Atmospheric CO_2 inversion validation using vertical profile measurements: Analysis of four independent inversion models

C. A. Pickett-Heaps,[1,2] P. J. Rayner,[1,3] R. M. Law,[4] P. Ciais,[1] P. K. Patra,[5] P. Bousquet,[1] P. Peylin,[1] S. Maksyutov,[5,6] J. Marshall,[7] C. Rödenbeck,[7] R. L. Langenfelds,[4] L. P. Steele,[4] R. J. Francey,[4] P. Tans,[8] and C. Sweeney[8]

Received 9 August 2010; revised 21 February 2011; accepted 18 March 2011; published 25 June 2011.

[1] We present the results of a validation of atmospheric inversions of CO_2 fluxes using four transport models. Each inversion uses data primarily from surface stations, combined with an atmospheric transport model, to estimate surface fluxes. The validation (or model evaluation) consists of running these optimized fluxes through the forward model and comparing the simulated concentrations with airborne concentration measurements. We focus on profiles from Cape Grim, Tasmania, and Carr, Colorado, while using other profile sites to test the generality of the comparison. Fits to the profiles are generally worse than to the surface data from the inversions and worse than the expected model-data mismatch. Thus inversion estimates are generally not consistent with the profile measurements. The TM3 model does better by some measures than the other three models. Models perform better over Tasmania than Colorado, and other profile sites bear out a general improvement from north to south and from continental to marine locations. There are also errors in the interannual variability of the fit, consistent in time and common across models. This suggests real variations in sources visible to the profile but not the surface measurements.

Citation: Pickett-Heaps, C. A., et al. (2011), Atmospheric CO_2 inversion validation using vertical profile measurements: Analysis of four independent inversion models, *J. Geophys. Res.*, *116*, D12305, doi:10.1029/2010JD014887.

1. Introduction

[2] Atmospheric CO_2 is increasing as a direct result of anthropogenic activities, although the increase is approximately half the rate expected from fossil fuel emission estimates [*Keeling et al.*, 1995; *Francey et al.*, 1995] due to carbon uptake by the biosphere and oceans. It is therefore vital that we understand this mitigating effect and how it might change in the future with continued changes in the global climate. Atmospheric inversion models are used to infer regional spatiotemporal CO_2 fluxes from observed atmospheric CO_2 measurements using atmospheric transport models (ATMs). Examples include work by *Rayner et al.* [1999], *Bousquet et al.* [2000], *Peylin et al.* [2001], *Rödenbeck et al.* [2003], *Michalak et al.* [2004, 2005], *Peylin et al.* [2005], *Patra et al.* [2005a, 2005b], *Peters et al.* [2007] and *Rayner et al.* [2008].

[3] While inversion techniques have advanced our understanding of the global carbon cycle, important challenges remain. The spatiotemporal distribution of atmospheric CO_2 data is limited, resulting in large flux uncertainty. Reliable flux estimates also depend on the accuracy of transport modeling, with vertical transport being particularly critical and yet displaying substantial variation between models [*Gurney et al.*, 2002]. *Gurney et al.* [2002] and subsequent similar studies could not assess the quality of the inversion flux estimates beyond their ability to fit the data used in the inversion. This paper considers a methodology and data set to allow for such an assessment.

[4] In an effort to assess inverse flux estimates, one can look for validation against independent data such as airborne data [*Stephens et al.*, 2007; *Peters et al.*, 2007] or column-integrated measurements [*Yang et al.*, 2007]. *Stephens et al.* [2007] performed a validation of inversion models using vertical profiles. A climatology of the northern hemispheric vertical CO_2 distribution from 10 aircraft sites was compared with the same vertical CO_2 distribution in the TRANSCOM models used by *Gurney et al.* [2004]. Nearly all the models exhibited systematic biases with the observed vertical gradients, in particular during winter that implied too little

[1]Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA/CNRS/UVSQ, Gif sur Yvette, France.
[2]CSIRO Marine and Atmospheric Research, Canberra, ACT, Australia.
[3]School of Earth Sciences, University of Melbourne, Parkville, Victoria, Australia.
[4]Centre for Australian Weather and Climate Research, CSIRO Marine and Atmospheric Research, Aspendale, Victoria, Australia.
[5]Research Institute for Global Change, JAMSTEC, Yokohama, Japan.
[6]National Institute for Environmental Science, Tsukuba, Japan.
[7]Max Plank Institute for Biogeochemistry, Jena, Germany.
[8]Global Monitoring Division, ESRL, NOAA, Boulder, Colorado, USA.

0148-0227/11/2010JD014887

vertical mixing. Moreover, a clear relationship between the extent of model biases in the vertical gradient and the balance between tropical and temperate carbon uptake was found. Thus, models could be differentiated by having greater consistency with the observed CO_2 vertical gradient and potentially more reliable flux estimates. *Chevallier et al.* [2009, 2010] also demonstrated the utility of model differentiation through validation against airborne data.

[5] We present the results of a new inversion model validation study using CO_2 profile data. Our study differs from that of *Stephens et al.* [2007] because we focus on two profile sites and perform a detailed evaluation of the models including an assessment of model consistency (including intermodel and model-data consistency). In particular, we consider interannual and seasonal variations in our validation analysis. Our profile sites are located in northwest Tasmania (Cape Grim) and northeast Colorado (Carr), although we include results from other profile sites for comparison. We present results from four independent inversion models that, while differing greatly in technical detail, are all based on the Bayesian synthesis inversion formulation [*Enting*, 2002]. We address an important question of whether the profile fit of any of the four models is consistent with the surface inversion fit. Such model consistency has important implications on whether a model can reconcile two observation data sets (surface data and aircraft data) and whether our inversion flux estimates may be improved through inclusion of more data (including aircraft data) using existing transport models or whether the transport models themselves need to be improved.

[6] In this study, we use the word 'validation' in the same sense as that used in numerical weather prediction: keeping some data from the inversion to test its quality. This is not intended as a performance test of the underlying models. Even as a validation of the inversion, it is necessarily partial since the validation only occurs at a few points in the atmosphere. We may therefore attach more significance to failures of the inversions to match the independent data than to success. An equally valid term to model validation is model evaluation.

[7] The outline of this paper is as follows. Section 2 outlines the procedure to cross-validate inversion models against airborne data including a description of the two vertical profile sites, details of the four independent inversion models and forward simulations and the method of comparing model CO_2 vertical profiles and observed profiles (including statistical analyses). Sections 3 and 4 present the results of the comparisons with observed profiles and an analysis of model consistency. Analysis of types of model error such as representation error, errors in the vertical gradient and errors on seasonal and interannual time scales is also provided.

2. Methods: Comparisons Between the Observed and Model Concentration Fields

[8] The vertical distribution of atmospheric carbon dioxide (CO_2) represented by an observed vertical profile is dependent on CO_2 exchange at the Earth's surface and atmospheric transport. Modeling this vertical distribution requires a two-step process: an optimized flux estimation by an atmospheric CO_2 inversion, followed by a forward simulation. The resulting simulation can be compared directly with the observed profile. The model fit to the observed vertical CO_2 distribution reflects on the quality of both the fluxes and transport, i.e., the whole inversion system.

2.1. Airborne Profile Data Used in This Study

[9] Two profile data sets were predominantly used in this study (Table 1): the Cape Grim profile data (AIA) and the profile data from Carr in northeastern Colorado (CAR). The Cape Grim data was collected as part of the Cape Grim Overflight Program (1991–2000, The CSIRO-Marine and Atmospheric Research) and the Carr data are collected as part of the NOAA/GMD aircraft profile measurement program (C. Sweeney et al., Carbon dioxide climatology of the NOAA/ESRL Greenhouse Gas Aircraft Network, manuscript in preparation, 2011), available at http://www.esrl.noaa.gov/gmd/ccgg/aircraft/index.html. These two data sets were chosen because (1) at the time of the study, both data sets consist of a long time series of profile measurements (>5 years), (2) profiles were flown at frequent, regular intervals following strict protocols, (3) profiles were flown from the surface to the upper troposphere (~8000 m) and (4) both profile sites are relatively 'clean air' measurement sites (although they significantly differ from each other.)

[10] The Cape Grim profile site (and associated surface station) is a well-known maritime observation site sampling Southern Ocean air under so-called 'baseline conditions'. Baseline or background conditions occur when the sampled air mass contains integrated signals of large-scale flux variation from distant regions and little or no influence from local flux variation. At Cape Grim, such conditions typically occur during an established southwesterly airstream originating from the Southern Ocean [*Pak et al.*, 1996; *Pak*, 2000; *Steele et al.*, 2003; *Francey*, 2005]. Most profiles were flown 2–3 days after the passage of a cold front [*Pak*, 2000], ensuring the presence of a southwesterly airstream. This potentially introduces a weather sampling bias and points to the need for 'real winds' (or analyzed winds not consisting of a climatology) within the model forward simulation and synchronized sampling for reliable comparisons with observations.

[11] The Carr profile site is a relatively high-altitude (~1700 m) continental observation site. While some profiles at Carr may be influenced by local emissions and/or atmospheric conditions inconsistent with a well-mixed atmosphere, particularly in the lower troposphere; most profiles indicate that air sampled at the Carr site is not much different from observations at Mona Loa (Sweeney et al., manuscript in preparation, 2011). Profiles made over Carr, CO are generally free from continental processes observed at other sites because it is a relatively high altitude site removed from large urban centers. Most of the air masses traveling to Carr, CO travel over the semiarid southwest region of U.S. where local ecosystem CO_2 fluxes are relatively small. It should be noted that this site lies in close proximity to the Rocky Mountains which are distinguished by topographic features at the kilometer scale. Such complex terrain can pose difficulties for global transport models.

[12] Results from other profile sites are briefly presented for a comparative analysis to Cape Grim and Carr. These NOAA/ESRL-GMD sites include Harvard Forest, Massachusetts (HFM), Park Falls, Wisconsin (LEF), Poker Flat, Alaska (PFA), Molokai Island, Hawaii (HAA) and Rarotonga, Cook

Table 1. Summary of the Profile Measurement Sites at Cape Grim (AIA) and Carr (CAR)[a]

	Cape Grim (AIA)	Carr (CAR)
Number of profiles	94	286
Archive period	1991–2000	1992–2002
Profile base	100 m	2100 m
Profile top	~8000 m	~8000 m
Frequency	approx. monthly	weekly
Location	−40° 33′, 144° 18′ NW Tasmania, AUS	40° 54′, −104° 48′ NE Colorado, USA
Measurement technique	flask meas., sample pairs	flask meas., sample pairs
Autonomous measurement	NO	YES
Lab responsible for analyses	GASLAB CSIRO-Marine and Atmospheric Research	Global Monitoring Division (GMD) NOAA/ESRL
Time of day	midday	midday
Contact person	Paul Steele *paul.steele@cmar.csiro.au*	Colm Sweeney, Pieter Tans *pieter.tans@noaa.gov*
Data availability	*CSIRO - MAR (Paul Steele)* *http://geomon-wg.ipsl.jussieu.fr/* World Data Centre for Greenhouse Gases	*http://www.esrl.noaa.gov/gmd/* *ftp://ftp.cmdl.noaa.gov/ccg/vp/* *http://geomon-wg.ipsl.jussieu.fr/*
Relevant citations	*Pak et al.* [1996], *Pak* [2000], *Langenfelds et al.* [1996a, 1996b, 1999]	*Stephens et al.* [2007]
Altitude layer		
Lower Troposphere	0 m (surface) to 1500 m	1600 m (surface) to 4000 m
Mid Troposphere	1500–4000 m	4000–6000 m
Upper Troposphere	4000–8000 m	6000–8000 m

[a]The altitude bins which the aircraft data were separated into at each profile site are also listed.

Islands (RTA). The sparser and shorter time series (3–5 years) afforded by these sites in comparison to the flux inversion time period (1988–2004) inhibits a more detailed study.

2.2. Atmospheric CO_2 Inversion Systems

[13] We use four independent inversion models: (1) CCAM inversion model, CSIRO Marine and Atmospheric Research (Australia); (2) LMDz inversion model, Laboratoire des Sciences du Climat et de l'Environnement, CEA/CNRS/UVSQ (France); (3) FRCGC inversion model, FRCGC/JAMSTEC (Japan); and (4) the Jena Inversion (version s96_v3.0) using the TM3 atmospheric transport model, Max Plank Institute for Biogeochemistry (Germany).

[14] Figure 1 provides a schematic diagram of the atmospheric CO_2 inversion model (and subsequent forward simulation). Inversion models are useful where a relation exists (in this context, defined by atmospheric transport) between a state of unknown parameters (flux estimates) and observations (of atmospheric concentration). This relation, or forward model, must be inverted to determine the state of parameters that best fit the observed quantities. The four inversion models are based on the classical Bayesian formulation [*Rodgers*, 2000; *Enting*, 2002]. We minimize the cost function in equation (1). This jointly minimizes the mismatch between simulated and observed CO_2 concentrations and the match to prior flux estimates. The optimized flux estimates are obtained using equation (2).

$$\Phi(x) = -\frac{1}{2}\left[(y-Kx)^T S_e^{-1}(y-Kx) + (x-x_0)^T S_0^{-1}(x-x_0)\right] \quad (1)$$

$$\hat{x} = x_0 + S_0 \cdot K^T\left(KS_0K^T + S_e\right)^{-1}\cdot(y - Kx_0) \quad (2)$$

$\hat{x}$ a posteriori parameter estimates;
x_0 a priori parameter estimates;
S_0 a priori parameter covariance matrix;
K transport (Jacobian) matrix;
y vector of observations;
S_e observation error covariance matrix.

[15] Technical settings of each inversion are however, substantially different (auxiliary material Text S1).[1] Two important differences are discussed in later sections of the paper. First, the CCAM response functions (concentrations arising from a unit flux) are calculated using only one year of winds (1999) used repeatedly whereas the other model response functions use interannually varying winds. Second, the TM3 response functions are matched to the time of individual observations where the other inversions use responses calculated from the monthly mean. Other notable differences are as follows.

[16] 1. Variation in the spatial resolution of the flux estimates (either region-defined or grid-based).

[17] 2. Method of generating response functions, either using a forward simulation (FRCGC, CCAM) or Eulerian retrotransport (LMDz) [*Hourdin and Talagrand*, 2006]. TM3 uses forward and adjoint model runs in an iterative cost function minimization technique [*Rödenbeck*, 2005] (see auxiliary material Text S1) similar to current 4D-variational data assimilation models.

[18] 2. Variation in the surface network used in the inversion and the extent to which processed data from the GLOBALVIEW-CO_2 product (Co-operative Atmospheric Data Integration Project-Carbon Dioxide, 2005, ftp.cmdl.noaa.gov/ccg/co2/globalview/) was used (e.g., data smoothing, interpolation and extrapolation, use of raw measurements).

[19] 3. In addition to the surface network, TM3 also makes use of aircraft data from the EOMCA data set that consists of high-altitude (10–12 km) measurements of CO_2

[1]Auxiliary materials are available in the HTML. doi:10.1029/2010JD014887.

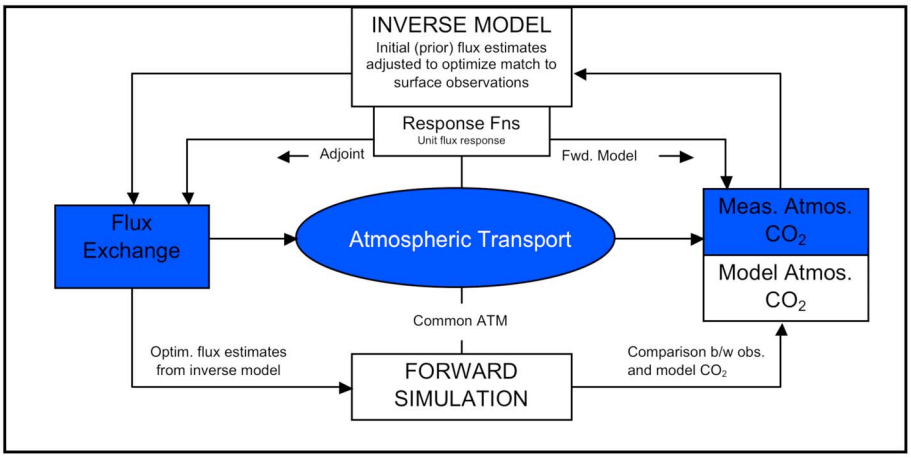

Figure 1. A schematic diagram illustrating the relationship between flux exchange and observed atmospheric CO_2 concentration that is determined by atmospheric transport. Also depicted is the inverse model of this relationship as well as the forward model (forward simulation). In both cases, a common transport model is used to model atmospheric transport.

on passenger flights between Tokyo and Sydney [*Matsueda et al.*, 2002]. The other three models only use surface data.

[20] 4. Variation in the data weights applied to surface measurements.

[21] 5. Variation in the a priori flux PDFs (equation (1)) and in particular, the use of a priori spatial and temporal error correlations.

[22] None of the four inversion models invert the vertical profile data (or any other nonsurface data) so the aircraft data remain independent of all inversion models. In regards to the TM3 inversion, the inclusion or exclusion of the EOMCA aircraft data had a negligible effect on the validation results presented in this study.

2.3. Forward Simulations

[23] For each inversion model, a corresponding forward simulation using the same transport model was run to calculate an optimized CO_2 concentration field, driven by the optimized flux estimates from the inversion model. This optimized concentration field is then compared to the observed profiles (Figure 1). Forward transport model simulations were run for the entire duration of the flux inversions and 'nudged' or 'driven' (depending on whether the model is a global circulation model or atmospheric transport model) using analyzed meteorology. Details of the forward simulations are summarized in auxiliary material Text S2. While the same transport model was used for each inversion model/forward simulation combination, the actual 'transport' used in the inversion may be different. All forward simulations model the actual atmospheric transport (or atmospheric forcing) at a given location in time and space. However, the transport used in the inversion models may instead represent monthly averaged transport fields. This difference in 'transport' is discussed later in the paper.

2.4. Comparison of the Model and Observed Concentration Fields

[24] An accurate model comparison requires appropriate sampling of the optimized CO_2 concentration field coincident in time and space with the observed profiles. The LMDz, FRCGC and TM3 optimized concentration fields were sampled using spatiotemporal interpolation between adjacent model grid points and time steps. For every individual profile flask measurement, a corresponding model data value at the same spatiotemporal location was obtained. The CCAM model was sampled using a different technique. A spatiotemporal domain defined by a 4 h time window (approximately the flight duration) and 4 surrounding grid points was used to generate a vertical CO_2 model profile. This model profile was then compared to the observed profile obtained during the same 4 h time window. Tests showed comparable results to the simpler sampling method used with the other models. In all models, interpolated surface pressure values were used to convert from sigma coordinates to pressure level coordinates. The coincident model temperature profile allowed for the calculation of the corresponding geopotential height of each model level. Linear vertical interpolation (between model levels above and below a profile observation) was then used to estimate the model CO_2 concentration at the altitude of the profile observations. Neither profile data set included the pressure level corresponding to the reported altitude of each measurement, thus necessitating the calculation of geopotential height of each model level at the time and location of each profile observation.

[25] The differences between observed and modeled CO_2 concentrations define the mismatch between the observed and model profiles. For each profile, two error estimates were obtained: the root mean square (RMS) error and the mean model bias. The RMS error is the square root of the sum of the squares of the model-observation residuals divided by the number of observations and is always positive (equation (3)). The mean bias is the mean of the residuals and can be positive or negative (equation (4)).

RMS error

$$RMS = \sqrt{\frac{\sum_{i=1}^{M} (m_i - o_i)^2}{M}} \tag{3}$$

m model concentration estimate;
o observed concentration;

M number of obs. in each profile;
i i^{th}observation of each profile.
RMS error calculated for each profile

Mean Bias

$$MB = \frac{\sum_{i=1}^{M}(m_i - o_i)}{M} \tag{4}$$

[26] Mean Bias is calculated for each profile.

[27] For each profile site, we calculated summary statistics of the RMS error (RMS) and model bias (MB) estimate, i.e., the sample mean (x_{RMS}, x_{MB}), sample standard deviation (s_{RMS}, s_{MB}) and standard error (se_{RMS}, se_{MB}).

Sample mean and standard deviation

$$x = \tfrac{1}{N} \cdot \sum_{i=0}^{N} x_i \tag{5}$$

$$s = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^{N-1} (x_i - \bar{x})^2} \tag{6}$$

where N is sample size (number of profiles).

Standard error of $\bar{x}$

$$se = \frac{s}{\sqrt{N}} \tag{7}$$

95% Confidence interval (CI)of μ

$$x - 2 \cdot se < \mu < x + 2 \cdot se \tag{8}$$

Note that μ represents the unknown mean value of the population distribution and is estimated by x. The confidence interval represents the uncertainty in the estimate x (or the likelihood that μ lies within the CI of x).

[28] Summary statistics, also calculated depending on season and/or altitude, were used to estimate 95% (~2.se) confidence intervals (CIs) of mean RMS and model bias estimates. We use these statistics, and particularly the CIs of the estimated parameters (mean model bias or mean RMS error, equation (7)) to assess the significance of model bias and significant differences between two error estimates from two different sample populations (e.g., two different inversion models, seasons or altitude layers). For example, we use the estimated CIs of model bias between different atmospheric layers or seasons to determine the significance of vertical or seasonal changes in model bias and to determine whether model bias is significantly different from zero.

[29] Such statistical analysis identifies patterns in model error that may be hidden due to model scatter and is advantageous as the number of the degree of freedom (DoF) is small. A major theme of the analysis presented is the determination of the significance of various error estimates (either RMS or mean bias). There will always be a difference between an observed and model profile. The real question relates to whether these differences are significant (i.e., large compared to the uncertainty represented by the confidence interval) and whether they significantly change depending on season, altitude or between inversion models.

[30] Error estimates were calculated for three different altitude layers (or bins) of the atmosphere at each profile location. Three altitude bins were defined at both profile sites (Table 1) primarily due to the data density of the observed profiles that consisted of flask measurements. Take note that there exist variations in the maximum altitude to which the profile measurements were obtained at both profile sites. For example, there are a different number of profiles (i.e., sample size) that contain measurements from 4000 to 6000 m and 6000 to 8000 m at Carr. Such changes are taken into account when calculating summary statistics and in the overall analysis of the model profile fit at both profile sites.

2.5. Estimating the Uncertainty in the Observed Concentration Field

[31] Even if the inversion model and subsequent forward simulation were free of any error, there would still be a model versus observation mismatch owing to uncertainty in the profile measurements arising from instrument error (usually small). Estimates of measurement uncertainty were obtained from personnel responsible for the measurement analyses. In addition, data uncertainty usually takes into account high-frequency variability in observed CO_2 that is not resolved in the model [*Gerbig et al.*, 2003]. The true concentration field is continuous and is sampled by a finite set of observations. The models represent the field by a series of grid box values. We must incorporate this inherent mismatch, often referred to as representation error, when judging a model-observation comparison. However, representation error may also produce systematic biases and therefore cannot be considered as influencing model error that is solely random in nature. The quantity we want to measure is the random variation in the profile observations (observation scatter) that we assume the model cannot represent. We then compare estimates of observation scatter with errors from the model fit to the observed profile.

[32] Estimates of observation scatter error were obtained by fitting low-degree polynomials to each observed profile from Cape Grim and Carr. Statistical tests between polynomials of different degrees (up to a maximum degree of 4) were used in selecting a suitable polynomial that captured the large-scale vertical trend of the observed profile without overfitting. Most vertical profiles were modeled with polynomials of degrees 1–2 and are thus unlikely to have been overfitted. Such vertical structure was considered resolvable by the inversion models. Scatter around each polynomial fit was used to define the level of observed profile scatter and was subsequently compared to the model-observation RMS error estimates. Similar orders of magnitude between RMS errors and profile random scatter suggest an adequate model fit to the observed profile, particularly in the presence of a small model bias.

3. Results

3.1. Model Surface Fit

[33] Assessment of the inversion surface fit, preferably at surface stations in close proximity to the profile site, should precede analysis of the model profile fit. The Cape Grim observation station (CGO) is an obvious choice for the Cape Grim profile data. However, with no surface station located

Table 2. Comparison of the Inversion Surface Fit and Profile Fit for Each Inversion Model[a]

	Model Fit (RMS Error, ppm)					
	CGO Surface		AIA Profile	NWR Surface		CAR Profile
	Inversion Fit	Sfc. Uncert.	Inversion Fit	Inversion Fit	Sfc. Uncert.	Inversion Fit
CCAM	0.03	0.3–0.4	0.62	0.30	0.90	1.29
LMDz	0.04	0.40	0.63	0.90	0.90	1.35
FRCGC	0.25	0.35	0.54	0.80	1.12	1.29
TM3	0.29	0.66–0.92	0.39	0.74	0.85–1.34	1.32

[a]The Cape Grim profile site overlies the Cape Grim (CGO) surface station. The Carr profile site is located near the Niwot Ridge (NWR) surface station. Similar RMS error estimates indicate consistency between the surface and profile model fit. The surface data uncertainties (Sfc. Uncert.) used in the respective inversion models are also included.

at Carr, Niwot Ridge (NWR, located 115 km SW of Carr) is used. Both CGO and NWR are included in the observation network of all four inversion models.

[34] Table 2 provides the RMS error of the surface fit of all four models at CGO and NWR. At both surface sites, the fit from all four inversion models lies within the respective data uncertainties prescribed in each inversion for these sites. Both CCAM and LMDz show a very close surface fit that ordinarily suggests the data uncertainties could be reduced, allowing for greater parameter constraint. TM3 shows more consistency between the surface data uncertainties and the inversion fit. At both sites for all models, the surface fits show negligible model bias.

3.2. Model Profile Fit: Cape Grim

[35] Summary statistics of the model profile comparisons at Cape Grim are shown in Table 3. For three of the four models, the profile fit is significantly worse than the fit at the surface (CGO), indicated overall by larger profile-fit RMS errors compared to the surface (Table 2). The RMS errors from the profile fit are also larger than the data uncertainties of the surface site. Comparisons for each season (not tabulated) indicated that TM3 shows consistency between the surface and profile fit from December-May. At other times, the difference in RMS error between the TM3 profile fit and surface fit is small (~0.2 ppm).

[36] Figures 2 and 3 present seasonal climatologies of RMS error and model bias at Cape Grim. All models display little variation in RMS error with season (or altitude, not shown). In statistical terms, RMS error effectively remains constant throughout the atmosphere for the entire year. This uniformity suggests that large, integrated flux signals are influencing the Cape Grim observed profiles, a result consistent with most profile flights taking place during baseline conditions. The models clearly fail to adequately capture these flux signals. Model error of a large-scale nature may also be playing a part.

[37] Systematic error significantly contributes to the RMS error at Cape Grim for all models (compare model bias and RMS estimates in Table 3), with CCAM showing significant seasonality in model bias (Figure 3, seasonal climatology). This seasonality extends throughout the atmosphere and is therefore not a result of errors in modeling the vertical gradient within the atmospheric profile. The flux seasonality that has been optimized to match the surface data at the CGO surface station is inconsistent with the observed seasonality in the Cape Grim profiles.

[38] Errors in inverse transport are perhaps a more likely source of bias at the vertical profile sites than transport errors in the forward simulation. Response functions from CCAM, FRCGC and LMDz consist of monthly mean response estimates (auxiliary material Text S1). By calculating mean estimates, the influence of model scatter may be averaged out relative to the forward simulation but any systematic errors in transport will remain. It is unlikely that the surface data themselves are in error, although a mismatch between data-selected surface observations with non-data-selected response functions has the potential to introduce systematic errors within the flux estimates. In addition, fixed response functions from a single year (1999) used in the CCAM inversion model could result in the repeated appearance of any seasonal errors and/or transport anomalies unique to 1999 in consecutive years of the inversion model. FRCGC and LMDz use the relevant analyzed winds for each year when generating monthly response functions. However, given similar profile fits near the surface at Cape Grim overall between CCAM, FRCGC and LMDz, the use of fixed response functions from a single year does not appear detrimental at least at Cape Grim.

[39] TM3 fits the Cape Grim profiles better than the other models (Figure 2). It has a small but significant negative bias that does not vary with season or altitude. TM3 uses the modeled concentration relevant for each observation (rather than the monthly mean) to generate inversion response functions and consequently may better represent the baseline observations used in the inversion. The TM3 inversion model does include high-altitude aircraft data from the EOMCA data set [*Matsueda et al.*, 2002]. However, excluding these high-altitude data from the TM3 inversion has a negligible effect on the TM3 model profile fit at Cape Grim.

Table 3. Summary Statistics of the Model Fit to the Cape Grim Profiles for All Four Models[a]

	Mean RMS Error			Mean Model Bias		
	x_{RMS}	s_{RMS}	se_{RMS}	x_{MB}	s_{MB}	se_{MB}
CCAM	0.62	0.21	0.02	−0.29	0.39	0.04
LMDz	0.63	0.39	0.04	−0.31	0.56	0.06
FRCGC	0.54	0.23	0.02	−0.12	0.43	0.04
TM3	0.39	0.20	0.02	−0.22	0.19	0.02

[a]The calculated mean estimate (x_{RMS}) of the model RMS error from each individual profile (and corresponding standard deviation (s_{RMS}) and standard error (se_{RMS}) of this estimate) is provided. Similar estimates for the mean model bias (x_{MB}, s_{MB} and se_{MB}) are provided. These calculated statistics of the model error parameters (RMS error and model bias) are essential in determining if significant differences in model error parameters between different models exist. Data from the full altitude range of the Cape Grim profiles (Table 1) have been used in the generation of these error estimates.

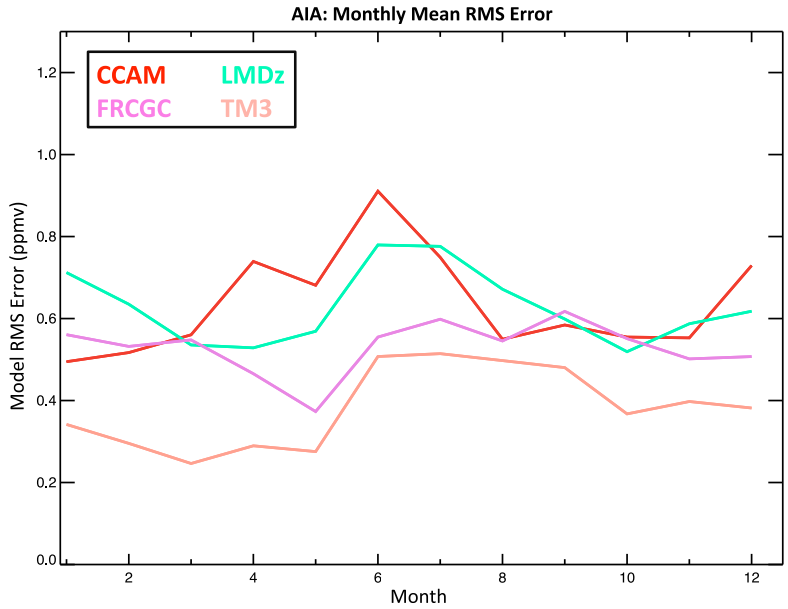

Figure 2. The mean monthly RMS error estimated from the model profile fit to the Cape Grim observed profiles throughout the Cape Grim time series. RMS error estimates are calculated from the model fit of the entire atmospheric column.

3.3. Model Profile Fit: Carr

[40] Summary statistics of the model profile comparisons at Carr are shown in Table 4. The profile fit at Carr for three of the four models is significantly worse than their respective surface fit at Niwot Ridge (compare the profile-fit RMS error with the surface-fit RMS error and surface uncertainties at NWR, Table 2). FRCGC approaches consistency from December to February in the free troposphere and TM3 appears consistent throughout most of the year except in the lower troposphere from June to August (not tabulated). The RMS error of the TM3 profile fit lies within the surface data uncertainties at NWR, despite the difference in RMS error between the surface and profile fits.

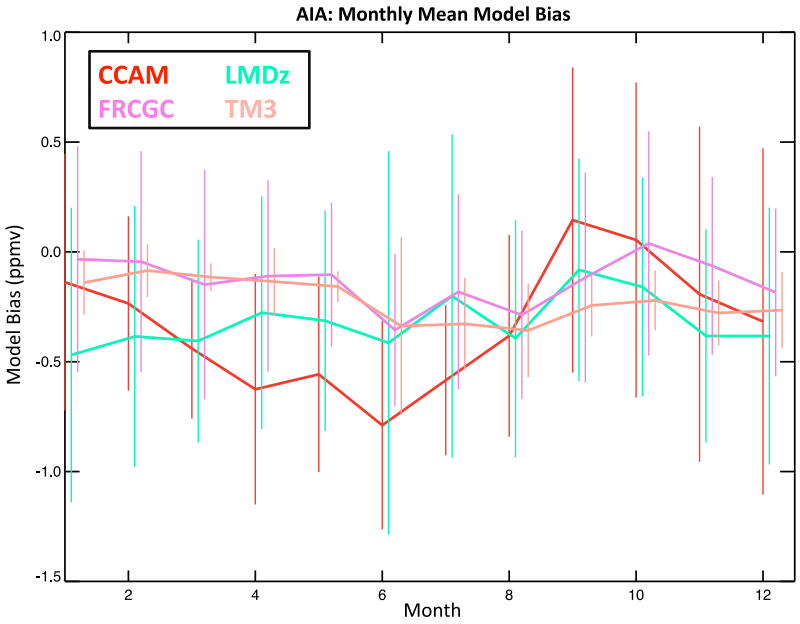

Figure 3. The mean monthly model bias estimated from the model profile fit to the Cape Grim observed profiles throughout the Cape Grim time series. Model bias estimates are calculated from the model fit of the entire atmospheric column. Also depicted is the estimated standard deviation of the monthly model bias (vertical bars). Each model is offset horizontally to improve clarity.

Table 4. Summary Statistics of the Model Fit To the Carr Profiles for All Four Models[a]

	Mean RMS Error			Mean Model Bias		
	x_{RMS}	s_{RMS}	se_{RMS}	x_{MB}	s_{MB}	se_{MB}
CCAM	1.29	0.78	0.05	−0.25	0.98	0.06
LMDz	1.35	0.90	0.05	0.10	1.17	0.07
FRCGC	1.29	0.78	0.05	−0.26	1.06	0.06
TM3	1.32	0.76	0.04	0.20	1.12	0.06

[a]The calculated mean estimate (x_{RMS}) of the model RMS error from each individual profile (and corresponding standard deviation (s_{RMS}) and standard error (se_{RMS}) of this estimate) is provided. Similar estimates for the mean model bias (x_{MB}, s_{MB} and se_{MB}) are provided. Data from the full altitude range of the Carr profiles (Table 1) have been used in the generation of these error estimates.

[41] For all models, the RMS error from the profile fit is much larger at Carr than Cape Grim. Magnitudes of the mean model bias estimates are similar to those at Cape Grim, suggesting random error is more prominent at Carr. There are large seasonal differences in RMS errors for all models (Figure 4), with lower values in winter (DJF, ~1ppm) and higher values in summer (JJA, ~2 ppm). Table 5 provides RMS error estimates for each season averaged throughout the atmosphere.

[42] Stratifying by altitude also produces large differences in RMS errors between adjacent altitude layers (not shown). Annually averaged RMS errors in the lower troposphere approach 2 ppm whereas above 4000 m they are less than 1ppm. The largest differences in RMS error occur from June to August, with RMS errors in the lower troposphere approaching ~2.5 ppm. From December to February, the difference in RMS error between altitude levels is reduced, with RMS errors in the free troposphere approaching those at Cape Grim. Large differences in RMS error between altitude layers may indicate errors in modeling the vertical gradient but they may also simply reflect increased observed concentration variability occurring in the lower troposphere that is not being captured by the transport model.

[43] Seasonality in model bias at Carr is evident in all models (Figure 5). When stratifying by season, all models were found to have significantly negative biases from June to August (Table 5). Stratifying by altitude (Figure 6) reveals this bias as centered in the lower troposphere from June to August, pointing to systematic errors in modeling the vertical gradient (see detailed discussion in section 3.5). In addition to errors in the vertical gradient, LMDz, FRCGC and TM3 show significant seasonality in bias extending throughout the atmospheric column. TM3 shows significantly positive and negative bias between December and February and between June and August, respectively. CCAM does not show this seasonality, suggesting an improved consistency with the observed seasonal cycle above 4000 m.

[44] Given the large RMS errors relative to model biases in Table 4 (and the vertical and seasonal variations therein), sources contributing to random error are likely to dominate at Carr. Representation error is one candidate and results from the inability to resolve small-scale flux and transport variability that give rise to observed high-frequency CO_2 variability. Thus, increased representation errors occur in the presence of increased concentration variability [*Gerbig et al.*, 2003]. However, representation error can also contribute to systematic errors (i.e., model bias). Local flux variability is likely to increase during the growing season at Carr (principally between June and August) and reduce during other times of the year due to the high-altitude, cold winter climate. Likewise, small-scale transport is usually centered in the lower troposphere and increased vertical

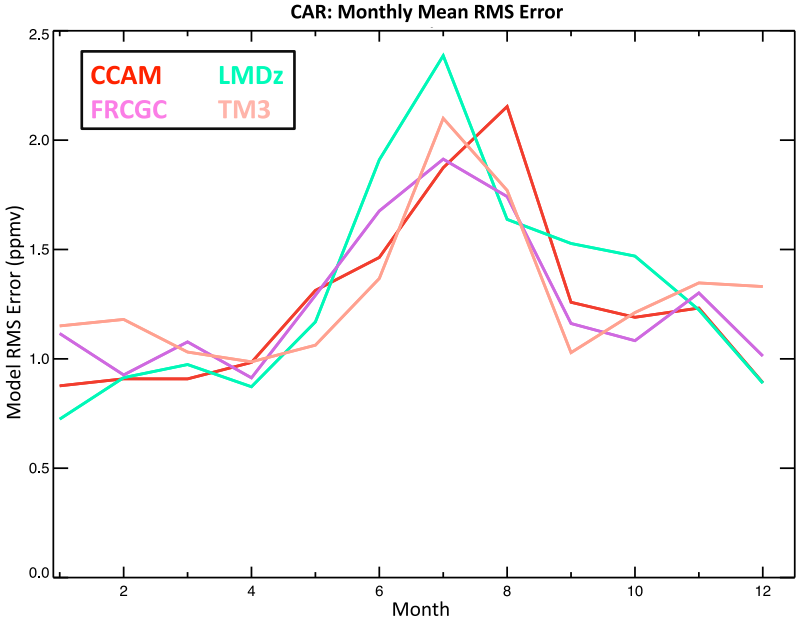

Figure 4. The mean monthly RMS error estimated from the model profile fit to the Carr observed profiles throughout the Carr time series. RMS error estimates are calculated from the model fit of the entire atmospheric column.

Table 5. Seasonal Statistics of the Model Fit to the Carr Profiles for Each Inversion Model[a]

	DJF	MAM	JJA	SON
CCAM				
RMS Error	0.89	1.11	1.85	1.22
Model Bias	−0.12	0.15	−0.70	−0.25
LMDz				
RMS Error	0.84	1.03	2.00	1.42
Model Bias	0.24	0.19	−0.79	0.80
FRCGC				
RMS Error	1.02	1.13	1.79	1.17
Model Bias	0.06	−0.46	−0.62	0.00
TM3				
RMS Error	1.21	1.03	1.78	1.19
Model Bias	0.87	0.42	−0.89	0.52

[a]Mean estimates of seasonal RMS error (x_{RMS}) and model bias (x_{MB}) are shown.

mixing, including increased convective activity, usually occurs in summer (JJA).

3.4. Internal Consistency: Comparison of the Inversion Profile Fit and Surface Fit, Instrument Error, and Observed Profile Scatter

[45] An inconsistency between the inversion surface fit and profile fit is problematic, highlighting the inability of the model to reconcile both data sets despite flux estimates optimized to match the surface data.

[46] The surface fit of all models was found to be acceptable. However, the profile fit from three of the four models was substantially worse than the fit at the surface at both profile sites (in terms of RMS error). This inconsistency casts doubt on the resulting flux estimates due to either a flux that influences the profile data but not the data within the surface network (Table 2), or to errors in model transport. This inconsistency remains when comparing the profile RMS errors to surface data uncertainties. Only TM3 shows consistency between the surface and profile errors. TM3 uses the raw flask observations at Cape Grim and the associated data weights are therefore larger because of the increased difficulty for the model to match these observations. All models show negligible model bias (or systematic variations therein) in the surface fit. In contrast, model bias from the profile fit is significant, including substantial seasonal and interannual components. Therefore, the profile fit from all models appears inconsistent with the surface fit at both profile sites in terms of model bias.

[47] Instrumental error of the Cape Grim profiles (0.1 ppm, L. P. Steele, personal communication, 2004) is significantly less than that of model error for all models. The low-order polynomial fit to observed profiles produces a mean scatter estimate of 0.2 ppm ± 0.1 ppm (SD), occasionally reaching 0.5 ppm. For about 80% of the Cape Grim profiles, RMS error from all models is significantly larger than, and consequently inconsistent with, estimated observation scatter (Table 3). Only TM3 shows some consistency during certain times of the year. The analytical accuracy and precision of measurements at CAR are reported to be 0.03 ppm but short-term storage and gas handling tests with a 12 flask sampling system suggest potential biases on the order of −0.09 ppm (see http://www.esrl.noaa.gov/gmd/ccgg/aircraft/pc.html). Low-order polynomial fits result in a mean scatter of ~0.5 ppm. While the level of scatter in many profiles approaches 1.0 ppm, model error is significantly larger than the estimated scatter for ~80% of the profiles (Table 4).

3.5. Modeling the Vertical Gradient

[48] The vertical gradient in CO_2 is an important indicator of the coupling between flux and transport processes

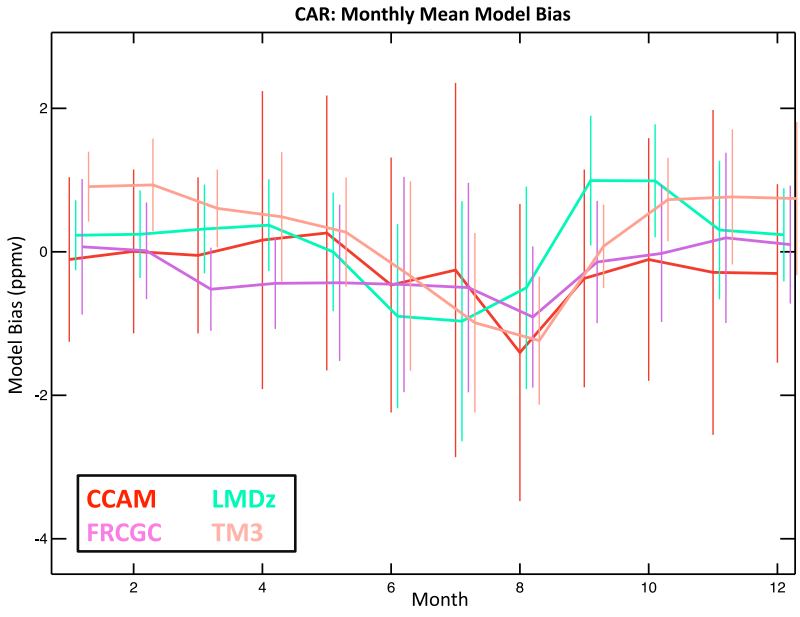

Figure 5. The mean monthly model bias estimated from the model profile fit to the Carr observed profiles throughout the Carr time series. Model bias estimates are calculated from the model fit of the entire atmospheric column. Also depicted is the estimated standard deviation of the monthly model bias (vertical bars). Each model is offset horizontally to improve clarity.

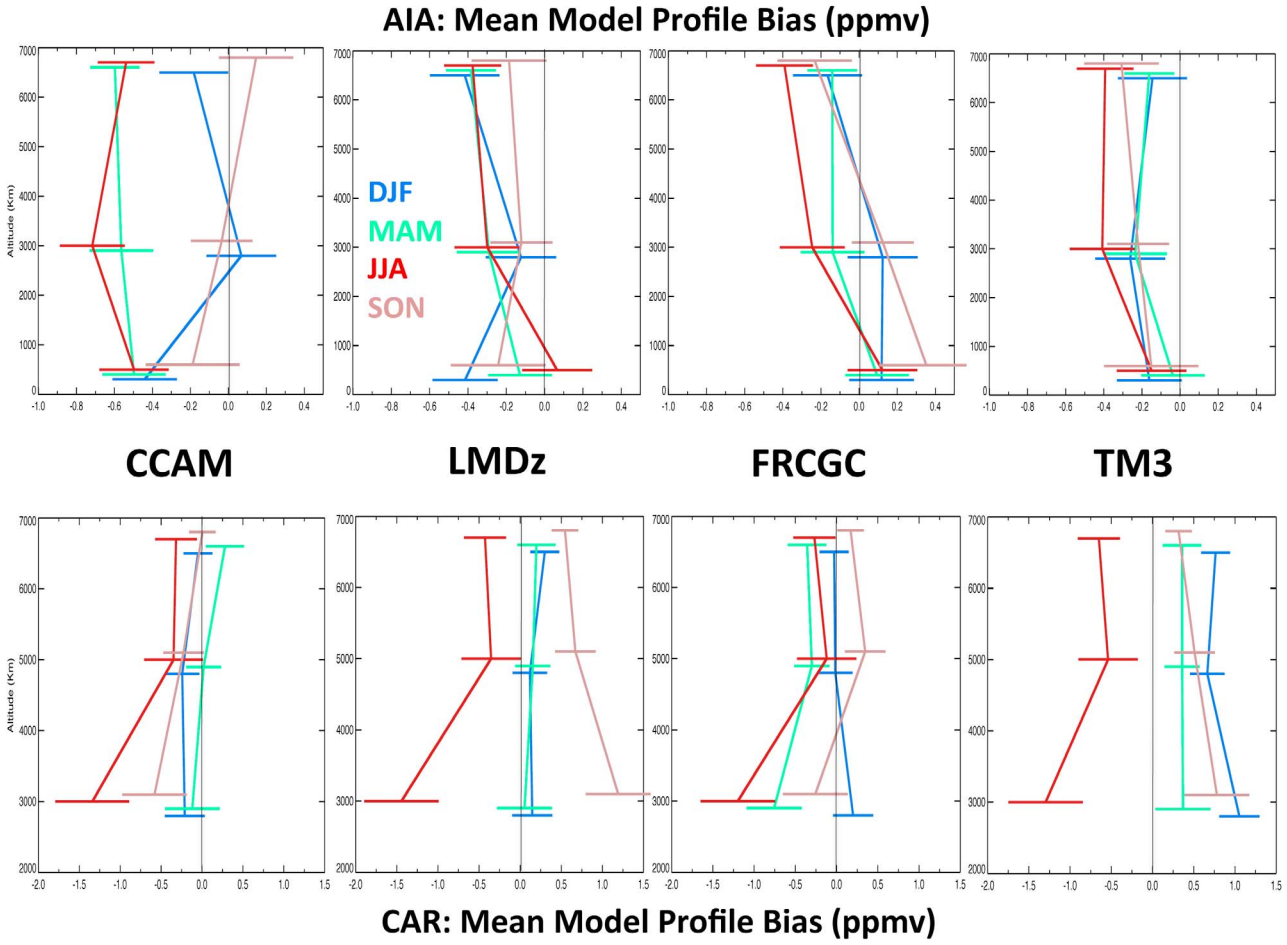

Figure 6. The mean seasonal model profile bias estimated from the model fit to the Cape Grim and Carr observed profiles respectively within three different atmospheric altitude bins (i.e., seasonal profiles of model bias). (top) Results from the model fit to the Cape Grim profiles and (bottom) results from the model fit to the Carr profiles. The horizontal bars for each inversion within each atmospheric region represent the confidence interval (CI) of the estimated seasonal mean model bias. Overlapping CIs indicate the model biases between different inversion models are insignificantly different from each other. CIs intercepting the *y* axis indicate a model bias insignificantly different from zero. Overlapping CIs of model bias from any single inversion model between different atmospheric regions suggest insignificant changes in model bias with altitude (i.e., no significant vertical gradient errors). The CIs are calculated from the estimated standard error of the mean model bias estimate.

influencing the CO_2 concentration field. Errors in matching the observed CO_2 vertical gradient point to a seasonal bias in model transport and cast doubt on subsequent flux estimates [*Stephens et al.*, 2007]. At individual profile sites, a close model fit to the surface and a poor fit to the observed CO_2 gradient suggests errors in vertical mixing.

[49] Analyzing the consistency of the modeled vertical gradient with the observed vertical gradient involves first dividing the observed and modeled vertical profiles into separate altitude bins. The calculated mean concentration estimates between the adjacent altitude bins from the observations and models can then be compared. Due to limits in the number of measurements in any given vertical profile, we limited the number of altitude bins to three at both Cape Grim and Carr (Table 1). Figure 6 provides seasonal plots of mean residual profiles from all four models at both profile sites.

[50] Significant differences in model bias between adjacent altitude bins hint at errors in the vertical gradient. However, to determine more exactly errors in these model gradients, we calculated both the observed and modeled vertical gradients between the adjacent altitude bins and compared them directly. Estimates of the vertical gradient within two different regions of the atmosphere were calculated: one between the first and second altitude bin (lower troposphere) and one between the second and third bin (upper troposphere).

[51] Annual and seasonal mean observed and modeled vertical gradients were estimated along with corresponding standard errors. We then calculated the difference between the observed and model vertical gradients and the corresponding uncertainty in this difference through the propagation of variance. This uncertainty in the gradient error was used to determine whether the gradient errors were

significantly different from zero. This methodology is similar to that of *Stephens et al.* [2007] who generated a climatology of both the observed and modeled vertical gradients before comparing the two. However, *Stephens et al.* [2007] did not present a statistical analysis of the significance in differences between the observed and modeled vertical gradients.

[52] At Carr, all four models were found to significantly overestimate the observed positive summertime (June–August) CO_2 vertical gradient in the lower troposphere. This overestimation also carried over into the next season for two of the four models (September–November). This result is consistent with all models exhibiting significantly larger negative model bias in the lower troposphere from June–August (Figure 6). An overestimation of the observed positive vertical gradient at Carr implies either an overestimate of summertime carbon uptake across all four models or too little vertical mixing. These effects could be separated if the surface inversion fit at Carr was known but no such data exists to verify the surface fit.

[53] However, summertime undermixing at Carr is inconsistent with the results of *Stephens et al.* [2007]. They found that the 12 TRANSCOM models from *Gurney et al.* [2004] tended to overestimate summertime vertical mixing in the northern hemisphere and we would therefore expect a reduced model gradient compared to the observed gradient. It would be surprising that the models presented here should all contradict the result from *Stephens et al.* [2007]. NWR is approximately 115 km SW of Carr. If the surface fit from each inversion model at NWR (known to be acceptable) is representative of the surface fit at Carr, the models are indeed underestimating vertical mixing. However, the extent to which the inversion fit at NWR can be used as a proxy for the fit at Carr is unclear. This is in part because (1) the model representation of mountain sites such as NWR can be particularly problematic and (2) the model footprint of such mountain sites can be large relative to other surface sites. Such large footprints are potentially unrealistic given local affects such as upslope breezes that bring stagnant mountain valley air up to the surface observation site during the morning hours (C. Sweeney, personal communication, 2011).

[54] Consideration of additional profile sites indicated that the inconsistency in vertical gradient errors between our study and that of *Stephens et al.* [2007] was only apparent at the Carr profile site. This finding is discussed in more detail in section 3.7.

[55] All models show reduced and statistically insignificant vertical gradient errors at other times of the year. This is particularly true from December–February where only TM3 indicated significant errors in the vertical gradient within the lower troposphere. *Stephens et al.* [2007] instead identified significant errors in the northern hemispheric climatological gradient among the TRANSCOM models from December–February. Consistent and systematic trends in modeling the vertical gradient higher up in the atmosphere across all four models were not identified. Significant differences in the vertical gradient in the upper atmosphere were only evident for TM3.

[56] Inferring significant errors in the vertical gradient from analysis of the differences in model bias between adjacent altitude layers is generally consistent with the results from the analysis of the vertical gradient estimates themselves. That is, significantly larger negative model bias in the lower troposphere relative to the free troposphere predominantly occurs from June to August only at Carr (Figure 6). However, the analysis of model bias within different altitude bins also reveals seasonal errors extending throughout the atmospheric column that are not associated with errors in modeling the vertical gradient. Three of the four models (LMDZ, FRCGC and TM3 but not CCAM) exhibit this seasonality in model bias which suggests a general model inconsistency with the observed seasonality in CO_2 over Colorado/western USA despite matching the observed seasonality at the surface. Therefore, two types of model inconsistency with the observed profiles are apparent: (1) inconsistencies between the model and observed vertical gradients calculated directly and (2) an inconsistency with the seasonal cycle in the column averages, a result also identified by *Yang et al.* [2007].

[57] The implications of errors in modeling the vertical gradient at Cape Grim are different to those at Carr. Unlike Carr, Cape Grim is predominantly a marine site and is far removed from any large-scale (continental) flux variations. Furthermore, the significance of errors in the modeled vertical gradients is somewhat dependant on the exact specification of the altitude bins. Three of the four models show consistent errors in modeling the vertical gradient for both levels.

[58] FRCGC shows vertical gradient errors extending throughout the atmosphere from June to November. These errors are likely related to large-scale horizontal/vertical transport errors. TM3 shows consistent errors in the lower tropospheric vertical gradient from March to May, despite showing better consistency with the Cape Grim profiles overall (Figure 3). CCAM also consistently shows errors in the lower troposphere from December to February (southern hemispheric summer). This seasonal error is separate from the more large-scale seasonal error in model bias that extends throughout the atmospheric column (evident in Figures 2, 3 and 6). As with the other three models at Carr, two types of model error influence inconsistencies in the CCAM profile fit at Cape Grim: (1) errors in the low-tropospheric vertical gradient (occurring from December to February) and (2) a general inconsistency with the observed seasonal variation at Cape Grim (reaching a maximum from March to August).

[59] An explanation for the summertime low-tropospheric vertical gradient errors within CCAM may be similar to that at CAR, namely vertical mixing errors. However, given that Cape Grim is an important baseline observation station, a mismatch between data-selected surface observations with non-data-selected response functions may introduce systematic changes in flux estimates that are revealed as inconsistent with nearby profile measurements. Recall from section 3.2 that CCAM closely matches the Cape Grim surface data used in the inversion model.

3.6. Interannual Variations in Model Bias

3.6.1. Cape Grim

[60] Figure 7 displays the annual mean model profile bias for each year of the Cape Grim profile time series. Error bars indicate an estimated $\pm 2.\sigma$ range of model bias (σ is the estimated standard deviation of individual model bias estimates and not the standard error). Given the similarity in

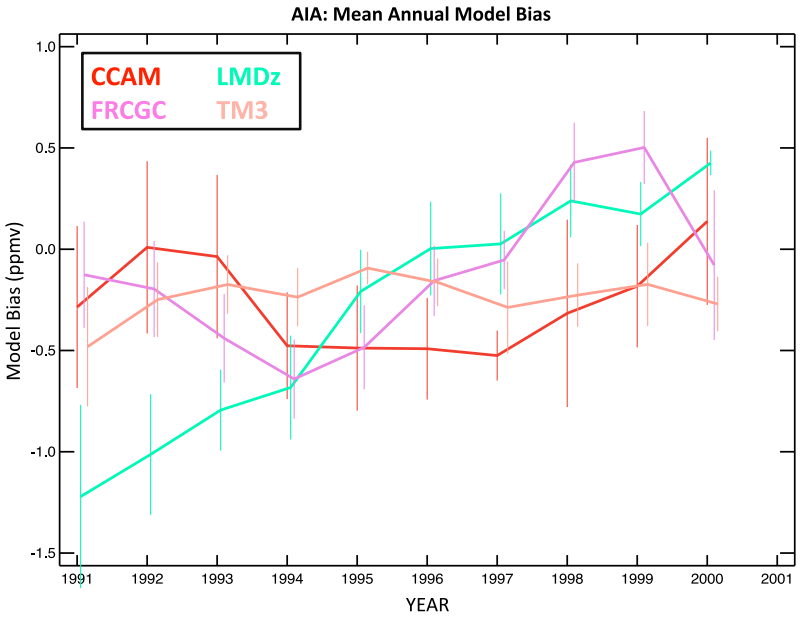

Figure 7. The mean annual model bias estimated from the model profile fit to the Cape Grim observed profiles throughout the Cape Grim time series. Model bias estimates are for the entire atmospheric column. Also depicted is the estimated standard deviation of the annual model bias (vertical bars). Each model is offset horizontally to improve clarity.

model error statistics between CCAM, LMDz and FRCGC, it is surprising to find large differences in the time series of model bias for these models at Cape Grim. There appear to be systematic shifts in model bias for the CCAM model from 1991 to ~1994 (revealing negative-positive-negative trend). From 1994 to 1997, the model bias remains negative (~−0.5 ppm) and after 1998, there is a positive trend. FRCGC also shows a significant negative trend in model bias up to 1994. The trend then becomes positive until reaching a maximum bias in 1999 before decreasing again.

[61] LMDz shows a considerable positive trend in model bias throughout the entire time series (Figure 7). Such an increasing trend might indicate a problem within the inversion system (e.g., conservation of mass) if it was not a regional occurrence. However, Carr does not show similar characteristics. The forward simulation is therefore not likely to be in error. This leaves two further possibilities. Either there is an inconsistency in the interannual variability required of fluxes to fit both the surface and profile data or the bias in the upper air simulation arises from fluxes unconstrained by the surface network. To explore this possibility we analyzed flux trends in large regions that may impact the atmosphere above Cape Grim. The southern oceanic regions surrounding Cape Grim were found not to contain any significant CO_2 flux trends for the duration of the Cape Grim profiles. However, Australia, Tropical Asia and South America (using the standard TRANSCOM region definitions) were identified as having a positive trend in CO_2 emissions to the atmosphere. While none of these three regions individually produce a trend likely to be responsible for the ~+0.15 ppm yr^{-1} change in bias of LMDz, together the three regions emit approximately 0.25–0.3 GtC yr^{-1}. An integrated total emission over 10 years of 2.5–3 GtC is consistent with an integrated change in bias of +1.5 ppm at Cape Grim. The trend in emissions within these three regions appears to diminish post 2000 and we might therefore expect a decrease in the positive LMDz trend at Cape Grim. However, no profile data after 2000 is available to test this hypothesis. Of the three land regions mentioned above, Tropical Asia (lying predominantly north of the equator) appears to be the least likely to have any influence on the Cape Grim profiles. However, there is evidence of interhemispheric mixing and biomass burning influencing the Cape Grim profiles [*Pak*, 2000].

[62] All of the above mentioned flux regions are particularly poorly constrained by the surface network. These positive trends in emissions may also be a result of a coupling with other regions containing negative trends in emissions that are not detected at Cape Grim. However, only the southern Atlantic Ocean was found to have a significantly negative trend in CO_2 fluxes, the magnitude of which was not large enough to fully counteract the increasing trend of the other three land regions. It is also worth noting that CCAM and FRCGC also contain positive and negative trends of similar magnitude to that of LMDz, but for shorter durations within the period of the Cape Grim profiles. While the consistency of the LMDz trend throughout the Cape Grim profile archive points to a systematic trend in the LMDz flux field, the trend of 0.15ppm yr^{-1} is still small enough that it may well lie within the concentration space spanned by uncertainties in the LMDz flux estimates. The LMDz inversion model has subsequently been improved, and the trends in flux estimates discussed above have been reduced.

[63] The model bias from TM3 is relatively constant throughout the whole period, again reflecting the overall improved fit to the Cape Grim profiles provided by TM3. It is also worth noting the differences in the model data bias ranges (error bars) between successive years for individual

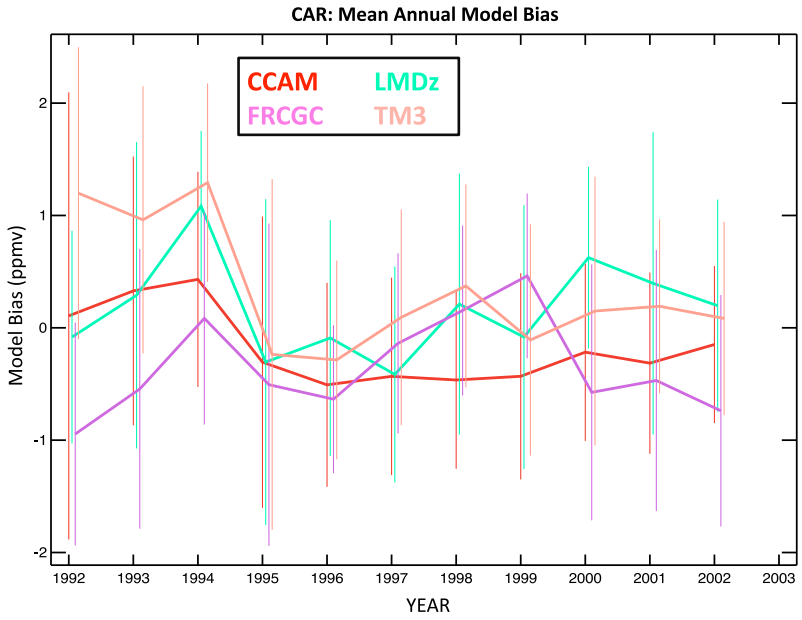

Figure 8. The mean annual model bias estimated from the model profile fit to the Carr observed profiles throughout the Carr time series. Model bias estimates are for the entire atmospheric column. Also depicted is the estimated standard deviation of the annual model bias (vertical bars). Each model is offset horizontally to improve clarity.

models as well as between models in individual years. These differences in error bars could be related to different levels of model scatter between the different models.

3.6.2. Carr

[64] Figure 8 displays the annual mean model profile bias for each year of the Carr profile time series. Interannual variation in model bias for all models is evident. All models show similar characteristics in the beginning of the study period, with an increasing and then decreasing model bias from 1993 to 1996 and a particularly marked negative shift in 1995. The CCAM model then shows relatively constant negative model bias through to the end of the time series. The other models show interannual variations that at times appear to be following a similar trend (three of the four models show an increasing trend between 1996 and 2000). FRCGC diverges from the other models at the end of the study period with a sharp change in trend in 1999.

[65] Such systematic changes in model bias with time are likely to arise from errors in interannual flux variation or model transport. The cause of such systematic variations is difficult to identify in any single model. However, common variations in model bias highlight flux interannual variability evident in airborne data but not at the surface. The fact that all inversions exhibit similar behavior at Carr from 1993 to 1996, including a marked shift in model bias in 1995 strengthens the case for hitherto unobserved flux variations at the surface but impacting the observed profiles. A similar common shift over a much larger sample of models (12) was also identified by *Baker et al.* [2006]. This early period in the time series was an anomalous period in the global carbon cycle usually linked to the Mount Pinatubo eruption in 1991 [*Peylin et al.*, 2005; *Rödenbeck et al.*, 2003; *Lucht et al.*, 2002]. Correlated trends in flux estimates with interannual variability in model bias do not necessarily occur in every instance. Such correlations indicate that the flux estimates are indeed unreliable as they are not consistent with the aircraft observations (or else there should be no significant model bias or variations therein). Errors in transport notwithstanding, a lack of correlation between flux estimates and model bias variability suggests the potential for flux variation that is not evident in the surface observation record and therefore absent in the flux estimates.

[66] Note that changes in the sampling protocol exist at both profile sites. The sampling altitude at Carr was extended to ~8000 m from ~6000 m in the second half of 1995. However, no obvious change in the model fit at Carr is coincident with the change in sampling protocol. The model profile fit was also analyzed (in time) within different layers of the atmosphere (4000–6000 m, 6000–8000 m, 4000–8000 m and the entire profile column) and all show similar features during periods in which there are measurements. It is instead the lower troposphere (2000–4000 m) that is different from the free troposphere, reflecting errors in the vertical gradient in the northern hemispheric summer. Similarly, no association between minor changes in the Cape Grim sampling protocol and significant changes in model fit were found.

3.7. Consideration of Other Profile Sites

[67] The CCAM model results at several other sites (NOAA/ESRL-GMD network: HAA, HFM, LEF, PFA and RTA) have been examined to test the robustness of the findings of this study. The available records are shorter at these sites so we have used them mainly to test the generality of results at Cape Grim and Carr.

[68] Figure 9 and Table 6 provide a general summary of the results from these model comparisons. It is immediately

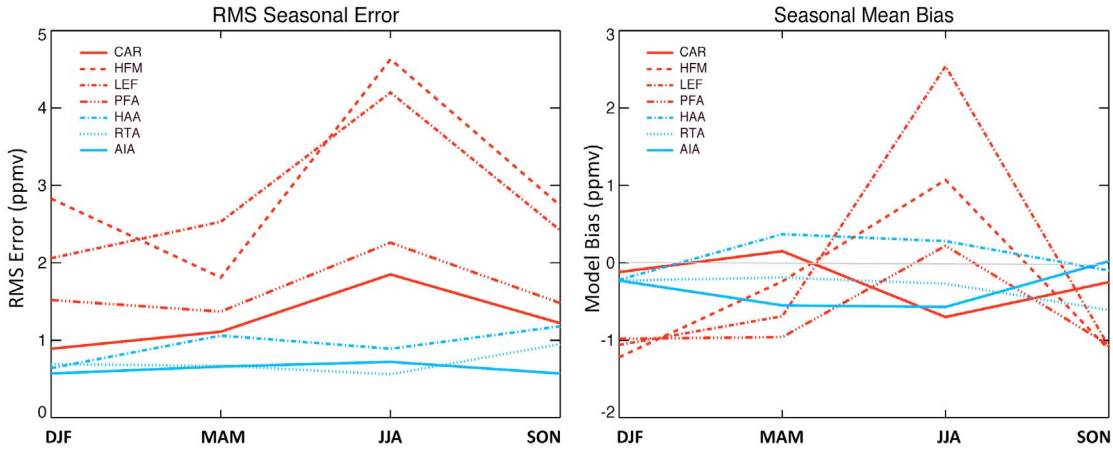

Figure 9. Full column seasonal RMS error and mean bias estimates for additional NOAA profile sites. Profile sites depicted in red are primarily continental observation sites. Profile sites depicted in blue are primarily marine observational sites. The marine profile sites and in addition, PFA, lie along an approximate longitudinal transect of the Pacific Ocean.

obvious that both Cape Grim and Carr are not sites that are particularly difficult to model relative to other sites. In addition, the following observations can be made.

[69] 1. Both LEF and HFM show significantly larger errors relative to Carr. Carr instead shows similar results to PFA, including seasonal variation in RMS errors.

[70] 2. PFA, LEF and HFM all show large seasonality in model error that is in part a result of errors in modeling the vertical gradient in June–August.

[71] 3. The results from PFA, LEF and HFM all hint at positive gradient errors in summer (implying overmixing) whereas Carr shows negative biases (implying undermixing). However, due to limited data availability at the time of this study, these gradient errors were not identified as statistically significant.

[72] 4. Cape Grim remains the easiest site to model, although there is little difference between Cape Grim and RTA. The improved model performance at both these sites may be attributed to their location in the Southern Hemisphere where there exists smaller seasonal variation in observed CO_2 relative to the Northern Hemisphere.

[73] 5. RTA also shows significant seasonality in model bias, although the nature of this seasonality (magnitude and phase) is different to that of Cape Grim and therefore, so is the likely cause(s).

[74] 6. Among the four profile sites running along the Pacific (from PFA, HAA, RTA and Cape Grim), there is a general increasing trend in model skill from north to south.

[75] 7. At HAA, CCAM shows increased error in the upper troposphere relative to the lower troposphere.

[76] The apparent contradiction between the summertime vertical gradient errors observed at Carr (implying undermixing) and those observed for the northern hemisphere [*Stephens et al.*, 2007] does not appear to be significant. Because results from other sites such as LEF, HFM and PFA do not imply undermixing, the results at Carr appear to be relatively local. We suspect that aggregation error of some form provides an explanation for why the vertical gradient errors at Carr are different. Flux emissions influencing the Carr profile site may in turn be influenced by observations at surface stations other than NWR, resulting in locally increased CO_2 uptake causing the Carr model profile gradients to be in error. Alternatively, there may be a coupling between adjacent flux regions that are both constrained by the surface network, but only one of which significantly influences the model profile at Carr.

[77] It might also be possible to classify Carr as a continental background site: modeling results are significantly better than at other continental profile sites but are nonetheless not comparable to marine background sites. Furthermore, modeling complexities related to the complex

Table 6. Comparison of the CCAM Model Profile Fit to Other NOAA Profile Sites[a]

Profile Site	Annual RMS Error	Annual Mean Bias	Longitude	Latitude	Sfc. Alt (m)	Profile Altitude Range (km)
CAR	1.29	−0.25	−104.30	40.37	1740	2.1–8.0
HFM	2.84	−0.52	−72.17	42.54	340	0.6–8.1
LEF	2.83	−0.08	−90.27	45.95	472	0.5–5.0
PFA	1.67	−0.67	−147.29	65.07	210	0.2–7.6
HAA	0.94	0.10	−158.95	21.23	3	0.0–8.1
RTA	0.70	−0.31	−159.83	−21.25	3	0.1–6.5
AIA	0.62	−0.29	144.69	−40.68	0	0.1–8.0

[a]Annual estimates of the mean RMS error (x_{RMS}) and mean model bias (x_{MB}) are shown. The seasonal variation of these error estimates is depicted in Figure 9. The profile site location and surface altitude (Sfc. Alt) are also indicated as well as the altitude range of the aircraft profiles.

terrain in the vicinity of Carr (i.e., the Rocky Mountains) appear not to be as important relative to resolving substantial concentration variations evident at other profile sites (HFM, LEF) where model error is more significant. HFM and LEF lie in relatively flat terrain but observe large concentration variations that are likely due to significant flux variation. The mixed landscape surrounding LEF includes wetlands and upland forests. HFM lies within a mixed-deciduous forest. These biomes/ecosystems are significantly different to those at Carr, a site located in a relatively arid climate where reduced concentration variability due to reduced flux exchange relative to LEF/HFM is likely to exist.

4. Discussion

[78] While validation of inversion models is useful in testing the reliability of inversion flux estimates, the technique does not itself establish that an improved set of flux estimates has been obtained through the application of a particular inverse model. Validation only serves to demonstrate consistency of an inversion model. Significant differences in model consistency can be useful in the differentiation between a collection of inverse models based on an assessment of greater reliability. However, similarity in model consistency does not suggest the two inverse models are the same.

[79] Validation tests also require the coupling of the inversion model with a forward simulation of optimized fluxes. The resulting model consistency relates to the coupled inverse system as a whole and not the individual components. Two main factors can cause inconsistencies with the validation data: (1) The inversion model, including transport modeling in the form of response functions and the inversion setup (prior flux estimate, surface network, surface data uncertainties, flux region resolution and distribution) and (2) the forward simulation.

[80] Analysis presented here does not yet provide a definitive way to separate errors in flux estimates from transport errors in the forward simulation. Furthermore, inversion flux estimates are fundamentally a function of model transport and observation constraints. In both cases, the same transport model is used. While transport errors in the forward simulation and adjoint mode (if it exists) should be exactly the same, in reality the transport used in the inversion model and forward simulation are not always the same. This is dependant on the inversion setup. In the case of TM3, the errors should theoretically be the same because the exact same atmospheric forcing used in the forward simulation is also used in reverse within the inversion. In contrast, both LMDz and FRGCG use monthly averaged atmospheric forcing within their respective inversion models. CCAM uses monthly averaged atmospheric forcing from only one year but all three models use the correct atmospheric forcing (or transport) in their respective forward simulations.

[81] However, analysis of error from the model profile fit can indicate possible sources of this error (transport error or otherwise). The systematic errors at Cape Grim are dominant, affecting the entire model profile fit, and suggest model inconsistency with large-scale fluxes and/or transport. In contrast, random model error (including the effects of representation error) has a more dominant influence at Carr. Multimodel validation experiments help identify interannual flux variation absent in the surface observation network (Figure 7 and Figure 8). Finally, the consistency of the model profile fit against different criteria can help determine the reliability of flux estimates. Demonstrating model consistency (either internal consistency or improved consistency relative to other models) is an important objective of this study.

[82] In consideration of the model profile fit alone, no one model is routinely more consistent with the profile data at both profile sites (i.e., provides a better profile fit). This is despite significant differences between the inversion model setup and model transport (auxiliary material Texts S1 and S2). Such differences include the spatial resolution of the flux estimates (region-based flux estimates versus model-grid flux estimates), consideration of interannual variation in winds, transport model resolution and use of a priori flux spatial correlations in the inversion setup.

[83] TM3 appears both significantly more consistent at Cape Grim and significantly less consistent at Carr compared to the other three models. The opposite is true for CCAM, showing relatively good consistency with the Carr profiles in the free troposphere. The reasons for this are not clear. Given the relatively small difference in RMS error between CCAM and the other models at Carr, decreased random error (or model scatter) possibly from the improved ability to resolve flux or transport variation is unlikely to be the reason. Instead, CCAM appears consistent with the observed seasonal variation in CO_2 above 4000 m over Carr (i.e., within the free troposphere), implying consistency with optimized seasonal flux variations influencing the free troposphere over Carr. The comparison with profile data across the four models thus allows for the differentiation between the different models in terms of consistency and some explanation for the differences. A similar argument can be applied to TM3 at Cape Grim. The TM3 model shows a small and constant model bias and the RMS errors are (1) less than those of the other models and (2) similar to that of the TM3 surface fit at Cape Grim.

[84] The assessment of model consistency described above relies on analysis of the model profile fit alone (and significant differences between the models such as reduced seasonal bias). A different conclusion arises if one considers instead the consistency between the profile and surface fit. In this case, TM3 appears the most consistent model at both sites. The surface fit is both statistically consistent with the surface data uncertainties as well as the profile fit, despite the significant seasonality in the Carr profile fit described above. These different assessments of inversion consistency are not contradictory. Rather, this analysis addresses the question whether a particular inversion fits the profile data as well as can be expected given measurement uncertainties and the difficulty of representing complex vertical structures in coarse-resolution models. In the case of TM3, the Carr profile fit is as good as can be expected because it is statistically consistent with the surface fit. Statistical consistency between the surface and profile fit therefore indicates the profile observations may be used as a further data constraint. For CCAM, LMDz and FRCGC, the model profile fits are inconsistent with the surface fit as well as the uncertainties attached to both the surface data and profile data (instrument error + representation error). These three

models cannot reconcile both the surface and aircraft data. The assessment of model consistency discussed here is however only based on two profile sites. A more thorough evaluation of model consistency either across the globe or even within large geographical regions requires the use of a much larger aircraft data set consisting of multiple profile sites.

[85] What is not addressed here is whether the mismatch with the profiles is consistent with the uncertainties in simulated concentrations arising from the flux uncertainties estimated by the inversion. This is of particular importance due to the inadequate constraints provided by the surface network. Answering this question requires an analysis of such uncertainties and is the subject of a forthcoming paper.

[86] Another important consideration is the significance of intermodel variation on the model profile fit, either in the context of flux estimation or forward simulations. Previous studies such as the TRANSCOM-3 experiments have focused on the significance of intermodel variation in flux estimates using surface network data. A logical extension of this would be to analyze the significance of between-model variations in flux estimates on the model profile fit. A similar case can be made for determining the significance of the forward simulation on the model profile fit through an intercomparison of forward simulations using a common flux field. Both lines of enquiry are currently being undertaken.

[87] The inclusion of more airborne data would also be beneficial. There are large amounts of airborne data collected as horizontal and vertical transects during intensive measurement campaigns. One collection of such data is available at http://geomon-wg.ipsl.jussieu.fr/ for CO_2 and CH_4 from a number of different measurement campaigns.

5. Conclusion

[88] This study explores the utility of airborne profile measurements for the validation of atmospheric CO_2 inversion models. As well as determining the consistency of an inversion model against various criteria, validation highlights errors within the inversion models that are not apparent from an analysis of the flux estimates or the fit to the surface data used in the inversion model. Therefore, results from validation studies help to either support or challenge the description of carbon cycle dynamics given by an inversion model.

[89] We can summarize the conclusions of the study as follows.

[90] 1. For three of the four models at both sites, the fit to the profile data is not statistically consistent with the fit to the surface data or to reasonable measurement and representation errors. However, TM3 demonstrates consistency between the profile and surface fit at both profile sites.

[91] 2. Errors at Cape Grim are dominated by model bias while at Carr the model profile fits also include considerable scatter and errors in fitting vertical gradients.

[92] 3. All models show considerable error in fitting the observed seasonal cycle throughout the atmospheric column at one or both profile sites.

[93] 4. There is interannual variability in the mismatch to the profile data. The consistency of this behavior across models suggests flux variability not observed by the surface network is influencing the observed profiles. The most striking example is the early 1990s at Carr.

[94] 5. Analysis of other sites suggests Cape Grim and Carr are part of a consistent picture of increasing difficulty to fit data as one moves north.

[95] 6. All models show errors in modeling the low-tropospheric summertime (JJA) vertical gradient at Carr.

[96] 7. The apparent inconsistency in modeling the vertical gradient at Carr with that of the Northern Hemisphere (as reported by *Stephens et al.* [2007]) is likely to only exist at Carr. Modeling results at other profile sites, although not shown as statistically significant, suggest that no inconsistency exists at these additional profile sites.

[97] TM3 is the most consistent and reliable model of the four models presented in this manuscript because the TM3 profile fit is consistent with the surface fit at both profile sites. Any significant errors identified in the TM3 profile match are to be expected as they remain consistent with the model fit to surface data included in the inversion. In addition, TM3 clearly provides a better profile fit at Cape Grim relative to the other models. TM3 does not provide a superior profile fit at Carr relative to the other four models (particularly CCAM) and this highlights errors in the TM3 flux estimates (predominantly the seasonality in flux estimates). Nevertheless, consistency between the TM3 profile and surface fit at Carr suggest these errors may be reduced with additional data constraints. However, a more thorough evaluation of model consistency requires a much larger data set consisting of several aircraft profiling sites.

[98] **Acknowledgments.** We acknowledge the support of the European Commission through the GEOmon (Global Earth Observation and Monitoring) Integrated Project under the 6th Framework Program (contract FP6-2005-Global-4-03667). We also acknowledge the support of the CSIRO-Office of the Chief Executive Postdoctoral fellowship award. Finally, we acknowledge the support of CSIRO-Marine and Atmospheric Research and the Laboratoire des Sciences du Climat et de l'Environnement.

References

Baker, D. F., et al. (2006), TransCom 3 inversion intercomparison: Impact of transport errors on interannual variability of regional CO_2 fluxes, 1988–2003, *Global Biogeochem. Cycles*, *20*, GB1002, doi:10.1029/2004GB002439.

Bousquet, P., P. Peylin, P. Ciais, C. Quere, P. Friedlingstein, and P. Tans (2000), Regional changes in carbon dioxide fluxes of land and oceans since 1980, *Science*, *290*, 1342–1346, doi:10.1126/science.290.5495.1342.

Chevallier, F., R. J. Engelen, C. Carouge, T. J. Conway, P. Peylin, C. Pickett-Heaps, M. Ramonet, P. J. Rayner, and I. Xueref-Remy (2009), AIRS-based versus flask-based estimation of carbon surface fluxes, *J. Geophys. Res.*, *114*, D20303, doi:10.1029/2009JD012311.

Chevallier, F., et al. (2010), CO_2 surface fluxes at grid point scale estimated from a global 21-year reanalysis of atmospheric measurements, *J. Geophys. Res.*, *115*, D21307, doi:10.1029/2010JD013887.

Enting, I. G. (2002), *Inverse Problems in Atmospheric Constituent Transport*, Cambridge Univ. Press, Cambridge, U. K., doi:10.1017/CBO9780511535741.

Francey, R. J. (2005), Record growth in CO_2 levels, *Environ. Chem.*, *2*, 3–5, doi:10.1071/EN05013.

Francey, R. J., P. P. Tans, C. E. Allison, I. G. Enting, J. W. C. White, and M. Trolier (1995), Changes in oceanic and terrestrial carbon uptake since 1982, *Nature*, *373*, 326–330, doi:10.1038/373326a0.

Gerbig, C., J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin, and C. A. Grainger (2003), Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, *J. Geophys. Res.*, *108*(D24), 4756, doi:10.1029/2002JD003018.

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, *415*, 626–630, doi:10.1038/415626a.

Gurney, K. R., et al. (2004), TRAHSCOM 3 inversion inter-comparison: Model mean results for the estimation of seasonal carbon sources

and sinks, *Global Biogeochem. Cycles*, *18*, GB1010, doi:10.1029/2003GB002111.

Hourdin, G., and O. Talagrand (2006), Eulerian backtracking of atmospheric tracers. I: Adjoint derivation and parameterization of sub-grid scale transport, *Q. J. R. Meteorol. Soc.*, *132*(615), 567–583, doi:10.1256/qj.03.198.A.

Keeling, C. D., T. P. Whorf, M. Wahlen, and J. van der Plicht (1995), Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980, *Nature*, *375*, 666–670, doi:10.1038/375666a0.

Langenfelds, R. L., R. J. Francey, L. P. Steele, P. J. Fraser, S. A. Coram, M. R. Hayes, D. J. Beardsmore, M. P. Lucarelli, and F. R. de Silva (1996a), Improved vertical sampling of the trace gas composition of the troposphere above Cape Grim since 1991, in *Baseline Atmospheric Program (Australia) 1993*, edited by R. J. Francey, A. L. Dick, and N. Derek, pp. 45–56, Bur. of Meteorol., Melbourne, Victoria, Australia.

Langenfelds, R. L., R. J. Francey, L. P. Steele, P. J. Fraser, S. A. Coram, M. R. Hayes, D. J. Beardsmore, M. P. Lucarelli, and F. R. de Silva (1996b), Program reports 4.11: Flask sampling from Cape Grim overflights, in *Baseline Atmospheric Program (Australia) 1994–95*, edited by R. J. Francey, A. L. Dick, and N. Derek, pp. 112–117, Bur. of Meteorol., Melbourne, Victoria, Australia.

Langenfelds, R. L., R. J. Francey, L. P. Steele, D. A. Spencer, and M. P. Lucarelli (1999), Program reports 4.10: Flask sampling from Cape Grim overflights, in *Baseline Atmospheric Program (Australia) 1996*, edited by J. L. Gras, N. Derek, N. W. Tindale, and A. L. Dick, pp. 96–97, Bur. of Meteorol., Melbourne, Victoria, Australia.

Lucht, W., I. C. Prentice, R. B. Myneni, S. Sitch, P. Friedlingstein, W. Cramer, P. Bousquet, W. Buermann, and B. Smith (2002), Climate control of the high-latitude vegetation greening trend and Pinatubo effect, *Science*, *296*, 1687–1689, doi:10.1126/science.1071828.

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

Michalak, A. M., L. Bruhwiler, and P. P. Tans (2004), A geo-statistical approach to surface flux estimation of atmospheric trace gases, *J. Geophys. Res.*, *109*, D14109, doi:10.1029/2003JD004422.

Michalak, A. M., A. Hirsch, L. Bruhwiler, K. R. Gurney, W. Peters, and P. P. Tans (2005), Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *J. Geophys. Res.*, *110*, D24107, doi:10.1029/2005JD005970.

Pak, B. C. Y. (2000), Vertical structure of atmospheric trace gases over southeast Australia, Ph.D. thesis, Dep. of Earth Sci., Univ. of Melbourne, Melbourne, Victoria, Australia.

Pak, B. C., R. L. Langenfelds, R. J. Francey, L. P. Steele, and I. Simmonds (1996), A climatology of trace gases from the Cape Grim overflights, 1992–1995, in *Baseline Atmospheric Program (Australia) 1994–95*, edited by R. J. Francey, A. L. Dick, and N. Derek, pp. 41–52, Bur. of Meteorol., Melbourne, Victoria, Australia.

Patra, P. K., S. Maksyutov, M. Ishizawa, T. Nakazawa, T. Takahashi, and J. Ukita (2005a), Interannual and decadal changes in the air-sea CO_2 flux from atmospheric CO_2 inverse modeling, *Global Biogeochem. Cycles*, *19*, GB4013, doi:10.1029/2004GB002257.

Patra, P. K., S. Maksyutov, and T. Nakazawa (2005b), Analysis of atmospheric CO_2 growth rates at Mauna Loa using CO_2 fluxes derived from an inverse model, *Tellus, Ser. B*, *57*, 357–365, doi:10.1111/j.1600-0889.2005.00159.x.

Peters, W., et al. (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *Proc. Natl. Acad. Sci. U. S. A.*, *104*(48), 18,925–18,930, doi:10.1073/pnas.0708986104.

Peylin, P., P. Bousquet, and P. Ciais (2001), Inverse modeling of atmospheric carbon dioxide fluxes—Responses, *Science*, *294*, 259, doi:10.1126/science.294.5541.259a.

Peylin, P., P. Bousquet, C. Le Quéré, S. Sitch, P. Friedlingstein, G. McKinley, N. Gruber, P. Rayner, and P. Ciais (2005), Multiple constraints on regional CO_2 flux variations over land and oceans, *Global Biogeochem. Cycles*, *19*, GB1011, doi:10.1029/2003GB002214.

Rayner, P. J., I. G. Enting, R. J. Francey, and R. L. Langenfelds (1999), Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, *Tellus, Ser. B*, *51*, 213–232.

Rayner, P. J., R. M. Law, C. E. Allison, R. J. Francey, C. M. Trudinger, and C. Pickett-Heaps (2008), The interannual variability of the global carbon cycle (1992–2005) inferred by inversion of atmospheric CO_2 and $\delta^{13}CO_2$ measurements, *Global Biogeochem. Cycles*, *22*, GB3008, doi:10.1029/2007GB003068.

Rödenbeck, C. (2005), Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, *Tech. Rep. 6*, Max-Plank-Institut für Biogeochemie, Jena, Germany.

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), Time-dependent atmospheric CO_2 inversions based on inter-annually varying tracer transport, *Tellus, Ser. B*, *55*, 488–497.

Rodgers, C. (2000), *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Sci., Hackensack, N. J.

Steele, L. P., P. B. Krummel, G. A. Da Costa, D. A. Spencer, L. W. Porter, S. B. Baly, R. L. Langenfelds, and L. N. Cooper (2003), Baseline carbon dioxide monitoring, in *Baseline Atmospheric Program (Australia) 1999–2000*, edited by N. W. Tindale, N. Derek, and P. J. Fraser, pp. 80–84, Bur. of Meteorol., Melbourne, Victoria, Australia.

Stephens, B. B., et al. (2007), The vertical distribution of atmospheric CO_2 defines the latitudinal partitioning of global carbon fluxes, *Science*, *316*, 5832, doi:10.1126/science.1137004.

Yang, Z. A., R. A. Washenfelder, G. Keppel-Aleks, N. Y. Krakauer, J. T. Randerson, P. P. Tans, C. Sweeney, and P. O. Wennberg (2007), New constraints on Northern Hemisphere growing season net flux, *Geophys. Res. Lett.*, *34*, L12807, doi:10.1029/2007GL029742.

P. Bousquet, P. Ciais, and P. Peylin, Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CEA/CNRS/UVSQ, Bat. 701, Orme des Merisiers, F-91191 Gif sur Yvette, France.

R. J. Francey, R. L. Langenfelds, R. M. Law, and L. P. Steele, Centre for Australian Weather and Climate Research, CSIRO Marine and Atmospheric Research, Aspendale, Victoria, Australia.

S. Maksyutov and P. K. Patra, Research Institute for Global Change, JAMSTEC, Yokohama 236001, Japan.

J. Marshall and C. Rödenbeck, Max Plank Institute for Biogeochemistry, Postfach 100164, D-07701 Jena, Germany.

C. A. Pickett-Heaps, CSIRO Marine and Atmospheric Research, GPO Box 3023, Canberra, ACT 2601, Australia. (christopher.pickett-heaps@csiro.au)

P. J. Rayner, School of Earth Sciences, University of Melbourne, Parkville, VIC 3010, Australia.

C. Sweeney and P. Tans, Global Monitoring Division, NOAA/ESRL, Boulder, CO 80305, USA.

Author/s:
Peylin, P.; Maksyutov, S.; Marshall, J.; Rödenbeck, C.; Langenfelds, R.L.; Steele, L.P.; Francey, R.J.; Tans P.; Sweeney C.; Pickett-Heaps, C. A.; Rayner, P. J.; Law, R. M.; Ciais, P.; Patra, P. K.; Bousquet, P.

Title:
Atmospheric CO2 inversion validation using vertical profile measurements: analysis of four independent inversion models

Date:
2011

Citation:
Pickett-Heaps, C. A., Rayner, P. J., Law, R. M., Ciais, P., Patra, P. K., Bousquet, P. et al. (2011). Atmospheric CO2 inversion validation using vertical profile measurements: analysis of four independent inversion models. Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2010JD014887.

Publication Status:
Published

Persistent Link:
http://hdl.handle.net/11343/32766