Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018


Geoscientific Model Development Discussions

# Assimilating Compact Phase Space Retrievals (CPSRs): Comparison with Independent Observations (MOZAIC *in situ* and IASI Retrievals and Extension to Assimilation of Truncated Retrieval Profiles

Arthur P. Mizzi[1], David P. Edwards[1], Jeffrey L. Anderson[2]

[1]National Center for Atmospheric Research, Atmospheric Chemistry Observations and Modelling Laboratory, Boulder, CO 80305, USA
[2]National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO, 80305, USA

*Correspondence to*: Arthur P. Mizzi (mizzi@ucar.edu)

**Abstract**

Assimilation of atmospheric composition retrievals presents computational challenges due to their high data volume and often sparse information density. Assimilation of compact phase space retrievals (CPSRs) meets those challenges and offers a promising alternative to assimilation of raw retrievals at reduced computational cost (Mizzi et al., 2016). This paper compares analysis and forecast results from assimilation of Terra/Measurement of Pollution in the Troposphere (MOPITT) carbon monoxide (CO) CPSRs with independent observations. We use MetOp-A/Infrared Atmospheric Sounding Interferometer (IASI) CO retrievals and Measurement of OZone, water vapor, carbon monoxide, and nitrogen oxides by in-service AIrbus airCraft (MOZAIC) *in situ* CO profiles for our independent observation comparisons. Generally, the results confirm that assimilation of MOPITT CPSRs improved the WRF-Chem/DART analysis fit and forecast skill at a reduced computational cost (~35% reduction) when compared to assimilation of raw or quasi-optimal retrievals (QORs). Comparison with the independent observations shows that assimilation of MOPITT CO generally improved the analysis fit and forecast skill in the lower troposphere but degraded it in the upper troposphere. We attribute that degradation to assimilation of MOPITT CO retrievals with a possible bias of ~14% above 300 hPa. To discard the biased retrievals, in this paper we also extend CPSRs to assimilation of truncated retrieval profiles (as opposed to assimilation of full retrieval profiles). Those results show that not assimilating the biased retrievals: (i) resolves the upper tropospheric analysis fit degradation issue, (ii) has commensurate reductions in assimilation computation cost, and (iii) reduces the impact of assimilating the remaining unbiased retrievals because the total information content and vertical sensitivities are changed.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




## 1 Introduction

Mizzi et al. (2016) introduced the assimilation of "compact phase space retrievals" (CPSRs) to address challenges associated with assimilating retrievals of atmospheric composition. They showed that assimilation of CPSRs reduced computation costs while maintaining or improving analysis fit and forecast skill. They reached that conclusion by comparing their results with assimilated observations. In the first part of this paper, we compare the results of assimilating CPSRs with independent observations. As in Mizzi et al. (2016), we assimilate conventional meteorological observations and Terra/Measurement of Pollution in the Troposphere (MOPITT) CO retrievals, but here we compare our analysis and forecast results with MetOp-A/Infrared Atmospheric Sounding Interferometer (IASI) CO retrievals and Measurement of OZone, water vapor, carbon monoxide, and nitrogen oxides by in-service AIrbus airCraft (MOZAIC) *in situ* CO profiles. Those comparisons generally show improved analysis fit and forecast skill from assimilating MOPITT CO retrievals but in the upper troposphere there is degraded skill possibly due to assimilation of retrievals with a positive bias of ~14% (Deeter et al., 2013 and Martinez-Alonso et al., 2014). In the second part of this paper, we extend the CPSR algorithm to assimilate truncated retrieval profiles and discard (i.e., do not assimilate) the biased retrievals. The rest of this paper is organized as follows: In the next section – Section II, we describe the forecast/data assimilation system together with the assimilated meteorological and chemistry observations. Section III describes the independent IASI and MOZAIC observations. Section IV presents descriptions of our experiments, retrieval pre-processing methods, and extension of CPSRs to truncated retrieval profiles. Section V compares the results of assimilating MOPITT CO retrievals (full and truncated profiles) with the IASI and MOZAIC CO observations. Finally, Section VI presents a summary of our results and conclusions.

## 2 WRF-Chem/DART Regional Forecasting Ensemble Data Assimilation System: Set-Up and Assimilated Observations

For the experiments reported here, we use the WRF-Chem/DART regional chemical transport/ensemble Kalman filter data assimilation system introduced by Mizzi et al. (2016). WRF-Chem/DART is made up of the Weather Research and Forecasting (WRF) model with chemistry (WRF-Chem) (www2.acd.ucar.edu/wrf-chem) coupled to the ensemble Kalman filter data assimilation utility from the Data Assimilation Research Testbed (DART) (www.image.ucar.edu/DAReS/DART; Anderson et al., 2009). WRF-Chem is a regional model that predicts conventional weather together with the transport, mixing, and chemical



transformation of atmospheric trace gases and aerosols. DART is an ensemble data assimilation system that uses the ensemble adjustment Kalman filter of Anderson (2001, 2003) together with adaptive inflation and localization.

We conduct continuous cycling experiments with 6-hr cycling (00, 06, 12, and 18 UTC) for the period 1 June 2008 00 UTC to 9 June 2008 18 UTC. To facilitate a large number of experiments, we use a reduced ensemble size of 20 members, a horizontal resolution of 100km (101 x 41 grid points), and an abbreviated 9-day study period (compared to the 30-day period used in Mizzi et al. (2016)). The reduced study period is not thought to negatively impact our results because the WRF-Chem/DART spin-up occurs within the first 48 to 72 hours. The WRF-Chem domain extends from ~176 W to ~50 W and ~7 N to ~54 N. We use 34 vertical levels with a model top at 10 hPa and ~15 levels below 500 hPa. We use DART adaptive prior covariance inflation with the recommended settings and DART three-dimensional Gaspari-Cohn localization with a localization radius half-width of ~300 km in the horizontal. (Anderson, 2008). Vertical localization is not used. These are the same settings as used by Mizzi et al. (2016).

The WRF-Chem initial and boundary conditions are derived from the National Oceanic and Atmospheric Administration/National Center for Environmental Prediction (NOAA/NCEP) Global Forecast Model (GFS) 0.5º six-hour forecasts. The WRF Preprocessing System (WPS) interpolates the GFS forecasts to our domain and generates the deterministic boundary conditions. We use the WRF Data Assimilation System (WRFDA) (*http://www2.mmm.ucar.edu/wrf/users/docs/user_guide/users_guide_chap6*; Barker et al., 2012) to generate the initial meteorological ensemble. The chemistry initial and boundary conditions are derived from the Model for Ozone and Related Chemical Tracers: MOZART-4 (MOZART) forecasts, and WRF-Chem utilities are used to interpolate those forecasts to our domain and generate the deterministic chemistry boundary conditions. The emissions and initial chemistry ensembles are generated as described in Mizzi et al. (2016). The ensemble distributions are Gaussian with a specified mean and standard deviation. The tails of those distributions are truncated to include 95% of the distribution and exclude anomalous outliers. That strategy ensures that the emissions and initial chemistry variable concentrations are positive definite. We do not include horizontal correlations for the emission perturbations because they are not relevant to the focus of this paper.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018


Geoscientific Model Development Discussions

At each cycle time depending on the experiment, we assimilate conventional meteorological and chemistry observations with DART and advance the analysis ensemble to the next cycle time with WRF-Chem. The resulting 6-hr forecast ensemble is then used as the first guess in the next assimilation step. Our conventional meteorological observations are NCEP automated data processing (ADP) upper air and surface observations (PREPBUFR observations), and our chemistry observations are MOPITT CO mixing ratio retrieval profiles. MOPITT is an instrument on the National Aeronautics and Space Administration's (NASA's) Earth Observing System Terra satellite. Its spatial resolution is 22 km at nadir over a swath width of 640 km. Its thermal infra-red (TIR) measurements are sensitive to CO in the middle and upper troposphere, while its near infra-red (NIR) measurements are sensitive to total column CO. We assimilate the MOPITT V5 thermal-infrared/near-infrared (TIR/NIR) retrieval products described by Deeter et al. (2013).

The horizontal resolution of the MOPITT data is much greater than that at which we run WRF-Chem. That difference translates to representativeness errors due to the smaller spatial scales that are resolved in the satellite data and not in the model. To address those errors, we construct super-observations as follows, we: (i) sort the retrievals, retrieval priors, averaging kernels, and retrieval error covariances into bins that are ~90 km square, (ii) calculate the bin-average for each of those variables, and (iii) assimilate the bin-average retrievals. We use an arithmetic average (as opposed to error covariance weighted average) when calculating the super-observation and do not apply a correction to the retrieval error covariance super-observation because we are interested in the impact of the reported errors and can apply an error tuning factor to adjust the errors and balance the observation fit as needed, however, the results reported in this paper are not tuned. Other studies e.g., Eskes et al. (2003), Miyazki et al. (2012 a and b, 2015), and Barre et al. (2016) have used similar super-observation strategies. We do not expect that tuning the observation errors would significantly impact our results because diagnostic analyses showed that the CO total error and forecast root-mean square error (RMSE) were properly balanced.

## 3 Independent Observations for Verification: MOZAIC *in situ* and IASI CO Retrieval Profiles

In the first part of this paper, we compare the results of assimilating MOPITT CO with independent observations (IASI CO



retrievals and MOZAIC *in situ* CO profiles). IASI is an instrument on the EUMETSAT (European Organization for the Exploitation of Meteorological Satellites) polar orbiting MetOp-A satellite. Clerbaux et al. (2009). It measures temperature, water vapor, fractional cloud cover, cloud top temperature, ozone, carbon monoxide, and methane. IASI has been operating from 2006 to the present. Its mission is to provide observational support for numerical weather prediction. IASI measures CO radiances under cloud-free conditions with a horizontal resolution of 25 km over a swath width of ~2,200 km. The IASI measurements are sensitive to CO in the mid- to lower troposphere. For more information see www.eumetsat.int.

MOZAIC was a European Research Infrastructure (ERI) project that collected long-term, global-scale measurements of atmospheric composition on international commercial airline flights from August 1994 to November 2014. Marenco et al. (1998). MOZAIC collected *in-situ* measurement of ozone, water vapor, carbon monoxide, and total nitrogen oxides. The available data products are geo-located (come with longitude, latitude, and pressure coordinates) and include simultaneous meteorological observations. During MOZAIC data acquisition was automatically performed on the ascent, descent, and cruise phases of round-trip international flights between Europe and America, Africa, the Middle East, and Asia. For more information see www.iagos.fr.

## 4 Experimental Design

We conduct WRF-Chem/DART forecast/assimilation cycling experiments that are similar to those of Mizzi et al. (2016). The primary differences are the: (i) use of super-observations, (ii) extension of CPSRs to truncated retrieval profiles, and (iii) use of localization to preclude the assimilated MOPITT CO observations from impacting any state variable other than CO. We performed a control experiment where we assimilated only conventional meteorological observations (the MET experiment), and we performed a series of chemical data assimilation experiments. In those experiments, we studied assimilation results from four types of retrieval pre-processing strategies: (i) Volume Mixing Ratio retrievals (VMRRs, the associated experiment is called the VMRR experiment), (ii) $Log_{10}$(VMRR) retrievals (L10VMRRs, called the L10VMRR experiment), (iii) Compact Phase Space Retrievals (CPSRs, called the CPSR experiment), and (iv) Quasi-Optimal Retrievals (QORs, called the QOR experiment). The CPSR and QOR experiments (as applied to assimilation of retrieval full profiles) were studied by Mizzi et

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




al. (2016). The VMRR experiment and the L10VMRR and CPSR experiments as applied to assimilation of truncated retrieval profiles are new. We include the L10VMR and QOR experiments as applied to retrieval full profiles because, as discussed in the Introduction, our comparison of those experiments with independent observations (discussed below in Section V.A) suggests that it may be beneficial to not assimilate MOPITT CO retrievals in the upper troposphere due to their possible bias. That concern motivates application of the L10VRR and CPSR experiments to the assimilation of truncated retrieval profiles (retrieval profiles obtained after discarding the potentially biased retrievals). The rest of this section describes those experiments. It should be noted that the different retrieval pre-processing methods (making up the different experiments) are applied after the customary quality assurance/quality control (QA/QC) checks that might discard entire retrieval profiles. Those forecast/assimilation experiments are summarized in Table 1.

### 4.1 The VMRR and L10VMRR Experiments

The MOPITT CO retrieval, averaging kernel, and error covariance products are reported in units of $\log_{10}$(VMR). The IASI CO products are in VMR. For ease of comparison and interpretation, it is convenient to convert the MOPITT data to VMR. While it is possible to convert the retrievals and error covariance, it is not possible to convert the averaging kernels. Consequently, the DART forward operator for MOPITT CO in VMR converts the state space CO profile from VMR to $\log_{10}$(VMR), applies the averaging kernel, and then converts the resulting expected observation (the expected retrieval profile) to VMR. Conceptually, we expect little difference between results from assimilating VMRRs and L10VMRRs due to an underlying assumption that the L10VMRRs have a Gaussian distribution (Deeter et al., 2007). However, non-linearity of the base-ten exponential operator that relates the L10VMRRs to the VMRRs and the extent to which the VMRR distributions are non-Gaussian may introduce differences. So, one goal of the related experiments is to determine whether those differences are significant. Another reason is to include pre-processing methods that enable us to not assimilate selected retrievals so we can compare the assimilation/forecast results with those from applying CPSRs to truncated retrieval profiles.

### 4.2 The QOR Experiment

The assimilation of QORs was discussed in Mizzi et al. (2016). We include QOR assimilation/forecast experiments for completeness and to provide a reference against which to compare the other retrieval pre-processing experiments. In addition

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




(although not discussed herein), QOR pre-processing can be applied to truncated retrieval profiles using the extension discussed in the next section on the CPSR experiment.

QORs are phase space retrievals introduced by Migliorini et al. (2008). They are derived by writing the retrieval equation as

$$\boldsymbol{y}_r - (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{y}_a - \boldsymbol{\varepsilon} = \boldsymbol{A}y_t \qquad (1)$$

where $\boldsymbol{y}_r$ is the retrieval profile (column vector, dimension $n$ – the number of observations in a full retrieval profile), $\boldsymbol{I}$ is the identity matrix (square matrix, dimension $n$ x $n$), $\boldsymbol{A}$ is the averaging kernel (square matrix, dimension $n$ x $n$, and rank $k$, where $k < n$), $\boldsymbol{y}_a$ is the retrieval prior profile (column vector, dimension $n$), $\boldsymbol{\varepsilon}$ is the measurement error in retrieval space (column vector, dimension $n$) with error covariance $\boldsymbol{E}_m$ (square matrix, dimension $n$ x $n$), and $\boldsymbol{y}_t$ is the unknown true atmospheric profile (column vector, dimension $n$). In this paper we transform Eq. (1) with the left singular vectors from a Singular Value Decomposition (SVD) of $\boldsymbol{E}_m$. If a SVD of $\boldsymbol{E}_m$ is $\boldsymbol{E}_m = \boldsymbol{\phi}\boldsymbol{\sigma}\varphi^T$, then the QOR profile is defined as

$$\boldsymbol{\sigma}^{-1/2}\boldsymbol{\phi}^T(\boldsymbol{y}_r - (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{y}_a - \boldsymbol{\varepsilon}) = \boldsymbol{\sigma}^{-1/2}\boldsymbol{\phi}^T\boldsymbol{A}y_t \qquad (2)$$

and the transformed $\boldsymbol{E}_m$ is the identity matrix. That diagonalization transform is similar to the CPSR diagonalization transform described in the next section except the QOR transform is applied to the raw averaging kernel and error covariance while the CPSR diagonalization transform is applied to the compressed averaging kernel and error covariance.

### 4.3 The CPSR Experiment and the Extension of CPSRs to Assimilation of Truncated Retrieval Profiles

The derivation and assimilation of CPSRs was first introduced by Mizzi et al. (2016). They derived CPSRs by applying two transforms to Eq. 1: (i) a compression transform based on the SVD of $\boldsymbol{A}$, and (ii) a diagonalization transform based on the SVD of the compressed $\boldsymbol{E}_m$. Their application can be characterized as CPSRs applied to full retrieval profiles (because none of the elements in the retrieval profile were discarded) or to square systems (because $\boldsymbol{A}$ is a square matrix). If we discard one or more elements of $\boldsymbol{y}_r$, then we must also discard the corresponding rows of $\boldsymbol{A}$ (call the modified forms $\hat{\boldsymbol{y}}_r$ and $\hat{\boldsymbol{A}}$ respectively)**.** The resulting $\hat{\boldsymbol{A}}$ is not a square matrix. Note that we must also discard the corresponding rows and columns of $\boldsymbol{E}_m$, so it remains square but its dimension is reduced. This application can be characterized as CPSRs applied to truncated retrieval profiles (because some of the elements of the retrieval profile have been discarded) or to rectangular systems (because $\hat{\boldsymbol{A}}$ is a non-square rectangular matrix). The mathematical formalism for CPSRs applied to rectangular systems is the same as that for

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




square systems because Mizzi et al. (2016) used a SVD (as opposed to an eigenvalue decomposition) in their CPSR derivation. In the remainder of this section, we extend the derivation of CPSRs from Mizzi et al. (2016) to rectangular systems.

We begin by conceptually discarding $q$ elements of $\boldsymbol{y}_r$. Generally, we discard the elements of the full retrieval profile $\boldsymbol{y}_r$ that are known to be systematically bad observations. If we discard multiple elements, they need not be sequential. The resulting truncated retrieval profile is denoted $\hat{\boldsymbol{y}}_r$ and its dimension is $\hat{n} = n - q$. We must also discard: (i) the corresponding elements of $\boldsymbol{\varepsilon}$ to get $\hat{\boldsymbol{\varepsilon}}$ with dimension $\hat{n}$, (ii) the corresponding rows of $\boldsymbol{A}$ to get $\hat{\boldsymbol{A}}$ with dimension $\hat{n} \times n$, and (iii) the corresponding rows and columns of $\boldsymbol{E}_m$ to get $\hat{\boldsymbol{E}}_m$ with dimension $\hat{n} \times \hat{n}$. Without loss of generality, we can drop the $\hat{}$ notation for the remainder of this paper and let $\boldsymbol{y}_r$, $\boldsymbol{\varepsilon}$, $\boldsymbol{A}$, and $\boldsymbol{E}_m$ represent their respective terms before and after discarding the retrieval elements that will not be assimilated. The rest of the derivation is the same as in Mizzi et al. (2016).

First, we apply a compression transform based on the leading left singular vectors of $\boldsymbol{A}$. If $\boldsymbol{A} = \boldsymbol{USV^T}$ is a SVD and $A_0 = U_0 S_0 V_0^T$ is the truncated SVD where the trailing singular vectors (those whose singular values are less than an *ad hoc* threshold of $1.0 \times 10^{-4}$) are replaced with zero vectors and the trailing singular values are set to zero, then the compressed form of Eq. 1 is

$$\boldsymbol{U}_0^T(\boldsymbol{y}_r - (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{y}_a - \boldsymbol{\varepsilon}) = \boldsymbol{S}_0\boldsymbol{V}_0^T\boldsymbol{y}_t \qquad (3)$$

and the compressed error covariance is

$$\boldsymbol{U}_0^T\boldsymbol{E}_m\boldsymbol{U}_0. \qquad (4)$$

Next, we apply a diagonalization transform. If the SVD of the compressed error covariance in (4) is $\boldsymbol{U}_0^T\boldsymbol{E}_m\boldsymbol{U}_0 = \boldsymbol{\Phi\Sigma\Psi}^T$, then the diagonalized and conditioned form of Eq. 3 is

$$\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Phi}^T\boldsymbol{U}_0^T(\boldsymbol{y}_r - (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{y}_a - \boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Phi}^T\boldsymbol{S}_0\boldsymbol{V}_0^T\boldsymbol{y}_t \qquad (5)$$

and that of (4) is the identity matrix. Eqs. 3 – 5 and the fully transformed error covariance are the same as in Mizzi et al. (2016) except that unwanted retrieval elements have been discarded.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




Finally, we note that the rank of $\mathbf{A}$ and the rank of $\hat{\mathbf{A}}$ are the same provided the difference between the dimension of $\mathbf{A}$ and the rank of $\boldsymbol{A}$ is greater than or equal to the number of discarded elements from the retrieval profile i.e., $n - k \geq q$. We also note that the $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Phi}^T\boldsymbol{S}_0\boldsymbol{V}_0^T$ on the right side of Eq. 5 is the transformed averaging kernel. It represents the sensitivity of changes in the phase space observations (the CPSR) to changes in the true CO concentrations at each vertical level. Unlike the raw averaging kernel, which included sensitivities to the null space contributions to the retrieval (the linearly dependent contributions from the right side of Eq. 1), the transformed averaging kernel contains only sensitivities for the measurement contributions to the retrieval (the linearly independent contributions from the right side of Eq. 1).

## 5 Results

### 5.1 Assimilation of Full Retrieval Profiles

In this section, we look at assimilation/forecast results from the experiments described in Section 4. The reader should note that the CPSR and QOR experiments are the same as the MOP CPSR and MOP QOR experiments from Mizzi et al. (2016) except: (i) the study period is shorter (nine days as opposed to one month), and (ii) we assimilate MOPITT super-observations.

The upper panels of Fig. 1 show forecast verification statistics (RMSE and Bias) for the different experiments when compared against MOPITT CO retrievals on the left and IASI CO retrievals on the right. For the MOPITT comparison, the MOPITT CO forward operator has been applied to the WRF-Chem results so the comparison is made in MOPITT CO retrieval space. Similarly, for the IASI comparison the IASI CO forward operator has been applied so the comparison is made in IASI CO retrieval space. The left panel can be compared with Fig. 8 from Mizzi et al. (2016). Qualitatively that comparison shows that the two figures are similar. The MET experiment yields the higher RMSE and bias while the CPSR and QOR experiments yield the lower RMSE and bias. Similar results are seen in the IASI CO comparison. It is interesting that for both comparisons: (i) The VMRR experiment shows a slight degradation when compared to the MET experiment, and (ii) the L10VMRR experiment is similar to the MET experiment and shows a small improvement when compared to the VMRR experiment. We suspect that result (i) is a consequence of the non-linearity of the base-ten log function and the non-Gaussianity of the VMRR distributions, and result (ii) is a consequence of the magnitude of the MOPITT observation errors used in the L10VMRR



experiment (discarding the observation error cross-covariance produced observation error variance that is large compared to that produced by the CPSR diagonalization transform).

Both upper panels of Fig. 1 generally show increasing improvement when moving from the MET ➔ VMRR ➔ L10VMRR ➔ CPSR and QOR experiments. We hypothesize that the improvement of the CPSR and QOR experiments over the VMRR and L10VMRR experiments is due to the phase space transforms which truncate and reduce the effective phase space observation errors resulting in the improved analysis fit and forecast skill. The similarity of the CPSR and QOR results was not found in Mizzi et al. (2016). We have investigated our results and conclude they are correct. There are two explanations for this discrepancy. First, as explained in Mizzi et al. (2016), we use the retrieval error covariance in retrieval space ($\boldsymbol{E_r}$) as the observation error covariance for assimilation purposes to account for other unquantified error sources, and $\boldsymbol{E_r} = (\boldsymbol{I} - \boldsymbol{A})\boldsymbol{E_a}$ where $\boldsymbol{E_a}$ is the retrieval *a priori* error covariance. If the singular vectors of $\boldsymbol{E_r}$ are equivalent to those of $\boldsymbol{A}$, we would get similar results from the CPSR and QOR experiments. However, $\boldsymbol{E_a}$ is specified in the retrieval algorithm as a covariance matrix, and generally there is no reason to suspect that it is such that the singular vectors of $\boldsymbol{E_r}$ are equivalent to those of $\boldsymbol{A}$ (for MOPITT CO they are not equivalent because their respective singular vectors are not orthogonal). Second, in the QOR experiment the diagonalization transform rotates the QOR equation so that the observation error cross-covariance contributions for each mode are included in their corresponding observation error variance. However, those modes are linearly dependent in the domain space defined by the rotated averaging kernel because the rotated averaging kernel is still singular. When those linearly dependent modes are assimilated, there is very little adjustment to the analysis. Consequently, the CPSR and QOR experiments yield similar results because: (i) the QOR experiment apportions the error and assimilates the linearly dependent modes (which have little or no impact), while (ii) the CPSR experiment apportions the error and does not assimilate the linearly dependent modes. Those results differ from the VMRR and L10VMRR experiments because the observation error variance used in the retrieval space experiments does not account for the error cross-covariance contributions, and the linearly independent portion of that error is different from that in the CPSR and QOR experiments.

In the lower panel of Fig. 1, we compare the chemical data assimilation experiments (the VMRR, L10VMRR, CPSR, and

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




QOR experiments) with the MET experiment and the MOZAIC ascent and descent sounding observations for Dallas, TX on 4 and 7 June 2008; Portland, OR on 3 June 2008; and Philadelphia, PA on 6 June 2008 (no other MOZAIC data were available for our study period and domain). We linearly interpolated the WRF-Chem forecasts to the MOZAIC observation locations and then composited the results. We did not plot the composited MOZAIC profile below 550 hPa because those data were more representative of the lower troposphere over urban areas than were our model grid and MOPITT super-observations. The MOZAIC comparison results are similar to those from the upper row of Fig. 1. They show that for the VMRR and L10VMRR experiments there is little improvement over the MET experiment when compared to MOZAIC. As mentioned earlier we suspect that occurs because the observation errors are too large. We ran similar experiments with reduced observation errors (not shown here) and found improved agreement. Results for the CPSR and QOR experiments show that: (i) for both experiments assimilation of phase space retrievals improves the 6-hr forecast skill in the mid- and lower troposphere compared to the MET experiment, (ii) the improvement is nearly the same for the CPSR and QOR experiments for the reasons discussed earlier, and (iii) there is little or no improvement near the surface. The upper tropospheric degradation in result (i) may be related to the bias of MOPITT CO retrievals in the upper troposphere discussed earlier (Deeter et al., 2013 and Martine-Alonso et al., 2014). Result (iii) is somewhat unexpected because MOPITT retrievals are documented to have sensitivity to CO in the upper and lower troposphere (Deeter et. al. 2007). We suspect result (iii) occurs because MOPITT's upper tropospheric sensitivities dominate its lower tropospheric sensitivities.

To test that hypothesis, we plot a histogram of the MOPITT degrees of freedom for signal (DOFS) for all terrestrial profiles in our domain during our study period in Fig. 2. The MOPITT DOFS is a measure of the amount of independent observed information in a retrieval profile. If a profile has independent information for the upper and lower troposphere, its DOFS must be ~2.0 or greater. The central histogram of Fig. 2 shows that the mean, median, and mode DOFS during this period are ~1.5 and that DOFS of ~2.0 or greater are relatively rare (< 5%). To gain a better understanding of the vertical structure of the MOPITT retrieval information content, we present a composite analysis for averaging kernel profiles in the neighborhood of different DOFS values in the lower row of Fig. 2 where panel (a) is the composite averaging kernels for all DOFS, (b) is for (0.9 < DOFS < 1.1, ~10%), (c) is for (1.4 < DOFS < 1.6, ~26%), and (d) is for (1.9 < DOFS < 2.1, ~4%). Those panels show



that the dominant sensitivity appears to be to the upper troposphere and that as the DOFS approaches 2.0 the sensitivity to the lower troposphere increases. That DOFS sensitivity distribution could explain the improvement drop off for the lower troposphere in Fig. 1 because retrievals with sensitivity to the lower troposphere are relatively rare. However, linear dependencies in the composite averaging kernels of Fig. 2 can mask the significance of the sensitivities in the lower troposphere in the more common DOFS categories.

To unmask those sensitivities, Fig. 3 presents a composite analysis of the different DOFS sensitivities based on the CPSR compression and diagonalization transforms. The upper row of Fig. 3 shows composite vertical profiles of the leading left singular vectors of the averaging kernel. Those singular vectors: (i) span the range of the averaging kernel (QOR space), (ii) are ranked such that the first singular vector explains the greatest amount of vertical variability in the QOR profile, the second singular vector explains the next greatest amount of variability, and so forth, and (iii) the sign is arbitrary, so we chose the sign that appears to have physical meaning, a scaling of −1.0 (that scaling has been included in Fig. 3). Note too that the number of leading singular vectors depends on the DOFS. For DOFs ≤ 1.0 there is at most one leading singular vector, for 1.0 > DOFS ≤ 2.0 there are at most two leading singular vectors, and for DOFS > 3 there are at most three leading singular vectors. The third vector is associated with a fractional DOFS and is an artifice of the retrieval process since the MOPITT instrument collects only two independent observations. In the analysis for Fig. 3, we retained three singular vectors for completeness, but it should be remembered the third vector (and sometimes the second vector) maps information to the null space of the averaging kernel. The middle row of Fig. 3 shows composite vertical profiles for the compressed averaging kernels. These profiles show the sensitivity of compressed QORs to the true atmospheric state. The bottom row shows the composite vertical profiles of the compressed and rotated averaging kernels (the profiles after the full CPSR transformation). These profiles show the sensitivity of CPSRs to the true atmospheric state.

Figure 3 shows some very interesting results. The upper row of Fig. 3 shows that for DOFS ≈ 1.0 the first leading singular vector has positive variability near the surface and negative variability in the upper troposphere (remember that the

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




second and third leading vectors map to the null space for DOFS ≈ 1.0). As the DOFS increases to 1.5, the first and second leading vectors have positive variability near the surface and weakly negative variability in the upper troposphere, and for DOFS of 2.0, the first leading vectors has positive variability throughout the troposphere while the second leading vectors has positive variability near the surface and negative variability in the upper troposphere.

Those DOFS-dependent variability profiles impact the sensitivities of the compressed averaging kernels in the second row of Fig. 2. These panels show that for all DOFS (the left most panel in column (a)) the first leading mode has its greatest sensitivity near the surface and the sensitivity decreases to a positive minimum near zero in the upper troposphere. Similarly the second leading mode has it greatest positive sensitivity near the surface but has strong negative sensitivity in the upper troposphere. The right three panels of the middle row in Fig. 3 (columns (b) –(d)) show the DOFS dependent changes of the compressed QORs. As seen with the singular vectors in the upper row, as the DOFS increases the sensitivity changes from weak positive sensitivity near the surface and strong negative sensitivity in the upper troposphere to strong positive sensitivity throughout the troposphere for the first leading mode and positive sensitivity near the surface and strong negative sensitivity in the upper troposphere for the second leading mode. Those results suggest that the MOPITT retrievals should be most sensitive to CO in the lower troposphere. However, an interesting thing happens when we account for the observation error covariance contributions. The lower row of Fig. 3 shows the compressed and rotated averaging kernel profiles. Here the negative scaling cancels each other because the SVD has been applied twice. The results show that when the error covariance is considered the significance of the leading modes become reversed due to scaling by the inverse square root of the compressed and rotated error variance. This does not mean that the third leading mode from first two rows of Fig. 3 becomes a dominant mode because it is still mapping to the null space. It means that the second leading mode from rows one and two effectively becomes the first leading mode in row three (similarly the first leading mode from rows one and two becomes the second leading mode in row three). The results in row three of Fig. 3 show that after removing the linear dependencies and accounting for the observation errors, the compressed and rotated averaging kernel has its greatest sensitivity in the upper troposphere for DOFS < 2.0 and weak sensitivity near the surface for DOFS ≈ 2.0. This analysis explains why



our comparison of the CPSR assimilation results with the MOZAIC *in situ* observations in the lower row of Fig. 1 showed assimilation impacts throughout the troposphere except near the surface. Other researchers who have assimilated MOPITT CO have not found this result because they do not adjust for the averaging kernel linear dependencies or for the observation error covariance.

Figures 4 and 5 show contour maps comparing the MET and CPSR experiments for 9 June 2008 18 UTC (Fig. 4) as well as the assimilated MOPITT CO retrievals and the corresponding independent IASI CO retrievals (Fig. 5). Examination of the forecast maps in the upper panel and the forecast difference map (CPSR experiment minus MET experiment) in the lower left panel of Fig. 4 shows that assimilation of MOPITT CO retrievals increased the CO concentrations over some areas (southern California, southern Baja, and northern Atlantic east of New England) and decreased the concentrations over broader areas (mid- to northeastern United States, southeastern United States, and southern Gulf of Mexico). Comparison of the MOPITT CO retrievals in the upper panels of Fig. 5 (the assimilated retrievals) with Fig. 4 shows that the analysis and forecast impacts are generally consistent with the observations. One area of difference is increased CO in the central United States, over Kansas and Nebraska (highlighted by the analysis increment map in the lower right panel of Fig. 4) that does not appear in the CPSR experiment forecast map (upper right panel of Fig. 4). We suspect that difference is due to: (i) the forecast advection of low CO into the regions of high CO in the analysis, and (ii) a low-bias in CO emissions used during the forecast that cannot support the high CO in the analysis. The assimilation of MOPITT retrievals increases CO in the analysis but during the forecast the advection of low CO or a low-bias in the CO emissions cannot support those CO increases so the forecast shows relatively low CO. Comparison of the analysis increments, the assimilated MOPITT CO retrievals, and the independent IASI CO retrievals (lower panels of Fig. 4) confirms that the assimilation of MOPITT retrievals generally improved the analysis and forecast agreement with the IASI retrievals compared to the MET experiment.

Figure 6 shows horizontal domain average vertical profiles for the MET and CPSR experiments compared against horizontal domain average profiles for MOPITT and IASI. The WRF-Chem profiles are plotted in retrieval space (after accounting for the averaging kernel and assimilation prior) and state space (without accounting for the averaging kernel and assimilation prior



– in physical space). Comparison of the model and MOPITT profiles (left two panels of Fig. 6) shows that the CPSR experiment generally draws the forecast and analysis retrieval space profiles closer to MOPITT than does the MET experiment. The same comparisons with the IASI profiles (right two panels of Fig. 6) shows a different result: (i) in the upper (pressure ($p$) < 300 hPa) and upper-mid (300 hPa < $p$ ≤ 500 hPa) the MET experiment draws the forecast and analysis profiles closer to IASI than does the CPSR experiment, and (ii) in the lower-mid (500 hPa ≤ $p$ < 600 hPa) and lower troposphere (600 hPa) the CPSR experiment draws the profiles closer to IASI. Those results highlight the previously discussed problem with assimilating potentially biased MOPITT CO retrievals in the upper troposphere (above 300 hPa). To address that problem, we propose to discard the biased retrievals and assimilate the unbiased truncated retrieval profiles with the extended CPSR method described in Section IV.

In summary, this section shows that assimilation of MOPITT CO retrievals improves analysis fit and forecast skill when compared to MOPITT as well as when compared to the independent (not assimilated) IASI and MOZAIC observations. It shows that: (i) assimilation of phase space retrievals (CPSR and QOR) improves analysis fit and forecast skill when compared to assimilation of raw retrievals (VMRR and L10VMRR) because the phase space transformation reduces the phase space observation errors, and (ii) the CPSR and QOR experiments yield similar results because they account for the observation error cross-covariance contribution in the same way (the diagonalization transform) and because the linearly dependent portion of the transformed retrievals do not contribute to the analysis increment (explicitly with CPSRs and implicitly through the assimilation algorithm for compressed QORs). It also shows that CPSR and QOR experiments did not improve the skill in the lower troposphere near the surface because: (i) MOPITT CO profiles with sufficient DOFS to resolve the lower tropospheric CO signal are relatively rare (for this domain and study period), and (ii) an analysis of the impact of the CPSR compression and diagonalization transforms shows that the upper tropospheric CO signal dominates the MOPITT CO sensitivities. Finally, this section shows that in the upper troposphere assimilation of potentially biased MOPITT observations introduced analysis and forecast error relative to the IASI observations.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




**5.2 Assimilation of Truncated Retrieval Profiles**

In this section, we test two methods for assimilating truncated retrieval profiles: (i) assimilate L10VMRR retrievals after discarding the biased retrievals (the L10VMRR-RJ3 experiment) and (ii) assimilate CPSRs extended to truncated retrieval profiles as described in Section V.C (the CPSR-RJ3 experiment). The L10VMRR-RJ3 experiment uses a retrieval space assimilation algorithm that provides a control against which we compare the CPSR-RJ3 experiment. The L10VMRR-RJ3 experiment is included only for comparison purposes. If the L10VMRR-RJ3 and CPSR-RJ3 experiments give similar results then the CPSR-RJ3 approach is preferred because it is computationally less expensive, removes linear dependencies, and accounts for the observation error covariance.

Figure 7 shows vertical profiles for the L10VMRR-RJ3 and CPSR-RJ3 experiments with results from the full retrieval profile assimilation experiments included for reference. For these experiments, we are assuming that in the upper troposphere the MOPITT CO retrievals are positively biased (as discussed earlier) and that the IASI CO retrieval more accurately reflect the true concentrations. Comparisons against the assimilated MOPITT observations in the upper panels show that discarding the biased observations had the desired effect – in the upper troposphere the analysis profile is drawn closer to that of the MET experiment for the L10VMRR-RJ3 experiment than for the L10VMRR experiment. Similar results are seen for comparison of the CPSR-RJ3 and CPSR experiments in the last two panels of the upper row. Unexpectedly for both experiments not assimilating observations in the upper troposphere there was a negative impact in the lower troposphere for the upper row of Fig. 7. Comparison with IASI CO retrievals in the lower row of Fig. 7 shows similar results: (i) the L10VMRR-RJ3 and CPSR-RJ3 retrieval space profiles are drawn closer to the IASI profile than the L10VMRR and CPSR profiles in the upper troposphere, and (ii) the agreement of the CPSR-RJ3 profiles with IASI is degraded in the middle and lower troposphere. We investigate the cause of those lower tropospheric results later in this section, but next we discuss the horizontal impacts of the truncated retrieval assimilation experiments.

Figures 8 and 9 show contour maps from the CPSR and CPSR-RJ3 experiments. Figure 8 shows the near surface impacts of not assimilating the biased retrievals. The CO 6-hr forecast contour maps in the upper row of Fig. 8 suggest that the lower



tropospheric impacts of not assimilating CO retrievals above 300 hPa are small. However, the forecast difference maps in the lower row highlight the changes. The CPSR-RJ3 experiment has small large-scale decreases in CO over the oceans and eastern United States. Also, the magnitude of positive forecast differences over CO hot spots over Southern California, Baja, and the northeastern United State has decreased. Those lower tropospheric difference maps highlight the temporally integrated impacts of these experiments. The upper tropospheric impacts are shown in Fig. 9 where we see similar results except that the CO hot spot reductions are over the southeastern United States. That was the goal of CPSR-RJ3 experiment (to remove the impact of assimilation potentially biased MOPITT CO retrievals in the upper troposphere).

A verification analysis for the L10VMRR-RJ3 and CPSR-RJ3 experiments is presented in the upper panel of Fig. 1. The L10VMRR-RJ3 and CPSR-RJ3 experiments have degraded forecast skill compared to the full profile assimilation experiments (the VMRR, L10VMRR, CPSR, and QOR experiments), but the CPSR-RJ3 experiment has improved skill compared to the L10VMRR-RJ3 experiment. The CPSR-RJ3 experiment skill improvement (compared to the L10VMRR-RJ3 experiment) is likely due to the observation error covariance reduction resulting from the CPSR transforms discussed earlier.

In summary, not assimilating the biased observations had positive impacts in the upper troposphere and negative impacts in the lower troposphere. We suspect the middle to lower tropospheric result occurs for two reasons. Discarding selected retrievals: (i) reduces the total information content of the assimilated retrievals; and (ii) reduces the sensitivity of the transformed averaging kernel. Those reductions combine to reduce the ensemble state variable correlations. To test that explanation (i) we compare the trace of the raw averaging kernel for the CPSR experiment with that for the CPSR-RJ3 experiment. The results are shown in the first two rows of Table 2 where “Full Profile” is from the CPSR experiment, and “Reject Top Three” is from the CPSR-RJ3 experiment. Comparison of those results shows a 25% reduction in the trace indicating that the total information content of the assimilated retrievals for the CPSR-RJ3 experiment is less than that for the CPSR experiment. For comparison purposes, Table 2 also shows trace reductions from not assimilating retrievals in the mid-troposphere (23%) and lower troposphere (9%). Those results suggest that most of the information in the MOPITT CO retrievals is from the upper troposphere, the second greatest amount is from the middle troposphere, and the smallest amount



is from the lower troposphere. To test explanation (ii) we plot the rows of the compressed and fully transformed averaging kernels in Fig. 10 where column (a) is for the CPSR experiment and column (b) is for the CPSR-RJ3 experiment. Figure 10 is similar to the last two rows of Fig. 3. Recall that the first row represents the sensitivity of changes in the compressed QORs to changes in the true CO concentrations, and the second row represents the sensitivity of changes in the CPSRs to changes in the true CO concentrations. Comparison of columns (a) and (b) shows that the CPSR-RJ3 experiment leading mode sensitivities are significantly reduced when compared to the CPSR experiment. The state variable correlations are proportional to the sensitivities so the reduced correlations result in analysis increment reductions. For comparison purposes columns (c) and (d) of Fig. 10 show results from experiments that discard retrievals in the middle and lower troposphere. Those profiles in combination with Table 2 show that most of the information and sensitivity is associated with the upper and mid-tropospheric retrievals. Discarding upper tropospheric retrievals alters the sensitivity magnitudes while discarding middle tropospheric retrievals alters the magnitudes and vertical structure. One interesting result is that most of the sensitivity loss in column (c) - the "Reject Middle Three" experiment - appears to be associated with the CPSR diagonalization transform. That suggests that the sensitivity loss is dependent on specification of the retrieval *a priori* error covariance.

Those changes occur because as different rows of the averaging kernel are discarded: (i) the amount of observed information in the modified averaging kernel changes, and (ii) the vertical structure of the bases for the range and domain of the modified averaging kernel changes. The impact of changes in the information content in (i) were discussed earlier. The impact of changes to the bases in (ii) has important consequences. The non-zero left singular vectors of the modified averaging kernel span the range of the modified averaging kernel but their dimension and vertical structure change when retrievals are discarded. That means the phase space observations change because the basis vectors used in the compression transform are different, and their sensitivity to the truncated retrieval profile vector is different. Similarly, the non-zero right singular vectors of the modified averaging kernel span the domain of the modified averaging kernel but their vertical structure changes when retrievals are discarded. Those changes occur solely because the information content of the modified averaging kernel is reduced (since the dimension of its domain – the space where the true CO profiles reside – is unchanged). Those changes are significant because they alter the elements (or levels) of the true profile to which the modified averaging kernel is sensitive. To summarize



not assimilating elements of the full retrieval profile alters the levels of the retrieval profile to which the phase space observations are sensitive. Discarding those elements also alters the levels of the true CO profile to which the modified averaging kernel is sensitive. Those sensitivity changes occur regardless of whether the assimilation is done in phase space as in the CPSR–RJ3 experiment or in retrieval space as in L10VMRR–RJ3 experiment. Consequently, results from the L10VMRR-RJ3 and CPSR-RJ3 experiments are similar.

This section shows that CPSRs can be extended to the assimilation of truncated retrieval profiles but that discarding upper tropospheric observations for MOPITT significantly reduces the total information content of the assimilated observations and the vertical sensitivities of the transformed averaging kernel profiles. Those reductions likely translate to reductions in the state variable correlations and commensurate reductions in the analysis increments. We are studying modification of the CSPR extension to truncated retrieval profiles to address the non-local impacts.

## 6 Summary and Conclusions

This paper had two goals: (i) compare the results of assimilating CPSRs with independent observations (we used MOZAIC *in situ* observations and IASI CO retrievals as the independent observations), and (ii) extend CPSRs to the assimilation of truncated retrieval profiles. The comparison against independent observations showed that: (i) assimilation of raw retrievals (VMRRs and L10VMRRs) had little impact on the analysis fit and forecast skill due to the magnitude of the observation errors, and (ii) the assimilation of phase space retrievals (CPSRs and QORs) improved both fit and skill. Conceptually, we expect the assimilation of raw retrievals and phase space retrievals to yield similar results. However, phase space transformation of the observation error covariance truncated the observation errors so that the CPSR and QOR experiments produced closer agreement with the assimilated and independent observations.

Comparison with the MOZAIC *in situ* CO observations for the CPSR and QOR experiments showed improved agreement with the MOZAIC observations in the upper and middle troposphere when compared to the MET experiment. There was little improvement near the surface. The lack of a near surface improvement was unexpected but our DOFS analysis in the



discussion of Figs. 2 and 3 showed that assimilation of MOPITT CO CPSRs and QORs does impact the near surface CO because after accounting for the MOPITT CO retrieval observation error covariance, the compressed and rotated averaging kernels have very little sensitivity to the CO near the surface.

Comparison with IASI CO retrievals also showed degraded skill in the upper troposphere due to assimilation of biased retrievals illustrating the need to extend CPSRs to assimilation of truncated retrieval profiles. Section 4.3 explained the extension, and Section 5.2 compared the L10VMRR-RJ3 (assimilation of raw truncated retrieval profiles) and CPSR-RJ3 (assimilation of phase space truncated retrieval profiles) experiments where we did not assimilate MOPITT CO retrievals above 300 hPa. That comparison showed that the L10VMRR-RJ3 and CPSR-RJ3 experiments produced similar results confirming the extension of CPSRs to truncated retrieval profiles.

Comparison of the L10VMRR-RJ3 and CPSR-RJ3 experiments highlighted an important characteristic of assimilating mixing ratio retrieval profiles. Not assimilating observations can significantly alter the: (i) information content of the assimilated observations; and (ii) the amplitude of the averaging kernel sensitivities. Those reductions combine to reduce the state variable correlations and the corresponding analysis increments. We are researching modifications of the CPSR extension to truncated retrieval profiles to address the remote impacts from not assimilating retrievals from selected levels.

**Code and Data Availability**

The current versions of the WRF-Chem, WRF, WRFVAR, and WPS codes are available from the WRF download site at http://www2.mmm.ucar.edu/wrf/users/download/get_sources.html. The current version of the DART code is at available at https://www.image.ucar.edu/DAReS/DART/DART2_Starting.php#download, and the current version of WRF-Chem/DART branch is available at https://www.image.ucar.edu/DAReS/DART/DART2_Starting.php#download. The WRF-Chem/DART branch is same as the DART code except for inclusion of the WRF-Chem/DART system. There is no need to down load both codes. Presently, there is no users guide available for WRF-Chem/DART. However, the authors have prepared a slide presentation that describes much of the chemical data assimilation script function, variables, and organization. Interested



readers should contact the authors for a copy of that presentation and assistance with using WRF-Chem/DART. The large scale models forecast and observational data used to run the ensemble forecast/data assimilation cycling experiments descried in the paper are generally available from the respective data distribution sites. That data set has not been posted to a public site due to its size but is available from the authors upon request.

**Acknowledgements**

NCAR is sponsored by the National Science Foundation (NSF). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of NSF. This research is also sponsored by National Aeronautics and Space Administration (NASA) grant NNX11AI51G. We gratefully acknowledge Chris Snyder and Avellino Arellano for discussions that led to a better understanding of the transformed averaging kernels associated with truncated retrieval profiles. We also acknowledge Louisa Emmons and Benjamin Gaubert for their thoughtful reviews and helpful suggestions both which improved the quality of this manuscript. We also acknowledge the use of data products from MOPITT, IASI, and MOZAIC/IAGOS programs.

**References**

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Mon. Wea. Rev., 129, 2884-2903, https://doi.org/10.1175/1520-0493(2001129<2884:AEAKFF>2..CO:2, 2001.

Anderson, J. L.: A local least squares framework for ensemble filtering, Mon. Wea. Rev., 131, 634-642, https://doi.org/10.1175/1520-0493(2003)<0634:ALLSFF>2.0.CO:2, 2003.

Anderson, J. L.: Spatially and temporally varying adaptive covariance inflation for ensemble filters, Tellus, 61, 72-83, https://doi.org/10.1111/j.1600-0870.2008.00361.x, 2008.

Anderson, J. L., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., and Arellano, A.: The Data Assimilation Research Testbed: A community facility, Bull. Amer. Meteor. Soc., 90, 1283-1296, https://doi.org/10.1175/2009BAMS2618.1, 2009.

Barker, D., X.-Y. Huang, Z. Liu, T. Auligné, X. Zhang, S. Rugg, R. Ajjaji, A. Bourgeois, J. Bray, Y. Chen, M. Demirtas, Y.-R. Guo, T. Henderson, W. Huang, H.-C. Lin, J. Michalakes, S. Rizvi, and X. Zhang: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. Bull. Amer. Meteor. Soc., 93, 831–843, https://doi.org/10.1175/BAMS=D-11-00167.1, 2012.

Barre, J., B. Gaubert, A. F. Arellano, H. M. Worden, D. P. Edwards, M. N. Deeter, J. L. Anderson, K. Raeder, N. Collins, S. Tilmes, G. Francis, C. Clerbaux, L. K. Emmons, G. G. Pfister, P.-F. Coheur, and D. Hurtmans: Assessing the impacts of assimilating IASI and MOPITT CO retrievals using CESM-CAM-chem and DART, J. Geophys. Res. Atmos., 120, 10501–



10529, https://doi.org/10.1002/2015JD023467, 2015.

Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal infrared IASI/MetOp sounder, Atmos. Chem. Phys., 9, 6041-6054, https://doi.org/10.5194/acp-9-6041-2009, 2009.

Deeter, M. N., D. P. Edwards, and J. C. Gille, Retrievals of carbon monoxide profiles from MOPITT observations using lognormal a priori statistics, J. Geophys. Res., 112, D11311, https://doi.org/10.1029/2006JD007999, 2007.

Deeter, M. N., D. P. Edwards, J. C. Gille, and J. R. Drummond, Sensitivity of MOPITT observations to carbon monoxide in the lower troposphere, J. Geophys. Res., 112, D24306, https://doi.org/10.1029/2007JD008929, 2007.

Deeter, M. N., S. Martinez-Alonso, D. P. Edwards, L. K. Emmons, J. C. Gille, H. M. Worden, J. V. Pittman, B. C. Daube, and S. C. Wofsy: 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, https://doi.org/10.1002/jgrd.50272, 2013.

Eskes, H. J., and K. F. Boersma: Averaging kernels of DOAS total-column satellite retrievals, Atmos. Chem. Phys., 3, 1285–1291, https://doi.org/10.5194/acp-3-1285-2003, 2003.

Miyazaki, K., H. J. Eskes, and K. Sudo: A tropospheric chemistry reanalysis for the years 2005–2012 based on an assimilation of OMI, MLS, TES, and MOPITT satellite data, Atmos. Chem. Phys., 15, 8315–8348, https://doi.org/10.5194/acp-15-8315-2015, 2015.

Marenco, A., V. Thouret, P. Nedelec, H. Smit, M. Helten, D. Kley, F. Karcher, P. Simon, K. Law, J. Pyle, G. Poschmann, R. Von Wrede C. Hume, and T. Cook: Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, an overview, J. Geophys. Res., 103(D19), 25631–25642, https://doi.org/10.1029/98jD00977, 1998.

Martinez-Alonso, S., M. N. Deeter, H. M. Worden, J. C. Gille, L. K. Emmons, L. L. Pan, M. Park, G. L. Manney, P. F. Bernath, C. D. Boone, K. A. Walker, F. Kolonjari, S. C. Wofsy, J. Pittman, and B. C. Daube, Comparison of upper tropospheric carbon monoxide from MOPITT, ACD-FTC, and HIPPO-QCLS, J. Geophys. Res. Atmos., 119, 14164-14164, https://doi.org/10.1002/2014JD022397.

Miyazaki, K., H. J. Eskes, and K. Sudo: Global $NO_x$ emission estimates derived from an assimilation of OMI tropospheric NO2 columns, Atmos. Chem. Phys., 12, 2263–2288, https://doi.org/10.5194/acp-12-2263-2012, 2012a.

Miyazaki, K., H. J. Eskes, K. Sudo, M. Takigawa, M. van Weele, and K. F. Boersma, Simultaneous assimilation of satellite $NO_2$, $O_3$, CO, and $HNO_3$ data for the analysis of tropospheric chemical composition and emissions, Atmos. Chem. Phys., 12, 9545-9579, https://doi.org/10.5194/acp-12-9545-2012, 2012b.

Mizzi, A. P., A. F. Arellano, D. P. Edwards, J. L. Anderson, and G. G. Pfister: Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: a regional chemical transport/ensemble Kalman filter data assimilation system, Geosci. Model Dev., 9, 965-978, https://doi.org/10.5194/gmd-9-965-2016, 2016.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018


Geoscientific Model Development Discussions

1

2

| Experiment | Assimilate meteorology observations | Assimilate MOPITT CO raw retrievals | Assimilate MOPITT CO CPSRs | Assimilate MOPITT CO QORs | Assimilate retrieval full profiles | Assimilate truncated retrieval profiles |
|---|---|---|---|---|---|---|
| **MET** | X | - | - | - | - | - |
| **VMRR** | X | X | - | - | X | - |
| **L10VMRR** | X | X | - | - | X | - |
| **CPSR** | X | - | X | - | X | - |
| **QOR** | X | - | - | X | X | - |
| **L10VMRR-RJ3** | X | X | - | - | - | X |
| **CPSR-RJ3** | X | - | X | - | - | X |

3

4 Table 1. Summary of the WRF-Chem/DART Forecast/Data Assimilation Experiments. An X indicates that an experiment
5 (denoted by different rows) assimilated a particular type of observation (denoted by different column). An – indicates that it
6 did not assimilate that type of observation.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018


Geoscientific Model Development Discussions

| | CompAK 1 | CompAK 2 | CompAK 3 | Trace |
|---|---|---|---|---|
| **Full Profile** | .6890 | .3058 | .0161 | 1.441 |
| **Reject Top Three** | .7530 | .2439 | .0096 | 1.076 |
| **Reject Middle Three** | .6796 | .3153 | .0157 | 1.106 |
| **Reject Bottom Three** | .7274 | .2685 | .0127 | 1.311 |

Table 2. Average total and fractional information content for each mode of the averaging kernel for the entire study period. CompAK 1 denotes the average fractional information in mode 1, CompAK 2 is for mode 2, and so forth. Trace denotes the total information content. "Full Profile" means all retrievals were assimilated (i.e., none were discarded). "Reject Top Three" means that retrievals at pressure levels < 300 hPa were discarded. "Reject Middle Three" means that retrievals between 300 hPa and 600 hPa were discarded. "Reject Bottom Three" means that retrievals below 700 hPa were discarded.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018

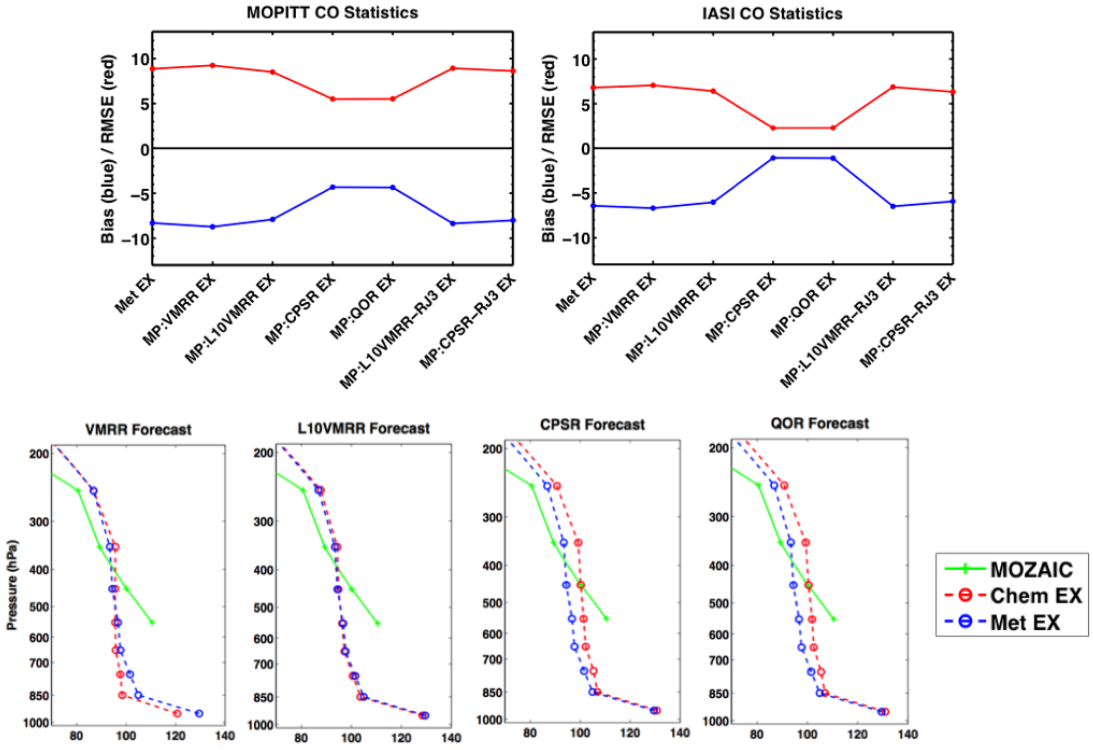




Figure 1. Forecast (assimilation prior) verification statistics for all experiments in MOPITT retrieval space on the upper left and IASI retrieval space on the upper right. For those panels the red curve is root mean square error (RMSE), and the blue curve is bias (model – observation). The experiments are described in the text and summarized in Table 1. The lower panel shows comparisons against the IAGOS/MOZAIC in situ CO profiles in ppb composited for 1 June 2008 00 UTC to 9 June 2008 18 UTC in state space. In the lower legend, the Chem EX refers to the VMRR, L10VMRR, CPSR, or QOR experiments depending on the panel.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018



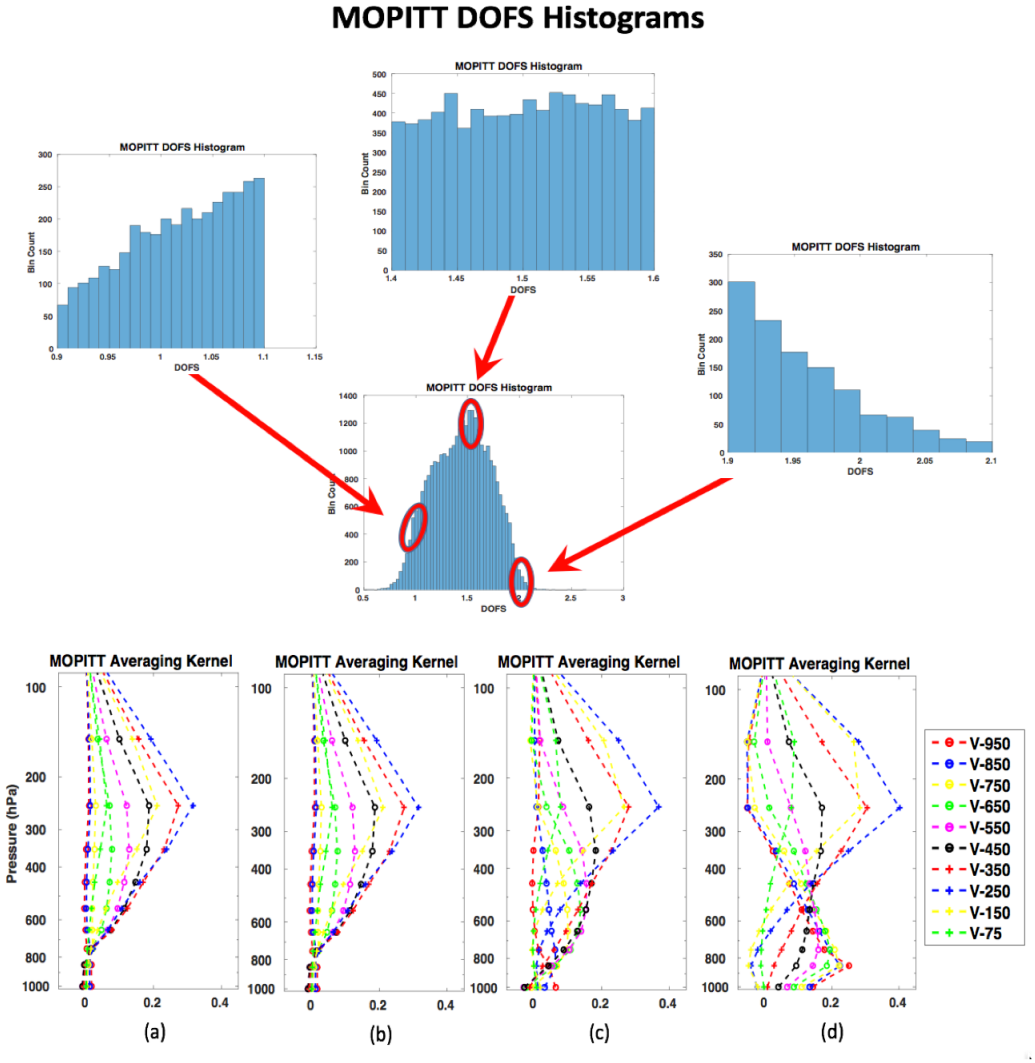
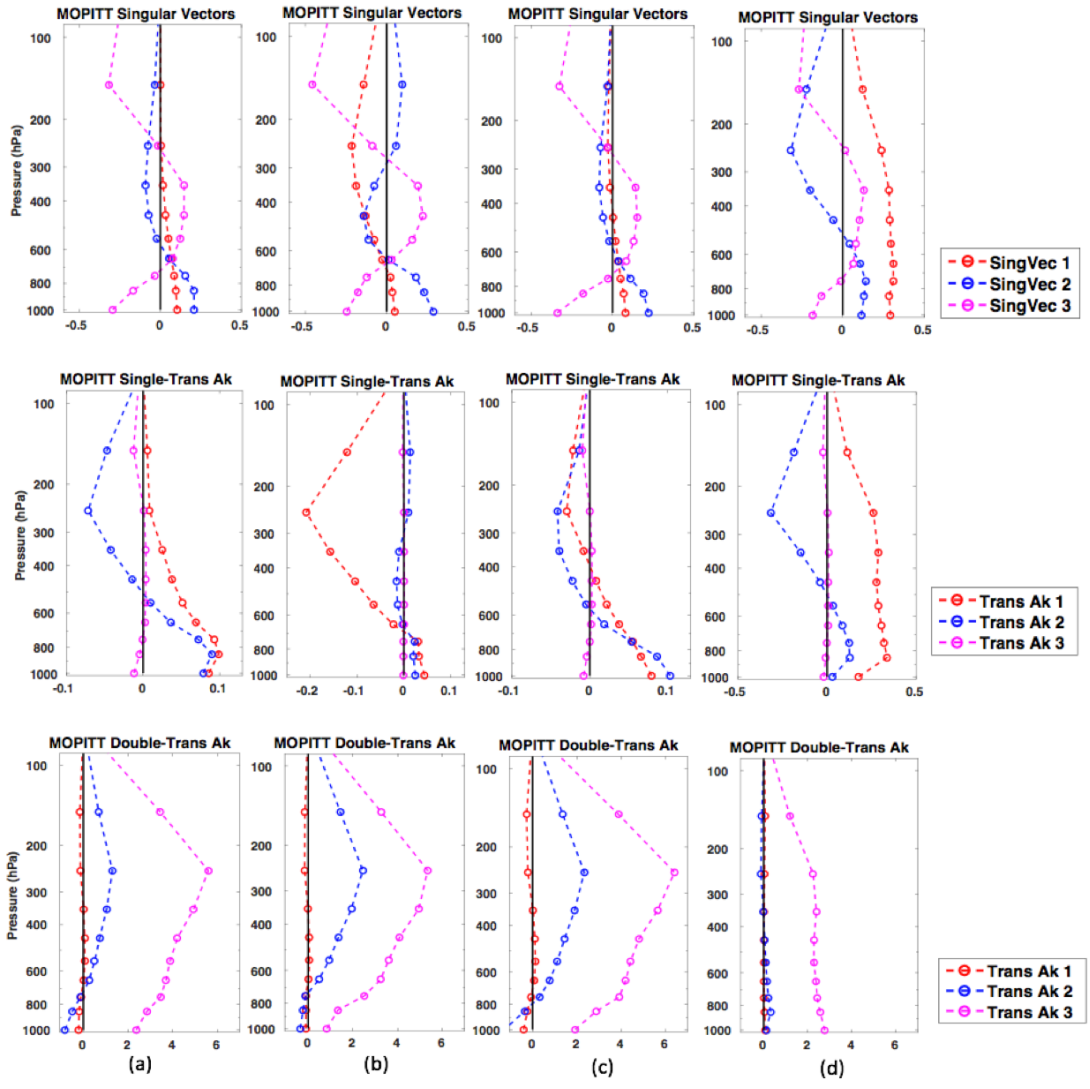




Figure 2. Histogram of MOPITT CO "degrees of freedom of signal" (DOFS) with blow-up histograms for selected DOFS ranges in the upper panels. The lower panels show composite MOPITT CO averaging kernel profiles for: (a) all DOFS, (b) (0.9 ≤ DOFS ≤ 1.1), (c) (1.4 ≤ DOFS ≤ 1.6), and (d) (1.9 ≤ DOFS ≤ 2.1). The averaging kernel identifiers are V-xxx where xxx is the approximate pressure level mid-point in hPa for the associated averaging kernel profile.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018


Geoscientific Model Development Discussions

1



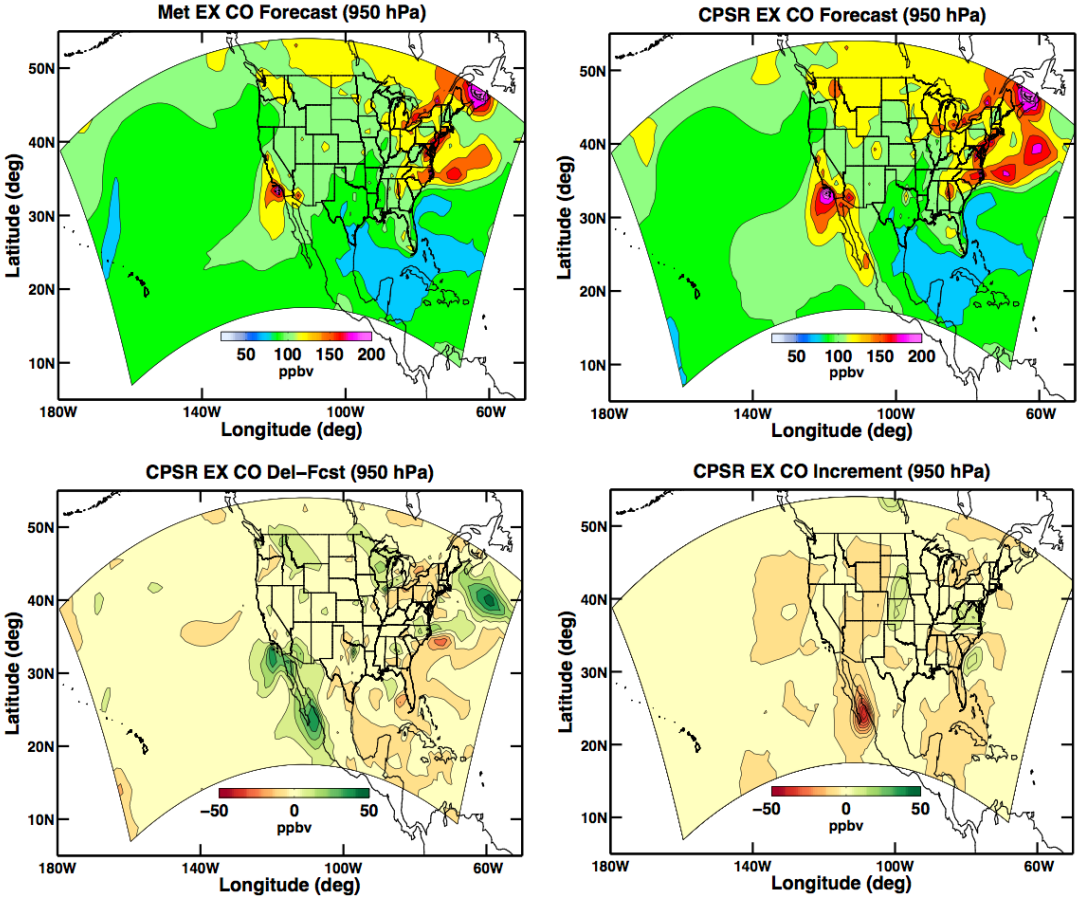

2

3 Figure 3. Composite vertical profiles for the: (i) leading left singular vectors of the MOPITT CO averaging kernels in the upper
4 row, (ii) compressed averaging kernels in the middle row, and (iii) rotated and compressed averaging kernels in the lower row.
5 The DOFS ranges are the same as defined for Fig. 2. For the profile labels "SingVec x" refers to ranked singular vectors where
6 x = 1 is the first leading singular vectors, x = 2 is the second leading singular vector, and so forth. "Trans Ak x" refers to the
7 compressed or rotated and compressed averaging kernel profile associated with the QOR and CPSR mode x respectively.
8

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




1



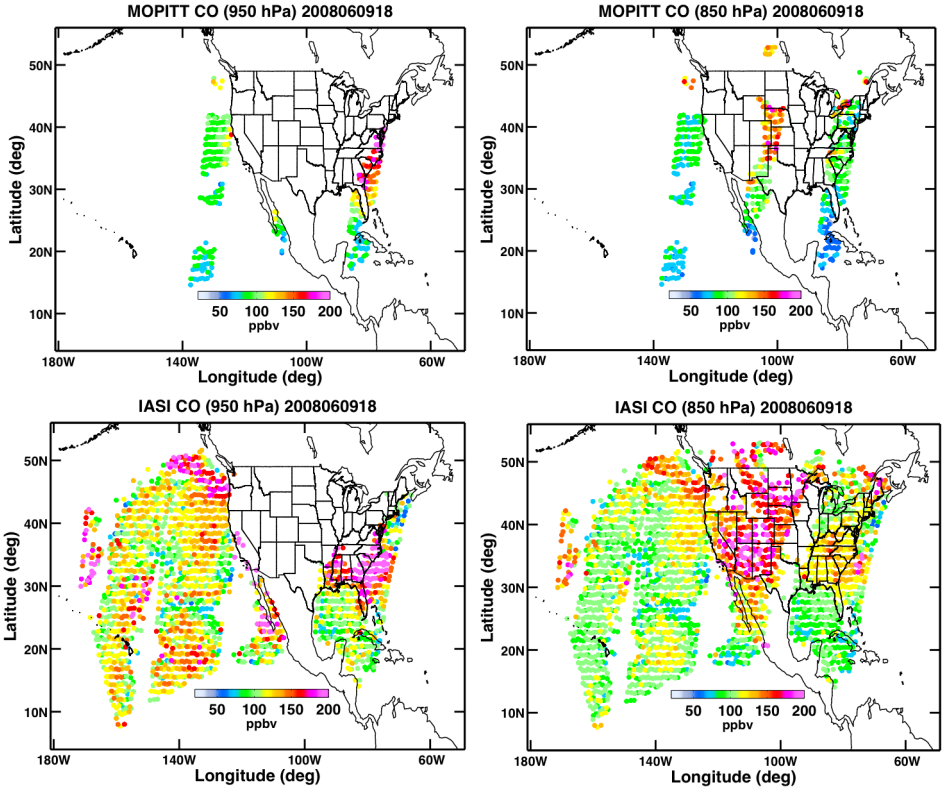

2

3 Figure 4. Shaded contours of CO in ppb for the MET and CPSR experiment 6-hr forecasts valid at this cycle time in the left
4 and right upper panels respectively. The lower row presents the difference between the CPSR and MET forecasts (the CPSR
5 experiment 6-hr forecast minus the MET experiment 6-hr forecast) in the left panel and the assimilation increment for analysis
6 at this cycle time in the right panel. All figures are for ~950 hPa and the 9 June 2008 18 UTC cycle. The curved rectangle
7 represents the WRF-Chem domain.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018


Geoscientific Model Development Discussions
Open Access
EGU

1




2

3 Figure 5. The assimilated MOPITT CO retrievals in the upper panels and the corresponding IASI CO retrievals (not
4 assimilated) in the lower panels. The left figures are for ~950 hPa, and the right figures are for ~850 hPa. All figures are for
5 the 9 June 2008 18 UTC cycle. The retrievals are in ppb.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018



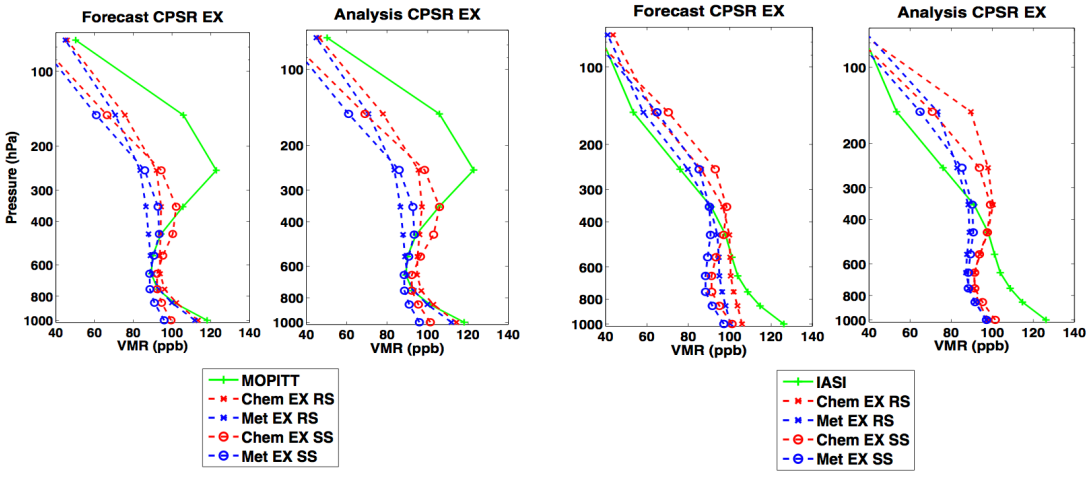




Figure 6. Vertical profiles of the time/horizontal domain average CO in ppb from the CPSR and MET experiments for 9 June 2008 18 UTC. The "RS" denotes results in retrieval space and the "SS" denotes results in state space. "Forecast" is the assimilation prior, and "Analysis" is the assimilation posterior. The left two panels compare the forecast/assimilation results against MOPITT CO retrievals (assimilated), and the right two panels compare those results against IASI CO retrievals (not assimilated). In the legends, Chem EX refers to the CPSR experiment.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018

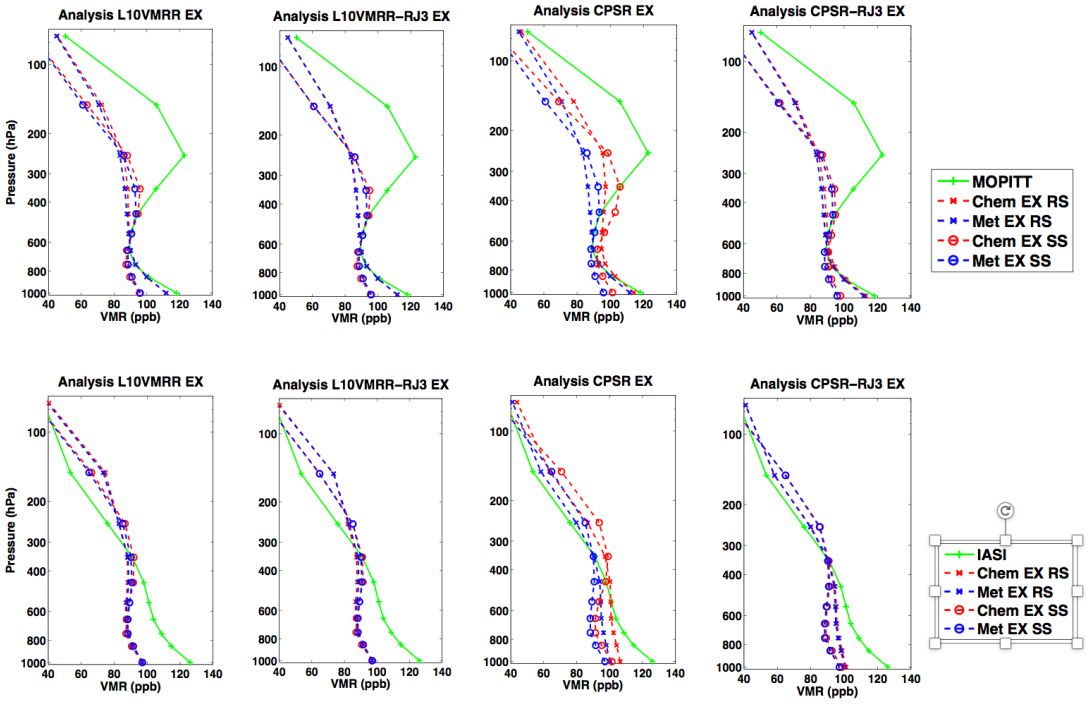
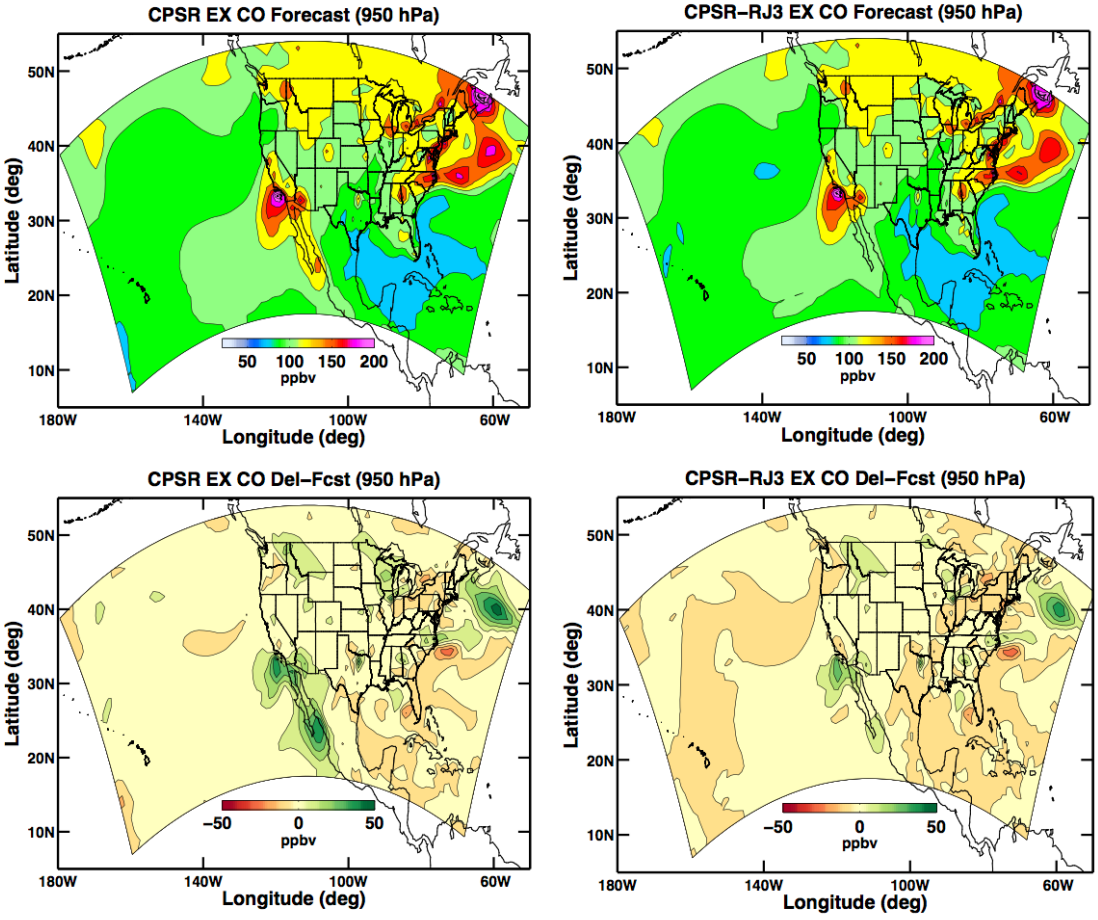




Figure 7. Same as Fig. 6 except this figure compares the L10VMRR, L10VMRR-RJ3, CPSR, and CPSR-RJ3 experiments. The upper panels compare the forecast/assimilation results against MOPITT CO retrievals (assimilated) and the lower panels compare those results against IASI CO retrievals (not assimilated). In the legends, Chem EX is a placeholder for the L10VMRR-RJ3, L10VMRR, CPSR, and CPSR-RJ3 experiments depending on the panel.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018




1




2

3 Figure 8. Shaded contours of CO in ppb for the CPSR and CPSR-RJ3 experiment assimilation priors in the left and right upper
4 panels respectively and for the CPSR and MET experiment difference (the CPSR – MET experiment) and the CPSR-RJ3 and
5 MET experiment difference (the CPSR-RJ3 – MET experiment) assimilation priors in the left and right lower panels
6 respectively. The CPSR experiments maps in this figure are the same as in Fig. 4 and included for reference. All figures are
7 for ~950 hPA at 9 June 2008 18 UTC.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018

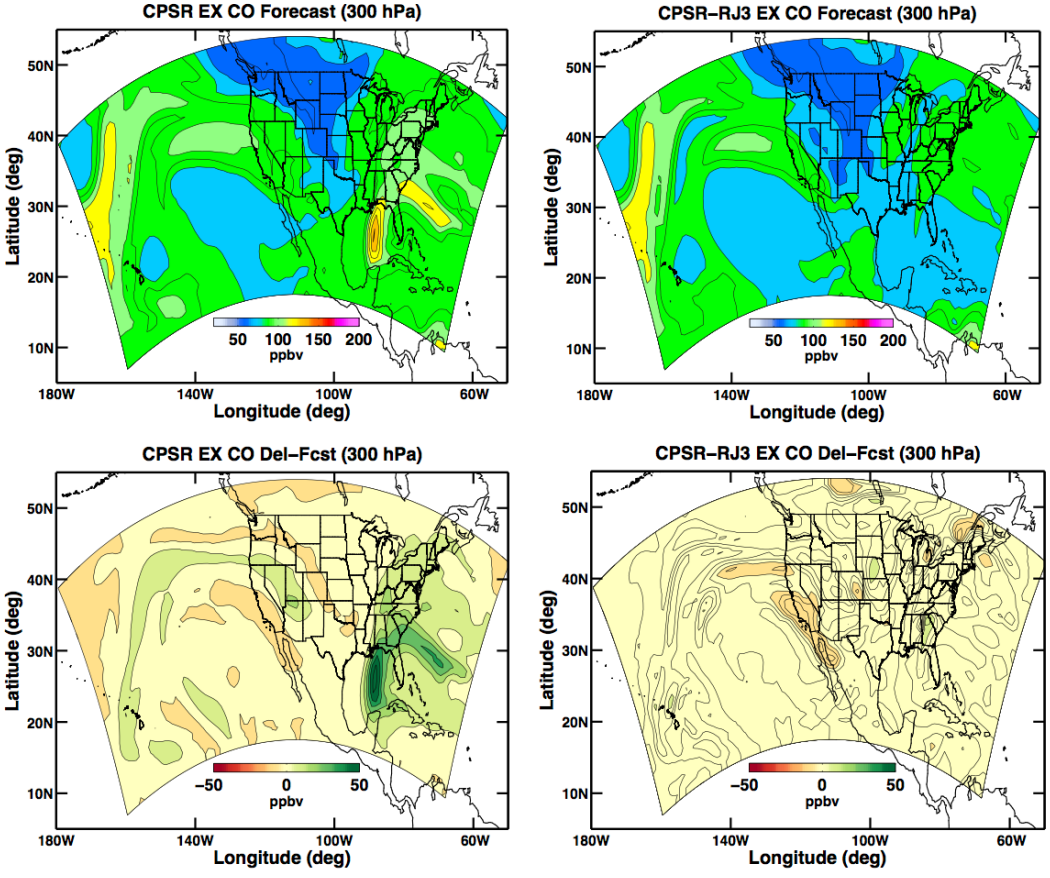
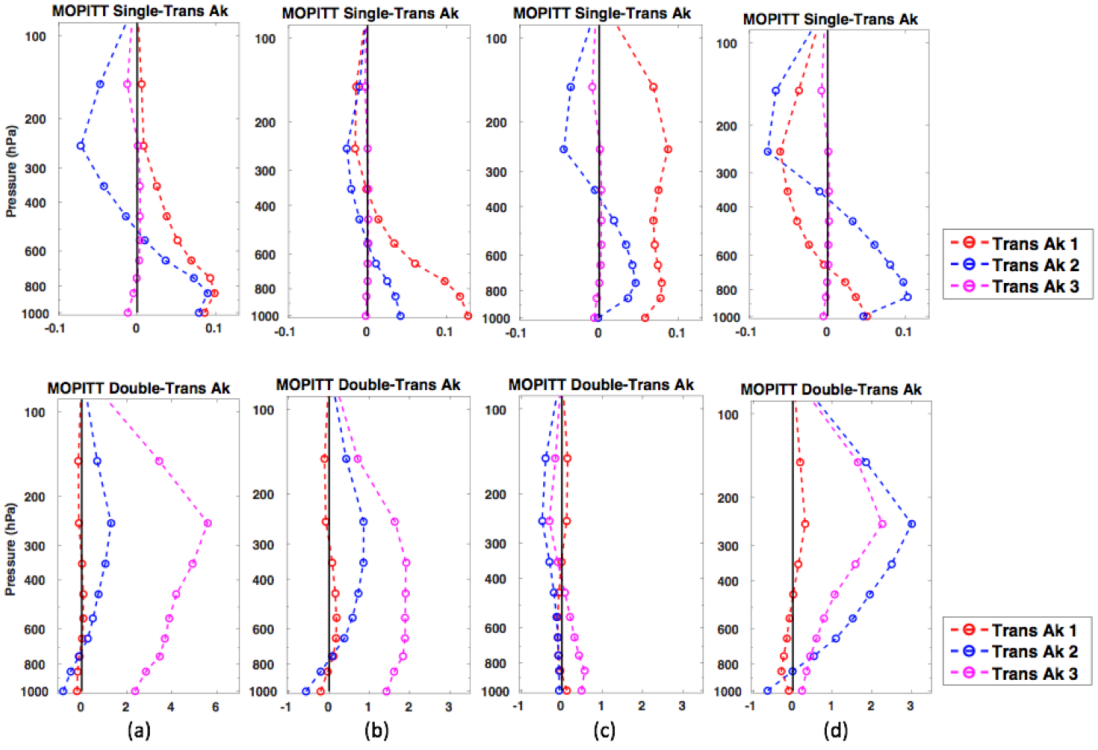




Figure 9. Same as Fig. 8 except for ~300 hPa.

Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-19
Manuscript under review for journal Geosci. Model Dev.
Discussion started: 2 March 2018


Geoscientific Model Development Discussions
Open Access
EGU

1




2

3 Figure 10. Same as the lower two rows of Fig. 3 except this figure is for the retrieval discard experiments. Column (a) is for
4 the full retrieval profile assimilation experiment and is the same as column (a) in Fig. 3. Column (b) is for the "Reject Top
5 Three" experiment in Table 2. Column (c) is for the "Reject Middle Three" experiment. Column (d) is for the "Reject Bottom
6 Three" experiment. Notice that the range of the abscissa is reduced from column (a) to columns (b) – (d).
7