Microwindow selection for high-spectral-resolution sounders

Anu Dudhia, Victoria L. Jay, and Clive D. Rodgers

The recent development of satellite instruments that obtain spectrally resolved measurements of the atmosphere has highlighted the problem of how to determine the best subsets, or microwindows, of such spectra for retrievals of temperature and composition. A technique is described that maximizes the information content (or some other figure of merit) based on the modeling of the propagation of systematic as well as random error terms through the retrieval process. Apart from selecting microwindows, this technique can also prioritize existing microwindows for different circumstances and provides a full error analysis of the retrieval. A practical application is demonstrated for the Michelson Interferometer for Passive Atmospheric Sounding limb-viewing interferometer, but the technique is equally applicable to nadir-viewing instruments.

OCIS codes: 000.4430, 010.1280, 280.0280, 300.6340.

1. Introduction

Several current generation of Earth-observing satellite instruments are designed to make continuous measurements of atmospheric IR spectra using either grating spectrometers (Atmospheric Infrared Sounder[1]) or interferometers [Interferometric Monitor for Greenhouse Gases,[2] Michelson Interferometer for Passive Atmospheric Sounding[3] (MIPAS), Infrared Atmospheric Sounding Interferometer,[4] and Tropospheric Emission Spectrometer[5]]. Such spectra typically provide thousands of radiance measurements every second, too many to incorporate into a real-time retrieval scheme, given present computing speeds. Consequently, attention has been focused on methods for the determination of the optimum subsets of such spectra (microwindows) that contain most of the potential information.

One approach is to simulate the propagation of random noise through the retrieval and select measurements that maximize the information content or degrees of freedom of the signal,[6,7] or that best satisfy other criteria.[8] Although this is reasonable if the retrieval errors are predominantly due to random measurement errors and/or errors in the *a priori* estimate, it does not allow for other (systematic) sources of error associated with unretrieved parameters. For example, it does not allow for the uncertainties in the modeling of the concentrations of the various contaminant species, that is, other molecules whose spectral signatures overlap those of the retrieved species but whose concentrations are not themselves retrieved.

An alternative approach addresses this problem by selecting microwindows that minimize the total error, including both random and systematic components.[9,10] However, this is achieved only through an approximation of the profile retrieval as a set of independent, single-layer retrievals. As a result, the effect of interlevel correlations in the actual retrieval is ignored. There is also the practical difficulty of the consolidation of the microwindows derived independently for each profile level into single microwindows applicable over the whole profile.

In this paper a scheme is presented that selects microwindows by simulating a full-profile retrieval, including the propagation of systematic error terms. As such, it can be considered either as an extension of the first (random error) method to include systematic errors, or as a multilayer version of the second (single-layer) method.

As well as selecting microwindows, this scheme also allows existing microwindows to be ordered ac-

When this research was performed, A. Dudhia (dudhia@atm.ox.ac.uk), V. L. Jay, and C. D. Rodgers were with Atmospheric, Oceanic, and Planetary Physics, Department of Physics, Oxford University, Parks Road, Oxford OX1 3PU, United Kingdom. V. L. Jay is now with the Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, United Kingdom.

Received 6 August 2001; revised manuscript received 30 January 2002.

0003-6935/02/183665-09$15.00/0

cording to different criteria and provides a complete error analysis of the resulting retrieval products.

As a practical demonstration, microwindows are selected for the retrieval of a methane profile from MIPAS measurements. Although MIPAS is a limb-viewing instrument, the technique is equally applicable to nadir-viewing geometries.

2. Theory

A. Optimal Estimation

Optimal estimation theory (e.g., Ref. 11) provides a linearized form for an estimate $\mathbf{x}$ of the state vector (atmospheric profile) that is based on a combination of a set of measurements $\mathbf{y}$ and a prior estimate $\mathbf{a}$ of the state

$$\mathbf{x} = \mathbf{G}\mathbf{y} + (\mathbf{I} - \mathbf{G}\mathbf{K})\mathbf{a}, \tag{1}$$

where $\mathbf{K}$ is the Jacobian (or weighting function) matrix ($K_{jl} = \partial y_j/\partial x_l$) and $\mathbf{G}$ is the gain (or contribution function) matrix given by

$$\mathbf{G} = \mathbf{S}_a\mathbf{K}^T(\mathbf{S}_y + \mathbf{K}\mathbf{S}_a\mathbf{K}^T)^{-1}. \tag{2}$$

$\mathbf{S}_a$ is the covariance of $\mathbf{a}$ about the true state, and $\mathbf{S}_y$ the covariance of $\mathbf{y}$ about the perfect measurements that would arise from this true state.

We obtain the covariance of the retrieval error by taking the covariance of Eq. (1) about the true state:

$$\mathbf{S}_x = \mathbf{G}\mathbf{S}_y\mathbf{G}^T + (\mathbf{I} - \mathbf{G}\mathbf{K})\mathbf{S}_a(\mathbf{I} - \mathbf{G}\mathbf{K})^T \tag{3}$$

$$= \mathbf{S}_a - \mathbf{S}_a\mathbf{K}^T(\mathbf{S}_y + \mathbf{K}\mathbf{S}_a\mathbf{K}^T)^{-1}\mathbf{K}\mathbf{S}_a \tag{4}$$

$$= (\mathbf{I} - \mathbf{G}\mathbf{K})\mathbf{S}_a. \tag{5}$$

For a set of m measurements, Eq. (2) requires the inversion of a matrix of size $m \times m$, which becomes impractical for large m. If $\mathbf{S}_y$ on its own can be inverted simply, a standard approach is to reformulate Eq. (2) as

$$\mathbf{G} = (\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K} + \mathbf{S}_a^{-1})^{-1}\mathbf{K}^T\mathbf{S}_y^{-1}. \tag{6}$$

However, if $\mathbf{S}_y$ is diagonal, an even more convenient solution is available, outlined in the next section.

B. Sequential Estimation

If the measurement errors are uncorrelated, i.e., $\mathbf{S}_y$ is diagonal (variances $S_{jj} \equiv \sigma_j^2$), Eqs. (1)–(5) can be applied to the individual measurements y_j, $j = 1 \ldots m$ sequentially:

$$\mathbf{x}^j = \mathbf{g}^j y_j + (\mathbf{I} - \mathbf{g}^j\mathbf{k}_j^T)\mathbf{a}^j, \tag{7}$$

$$\mathbf{g}^j = \mathbf{S}_a^j\mathbf{k}_j(\sigma_j^2 + \mathbf{k}_j^T\mathbf{S}_a^j\mathbf{k}_j)^{-1}, \tag{8}$$

$$\mathbf{S}_x^j = (\mathbf{I} - \mathbf{g}^j\mathbf{k}_j^T)\mathbf{S}_a^j, \tag{9}$$

where $\mathbf{k}_j^T$ is the row of the Jacobian matrix associated with y_j; superscript j denotes quantities progressively modified for $j = 1 \ldots m$ measurements, and $\mathbf{a}^j$, $\mathbf{S}_a^j \equiv \mathbf{x}^{j-1}$, $\mathbf{S}_x^{j-1}$. The original, independent *a priori* information is used to set $\mathbf{a}^1$, $\mathbf{S}_a^1$, but thereafter the prior estimate also incorporates information from previous measurements. We will use the terms "*a priori*" and "previous" estimates to distinguish the two cases.

This replaces the inversion of the $m \times m$ matrix in Eq. (2) with m inversions of a scalar quantity in Eq. (8). It is, of course, possible to apply sequential estimation to larger subsets of $\mathbf{y}$ than to individual measurements, although less computationally efficient. The replacement of the variance σ_j^2 with the equivalent subset of the full covariance matrix $\mathbf{S}_y$ allows for error correlations within each subset. In Subsection 2.G a hybrid approach is described that deals with apodized measurements, initially treating them as single uncorrelated measurements but including the correlations when established as part of a microwindow.

C. Forward-Model Errors

In reality, the matrix $\mathbf{S}_y$ represents not only the covariance $\mathbf{S}_\epsilon$ of the difference between the real and the perfect measurements, but also the covariance $\mathbf{S}_f$ of the difference between the computational forward model used in the retrieval and the physical radiative transfer process

$$\mathbf{S}_y = \mathbf{S}_\epsilon + \mathbf{S}_f. \tag{10}$$

While the measurement component of the error may be dominated by random noise ($\mathbf{S}_\epsilon \sim$ diagonal), forward-model errors, e.g., the effect of underestimating a contaminant, tend to be correlated. For convenience, $\mathbf{S}_\epsilon$ will be used to refer to just the random or locally correlated (e.g., apodized noise) component of the measurement error, whereas more widely correlated components (e.g., calibration errors) will be considered part of $\mathbf{S}_f$.

Practical retrieval schemes, for various reasons, often assume $\mathbf{S}_y \equiv \mathbf{S}_\epsilon$ so that the weight assigned to each measurement in the gain matrix [Eq. (2)] depends only on the random component. Even if $\mathbf{S}_f$ is not negligible, this substitution can still be used within the gain matrix (although, technically, the retrieval is no longer optimal estimation[12]); however, the true covariance of such a retrieval [obtained with substitution of Eq. (10) in Eq. (3)] has an additional term $\mathbf{G}\mathbf{S}_f\mathbf{G}^T$ on the right hand side of Eq. (5). While mathematically simple, this extra term, representing the propagation of correlated measurements errors through the retrieval, can be too large a matrix ($m \times m$) to be evaluated directly, and too widely correlated to be decomposed into block diagonals and evaluated sequentially. In the next section, $\mathbf{S}_f$ is mapped into a form that can be propagated through the sequential estimator.

D. Systematic Errors

Systematic errors are defined here as errors that are correlated between elements of the prior estimate $\mathbf{a}$ and the measurement vector $\mathbf{y}$, and that therefore are not described by either of the associated covariance matrices $\mathbf{S}_a$, $\mathbf{S}_y$. Normally such errors are neglected either because the prior estimate is assumed to be independent *a priori* information, as in optimal esti-

mation, or because measurement errors are assumed random, as in sequential estimation. However, for the purposes of microwindow selection, we require a scheme for handling systematic errors within the framework of sequential estimation.

In addition to the random error, a set of m measurements $\mathbf{y}$ will also contain error components $\delta\mathbf{y}^i$ due to a variety of error sources i. The magnitude of the error vector $\delta\mathbf{y}^i$ corresponds to a 1σ perturbation in the error source, and the shape is characteristic of the particular error source. Assuming that all of these errors are independent, the total error covariance of $\mathbf{y}$ [see Eq. (10)] becomes

$$\mathbf{S}_y^{\text{tot}} = \mathbf{S}_y^{\text{rnd}} + \sum_i E\{(\delta\mathbf{y}^i)(\delta\mathbf{y}^i)^T\} \tag{11}$$

$$= \mathbf{S}_y^{\text{rnd}} + \sum_i \mathbf{S}_y^i, \tag{12}$$

where $E\{\ldots\}$ denotes expectation value. In practice, each vector $\delta\mathbf{y}^i$ can be computed as the difference between the nominal set of measurements and a set where the error source is perturbed by 1σ, and the corresponding covariance is then given directly by $(\delta\mathbf{y}^i)(\delta\mathbf{y}^i)^T$. Thus the correlation information in the $m \times m$ total covariance can be decomposed into a set of error vectors $\delta\mathbf{y}^i$, each of length m.

Equation (1) describes the mapping of the measurement vector into the retrieval, so the measurement error vectors $\delta\mathbf{y}^i$ will map into corresponding retrieval error vectors $\delta\mathbf{x}^i$ in the same way:

$$\delta\mathbf{x}^i = \mathbf{G}\delta\mathbf{y}^i + (\mathbf{I} - \mathbf{GK})\delta\mathbf{a}^i, \tag{13}$$

where $\delta\mathbf{a}^i$ allows for any equivalent systematic error components in the prior estimate.

The retrieval total error covariance $\mathbf{S}_x^{\text{tot}}$ is then given by

$$\mathbf{S}_x^{\text{tot}} = \mathbf{S}_x^{\text{rnd}} + \mathbf{S}_x^{\text{sys}}, \tag{14}$$

where

$$\mathbf{S}_x^{\text{rnd}} = (\mathbf{I} - \mathbf{GK})\mathbf{S}_a^{\text{rnd}}, \tag{15}$$

$$\mathbf{S}_x^{\text{sys}} = \sum_i E\{(\delta\mathbf{x}^i)(\delta\mathbf{x}^i)^T\} \equiv \sum_i \mathbf{S}_x^i. \tag{16}$$

The same total error covariance could be obtained if *all* of the measurements are incorporated simultaneously with the full-error covariance matrix for $\mathbf{S}_y$ in Eq. (3) (e.g., by adding $\mathbf{GS}_f\mathbf{G}^T$), but the advantage of propagating the systematic errors as a set of vectors rather than as a covariance matrix is that no information is lost when sequential estimation is applied.

As noted for the single-layer approach[9] Eq. (16) strictly only applies to so-called forward model parameter errors, which are expected to introduce zero bias when averaged over a large number of retrievals. Errors in the formulation of the forward model itself, typically due to neglected physics, may lead to known and persistent retrieval biases. Consequently these are not independent of each other, and all such terms $\delta\mathbf{x}^i$ should be first added together, and the expectation value should be calculated for their sum.

To put the two previous methods of microwindow selection into the context of Eqs. (14)–(16), the random error approach considers only Eq. (15), while the single-layer approach treats $\mathbf{x}$ as a scalar so the $\mathbf{S}_x$ covariance matrices become simple variances (actually 2×2 matrices, since a background continuum-like term is retrieved for each microwindow as well as the target species).

E. Figure of Merit

To compare covariances it is necessary to define a scalar quantity, the Figure of Merit, for the retrieval.

The Shannon information content is one measure that can be used. The information content H, measured in bits, is defined by

$$H = -\log_2(F_x/F_a), \tag{17}$$

where $F_x = |\mathbf{S}_x^{\text{tot}}|$, the determinant of the retrieval error covariance, and $F_a = |\mathbf{S}_a^{\text{tot}}|$, the determinant of the *a priori* error covariance. For example, if the covariance matrices were diagonal, 1 bit of information would correspond to a factor-4 reduction in the variance of one element (or a factor-2 reduction in σ). If the retrieval is regarded as a sequential combination of j measurements (either singly or as microwindows), then an increment ΔH_j can be assigned to each:

$$\Delta H_j = -\log_2(F_j/F_{j-1}), \tag{18}$$

where F_{j-1}, F_j are the determinants of the retrieval covariance before and after the jth measurements are included, and $F_0 \equiv F_a$.

Although information content seems a natural choice, there are other alternatives. The formulation of Eqs. (17) and (18) can still be used, but the (scalar) function F can be redefined according to priorities. Some examples are listed below.

Weighting against systematic errors

$$F = |\mathbf{S}^{\text{rnd}} + \alpha\mathbf{S}^{\text{sys}}|, \tag{19}$$

where $\alpha > 1$. This favors measurements that introduce relatively low systematic errors compared with random errors.

Weighting toward an accuracy requirement

$$F = |\mathbf{S}^{\text{tot}} + \mathbf{S}^{\text{req}}|, \tag{20}$$

where the requirement covariance $\mathbf{S}^{\text{req}}$ contains target accuracies for each profile level. This favors measurements that improve the retrieval where the requirement accuracy is not yet met.

Weighting to minimize variances

$$F = \prod_k S_{kk}^{\text{tot}}, \tag{21}$$

equivalent to evaluating the determinant ignoring off-diagonal elements. This favors measurements that maximize the accuracy of each individually re-

trieved point, which might be preferred if the end user cannot use the covariance information.

Weighting with a CPU cost

$$F = |\mathbf{S}^{\text{tot}}|C^n, \tag{22}$$

where C is the estimate of the processing time required and n is some power controlling its relative weight. This favors measurements that produce more information for a given factor increase in CPU time. An example of the influence of CPU cost on microwindow selection is given in Subsection 4.C.

F. Selecting Individual Measurements

Having established the theoretical background, we now detail the procedure for selecting individual measurements for a retrieval.

First, error vectors $\delta\mathbf{y}^i$ are precomputed for every relevant source i for the entire measurement domain. Also required are the Jacobians $\mathbf{k}^T$ for each measurement: Each is a row vector with length matching that of the retrieved state vector $\mathbf{x}$.

A suitable *a priori* covariance $\mathbf{S}_a$ is established, corresponding to any initial knowledge of the elements of $\mathbf{x}$ to be retrieved, such as the covariance of a numerical model or climatological uncertainties. For retrievals that use no explicit *a priori* information, diagonal elements of $\mathbf{S}_a$ can be set to values that are large values compared with the expected retrieval accuracy. Assuming that the *a priori* estimate contains no correlations with measurement errors, $\delta\mathbf{a}^i$ can be initialized to zero for each error source.

Then, for each measurement y, the scalar versions of Eqs. (13) and (14) are computed:

$$\mathbf{g} = \mathbf{S}_a^{\text{rnd}}\mathbf{k}(\sigma_y{}^2 + \mathbf{k}^T\mathbf{S}_a^{\text{rnd}}\mathbf{k})^{-1}, \tag{23}$$

$$\mathbf{S}_x^{\text{rnd}} = (\mathbf{I} - \mathbf{g}\mathbf{k}^T)\mathbf{S}_a^{\text{rnd}}, \tag{24}$$

$$\delta\mathbf{x}^i = \mathbf{g}\delta y^i + (\mathbf{I} - \mathbf{g}\mathbf{k}^T)\delta\mathbf{a}^i, \tag{25}$$

$$\mathbf{S}_x^{\text{tot}} = \mathbf{S}_x^{\text{rnd}} + \sum_i (\delta\mathbf{x}^i)(\delta\mathbf{x}^i)^T. \tag{26}$$

Here it is assumed that the retrieval weights measurements according to noise variance $\sigma_y{}^2$. Note that the gain matrix $\mathbf{g}$ for a single measurement is a column vector, and the inverted quantity is a scalar, so the above operations can be performed rapidly for a large number of measurements, provided that δy_i and $\mathbf{k}^T$ are easily obtained.

The Figure of Merit is computed for each measurement, and the best measurement is then selected. The values $\mathbf{S}_x^{\text{rnd}}$, $\delta\mathbf{x}^i$ are copied to $\mathbf{S}_a^{\text{rnd}}$, $\delta\mathbf{a}^i$ (as in sequential estimation) so that the retrieval incorporating the best measurement now forms the previous estimate for the selection of the next measurement. This process is then repeated until as many measurements are selected as required, or until the addition of any further measurements reduces the Figure of Merit.

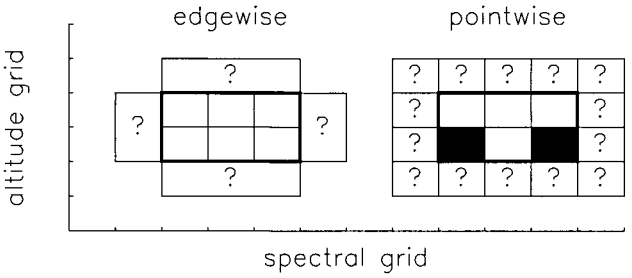

Fig. 1. Sketch indicating two possible methods of growing microwindows. A microwindow, currently bounded by the thick line, contains six measurements (3 spectral points × 2 tangent altitudes). Growing edgewise all points along one boundary are tested together, and if this improves the Figure of Merit, the microwindow is expanded in that direction. Growing pointwise all points along the boundary are tested individually, and only those points with positive impact on the Figure of Merit are included. This can leave masked points within the microwindow, as indicated by solid squares, corresponding to measurements that are excluded from the retrieval.

G. Growing Microwindows

The domain of available measurements for a limb sounder can be regarded as a two-dimensional grid, with each measurement defined by its position along the spectral and tangent altitude axes. For nadir-sounders there is just the spectral axis.

Rather than use isolated measurements, using clusters of adjacent measurements (microwindows) usually allows improved computationally efficiency in the forward model.[13] For example, the use of the same spectral grid point for successive tangent altitudes allows a reduction in the average number of independent path elements per measurement required to model the atmosphere at higher levels and represent the field-of-view convolution, while the use of adjacent spectral grid points at any one tangent altitude allows a reduction in the average number of spectral lines modeled and monochromatic calculations required to represent the apodized instrument line shape. In the retrieval domain, the use of microwindows also allows spectrally smooth but unknown continuumlike terms to be retrieved and/or removed within each microwindow.

The method for the selection of individual measurements can be extended to select coherent subsets of measurements, but the number of permutations makes this unfeasible for subsets larger than 3 or 4. One alternative is to incorporate a factor into the Figure of Merit (e.g., through the CPU cost) to favor a selection of measurements that is adjacent to existing measurements, although this does not guarantee that all individual measurements will grow into microwindows.

Instead, a two-stage technique is suggested. The first stage is to evaluate each individual measurement as described in the previous section. Then, with use of the best measurements as starting points, "grow" a number of trial microwindows. The best microwindow is retained and incorporated into the

retrieval. The process is then repeated, starting with a reevaluation of the remaining individual measurements.

Microwindows can be grown either edgewise or pointwise (Fig. 1). Edgewise, all of the measurements along each edge are tested together, and the microwindow is expanded in the direction that provides the most improvement. This will always generate rectangular microwindows, that is, microwindows that use all measurements within the four boundaries. Pointwise, the measurements around the current microwindow boundary are tested individually, and only those resulting in some improvement are added. This usually leaves unselected measurements, so-called masked points. Another possibility, given a microwindow grown by either of the above techniques, is to test each individual measurement within the microwindow and see whether the Figure of Merit improves when the measurement is included or excluded and invert the logical status of its mask if appropriate.

Assuming selection according to information content, measurements are masked (excluded) when they contribute negative information. This arises because the weight (gain matrix) is determined solely by the measurement's capacity to reduce the random error, whereas the information content assesses the total error. While it may be impractical to assign weights according to total error (although, see Ref. 12), allowing points to be excluded in this way at least gives the retrieval the option of assigning zero weight instead of the random error weight.

Apodization introduces correlations of the random noise between spectrally adjacent points. This is ignored when isolated measurements are considered, but when microwindows are grown this can be incorporated with the vector update [Eqs. (13) and (14)] to include the subset $\mathbf{y}$ of affected measurements within a microwindow. The correlations are then expressed in the off-diagonal elements of the associated random error covariance matrix $\mathbf{S}_y^{\mathrm{rnd}}$. This method is slower than the inclusion of measurements individually using the scalar update [Eqs. (23)–(26)], but the subset $\mathbf{y}$ is usually small enough that the overhead is not excessive.

Growth stops either when the inclusion of any further points results in a lower Figure of Merit or when some predefined size limit is reached.

3. Implementation for MIPAS

As a practical demonstration, microwindows are selected for methane retrievals from the MIPAS instrument.[13] The MIPAS is a limb-scanning instrument, part of the payload of the European Space Agency ENVISAT satellite due to be launched in 2002. Here it is assumed that a limb-scan sequence consists of a set of 16 spectra from tangent heights from 8–53 km in 3-km steps, and the retrieved profile is on the same levels. MIPAS spectra cover 685–2410 cm^{-1} on a grid spacing of 0.025 cm^{-1}, but for this exercise selection will be confined to the region 1215–1400 cm^{-1}, where most of the strong methane lines occur. Thus the measurement domain consists of $16 \times 7401 = 118{,}416$ points.

A. State Vector

The state vector $\mathbf{x}$ consists of 16 elements representing CH_4 volume mixing ratio, plus additional elements for each microwindow representing the retrieved atmospheric continuum for each tangent altitude and a radiometric offset. It is assumed that the continuum and offset are independent for each microwindow, so only the state vector terms related to the current microwindow have to be retained. These additional elements are required to model the behavior of the operational retrieval but are not included in the Figure of Merit calculations.

B. *A Priori* Covariance

The MIPAS operational retrieval does not use any explicit climatological constraint, so large *a priori* uncertainties of 100% are assumed for each volume mixing ration element, 1000% for continuum elements, and the instrument noise for the offset. No correlations or systematic errors are included in the *a priori* covariance.

C. Jacobians

For each point along the spectral axis, there are potentially 256 methane Jacobian matrix elements to consider, corresponding to 16 tangent altitudes $\times$ 16 profile levels (and the same again for continuum elements). To save space and computation, only the elements k_{ll} are precomputed, corresponding to the sensitivities of measurements at each tangent altitude l to perturbations in the retrieved quantities at the same profile level. Sensitivities k_{jl} of measurements at other tangent heights j to perturbations at level l are derived from

$$k_{jl} = k_{ll}(u_{jl}/u_{ll}), \tag{27}$$

where u_{jl}, u_{ll} represents the perturbed airmasses from level l in the tangent path of measurements j and l.

D. Error Sources

Random noise values are taken from test results on the flight instrument. The apodization introduces spectral correlations that are included in $\mathbf{S}_y^{\mathrm{rnd}}$ while the microwindow is grown.

Error spectra $\delta\mathbf{y}^i$ are precomputed on the basis of the perturbation of a standard midlatitude, daytime profile assuming the following error sources[14]:

Climatological variabilities of 27 different species (excluding CH_4);

Uncertainties in temperature (1 K) and pressure (2%) from outputs of the MIPAS p, T retrieval;

Uncertainties in radiometric gain (2%), spectral calibration (0.001 cm^{-1}), and instrument line shape[15];

Spectroscopic uncertainties in the HITRAN database[16] and the parameterization of molecular continua; and

Other deficiencies in the forward model include horizontal temperature gradients (±1 K/100 km), emission that is not in local thermodynamic equilibrium (non-LTE), CO_2 line-mixing

The forward model for the operational CH_4 retrieval will use, as input, profiles of temperature, pressure, and any contaminant gases that have already been retrieved by MIPAS. However, for the purposes of microwindow selection, we assume that only climatological mean profiles are available for the contaminant species. The error associated with this assumption is given by the atmospheric variability about the climatological mean, so for each species an error spectrum is computed from the radiance differences expected between a nominal atmosphere and an atmosphere with the species profile perturbed by its estimated 1σ variability. Each contaminant species error is assumed to be correlated fully (i.e., with a correlation coefficient of +1) across both the spectral and the altitude domains. The spectral correlation is clearly valid, but the altitude correlation assumes, for example, that if the concentration is underestimated at one altitude it is underestimated at all altitudes. Whereas this may not be particularly realistic, it is probably adequate for these purposes.

The forward model calculation for all measurements at a particular tangent altitude uses the same retrieved temperature, so this error is fully correlated in the spectral domain. However, it is assumed that the temperature retrieval errors themselves are largely random between profile levels, so there is no correlation in the altitude domain. Therefore the temperature errors are represented by 16 separate error sources, each generated from the nominal atmosphere with a 1-K perturbation at a single tangent level. The same applies for the pressure errors.

The gain and HITRAN uncertainties are assumed to be correlated fully in the altitude domain, but only correlated over short distances in the spectral domain. Each therefore is regarded as fully correlated within one microwindow but random between microwindows. In principle, these are also regarded as a separate error sources for each microwindow, although spectrally decorrelated errors do not require separate error vectors, since no two microwindows use the same spectral points (unlike vertically decorrelated errors such as temperature).

The operational forward model estimates line mixing but makes no allowance for horizontal gradients or non-LTE effects. However, since ignoring horizontal gradients is equivalent to assuming a zero gradient, this only leaves non-LTE as the single forward-model error introducing any bias, in which case it can be treated as the other model parameter errors using Eq. (16).

4. Results

A. Selected Microwindows

Sets of ten microwindows were selected with two different algorithms. Microwindows 1–10 (Table 1) are rectangular microwindows, using all measurements within the boundaries. Microwindows A–J (Table 2) are masked microwindows where, on average, only 49% of measurements within the boundaries are used. These are plotted in Fig. 2. Both algorithms select the 1227–1230 cm^{-1} region for the first (and largest) microwindow but differ subsequently. The masked microwindows are larger; most cover either

Table 1. First 10 Rectangular Microwindows, with Their Wave Number (cm^{-1}) and Altitude (km) Boundaries, and Number of Measurements in Each

MW	Wave number	Alt.	Meas.
1	1227.550–1229.875	8–53	1504
2	1340.725–1343.725	23–53	1331
3	1247.800–1248.650	11–53	525
4	1306.250–1306.875	23–53	286
5	1215.675–1216.575	8–53	592
6	1260.275–1260.875	14–53	350
7	1230.200–1231.325	8–53	736
8	1326.250–1327.875	23–47	594
9	1305.475–1306.075	44–53	100
10	1223.275–1224.125	8–50	525
		Total:	6543

Table 2. First 10 Masked Microwindows, with Their Wave Number (cm^{-1}) and Altitude (km) Boundaries, Total Number of Measurements within Boundaries, and Number Actually Used

MW	Wave number	Alt.	Meas.	Used
A	1227.900–1230.900	8–53	1936	1494
B	1304.300–1307.300	8–53	1936	392
C	1346.650–1349.650	8–47	1694	855
D	1269.725–1271.550	8–23	444	215
E	1299.925–1302.925	44–53	484	321
F	1327.275–1330.250	8–53	1920	900
G	1232.775–1235.775	8–41	1452	712
H	1260.175–1261.000	8–35	340	139
I	1252.350–1255.350	44–53	484	339
J	1261.700–1263.250	8–38	693	191
		Total:	11383	5558

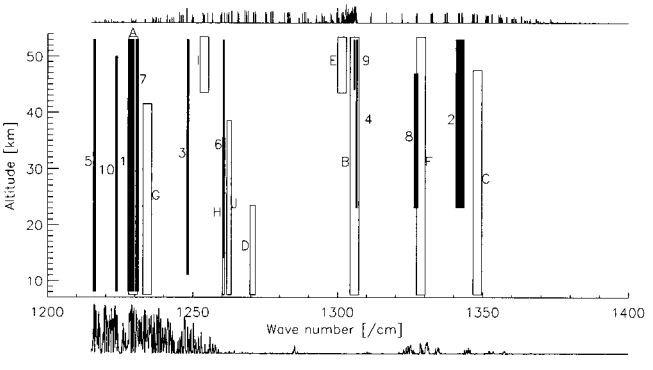

Fig. 2. Solid boxes and numbers show the rectangular microwindows as listed in Table 1. The open boxes and letters show the masked microwindows extended slightly in altitude for clarity ("A" overlaps "1" and "7"). The Jacobian spectra (arbitrary units) for methane at 53 km (top) and 8 km (bottom) are also plotted.

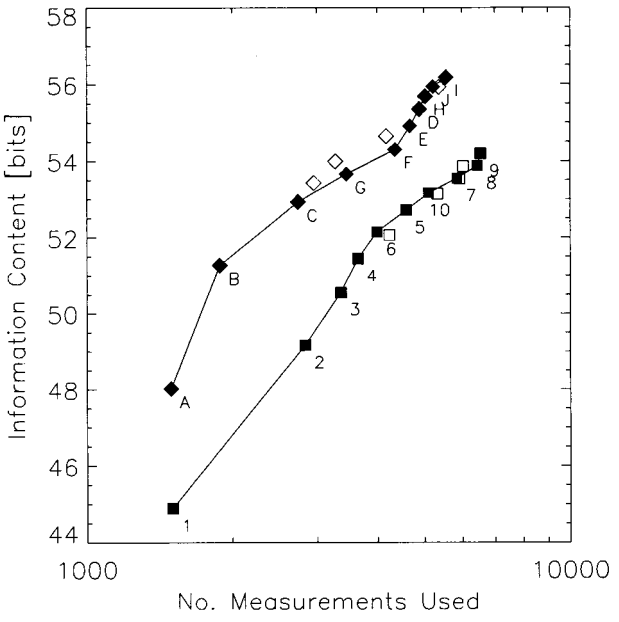

Fig. 3. Growth of information content with number of measurements for the microwindows listed in Tables 1 and 2 after reordering. Open symbols show original selection sequence (ordered A–G and 1–9 from left to right across the graph).

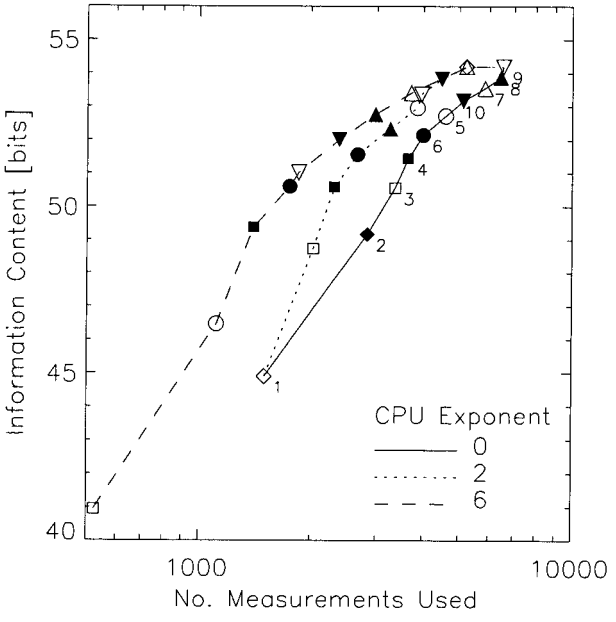

Fig. 4. Growth of information content with number of measurements for different CPU cost functions. Microwindows are distinguished by different symbols, as labeled for the $n = 0$ curve.

the full altitude range or reach the maximum spectral width of 3 cm^{-1}. This is expected since, as these microwindows are grown, the negative information points are bypassed with a mask, whereas a concentration of such points would prevent a rectangular microwindow from expanding further. Another consequence is that masked microwindows can sometimes include two rectangular microwindows: "A" covers "1" and "7"; "B" covers "4" and "9". Only the rectangular microwindows below 1250 cm^{-1} extend to lower altitudes (where the methane Jacobian is strongest). At higher wavenumbers the rectangular microwindows are limited to 23 km and above, presumably due to increased interference from lines of other species below, whereas the masked microwindows in this region appear to extract methane information down to 8 km by masking out these lines.

B. Information Content

The original selection attempted to find the microwindows in a sequence of decreasing information content, although only through a trial-and-error approach, so there is no guarantee that all of the best ten microwindows are found, or that they are in the correct sequence. The sequence, at least, can be resorted by evaluating each microwindow in turn to find the one that contributes most information to the *a priori*, include it in the retrieval, and repeat the process for the remaining microwindows, and so on.

The results of sorting the two lists are shown in Fig. 3. In both cases the first three microwindows are unchanged. Subsequent changes mainly serve to promote the larger microwindows up the sequence without significantly affecting the slope. The slope itself corresponds to around 16 bits (or a factor-2 improvement in standard deviation at all profile levels) for a factor-10 increase in the number of measurements.

The offset of the two curves suggests that the masked, compared with rectangular, microwindows appear to yield approximately three extra bits of information for a given number of measurements used or require about half as many measurements to obtain a given accuracy. However, if computing effort depends on the total number of measurements, including those points that are masked out, the two are fairly similar in this example. In this case the only advantage of using masked microwindows is that fewer microwindows, rather than measurements, are required for a given accuracy.

C. Sorting by CPU Cost

Microwindow lists can be selected with one criterion then sorted with another, such as attempting to maximize the information that can be retrieved given limited processing time. Assuming a simple CPU cost function of the form

$$C = 1 + N, \tag{28}$$

where N is the number of measurements, this is incorporated into the Figure of Merit [Eq. (22)], with various exponents n and the ten rectangular microwindows resorted according to the new criterion. The results are shown in Fig. 4.

The use of $n = 0$ is the equivalent of no CPU cost, and so is the same curve as in Fig. 3. The other curves start off more steeply, achieving a given accuracy with fewer measurements (although more microwindows) before converging to the same point at

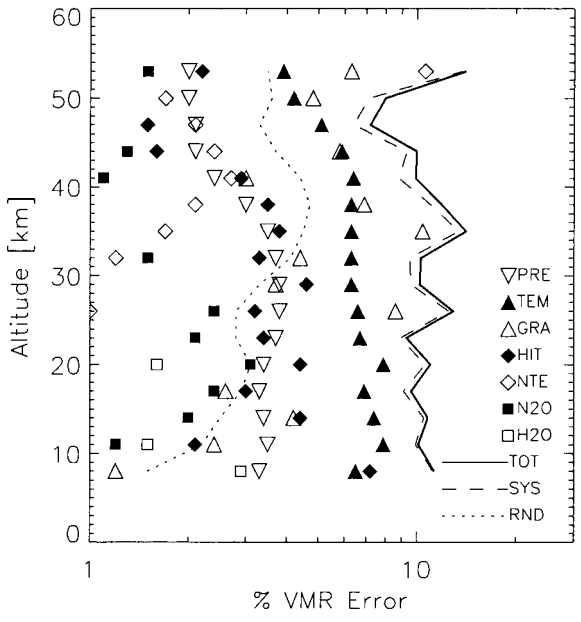

Fig. 5. Major components of the retrieval error from the rectangular microwindows listed in Table 1. Symbols indicate pressure, temperature, horizontal gradients, HITRAN, non-LTE, N_2O, and H_2O uncertainties. The combination of these gives the net systematic component, and combined with the random component gives the total error.

which all ten microwindows are used. At $n \sim 6$, the much smaller MW#3 replaces MW#1 as the first-choice microwindow.

D. Error Analysis

The random covariance $\mathbf{S}_x^{\mathrm{rnd}}$ and the retrieval error vectors $\delta\mathbf{x}^i$ contain information on the contribution of each term in the total error of the retrieval. These are plotted as profiles in Fig. 5.

This plotting shows an absolute accuracy of about 10% in retrieved CH_4 mixing ratio. The random error, or precision, contributes a relatively small component: 1–3%. The major contributions are the 1 K temperature and 2% pressure errors from the p, T retrieval, non-LTE effects at high altitudes, horizontal temperature gradients at middle altitudes, and HITRAN database uncertainties at low altitudes. N_2O is the only contaminant with any significant error contribution, peaking at 4% at 20 km (although the N_2O concentration is larger at lower altitudes, its mixing ratio is more predictable hence contributes a smaller error). H_2O contributions in the troposphere are probably eliminated by the simultaneous continuum fit in the retrieval model.

5. Conclusions

A scheme has been developed for selecting microwindows by a modeling of the retrieval propagation of systematic as well as random errors. This provides a mathematical basis for the weighting of signal from the target species against errors introduced by uncertainties in other parameters such as interfering species.

The technique will be used for the selection of microwindows for the operational processing of MIPAS data, but is equally applicable to other limb or nadir-sounding instruments that obtain atmospheric spectra. External information can also be included by way of the *a priori* covariance, allowing, for example, microwindows to be selected which best complement the output of a numerical model.

Instead of minimizing the total error, other priorities may also be defined by way of a Figure of Merit. The microwindows are generated in order of diminishing merit, so the sequence may be optimally truncated at any point. The scheme can also be used to reorder microwindows from an existing list according to different priorities or circumstances, such as a changed set of retrieval levels or missing data, or revised estimates of some error source.

The method relies on a representation of all error sources in measurement space (i.e., as perturbation spectra), which are then mapped into state vector space (i.e., retrieved profiles). This can be used to provide a full error analysis of the retrieval.

Some important limitations should be noted. First, the process is not optimal in the sense that it can determine the best possible set of microwindows. This is partly due to the trial-and-error nature of the growing of microwindows from individual measurements, although taking a large number of trials improves the chances of the best microwindow being found. More fundamentally, the selection is strictly sequential in that it maximizes the increase in the chosen Figure of Merit at each opportunity: It does not allow for combinations of individually less beneficial microwindows that may give greater net improvement (the given example of weighting against systematic errors attempts to address this issue).

Second, the above selection is based on the definition of error spectra and Jacobians for a single atmospheric profile and would therefore be less ideal for different profiles. For the selection of the MIPAS operational microwindows, an ensemble of five different atmospheric profiles has been used, each with its own set of spectra, and the Figure of Merit represents an average of the five different obtained values. This produces a single set of globally optimized microwindows that should be suitable for all latitudes and seasons.

During this work A. Dudhia was employed under an European Space Agency contract, and V. Jay was supported under a Natural Environment Research Council studentship.

References

1. "AIRS—Atmospheric Infrared Sounder Homepage," (Jet Propulsion Laboratory, Pasadena, Calif., 2001), http://www-airs.jpl.nasa.gov/
2. "Interferometric Monitor for Greenhouse Gases (IMG)," (Earth Observation Research Center, Tokyo, Japan, 2001), http://www.eorc.nasda.go.jp/ADEOS/Project/Img.html.
3. "MIPAS Introduction," (European Space Agency, Paris, France, 2001), http://www.envisat.esa.int/instruments/mipas/.

4. "IASI Welcome!," (Centre National d'Etudes Spatiales, Toulouse, France, 2001), http://smsc.cnes.fr/IASI/welcome.html.
5. "Welcome to TES," (Jet Propulsion Laboratory, Pasadena, Calif., 2001), http://tes.jpl.nasa.gov/.
6. C. D. Rodgers, "Information content and optimisation of high spectral resolution measurements," in *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II*, P. B. Hays and J. Wang, eds., Proc. SPIE **2830,** 136–147 (1996).
7. C. D. Rodgers, "Information content and optimisation of high spectral resolution remote measurements," Adv. Space Res. **21,** 361–367 (1998).
8. F. Rabier, N. Fourrié, D. Chafaï, and P. Prunet, "Channel selection methods for infrared atmospheric sounding interferometer radiances," Q. J. R. Meteorol. Soc. **128,** 1–15 (2001).
9. T. von Clarmann and G. Echle, "Selection of optimized microwindows for atmospheric spectroscopy," Appl. Opt. **37,** 7661–7660 (1998).
10. G. Echle, T. von Clarmann, A. Dudhia, J.-M. Flaud, B. Funke, N. Glatthor, B. J. Kerridge, M. López-Puertas, F. J. Martín-Torres, and G. P. Stiller, "Optimized spectral microwindows for MIPAS-Envisat data analysis," Appl. Opt. **39,** 5531–5540 (2000).
11. C. D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice* (World Scientific, Singapore, 2000).
12. T. von Clarmann, U. Grabowski, and M. Kiefer, "On the role of non-random errors in inverse problems in radiative transfer and other applications," J. Quant. Spectrosc. Radiat. Transfer **71,** 39–46 (2001).
13. M. Ridolfi, B. Carli, M. Carlotti, T. von Clarmann, B. M. Dinelli, A. Dudhia, J. M. Flaud, M. Höpfner, P. E. Morris, P. Raspollini, G. Stiller, and R. J. Wells, "Optimized forward model and retrieval scheme for MIPAS near-real-time data processing," Appl. Opt. **39,** 1323–1340 (2000).
14. T. von Clarmann, A. Dudhia, G. Echle, J.-M. Flaud, C. Harrold, B. Kerridge, K. Koutoulaki, A. Linden, M. López-Puertas, M. A. López-Valverde, F. J. Martín-Torres, J. Reburn, J. Remedios, C. D. Rodgers, R. Siddans, R. J. Wells, and G. Zaragoza, "Study on the simulation of atmospheric infrared spectra," ESA contract 12054/96/NL/CN (European Space Agency, Paris, 1998).
15. European Space Agency, "ENVISAT-MIPAS: an instrument for atmospheric chemistry and climate research," ESA SP-1229 (European Space Agency Research and Technology Centre, Noordwijk, The Netherlands, 2000).
16. L. S. Rothman, C. P. Rinsland, A. Goldman, S. T. Massie, D. P. Edwards, J.-M. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J.-Y. Mandin, J. Schroeder, A. McCann, R. R. Gamache, R. B. Wattson, K. Yoshino, K. V. Chance, K. W. Jucks, L. R. Brown, V. Nemtchinov, and P. Varanasi, "The HITRAN Molecular Spectroscopic Database and HAWKS (HITRAN Atmospheric Workstation): 1996 Edition," J. Quant. Spectrosc. Radiat. Transfer **60,** 665–710 (1998).