Evaluation of single-footprint AIRS CH_4 Profile Retrieval Uncertainties Using Aircraft Profile Measurements

Susan S. Kulawik[1], John R. Worden[2], Vivienne H. Payne[2], Dejian Fu[2], Steve C. Wofsy[3], Kathryn McKain[4,5], Colm Sweeney[4], Bruce C. Daube, Jr[6], Alan Lipton[7], Igor Polonsky[7], Yuguang He[7], Karen E. Cady-Pereira[7], Edward J. Dlugokencky[4], Daniel J. Jacob[6], Yi Yin[8]

[1]BAER Institute, 625 2nd Street, Suite 209, Petaluma, CA, USA
[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[3]School of Engineering and Applied Sciences and Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA
[4]National Oceanic and Atmospheric Administration, Global Monitoring Laboratory, Boulder, CO, USA
[5]University of Colorado, Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA
[6]Harvard University, Cambridge, MA 02138, USA
[7]Atmospheric and Environmental Research, Inc., Lexington, Massachusetts, USA
[8]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

Correspondence to: Susan S. Kulawik (susan.s.kulawik@nasa.gov)

Abstract.

We evaluate the uncertainties of methane optimal estimation retrievals from single footprint thermal infrared observations from the Atmospheric Infrared Sounder (AIRS). These retrievals are primarily sensitive to atmospheric methane in the mid-troposphere through the lower stratosphere (~2 to ~17 km). We compare to in situ observations made from aircraft during the Hiaper Pole to Pole Observations (HIPPO) and Atmospheric Tomography Mission (ATom) campaigns, and from the NOAA GML aircraft network, between the surface and 5-13 km, across a range of years, latitudes between 60 S to 80 N, and over land and ocean. After a global, pressure dependent bias correction, we find that the land and ocean have similar biases and that the reported observation error (combined measurement and interference errors) of ~27 ppb is consistent with the standard deviation between aircraft and individual AIRS observations. A single observation has measurement (noise related) uncertainty of ~17 ppb, a ~20 ppb uncertainty from radiative interferences (e.g. from water, temperature, etc.), and ~ 30 ppb due to "smoothing error", which is partially removed when making comparisons to in situ measurements or models in a way that account for this regularization. We estimate a 10 ppb validation uncertainty because the aircraft typically did not measure methane at altitudes where the AIRS measurements have some sensitivity, e.g. the stratosphere, and there is uncertainty in the truth that we validate against. Daily averaging only partly reduces the difference between aircraft and satellite observation, likely because of correlated errors introduced into the retrieval from temperature and water vapor. For example, averaging 9 observations only reduces the aircraft/model difference to ~17 ppb versus the expected ~10 ppb. Seasonal averages can reduce

this ~17 ppb uncertainty further to ~10 ppb, as determined through comparison with NOAA aircraft, likely because uncertainties related to radiative effects of temperature and water vapor are reduced when averaged over a season.

1 Introduction

Advances in remote sensing, global transport modeling, and an increasingly dense network of surface measurements have led to substantive advances in evaluating the components and error structure of the global methane budget and the processes controlling this budget. For example, Frankenberg et al. (2005, 2011) showed that total column methane estimates could be derived from near infrared (NIR) radiances at ~1.6 microns measured by the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY). Since then, methane retrievals have also been applied to NIR radiances from the Greenhouse Gases Observing Satellite (GOSAT) instrument (e.g. Parker et al., 2011; Schepers et al., 2012), launched in 2009, and the TROPOspheric Monitoring Instrument (TROPOMI, e.g. Hu et al., 2018), launched in 2017. These data have sufficient accuracy to map regional surface methane enhancements (e.g. Kort et al., 2014; Wecht et al., 2014) and point source anomalies (Varon et al., 2019; Pandey et al., 2019). Estimates of the free-tropospheric methane concentrations from spaceborne measurements in the thermal infrared (TIR) at ~8 microns were demonstrated using radiances from the Aura Tropospheric Emission Spectrometer (TES, Worden et al., 2012; 2013b), the Atmospheric Infrared Sounder (AIRS, e.g. Xiong et al., 2013), the Infrared Atmospheric Sounding Interferometers (IASI, e.g. Ravazi et al., 2009; De Wachter et al., 2017; Siddans et al., 2017), the Cross-Track Infrared Sounders (CrIS, e.g. Smith and Barnet, 2019) and TIR GOSAT measurements (de Lange and Landgraf, 2018). TIR methane measurements have been used to evaluate the role of fires (e.g. Worden et al., 2013b; 2017a), Asian emissions and stratospheric intrusions (e.g. Xiong et al., 2009; 2013) on the global methane budget.

The goal of this paper is to evaluate the uncertainties of new methane retrievals from AIRS single footprint, original (non-cloud-cleared) radiances using aircraft measurements from the HIAPER Pole-to-Pole Observations (HIPPO) and Atmospheric Tomography Mission (ATom) campaigns and National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Laboratory (GML) aircraft network, taken between 2006 and 2017. Evaluation of these uncertainties are needed to determine if AIRS methane data can characterize and improve errors in global chemistry transport models. For example, a recent paper by Zhang et al. (2018) combined synthetic CrIS and TROPOMI methane retrievals and a global inversion system to show that it would be possible to infer the north-south gradient of OH, the primary methane sink, to within 10%, and temporal variations of OH concentrations. However, knowing the accuracy of the methane data is important for inferring the uncertainty in the spatio-temporal variability of OH. Over decadal time scales, OH can vary by 3-5% (e.g. Turner et al., 2018a, 2018b, 2019; Rigby et al., 2017). Therefore, to be useful for understanding OH, monthly or seasonally averaged AIRS data should have an uncertainty that is less than 3-5% (55-99 ppb).

In this paper we present an evaluation of methane retrievals derived from AIRS single footprint radiances. We follow an optimal estimation approach (Rodgers, 2000), based on the heritage of the Aura Tropospheric Emission Spectrometer (TES) algorithm (Bowman et al., 2006), now called the MUlti-SpEctra, MUlti-SpEcies, MUlti-Sensors (MUSES) algorithm (Worden et al., 2006, 2013b; Fu et al., 2013, 2016, 2018, 2019). The MUSES algorithm uses radiances from one or multiple instruments to quantify and characterize geophysical parameters derivable from those radiances. The optimal estimation method provides the vertical sensitivity (i.e., the averaging kernel matrix) and estimates of the uncertainties due to noise and to radiative interferences such as temperature, N2O, and water vapor. We compare AIRS retrievals with corresponding aircraft data over a range of latitudes and longitudes in order to evaluate the calculated uncertainties over ocean and land. Much of the description of the forward model and retrieval approach is provided in Worden et al. (2012, 2019). We therefore refer the reader to these papers for a more in-depth description of the retrieval approach and only summarize aspects here that are relevant for comparing the AIRS methane retrievals to aircraft data.

2 Datasets used in this paper

The quantities of interest that we validate in this paper are a) the AIRS CH_4 dry volume mixing ratio (VMR) at particular pressure values between 750 hPa and 300 hPa, or b) the AIRS CH_4 dry VMR partial column XCH_4 covering the same pressure range that is measured by the aircraft. We use aircraft profiles which span the pressure range that contains at least 0.20 degrees of freedom for the AIRS CH_4 partial column. The retrieval estimates AIRS CH_4 dry volume mixing ratio (VMR) profile. When a "partial column quantity" is validated, the retrieved CH_4 profile is post-processed into partial column XCH_4 VMR relative to dry air, with methodology from Connor et al. (2008) and Kulawik et al. (2017), where the VMR's at the pressure levels are weighted according to a pressure weighting function, resulting in a partial column volume mixing ratio (VMR).

2.1 Description of AIRS

The AIRS instrument is a nadir-viewing, scanning infrared spectrometer (Aumann et al., 2003; Pagano et al., 2003; Irion et al., 2018; DeSouza-Machado et al., 2018) that is onboard the NASA Aqua satellite and was launched in 2002. AIRS measures the thermal radiance between approximately 3-12 microns with a resolving power of approximately 1200. For the 8 micron spectral range used for the HDO/H_2O/CH_4 retrievals, the spectral resolution is ~1 wavenumber (cm^{-1}), with a gridding of ~0.5 cm^{-1}, and the signal-to-noise (SNR) ranges from ~400 to ~1000 over the 8 micron region for a typical tropical scene. A single footprint has a diameter of ~15 km in the nadir; given the ~1250 km swath, the AIRS instrument can measure nearly the whole globe in a single day. The Aqua satellite is part of the “A-Train” that consists of multiple satellites and instruments, including TES, in a sun-synchronous orbit at 705 km with an approximately 1:30 am and 1:30 pm equator crossing-time. In this paper, we use only daytime data to match the validation observations.

2.2 Overview of Aircraft Data

Measurements from the HIPPO (Wofsy et al., 2012) and ATom (Wofsy et al., 2018) aircraft campaigns provide excellent datasets for satellite validation, due to their wide latitudinal coverage, the large vertical extent of the profiles (up to 9-12 km), and the availability of campaigns over a wide range of months. Each of the five HIPPO campaigns flew south, then north over a period of weeks, often using a different path for the northern and southern legs, with campaign dates in 2009 - 2011. Atmospheric methane concentrations were measured with a quantum cascade laser spectrometer (QCLS) at 1 Hz frequency with accuracy of 1.0 ppb and precision of 0.5 ppb (Santoni et al., 2014). HIPPO methane data are reported on the WMO X2004 scale and have been used in several other studies to evaluate satellite retrievals of methane (e.g. Alvarado et al., 2015; Wecht et al., 2012; Crevoisier et al., 2013). Comparisons with NOAA flask data showed a mean positive bias of 0.85 ppb for the QCLS during the HIPPO campaigns, which is consistent with the estimated QCLS accuracy of 1.0 ppb (Santoni et al., 2014; Kort et al., 2011). We used 396 QCLS CH_4 profiles from the HIPPO campaigns. Using coincidence criteria of ±9 hours, ±50 km, 22,271 AIRS observations were processed, of which 5537 passed quality flags. The latitude of the matches ranges from 57S to 81N.

We compare AIRS to observations from the ATom aircraft campaigns 1-4 (Wofsy et al., 2018). This comparison provides validation ~7 years after HIPPO, between 2016 and 2018. Similar to HIPPO, these observations include observations in the Pacific Ocean, but ATom also includes observations in the Atlantic (as seen in Table A.1 and Fig. 1). ATom methane data are reported on the WMO X2004A scale. We used 289 profiles from the ATom campaigns from the NOAA Picarro instrument (Karion et al., 2013). For more information on the instrument, see https://espo.nasa.gov/sites/default/files/archive_docs/NOAA-Picarro_ATom1234_readme.pdf. Using coincidence criteria of ±9 hours, ±50 km, 21,225 AIRS observations were processed, of which 4913 passed quality flags. The latitude of the matches ranges from 65S to 65N.

The NOAA GML aircraft network observations (Cooperative Global Atmospheric Data Integration Project, 2019) are taken twice per month at fixed sites primarily in North America, and also Rarotonga (RTA) at 21S (Sweeney et al., 2015). NOAA aircraft network methane data are reported on the WMO X2004A scale. Although HIPPO data are not reported on the same scale as ATom and NOAA aircraft network data, differences in values of calibration tanks used for HIPPO (Santoni et al., 2014) on the two different scales are < 1 ppb. We match AIRS and NOAA aircraft observations between 2006 and 2017, with coincidence criteria of 50 km and 9 hours, finding ~43,000 matches, and 18,000 good quality matches following the retrieval, to 719 aircraft measurements, at sites ACG (67.7N, 164.6E, 401 matches), ESP (49.4N, 126.5E, 2743 matches), NHA (43.0N, 70.6E, 2682 matches), THD (41.1N, 124.2E, 1551 matches), CMA (38.8N, 74.3E, 3269 matches), TGC (27.7N, 96.9E, 1944 matches), and RTA (21.2S, 159.8E, 810 matches).

Figure 1 shows the locations of all the aircraft data used for the comparisons described in this paper. Most of the ocean measurements are from the HIPPO and ATom campaigns that spans a range of latitudes, whereas most of the land measurements are taken over North America.

3 MUSES-AIRS Optimal Estimation of CH_4 from single-footprint, original (non-cloud-cleared) AIRS radiances

Worden et al. (2012, 2019) describe in detail the forward model and retrieval approach used for estimating methane from TES and AIRS radiances. The radiative transfer forward model used for this work is the Optimal Spectral Sampling (OSS) fast radiative transfer model (RTM) (Moncet et al., 2005, 2008, 2015). In particular, radiances from the thermal infrared bands at 8 and 12 microns are used to quantify profiles of atmospheric concentrations of CH_4, HDO, H_2O, N_2O, as well as temperature, emissivity, and cloud properties. The atmospheric parameters are retrieved as vertical profiles. Since we use optimal estimation, or OE, (e.g. Rodgers, 2000; Bowman et al., 2006) to estimate these quantities we can characterize the vertical resolution and uncertainties of these retrievals, which allows us to compare them to models and independent data sets, while accounting for the regularization used for the retrieval. We follow the OE approach for the Aura TES instrument (e.g. Bowman et al., 2006; Worden et al., 2006, 2012) but with some differences. First, methane retrievals using the TES radiances are obtained using only the 8 micron band, because of slight calibration differences between the detectors that measure the 12 and 8 micron bands (e.g. Shephard et al., 2008; Connor et al., 2011). For the AIRS retrievals, we use both the 8 and 12 micron bands in order to better constrain temperature in the troposphere and stratosphere. Secondly, the TES based retrieval uses the ratio of a jointly-retrieved N2O profile to the CH_4 profile in order to help correct biases related to temperature variations in the (UTLS) upper-troposphere lower-stratosphere (Worden et al., 2012). However, the N_2O correction is not used for the AIRS retrievals because we can jointly estimate temperature in the UTLS region using the 12 micron band. We use similar quality flags as the TES retrievals such as checks on the radiance residual, residual signal, and cloud optical depth as discussed in Kulawik et al. (2006a, 2006b), except that we screen out cloudy and low-sensitivity cases, resulting in about 1/4 of the data passing screening. Quality flags are discussed in more detail in the Aura-TES user's guide (pp 27-30, Herman et al., 2018). The specific flags used for AIRS CH_4 are as follows, which were set by minimizing the standard deviation of small clusters of retrievals and to standardize the sensitivity:

Good quality and sensitivity flagging for AIRS CH_4:

- Radiance residual rms < 1.5. This parameter is the mean difference between the observed and fit radiance normalized by the radiance measurement error.
- The absolute value of Radiance residual mean < 0.15. This parameter screens off the mean difference of the radiance residual.
- The absolute value of KdotdL < 0.23. This parameter is the mean difference of the dot product of the Jacobians and the radiance residual normalized by the radiance measurement error, and smaller values indicate that there is little remaining information in the signal.
- The surface temperature minus the near-surface atmospheric temperature value < 30K. This ensures that the thermal gradient is less than 30K between the surface and lowest atmospheric temperature.
- Cloudtop pressure > 90 hPa. This ensures that the retrieved cloudtop pressure is in or near the Troposphere.

- Cloud optical depth < 0.3. This ensures that the cloud is not opaque and there is fairly uniform sensitivity so that the bias correction is fairly consistent. The bias versus cloud optical depth is shown in the supplement.
- Cloud variability versus wavenumber < 1.5 * cloud OD. This ensures that the cloud optical depth does not vary too much over the retrieval window.
- Degrees of freedom > 1.1, defined following Eq. 2. This ensures a minimum sensitivity so that the bias correction is fairly consistent.
- Tropospheric degrees of freedom > 0.7, defined following Eq. 2. This ensures a consistent Tropospheric sensitivity, so that the bias correction is fairly consistent.
- Stratospheric degrees of freedom < 0.5, defined following Eq. 2. This ensures that there is a consistent stratospheric sensitivity, so that the bias correction is fairly consistent.
- Predicted error on the column above 750 hPa < 53 ppb. The predicted error is the total error from the linear estimate, Eq. 7b, and is included in the output product. This ensures that the predicted error, which is correlated to the actual error, is not too large.

3.1 Retrieval Error Characteristics

Detailed descriptions of the use of optimal estimation (OE) to infer trace gas profiles from remote sensing radiance measurements retrieval is included in numerous publications (e.g. Rodgers, 2000; Worden et al., 2006; Bowman et al., 2006). However, we present a partial description here as it is relevant for comparing the AIRS methane retrievals and aircraft profile measurements. As discussed in Rodgers (2000), the estimate for a trace gas profile inferred (or inverted) from a radiance spectrum is described by the following linear equation:

$$\hat{x} = x_a + A(x - x_a) + G K^b b_{error} + Gn \quad \textbf{(1)}$$

where $\hat{x}$ is the estimate of Log(VMR), x_a is the log of the a priori concentration profile used to regularize the inversion, G is the gain matrix, $\boldsymbol{b_{error}}$ represents errors in systematic parameters, with K^b the sensitivity of the radiance to changes in **b**. We split x into [x,y], where x is the quantity of interest, the methane profile, and y are the jointly estimated quantities (such as temperature, water vapor, clouds, and surface properties), which results in the cross-state error (Worden et al., 2004; Connor et al., 2008).

$$\hat{x} = x_a + A_{xx}(x - x_a) + G K^b b_{error} + A_{xy}(y - y_a) + Gn \quad \textbf{(2)}$$

For the AIRS (and TES) OE methane retrievals, $\mathbf{x}_a$ comes from the MOZART atmosphere chemistry model (e.g. Brasseur *et al*., 1998). The vector **x** is the "true state", or in this case the (log) concentration profile. The matrix **A** is the averaging kernel matrix or $A = \frac{\partial \hat{x}}{\partial x}$ and describes the vertical sensitivity of the measurement. A_{xx} describes the dependence of $\hat{x}$ on the true state x, and A_{xy} describes the dependence of $\hat{x}$ on the true state y, which is non-zero because of correlations in the Jacobians, **K**, for x and y. The matrix **G** relates changes in the radiance (**L**) to perturbations in **x**, $G = \frac{\partial x}{\partial L}$. The vector **n** is the noise vector, the matrix **K** is the sensitivity of the radiance to changes in (log) concentration $K = \frac{\partial L}{\partial x} = \frac{\partial L}{\partial log(VMR)}$, and the set of vectors $\mathbf{b}_i$ represent interference errors not estimated from the observed radiances. The true state, noise vector, and interference errors as described here are the "true" values and are therefore not actually known but are represented in this form so that we can calculate how their uncertainties affect the estimate $\hat{x}$. An example averaging kernel matrix is shown in Figure 2 and shows that AIRS based estimates of methane are most sensitive to methane in the free-troposphere and lower-stratosphere as demonstrated previously for AIRS and other TIR based estimates of tropospheric methane (e.g. Xiong *et al.*, 2016; de Lange and Landgraf, 2018).

The degrees of freedom, DOFs, describing the sensitivity of $\hat{x}$ to the true state, and is equal to the trace of A_{xx}. The degrees of freedom in the troposphere is equal to the trace of the averaging kernel corresponding to the troposphere, and the degrees of freedom in the stratosphere is equal to the trace of the averaging kernel corresponding to the stratosphere. The troposphere is defined using the tropopause height parameter from version 5 of the NASA Global Modeling and Assimilation Office (GMAO) Goddard Earth Observing System (GEOS-5) model (Molad et al., 2012).

Finally, we look at the quantity of interest, $\hat{x} =$ **hx.** The vector **h** combines all the necessary operations that maps the (log) concentration profiles to whatever quantity is needed such as selecting one particular pressure level (e.g. **h**= [0,0,0,1,0,0,0, …], selecting a column average, (**h** = pressure weighting function) – see Connor et al., 2008 or Kulawik et al., 2017) or selecting the VMR mean (e.g. **h[:]**=1/m, where m is the number of pressure levels to average).

$$\hat{x} = h\hat{x} \tag{3a}$$

$$\hat{x} = hx_a + hA_{xx}(x - x_a) + hG_x K^b b_{error} + hA_{xy}(y - y_a) + hG_x n \tag{3b}$$

In Eq. 3a, the vector $\hat{x}$ (denoted in bold) is converted to the scalar of interest, $\hat{x}$ (non-bold, italic). In our validation comparisons, **h** is used to select 1) a specific pressure level that is measured by the aircraft, 2) the partial column XCH_4 VMR within the pressure levels measured by the aircraft, and 3) the partial column XCH_4 between 750 hPa and the top of the atmosphere.

3.2 Approach for Comparing AIRS measurements to aircraft profiles

A challenge in comparing the satellite-based AIRS measurements to aircraft data is that the aircraft will typically measure only a section of the atmosphere (e.g. the troposphere), whereas the AIRS measurements are sensitive, to varying degrees (see Fig. 2), to the entire atmosphere. To account for these differences, we divide the atmosphere into two parts $\mathbf{x} = [\mathbf{x}_c, \mathbf{x}_s]$: where $\mathbf{x}_c$ is the part measured by the aircraft (denoted **c** for airCraft), and $\mathbf{x}_s$ is the part not measured by the aircraft (denoted **s** for Stratospheric):

$$\hat{x}_c = h_c x_a + h_c A_{cc}(x_c - x_a^c) + hG_c K^b b_{error} + h_c A_{cy}(y - y_a) + h_c A_{cs}(x_s - x_a^s) A_{cs}(x_s - x_a^s) + h_c G_c n \quad \textbf{(4)}$$

where the term $\mathbf{A_{cs}}$ is the cross-term in the averaging kernel that describes the partial derivatives of the aircraft-measured levels (e.g. the troposphere) to the un-measured levels (e.g. the stratosphere). Equation 4 describes how the AIRS measurement $\hat{x}_c$ responds to the true state $[\mathbf{x}_c, \mathbf{x}_s]$. So, if for example, the aircraft measured indices [0:9], and did not measure pressure levels [10:*], then $A_{cc} = A[0{:}9, 0{:}9]$ and $A_{cs} = A[0{:}9, 10{:}65]$, where A is the full averaging kernel.

We compare our AIRS observation, $\hat{x}_c$ in Eq. 4, to our aircraft observation, $x_{aircraft}$. To compare directly to the aircraft observation (without accounting for AIRS sensitivity) we would compare to $\hat{x}_{aircraft}^c = h_c x_{aircraft}$. The expected total error includes the smoothing error, which the covariance of the $h_c A_{cc}(x_c - x_a^c)$ (Rodgers, 2000), where the covariance of $(x_c - x_a^c)$ is the a priori covariance, S_a^{xx}. The smoothing error is:

Smoothing error = $h_c A_{cc} S_a^{xx} A_{cc}^T h_c^T$ **(5)**

We estimate the smoothing error for the partial column XCH_4 VMR within the pressure levels measured by the aircraft to be 30 ppb, using Eq. 5. This estimate strongly depends on S_a^{xx}, the a priori covariance, which is the same as in Worden et al. (2012); briefly 5% diagonal variability with correlations in pressure set from the MOZART model. In Equation 6a, we apply the AIRS Averaging kernel to the aircraft measurement to fully account for the AIRS sensitivity:

$$\hat{x}_{aircraft}^c = h_c x_a + h_c A_{cc}\left(x_{aircraft}^c - x_a^c\right) + h_c A_{cs}\left(x_{aircraft}^s - x_a^s\right) \quad \textbf{(6a)}$$

$$\hat{x}_{aircraft}^c = h_c x_a + h_c A_{cc}\left(x_{aircraft}^c - x_a^c\right) \quad \textbf{(6b)}$$

One issue is that we do not actually have aircraft observations in the "s" part of the atmosphere, $x_{aircraft}^s$, which is used in the second term of Eq. 6a. We have aircraft observations in the "c" part of the atmosphere only, so we apply the Averaging Kernel to this part of the atmosphere only. Equation 6a accounts for all of the AIRS smoothing error, whereas Equation 6b (the equation used in this work, other than Section 3.3) only accounts for the smoothing error from the part of the atmosphere measured by the aircraft profile. The difference from Eqs. 6a and 6b is discussed in Section 3.3.

Equation 7a is the predicted bias between $\hat{x}_c$ (the measured AIRS value) and $\hat{x}^{c}_{aircraft}$ (the aircraft value with the AIRS Averaging kernel applied), and is the expected difference of Eqs. 4 and 6b. Equation 7b is the covariance of Eq. 7a, and estimates the predicted error:

$$E(\hat{x}_c - \hat{x}^{c}_{aircraft}) = h_c G_c K^b E(b_{error}) + h_c A_{cy} E(\hat{y} - y_a) + h_c A_{cs} E(\hat{x}_s - x_s) + h_c G_c E(n_{error}) \quad \textbf{(7a)}$$

$$E|| \ (\hat{x}_c - \hat{x}^{c}_{aircraft})|| = h_c (G_c K^b S_a^{bb} K_b^T G_c^T + A_{cy} S_a^{yy} A_{cy}^T + A_{cs} S_a^{ss} A_{cs}^T + S_m^{cc}) h_c^T \quad \textbf{(7b)}$$

Eq. 7a represents the propagation of mean biases from: (1) non-retrieved parameters and assumptions, e.g. spectroscopy (b), (2) jointly retrieved parameters, e.g. temperature, (y), (3) "stratospheric", describing the impact of the part of the atmosphere not covered by the aircraft on the measured part (x_s), or (4) measurement errors (n) into biases of $\hat{x}_c$. The mean bias from 7a is difficult to characterize theoretically and is characterized during validation, and assumed to be primarily from the first term (e.g. spectroscopy). Equation 7b is the covariance of the terms in 7a, where e.g. the covariance of b_{error} is S_a^{bb}. Equations 7b represent the "observation covariance". The square root of 7b is the predicted observation error. Although Eq. 7b has overall zero bias, it can produce regional and temporal biases, e.g. as seen in Connor et al. (2016), where these biases approach zero over long enough spatial or temporal scales. The error covariances all represent fractional errors, in log(VMR). Because the retrieved quantity log(VMR), the error in ppb is approximately the fractional error times the methane value in ppb.

For the purpose of evaluating the AIRS methane measurement uncertainties and comparing the AIRS methane to aircraft in situ measurements we refer to the four terms on the right side of Eq. 7b as:

1) S_b^{cc} is the systematic error due to terms that are not accounted for in the retrieval state vector, such as spectroscopy and calibration; these terms are estimated by comparisons with the aircraft data. A pressure-dependent bias correction, described in Section 3.4, of approximately -60 ppb is used to correct this systematic bias.
2) $A_{cy} S_a^{yy} A_{cy}^T$, the “cross-state”, which is included in the MUSES-AIRS methane estimate product files, and is the propagation of temperature, water vapor, and cloud errors into AIRS. The errors in the retrieved temperature and water vapor at nearby location are correlated over short spatio-temporal scales, as described in Section 4, and so this error does not reduce with averaging nearby observations. However, monthly or seasonal averages reduces cross-state error, because systematic errors from temperature / water / cloud can be assumed to vary pseudo-randomly over larger time scales.

3) $A_{cs}S_a^{ss}A_{cs}^T$ is the "validation uncertainty" due to knowledge uncertainty of the stratosphere although this may also contain other levels that are also not measured by the aircraft. This is the smoothing error which cannot be removed from the comparisons because the aircraft does not make measurements at the "**s**" ("stratospheric") levels. We estimate this validation uncertainty as ~10 ppb (estimated in Section 3.3). This estimate depends on the accuracy of the model used to extend the aircraft profile during the validation process and was estimated for the model that we used in validation.

4) S_m^{cc}, the "measurement" error, which is included in the AIRS methane estimate product files. The measurement error is random and is expected to reduce as the inverse square root of the number of observations averaged. We estimate this error as ~18 ppb (using the last term of Eq. 7b and shown in Fig. 3) and find it to be a random error that reduces with averaging.

Figure 3 shows the predicted errors for the AIRS partial column XCH_4 VMR within the pressure levels measured by the aircraft. The measurement error (light green) is 18 ppb (from the last term of Eq. 7b) , and the total error for a single observation (including smoothing error) is 41 ppb. A component of the total error, the cross-state error, is estimated as 21 ppb (from Eq. 7b).

3.3 Estimating validation uncertainty due to aircraft not measuring the stratosphere

A typical aircraft profile will only measure part of the troposphere and rarely measure into the stratosphere. However, the AIRS methane profile measurements are sensitive to methane variations over the whole atmosphere as shown by the averaging kernel matrix in Figure 2. Similarly, the true state in the troposphere influences retrieved values in the stratosphere. Options for dealing with this are a) extending the true with the AIRS prior or b) extending the true with a model profile value.

This section estimates this uncertainty by calculating the difference of $x^{c}_{aircraft}$ for Eq. 6a minus Eq. 6b when extending the aircraft using two different "true" profiles taken from two different global atmospheric chemistry models, the Laboratoire de Météorologie Dynamique (LMDz) model (e.g. Folberth et al., 2006) model and the Goddard Earth Observing System (GEOS-Chem) model (e.g. Maasakkers et al., 2019). So, if the model value equaled the AIRS prior in the stratosphere, this difference would be zero. The differences for $x^{c}_{aircraft}$ from LMDz model and GEOS-Chem are shown in Figure 4 for all HIPPO ocean and land data; these differences show that model/model differences in the stratosphere can contribute significantly to the differences between AIRS and aircraft validation.

These differences provide an estimate for how knowledge error in the stratosphere projects to uncertainties in our methane retrievals. For example, this uncertainty varies with latitude, similar to the residual bias between the AIRS estimate and aircraft (next section). Furthermore, the variability over small latitudinal ranges of 10 degrees or less suggests that the random part of the stratospheric error is smaller than this latitudinal variability. Our estimate for this error is the average of these two errors, 10 ppb, and places an upper bound on the ability to validate AIRS CH_4. Our estimate for this error agrees with the 10 ppb estimate for the impact of stratospheric uncertainty on column estimates from aircraft profiles (Wunch et al., 2010). Appendix A shows further analysis of mean differences of AIRS minus aircraft for different profile extension choices. The bias varies by ~ 5 ppb for different profile extension choices when comparing at 700 hPa, ~10 ppb for different profile extension choices when comparing at 500 hPa, and ~11 ppb for different profile extension choices when comparing the column above 750 hPa.

The methane profile has a strong variable negative vertical gradient in the stratosphere. Models in general have a positive bias in the extratropical stratosphere (Patra et al., 2011). In GEOS-Chem 4x5, the column bias is shown in Figure 2c of Turner et al. (2015) and further discussed in Maasakkers (2019), which resolves the bias to the stratosphere, and model stratospheric accuracy is an active research area (Ostler et al., 2016; Maasakkers et al., 2019).

3.4 Bias Correction

AIRS CH_4 shows a persistent high bias of 25 to 90 ppb versus aircraft observations in Fig. 5. Previous studies using remotely sensed measurements suggest that a bias correction to the AIRS methane profile measurement must account for the vertical sensitivity (e.g. Worden et al., 2011). For example, in the limit where the AIRS measurement is perfectly sensitive to the vertical distribution of methane, the bias correction could be a simple scaling factor. However, in the limit where the AIRS measurement is completely insensitive (e.g. DOFS = 0.0) then the bias correction is zero. We therefore use the bias correction approach described in Worden et al. (2011), where a bias profile (which varies by pressure) is passed through the averaging kernel to account for the AIRS sensitivity, as seen in Eq. 8. The form of the bias profile, δ_{bias} is set in Eq. 9.

We use HIPPO-4 observations to set a bias correction which we then evaluate with the other HIPPO campaigns and NOAA aircraft network data. HIPPO-4 was selected as it covers a wide range of latitudes and so that the bias correction can be set and tested with two independent datasets. To set the bias, we use Eq. 6b to estimate the aircraft observation as seen by AIRS, then compare this to AIRS observations. The result (by pressure level) is shown in Table 1. Then a bias was applied to AIRS using Eq. 8, with the bias term δ_{bias} in the form of Eq. 9.

$$\hat{x}_{corrected} = \hat{x}_{orig} + A\delta_{bias} \qquad \textbf{(8)}$$

Where $\hat{x}$ = ln(VMR), because the retrieved quantity is ln(VMR), δ_{bias} is a vector, and A is the averaging kernel matrix for $\hat{x}$ = ln(VMR). We fit a single bias function for all AIRS measurements by minimizing the difference between the AIRS and HIPPO-4 with δ_{bias} constrained to have a slope with pressure, and two pressure domains. We specify that δ_{bias} cannot jump more than 0.05 (5%) between the two domains.

$$\delta_{bias} = c + dP \text{ (P > P}_o\text{)}$$
$$\delta_{bias} = e + fP \text{ (P < P}_o\text{)} \quad \textbf{(9)}$$

where P is pressure in hPa. The optimized bias correction parameters were: c = 0.0; d = - 6.1e-5; P_0=400 hPa; e=-0.09; f=0.00018. This bias correction results are shown for HIPPO-4; HIPPO-1,2,3,5; and NOAA observations in Table 1. The remainder of the paper, unless specified, uses data bias-corrected by Eqs. 8 and 9.

Figure 6 shows the effect of bias correction on the average of all HIPPO 1,2,3,5 AIRS profiles. The bias correction improves the mean AIRS / aircraft difference and improves the pressure-dependent skew in the bias (Table 1). The HIPPO data is shown before and after the AIRS averaging kernel is applied (using Eq. 6b), which has the effect of bringing the HIPPO observations towards the AIRS prior. This is to match the imperfect sensitivity of satellite-based observations, which are similarly influenced by the prior.

4 Evaluation against aircraft data versus latitude

4.1 Comparison of aircraft observations with and without bias correction

Figure 5 shows a comparison between all AIRS measurements within 50 km and 9h of an aircraft measurement and the aircraft measurement. The quantity compared is the partial column XCH_4 VMR within the pressure levels measured from the aircraft. There is a mean bias of 57 ppb overall, ~52 ppb for ocean and ~76 ppb for the land. The RMS difference is ~26 ppb. Furthermore, there appears to be latitudinal variations in the bias. For example, the mean difference between the AIRS and aircraft over the ocean for latitudes less than 20 S is ~74 ppb and for latitudes between 20 S and 20 N this bias is ~ 56 ppb.

Figure 7 shows the same comparisons as Fig. 5 after bias correction (described in Section 3.4). The mean bias is 1 ppb, and the RMS difference is 24 ppb. The overall land bias is 12 ppb, and the overall ocean bias is -2 ppb. The bias calculated in Fig. 7 is weighs every point equally. Table A.1 shows a slightly different result for these biases, where the bias is calculated by campaign, then averaged over all campaigns. In Table A.1 the partial column XCH_4 VMR within the pressure levels measured by the aircraft has a bias of 16 ppb for land, and -2 ppb for ocean. Note that the HIPPO land observations are primarily in Australia, New Zealand, and North America, whereas the ocean comparisons are in the mid-Pacific, as seen in Fig. 1. We expect the RMS difference to be similar to the observation error, as the terms that make up the observation error are the primary source of variability in the observations (e.g. Worden et al., 2017b). The predicted observation error from Fig. 3, is 27 ppb, and is consistent with the RMS difference seen here, 23 ppb. However, knowledge of the stratosphere / validation uncertainty is potentially a large component of the latitudinal variability in the difference seen in the bottom panel of Fig. 7.

We also compare to NOAA aircraft network and ATom observations and find similar results as HIPPO. Figure 8, discussed in Section 4.2, shows ATom results, and Figure 9, discussed in Section 4.2, shows comparisons to a NOAA aircraft time series. The biases for different pressure ranges, campaigns, and surfaces is shown in Table A.1. Table A.3 shows the standard deviation of AIRS minus validation by pressure and surface type, for single observations, daily, and seasonal averages.

4.2 Errors in averaged AIRS data

Satellite data are typically averaged in order to improve the precision of a comparison between data and model. However, as shown in the previous figure, these data contain errors that vary with latitude. For example, knowledge error of the true profile in the stratosphere as well as errors in the jointly retrieved AIRS temperature and water vapor retrievals have both a random and a bias component, both of which vary with latitude. The bias component is approximately the same for all AIRS methane measurements taken on the same day within 50 km, as we do not expect large variations in temperature and water vapor errors over these scales, which we presume to be a driver of these correlated errors. To quantify the component of the accuracy that cannot be reduced by averaging, we compare averages of AIRS measurements to HIPPO and ATom measurements. We average over 1 day, the AIRS observations matching a single HIPPO or ATom measurement, within +-50 km and 9 hours of the measurement. We specify that there needs to be at least 9 AIRS observations for each comparison so that the systematic error, and not the precision (or measurement error), is the dominant term. These daily AIRS averages contain, on average, 20 AIRS observations. Figure 8 shows the predicted error, assuming that the error is random, which is calculated by dividing the single observation error (24 ppb RMS shown in Fig. 7) by the square root of the number of observations that are averaged. The mean predicted error for the averaged data, assuming random errors, is 6 ppb. The actual standard deviation between the averaged AIRS and HIPPO or ATom data is ~17 ppb, which is much larger and indicates that the errors within 1 day and 50 km are correlated. Note that same-colored adjacent points (i.e. adjacent observations from the same campaign) often show similar biases. Because this RMS difference is much larger than what is expected if the errors were purely random, this shows the presence of systematic errors, either in the AIRS data or in the validation uncertainty. We therefore report 17 ppb as the limiting error when averaging AIRS data within one-degree grids and 1 day for the purpose of comparing to models or other methane profiles.

On the other hand, averaging AIRS data seasonally can reduce the error further because geophysical errors from as temperature and water vapor vary over longer time scales. We demonstrate this aspect of the AIRS uncertainties by comparing averaged AIRS data to the NOAA aircraft methane profiles taken off the coast near Corpus Christi, Texas (27.7N, 96.9W, site TGC). We screen for at least 3 observations per day, less than the 9 observations/day used for HIPPO / ATom daily averages in order to get enough daily averages to explore how the errors reduce with monthly and seasonal averages, since the aircraft make 1-2 measurements per month. Figure 9 shows daily, monthly, 90-day, and seasonal averages of the partial column XCH_4 VMR within the pressure levels measured from the aircraft at TGC. The seasonal averages are created by converting all AIRS/aircraft matched pairs to 2012 by adding 5.4 ppb per year multiplied by (year minus 2012) to account for the mean annual growth rate.

The growth rate of 5.4 ppb/year is the mean increase during the AIRS record time period (2002-2019) estimated from the NOAA Global Monitoring Laboratory global surface measurements (https://esrl.noaa.gov/gmd/ccgg/trends_ch4/). Since we are converting matched pairs of aircraft and AIRS to 2012, the differences between these matched pairs is unaffected by the accuracy of the conversion to 2012.

4.2.1 Daily average errors at TGC

We look at daily averages versus aircraft data and find a similar result as that found with comparisons to ATom and HIPPO: daily averages have much larger errors than would be predicted if random errors are assumed. The standard deviation of (AIRS minus aircraft) at TGC is 24 ppb, the standard deviation for (daily AIRS average minus aircraft) is 11.5 ppb, as seen in Fig. 9a, and the predicted error for daily averages, assuming randomness in the error, is 6.0 ppb. Therefore, similarly to ATom and HIPPO, errors within 1 day and 50 km contain 11.5 ppb correlated error.

4.2.2 Monthly average errors at TGC

The NOAA aircraft measurements are usually taken about twice per month. The standard deviation of (monthly AIRS average minus aircraft) is 8.2 ppb (Figure 9b) for months containing more than 1 aircraft observation. This is compared to the daily error divided by the square root of the number of days averaged, 8.0 ppb. Therefore, errors for observations ~2 weeks apart are uncorrelated.

4.2.3 3-month average errors at TGC

We average over 3-month scales, where averages must have at least 3 days. The standard deviation of (3-month AIRS average minus aircraft) is 6.2 ppb. The predicted error, taking the 11.5 ppb daily error and dividing by the square root of the number of days averaged, is 6.0 ppb. Therefore, errors for 3-month averages are ~uncorrelated.

4.2.3 Seasonal cycle average errors at TGC

We average matched pairs within each month from any year. (AIRS minus aircraft) for these averages, have a standard deviation of 5.9 ppb, whereas the predicted error, from the daily average divided by the square root of number of observations, is 4.2 ppb.

4.2.4 Summary of average errors at TGC

To summarize, averaging AIRS observations within one day reduces the error versus aircraft, but correlated errors prevent daily averaged errors from dropping below 11.5 ppb. Averaging daily averages over 1 or 3 months equals the daily error divided by the square root of the number of days averaged, indicating that errors are random in this domain. However, averaging months from multiple years, does not reduce the error below 6 ppb, either due to correlated errors, or validation uncertainty.

4.2.5 Summary of errors at all NOAA aircraft sites

Table A.3 in Appendix A shows the single-observation standard deviation for all NOAA aircraft sites. The ocean vs. land observations show similar values, with land and ocean standard deviations within 2 ppb. A single land observation has a standard deviation versus aircraft observations of 23 ppb for the partial column XCH_4 VMR within the pressure levels measured from the aircraft, in agreement with predicted observation error of 23 ppb. The standard deviation for daily averages is 15.2 ppb. This can be compared to the predicted error for the daily averages, assuming randomness, of 5.9 ppb. This indicates that there are correlated (non-random) errors on the order of 15 ppb when averaging observations within 50 km and 1 day. The monthly standard deviation is 10.9, in reasonable agreement with the predicted of 9.4 ppb (from the daily average standard deviation divided by the number of observations averaged). The seasonal cycle average, which is a monthly average of all matched pairs from all years, has a standard deviation of 7.7 ppb, which is similar to the predicted error of 6.9 ppb (from the daily average divided by the square root of number of observations). We find that estimating the error as the daily standard deviation divided by the square root of the number of days averaged is a reasonable estimate of the actual error.

4.2.6 The bias and bias uncertainty

The bias is estimated by calculating the mean bias for each campaign or station separately, then calculating the mean and standard deviation for all campaigns / stations. The bias versus HIPPO is 0 ± 4 ppb. The bias versus ATom is 3 ± 4 ppb. The bias versus NOAA measurements is 9 ± 7 ppb.

5 Discussion and Conclusions

We validate single-footprint AIRS methane by comparing 27,000 AIRS methane retrievals to 396 aircraft profiles from the HIPPO campaign, 719 profiles from the NOAA GML aircraft network, and 289 aircraft profiles from the ATom campaign, taken across a range of latitudes, longitudes, and times. The AIRS methane retrievals are derived using the MUSES optimal estimation algorithm that has previously been applied to Aura TES radiances (e.g. Fu et al., 2013). After adjusting the aircraft profile to account for the AIRS sensitivity (using the averaging kernel and a priori profile), we compare the mean methane value over the aircraft profile to the mean methane from the AIRS profile over the same altitude (or pressure) range. We use a subset of validation data to derive a pressure-dependent bias correction on the order of -60 ppb, and test this on an independent set of validation data. After the bias correction, we report a bias of 0 +/- 10 ppb. The bias between AIRS and aircraft varies with pressure and location, as seen in Appendix A.

After applying the bias correction, from Eq. 8 and 9, the RMS difference between the AIRS and aircraft data of the partial column XCH_4 VMR within the pressure levels measured by the aircraft of ~22 ppb is consistent with the mean observation error, composed of random error from noise and the cross-state errors from jointly retrieved temperature, water vapor, clouds, and surface parameters that are projected onto the AIRS methane retrieval. The extent to which the aircraft profiles used here can be utilized as "truth" for the purposes of validation is limited by knowledge of the methane profile above the aircraft profile (referred to here as "validation uncertainty"), which limits our knowledge of "truth" to within about 10 ppb. This uncertainty is consistent with the location-dependent bias in the satellite/aircraft comparisons which can vary by ~10 ppb.

We quantify the AIRS minus validation standard deviation for single observations, daily averages (within 50 km of the validation location), monthly averages, and seasonal averages for data bias corrected using Eqs. 7 and 8. The AIRS minus validation standard deviations are: 24 ppb (single AIRS footprint), 17 ppb (daily AIRS averages within 1 degree latitude and longitude), 10 ppb ("monthly" AIRS averages), 9 ppb (3-month AIRS average), and 7 ppb (seasonal cycle average). The errors on averaged AIRS data are likely an upper bound on the AIRS error, due to the uncertainty in the validation. The single-footprint and daily average standard deviations for different pressure ranges and surface types are shown in Appendix A. We recommend using the standard deviations in this paragraph as the error budget for the specified averaged quantities.

These results can be compared to AIRS v6 validation by Xiong et al. (2015), which validated AIRS CH_4 retrieved from cloud-cleared radiances on the 9-footprint 45 km field of regard. Xiong et al. (2015) finds AIRS standard deviations versus HIPPO of 0.9% (16 ppb) for pressures between 575 and 777 hPa, 1.2 % (18 ppb) standard deviation for pressures between 441 and 575 hPa, and 1.6% (29 ppb) between 343 and 441 hPa. Xiong et al. (2015) also found a pressure-dependent bias, with a -25 ppb bias near the top of the troposphere, and a high 5 ppb bias near the mid-Troposphere.

5 Appendix A: Biases and standard deviations for different stations, campaigns, pressures, and surface types

We characterize the bias versus validation data by station, campaign, and pressure level. Table A.1 shows biases versus validation data, after bias correction with Eqs. 8 and 9. In the HIPPO comparisons, the biases are generally smaller than about 10 ppb. There is no overall pattern in the bias by season. The land data is biased higher than ocean for HIPPO comparisons (about +20 ppb). However, note that the land observations versus HIPPO are primarily in Australia and New Zealand, whereas the ocean comparisons are in the mid-Pacific.

The NOAA aircraft network comparisons are sorted by site. Many NOAA aircraft locations are at land/ocean interfaces, allowing a more direct comparison of the land/ocean biases. On average, the AIRS land observations are 0-5 ppb higher than AIRS ocean observations at the different pressures and pressure ranges. The overall bias of AIRS versus NOAA aircraft is +7.1 ppb, whereas AIRS versus HIPPO is 4.4 ppb for the partial column XCH_4 VMR within the pressure levels measured by the aircraft. This is consistent with AIRS land having a high bias versus ocean of 0-5 ppb. The standard deviation of the bias for the different campaigns is a useful quantity as it is an indication of systematic error. The standard deviation of the bias varies from 4 ppb to 9 ppb for the different vertical quantities.

Table A.2 shows the mean bias for AIRS minus NOAA GML aircraft for land and ocean AIRS observations. The different rows extend the aircraft using the AIRS prior, the CarbonTracker model (from https://www.esrl.noaa.gov/gmd/ccgg/carbontracker-ch4/ or the GEOS-Chem model (both are extended through 2018 using 2.5% secular increase). The goal of this table is to approximate the influence of the profile extension on the validation accuracy.

Table A.3 shows the standard deviation for AIRS observations minus validation data for land / ocean for different pressure ranges for both single observations and AIRS averages. The mean bias at each site is subtracted prior to calculating the standard deviation. This table shows the standard deviations for single observations and averaged quantities. The predicted error for the daily average is the observation error divided by the square root of the number of observations, and is much smaller than the actual standard deviation, indicating correlated errors. The predicted error for the monthly, 3-month, and seasonal cycle averages is the daily standard deviation divided by the square root of the number of days averaged and ~agrees with the actual standard deviation for the partial column XCH_4 VMR within the pressure levels measured by the aircraft. The location-dependent biases are subtracted from AIRS prior to calculating the standard deviation in all but the last two rows. The last two rows shows the standard deviations without subtracting the location-dependent biases, which increases the standard deviation from about 8 ppb to about 9 ppb.

Author contributions: SSK and JRW are responsible for the study design, data analysis, and manuscript writing; VHP was responsible for data analysis and manuscript editing; DF was responsible for implementing AIRS into the MUSES retrieval system; SCW and BCD were responsible for HIPPO CH_4 data; KM and CS were responsible for the ATom CH_4 data; EJD, CS, and KM were responsible for NOAA GML aircraft data; AL, IP, YH, and KC were responsible for implementation of the fast RT, OSS, used in this work. YY provided LMDZ model runs. DJ provided guidance on GEOS-Chem model runs.

Acknowledgements: This work is supported by NASA ROSES Aura Science Team NNN13D455T. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The NOAA GML aircraft observations were obtained from http://dx.doi.org/10.25925/20190108 and supported by NOAA. The HIPPO aircraft data were obtained from http://www.eol.ucar.edu/projects/hippo/ and supported by NOAA and NSF. Thanks to Bruce Daube, Eric Kort, Jasna Pittman, Greg Santoni and others for QCLS CH_4 data collection/processing. The ATom aircraft data were obtained from https://daac.ornl.gov/ATOM/ and supported by NASA. The GEOS-Chem model output is described in Worden et al. (2013a). Thank you to helpful comments and feedback from J.D. Maasakkers.

Data Availability: AIRS methane data are available at: https://avdc.gsfc.nasa.gov/pub/data/satellite/Aura/TES/.AIRS/TROPESS/YEAR/.... Note that the field "original_species" should be used with the bias correction described in this paper. The specific datasets used in this work are archived at: https://drive.google.com/file/d/1crNs-QcOzbjiZUiTyRiTEsFORFTbODAW/view?usp=sharing.

References

Alvarado, M. J., Payne, V. H., Cady-Pereira, K. E., Hegarty, J. D., Kulawik, S. S., Wecht, K. J., Worden, J. R., Pittman, J. V. and Wofsy, S. C.: Impacts of updated spectroscopy on thermal infrared retrievals of methane evaluated with HIPPO data, Atmospheric Measurement Techniques, 8(2), 965–985, doi:10.5194/amt-8-965-2015, 2015.

Brasseur, G. P., Hauglustaine, D. A., Walters, S., Rasch, P. J., Muller, J. F., Granier, C., and Tie, X. X.: MOZART, a global chemical transport model for ozone and related chemical trac- ers 1. Model description, J. Geophys. Res.-Atmos., 103, 28265–28289, 1998.

Bowman, K. W., Rodgers, C. D., Kulawik, S. S., Worden, J., Sarkissian, E., Osterman, G., Steck, T., Lou, M., Eldering, A. and Shephard, M.: Tropospheric emission spectrometer: Retrieval method and error analysis, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 44(5), 1297–1307, 2006.

Connor, B. J., B ̈osch, H., Toon, G., Sen, B., Miller, C., andCrisp, D.: Orbiting Carbon Observatory: Inverse method andprospective error analysis, J. Geophys. Res., 113, A05305,doi:10.1029/2006JD008336, 2008.

Connor, T. C., Shephard, M. W., Payne, V. H., Cady-Pereira, K. E., Kulawik, S. S., Luo, M., Osterman, G., and Lampel, M.: Long-term stability of TES satellite radiance measurements, Atmos. Meas. Tech., 4, 1481–1490, https://doi.org/10.5194/amt-4-1481-2011, 2011

Connor, B., Bösch, H., McDuffie, J., Taylor, T., Fu, D., Frankenberg, C., O'Dell, C., Payne, V. H., Gunson, M., Pollock, R., Hobbs, J., Oyafuso, F., and Jiang, Y.: Quantification of uncertainties in OCO-2 measurements of XCO2: simulations and linear error analysis, Atmos. Meas. Tech., 9, 5227–5238, https://doi.org/10.5194/amt-9-5227-2016, 2016.

Cooperative Global Atmospheric Data Integration Project: Multi-laboratory compilation of atmospheric methane data for the period 1957-2017; obspack_ch4_1_GLOBALVIEWplus_v1.0_2019_01_08; NOAA, Earth System Research Laboratory, Global Monitoring Division, http://dx.doi.org/10.25925/20190108, 2019.

Crevoisier, C., Nobileau, D., Armante, R., Crépeau, L., Machida, T., Sawa, Y., Matsueda, H., Schuck, T., Thonat, T., Pernin, J., Scott, N. A. and Chedin, A.: The 2007–2011 evolution of tropical methane in the mid-troposphere as seen from space by MetOp-A/IASI, Atmospheric Chemistry and Physics, 13(8), 4279–4289, doi:10.5194/acp-13-4279-2013, 2013.

DeSouza-Machado, S., Strow, L. L., Tangborn, A., Huang, X., Chen, X., Liu, X., Wu, W. and Yang, Q.: Single-footprint retrievals for AIRS using a fast TwoSlab cloud-representation model and the SARTA all-sky infrared radiative transfer algorithm, Atmos. Meas. Tech., 11(1), 529–550, doi:10.1029/2005GL023211, 2018.

de Lange, A. and Landgraf, J.: Methane profiles from GOSAT thermal infrared spectra, Atmos. Meas. Tech., 11, 3815–3828, https://doi.org/10.5194/amt-11-3815-2018, 2018

De Wachter, E., Kumps, N., Vandaele, A. C., Langerock, B., and De Mazière, M.: Retrieval and validation of MetOp/IASI methane, Atmos. Meas. Tech., 10, 4623–4638, https://doi.org/10.5194/amt-10-4623-2017, 2017.

Folberth, G. A., Hauglustaine, D. A., Lathière, J., and Brocheton, F., Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: model description and impact analysis of biogenic hydrocarbons on tropospheric chemistry. Atmospheric Chemistry and Physics, 6(8), 2273–2319. https://doi.org/10.5194/acp-6-2273-2006, 2006.

Frankenberg, C., Meirink, J., Van Weele, M., Platt, U. and Wagner, T.: Assessing methane emissions from global space-borne observations, Science, 308(5724), 1010–1014, doi:10.1126/science.1106644, 2005.

Frankenberg, C., Aben, I., Bergamaschi, P., Dlugokencky, E. J., van Hees, R., Houweling, S., van der Meer, P., Snel, R. and Tol, P.: Global column-averaged methane mixing ratios from 2003 to 2009 as derived from SCIAMACHY: Trends and variability, J. Geophys. Res, 116(D4), D04302, doi:10.1029/2010JD014849, 2011.

Fu, D., Worden, J. R., Liu, X., Kulawik, S. S., Bowman, K. W. and Natraj, V.: Characterization of ozone profiles derived from Aura TES and OMI radiances, Atmospheric Chemistry and Physics, 13(6), 3445–3462, doi:10.5194/acp-13-3445-2013, 2013.

Fu, D., Bowman, K. W., Worden, H. M., Natraj, V., Worden, J. R., Yu, S., Veefkind, P., Aben, I., Landgraf, J., Strow, L., and Han, Y.: High-resolution tropospheric carbon monoxide profiles retrieved from CrIS and TROPOMI, Atmos. Meas. Tech., 9, 2567-4572579, https://doi.org/10.5194/amt-9-2567-2016, 2016.

Fu, D., Kulawik, S. S., Miyazaki, K., Bowman, K. W., Worden, J. R., Eldering, A., Livesey, N. J., Teixeira, J., Irion, F. W., Herman, R. L., Osterman, G. B., Liu, X., Levelt, P. F., Thompson, A. M. and Luo, M.: Retrievals of tropospheric ozone profiles from the synergism of AIRS and OMI: methodology and validation, Atmospheric Measurement Techniques, 11(10), 5587–5605, doi:10.5194/amt-11-5587-2018-supplement, 2018.

Fu D., Millet D.B., Wells K.C., Payne V.H., Yu S., Guenther A., and Eldering A.: Direct retrieval of isoprene from satellite-based infrared measurements, Nature Communication, 10.3811, doi:10.1038/s41467-019-11835-0. 2019.

Herman, Robert, Kulawik, S. S., Kevin Bowman, Karen Cady-Pereira, Annmarie Eldering, Brendan Fisher, Dejian Fu, Robert Herman, Daniel Jacob, Line Jourdain, Susan Kulawik, Ming Luo, Ruth Monarrez, Gregory Osterman, Susan Paradise, Vivienne Payne, Sassaneh Poosti, Nigel Richards, David Rider, Douglas Shepard, Mark Shephard, Felicia Vilnrotter, Helen Worden, John Worden, Hyejung Yun, Lin Zhang, Level 2 (L2) Data User’s Guide (Up to & including Version 7 data), found at https://tes.jpl.nasa.gov/pdf/TES_Level_2_Data_Users_Guide_V7.0.pdf, 2018.

Hu, H., Landgraf, J., Detmers, R., Borsdorff, T., Aan de Brugh, J., Aben, I., et al. (2018). Toward global mapping of methane with TROPOMI: First results and intersatellite comparison to GOSAT. Geophysical Research Letters, 45, 3682– 3689. https://doi.org/10.1002/2018GL077259

Irion, F. W., Kahn, B. H., Schreier, M. M., Fetzer, E. J., Fishbein, E., Fu, D., Kalmus, P., Wilson, R. C., Wong, S. and Yue, Q.: Single-footprint retrievals of temperature, water vapor and cloud properties from AIRS, Atmos. Meas. Tech., 11(2), 971–995, doi:10.1117/12.615244, 2018.

Kort, E. A., Patra, P. K., Ishijima, K., Daube, B. C., Jiménez, R.,Elkin, J., Hurst, D., Moore, F. L., Sweeney, C., and Wofsy,S. C.: Tropospheric distribution and variability of N2O: Evidencefor strong tropical emissions, Geophys. Res. Lett., 38, L15806,https://doi.org/10.1029/2011GL047612, 2011.

Kort, E. A., Wofsy, S. C., Daube, B. C., Diao, M., Elkins, J. W., Gao, R. S., Hintsa, E. J., Hurst, D. F., Jiménez, R., Moore, F. L., Spackman, J. R. and Zondlo, M. A.: Atmospheric observations of Arctic Ocean methane emissions up to 82° north, Nature Geoscience, 5(5), 318–321, doi:10.1038/ngeo1452, 2012.

Kort, E. A., Frankenberg, C., Costigan, K. R., Lindenmaier, R., dubey, M. K. and Wunch, D.: Four corners: The largest US methane anomaly viewed from space, Geophysical Research Letters, 10, 6898, doi:10.1002/2014GL061503, 2014.

Kulawik, S. S., Worden, H., Osterman, G., Ming Luo, Beer, R., Kinnison, D. E., Bowman, K. W., Worden, J., Eldering, A., Lampel, M., Steck, T. and Rodgers, C. D.: TES atmospheric profile retrieval characterization: an orbit of simulated observations, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 44(5), 1324–1333, doi:10.1109/TGRS.2006.871207, 2006a.

Kulawik, S. S., Worden, J., Eldering, A., Bowman, K., Gunson, M., Osterman, G. B., Zhang, L., Clough, S. A., Shephard, M. W. and Beer, R.: Implementation of cloud retrievals for Tropospheric Emission Spectrometer (TES) atmospheric retrievals: part 1. Description and characterization of errors on trace gas retrievals, Journal of Geophysical Research-Atmospheres, 111, D24204, doi:10.1029/2005JD006733, 2006b.

Kulawik, S. S., O'Dell, C., Payne, V. H., Kuai, L., Worden, H. M., Biraud, S. C., Sweeney, C., Stephens, B., Iraci, L. T., Yates, E. L., and Tanaka, T.: Lower-tropospheric CO2 from near-infrared ACOS-GOSAT observations, Atmos. Chem. Phys., 17, 5407–5438, https://doi.org/10.5194/acp-17-5407-2017, 2017.

Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Scarpelli, T. R., Nesser, H., Sheng, J.-X., Zhang, Y., Hersher, M., Bloom, A. A., Bowman, K. W., Worden, J. R., Janssens-Maenhout, G., and Parker, R. J.: Global distribution of methane emissions, emission trends, and OH concentrations and trends inferred from an inversion of GOSAT satellite data for 2010–2015, Atmospheric Chemistry and Physics, 19(11), 7859–7881, doi:10.5194/acp-19-7859-2019, 2019.

Molod, A. Takacs, Lawrence, Suarez, Max, Bacmeister, Julio, Song, In-Sun, Eichmann, Andrew. The GEOS-5 atmospheric general circulation model: mean climate and development from MERRA to Fortuna. Tech. Rep. 28, Goddard Space Flight Center, NASA/TM–2012-104606 28, 1–124 (2012). https://ntrs.
nasa.gov/search.jsp?R=20120011790.

Moncet, J-L. et al: Algorithm theoretical basis document for the Cross Track Infrared Sound er (CrIS). Volume II, Environmental Data Records (EDR), version 4.2. AER Tech. Doc. P1187-TR-I-08, 298 pp. [Available online at http://npp.gsfc.nasa.gov/sciencedocuments/2013-01/474-00056_RevABaseline.pdf.], 2005

Moncet, J-L., G. Uymin, A. E. Lipton, and H. E. Snell: Infrared radiance modeling by optimal spectral sampling. J. Atmos. Sci., 65, 3917–3934, doi: 10.1175/2008JAS2711.1, 2008

Moncet, J-L., Uymin, G., Liang, P. and Lipton, A.E: Fast and accurate radiative transfer in the thermal regime by simultaneous optimal spectral sampling over all channels. Journal of the Atmospheric Sciences, vol 72, 2622-2641, doi: 0.1175/JAS-D-14-0190.1, 2015

Ostler, A., Sussmann, R., Patra, P. K., Houweling, S., De Bruine, M., Stiller, G. P., Haenel, F. J., Plieninger, J., Bousquet, P., Yin, Y., Saunois, M., Walker, K. A., Deutscher, N. M., Griffith, D. W. T., Blumenstock, T., Hase, F., Warneke, T., Wang, Z., Kivi, R., and Robinson, J.: Evaluation of column-averaged methane in models and TCCON with a focus on the stratosphere, Atmos. Meas. Tech., 9, 4843–4859, https://doi.org/10.5194/amt-9-4843-2016, 2016.

Pagano, T. S., Aumann, H. H., Hagan, D. E., and Overoye, K.: Prelaunch and in-flight radiometric calibration of the Atmospheric Infrared Sounder (AIRS), IEEE T. Geosci. Remote, 41, 265–273, 2003.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O. and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations, Geophys. Res. Lett, 38(15), L15807, doi:10.1029/2011GL047871, 2011.

Rigby, M., Montzka, S. A., Prinn, R. G., White, J. W. C., Young, D., O'Doherty, S., Lunt, M. F., Ganesan, A. L., Manning, A. J., Simmonds, P. G., Salameh, P. K., Harth, C. M., Muehle, J., Siddans, R., Knappett, D., Kerridge, B., Waterfall, A., Hurley, J., Latter, B., Boesch, H., and Parker, R.: Global height-resolved methane retrievals from the Infrared Atmospheric Sounding Interferometer (IASI) on MetOp, Atmos. Meas. Tech., 10, 4135–4164, https://doi.org/10.5194/amt-10-4135-2017, 2017.

Pandey, S., Gautam, R., Houweling, S., van der Gon, H. D., Sadavarte, P., Borsdorff, T., Hasekamp, O., Landgraf, J., Tol, P., van Kempen, T., Hoogeveen, R., van Hees, R., Hamburg, S. P., Maasakkers, J. D., and Aben, I.: Satellite observations reveal extreme methane leakage from a natural gas well blowout, Proc Natl Acad Sci USA, 116, 26376-26381, DOI: 10.1073/pnas.1908712116, 2019.

Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P., Ito, A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: TransCom model simulations of CH4 and related species: linking transport, surface flux and chemical loss with CH4 variability in the troposphere and lower stratosphere, Atmos. Chem. Phys., 11, 12813–12837, https://doi.org/10.5194/acp-11-12813-2011, 2011.

Razavi, A., Clerbaux, C., Wespes, C., Clarisse, L., Hurtmans, D., Payan, S., Camy-Peyret, C., and Coheur, P. F.: Characterization of methane retrievals from the IASI space-borne sounder, Atmos. Chem. Phys., 9, 7889–7899, https://doi.org/10.5194/acp-9-7889-2009, 2009

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: The- ory and Practice, World Sci., Tokyo, 2000.

Santoni, G. W., Daube, B. C., Kort, E. A., Jiménez, R., Park, S., Pittman, J. V., Gottlieb, E., Xiang, B., Zahniser, M. S., Nelson, D. D., McManus, J. B., Peischl, J., Ryerson, T. B., Holloway, J. S., Andrews, A. E., Sweeney, C., Hall, B., Hintsa, E. J., Moore, F. L., Elkins, J. W., Hurst, D. F., Stephens, B. B., Bent, J., and Wofsy, S. C.: Evaluation of the airborne quantum cascade laser spectrometer (QCLS) measurements of the carbon and greenhouse gas suite – CO2, CH4, N2O, and CO – during the CalNex and HIPPO campaigns, Atmos. Meas. Tech., 7, 1509–1526, https://doi.org/10.5194/amt-7-1509-2014, 2014.

Schepers, D., Guerlet, S., Butz, A., Landgraf, J., Frankenberg, C., Hasekamp, O., Blavier, J.-F., Deutscher, N. M., Griffith, D. W. T., Hase, F., Kyro, E., Morino, I., Sherlock, V., Sussmann, R., and Aben, I.: Methane retrievals from Greenhouse Gases Observing Satellite (GOSAT) shortwave infrared measurements: performance comparison of proxy and physics retrieval algorithms, J. Geophys. Res., 117, D10307, doi:10.1029/2012JD017549, 2012.

Shephard, M.W., Worden, H.M., Cady-Pereira, K.E., Lampel, M., Luo, M., Bowman, K.W., Sarkissian, E., Beer, R., Rider, D.M., Tobin, D.C., Revercomb, H.E., Fisher, B.M., Tremblay, D., Clough, S.A., Osterman, G.B., and Gunson, M., Tropospheric Emission Spectrometer nadir spectral radiance comparisons, J. Geophys. Res., 113, D15S05, doi:10.1029/2007JD008856, 2008.

Siddans, R., Knappett, D., Kerridge, B., Waterfall, A., Hurley, J., Latter, B., Boesch, H., and Parker, R.: Global height-resolved methane retrievals from the Infrared Atmospheric Sounding Interferometer (IASI) on MetOp, Atmos. Meas. Tech., 10, 4135–4164, https://doi.org/10.5194/amt-10-4135-2017, 2017.

Smith, N. and Barnet, C. D. (2019), Uncertainty characterization and propagation in the Community Long-Term Infrared Microwave Combined Atmospheric Product System (CLIMCAPS), Remote Sens., 11, doi:10.3390/rs11101227

Turner, A. J., Frankenberg, C. and Kort, E. A.: Interpreting contemporary trends in atmospheric methane, Proceedings of the National Academy of Sciences of the United States of America, 8(8), 201814297–9, doi:10.1073/pnas.1814297116, 2019.

Turner, A. J., Fung, I., Naik, V., Horowitz, L. W. and Cohen, R. C.: Modulation of hydroxyl variability by ENSO in the absence of external forcing, Proceedings of the National Academy of Sciences of the United States of America, 115(36), 8931–8936, doi:10.1073/pnas.1807532115, 2018a.

Turner, A. J., Jacob, D. J., Benmergui, J., Brandman, J., White, L. and Randles, C. A.: Assessing the capability of different satellite observing configurations to resolve the distribution of methane emissions at kilometer scales, Atmospheric Chemistry and Physics, 18(11), 8265–8278, doi:10.5194/acp-18-8265-2018, 2018b.

Turner, A.J., D.J. Jacob, K.J. Wecht, J.D. Maasakkers, E. Lundgren, A.E. Andrews, S.C. Biraud, H. Boesch, K.W. Bowman, N.M. Deutscher, M.K. Dubey, D.W.T. Griffith, F. Hase, A. Kuze, J. Notholt, H. Ohyama, R. Parker, V.H. Payne, R. Sussmann, C. Sweeney, V.A. Velazco, T. Warneke, P.O. Wennberg, and D. Wunch Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, Atmos. Chem. Phys., 15, 7049-7069, 2015.

Varon, D. J., McKeever, J., Jervis, D., Maasakkers, J. D., Pandey, S., Houweling, S., et al. (2019). Satellite discovery of anomalously large methane point sources from oil/gas production. Geophysical Research Letters, 46, 13507– 13516. https://doi.org/10.1029/2019GL083798

Wecht, K. J., Jacob, D. J., Wofsy, S. C., Kort, E. A., Worden, J. R., Kulawik, S. S., Henze, D. K., Kopacz, M. and Payne, V. H.: Validation of TES methane with HIPPO aircraft observations: implications for inverse modeling of methane sources, Atmospheric Chemistry and Physics, 12(4), 1823–1832, doi:10.5194/acp-12-1823-2012, 2012.

Wecht, K. J., Jacob, D. J. and Sulprizio, M. P.: Spatially resolving methane emissions in California: constraints from the CalNex aircraft campaign and from present (GOSAT, TES) and future (TROPOMI) missions, Atmospheric Chemistry and Physics, doi:10.5194/acp-14-8173-2014, 2014.

Wofsy, S. C., B. C. Daube, R. Jimenez, E. Kort, J. V. Pittman, S. Park, R. Commane, B. Xiang G. Santoni, D. Jacob, J. Fisher, C. Pickett-Heaps, H. Wang, K. Wecht, Q.-Q. Wang, B. B. Stephens, S. Shertz, A.S. Watt, P. Romashkin, T. Campos, J. Haggerty, W. A. Cooper, D. Rogers, S. Beaton, R. Hendershot, J. W. Elkins, D. W. Fahey, R. S. Gao, F. Moore, S. A. Montzka, J. P. Schwarz, A. E. Perring, D. Hurst, B. R. Miller, C. Sweeney, S. Oltmans, D. Nance, E. Hintsa, G. Dutton, L. A. Watts, J. R. Spackman, K. H. Rosenlof, E. A. Ray, B. Hall, M. A. Zondlo, M. Diao, R. Keeling, J. Bent, E. L. Atlas, R. Lueb, M. J. Mahoney. HIPPO Merged 10-second Meteorology, Atmospheric Chemistry Aerosol Data (R_20121129), Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee U.S.A., 2012.

Wofsy, S.C., S. Afshar, H.M. Allen, E.C. Apel, E.C. Asher, B. Barletta, J. Bent, H. Bian, B.C. Biggs, D.R. Blake, N. Blake, I. Bourgeois, C.A. Brock, W.H. Brune, J.W. Budney, T.P. Bui, A. Butler, P. Campuzano-Jost, C.S. Chang, M. Chin, R. Commane, G. Correa, J.D. Crounse, P. D. Cullis, B.C. Daube, D.A. Day, J.M. Dean-Day, J.E. Dibb, J.P. DiGangi, G.S. Diskin, M. Dollner, J.W. Elkins, F. Erdesz, A.M. Fiore, C.M. Flynn, K.D. Froyd, D.W. Gesler, S.R. Hall, T.F. Hanisco, R.A. Hannun, A.J. Hills, E.J. Hintsa, A. Hoffman, R.S. Hornbrook, L.G. Huey, S. Hughes, J.L. Jimenez, B.J. Johnson, J.M. Katich, R.F. Keeling, M.J. Kim, A. Kupc, L.R. Lait, J.-F. Lamarque, J. Liu, K. McKain, R.J. Mclaughlin, S. Meinardi, D.O. Miller, S.A. Montzka, F.L. Moore, E.J. Morgan, D.M. Murphy, L.T. Murray, B.A. Nault, J.A. Neuman, P.A. Newman, J.M. Nicely, X. Pan, W. Paplawsky, J. Peischl, M.J. Prather, D.J. Price, E. Ray, J.M. Reeves, M. Richardson, A.W. Rollins, K.H. Rosenlof, T.B. Ryerson, E. Scheuer, G.P. Schill, J.C. Schroder, J.P. Schwarz, J.M. St.Clair, S.D. Steenrod, B.B. Stephens, S.A. Strode, C. Sweeney, D. Tanner, A.P. Teng, A.B. Thames, C.R. Thompson, K. Ullmann, P.R. Veres, N. Vieznor, N.L. Wagner, A. Watt, R. Weber, B. Weinzierl, P.O. Wennberg, C.J. Williamson, J.C. Wilson, G.M. Wolfe, C.T. Woods, and L.H. Zeng. 2018. ATom: Merged Atmospheric Chemistry, Trace Gases, and Aerosols. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1581

Worden, J., Kulawik, S., Shepard, M., Clough, S., Worden, H., Bowman, K. and Goldman, A.: Predicted errors of tropospheric emission spectrometer nadir retrievals from spectral window selection, Journal of Geophysical Research-Atmospheres, 109(D9), D09308, doi:10.1029/2004JD004522, 2004.

Worden, J., Noone, D., Galewsky, J., Bailey, A., Bowman, K., Brown, D., Hurley, J., Kulawik, S., Lee, J. and Strong, M.: Estimate of bias in Aura TES HDO/H2O profiles from comparison of TES and in situ HDO/H2O measurements at the Mauna Loa observatory, Atmospheric Chemistry and Physics, 11(9), 4491–4503, doi:10.5194/acp-11-4491-2011, 2011.

Worden, J., Kulawik, S., Frankenberg, C., Payne, V., Bowman, K., Cady-Peirara, K., Wecht, K., Lee, J.-E., and Noone, D.: Profiles of CH4, HDO, H2O, and N2O with improved lower tropospheric vertical resolution from Aura TES radiances, Atmos. Meas. Tech., 5, 397-411, https://doi.org/10.5194/amt-5-397-2012, 2012.

Worden, J., K. Wecht, C. Frankenberg, M. Alvarado, K. Bowman, E. Kort, S. Kulawik, M. Lee, V. Payne, and H. Worden, CH4 and CO distributions over tropical fires during October 2006 observed by the Aura TES satellite instrument and modeled by GEOS-Chem, Atmospheric Chemistry and Physics, 13(7), 3679–3692, doi:10.5194/acp-13-3679-2013, 2013a.

Worden, J., Jiang, Z., Jones, D. B. A., Alvarado, M., Bowman, K., Frankenberg, C., Kort, E. A., Kulawik, S. S., Lee, M., Liu, J., Payne, V., Wecht, K. and Worden, H.: El Nino, the 2006 Indonesian Peat Fires, and the distribution of atmospheric methane, Geophys. Res. Lett, 40, 1, doi:10.1002/grl.50937, 2013b.

Worden, J. R., Turner, A. J., Bloom, A., Kulawik, S. S., Liu, J., Lee, M., Weidner, R., Bowman, K., Frankenberg, C., Parker, R., and Payne, V. H.: Quantifying lower tropospheric methane concentrations using GOSAT near-IR and TES thermal IR measurements, Atmos. Meas. Tech., 8, 3433-3445, https://doi.org/10.5194/amt-8-3433-2015, 2015.

Worden, J. R., Bloom, A. A., Pandey, S., Jiang, Z., Worden, H. M., Walker, T. W., Houweling, S. and Röckmann, T.: Reduced biomass burning emissions reconcile conflicting estimates of the post-2006 atmospheric methane budget, Nat Commun, 1–11, doi:10.1038/s41467-017-02246-0, 2017a.

Worden, J. R., Doran, G., Kulawik, S., Eldering, A., Crisp, D., Frankenberg, C., O amp apos Dell, C. and Bowman, K.: Evaluation and attribution of OCO-2 XCO2 uncertainties, Atmospheric Measurement Techniques, 10(7), 2759–2771, doi:10.5194/amt-10-2759-2017, 2017b.

Worden, J. R., Kulawik, S. S., Fu, D., Payne, V. H., Lipton, A. E., Polonsky, I., He, Y., Cady-Pereira, K., Moncet, J.-L., Herman, R. L., Irion, F. W. and Bowman, K. W.: Characterization and evaluation of AIRS-based estimates of the deuterium content of water vapor, Atmospheric Measurement Techniques, 12(4), 2331–2339, doi:10.5194/amt-12-2331-2019, 2019.

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., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech., 3, 1351–1362, https://doi.org/10.5194/amt-3-1351-2010, 2010.

Xiong, X., Barnet, C., Maddy, E., Wofsy, S. C., Chen, L. A., Karion, A. and Sweeney, C.: Detection of methane depletion associated with stratospheric intrusion by atmospheric infrared sounder (AIRS), Geophysical Research Letters, 40(10), 2455–2459, doi:10.1002/grl.50476, 2013.

Xiong, X., Houweling, S., Wei, J., Maddy, E., Sun, F. and Barnet, C.: Methane plume over south Asia during the monsoon season: satellite observation and model simulation, Atmospheric Chemistry and Physics, 9(3), 783–794, 2009.

Xiong X, Han Y, Liu Q and Weng F, Comparison of atmospheric methane retrievals from AIRS and IASI IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 3297–303, 2016.

Xiong, X., Weng, F., Liu, Q., and Olsen, E.: Space-borne observation of methane from atmospheric infrared sounder version 6: validation and implications for data analysis, Atmos. Meas. Tech. Discuss., 8, 8563–8597, https://doi.org/10.5194/amtd-8-8563-2015, 2015.

Zhang, Y., Jacob, D. J., Maasakkers, J. D., Sulprizio, M. P., Sheng, J.-X., Gautam, R. and Worden, J.: Monitoring global tropospheric OH concentrations using satellite observations of atmospheric methane, Atmospheric Chemistry and Physics, 18(21), 15959–15973, doi:10.1029/2007JG000500, 2018.

Table 1. Bias versus pressure with and without bias correction. The bias correction was developed on HIPPO-4 and tested on HIPPO-4; HIPPO-1,2,3,5; and NOAA aircraft network.

Pressure (hPa)	AIRS minus aircraft_AK (HIPPO-4) (ppb)	After bias correction (HIPPO-4) (ppb)	After bias correction (all HIPPO except HIPPO-4) (ppb)	After bias correction (all NOAA) (ppb)
1000	24	-1	-3	1
824	36	0	-4	1
681	48	1	-5	2
562	58	1	-4	2
464	60	-5	-3	3
383	67	-5	-2	2
316	81	1	4	-
261	86	1	4	-
215	89	1	3	-
161	-	-	4	-

Table A.1 Bias by campaign, station, land/ocean, and pressure.

Station/campaign	Location	Time period	Bias 700 hPa (ppb)	Bias 500 hPa (ppb)	Bias 300 hPa (ppb)	Bias column matching aircraft (ppb)	Bias column above 750 hPa (ppb)
HIPPO 1S	Pacific	Jan, 2009	-6.2	2.4	11.0	4.2	6.3
HIPPO 1N	Pacific	Jan, 2009	-3.2	3.7	12.5	-0.1	4.8
HIPPO 2S	Pacific	Nov, 2009	-9.0	-0.4	9.8	-4.4	5.0
HIPPO 2N	Pacific	Nov, 2009	-4.3	-3.3	-3.1	-4.0	-4.0
HIPPO 3N	Pacific	Apr, 2010	-8.5	1.1	16.5	-2.6	2.6
HIPPO 4S	Pacific	Jun, 2011	-0.7	-2.0	9.5	1.8	10.2
HIPPO 4N	Pacific	Jul, 2011	8.7	11.8	0.7	8.7	7.3
HIPPO 5S	Pacific	Aug, 2011	1.2	7.6	13.3	4.5	9.3
HIPPO 5N	Pacific	Sep, 2011	-5.2	0.5	1.2	-2.0	2.2
HIPPO all land	**-**	**-**	**10.9**	**18.2**	**17.8**	**16.1**	**14.8**
HIPPO all ocean	**-**	**-**	**-5.2**	**-0.9**	**4.3**	**-1.7**	**3.1**
HIPPO all (mean)	**-**	**-**	**-2.9**	**2.1**	**7.9**	**0.7**	**4.9**
HIPPO all (stdev)	**-**	**-**	**5.9**	**5.2**	**6.7**	**4.4**	**4.3**
ACG	68N, 152W	-	21.4	-	-	18.6	26.7
ESP	49N, 126W	-	9.7	-	-	8.2	13.8
NHA	43N, 71W	-	15.7	23.8	-	15.7	19.3
THD	41N, 124W	-	13.6	21.7	-	14.0	21.2
CMA	39N, 74W	-	-0.2	5.7	-	0.9	3.6
TGC	28N, 97W	-	1.0	7.9	-	2.3	6.5
RTA	21S, 160W	-	3.7	11.5	-	3.9	12.8
NOAA all land	**-**	**-**	**9.2**	**16.8**	**-**	**9.4**	**14.3**
NOAA all ocean	**-**	**-**	**9.0**	**12.8**	**-**	**8.7**	**15.4**

NOAA all (mean)	**-**	**-**	**9.3**	**14.1**	**-**	**9.1**	**14.8**
NOAA all (stdev)	**-**	**-**	**8.1**	**8.2**	**-**	**7.1**	**8.2**
ATom 1S	Pacific	Aug, 2016	-0.2	4.5	7.7	2.0	3.5
ATom 1N	Atlantic	Aug, 2016	0.2	3.2	13.2	2.8	6.9
ATom 2S	Pacific	Feb, 2017	-6.8	0.7	8.4	-2.5	5.2
ATom 2N	Atlantic	Feb, 2017	5.7	12.3	25.3	8.3	12.5
ATom 3S	Pacific	Oct, 2017	-2.5	3.0	9.1	0.9	5.9
ATom 3N	Atlantic/Pacific	Oct, 2017	6.5	13.0	21.9	9.3	13.8
ATom 4S	Pacific	April/May, 2018	-0.1	3.9	9.4	2.3	6.0
ATom 4N	Atlantic	May, 2018	-1.4	5.9	23.4	3.4	13.2
ATom all land	**-**	**-**	**16.7**	**23.6**	**26.2**	**17.0**	**18.2**
ATom all ocean	**-**	**-**	**-3.2**	**2.4**	**13.4**	**0.6**	**6.5**
ATom all (mean)	**-**	**-**	**0.1**	**5.8**	**14.7**	**3.2**	**8.3**
ATom all (stdev)	**-**	**-**	**4.3**	**4.5**	**7.5**	**3.8**	**4.1**

Table A.2 Change in the mean bias of the partial column matching the NOAAaircraft observation using different aircraft profile extensions from the top aircraft measurement to the top of the atmosphere.

Quantity	Profile extension	Bias 700 hPa (ppb)	Bias 500 hPa (ppb)	Bias 300 hPa (ppb)	Bias column matching aircraft (ppb)	Bias column above 750 hPa (ppb)
Land NOAA	CT	6.0	10.3	-	6.1	3.8
Ocean NOAA	CT	4.5	5.7	-	4.3	4.0
Land NOAA	prior	9.2	16.8	-	9.4	14.3
Ocean NOAA	prior	9.0	12.8	-	8.7	15.4
Land NOAA	GEOS-Chem	6.4	11.7	-	6.7	6.4
Ocean NOAA	GEOS-Chem	4.4	7.7	-	4.5	6.4

Table A.3 Standard deviation of AIRS minus validation for land / ocean observations and different pressures / pressure ranges. Rows 1-2 show the standard deviation for single observation, rows 3-4 show the predicted observation error, rows 5-8 show the standard deviation for daily averages, rows 9-10 show the predicted error for daily averages (assuming random error), rows 11-12 show the standard deviation for 3-month averages, rows 13-14 show the standard deviation for seasonal cycle averages (average the same month of all years), rows 15-16 show the predicted error for the seasonal cycle averages, and rows 17-18 show the standard deviation without bias subtraction. The site-dependent biases from Table A.1 are subtracted prior to calculating the standard deviation.

Quantity	Stdev 700 hPa (ppb)	Stdev 500 hPa (ppb)	Stdev 300 hPa (ppb)	**Stdev column matching aircraft (ppb)**	Stdev column above 750 hPa (ppb)
Land single	26	29	26	**23**	25
Ocean single	25	27	26	**22**	24
Land observation error	26	26	19	**23**	19
Ocean observation error	28	28	20	**24**	19
Land daily (≥3 obs/day)	17	21	16	**15**	20
Ocean daily (≥3 obs/day)	18	21	21	**16**	20
Land daily (≥9 obs/day)	16	20	16	**14**	20
Ocean daily (≥9 obs/day)	17	19	21	**15**	18
Land daily (≥9 obs/day) pred.	9.7	9.9	5.7	**8.5**	7.0
Ocean daily (≥9 obs/day) pred.	8.4	7.9	4.6	**7.0**	5.7
Land 3-month (≥3 obs/day, ≥3 days)	9.5	13.3	-	**8.8**	12.9

Ocean 3-month (≥3 obs/day, ≥3 days)	9.0	11.8	-	**8.3**	11.8
Land monthly (average all years)	8.3	11.8	-	**7.7**	10.7
Ocean monthly (average all years)	8.3	10.4	-	**7.5**	10.1
Land monthly (average all years) pred.	7.7	9.9	-	**6.9**	9.3
Ocean monthly (average all years) pred.	8.0	9.8	-	**7.2**	9.5
Land monthly (average all years) without bias subtraction	9.9	13.7	-	**9.1**	12.2
Ocean monthly (average all years) without bias subtraction	10.4	12.3	-	**9.4**	11.6

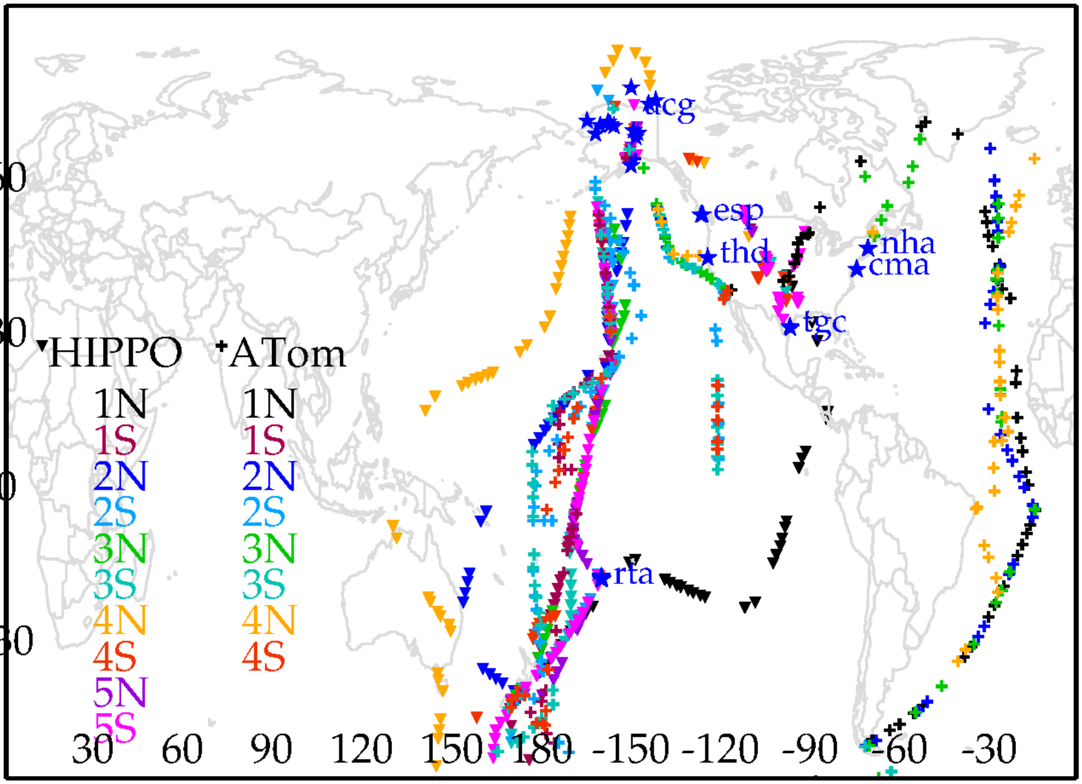

Figure 1: Location of aircraft profile measurements used for validation. The upside-down triangles show HIPPO, + show ATom, and blue stars show NOAA ESRL aircraft validation locations.

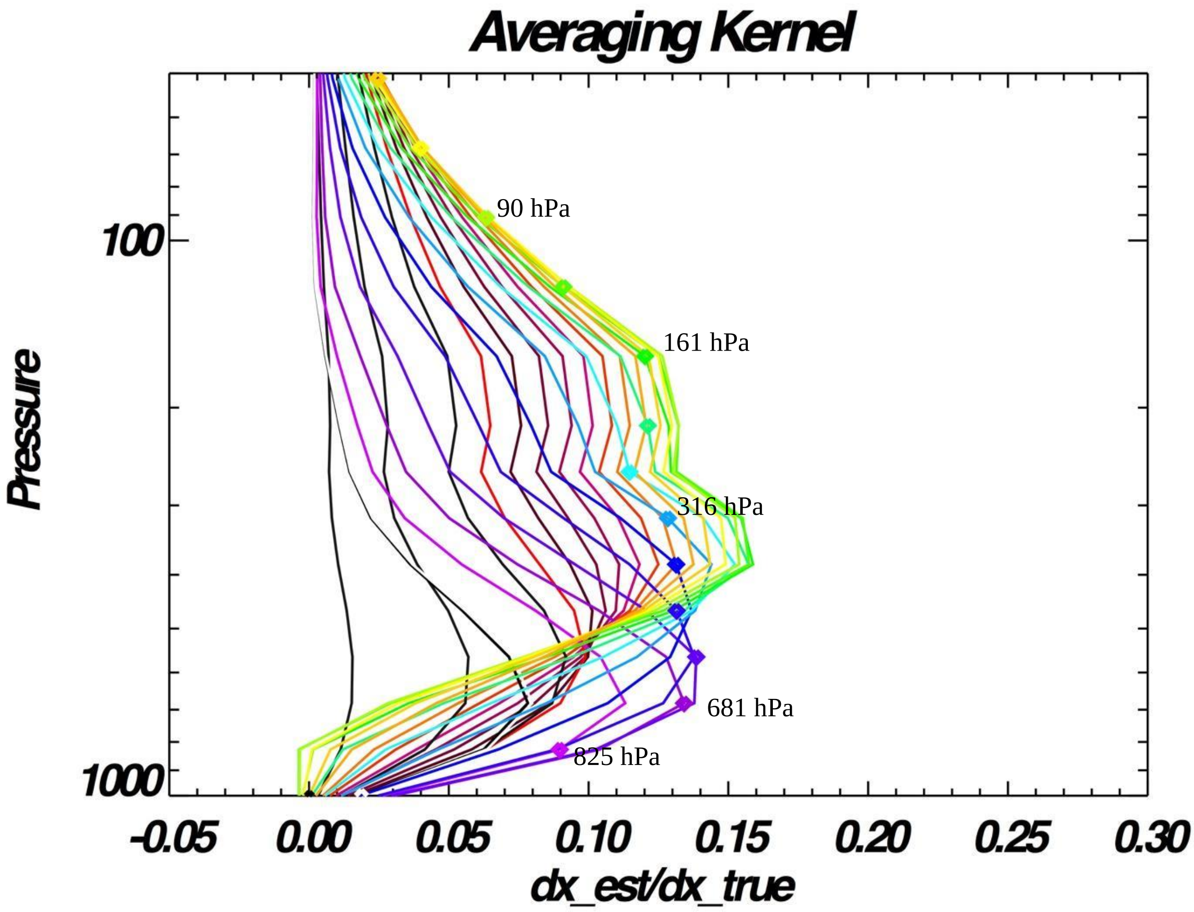

Figure 2: The rows of an averaging kernel for CH_4 for a tropical scene. The colors help for visualization of the pressure levels for each row of the averaging kernel. The diamonds indicate the pressure level corresponding to the row of the averaging kernel, for pressures 1012, 825, 681, 562, 464, 383, 316, 261, 215, 161, 121, 90, 68, 51 hPa.

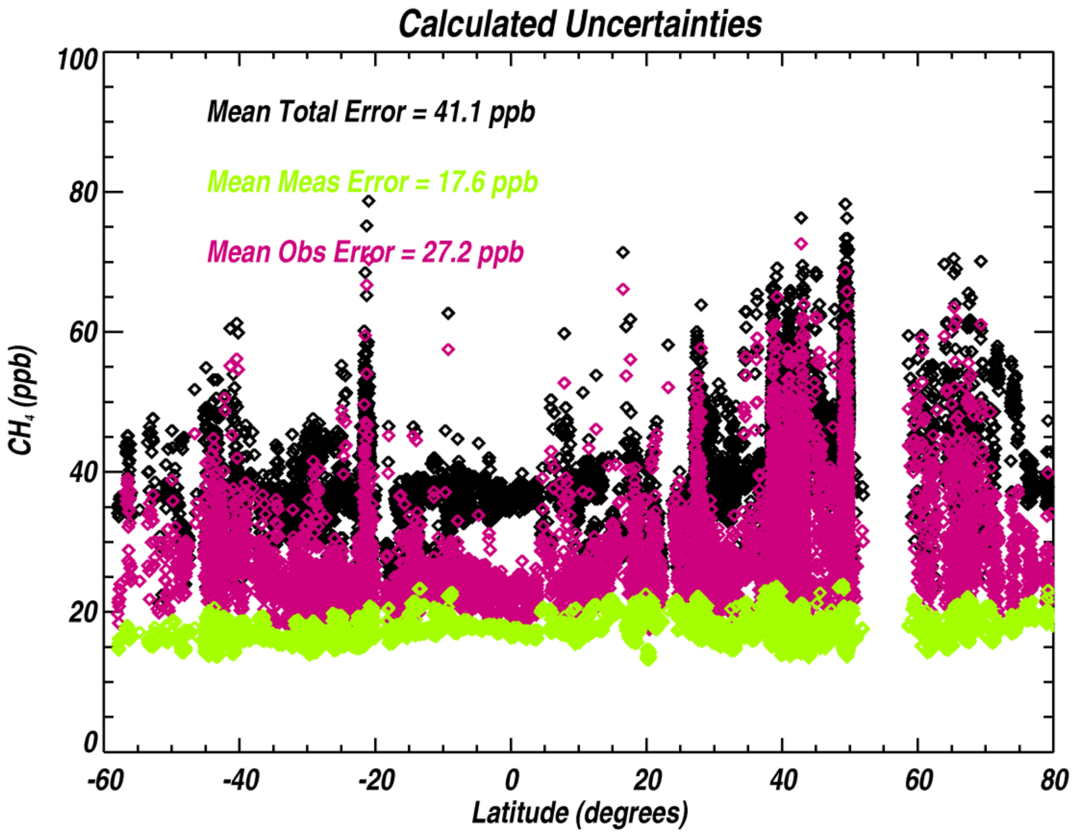

Figure 3: Calculated errors for AIRS measurements shown in this paper. The total error shown is the smoothing error (Eq. 5) plus the observation error (Eq. 7b). The measurement error is the last term of Eq. 7b, and the only fully random error.

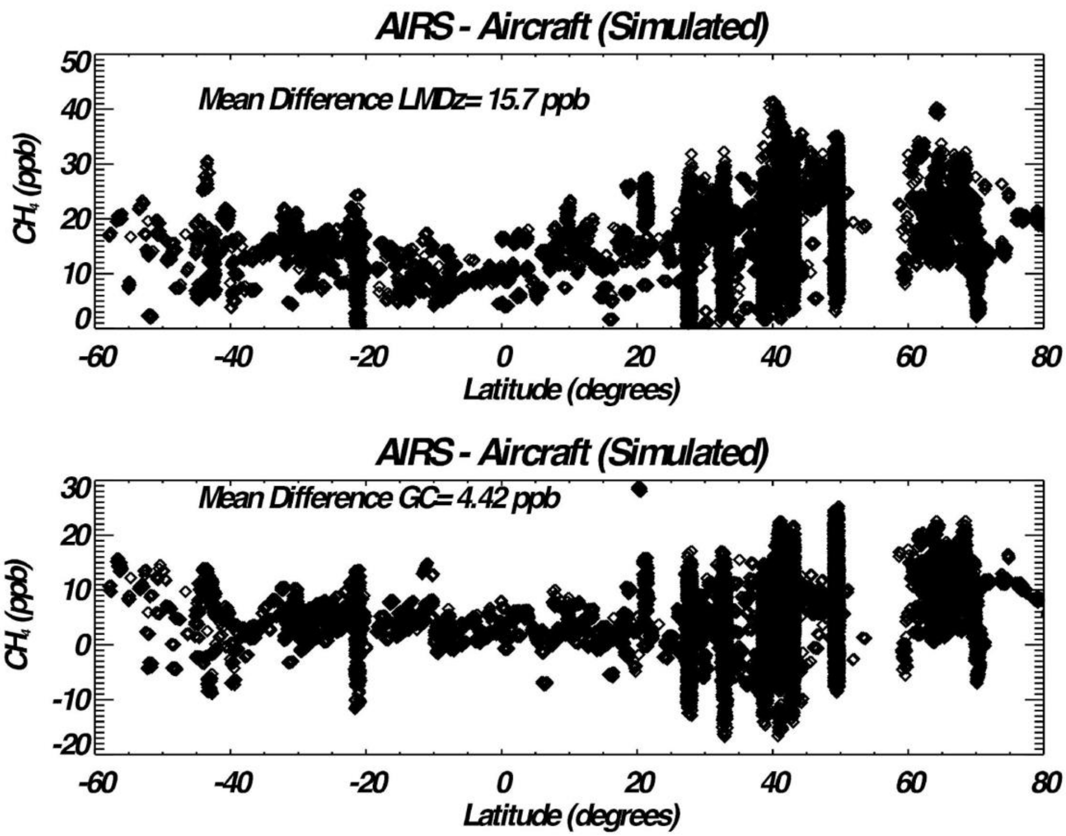

Figure 4: Simulated comparison between AIRS and Aircraft in which the LMDz model (top) and GEOS-Chem model (bottom) are used for the simulation. This represents uncertainty in the true state that we validate against.

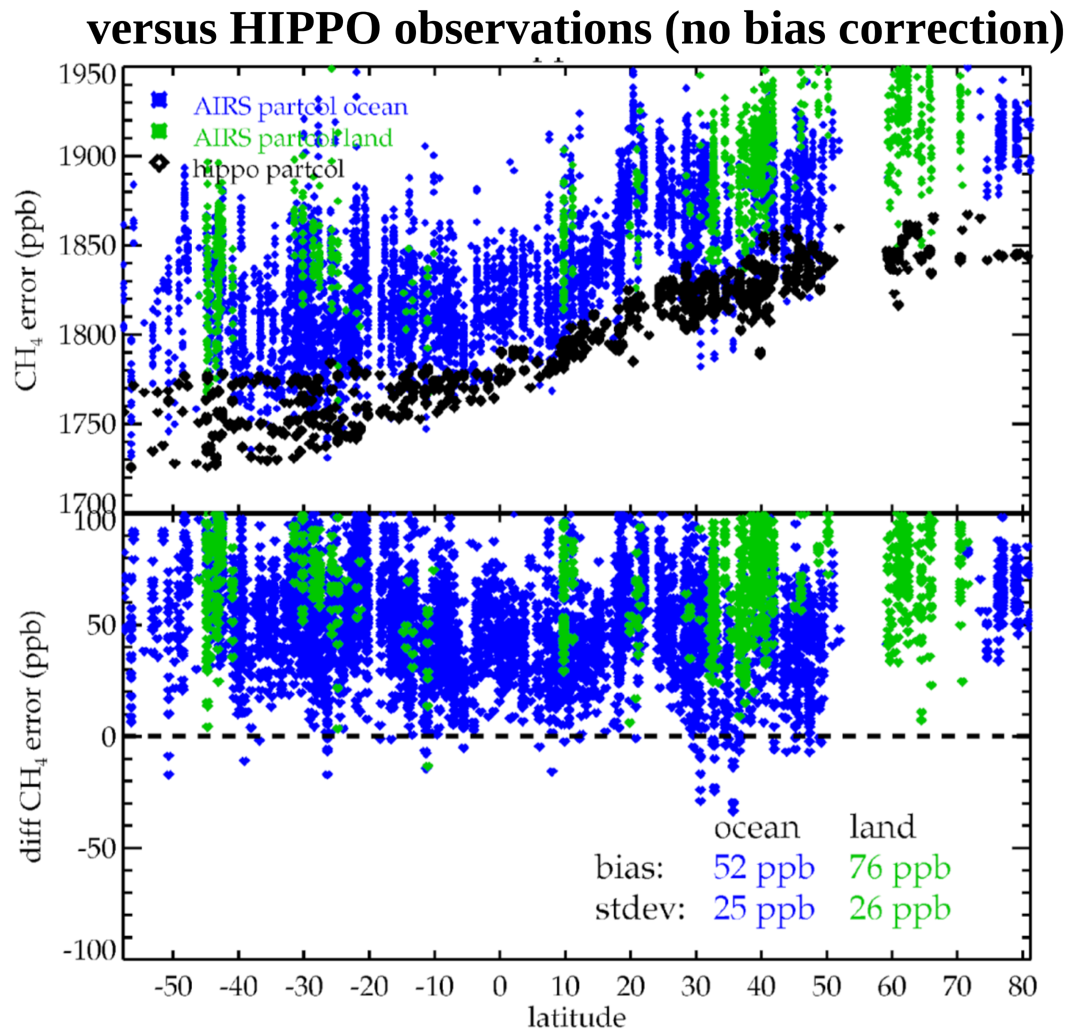

Figure 5: Comparison of AIRS methane VMR to aircraft for all HIPPO comparisons over the partial column XCH_4 VMR within the pressure levels measured by the aircraft. Blue shows AIRS ocean observations and green shows AIRS land observations.

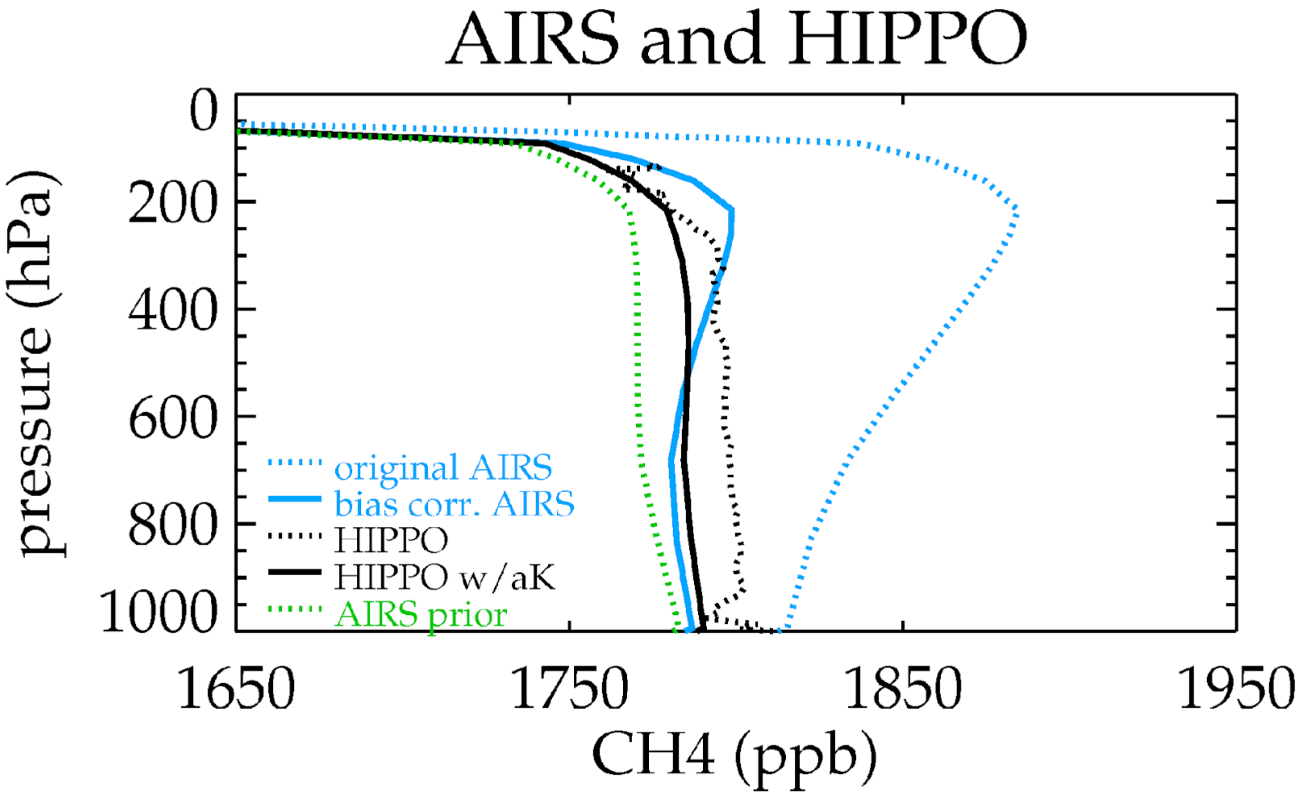

Figure 6: Example of the effect of bias correction on the AIRS profile from averaged HIPPO-1,2,3,5. The blue lines shows the AIRS methane profile before (dotted) and after (solid) bias correction. The black line shows the HIPPO measurements before (dotted) and after averaging kernel is applied (solid).

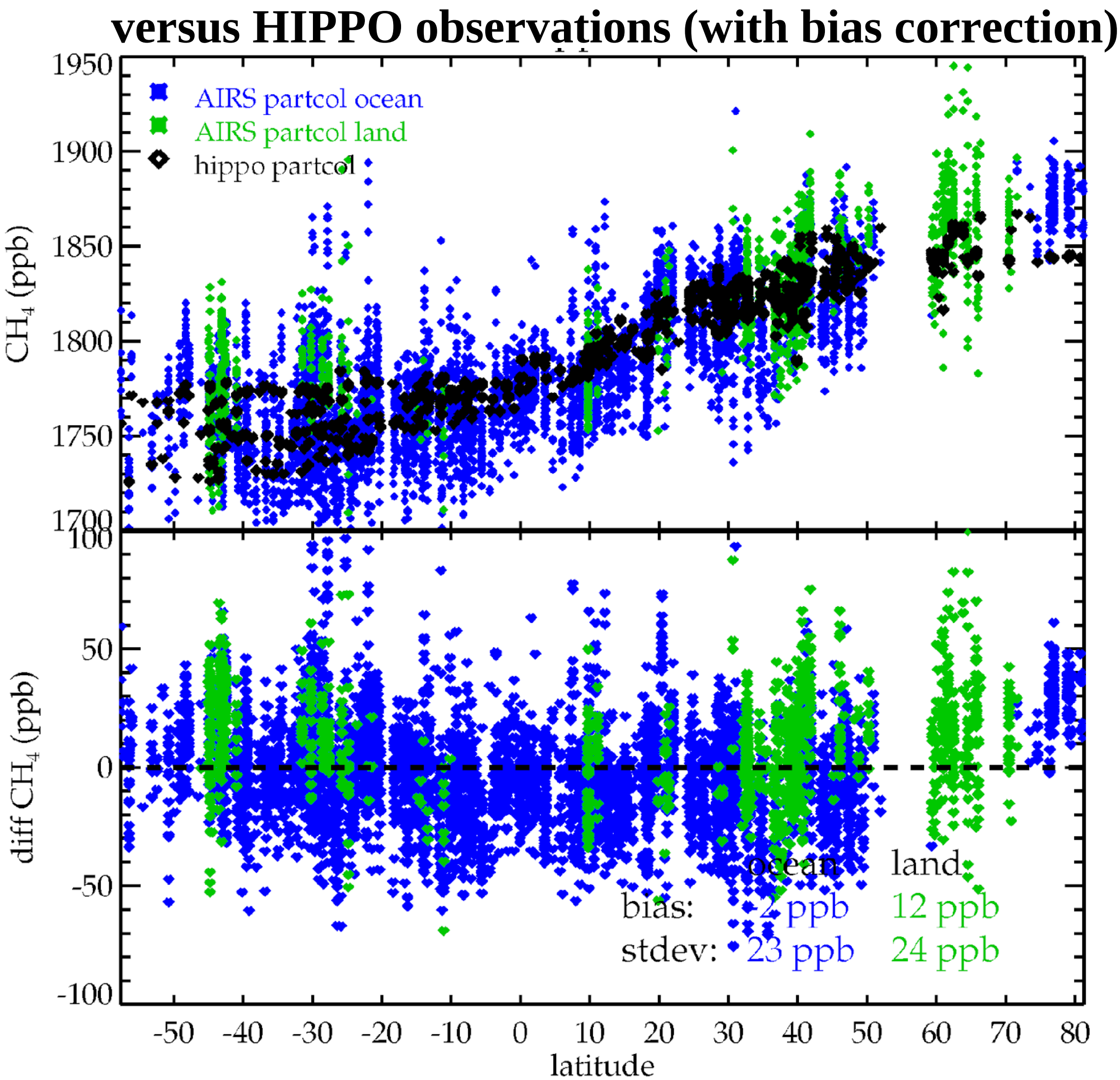

Figure 7: Same as Figure 5, but after bias correction. The ocean has -2 ppb bias, 23 ppb stdev, and the land has 12 ppb bias, 24 ppb stdev.

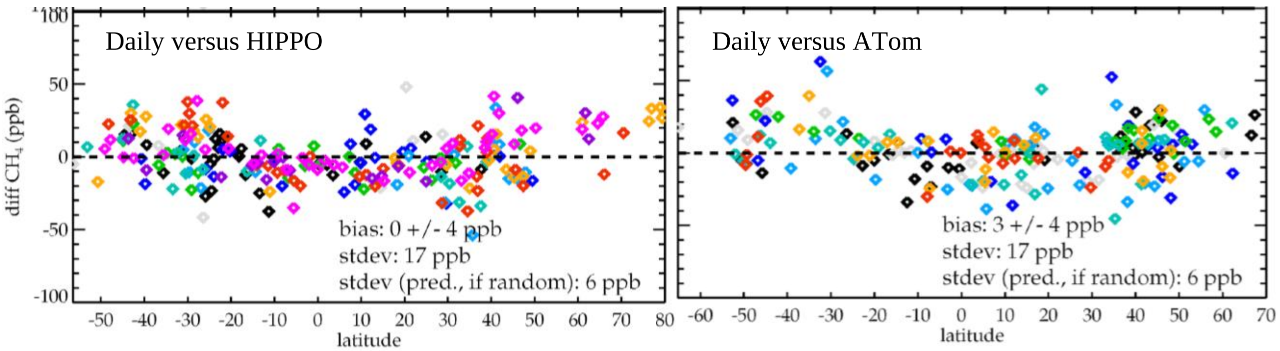

Figure 8: Comparison of daily averaged AIRS to HIPPO measurements (left) and ATom measurements (right) for the partial column observed by the aircraft. The different colors correspond to the campaigns shown in Fig. 1.

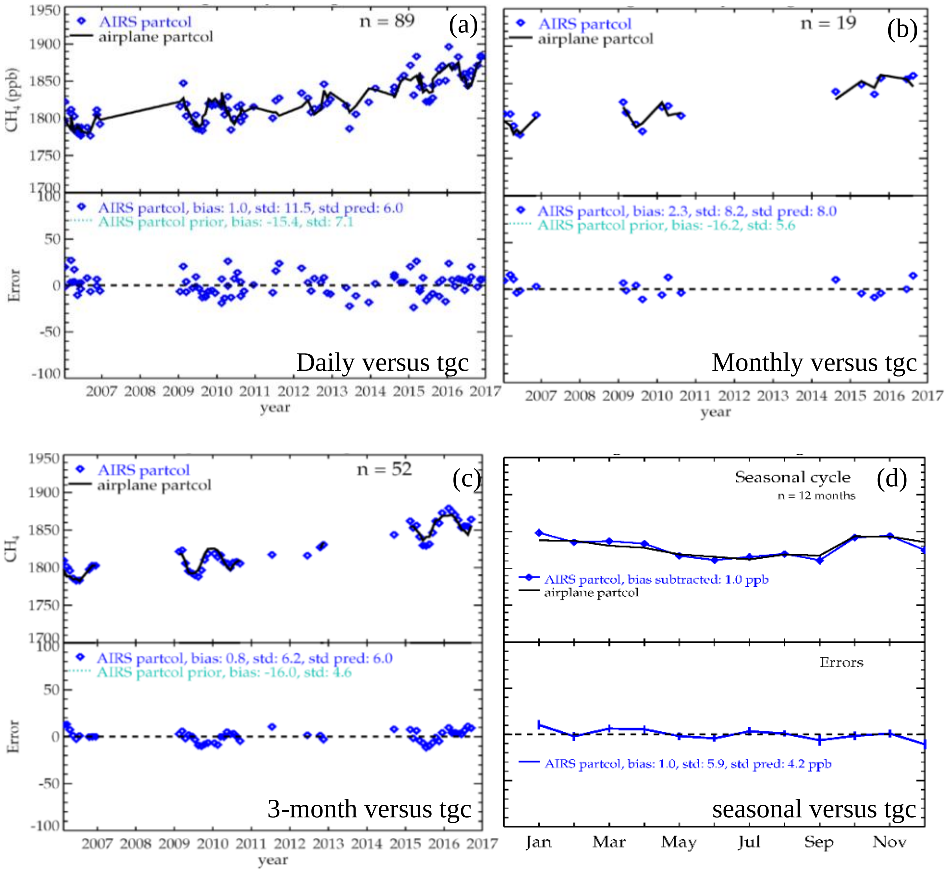

Figure 9: Comparison at TGC (27.7N, 96.9E). (Top) Comparison of AIRS and co-located NOAA aircraft flights in SE Texas for the partial column XCH_4 VMR within the pressure levels measured by the aircraft. Data are averaged over (a) 1 day, (b) 1 month, (c) 90-days, and (d) averaged by month from all years. (Bottom) Difference from the aircraft. The predicted error for daily observations is the observation error (27 ppb) divided by the square root of the number of observations. The predicted monthly or seasonal error is the mean daily error (11.5 ppb) divided by the square root of the number of days averaged.