Atmos. Meas. Tech. Discuss., 4, 7159–7183, 2011
www.atmos-meas-tech-discuss.net/4/7159/2011/
doi:10.5194/amtd-4-7159-2011


Atmospheric Measurement Techniques Discussions



# How to average logarithmic retrievals

**B. Funke**[1] **and T. von Clarmann**[2]

[1]Instituto de Astrofísica de Andalucía, CSIC, Granada, Spain
[2]Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research, Karlsruhe, Germany

Received: 18 November 2011 – Accepted: 23 November 2011 – Published: 5 December 2011

Correspondence to: B. Funke (bernd@iaa.es)

Published by Copernicus Publications on behalf of the European Geosciences Union.

## Abstract

Calculation of mean trace gas contributions from profiles obtained by retrievals of the logarithm of the abundance rather than retrievals of the abundance itself are prone to biases. By means of a system simulator, biases of linear versus logarithmic averaging were evaluated for both maximum likelihood and maximum a priori retrievals, for various signal to noise ratios and atmospheric variabilities. These biases can easily reach several ten percent. As a rule of thumb we found for maximum likelihood retrievals that linear averaging better represents the true mean value in cases of large local natural variability and high signal to noise ratios, while for small local natural variability logarithmic averaging often is superior. In the case of maximum a posteriori retrievals, the mean is dominated by the a priori information used in the retrievals and the method of averaging is of minor concern. For larger natural variabilities, the appropriateness of the one or the other method of averaging depends on the particular case because the various biasing mechanisms partly compensate in a hardly predictable manner. This complication arises mainly because of the fact that in logarithmic retrievals the weight of the prior information depends on abundance of the gas itself. No simple rule was found on which kind of averaging is superior, and instead of suggesting simple recipes we cannot do much more than to create awareness of the traps related with averaging of mixing ratios obtained from logarithmic retrievals.

## 1 Introduction

Retrieval of mixing ratios or concentrations of atmospheric trace species from remote radiance or transmission measurements involves inverse modelling of radiative transfer. In order to avoid to retrieve negative thus unphysical mixing ratios of trace species, to better cope with the large dynamical range of possible values, or to better reflect the assumed natural distribution of the species under assessment, often the logarithm of the concentration is retrieved instead of the concentration itself (e.g., von Clarmann



AMTD

4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann

et al., 2009; Funke et al., 2009; Papandrea et al., 2005; Bowman et al., 2006; Schneider et al., 2006; Urban et al., 2005). While negative concentrations certainly are unphysical, their removal by the logarithmic retrieval may bias averages of retrieved concentrations high due to the asymmetric error propagation. In this paper we assess if it is appropriate to average results in the logarithmic domain in order to reduce these biases. This analysis is done by means of a system simulator which propagates signal and noise through an idealized retrieval and which is described in Sect. 2. In Sect. 3 we analyze related case studies, and in Sect. 4 we give recommendations which kind of averaging is advisable in which context and critically discuss to which degree the conclusions of this study can be generalized towards a wider range of applications beyond the idealized cases analyzed in this paper.

## 2 The system simulator

The averaging procedures involving concentrations or their logarithms are assessed by a methodical numerical Monte Carlo-type experiment based on a system simulator by which measurement signal and noise are propagated through an idealized retrieval system and finally averaged. The system simulator is idealized in a sense that (a) we assume locally linear radiative transfer, and (b) we restrict the problem to one dimension, i.e. to a scalar signal $y$ which depends on a scalar concentration $x$. In order to avoid any dispute by which probability density function (normal, lognormal, inversely normal etc) the true atmospheric state is best represented, we use a modeled distribution of atmospheric concentrations as reference ensemble $\boldsymbol{x}$ of concentrations $x_n$, $n = 1, \ldots N$, expressed as volume mixing ratios (vmrs).

For each concentration $x_n$, the related measurement signal $y_n$ is simulated as

$$y_n = y_0 + k x_n + \epsilon_n, \tag{1}$$

where $y_0$ is a constant background signal, $k$ is the sensitivity $\mathrm{d}y/\mathrm{d}x$ of the measurement system and $\epsilon_n$ is the measurement error associated with the $n$th measurement.



Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

AMTD

4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann

The latter is obtained from a pseudo-random number generator providing normally distributed random numbers of zero expectation and variance $s^2$. Without loss of generality, we set $y_0 = 0$ and $k = 1$. The measurement variance is then set to $s^2 = \bar{x}^2/r^2$, where $\bar{x}$ is the mean value of the concentrations $x_n$ and $r$ is a tunable average signal to noise ratio (SNR).

The simulated measurements $y_n$ are then propagated through an iterative retrieval simulator operating in the ln(vmr) domain. With

$$\frac{\mathrm{d}y_n}{\mathrm{d}\ln x_n} = k\frac{\mathrm{d}x_n}{\mathrm{d}\ln x_n} = x_n \tag{2}$$

we have for iterative unconstrained maximum likelihood retrievals

$$\ln x_{n,i+1} = \ln x_{n,i} + x_{n,i}^{-1}(y_n - x_{n,i}), \tag{3}$$

where $x_{i,n}$ is the concentration retrieved in the $i$th iteration from the $n$th simulated measurement. For positive signals $y_n$, Eq. (3) converges towards $x_{n,I} = x_n + \epsilon_n$, where $I$ denotes the final iteration of sample $n$ and $\epsilon_n$ is the measurement error propagated into the $x$-space, as in the case of linear maximum likelihood retrievals.

For maximum a posteriori (optimal estimation) retrievals (Rodgers, 2000), we have

$$\ln x_{n,i+1} = \ln x_{n,i} + \frac{\frac{x_{n,i}}{s^2}(y_n - x_{n,i}) - \sigma_a^{-2}(\ln x_{n,i} - \ln x_a)}{\sigma_a^{-2} + x_{n,i}^2 s^{-2}} \tag{4}$$

where $x_a$ is the prior information on $x$ with variance $\sigma_a^2$ in the logarithmic domain. Since biases caused by inappropriate prior information and variance are beyond the scope of this study, we make two further idealizing assumptions:

$$x_a = \exp\left[\frac{\sum_{n=1}^{N}\ln x_n}{N}\right] \tag{5}$$





AMTD
4, 7159–7183, 2011

and

$$\sigma_a^2 = \frac{\sum_{n=1}^{N} (\ln x_n - \ln x_a)^2}{N} \tag{6}$$

As a variation of this scheme, we have also performed simulations with a constant a priori variance $\sigma_a^2 = \ln 2$, corresponding to a 100 % variance in the linear space. This is motivated by the lack of reliable information on climatological variances in real applications. A 100 % variance is thus often assumed in optimal estimation schemes applied to remote sensing data. The rationale for choosing this value is to reduce the a priori content of the results while guaranteeing reasonable and stable retrievals. In the following, we refer to this ad hoc variant of optimal estimation as "modified" maximum a posteriori approach.

Convergence of the iterative retrieval scheme is reached when the absolute difference of $\ln x_{n,i+1}$ and $\ln x_{n,i}$ is smaller than $0.001\sigma_x$, i.e., a fraction of the estimated retrieval noise error in the logarithmic retrieval space, which can be expressed using $k = 1$ by

$$\sigma_x = \frac{s x_{n,i}}{x_{n,i}^2 + s^2 \sigma_a^{-2}}, \tag{7}$$

with $\sigma_a^{-2} = 0$ for maximum likelihood retrievals, $\sigma_a^{-2} = 1/(\ln 2)$ for modified maximum a posteriori retrievals, and $\sigma_a^{-2}$ inferred from the actual variability of the true state in the logarithmic domain for Bayesian maximum a posteriori retrievals.

Logarithmic retrievals do not allow zero-residual retrievals in the case when the linear retrieval would give a negative concentration, although the inverse problem is by no means algebraically overconstrained. As a consequence, convergence is not reached in unconstrained retrievals (i.e. maximum likelihood) if the signal is negative. Here, we reject unconverged maximum likelihood retrievals (i.e., if convergence is not reached after 20 iterations) from the retrieved ensemble before averaging the results, as done in many remote sensing applications. It should be noted that also other treatments





AMTD
4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann

of unconverged retrievals (i.e., consideration of these data after a maximum number of iterations) are occasionally applied in practice, and averaging results for maximum likelihood retrievals depend strongly on its choice (see discussion in Sect. 3.1). Non-convergence, however, should not occur in the case of maximum a posteriori retrieval since the constraint introduces a well defined minimum in the cost function even in the case of negative signals.

The assessment of averaging procedures is then based on the comparison of $N^{-1}\sum_{n=1}^{N} x_{n,l}$ to $N^{-1}\sum_{n=1}^{N} x_n$ and $N^{-1}\sum_{n=1}^{N} \ln x_{n,l}$ to $N^{-1}\sum_{n=1}^{N} \ln x_n$, respectively. Comparison of like with like certainly is idealistic, because in practice climatological data are often compared among each other without questioning how the climatologies have been generated. Nevertheless, we think that comparison of logarithmic averages of logarithmic retrievals with linear averages of the true state would not be fair.

It is evident that averages of maximum a posteriori retrievals depend on the a priori information. A more rigorous approach is thus to compare to averages of

$$\tilde{x}_n = \exp[A_n \ln(x_n) + (1 - A_n)\ln(x_a)], \tag{8}$$

with the averaging kernel

$$A_n = \frac{x_{n,l}^2}{x_{n,l}^2 + s^2\sigma_a^{-2}}. \tag{9}$$

$\tilde{x}_n$ is the retrieval response to the "true" state $x_n$ (i.e., the retrieval result for $\epsilon = 0$, i.e. no measurement noise considered). Note that $1 - A_n$ describes also the a priori contribution to the solution $x_{n,l}$ which, contrary to the case of linear retrievals, depends on the solution itself. This approach is commonly used in point-to-point comparisons of remotely sensed data with model results or independent measurements. However, in many averaging applications (i.e. comparisons of trace gas climatologies), averaging kernels of individual ensemble members are not available. For this reason, we compare to averages of both $\tilde{x}_n$ and $x_n$ in the case of maximum a posteriori retrievals.





AMTD
4, 7159–7183, 2011

While averaging of maximum a posteriori retrievals without re-adjustment of the content of a constant priori information is questionable because the optimal mean is not identical to the mean of optimal estimates (c.f. Chapt. 10.4.1 in Rodgers, 2000), and because prior knowledge of a single atmospheric state is less reliable than priori knowledge on the mean atmospheric state, we ignore the re-adjustment of the a priori content to produce an optimal average, because this is rarely done in practice and beyond the scope of this paper.

## 3 Case studies

The case studies are organized in a way that first unconstrained maximum likelihood retrievals and then standard and modified maximum a posteriori retrievals are discussed. For each of these retrieval schemes we perform simulations for different SNRs covering values from 0.5 to 10.

In order to provide realistic examples, the ensembles $x_n$ represent zonal distributions of CO and $H_2O$ taken from WACCM model simulations described in Jackman et al. (2008) for the period of November 2003 in a vertical range from 1000–0.001 hPa with global latitudinal coverage. Concentrations of CO and $H_2O$ are retrieved in the ln(vmr) space in many atmospheric remote sensing applications (e.g., von Clarmann et al., 2009; Funke et al., 2009; Schneider et al., 2006; Deeter et al., 2007). Each ensemble includes about 10 000 members. The global distribution of the corresponding averages (i.e., zonal means) is shown in Fig. 1 (top).

Highest standard deviations $\sigma_m$ of the modeled distributions are found where spatial gradients are strongest, i.e. in regions of transport barriers, vertical transport, etc. In the case of CO, this occurs in the polar regions in the mid-stratosphere and is related to vertical transport by the meridional circulation. $H_2O$ variability is highest in the UTLS (see Fig. 1, middle). These standard deviations represent the local natural variability of the atmospheric state, as opposed to any scatter being caused by measurement noise.

The magnitudes of differences of linear and logarithmically averaged zonal means correlate spatially with the standard deviation of the distributions for both CO and $H_2O$ (Fig. 1, bottom). This correlation is quite compact (see Fig. 2). The differences between linear and logarithmic averaging are somewhat more pronounced for $H_2O$ than for CO. For a local natural variability of 100 % in terms of standard deviation, the differences reach 40 % for $H_2O$ but only 30 % for CO.

### 3.1 Maximum likelihood retrievals

In this section we assume that direct inversion of the radiative transfer equation (Eq. 1) is used in the logarithmic domain, without application of any constraint beyond that implied by the use of the logarithm of the concentration. For positive signals, the logarithmic maximum likelihood retrieval yields the same result as the retrieval in the linear domain. However, the rejection of unconverged logarithmic retrievals with a negative signal from the averages leads to positive bias compared to averaged results retrieved in the linear domain.

As a consequence, linear averages of logarithmic retrievals are biased high compared to the corresponding averages of the "true" distribution. For a constant SNR, this bias is correlated with the dispersion of the "true" ensemble. Assuming a SNR of 2, the bias between the retrieved and the true zonal mean distributions of CO and $H_2O$ varies from 2 % to 25 %, being highest in regions with most pronounced atmospheric variability (see Fig. 3, upper panels).

In the case of logarithmic averaging, this behavior is partly compensated by the asymmetric mapping of the normal-distributed noise into the ln(vmr) parameter space, introducing a negative bias compared to the corresponding averages of the "true" distributions. This negative bias dominates for low atmospheric variability. Figure 3 (lower panels) shows this behavior for a SNR of 2. Here, the overall bias between retrieved and true zonal mean distributions of CO and $H_2O$ varies from −15 % to 30 % in dependence of the "true" distributions' dispersion.





AMTD
4, 7159–7183, 2011



B. Funke and T. von Clarmann

Interestingly, logarithmically averaged retrievals of some particular latitude/altitude boxes become virtually zero. This random-like behavior is introduced by single retrievals of signals being infinitesimally close to zero, leading to high negative values in the ln(vmr) space. The occurrence of such artifacts is ruled by the probability density of such small signals which, in turn, is linked to the variance of the measurement noise. For high SNR, this probability density is small, because the relevant interval of signals is located on the tail of the Gaussian distribution. It is also small for low SNR, because due to the broader probability distribution function most negative signal values are so negative that the retrieval does not reach convergence and related results are discarded, and only very few measurements hit the small interval where the measurements are negative to cause problems but their absolute values are small enough to still allow convergence. Most frequent occurrences of this peculiarity are found for intermediate SNRs of around 2.

Figure 4 shows the relative differences between retrieved and "true" averages as function of the relative standard deviation of "true" distribution $\sigma_m$ for a SNR of 2, summarizing the behavior discussed above. The correlation of differences and the local natural variability is quite compact. Therefore, in the following we restrict our analysis to its average dependence (indicated by solid lines in Fig. 4).

It is interesting to notice that, in the case of $H_2O$ distributions, the dependence on $\sigma_m$ is more pronounced for logarithmic than for linear averages, while this is not the case for CO. This is most likely related to differences in the shapes of the PDFs of both species in regions with high local natural variability.

Figure 5 summarizes the results for maximum likelihood retrievals for a variety of SNRs. In general, linear averaging is superior in the case of high SNR and large local natural variability, while logarithmic averaging is superior in the case of small SNR and small natural variability, although exceptions exist.

It should be noted that this evaluation is only valid in the case of rejection of unconverged retrievals. If the results of final unconverged iterations would have been included or those results would have been set to an arbitrary small number, the positive





AMTD
4, 7159–7183, 2011



B. Funke and T. von Clarmann

bias found in linear averages was reduced by the fraction of converged retrievals with positive signal, while a pronounced low bias of logarithmic averages was introduced due to the increased contribution of very small values of $x_{n,l}$. The latter depends then strongly on the choice of the maximum number of iterations or on the choice of the fake value, respectively. In our case (maximum number of iterations of 20), the inclusion of these retrieval results would introduce a low bias of logarithmic averages of up to 70 % for low SNR, completely disabling its meaningful interpretation.

### 3.2 Maximum a posteriori retrievals

The same kind of analysis also has been performed for maximum a posteriori retrievals. Comparisons of the linear and logarithmic averages with the "true" averages as a function of natural variability are shown in Fig. 6 for various signal to noise ratios. In the case of low local natural variability – and thus also low a priori variance – the differences are small because the content of a priori information in the retrieval is large. As already mentioned in Sect. 2, averaging of retrievals containing a constant prior information does not produce an optimal average, since the prior information is systematically overrepresented in the mean. On the other hand, the prior information characterizes the mean state of the atmosphere better than an actual state. The problem of the need of re-adjustment of the weight of the priori information in the mean, however, is beyond the scope of this study.

For intermediate local natural variability and a priori variance, linear averages of logarithmically retrieved mixing ratios are biased low. This is, because the a priori state of the atmosphere is calculated by logarithmic averaging of the true distribution. This low bias is not more than the bias between the logarithmic and linear averages of the "true" distribution, which is, via the content of priori information in the retrieval, propagated to the averages of the retrievals. Results for less ideal a priori information may be different. In the case of large local natural variability along with large a priori variance, the bias of the linear average turns high, similar to the case of maximum likelihood retrievals but with a considerably smaller amplitude (less than 5 % even for





AMTD
4, 7159–7183, 2011

low SNRs). The positive bias of maximum a posteriori retrieval averages, however, is not related to the rejection of unconverged retrievals of negative signals (convergence is in our simplified one-dimensional retrieval always achieved due to the constraint), but to the dependence of the a priori contribution to the retrieval solution on the solution itself (see Sect. 2). As a consequence, retrievals of low signals (either due to low values of $x_n$ or negative values of $\epsilon_n$) have a higher a priori contribution than those of high signals, resulting in a high bias. The turning point, i.e., where the high bias due to asymmetric a priori mapping starts to overcompensate the low bias introduced by $x_a$ itself, depends strongly on the SNR. For SNRs greater than 5, differences start to increase already at standard deviations below 20 %, while for SNRs lower than 1 the turning point is located at standard deviations greater than 50 %.

Logarithmic averages of logarithmic maximum a posteriori retrievals are generally higher than the logarithmic averages of the "true" distribution for intermediate to high values of $\sigma_\mathrm{m}$ (i.e., when there is a substantial contribution of the measurement to the retrieval solution). Contrarily to the linear averaging case, no negative bias due to the a priori contribution is introduced since, in or idealized case, the prior information is identical the the "true" logarithmic average. Logarithmic averaging performs apparently worse compared to linear averaging for high values of $\sigma_\mathrm{m}$. This, however, is related to the compensation effect of the a priori in the linear averaging (see discussion above) and hence depends strongly on the choice of the a priori.

In addition, we have also compared the linear and logarithmic averages to the averaged linear retrieval response to the "true" distribution $x_n$, the latter obtained by applying the averaging kernels $A_n$ (Rodgers, 2000) to $x_n$ according to Eq. (8) (see Fig. 7). In first order, these comparisons show the isolated effect of the asymmetric mapping of noise in the constrained retrieval, that is, the influence of the a priori information on the difference between the retrieved and the "true" mean, is removed. Now, the bias between linear averages of the retrieval results and the linear retrieval response to the "true" distribution is generally positive. This high bias is increased compared to Fig. 6 since the compensation effect of the a priori contribution is removed. For logarithmic





AMTD
4, 7159–7183, 2011

averages, these differences are smaller. The generally better performance of logarithmic averages is related to the compensation of the positive bias due to the dependence of the a priori contribution on $x_{n,l}$ by the negative bias caused by asymmetric mapping of normal-distributed noise in the ln(vmr) space. The latter dominates for high SNRs, giving raise to a negative overall bias of logarithmic averages for SNRs greater than 5 in Fig. 7.

In summary, except for high SNRs (> 5), logarithmic averaging of logarithmic maximum a posteriori retrievals is recommended in validation exercises or point-to-point model-data comparisons whenever averaging kernels related to individual measurements are applied to the corresponding data to be compared.

### 3.3 Modified a posteriori retrievals

Since in practical applications of the optimal estimation retrieval scheme often ad hoc choices of a priori variances are made, we also have studied averaging of logarithmic retrievals where the a priori standard deviation was set to 100 % (see Fig. 8). For high value of $\sigma_m$, the behavior is similar to the classical maximum a posteriori retrievals (Fig. 7). For lower values of $\sigma_m$, however, the measurement contribution to the retrieval solution is much higher than in the classical maximum a posteriori case and differences of retrieved and "true" averages are increased.

The comparison of the linear and logarithmic averages to the averaged linear retrieval response to the "true" distribution $x_n$ (see Fig. 9) shows that the use of a constant a priori variance removes the dependence of the biases on $\sigma_m$ and they remain approximately constant. Again, better performance is achieved with logarithmic averaging for SNR < 5 due to the compensation of the positive bias due to the dependence of the a priori contribution on $x_{n,l}$ by the negative bias caused by asymmetric mapping of normal-distributed noise in the ln(vmr) space. For higher SNRs, however, linear averaging yields smaller biases.



AMTD
4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann

## 4 Conclusions

Ideally, the average of concentration shall be mass-conservative in a sense that the average concentration times the airmass equals the total amount of the target gas in the airmass. This can only be achieved with linear averaging. Both linear and logarithmic averaging of logarithmic retrievals can lead to biases of several ten percent, which are typically larger for larger local atmospheric variability. Biases caused by the impossibility of logarithmic retrievals of mapping negative measured signals into the atmospheric state space can be remedied by neither of the averaging schemes.

Usually, for maximum likelihood retrievals linear averaging better represents the true mean value in cases of large local natural variability and high signal to noise ratios, while for small local natural variability logarithmic averaging often is superior. For maximum a posteriori retrievals, the dependence of the weight of the priori information on the state value itself causes some hardly predictable interaction between the effect of the constraint on the retrieval and the characteristics of the averaging procedure. Since in logarithmic retrievals the priori information is ideally chosen as the expectation value of the logarithm of the atmospheric state variable, logarithmic averaging of results better reproduces the logarithmic average of the true atmospheric state in cases when the retrieval is dominated by the prior information. For higher atmospheric variability, which in a truly Bayesian maximum a posteriori retrieval goes along with a lesser weight of prior information, the bias of the average is composed of the superposition of the effects of multiple biasing processes of positive and negative sign. The assumption that the prior information is identical with the true local mean of the atmospheric state is certainly an ideal case and more realistic cases where the a priori information differs from the true mean state of the atmosphere may lead to even different results but could not be assessed here.

Further, these investigations refer to an ideal world where logarithmic averages of logarithmic retrievals are compared to logarithmic averages of the true atmospheric state values. In the real world, however, the user of climatologies may not ask about





AMTD
4, 7159–7183, 2011



B. Funke and T. von Clarmann

the procedure how a climatology has been generated and may unintentionally compare a climatology based on logarithmic averages of logarithmic retrievals from one instrument with linear averages of linear retrievals of another instrument, which adds an additional bias of up to 40 % (see Fig. 2).

In summary, averaging of logarithmically retrieved abundances of atmospheric species contains a lot of traps which cannot be avoided by application of a simple recipe. Particularly, biasing can never be systematically avoided by using a superior averaging scheme. At best, limitation of damage can be aimed at. While logarithmic averaging in some cases indeed performs better than linear averaging, particularly in some cases of Bayesian or modified maximum a priori retrievals, related biases are by no means fully compensated.

Although our simulations have been carried out for an one-dimensional retrieval problem under the idealized assumption of locally linear radiative transport, the conclusions of this study can be generalized in a qualitative manner to more realistic retrieval problems. For example, multi-dimensional profile retrievals, typically performed in remote sensing applications, would suffer the same problems as described here with the added complexity of correlations between different profile points. These correlations are typically introduced by instrumental and/or geometrical limitations in vertically resolving the profiles, i.e., the line of sight of a remote sounder travels through multiple atmospheric layers. Thus a single measurement error cannot be assigned to a single profile point but to positively or negatively correlated errors at various altitudes. Inclusion of constraints (e.g. maximum a posteriori retrievals) further contributes to these correlations.

The inclusion of non-linear radiative transport would not alter the presented results for unconstrained maximum likelihood retrievals, however, results for maximum a posteriori retrieval averages might differ due to an amplification (or reduction) of the positive bias related to the dependence of the a priori contribution to the retrieval solution on the solution itself. The latter effect, as described in Sect. 3.2, is already caused by the "artificial" non-linearity introduced by the retrieval of ln(vmr). Additional non-linearity



AMTD
4, 7159–7183, 2011

related to radiative transfer leads only to its modification. In consequence, we also expect biases related to the dependence of the a priori contribution to the retrieval solution on the solution itself in the case of averaging linear maximum a posteriori retrievals whenever non-linear radiative transfer occurs or, if the observed signal depends on additional quantities (e.g. temperature in the case of emission measurements) being correlated to the retrieval quantity.

*Acknowledgements.* The authors like to thank the Toronto SPARC office and the International Space Science Institute (ISSI) in Berne for hosting team meetings where the issue of this paper became evident and for providing generous travel support. The authors would also like to thank the team members of the SPARC Data Initiative led by Michaela Hegglin and Susann Tegtmeier for triggering our interest in this interesting problem, and Charles Jackman for kindly providing WACCM model results. BF was supported by by the Spanish MICINN under project AYA2008-03498/ESP and project 200950I081 of CSIC.

## References

Bowman, K. W., Rodgers, C. D., Kulawik, S. S., Worden, J., Sarkissian, E., Osterman, G., Steck, T., Lou, M., Eldering, A., Shephard, M., Worden, H., Lampel, M., Clough, S., Brown, P., Rinsland, C., Gunson, M., and Beer, R.: Tropospheric emission spectrometer: retrieval method and error analysis, IEEE T. Geosci. Remote, 44, 1297–1307, 2006. 7161

von Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U., Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G. P., and Versick, S.: Retrieval of temperature, $H_2O$, $O_3$, $HNO_3$, $CH_4$, $N_2O$, $ClONO_2$ and ClO from MIPAS reduced resolution nominal mode limb emission measurements, Atmos. Meas. Tech., 2, 159–175, doi:10.5194/amt-2-159-2009, 2009. 7160, 7165

Deeter, M. N., Edwards, D. P., and Gille, J. C.: Retrievals of carbon monoxide profiles from MOPITT observations using lognormal a priori statistics, J. Geophys. Res., 112, D11311, doi:10.1029/2006JD007999, 2007. 7165

Funke, B., López-Puertas, M., García-Comas, M., Stiller, G. P., von Clarmann, T., Höpfner, M., Glatthor, N., Grabowski, U., Kellmann, S., and Linden, A.: Carbon monoxide distributions from the upper troposphere to the mesosphere inferred from 4.7 µm non-local thermal equi-





AMTD
4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann

librium emissions measured by MIPAS on Envisat, Atmos. Chem. Phys., 9, 2387–2411, doi:10.5194/acp-9-2387-2009, 2009. 7161, 7165

Jackman, C. H., Marsh, D. R., Vitt, F. M., Garcia, R. R., Fleming, E. L., Labow, G. J., Randall, C. E., López-Puertas, M., Funke, B., von Clarmann, T., and Stiller, G. P.: Short- and medium-term atmospheric constituent effects of very large solar proton events, Atmos. Chem. Phys., 8, 765–785, doi:10.5194/acp-8-765-2008, 2008. 7165

Papandrea, E., Dudhia, A., Grainger, R. G., Vancassel, X., and Chipperfield, M. P.: Retrieval of global hydrogen peroxide ($H_2O_2$) profiles using ENVISAT–MIPAS, Geophys. Res. Lett., 32, L14809, doi:10.1029/2005GL022870, 2005. 7161

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, vol. 2 of Series on Atmospheric, Oceanic and Planetary Physics, edited by: Taylor, F. W., World Scientific, Singapore, 2000. 7162, 7165, 7169

Schneider, M., Hase, F., and Blumenstock, T.: Ground-based remote sensing of HDO/$H_2O$ ratio profiles: introduction and validation of an innovative retrieval approach, Atmos. Chem. Phys., 6, 4705–4722, doi:10.5194/acp-6-4705-2006, 2006. 7161, 7165

Urban, J., Lautié, N., Flochmoën, E. L., Jiménez, C., Eriksson, P., de La Noë, J., Dupuy, E., Ekström, M., El Amraoui, L., Frisk, U., Murtagh, D., Olberg, M., and Ricaud, P.: Odin/SMR limb observations of stratospheric trace gases: level 2 processing of ClO, $N_2O$, $HNO_3$, and $O_3$, J. Geophys. Res., 110, D14307, doi:10.1029/2004JD005741, 2005. 7161

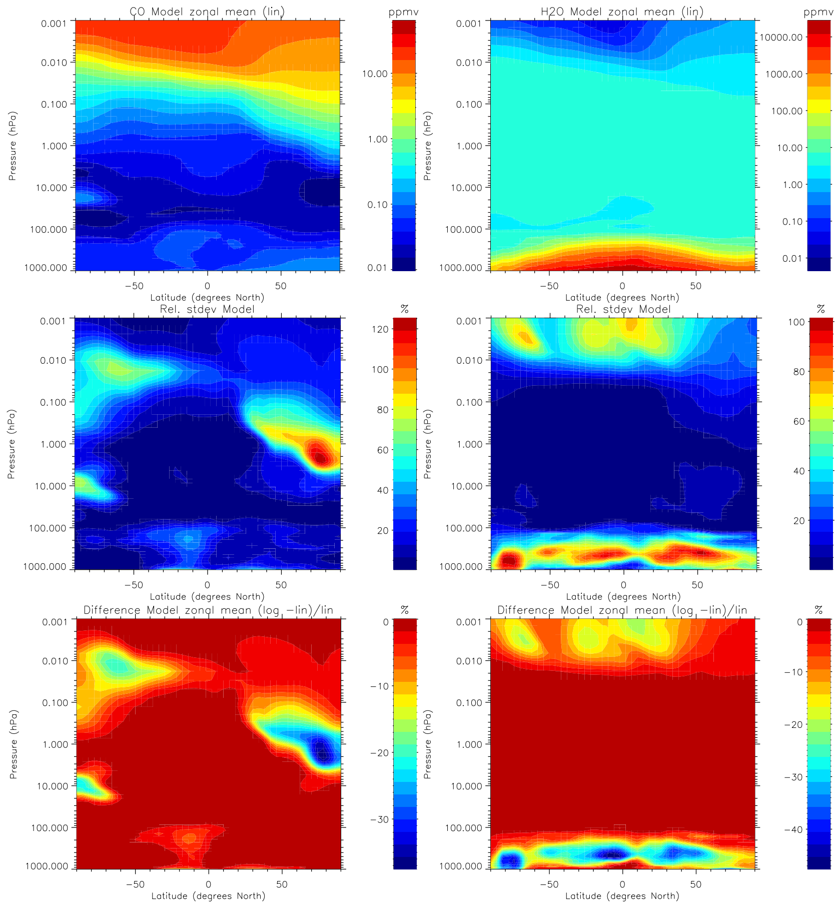


**Fig. 1.** Model zonal mean distributions of zonal mean vmrs (top), standard deviations (middle), and differences between linear and logarithmically averaged means (bottom) for CO (left) and $H_2O$ (right).





AMTD

4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann

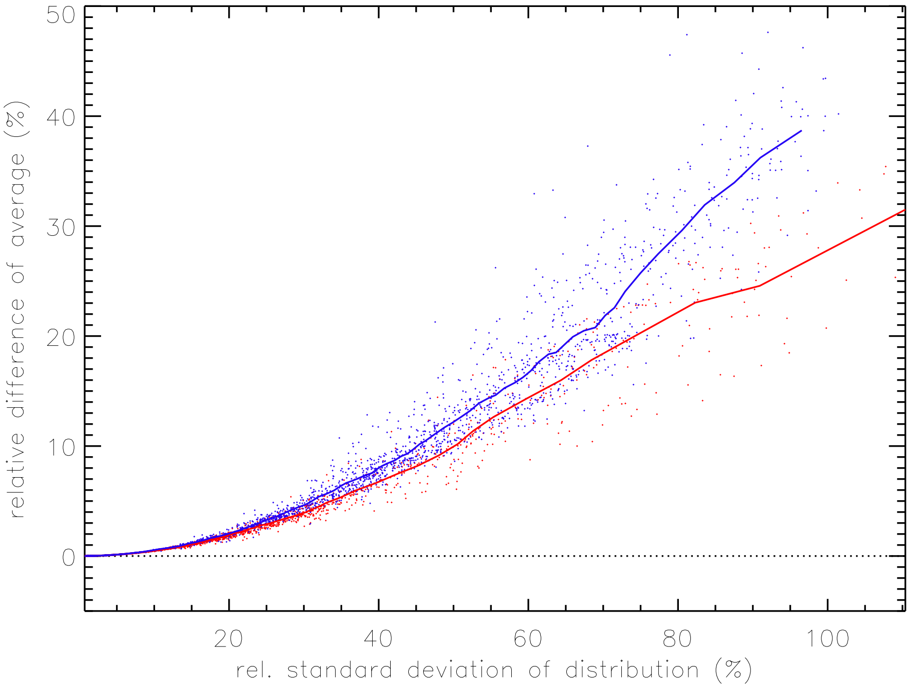


**Fig. 2.** Differences of linearly and logarithmically averaged zonal means (relative to linearly averaged zonal means) versus standard deviation of the distributions for CO (red) and $H_2O$ (blue). Data points represent single latitude/altitude boxes.





AMTD
4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and
T. von Clarmann

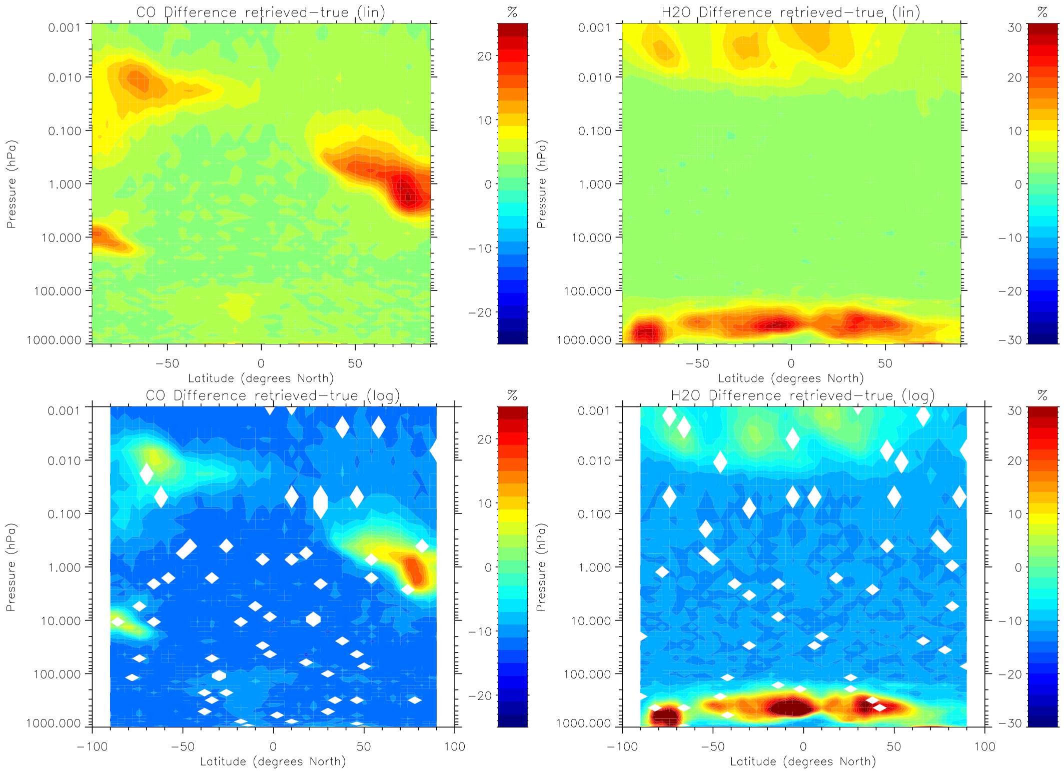


**Fig. 3.** Zonal mean differences of retrieved and "true" distributions (relative to the latter) for CO (left) and $H_2O$ (right). Top: linear averages, bottom: logarithmic averages. Results are shown for maximum likelihood ln(vmr)-retrieval simulations with a signal to noise ratio of 2. White regions in the lower panels reflect unreasonable results ("zero" averages, see text for further explanations).





AMTD
4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann

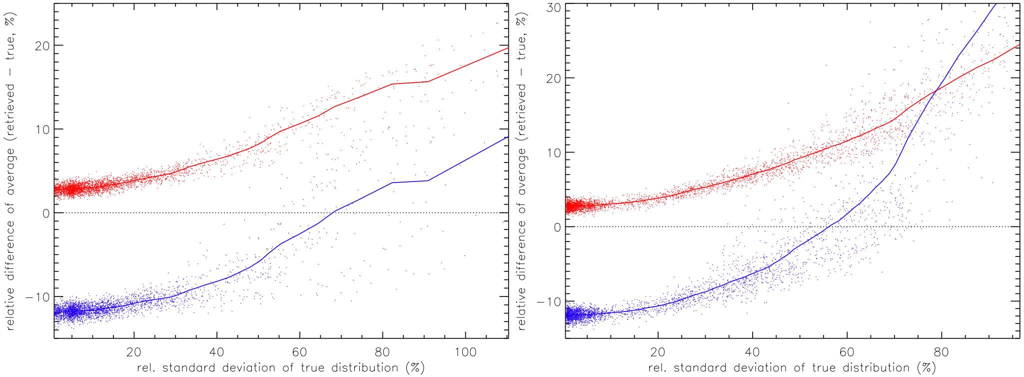


**Fig. 4.** Differences of retrieved and "true" zonal means (relative to the latter) versus local natural variability in terms of standard deviation for CO (left) and $H_2O$ (right). Red: linear averages, blue: logarithmic averages. Data points reflect individual latitude/altitude boxes. Results are shown for maximum likelihood ln(vmr)-retrieval simulations with a signal to noise ratio of 2.





AMTD
4, 7159–7183, 2011

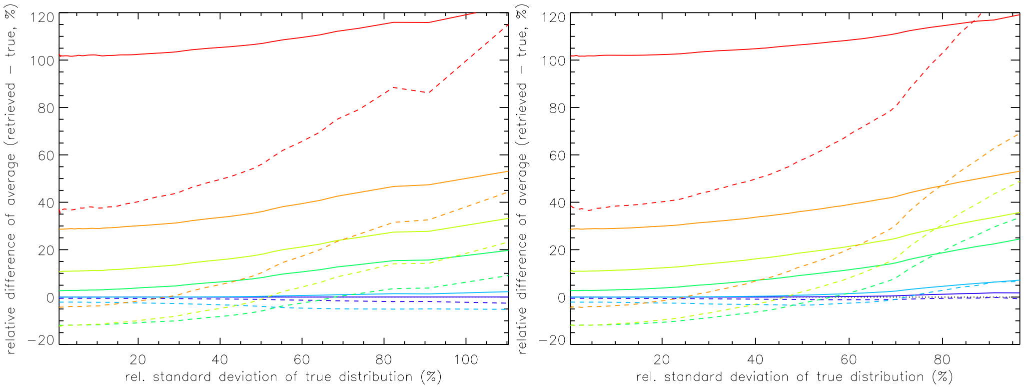


**Fig. 5.** Differences of retrieved and "true" zonal means (relative to the latter) versus local natural variaility in terms of standard deviation for the CO (left) and $H_2O$ (right). Solid: linear averages, dashed: logarithmic averages. Results are shown for maximum likelihood ln(vmr)-retrieval simulations with signal to noise ratios of 10 (dark blue), 5 (light blue), 2 (dark green), 1.43 (light green), 1.0 (orange), and 0.5 (red).





AMTD
4, 7159–7183, 2011



B. Funke and T. von Clarmann

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

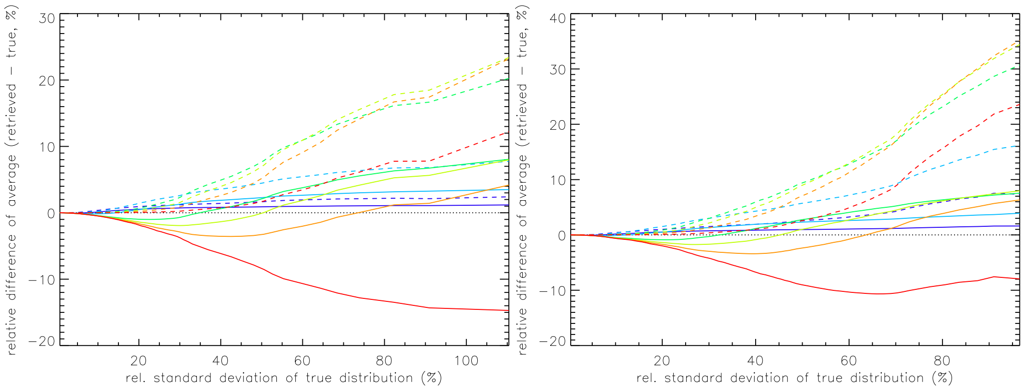


**Fig. 6.** Differences of retrieved and "true" zonal means (relative to the latter) versus local natural variability in terms of standard deviation for the CO (left) and $H_2O$ (right). Solid: linear averages, dashed: logarithmic averages. Results are shown for maximum a posteriori ln(vmr)-retrieval simulations with SNRs of 10 (dark blue), 5 (light blue), 2 (dark green), 1.43 (light green), 1.0 (orange), and 0.5 (red). Note that the a priori contribution to the retrievals is greater than 50 % for standard deviations below 10, 20, 50, 70, 100, and 200 %, respectively.





AMTD
4, 7159–7183, 2011



B. Funke and
T. von Clarmann

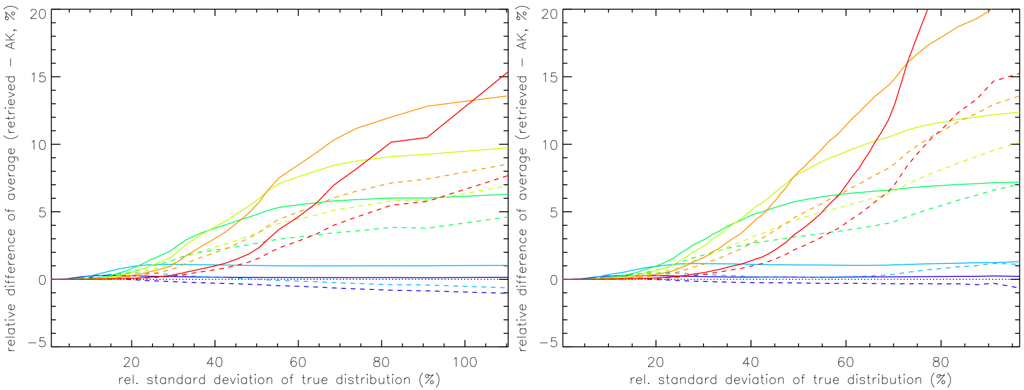


**Fig. 7.** Differences between the mean values and the mean "true" atmospheric state after application of the averaging kernel to the latter for CO (left) and $H_2O$ (right). Solid: linear averages, dashed: logarithmic averages. Results are shown for maximum a posteriori ln(vmr)-retrieval simulations with SNRs of 10 (dark blue), 5 (light blue), 2 (dark green), 1.43 (light green), 1.0 (orange), and 0.5 (red).





AMTD
4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and
T. von Clarmann

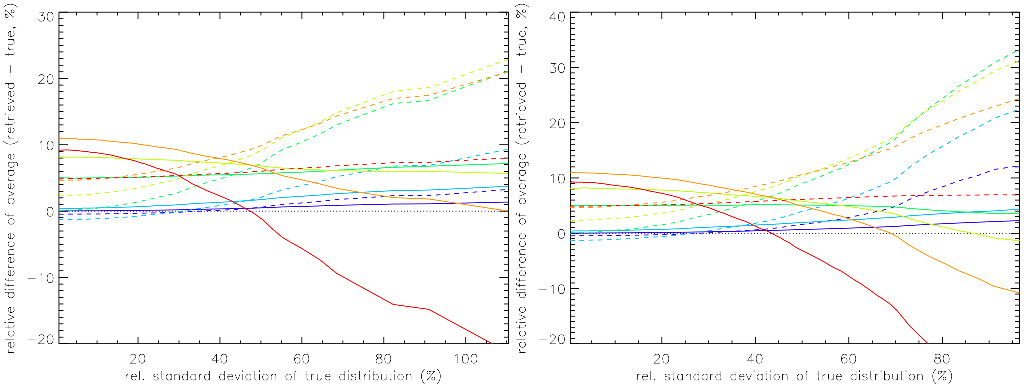


**Fig. 8.** Differences of retrieved and "true" zonal means (relative to the latter) versus local natural variability in terms of standard deviation for the CO (left) and $H_2O$ (right). Solid: linear averages, dashed: logarithmic averages. Simulations are performed with a modified maximum a posteriori ln(vmr)-retrieval simulations using a fixed constraint corresponding to climatological variance of 100 %. Results for signal to noise ratios of 10 (dark blue), 5 (light blue), 2 (dark green), 1.43 (light green), 1.0 (orange), and 0.5 (red) are shown. The resulting a priori contributions are 1, 3.8, 20, 33, 50, and 80 %, respectively.

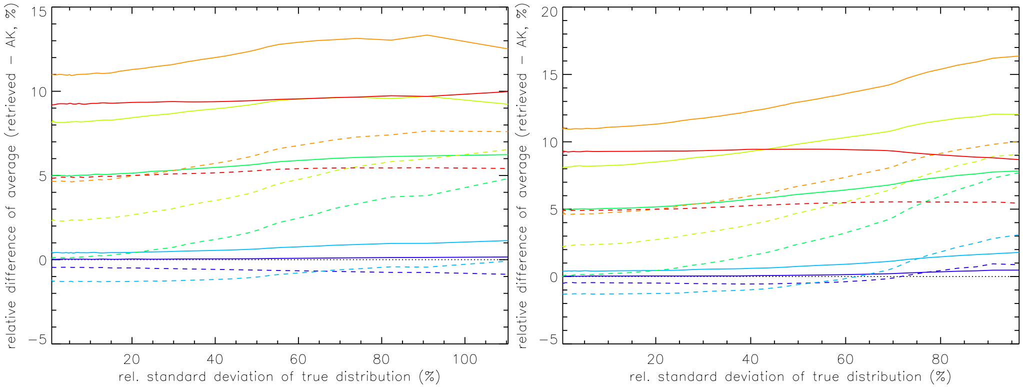


**Fig. 9.** Differences of retrieved and linear retrieval response (i.e. averaging kernels applied to the "true" profiles) zonal means (relative to the latter) versus standard deviation of distributions for the CO (left) and $H_2O$ (right). Solid: linear averages, dashed: logarithmic averages. Simulations are performed with a modified maximum a posteriori ln(vmr)-retrieval simulations using a fixed constraint corresponding to climatological variance of 100 %. Results for signal to noise ratios of 10 (dark blue), 5 (light blue), 2 (dark green), 1.43 (light green), 1.0 (orange), and 0.5 (red) are shown.





AMTD
4, 7159–7183, 2011

How to average logarithmic retrievals

B. Funke and T. von Clarmann