Water vapour profiles by ground-based FTIR spectroscopy: study for an optimised retrieval and its validation

M. Schneider, F. Hase, T. Blumenstock

Atmos. Chem. Phys. Discuss., 5, 9493–9545, 2005
www.atmos-chem-phys.org/acpd/5/9493/
SRef-ID: 1680-7375/acpd/2005-5-9493
European Geosciences Union

Atmospheric Chemistry and Physics Discussions

Water vapour profiles by ground-based FTIR spectroscopy: study for an optimised retrieval and its validation

M. Schneider, F. Hase, and T. Blumenstock

IMK-ASF, Forschungszentrum Karlsruhe and Universität Karlsruhe, Germany

Received: 7 July 2005 – Accepted: 31 August 2005 – Published: 29 September 2005

Correspondence to: M. Schneider (matthias.schneider@imk.fzk.de)

Abstract

The sensitivity of ground-based instruments measuring in the infrared with respect to tropospheric water vapour content is generally limited to the lower and middle troposphere. The large vertical gradients and variabilities avoid a better sensitivity for the upper troposphere/lower stratosphere region. In this work an optimised retrieval is presented and it is demonstrated that compared to a commonly applied method, it widely improves the performance of the FTIR technique with respect to upper tropospheric water vapour. Within a realistic error scenario it is estimated that the optimised method reduces the upper tropospheric uncertainties by about 25–30%, leading to a noise to signal ratio of 50%. The reasons for this improvement and the possible deficiencies of the method are discussed. The estimations are confirmed by a comparison of retrieval results based on real FTIR measurements with coinciding measurements of synoptical meteorological radiosondes.

1. Introduction

The composition of the Earth's atmosphere has been profoundly modified throughout the last decades mainly by human activities. Prominent examples are the stratospheric ozone depletion and the upward trend in the concentration of greenhouse gases. While studies about the stratospheric composition have progressed rather well, there still exists a considerable deficiency for data from the free troposphere. Knowing the composition and evolution of these altitude regions is essential for the scientific verification of the Kyoto and Montreal Protocols and Amendments and for global climate modelling. Water vapour is the dominant greenhouse gas in the atmosphere, and in particular its concentration and evolution in the upper troposphere and lower stratosphere (UT/LS) are of great scientific interest for climate modelling (Harries, 1997; Spencer and Braswell, 1997). Currently there is no outstanding routine technique for measuring water vapour in the UT/LS. The quick changes of atmospheric water vapour

ACPD

5, 9493–9545, 2005

concentrations with time, their large horizontal gradients, and their decrease of several orders of magnitude with height makes their accurate detection a challenging task for any measurement technique. Traditionally tropospheric water vapour profiles are measured by synoptical meteorological radiosondes. However, this method has some deficiencies at altitudes above 6–8 km, which are mainly due to uncertainties in the pre-flight calibration and temperature dependence (Miloshevich, 2001; Leiterer et al., 2004). Other applied techniques are remote sensing from the ground by Lidar or Microwave instruments. Both are limited in their sensitivity: the Lidar generally to below 8–10 km, and the microwave measurements to above 15 km (SPARC, 2000). Satellite instruments also struggle to reach below this altitude. In this context the suggested formalism of retrieving upper tropospheric water vapour amounts from ground-based FTIR measurements aims to support efforts to obtain quality UT/LS water vapour data for research. To our knowledge, it is the first time that water vapour profiles measured by this technique are presented. A great advantage is that high quality ground-based FTIR measurements have already been performed during the last 10–15 years within the Network for Detection of Stratospheric Change (Kurylo, 1991, 2000; NDSC, web site). Therefore a long-term record of water vapour could be made available, with both temporal and to some extent, spatial coverage.

The structure of the article is as follows: first it is argued how the suggested optimisation acts in the context of inversion theory. Its advantages and deficiencies compared to a method, commonly used for trace gas retrievals, are discussed. In the third section an error assessment adds precise quantitative estimations about the expected improvements to these qualitative considerations. It is also shown how possible deficiencies of the optimised method can be eliminated. Finally, these estimations are validated by a comparison of retrieval results based on real measurements with coinciding in-situ measurements.

2. Optimised water vapour retrieval

An inversion problem is generally under-determined. Many state vectors ($\boldsymbol{x}$) are consistent with the measurement vector ($\boldsymbol{y}$). If one also considers measurement noise ($\boldsymbol{\epsilon_y}$), there is an even wider range of possible solutions within $\boldsymbol{\epsilon_y}$, in accordance to the measurement vector: in the equation,

$$\hat{\boldsymbol{y}} = \boldsymbol{y} + \boldsymbol{\epsilon_y} = \mathbf{K}\boldsymbol{x} \tag{1}$$

the matrix $\mathbf{K}$ is ill-conditioned. Its effective rank is smaller than the dimension of state space, i.e. it is singular and cannot be simply inverted. To come to an unique solution of $\boldsymbol{x}$, the state space is constrained by requiring:

$$\mathbf{B}\boldsymbol{x} = \mathbf{B}\boldsymbol{x_a} \tag{2}$$

where $\boldsymbol{x_a}$ is a 'typical' or a-priori state and the matrix $\mathbf{B}$ determines the kind of required similarity of $\boldsymbol{x}$ with $\boldsymbol{x_a}$. This equation constrains the solution independently from the measurement, i.e. before the measurement is made. Therefore $\mathbf{B}$ and $\boldsymbol{x_a}$ contain the kind of information known about the state prior to the measurement. Subsequently, the state vector $\boldsymbol{x}$ for which the whole system of equations (Eqs. 1 and 2) is fulfilled in a least squares sense is selected as the solution, i.e. one has to minimise the cost function:

$$\sigma^{-2}(\boldsymbol{y} - \mathbf{K}\boldsymbol{x})^T(\boldsymbol{y} - \mathbf{K}\boldsymbol{x}) + (\boldsymbol{x} - \boldsymbol{x_a})^T\mathbf{B}^T\mathbf{B}(\boldsymbol{x} - \boldsymbol{x_a}) \tag{3}$$

Here $(\boldsymbol{\epsilon_y}^T\boldsymbol{\epsilon_y})^{-1}$ was identified by σ^{-2}. It is obvious that the applied a-priori information ($\mathbf{B}$ and $\boldsymbol{x_a}$) influences the solution. For water vapour the large amount of synoptical meteorological sonde (ptu-sonde) data allows a detailed study of the a-priori state. In the following it is discussed whether the extensive a-priori information can be used to optimise the performance of the retrieval. The study of a-priori data is done for the island of Tenerife, where ptu-sondes are launched twice daily (at 00:00 and 12:00 UT) within the global radiosonde network and where an FTIR instrument has been operating since 1999 at a mountain observatory (Izaña Observatory, Schneider et al., 2005).

ACPD
5, 9493–9545, 2005

2.1. Characterisation of a-priori data

The study is based on the daily 12:00 UT soundings performed from 1999 to 2003. It has been observed that an in-situ instrument – located at the mountain observatory – and the sonde, when measuring at the observatory's altitude, detect quite different humidities because of their different locations, i.e. on the surface and in the free troposphere (see Sect. 4). For this reason the analysed profiles are built up by a combination of the in-situ measurements at the instrument's site (for the lowest layer), and sonde measurements (for all other layers below 16 km). For higher altitudes a mean mixing ratio of 2.5 ppmv and covariances like those at 16 km are applied. The left panel of Fig. 1 shows the correlation matrix $\mathbf{\Gamma_a}$ determined from these a-priori profiles. It demonstrates how variabilities at different altitudes typically correlate with each other. In the real atmosphere the mixing ratios for different altitudes show correlation coefficients of at least 0.5 within a layer of around 2.5 km. The a-priori covariance matrix $\mathbf{S_a}$ is calculated from $\mathbf{\Gamma_a}$ by $\mathbf{S_a}=\mathbf{\Sigma_a\Gamma_a\Sigma_a}^T$, where $\mathbf{\Sigma_a}$ is a diagonal matrix containing the a-priori variabilities at a certain altitude. These variabilities are depicted as a red line in the right panel of Fig. 1. The black line shows the mean mixing ratios. The determined mean and covariances only describe the whole ensemble completely if mixing ratios are normally distributed. This is generally assumed and often justified by the fact that entropy is then maximised: if only the mean and the covariance are known a supposed normal distribution is thus the least restricting assumption about the a-priori state (Sect. 10.3.3.2 in Rodgers, 2000). However, this does not necessarily reflect the real situation!

A further examination of the sonde data reveals that the mixing ratios at a certain altitude are not normally but log-normally distributed. Their pdf is:

$$P_x = \frac{1}{x\sigma\sqrt{2\pi}} \exp -\frac{(\ln x - \ln m)^2}{2\sigma^2} \quad (4)$$

with a shape parameter σ ranging from 1.15 ppmv in the middle troposphere to 0.55 ppmv above 10 km, and a median m between 5000 ppmv close to the surface

and 1.5 ppmv in the stratosphere. The only exception of this distribution is the first ≈100 m above the surface, where the mixing ratios are more normally distributed. It is possible to sample all this additional information in a simple mean state vector and a covariance matrix. This is achieved by transforming the state on a logarithmic scale, which transforms the log-normal pdf to a normal pdf. A normal pdf can be completely described by its covariance and its mean. A χ^2-test reveals how the description of the a-priori state is improved by this transformation. This test determines the probability of a particular random vector of belonging to an assumed normal distribution. If a vector $\boldsymbol{x}$ is supposed to be a member of a Gaussian ensemble with zero mean and covariance $\mathbf{S}$ the quantity considered is:

$$\chi^2 = \boldsymbol{x}^T \mathbf{S}^{-1} \boldsymbol{x} \tag{5}$$

The χ^2 test clearly rejects a normal distribution of the mixing ratios. This can be seen by comparing the theoretical cumulative distribution function (cdf) of χ^2 with the one determined by Eq. (5). Figure 2 demonstrates that the theoretical χ^2 cdf differs clearly from the cdf obtained from the ensemble's state vectors if they are assumed to be normally distributed (difference between black line and black squares). More than 95% of the ensemble's state vectors are not consistent with this assumption. On the other hand, a prior log-normal pdf is well confirmed. If the mixing ratios and the covariances are transformed to a logarithmic scale, only approximately 10% of the ensemble's states fail the test (compare black line and red circles).

2.2. Discussion of two retrieval methods

This section discusses the differences between an inversion performed on a linear scale, which is the method commonly used for trace gas retrievals, and one performed on a logarithmic scale. The logarithmic retrieval is occasionally applied as a positivity constraint, since it avoids negative components in the solution vector. In the case of water vapour it has a further advantage. It converts the state for which Eq. (3) minimises in a statistically optimal solution: on a logarithmic scale the a-priori state can be

ACPD
5, 9493–9545, 2005

described correctly in the form of a mean and covariance. Under these circumstances, substituting $\mathbf{B}^T\mathbf{B}$ and $\boldsymbol{x_a}$ in Eq. (3) by the inverse of the logarithmic a-priori covariance ($\mathbf{S_a}^{-1}$) and the median state vector, leads to a cost function, which is directly proportional to the negative logarithm of the a-posteriori probability density function (pdf) of the Bayesian approach. This posterior pdf is the conditional pdf of the state given the measurement, or in other words, the a-priori pdf of the state updated by the information given in the measurement. The minimisation of Eq. (3) thus yields the maximum a-posteriori solution, i.e. it is the most probable state given the measurement.

To the contrary, on a linear scale setting $\mathbf{B}^T\mathbf{B}$ as $\mathbf{S_a}^{-1}$ and $\boldsymbol{x_a}$ as mean state in Eq. (3) does not lead to a statistically optimal solution. It is not related to the a-posteriori pdf in the Bayesian sense. On a linear scale the a-priori state is log-normally distributed. Therefore, seen from a statistical point of view, the second term of the cost function over-constrains states above the mean and under-constrains states below the median. As a consequence, the probability of states above the mean is underestimated and below the median overestimated – the overestimation is greater the further away it is from the centre of the a-priori distribution. Thus, if compared to a correct maximum a-posteriori solution, the retrieval tends to underestimate the values of the real state both far above and far below the mean state.

However, the transformation on a logarithmic scale introduces some other problems: it widely increases the non-linearity of the forward model, which requires decreasing the differences between each iteration step, and thus lowers the speed of convergence. Furthermore, in the retransformed linear scale the constraints now depend on the solution, which may cause misinterpretations of the spectra. To assess whether the linear or logarithmic retrieval performs better both retrieval approaches are extensively examined first by a theoretical (Sect. 3) and second by an empirical validation (Sect. 4).

2.3. Applied inversion code and spectral region

PROFFIT (Hase et al., 2004) is the inversion code used. It applies the Karlsruhe Optimised and Precise Radiative Transfer Algorithm (KOPRA, Höpfner et al., 1998; Kuntz

et al., 1998; Stiller et al., 1998) as the forward model, which was developed for the analysis of MIPAS-Envisat limb sounder spectra. PROFFIT enables the inversion on a linear and logarithmic scale. Hence, in the case of water vapour, it enables the correct application of prior information to obtain a statistically optimal solution. PROFFIT does not employ a fixed a-priori value for the measurement noise (σ of Eq. 3). This value is taken from the residuals of the fit itself, performing an automatic quality control of the measured spectra. Furthermore, if the observed absorptions depend on temperature, PROFFIT allows the retrieval of temperature profiles. For both the linear and logarithmic retrieval, the same fit strategy is applied: three microwindows between 1110 and 1122 cm^{-1} are fitted. Figure 3 shows a typical situation for an evaluation of a real measurement. The black line represents the measurement, the red dotted line the simulated spectrum and the green line the difference between both. The H_2O signatures are marked in the Figure. One can observe that two stronger lines (at 1111.5 and 1121.2 cm^{-1}) and two relatively weak lines (at 1117.6 and 1120.8 cm^{-1}) lie within these spectral regions, where additionally O_3 is an important absorber (numerous thin strong signatures). The profile of this species is thus simultaneously retrieved. Other interfering gases are CO_2, N_2O, and CH_4, whereby the latter two are also simultaneously retrieved by scaling their respective climatological profiles, the former is kept fixed to a climatological profile. Spectroscopic line parameters are taken from the HITRAN 2000 database Rothman et al. (2003), except for O_3, where parameters from Wagner et al. (2002) are applied.

3. Error analysis and sensitivity assessment

Assuming linearity for the forward model $\boldsymbol{F}$ and the inverse model $\boldsymbol{I}$ within the uncertainties of the retrieved state and the model parameters it is (Rodgers, 2000):

$$\hat{\boldsymbol{x}} - \boldsymbol{x} = \left(\frac{\partial \boldsymbol{I}[\boldsymbol{F}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}), \hat{\boldsymbol{p}}]}{\partial \boldsymbol{y}} \frac{\partial \boldsymbol{F}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}})}{\partial \boldsymbol{x}} - \mathbf{I} \right) (\boldsymbol{x} - \boldsymbol{x_a})$$

$$+ \frac{\partial I[F(\hat{x},\hat{p}),\hat{p}]}{\partial y}\frac{\partial F(\hat{x},\hat{p})}{\partial p}(p-\hat{p})$$
$$+ \frac{\partial I[F(\hat{x},\hat{p}),\hat{p}]}{\partial y}(y-\hat{y})$$
$$= (\hat{\mathbf{A}}-\mathbf{I})(x-x_a) + \hat{\mathbf{G}}\hat{\mathbf{K}}_p(p-\hat{p}) + \hat{\mathbf{G}}(y-\hat{y}) \quad (6)$$

i.e. the difference between the retrieved and the real state ($\hat{x}-x$) – the error – can be linearised about a mean profile x_a, the estimated model parameters $\hat{p}$, and the measured spectrum $\hat{y}$. Here $\mathbf{I}$ is the identity matrix, $\hat{\mathbf{A}}$ the averaging kernel matrix, $\hat{\mathbf{G}}$ the gain matrix, and $\hat{\mathbf{K}}_p$ a sensitivity matrix to model parameters:

$$\hat{\mathbf{A}} = \hat{\mathbf{G}}\hat{\mathbf{K}}$$
$$\hat{\mathbf{G}} = \frac{\partial I[F(\hat{x},\hat{p},\hat{p}]}{\partial y}$$
$$\hat{\mathbf{K}} = \frac{\partial F(\hat{x},\hat{p})}{\partial x}$$
$$\hat{\mathbf{K}}_p = \frac{\partial F(\hat{x},\hat{p})}{\partial p} \quad (7)$$

whereby $\hat{\mathbf{K}}$ is the Jacobian. Equation (6) identifies three principle error sources. These are the inherent finite vertical resolution, the input parameters applied in the inversion procedure, and the measurement noise. This analytic error estimation may be applied if the inversion is performed on a linear scale. In this case, the constraints and consequently $\hat{\mathbf{G}}$ are constant within the uncertainty of $\hat{x}$. However, if the inversion is performed on a logarithmic scale the constraints are constant on this scale, but variable on the retransformed linear scale. Changes of the state vector towards values above the a-priori value are only weakly constrained, while changes towards smaller values are more strongly constrained. As a consequence $\hat{\mathbf{G}}$ cannot necessarily be considered constant within the uncertainty of the retrieved state and some model parameters. The latter is particularly problematic for water vapour. The phase error of the instrumental

line shape and the temperature profile have a large impact on the spectra. This is due to the broad and strong absorption signatures of water vapor. Consequently, all these errors can only be estimated by a full treatment, i.e. forward modelling to determine the impact of the parameter error on the spectrum and its subsequent inversion. In this work all errors, for linear as well as logarithmic retrieval, are estimated by a full treatment for consistency reasons. All random errors are expressed as ratios between the variability of the error and the variability of the retrieved value (noise to signal). It thus gives information about the quantity of the observed variabilities that does not correspond to a real atmospheric variability. Any amount with a noise to signal ratio above 100% is thus not observable. The systematic errors are expressed as a ratio of the mean of the error and the mean of the retrieved value. First the random errors are estimated and systematic errors are briefly discussed at the end of this section.

Together with the error estimation a sensitivity assessment are performed. Generally the averaging kernels (columns of $\hat{\mathbf{A}}$) are used to estimate the sensitivity of the retrieval at certain altitudes. They document by how much ppmv the retrieved solution will change due to a variability of 1 ppmv in the real atmosphere. They may inform that 1 ppmv more at 5 km is reflected in the retrieval by an extra of 0.1 ppmv at 8 km. However, the typical real atmospheric variabilities at different altitudes are not considered and hence to what extent the typical variability as retrieved at 8 km is disturbed by typical variabilities at 5 km. This is a minor problem if the mixing ratio variabilities have the same magnitude throughout the atmosphere. The variabilities of water vapour decrease by 3–4 orders of magnitude from the surface to the tropopause (see Fig. 1), thus the interpretation of the averaging kernels is quite limited. Alternatively, one may produce adequately normed kernels to address this deficiency. Here a full treatment, consisting of forward calculation of assumed real states and subsequent inversion, is used to estimate the response of the retrieval on real atmospheric variabilities. Therefore, the real state vectors are correlated to their corresponding retrieved vectors. The correlation coefficient (ρ) considers the different magnitudes of the variabilities. For instance, ρ between the real state at 5 km and the retrieved state at 8 km gives the typ-

ACPD
5, 9493–9545, 2005

ical fraction of the retrieved variabilities at 8 km due to disturbances from 5 km. These correlation matrices give a good overview of the relation between real atmospheric variabilities and the retrieved variabilities.

Error estimation and sensitivity assessment are performed for the whole ensemble (the ensemble used for calculating the a-priori mean and covariances), and for a sub-ensemble of selected conditions, when especially good upper tropospheric sensitivity is expected. The trace of the averaging kernel (tr($\hat{\mathbf{A}}$)) determines the degree of freedom (DOF) for the whole posterior state space. It communicates the amount of information present in the spectra used by the retrieval for updating the a-priori state. The higher the value of tr($\hat{\mathbf{A}}$) the more information comes from the measurement and the less from the a-priori assumptions. Similarly, the DOF for the sub posterior state space, corresponding to the upper troposphere, can be calculated by adding only the diagonal elements of $\hat{\mathbf{A}}$, which represent these altitudes (subsequently named DOF-UT). Here the region above 7.6 km and below 12.4 km is defined as upper troposphere and all situations with a DOF-UT above 0.2 are classified as days on which the retrieval suggests good upper tropospheric sensitivity. The choice is somehow arbitrary, but it appeared to be a good compromise between good sensitivity and size of the ensemble (more than 30% of all days fulfill this criterion). The relevant days are coupled to low slant columns in the lower troposphere, i.e. to unsaturated water vapour signatures. The DOF criterion also considers other unfavourable measurement conditions such as low spectra intensity due to high aerosol loading. This situation occurs occasionally at Izaña owing to Saharan dust intrusion events. In this work, the DOF values according to the logarithmic retrieval are used. This ensures that the prior information is applied correctly and, as a consequence, optimal use of the information present in the spectrum is made. Constructing an ensemble with high DOF-UT values according to the linear retrieval would impose an additional filter, since days with high mixing ratios would be excluded in advance (over-constraint case). This means that the observed spectral signatures are not optimally exploited. On the other hand, some days with low ratios are admitted although spectral signatures are quite uncertain, i.e. spectral signa-

ACPD
5, 9493–9545, 2005

tures are over-interpreted (under-constraint case). The effect of this additional filtering is demonstrated in Fig. 4. The black line shows the pdf of real upper tropospheric column amount for the whole ensemble: the a-priori pdf for this amount. The red curve shows the pdf for DOF-UT values above 0.2 according to the linear retrieval, and the green curve when classification is performed according to the logarithmic retrieval. Applying the logarithmic DOF-UT criterion leaves the a-priori pdf unchanged. It works independently from the actual state of the upper troposphere. On the other hand, classification with the linear DOF-UT criterion changes the a-priori pdf. This sub-ensemble would not be representative for the a-priori state of the upper troposphere.

3.1. Smoothing error

The smoothing error has the form of a covariance matrix, with large outer diagonal elements: the errors at different altitudes are strongly correlated. Disregarding the correlations overestimates the importance of the smoothing error and lead to an incorrect conclusion concerning the retrieval's performance. They can be presented in the form of error patterns (Rodgers, 2000), whose interpretation is however not straightforward. Here the errors are assessed for layers and not for a single altitude, which has the advantage that the interlevel correlations within the layers are considered automatically. Moreover, bearing in mind the modest vertical resolution of trace gas profiles determined by ground-based FTIR spectroscopy, the objective of this technique should consist of retrieving the amount of a certain layer rather than a concentration at a single altitude. Thus, an error estimation for layers avoids extensive explanations about the interlevel correlations and is more interesting. Furthermore, the sensitivity assessment in the form of correlation matrices will already give some valuable insight into the correlation length of the errors.

Figure 5 shows correlation matrices in the absence of parameter errors. They document the sensitivity of the retrieval if the smoothing error alone is taken into account. The left panels show the linear retrieval, the right panels the logarithmic retrieval, the upper panels the whole ensemble, and the lower the high DOF-UT sub-ensemble. Con-

sidering the whole ensemble the sensitivity is limited to altitudes below 10 km. Furthermore, the upper tropospheric mixing ratios of the linear retrieval tend to depend more on variabilities at lower altitudes. For example, the value retrieved at 9 km is mainly influenced by the real atmospheric situation at 7 km. This incorrect altitude attribution is less pronounced in the logarithmic retrieval. For the DOF-UT sub-ensemble the sensitivity is extended by 1–2 km towards higher altitudes. In this case, the observing system provides good information about the atmospheric water vapour variabilities up to 12 km (ρ at the diagonal above 0.5). As before, for the linear retrieval, the amounts at higher altitudes are strongly disturbed by the real states at lower altitudes, while, for the logarithmic retrieval, high correlation coefficients are more concentrated around the diagonal of the matrix. The correlation length of the smoothing error interlevel correlations is smaller, resulting in an improved vertical resolution when compared to the linear retrieval. For example, the mixing ratio retrieved at 11 km for the high DOF-UT sub ensemble has a ρ value for the correlation with its corresponding real value of 0.57 and 0.6 for the linear and logarithmic retrieval, respectively. Apparently, the logarithmic retrieval performs only marginally better. However, for the linear case large correlation to real values at lower altitudes are found (e.g. ρ of 0.82 at 7 km). These disturbances are significantly reduced in the logarithmic case (ρ of 0.69 at 7 km). Here the state retrieved at 11 km is much less influenced by variabilities at 7 km. Thus the precision of the state retrieved at 11 km can already be sufficiently improved by considering the disturbances originating from altitudes down to about 7.5 km only. The linear retrieval, on the other hand, should very likely take into account values from further down in order to reach a similar precision. This means that the correlation length of the smoothing error is larger for the linear retrieval. To determine the amount of a layer with a certain precision the layer must be broader for the linear retrieval if compared to the logarithmic retrieval.

Figure 5 suggests that the observing system is capable of resolving quite fine structures for the lowest atmospheric layer. At increasing altitude the vertical resolution decreases. This is considered when presenting the errors. Figure 6 depicts the smoothing

error of several layers throughout the troposphere. The altitude region of each layer is indicated by the error bars, which increase with altitude. The black-filled squares represent the typical error for the whole ensemble: left panel for the linear and right for the logarithmic retrieval. It confirms the observation made in Fig. 5 that the logarithmic retrieval performs better due to its finer structured interlevel correlations. This is especially true for the upper troposphere. However, even for the logarithmic retrieval the noise to signal ratio exceeds 80% for the 7.6–12.4 km layer. The red-filled squares show the same but for the DOF-UT ensemble. For this sub-ensemble the smoothing error of the 7.6–12.4 km layer is significantly reduced (logarithmic retrieval: from 84% to 54%). For the linear retrieval the ratio remains around 80%. This confirms the better performance expected for the logarithmic retrieval.

3.2. Model parameter error

In this subsection errors due to measurement noise, uncertainties in solar angle, instrumental line shape (ILS: modulation efficiency and phase error Hase et al., 1999), temperature profile, and spectroscopic parameters (line intensity and pressure broadening coefficient) are estimated. Although line intensity and pressure broadening coefficients are systematic uncertainties they may produce random errors. This is due to the nonlinearity of the problem ($\hat{\mathbf{K}}_p$ depends on the state). The assumed parameter uncertainties are listed in Table 1. Two sources are considered as errors in the temperature profile: first, the measurement uncertainty of the sonde, which is assumed to be 0.5 K throughout the whole troposphere and to have no interlevel correlations. Second, the temporal differences between the FTIR and the sonde's temperature measurements, which are estimated to be 1.5 K at the surface and 0.5 K in the rest of the troposphere, with 5 km correlation length for the interlevel correlations.

Errors due to measurement noise, uncertainties in the modulation efficiencies, the solar angle and the line intensity are situated below or around 5%. They may thus be neglected if compared to the errors caused by phase error, temperature profile, or pressure broadening coefficient uncertainties. Figure 7 shows the latter errors for the

ACPD
5, 9493–9545, 2005

whole ensemble (upper panels) and for the DOF-UT ensemble (lower panels). Considering the whole ensemble the temperature uncertainty provides the largest errors (red crosses). The errors are generally larger for the logarithmic retrieval, in particular the temperature error. Here the 45% at 10 km for the linear retrieval is much lower than the 70% at 8 km for the logarithmic retrieval. This is due to the retrieval's misinterpretation of spectral signatures arising from errors in the temperature profile. Since $\hat{\mathbf{G}}\hat{\mathbf{K}}_\mathbf{p}$ from Eq. (6) is generally not equal to zero, the parameter error in the measurement space may be transformed into the state space. This is a minor problem when the minimisation of the cost function (Eq. 3) is performed on a linear scale. Then changes of the state vector with respect to its a-priori state and the magnitude of the constraining term are linearly correlated. A misinterpretation would thus mean a large value of the constraining term and consequently Eq. (3) would never be minimised. On a logarithmic scale, however, a linear increase of the constraining term is related to an exponential increase of the retransformed state vector. Hence, a significant change of the state vector is not avoided by the constraining term. The problem can be reduced by a simultaneous retrieval of the temperature profile, which adds two terms to the cost function:

$$\sigma^{-2}(\boldsymbol{y}-\mathbf{K}\boldsymbol{x})^T(\boldsymbol{y}-\mathbf{K}\boldsymbol{x})+(\boldsymbol{x}-\boldsymbol{x}_a)^T{\mathbf{S}_\mathbf{a}}^{-1}(\boldsymbol{x}-\boldsymbol{x}_a)$$
$$+\sigma^{-2}(\boldsymbol{y}-\mathbf{K}_\mathbf{t}\boldsymbol{t})^T(\boldsymbol{y}-\mathbf{K}_\mathbf{t}\boldsymbol{t})+(\boldsymbol{t}-\boldsymbol{t}_a)^T{\mathbf{S}_{\epsilon\mathbf{t}}}^{-1}(\boldsymbol{t}-\boldsymbol{t}_a) \quad (8)$$

Here $\boldsymbol{t}$ and $\boldsymbol{t}_a$ are the real and the assumed temperature state vector, $\mathbf{K}_\mathbf{t}$ the sensitivity (or Jacobian) matrix for the temperature, and $\mathbf{S}_{\epsilon\mathbf{t}}$ the error covariance matrix for the temperature. Thus a temperature error does not lead to an adjustment of the first term – a misinterpretation of spectral information –, but to an adjustment of the third term in Eq. (8). This reduces the probability of misinterpreting the temperature error. In the upper troposphere, for example, the simultaneous fitting of the temperature profile reduces the error from over 70% to around 20%. This is seen by comparing the red crosses with the red squares in Fig. 7. This strategy leaves the uncertainty in phase error and pressure broadening coefficient as the most important error sources.

ACPD
5, 9493–9545, 2005

For the DOF-UT sub-ensemble (lower panels of Fig. 7), the errors for the middle and upper troposphere are much smaller (below 30%). Above 5 km at least, errors for the linear and logarithmic retrieval are now similar. For this ensemble a misinterpretation of spectral signatures is less probable. Apparently, the condition for high DOF-UT values simultaneously eliminates days predestined for misinterpretation. However, a simultaneous retrieval of the temperature further improves the retrievals by reducing the temperature error for the lower troposphere in particular. The most important errors are uncertainties in pressure broadening coefficient and the phase error.

3.3. Total random errors

Due to the strong non-linearity of $\hat{\mathbf{G}}$ the total error cannot be deduced from the smoothing and parameter errors presented above. It has to be simulated separately by a full treatment. Figure 8 shows the correlation matrices for consideration of parameter errors according to Table 1 and for retrievals without simultaneous fitting of the temperature profile. It is the same as Fig. 5 but in the presence of parameter errors. The matrices for the whole ensemble (upper panels) show that the parameter errors reduce the sensitivity of both retrievals in the middle and upper troposphere. Additionally the logarithmic retrieval performs quite badly in the lower troposphere. For the high DOF-UT sub-ensemble the situation is similar. For both retrievals the correlations are much smaller than those observed in Fig. 5, with degradation being more pronounced in the logarithmic case. The total errors for this kind of retrievals are depicted in Fig. 9. If the whole ensemble is considered (black squares) even the retrieval of the 6.4–8.8 km layer becomes uncertain (noise/signal of 75% for linear and logarithmic retrieval). In case of the logarithmic retrieval the large error in the lower troposphere stands out. For the DOF-UT sub-ensemble the error in the 6.4–8.8 km layer is reduced to 45%. This realistic error scenario suggests that, even when only favourable days are considered, sensitivity is limited to below 8 km, and that the linear retrieval performs better in the middle and lower troposphere.

The reason for the worse performance of the logarithmic retrieval is due to the mis-

interpretations of spectral signatures as discussed above. There it was shown that the misinterpretation of a temperature error is strongly reduced by simultaneously fitting this parameter. Figures 10 and 11 show that this strategy is also successful concerning the total error. For the logarithmic retrieval the respective correlation matrices (Fig. 10) are very similar to those without additional parameter errors (Fig. 5). For the DOF-UT sub-ensemble it should now be possible to retrieve upper tropospheric water vapour amount independently from the humidity below 6 km. The linear retrieval only profits marginally from a simultaneous temperature fit. Here the slightly better correlations along the diagonal are counterbalanced by outer diagonal correlations: very high error correlation length and thus bad vertical resolution. Figure 11 depicts the respective total errors. The results demonstrate that, for a realistic error scenario and a simultaneous fit of temperature, the logarithmic retrieval performs better than the linear one. For days with high DOF-UT the water vapour content can be retrieved up to 10 km with an acceptable noise to signal ratio of 58%.

The error and sensitivity assessment reveals that a simultaneous fit of the temperature profile improves the precision of the retrievals, slightly in the linear case and strongly in the logarithmic case. Tables 2 and 3 summarize random errors for column amounts of 3 layers representing the lower troposphere (LT, 2.3–3.3 km), the middle troposphere (MT, 4.3–6.4 km), and the upper troposphere (UT, 7.6–12.4 km), and for the total column amount. Figure 12 depicts the correlations between real (assumed) amount and retrieved amount of the 3 representative layers. The correlation coefficient (ρ) and the slope (m) of the regression line are given in the panels. Black squares and black lines represent linear retrieval, and red circles and red lines logarithmic retrieval. While correlation coefficients are quite similar, the better performance of the logarithmic retrieval manifests itself by higher sensitivity (higher values of slopes m), in particular for the upper troposphere, where DOF values are generally low, and a-priori assumptions are important. The right panel of Fig. 12 illustrates the above mentioned underestimation of the linear retrieval of high water vapour amounts.

ACPD
5, 9493–9545, 2005

3.4. Systematic errors

The only systematic error sources are the spectroscopic line parameters. Thus, if the retrieval works correctly only they should provide a systematic error. This is not the case for the linear retrieval: it occasionally over- or under-constrains the solution and as a consequence it systematically underestimates both very large and very low mixing ratios. This error may be interpreted as systematic smoothing error. The systematic errors for the linear retrieval are listed in Table 4 for the three partial column amounts representing the LT, MT, and UT, and for the total column amount. They are expressed as the ratio of the mean of the error and the mean of the retrieved value. For the logarithmic retrieval care has to be taken when calculating systematic errors. Since its posterior ensemble is log-normally distributed (see also Sect. 3.5), the median rather than the mean should be considered. In this case it is more appropriate to express the systematic errors as a ratio of the median of the error and the median of the retrieved value. Table 5 lists the respective estimations. They are very similar to the linear retrieval. The systematic underestimation of high upper tropospheric amounts in the case of the linear retrieval becomes visible for the DOF-UT sub-ensemble (−10%). Under the same circumstances the systematic median for the logarithmic retrieval is −1% only.

3.5. Characterisation of posterior ensembles

On a logarithmic scale all involved pdfs are Gaussian distributions. A correctly working retrieval should therefore produce a normal pdf for the posterior ensemble, or if referred to the retransformed linear scale, a log-normal pdf. It should not change the principle distribution characteristics of the a-priori ensemble. The situation of the linear retrieval is different because it involves normal and log-normal pdfs. Consequently the posterior pdf may be something between a log-normal and normal pdf. A χ^2 test can check this issue. The posterior covariance matrix is $\mathbf{S}_{\hat{\mathbf{x}}} = \epsilon\{\hat{\boldsymbol{x}}\hat{\boldsymbol{x}}^T\}$. In contrast to the a-priori covariance matrix $\mathbf{S}_\mathbf{a}$, the matrix $\mathbf{S}_{\hat{\mathbf{x}}}$ is singular, since the solution space has fewer dimensions

ACPD
5, 9493–9545, 2005

than the a-priori space. The calculation of the χ^2 values according to Eq. (5) is thus not straightforward. However, since the covariance matrix is symmetric its singular value decomposition leads to $\mathbf{L}\mathbf{\Lambda}\mathbf{L}^T$, with the columns of $\mathbf{L}$ containing its eigenvectors and the diagonal matrix $\mathbf{\Lambda}$ its corresponding eigenvalues. As $\mathbf{S}^{-1}$ in Eq. (5) a pseudoinverse is applied, which only considers the 3 largest eigenvalues. The χ^2 calculated with this inverse would thus have 3 degrees of freedom. The test is performed for all aforementioned retrievals: with/without parameter errors and with/without simultaneous fitting of temperature. The calculations have to be performed on a logarithmic scale for the logarithmic and on a linear scale for the linear retrieval. In Fig. 13 the theoretical χ^2 cumulative distribution function (cdf) for 3 degrees of freedom (black line) is compared to the χ^2 cdf derived from the different posterior ensembles. The upper panel shows the comparison for the linear retrieval. The black squares (in the graph partially hidden by the red circles) represent the posterior ensemble when no parameter errors are assumed. 90% of all posterior vectors are now consistent with a normal distribution. This means that the linear retrieval forces the originally log-normally distributed ensemble into a Gaussian ensemble. If additional errors are present the solutions are more constrained and the empirical χ^2 values lie generally below the theoretical χ^2 values for 3 degrees of freedom (black circles). A simultaneous retrieval of the temperature enables a better exploitation of the information present in the spectra. In this case the empirical χ^2 cdf is once again close to the theoretical for 3 degrees of freedom. The lower panel of Fig. 13 shows the same for the logarithmic retrieval. In the absence of parameter errors, the characteristics of the a-priori distribution do not change. It is still a log-normal distribution (black squares). In the presence of parameter errors approximately 10% of χ^2 values of the posterior vectors are too large (black circles). Apparently, the degree of freedom is enhanced for these vectors: in the event of misinterpretation of spectral signatures the logarithmic retrieval over-interprets the information present in the spectra. Applying a retrieval with simultaneous temperature fitting reduces the difference between the empirical and the theoretical χ^2 cdf: the respective posterior ensemble is quite well described by a log-normal distribution (red circles).

ACPD
5, 9493–9545, 2005

In the case of misinterpretation of spectral signatures the logarithmic retrieval overinterprets spectral signatures. This becomes apparent by comparing the DOF values (tr($\hat{\mathbf{A}}$)) for a retrieval with and without additional parameter errors. If the retrieval is working correctly adding further errors should reduce the DOF value, since the information in the spectra is more uncertain. However, on a logarithmic scale occasionally the contrary is observed. Figure 14 compares the DOF values for the logarithmic retrievals with and without additional errors. If the temperature profile is not simultaneously fitted (left panel) occasionally more information is retrieved from the erroneous spectra than from the spectra with only white noise, which means that errors in the spectra are misinterpreted as information. This problem disappears by fitting the temperature profile simultaneously (right panel).

4. Comparison of retrieval results to ptu-sonde measurements

4.1. The FTIR measurements

Since March 1999 measurements of highly-resolved infrared solar absorption spectra are routinely performed at the Izaña Observatory, situated on the Canary Island of Tenerife (28°18′ N, 16°29′ W) at 2370 m a.s.l. Its position in the Atlantic Ocean and above a stable inversion layer, typical for subtropical regions, provides clean air and clear sky conditions most of the year. This offers good conditions for atmospheric observations by remote sensing techniques. The spectra are obtained by a Bruker IFS 120M applying a resolution of 0.0036 to 0.005 cm^{-1} and no numerical apodisation. The spectral intensities are determined by a liquid-nitrogen cooled HgCdTe detector, which, in order to ensure linearity, is operated in a photovoltaic mode. During short periods in 1999 and 2001 a photoconductive detector was applied whose nonlinearities were corrected. The spectra are typically constructed by co-adding up to 8 scans recorded in about 10 or 13 min, depending on their resolution. Analysing the shape of the absorption lines (lines are widened by pressure broadening) and their different temperature

ACPD
5, 9493–9545, 2005

sensitivities enables the retrieval of the absorbers' vertical distribution. Since the instrumental line shape (ILS) also affects the shape of the measured absorption lines, this instrumental characteristic should be determined independently from the atmospheric measurements. This is done on average every two months using cell measurements and LINEFIT software as described in Hase et al. (1999). The temperature and pressure profiles, necessary for the inversion, are taken from the synoptical meteorological 12:00 UT sondes. Above 30 km data from the Goddard Space Flight Center's automailer system are applied. Some results of these measurements are presented in Schneider et al. (2005) and references therein.

4.2. The radiosonde measurements

Until September 2002 the meteorological soundings were launched from Santa Cruz de Tenerife, 35 km northeast of the observatory, and since October 2002 in an automised mode from Güimar, 15 km southeast of the observatory. The sondes are equipped with a Vaisala RS80-A thin-film capacitive sensor which determines relative humidity. The sonde data are corrected by a method suggested by Leiterer et al. (2004), who reported a remaining random error of less than 5% throughout the troposphere. Other authors report correction methods with a remaining uncertainty of over 10% (Miloshevich, 2001). Furthermore, the precision of the water vapour measured by the RS80-A sensor may be degraded due to chemical contamination during storage. To avoid sondes with iced detectors, sondes that passed through clouds are not taken into account. Therefore sondes which detect a vapour pressure close to the liquid or ice saturation pressure are disregarded. Furthermore, sondes with unrealistic high humidities above 10 km, which may indicate an iced detector, are excluded. The corrected sonde mixing ratios are finally sampled on the altitude grid of the retrieval by requiring that linear interpolation of the mixing ratios between two grid levels yield the same partial columns as the original highly-resolved data.

ACPD
5, 9493–9545, 2005

4.3. Temporal and spatial variability

The large temporal and spatial variabilities of atmospheric water vapour are problematic when measurements conducted from different platforms are to be compared. Both experiments should be conducted at the same time and sound the same atmospheric location. For this reason only sonde measurements coinciding within 2 h of the FTIR measurements are used for the comparison. Spatial coincidence is difficult to achieve. The sonde measures in-situ and will always be situated at a certain distance from the imaginary line between the FTIR instrument and the sun. This is particularly problematic for the lowest layer above the FTIR instrument as, while the FTIR instrument is located at the surface the sonde is typically floating around 30 km south of the observatory in the free troposphere. A comparison between the humidity measured in-situ at the observatory and the sonde's humidity demonstrated that the water vapour amounts close to the surface are more variable and on average 40% larger than those in the free troposphere.

4.4. Comparison

Within the comparison period, ranging from March 1999 to January 2004, the criterions for sonde quality (no clouds, realistic humidity above 10 km) and temporal coincidence with FTIR measurements are fulfilled in 157 occasions only. 50 of them belong additionally to the high DOF sub-ensemble. In Fig. 15 correlation matrices of FTIR and sonde profiles are presented. They are the experimental analogue to the simulated correlations shown in Fig. 10. The upper panels show the situation for all coincidences and the lower panels for those when additionally favourable upper tropospheric conditions are expected. Considering all situations the linear retrieval is apparently more consistent with the sonde measurements than the logarithmic retrieval, since it has larger ρ values along the diagonal of the matrix. The degraded performance of the logarithmic retrieval may be due to a slight misinterpretation of an incorrect ILS characterisation. As seen in Fig. 7 the phase error is similar to the temperature error and may

ACPD
5, 9493–9545, 2005

cause similar problems if the assumptions of Table 1 are too optimistic for the applied Bruker IFS 120M spectrometer. However, it should be considered that the linear retrieval has large outer diagonal elements, in particular above 5 km. For the logarithmic retrieval, on the other hand, large correlation coefficients are quite well centred around the diagonal, which counterbalances the lower diagonal values, since it means that the correlation lengths towards sonde mixing ratios are smaller compared to those of the linear retrieval. This is a consequence of the poorer vertical resolution of the latter (see explanations about smoothing error in Sect. 3), and even more important considering the situation of the upper troposphere for the DOF-UT coincidences. Here the ρ values on the diagonal are quite similar for the linear and logarithmic retrieval. However, the logarithmic solutions above 9 km are much less correlated with the sonde measurements around 7 km. For example, the state retrieved at 10 km has a ρ value with the real state at 7 km of 0.78 in the linear and 0.65 in the logarithmic case only. Therefore, the variabilities of the amount detected by the sonde for the UT layer (7.6–12.4 km) should be more consistent with the variabilities of the logarithmic retrieval than with those of the linear retrieval. This is confirmed by Tables 6 and 7, which list the mean and standard deviation of the difference between sonde and FTIR measurements. The standard deviation describes the level of consistency between the variabilities detected by the sonde and the FTIR measurements. It may also be seen as overall precision of FTIR and sonde experiments together. For the UT layer and considering the coincidences with high DOF-UT values only, it is 76% for the linear and 50% for the logarithmic retrieval. These calculations even disregard the random errors of the sonde measurements and temporal and spatial mismatching of both measurements. The values are – at least qualitatively – well consistent with the simulations in Sect. 3, where the total precision of the FTIR measurements is estimated as 89% for the linear and 58% for the logarithmic retrieval (see total error in Tables 2 and 3). The measurements are made with a Bruker IFS 120M. Since the ILS of this instrument is somehow instable, the logarithmic retrieval may be improved even further by a simultaneous retrieval of the ILS. This would eliminate possible misinterpretations of ILS errors, in a similar

ACPD
5, 9493–9545, 2005

way as the simultaneous temperature retrieval prevents the misinterpretation of temperature errors. Only the logarithmic retrieval enables detection of UT water vapour variabilities. The retrieval on a linear scale performs too poorly for this objective.

An outstanding difference between Tables 6 and 7 and Tables 2 and 3 is the poorer consistency for the LT layer of FTIR when compared to sonde than when compared within the simulations: empirical standard deviation of ≈45% compared to the estimated values of ≈22%. This is due to the aforementioned different conditions in the lowermost layer above the instrument (surface influences) and the corresponding layer at the sonde (free troposphere). For the same reason in the LT layer the experimentally observed systematic differences are much larger than the simulated systematic errors: the LT at the site of the instrument is more humid than the free tropospheric LT. Since the LT mainly determines the total column amount, the latter is largely affected by these differences. The estimated and empirically observed precision for the MT are highly consistent (noise to signal around 30%).

The different conditions in the lowermost layer for both experiments make it difficult to decide whether the estimations about the systematic errors – accuracies (Tables 4 and 5) – are consistent with the mean differences in Tables 6 and 7 or not. The observation of an increased positive difference for FTIR-sonde in the upper troposphere compared to the middle troposphere may have two explanations. First it may manifest the known dry bias of sonde measurements (Turner et al., 2003) or second it may be due to a pressure broadening coefficient, which is systematically too low.

Figure 16 shows the correlation between LT, MT, and UT partial column amounts of FTIR and sonde measurements. The greatest differences with Fig. 12 are observed for the LT (as discussed above). For the MT and UT the consistency between the simulations and the empirical observations is excellent, even though the errors of the sonde measurements and temporal and spatial mismatching are still disregarded. The UT amounts of the sonde and the logarithmic retrieval show a good linear correlation with a slope of the regression line of 1.04, whereby for the linear retrieval a systematic underestimation of high water vapour amounts is observed (right panel). This is con-

ACPD
5, 9493–9545, 2005

sistent with the simulations (right panel of Fig. 12). The empirical comparison of FTIR and sonde data suggests that the FTIR system is even more sensitive than proposed by the theoretical study performed in Sect. 3. This is reflected in the larger slopes m of the regression lines. The differences are especially pronounced in the UT when only days with high DOF-UT are considered (for logarithmic retrieval simulated m of 0.68 and observed m of 1.04). An explanation is that, at instrument altitude (2.3 km), a mixing ratio determined by an in-situ instrument was applied for the simulation. This relatively high value is then spread out up to the next grid point (3.3 km). However, the enhanced humidity due to surface conditions is very likely limited to the lowest 100 m of the atmosphere. This overestimation of simulated LT amounts reduces the estimated sensitivity in the UT (high DOF-UT values are correlated to lower tropospheric slant columns).

5. Subtropical water vapour time series

Figure 17 depicts a more than six year record of tropospheric water vapour amounts as determined by the logarithmic retrieval with simultaneous fitting of the temperature. While for the lower and middle tropospheric values all measurement days are depicted, the upper tropospheric values are presented only when the DOF-UT values exceeded 0.2. For the lower and middle troposphere a well pronounced seasonal cycle is observed. Values are highest at the end of summer and lowest in the winter months. A similar clear seasonal dependence is not observed for the upper tropospheric amounts. Values are sometimes even especially high in autumn/winter, which demonstrates their independence from lower tropospheric levels. A quick view may give the impression of increasing water vapour contents in the upper troposphere; however, for a serious trend analysis a longer time series would be needed.

ACPD
5, 9493–9545, 2005

6. Summary and conclusions

Compared to other atmospheric components, the retrieval of atmospheric water vapour from ground-based FTIR measurements has additional difficulties. Water vapour has very large vertical gradients and variabilities, which generally limit the sensitivity of the ground-based technique to the lower and middle troposphere. The spectral signatures originating from the upper troposphere are rather weak and thus their retrieved values depend to an important extent on a-priori assumptions. Water vapour mixing ratios are log-normally distributed and an inversion on a logarithmic scale enables the correct application of this a-prior knowledge and consequently leads to a statistically optimal retrieval. However, this method introduces the risk of misinterpreting spectral signatures produced by errors in assumed model parameters. It is shown that the misinterpretations can be controlled by simultaneously fitting the temperature profile. A logarithmic retrieval should therefore perform better than the commonly applied linear retrieval, in particular in the upper troposphere. For a realistic error scenario an improvement of around 30% for the respective noise to signal ratio is estimated. This is confirmed by a comparison to sonde measurements. While the linear retrieval leads to a noise to signal ratio of around 75%, the logarithmic retrieval provides a ratio of 50%. Lower and middle tropospheric amounts are detectable with precisions (noise to signal ratio) of 20 and 30%, respectively.

The advantage of the FTIR technique compared to the meteorological sondes is that the errors are well understood and water isotope evaluation is possible. This may allow a study of hydrometeorological processes in the atmosphere.

The suggested method can be applied to other dataset of highly-resolved infrared spectra (e.g. to measurements made within the Network for Detection of Stratospheric Change). However, the capability of the method would have to be investigated for each measurement site individually. For upper tropospheric sensitivity, a good characterisation of the instrumental line shape (phase error), reliable temperature profile data, and the absorption lines being unsaturated are required. The upper tropospheric sensitivity

ACPD
5, 9493–9545, 2005

is expected to be better the lower the water vapour content in the lowest layers and the stabler the instrumental line shape. In this context the subtropical site of Izaña, located on an island, and the application of a Bruker IFS 120M are surely not the best conditions. For mid-latitudinal alpine stations or subpolar and polar stations equipped, for instance, with a Bruker IFS 120HR even better sensitivities should be expected.

Acknowledgements. We would like to thank the Bundesministerium für Bildung und Forschung for funding via the DLR (contracts 50EE0008 and 50EE0203). Furthermore, we are grateful to the Izaña Observatory for facilitating the sonde data and for allowing us to use its infrastructure and to the Goddard Space Flight Center for providing the temperature and pressure profiles of the National Centers for Environmental Prediction via the automailer system.

References

Harries, J. E.: Atmospheric radiation and atmospheric humidity, Q. J. R. Meteorol. Soc., 123, 2173–2186, 1997. 9494

Hase, F., Blumenstock, T., and Paton-Walsh, C.: Analysis of the instrumental line shape of high-resolution Fourier transform IR spectrometers with gas cell measurements and new retrieval software, Appl. Opt., 38, 3417–3422, 1999. 9506, 9513

Hase, F., Hannigan, J. W., Coffey, M. T., Goldman, A., Höpfner, M., Jones, N. B., Rinsland, C. P., and Wood, S. W.: Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements, J. Quant. Spectrosc. Radiat. Transfer, 87, 25–52, 2004. 9499

Höpfner, M., Stiller, G. P., Kuntz, M., Clarmann, T. v., Echle, G., Funke, B., Glatthor, N., Hase, F., Kemnitzer, H., and Zorn, S.: The Karlsruhe optimized and precise radiative transfer algorithm, Part II: Interface to retrieval applications, SPIE Proceedings 1998, 3501, 186–195, 1998. 9499

Kuntz, M., Höpfner, M., Stiller, G. P., Clarmann, T. v., Echle, G., Funke, B., Glatthor, N., Hase, F., Kemnitzer, H., and Zorn, S.: The Karlsruhe optimized and precise radiative transfer algorithm, Part III: ADDLIN and TRANSF algorithms for modeling spectral transmittance and radiance, SPIE Proceedings 1998, 3501, 247–256, 1998. 9499

Kurylo, M. J.: Network for the detection of stratospheric change (NDSC), Proc. SPIE–Int. Co. Opt. Eng. 1991, 1491, 168–174, 1991. 9495

Kurylo, M. J. and Zander, R.: The NDSC – Its status after 10 years of operation, Proceedings of XIX Quadrennial Ozone Symposium, Hokkaido University, Sapporo, Japan, 167–168, 2000. 9495

Leiterer, U., Dier, H., Nagel, D., Naebert, T., Althausen, D., Franke, K., Kats, A., and Wagner, F.: Correction Method for RS80-A Humicap Humidity Profiles and their Validation by Lidar Backscattering Profiles in Tropical Cirrus Clauds, J. Atmos. Oceanic Technol., 22, 18–29, 2005. 9495, 9513

Miloshevich, L. M., Vömel, H., Paukkunen, A., Heymsfield, A. J., and Oltmans, S. J.: Characterization and correction of relative humidity measurements from Viasalla RS80-A radiosondes at cold temperatures, J. Atmos. Oceanic Technol., 18, 135–155, 2001. 9495, 9513

NDSC: http://www.ndsc.ws/, 2005. 9495

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Praxis, World Scientific Publishing Co., Singapore, 2000. 9497, 9500, 9504

Rothman, L. S., Barbe, A., Benner, D. C., Brown, L. R., Camy-Peyret, C., Carleer, M. R., Chance, K. V., Clerbaux, C., Dana, V., Devi, V. M., Fayt, A., Fischer, J., Flaud, J.-M., Gamache, R. R., Goldman, A., Jacquemart, D., Jucks, K. W., Lafferty, W. J., Mandin, J.-Y., Massie, S. T., Newnham, D. A., Perrin, A., Rinsland, C. P., Schroeder, J., Smith, K. M., Smith, M. A. H., Tang, K., Toth, R. A., Vander Auwera, J., Varanasi, P., and Yoshino, K.: The HITRAN Molecular Spectroscopic Database: Edition of 2000 Including Updates through 2001, J. Quant. Spectrosc. Radiat. Transfer, 82, 5–44, 2003. 9500

Schneider, M., Blumenstock, T., Chipperfield, M., Hase, F., Kouker, W., Reddmann, T., Ruhnke, R., Cuevas, E., and Fischer, H.: Subtropical trace gas profiles determined by ground-based FTIR spectroscopy at Izaña (28°, 16°): Five year record, error analysis, and comparison with 3D-CTMs, Atmos. Chem. Phys., 5, 153–167, 2005, SRef-ID: 1680-7324/acp/2005-5-153. 9496, 9513

SPARC: Assessment of Upper Tropospheric and Stratospheric Water Vapour, edited by: Kley, D., Russell III, J. M., and Phillips, C., WCRP-113, WMO/TD-No. 1043, SPARC report No. 2, December, 2000. 9495

Spencer, R. W. and Braswell, W. D.: How dry is the tropical free troposphere? Implications for global warming theory, Bull. Am. Meteorol. Soc., 78, 1097–1106, 1997. 9494

Stiller, G. P., Höpfner, M., Kuntz, M., Clarmann, T. v., Echle, G., Fischer, H., Funke, B., Glatthor,

N., Hase, F., Kemnitzer, H., and Zorn, S.: The Karlsruhe optimized and precise radiative transfer algorithm, Part I: Requirements, justification and model error estimation, SPIE Proceedings 1998, 3501, 257–268, 1998. 9500

Turner, D. D., Lesht, B. M., Clough, S. A., Liljegren, J. C., Revercomb, H. E., and Tobin, D. C.: Dry Bias and Variability in Vaisala RS80-H Radiosondes: The ARM Experience, J. Atmos. Oceanic Technol., 20, 117–132, 2003. 9516

Wagner, G., Birk, M., Schreier, F., and Flaud, J.-M.: Spectroscopic database for Ozone in the fundamental spectral regions, J. Geophys. Res., 107, 4626–4643, 2002. 9500

Table 1. Assumed uncertainties.

error source	uncertainty
measurement noise	S/N of 500
phase error	0.02 rad
modulation eff.	2%
T profile[a]	up to 2.5 K at surface 1 K rest of troposphere
solar angle	0.1°
line intensity	5%
pres. broad. coef.	1%

[a] detailed description see text

ACPD
5, 9493–9545, 2005

Table 2. Estimated noise/signal of linear retrieval with simultaneous fitting of temperature [%]. The value in brackets for the 7.6–12.4 km layer corresponds to the high DOF-UT sub-ensemble.

error source	2.3–3.3 km	4.3–6.4 km	7.6–12.4 km	total
smoothing	14	24	87 (81)	3
meas. noise	3	2	10 (9)	1
pha. err.	11	8	27 (26)	2
mod eff.	1	1	<1 (<1)	<1
T. profile	4	2	14 (11)	2
solar angle	<1	1	<1 (<1)	1
line int.	5	5	4 (3)	5
pres. coef.	14	10	21 (13)	1
total	**22**	**31**	**123 (89)**	**6**

ACPD
5, 9493–9545, 2005

Table 3. Same as Table 2, but for logarithmic retrieval.

error source	2.3–3.3 km	4.3–6.4 km	7.6–12.4 km	total
smoothing	11	21	85 (55)	2
meas. noise	4	3	10 (9)	1
pha. err.	18	10	33 (29)	2
mod eff.	1	1	1 (<1)	<1
T. profile	8	6	19 (14)	1
solar angle	<1	1	<1 (<1)	1
line int.	4	5	4 (4)	5
pres. coef.	28	14	38 (16)	2
total	**23**	**29**	**128 (58)**	**7**

Table 4. Estimated systematic errors of linear retrieval [%]. The value in bracket for the 7.6–12.4 km layer corresponds to the high DOF-UT sub-ensemble.

error source	2.3–3.3 km	4.3–6.4 km	7.6–12.4 km	total
smoothing	1	−1	0 (−10)	0
line int.	−4	−5	−1 (−3)	−5
pres. coef.	−16	+14	−19 (−7)	−1

ACPD

5, 9493–9545, 2005

Table 5. Same as Table 4, but for logarithmic retrieval.

error source	2.3–3.3 km	4.3–6.4 km	7.6–12.4 km	total
smoothing	0	+3	+3 (−1)	0
line int.	−4	−5	−1 (−3)	−5
pres. coef.	−16	+13	−19 (−9)	−1

Table 6. Differences between sonde and FTIR column amounts for linear retrieval ((FTIR-sonde)/FTIR)[%]. The value in bracket for the 7.6–12.4 km layer corresponds to the high DOF-UT sub-ensemble.

	2.3–3.3 km	4.3–6.4 km	7.6–10.0 km	total
std	41	32	121 (76)	26
mean	+27	+3	+22 (+22)	+14

ACPD
5, 9493–9545, 2005

Table 7. Same as Table 6 but for logarithmic retrieval.

	2.3–3.3 km	4.3–6.4 km	7.6–10.0 km	total
std	47	33	100 (50)	26
mean	+23	+8	+25 (+32)	+14

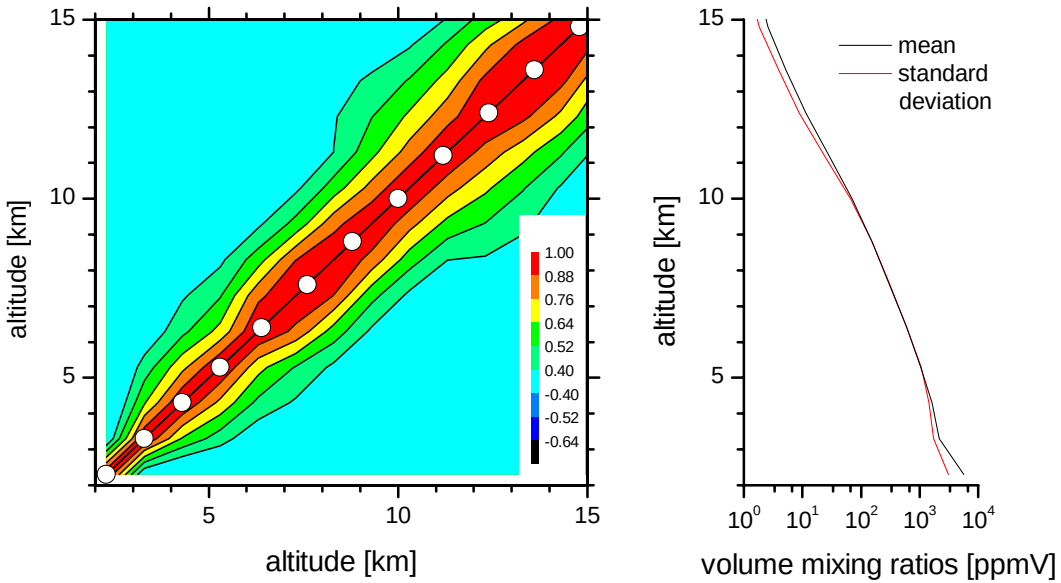

Fig. 1. Description of a-priori state. Left panel: correlation matrix. Right panel: black line: mean state; red line: standard deviation of mean state.

ACPD

5, 9493–9545, 2005

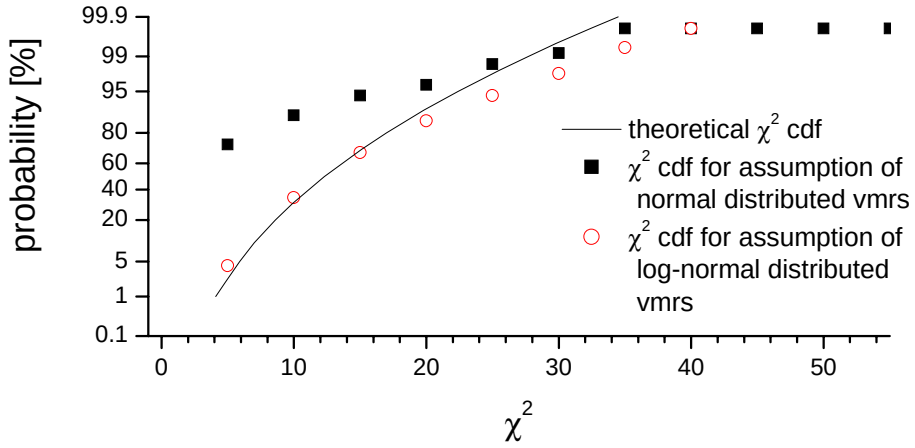

Fig. 2. χ^2 test for different of a-priori assumptions. Black line: theoretical χ^2 cumulative distribution function (cdf); black filled squares: χ^2 cdf of ensemble for assumed normal pdf on a linear scale; red circles: χ^2 cdf of ensemble for assumed normal pdf on a logarithmic scale.

ACPD
5, 9493–9545, 2005

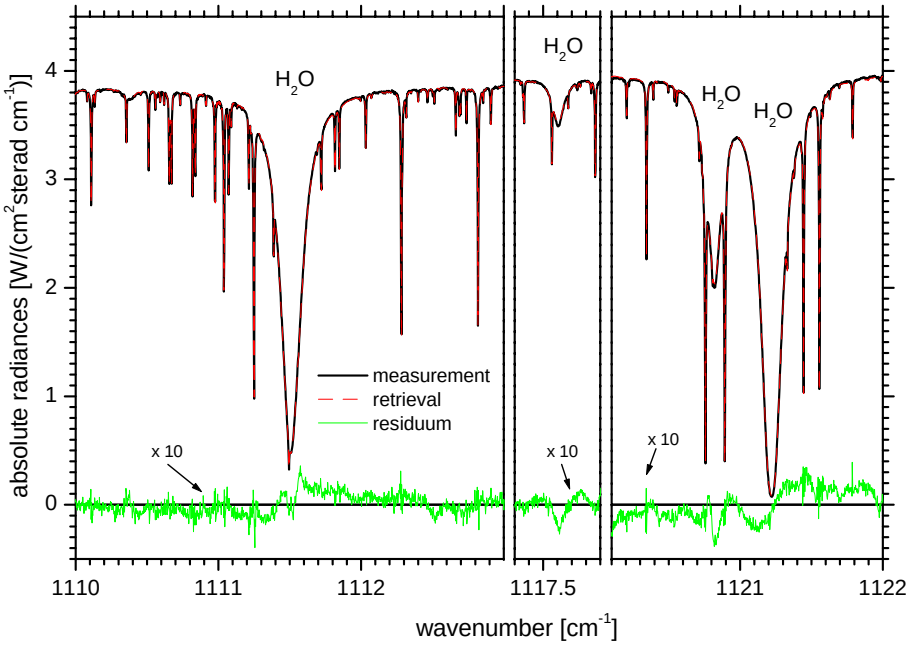

Fig. 3. Spectral regions applied for retrieval. Plotted is the situation for a real measurement taken on 10 March of 2003 (solar elevation angle 50°). Black line: measured spectrum; red line: simulated spectrum; green line: difference between simulation and measurement.

ACPD
5, 9493–9545, 2005

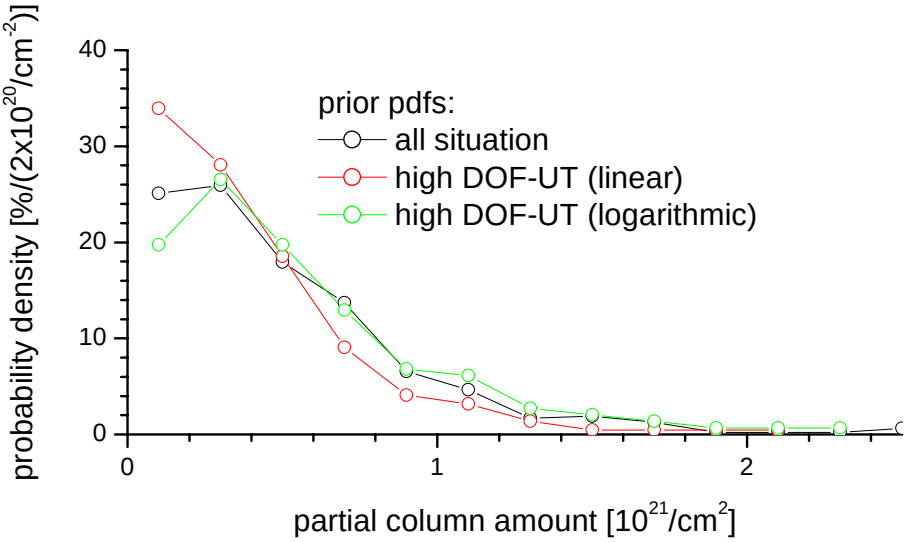

Fig. 4. Prior pdfs for the upper tropospheric column amounts (7.6–12.4 km) for different ensembles. Black line: whole ensemble, i.e. a-priori pdf. Red line: high DOF-UT ensemble according to linear retrieval Green line: high DOF-UT ensemble according to logarithmic retrieval.

ACPD
5, 9493–9545, 2005

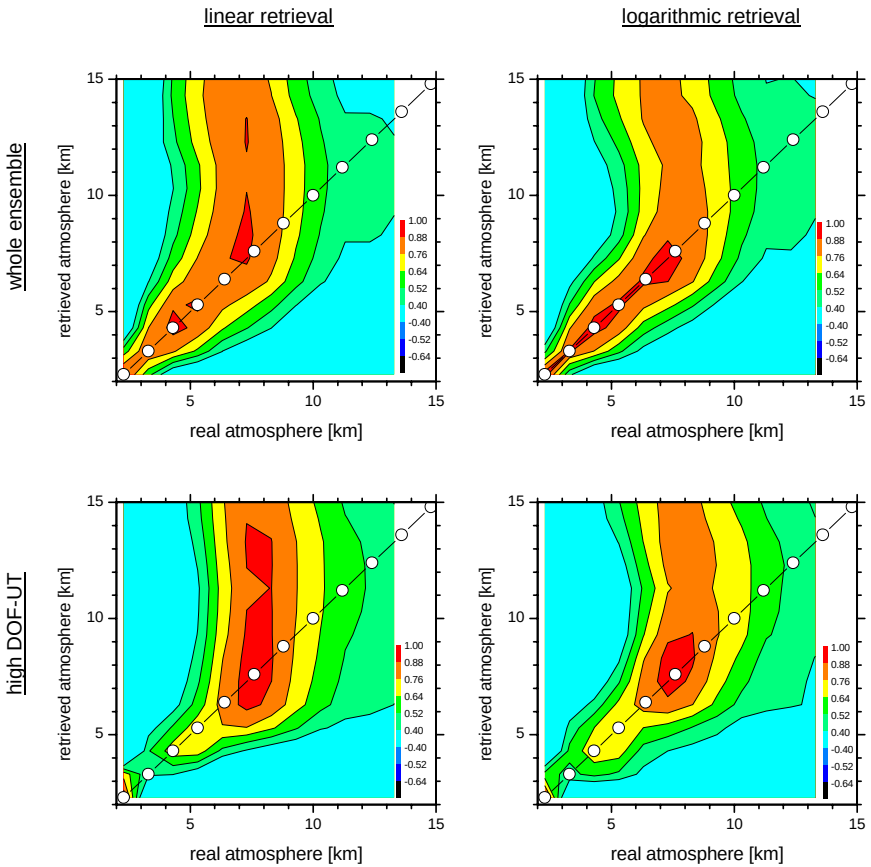

Fig. 5. Sensibility of observing system in the absence of parameter error. Depicted are correlation matrices between assumed real state and retrieved state. Left panels for retrieval on a linear scale, right panels for retrieval on a logarithmic scale. Upper panels for the whole ensemble, lower panels for the high DOF-UT sub-ensemble. Colors mark the values of the correlation coefficients (ρ) as given in legend.

ACPD
5, 9493–9545, 2005

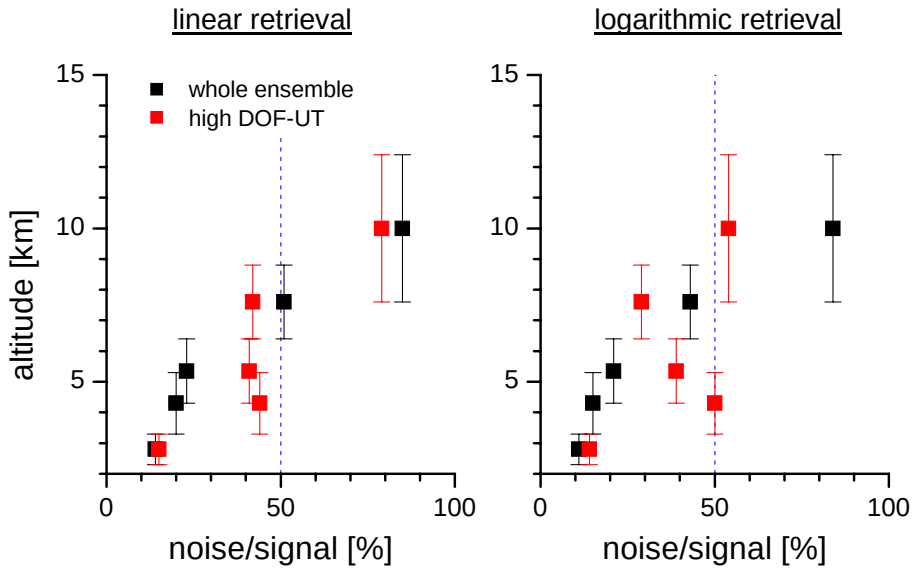

Fig. 6. Smoothing errors in the retrieved profiles. Left panel: linear retrieval. Right panel: logarithmic retrieval. Colors as described in legend.

ACPD
5, 9493–9545, 2005

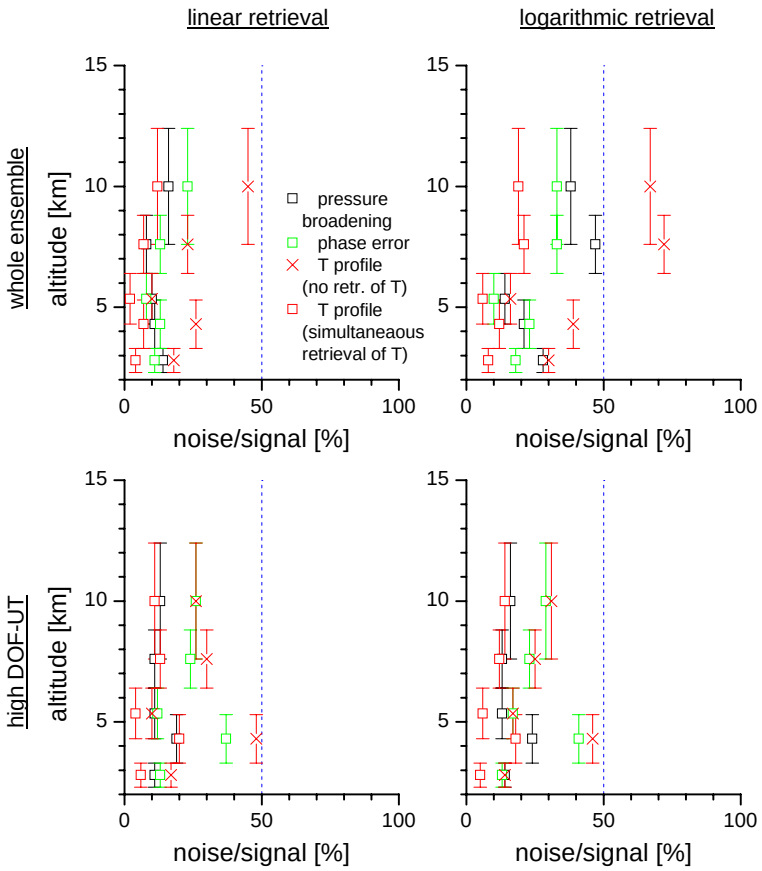

Fig. 7. Parameter errors in the retrieved profiles. Upper panels: for the whole ensemble. Bottom panels: for the DOF-UT sub-ensemble. Left panels: for the linear retrieval. Right panels: for the logarithmic retrievals. Symbols as described in the legend.

ACPD
5, 9493–9545, 2005

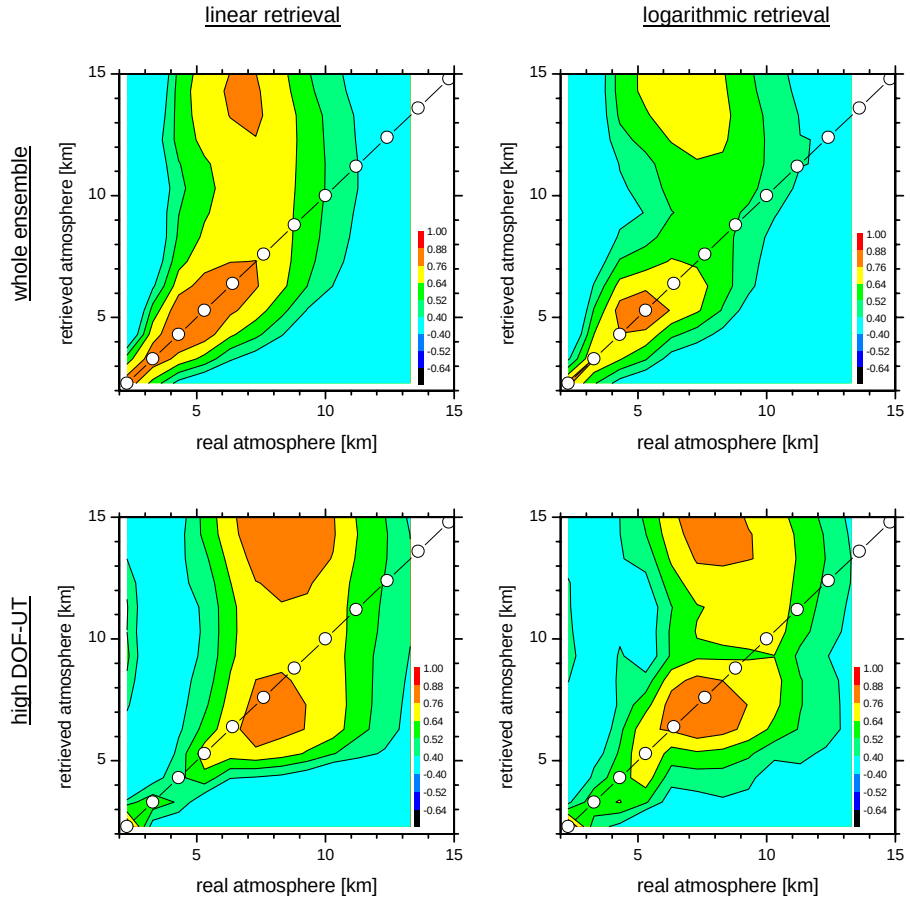

Fig. 8. Same as Fig. 5 but in the presence of parameter error as listed in Table 1.

ACPD
5, 9493–9545, 2005

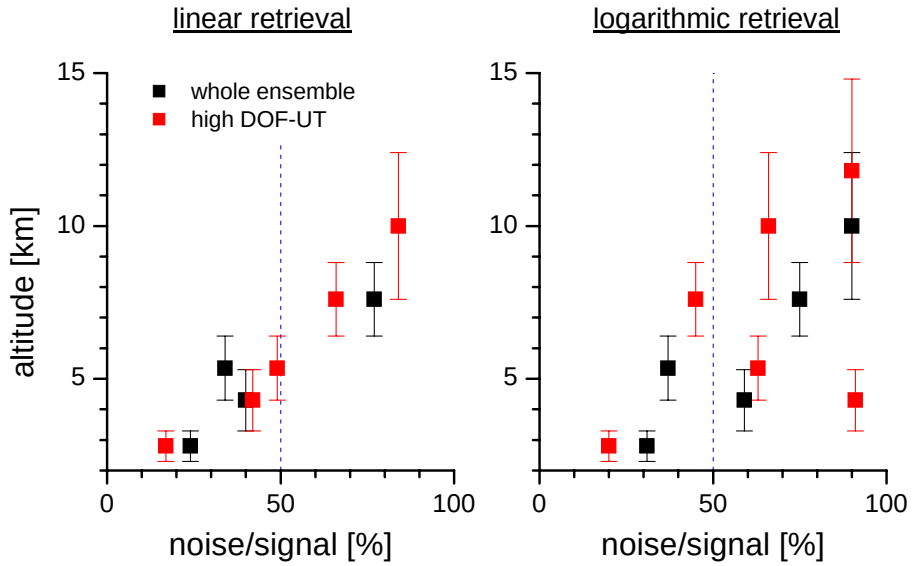

Fig. 9. Same as Fig. 6 but in the presence of parameter error as listed in Table 1.

ACPD
5, 9493–9545, 2005

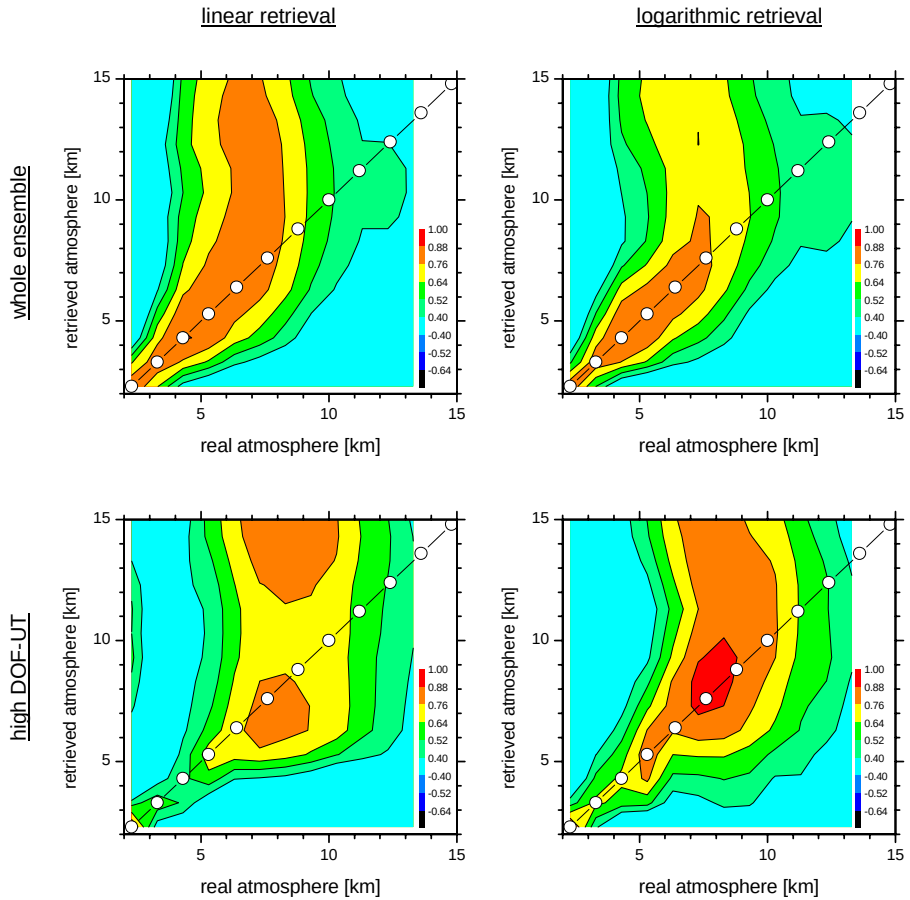

Fig. 10. Same as Fig. 8 but with simultaneous retrieval of temperature profile.

ACPD
5, 9493–9545, 2005

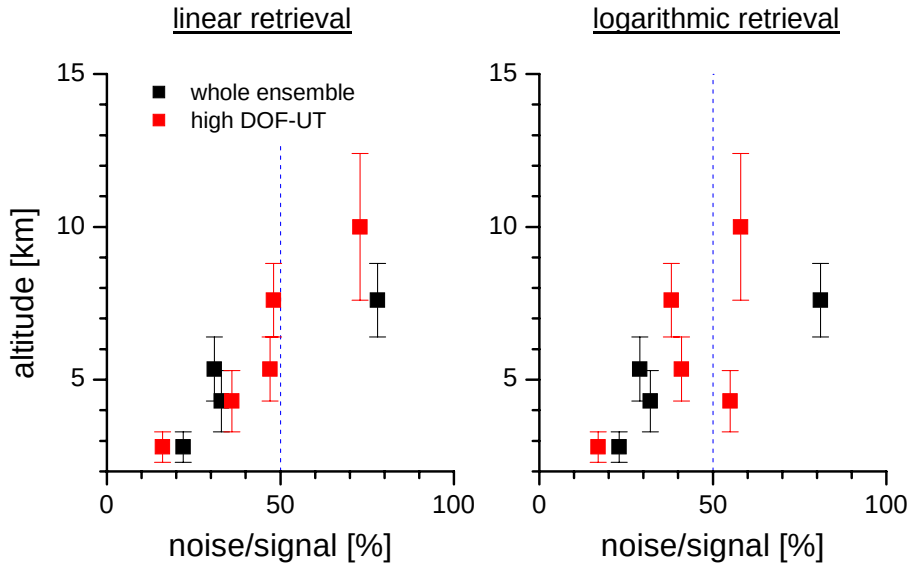

Fig. 11. Same as Fig. 9 but with simultaneous retrieval of temperature profile.

ACPD
5, 9493–9545, 2005

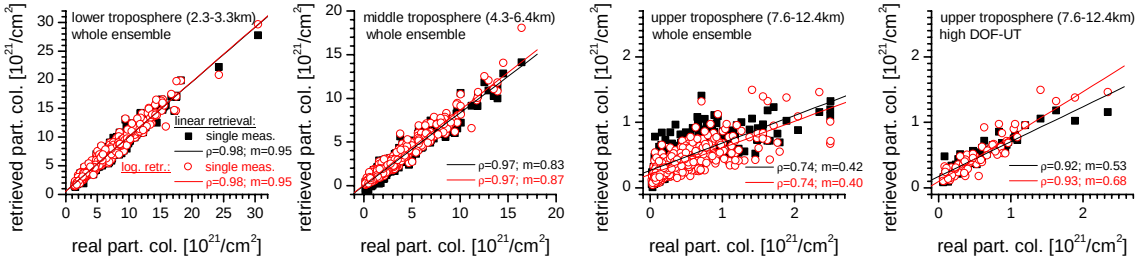

Fig. 12. Correlations between assumed real partial column amounts and their corresponding retrieved amounts for simultaneous retrieval of temperature profile. Shown are from the left to the right: lower troposphere, middle troposphere, upper troposphere (all for whole ensemble), and upper troposphere (for high DOF-UT sub-ensemble). Black squares and black lines: retrieval on a linear scale and corresponding regression line. Red circles: retrieval on a logarithmic scale and corresponding regression line.

ACPD
5, 9493–9545, 2005

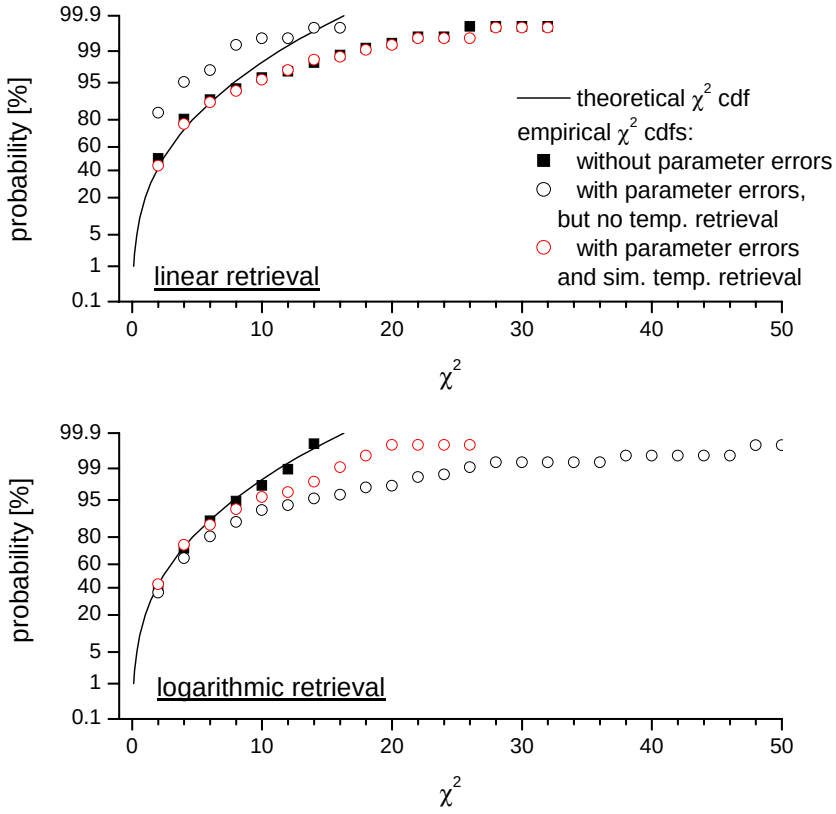

Fig. 13. χ^2 test for posterior ensembles. Upper panel: linear retrieval. Lower panel: logarithmic retrieval. Black line: theoretical theoretical χ^2 cumulative distribution function (cdf) for 3 degrees of freedom; black filled squares: empirical χ^2 cdf of ensemble in absence of parameter errors; black circles: empirical χ^2 cdf of ensemble in the presence of parameter errors and without retrieval of temperature profile; red circles: empirical χ^2 cdf of ensemble in the presence of parameter errors and simultaneous retrieval of temperature profile.

ACPD
5, 9493–9545, 2005

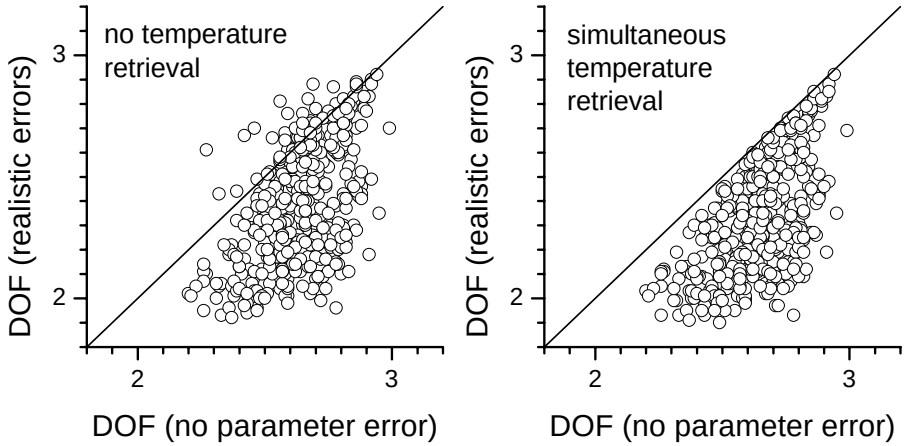

Fig. 14. DOF values for logarithmic retrievals with realistic error assumptions compared to DOF values of logarithmic retrieval in the absence of errors. Left panel: no retrieval of temperature profile. Right panel: simultaneous retrieval of temperature profile.

ACPD
5, 9493–9545, 2005

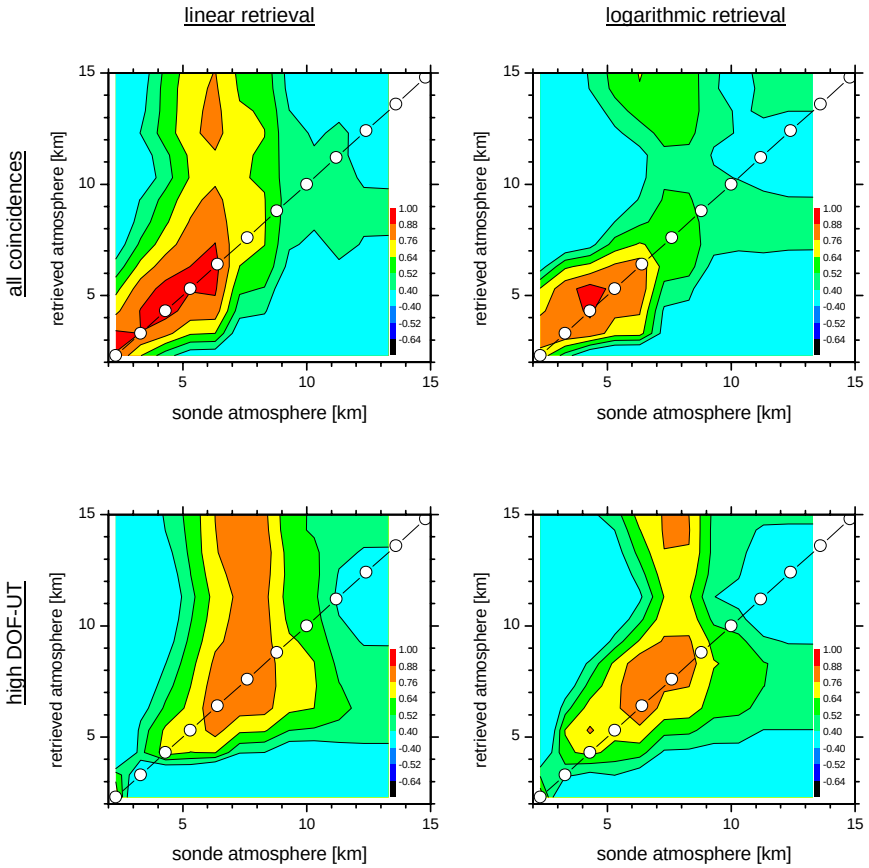

Fig. 15. Correlation between FTIR and sonde profiles for retrievals with simultaneous fitting of temperature. Depicted are correlation matrices. Left panels for retrieval on a linear scale, right panels for retrieval on a logarithmic scale. Upper panels: all coincidences. Lower panels: Coincidences with high DOF-UT. The colours mark the values of the correlation coefficients (ρ) as given in the legend.

ACPD
5, 9493–9545, 2005

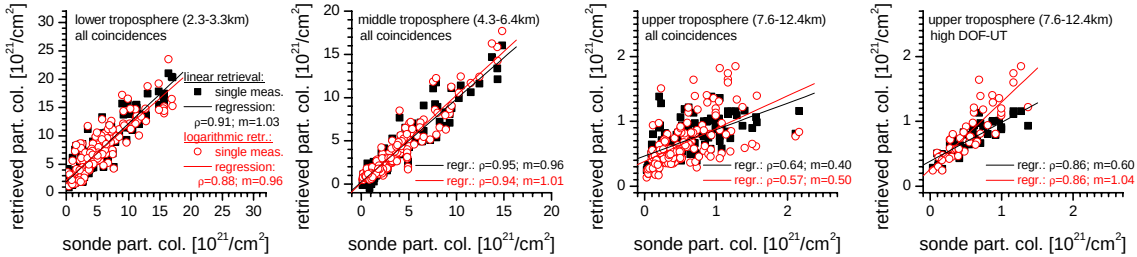

Fig. 16. Correlations between sonde and FTIR partial column amounts for simultaneous retrieval of temperature profile. Shown from left to right: lower troposphere, middle troposphere, upper troposphere (for all coincidences), and upper troposphere (for coincidences with high DOF-UT). Black squares and black lines: retrieval on a linear scale and corresponding regression line. Red circles: retrieval on a logarithmic scale and corresponding regression line.

ACPD
5, 9493–9545, 2005

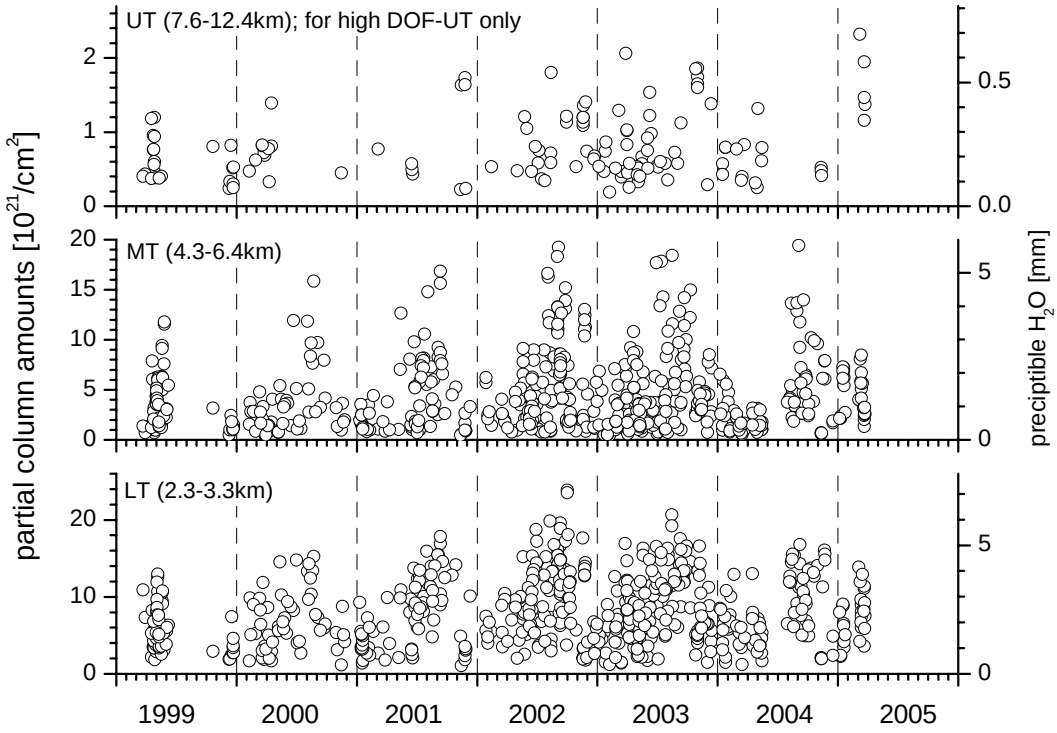

Fig. 17. Time series of water vapour above Tenerife island determined from FTIR measurements. Upper panel: upper tropospheric column amounts (7.6–12.4 km). Middle panel: middle tropospheric column amounts (4.3–6.4 km). Bottom panel: lower tropospheric column amounts (2.3–3.3 km).

ACPD
5, 9493–9545, 2005