https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions
Open Access
EGU

Validation of MOPITT Carbon Monoxide (CO) retrievals over urban regions

Wenfu Tang[1,2], Helen M. Worden[2], Merritt N. Deeter[2], David P. Edwards[2], Louisa K. Emmons[2], Sara Martínez-Alonso[2], Benjamin Gaubert[2], Rebecca R. Buchholz[2], Glenn S. Diskin[3], Russell R. Dickerson[4], and Yutaka Kondo[5]

[1]Advanced Study Program, National Center for Atmospheric Research, Boulder, CO, USA

[2]Atmospheric Chemistry Observations and Modeling, National Center for Atmospheric Research, Boulder, CO, USA

[3]NASA Langley Research Center, Hampton, VA, USA

[4]Department of Atmospheric and OceanicScience, University of Maryland, College Park, MD, USA

[5]National Institute of Polar Research, Tachikawa, Japan

Correspondence: Wenfu Tang (wenfut@ucar.edu)

Abstract

The performance of the Measurements of Pollution in the Troposphere (MOPITT) retrievals over urban regions has not been validated systematically, even though MOPITT observations are widely used to study CO over urban regions. Here we validate MOPITT products over urban regions using aircraft measurements from DISCOVER-AQ, SEAC[4]RS, ARIAs, A-FORCE, and KORUS-AQ campaigns. Overall, MOPITT performs reasonably well over both urban and non-urban regions, overall biases for V8J and V8T vary from -0.7% to 0.0%, and from 2.0% to 3.5%, respectively. The evaluation statistics of MOPITT V8J and V8T over non-urban regions are better than that over urban regions with smaller biases and higher correlation coefficients. We find that the performance of MOPITT V8J and V8T at high CO concentrations is not as good as that at low CO concentrations, although CO variability may tend to exaggerate retrieval biases in heavily-polluted scenes. We test the sensitivities of validation results to

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

assumptions and data filters applied during the comparisons of MOPITT retrievals and in-situ profiles. The results at the surface are insensitive to the model-based profile extension (required due to aircraft altitude limitations) whereas the results at levels with limited aircraft observations are more sensitive to the model-based profile extension. The validation results are insensitive to the allowed maximum time difference as criteria for co-location (12 hours, 6 hours, 3 hours, and 1 hour), and are generally insensitive to the radius for co-location, except for the case where the radius is small (25 km) and hence the MOPITT retrievals included in the validation become very small. Daytime MOPITT products have overall smaller biases than nighttime MOPITT products when comparing both MOPITT daytime and nighttime retrievals to the daytime aircraft observations. However, it would be premature to draw conclusions on the performance of MOPITT nighttime retrievals without nighttime aircraft observations. Applying signal-to-noise ratio (SNR) filters does not necessarily improve the overall agreement between MOPITT retrievals and in-situ profiles, likely due to the reduced number of MOPITT retrievals that result for comparison. Comparisons of MOPITT retrievals and in-situ profiles over complex urban or polluted regimes are inherently challenging due to spatial and temporal variabilities of CO within MOPITT retrieval pixels (i.e., footprints). We demonstrate the some of that errors are due to CO representativeness with these sensitivity tests, but further quantification of validation errors due to CO variability within the MOPITT footprint will require future work.

1. Introduction

The Measurements of Pollution in the Troposphere (MOPITT) instrument onboard the NASA Terra satellite has been retrieving total column amounts and volume mixing ratio (VMR) profiles of carbon monoxide (CO) using both thermal-infrared (TIR) and near-infrared (NIR) measurements since March, 2000. Besides the TIR-only and NIR-only products, MOPITT also provides the multispectral TIR-NIR product, which has enhanced the sensitivity to near-surface CO (Deeter et al., 2011, 2013; Worden et al., 2010). Since the start of the mission, the MOPITT CO retrieval algorithm has been improved and enhanced continuously (Worden et al., 2014). For example, the Version 6 product improvements included the reduction of both a geolocation bias and a significant latitude-dependent retrieval bias in the upper troposphere (Deeter et al., 2014). In the Version 7 products, a new strategy for radiance-bias correction and an improved method for

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

calibrating MOPITT's NIR radiances were included (Deeter et al., 2017). For the recently released MOPITT Version 8 products, enhancements include a new radiance bias correction method (Deeter et al., 2019). Meanwhile, the MOPITT products have been extensively evaluated and validated with in-situ measurements, though this has been done primarily over non-urban areas (Deeter et al., 2010, 2012, 2013, 2014, 2016, 2017, 2019; Emmons et al., 2004, 2007, 2009). For the past two decades, MOPITT CO products have been widely used for various applications including understanding atmospheric composition, evaluating atmospheric chemistry models, and constraining inverse analyses of CO emissions (e.g., Arellano et al., 2004, 2006, 2007; Chen et al., 2009; Edwards et al., 2006; Emmons et al., 2010; Fortems-Cheiney et al., 2011; Gaubert et al., 2016; Heald et al., 2004; Jiang et al., 2018; Kopacz et al., 2009, 2010; Kumar et al., 2012; Lamarque et al., 2012; Tang et al., 2018; Yurganov et al., 2005).

MOPITT products are particularly useful for monitoring and analyzing air pollution over urban regions because of the enhanced retrieval sensitivity to near-surface CO and the long-term record (e.g., Clerbaux et al., 2008; Girach and Nair, 2014; Jiang et al., 2015, 2018; Kar et al., 2010; Tang et al., 2019; Worden et al., 2010; Li and Liu, 2011; He et al., 2013; Aliyu and Botai, 2018; Kanakidou et al., 2011). However, the performance of MOPITT retrievals over urban regions has not yet been validated systematically. Furthermore, in situ observations of CO profiles over urban areas are limited, especially in Asia. Indeed, along with the non-urban validation exercises mentioned above, development and validation of the MOPITT retrieval algorithm relies heavily on in-situ measurements over remote regions, such as measurements from the HIAPER Pole-to-Pole Observations (HIPPO) and the Atmospheric Tomography Mission (ATom) campaigns (e.g., Deeter et al., 2013, 2014, 2017, 2019). Comparisons of MOPITT products to measurements with aircraft profiles during the Korea United States Air Quality (KORUS-AQ) campaign over South Korea have only recently been made in Deeter et al. (2019), but without explicitly analyzing MOPITT performance over urban regions.

In this study, we validate MOPITT version 8 and 7 products over urban regions by comparing with aircraft profiles that are over urban regions (as well as non-urban regions) from campaigns including: Deriving Information on Surface conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ); the Studies of Emissions and Atmospheric Composition, Clouds, and Climate Coupling by Regional Surveys (SEAC^{4}RS); the

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions

Air Chemistry Research In Asia (ARIAs); the Aerosol Radiative Forcing in East Asia (A-FORCE); and KORUS-AQ. These campaigns are introduced in Section 2, along with a brief introduction of the MOPITT products and the validation methodology used. We present the validation results and discuss the impacts of key factors in the retrieval process on the retrieval results in Section 3. In Section 4, we discuss the sensitivities of results to the assumptions and data filters made for aircraft-satellite comparisons not only in this study, but also in previous evaluation studies of MOPITT and other satellite products. Section 5 gives the conclusions of the study.

2. Data and methods

2.1 MOPITT retrievals and products

MOPITT is a nadir sounding satellite instrument flying on the NASA Terra satellite. It uses a gas filter correlation radiometer and measures at both the TIR band near 4.7 μm and the NIR band near 2.3 μm. These retrievals have a spatial resolution of about 22 km × 22 km with satellite overpass time at approximately 10:30 and 22:30 (local time). To determine a unique CO concentration profile from the MOPITT measured radiances, an optimal estimation-based retrieval algorithm, and a fast radiative transfer model are used (Deeter et al., 2003; Edwards et al., 1999). The retrieved state vector (x_{rtv}) for optimal estimation-based retrievals can be expressed as

$$x_{rtv} = x_a + \boldsymbol{A}(x_{true} - x_a) + \epsilon \quad (1)$$

x_a and x_{true} are the a priori state vector and the true state vector, respectively. $\boldsymbol{A}$ (which has a size of 10×10) is the retrieval averaging kernel matrix (AK) that represents the sensitivity of retrieved profiles to actual profiles and ϵ is the random error vector. Note that CO profiles are retrieved as $\log_{10}$(VMR) quantities.

We focus on evaluating the recently released version 8, as well as the version 7, of the MOPITT TIR, NIR, and multispectral TIR-NIR products. The two versions of MOPITT products were introduced in detail in Deeter et al. (2017) and Deeter et al. (2019).

2.2 Aircraft measurements used for comparisons

Aircraft-sampled profiles of CO concentrations during the DISCOVER-AQ, SEAC^{4}RS, ARIAs, A-FORCE, and KORUS-AQ campaigns are used for comparisons with MOPITT-retrieved profiles. DISCOVER-AQ, and SEAC^{4}RS were conducted over the US, while ARIAs, A-FORCE, and KORUS-AQ were conducted over East Asia (EA). Locations of the aircraft profiles from these campaigns are compared with the MODIS (Moderate Resolution Imaging Spectroradiometer) Terra+Aqua Land Cover Type Climate Modeling Grid Yearly Level 3 version 6 0.05°×0.05° Global product (MCD12C1 v006) (Friedl and Sulla-Menashe, 2015) to determine if a profile is sampled over urban or non-urban regions. Specifically, for each aircraft profile, a 0.5°×0.5° box centered over the location of the aircraft profile (average of latitude and longitude of aircraft observations in the profile) is selected. If the urban and built-up fraction in the box is larger than 10%, the profile is determined to be an urban profile. Overall, for each campaign, the averaged aircraft profile over urban regions has higher CO concentrations compared to that over non-urban regions, especially near the surface (see Figure S1). Profiles during ARIAs are the exception, as the averaged profile over non-urban regions has higher CO concentrations especially near the surface. We also notice for aircraft profiles sampled during KORUS-AQ, even though the averaged profile over urban regions has slightly higher CO concentration near the surface, the profiles over urban and non-urban are close. This is largely due to the fact that many of the non-urban aircraft profiles are sampled over the Taehwa forest site, which is impacted by CO transported from the nearby Seoul urban region. Urban regions do not always have higher CO concentrations than non-urban regions. Therefore, because of the complexity of urban regions and their connection with non-urban regions nearby, we also provide analysis of validation at high CO concentrations regardless of landcover type.

The campaigns and profiles are summarized in the Table 1 and Figure 1. During DISCOVER-AQ, SEAC^{4}RS, and KORUS-AQ, CO concentrations were measured by the NASA Differential Absorption Carbon monOxide Measurement (DACOM), whereas during ARIAs and A-FORCE, CO concentrations were measured by different instruments, a Picarro G2401-I and Aero-Laser GmbH AL5002, respectively. Note that the primary goal of DISCOVER-AQ was to provide aircraft observation methodologies for satellite validation (e.g., Lamsal et al. (2014)). DISCOVER-AQ provides 121 profiles over four urban regions, making it particularly useful for the goal of this study. Because of this, our validation results are heavily driven by aircraft profiles from DISCOVER-AQ. Even though there are only two profiles sampled over urban regions, the A-FORCE campaign provides in total 45 profiles sampled over EA during Spring 2009, Winter 2013, and Summer 2013. The seasonal and spatial coverage of the dataset makes it representative

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions

of the region. The ARIAs campaign provides 19 profiles and three of these were sampled over Chinese urban regions. Only few previous studies have validated MOPITT products over China (e.g., Hedelius et al., 2019), so aircraft profiles from ARIAs have also been included in this study.

2.3 Method for comparing aircraft measurements and MOPITT profiles

We generally follow the method that has been used in previous MOPITT evaluation and validation studies (Deeter et al., 2010, 2012, 2013, 2014, 2016, 2017, 2019; Emmons et al., 2004, 2007, 2009). There are four main steps in aircraft versus MOPITT comparisons.

(1) Because of aircraft altitude limitations, in-situ data from field campaigns do not typically reach the highest altitudes at which MOPITT radiances are sensitive. Therefore, to obtain a complete vertical profile as required for comparison with MOPITT retrievals, each in-situ profile is extended vertically using the following steps: (i) the aircraft measurements are interpolated to the 35-level vertical grid used in MOPITT forward model calculations (0.2–1060 hPa); (ii) the levels from the surface to the lowest-altitude aircraft measurement are filled with the value of the in-situ measurement at the lowest-altitude aircraft measurement; (iii) for levels above a certain pressure level P_{interp} (e.g., 200 hPa), model or reanalysis data are used directly; (iv) for levels between the highest-altitude aircraft measurement and below P_{interp}, values are linearly interpolated. Unlike the previous MOPITT evaluation studies that used monthly model results from MOZART (Model for OZone And Related chemical Tracers) (Emmons et al., 2010) or CAM-chem (Community Atmosphere Model with chemistry) (Lamarque et al., 2012), here we use 3-hourly Copernicus Atmosphere Monitoring Service (CAMS) reanalysis of CO produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). CAMS CO reanalysis has a horizontal resolution of 80 km × 80 km, and 60 vertical grids (from surface to 0.1 hPa). Satellite retrievals of atmospheric composition including MOPITT TIR Version 6 total column CO retrievals are assimilated in the CAMS reanalysis (Inness et al., 2019; https://confluence.ecmwf.int/pages/viewpage.action?pageId=83396018). The final CO profile at the 35-level vertical grid is then regridded onto a coarser 10-level grid (for consistency with the actual MOPITT retrieval grid) by averaging the fine-grid VMR values in the layers immediately above the corresponding levels in the retrieval grid. We have conducted further calculations to investigate the sensitivity of validation results to P_{interp} in Section 4.1.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions

(2) For a given in-situ profile, only MOPITT profiles retrieved within the radius of 100 km and within 12 hours of the acquisition of the aircraft profile are considered co-located with the aircraft profile and are selected for comparisons. Sensitivities of validation results to the radius and time criteria for co-location selection have been further investigated in Section 4.2.

(3) For each pair of co-located MOPITT retrieval and in-situ profiles, we apply the MOPITT a priori profile and averaging kernel to the in-situ profile,

$$x_{transformed} = x_a + A(x_{in-situ} - x_a) \tag{2}$$

so that the transformed in-situ profile ($x_{transformed}$) has the same degree of smoothing and a priori dependence as the MOPITT profile.

(4) For each in-situ profile, there are likely to be multiple MOPITT retrievals that meet the above co-location criteria. If an in-situ profile is co-located with fewer than five MOPITT retrievals, the in-situ profile is not used in the following study and analysis. If an in-situ profile is co-located with five or more MOPITT retrievals, these co-located MOPITT profiles are averaged as $\log_{10}$(VMR). Applying these corresponding different MOPITT a priori profiles and averaging kernels to the same in-situ profile results in different transformed in-situ profiles. These transformed in-situ profiles that are generated from the same in-situ profile are also averaged.

Figure 2 shows an example of profile comparisons (the original aircraft profile, aircraft profile extended with CAMS reanalysis data and regridded to 35-level grid, $x_{in-situ}$, x_a, $x_{transformed}$, and x_{rtv}) in VMR for an aircraft profile sampled on July 22, 2011 during DISCOVER-AQ DC. Figure 2 also demonstrates what to expect within a MOPITT retrieval pixel and vertical level. The MOPITT retrievals have a spatial resolution of about 22 km × 22 km, and each MOPITT retrieval level corresponds to a uniformly-weighted layer immediately above that level. The vertical and horizontal variability of the original aircraft CO observations in each MOPITT layer (represented by standard deviation) are also shown. Taking the level of 800 hPa as an example, the variability of the original aircraft CO observations in the level is 21.4 ppb, which is larger than the difference between $x_{transformed}$ and x_{rtv} at that level. We also show the relative scale of the aircraft profile (3 km × 5 km) and a MOPITT retrieval pixel (22 km × 22 km) in Figure 2. We expect the variability of CO within a MOPITT retrieval pixel to be even larger than the CO variability within the scale of 3 km × 5 km. The variability within a satellite pixel and the

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

representativeness error in the satellite retrieval and aircraft profile comparisons make it very challenging to validate satellite retrievals against aircraft observations. This is one of the major reasons that MOPITT has yet to be validated over urban regions. The representativeness error has been discussed in previous studies (Fishman et al., 2011; Follette-Cook et al., 2015; Judd et al., 2019). In this study, we demonstrate this challenge with an example in Figure 2. We also show in Section 4 the sensitivity analysis to provide perspectives on how the spatial and temporal representativeness may change the validation results. Further quantification of the variability within MOPITT pixels would be very challenging (partially due to limited coverage of the observational data), and we will elaborate more on this issue in Section 5.

3. MOPITT validation over urban regions

In this section, the MOPITT validation results are provided for only daytime retrievals (i.e., solar zenith angle < 80° in the retrieval), because (1) MOPITT retrievals generally contain more CO profile information in daytime, which is reflected in AKs and Degrees of Freedom for Signal (DFS) in Figure 3, and (2) most aircraft profiles are sampled during daytime. In Section 4.3, we discuss the sensitivity to the inclusion of MOPITT nighttime retrievals in the validation process. In addition, many aircraft profiles, especially those from DISCOVER-AQ, lack observations above 600 hPa. Even though we extended the aircraft profiles vertically with reanalysis data (as discussed in Section 2.3), this still prevents the use of these profiles for validating MOPITT retrievals at upper levels against observations. In this paper, we only focus on validating MOPITT retrievals below 600 hPa. Nevertheless, since the CO retrievals below 600 hPa are still weakly impacted by CO fields in the upper levels (as shown by the AKs in Figure 3), in Section 4.1 we perform sensitivity tests on how augmenting the aircraft profiles with reanalysis fields affects the validation results.

3.1 Overall statistics

The overall validation results are presented in Table 2. Following Deeter et al. (2017), retrieval biases and standard deviation (SD) are calculated based on mean x_{rtv} and $x_{transformed}$ for each in-situ profile, and converted from log(VMR) to percent. The correlation coefficient (r) is quantified based on $(x_{rtv} - x_a)$ and the corresponding $\left(x_{transformed} - x_a\right)$ to avoid correlations

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

which mainly result from the variability of the a priori. x_{rtv}, $x_{transformed}$, and x_a are in $\log_{10}$(VMR) space in order to apply the AKs, which are computed for x_{rtv} in $\log_{10}$(VMR). Corresponding results for MOPITT Version 8 TIR-only (V8T) and Version 8 TIR-NIR (V8J) are shown in Figures 4 (for all profiles) and 5 (for urban profiles). Overall biases for V8J products (averaged over all campaigns in Table 1) vary from -0.7% to 0.0%, which are lower than biases for V8T (from 2.0% to 3.5%). Overall biases for V8J products are also lower than biases for V7J (from -0.5% to -5.4%). For V8J and V7J, biases over urban regions vary from -0.2% to -0.8% and from -8.9% to -1.4%, respectively, which are generally higher than biases over non-urban regions (-0.3%~1.1% and -3.3%~0.1%). Correlation coefficients over non-urban regions are generally higher than those over urban regions for all six products (V7T, V8T, V7N, V8N, V7J, V8J) at all three levels (surface, 800 hPa, 600 hPa). For example, for the V8J product, correlation coefficients over urban regions are 0.53, 0.57, and 0.53 at the surface, 800 hPa, and 600 hPa, respectively, while over non-urban regions, the corresponding correlation coefficients are 0.76, 0.73 and 0.67. We also notice that V8 products generally have higher correlation coefficients with in-situ measurements than V7 over non-urban regions, whereas over urban regions, V8 products generally have lower correlation coefficients than V7. Overall, MOPITT products (especially V8J) perform reasonably well over both urban and non-urban regions. Performance over non-urban regions is better than that over urban regions in terms of correlation coefficients and biases for V8J and V7J.

3.2 Discussions on individual campaigns

We also provide MOPITT V8J evaluation against individual field campaigns in Figure 6. The corresponding results for MOPITT V8T are summarized in Figure S2. The patterns of biases are very similar for MOPITT V8J and V8T. Thus, in this sub-section, we focus on V8J unless stated otherwise. Overall, besides comparisons with A-FORCE and ARIAs, biases over urban regions and non-urban regions do not have a significant difference. Neither do biases determined for campaigns over the US and EA differ significantly, either. When compared to DISCOVER-AQ CA, MOPITT CO values are generally higher than in-situ profiles at 600 hPa but not at the surface. This is likely related to the fact that the DISCOVER-AQ CA aircraft profiles are mostly below 600 hPa, and hence CO values of these in-situ profiles at 600 hPa and above are filled with CAMS reanalysis data. In addition, DISCOVER-AQ CA was conducted in the winter when boundary layer height is at lower altitudes, which could also explain the difference, in particular

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions

since most of the other campaigns are in more favorable weather conditions. The lack of aircraft observations at 600 hPa and above also has a smaller impact on the biases at 800 hPa through applying AK (see Figure 3). During the A-FORCE campaign, only 2 in-situ profiles out of 45 were sampled over urban regions. The locations of the two profiles are close to each other and they are both sampled on/near the coast of South Korea (Figure 1). MOPITT has large negative biases (-30%~-40%) when compared to these two profiles. The averaged $x_{in-situ}$, x_a, $x_{transformed}$, and x_{rtv} over non-urban regions during A-FORCE and the $x_{in-situ}$, x_a, $x_{transformed}$, and x_{rtv} of the two profiles over urban regions are shown in Figure S3. Compared to the averaged $x_{in-situ}$, the $x_{in-situ}$ for the two profiles over the urban regions have large enhancements near the surface and between 600~800 hPa. Even though the x_a and x_{rtv} for the two profiles have higher CO concentrations (~400 ppb at the surface) than the averaged x_a and x_{rtv} (~200 ppb at the surface), they are still lower than the $x_{transformed}$. As for KORUS-AQ, MOPITT also has a negative bias (though smaller) when compared to the profiles over urban regions. Most of these KORUS-AQ profiles were located near the two profiles from A-FORCE but farther from the coast. The negative bias is not seen over non-urban regions during KORUS-AQ at the surface. When compared to the in-situ profiles from ARIAs, MOPITT has a large positive bias, especially over urban regions (20%~30%). During ARIAs, in-situ profiles over urban regions have lower CO values (~200 ppb at the surface) than those in-situ profiles over non-urban regions (~ 400 ppb at the surface; Figure S4). We note there are only a small number of in-situ profiles over urban regions in EA used in this study, compared to what is provided by DISCOVER-AQ in the US. The large negative biases against A-FORCE and large positive biases against ARIAs point to the need for more in-situ observations over EA.

3.3 Validation at high CO concentrations

Urban regions are often associated with high CO concentrations. But this is not always the case (e.g., Figure S4). Here we separate the in-situ profiles at the surface, 800 hPa, and 600 hPa into lower 50% CO values and higher 50% CO values based on CO values at each level to demonstrate the impact of CO concentrations on the MOPITT product validation (Figure 7). For V8J, MOPITT has smaller biases at higher 50% CO concentrations all three levels, whereas for V8T, MOPITT has larger biases at the surface and 600 hPa at higher 50% CO concentrations. For both V8J and V8T, MOPITT has larger SDs and lower correlation coefficients at the surface, 800 hPa, and 600 hPa if only the upper 50% of measured CO mixing ratios are considered, suggesting that this validation of MOPITT at higher CO concentrations is not as good as that at lower CO concentrations. In contrast, Deeter et al. (2016) found that the retrieval biases do not visibly increase at the upper range of CO concentrations when compared to aircraft measurements over the Amazon basin. The vertical error bars in Figure 7 (caused by the multiple co-located MOPITT profiles with one in-situ profile) represent the variability (standard deviation) of the MOPITT data used to calculate each of the plotted mean values. For an in-situ profile, the variability of the MOPITT data located within its radius of 100 km and within 12 hours is larger when the in-situ profile has higher CO values, indicated by larger error bars at higher 50% CO concentrations. However, it is unclear whether the larger apparent bias at high CO concentration actually represents larger retrieval uncertainties or could be related to greater CO variability and representativeness of the in situ profile within the co-location radius used for analyzing the MOPITT data. We will discuss the sensitivity of radius and time difference for the selection of co-located data in Section 4. The difference in the variability at different CO concentrations was not found in Deeter et al. (2016). It could be partially due to the fact that the aircraft profiles over the Amazon basin used in Deeter et al. (2016) were sampled in more geographically homogeneous conditions, whereas the profiles used in this study are from different campaigns, and high CO concentrations over and near urban regions might be associated with more complex and inhomogeneous conditions.

4. Sensitivities to assumptions made for aircraft-satellite comparisons

4.1 Sensitivity to the in-situ profile extension

As discussed in Section 2.3, the in-situ profiles must be vertically extrapolated or extended for use in MOPITT validation due to aircraft altitude limits. Thus, model or reanalysis data must be merged with the in-situ data to generate a complete CO profile for comparisons with MOPITT satellite retrievals. The use of model or reanalysis data may introduce uncertainties in the validation results as they are not measured directly. The parameter P_{interp} controls the impact of the model-based profile extension on the shape and value of in-situ profiles (see Figure S5). Here we test the sensitivity of validation results to various P_{interp} values (100 hPa, 200 hPa, 300 hPa, 400 hPa, 500 hPa) to demonstrate the potential impact of the profile extension on the validation results.

Note that the model-based profile extension and the value of P_{interp} impacts the validation results through changing the augmented observational profile, which is different from the other sensitivity tests in this study that change the selection of MOPITT data. The validation results at the surface are insensitive to the selection of P_{interp} (Figure 8). The overall validation results at the 800 hPa are also not sensitive to P_{interp}, except for the validation results against DISCOVER-AQ CA which have slightly larger biases when P_{interp} is 200 hPa or 100 hPa. As mentioned in Section 3.2, the DISCOVER-AQ CA aircraft profiles are mostly below 600 hPa, and hence CO values of these in-situ profiles at 600 hPa and above are extended using reanalysis data. Therefore, the validation results against DISCOVER-AQ CA are more likely to be affected by P_{interp} compared to other campaigns which typically obtained higher maximum aircraft altitudes. At 600 hPa, the validation results are more affected by P_{interp} compared to the those at the surface and 800 hPa. The validation results using 100 hPa as P_{interp} have larger biases. The validation results using 300, 400, or 500 hPa as P_{interp} are not significantly different for the validation results against DISCOVER-AQ CA. The validation results against DISCOVER-AQ CA using 200 hPa as P_{interp} show similar results as those using 100 hPa as P_{interp}. The validation results to the P_{interp} at 400 hPa and 200 hPa are even more sensitive with larger biases (Figure S6). As mentioned in Section 3.2, the DISCOVER-AQ CA aircraft measurements concentrate below 600 hPa, so CO values in the in-situ profiles at 600 hPa and above are filled with and are more sensitive to CAMS reanalysis data. The CAMS 3-hourly reanalysis data are constrained by observations, but its usage may still introduce the uncertainties in the validation results especially at upper pressure levels (e.g., 200 hPa and 400 hPa). Previous MOPITT evaluation results may be subject to larger uncertainties by using CAM-chem monthly CO fields that are not constrained by observations.

4.2 Sensitivity to the radius and allowed maximum time difference as criteria for co-location

The criteria for co-location in this study (within the radius of 100 km and within 12 hours of the acquisition of the aircraft profile) generally follow previous MOPITT validation studies (e.g., Deeter et al., 2016, 2019) and are chosen empirically. They are selected based on a trade-off between uncertainties generated from CO spatial and/or temporal variability, and the number of included MOPITT retrievals that impacts the statistical robustness. Here we test the sensitivity of the validation results to the two criteria for co-location. The boxplot of biases calculated with different radii (200 km, 100 km, 50 km, and 25 km) at the surface, 800 hPa, and 600 hPa are shown in Figure 9. Overall, the biases calculated with radius of 200 km, 100 km and 50 km are close,

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

whereas the biases calculated with the radius of 25 km are different from others. The validation results using the radius of 25 km generally have larger biases and SD, due to a smaller number of included MOPITT retrievals. In some cases, there are no matched MOPITT retrievals within the radius of 25km of the aircraft profile (e.g., DISCOVER-AQ CA and ARIAs). In addition, representativeness errors would be expected to go up if there are only a few retrievals over a a more more polluted and perhaps heterogeneous area. We note that the usage of the largest radius (200 km) in this paper does not appear to degrade the results through introducing representativeness errors generated from CO spatial and/or temporal variability, whereas use of the smallest radius (25 km) degrades the results by reducing the number of included MOPITT retrievals.

The boxplot of biases calculated with four sets of allowed maximum time difference (12 hours, 6 hours, 3 hours, and 1 hours) are shown in Figure 10. The overall validation results are not sensitive to the selection of allowed maximum time difference, especially at the surface. One exception is the validation results against the SEAC^{4}RS campaign at 600 hPa, due to a smaller number of MOPITT retrievals in the shorter time window. We note that when validated against the ARIAs campaign, the biases at the surface, 800 hPa and 600 hPa are smaller with the allowed maximum time difference as 1h, indicating the temporal variability is relatively large in the region. And the improvement observed for ARIAs for the shortest time also points to the possibility that short term emission sources might be responsible for the large biases there. On the other hand, when the allowed maximum time difference equals 1 hour, there are only 6 aircraft profiles that have matched MOPITT retrievals.

4.3 Sensitivity to the inclusion of MOPITT nighttime retrievals

Previous MOPITT validation studies have only included MOPITT daytime observations. Over land, MOPITT retrievals for daytime and nighttime overpasses are characterized by significantly different averaging kernels (Figure 3), and may be subject to different types of retrieval error (Deeter et al., 2007). CO has a long enough lifetime in the free troposphere that nighttime observations could be potentially comparable, in general, to the daytime flights for remote sites. However, for urban regions where the spatiotemporal variability of the emissions and evolution of the planetary boundary layer drives large changes in the measured CO, comparisons of MOPITT nighttime observations to aircraft profiles sampled during daytime may introduce

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

representative uncertainties. It is difficult to disentangle the effects of the MOPITT daytime/nighttime performance and the uncertainty from the temporal representativeness, based on the comparison of the MOPITT daytime/nighttime retrievals with daytime aircraft profiles. Therefore, we only include the results in Figure S7 and briefly describe the results here without drawing any further conclusions. Overall, MOPITT nighttime retrievals have larger biases than daytime retrievals, which could be expected since most of the aircraft profiles are sampled during daytime. Flight campaigns with nighttime observations are needed to validate MOPITT nighttime retrievals.

4.4 Sensitivity to the signal-to-noise ratio (SNR) filters

The MOPITT Level 3 data are generated from Level 2 data, and are available as gridded daily-mean and monthly-mean files. Pixel filtering and signal-to-noise ratio (SNR) thresholds for Channel 5 and 6 Average radiances are used when averaging Level 2 data into Level 3 data, and this increases overall mean DFS values (details can be found in the MOPITT user guide; https://www2.acom.ucar.edu/sites/default/files/mopitt/v8_users_guide_201812.pdf). Taking MOPITT V8J daytime product as an example, Level 3 data product excludes all observations from Pixel 3 (one of the four elements of MOPITT's linear detector array that has highly variable Channel 7 SNR values), or observations where both the Channel 5 Average radiances SNR < 1000 and the Channel 6 Average radiances SNR < 400. In Figure 11, we test the impact of applying the aforementioned SNR filters on the validation results. We find that applying the SNR filters does not improve the overall agreement between MOPITT retrievals and in-situ profiles. In some cases, applying the SNR filters degrades the validation results (e.g., DISCOVER-AQ DC at the surface, DISCOVER-AQ CA at the surface, KORUS-AQ at 600 hPa, and ARIAs at the surface, 800 hPa, and 600 hPa). This is mostly because applying the SNR filters reduces the number of MOPITT retrievals included in the comparisons. This effect is particularly important if there are not many MOPITT retrievals to begin with (such as our comparisons with in-situ profiles in this study). However, when generating Level 3 data from Level 2 data, the circumstance is different as there are usually much more data to perform the filter and averaging process.

5. Discussion and conclusions

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

MOPITT products are widely used for monitoring and analyzing CO over urban regions. However, systematic validation against observations over urban regions has been lacking. In this study, we compared MOPITT products over urban regions to aircraft measurements from DISCOVER-AQ, SEAC^{4}RS, ARIAs, A-FORCE, and KORUS-AQ campaigns. The DISCOVER-AQ campaign was designed primarily with satellite validation in mind, and the campaign over DC, CA, TX, and CO together contributes 64.8% (232 out of 358) of the aircraft profiles and 91.0% (121 out of 133) of the aircraft profiles over the urban regions (Table 1). Therefore, the DISCOVER-AQ campaign largely contributes to the validation results and the statistics in this study. We found that MOPITT biases are well within the 10% required accuracy for both urban and non-urban regions (overall biases for V8J and V8T vary from -0.7% to 0.0%, and from 2.0% to 3.5%). The performance over non-urban regions is better than that over urban regions in terms of correlation coefficients for the 6 products in Table 2, and biases of V8J and V7J. However, the in-situ profiles over EA used in this study are limited, especially over urban regions (only 11 profiles). The large biases against aircraft profiles from the A-FORCE and ARIAs campaigns point to the need for more in-situ observations over EA. We also studied the impact of CO concentrations on the MOPITT product validation by dividing the aircraft profiles of CO to two groups of high CO (upper 50%) and low CO (lower 50%). We found that MOPITT retrievals at high CO concentrations have higher biases and lower correlations compared low CO concentrations, although CO variability may tend to exaggerate retrieval biases in heavily-polluted scenes.

In addition, the assumptions and data filters made during aircraft-satellite comparisons may impact the validation results. We tested the sensitivities of validation results to assumptions and data filters, including the model-based extension to the in-situ profile, radius and allowed maximum time difference as criteria for the selection of co-located data, the inclusion of nighttime MOPITT data, and the SNR filters. The validation results at the surface are insensitive to the model-based profile extension, whereas the validation results at upper levels (e.g., 400 hPa and 200 hPa) are more sensitive to the profile extension, as there are very limited aircraft observations. The validation results are insensitive to the allowed maximum time difference as co-location criteria, and are generally insensitive to the radius for co-location except for the case with a radius of 25 km, where a small number of MOPITT retrievals are included in the validation. Overall, daytime MOPITT products overall have smaller biases than nighttime MOPITT products. However, conclusions regarding the performance of MOPITT daytime and nighttime retrievals

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

cannot be drawn due to the fact that most of the aircraft profiles are sampled during daytime. As we mentioned earlier, MOPITT daytime and nighttime retrievals may be subject to different retrieval errors. In addition, previous studies suggest pollutants themselves may have different characteristics during daytime and nighttime (e.g., Yan et al., 2018). Therefore, validation of MOPITT nighttime retrievals, with a sufficient number of nighttime airborne profiles, is needed in order to study nighttime CO characteristics and trends. Applying SNR filters does not necessarily improve the overall agreement between MOPITT retrievals and in-situ profiles, and this may be partially caused by the smaller number of MOPITT retrievals in the validation process after the SNR filters, which is unlikely to happen when generating Level 3 data. We note that validation results against ARIAs are an exception in a few sensitivity tests due to rather a limited number of aircraft measurements. Given the large biases against aircraft profiles from the ARIAs campaign, more in-situ observations over EA especially China are needed in order to validate MOPITT products in the region.

Validation and evaluation of satellite retrievals with aircraft observations are very challenging, and assumptions have to be made for the comparisons. As discussed in Section 2, the CO spatial variability within MOPITT retrieval pixels and the representativeness error of aircraft profiles when compared to MOPITT retrievals may introduce uncertainties in the validation results. This issue is difficult to address and quantify due to the limited spatial coverage of dense aircraft observations. Follette-Cook et al. (2015) quantified spatial and temporal variability of column integrated air pollutants, including CO, during DISCOVER-AQ DC from modeling perspective (using the Weather Research and Forecasting model coupled with Chemistry - WRF-Chem). They found that during the July 2011 DISCOVER-AQ campaign, the mean CO difference at the distance of 20-24 km is ~30 ppb (derived from the aircraft observations) and ~40 ppb (derived from co-located WRF-Chem output), based on structure function analyses. Judd et al. (2019) explored the impact of spatial resolution on tropospheric NO_2 column retrievals with NASA Geostationary Trace Gas and Aerosol Sensor Optimization (GeoTASO). We expect CO to have a smaller spatial and temporal variability than NO_2 due primarily to its relatively longer lifetime, though future analyses of NO_2 variability within urban regions using GeoTASO could provide an upper estimate on CO variability. In addition, the variability of Tropospheric Monitoring Instrument (TROPOMI) CO retrievals, with a pixel size of 7 km×7 km (Landgraf et al., 2016). within the larger MOPITT footprint might also provide information on MOPITT sub-pixel variability. Further research on trace gas spatial variability within satellite retrieval pixels, and quantification of the representativeness error incurred by using individual aircraft profiles in validation comparisons is needed, and will be the subject of a follow-up study.

Acknowledgements

MOPITT products are available at https://www2.acom.ucar.edu/mopitt. MOPITT data can be downloaded at https://earthdata.nasa.gov/. The NCAR MOPITT project is supported by the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Program. The authors thank the DISCOVER-AQ, SEAC^{4}RS, ARIAs, A-FORCE, and KORUS-AQ Science Teams for the valuable in-situ observations. DISCOVER-AQ data can be accessed at https://www-air.larc.nasa.gov/missions/discover-aq/discover-aq.html. SEAC^{4}RS data can be accessed at https://www-air.larc.nasa.gov/missions/seac4rs/. We thank Drs. Naga Oshima and Makoto Koike for the A-FORCE data. KORUS-AQ data can be accessed at https://www-air.larc.nasa.gov/missions/korus-aq/index.html. The authors thank Dr. Frank Flocke for helpful comments on the manuscript. Wenfu Tang thanks Dr. Cenlin He for helpful discussions. The National Center for Atmospheric Research (NCAR) is sponsored by the National Science Foundation. W. Tang is supported by a NCAR Advanced Study Program Postdoctoral Fellowship.

References

Al-Saadi, J., Carmichael, G., Crawford, J., Emmons, L., Song, C. K., Chang, L. S., ... & Park, R. (2015). NASA contributions to KORUS-AQ: An international cooperative air quality field study in Korea. NASA White Paper available at: https://goo. gl/VhssdX (last access: 3 May 2016).

Aliyu, Y. A., & Botai, J. O.: Appraising city-scale pollution monitoring capabilities of multi-satellite datasets using portable pollutant monitors. *Atmospheric environment*, *179*, 239-249, 2018.

Arellano, A. F., Kasibhatla, P. S., Giglio, L., van der Werf, G. R., and Randerson, J. T.: Topdown estimates of global CO sources using MOPITT measurements, Geophys. Res. Lett., 31, L01104, doi:10.1029/2003GL018609, 2004.

Arellano, A. F., Kasibhatla, P. S., Giglio, L., van der Werf, G. R., Randerson, J. T., and Collatz, G. J.:, Time-dependent inversion estimates of global biomass-burning CO emissions using Measurement of Pollution in the Troposphere (MOPITT) measurements, J. Geophys. Res., 111, D09303, doi:10.1029/2005JD006613, 2006.

Arellano A. F. Jr., Raeder, K., Anderson, J. L., Hess, P. G., Emmons, L. K., Edwards, D. P., Pfister, G. G., Campos, T. L., and Sachse, G. W.: Evaluating model performance of an ensemble-based chemical data assimilation system during INTEX-B field mission, Atmos. Chem. Phys., 7, 5695-5710, https://doi.org/10.5194/acp-7-5695-2007, 2007.

Chen, D., Wang, Y., McElroy, M. B., He, K., Yantosca, R. M., and Le Sager, P.: Regional CO pollution and export in China simulated by the high-resolution nested-grid GEOS-Chem model, Atmos. Chem. Phys., 9, 3825-3839, https://doi.org/10.5194/acp-9-3825-2009, 2009.

Clerbaux, C., Edwards, D. P., Deeter, M., Emmons, L., Lamarque, J.-F., Tie, X. X., Massie, S. T., and Gille, J.: Carbon monoxide pollution from cities and urban areas observed by the Terra/MOPITT mission, Geophys. Res. Lett., 35, 3817, doi:10.1029/2007GL032300, 2008.

Deeter, M. N., Emmons, L. K., Francis, G. L., Edwards, D. P., Gille, J. C., Warner, J. X., Khattatov, B., Ziskin, D., Lamarque, J.-F., Ho, S.-P., Yudin, V., Attie, J.-L., Packman, D., Chen, J., Mao, D., and Drummond, J. R.: Operational carbon monoxide retrieval algorithm and selected results for the MOPITT instrument, J. Geophys. Res., 108(D14), 4399, doi:10.1029/2002JD003186, 2003.

Deeter,M. N., Edwards, D. P., Gille, J. C., Emmons, L. K., Francis, G., Ho, S.-P., Mao, D., Masters, D., Worden, H., Drummond, J. R., and Novelli, P. C.: The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, J. Geophys. Res., 115, D07306, doi:10.1029/2009JD013005, 2010.

Deeter, M. N., Worden, H. M., Gille, J. C., Edwards, D. P., Mao, D., and Drummond, J. R.: MOPITT multispectral CO retrievals: Origins and effects of geophysical radiance errors, J. Geophys. Res., 116, D15303, doi:10.1029/2011JD015703, 2011.

Deeter, M. N., Worden, H. M., Edwards, D. P., Gille, J. C., and Andrews, A. E.: Evaluation of MOPITT Retrievals of Lowertropospheric Carbon Monoxide over the United States, J. Geophys. Res., 117, D13306, doi:10.1029/2012JD017553, 2012.

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

Deeter, M. N., Martínez-Alonso, S., Edwards, D. P., Emmons, L. K., Gille, J. C., Worden, H. M., Sweeney, C., Pittman, J. V., Daube, B. C., and Wofsy, S. C.: The MOPITT Version 6 product: algorithm enhancements and validation, Atmos. Meas. Tech., 7, 3623–3632, https://doi.org/10.5194/amt-7-3623-2014, 2014.

Deeter, M. N., Martínez-Alonso, S., Gatti, L. V., Gloor, M., Miller, J. B., Domingues, L. G., and Correia, C. S. C.: Validation and analysis of MOPITT CO observations of the Amazon Basin, Atmos. Meas. Tech., 9, 3999–4012, https://doi.org/10.5194/amt-9-3999-2016, 2016.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions

Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Martínez-Alonso, S., Worden, H. M., and Sweeney, C.: A climate-scale satellite record for carbon monoxide: the MOPITT Version 7 product, Atmos. Meas. Tech., 10, 2533–2555, https://doi.org/10.5194/amt-10-2533-2017, 2017.

Deeter, M. N., Edwards, D. P., Francis, G. L., Gille, J. C., Mao, D., Martínez-Alonso, S., Worden, H. M., Ziskin, D., and Andreae, M. O.: Radiance-based retrieval bias mitigation for the MOPITT instrument: the version 8 product, Atmos. Meas. Tech., 12, 4561–4580, https://doi.org/10.5194/amt-12-4561-2019, 2019.

Edwards, D., Halvorson, C., and Gille, J.: Radiative transfer modeling for the EOS Terra satellite Measurement of Pollution in the Troposphere (MOPITT) instrument, J. Geophys. Res., 104, 16755–16775, 1999.

Edwards, D. P., Petron, G., Novelli, P. C., Emmons, L. K., Gille, J. C., and Drummond, J. R.: Southern Hemisphere carbon monoxide interannual variability observed by Terra/Measurement of Pollution in the Troposphere (MOPITT), J. Geophys. Res., 111, D16303, doi:10.1029/2006JD007079, 2006.

Emmons, L. K., Deeter, M. N., Gille, J. C., et al.: Validation of Measurements of Pollution in the Troposphere (MOPITT) CO retrievals with aircraft in situ profiles, J. Geophys. Res., 109, D03309, doi:10.1029/2003JD004101, 2004.

Emmons, L. K., Pfister, G. G., Edwards, D. P., Gille, J. C., Sachse, G., Blake, D., Wofsy, S., Gerbig, C., Matross, D., and N´ed´elec, P.: Measurements of Pollution in the Troposphere (MOPITT) validation exercises during summer 2004 field campaigns over North America, J. Geophys. Res., 112, D12S02, doi:0.1029/2006JD007833, 2007.

Emmons, L. K., Edwards, D. P., Deeter, M. N., Gille, J. C., Campos, T., Nédélec, P., Novelli, P., and Sachse, G.: Measurements of Pollution In The Troposphere (MOPITT) validation through 2006, Atmos. Chem. Phys., 9, 1795-1803, https://doi.org/10.5194/acp-9-1795-2009, 2009.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), Geosci. Model Dev., 3, 43–67, https://doi.org/10.5194/gmd-3-43-2010, 2010.

Fishman, J., Silverman, M. L., Crawford, J. H., & Creilson, J. K.: A study of regional-scale variability of in situ and model-generated tropospheric trace gases: Insights into observational requirements for a satellite in geostationary orbit. Atmospheric environment, 45(27), 4682-4694, 2011.

Follette-Cook, M., Pickering, K., Crawford, J., Duncan, B., Loughner, C., Diskin, G., Fried, A., and Weinheimer, A.: Spatial and temporal variability of trace gas columns derived from WRF/Chem regional model output: Planning for geostationary observations of atmospheric composition, Atmos. Environ., 118, 28–44, doi:10.1016/j.atmosenv.2015.07.024, 2015.

Fortems-Cheiney, A., Chevallier, F., Pison, I., Bousquet, P., Szopa, S., Deeter, M. N., and Clerbaux, C.: Ten years of CO emissions as seen from Measurements of Pollution in the Troposphere (MOPITT), J. Geophys. Res., 116, D05304, doi:10.1029/2010JD014416, 2011.

Friedl, M., Sulla-Menashe, D. (2015). MCD12C1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2019-08-12 from https://doi.org/10.5067/MODIS/MCD12C1.006.

Gaubert, B., Arellano, A. F., Barré, J., Worden, H. M., Emmons, L. K., Tilmes, S., Buchholz, R. R.,Vitt, F., Raeder, K., Collins, N., Anderson, J. L., Wiedinmyer, C., Martinez Alonso, S., Edwards, D. P., Andreae, M. O., Hannigan, J. W., Petri, C., Strong, K., and Jones, N.: Toward a chemical reanalysis in a coupled chemistry-climate model: An evaluation of MOPITT CO assimilation and its impact on tropospheric composition, J. Geophys. Res.-Atmos., 121, 7310–7343, https://doi.org/10.1002/2016JD024863, 2016.

Girach, I. A. and Nair, P. R.: Carbon monoxide over Indian region as observed by MOPITT, Atmos. Environ., 99, 599–609, 2014.

He, H., Stehr, J. W., Hains, J. C., Krask, D. J., Doddridge, B. G., Vinnikov, K. Y., Canty, T. P., Hosley, K. M., Salawitch, R. J., Worden, H. M., and Dickerson, R. R.: Trends in emissions and concentrations of air pollutants in the lower troposphere in the Baltimore/Washington airshed from 1997 to 2011, Atmos. Chem. Phys., 13, 7859-7874, https://doi.org/10.5194/acp-13-7859-2013, 2013.

Heald, C. L., Jacob, D. J., Jones, D. B. A., Palmer, P. I., Logan, J. A., Streets, D. G., Sachse, G. W., Gille, J. C., Hoffman, R. N., and Nehrkorn, T.: Comparative inverse analysis of satellite (MOPITT) and aircraft (TRACE-P) observations to estimate Asian sources of carbon monoxide, J. Geophys. Res., 109, D23306, doi:10.1029/2004JD005185, 2004.

Hedelius, J. K., He, T.-L., Jones, D. B. A., Baier, B. C., Buchholz, R. R., De Mazière, M., Deutscher, N. M., Dubey, M. K., Feist, D. G., Griffith, D. W. T., Hase, F., Iraci, L. T., Jeseck, P., Kiel, M., Kivi, R., Liu, C., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Pollard, D. F., Rettinger, M., Roche, S., Roehl, C. M., Schneider, M., Shiomi, K., Strong, K., Sussmann, R., Sweeney, C., Té, Y., Uchino, O., Velazco, V. A., Wang, W., Warneke, T., Wennberg, P. O., Worden, H. M., and Wunch, D.: Evaluation of MOPITT Version 7 joint TIR–NIR XCO retrievals with TCCON, Atmos. Meas. Tech., 12, 5547–5572, https://doi.org/10.5194/amt-12-5547-2019, 2019.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmos. Chem. Phys., 19, 3515-3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Jiang, Z., Worden, J. R., Jones, D. B. A., Lin, J.-T., Verstraeten, W. W., and Henze, D. K.: Constraints on Asian ozone using Aura TES, OMI and Terra MOPITT, Atmos. Chem. Phys., 15, 99-112, https://doi.org/10.5194/acp-15-99-2015, 2015.

Jiang, Z., McDonald, B. C., Worden, H., Worden, J. R., Miyazaki, K., Qu, Z., Henze, D. K., Jones, D. B. A., Arellano, A. F., Fischer, E. V., Zhu, L. Y., and Boersma, K. F.: Unexpected slowdown of US pollutant emission reduction in the past decade, P. Natl. Acad. Sci. USA, 115, 5099–5104, https://doi.org/10.1073/pnas.1801191115, 2018.

Judd, L. M., Al-Saadi, J. A., Janz, S. J., Kowalewski, M. G., Pierce, R. B., Szykman, J. J., Valin, L. C., Swap, R., Cede, A., Mueller, M., Tiefengraber, M., Abuhassan, N., and Williams, D.: Evaluating the impact of spatial resolution on tropospheric NO2 column comparisons within urban areas using high-resolution airborne data, Atmos. Meas. Tech., https://doi.org/10.5194/amt-2019-161, 2019.

Kanakidou, M., Mihalopoulos, N., Kindap, T., Im, U., Vrekoussis, M., Dermitzaki, E., Gerasopoulos, E. , Unal, A., Koc¸ak, M., Markakis, K., Melas, D., Youssef, A. F., and Moubasher, H.: Megacities as hot spots of air pollution in the East Mediterranean, doi:10.1016/j.atmosenv.2010.11.048, 2011.

Kar, J., Deeter, M. N., Fishman, J., Liu, Z., Omar, A., Creilson, J. K., Trepte, C. R., Vaughan, M. A., and Winker, D. M.: Wintertime pollution over the Eastern Indo-Gangetic Plains as observed from MOPITT, CALIPSO and tropospheric ozone residual data, Atmos. Chem. Phys., 10, 12273-12283, https://doi.org/10.5194/acp-10-12273-2010, 2010.

Kopacz, M., Jacob, D. J., Henze, D. K., Heald, C. L., Streets, D. G., and Zhang, Q.: Comparison of adjoint and analytical Bayesian inversion methods for constraining Asian sources of carbon monoxide using satellite (MOPITT) measurements of CO columns, J. Geophys. Res., 114, D04305, doi:10.1029/2007JD009264., 2009.

Kopacz, M., Jacob, D. J., Fisher, J. A., Logan, J. A., Zhang, L., Megretskaia, I. A., Yantosca, R. M., Singh, K., Henze, D. K., Burrows, J. P., Buchwitz, M., Khlystova, I., McMillan, W. W., Gille, J. C., Edwards, D. P., Eldering, A., Thouret, V., and Nedelec, P.: 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, https://doi.org/10.5194/acp-10-855-2010, 2010.

Kumar, R., Naja, M., Pfister, G. G., Barth, M. C., Wiedinmyer, C., and Brasseur, G. P.: Simulations over South Asia using the Weather Research and Forecasting model with Chemistry (WRF-Chem): chemistry evaluation and initial results, Geosci. Model Dev., 5, 619–648, https://doi.org/10.5194/gmd-5-619-2012, 2012.

Lamarque, J.-F., Emmons, L. K., Hess, P. G., Kinnison, D. E., Tilmes, S., Vitt, F., Heald, C. L., Holland, E. A., Lauritzen, P. H., Neu, J., Orlando, J. J., Rasch, P. J., and Tyndall, G. K.: CAM-chem: description and evaluation of interactive atmospheric chemistry in the Community Earth System Model, Geosci. Model Dev., 5, 369–411, https://doi.org/10.5194/gmd-5-369-2012, 2012.

Lamsal, L. N., Krotkov, N. A., Celarier, E. A., Swartz, W. H., Pickering, K. E., Bucsela, E. J., Gleason, J. F., Martin, R. V., Philip, S., Irie, H., Cede, A., Herman, J., Weinheimer, A., Szykman, J. J., and Knepp, T. N.: Evaluation of OMI operational standard NO2 column retrievals using in situ and surface-based NO2 observations, Atmos. Chem. Phys., 14, 11587-11609, https://doi.org/10.5194/acp-14-11587-2014, 2014.

Landgraf, J., aan de Brugh, J., Scheepmaker, R., Borsdorff, T., Hu, H., Houweling, S., Butz, A., Aben, I., and Hasekamp, O.: Carbon monoxide total column retrievals from TROPOMI shortwave infrared measurements, Atmos. Meas. Tech., 9, 4955–4975, https://doi.org/10.5194/amt-9-4955-2016, 2016.

Li, L. and Liu, Y.: Space-borne and ground observations of the characteristics of CO pollution in Beijing, 2000–2010, Atmos. Environ., 45, 2367, doi:10.1016/j.atmosenv.2011.02.026, 2011.

Tang, W., Arellano, A. F., DiGangi, J. P., Choi, Y., Diskin, G. S., Agustí-Panareda, A., Parrington, M., Massart, S., Gaubert, B., Lee, Y., Kim, D., Jung, J., Hong, J., Hong, J.-W., Kanaya, Y., Lee, M., Stauffer, R. M., Thompson, A. M., Flynn, J. H., and Woo, J.-H.: Evaluating high-resolution forecasts of atmospheric CO and CO2 from a global prediction system during KORUS-AQ field campaign, Atmos. Chem. Phys., 18, 11007–11030, https://doi.org/10.5194/acp-18-11007-2018, 2018.

Tang, W., Arellano, A. F., Gaubert, B., Miyazaki, K., and Worden, H. M.: Satellite data reveal a common combustion emission pathway for major cities in China, Atmos. Chem. Phys., 19, 4269–4288, https://doi.org/10.5194/acp-19-4269-2019, 2019.

Toon, O. B., Maring, H., Dibb, J., Ferrare, R., Jacob, D. J., Jensen, E. J., Luo, Z. J., Mace, G. G., Pan, L. L., Pfister, L., and Rosenlof, K. H.: Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC4RS) field mission, J. Geophys. Res.-Atmos., 121, 4967–5009, doi:10.1002/2015JD024297, 2016.

Wang, F., Li, Z., Ren, X., Jiang, Q., He, H., Dickerson, R. R., Dong, X., and Lv, F.: Vertical distributions of aerosol optical properties during the spring 2016 ARIAs airborne campaign in the North China Plain, Atmos. Chem. Phys., 18, 8995–9010, https://doi.org/10.5194/acp-18-8995-2018, 2018.

Worden, H. M., Deeter, M. N., Edwards, D. P., Gille, J. C., Drummond, J. R., and Nédélec, P.: Observations of near-surface carbon monoxide from space using MOPITT multispectral retrievals, J. Geophys. Res., 115, D18314, https://doi.org/10.1029/2010JD014242, 2010.

Worden, H. M., Deeter, M. N., Edwards, D. P., Gille, J., Drummond, J., Emmons, L. K., Francis, G., and Martínez-Alonso, S.: 13 years of MOPITT operations: lessons from MOPITT retrieval algorithm development, Ann. Geophys., 56, 1–5, https://doi.org/10.4401/ag-6330, 2014.

Yurganov, L. N., Duchatelet, P., Dzhola, A. V., Edwards, D. P., Hase, F., Kramer, I., Mahieu, E., Mellqvist, J., Notholt, J., Novelli, P. C., Rockmann, A., Scheel, H. E., Schneider, M., Schulz, A., Strandberg, A., Sussmann, R., Tanimoto, H., Velazco, V., Drummond, J. R., and Gille, J. C.: Increased Northern Hemispheric carbon monoxide burden in the troposphere in 2002 and 2003 detected from the ground and from space, Atmos. Chem. Phys., 5, 563-573, https://doi.org/10.5194/acp-5-563-2005, 2005.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques
Discussions

Table 1. In-situ datasets of CO used for MOPITT products validation in this study.

	Period	Region	Number of profiles	Number of profiles over urban	Technique	Reference
DISCOVER-AQ DC	Jul, 2011	Baltimore-Washington, D.C., US	80	36	NASA DACOM	https://www-air.larc.nasa.gov/missions/discover aq/
DISCOVER-AQ CA	Jan-Feb, 2013	California, US	35	12	NASA DACOM	
DISCOVER-AQ TX	Sep, 2013	Texas, US	61	37	NASA DACOM	
DISCOVER-AQ CO	Jul-Aug, 2014	Colorado, US	56	36	NASA DACOM	
SEAC^{4}RS	Aug-Sep, 2013	US	15	1	NASA DACOM	Toon et al. (2016)
A-FORCE	Mar-Apr, 2009; Feb-Mar, 2013; Jun-Jul, 2013	Japan, South Korea, Pacific Ocean	45	2	AL5002, Aero-Laser GmbH	Oshima et al. (2012); Kondo et al. (2016)
KORUS-AQ	May-Jun, 2016	South Korea	47	6	NASA DACOM	Al-Saadi et al. (2015)
ARIAs	May-Jun, 2016	Hebei, East China	19	3	Picarro G2401-i	Wang et al. (2018)

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions

Table 2. Summarized validation results for V7 and V8 TIR-only (V7T and V8T), NIR-only (V7N and V8N) and TIR-NIR (V7J and V8J) products based on in-situ profiles from DISCOVER-AQ, SEAC[4]RS, A-FORCE, KORUS-AQ, and ARIAs.

		Surface			800 hPa			600 hPa		
		All	Urban	Non-urban	All	Urban	Non-urban	All	Urban	Non-urban
V7T	Bias (%)	0.1	-1.7	1.1	0.8	-0.6	1.7	4.0	3.9	4.0
	SD (%)	9.5	8.6	9.8	11.0	9.0	11.9	11.4	9.0	12.7
	r	0.71	0.67	0.72	0.66	0.65	0.66	0.63	0.58	0.64
V8T	Bias (%)	2.0	0.9	2.7	2.2	1.4	2.7	3.5	3.5	3.5
	SD (%)	9.3	9.6	9.0	10.7	9.7	11.2	11.7	10.0	12.6
	r	0.70	0.58	0.75	0.66	0.58	0.69	0.63	0.54	0.66
V7N	Bias (%)	-2.0	-2.8	-1.5	-1.6	-2.1	-1.1	-1.6	-1.9	-1.3
	SD (%)	6.7	6.4	6.9	5.7	5.2	6.0	4.3	4.2	4.4
	r	0.62	0.54	0.67	0.56	0.45	0.61	0.61	0.48	0.68
V8N	Bias (%)	1.4	0.4	2.2	1.6	0.9	2.1	1.2	0.8	1.5
	SD (%)	6.9	6.7	6.9	6.0	5.8	6.1	4.6	4.7	4.5
	r	0.60	0.52	0.67	0.54	0.40	0.62	0.59	0.42	0.68
V7J	Bias (%)	-5.4	-8.9	-3.3	-3.9	-6.5	-2.4	-0.5	-1.4	0.1
	SD (%)	13.5	12.1	13.9	14.2	12.4	15.0	13.6	11.0	14.8
	r	0.68	0.63	0.70	0.64	0.58	0.66	0.60	0.52	0.62
V8J	Bias (%)	0.0	-2.0	1.1	-0.7	-1.6	-0.1	-0.5	-0.8	-0.3
	SD (%)	12.7	13.7	12.0	12.9	12.5	13.1	12.8	10.9	13.8
	r	0.69	0.53	0.76	0.69	0.57	0.73	0.65	0.53	0.67

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions

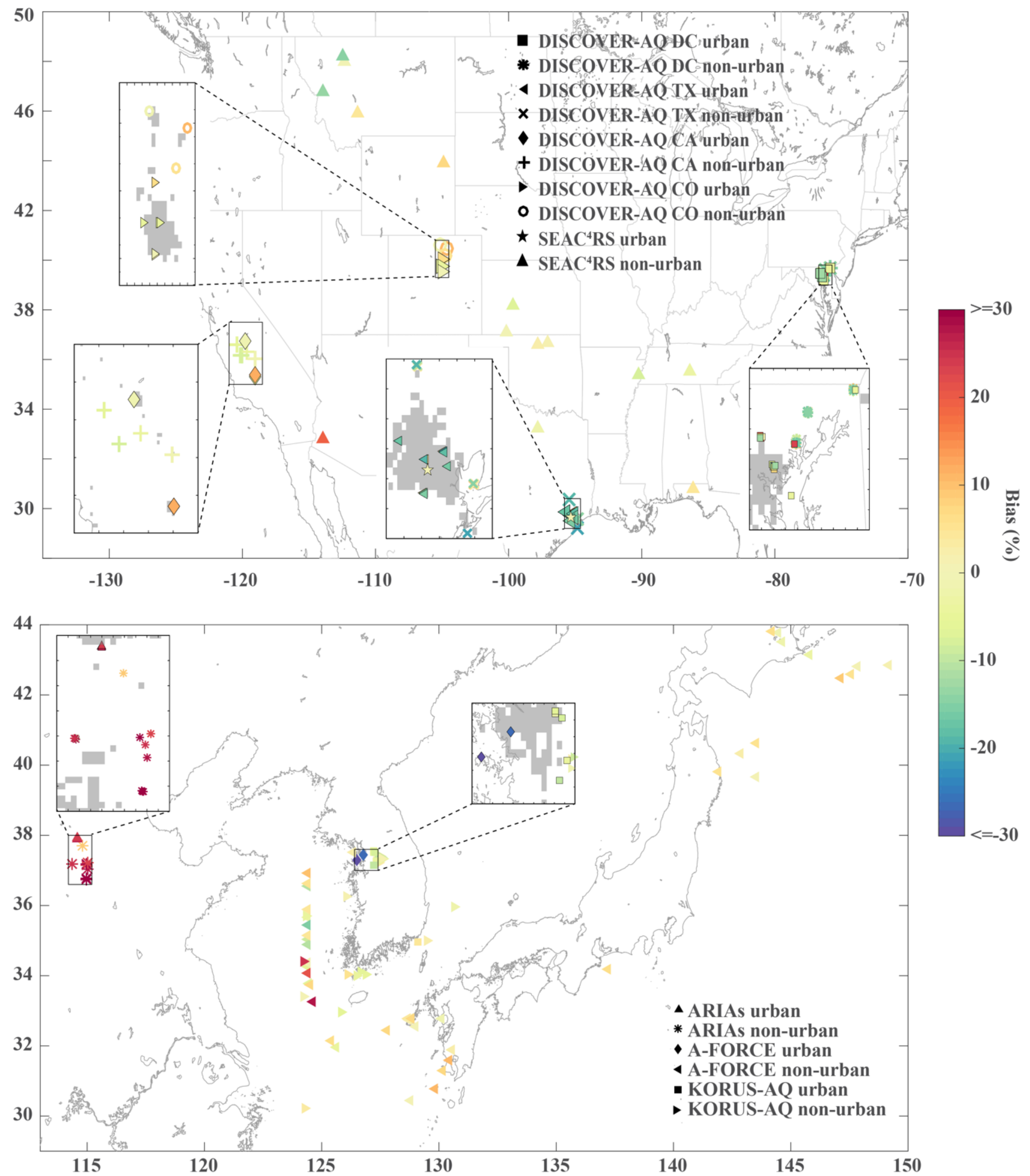

Figure 1. Spatial distributions of aircraft profiles from the DISCOVER-AQ, SEAC^{4}RS, ARIAs, A-FORCE, and KORUS-AQ campaigns. Urban and built-up land cover (from MCD12C1 v006) are shown by gray shade in the boxes. Bias of MOPITT V8J comparing to the aircraft profile at the surface level are shown by the color of the profile.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques
Discussions

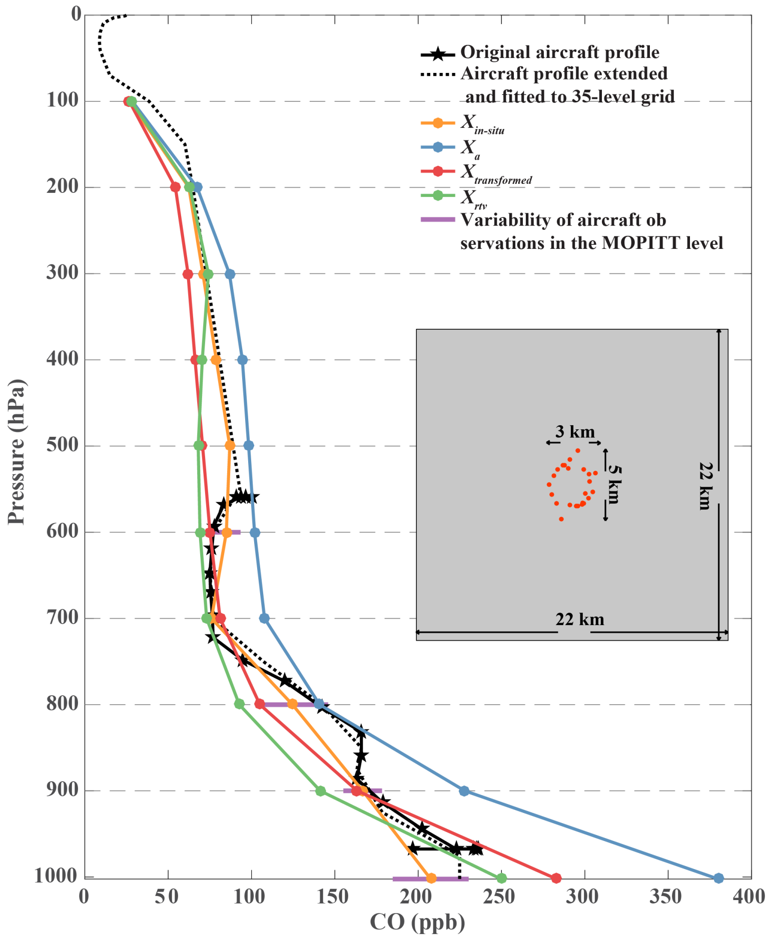

Figure 2. Example of profile comparisons for an aircraft profile sampled on July 22, 2011 during DISCOVER-AQ DC. The black solid line represents the original aircraft profile and the stars represent the original aircraft observations, the black dotted line is the aircraft profile extended with CAMS reanalysis data, and regridded to 35-level grid. The in-situ profile regridded at 10-level grid ($x_{in-situ}$), the MOPITT a priori profile (x_a), the in-situ profile transformed with the MOPITT a priori and AK ($x_{transformed}$), and the MOPITT retrieved profile (x_{rtv}) are shown in colored lines with dots. The purple bars centered at the $x_{in-situ}$ at each MOPITT retrieval level show the vertical and horizontal variability of the original aircraft observations in the MOPITT layer, indicated by standard deviation. Note that each MOPITT retrieval level corresponds to a uniformly-weighted layer immediately above that level. Superimposed gray box shows the horizontal scale of the profile (each aircraft observation is represented by a red dot) and a MOPITT pixel (gray box).

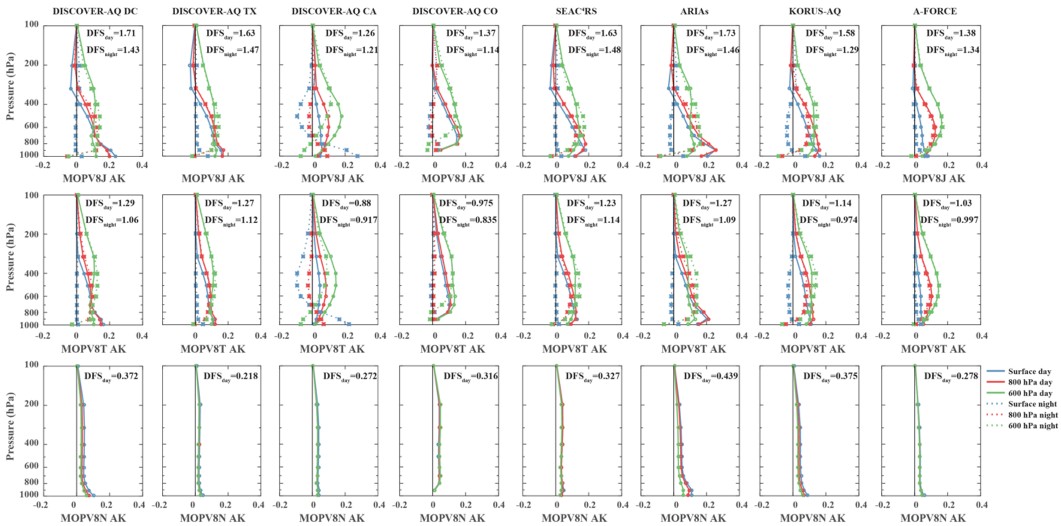

Figure 3. Mean retrieval averaging kernels for the MOPITT V8J, V8T, and V8N for the corresponding in-situ profiles from the DISCOVER-AQ, SEAC^{4}RS, ARIAs, KORUS-AQ, and A-FORCE at daytime (solid lines) and nighttime (dashed lines).

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

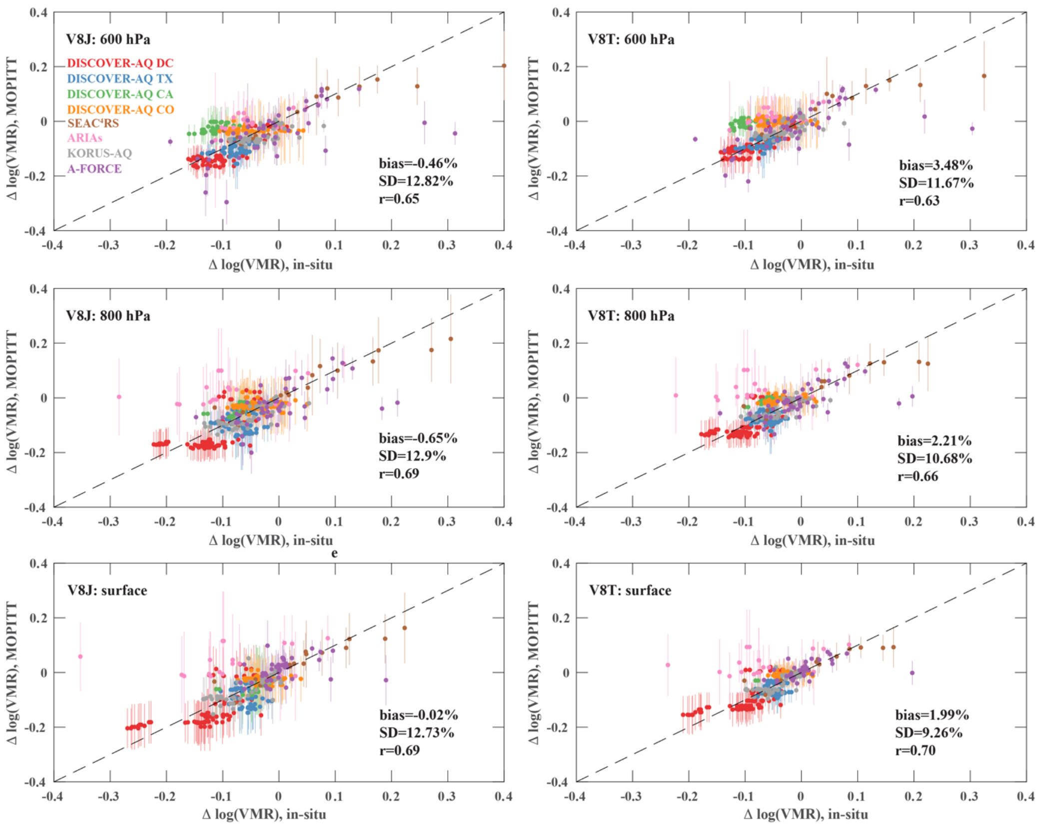

Figure 4. MOPITT V8J and V8T validation results over both urban and non-urban regions at 600 hPa, 800 hPa, and the surface in terms of Δlog (VMR). The variability of the MOPITT data used to calculate each of the plotted mean values are represented by the vertical error bars.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques
Discussions

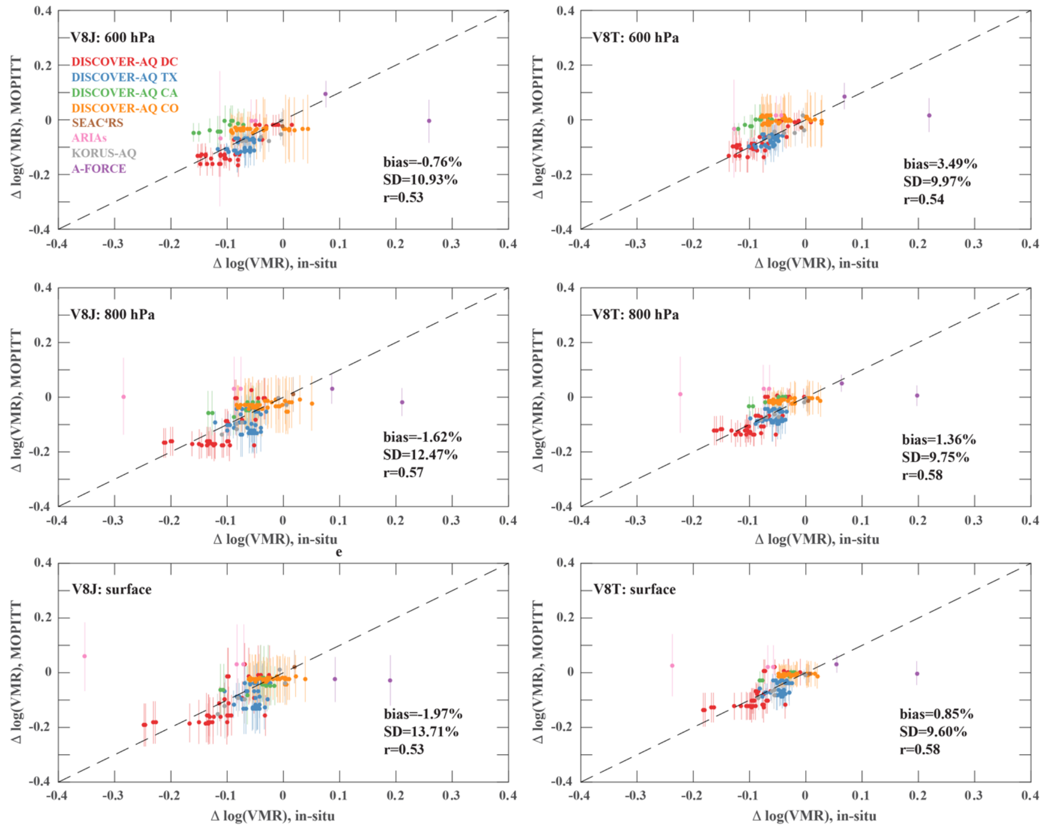

Figure 5. MOPITT V8J and V8T validation results against aircraft profiles over urban regions at 600 hPa, 800 hPa, and the surface in terms of Δlog (VMR). See caption to Figure 2.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques Discussions
Open Access

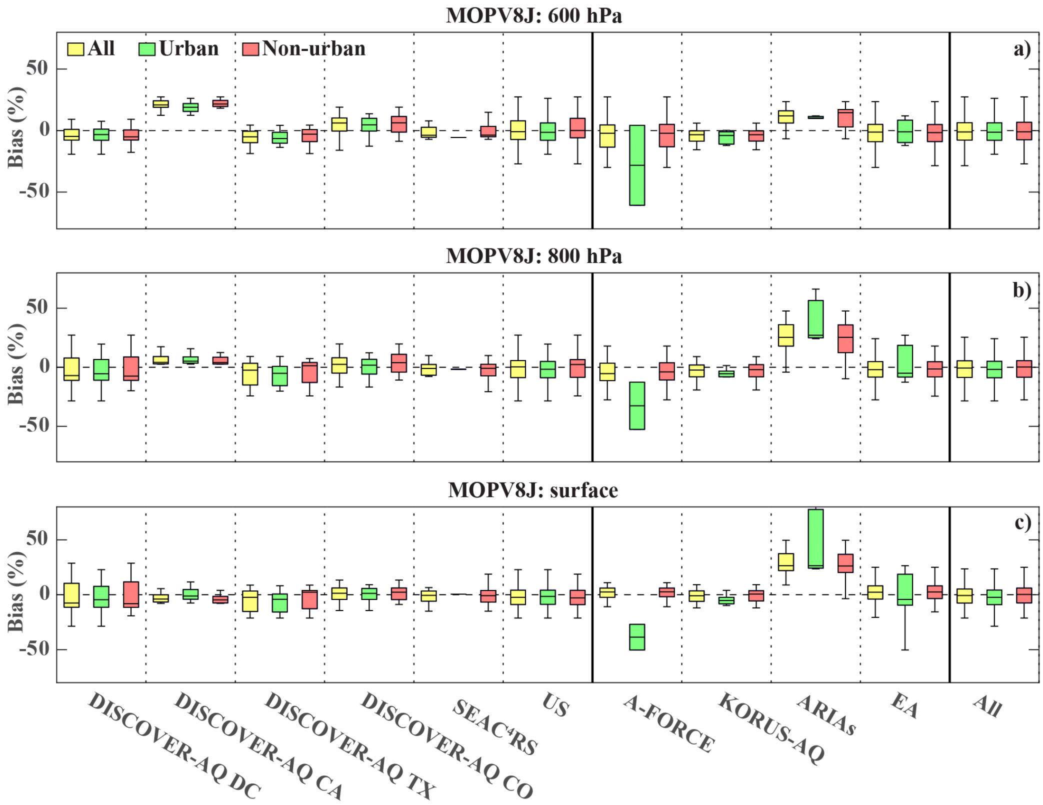

Figure 6. Boxplot (with medians represented by middle bars, interquartile ranges between 25th and 75th percentiles represented by boxes, and the most extreme data points not considered outliers represented by whiskers) for biases (%) for the profiles over both urban and non-urban regions (yellow), profiles over urban regions (green), and profiles over non-urban regions (red) at 600 hPa (panel a), 800 hPa (panel b), and the surface (panel c).

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques
Discussions

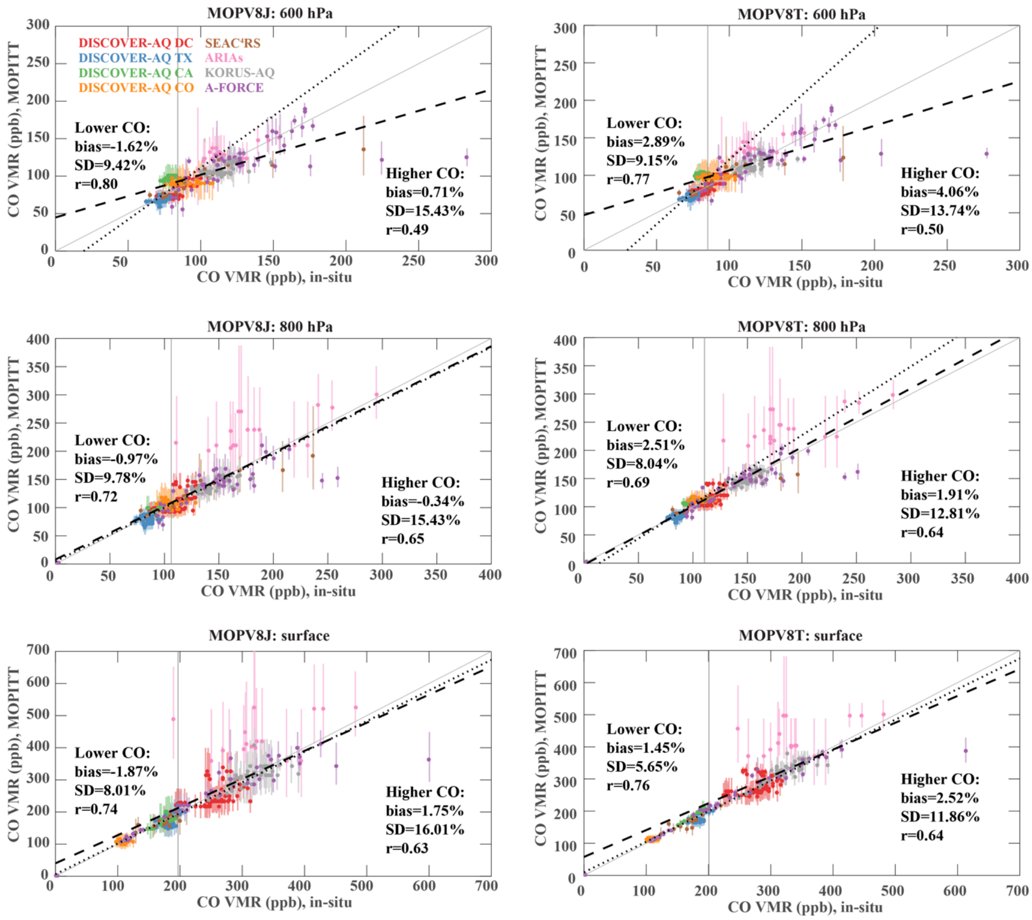

Figure 7. MOPITT V8J and V8T validation results at 600 hPa, 800 hPa, and the surface against the lower 50% in-situ profiles of CO and higher 50% in-situ profiles of CO. The variability of the MOPITT data used to calculate each of the plotted mean values are represented by the vertical error bars. Each panel shows the least-squares best-fit lines for the lower 50% CO concentrations (dotted line) and the higher 50% CO concentrations (dashed line).

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques
Open Access
EGU
Discussions

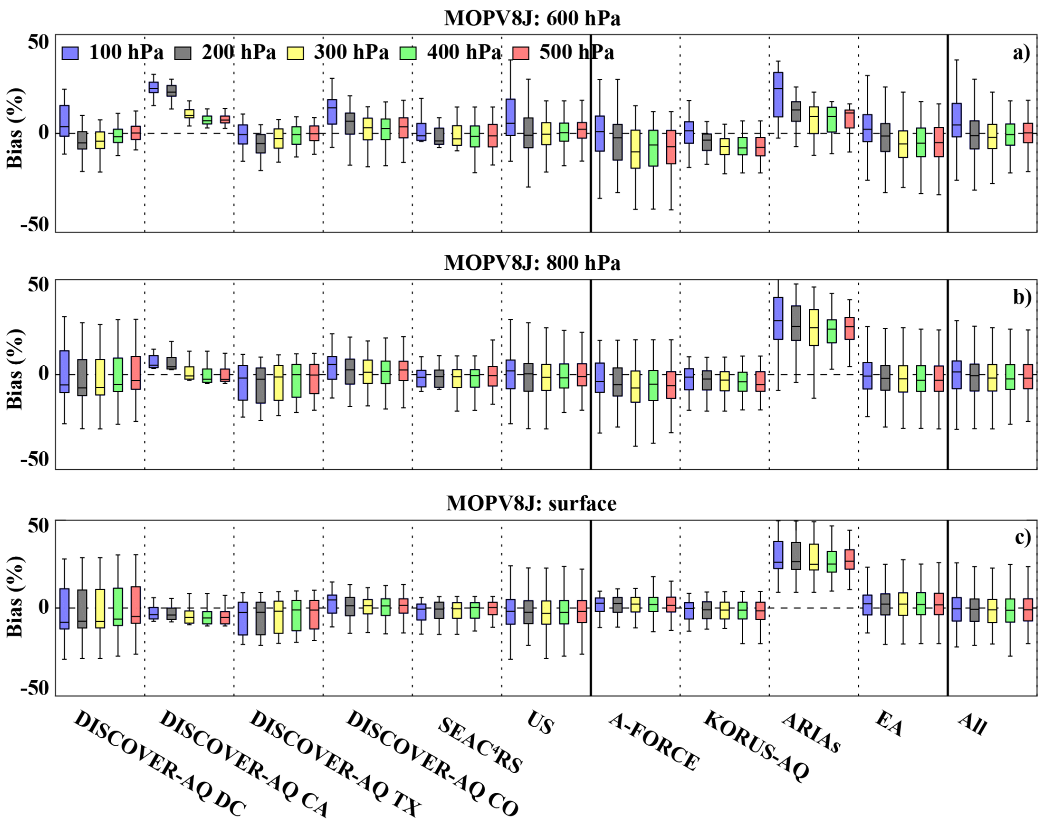

Figure 8. Sensitivity to P_{interp}. Biases (%) using 100 hPa (blue), 200 hPa (gray), 300 hPa (yellow), 400 hPa (green), and 500 hPa (red) as P_{interp} at 600 hPa (panel a), 800 hPa (panel b), and the surface (panel c) are shown by boxplot (with medians represented by middle bars, interquartile ranges between 25th and 75th percentiles represented by boxes, and the most extreme data points not considered outliers represented by whiskers).

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

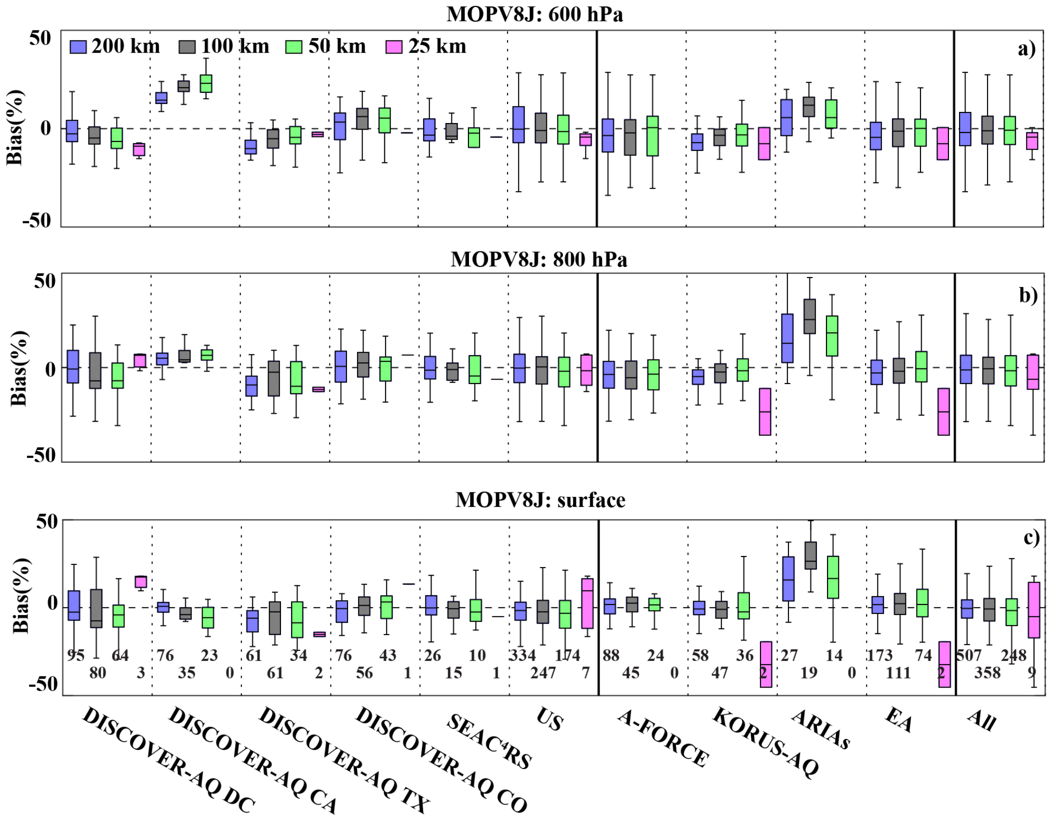

Figure 9. Sensitivity to the radius as criteria for co-location. Biases (%) using 200 km (blue), 100 km (gray), 50 km (green), and 25 km (pink) as the radius for co-location at 600 hPa (panel a), 800 hPa (panel b), and the surface (panel c) are shown by boxplot (with medians represented by middle bars, interquartile ranges between 25th and 75th percentiles represented by boxes, and the most extreme data points not considered outliers represented by whiskers). The numbers in panel c correspond to the number of in-situ profiles qualified for validation within the given radius.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques
Discussions

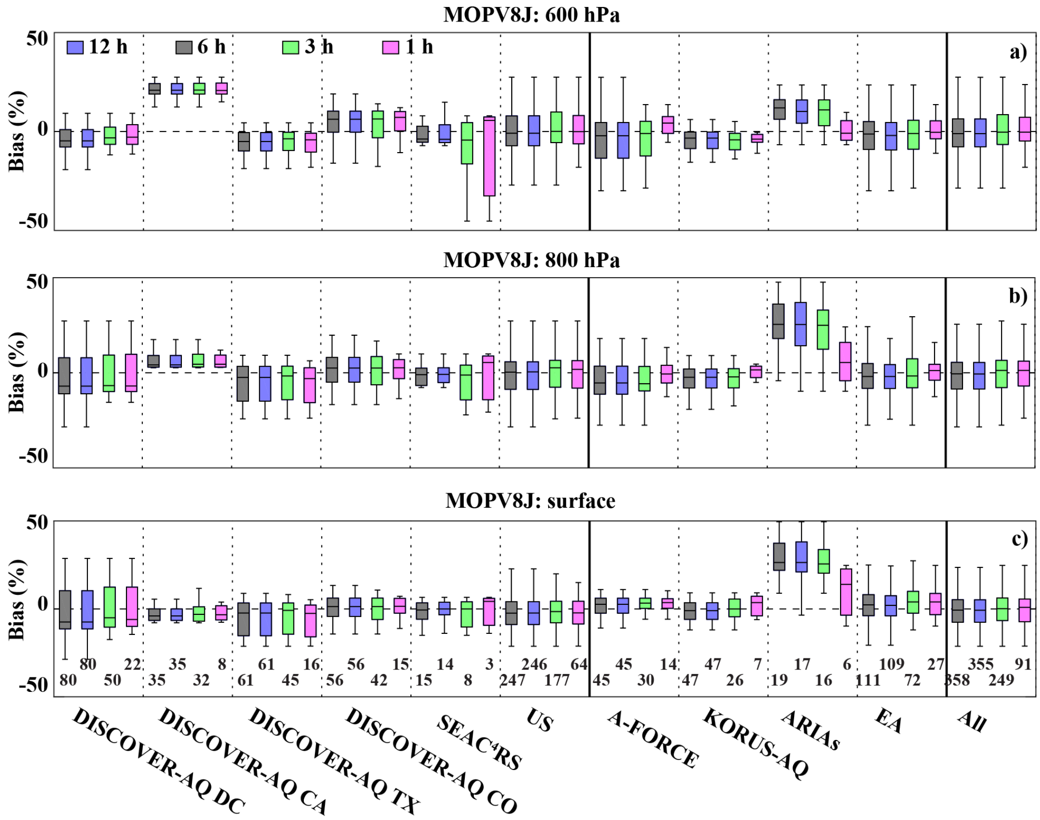

Figure 10. Sensitivity to the allowed maximum time difference as criteria for co-location. Biases (%) using 12 hour (blue), 6 hour (gray), 3 hour (green), and 1 hour (pink) as the allowed maximum time difference for co-location at 600 hPa (panel a), 800 hPa (panel b), and the surface (panel c) are shown by boxplot (with medians represented by middle bars, interquartile ranges between 25th and 75th percentiles represented by boxes, and the most extreme data points not considered outliers represented by whiskers). The numbers in panel c correspond to the number of in-situ profiles qualified for validation within the given allowed maximum time difference.

https://doi.org/10.5194/amt-2019-419
Preprint. Discussion started: 28 November 2019

Atmospheric Measurement Techniques
Discussions

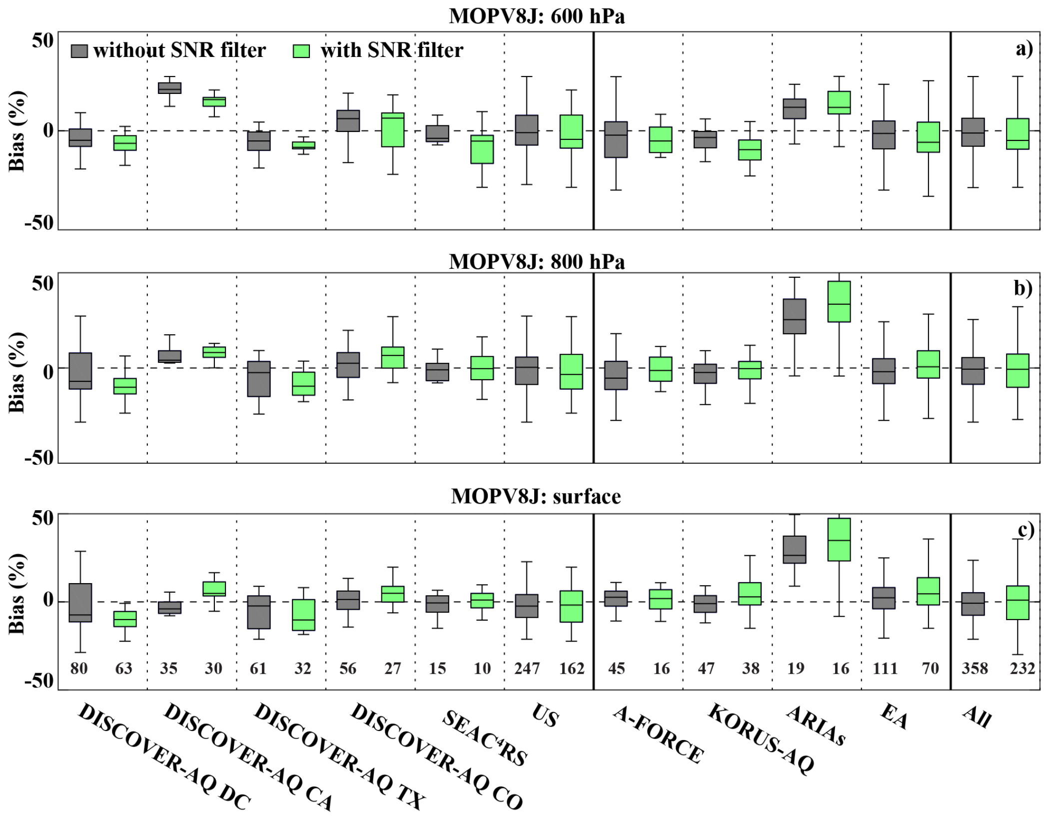

Figure 11. Sensitivity to the signal-to-noise ratio (SNR) filters. Biases (%) for MOPITT retrievals without SNR filters (gray), and MOPITT retrievals with SNR filters (green) at 600 hPa (panel a), 800 hPa (panel b), and the surface (panel c) are shown by boxplot (with medians represented by middle bars, interquartile ranges between 25th and 75th percentiles represented by boxes, and the most extreme data points not considered outliers represented by whiskers). The numbers in panel c correspond to the number of in-situ profiles qualified for validation without or with SNR filters.