Atmos. Chem. Phys. Discuss., 12, 6283–6329, 2012
www.atmos-chem-phys-discuss.net/12/6283/2012/
doi:10.5194/acpd-12-6283-2012
© Author(s) 2012. CC Attribution 3.0 License.

Atmospheric Chemistry and Physics Discussions

This discussion paper is/has been under review for the journal Atmospheric Chemistry and Physics (ACP). Please refer to the corresponding final paper in ACP if available.

# Comparison of improved Aura Tropospheric Emission Spectrometer (TES) $CO_2$ with HIPPO and SGP aircraft profile measurements

**S. S. Kulawik[1], J. R. Worden[1], S. C. Wofsy[2], S. C. Biraud[3], R. Nassar[4], D. B. A. Jones[5], E. T. Olsen[1], and G. B. Osterman, and the TES and HIPPO teams[1]**

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[2]Harvard University, Cambridge, MA 02138, USA
[3]Lawrence Berkeley National Laboratory, Earth Sciences Division, Berkeley, CA 94720, USA
[4]Environment Canada, Toronto, Ontario, Canada
[5]University of Toronto, Department of Physics, Toronto, Ontario, M5S 1A7, Canada

Received: 18 January 2012 – Accepted: 1 February 2012 – Published: 29 February 2012

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

Published by Copernicus Publications on behalf of the European Geosciences Union.

## Abstract

Comparisons are made between mid-tropospheric Tropospheric Emission Spectrometer (TES) carbon dioxide ($CO_2$) satellite measurements and ocean profiles from three Hiaper Pole-to-Pole Observations (HIPPO) campaigns and land aircraft profiles from the United States Southern Great Plains (SGP) Atmospheric Radiation Measurement (ARM) site over a 4-yr period. These comparisons are used to characterize the bias in the TES $CO_2$ estimates and to assess whether calculated and actual uncertainties and sensitivities are consistent. The HIPPO dataset is one of the few datasets spanning the altitude range where TES $CO_2$ estimates are sensitive, which is especially important for characterization of biases. We find that TES $CO_2$ estimates capture the seasonal and latitudinal gradients observed by HIPPO $CO_2$ measurements; actual errors range from 0.8–1.2 ppm, depending on the campaign, and are approximately 1.4 times larger than the predicted errors. The bias of TES versus HIPPO is within 0.85 ppm for each of the 3 campaigns; however several of the sub-tropical TES $CO_2$ estimates are lower than expected based on the calculated errors. Comparisons of aircraft flask profiles, which are measured from the surface to 5 km, to TES $CO_2$ at the SGP ARM site show good agreement with an overall bias of 0.1 ppm and rms of 1.0 ppm. We also find that the predicted sensitivity of the TES $CO_2$ estimates is too high, which results from using a multi-step retrieval for $CO_2$ and temperature. We find that the averaging kernel in the TES product corrected by a pressure-dependent factor accurately reflects the sensitivity of the TES $CO_2$ product.

## 1 Introduction

Over the past decade, measurements of carbon dioxide ($CO_2$) from space have become increasingly prevalent, with $CO_2$ measurements from SCIAMACHY, AIRS, TES, IASI, ACE, and GOSAT (e.g., Reuter et al., 2011; Chahine et al., 2008; Kulawik et al., 2010; Crevosier et al., 2009; Foucher et al., 2011; Yoshida et al., 2011; Crisp et al.,





ACPD
12, 6283–6329, 2012

2011; Butz et al., 2011). Because the variability of $CO_2$ is smaller than other trace gases, it is more critical to characterize errors in spectroscopy, calibration, atmospheric temperature and water, surface parameters, clouds and aerosol parameters for $CO_2$ than for more variable products. There is also a need to understand and validate biases and errors with great accuracy for the data to be useful for estimating $CO_2$ sources and sinks. Consistent validation and intercomparisons for satellite data, necessary for combining or utilizing multiple satellite results, are challenging since the different products have different coverage, vertical sensitivity, and averaging strategies (as summarized in Table 1). In this paper, we present comparisons of TES $CO_2$ to aircraft profile data from the HIPPO campaigns and from the Southern Great Plains ARM site to quantify errors, biases, and correlations between TES and the validation data. The techniques and methods shown in this paper are applicable to validation of other instruments with coincident aircraft profiles. Daytime validation profiles that go as high as possible over land (at least up to 10 km) would be ideal for the cross-validation of all $CO_2$ instruments because of the measurement characteristics shown in Table 1.

Multiple studies have estimated the precision and bias required to utilize atmospheric $CO_2$ measurements for source and sink estimates. Using simulated observations, Rayner and O'Brien (2001) showed that satellite measurements of $CO_2$ column abundances with a precision of 2.5 ppm, averaged monthly on spatial scales of $8^\circ \times 10^\circ$, would offer more information on $CO_2$ fluxes than can be obtained from the existing surface network. Houweling et al. (2004) also carried out simulations suggesting that latitude-dependent biases of less than 0.3 ppm are necessary for upper tropospheric $CO_2$ data to be useful for estimating sources and sinks. Recently, Nassar et al. (2011) showed that $5^\circ \times 5^\circ$ monthly-averaged TES observations at 500 hPa over ocean with mean errors of 4.7 ppm between 40° S and 40° N provided information that was complementary to flask data and especially helped constrain tropical land regions. Nassar et al. (2011) mitigated latitude and seasonally dependent biases of 1–2 ppm using 3 different correction methods to estimate sources and sinks from combined TES mid-tropospheric $CO_2$ and surface flask $CO_2$. Although the exact magnitude of regional





ACPD
12, 6283–6329, 2012

fluxes differed based on the bias correction approach used, key results are generally robust within the predicted errors. This indicates a strategy for assessing the robustness of flux estimates with spatially or temporally varying biases; however smaller biases would of course be perferable.

Kulawik et al. (2010) showed that the TES $CO_2$ prototype results compared well to aircraft data over Northern Hemisphere ocean sites but showed less reliable results over Southern Hemisphere ocean sites in some months and over land. The peak sensitivity of TES $CO_2$ was seen to be about 500 mb with sensitivity between approximately 40° S and 45° N. Based on the findings of Kulawik et al. (2010), updates were made to the retrieval strategy which significantly improved the accuracy of the TES $CO_2$ retrieval over land and changed the overall bias of TES $CO_2$ from a 1.8 % low bias to a 0.13 % high bias. Results with the new version, processed with the TES prototype, are shown in this paper. Much of the TES data record has been processed with the prototype, and the TES production code has recently started producing v005 products with this strategy, which will process the complete TES dataset by late 2012. The TES v005 $CO_2$ products are about 0.1 % higher than the prototype; an investigation showed that this results from a slight difference in the altitude grid calculation.

## 2 Measurements

### 2.1 The TES instrument

TES is on the Earth Observing System Aura (EOS-Aura) satellite and makes high spectral resolution nadir measurements in the thermal infrared (660 $cm^{-1}$–2260 $cm^{-1}$, with unapodized resolution of 0.06 $cm^{-1}$, apodized resolution of 0.1 $cm^{-1}$). TES was launched in July 2004 in a sun-synchronous orbit at an altitude of 705 km with an equatorial crossing time of 13:38 (local mean solar time) and with a repeat cycle of 16 days. In standard "global survey" mode, 2000–3000 observations are taken every other day (Beer, 2006). There are additional targeted "special observations", which are





ACPD
12, 6283–6329, 2012

not used in this analysis as they are less spatially and temporally uniform. For details on the TES instrument, see Beer (2006), and for information on the retrieval process see Bowman et al. (2006) and Kulawik et al. (2006, 2010).

### 2.2 HIPPO aircraft measurements

For validation of observations over oceans, we compare to the HIPPO 1, 2, and 3 campaigns (Wofsy, 2011) over the Pacific from 80° N to 70° S for January, 2009, November, 2009, and April 2010, respectively. The profiles are measured between 0.3 km and 10 km (with some extending up to 14 km), covering a large fraction of the TES vertical sensitivity with data traceable to WMO standards with an accuracy better than 0.1 ppm. For the TES latitude range of 40° S to 40° N the number of HIPPO-TES profile matches are 37, 64, and 29 for the HIPPO 1, 2, and 3 campaigns, respectively. (TES matching means averaging TES profiles coincident with the aircraft profiles within 10° longitude, 4° latitude, and 14 days). We use the HIPPO-identified profiles and the CO2_X field, based on 1 s data averaged to 10 s, from two (harmonized) sensors: CO2-QCLS and CO2-OMS. Note we do not use $CO_2$ profiles from HIPPO 1 flights 8–11, when the $CO_2$ instrumentation received a small fraction of air contaminated by the aircraft cabin. Contamination is assessed by comparison with flask samples, at levels of 2 ppm or more. Flight 7 possibly also had a small amount of contamination above 3000 m, but of less than 1 ppm. Changes to the aircraft sampling system were made after HIPPO-1 and no contamination was detected thereafter in the reported data.

### 2.3 SGP aircraft measurements

For validation of observations over land, we compare to aircraft flask measurements taken at the Southern Great Plains (SGP) Atmospheric Radiation Measurement (ARM) site with up to 12 measurements between 0.3 and 5.3 km altitude up to 8 times per month (Riley et al., 2009). Measurements with quality flag "..P" are used. This site is located in the Southern United States at 36.8° N, 97.5° W, and has data starting in





ACPD
12, 6283–6329, 2012

2002. Additionally, starting in late 2010, coincident aircraft measurements have been coordinated with TES stare observations consisting of up to 32 observations at the same ground location; the stare observations will be analyzed in a future paper.

### 2.4 CONTRAIL aircraft and AIRS satellite measurements

Because the SGP aircraft measurements cover only part of the altitude range of TES sensitivity to $CO_2$, we test extending these aircraft measurements with measurements from the Comprehensive Observation Network for TRace gases by AIrLiner (CONTRAIL) aircraft (Matsueda et al., 2002, 2008; Machida et al., 2008) or co-located Atmospheric Infrared Sounder (AIRS) $CO_2$ measurements, which have peak vertical sensitivity at ~9 km. The CONTRAIL measurements are between 9 and 11 km and are located over the western Pacific Ocean (between Japan and Australia); these are matched by latitude to the SGP site. For AIRS, the Level 3 calendar monthly v5 product was used with spatial averaging to match the TES spatial averaging.

## 3 Description of the TES $CO_2$ product

### 3.1 Retrieval strategy

#### 3.1.1 Updates from the previous version

The retrieval strategy for the TES $CO_2$ esimates was updated from Kulawik et al. (2010) to address issues found through validation of the prototype $CO_2$ data. The previous version compared well to validation data in Northern Hemisphere ocean, but less well to observations over land and in the Southern Hemisphere ocean. Observations over land showed a high bias and higher than expected rms differences compared with aircraft data, and observations over ocean in the Southern Hemisphere showed some latitudinal and seasonal biases (see Kulawik et al., 2010, Figs. 9, 10, and 12). One known issue in the TES retrieval is the spectroscopic inconsistency between the $CO_2$





ACPD
12, 6283–6329, 2012

$\nu_2$ and laser bands used for the $CO_2$ retrieval (Kulawik et al., 2010); consequently a retrieval using both bands simultaneously will have inconsistent biases depending on the relative weights of the two bands.

The laser bands yield the best results when temperature and water profiles are known, and the $\nu_2$ band is essential for constraining temperature and water. So, to address the need for the $\nu_2$ band, but to mitigate the effects of the inconsistent spectroscopy, a 2-step retrieval is used. In the first step, atmospheric temperature, water, ozone, carbon dioxide, surface temperature, cloud pressure, cloud optical depth, and emissivity (over land) are retrieved for windows covering both the $\nu_2$ and laser bands. This uses the 5-level $CO_2$ retrieval grid (suface, 511 hPa, 133 hPa, 10 hPa, 0.1 hPa). The 511 hPa result is biased low by about 6 ppm, with the surface result tending to be biased even more and the 133 hPa result tending to be biased less. Adding more retrieval levels to this step resulted in increased altitude-dependent biases. The second step retrieves only $CO_2$ and surface temperature in the 980 $cm^{-1}$ laser band keeping atmospheric temperature, water, etc. from Step 1 and using a 14-level retrieval vector for $CO_2$ (suface, 909, 681, 511, 383, 287, 215, 161, 121, 91, 51, 29, 4.6, 0.1 hPa). Because of overlapping sensitivity in the mid-troposphere between the $\nu_2$ and laser bands to $CO_2$, we recommend that users of this data set primarily use the level at 511 hPa, as other levels can show biases and errors likely due to inconsistencies between the bands.

We found that ozone has about 2 Degrees of Freedom (DOF) in the windows selected, almost exclusively in the stratosphere, and so we included ozone in the updated strategy. We also found that the 1080 $cm^{-1}$ laser band has a large silicate emissivity feature which has a significant impact on the radiance and affects land retrievals over arid regions, so this spectral region is not used in Step 2. We also found that extending the window used for the $\nu_2$ band from 671–725 $cm^{-1}$ to 660–775 $cm^{-1}$ improved results, likely due to increased sensitivity. Finally, we removed some spectral regions contaminated by minor interferent species, such as formic acid and formaldehyde, as well as some spectral regions with unknown but persistent residual features. Formic





ACPD
12, 6283–6329, 2012

acid and formaldehyde typically exist at very low levels in the atmosphere, but appear at significant concentrations in biomass burning plumes, which could lead to spatially dependent biases in $CO_2$ if included in the retrieval. The spectral ranges used for the two steps are shown in Table 2. The resulting strategy is implemented in the TES products for v5 data.

### 3.1.2 A priori and values and assumptions

The a priori covariance and the constraint for the 5-level $CO_2$ retrieval in Step 1 are described in Kulawik et al. (2010). The constraint for the 14-level $CO_2$ retrieval in Step 2 was created with a similar process as the 5-level constraint described in Kulawik et al. (2010) and was geared towards minimizing the error for a 10-target average. The TES radiative transfer forward model and spectroscopic parameters are the same as in Kulawik et al. (2010).

The TES initial guess and a priori states are taken from the chemical transport model MATCH (Nevison et al., 2008) used in conjunction with a variety of other models to provide $CO_2$ surface fluxes based on 2004 (D. Baker, personal communication, 2008). The key surface $CO_2$ fluxes are derived from models including the Carnegie-Ames-Stanford Approach (CASA) land biosphere model (Olsen and Randerson, 2004), ocean fluxes from the Wood's Hole Oceanographic Institute (WHOI) model (Moore et al., 2004) and a realistic, annually varying fossil fuel source scheme (Nevison et al., 2008). The $CO_2$ fields generated by the model compared well to GLOBALVIEW atmospheric $CO_2$ data (Osterman, TES Design File Memo). The initial guess and a priori are binned averages of the model for every 10° latitude and 180° longitude (i.e. 18 latitude bins and 2 longitude bins, 0–180° E, 180E-360° E). This binned monthly mean climatology for 2004 was then scaled upward yearly (by 1.0055) to best match the annual increase in $CO_2$.





ACPD
12, 6283–6329, 2012

### 3.2 Characterizing and validating TES errors and sensitivity

Predicted errors and sensitivity are important to characterize for application of the data, particularly when errors and sensitivity vary from target to target because of variability of clouds and surface properties. For error analysis and sensitivity characterization, the non-linear retrieval process is assumed to be represented by the linear estimate (e.g., Rodgers, 2000; Connor et al., 2008):

$$\mathbf{x}_{\mathrm{est}} = \mathbf{x}_{\mathrm{a}} + \mathbf{A}(\mathbf{x}_{\mathrm{true}} - \mathbf{x}_{\mathrm{a}}) + \mathbf{Gn} + \mathbf{GK}_{\mathrm{b}}\Delta\mathbf{b} \tag{1}$$

where $\mathbf{x}_{\mathrm{est}}$ is the log of the estimate, $\mathbf{A}$ is the averaging kernel (sensitivity of the retrieved state to the true state) $= \mathrm{d}x_{\mathrm{est}}/\mathrm{d}x_{\mathrm{true}}$ $\mathbf{x}_{\mathrm{a}}$ is the log of the a priori constraint vector, $\mathbf{x}_{\mathrm{true}}$ is the log of the true state, $\mathbf{G}$ is the gain matrix (sensitivity of the measurement to radiance errors), $\mathbf{n}$ is the radiance error vector, $\mathbf{K}_{\mathrm{b}}$ is the interferent Jacobian (sensitivity of the radiance to each interferent parameter), and $\Delta\mathbf{b}$ are the errors in the interferent parameters.

Note that for TES, all parameters besides temperature and emissivity are retrieved in log, so that the retrieved parameter is $\mathbf{x} = \log(\mathbf{VMR})$. Connor et al. (2008) further separates the retrieval vector, $\mathbf{x}$, into retrieved $CO_2$ parameters (here denoted $\mathbf{x}$) and all other jointly retrieved parameters (here donoted $\mathbf{y}$).

$$\mathbf{x}_{\mathrm{est}} = \mathbf{x}_{\mathrm{a}} + \mathbf{A}_{\mathrm{xx}}(\mathbf{x}_{\mathrm{true}} - \mathbf{x}_{\mathrm{a}}) + \mathbf{A}_{\mathrm{xy}}(\mathbf{y}_{\mathrm{true}} - \mathbf{y}_{\mathrm{a}}) + \mathbf{Gn} + \mathbf{GK}_{\mathrm{b}}\Delta\mathbf{b} \tag{2}$$

where $\mathbf{A}_{\mathrm{xx}}$ is the sub-block of the averaging kernel corresponding to the impact of $CO_2$ on the retrieved $CO_2$ parameters, and the $\mathbf{A}_{\mathrm{xy}}$ is the sub-block of the averaging kernel corresponding to the impact of non-$CO_2$ parameters on retrieved $CO_2$.

Subtracting $\mathbf{x}_{\mathrm{true}}$ from the left and right side of Eq. (2) and taking the covariance gives the predicted error covariance:

$$\mathbf{S}_{\mathrm{err}} = \underbrace{\mathbf{GS}_{\mathrm{m}}\mathbf{G}^T}_{\text{Measurement}} + \underbrace{\mathbf{GK_bS}_{\mathrm{berr}}(\mathbf{GK})^T}_{\text{Interferent}} + \underbrace{(1-\mathbf{A}_{\mathrm{xx}})\mathbf{S}_{\mathrm{a,xx}}(1-\mathbf{A}_{\mathrm{xx}})^T}_{\text{Smoothing}} + \underbrace{\mathbf{A}_{\mathrm{xy}}\mathbf{S}_{\mathrm{a,yy}}(1-\mathbf{A}_{\mathrm{xy}})^T}_{\text{Cross state}} \tag{3}$$





ACPD
12, 6283–6329, 2012

where $\mathbf{S}_{err}$ is the total error covariance, $\mathbf{S}_m$ is the covariance of the radiance error, and $\mathbf{S}_a$ is the a priori covariance. The cross state error (described in Worden et al., 2004; Connor et al., 2008) is the $CO_2$ error resulting from jointly retrieved species, and the smoothing error results from the effects of the constraint matrix.

The cross-state component is due to the propagation of error from jointly retrieved species into $CO_2$; in this case, surface temperature. This error should decrease with target averaging over regional, monthly scales, as the surface temperature error will likely vary in sign and magnitude. Similarly, interferent and measurement errors should also decrease with target averaging over regional, monthly scales. However, averaging targets with the same $CO_2$ true state results in a bias for the smoothing term which does not decrease with averaging. The predicted total error covariance for an $n$ target average is:

$$\mathbf{S}_{err} = (\mathbf{S}_{meas} + \mathbf{S}_{int} + \mathbf{S}_{cross\text{-}state})/n + \mathbf{S}_{smooth}$$
$$\mathbf{S}_{err} = \mathbf{S}_{obs}/n + \mathbf{S}_{smooth} \quad (4)$$

The observation error ($\mathbf{S}_{obs}$) and smoothing error covariances in Eq. (4) are included in the TES products (Osterman et al., 2009). The predicted error for a particular level is the square-root diagonal of the predicted error covariance at that level, and the off-diagonal terms describe correlated errors between levels. Spectroscopic and calibration errors, which may contribute an additional bias and/or random error, are not included in Eq. (2), but could be added in, if known, as the gain matrix multiplied by the radiance error.

We validate with aircraft profile data, where the true state, $\mathbf{x}_{true}$, is known for at least portions of the atmosphere. To construct $\mathbf{x}_{true}$ on the TES pressure levels, the following steps are taken: (1) interpolate/extrapolate the aircraft profile to the 65-level TES pressure grid; (2) replace values below all aircraft measurements with the lowest altitude aircraft measurement value; (3) replace values above all aircraft measurements with the highest altitude aircraft measurement value. We then apply the "Observation





ACPD
12, 6283–6329, 2012

operator" to this profile to assess the effects of TES sensitivity:

$$\mathbf{x}_{\text{pred}} = \mathbf{x}_{\text{a}} + \mathbf{A}_{\text{xx}}(\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{a}}) \tag{5}$$

$\mathbf{x}_{\text{pred}}$ is what TES would see if it observed the airmass described by the aircraft profile in the absense of any other errors due to the vertical resolution and sensitivity of the TES instrument. Since we have applied the TES sensitivity to the aircraft profile, there is no smoothing error term when comparing $\mathbf{x}_{\text{pred}}$ and $\mathbf{x}_{\text{tes}}$. The predicted error for $\mathbf{x}_{\text{tes}}$ compared to $\mathbf{x}_{\text{pred}}$ is the observation error, which is significantly smaller than the smoothing error when averaging over ~ 40 profiles.

The SGP aircraft data go up to ~ 5 km, covering only part of the range of TES sensitivity and so the choice of the value for $\mathbf{x}_{\text{true}}$ above 5 km could have an impact on $\mathbf{x}_{\text{pred}}$. We set $\mathbf{x}_{\text{true}}$ to carbon dioxide values either from AIRS, CONTRAIL, or the highest altitude aircraft measurement; the differences in these results characterize the size of the uncertainty introduced from uncertainty in the true profile.

The predicted sensitivity and retrieval non-linearity can be validated, as described in Kulawik et al. (2008, 2010), by running non-linear retrievals using two different a priori vectors, $\mathbf{x}_{\text{a}}$ and $\mathbf{x}'_{\text{a}}$, resulting in the non-linear retrievals, $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}}'$, respectively. $\hat{\mathbf{x}}'$ is then converted via a linear transformation to use $\mathbf{x}_{\text{a}}$ using the following linear equation:

$$\hat{\mathbf{x}}_{\text{est}} = \hat{\mathbf{x}}' + \mathbf{A}(\mathbf{x}_{\text{a}} - \mathbf{x}'_{\text{a}}) \tag{6}$$

where recall $\mathbf{x}$ is log(VMR). $\hat{\mathbf{x}}_{\text{est}}$ from Eq. (6) is compared to $\hat{\mathbf{x}}$. If they compare within the predicted errors, it validates both the predicted sensitivity and the non-linearity of the system. The comparison between $\hat{\mathbf{x}}_{\text{est}}$ and $\hat{\mathbf{x}}$ answer two questions: (1) how sensitive are the results to the starting point of the retrieval? and (2) can we use the sensitivity to predict the results we expect to see?

### 3.3 Predicted errors for TES $CO_2$

Figure 1 shows the predicted errors for a single target and for a 40-target average for land and ocean scenes. At 500 hPa (about 5 km), the dominant error source for a single





ACPD
12, 6283–6329, 2012



CC BY

target is interferent error, at about 4–7 ppm, due to, in order of importance, temperature, cloud parameters, water, and ozone. Measurement error is also signficant, contributing nearly 4 ppm, followed by the smoothing error, which contributes about 1.5 ppm. Errors from the jointly retrieved surface temperature are small. The total error is about 6–7 ppm for a single target. However, when 40 targets are averaged, the interferent and measurement errors are taken to be quasi-random, and are reduced by 1/sqrt(40). The dominant error for the 40-target average is the smoothing error, resulting from imperfect sensitivity. Land targets in general have higher interferent error due to the uncertainty in emissivity.

### 3.4 Predicted sensitivity for TES $CO_2$

The predicted sensitivity is captured in the averaging kernel (AK), shown in Fig. 2. The left panel shows the AK for all levels for Step 1, which includes the joint retrieval of all interferents in both the $\nu_2$ and the laser band spectral regions. This AK shows the potential for resolving $CO_2$ at different pressure levels once the spectroscopy is addressed. The middle panel shows the predicted AK for the final $CO_2$ step. Note that all levels have very similar sensitivity but with more predicted sensitivity than the first step, mainly because the second step only retrieves $CO_2$ and surface temperature in a narrow spectral range. The right panel compares the averaging kernel row at 511 hPa for TES observations matching the HIPPO campaigns and observations near SGP. Note that the SGP-matched targets, over land, show more variability in the sensitivity because of seasonal and day/night variations in surface temperature. The averaging kernel on the far right panel of Fig. 2 has been corrected by a pressure dependent factor, shown in Table 3, to reflect the actual sensitivity (see Appendix for a description of how the averaging kernel was validated and the pressure dependent factor was calculated). We find that this ratio is very similar for all pressure levels (results not shown), so that this ratio can be used for all retrieval pressures. All remaining results in this paper have this factor applied to the predicted averaging kernel.





ACPD
12, 6283–6329, 2012

### 3.5 Bias correction for TES $CO_2$

TES has a small positive bias, of about 0.4 ppm, or 0.13 %. The bias correction factor was set by comparisons of TES data processed through the prototype retrieval to the HIPPO aircraft data using older revisions of both datasets. This value is difficult to calculate precisely because of the uncertainty introduced by errors, comparison thresholds, quality flags, and minor processing updates. The correction is carried out using the equation $\mathbf{x}_{\mathrm{corr},i} = \mathbf{x}_{\mathrm{raw},i} + \mathbf{A}_{ij}\mathbf{bias}_j\mathbf{x}_{\mathrm{raw},i}$, as discussed in Kulawik et al. (2010), where $\mathbf{bias}_j = -0.0013$ for all $j$. Here, $\mathbf{x}_{\mathrm{raw}}$ and $\mathbf{x}_{\mathrm{corr}}$ are the retrieved and corrected VMR values, respectively (not log(VMR)) for $CO_2$. All results shown in this paper have been corrected by this factor. Examination of the initial output from the TES production code for v5 finds no bias correction is needed for TES production code results; the final bias for TES v005 production results will be set when there is more complete output.

## 4 Actual and predicted errors compared with HIPPO and SGP

Figure 3 shows a plot of the matching locations for TES and HIPPO 1, HIPPO 2, HIPPO 3, and SGP. For HIPPO coincidences, TES results within 2 weeks, 10° longitude and 4° latitude are averaged for comparison with each HIPPO profile. Additionally, the mean time for the TES observations must be within 7 days of each HIPPO profile. This criteria only affects HIPPO-3, as TES was not taking measurements for part of this campaign. For SGP comparisons, TES and SGP data are both averaged within each month, and TES is averaged within 10° longitude and 5° latitude of the SGP observations.

### 4.1 Comparison of TES and HIPPO measurements

Figure 4 shows the comparisons between TES and HIPPO. The left panels show curtain plots of the HIPPO results with the TES result plotted at the altitude of TES peak sensitivity. The right panels show the comparisons between the TES result and the





ACPD
12, 6283–6329, 2012

HIPPO profiles with the TES observation operator applied (Eq. 5) versus latitude. Note that the green dashed line shows the a priori, which is too high in the Southern Hemisphere. The larger errors in the TES averages in April 2010 reflect the lower number of measurements made by TES during this time period. The average number of TES observations that are averaged for each HIPPO observation are 64, 60, and 28 for HIPPO 1, 2, and 3, respectively. In general, TES compares well to HIPPO, showing similar patterns in each time period. However, there are locations and times that show persistent errors larger than the predicted errors, most notably at $\sim 15^{\circ}$ S and $\sim 10^{\circ}$ N most obviously in HIPPO 2, but also seen in HIPPO 3. A histogram of the values composing the TES averages for the bad TES points (not shown) shows that the entire distribution of points is shifted lower, rather than a few outliers causing the low values. The correlation of errors in a particular region and preliminary analysis of the TES "Stare" observations at SGP indicates that likely these outliers result from a bias in the interferent errors, rather than the assumed quasi-random distribution of interferent error. Since averaging does not reduce a biased error, the error for the averaged product would be comparable to the single-target interferent error of 4–5 ppm. These differences appear in the subtropics at $\pm 10^{\circ}$ which is a region where there are stronger gradients in humidity and temperature; consequently, care should be taken when using TES data in these regions until more robust quality flags or retrievals are available.

Figure 5 shows the TES/HIPPO comparisons in the context of the overall patterns seen by TES. In Fig. 5b, the low TES values, compared HIPPO 2, in Fig. 4 at $\sim 10^{\circ}$ S and $\sim 15^{\circ}$ N can be seen as part of a larger spatial pattern seen by TES. Looking at the other TES retrieved values, a similar pattern can be seen in TES ozone, water, and HDO at 681 mb for November, 2009 (http://tes.jpl.nasa.gov/visualization/SCIENCE_PLOTS/TES_L3_Monthly.htm). As this pattern is persistent in TES from year to year (data not shown) but is not seen with the HIPPO data, it most likely indicates a problem in the retrieved TES $CO_2$ at these locations.

To validate the predicted sensitivity, runs were also performed for HIPPO comparisons with a uniform 385 ppm a priori and initial guess. We compare the difference





ACPD
12, 6283–6329, 2012

between the results obtained with the fixed 385 ppm prior to the variable prior results (as seen in Fig. 4), which are then linearly converted to a uniform 385 ppm prior via Eq. (6). When the differences are smaller or comparable to the observation error the sensitivity, as described by the averaging kernel, is validated. For HIPPO 1, the TES-TES comparisons (TES results with a variable prior converted to a fixed prior via Eq. (6), versus TES results with a fixed prior) have a 0.02 ppm bias and 0.16 ppm standard deviation compared to observation error of 0.8 ppm. For HIPPO 2, the TES-TES comparisons have a −0.03 ppm bias and 0.34 ppm standard deviation compared to an observation error of 0.6 ppm. For HIPPO 3, the TES-TES comparisons have a −0.45 ppm bias and 1.3 ppm standard deviation compared to an observation error of 0.9 ppm. The bias for all cases is less than the observation error, and in 2 of the 3 cases, the rms difference is less than the observation error. In HIPPO-3, the rms is 0.4 ppm larger than the observation error. This comparison validates the predicted sensitivity and linearity of the retrieval system for ocean targets.

Figure 6 presents a scatter plot between TES averages and associated HIPPO measurements, showing statistical comparisons nd predicted and actual errors. The TES biases for the HIPPO 1, 2, and 3 campaigns, respectively are $-0.85 \pm 0.10$, $-0.06 \pm 0.08$, and $-0.21 \pm 0.18$ ppm. HIPPO 1 shows a TES bias signficantly larger than the predicted error in the bias, which is the predicted observation error/sqrt(# of measurements). The actual errors are also larger than the predicted, by an average factor of 1.4, likely because the interferent errors are at least somewhat correlated, rather than random, for the averaged targets and/or the strong interference between $H_2O$, temperature, and $CO_2$ leads to unquantified non-linearity in the $CO_2$ retrieval (Boxe et al., 2010). Consistency between predicted and actual errors is critical for the scientific use of the data, especially data assimilation or $CO_2$ flux estimates.

#### 4.1.1 Correlations between TES and HIPPO

Correlations between HIPPO and TES are shown in Table 4. Because the coincidences are not perfect, and are therefore subject to the natural variability of the atmosphere





ACPD
12, 6283–6329, 2012

within the coincedence region and time, a correlation of 1.00 is not to be expected. This section also shows how the correlation degrades when the error is comparable or larger than the variability. The correlation between $x$ and $y$ (where $x$ and $y$ have mean of 0) is defined as:

$$c_o = \frac{\mathbf{x}\cdot\mathbf{y}}{\sqrt{\mathbf{x}\cdot\mathbf{x}}\sqrt{\mathbf{y}\cdot\mathbf{y}}} \tag{7}$$

Adding in errors for **x** and assuming that the errors are uncorrelated with **x** or **y**, the correlation $c$ is:

$$c = \frac{\mathbf{x}\cdot\mathbf{y}}{\sqrt{\mathbf{x}\cdot\mathbf{x}+\varepsilon_x\cdot\varepsilon_x}\sqrt{\mathbf{y}\cdot\mathbf{y}}} = c_o\frac{1}{\sqrt{1+\varepsilon_x^2/\sigma_x^2}} \tag{8a}$$

$$c_o = c\sqrt{1+\varepsilon_x^2/\sigma_x^2} \tag{8b}$$

where the variability of **x** is denoted $\sigma_x$ and the error in **x** is denoted $\varepsilon_x$. From Eq. (8), it is apparent that errors that are equal to or larger than the variability will significantly degrade the observed correlations; for a more detailed discussion of how errors affect correlations, see Zhang et al. (2008). Using the predicted errors, variability, and observed correlations, and using Eq. (8b), we can calculate the underlying correlation in the absence of error, with results shown in Table 4. The raw correlations range from 0.56–0.80, and the "error free" correlations range from 0.7–0.95 when the predicted error is used in Eq. (8), and 0.93–1.0 when the actual error is used in Eq. (8). We find that the correlation is lower, as expected, for the HIPPO comparisons which have a lower variability/error ratio, and when the actual error is accounted for, we find correlations of 0.9–1.0 between TES and HIPPO.

#### 4.1.2 Coincidence criteria and effect on errors

Differences between the air parcels measured by HIPPO and those measured by TES will impart an error in the comparison between these two data sets. We estimate this





ACPD
12, 6283–6329, 2012

error as follows. The averaging of the TES data shown in Fig. 3 is within 4° latitude, 10° longitude, and 2 weeks of the HIPPO data. Because of the range of spatio-temporal locations, we expect that most interferent errors contribute quasi-randomly to the total error budget, with errors scaling as 1/sqrt(# observations). Table 5a shows the effects of averaging within 5°, 10°, or 15° longitude (keeping the latitude and time coincidence specified as above). The actual errors are approximately ~40 % larger than predicted for most cases, with the actual errors scaling sub-linearly with 1/sqrt(# observations). The correlations, predicted errors, and actual errors show consistent improvement between 5° and 10° averaging, but between 10° and 15° there is no improvement in correlation for HIPPO 1 and 3, and little improvement in the actual errors. The longitudinal offset comparisons (Sect. 4.1.2) shows that offsetting between the HIPPO and TES $CO_2$ by 15° results in more erratic biases, marginally higher errors and lower correlations.

Table 5b shows the results for averaging of the TES data within 7, 14, and 21 days; 2°, 4°, and 6° latitude; and cloud cutoffs of 0.1 or 0.5 optical depth. Results shown are averages of the 3 HIPPO results. The predicted errors scale according to 1/sqrt(# observations). Similar to the longitudinal conclusions, there is improvement between the tight and medium criteria, particularly in the predicted and actual errors. Although there is improvement in the correlations for the loose criteria, the actual errors do not improve.. Based on Table 5b, we recommend using the medium criteria when utilizing TES $CO_2$ for plotting or comparisons to other datasets.

#### 4.1.3 Logitudinal offset and effect on errors

TES $CO_2$ shows more longitudinal variability than models (e.g., Kulawik et al., 2010; Nassar et al., 2011). There is still some uncertainty as to how much of the longitudinal variability seen by TES represents locally correlated errors versus true variability. In Kulawik et al. (2010), TES was offset by longitude and compared to ocean surface sites and TES was shown to correlate best when the longitude was not offset. In Table 6 we show results when TES is offset by −15°, 0°, and +15° prior to co-locating





ACPD
12, 6283–6329, 2012

to the HIPPO data. Note that the −15° shift resulted in significantly more matches, likely a result of fewer cloudy scenes or more targets, and this affects the predicted errors which go with 1/sqrt(# observations). To gauge the quality of the match, we looked at (1) correlations, (2) actual error relative to the predicted error, and (3) biases. For all shift cases, the biases are farther from zero, and on average the actual errors and correlations are worse for the shift cases, but it is not worse for all cases. From Table 6, TES overall agrees best with HIPPO when it is aligned, rather than offset by 15 degrees in longitude. This indicates that at least some of the longitudinal variability observed by TES reflects variability in the underlying true state.

## 4.2 Comparison to aircraft data from the Southern Great Plains (SGP) ARM site

For comparisons between TES and SGP aircraft profile data, both datasets are monthly averaged, and TES is also averaged within 5° latitude, and 10° longitude of the SGP site (see Fig. 3). On the plots, the average of all aircraft data above 2 km is shown in orange labeled "ave SGP" (e.g., Fig. 7a,b). Aircraft profiles with the TES observation operator applied are shown in green labeled "SGP w/obs".

### 4.2.1 Effects of the validation profile above 5 km

As shown in Fig. 2, TES has significant sensitivity from 1–10 km. Since the aircraft profiles range between the surface to ∼ 5 km, we test three methods for extending the aircraft profiles to the upper range of TES sensitivity (1) extend the top aircraft value upwards indefinitely, (2) interpolate from the top SGP value to the AIRS value at 9 km, (3) interpolate from the top SGP value to CONTRAIL value at ∼ 10 km. Figure 8 shows results for the first two methods. Although the results are similar, extension of the SGP aircraft data with AIRS $CO_2$ values improves the bias from 0.34 to 0.13, and slightly improves the rms from 0.97 to 0.96 ppm. The use of CONTRAIL aircraft data to extend the SGP profile improves the rms from 0.97 to 0.96 and increases the bias slightly from 0.34 to 0.40. Note that the CONTRAIL data are flask measurements taken in





ACPD
12, 6283–6329, 2012

the Pacific matched by latitude to the SGP latitude and are not co-located with the SGP observations. As seen in Fig. 8, all datasets show similar seasonal cycles and yearly increases, with the amplitude on the AIRS cycle somewhat less and with the CONTRAIL data averaging somewhat lower (again note that the CONTRAIL data are at a different longitude).

Since extending SGP with AIRS above the SGP measurements gives somewhat better results, this method will be used to extend the SGP data for the remainder of this paper. This study, however, shows that missing validation data above 5 km results in at least a 0.01 ppm rms error and a bias uncertainty on the order of 0.3 ppm for the validation data.

#### 4.2.2 Results for different a priori and initial guesses

In this next section we evaluate TES $CO_2$ using both the standard TES prior, and using a constant prior, i.e. without any a priori knowledge of the $CO_2$ values. We compare results using these two different priors, linearly transforming the variable prior results to use the constant prior using Eq. (6), to determine whether the TES $CO_2$ retrieval strategy is linear and that the predicted sensitivity is correct, as well as verify that TES can capture the seasonal and yearly trends in the absence of a priori knowledge of $CO_2$.

Figure 7 shows time trend comparisons between TES and SGP aircraft measurements. The top two plots show the monthly averages of SGP data above 2 km and the SGP data with the TES observation operator with two a priori choices. The constant a priori choice will dampen the expected results, though in a predictable manner. Since the sum of the row of the 511 hPa averaging kernel for SGP averages about 0.65, about 2/3 of the variability should be captured when using a fixed prior.

The lower 3 panels of Fig. 7 show results when TES is started at the standard TES initial guess and prior (Fig. 7c), when TES is converted to a fixed prior after retrievals using Eq. (6) (Fig. 7d), and when TES is started at a uniform initial guess and prior (Fig. 7e). To validate sensitivity and retrieval non-linearity it is important that the results





ACPD
12, 6283–6329, 2012

in Fig. 7d,e agree, as discussed in Sect. 3.4. Comparing TES and the validation data, TES shows expected seasonal and yearly patterns over the four years of comparisons, both when TES is started at a "good" initial guess and prior, and when TES uses a uniform initial guess and prior for $CO_2$ which gives the TES retrieval system no a priori knowledge of $CO_2$. As seen in Table 7, the correlation between TES and the aircraft is 0.96 when a variable prior is used, 0.89 when the results are linearly converted to a constant prior using Eq. (6), and 0.88 when a constant prior is used for the non-linear retrievals. The similarity between the last two results indicates that the TES $CO_2$ retrieval strategy is predictably linear and that the predicted sensitivity is correct. The underlying correlations of these results are consistent when the errors and variability are considered, as seen in Table 4. If the error is taken as the actual, rather than predicted error, the underlying correlation between TES and SGP is $\sim 1$ (using Eq. 8).

The 4th line in Table 7 shows TES converted to a constant prior, using Eq. (6). The reduction in the correlation from 0.96 to 0.89 is a result of the reduced variability in the SGP data with the observation operator when a constant prior is used (see discussion following Eq. 8). The 5th line in Table 7 shows the non-linear retrievals using the constant prior which has a comparable correlation of 0.88, and slightly worse standard deviation (1.01 vs. 0.96 ppm). Finally, a straight comparison between TES and SGP *without* the TES observation operator shows a degraded correlation with significantly higher errors and bias. This shows the importance of properly accounting for the TES sensitivity using the TES observation operator in applications of the TES $CO_2$ data.

There is significant improvement over the previous data version (described in Kulawik et al., 2010): the previous data version (with a 385 ppm prior) showed a correlation of 0.7, actual rms errors of 1.5 ppm, and a bias of 2.50 ppm, the current results (for the same uniform prior) show a correlation of 0.88, actual rms errors of 1.01, and a bias of 0.01 ppm. From this analysis, we expect that the land data in this version are well-characterized with respect to the validation data and are therefore sufficiently reliable for use in scientific analyses.





ACPD
12, 6283–6329, 2012

### 4.3 Comparisons at additional pressure levels

In addition to the 511 hPa level, we also validate pressure levels 681 hPa and 908 hPa versus SGP, and these plus the 383 hPa level for HIPPO. SGP is not validated at 383 hPa because the SGP data is collected up through about 5 km (500 hPa). As seen in Table 8, the validation data with the TES observation operator applied (column "Obs. Var.") shows the most variability for the 681 hPa level. This is a combination of the degrees of freedom and the larger variability of the true state near the surface. Similarly, the 383 hPa level shows significantly smaller variability. All levels show skill, however, comparing the TES averaged results for 511 and 908 hPa, we find that the correlation is about 0.95, and rms difference varying between 0.3 and 0.7 ppm, and bias difference varying between −0.1 and 0.9 ppm for the different datasets. Because of the similarity of the TES values at the different pressure levels and because of the similarities in the averaging kernel as seen in the middle panel of Fig. 2, it is unclear how much independent information is added by the additional levels. However, all levels show skill compared with validation data and can be used for scientific analysis.

## 5 Conclusions

TES $CO_2$ estimates capture the latitudinal gradients and seasonal patterns found in the HIPPO and SGP aircraft data. The comparison with HIPPO and SGP data reveal biases $\leq 0.85$ ppm, errors for monthly-averaged data on the order of 0.8–1.2 ppm, using an overall bias correction of 0.13 %. Improvements from the previous TES $CO_2$ product are remarkable over land, and both land and ocean data for all pressure levels in this version of TES $CO_2$ can be used for scientific analyses. Comparisons of averaged TES to both HIPPO and SGP aircraft profile data show the actual errors averaging $\sim 1.4$ times the predicted errors for monthly averages for ocean validation and $\sim 1.7$ times the predicted errors for land validation. We find some locations and times with persistent biases on the order of the single-target predicted errors, which indicate areas





ACPD
12, 6283–6329, 2012



CC BY

where TES can be improved or needs to be flagged as bad. The biases for TES for the different HIPPO datasets were −0.85, −0.05, and −0.21 ppm, with the predicted errors for the bias to be 0.1–0.2 ppm. For the SGP dataset, the bias over a 4-yr period is 0.34, with a predicted bias error of 0.15 ppm. Because the SGP aircraft data only go up to 5 km, we find that using AIRS values above 5 km improves the TES bias to 0.13 and very slightly improves the rms error. The correlations between TES and validation data are 0.8, 0.6, and 0.9 for HIPPO and 0.9 for SGP comparisons. For the HIPPO comparisons, we show that the lower correlations can be explained by a predictable degradation of the correlation due to the error/variability ratio and the observed correlations are consistent with underlying correlations of 0.9–1.0.

In the appendix, we find that the senstivity reported in the TES products over-predicts the actual sensitivity because of the 2-step retrieval strategy that is used for the $CO_2$ estimates. For the current product release, we find that a pressure-dependent multiplicative factor applied to the sensitivity in the TES product results in an accurate prediction for the TES sensivitity to the vertical distribution of $CO_2$. The TES sensivity peaks at ~500 hPa with some sensitivity in the upper troposphere. On average, ocean targets show greater sensitivity than land targets, but both capture seasonal and yearly cycles in $CO_2$. We validate the sensivity by comparing results with a very good initial guess and prior compared with a fixed initial guess and prior. Both show the same seasonal and yearly patterns and agree within the observation error when converted to use the same prior using a linear transform.

Averaging TES within 5° latitude, 10° longitude, and 14 days gives the best results when comparing to validation data, considering correlations, errors, and biases. Averaging over larger times or distances does not significantly improve results and averaging over smaller times or distances results in higher predicted and actual errors. For assimilation, averaging over smaller areas, or no averaging, is fine, however, because the assimilation scheme accounts for the tradeoff between number of observations and error (e.g., see Kulawik et al., 2010). When averaging over the above times and distances the actual errors for comparisons to validation data are 0.8–1.2 ppm, with the





ACPD
12, 6283–6329, 2012

predicted errors about a factor of 1.4–1.7 too small for averaged data. Offsetting results by longitude degrades results, particularly for the bias, indicating that at least some of the longitudinal variability observed by TES represents true variations in $CO_2$.

## Appendix A

### Accurate sensitivity calculation for the multi-step retrieval

Careful characterization of the uncertainties and sensitivity of the $CO_2$ estimates are critical for using these data in scientific analysis. We find through comparisons of the TES $CO_2$ retrievals using different a priori vectors that the calculated sensitivity is inconsistent with the actual sensitivity. We calculate the averaging kernel theoretically by propagating perturbations in the true state at every level through the retrieval system and show how effects from retrieved parameters from previous steps need to be included in the calculated Averaging Kernel.

Considering a single-step retrieval, the effect of a perturbation in the true state on the retrieved state is captured by the following steps:

1. $CO_2$ is changed by a small amount at a particular pressure level, $\Delta\mathbf{x}$
2. This results in a change in the radiance, which is the Jacobian multiplied by the state change: $\Delta\mathbf{L} = \mathbf{K}^x \Delta\mathbf{x}$
3. When the perturbed radiance is used in a retrieval, it results in a change in the retrieved $CO_2$: $\Delta\hat{\mathbf{x}} = \mathbf{G}^x \Delta\mathbf{L}$
4. The AK is defined as $\Delta\hat{\mathbf{x}}/\Delta\mathbf{x}$, and this is:

$$\Delta\hat{\mathbf{x}}/\Delta\mathbf{x} = \mathbf{G}^x \mathbf{K}_x = \mathbf{A} \tag{A1}$$

**A** in Eq. (A1) is the standard equation for the averaging kernel. However, since in our case $CO_2$ is retrieved in 2 steps, the radiance perturbation first passes through step 1





ACPD
12, 6283–6329, 2012

of the retrieval. The theoretical averaging kernel resulting from the two-step retrieval, using a similar analysis to the above, but propagating through both steps, is:

$$\Delta\hat{\mathbf{x}}/\Delta\mathbf{x} = \mathbf{G}_2^x\mathbf{K}_2^x - \mathbf{G}_2^x\mathbf{K}_2^y\mathbf{G}_1^y\mathbf{K}_1^x \tag{A2}$$

where $\mathbf{G}_2^x$ is the gain matrix for $CO_2$ for step 2, $\mathbf{K}_2^x$ is the Jacobian for $CO_2$ for Step 2, $\mathbf{K}_2^y$ is the Jacobian for all non-retrieved and interferent parameters in the step 2 windows, $\mathbf{G}_1^y$ is the gain matrix for all non-$CO_2$ parameters for step 1, and $\mathbf{K}_1^x$ is the Jacobian for $CO_2$ parameters for step 1. The first term is the standard averaging kernel predicted for Step 2, as seen above, and is the AK included in the TES products, and the second term is a result of both non-$CO_2$ parameters and $CO_2$ parameters being active in both step windows. If $CO_2$ were not active in the Step 1 windows, then $\mathbf{K}_1^x$ would be zero, or if temperature, for example, had no influence in the Step 2 windows, then $\mathbf{K}_2^y$ would be zero.

We calculate the averaging kernel both from Eq. (A2) and from using perturbations in the $CO_2$ state vector propagated through the full non-linear least squares (NLLS) retrieval system. For the latter, a retrieval is performed on noise-free simulated radiances with the initial and priori state both set to the true state. Following this, the true state $CO_2$ at a single pressure is increased by 0.08 %, a new radiance is calculated, and a new retrieval is performed. The change in the retrieved state at 511 hPa divided by the change in the true state at each pressure gives the actual 511 hPa row of the averaging kernel. Different perturbation strengths were tried, resulting in the same AK results. Both Eq. (7) and the perturbation approach show enhanced $CO_2$ sensitivity higher in the atmosphere than predicted in the Step 2-only averaging kernel (Eq. A1). Figure A1 shows the TES products AK and actual AK (using retrievals on perturbed data) for both steps. Since the TES products do not contain the information to calculate the full AK using Eq. (A2), the ratio between the actual and predicted sensitivity is calculated as a function of pressure. This scale factor, shown in Fig. A1b and Table 3, is applied to correct the calculated averaging kernel. This corrected averaging kernel is used and validated in this paper.





ACPD
12, 6283–6329, 2012

We compare retrievals using the methodology described near Eq. (6) of the main paper. For our comparisons, $\mathbf{x}_a{}'$ was set to the TES operational a priori value, and $\mathbf{x}_a$ was set to a constant 385 ppm value. Note that the initial guess is also set to the same a priori value – the initial guess should not matter but, depending on the non-linearity, can influence the answer. So the constant 385 ppm retrievals contain no prior knowledge of $CO_2$ in either the initial state or the prior.

The left panel of Fig. A2 shows a histogram of the difference between $\mathbf{x}_a{}'$ and $\mathbf{x}_a$ (green), the difference between $\hat{\mathbf{x}}_{est}$ and $\hat{\mathbf{x}}$ (black), and the predicted total error (dashed) for monthly average TES values near SGP. The prior distribution has a standard deviation of 3.6 ppm and a bias of −0.86 ppm, and the final difference has a standard deviation of 0.49 ppm and a bias of +0.11 ppm. For comparison, the observation error is 0.5 ppm for these monthly averages. The time series panel (Fig. A2) shows that even when the initial guess and prior are set to a constant value, with no a priori knowledge of the $CO_2$ concentrations, the correct seasonal and yearly cycles are seen with the predicted sensitivity. When this same test was performed with the *original* averaging kernels in the TES products, the analysis showed that the sensitivity was over-predicted because differences between the red and black lines were correlated with the differences between the the variable prior and the fixed prior. This correlation indicated that the predicted averaging kernels were not correct, leading to the correction discussed in this section.

*Acknowledgement.* CONTRAIL data were obtained from Toshinobu Machida at the National Institute for Environmental Studies, Tsukuba, Japan.

SGP data was supported by the Office of Biological and Environmental Research of the US Department of Energy under contract No. DE-AC02-05CH11231 as part of the Atmospheric Radiation Measurement Program.

Work at the Jet Propulsion Laboratory, California Institute of Technology, was performed under a contract with the National Aeronautics and Space Administration and funded through NASA ACMAP, 2007.

## References


Beer, R.: TES on the Aura mission: Scientific objectives, measurements, and analysis overview, IEEE T. Geosci. Remote, 44, 1102–1105, 2006.

Bowman, K. W., Rodgers, C. D., Kulawik, S. S., Worden, J., Sarkissian, E., Osterman, G., Steck, T., Lou, M., Eldering, A., Shephard, M., Worden, H., Lampel, M., Clough, S., Brown, P., Rinsland, C., Gunson, M., and Beer, R.: Tropospheric emission spectrometer: retrieval method and error analysis, IEEE T. Geosci. Remote, 44, 1297–1307, 2006.

Boxe, C. S., Worden, J. R., Bowman, K. W., Kulawik, S. S., Neu, J. L., Ford, W. C., Osterman, G. B., Herman, R. L., Eldering, A., Tarasick, D. W., Thompson, A. M., Doughty, D. C., Hoffmann, M. R., and Oltmans, S. J.: Validation of northern latitude Tropospheric Emission Spectrometer stare ozone profiles with ARC-IONS sondes during ARCTAS: sensitivity, bias and error analysis, Atmos. Chem. Phys., 10, 9901–9914, doi:10.5194/acp-10-9901-2010, 2010.

Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., Kuze, A., Keppel-Aleks, G., Toon, G., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., and Warneke, T.: Toward accurate $CO_2$ and $CH_4$ observations from GOSAT, Geophys. Res. Lett., 38, L14812, doi:10.1029/2011GL047888, 2011.

Chahine, M. T., Chen, L., Dimotakis, P., Jiang, X., Li, Q. B., Olsen, E. T., Pagano, T., Randerson, J., and Yung, Y. L.: Satellite remote sounding of mid-tropospheric $CO_2$, Geophys. Res. Lett., 35, L17807, doi:10.1029/2008GL035022, 2008.

Connor, B. J., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: inverse method and prospective error analysis, J. Geophys. Res., 113, D05305, doi:10.1029/2006JD008336, 2008.

Crevoisier, C., Chédin, A., Matsueda, H., Machida, T., Armante, R., and Scott, N. A.: First year of upper tropospheric integrated content of $CO_2$ from IASI hyperspectral infrared observations, Atmos. Chem. Phys., 9, 4797–4810, doi:10.5194/acp-9-4797-2009, 2009.

Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., Brown, L. R., Castano, R., Connor, B., Deutscher, N. M., Eldering, A., Griffith, D., Gunson, M., Kuze, A., Mandrake, L., McDuffie, J., Messerschmidt, J., Miller, C. E., Morino, I., Natraj, V., Notholt, J., O'Brien, D., Oyafuso, F., Polonsky, I., Robinson, J., Salawitch, R., Sherlock, V., Smyth, M., Suto, H., Taylor, T., Thompson, D. R., Wennberg, P. O., Wunch, D., and Yung, Y. L.: The






ACPD
12, 6283–6329, 2012

ACOS $X_{CO_2}$ retrieval algorithm, Part 2: Global $X_{CO_2}$ data characterization, Atmos. Meas. Tech. Discuss., 5, 1–60, doi:10.5194/amtd-5-1-2012, 2012.

Foucher, P. Y., Chédin, A., Armante, R., Boone, C., Crevoisier, C., and Bernath, P.: Carbon dioxide atmospheric vertical profiles retrieved from space observation using ACE-FTS solar occultation instrument, Atmos. Chem. Phys., 11, 2455–2470, doi:10.5194/acp-11-2455-2011, 2011.

Houweling, S., Breon, F.-M., Aben, I., Rödenbeck, C., Gloor, M., Heimann, M., and Ciais, P.: Inverse modeling of $CO_2$ sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time, Atmos. Chem. Phys., 4, 523–538, doi:10.5194/acp-4-523-2004, 2004.

Kulawik, S. S., Osterman, G., Jones, D. B. A., and Bowman, K. W.: Calculation of altitude-dependent Tikhonov constraints for TES nadir retrievals, IEEE T. Geosci. Remote, 44(5), 1334–1342, 2006.

Kulawik, S. S., Bowman, K. W., Luo, M., Rodgers, C. D., and Jourdain, L.: Impact of nonlinearity on changing the a priori of trace gas profile estimates from the Tropospheric Emission Spectrometer (TES), Atmos. Chem. Phys., 8, 3081–3092, doi:10.5194/acp-8-3081-2008, 2008.

Kulawik, S. S., Jones, D. B. A., Nassar, R., Irion, F. W., Worden, J. R., Bowman, K. W., Machida, T., Matsueda, H., Sawa, Y., Biraud, S. C., Fischer, M. L., and Jacobson, A. R.: Characterization of Tropospheric Emission Spectrometer (TES) $CO_2$ for carbon cycle science, Atmos. Chem. Phys., 10, 5601–5623, doi:10.5194/acp-10-5601-2010, 2010.

Machida, T., Matsueda, H., Sawa, Y., Nakagawa, Y., Hirotani, K., Kondo, N., Goto, K., Nakazawa, T., Ishikawa, K., and Ogawa, T.: Worldwide measurements of atmospheric $CO_2$ and other trace gas species using commercial airlines, J. Atmos. Ocean. Technol., 25(10), 1744–1754, 2008.

Matsueda, H., Inoue, H. Y., and Ishii, M.: Aircraft observation of carbon dioxide at 8–13 km altitude over the Western Pacific from 1993 to 1999, Tellus B, 54(1), 1–21, 2002.

Matsueda, H., Machida, T., Sawa, Y., Nakagawa, Y., Hirotani, K., Ikeda, H., Kondo, N., and Goto, K.: Evaluation of atmospheric $CO_2$ measurements from new flask air sampling of JAL airliner observation, Pap. Meteorol. Geophys., 59, 1–17, 2008.

Moore, J. K., Doney, S. C., and Lindsay, K.: Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model, Global Biogeochem. Cy., 18, GB4028, doi:10.1029/2004GB002220, 2004.

Nassar, R., Jones, D. B. A., Kulawik, S. S., Worden, J. R., Bowman, K. W., Andres, R. J.,

Suntharalingam, P., Chen, J. M., Brenninkmeijer, C. A. M., Schuck, T. J., Conway, T. J., and Worthy, D. E.: Inverse modeling of $CO_2$ sources and sinks using satellite observations of $CO_2$ from TES and surface flask measurements, Atmos. Chem. Phys., 11, 6029–6047, doi:10.5194/acp-11-6029-2011, 2011.

Nevison, C., Mahowald, N., Doney, S., Lima, I., van der Werf, G., Randerson, J., Baker, D., Kasibhatla, P., and McKinley, G.: Contribution of ocean, fossil fuel, land biosphere and biomass burning carbon fluxes to seasonal and interannual variability in atmospheric $CO_2$, J. Geophys. Res., 113, G01010, doi:10.1029/2007JG000408, 2008.

Olsen, S. C. and Randerson, J. T.: Differences between surface and column atmospheric $CO_2$ and implications for carbon cycle research, J. Geophys. Res.-Atmos., 109, D0230, doi:10.1029/2003JD003968, 2004.

Osterman, G. B., et al.: Tropospheric Emission Spectrometer TES L2 Data User's Guide, Version 4.0, Pasadena, Jet Propulsion Laboratory/California Institute of Technology, 2009.

Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed $CO_2$ concentration data in surface source inversions, Geophys. Res. Lett., 28(1), 175–178, 2001.

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P., Connor, B. J., Deutscher, N. M., Griffith, D. W. T., Heymann, J., Keppel-Aleks, G., Messerschmidt, J., Notholt, J., Petri, C., Robinson, J., Schneising, O., Sherlock, V., Velazco, V., Warneke, T., Wennberg, P. O., and Wunch, D.: Retrieval of atmospheric $CO_2$ with enhanced accuracy and precision from SCIAMACHY: validation with FTS measurements and comparison with model results, J. Geophys. Res.-Atmos., 116, D04301, doi:10.1029/2010JD015047, 2011.

Riley, W. J., Biraud, S. C., Torn, M. S., Fischer, M. L., Billesbach, D. P., and Berry, J. A.: Regional $CO_2$ and latent heat surface fluxes in the Southern Great Plains: measurements, modeling, and scaling, J. Geophys. Res.-Biogeosci., 114, D04301, doi:10.1029/2010JD015047, 2009.

Rodgers, C.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing Co, Singapore, 2000.

Wofsy, S. C.. HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, Philos. T. R. Soc. A, 369(1943), 2073–2086, doi:10.1098/rsta.2010.0313, 2011.

Worden, J., Kulawik, S. S., Shephard, M. W., Clough, S. A., Worden, H., Bowman, K., and Goldman, A.. Predicted errors of tropospheric emission spectrometer nadir retrievals from spectral window selection, J. Geophys. Res.-Atmos., 109, D09308, doi:10.1029/2004JD004522, 2004.





ACPD
12, 6283–6329, 2012

Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: Retrieval algorithm for $CO_2$ and $CH_4$ column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite, Atmos. Meas. Tech., 4, 717–734, doi:10.5194/amt-4-717-2011, 2011.

Zhang, L., Jacob, D. J., Boersma, K. F., Jaffe, D. A., Olson, J. R., Bowman, K. W., Worden, J. R., Thompson, A. M., Avery, M. A., Cohen, R. C., Dibb, J. E., Flock, F. M., Fuelberg, H. E., Huey, L. G., McMillan, W. W., Singh, H. B., and Weinheimer, A. J.: Transpacific transport of ozone pollution and the effect of recent Asian emission increases on air quality in North America: an integrated analysis using satellite, aircraft, ozonesonde, and surface observations, Atmos. Chem. Phys., 8, 6117–6136, doi:10.5194/acp-8-6117-2008, 2008.

**Table 1.** Comparisons between $CO_2$ datasets. Summary of coverage, sensitivity, averaging strategies, and errors for several different $CO_2$ products. The averaging and precision are somewhat subjective estimates, with information provided through communication with Ed Olsen (AIRS), Max Reuter (SCIAMACHY) algorithm (Reuter et al., 2011), Greg Osterman (GOSAT and OCO-2), and from Crevosier et al. (2009, IASI).

| | Launch | Spectral region | Peak sens. | Day/ night | Land/ ocean | Latitude | Cloud OD | Obs/ day | Averaging | Precis (ppm) |
|---|---|---|---|---|---|---|---|---|---|---|
| AIRS | 2002 | IR | 6–9 km | Both | Both | 60° S–90° N | All | ~ 15 000 | ≥ 9 targets | 2 |
| SCIAMACHY | 2002 | UV-VIS-IR | Column (col) | Day | Land | 80° S–80° N | ~ 0 | < 10 000 | 5° × 2 month | ~ 1.4 |
| TES | 2004 | IR | 5 km | Both | Both | 40° S–40° N | < 0.5 | ~ 500 | 15° × 1 month | ~ 1.2 |
| IASI | 2006 | IR | 11–13 km | Both | Both | 20° S–20° N | Clear | | 5° × 1 month | 2.0 |
| GOSAT | 2009 | Near IR | Col. | Day | Both | 80° S–80° N | < 0.2 | ~ 2000 | None | 3 |
| OCO-2 | 2013 | Near IR | Col. | Day | Both | 80° S–80° N | < 0.2 | ~ 200 000 | None | < 2 |





ACPD
12, 6283–6329, 2012



CC BY

**Table 2.** The spectral ranges used for TES $CO_2$ with the filter name characteristic of the TES instrument. These spectral ranges have many small spectral regions removed to avoid minor interferent species and persistent spectral residuals. The species included in the forward model were $H_2O$, $CO_2$, $O_3$, $HNO_3$ for the 2B1 filter and $H_2O$, $CO_2$, $O_3$, CFC-11, CFC-12, $NH_3$ for the 1B2 filter.

| | Step 1 | | | Step 2 | |
|---|---|---|---|---|---|
| TES filter | Start ($cm^{-1}$) | End ($cm^{-1}$) | TES filter | Start ($cm^{-1}$) | ($cm^{-1}$) |
| 2B1 | 660.04 | 775.00 | 1B2 | 968.06 | 989.66 |
| 1B2 | 968.06 | 1003.28 | | | |
| 1B2 | 1070.000 | 1100.00 | | | |
| 1B2 | 1110.00 | 1117.40 | | | |





ACPD
12, 6283–6329, 2012

**Table 3.** Sensitivity factor to multiply the 511 hPa averaging kernel row on the $CO_2$ retrieval pressure grid. This ratio is also valid for the 908 hPa, 681 hPa, and 383 hPa averaging kernel rows.

| Pressure (hPa) | Ratio |
|---|---|
| 1000.00 | 0.351038 |
| 908.514 | 0.513463 |
| 681.291 | 0.635048 |
| 510.898 | 0.616426 |
| 383.117 | 0.649254 |
| 287.298 | 0.787116 |
| 215.444 | 1.15804 |
| 161.561 | 1.69716 |
| 121.152 | 2.34417 |
| 90.8518 | 1.99004 |
| 51.0896 | 0.753712 |
| 28.7299 | 0.745675 |
| 4.6416 | 0.365056 |
| 0.1000 | 1.000000 |





ACPD
12, 6283–6329, 2012

**Table 4.** Correlations: calculated and error corrected (ppm). The calculated correlations, $c$, between TES, HIPPO, and SGP are shown, as well as the correlations corrected by the degradation effects of errors, $c_o$. "Variability" is the standard deviation of the aircraft data with the TES observation operator, and "Error" is the predicted observation error, which is about 1.5 times smaller than the actual error. Calculating $c_o$ using the actual error (see Fig. 6 and Table 7) rather than the predicted error gives $c_o$ of about 1 for SGP and 0.9–1.0 for HIPPO.

| | Variability | Pred. error | $c$ | $c_o$ (pred error) | $c_o$ (actual error) |
|---|---|---|---|---|---|
| HIPPO 1 | 1.23 | 0.79 | 0.80 | 0.95 | 0.93 |
| HIPPO 2 | 0.84 | 0.62 | 0.56 | 0.69 | 0.99 |
| HIPPO 3 | 2.31 | 0.93 | 0.86 | 0.95 | 1.00 |
| SGP – var prior | 3.34 | 0.55 | 0.96 | 0.97 | 0.99 |
| SGP – conv. const prior | 2.04 | 0.55 | 0.89 | 0.92 | 0.98 |
| SGP – const prior | 2.04 | 0.51 | 0.88 | 0.91 | 0.98 |





ACPD
12, 6283–6329, 2012

**Table 5.** Detailed effects of longitude coincidence criteria on results. **(a)** Calculated correlations, predicted, and actual errors, and biases when averaging within 5°, 10°, and 15° longitude for each of the HIPPO campaigns and their mean. *n* refers to the # of TES targets averaged per HIPPO comparison, and the column "actl" is the actual error divided by the predicted; a value of 1 indicates the predicted and actual errors agree. The average actual errors (pred*actl) are 1.64, 1.10, and 1.06 ppm for ±5°, ±10°, and ±15°, respectively. **(b)** Average effects of all coincidence criteria. Average results over the 3 HIPPO campaigns for tighter and looser coincidence criteria for longitude (within 5°, 10°, and 15°), time (within 7, 14, and 21 days), latitude (within 2°, 4°, 6°), and clouds (0.1 average OD, 0.5 average OD. Thick cloud targets were not processed, so there is no "loose criteria" case for clouds). The *medium* case is the same for all tests.

| **(a)** | Longitude ±5 | | | | | Longitude ±10 | | | | | Longitude ±15 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Corr | Pred | Actl | Bias | *n* | Corr | Pred | Actl | Bias | *n* | Corr | Pred | Actl | Bias | *n* |
| HIPPO 1 | 0.71 | 1.11 | 0.91 | −0.18 | 32 | 0.80 | 0.79 | 0.96 | −0.85 | 64 | 0.82 | 0.62 | 1.13 | −0.93 | 100 |
| HIPPO 2 | 0.38 | 0.94 | 1.63 | 0.17 | 27 | 0.56 | 0.62 | 1.99 | −0.05 | 60 | 0.66 | 0.49 | 2.28 | −0.14 | 95 |
| HIPPO 3 | 0.81 | 1.53 | 1.47 | 0.06 | 13 | 0.88 | 0.93 | 1.30 | −0.21 | 30 | 0.87 | 0.78 | 1.67 | −0.32 | 43 |
| Mean | 0.64 | 1.19 | 1.38 | 0.02 | 24 | 0.75 | 0.78 | 1.42 | −0.37 | 51 | 0.78 | 0.63 | 1.69 | −0.46 | 79 |
| **(b)** | Tight criteria | | | | | Medium | | | | | Loose criteria | | | | |
| | Corr | Pred | Actl | Bias | *n* | Corr | Pred | Actl | Bias | *n* | Corr | Pred | Actl | Bias | *n* |
| Longitude | 0.64 | 1.19 | 1.28 | 0.02 | 24 | 0.75 | 0.78 | 1.42 | −0.37 | 51 | 0.78 | 0.63 | 1.67 | −0.46 | 82 |
| Time | 0.54 | 1.13 | 1.40 | −0.46 | 28 | 0.75 | 0.78 | 1.42 | −0.37 | 51 | 0.81 | 0.63 | 1.58 | −0.53 | 78 |
| Latitude | 0.67 | 1.13 | 1.28 | −0.43 | 26 | 0.75 | 0.78 | 1.42 | −0.37 | 51 | 0.80 | 0.62 | 1.47 | −0.48 | 79 |
| Clouds | 0.71 | 1.08 | 1.71 | −0.27 | 28 | 0.75 | 0.78 | 1.42 | −0.37 | 51 | – | – | – | – | – |





ACPD
12, 6283–6329, 2012

**Table 6.** Effects of longitudinal *shift* in TES on TES-HIPPO comparisons. TES is offset by +15°, 0°, or −15° for comparisons with with HIPPO. The column "actl" is the actual error divided by the predicted; a value of 1 would mean the predicted and actual errors agree. The unshifted case has the best average results. Latitude +15 means that the mean TES location is +15° from HIPPO.

| | Longitude + 15 | | | | | Longitude + 0 | | | | | Longitude − 15 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Corr | Pred | Actl | Bias | *n* | Corr | Pred | Actl | Bias | *n* | Corr | Pred | Actl | Bias | *n* |
| HIPPO 1 | 0.66 | 0.80 | 1.30 | 1.03 | 60 | 0.80 | 0.79 | 0.96 | −0.85 | 64 | 0.40 | 0.69 | 1.85 | −1.51 | 83 |
| HIPPO 2 | 0.74 | 0.60 | 1.86 | 0.47 | 65 | 0.56 | 0.62 | 1.99 | −0.05 | 60 | 0.57 | 0.58 | 2.02 | 0.26 | 74 |
| HIPPO 3 | 0.65 | 0.97 | 2.07 | 0.30 | 25 | 0.88 | 0.93 | 1.30 | −0.21 | 30 | 0.88 | 0.81 | 1.94 | −0.76 | 36 |
| Mean | 0.68 | 0.79 | 1.74 | 0.60 | 51 | 0.75 | 0.78 | 1.42 | 0.37 | 51 | 0.62 | 0.72 | 1.94 | −0.67 | 64 |





ACPD
12, 6283–6329, 2012

**Table 7.** Characteristics of TES-SGP comparisons. Comparisons between the TES initial guess, TES results and SGP for different configurations. The TES results compared to SGP with the TES observation operator are shown in black, and other comparisons are shown in gray. The number of comparison points is 53 and the average number of TES observations per comparison is 88.

| Quantity | Compared to | Correlation | Pred. error | Actual error | Bias |
|---|---|---|---|---|---|
| Var. prior | SGP | 0.97 | – | 0.82 | −0.72 |
| Constant prior | SGP | 0.00 | – | 3.53 | −0.86 |
| TES w/var. prior | SGP w/obs | 0.96 | 0.55 | 0.96 | 0.13 ± 0.1 |
| TES conv. const prior | SGP w/obs | 0.89 | 0.55 | 0.96 | 0.13 ± 0.1 |
| TES w/const. prior | SGP w/obs | 0.88 | 0.51 | 1.01 | 0.01 ± 0.1 |
| TES w/const. prior | SGP | 0.82 | 0.51 | 2.22 | −0.33 ± 0.1 |





ACPD
12, 6283–6329, 2012

**Table 8.** Comparisons for different pressure levels using a 385 ppm initial $CO_2$ value. Comparisons between and SGP or HIPPO at different pressures, starting with a uniform initial guess and prior. "Obs. var." is the variability of the validation data with the TES observation operator applied, DOFs is the total of the averaging kernel row for the specified pressure level, and the "actl. error" column is the actual divided by predicted error.

| Comparison | Pressure (hPa) | Obs. var. (ppm) | DOFs | Correlation | Pred. error (ppm) | Actl. error | Bias (ppm) |
|---|---|---|---|---|---|---|---|
| SGP w/constant prior | 908 | 1.80 | 0.62 | 0.80 | 0.56 | 2.29 | −0.10 ± 0.1 |
| | 681 | 2.24 | 0.63 | 0.85 | 0.59 | 2.22 | −0.02 ± 0.1 |
| | 511 | 2.02 | 0.62 | 0.88 | 0.51 | 1.98 | 0.01 ± 0.1 |
| HIPPO-1 w/constant prior | 908 | 1.17 | 0.79 | 0.49 | 0.70 | 1.56 | −0.74 ± 0.1 |
| | 681 | 1.27 | 0.88 | 0.59 | 0.76 | 1.42 | −0.88 ± 0.1 |
| | 511 | 1.02 | 0.73 | 0.62 | 0.61 | 1.39 | −0.86 ± 0.1 |
| | 383 | 0.75 | 0.54 | 0.63 | 0.44 | 1.41 | −0.72 ± 0.1 |
| HIPPO-2 w/constant prior | 908 | 0.56 | 0.80 | 0.61 | 0.77 | 1.56 | 0.16 ± 0.1 |
| | 681 | 0.69 | 0.87 | 0.56 | 0.81 | 1.68 | 0.00 ± 0.1 |
| | 511 | 0.48 | 0.71 | 0.52 | 0.63 | 1.83 | −0.21 ± 0.1 |
| | 383 | 0.35 | 0.52 | 0.49 | 0.45 | 1.96 | −0.28 ± 0.1 |
| HIPPO-3 w/constant prior | 908 | 1.97 | 0.75 | 0.83 | 1.18 | 1.14 | 0.92 ± 0.2 |
| | 681 | 2.02 | 0.85 | 0.73 | 1.30 | 1.33 | 0.69 ± 0.2 |
| | 511 | 1.69 | 0.72 | 0.68 | 1.06 | 1.46 | 0.26 ± 0.2 |
| | 383 | 1.26 | 0.54 | 0.65 | 0.78 | 1.55 | 0.03 ± 0.2 |





ACPD
12, 6283–6329, 2012

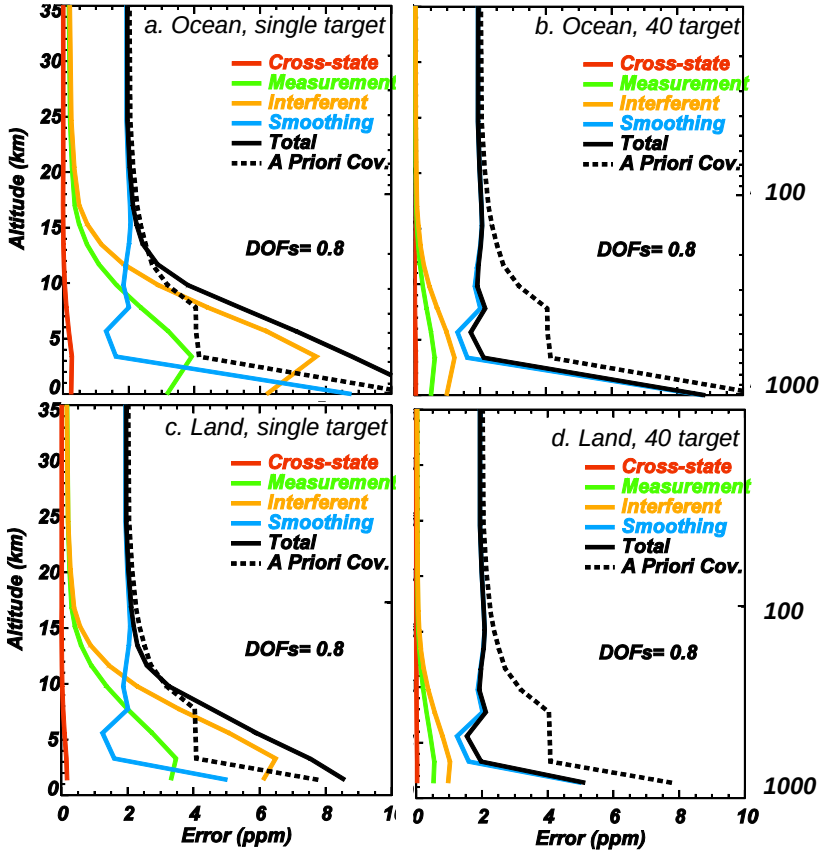


**Fig. 1.** Errors for an ocean scene **(a,b)** and land scene **(c,d)**. Left panels show single target errors and right panels show errors for 40-target averages assuming a random distribution of measurement, interferent, and cross-state errors.





ACPD
12, 6283–6329, 2012

Comparison of improved Aura TES $CO_2$

S. S. Kulawik et al.

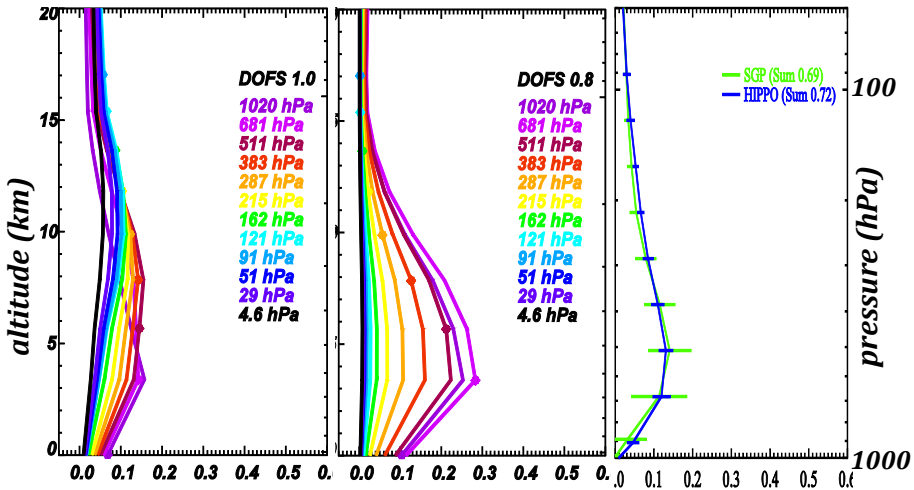


**Fig. 2.** Averaging kernel for the initial $CO_2$ step (left), the final $CO_2$ step (center), and the corrected Averaging Kernel row for 511 hPa for the final $CO_2$ step (right) for SGP and HIPPO cases. Note that the predicted sensitivity in the lower troposphere is less for the initial step because temperature, water, and cloud properties are jointly retrieved.





ACPD
12, 6283–6329, 2012

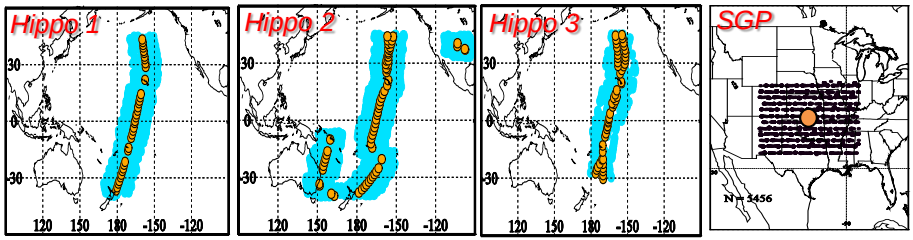


**Fig. 3.** HIPPO-1, HIPPO-2, and HIPPO-3, SGP and TES coincident observation locations. For HIPPO, each orange dot shows a $CO_2$ profile location. The blue values show the TES observations which are averaged for comparisons. Note that for plots versus latitude, there can be multiple longitudes or times as seen on the above plots. For SGP comparisons, the aircraft measurements are located at the orange dot, and the black dots show the TES observation locations which are averaged for comparisons.





ACPD
12, 6283–6329, 2012

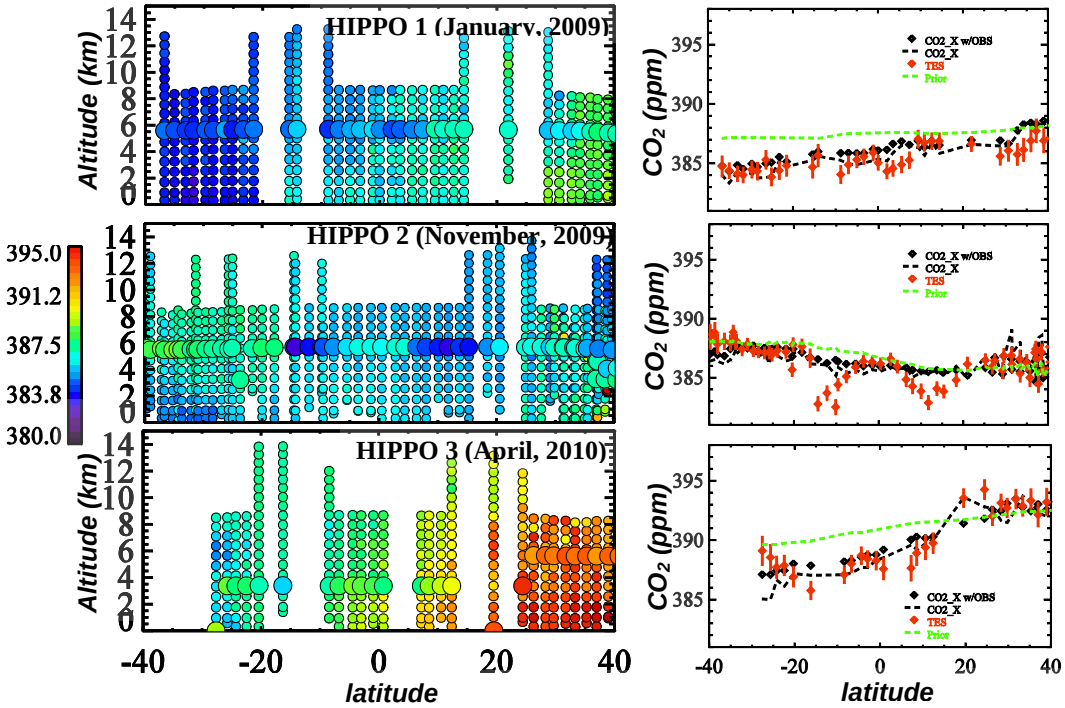


**Fig. 4.** (left panels) Curtain plots comparing TES and HIPPO 1 (top), HIPPO 2 (middle) and HIPPO 3 (bottom) versus latitude. The small dots show the HIPPO values and the large dots show the TES value plotted at the altitude of maximum sensitivity. (right panels) Plots versus latitude now showing the TES value with error bars (red), HIPPO with the TES observation operator applied (black dots), and TES prior and initial guess (green dashed).





ACPD
12, 6283–6329, 2012

Comparison of improved Aura TES $CO_2$

S. S. Kulawik et al.

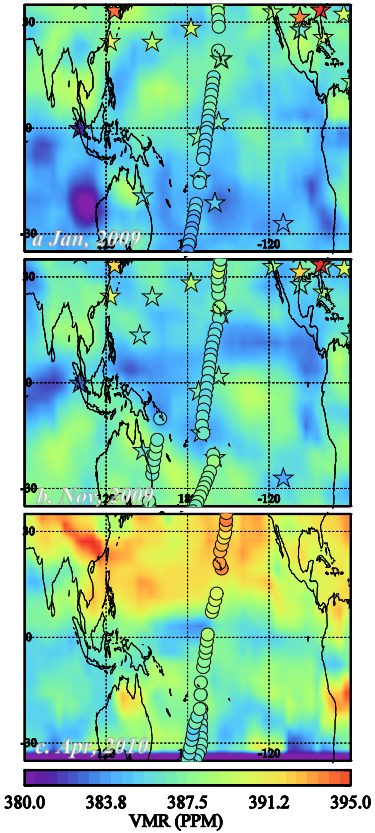


**Fig. 5.** TES monthly averaged results with a moving box ±4° latitude, ±10° longitude, at 511 hPa. HIPPO values at 5 km measured in each month shown as circles. Panel **(a)** corresponds to HIPPO 1, panel **(b)** to HIPPO 2, and panel **(c)** to HIPPO 3. The monthly averaged GLOBALVIEW station values (stars) are shown for context; these surface measurements not necessarily expected to agree with mid-Tropospheric values.





ACPD
12, 6283–6329, 2012

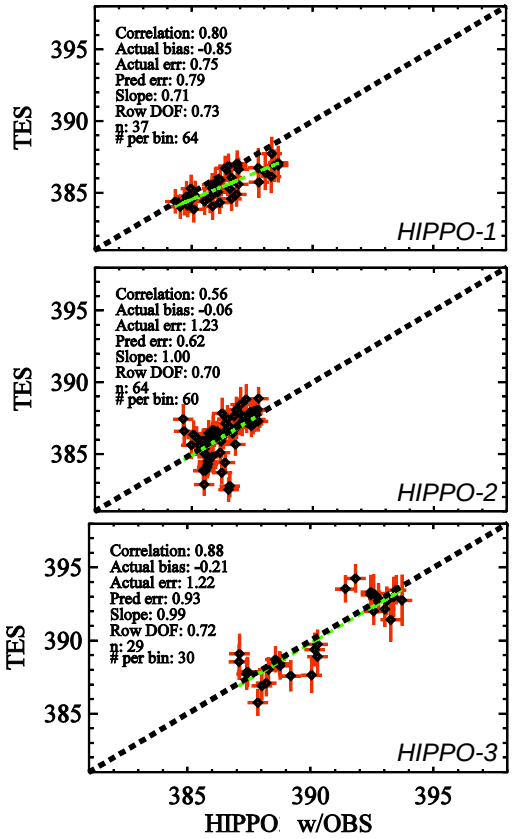


**Fig. 6.** Comparison of TES and HIPPO. The TES results are averaged within 2 weeks, 4° latitude, and 10° longitude for each of the HIPPO campaigns. The HIPPO results have the TES observation operator applied to account for TES sensitivity.





ACPD
12, 6283–6329, 2012

Comparison of improved Aura TES $CO_2$

S. S. Kulawik et al.

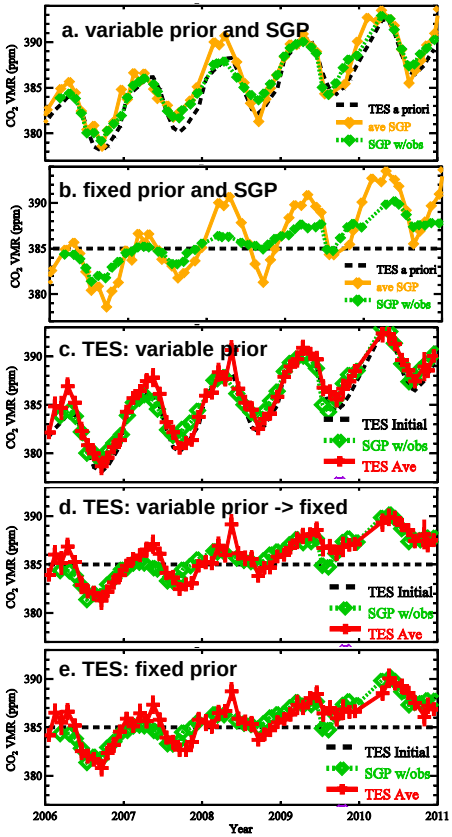


**Fig. 7.** TES compared to SGP aircraft profile data. Monthly averages of all SGP aircraft measurements above 2 km (orange) and SGP aircraft profiles measurements with TES observation operator applied (SGP w/obs, green) are shown in panel **(a)** with variable prior and panel **(b)** using a fixed prior. The observation operator applied to the SGP profiles shows what TES is expected to measure given the prior used and the TES sensitivity. TES results (red) are compared to SGP w/obs (green) in panels **(c–e)**. TES results are shown for a variable prior **(c)**, a variable prior converted linearly to a fixed prior using Eq. 6d, and for non-linear retrievals using a fixed prior and initial guess **(e)**. The bottom two panels agree well with each other (more detailed comparisons shown in Fig. 4). Note: for the first few months of 2010, TES did not collect data.





ACPD
12, 6283–6329, 2012

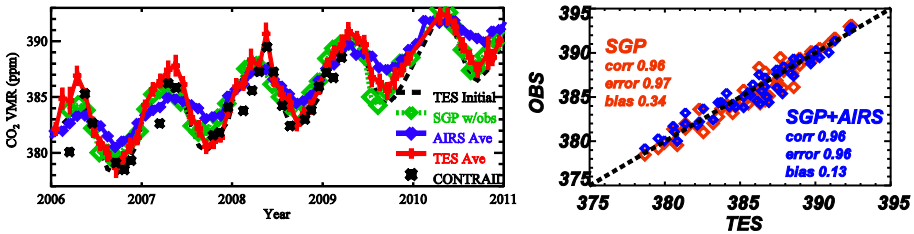


**Fig. 8.** TES compared to SGP aircraft profile data either extending the aircraft data with the top value or transitioning to AIRS $CO_2$ measurements at 9 km. Left panel: A time series showing monthly averages for TES (red), AIRS (blue), SGP with the TES observation operator (green), and CONTRAIL aircraft measurements (black x). Right panel: statistics for SGP and SGP + AIRS results. Adding AIRS in the upper troposphere results in a modest improvement in the bias.





ACPD
12, 6283–6329, 2012

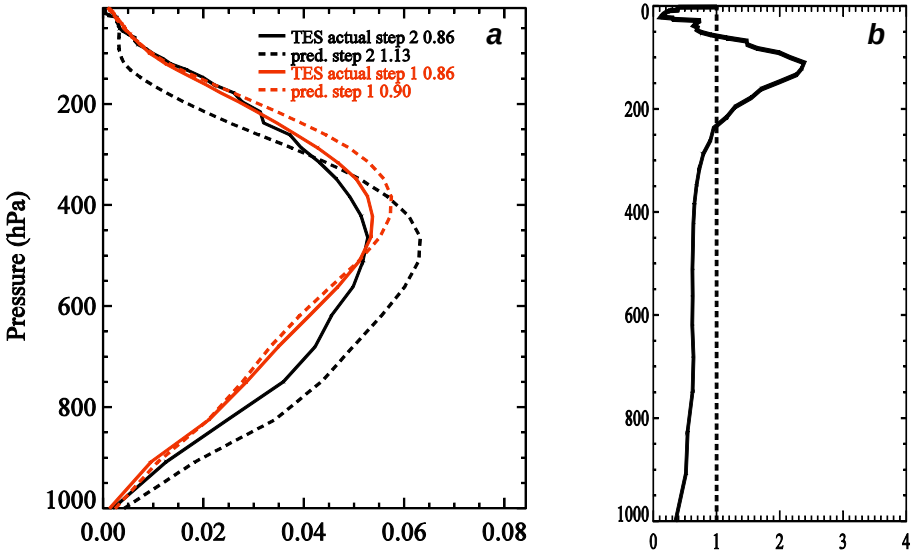


**Figure A1.** **(a)** Step 1 predicted and actual averaging kernel row for 511 hPa. Step 1 actual and predicted are very similar (red), however the step 2 actual averaging kernel (black solid) is significantly different than predicted (black dashed). The peak sensitivity is ~500 hPa, similar to predicted, however the sensitivity is less, and the sensitivity above 200 hPa follows the Step 1 sensitivity. **(b)** The ratio of actual divided by predicted sensitivity for 511 hPa. This ratio is used to correct the averaging kernel provided in the TES product.





ACPD
12, 6283–6329, 2012

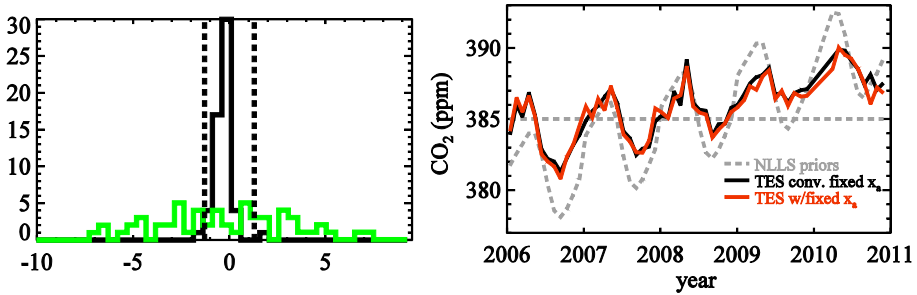


**Figure A2.** Validation of the predicted sensitivity and non-linearity using a constant prior versus a variable prior converted to a constant prior via a linear transform following the nonlinear estimate. Green shows a histogram of the differences of the prior values and black shows a histogram of the result values for TES monthly averages near the SGP site. The dashed lines are the predicted total errors (left panel). Shows a time series, with the fixed prior shown in red and the variable prior converted to a fixed prior shown in black. The gray dashed lines show the two prior values. The red and black lines show excellent agreement, validating the predicted sensitivity and non-linearity (right panel).





ACPD
12, 6283–6329, 2012