1. INTRODUCTION

Anthropogenic activities have significant impacts on biogeochemical cycles. To better understand the changes occurring in the atmosphere and to clearly distinguish natural from anthropogenic influences, it is extremely important to monitor the temporal and spatial distributions of gases and identify their sources and sinks. Carbon monoxide (CO) is one of the key tropospheric trace species. With roughly a 2 month lifetime, and with diverse sources, both natural and anthropogenic (CH$_4$ oxidation, NMHC oxidation, biomass burning, fossil fuel burning etc.), CO can serve as a useful tracer of atmospheric transport. CO also affects the concentration of the hydroxyl radical (OH), which is involved in much of the chemistry in the troposphere. However, OH has an extremely short lifetime and is difficult to measure. Therefore, the ability to continuously monitor CO from space should provide an important window on tropospheric chemistry.

To measure the spatial and temporal variation of the CO profile and total column amount in the troposphere, the Measurements of Pollution In The Troposphere (MOPITT) instrument was launched in 1999 on board the NASA Terra satellite. MOPITT is an eight-channel gas correlation radiometer; each channel generates an average (A) signal and a difference (D) signal (Drummond, 1992). The A signals are sensitive to the background emissions, while the D signals are sensitive to the target gas vertical distribution. MOPITT measurements can resolve the vertical distribution of tropospheric CO in 3-4 layers with a 22x22 km horizontal resolution.

The MOPITT operational retrieval is based on the Maximum Likelihood (ML) method (Pan et al., 1998; Rodgers, 2000). The ML retrieval algorithm seeks the statistically most likely CO profile consistent with both the observed radiances and a priori information. The role of the a priori mean profile and covariance matrix is to constrain the retrieved profile to fall within the range of physically realistic solutions (based on variability statistics of a selected set of observed in-situ profiles). The relative weighting of the a priori information and information from the measured radiances in the retrieved profile is controlled directly by the a priori covariance matrix and measurement error covariance matrix.

The operational MOPITT CO retrieval currently uses a fixed global a priori. This approach was adopted initially to ensure that observed geographical variations in the retrieval results are due to information in the measured radiances rather than features of the a priori. Studies (Hansen et al., 1995; Pan, et al., 1998) have confirmed that the choice of a priori affects the accuracy of the retrievals. There is some question as to whether the use of a global a priori is an adequate representation of the seasonal variation in the CO profiles for diverse locations. Furthermore, only limited regional surface observations and aircraft measurements from field experiments are available for the construction of the MOPITT global a priori. The purpose of this paper is to quantify the sensitivity of the retrieval to the use of a fixed global a priori in the ML method. We conduct simulation experiments to explore the impact of using a fixed global a priori on the MOPITT CO retrievals. This is further illustrated using an alternative criterion to dynamically choose an a priori error covariance matrix (the DAP method) which constrains the CO retrievals using primarily measurement errors. The method will be described in section 2. The simulated CO profiles retrieved from the DAP method are compared to CO profiles retrieved from the ML method in section 3.

2. DATA and the DYNAMIC A PRIORI Method

MOPITT has four CO thermal channels, two CO solar reflectance channels and two CH$_4$ solar reflectance channels. Characteristics of the MOPITT channels are described in Drummond (1992). The radiative transfer equation (RTE) for the upwelling MOPITT A and D signals for each channel is described in Pan et al., (1998). The MOPITT transmittance model (Edwards, et al., 1999) is used in this study to simulate MOPITT radiances while the mixing ratio of target and interfering gases, viewing geometry, vertical thermal profiles and surface emissivity and temperature are
used as inputs. If the measurement errors can be described approximately by a Gaussian distribution, then a regularized estimate of the state vector $X$ (CO vertical profile, surface skin temperature and emissivity) is given by the minimizer of the penalty function (O’Sullivan and Wahba 1985; Wahba, 1990):

$$ J(X) = [Y^m - F(X)]^T [\Sigma_m]^{-1} [Y^m - F(X)] + \gamma [X - X^b] \Sigma^{-1} [X - X^b] \quad (1) $$

In Eq. (1), $Y^m$ is a vector of MOPITT radiances, $X^b$ is the initial guess profile; $F(X)$ is the radiative transfer model described in Pan et al., (1998); $\Sigma$ is the expected covariance matrix of measurement errors and the forward model errors; $C_{\gamma}$ is the first guess expected error covariance matrix used to constrain the solution, and is defined as $\gamma C$ where $C$ is the a priori error covariance matrix for CO profiles, surface skin temperature and emissivity. $C$ is generated from CO profiles collected from field campaigns (see Section 3), while $\gamma$ is a smoothing parameter, which balances the fit to the observations (first term in Eq. (1)) and the fit to the a priori (the second term in Eq. (1)). In the case of the ML method, the $C$ is a statistical constraint based on other measurements and does not vary. This is equivalent to setting $\gamma$ equal to unity.

Since the RTE is nonlinear, the optimal state parameters must be found iteratively. The Newtonian nonlinear iteration method (O’Sullivan and Wahba, 1985) is applied here such that

$$ X_{n+1} = X_n + C_p K_p (K_p C_p K_p + E)^{-1} \delta Y^m + K_p (X_n - X_n) \quad (2) $$

where $\delta Y^m$ indicates the differences between the observed and calculated radiances, $X_n$ is a vector of the first guess state profile generated from the mean of all CO profiles, and $K$ is the weighting function matrix defined by $\partial Y = K \partial X$. The normalized radiance residual is defined as

$$ R_{n+1} = \frac{1}{nch \Sigma} \sum_{k=1}^{nch} |Y^m_k - F(X_{n+1})| / \text{var} \quad (3) $$

where $k$ is an index of channel number, $nch$ is the total number of channels used in the retrieval, and $\text{var}_k$ is the diagonal term of the error covariance matrix $E$, which is specified from differences between calculated and measured radiances. In each iteration, the discrepancy principle (Morozov, 1966) provides smoothing parameters so that the residual norms are close to a priori upper bounds for the $A$ and $D$ signals respectively. The a priori upper bound for $A$ or $D$ signals is defined as

$$ \sigma = \left( \frac{1}{nch \Sigma} \sum_{k=1}^{nch} \eta_k^2 / \text{var}_k \right)^{1/2} \quad (4) $$

where $\eta_k^2$ is the square of the instrument noise plus the square of the forward model errors in channel $k$.

The error is assumed to be random for each channel. Here the measurement noise is known from the calibration. The forward model error is assumed to be 2% of the instrument noise for each channel for the determination of $\sigma$. If $R_{\text{sn+1}} > \sigma$, then the solution is assumed to be over-constrained, and $\gamma$ will be increased for the next iteration. If $R_{\text{sn+1}} < \sigma$, then the solution is assumed to be under-constrained, and $\gamma$ will be decreased. For each iterative step, a value of $\gamma$ for $A$ signals and a value of $\gamma$ for $D$ signals are determined. The iteration stops when the absolute value of the difference between $R_{\text{sn+1}}$ and $\sigma$ is less than a reasonable small number, assumed to be 0.05$s$. However, because the CO vertical distribution may not be completely resolved by MOPITT, retrieval results may not meet the required criterion after several iterations. In this case, the solution having the smallest difference between $R_{\text{sn+1}}$ and $\sigma$ is selected. The retrieval results with $R_{\text{sn+1}}$ larger than 0.5$s$ are discarded. The initial value of $\gamma$, $\gamma_0$, is taken as 10. By using this method, reasonable $\gamma$ can be found, and the retrieved CO profiles and total column will be an inverse solution of the radiative transfer equation in the sense of minimizing the penalty function. In summary, in ML, we assume $C$ is known completely and is held constant in the retrieval. In DAP, $C$ is allowed to vary but the forward model and instrument noise become the constraint.

3. RETRIEVAL RESULTS

To show how different a priori background covariance matrices affect the retrieval results, simulated retrieval experiments have been conducted. A total of 525 CO profiles with ancillary temperature and water vapor vertical distributions were collected from 10 different geophysical locations for varied seasons (Wang, et al., 1999). The CO profiles with odd index numbers are separated from those with even index numbers. The 263 CO profiles with odd index number are used as a training set and the other 262 CO profiles are used as a test set. The a priori covariance matrix ($C$) is generated from the training set. The square roots of the diagonal terms of $C$ for the training set are shown in Fig. 1. In real MOPITT retrievals, because of various limitations of the training set profiles (poor representation of some geographical areas and
seasonal variability, etc.) this a priori matrix cannot be assumed to perfectly represent global CO variability. In this study, CO profiles measured from Carr, Colorado and during the Pacific Exploratory Mission in the Western Pacific Region (PEMWEST) phase B (Wu, et al., 1997) are used to generate alternative a priori matrices (CARR C and PEM-B C). As shown in Fig. 1,

Fig. 1. The square roots of the diagonal terms of CO vertical profiles for a priori covariance matrices C, CARR C and PEM-B C.

the CO variations for different vertical levels over Carr (CARR C) is less than those for PEM-B and is used to represent the model CO ensemble or local observed CO ensemble. The a priori matrix, PEM-B C, is used to represent cases of over-estimated background errors for both the ML and DAP algorithms. In these studies, the ML a priori mean profile is held constant.

The retrieval results for all three a priori matrices for both ML and DAP methods are shown in Figs. 2(a) and 2(b) respectively. For all retrieval experiments, the instrument noises for the A and D signals and forward model error are randomly added to the forward calculated MOPITT radiances. The mean of the training set is used as the initial profile for all retrievals. Because the CO profiles are divided into the testing set and the training set, the STD of the CO training set (Fig. 1) could quantify the CO variations for different vertical levels for the testing set (not shown). Under the condition that the measurement errors and the a priori distribution are well known and are well-represented by a Gaussian distribution, the ML method will provide the best estimates of the state parameters (ML with C cases) by minimizing the penalty function (Wahba, 1985) (Fig. 2). However, in real global MOPITT retrievals, this detailed knowledge of the background error covariance matrix may not be available. Some retrieval results may be under-constrained by the a priori matrix like CARR C (over-represented cases). Some retrieval results may be

over-constrained by the a priori matrix like PEM-B C (under-represented cases), where a small a priori matrix (smaller standard deviation) will decrease the relative magnitude between $KK^{-1}$ and $E$ in Eq. (2) and the CO retrieval will be more tightly constrained by the a priori. The RMSE (Root Mean Square of Errors) of all CO profiles, and the RMS (Root Mean Square) values of true fractional CO column error for ML with CARR C, ML with C, DAP with C and PEM-B C are listed in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>ML with CARR C</th>
<th>ML with C</th>
<th>DAP with C</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE of all CO levels (ppbv)</td>
<td>18.8004</td>
<td>16.3173</td>
<td>13.7083</td>
</tr>
<tr>
<td>RMSE of true fractional CO column error (%)</td>
<td>6.8</td>
<td>6.18</td>
<td>4.72</td>
</tr>
</tbody>
</table>

Table 1. RMSE (Root Mean Square of Errors) of all CO profiles of the RMS (Root Mean Square) values of true fractional CO column error for ML with PEM-B C, ML with CARR C, ML with C and DAP with C cases.

On the other hand, CO profiles retrieved with the DAP method are less affected by a pre-chosen a priori covariance matrix (Fig. 2(b)) than those retrieved
from the ML method. In the DAP CO retrieval, the relative magnitude between $K/C$ and $E$ in Eq. (2) is dynamically adjusted to satisfy the criteria mentioned in Section 2 according to the initial profile ($X_0$), the instrument and forward model noises ($E$), the estimated background error ($C$), and the true CO profile. The $\gamma$ found from the smallest difference between radiance residuals (defined in Eq. (3)) and a priori upper bound (defined in Eq. (4)) (Fig. 3(b)) can provide reasonably good CO retrievals (Fig. 3(a)). However, to be useful practically in an operational sense, the DAP method would require greater characterization of the measurement errors (both radiance noise and forward model) than is currently available for the MOPITT instrument.

4. CONCLUDING REMARKS

The role of the a priori covariance matrix in the retrieval algorithm is to provide extra background information about the retrieval parameters, to constrain the retrieval results by truncating the instrument errors and forward model errors from the measurement information, and to stabilize the matrix inversion. Our studies show that ML retrievals produce the smallest retrieval errors when the simulation profiles are statistically consistent with the selected a priori matrix. When processing MOPITT real operational data by using the ML method, however, the fixed covariance generated from limited regional surface observations and aircraft measurements from field experiments may not be optimum. Compared with the ML, retrievals from the DAP method are less affected by the pre-chosen a priori covariance matrix. However, this method requires good characterization of the measurement error. In the future, when MOPITT processing has stabilized and we better understand the measurements and associated error sources, work will begin on the implementation of a variable a priori that depends on location and season. This will be tested thoroughly with data from validation sites and field campaigns.

5. REFERENCE


