Tropospheric Emission Spectrometer: Retrieval Method and Error Analysis

Kevin W. Bowman, Clive D. Rodgers, Susan Sund Kulawik, John Worden, Edwin Sarkissian, Greg Osterman, Tilman Steck, Ming Lou, Annmarie Eldering, Mark Shephard, Helen Worden, Michael Lampel, Shepard Clough, Pat Brown, Curtis Rinsland, Michael Gunson, and Reinhard Beer

***Abstract*—We describe the approach for the estimation of the atmospheric state, e.g., temperature, water, ozone, from calibrated, spectral radiances measured from the Tropospheric Emission Spectrometer (TES) onboard the Aura spacecraft. The methodology is based on the *maximum a posteriori* estimate, which mathematically requires the minimization of the difference between observed spectral radiances and a nonlinear model of radiative transfer of the atmospheric state subject to the constraint that the estimated state must be consistent with an *a priori* probability distribution for that state. The minimization techniques employed here are based on the trust-region Levenberg–Marquardt algorithm. An analysis of the errors for this estimate include smoothing, random, spectroscopic, "cross-state," representation, and systematic errors. In addition, several metrics and diagnostics are introduced that assess the resolution, quality, and statistical significance of the retrievals. We illustrate this methodology for the retrieval of atmospheric and surface temperature, water vapor, and ozone over the Gulf of Mexico on November 3, 2004.**

***Index Terms*—Atmospheres, constituents, inverse methods, remote sounding, temperature.**

I. Introduction

THE Tropospheric Emission Spectrometer (TES) [1], launched on July 2004 on the EOS Aura mission, will provide the first global view of tropospheric ozone. The investigation will focus on mapping the global distribution of tropospheric ozone and on understanding the factors that control ozone concentrations.

TES is an infrared, high-resolution, Fourier transform spectrometer covering the spectral range 650–3050 cm^{-1} (3.3–15.4 μm) at a spectral resolution of 0.1 cm^{-1} (nadir viewing) or 0.025 cm^{-1} (limb viewing). TES observed spectral radiances are used to estimate the atmospheric state, which includes vertical distributions of trace gases along with related atmospheric and surface parameters such as effective cloud optical depth and surface emissivity. The relationship between the state of the atmosphere and observed spectra can be described by a nonlinear forward model of the radiative transfer from the atmosphere through the instrument. The atmospheric state is estimated by minimizing the L^2 norm difference between the observations and forward model subject to constraints on the first and second order moments of those parameters [2] by means of a constrained nonlinear least squares fitting procedure (alternative approaches have also been explored [3], [4]). A critical outcome of this process is a detailed characterization of the smoothing, random, and systematic errors for the target parameters along with important retrieval metrics such as degrees of freedom (DOF), information content, and vertical resolution. However, for the data user to fully understand the nature of the product and its errors, it is necessary to understand the details of this process.

Manuscript received May 2, 2005; revised September 29, 2005. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

K. W. Bowman, S. S. Kulawik, J. Worden, E. Sarkissian, G. Osterman, M. Lou, A. Eldering, H. Worden, M. Gunson, and R. Beer are with the Jet Propulsion Laboratory Pasadena, CA 91109 USA (e-mail: kevin.bowman@jpl.nasa.gov).

C. D. Rodgers is with the Department of Physics, University of Oxford, Oxford OX1 2JD, U.K.

T. Steck is with the Institut for Meterologie und Klimaforschung, Forschungszentrum Karlsruhe 76021, Germany.

M. Shepard and S. Clough are with Atmospheric and Environmental Research, Lexington, MA 02421 USA.

M. Lampel is with Raytheon, ITSS, Pasadena, CA 91101 USA.

P. Brown is with Hartford Financial Services, Hartford, CT 06115 USA.

C. Rinsland is with NASA Langley Research Center, Hampton, VA 23681 USA.

Digital Object Identifier 10.1109/TGRS.2006.871234

II. Retrieval Methodology

A. Additive Noise Model

Measured radiances in TES can be related to a forward model through the following additive noise model:

$$\mathbf{y} = \mathbf{L}(\mathbf{x}, \mathbf{b}) + \boldsymbol{\epsilon} \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^M$ (a real vector of length M), is the calibrated, measured spectrum and $\mathbf{x} \in \mathbb{R}^N$ is the "full" state vector, and $\mathbf{b}$ is a vector of nonretrieved forward model parameters. A *state vector* specifies the atmospheric state and instrumental parameters that influence the measurement. We will use the term *full state vector* to indicate the set of parameters required by the forward model to simulate a measurement to the necessary accuracy, and the term *retrieval state vector* or simply *retrieval vector* (usually $\mathbf{z}$) to indicate a reduced resolution state of the full state vector that will be estimated. The remaining elements, which are encapsulated in $\mathbf{b}$, are predetermined, e.g, surface height, geometry and trajectory of the spacecraft, etc. The forward model operator $\mathbf{L} : \mathbb{R}^N \rightarrow \mathbb{R}^M$: is a discretized version of the equation of radiative transfer that simulates the spectrum resulting from the propagation of radiation through the atmosphere from the Earth to the spacecraft. The noise term $\boldsymbol{\epsilon} \in \mathbb{R}^M$ is assumed to be zero mean, Gaussian noise with the error covariance, $\mathbf{S}_\epsilon \triangleq E[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T]$ where $E[\cdot]$ is the expectation operator

[5]. Operationally, the spectral error covariance is assumed to be diagonal with elements $[\mathbf{S}_\epsilon]_{ii} = \sigma_i^2$.

1) Forward Model: The TES forward model is described in detail by [6]. The radiative transfer is referenced with respect to the logarithm of pressure. For nadir viewing, surface temperature and emissivity along with clouds are included in the forward model. In the limb view, the trace gas distribution is modeled as spherically symmetric. The TES instrument line shape is applied to the modeled spectra producing a spectral resolution and spacing of 0.06 cm^{-1} for nadir spectra, and 0.02 cm^{-1} for limb spectra. In addition the spectra in apodized with Norton–Beer medium apodization [7], [8]. The whole spectrum is not exploited for the retrieval. Spectral windows are carefully selected in order to reduce the computational load and minimize systematic errors from nonretrieved atmospheric parameters. Objective methods for microwindow selection have been discussed by [9]–[12]. Candidate regions are selected for each species based on known absorption and interferences. The information content H for each spectral element in the interval is evaluated given an expected *a priori* covariance, data error, constraints, and atmospheric Jacobians, i.e., following Shannon [13]:

$$H = \frac{1}{2}\log_2\left(\frac{|\mathbf{S}_a|}{|\hat{\mathbf{S}}_{\tilde{\mathbf{x}}}|}\right) \tag{2}$$

where

$$\mathbf{S}_a = E[(\mathbf{x}-\mathbf{x}_a)(\mathbf{x}-\mathbf{x}_a)^\top] \tag{3}$$

is the *a priori* covariance matrix, $\mathbf{x}_a = E[\mathbf{x}]$ is the *a priori* vector, and $\hat{\mathbf{S}}$ is *a posteriori* covariance matrix, which is defined in (44). Spectral elements are selected to maximize the information. We have used the approach described in [12].

The spectral size of the microwindows ranges from a few wavenumbers to one hundred or more wavenumbers, depending upon the species to be retrieved. In order to achieve the maximum reduction in computing time, both the calculated and observed spectrum are filtered with the Norton–Beer medium apodization [7]. This reduces the ringing of lines, and thus the width of the spectrum required.

Different microwindows provide information about different altitude regions and are different for nadir and limb views. Generally, strong lines are used for high altitudes, while weak lines will give more information at low altitudes.

2) Measurement Error: The measurement error vector $\boldsymbol{\epsilon}$ in (1) comprises both random sources (primarily detector noise) and systematic sources from uncertain forward model parameters. Only the error covariance matrix of the random error is used in the retrieval. The effect of systematic errors on the retrieval are assessed *a posteriori* as discussed in Section V. This approach is taken to maintain a diagonal covariance matrix, which reduces computational processing time. Noise equivalent spectral radiance (NESR) has been obtained from the preflight laboratory calibration of TES [14], and is summarized in Table I. Table I gives the bandpass frequency average of the modeled NESR expected for clear sky nadir views with a surface temperature around 300 K (this will tend to slightly overestimate noise from polar measurements) and for moderate Norton–Beer apodization (which includes an additional multiplicative factor of 0.665).

TABLE I
BAND AVERAGE NESR VALUES OBTAINED FROM PREFLIGHT CALIBRATION

TES Filter ID	Frequency Range cm^{-1}	Nadir NESR nW cm^{-2} $\text{sr}^{-1}/\text{cm}^{-1}$ 16 pixel average
2B1	650-900	133
1B2	950-1150	17
2A1	1100-1325	17
1A1	1900-2250	23

B. Constrained Least Squares Estimation

Given the additive noise model of the measured radiances, we find an estimate $\hat{\mathbf{z}}$ of the retrieval vector $\mathbf{z}$ by minimizing the difference between the observed and the forward model spectral radiances, subject optionally to a quadratic constraint. That is, we minimize the cost function

$$C(\mathbf{z}) = \|\mathbf{y}-\mathbf{L}(\mathbf{x}))\|^2_{\mathbf{S}_\epsilon^{-1}} + \|\mathbf{z}-\mathbf{z}_c\|^2_{\boldsymbol{\Lambda}} \tag{4}$$

where the full state vector is restricted to subspace defined by the interpolation or mapping $\mathbf{x} = \mathfrak{M}(\mathbf{z})$. The second quadratic term is optional, and can be used to provide a constraint on the solution if the problem is ill-posed or ill-conditioned. The state vector $\mathbf{z}_c$ can be regarded as $\mathbf{x}_a = E[\mathbf{x}]$, the expected state of the atmosphere in the absence of a measurement, i.e., the *a priori*, and $\boldsymbol{\Lambda}$ as the inverse of the covariance of the uncertainty $\mathbf{S}_a = E[(\mathbf{x}-\mathbf{x}_a)(\mathbf{x}-\mathbf{x}_a)^\top]$, or the term can just be regarded as an *ad hoc* constraint providing e.g., smoothness.

Minimization proceeds iteratively from an initial guess by means of an appropriate nonlinear least squares solver until convergence to obtain the retrieval vector from $\hat{\mathbf{z}} = \min_{\mathbf{z}}(C(\mathbf{z}))$ and the retrieved full state vector from $\hat{\mathbf{x}} = \mathfrak{M}(\hat{\mathbf{z}})$. In practice the mapping is linearised as described in Section III-A.

When the constraint term is based on an *a priori* Gaussian probability density function for the state, the constrained least squares estimate becomes a *maximum a posteriori* (MAP) estimate. If the constraint term is absent, and the only constraint on $\hat{\mathbf{x}}$ is provided by the mapping, the estimate is a *maximum likelihood* (ML) one. In the latter case, the second quadratic term in (4) is dropped.

III. CONSTRAINTS

A. Mapping

1) Mapping Between Retrieval and Full State Vectors: The discretization of the full state vector is chosen to be fine enough to accurately calculate the equation of radiative transfer. However, for the purposes of a retrieval, this grid may be too fine. A mapping is applied that restricts that solution space of the retrieval to a resolution that is more representative of TES sensitivities but still can represent vertical atmospheric variability [15]. The retrieval vector, $\mathbf{z} \in \mathbb{R}^{N'}$, is related to the full state vector by a mapping $\mathfrak{M} : \mathbb{R}^{N'} \rightarrow \mathbb{R}^N$ so that $\mathbf{x} = \mathfrak{M}(\mathbf{z})$. In general, $N' < N$. For simplicity, we shall restrict our attention initially to linear maps so that

$$\mathbf{x} = \mathbf{M}\mathbf{z}. \tag{5}$$

The full state vector is expressed as a linear combination of basis vectors $\mathbf{m}_i$ (the columns of $\mathbf{M}$) weighted by retrieval parameters z_i. The mapping matrix is typically a piecewise linear interpolation function of pressure though the matrix could represent

any general linear transformation, e.g., singular value decomposition (SVD), wavelets, etc. It is important to note that (5) is restriction on the full state vector for the retrieval, not for the actual atmospheric state.

In order to calculate an initial guess of the atmospheric state, an inverse mapping from a full state vector to a retrieval vector is needed

$$\tilde{\mathbf{z}} = \mathbf{M}^* \mathbf{x}. \tag{6}$$

If the rank of $\mathbf{M}$ is N', then the inverse mapping is chosen to be least-squares inverse of the map given by

$$\mathbf{M}^* = (\mathbf{M}^\top \mathbf{M})^{-1} \mathbf{M}^\top. \tag{7}$$

The Jacobian of the retrieval state vector with respect to the radiances is expressed as

$$\mathbf{K}_z = \frac{\partial \mathbf{L}(\mathbf{Mz})}{\partial \mathbf{z}}. \tag{8}$$

The chain rule can be applied to (8) in order to relate the retrieval Jacobian to the full-state Jacobian

$$\frac{\partial \mathbf{L}}{\partial \mathbf{z}} = \frac{\partial \mathbf{L}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{z}} \tag{9}$$

or

$$\mathbf{K}_z = \mathbf{K}_x \mathbf{M}. \tag{10}$$

The full-state Jacobian is calculated in the forward model and then mapped to the retrieval Jacobian through (10) for use in the retrieval algorithm.

For several atmospheric species, e.g., ozone, H_2O, the retrieval of the natural log of the volume mixing ratio (VMR) is preferable. In this case

$$\mathbf{x} = \ln \mathbf{q} = \mathbf{Mz} \tag{11}$$

where $\mathbf{q}$ is a state vector whose elements are the volume mixing ratio of the atmospheric state as a function of pressure. The nonlinearity introduced by the natural log has no impact on the linearity of the mapping, but does have an impact on the error analysis.

B. A Priori Constraints

A priori is a description of what is known or believed about the state before the measurement is considered. Typically it may comprise a climatological estimate of the state plus some measure of its uncertainty, most conveniently expressed as a covariance matrix. This covariance matrix is calculated from chemistry and transport models (CTM), sondes, and meterological assimilation systems. For temperature and H_2O, the atmospheric profiles are calculated from the GEOS global transport model maintained at NASA's Global Modeling and Assimilation Office (GMAO). For GEOS-4, these profiles are discretized on $1.25^\circ \times 1.00^\circ$ (Lon $\times$ Lat) grid. The initial guess for temperature and water take the eight closest profiles in latitude, longitude, and universal time coordinate (UTC) and interpolates them to the location and time of the TES footprint. Surface temperature initial guess is arrived at by the same process using the GMAO skin temperature fields. Initial guess atmospheric profiles for all other chemical species are determined using MOZART [16] model results and are interpolated to the TES footprint in a similar manner. In addition, *ad hoc* smoothness can be imposed by adding a quadratic form to the cost function that penalizes the first or second difference of the profile [17], [18] or by defining correlations between elements of the profile that decay exponentially as a function of distance [19].

Two approaches are used to calculate constraints. The first approach fits the row of a climatological covariance matrix calculated from a global CTM, e.g., MOZART, to an exponentially declining function

$$[\mathbf{S}'_a]_i = \hat{\beta} \exp(-\hat{\alpha}(\ln p_i - \ln p) \tag{12}$$

where α and β are determined by

$$(\hat{\alpha}, \hat{\beta}) = \min_{\alpha,\beta} \|[\mathbf{S}_a]_i - \beta \exp(-\alpha(\ln p_i - \ln p)\|^2 \tag{13}$$

where p is pressure, i is a pressure level index, $\alpha > 0$ scales the rate of decrease, and β adjusts the overall magnitude of the covariance. This approach permits the construction of a constraint matrix that includes most of the atmospheric variability but avoids numerical problems such as singularity and physical problems such as nonphysical correlations between troposphere and stratosphere that can arise from CTMs.

The second approach is based on an altitude dependent Tikhonov constraint where a combination of first and second derivative norm constraints are imposed on a retrieved vertical profile. However, the strength of the constraints relative to the measurement are allowed to vary as a function of altitude [20].

IV. Numerical Approach

A. Structured Least Squares

There is a wide range of numerical methods for minimizing nonlinear least-squares functions such as the cost function in (4) [21], [22]. The basic methods available in the retrieval are the Gauss–Newton method and the Levenberg–Marquardt (LM) method. Both methods minimize the cost function in (4) by iterative minimization with respect to the retrieval vector. We define a step in the retrieval vector as

$$\delta \mathbf{z} = \mathbf{z}_{i+1} - \mathbf{z}_i. \tag{14}$$

The Gauss–Newton method consists of iteratively minimizing the cost function linearized about $\mathbf{z}_i$. For the MAP cost function, a Gauss–Newton iteration can be calculated by solving the following equation [23]:

$$\mathbf{K}' \delta \mathbf{z} = \delta \mathbf{y}' \tag{15}$$

where

$$\mathbf{K}' = \begin{pmatrix} \mathbf{S}_\epsilon^{-1/2} \mathbf{K}_z(\mathbf{z}_i) \\ \mathbf{\Lambda}^{1/2} \end{pmatrix} \tag{16}$$

and

$$\delta \mathbf{y}' = \begin{pmatrix} -\mathbf{S}_\epsilon^{-1/2} (\mathbf{y} - \mathbf{L}(\mathbf{Mz}_i) \\ \mathbf{\Lambda}^{1/2} (\mathbf{z} - \mathbf{z}_c) \end{pmatrix}. \tag{17}$$

The Jacobian, $\mathbf{K}_z$, in (16) is defined in (8) and related to the full state Jacobian in (10). A numerically robust method for solving (15) is through the factorization of the augmented Jacobian $\mathbf{K}'$ into the QR decomposition

$$\mathbf{QR} \delta \mathbf{z} = \delta \mathbf{y}' \tag{18}$$

where $\mathbf{Q} \in \mathbb{R}^{M\times M}$ is an orthogonal matrix and $\mathbf{R} \in \mathbb{R}^{M\times N}$ is an upper triangular matrix. The orthogonal matrix can be inverted implicitly and the retrieval vector can be calculated by back substitution from $\mathbf{R}$ [22]. The forward model is evaluated by the updated full state vector $\mathbf{x}_{i+1} = \mathbf{M}\mathbf{z}_{i+1}$ and (15) is solved again until convergence.

Alternatively, the least squares solution to (15) may also be written as

$$\delta\mathbf{z} = (\mathbf{K}'^T\mathbf{K}')^{-1}\mathbf{K}'^T\delta\mathbf{y}' \tag{19}$$

or

$$\mathbf{z}_{i+1} = \mathbf{z}_i + (\mathbf{\Lambda} + \mathbf{K}_i^T\mathbf{S}_\epsilon^{-1}\mathbf{K}_i)^{-1} (\mathbf{K}_i^T\mathbf{S}_\epsilon^{-1}[\mathbf{y} - \mathbf{L}(\mathbf{M}\mathbf{z}_i)] - \mathbf{\Lambda}[\mathbf{z}_i - \mathbf{z}_c]). \tag{20}$$

Note that the efficiency of the iteration is enhanced if $\mathbf{S}_\epsilon$ is diagonal, particularly for cases where the measurement vector is large, as in the case of TES. However, the apodization of microwindows will introduce off-diagonal elements in $\mathbf{S}_\epsilon$ and other sources such as calibration may also cause correlations in the measurement errors. The off-diagonal elements due to apodization can be calculated. For numerical efficiency, however, the measurement error covariance is assumed to be diagonal.

The Gauss–Newton method is satisfactory for *small residual problems* [21]. For these problems, the initial guess is in a region sufficiently close to the solution such that second-order derivative of the cost function are small. If the initial guess does not satisfy this condition, then the LM method is used. This algorithm is implemented as a *trust-region* method [24]. In this method, the cost function in (4) is minimized subject to the trust region

$$\|(\mathbf{z}_{i+1} - \mathbf{z}_i)\|^2_{\mathbf{W}^\top\mathbf{W}} \leq \Delta_i. \tag{21}$$

The *trust region radius* Δ_i defines the radius of a hyper-ellipsoid over which the cost function is considered linear with respect to the step. An LM step is calculated by solving

$$\begin{pmatrix} \mathbf{K}' \\ \gamma_i^{1/2}\mathbf{W} \end{pmatrix} \delta\mathbf{z} = \begin{pmatrix} \delta\mathbf{y}' \\ \mathbf{0} \end{pmatrix} \tag{22}$$

with the QR decomposition or alternatively as

$$\mathbf{z}_{i+1} = \mathbf{z}_i + (\gamma_i\mathbf{W}^T\mathbf{W} + \mathbf{\Lambda} + \mathbf{K}_i^T\mathbf{S}_\epsilon^{-1}\mathbf{K}_i)^{-1} (\mathbf{K}_i^T\mathbf{S}_\epsilon^{-1}[\mathbf{y} - \mathbf{L}(\mathbf{M}\mathbf{z}_i)] - \mathbf{\Lambda}[\mathbf{z}_i - \mathbf{z}_c]) \tag{23}$$

where the parameter γ_i is called the LM parameter and $\mathbf{W}$ is a nonzero scaling matrix. The LM parameter is varied from step to step according to the strategy described by Moré [24]:

```
1: Find γ_i such that the step size is within the
trust region radius.
2: Check that C(z_{i+1}) < C(z_i). If the update cost
function has increased, then reduce the trust region
radius and return to step 1.
3: Calculate the linearity measure, ρ, from (24).
4: Increase the trust region radius if ρ > 0.01 and
decrease otherwise.
5: Return to step 1 for next iteration.
```

The linearity measure is defined as

$$\rho = \frac{1 - \left(\frac{\|\boldsymbol{\delta}\mathbf{y}_{i+1}\|}{\|\boldsymbol{\delta}\mathbf{y}_i\|}\right)^2}{\left(\frac{\|\mathbf{K}_i'\delta\mathbf{z}\|}{\|\boldsymbol{\delta}\mathbf{y}_i\|}\right)^2 + 2\left(\frac{\sqrt{\gamma_i}\|\mathbf{W}_i\delta\mathbf{z}\|}{\|\boldsymbol{\delta}\mathbf{y}_i\|}\right)^2} \tag{24}$$

where $\boldsymbol{\delta}\mathbf{y}_i$ is the residual for the ith iteration. This measure compares the difference between the expected reduction in the cost function relative under the assumption of a perfectly linear system to the actual reduction in the cost function. Perfect agreement between the predicted and actual reduction occurs when $\rho = 1$, less reduction than expected for $\rho \in (0, 1)$, greater than expected reduction for $\rho > 1$, and an increase in the cost function for $\rho < 0$.

B. Stopping and Convergence Criteria

Stopping criteria provide a set of sufficient conditions for which an iterative minimization scheme, such as the Gauss–Newton algorithm, should terminate. These criteria are distinguished from convergence criteria in that satisfying one or more of these conditions do not necessarily imply convergence. Stopping criteria discussed here are one-point and two-point criteria, which depend on the current iteration or the current iteration and the previous iteration, respectively.

1) Sufficient Condition Criteria: The first one-point criterion is the maximum number of iterations

$$i \leqslant N_{\max} \tag{25}$$

where i is the iteration number. Clearly, this criterion is a poor indicator of convergence. However, computational restrictions require the use of (25). The quality of the retrieval can be assessed afterwards. The second one-point criterion is based on the cost function in (4)

$$\frac{\sqrt{C(\mathbf{z})}}{\sqrt{\mathbf{N}}} \leqslant 1 - \delta. \tag{26}$$

The value of δ is usually set to be equal to the normalized variance of the χ^2 distribution of the cost function in (4). If the estimates of the covariance of the measurement error are correct, then continuing to iterate after (26) has been satisfied will only "fit the noise." Nevertheless, this condition does not indicate whether a local minimum has been reached and hence does not test convergence.

2) Tests For Convergence: The following criteria are used in conjunction to test for convergence [25], [26]:

$$\frac{\left\|\mathbf{K}'^T\boldsymbol{\delta}\mathbf{y}_\mathbf{i}'\right\|_2}{1 + C(\mathbf{z}_i)} \leqslant \sqrt{\varepsilon} \tag{27}$$

$$\frac{\|\mathbf{z}_i - \mathbf{z}_{i-1}\|_2}{1 + \|\mathbf{z}_i\|_2} \leqslant \sqrt{\varepsilon} \tag{28}$$

$$\frac{|C(\mathbf{z}_i) - C(\mathbf{z}_{i-1})|}{1 + C(\mathbf{z}_i)} \leqslant \varepsilon \tag{29}$$

where ε is the threshold value. Equation (27) insures that the derivative of the cost function, $\nabla C(\mathbf{z}) = -\mathbf{K}'^T\boldsymbol{\delta}\mathbf{y}'$ is close to zero, which is a necessary condition for a local minimum. Equations (28) and (29), which are two-point stopping criteria, check that the fractional change in the state vector and cost function are going to zero. The denominator is augmented by "1+" in order

to avoid a divide-by-zero situation. The value of the threshold is based on the accuracy of the Jacobian calculation. These three condition must be simultaneously satisfied in order for the iterations in the minimization algorithm to terminate.

V. Error Characterization

A. Linear Retrieval

If the estimated state calculated from (4) is "close" to the actual state, then the estimated state can be expressed in terms of the actual state through the linear retrieval [2]

$$\hat{\mathbf{x}} = \mathbf{x}_c + \mathbf{A}_{xx}(\mathbf{x} - \mathbf{x}_c) + \mathbf{M}\mathbf{G}_z\mathbf{n} + \sum_i \mathbf{M}\mathbf{G}_z\mathbf{K}_{\mathbf{b}}^i(\mathbf{b}^i - \mathbf{b}_a^i) \tag{30}$$

where $\mathbf{M}$ is the mapping matrix defined in (5), $\mathbf{n}$ is the spectral noise vector defined in (1), $\mathbf{x}$ is the full state vector of the actual state, and $\mathbf{x}_c = \mathbf{M}\mathbf{z}_c$ is the constraint vector. The vector $\mathbf{b}$ is the true state for those parameters that also affect the modeled radiance, e.g., concentrations of interfering gases, calibration, etc. The vector $\mathbf{b}_a$ is the corresponding *a priori* values for the vector $\mathbf{b}$. The Jacobian, $\mathbf{K}_b = \partial\mathbf{L}/\partial\mathbf{b}$, describes the dependency of the forward model radiance, $\mathbf{L}$, on the vector $\mathbf{b}$. The $\mathbf{G}_z$ is the gain matrix, which is defined by

$$\mathbf{G}_z = \frac{\partial \mathbf{z}}{\partial \mathbf{L}} = \left(\mathbf{K}_z^T\mathbf{S}_n^{-1}\mathbf{K}_z + \Lambda_z\right)^{-1}\mathbf{K}_z^T\mathbf{S}_n^{-1}. \tag{31}$$

The retrieval Jacobian, $\mathbf{K}_z$, is defined in (8). Equation (30) is a valid approximation to the minimization of (4) when the estimate is close to the true state

$$\mathbf{K}_x[\mathbf{x} - \hat{\mathbf{x}}] \approx \mathbf{L}(\mathbf{x}, \mathbf{b}) - \mathbf{L}(\hat{\mathbf{x}}, \mathbf{b}). \tag{32}$$

The averaging kernel matrix or resolution matrix, $\mathbf{A}_{xx} = \partial\hat{\mathbf{x}}/\partial\mathbf{x}$ is the sensitivity of the retrieval to the true state of the atmosphere and is computed by the following equation:

$$\mathbf{A}_{xx} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \mathbf{L}}\frac{\partial \mathbf{L}}{\partial \mathbf{x}} = \mathbf{M}\mathbf{G}_z\mathbf{K}_x. \tag{33}$$

The retrieved state vector may be a combination of several different atmospheric and surface parameters, e.g., surface temperature, atmospheric temperature, and water vapor. In this case, it is useful to quantify the effect of jointly retrieved parameters on a given state, e.g., the effect of the retrieval of temperature on water vapor. For jointly retrieved parameters, the averaging kernel matrix maybe partitioned into submatrices that describe the sensitivity of a particular retrieved state, e.g, water vapor, to the perturbation of another true state, e.g., temperature

$$\mathbf{A}_{i,j} = \frac{\partial \hat{\mathbf{x}}_i}{\partial \mathbf{x}_j} \tag{34}$$

where i is an index for a subset of the state vector, e.g, vertical water vapor profile, and j is the index to another subset of the state vector, e.g., vertical temperature profile. Equation (30) can be modified with an additional "cross-state" term to be

$$\hat{\mathbf{x}}^j = \mathbf{x}_c^j + \mathbf{A}_{jj}(\mathbf{x}^j - \mathbf{x}_c^j) + \sum_{i=1, i\neq j}^{Q} \mathbf{A}_{ij}(\mathbf{x}^i - \mathbf{x}_c^i) + \text{error terms} \tag{35}$$

where Q is the number of components of the state vector associated with different surface and atmospheric parameters. The "error terms" in (35) are the random and systematic errors in (30).

B. Error Analysis of Linear Retrieval

1) Retrieval Covariance: Equations (30) and (35) becomes the basis for the calculation of the second-order statistics of the retrieval and its error. The mean of the estimate is simply

$$\bar{\hat{\mathbf{x}}} = E[\hat{\mathbf{x}}]. \tag{36}$$

The estimate will be unbiased, i.e., $E[\hat{\mathbf{x}}] = E[\mathbf{x}]$, if $\mathbf{x}_c = E[\mathbf{x}]$, $\mathbf{b}_a^i = E[\mathbf{b}^i]$, and $E[\mathbf{n}] = 0$. The covariance of the retrieval for the jth atmospheric parameter, i.e., $\mathbf{S}_{\hat{\mathbf{x}}_j} = E[(\hat{\mathbf{x}}_j - \bar{\hat{\mathbf{x}}}_j)(\hat{\mathbf{x}}_j - \bar{\hat{\mathbf{x}}}_j)^\top]$ may be written as a sum of the following covariance matrices:

$$\mathbf{S}_{\hat{\mathbf{x}}_j} = \hat{\mathbf{S}}_s^j + \mathbf{S}_{\text{cs}}^j + \mathbf{S}_m + \mathbf{S}_{\text{sys}} \tag{37}$$

where the smoothing covariance is

$$\hat{\mathbf{S}}_s^j = \mathbf{A}_{jj}\mathbf{S}_{\mathbf{x}_j}\mathbf{A}_{jj}^\top \tag{38}$$

and the "cross-state" error covariance is

$$\mathbf{S}_{\text{cs}}^j = \sum_{i=1, i\neq j}^{Q} \mathbf{A}_{ij}\mathbf{S}_{\mathbf{x}_i}(\mathbf{A}_{ij})^\top \tag{39}$$

where we have assumed $E[\mathbf{x}_i\mathbf{x}_j] = 0$. In addition, the measurement error covariance is

$$\mathbf{S}_m = \mathbf{M}\mathbf{G}_z\mathbf{S}_\epsilon\mathbf{G}_z^\top\mathbf{M}^\top \tag{40}$$

and the "systematic" error covariance is

$$\mathbf{S}_{\text{sys}} = \sum_i \mathbf{M}\mathbf{G}_z\mathbf{K}_{\mathbf{b}}^i\mathbf{S}_b^i(\mathbf{M}\mathbf{G}_z\mathbf{K}_{\mathbf{b}}^i)^\top. \tag{41}$$

The right-hand side of (37) is composed of four terms. The first term, the so-called smoothing term [2], results from applying constraints to the estimate of retrieval state vector. These constraints can be a combination of "mapping" or "hard" constraints (e.g., representing the profile on a coarse pressure grid via the mapping described in Section III-A), or "soft" constraints (e.g., adding a quadratic penalty function to (4) with a constraint matrix described in Section III-B) in order to ensure an acceptable regularization. Physically, the smoothing error describes the uncertainty due to un-resolved fine structure. The second term is the so-called "cross-state" error covariance that describes the effect of other jointly retrieved states on a particular state. This term disappears if the entire retrieval is considered or the term is moved into the systematic error if only one atmospheric parameter is retrieved. Nevertheless, if one is only characterizing some subset of an atmospheric parameter, e.g., tropospheric ozone, then the cross-state is a useful term to characterize the impact of the rest of the set of atmospheric parameters on the retrieved state, e.g., stratospheric ozone. The third term (measurement error covariance) transforms the random instrument spectral error covariance to an error on the state vector. The fourth term transforms the error covariance from forward model parameters to an error on the retrieved state vector. The systematic errors that have been considered are errors from previously retrieved atmospheric trace gas species, atmospheric temperature, surface parameters

(surface temperature and emissivity), and spectroscopic line errors. The quantitative description of expected errors for TES nadir retrievals are presented in [12]. In the case of previously retrieved parameters, the *a posteriori* error covariance is used for $\mathbf{S}_b$. We refer to this case as "error propagation."

2) Error Covariance: The error in the retrieval is defined as the difference between true state and the estimate

$$\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}. \tag{42}$$

The second-order statistics of the error, $\tilde{\mathbf{x}}$, are similar to the statistics of the retrieval. The mean of the error

$$E[\tilde{\mathbf{x}}] = \bar{\tilde{\mathbf{x}}} \tag{43}$$

is unbiased, i.e., $\bar{\tilde{\mathbf{x}}} = 0$, if $\mathbf{x}_c = E[\mathbf{x}]$, $\mathbf{b}_a^i = E[\mathbf{b}^i]$, and $E[\mathbf{n}] = 0$.

Substituting (30) and (35) into (42), the total error or *a posteriori* error covariance, $\mathbf{S}_{\tilde{\mathbf{x}}_j} = E[(\hat{\mathbf{x}}_j - \bar{\tilde{\mathbf{x}}}_j)(\hat{\mathbf{x}}_j - \bar{\tilde{\mathbf{x}}}_j)^\top]$, can be written as

$$\mathbf{S}_{\tilde{\mathbf{x}}_j} = \mathbf{S}_s^j + \mathbf{S}_{\text{cs}}^j + \mathbf{S}_m + \mathbf{S}_{\text{sys}} \tag{44}$$

where the smoothing error covariance is

$$\mathbf{S}_s^j = (\mathbf{I} - \mathbf{A}_{jj})\mathbf{S}_{\mathbf{x}_j}(\mathbf{I} - \mathbf{A}_{jj})^\top \tag{45}$$

and the other terms are the same as in (37).

C. Retrieval Metrics and Diagnostics

Retrieval metrics and diagnostics provide information on the performance of the retrievals that can be distilled to scalar quantities. In order to determine whether the LM algorithm has converged to a local minimum or detect the presence of systematic errors, the spectral χ^2 statistics are calculated

$$\chi^2 = \|\mathbf{y} - \mathbf{L}(\hat{\mathbf{x}})\|^2_{\mathbf{S}_\epsilon^{-1}}. \tag{46}$$

The covariance of the spectral residual, $\delta\mathbf{y} = \mathbf{S}_\epsilon^{-1/2}(\mathbf{y} - \mathbf{L}(\hat{\mathbf{x}}))$ is [2]

$$E[\delta\mathbf{y}\delta\mathbf{y}^\top] = (\mathbf{S}_\epsilon^{-1/2}\mathbf{K}\mathbf{S}_a\mathbf{K}^\top\mathbf{S}_\epsilon^{-1/2} + \mathbf{I})^{-1} \tag{47}$$

where $\mathbf{S}_a$ is defined in (3). Neglecting frequency-correlation from the apodization and spectral systematic errors, and assuming that the inequality

$$\|\mathbf{S}_\epsilon^{-1/2}\mathbf{K}\mathbf{S}_a\mathbf{K}^\top\mathbf{S}_\epsilon^{-1/2}\| \ll \|\mathbf{I}\| \tag{48}$$

holds, then the right-hand side of (46) should follow a χ distribution where a standard deviation in the mean of the residual is $0 \pm (1/M)^{1/2}$ and the root-mean-square (rms) of $1 \pm (2/M)^{1/2}$, where M is the number of spectral measurement points used in the retrieval. In addition to the spectral χ^2, the mean and standard deviation of the spectral residuals are also an important retrieval diagnostic.

There are three retrieval metrics that assess the performance of the retrieval: two are based on the averaging kernel and one is based *a posteriori* covariance matrix. The "resolution" of the retrieval can be defined from the averaging kernel matrix. The vertical resolution of an atmospheric retrieval, defined on a pressure (or altitude grid), can be derived from the rows of the averaging kernel matrix $\partial\hat{x}_i/\partial\mathbf{x}$ which define the relative contribution of each element of the true state to the estimate at pressure (or altitude) level i. The resolution is then defined as the full-width half-maximum of the rows of the averaging kernel.

In the case of a MAP estimate, the expected value of the cost function [(4)] is equal to the number of DOF in the measurement, i.e., the number of elements M in the measurement vector. This can be divided into two parts, corresponding to the two terms in the cost function. In particular, the degrees of freedom for signal (DOFS) d_s is defined as the expected value of the second term [2]

$$d_s = E[(\mathbf{x} - \mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x} - \mathbf{x}_a)]. \tag{49}$$

Note that d_s is not defined in the case of a ML estimate, or when an *ad hoc* constraint term is used. Degrees of freedom for signal is a measure of the number of independent pieces of information provided by the measurement, weighted according to the signal-to-noise variance of each piece of information. For example an independent quantity with a s/n ratio of unity would contribute 0.5 to d_s. It can also be thought of as a measure of the minimum number of parameters that could be used to define a retrieval vector without loss of information. The information content, which is defined in (2), is a measure of the reduction in uncertainty *a posteriori*. This reduction is based on the ratio of the error volume of the variability of the state to the *a posteriori* variability.

VI. Retrieval and Error Characterization of Single Nadir Scene

A. Strategy

In order to elucidate the mathematics discussed in the previous sections, we consider a TES nadir retrieval taken on November 3, 2005, over the Gulf of Mexico at 25.3° latitude and −95.82° longitude, which is roughly 75 miles southeast of the U.S.–Mexican border. This retrieval was taken as part of the Aura Validation Experiment (AVE) Campaign conducted from Ellington Field in Houston, TX [27]. For this campaign, the retrievals were taken as part of a special observation mode known as "step-and-stare" [1].

The strategy for nadir retrievals over oceans scenes is first to detect the existence of clouds by comparing the difference in surface brightness temperatures in spectral windows from 867.04 to 897.04 cm^{-1} and from 897.1 to 899.89 cm^{-1} with predicted surface temperature derived from GMAO meterological data. If this difference exceeds a threshold of 6 K then a cloud is presumed present. Clouds are modeled as a frequency-dependent effective optical depth localized at a single pressure level. Over ocean, the effective optical depth spectral distribution, cloud pressure height, calibration scale parameter, temperature, water vapor, and ozone are simultaneously estimated using spectral windows contained in the spectral regions of 990.02–1065.02 cm^{-1} and from 1172.56 to 1317.82 cm^{-1}; the specific microwindows within those regions can be found in [12]. For retrievals over land, spectral surface emissivity and

TABLE II
VALUES OF CONVERGENCE CRITERIA FOR THE FIRST AND LAST ITERATION OF THE SIMULTANEOUS ESTIMATE OF CLOUD, CALIBRATION SCALE, TEMPERATURE, WATER, AND OZONE

iteration	Equation 27	Equation 28	Equation 29
1	37.5	8.0×10^{-5}	34.0
15	0.27	2.2×10^{-8}	1.6×10^{-4}

surface temperature are included. Subsequent to this estimate, carbon monoxide is estimated followed by methane. At each step estimated errors in atmospheric profiles calculated from the previous step are propagated to the error analysis of the current step, e.g., the *a posteriori* error covariance calculated from the retrieval of water vapor is included in the systematic error for water vapor in the error analysis of methane.

The constraint vectors for temperature, water, and surface temperature are computed from interpolating 4 spatial grid points and one time step of GMAO reanalysis. The constraint matrix for temperature and surface temperature is calculated from a MOZART climatology scaled by predicted NCEP errors. The water vapor constraint matrix is based on the altitude-dependent Tikhonov constraint [20]. The constraint vector and matrix for ozone was obtained from the MOZART climatology using the fitting procedure outlined in Section III-B. The carbon monoxide and methane constraint matrices were also calculated from the altitude-dependent Tikhonov constraint while their constraint vectors were calculated from MOZART. For all retrievals, the constraint vector was chosen to be the initial guess for the trust-region LM algorithm.

B. Retrieval

In this section, we will restrict our attention to the simultaneous retrieval of temperature, water, and ozone. The LM algorithm ran for 15 iterations with an initial trust region of $\Delta_0 = 100$. Of those, iterations 3, 5, 8, and 13 were rejected because the cost function increased. Of those steps that were accepted, the median value of the linearity measure ρ [(24)] was .992 which indicates good agreement between the expected decrease in the cost function relative to the actual decrease. The trust region reached a maximum value at $\Delta_{10} = 170$, a minimum value at $\Delta_{14} = 17$ and the final trust region value of $\Delta_{15} = 35$. The limited increase in the value of the trust region indicates the presence of significant nonlinearities in the cost function and a consequent reduction in convergence rate. The values used for the convergence criteria are shown in Table II for the first and last iteration. The tolerance used for the criteria was $\varepsilon = 0.000\,45$. There was a five-orders-of-magnitude reduction in the change in the cost function, a three-orders-of-magnitude reduction in the change of the state vector, and a two-orders-of-magnitude reduction in the change of the cost function derivative. The change in cost function and state vector satisfied the convergence criteria but the cost function derivative did not. The mean and standard deviation of the normalized residuals were 0.0085 and 1.064, respectively, which is close to the expected variability. Thus, there appears to be a reasonable balance between the number of iterations and a strict adherence to the criteria. Nevertheless, further investigation is required to determine the best threshold for an ensemble of retrievals.

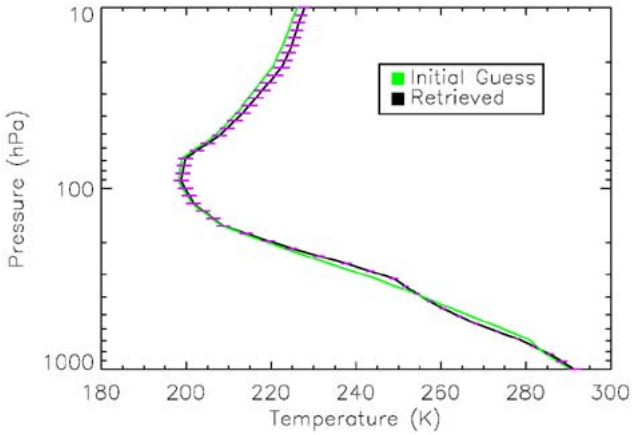


Fig. 1. Estimate of vertical distribution of atmospheric temperature for 25.3° latitude, −95.82° longitude, November 3, 2004. The initial guess was taken as the constraint vector. Error bars are calculated from the square root of the diagonal of the total error covariance matrix.

TABLE III
DIAGNOSTIC VALUES FOR A SUBSET OF RETRIEVED ATMOSPHERIC QUANTITIES

Diagnostic	TSUR	TATM	H_2O	O_3	CO
dofs	0.007	5.6	3.0	3.8	1.2
AvgVertRes (km)	NA	7.8	3.1	8.9	10.9
InfContent	0.007	8.2	5.39	9.06	0.45

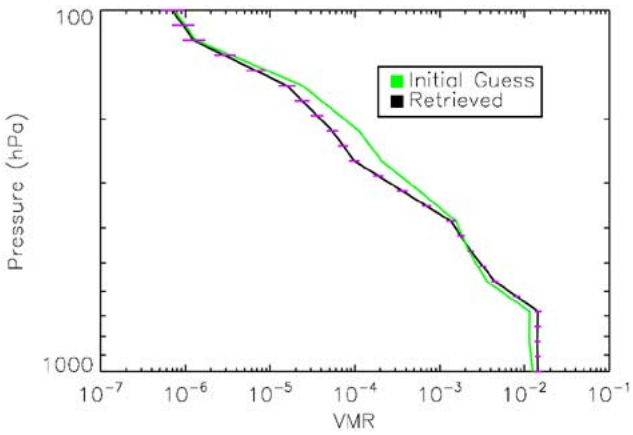


Fig. 2. Estimate of vertical distribution of water vapor for 25.3° latitude, −95.82° longitude, November 3, 2004. The initial guess was taken as the constraint vector. Error bars are calculated from the square root of the diagonal of the total error covariance matrix.

The temperature retrieval is shown in Fig. 1. The error bars listed in the figure are the square-root of diagonal of the total error covariance matrix denoted in (44). The estimate did not vary significantly from the constraint vector for pressures $p \in [50, 150]$ hPa. However, there was significant deviation from the prior for $p \in [200, 700]$ hPa. The estimate cloud top pressure was 779 hPa with an average effective optical depth of 0.78. The effect of the cloud is apparent on the retrieval below 700 mb because the retrieval does not vary from prior and the error increases significantly relative to the error between 200 and 700 hPa. The surface temperature was estimated at 300.82 where the prior is 300.99. Inspection of the DOFS in Table III for surface temperature indicates that the retrieval is not sensitive to the surface because of the cloud opacity. The average

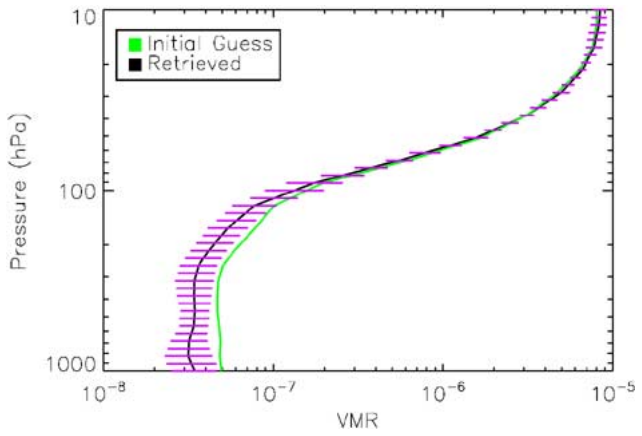


Fig. 3. Estimate of vertical distribution of ozone for 25.3° latitude, −95.82° longitude, November 3, 2004. The initial guess was taken as the constraint vector. Error bars are calculated from the square root of the diagonal of the total error covariance matrix.

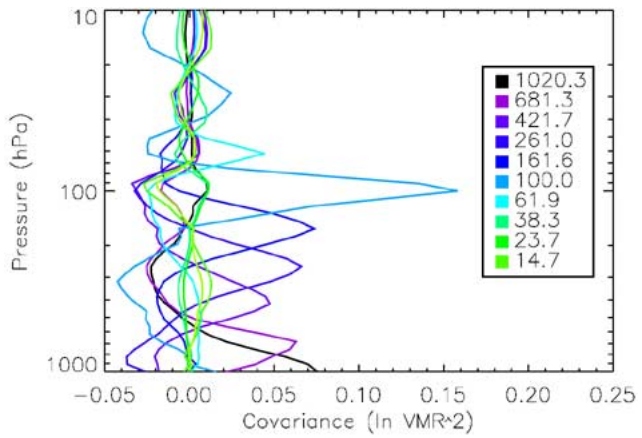


Fig. 4. Error covariance matrix for ozone estimate for 25.3° latitude, −95.82° longitude, November 3, 2004. Each row of the matrix is plotted with its corresponding pressure.

vertical resolution for temperature is 7.8 km. However, this resolution varies considerably as a function of altitude. In the region of $p \in [200, 700]$, for example, the vertical resolution was closer to 2.5 km. From both the reduced error and the increased vertical resolution, we expect the variations from the prior in this pressure region to be reflective of the actual state. The retrieval of water vapor is shown in Fig. 2. The average vertical resolution is 3.1 km and the DOFS is 3.1 with little sensitivity above 150 hPa. The water vapor profile is significantly more dry for $p \in [400, 150]$ hPa and more moist in $p \in [1000, 400]$ than the prior.

The estimate ozone profile is shown in Fig. 3 with an average vertical resolution of 9.4 km and a DOFS of 3.3. This vertical resolution is less than the predicted vertical resolution for midlatitudes described in [12]. The difference between these resolutions is due to the simultaneous estimate of ozone and temperature using the 10-μm band, the presence of clouds, and the retrieval of a calibration scaling factor. Future retrievals may no longer need the calibration scale factor and consequently could use the 6-μm CO_2 band for the temperature retrieval, which should yield vertical resolutions consistent with [12].

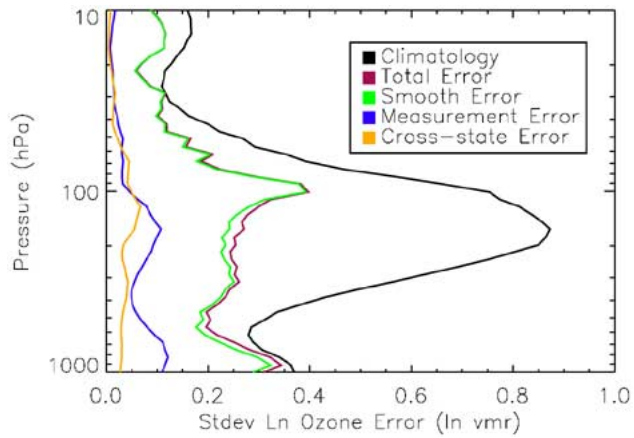


Fig. 5. Diagonals of the total, smoothing, random, and cross-state error covariance matrices for the ozone estimate.

The retrieval is within the error bars from $p \in [100, 10]$ hPa. Below 100 hPa, the ozone is uniformly less than the prior. Given the discretization of the full state vector at approximately 1 km/level and a vertical resolution of 9.4 km, the correlations between state vector elements will be significant as shown in the error covariance in Fig. 4. The peak error is approximately 0.15 at 100 hPa. This number corresponds to the variance in the logarithm of ozone. From the approximation $\ln \delta x \approx \delta x/x$, this error can be interpreted as fractional. Consequently, the standard deviation at 100 hPa is approximately 39%. As expected the error is maximal at the pressure level of the thermal tropopause. Furthermore, this error is strongly correlated with the error at 30 hPa and anticorrelated with the error at 300 hPa. The mean percentage error below the thermal tropopause is approximately 25% with a correlation length of roughly 6 km. Above the thermal tropopause to 10 hPa, the percentage error decreases to 7.3% at 23.7 hPa, consistent with the increase in temperature and a two-orders-of-magnitude increase in volume mixing ratio as well as an order-of-magnitude increase in number density. However, the distribution of the error can not be understood without considering the error terms in (44) that make up the total error. The standard deviation of the total, smoothing, random, and cross-state error are shown in Fig. 5. The total error is dominated by the smoothing error throughout the ozone profile. Consequently, the reduction in error near 20 hPa is also due to the reduction in climatological variability of ozone. Nevertheless, to the extent that the *a priori* covariance matrix, the diagonal of which is shown in Fig. 5, is an accurate representation of the statistical variability of ozone at that location, the balance between the smoothing, random, and cross-state errors are optimal. The greatest reduction in error is in the upper troposphere near 160 hPa where the percentage variability of 87% has been reduced to 27%. The overall reduction in error before and after the retrieval is quantified by the information content, which was defined in (2). This metric is based on the ratio of the error volume of the natural variability, which is described by the *a priori* covariance matrix, to the *a posteriori* error volume, which is described by the total error covariance matrix. From Table III, the information content for the ozone retrieval is 9.02 bits, which indicates an overall reduction in error volume.

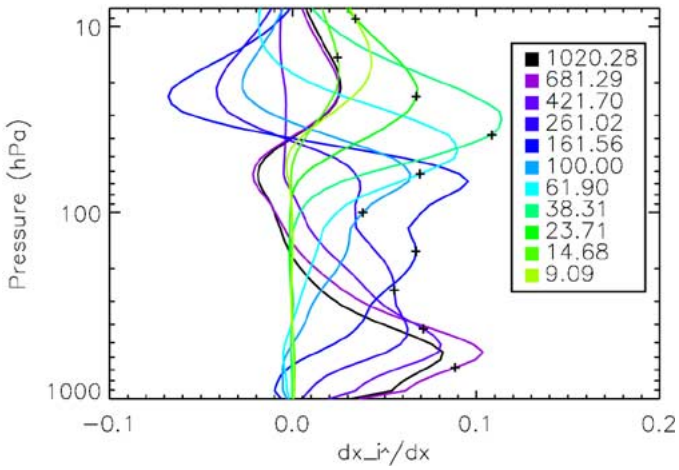


Fig. 6. O_3 averaging kernel matrix. The rows of this matrix, which are also called the averaging kernel are plotted for different pressure levels of the atmospheric state. The black cross-hairs are the pressure levels for the corresponding to each averaging kernel.

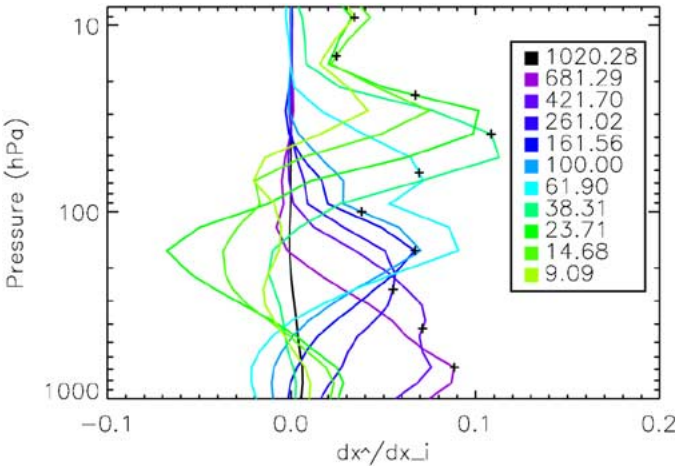


Fig. 7. O_3 averaging kernel matrix. The columns of this matrix, which are also called the delta-response functions, are plotted for different pressure levels of the atmospheric state. The black cross-hairs are the pressure levels corresponding to each delta-response function.

The dominance of the smoothing error indicates that the error in the estimate is limited primarily by the vertical resolution of the retrieval. The averaging kernels, which are shown in Fig. 6, can be defined as the change in an estimated state element given a change in the true state, i.e., $\partial[\hat{x}]_i/\partial\mathbf{x}$. The width of these averaging kernels are used to define the vertical resolution and are listed in Table III. For ozone the average is 9.4 km. The columns of the averaging kernel matrix, which are also called the δ-response functions, are shown in Fig. 7 and can be equally useful in interpreting a retrieval. These functions define the change in the estimated to state in response to a perturbation of the true state at a specific pressure level. Near 1020 hPa, the corresponding delta response function is less than .01, indicating little sensitivity at the surface due to the presence of moderately thick clouds. In the troposphere up to the thermal tropopause ($\approx$ 16 km), the δ-response functions are broadly unimodal. The response function at 62 hPa is particularly broad with a width at half-power of roughly 12 km with the dip in the response corresponding to the thermal tropopause. The response functions above the thermal tropopause show the influence on lower stratospheric ozone on the upper tropospheric ozone estimate. For the δ-response at 23.71 hPa, the fractional change in the estimate state is 0.11 at 31.6 hPa whereas there is a -0.067 fractional change in the estimate state at 161 hPa. This magnitude of this change is equal to the fractional change in the δ-response at 161 hPa. Consequently, the ozone estimate at 161 hPa is equally influenced by the true state at 161 hPa and 23.71 hPa but in opposite sign. This conclusion can be reached equally by examining the averaging kernel at 161 hPa, which additionally shows that the maximal contribution to the estimate comes from around 70 hPa. While the contribution of the ozone at multiple altitudes to the estimate of ozone at a particular altitude may make it difficult to interpret the ozone retrieval, these contributions can be explicitly addressed in formal comparisons with models or other measurements [28], [29].

VII. Conclusion

The methodology used for TES retrievals is based on a Bayesian framework for which the estimation of the atmospheric state is obtained through the solution of a constrained nonlinear least squares problem. A stable and unique solution is dependent on the choice of regularization and a robust minimization algorithm. The regularization is latitude-dependent and can be a combination of altitude-dependent Tikhonov or climatological constraint matrices as well as a restriction of the solution to a coarser vertical grid through mapping or "hard" constraints. The structured least squares trust-region algorithm combines the higher convergence rates of a Gauss–Newton solver with the robust conjugate gradient algorithm in a way that checks the linearity at each step. This algorithm has worked effectively with state vectors containing hundreds of elements.

Characterization of TES retrievals includes the quantification of smoothing, random, "cross-state," and systematic errors. Furthermore, the effect of constraint matrices, constraint vectors, and mapping on the estimate are characterized through the averaging kernel matrix. Given the dimensionality of these error covariances and averaging kernels, a number of scalar retrieval metrics were introduced such as information content, DOFS, and vertical resolution.

These techniques were applied to the analysis of a single TES nadir scene taken in the Gulf of Mexico in November 2004. The fractional error in the tropospheric ozone retrieval was substantially reduced relative to the *a priori*, particular in the upper troposphere. Moreover, this reduction is possible even in the presence of clouds in the field-of-view. However, analysis of the ozone averaging kernels suggest special care must be taken to interpret the upper tropospheric ozone profiles due to the influence of lower stratospheric ozone.

References

[1] R. Beer, T. Glavich, and D. Rider, "Tropospheric Emission Spectrometer for the Earth Observing System's Aura satellite," *Appl. Opt.*, vol. 40, pp. 2356–2367, 2001.

[2] C. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*. Singapore: World Scientific, 2000.

[3] S. Clough, C. Rinsland, and P. Brown, "Retrieval of tropospheric ozone from simulations of nadir spectral radiances as observed from space," *J. Geophys. Res.—Atmos.*, vol. 100, no. D8, pp. 16 579–16 593, 1995.

[4] S. Clough, J. Worden, P. Brown, C. Rinsland, and R. Beer, "Retrieval of tropospheric ozone from simulations of limb spectral radiances as observed from space," *J. Geophys. Res.*, vol. 107, no. D21, p. 4589, 2002.
[5] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.
[6] S. Clough, M. Shepard, J. R. Worden, P. D. Brown, H. M. Worden, M. Lou, C. Rodgers, C. Rinsland, A. Goldman, L. Brown, A. Eldering, S. S. Kulawik, K. Cady-Pereira, G. Osterman, and R. Beer, "Forward model and Jacobians for Tropospheric Emission Spectrometer retrievals," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1308–1323, May 2006.
[7] R. Norton and R. Beer, "New apodizing functions for Fourier spectrometry," *J. Opt. Soc. Amer.*, vol. 66, no. 259, pp. 259–263, 1976.
[8] ——, "Correction to new apodizing functions for Fourier spectrometry," *J. Opt. Soc. Amer.*, vol. 67, no. 419, pp. 145–149, 1977.
[9] C. Rodgers, "Information content and optimization of high spectral resolution measurements," in *Proc. SPIE Conf. Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II*, vol. 2830, P. B. Hays and J. Wang, Eds., 1996, pp. 136–147.
[10] T.v. Clarmann and G. Echle, "Selection of optimized microwindows for atmospheric spectroscopy," *Appl. Opt.*, vol. 37, no. 33, pp. 7661–7669, 1998.
[11] A. Dudhia, V. Jay, and C. Rodgers, "Microwindow selection for high-spectral-resolution sounders," *Appl. Opt.*, vol. 41, no. 18, pp. 3665–3673, 2002.
[12] J. Worden, S. S. Kulawik, M. Shepard, S. Clough, H. Worden, K. Bowman, and A. Goldman, "Predicted errors of Tropospheric Emission Spectrometer nadir retrievals from spectral window selection," *J. Geophys. Res.*, vol. 109, no. D09308, May 2004.
[13] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana: Univ. Illinois Press, 1949.
[14] H. Worden, R. Beer, K. Bowman, B. Fisher, M. Lou, E. Sarkissian, D. Tremblay, and J. Zong, "TES level1 algorithms: Interferogram processing, geolocation, radiometric, and spectral calibration," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1288–1296, May 2006.
[15] K. Bowman, J. Worden, T. Steck, H. Worden, S. Clough, and C. Rodgers, "Capturing time and vertical variability of tropospheric ozone: A study using TES nadir retrievals," *J. Geophys. Res.*, vol. 107, no. D23, p. 4723, 2002.
[16] G. Brasseur, D. Hauglustaine, S. Walters, P. Rasch, J. Muller, C. Granier, and X. Tie, "MOZART, A global chemical transport model for ozone and related chemical tracers 1. Model description," *J. Geophys. Res.—Atmos.*, vol. 103, pp. 28 265–28 289, 1998.
[17] S. Twomey, "On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature," *J. ACM*, vol. 10, pp. 97–101, 1963.
[18] T. Steck, "Methods for determining regularization for atmospheric retrieval problems," *Appl. Opt.*, vol. 41, no. 9, pp. 1788–1797, 2002.
[19] T. Steck and T. Clarmann, "Constrained profile retrieval applied to the observation mode of the Michelson interferometer for passive atmospheric sounding," *Appl. Opt.*, vol. 40, no. 21, pp. 3559–3571, 2001.
[20] S. S. Kulawik, G. Osterman, D. B. A. Jones, and K. W. Bowman, "Calculation of altitude-dependent Tikhonov constraints for TES nadir retrievals," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1334–1342, May 2006.
[21] R. Fletcher, *Practical Methods of Optimization*. New York: Wiley, 1987.
[22] A. Björck, *Numerical Methods for Least Squares Problems*. Philadelphia, PA: SIAM, 1996.
[23] E. Sarkissian, "The Levenberg–Marquardt algorithm for solving the nonlinear least squares problem: Theory, implementation and application," M.S. thesis, California State Univ., Los Angeles, 2001.
[24] J. J. Moré, "The Levenberg–Marquardt algorithm: Implementation and theory," in *Proc. Biennial Conf. Numerical Analysis*, 1977, pp. 105–116.
[25] R. Aster, B. Borchers, and C. Thurber, *Parameter Estimation and Inverse Problems*. New York: Academic, 2004.
[26] S. Nash and A. Sofer, *Linear and Nonlinear Programming*. New York: McGraw-Hill, 1996.
[27] (2005, Mar.) Aura Validation Experiment (AVE) Campaign. [Online]. Available: http://cloud1.arc.nasa.gov/ave-houston/index.cgi
[28] C. Rodgers and B. Connor, "Intercomparison of remote sounding instruments," *J. Geophys. Res.*, vol. 108, no. D3, p. 4119, 2003.
[29] D. B. A. Jones, K. W. Bowman, P. I. Palmer, J. R. Worden, D. J. Jacob, R. N. Hoffman, I. Bey, and R. M. Yantosca, "Potential of observations from the Tropospheric Emission Spectrometer to constrain continental sources of carbon monoxide," *J. Geophys. Res.—Atmos.*, vol. 108, no. D24, p. 4789, 2003.

Kevin W. Bowman received the B.E.E. degree in electrical engineering from Auburn University, Auburn, AL, the Diplome de Specialization en Traitement et Transmission des Informations from the Ecole Superieure d'Electricite (SUPELEC), Metz, France, the M.S. and Ph.D. in electrical engineering at the Georgia Institute of Technology, Atlanta, in 1991, 1992, 1993, and 1997, respectively.

He has been a Member of the Technical Staff at the Jet Propulsion Laboratory, Pasadena, CA, since 1997. His primary research interests are in the area of atmospheric remote sensing and tropospheric chemistry including inverse modeling and data assimilation from satellite observations, estimation and error analysis of trace gas profiles, and calibration algorithms for infrared Fourier transform spectrometers.

Clive D. Rodgers received the B.A., M.A., and Ph.D. degrees in physics from Cambridge University, Cambridge, U.K., in 1960, 1963, and 1964, respectively.

He was a University Lecturer in the Department of Physics at the University of Oxford from 1967 until his retirement in 2001. He has continued his research as a Senior Research Fellow at Jesus College, University of Oxford. His main research interests include radiative transfer, molecular spectroscopy, the application of estimation theory to the inverse problem of radiative transfer, and the application and assimilation of satellite data to stratospheric and mesospheric dynamics and chemistry. He is the author of *Inverse Methods for Atmospheric Sounding: Theory and Practice* (World Scientific, 2001).

Susan Sund Kulawik received the B.S. degree from The Ohio State University, Columbus, in 1991 and the Ph.D. degree in physics from the University of Michigan, Ann Arbor, in 1999.

She is a member of the Tropospheric Emission Spectrometer (TES) Science Team and has worked on TES since 1999, first at Raytheon ITSS in Pasadena, CA, and at the Jet Propulsion Laboratory, Pasadena, since 2002. Her research interests include the use of information theory to select constraints and spectral windows, error analysis, retrievals in the presence of clouds, and HNO3 retrievals from TES data.

John Worden received the B.S. degree in physics from the University of California, Santa Cruz, in 1990 and the Ph.D. degree in physics from the University of Colorado, Boulder, in 1996.

He has worked on the Tropospheric Emission Spectrometer (TES) since 1999, as a member of the TES science team at the Jet Propulsion Laboratory, Pasadena, CA (since 2002), at Raytheon ITSS in Pasadena, CA (2000–2002), and at Atmospheric and Environmental Research, Cambridge, MA (1999–2000). His research interests include retrievals of atmospheric composition, radiative transfer through the atmosphere, understanding the water cycle using measurements of water and its isotopes, and atmospheric pollution sources and sinks.

Edwin Sarkissian received the B.S. degree in computer science and the M.S. degree in applied mathematics from California State University, Los Angeles.

He joined the Jet Propulsion Laboratory, Pasadena, CA, in 1995. He is the Technical Lead for the Level 1B software of the Tropospheric Emission Spectrometer (TES) on the Earth Observing System Aura satellite launched on July 15, 2004.

Gregory Osterman received the B.S. degree in physics from Texas A&M University, College Station, in 1987, the M.S. degree in physics from Texas Tech University, Lubbock, and the Ph.D. degree in physics from the University of Texas at Dallas, Richardson, in 1994.

He is a member of the Tropospheric Emission Spectrometer (TES) Science Team, Jet Propulsion Laboratory, Pasadena, CA, and focuses primarily on validation of TES data. His research interests also include studying tropospheric ozone using TES and other data sources as well as models.

Tilman Steck received the Diploma in physics from the University of Tübingen, Tübingen, Germany, in 1996 and the Ph.D. degree in physics from the University of Karlsruhe (TH), Karlsruhe, Germany, in 2000.

He has worked in the area of atmospheric remote sensing since 1996, especially in the development of retrieval algorithms and selection of constraints for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on Envisat. He held a postdoctoral position at the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, during 2000 and 2001, on the TES Science Team.

Ming Luo received the Ph.D. degree in atmosphere and space sciences from the University of Michigan, Ann Arbor, in 1991.

She is a member of the Tropospheric Emission Spectrometer (TES) Science Team, Jet Propulsion Laboratory, Pasadena, CA, and works on L2 forward model and retrieval algorithms, data calibration and validation, and L3 algorithms including visualization of TES data.

Annmarie Eldering received the B.E. degree in chemical engineering from Cooper Union, New York, in 1988 and the Ph.D. degree in environmental engineering from the California Institute of Technology, Pasadena, in 1994.

She is currently a Research Scientist at the Jet Propulsion Laboratory, Pasadena, CA. Her research interests include cloud characterization from remote sensing measurements of clouds and the impact of clouds and water vapor on climate.

Mark Shephard received the B.S degree in physics from the University of Guelph, Guelph, ON, Canada, in 1995, and the M.S. degree in atmospheric and oceanic science from McGill University, Montreal, QC, Canada, in 1999.

He is currently working in the Radiation and Climate Group at Atmospheric and Environmental Research (AER), Inc., Lexington, MA. His general research interests include satellite remote sensing and atmospheric radiative transfer. He is a member of the Tropospheric Emission Spectrometer (TES) Science Team, and his current activities include the further development and validation of radiative transfer models (i.e., LBLRTM), spectroscopic line parameters, and atmospheric constituent retrievals.

Helen Worden received the Ph.D. degree in particle physics from Cornell University, Ithaca, NY, in 1991.

She has worked at the Jet Propulsion Laboratory since 1993, where she is a member of the Tropospheric Emission Spectrometer (TES) Science Team and works on calibration and retrieval algorithms.

Michael Lampel received the B.A. degree in physics from the College of Creative Studies, University of California, Santa Barbara in 1976, and the M.A. and Ph.D. degrees in physics from the University of California, Berkeley, in 1978 and 1984, respectively.

He is the Science Lead for Raytheon Information Solutions Science Systems and Analysis Department, Pasadena, CA, and has worked on the Tropospheric Emission Spectrometer (TES) since 1998. His primary work has been in applied physics and modeling: medical imaging, free electron lasers, particle beams, radiative transfer, and atmospheric remote sensing.

Shepard Clough received the B.Eng. degree from Cornell University, Ithaca, NY, and the M.A. degree from Columbia University, New York, in 1954 and 1958, respectively, both in physics.

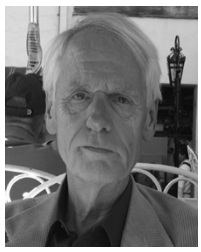
He is currently with Atmospheric and Environmental Research (AER), Inc., Lexington, MA. His principle areas of interest include atmospheric radiative transfer, molecular physics, and approaches to the retrieval of information from remotely sensed measurements. His current activities are focused on the improvement of radiative transfer modeling for general circulation models with application to climate change studies. He is a member of the Science Team on the Atmospheric Radiation Measurement (ARM) program and a Co-Investigator on the Tropospheric Emission Spectrometer (TES). He has been responsible for the development and the widely used atmospheric and radiative modeling codes including LBLRTM, RRTM, and CHARTS, and has made significant contributions to the HITRAN spectroscopic database.

Mr. Clough is currently a member of the International Radiation Commission.

Pat Brown, photograph and biography not available at time of publication.

Curtis Rinsland, photograph and biography not available at time of publication.

Michael Gunson, photograph and biography not available at time of publication.



Reinhard Beer received the Ph.D. degree in physics from the University of Manchester, Manchester, U.K., in 1960.

He joined the Jet Propulsion Laboratory, Pasadena, CA, in 1963, where he is the Principal Investigator for the Tropospheric Emission Spectrometer on the Earth Observing System Aura satellite launched on July 15, 2004.