieval algorithm for GOSAT SWIR XCO<sub>2</sub> and XCH<sub>4</sub> and their validation using TCCON data. *Atmos. Meas. Tech.* **2013**, *6*, 1533–1547.
18. Wunch, D.; Toon, G.C.; Blavier, J.F.L.; Washenfelder, R.A.; Notholt, J.; Connor, B.J.; Griffith, D.W.T.; Sherlock, V.; Wennberg, P.O. The Total Carbon Column Observing Network. *Phil. Trans. R. Soc. A* **2011**, *369*, 2087–2112.
19. Butz, A.; Hasekamp, O.P.; Frankenberg, C.; Aben, I. Retrievals of atmospheric CO<sub>2</sub> from simulated space-borne measurements of backscattered near-infrared sunlight: accounting for aerosol effects. *Atmos. Ocean* **2009**, *48*, 3322–3336.

20. Yoshida, Y.; Ota, Y.; Eguchi, N.; Kikuchi, N.; Nobuta, K.; Tran, H.; Morino, I.; Yokota, T. Retrieval algorithm for CO<sub>2</sub> and CH<sub>4</sub> column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite. *Atmos. Meas. Tech.* **2011**, *4*, 717–734.
21. Ishida, H.; Nakajima, T.Y. Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager. *J. Geophys. Res. Atmos.* **2009**, *114*, D07206.
22. Rodgers, C.D. *Inverse Methods for Atmospheric Sounding—Theory and Practise*; Series on Atmospheric, Oceanic and Planetary Physics; World Scientific: Singapore, 2000; Volume 2.
23. Rothman, L.; Gordon, I.E.; Barbe, A.; Benner, D.C.; Bernath, P.F.; Birk, M.; Boudon, V.; Brown, L.R.; Campargue, A.; Champion, J.P.; *et al.* The HITRAN 2008 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat. Transf.* **2009**, *110*, 533–572.
24. Takemura, H.; Egashira, M.; Matsuzawa, K.; Ichijo, H.; O'ishi, R.; Abe-Ouchi, A. A simulation of the global distribution and radiative forcing of soil dust aerosols at the Last Glacial Maximum. *Atmos. Chem. Phys.* **2015**, *9*, 3061–3073.
25. The GOSAT User Interface Gateway. Available online: <http://data.gosat.nies.go.jp/> (accessed on 9 May 2016).
26. Morino, I.; Uchino, O.; Inoue, M.; Yoshida, Y.; Yokota, T.; Wennberg, P.O.; Toon, G.C.; Wunch, D.; Roehl, C.M.; Notholt, J.; *et al.* Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra. *Atmos. Meas. Tech.* **2011**, *4*, 1061–1076.
27. Inoue, M.; Morino, I.; Uchino, O.; Miyamoto, Y.; Yoshida, Y.; Yokota, T.; Machida, T.; Sawa, Y.; Matsueda, H.; Sweeney, C.; *et al.* Validation of XCO<sub>2</sub> derived from SWIR spectra of GOSAT TANSO-FTS with aircraft measurement data. *Atmos. Chem. Phys.* **2013**, *13*, 9771–9788.
28. Inoue, M.; Morino, I.; Uchino, O.; Miyamoto, Y.; Saeki, T.; Yoshida, Y.; Yokota, T.; Sweeney, C.; Tans, P.P.; Biraud, S.C.; *et al.* Validation of XCH<sub>4</sub> derived from SWIR spectra of GOSAT TANSO-FTS with aircraft measurement data. *Atmos. Meas. Tech.* **2014**, *7*, 2987–3005.
29. Frankenberg, C.; Wunch, D.; Toon, G.; Risi, C.; Scheepmaker, R.; Lee, J.E.; Wennberg, P.; Worden, J. Water vapor isotopologue retrievals from high-resolution GOSAT shortwave infrared spectra. *Atmos. Meas. Tech.* **2013**, *6*, 263–274.
30. Boesch, H.; Deutscher, N.M.; Warneke, T.; Byckling, K.; Cogan, A.J.; Griffith, D.W.T.; Notholt, J.; Parker, R.J.; Wang, Z. HDO/H<sub>2</sub>O ratio retrievals from GOSAT. *Atmos. Meas. Tech.* **2013**, *6*, 599–612.
31. O'Dell, C.W.; Connor, B.; Bösch, H.; O'Brien, D.; Frankenberg, C.; Castano, R.; Christi, M.; Eldering, D.; Fisher, B.; Gunson, M.; *et al.* The ACOS CO<sub>2</sub> retrieval algorithm—Part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.* **2012**, *5*, 99–121.
32. Butz, A.; Guerlet, S.; Hasekamp, O.; Schepers, D.; Galli, A.; Aben, I.; Frankenberg, C.; Hartmann, J.M.; Tran, H.; Kuze, A.; *et al.* Toward accurate CO<sub>2</sub> and CH<sub>4</sub> observations from GOSAT. *Geophys. Res. Lett.* **2011**, *38*, L14812.
33. Wunch, D.; Toon, G.C.; Wennberg, P.O.; Wofsy, S.C.; Stephens, B.B.; Fischer, M.L.; Uchino, O.; Abshire, J.B.; Bernath, P.; Biraud, S.C.; *et al.* Calibration of the Total Carbon Column Observing Network using aircraft profile data. *Atmos. Meas. Tech.* **2010**, *3*, 1351–1362.
34. The TCCON Data Archive. Available online: <http://tccon.ornl.gov/> (accessed on 9 May 2016).
35. The TCCON-Wiki Website. Available online: <http://tccon-wiki.caltech.edu/> (accessed on 9 May 2016).
36. Kivi, R.; Heikkinen, P.; Kyrö, E. *TCCON Data from Sodankylä, Finland, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tccon.ggg2014.sodankyla01.R0/1149280.
37. Deutscher, N.; Notholt, J.; Messerschmidt, J.; Weinzierl, C.; Warneke, T.; Petri, C.; Grupe, P.; Katrynski, K. *TCCON Data from Bialystok, Poland, Release GGG2014R1*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tccon.ggg2014.bialystok01.R1/1183984.
38. Notholt, J.; Petri, C.; Warneke, T.; Deutscher, N.; Buschmann, M.; Weinzierl, C.; Macatangay, R.; Grupe, P. *TCCON Data from Bremen, Germany, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tccon.ggg2014.bremen01.R0/1149275.
39. Hase, F.; Blumenstock, T.; Dohe, S.; Groß, J.; Kiel, M. *TCCON Data from Karlsruhe, Germany, Release GGG2014R1*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tccon.ggg2014.karlsruhe01.R1/1182416.

40. Warneke, T.; Messerschmidt, J.; Notholt, J.; Weinzierl, C.; Deutscher, N.; Petri, C.; Grupe, P.; Vuillemin, C.; Truong, F.; Schmidt, M.; Ramonet, M.; Parmentier, E. *TCCON Data from Orléans, France, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.orleans01.R0/1149276.
41. Sussmann, R.; Rettinger, M. *TCCON Data from Garmisch, Germany, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.garmisch01.R0/1149299.
42. Wennberg, P.O.; Roehl, C.; Wunch, D.; Toon, G.C.; Blavier, J.F.; Washenfelder, R.; Keppel-Aleks, G.; Allen, N.; Ayers, J. *TCCON Data from Park Falls, Wisconsin, USA, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.parkfalls01.R0/1149161.
43. Iraci, L.; Podolske, J.; Hillyard, P.; Roehl, C.; Wennberg, P.O.; Blavier, J.F.; Landeros, J.; Allen, N.; Wunch, D.; Zavaleta, J.; et al. *TCCON Data from Indianapolis, Indiana, USA, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.indianapolis01.R0/1149164.
44. Wennberg, P.O.; Wunch, D.; Roehl, C.; Blavier, J.F.; Toon, G.C.; Allen, N.; Dowell, P.; Teske, K.; Martin, C.; Martin, J. *TCCON Data from Lamont, Oklahoma, USA, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.lamont01.R0/1149159.
45. Morino, I.; Matsuzaki, T.; Ikegami, H.; Shishime, A. *TCCON Data from Tsukuba, Ibaraki, Japan, 125 HR, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.tsukuba02.R0/1149301.
46. Iraci, L.; Podolske, J.; Hillyard, P.; Roehl, C.; Wennberg, P.O.; Blavier, J.F.; Landeros, J.; Allen, N.; Wunch, D.; Zavaleta, J.; et al. *TCCON Data from Armstrong Flight Research Center, Edwards, CA, USA, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.edwards01.R0/1149289.
47. Wennberg, P.O.; Roehl, C.; Blavier, J.F.; Wunch, D.; Landeros, J.; Allen, N. *TCCON Data from Jet Propulsion Laboratory, Pasadena, California, USA, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.jpl02.R0/1149297.
48. Wennberg, P.O.; Wunch, D.; Roehl, C.; Blavier, J.F.; Toon, G.C.; Allen, N. *TCCON Data from California Institute of Technology, Pasadena, California, USA, Release GGG2014R1*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.pasadena01.R1/1182415.
49. Shiomi, K.; Kawakami, S.; Ohyama, H.; Arai, K.; Okumura, H.; Taura, C.; Fukamachi, T.; Sakashita, M. *TCCON Data from Saga, Japan, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.saga01.R0/1149283.
50. Griffith, D.W.T.; Deutscher, N.; Velazco, V.A.; Wennberg, P.O.; Yavin, Y.; Keppel-Aleks, G.; Washenfelder, R.; Toon, G.C.; Blavier, J.F.; Murphy, C.; et al. *TCCON Data from Darwin, Australia, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.darwin01.R0/1149290.
51. De Mazière, M.; Desmet, F.; Hermans, C.; Scolas, F.; Kumps, N.; Metzger, J.M.; Duflot, V.; Cammas, J.P. *TCCON Data from Reunion Island (La Réunion), France, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.reunion01.R0/1149288.
52. Griffith, D.W.T.; Velazco, V.A.; Deutscher, N.; Murphy, C.; Jones, N.; Wilson, S.; Macatangay, R.; Kettlewell, G.; Buchholz, R.R.; Riggensbach, M. *TCCON Data from Wollongong, Australia, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.wollongong01.R0/1149291.
53. Sherlock, V.; Connor, B.; Robinson, J.; Shiona, H.; Smale, D.; Pollard, D. *TCCON Data from Lauder, New Zealand, 125 HR, Release GGG2014R0*; Carbon Dioxide Information Analysis Center; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2014; doi:10.14291/tcon.ggg2014.lauder02.R0/1149298.
54. The TCCON Sites: Dryden/Armstrong/Edwards. Available online: <http://tcon-wiki.caltech.edu/Sites/Dryden> (accessed on 10 May 2016).

55. Toon, G.C.; Blavier, J.F.; Sen, B.; Salawitch, R.J.; Osterman, G.B.; Notholt, J.; Rex, M.; McElroy, C.T.; Russell, J.M., III. Ground-based observations of Arctic ozone loss during spring and summer 1997. *J. Geophys. Res.* **1999**, *104*, 26497–26510.
56. Wunch, D.; Toon, G.C.; Sherlock, V.; Deutscher, N.M.; Liu, X.; Feist, D.G.; Wennberg, P.O. *The Total Carbon Column Observing Network's GGG2014 Data Version, 2015*; Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2015. Available online: <http://dx.doi.org/10.14291/tcon.ggg2014.documentation.R0/1221662> (accessed on 10 May 2016).
57. Messerschmidt, J.; Geibel, M.C.; Blumenstock, T.; Chen, H.; Deutscher, N.M.; Engel, A.; Feist, D.G.; Gerbig, C.; Gisi, M.; Hase, F.; *et al.* Calibration of TCCON column-averaged CO<sub>2</sub>: The first aircraft campaign over European TCCON sites. *Atmos. Chem. Phys.* **2011**, *11*, 10765–10777.
58. Geibel, M.C.; Messerschmidt, J.; Gerbig, C.; Blumenstock, T.; Chen, H.; Hase, F.; Kolle, O.; Lavrič, J.V.; Notholt, J.; Palm, M.; *et al.* Calibration of column-averaged CH<sub>4</sub> over European TCCON FTS sites with airborne in-situ measurements. *Atmos. Chem. Phys.* **2012**, *12*, 8763–8775.
59. Yoshida, Y.; Kikuchi, N.; Inoue, M.; Morino, I.; Uchino, O.; Yokota, T. Recent Topics about the GOSAT TANSO-FTS SWIR L2 Retrievals. *AGU Fall Meet. Abstr.* **2014**, *1*, 3154.
60. The TCCON Data Description: Site-Specific Notes. Available online: [https://tcon-wiki.caltech.edu/Network\\_Policy/Data\\_Use\\_Policy/Data\\_Description#SITE-SPECIFIC\\_NOTES](https://tcon-wiki.caltech.edu/Network_Policy/Data_Use_Policy/Data_Description#SITE-SPECIFIC_NOTES) (accessed on 10 May 2016).
61. Rodgers, C.D.; Connor, B.J. Intercomparison of remote sounding instruments. *J. Geophys. Res. Atmos.* **2003**, *108*, 4116–4129.
62. Wunch, D.; Wennberg, P.O.; Toon, G.C.; Connor, B.J.; Fisher, B.; Osterman, G.B.; Frankenberg, C.; Mandrake, L.; O'Dell, C.; Ahonen, P.; *et al.* A method for evaluating bias in global measurements of CO<sub>2</sub> total columns from space. *Atmos. Chem. Phys.* **2011**, *11*, 12317–12337.
63. Du Piesanie, A.; Piders, A.J.M.; Aben, I.; Schrijver, H.; Wang, P.; Noël, S. Validation of two independent retrievals of SCIAMACHY water vapour columns using radiosonde data. *Atmos. Meas. Tech.* **2013**, *6*, 2925–2940.
64. Wagner, T.; Andreae, M.O.; Beirle, S.; Dörner, S.; Mies, K.; Shaiganfar, R. MAX-DOAS observations of the total atmospheric water vapour column and comparison with independent observations. *Atmos. Meas. Tech.* **2013**, *6*, 131–149.
65. Palm, M.; Melsheimer, C.; Noël, S.; Heise, S.; Notholt, J.; Burrows, J.; Schrems, O. Integrated water vapor above Ny Ålesund, Spitsbergen: A multi-sensor intercomparison. *Atmos. Chem. Phys.* **2010**, *10*, 1215–1226.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).