Validation of MOPITT Version 5 thermal-infrared, near-infrared, and multispectral carbon monoxide profile retrievals for 2000–2011

M. N. Deeter,[1] S. Martínez-Alonso,[1] D. P. Edwards,[1] L. K. Emmons,[1] J. C. Gille,[1] H. M. Worden,[1] J. V. Pittman,[2] B. C. Daube,[2] and S. C. Wofsy[2]

[1]   Validation results are reported for the MOPITT (Measurements of Pollution in the Troposphere) "Version 5" (V5) product for tropospheric carbon monoxide (CO) and are compared to results for the "Version 4" product. The V5 retrieval algorithm introduces (1) a method for reducing retrieval bias drift associated with long-term instrumental degradation, (2) a more exact representation of the effects of random errors in the radiances and, for the first time, (3) the use of MOPITT's near-infrared (NIR) radiances to complement the thermal-infrared (TIR) radiances. Exploiting TIR and NIR radiances together facilitates retrievals of CO in the lowermost troposphere. V5 retrieval products based (1) solely on TIR measurements, (2) solely on NIR measurements and (3) on both TIR and NIR measurements are separately validated and analyzed. Actual retrieved CO profiles and total columns are compared with equivalent retrievals based on in situ measurements from (1) routine NOAA aircraft sampling mainly over North America and (2) the "HIAPER Pole to Pole Observations" (HIPPO) field campaign. Particular attention is focused on the long-term stability and geographical uniformity of the retrieval errors. Results for the retrieved total column clearly indicate reduced temporal bias drift in the V5 products compared to the V4 product, and do not exhibit a positive bias in the Southern Hemisphere, which is evident in the V4 product.

Citation: Deeter, M. N., S. Martínez-Alonso, D. P. Edwards, L. K. Emmons, J. C. Gille, H. M. Worden, J. V. Pittman, B. C. Daube, and S. C. Wofsy (2013), Validation of MOPITT Version 5 thermal-infrared, near-infrared, and multispectral carbon monoxide profile retrievals for 2000–2011, *J. Geophys. Res. Atmos.*, *118*, 6710–6725, doi:10.1002/jgrd.50272.

1.   Introduction

[2]   Satellite observations of tropospheric carbon monoxide (CO) are exploited in many atmospheric science applications including air quality studies, chemical weather forecasting and the characterization of CO emissions through inverse modeling. MOPITT's gas correlation radiometers observe CO simultaneously in both a thermal-infrared (TIR) band near 4.7 μm and a near-infrared (NIR) band near 2.3 μm. This is a unique feature of the MOPITT instrument compared to other tropospheric CO satellite instruments. For retrieving CO volume mixing ratio (VMR) in the lower troposphere, TIR and NIR observations are complementary: TIR radiances are often most sensitive to CO in the mid- and upper-troposphere, whereas NIR observations mainly provide information about the CO total column (with uniform sensitivity throughout the troposphere).

As recently demonstrated, the sensitivity to CO in the lower troposphere is significantly greater for retrievals exploiting simultaneous TIR and NIR measurements than for retrievals based on either spectral region alone [*Worden et al.*, 2010; *Deeter et al.*, 2011, 2012]. "Multispectral" CO retrieval products based on simultaneous TIR and NIR observations are one of three CO retrieval products available in the recent (2011) MOPITT Version 5 (V5) data release. V5 TIR-only and NIR-only CO retrieval products are also available. The MOPITT Version 4 (V4) product is based exclusively on TIR observations [*Deeter et al.*, 2010].

[3]   The MOPITT V4 product and V5 TIR-only product are comparable to TIR-only CO products from the AIRS (Atmospheric Infrared Sounder) [*McMillan et al.*, 2005], TES (Tropospheric Emission Spectrometer) [*Bowman et al.*, 2006] and IASI (Infrared Atmospheric Sounding Interferometer) [*George et al.*, 2009] instruments. An intercomparison of the MOPITT V5 TIR-only total column product with AIRS, TES and IASI total column products based on available data from 2000 to 2011 was recently reported [*Worden et al.*, 2013]. All of the products show reasonable agreement in seasonal variations and mean total column averages for the Northern Hemisphere while AIRS data exhibit a clear high bias in mean values for the Southern Hemisphere. Only MOPITT and AIRS had sufficient data records for trend determination and both show similar decreasing

[1]Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, Colorado, USA.

[2]Department of Earth and Planetary Science, Harvard University, Cambridge, Massachusetts, USA.

Corresponding author: M. N. Deeter, Atmospheric Chemistry Division, National Center for Atmospheric Research, Boulder, CO 80307, USA. (mnd@ucar.edu)

decadal trends in CO total column for both hemispheres and for specific regional averages in Europe, E. USA and E. China. The MOPITT V5 NIR-only product exploits the same spectral band as the SCIAMACHY CO product [*de Laat et al.*, 2007], but is only available in clear-sky daytime scenes over land. The MOPITT multispectral TIR/NIR product is available for all scenes (day/night, land/ocean), but for nighttime/land and all ocean scenes, the retrieval actually exploits only the TIR channels.

[4] This manuscript presents validation results and analysis for MOPITT V4 and V5 products using in situ CO profiles measured from an aircraft. The remainder of this paper is organized as follows. Significant enhancements in the V5 retrieval algorithm and V5 data quality issues are discussed in section 2. Section 3 describes the in situ data sets used for this work. Section 4 presents validation results in the form of scatter plots comparing operational retrieval results and expected retrieval results based on the in situ profile data, MOPITT averaging kernels and a priori. Results are shown for the MOPITT V4 product in addition to the three variants of the MOPITT V5 product. The data are analyzed further in Section 5 to quantify retrieval temporal bias drift and investigate the geographical variability of the retrieval bias. Conclusions are reported in section 6.

2. Version 5 Products

2.1. V5 Retrieval Algorithm Features

[5] The MOPITT V5 products are generated with an iterative optimal estimation-based retrieval algorithm, which is very similar to the algorithm used for the V4 product [*Deeter et al.*, 2010]. The V4 and V5 retrieval algorithms both perform CO retrievals of log(VMR) on the same 10-level retrieval grid (surface, 900 hPa, 800 hPa, ..., 100 hPa) and use the same a priori profiles and a priori covariance matrix. All MOPITT CO retrievals are based on some subset of the Average and Difference radiances from MOPITT channels 5, 6, and 7. Each channel is associated with a specific gas correlation radiometer [*Pan et al.*, 1998]. V5 TIR-only retrievals are based on the 5A, 5D, and 7D radiances, the same radiance subset used for the V4 product [*Deeter et al.*, 2010]. V5 NIR-only retrievals are based solely on the ratio of the 6D and 6A radiances. V5 TIR/NIR retrievals exploit the 5A, 5D, 7D, 6D, and 6A radiances. The NIR-only and TIR/NIR retrieval products exploit a new feature in the retrieval algorithm, which increases the influence of the NIR measurements at the expense of increased random retrieval error [*Deeter et al.*, 2012].

[6] Unlike earlier products, all V5 products are processed using a forward model in the retrieval algorithm which explicitly accounts for long-term instrumental changes. Whereas all V3 and V4 products were based on a static radiative transfer model assuming fixed instrumental parameters [*Edwards et al.*, 1999], the instrument state in the V5 operational radiative transfer model is represented differently for each month of the mission. Monthly mean operating temperatures and pressures of MOPITT's gas correlation cells are used to develop the regression coefficients used by the operational radiative transfer model ("MOPFAS") for each month. Gradual changes in these instrumental parameters were earlier identified as a source of long-term drift in MOPITT retrieval biases [*Emmons et al.*, 2009]. The decision to represent the instrument state with monthly-mean parameters was based on the finding that instrumental changes within a single month do not typically result in significant retrieval biases whereas changes over longer periods (e.g., a year or more) cause significant drift. Thus, temporal discontinuities in the retrieval results across monthly boundaries should not be evident. A consequence of this new modeling strategy is that the forward model coefficients for months before the MOPITT instrumental "anomaly" in 2001, caused by the failure of one of MOPITT's two coolers [*Deeter et al.*, 2004], are now generated in the same manner as for months after the anomaly. Thus, the distinction between the periods before and after the anomaly ("Phase 1" and "Phase 2") is not relevant for V5 products as it was for earlier products.

[7] In addition, all V5 retrieval products are based on a new method for calculating radiance uncertainties for MOPITT's length-modulation cell (LMC) channels [*Deeter et al.*, 2011]. The new method accounts for both instrumental noise and "geophysical noise," i.e., random errors in the calibrated radiances resulting from the combined effects of field of view motion and fine-scale spatial variability in surface radiative properties during each observation [*Deeter et al.*, 2011]. All earlier MOPITT retrieval products only accounted for instrumental noise. Over land, the magnitude of geophysical noise varies strongly, even for adjacent MOPITT pixels, and is often much greater than the instrumental noise. In the retrieval algorithm, the more rigorous method for determining V5 radiance uncertainties can change the effective weighting of the radiances and should yield more reliable Level 2 retrieval uncertainties.

[8] The cloud detection algorithm for V5 retrieval processing has also been modified. The MOPITT cloud detection algorithm exploits MOPITT thermal channel radiances as well as the MODIS cloud mask product to determine whether MOPITT radiances are affected by clouds in the field of view. Except for the case of scenes where the MODIS cloud mask indicates low clouds, only clear-sky observations are processed by the retrieval algorithm. For V5 products, the treatment of the MODIS cloud mask has been substantially improved. Whereas the previous cloud detection method (used in V3 and V4 processing) exploited a parameterization to identify the set of MODIS 1-km pixels within the boundaries of each MOPITT observation, the new method relies on a precalculated lookup table of MODIS relative offset indices for each MOPITT observation based on the scan indices (i.e., the "pixel" and "stare" indices) which uniquely identify its position in the scan pattern. The new cloud mask collocation method is numerically more efficient and reduces the number of incorrectly identified MODIS pixels by about a factor of two.

[9] Finally, a subtle difference between the V4 and V5 retrieval algorithms involves the association between retrieved levels and layers. For V5 products, each retrieval level corresponds to the layer immediately above that level. Within each layer VMRs are assumed constant. Therefore, for example, the V5 retrieval product for 700 hPa corresponds to the mean VMR for the layer between 700 and 600 hPa. The topmost retrieval level represents the layer between 100 and 50 hPa. In contrast, the V4 layering scheme employed non-uniformly weighted layers [*Deeter*, 2009]. Quantitatively, this change could be important when com-

Table 1. NOAA Validation Site Locations

NOAA Validation site	Latitude	Longitude	Observational period
Briggsdale, Colorado	40.37°N	104.3°W	13 Jan 2000–30 Aug 2011
Molokai, Hawaii	21.23°N	158.95°W	31 Jan 2000–22 Apr 2008
Poker Flat, Alaska	65.07°N	147.29°W	7 Feb 2000–26 Aug 2011
Harvard Forest, Massachusetts	42.54°N	72.17°W	8 Feb 2000–18 Nov 2007
Rarotonga, Cook Islands	21.25°S	159.83°W	17 Apr 2000–25 Jun 2011
Charleston, South Carolina	32.77°N	79.55°W	22 Aug 2003–28 Sep 2009
Norfolk, Virginia	36.70°N	75.50°W	28 Jul 2004–29 Apr 2007
Worcester, Massachusetts	42.95°N	70.63°W	5 Jan 2008–18 Aug 2011
Sinton, Texas	27.73°N	96.86°W	8 Jan 2008–16 Aug 2011
Cape May, New Jersey	38.83°N	74.32°W	10 Jan 2008–17 Aug 2011
Beaver Crossing, Nebraska	40.80°N	97.18°W	19 Jan 2008–17 Apr 2011
West Branch, Iowa	41.72°N	91.35°W	19 Jan 2008–31 Aug 2011
East Trout Lake, Saskatchewan	54.35°N	104.98°W	26 Jan 2008–19 Aug 2011
Dahlen, North Dakota	47.50°N	99.24°W	11 Mar 2008–28 Jun 2011
Trinidad Head, California	41.05°N	124.15°W	22 Mar 2008–8 Aug 2011

paring MOPITT products to model output or in situ data exhibiting strong vertical gradients.

2.2. V5 Data Quality

[10] Two MOPITT data quality issues have recently been identified and are described below. The first involves a systematic bias in reported MOPITT geolocation data. The second issue relates to the use of climatological water vapor data in the MOPITT Level 2 retrieval processing. These two issues affect V4 and V5 MOPITT products, although, as discussed in sections 4.1.5 and 4.2.3, they appear to produce a negligible effect on the validation results. Nevertheless, these issues could be significant in specific applications of the MOPITT data. For example, the geolocation problem should be considered in the analyses of CO distributions in urban regions.

[11] The error in the MOPITT geolocation values (latitude and longitude) appears to be the result of a small angular misalignment between the MOPITT instrument and the Terra spacecraft. This issue affects both current (V4 and V5) and previous (V3) MOPITT operational products. A report analyzing these errors and describing a first-order correction method is available at http://web3.acd.ucar.edu/mopitt/GeolocationBiasReport.pdf. In descending MOPITT overpasses (i.e., when the Terra satellite is heading south), reported longitudes contained in the Levels 1 and 2 data files appear to be systematically biased by 0.3–0.4° to the west of MOPITT's actual field of view. Reported latitudes in descending overpasses do not appear to be significantly biased. The magnitude of the geolocation error is similar to the size of a single MOPITT field of view (22 km). Geolocation errors for ascending overpasses are different than errors for descending overpasses because of the different orbital geometry. A comprehensive method for eliminating this geolocation error at the initial stage of data processing will be implemented in the next operational MOPITT product (Version 6). The effects of geolocation errors on the MOPITT validation results are considered in sections 4.1.5 and 4.2.3.

[12] To accurately simulate the dependence of the MOPITT radiances on CO concentration, the MOPITT
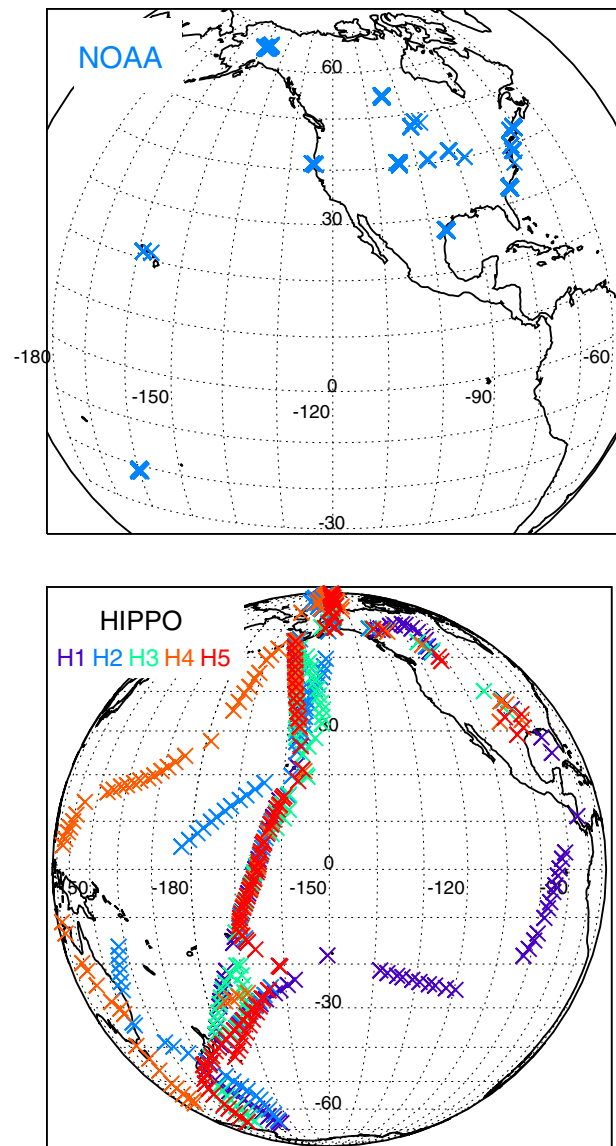

Figure 1. Maps showing locations of (top panel) NOAA profiles and (bottom panel) HIPPO profiles.
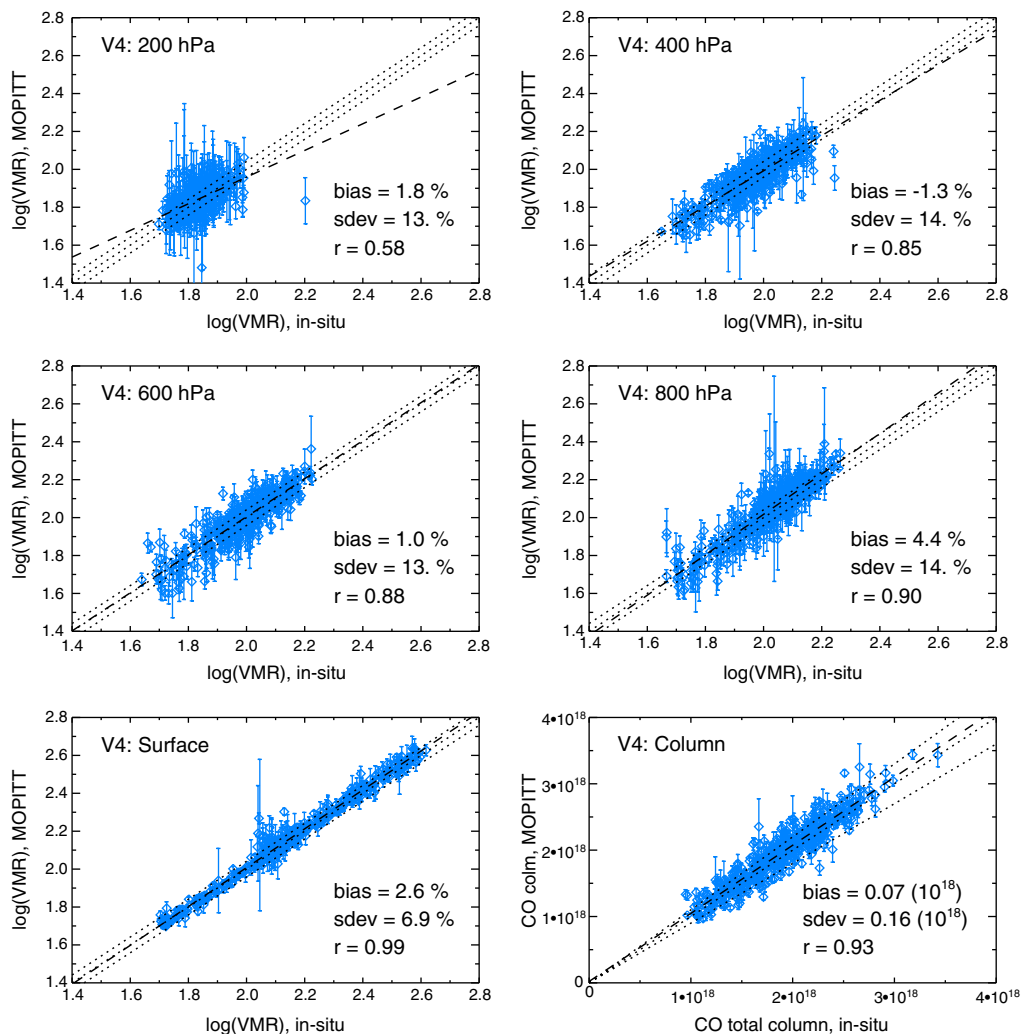
Figure 2. Scatter plots showing V4 validation results based on NOAA profiles. CO total column values in the bottom right panel are reported in units of molecules/cm^2. Dotted lines in each panel indicate the ideal one-to-one dependence and $\pm 10\%$ error boundaries. Error bars attached to each data point indicate the associated standard deviation of the retrieved log(VMR) values for each overpass. The dashed line in each panel shows the least-squares best fit.

retrieval algorithm requires ancillary data sources for atmospheric temperature and water vapor profiles. Current MOPITT products rely on NCEP analysis for these profiles [*Deeter et al.*, 2003]. However, as described in the MOPITT Version 4 User's Guide [*Deeter*, 2009], NCEP water vapor profiles occasionally include non-physical values, which prevent normal execution of the retrieval algorithm. In these cases, climatological NCEP water vapor concentrations are substituted into the operational NCEP water vapor profiles in place of the non-physical values. Retrievals based on climatological water vapor values can be identified using the standard MOPITT Level 2 diagnostic labeled "Water Vapor Climatology Content." While retrievals based on climatological water vapor profiles are usually physically reasonable, rare scenes have been observed where CO retrievals exhibit anomalously large concentrations near the surface. The effects of climatological water vapor values on the MOPITT validation results are discussed in sections 4.1.5 and 4.2.3.

3. Validation Data Sets

3.1. NOAA Profiles

[13] In situ CO vertical profiles produced through NOAA's flask sampling program have been exploited in several previous MOPITT validation papers [*Emmons et al.*, 2004; *Emmons et al.*, 2009; *Deeter et al.*, 2010]. Flask samples acquired on aircraft are processed by the Global Monitoring Division of NOAA's Earth System Research Laboratory (ESRL). Locations of NOAA stations used for V5 validation are listed in Table 1 and are plotted in the top panel of Figure 1. The large majority of the stations are located in North America. Many of the stations became operational within the last 5 years and were therefore not exploited in previous MOPITT validation studies. Flask samples are typically acquired from near the surface up to about 350–400 hPa. Typical in situ profiles are derived from approximately 12–15 flask samples. In order to obtain a complete validation profile for comparison with MOPITT

Table 2. Summarized Validation Results (Bias, Standard Deviation, and Correlation Coefficient)[a]

		Total Column	Surface	800 hPa	600 hPa	400 hPa	200 hPa
V4	bias	0.07	2.6	4.4	1.0	−1.3	1.8
	sdev	0.16	6.9	14.	13.	14.	13.
	r	0.93	0.99	0.90	0.88	0.85	0.58
	drift	0.022 +/− 0.002	0.05 +/− 0.10	0.55 +/− 0.21	1.36 +/− 0.20	2.20 +/− 0.19	1.33 +/− 0.20
V5T	bias	0.06	1.0	−0.1	0.1	6.0	7.9
	sdev	0.15	8.4	9.7	10.	13.	11.
	r	0.94	0.98	0.94	0.91	0.82	0.64
	drift	0.000 +/− 0.002	−0.44 +/− 0.12	−0.76 +/− 0.13	−0.53 +/− 0.14	0.59 +/− 0.19	0.81 +/− 0.17
V5N	bias	0.08	3.1	3.1	3.5	3.7	2.3
	sdev	0.17	8.1	6.9	7.8	8.4	5.7
	r	0.91	0.97	0.89	0.89	0.89	0.81
	drift	−0.003 +/− 0.004	−0.16 +/− 0.19	−0.13 +/− 0.16	−0.21 +/− 0.19	−0.22 +/− 0.20	−0.12 +/− 0.14
V5J	bias	0.08	2.7	−1.1	−4.6	4.0	14.
	sdev	0.19	16.	16.	11.	16.	19.
	r	0.90	0.92	0.80	0.84	0.68	0.48
	drift	0.001 +/− 0.003	−1.05 +/− 0.25	−1.63 +/− 0.23	−0.80 +/− 0.18	1.53 +/− 0.24	2.33 +/− 0.28

[a]Based on in situ data from NOAA validation sites. Bias and standard deviation statistics for the total column are in units of 10^{18} molecules/cm^2. Bias and standard deviations for retrieval levels are expressed in percent. Total column drift is in units of 10^{18} molecules/cm^2/yr. Drift for the retrieval levels is expressed in percent/yr.
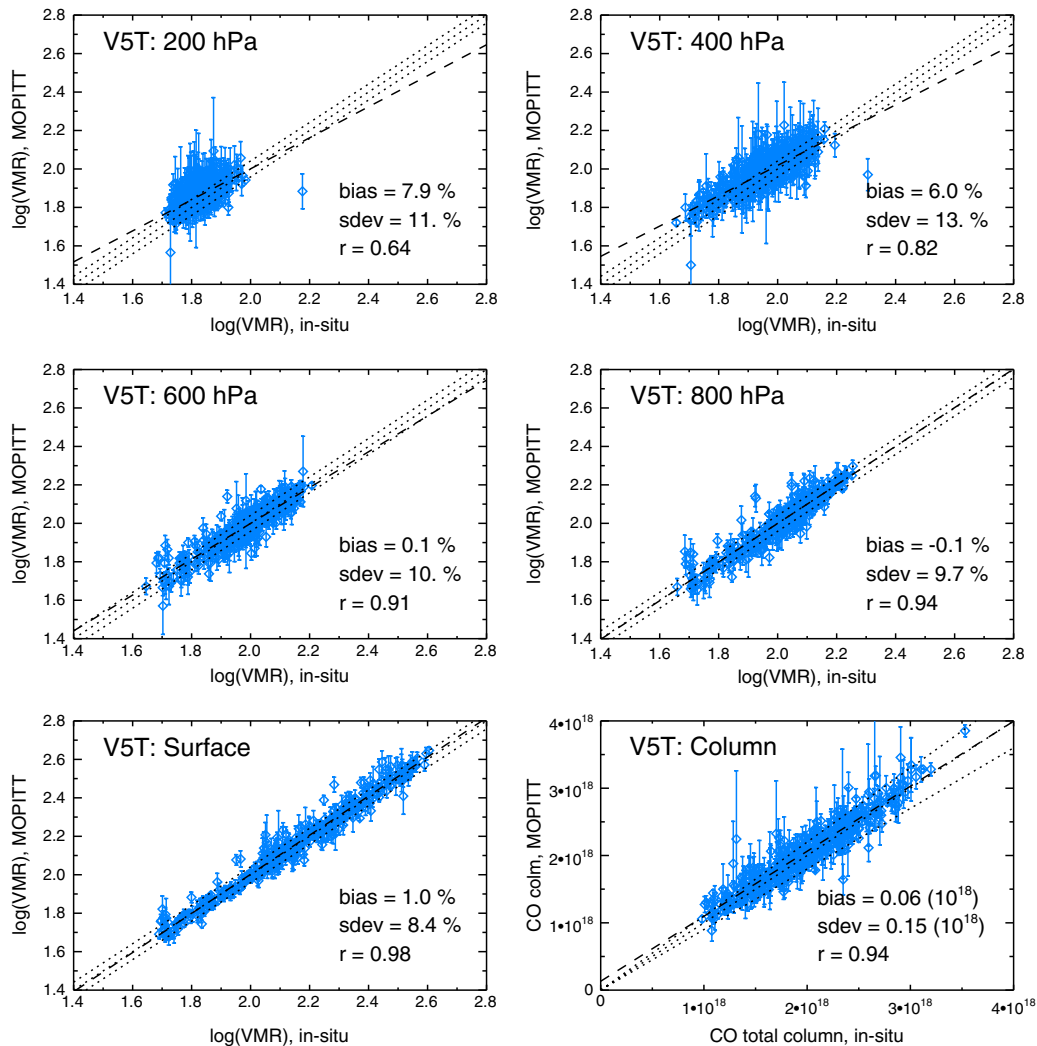

Figure 3. Scatter plots showing V5 TIR-only validation results based on NOAA profiles. See caption of Figure 2.
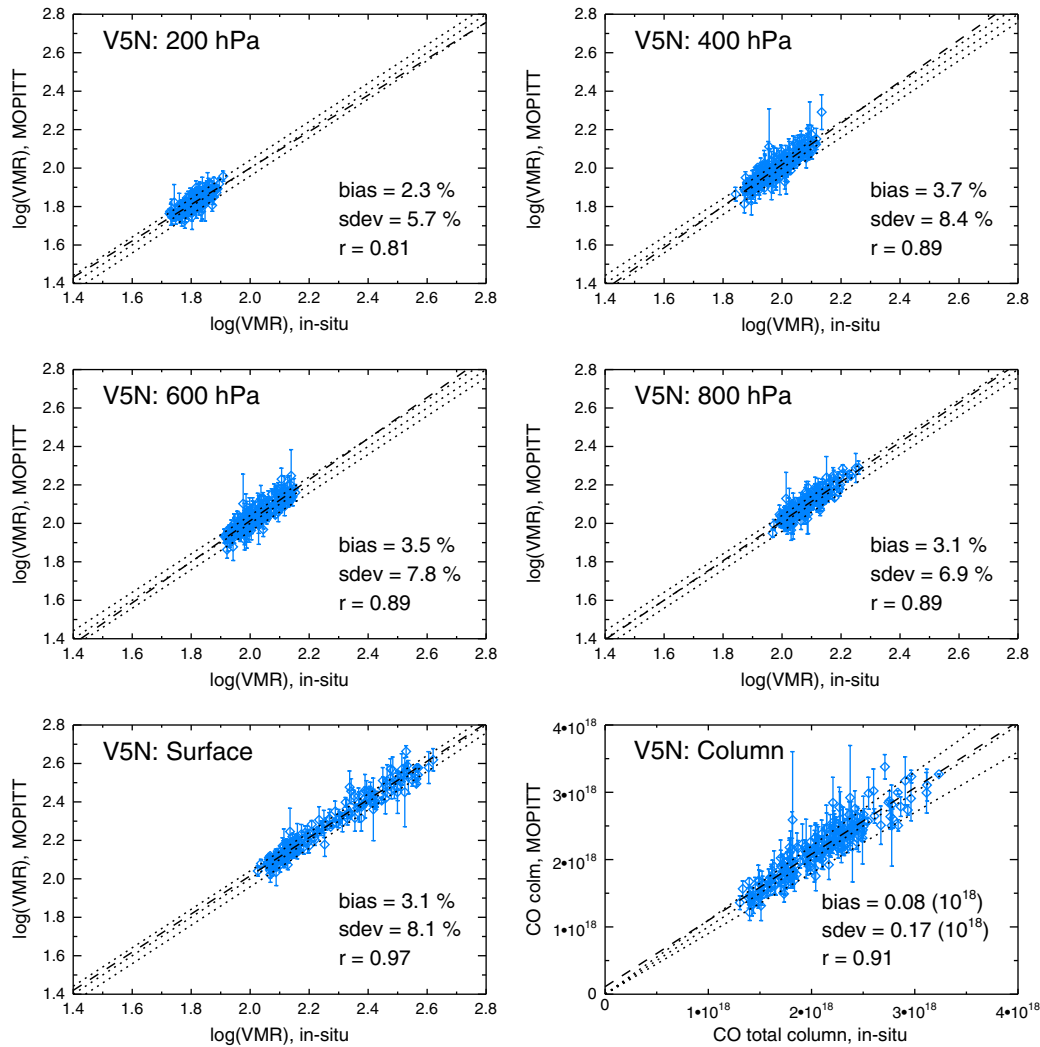
Figure 4. Scatter plots showing V5 NIR-only validation results based on NOAA profiles. See caption of Figure 2.

retrievals, each in situ profile is extended vertically above the highest in situ measurement using the MOZART chemical transport model and then resampled to the standard pressure grid used for the MOPITT operational radiative transfer model [*Emmons et al.*, 2004]. This implies that the NOAA profiles are probably less valuable for validating the highest MOPITT retrieval levels (e.g., 100 and 200 hPa) than for lower levels. The entire database of NOAA aircraft profiles acquired during the MOPITT mission currently includes more than 2000 CO profiles.

3.2. HIPPO Profiles

[14] The "HIAPER Pole to Pole Observations" (HIPPO) campaign included five phases of operations between 2009 and 2011 [*Wofsy et al.*, 2011]. Observations were made during January 2009 (Phase 1), November 2009 (Phase 2), April 2010 (Phase 3), June 2011 (Phase 4), and August/September 2011 (Phase 5). The focus of the campaign was on investigations of the carbon cycle and greenhouse gas distributions throughout the troposphere. In situ measurements of atmospheric composition were performed over a wide latitudinal range (from approximately 67°S

to 80°N) mostly over the Pacific Ocean, and over a wide altitude range (from the surface up to pressures of 150–300 hPa). As demonstrated in section 5.2, the extensive coverage of the HIPPO flights allows this data set to be used to analyze the geographical dependence of retrieval biases. Moreover, HIPPO observations in the upper troposphere (e.g., pressures between 150 and 300 hPa) are particularly valuable for validating the higher MOPITT retrieval levels (e.g., 200 hPa). CO VMR measurements were performed during HIPPO at 1 Hz sampling with the Quantum Cascade Laser Spectrometer (QCLS) instrument [*Jimenez et al.*, 2005; *McManus et al.*, 2010]. This instrument measures absorption from CO infrared transition lines at 4.59 μm using molecular line parameters from the HITRAN database [*Rothman et al.*, 2009]. CO measurements have one-sigma precision of 0.15 ppb and accuracy of 3.5 ppb. Additional information about the QCLS instrument can be found at http://hippo.ucar.edu/instruments/chemistry#qcls. Locations of the HIPPO in situ profiles are shown in the bottom panel of Figure 1. A total of 567 in situ CO profiles acquired during the five phases of HIPPO were used for MOPITT validation.
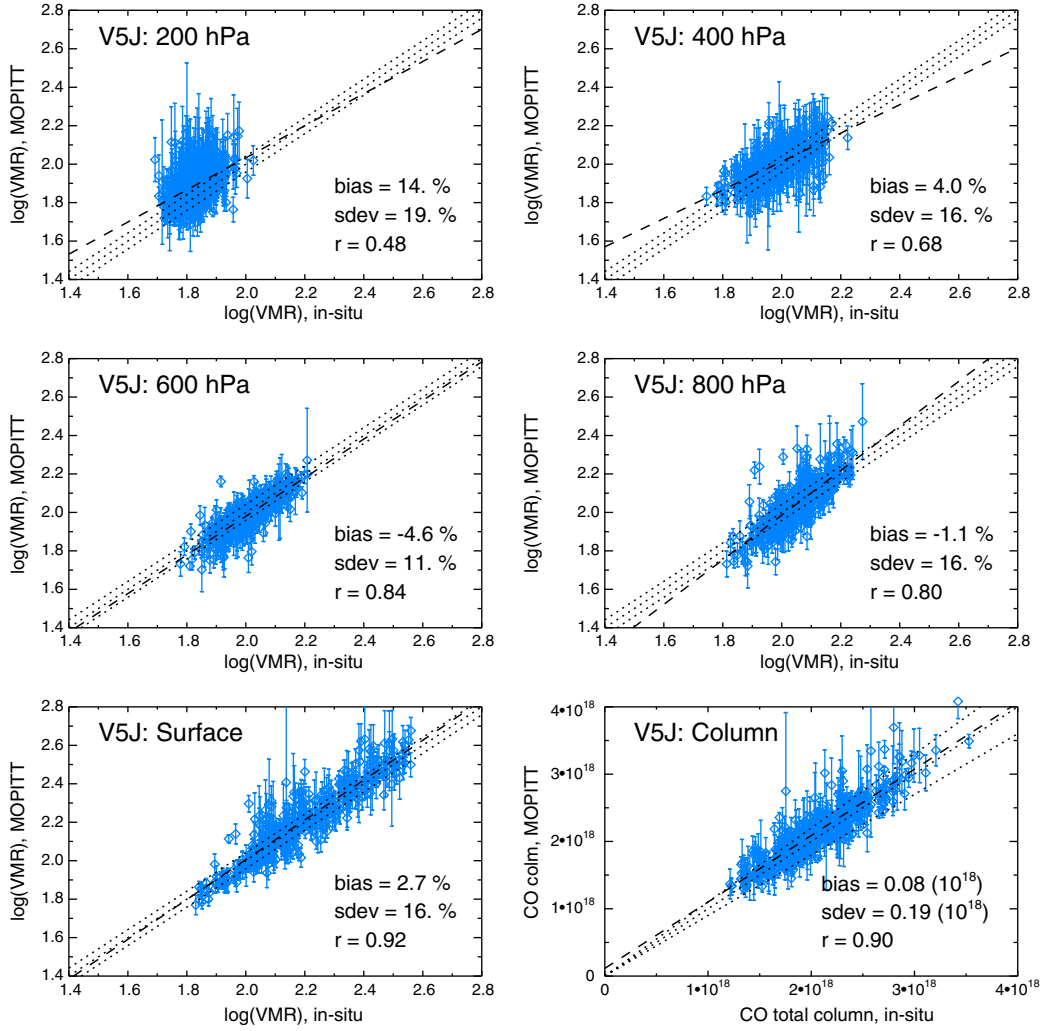
Figure 5. Scatter plots showing V5 TIR/NIR validation results based on NOAA profiles. See caption of Figure 2.

4. Validation Results

[15] Retrieval validation involves comparisons of MOPITT retrieval products (CO VMR profiles and total columns) with in situ measurements. For this purpose, we consider the in situ measurements to be exact and assume that the in situ vertical profiles are representative horizontally over an extended region (within 50 km for the NOAA profiles and 200 km for the HIPPO profiles) around the sampling location. Because of the coarseness of the radiance weighting functions (or "Jacobians") and the underconstrained nature of the retrieval process, retrieval products obtained with optimal estimation-type retrieval algorithms are constrained by a priori information as well as the measurements [*Pan et al.*, 1998; *Rodgers*, 2000]. A priori information is represented by (1) an a priori profile x_a and (2) an a priori covariance matrix, which determine the strength of the a priori constraint. The relationship between the true profile x_{true}, a priori profile and retrieved profile x_{rtv} is expressed by the equation

$$x_{rtv} = x_a + A(x_{true} - x_a) \qquad (1)$$

where A is the retrieval averaging kernel matrix. For V4 and V5 products, the vector quantities x_{true}, x_a and x_{rtv} are expressed in terms of log(VMR) rather than VMR. A quantifies the sensitivity of the retrieved profile to the true profile and is provided as a diagnostic for each retrieval in the MOPITT V4 and V5 products. A depends on the weighting functions, a priori covariance matrix and instrument error covariance matrix. (To be precise, the dependence of the weighting functions on x_{rtv} implies that A also depends on x_{rtv} and therefore that equation (1) is only valid to first-order.) Thus, when x_{true} is known (from in situ measurements, for example), equation (1) provides a formula for calculating equivalent retrievals, which account for the inclusion of a priori information and the smoothing effect of the averaging kernel matrix [*Rodgers*, 2000].

[16] The MOPITT-retrieved total column values are compared with equivalent total column values C_{sim} calculated according to

$$C_{sim} = C_a + a(x_{true} - x_a) \qquad (2)$$

where C_a is the a priori total column value and a is the total column averaging kernel. The method for calculating a
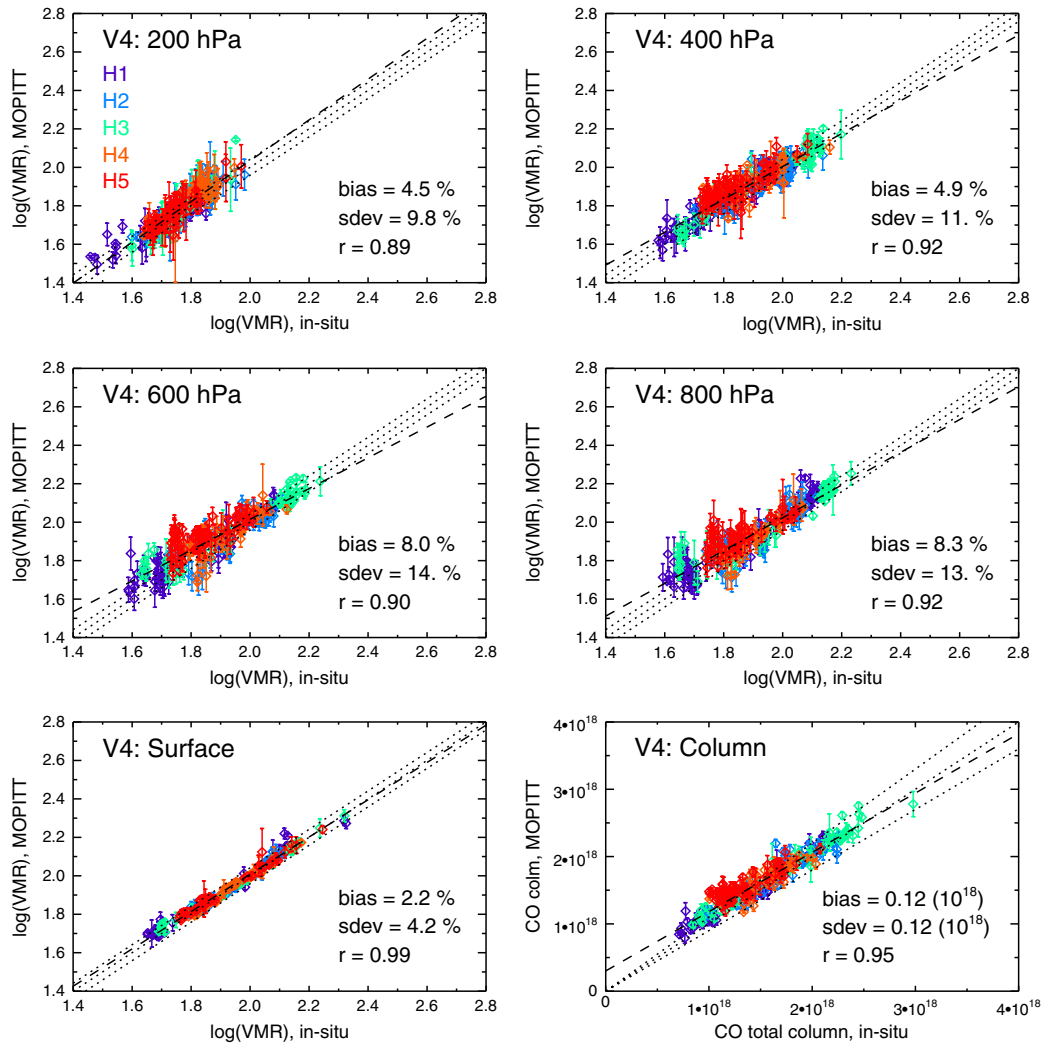
Figure 6. Scatter plots showing V4 validation results based on HIPPO profiles. See caption of Figure 2. Results for each of the five phases of HIPPO are color-coded as indicated in the top left panel.

from *A* is detailed in the MOPITT Version 4 User's Guide [*Deeter*, 2009].

4.1. NOAA Profiles

4.1.1. V4

[17] Scatter plots comparing MOPITT V4 retrieval results and corresponding equivalent retrievals based on the NOAA aircraft in situ profiles are shown in Figure 2. Separate panels present results for 200, 400, 600 and 800 hPa, the surface-level retrieval and the retrieved total column. Each plotted point indicates (1) the mean in situ-based log(VMR)

or total column value on the horizontal axis and (2) the mean retrieved log(VMR) or total column value on the vertical axis. For each overpass, plotted log(VMR) averages are calculated for all MOPITT observations acquired within 50 km of the in situ profile and within 12 h of the time at which the in situ profile was measured. Dotted lines in each panel indicate the ideal one-to-one dependence and ±10% error boundaries. Error bars attached to each data point indicate the associated standard deviation of the retrieved log(VMR) values for each overpass. The dashed line in each panel shows the least-squares best

Table 3. Summarized Validation Results Based on in Situ Data From Hippo Field Campaign. See Caption of Table 2

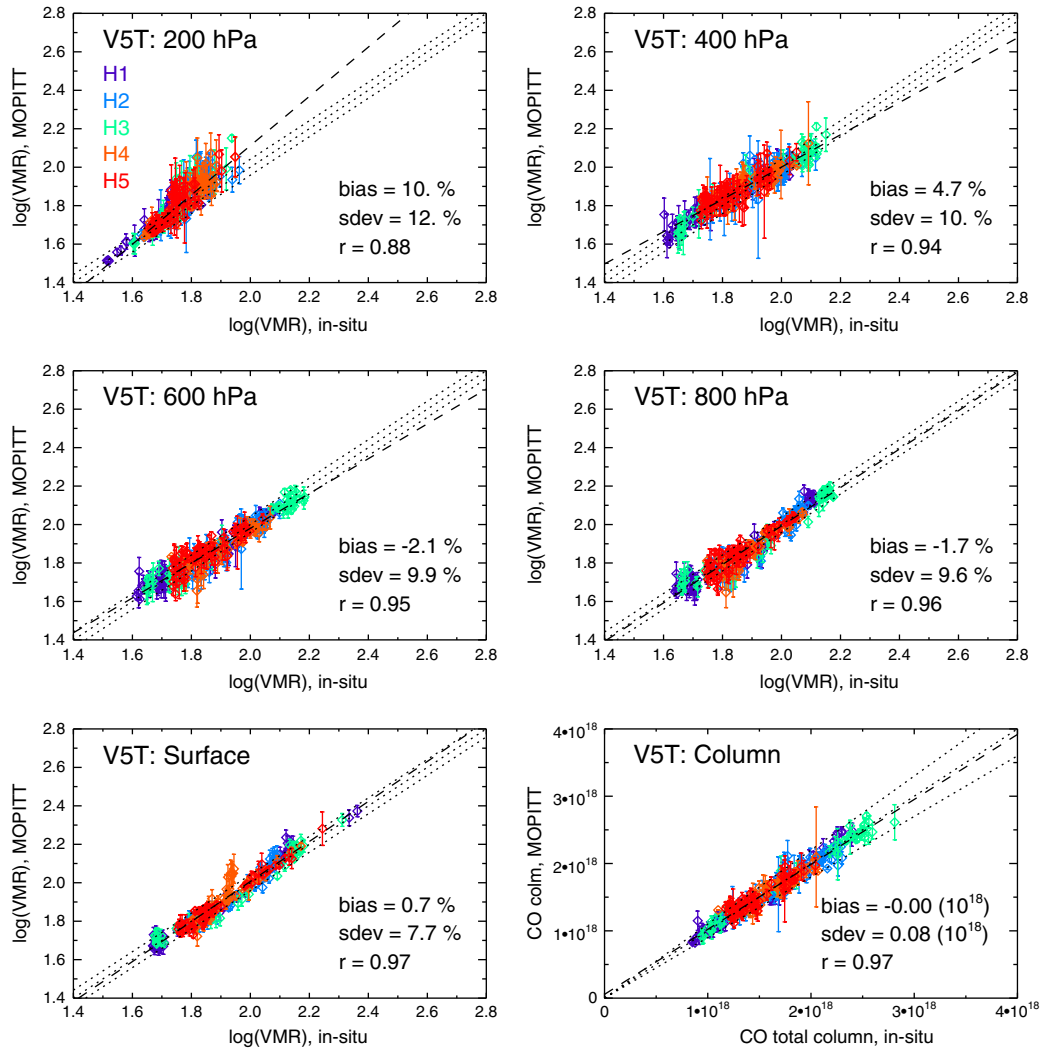|		Total Column	Surface	800 hPa	600 hPa	400 hPa	200 hPa
V4	bias	0.12	2.2	8.3	8.0	4.9	4.5
	sdev	0.12	4.2	13.	14.	11.	9.8
	r	0.95	0.99	0.92	0.90	0.92	0.89
V5T	bias	0.00	0.7	−1.7	−2.1	4.7	10.
	sdev	0.08	7.7	9.6	9.9	10.	12.
	r	0.97	0.97	0.96	0.95	0.94	0.88

Figure 7. Scatter plots showing V5 TIR-only validation results based on HIPPO profiles. See caption of Figure 2.

fit. Overall bias, standard deviation, and correlation coefficient are listed on each panel and are also summarized in Table 2.

[18] The VMR retrieval biases vary from –1.3% at 400 hPa to 4.4% at 800 hPa. The total column retrieval bias is about 0.07×10^{18} molecules/cm^2. Standard deviation values at most levels are approximately 13–14%, and correlation coefficients vary from 0.58 to 0.99. The lowest correlation coefficient is observed at 200 hPa. This is possibly due to the fact that the NOAA aircrafts typically do not reach altitudes higher than 300 hPa. Above this level, the in situ profiles used for validation rely on extrapolated in situ data and model climatology [*Emmons et al.*, 2004]. The significance of the small bias (2.2%), small standard deviation (6.9%) and large correlation coefficient (0.99) at the surface is unclear since TIR-only averaging kernels for this level depend on thermal contrast conditions and are highly variable [*Deeter et al.*, 2007, 2012]. Thus, the influence of the a priori on the surface-level validation results is also highly variable.

[19] Earlier reported biases for the V4 product based on some of the same NOAA validation sites from 2001 to 2007 [*Deeter et al.*, 2010] were –0.8% at 100 hPa, –5.9% at 400

hPa, 0.4% at 700 hPa and 0.6% at the surface. Although updated validation results for the 100 hPa and 700 hPa levels are not reported here, the biases listed in Figure 2 are clearly larger than in the earlier paper. At 400 hPa, for example, the updated bias is 4.6% larger than was reported in the earlier paper. This apparent discrepancy is expected as the result of positive bias drift in the V4 product [*Emmons et al.*, 2009] and the longer observational period analyzed here. Both simulations and earlier validation results demonstrated that retrieval bias drift in the V4 product is positive at all levels, and is strongest near 400 hPa. Bias drift for the V4 and V5 products is analyzed in Section 5.1.

4.1.2. V5 TIR-only

[20] The V5 TIR-only product is based on the same radiance subset as the V4 product, but exploits the time-dependent radiative transfer model and the more rigorous noise calculation method described in section 1. V5 TIR-only retrieval results are compared with corresponding simulated retrievals in Figure 3 and are summarized in Table 2. Compared to the V4 results, retrieval bias is significantly smaller for the lower tropospheric levels (i.e., 600 hPa, 800 hPa and surface), but is significantly larger at 200 and 400 hPa. Standard deviations decrease signifi-
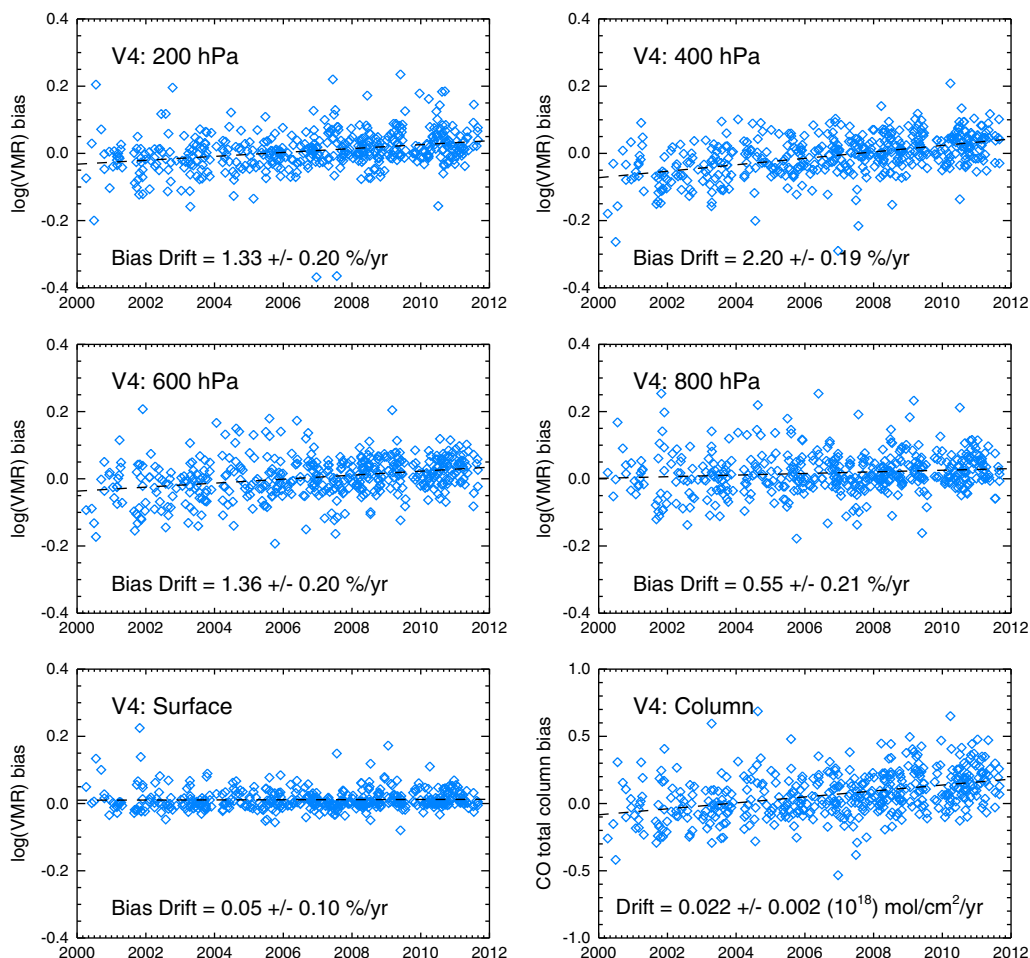
Figure 8. Time series plot showing V4 bias trends based on NOAA profiles.

cantly at 600 and 800 hPa, but increase somewhat at the surface. Correlation coefficients increase at 200, 600 and 800 hPa, but decrease at 400 hPa. Generally, the V5 TIR-only retrievals exhibit significantly improved validation statistics in the mid-troposphere, while the bias statistics for 200 and 400 hPa seem to worsen. Retrieved total column statistics are slightly better for the V5 TIR-only product than for V4.

4.1.3. V5 NIR-only

[21] The MOPITT NIR-only retrievals are based solely on the ratio of the Channel 6 Difference and Average signals [*Deeter*, 2009]. Since the weighting function for this ratio varies weakly with altitude [*Pan et al.*, 1998], NIR-only retrievals are mainly useful for constraining CO total column; NIR-only retrieved profiles contain no useful information about the CO vertical distribution. V5 NIR-only retrieval results are compared with corresponding simulated retrievals in Figure 4. Results are summarized in Table 2. The smaller data set evident in this figure (compared to Figure 3, for example) is due to the limitation of NIR-only retrievals to daytime scenes over land. NIR-only validation results indicate a positive bias at all levels ranging from about 2 to 4%. The total column retrieval bias is about 0.08×10^{18} molecules/cm^2, which is similar to the biases for the V4 and V5 TIR-only total column products. VMR standard deviations range from about 6 to 8%.

4.1.4. V5 TIR/NIR

[22] The TIR/NIR validation results shown in Figure 5 and summarized in Table 2 indicate biases ranging from about –5% at 600 hPa to 14% at 200 hPa. The total column retrieval bias is about 0.08×10^{18} molecules/cm^2, which is similar to the biases for the V5 TIR-only and NIR-only total column products. VMR standard deviations range from about 6 to 8%. Compared to the other TIR-only and NIR-only products, standard deviations are larger and correlation coefficients are smaller. As discussed previously, the TIR/NIR product exhibits relatively large random retrieval errors as the result of a strategy to increase the influence of the NIR radiances [*Deeter et al.*, 2012].

4.1.5. Data Filtering

[23] Data quality problems associated with geolocation errors and water vapor profiles were discussed in section 2.2. In particular scenes, either of these issues could contribute to the overall retrieval error, and thereby affect the retrieval validation statistics. To quantify the effect of water vapor climatology, V5 TIR-only validation results were recalculated after first discarding all retrievals where the water vapor climatology content diagnostic exceeded 0.1. For the NOAA validation results, this filter excluded approximately 17% of all MOPITT retrievals collocated with the NOAA in situ profiles, but changed the overall bias and standard deviation statistics at all levels by 1% or less.
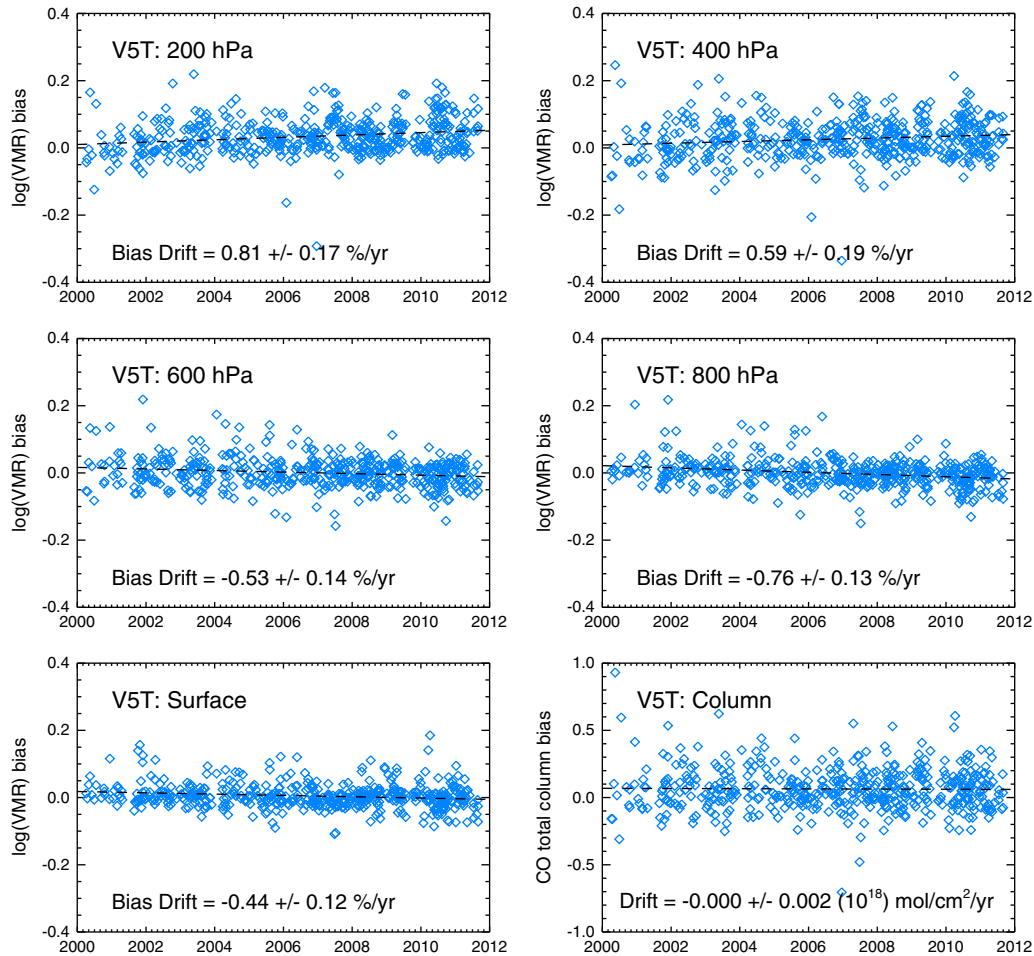
Figure 9. Time series plot showing V5 TIR-only bias trends based on NOAA profiles.

Thus, over North America, the use of climatological water vapor profiles does not significantly degrade the MOPITT validation statistics. Similarly, the validation statistics were also recalculated after applying a first-order correction to the latitude and longitude values in the MOPITT Level 2 product using a table of latitude-dependent corrections contained in the report available at http://web3.acd.ucar.edu/mopitt/GeolocationBiasReport.pdf. Again, it was found that the overall bias and standard deviation statistics changed by about 1% or less at all levels.

4.2. HIPPO Profiles

[24] As shown in Figure 1, the HIPPO campaign was primarily conducted over the Pacific Ocean. Since MOPITT NIR observations can only be exploited in daytime scenes over land, the HIPPO profiles are used here only to evaluate the V4 and V5 TIR-only retrieval products. Figure 1 also indicates that the HIPPO profiles typically were acquired far downwind from significant continental CO sources. In these remote regions, vertical and horizontal mixing should lead to relatively weak horizontal CO gradients in comparison with the regions of North America sampled by the NOAA profile data set. Therefore, whereas validation based on the NOAA profiles exploited MOPITT observations within 50 km of each in situ profile, we chose a maximum distance of 200 km for the HIPPO profiles. Relative to a 50 km radius threshold,

this choice more than doubles the number of HIPPO profiles actually exploited for validation (from 143 to 311). In addition, the larger radius results in more retrieval averaging than for the validation using the NOAA profiles, and a stronger reduction in the effects of random retrieval error.

4.2.1. V4

[25] The V4 validation results based on the HIPPO profiles are presented in Figure 6 and summarized in Table 3. Results for each of the five phases of HIPPO are color-coded. In addition to an overall shift towards weaker CO concentrations associated with remote oceanic regions, the HIPPO results in Figure 6 are clearly different than the V4 NOAA validation results presented in Figure 2 in two ways. First, at 200 hPa, the HIPPO results indicate a smaller standard deviation and a much larger correlation coefficient. The better statistics for HIPPO at this level are likely at least partially the result of the higher maximum altitude of the HIPPO in situ measurements. In addition, the HIPPO results for 600 hPa, 800 hPa, and the total column all exhibit significantly larger positive biases compared to the corresponding NOAA results. This effect could have two causes. First, as described in Section 5.1, V4 bias drift should produce larger biases over the HIPPO observational period (2009–2011) than for the NOAA observational period (2000–2011). Second, the geographical regions sampled by the NOAA and HIPPO
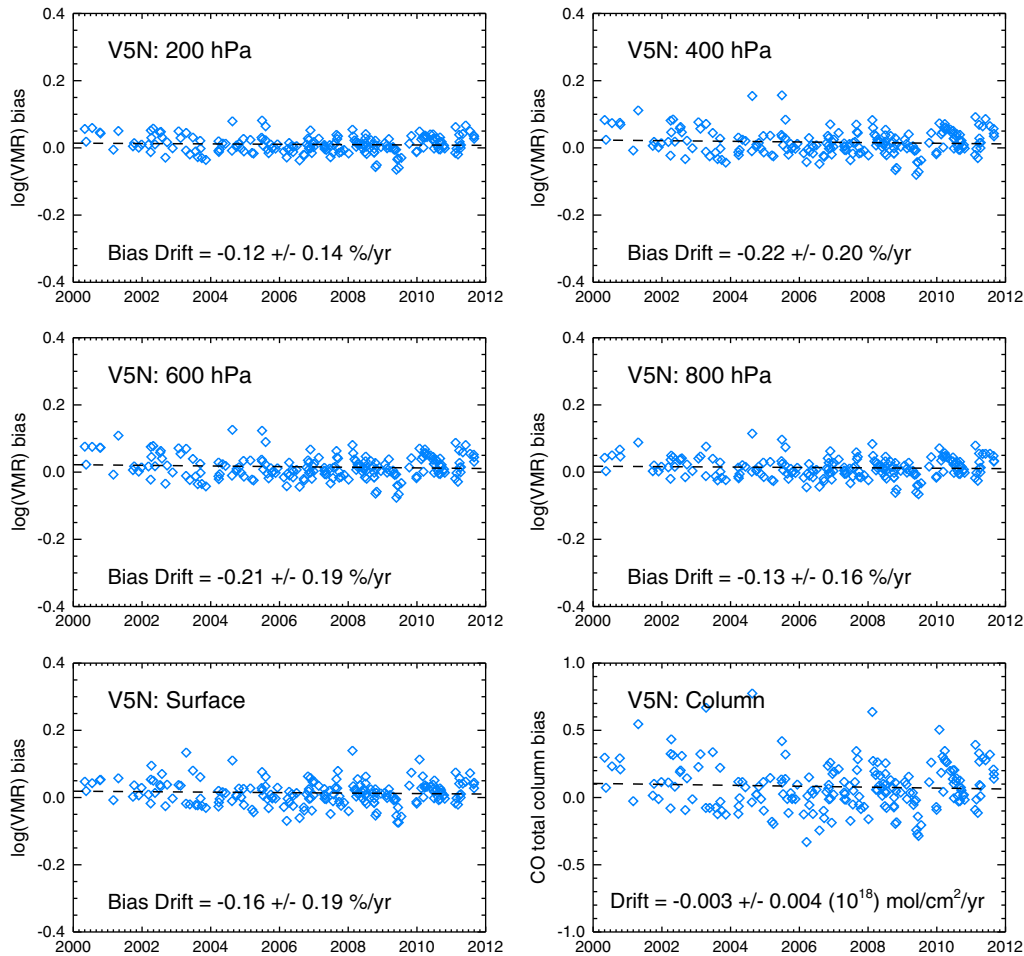
Figure 10. Time series plot showing V5 NIR-only bias trends based on NOAA profiles.

data sets are very different, and bias characteristics could be geographically dependent. This issue is addressed in section 5.2.

4.2.2. V5 TIR-only

[26] The HIPPO validation results for the V5 TIR-only product are presented in Figure 7 and summarized in Table 3. Compared to the V4 results, the V5 TIR-only validation statistics in Table 3 are clearly better for 600 hPa, 800 hPa, and the total column, and worse for 200 hPa. As indicated by the best-fit lines in Figures 6 and 7, the clearest difference between the V4 and V5 TIR-only validation results occurs for low CO VMRs. In the lower range of CO concentrations, the V4 results at 600 and 800 hPa in Figure 6 indicate a positive bias exceeding 10% whereas the V5 results in Figure 7 for the same levels indicate much weaker biases. This effect is also clearly apparent in the total column validation results. The HIPPO validation results for the V4 and V5 TIR-only products are analyzed further in Section 5.2.

4.2.3. Data Filtering

[27] The same experiments described in section 4.1.5 for characterizing the effects of water vapor climatology and geolocation errors were repeated using the HIPPO data set. As for the NOAA profiles, these experiments did not change the bias and standard deviation statistics by more than about 1% at any level. These data quality issues are therefore not considered significant with respect to validation.

5. Analysis

5.1. Long-Term Stability

[28] Throughout the MOPITT mission, in situ CO profiles over North America have been routinely acquired by NOAA using flask sampling and subsequent laboratory analysis. The continuity of this data set makes it ideal for analyzing the long-term stability of the MOPITT products and for justifying their use as climate data records. The time dependence of MOPITT V4 retrieval biases (i.e., retrieved VMR values subtracted by corresponding in situ based values) is shown in Figure 8. The dashed line shown in each panel is a least-squares best fit to the data. The slope of this line quantifies the long-term bias drift and is listed in each panel and in Table 2. For the V4 product, positive bias drift is clearly evident at all levels between 200 and 800 hPa and exceeds 2%/yr at 400 hPa. These long-term trends are roughly consistent with earlier reported V4 validation results for the period 2001–2007 [*Deeter et al.*, 2010] and bias simulations [*Emmons et al.*, 2009]. For example, the bias drift for the CO total column listed in Figure 8 is 0.022×10^{18} molecules/cm^2/yr whereas the earlier reported value was 0.018×10^{18} molecules/cm^2/yr.

[29] Bias drift time series for the V5 TIR-only, NIR-only and TIR/NIR products are shown in Figures 9, 10, and 11. Whereas the V4 product exhibits positive drift at all lev-
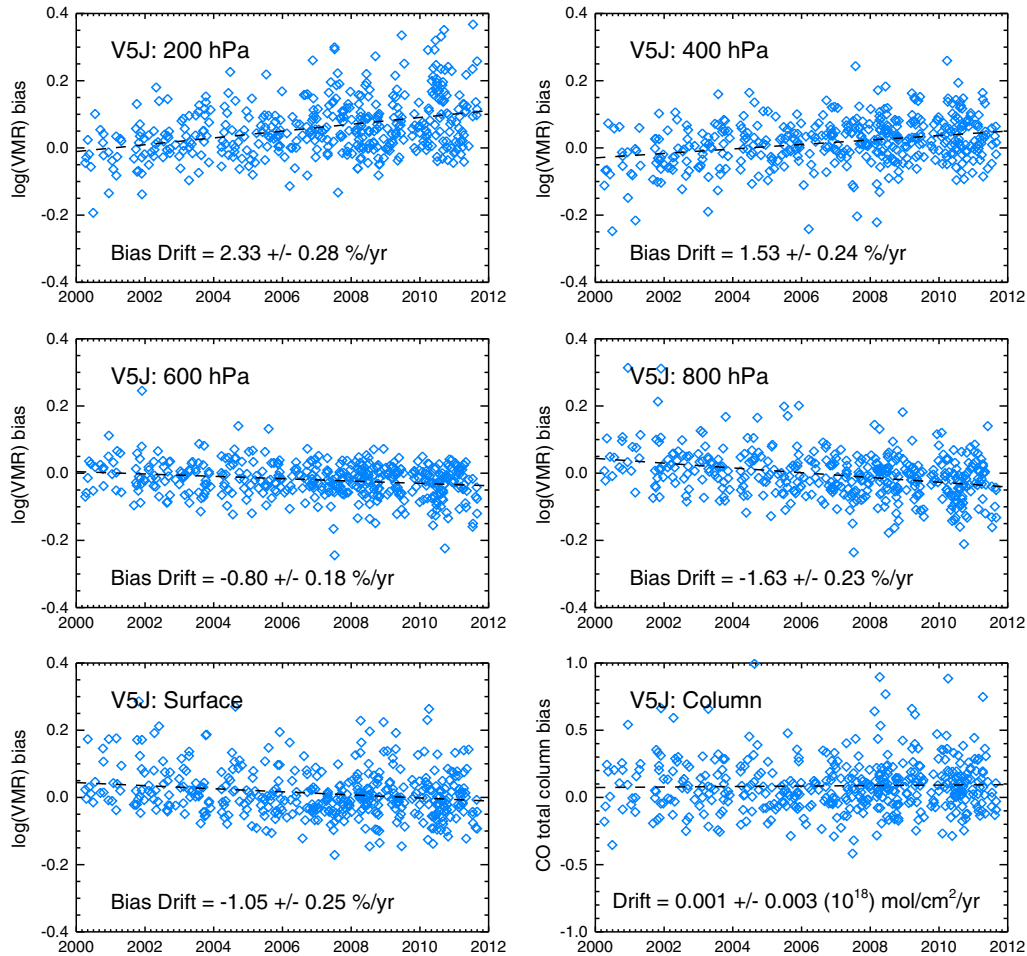
Figure 11. Time series plot showing V5 TIR/NIR bias trends based on NOAA profiles.

els except the surface, the V5 TIR-only product exhibits weaker positive drift at 200 and 400 hPa, and negative drift at 600 and 800 hPa and at the surface. The absolute values of the V5 TIR-only bias drifts are generally less than 1%/yr at all levels. Bias drift for the NIR-only product shown in Figure 10 varies from about –0.2 to –0.1%/yr. Bias drift for the TIR/NIR product varies from –1.6%/yr at 800 hPa to 2.3%/yr at 200 hPa. Bias drift in the retrieved total column appears negligible for all three types of V5 products.

[30] For both the V5 TIR-only and V5 TIR/NIR product, the negligible total column drift appears to be the result of compensating bias drifts in the upper and lower troposphere. Both the positive bias drift evident in the upper tropospheric levels and the negative bias drift indicated in the lower tropospheric levels are statistically significant, i.e., the slope uncertainty is smaller than the absolute value of the slope. This apparent "residual bias drift" is not well-understood, but could indicate either (1) some type of gradual instrumental changes unrelated to the correlation cells' temperature and pressure or (2) long-term bias drift in the NCEP meteorological data required by the MOPITT retrieval algorithm.

5.2. Geographical Variability

[31] Inverse modeling studies have found indirect evidence of a latitude-dependent bias in the V3 and V4

MOPITT products [*Kopacz et al.*, 2010; *Fortems-Cheiney et al.*, 2011], particularly in the Southern Hemisphere. With nearly complete latitudinal coverage, the HIPPO data set is well suited for investigating the geographical bias of MOPITT retrievals. MOPITT V4 retrieval biases (retrieved minus simulated) calculated with the HIPPO in situ profiles are plotted versus latitude in Figure 12. The large black diamonds and error bars in each panel indicate bias statistics (mean and standard deviation) calculated over each 30 degree-wide latitudinal zone. Results for the V4 total column, in the bottom right panel, exhibit a strong positive bias in the Southern Hemisphere, consistent with inverse modeling results [*Fortems-Cheiney et al.*, 2011]. Figure 12 also indicates that this geographical bias is even stronger at 600 and 800 hPa.

[32] Retrieval biases for the V5 TIR-only product are plotted versus latitude in Figure 13. At least three features distinguish these results from the V4 results in Figure 12. First, the observed positive biases in the V4 total column results are greatly reduced in the V5 TIR-only product. Second, as indicated by the error bars attached to the zonal-average black diamonds, the variability in the V5 bias results appears significantly smaller than for V4. This effect, which is most evident in the results for 600 hPa, 800 hPa, and total column, is probably at least partially the result of the reduced bias drift associated with V5. In the results for V4, the later
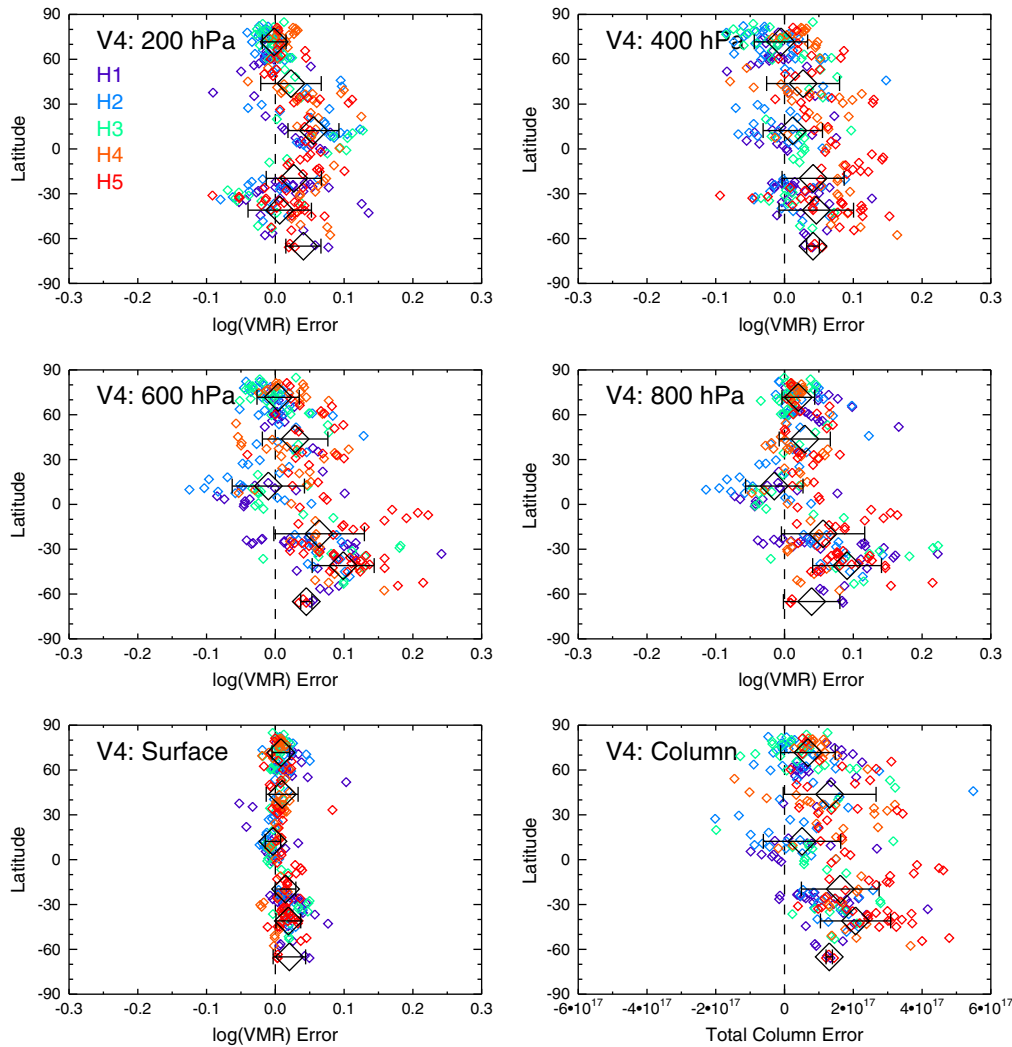
Figure 12. Latitudinal dependence of MOPITT V4 retrieval biases based on HIPPO profiles. Large black diamonds and error bars indicate mean biases and standard deviations within each 30 degree-wide latitudinal zone.

phases of HIPPO (e.g., Phase 5, plotted in red) clearly indicate larger biases than the earlier phases of HIPPO. For the V5 results, the retrieval biases do not appear to vary temporally. Third, V5 biases tend to be much larger in the Tropics than in midlatitude or polar regions. For example, the V5 validation results at 200 hPa exhibit a well-defined positive bias between the Equator and 30°N, which is much weaker in the V4 results. For the same latitudinal range, the V5 retrieval bias at 800 hPa exhibits a well-defined negative bias. The source of this latitude-dependent bias in the V5 results is not clear.

6. Conclusion

[33] Satellite remote sensing products for trace gas concentrations are subject to many potential sources of error, some of which may vary temporally or geographically. These errors are highly specific to particular instruments and may depend on the details of the retrieval algorithm used to produce the trace gas retrievals from the raw measurements.

Temporal stability is especially important to applications involving climate. Thus, rigorous validation is an essential prerequisite to the quantitative use of these products. Two complementary in situ data sets have been exploited to compare and analyze retrieval errors associated with the MOPITT Version 4 and Version 5 products for tropospheric CO. Vertical CO profiles produced under NOAA's flask sampling program permit the analysis of MOPITT retrievals over North America for the duration of the MOPITT mission. The HIPPO QCLS CO data set, acquired between 2009 and 2011, is exploited in the first global-scale validation of MOPITT products.

[34] Analysis of the V4 validation results based on the NOAA profiles reveals significant positive bias drift, exceeding 2%/yr at 400 hPa. This trend appears to mainly be the result of known gradual changes in the MOPITT gas correlation cells' operating parameters. In addition, V4 validation results based on the HIPPO profiles exhibit a substantial geographical bias in the Southern Hemisphere.

[35] The MOPITT V5 retrieval algorithm explicitly accounts for long-term changes in the gas correlation cells
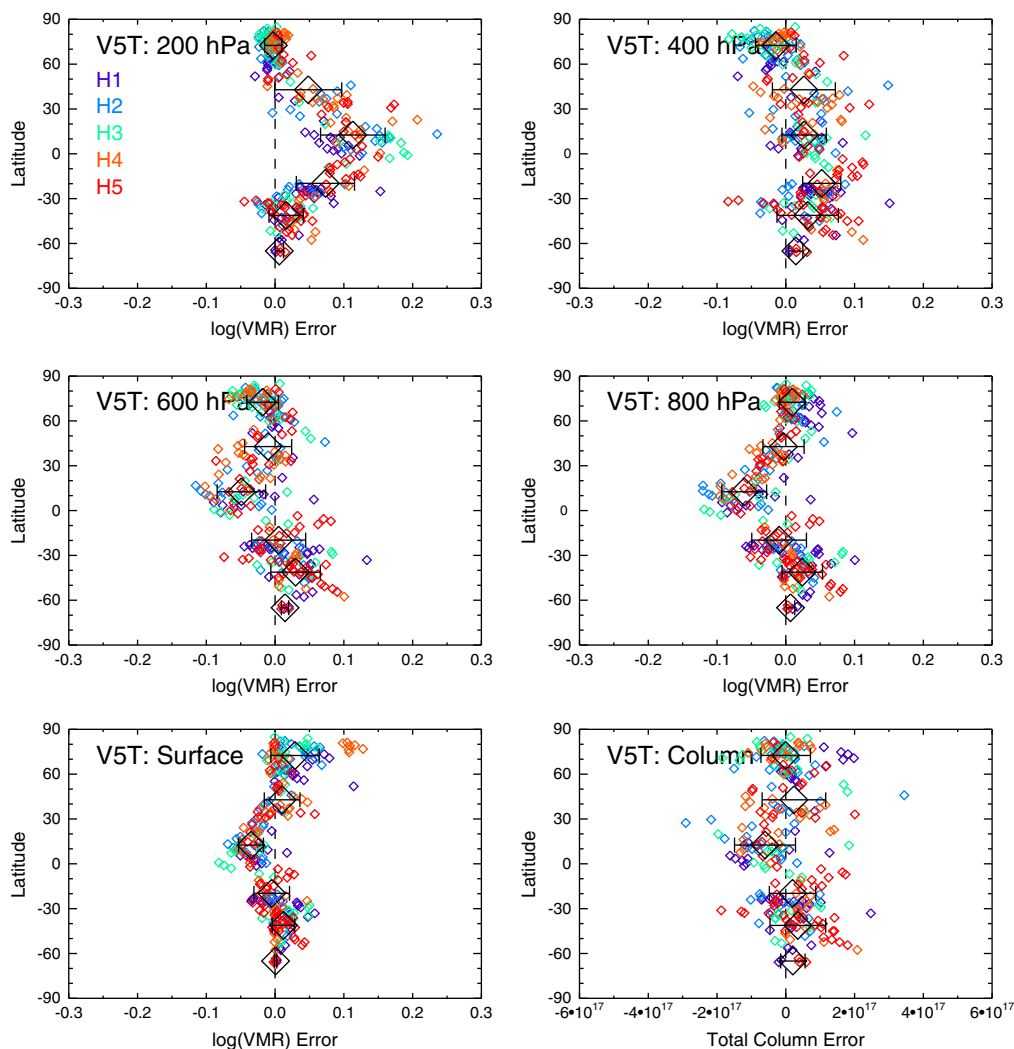
Figure 13. Latitudinal dependence of MOPITT V5 TIR-only retrieval biases based on HIPPO profiles. See caption of Figure 12.

in the MOPITT instrument, unlike the V4 retrieval algorithm. However, this feature appears to mitigate, but not eliminate, long-term bias drift in the MOPITT V5 TIR-only product. Based on the NOAA profile validation results, V5 TIR-only bias drift is less than 1%/yr at all levels, but is still statistically significant. Bias drift in the V5 TIR-only total column product appears to be negligible, as the result of opposing bias drifts in the upper and lower troposphere. HIPPO results for the V5 TIR-only product indicate much smaller biases in the total column compared to V4, especially in the Southern Hemisphere, but also indicate larger biases in the Tropics at 200 hPa. The observation that MOPITT retrieval biases exhibit a pronounced dependence on latitude and pressure (and temporal dependence) might be consistent with biases in the NCEP water vapor and temperature profiles used in retrieval processing. This effect will be investigated in the development of future MOPITT products.

[36] Validation results for the V5 NIR-only product using the NOAA profiles indicate negligible bias at all levels and no evidence of bias drift. Results for the V5 TIR/NIR product are similar to the V5 TIR-only results, but are exaggerated. The bias drift in this product exceeds 2%/yr at 200 hPa. Similar to the V5 TIR-only product, bias drift in the V5 TIR/NIR total column product appears to be negligible, due to compensating bias drifts in the upper and lower troposphere. The source of the residual bias drift in the V5 TIR-only and TIR/NIR is unclear.

[37] These results clearly indicate that the three variants of the MOPITT V5 product are not equally appropriate for all potential applications. The V5 TIR/NIR product offers the greatest vertical resolution, and particularly the greatest sensitivity to CO in the lower troposphere. This feature should benefit both inverse modeling analyses and air quality applications. However, this product also exhibits relatively large random retrieval errors and bias drift. Moreover, the main benefits of this product are only available for daytime MOPITT observations over land. Applications requiring the highest temporal stability and similar performance in variable observing situations (day and night, land and ocean) should rely on the V5 TIR-only product. The V5 NIR-only product is appropriate for the analysis of CO total columns, but is strictly limited to daytime observations over land.

References

Bowman, K. W., et al. (2006), Tropospheric emission spectrometer: Retrieval method and error analysis, *IEEE Trans. Geosci. Remote Sens.*, *44*, 1297–1307, doi:10.1109/TGRS.2006.871234.

Deeter, M. N., et al. (2003), Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, *J. Geophys. Res.*, *108*(D14), 4399, doi:10.1029/2002JD003186.

Deeter, M. N., et al. (2004), Vertical resolution and information content of CO profiles retrieved by MOPITT, *Geophys. Res. Lett.*, *31*, L15112, doi:10.1029/2004GL020235.

Deeter, M. N., D. P. Edwards, J. C. Gille, and J. R. Drummond (2007), Sensitivity of MOPITT observations to carbon monoxide in the lower troposphere, *J. Geophys. Res.*, *112*, doi:10.1029/2007JD008929.

Deeter, M. N. (2009), *MOPITT (Measurements of Pollution in the Troposphere) Validated Version 4 Product User's Guide*, National Center for Atmospheric Research. Available at http://web3.acd.ucar.edu/mopitt/v4_users_guide_val.pdf.

Deeter, M. N., et al. (2010), The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, *J. Geophys. Res.*, *115*, doi:10.1029/2009JD013005.

Deeter, M. N., H. M. Worden, D. P. Edwards, J. C. Gille, D. Mao, and J. R. Drummond (2011), MOPITT multispectral CO retrievals: Origins and effects of geophysical radiance errors, *J. Geophys. Res.*, *116*, doi:10.1029/2011JD015703.

Deeter, M. N., H. M. Worden, D. P. Edwards, J. C. Gille, and A. E. Andrews (2012), Evaluation of MOPITT retrievals of lower-tropospheric carbon monoxide over the United States, *J. Geophys. Res.*, *117*, doi:10.1029/2011JD017553.

Edwards, D. P., C. M. Halvorson, and J. C. Gille (1999), Radiative transfer modeling for the EOS Terra satellite measurements of pollution in the troposphere (MOPITT) instrument, *J. Geophys. Res.*, *104*, 16,755–16,775.

Emmons, L. K., et al. (2004), Validation of measurements of pollution in the troposphere (MOPITT) CO retrievals with aircraft in situ profiles, *J. Geophys. Res.*, *109*, doi:10.1029/2003JD004101.

Emmons, L. K., et al. (2009), Measurements of pollution in the troposphere (MOPITT) validation through 2006, *Atmos. Chem. Phys.*, *9*, 1795–1803, doi:10.5194/acp-9-1795-2009.

Fortems-Cheiney, A., et al. (2011), Ten years of CO emissions as seen from measurements of pollution in the troposphere MOPITT, *J. Geophys. Res.*, *116*, doi:10.1029/2010JD014416.

George, M., et al. (2009), Carbon monoxide distributions from the IASI/METOP mission: Evaluation with other space-borne remote sensors, *Atmos. Chem. Phys.*, *9*, 8317–8330, doi:10.5194/acp-9-8317-2009.

Jimenez, R., et al. (2005), Atmospheric trace gas measurements using a dual quantum-cascade laser mid-infrared absorption spectrometer, in *Novel In-Plane Semiconductor Lasers IV, Proc. SPIE vol. 5738, Carmen Mermelstein and David P. Bour, Eds.*, Bellingham, WA.

Kopacz, M., et al. (2010), Global estimates of CO sources with high resolution by adjoint inversion of multiple satellite datasets (MOPITT, AIRS, SCIAMACHY, TES), *Atmos. Chem. Phys.*, *10*, 855–876, doi:10.5194/acp-10-855-2010.

de Laat, A. T. J., et al. (2007), Scanning imaging absorption spectrometer for atmospheric chartography carbon monoxide total columns: Statistical evaluation and comparison with chemistry transport model results, *J. Geophys. Res.*, *112*, doi:10.1029/2007JD008256.

McManus, J. B., et al. (2010), Application of quantum cascade lasers to high-precision atmospheric trace gas measurements, *Opt. Eng.*, *49*, 111124, doi:10.1117/1.3498782.

McMillan, W. W., et al. (2005), Daily global maps of carbon monoxide from NASA's atmospheric infrared sounder, *Geophys. Res. Lett.*, *31*, L11801, doi:10.1029/2004GL021821.

Pan, L., J. Gille, D. P. Edwards, P. L. Bailey, and C. D. Rodgers (1998), Retrieval of tropospheric carbon monoxide for the MOPITT instrument, *J. Geophys. Res.*, *103*, 32,277–32,290.

Rodgers, C. D. (2000), *Inverse Methods for Atmospheric Sounding, Theory and Practice*, World Scientific, Singapore.

Rothman, L. S., et al. (2009), The HITRAN 2008 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer*, *110*, 533–572, doi:10.1016/j.jqsrt.2009.02.013.

Wofsy, S. C., et al. (2011), HIAPER pole-to-pole observations (HIPPO): Fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, *Phil. Trans. R. Soc. A*, *369*, 2073–2086, doi:10.1098/rsta.2010.0313.

Worden, H. M., M. N. Deeter, D. P. Edwards, J. C. Gille, J. R. Drummond, and P. P. Nedelec (2010), Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals, *J. Geophys. Res.*, *115*, doi:10.1029/2010JD014242.

Worden, H. M., et al. (2013), Decadal record of satellite carbon monoxide observations, *Atmos. Chem. Phys.*, *13*, 837–850, doi:10.5194/acp-13-837-2013.