



Merged ozone profiles from four MIPAS processors

Alexandra Laeng¹, Thomas von Clarmann¹, Gabriele Stiller¹, Bianca Maria Dinelli², Anu Dudhia³, Piera Raspollini⁴, Norbert Glatthor¹, Udo Grabowski¹, Viktoria Sofieva⁵, Lucien Froidevaux⁶, Kaley A. Walker⁷, and Claus Zehner⁸

¹Institut für Meteorologie und Klimaforschung, Karlsruhe Institute of Technology, Karlsruhe, Germany

²ISAC-CNR, Bologna, Italy

³Earth Observation Data Group, Oxford University, Oxford, UK

⁴IFAC-CNR, Florence, Italy

⁵Finnish Meteorological Institute, Helsinki, Finland

⁶Jet Propulsion Laboratory, California Institute of Technology, Pasadena, USA

⁷University of Toronto, Toronto, Canada

⁸ESA ESRIN, Frascati, Italy

Correspondence to: Alexandra Laeng (alexandra.laeng@kit.edu)

Received: 16 July 2016 – Discussion started: 17 October 2016

Revised: 14 March 2017 – Accepted: 31 March 2017 – Published: 21 April 2017

Abstract. The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) was an infrared (IR) limb emission spectrometer on the Envisat platform. Currently, there are four MIPAS ozone data products, including the operational Level-2 ozone product processed at ESA, with the scientific prototype processor being operated at IFAC Florence, and three independent research products developed by the Istituto di Fisica Applicata Nello Carrara (ISAC-CNR)/University of Bologna, Oxford University, and the Karlsruhe Institute of Technology–Institute of Meteorology and Climate Research/Instituto de Astrofísica de Andalucía (KIT–IMK/IAA). Here we present a dataset of ozone vertical profiles obtained by merging ozone retrievals from four independent Level-2 MIPAS processors. We also discuss the advantages and the shortcomings of this merged product. As the four processors retrieve ozone in different parts of the spectra (microwindows), the source measurements can be considered as nearly independent with respect to measurement noise. Hence, the information content of the merged product is greater and the precision is better than those of any parent (source) dataset.

The merging is performed on a profile per profile basis. Parent ozone profiles are weighted based on the corresponding error covariance matrices; the error correlations between different profile levels are taken into account. The intercorrelations between the processors' errors are evaluated statistically and are used in the merging. The height range of the

merged product is 20–55 km, and error covariance matrices are provided as diagnostics. Validation of the merged dataset is performed by comparison with ozone profiles from ACE-FTS (Atmospheric Chemistry Experiment–Fourier Transform Spectrometer) and MLS (Microwave Limb Sounder). Even though the merging is not supposed to remove the biases of the parent datasets, around the ozone volume mixing ratio peak the merged product is found to have a smaller (up to 0.1 ppmv) bias with respect to ACE-FTS than any of the parent datasets. The bias with respect to MLS is of the order of 0.15 ppmv at 20–30 km height and up to 0.45 ppmv at larger altitudes. The agreement between the merged data MIPAS dataset with ACE-FTS is better than that with MLS. This is, however, the case for all parent processors as well.

1 Introduction

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) was an infrared (IR) limb emission spectrometer onboard the ENVISAT platform. It measured during day and night at 6 to 70 km (up to 170 km in special modes), pole-to-pole, producing more than 1000 profiles per day. Around 30 species, temperature, and cloud composition could be derived from these measurements. In 2002–2004, the instrument operated in full spectral resolution, with a vertical resolution of about 3.5–6 km for the retrieved ozone

product; this period of MIPAS operations is referred to as the full resolution (FR) period. Due to a failure of the instrument's mirror slide in 2004, the operations were suspended during almost a year and were resumed in 2005 with reduced spectral, but improved vertical, resolution. The corresponding period, until the loss of communications with the ENVISAT platform in April 2012, is referred to as the reduced resolution (RR) period of MIPAS operations.

MIPAS Level-2 data are operationally processed at ESA, with the scientific prototype processor at IFAC Florence (Raspollini et al., 2013). Beyond this, there are three independent scientific Level-2 processors: at the Istituto di Fisica Applicata Nello Carrara (ISAC-CNR)/University of Bologna (Carlotti et al., 2006; Dinelli et al., 2010), at Oxford University (<http://www.atm.ox.ac.uk/MORSE/>), and the Institute of Meteorology and Climate Research/Instituto de Astrofísica de Andalucía (IMK/IAA) processor at KIT, Karlsruhe Institute of Technology (von Clarmann et al., 2003, 2009). Henceforth, the four processors will be referred to as the ESA, Bologna, Oxford, and KIT processors. The existence of these four products often leads to confusion in the scientific community about their differences and which one to use. The unified description of the four processors is given in Laeng et al. (2015). The main similarities and differences between the four processors can be summarized as follows:

- all four processors use the same Level-1b spectra provided by ESA, but the Level-2 retrieval algorithms are different;
- all four processors use microwindows instead of the full spectrum, but microwindow selection differs; for the rationale behind this approach see von Clarmann and Echle (1998), Echle et al. (2000), and Dudhia et al. (2002);
- all four processors apply a global fit approach in a sense that the tangent altitudes of a limb scan are processed simultaneously rather than sequentially (Carlotti, 1988); however, they use different regularization approaches;
- the Bologna processor uses a full 2-D-approach, that is, all measurements in a complete orbit are processed simultaneously; the horizontal variation of the atmospheric state within the orbit plane is considered; the KIT processor accounts for horizontal temperature gradients along the line of sight direction in the ozone retrieval; the other processors consider atmospheric variation in the altitude domain only.

In the frame of ESA's Ozone Climate Change Initiative project, a round robin evaluation of ozone products from the four MIPAS processors was performed. Comparison with ground-based instruments revealed that all four processors reproduce on average the correct ozone distribution in a similar way, with small differences in bias appearing most clearly in the troposphere. The KIT bias was shown to be

less than ± 25 ppbv, depending on latitude, compared to a systematic positive bias of roughly $+50$ ppbv for the other products. In the stratosphere, all four products are positively biased by 2–5 % relative to ozonesonde, lidar, Aura-MLS (Aura-Microwave Limb Sounder), and ACE-FTS (Atmospheric Chemistry Experiment-Fourier Transform Spectrometer) observations. Comparison with satellite instruments showed that the KIT product is biased high only above 35 km, while the other products start to be biased high at somewhat lower altitudes. This was tentatively attributed to the use of the MIPAS AB band ($1020\text{--}1170\text{ cm}^{-1}$) which is more restrictively used by the KIT processor below 35 km. The MIPAS datasets are generally in better mutual agreement than each of them with other ozone records. The good resemblance of the biases, for example, is a strong indication that the biases are not caused by one particular retrieval processor but come from the instrument itself, from the Level-1 processing or from the spectroscopic data used in the Level-2 retrievals. The details of this analysis are reported in Laeng et al. (2015). Given the availability of four independent MIPAS ozone data products, the question arose of how to optimize the use of all MIPAS data products. This gave rise to the title activity of this paper: to demonstrate the feasibility of merging the ozone data from the four MIPAS datasets. Contrary to the merging of data from multiple sensors, the following issues do not apply to the merging of multiple data products of a single sensor: sampling issues in space and time, and insufficient time overlap. Further, as follows from Kleinert et al. (2007), the degradation of the MIPAS instrument in the parts of the spectra used by different MIPAS processors is similar. Since different processors analyze different parts of the spectrum (microwindows), the source (parent) measurements can be considered as nearly independent with respect to the primary measurement errors. The microwindows used by the four MIPAS processors are reported in the Table 1.

The use of different data points by the four processors gives rise to the expectation for the merged product to have a better precision than the individual contributing datasets. However, it should be treated with caution: this expectation relies on the assumption that the dominating source of uncertainties coming with the data is measurement noise, or any other source of random error which is uncorrelated between the parent datasets. It is, of course, not expected that systematic error components will necessarily average out by the merging operation. The merging was performed on the MIPAS data provided for the round robin exercise only, namely ozone data for the years 2007 and 2008.

2 Merging approach

The merging is performed on a profile per profile basis. Parent ozone profiles are weighted based on their corresponding random error covariance matrices; the correlations be-

Table 1. Spectral analysis windows used by the four MIPAS processors for ozone retrieval.

	Microwindow (cm^{-1})	Altitude range (km)
ESA	729.25–732.25	15–42
	756.625–759.625	9–36
	1043.875–1046.875	27–68
	1117.000–1120.000	6–42
	1123.5625–1126.5625	9–68
Oxford	729.2500–732.25	15.0–46.0
	756.6250–759.625	7.5–37
	1043.8750–1046.875	27–70
	1117.0000–1120.000	6–46
	1123.5625–1126.5625	7.5–70
Bologna	686.688–689.688	60–68
	689.750–692.750	30–52
	731.188–734.188	6–36
	790.625–793.625	27–47
	1036.313–1039.313	33–68
	1071.875–1074.875	6–39
	1651.000–1654.000	6–36
	1682.688–1685.688	27–68
KIT	760.6875–764.3125	6–68
	766.8750–767.1875	6–68
	776.1875–777.9375	6–68
	781.0000–782.8750	6–68
	787.0000–788.0000	6–68
	1029.0000–1031.0000	36–68
1038.0000–1039.0000	36–68	

tween different altitude levels of the profiles are taken into account. The intercorrelations between the processors' errors are evaluated statistically and are used in the merging. The small sample size (four processors only) is an obstacle to the identification of outliers. It only takes one processor to significantly deviate from the true profile, and the merged product will be worse than any of the other three. Our choice is, however, to always use all four processors' values.

The merged profile is constructed as a weighted mean of the four parent profiles. For each processor, the errors at different height levels are correlated because of error propagation patterns typical for limb sounding. Therefore, the value of the merged profile at each level is a linear combination of all the levels of all four processors, with weights defined by corresponding error covariance matrices. This means that the weights depend on the uncertainties' size: the smaller the error of a processor are, the larger its contribution to the merged profile is. The merging is performed on a fixed pressure grid which corresponds approximately to the MIPAS RR nominal tangent altitude grid. At the upper and lower ends of the profiles, it occurs frequently that not all four processors provide data. The height range was hence limited to 62–0.8 hPa (~ 20 –55 km). We note n as the number of vertical levels in the profile. The merged profile is obtained as

$$\mathbf{x}_{\text{merged}} = \left(\begin{pmatrix} \mathbf{e} & \mathbf{e} & \mathbf{e} & \mathbf{e} \end{pmatrix} \mathbf{C}^{-1} \begin{pmatrix} \mathbf{e} \\ \mathbf{e} \\ \mathbf{e} \\ \mathbf{e} \end{pmatrix} \right)^{-1} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \end{pmatrix}, \quad (1)$$

where \mathbf{e} is $n \times n$ identity matrix, \mathbf{x}_i , $i = 1, 2, 3, 4$ is a column vector representing the profile from the processor i , and \mathbf{C} is the processor intercorrelation $4n \times 4n$ matrix defined as follows:

$$\mathbf{C} = \begin{pmatrix} \mathbf{S}_{x_1} & \mathbf{S}_{12}^T & \mathbf{S}_{13}^T & \mathbf{S}_{14}^T \\ \mathbf{S}_{12} & \mathbf{S}_{x_2} & \mathbf{S}_{23}^T & \mathbf{S}_{24}^T \\ \mathbf{S}_{13} & \mathbf{S}_{23} & \mathbf{S}_{x_3} & \mathbf{S}_{34}^T \\ \mathbf{S}_{14} & \mathbf{S}_{24} & \mathbf{S}_{34} & \mathbf{S}_{x_4} \end{pmatrix}, \quad (2)$$

where \mathbf{S}_{x_i} is the random retrieval error covariance matrix of processor i , \mathbf{S}_{ij} (i and j correspond to processors) are $n \times n$ matrices defined by

$$\mathbf{S}_{ij} = \mathbf{r}_{ij} \sqrt{\text{diag}(\mathbf{S}_{x_i})} \sqrt{\text{diag}(\mathbf{S}_{x_j})}, \quad (3)$$

with \mathbf{r}_{ij} being $n \times n$ matrices representing the correlation of errors on different levels of two processors; the calculation of \mathbf{r}_{ij} is discussed below in Sect. 3.

The random error covariance matrix of the merged profile is given by

$$\mathbf{S}_{\mathbf{x}_{\text{merged}}} = \left(\begin{pmatrix} \mathbf{e} & \mathbf{e} & \mathbf{e} & \mathbf{e} \end{pmatrix} \mathbf{C}^{-1} \begin{pmatrix} \mathbf{e} \\ \mathbf{e} \\ \mathbf{e} \\ \mathbf{e} \end{pmatrix} \right)^{-1}. \quad (4)$$

As the vertical resolutions of the four processors are very close (see Laeng et al., 2015 for details), our choice is not to take the averaging kernels into the merging formalism. See Ceccherini et al. (2015) for a merging formalism involving the averaging kernels but omitting the correlation between the random errors of the parent datasets.

3 Correlation coefficients

A retrieved atmospheric profile $\hat{\mathbf{x}}$ can be decomposed into the contribution from the true profile and the contribution from bias and random error:

$$\hat{\mathbf{x}} = \mathbf{x} + \epsilon_{\text{bias}} + \epsilon_{\text{random}}, \quad (5)$$

where $\hat{\mathbf{x}}$ is the retrieved profile, \mathbf{x} is the true profile, and ϵ_{bias} and ϵ_{random} are respectively the bias and random components

of the error. We pragmatically define “biases” as the average over time at a given altitude and latitude and “random components” as all other error components which survive the subtraction of the bias. Partly correlated errors contribute both to the bias and the random error component. We expect that the random errors in tendency cancel out by the merging, while the bias component of the error will survive the merging. Our approach does not require that the bias components are isolated beforehand, because we statistically evaluate the inter-processor error correlations using debiased differences, as outlined below.

We evaluate the intercorrelation of random errors of different processors by examining the statistics of differences between each pair of processors in the following way.

The random errors of processor i at height p and of processor j at height q , $i, j = 0, 1, 2, 3, p, q = 1, \dots, n$ are deduced from the Eq. (5) as

$$\begin{aligned}\epsilon_{\text{random},i}^p &= \hat{\mathbf{x}}_i^p - \mathbf{x}^p - \epsilon_{\text{bias},i}^p, \text{ and} \\ \epsilon_{\text{random},j}^q &= \hat{\mathbf{x}}_j^q - \mathbf{x}^q - \epsilon_{\text{bias},j}^q.\end{aligned}\quad (6)$$

By definition, the correlation coefficient $\rho_{X,Y}$ between random variables X and Y is

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}, \quad (7)$$

where cov stands for covariance, E stands for mathematical expectation operator, μ_X stands for the expectation values of X , μ_Y stands for the expectation values of Y , σ_X stands for the standard deviation of X , and σ_Y stands for the standard deviation of Y . In our case, the random variables would be the random error of the values of processor i at height p and the random error of the values of processor j at height q ($i = 0, 1, 2, 3$). Since the aim of our merging operation is to improve the precision, the weight of each parent profile shall be associated with the its precision and not with its accuracy. Thus, we use debiased differences to infer the inter-profile correlation coefficients. N is the number of profiles in the whole 2007–2008 sample. On each geolocation k , $k = 1, \dots, N$, we will use the best estimate of the truth \mathbf{x} that we have, namely the mean profile of the four processors on this geolocation; we note it by $\bar{\mathbf{x}}_k$, $k = 1, \dots, N$, and its p th level is noted by \bar{x}_k^p . The expectation of X (or Y) is estimated as the mean debiased difference between the retrieved profile and the truth over all geolocations:

$$\mu_X = \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{i,l}^p - \mathbf{x}_l^p) \text{ and } \mu_Y = \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{j,l}^q - \mathbf{x}_l^q), \quad (8)$$

where the vector $\hat{\mathbf{x}}_{i,l}$ is the profile retrieved by processor i on the l th geolocation and $\hat{\mathbf{x}}_{i,l}^p$ is its p th level; \mathbf{x}_l is the true profile on the l th geolocation and \mathbf{x}_l^p is its p th level. Then the realization of $X - \mu_X$ on the k th geolocation is

$$\begin{aligned}(X - \mu_X)_k &= \hat{\mathbf{x}}_{i,k}^p - \mathbf{x}_k^p - \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{i,l}^p - \mathbf{x}_l^p) \\ &= \hat{\mathbf{x}}_{i,k}^p - \bar{\mathbf{x}}_k^p - \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{i,l}^p - \bar{\mathbf{x}}_l^p)\end{aligned}\quad (9)$$

(the true profile \mathbf{x}_k was replaced by its estimate $\bar{\mathbf{x}}_k$), and similarly the realization of $Y - \mu_Y$ on the k th geolocation is

$$(Y - \mu_Y)_k = \hat{\mathbf{x}}_{j,k}^q - \bar{\mathbf{x}}_k^q - \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{j,l}^q - \bar{\mathbf{x}}_l^q). \quad (10)$$

We use hence the following estimator of the correlation between the random errors $\epsilon_{\text{random},i}^p$ and $\epsilon_{\text{random},j}^q$. The correlation coefficient r_{ij}^{pq} between the random errors of processor i at height p and the random errors of processor j at height q is

$$r_{ij}^{pq} = \frac{\sum_{k=1}^N (\hat{\mathbf{x}}_{i,k}^p - \bar{\mathbf{x}}_k^p - \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{i,l}^p - \bar{\mathbf{x}}_l^p)) (\hat{\mathbf{x}}_{j,k}^q - \bar{\mathbf{x}}_k^q - \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{j,l}^q - \bar{\mathbf{x}}_l^q))}{\sqrt{\sum_{k=1}^N (\hat{\mathbf{x}}_{i,k}^p - \bar{\mathbf{x}}_k^p - \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{i,l}^p - \bar{\mathbf{x}}_l^p))^2 (\hat{\mathbf{x}}_{j,k}^q - \bar{\mathbf{x}}_k^q - \frac{1}{N} \sum_{l=1}^N (\hat{\mathbf{x}}_{j,l}^q - \bar{\mathbf{x}}_l^q))^2}}. \quad (11)$$

In this formula, the third term in each outer sum is the bias of corresponding processor, by taking it out of the first term we obtain a debiased profile, and the second term in the bracket is the mean around which the variation of debiased profiles is calculated.

Note that the obtained matrices are not symmetric, which is to be expected: there is no reason why the random errors of Bologna at height 20 km and random errors of KIT at height 35 km would be correlated exactly as the random errors of Bologna at height 35 km and random errors of KIT at height 20 km.

Figure 1 shows that the errors are correlated for all six pairs; this correlation is not negligible: with a minimal value of -0.6 and maximum of 0.26 . This means that the coefficients r_{ij}^{pq} can not be assumed zero and must be taken into the merging formula. Figure 1 also demonstrates that error correlations between the retrieved profiles from different processors are not mostly due to error correlations in the measured spectra: Oxford and ESA use identical measurements (see Table 1), but the highest correlation is observed in Bologna–Oxford and KIT–Oxford cases. However, the similarities/differences in the retrieval algorithms seem to also play a role, and that could explain some high absolute values of the coefficients. A gain in precision of the merged product can be expected if these correlations are adequately taken into account.

4 Statistical covariance matrices

In the previous section we have discussed how the inter-processor correlations were diagnosed. Now we turn to the

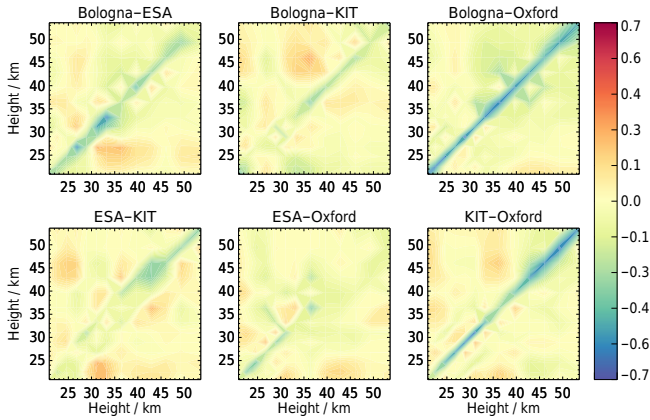


Figure 1. Correlation of errors of four processors calculated by Eq. (11). Obtained matrices are not symmetric, which is to be expected. The errors are non-negligibly correlated for all six pairs, which means that the coefficients can not be assumed zero and must be taken into the merging formula.

inter-level error covariances of each single processor. These error covariance matrices are needed for each single processor to construct the processor intercorrelation matrix as given by Eq. (2). Only two processors, ESA and KIT, provide the covariance matrices for each profile. For the other two processors, only statistical error covariance matrices can be evaluated empirically. The error covariance matrices are taken into the merging for controlling the weight of each processor in the average. Thus, it is more important to evaluate the covariance matrices in a consistent way than to have a particularly good covariance matrix for a subset of profiles. Therefore, we have decided to use statistical covariance matrices for all four processors. In order to reduce the correlation due to natural atmospheric variability, we calculate it on summer profiles in the 20° S–20° N latitude band. The mean of all summer tropical stratospheric profiles of a processor i is noted $\bar{x}_{i, \text{mean}}$. The error covariance matrix of a processor i is evaluated from the sample of debiased summer tropical stratospheric profiles:

$$\mathbf{S}_{\mathbf{x}_i} = \frac{1}{N} \sum_{k=1}^N \left(\mathbf{x}_{i,k}^p - \bar{\mathbf{x}}_{i, \text{mean}}^p - \frac{1}{N} \sum_{l=1}^N \left(\mathbf{x}_{i,l}^p - \bar{\mathbf{x}}_{i, \text{mean}}^p \right) \right) \times \left(\mathbf{x}_{i,k}^q - \bar{\mathbf{x}}_{i, \text{mean}}^q - \frac{1}{N} \sum_{l=1}^N \left(\mathbf{x}_{i,l}^q - \bar{\mathbf{x}}_{i, \text{mean}}^q \right) \right). \quad (12)$$

Note that this formula can also be obtained from the formula for the correlation of errors by taking $i = j$ and replacing the mean of four profiles at each geolocation by the mean of the processor in the summer tropical stratosphere. The obtained error covariance matrices are shown in Fig. 2. The results obtained are consistent with the error bars validation from Laeng et al. (2015).

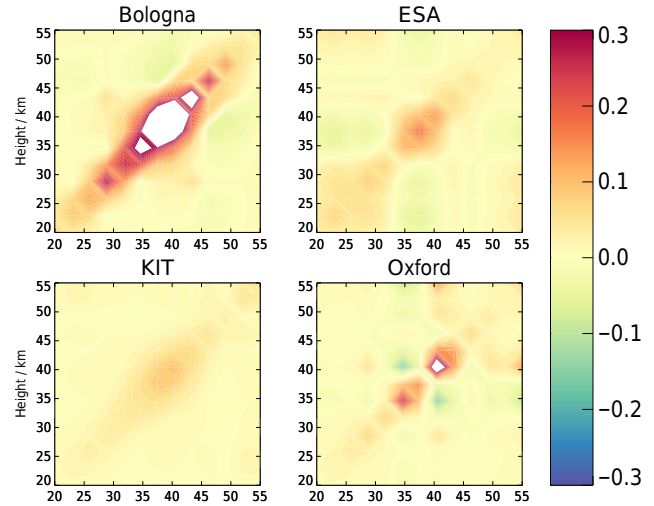


Figure 2. Statistical covariance matrices of four parent MIPAS processors. The white areas in Bologna and Oxford plots are for values bigger than 0.3: up to 0.82 for Bologna and up to 0.44 for Oxford processors.

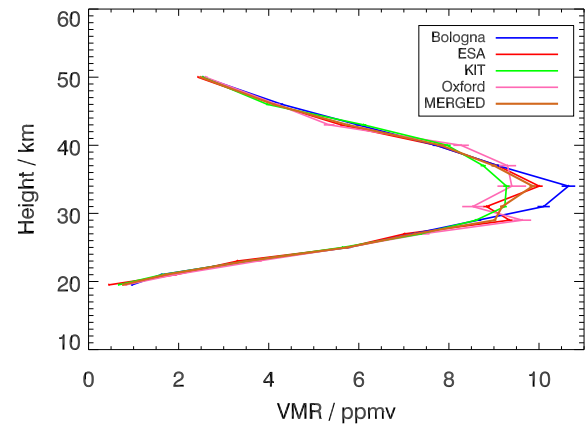


Figure 3. Parent MIPAS profiles and the resulting merged MIPAS profile on geolocation 33441_20080723T072843Z (0.2° S, 40.5° E). The profile from the Bologna processor is the blue line; from the ESA processor, red; from the KIT processor, green; and from the Oxford processor, pink. The merged profile is the brown line.

5 Merged profiles and their validation

Figure 3 shows an example of the merging of the four individual parent profiles into one merged profile for the tropical summer geolocation 33441_20080723T072843Z (0.2° S, 40.5° E).

Merging of various data products from the same instrument is not necessarily supposed to remove the bias of the parent datasets. Instead, it is supposed to ameliorate the precision of the product since the parent processors rely on different spectral information (different microwindows). At heights where the precision of the merged product is better

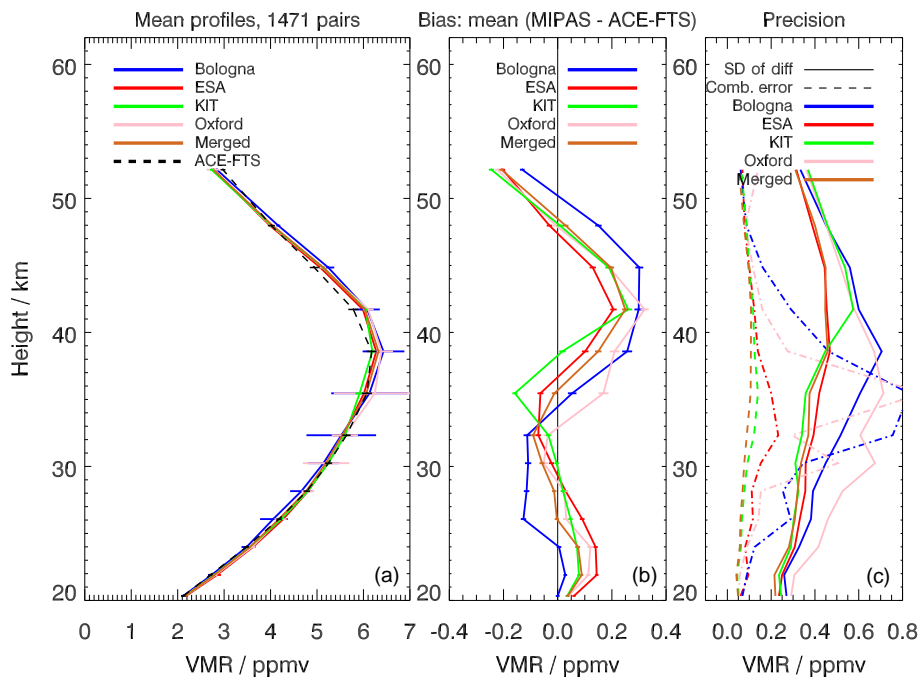


Figure 4. Mean profiles (a), bias (b), and precision validation (debiased standard deviation of the differences, c) of the four parent datasets and merged MIPAS dataset with respect to ACE-FTS ozone profiles in 2007–2008.

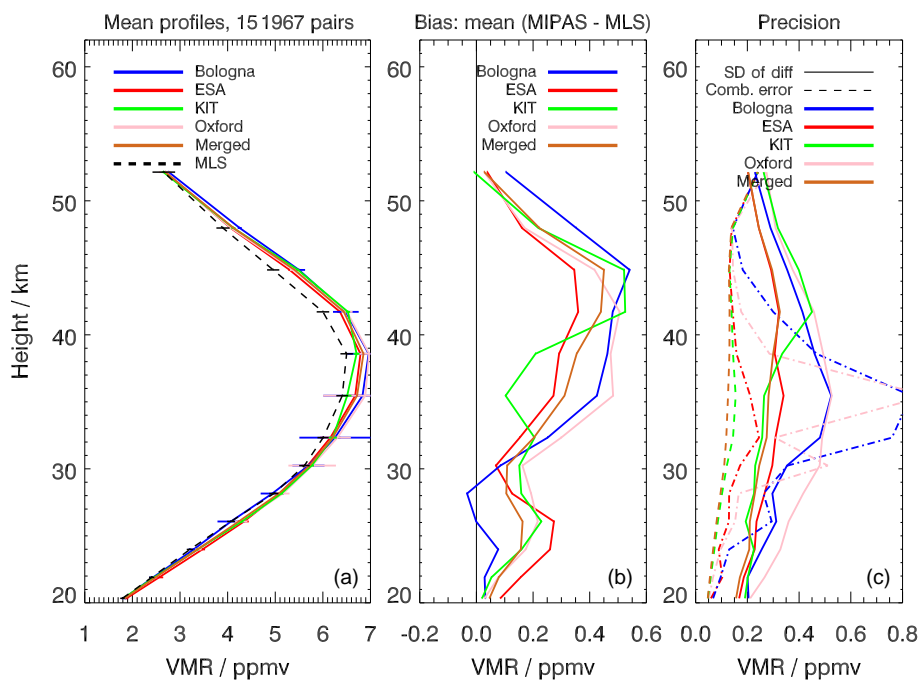


Figure 5. Mean profiles (a), bias (b), and precision validation (debiased standard deviation of the differences, c) of the four parent datasets and merged MIPAS dataset with respect to MLS ozone profiles in 2007–2008.

than the precision of any of the parent datasets, the merging is successful. Figure 4 shows simultaneous comparison of the four parent MIPAS datasets and the merged MIPAS dataset with ACE-FTS version 3.5 ozone dataset, for collocation criteria 5 h and 500 km. In terms of precision, hence, the merging is a success at 20–28 and 39–43 km. At 28–38 km, KIT's precision in terms of the standard deviation of the differences is better than the precision of the merged product. At 44–52 km, ESA's precision is better than the precision of the merged product. All these conclusions are conditional to the assumption that ACE represents the truth. Although the merging is not supposed to remove the bias, it can change the bias through the averaging, and depending on the signs of biases of parent profiles, this could lead to the improvement of the bias. This is what is happening in the comparison with ACE-FTS: at 24–28 and 33–37 km, the merged product agrees with ACE $\sim 1\%$ better than KIT, while at all other heights, KIT agrees better. Interestingly, in integrated view over the altitude range around the ozone VMR (volume mixing ratio) peak, where all four processors have a known positive bias (Laeng et al., 2015), the merged product is performing $\sim 0.8\%$ better than any of the four processors (Fig. 4b).

Figure 5 shows simultaneous comparison of the four parent MIPAS datasets and the merged MIPAS dataset with MLS version 3.3 dataset, for collocation criteria 4 h and 250 km. At 20–25 km, the precision of the merged product is better than the precision of any individual dataset. When looking at the whole height range, the overall precision of the merged product is better than the precision of any of the parent datasets. The overall agreement of the merged product with MLS is not as good as that with ACE-FTS; this is also the case for all parent datasets. In terms of bias, the merged product performs $\sim 1.5\text{--}2\%$ better at 24–33 and 41–45 km, while KIT performs better at the remaining heights. In particular, unlike for the comparison with ACE, around the ozone VMR peak, the agreement of KIT is $\sim 2.5\text{--}2.8\%$ better than the agreement of the merged dataset.

6 Conclusions

We created a 2-year dataset of merged ozone profiles from four independent MIPAS Level-2 processors. The novelty of the product is a mathematically clean way of performing the merging: the weighting of parent profiles is realized by corresponding inverse error covariance matrices, the correlations between different profile levels are considered, and the inter-correlations between processors' errors are evaluated statistically and are used in the merging. In comparison to the individual parent datasets, the merged product has a restricted height range (20–55 km) and only a statistical covariance matrix can be provided. Validation of the merged dataset is performed by comparing with ozone profiles from ACE-FTS and MLS. Comparison with ACE-FTS looks better than with MLS. This is, however, the case for all parent processors as

well. Despite the fact that the merging is not supposed to remove the bias, the high bias around the ozone VMR peak known for the parent profiles is reduced in comparison with ACE-FTS (but not with MLS). The overall precision of the merged product is better than that of any of the four processors. This product could therefore be of use in specific studies requiring improved precision of the MIPAS ozone record.

Data availability. The merged MIPAS data product is available at <http://www.esa-ozone-cci.org> (Laeng, 2016).

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was performed in the frame of European Space Agency (ESA) project Ozone_cci. All four MIPAS teams acknowledge ESA for providing MIPAS L1b data. The ACE mission is supported primarily by the Canadian Space Agency. Work at the Jet Propulsion Laboratory was performed under contract with the National Aeronautics and Space Administration.

We acknowledge support by Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of Karlsruhe Institute of Technology.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

Edited by: M. Weber

Reviewed by: two anonymous referees

References

- Carlotti, M.: Global-fit approach to the analysis of limb-scanning atmospheric measurements, *Appl. Optics*, 27, 3250–3254, 1988.
- Carlotti, M., Brizzi, G., Papandrea, E., Prevedelli, M., Ridolfi, M., Dinelli, B. M., and Magnani, L.: GMTR: Two-dimensional geofit multitarget retrieval model for Michelson Interferometer for Passive Atmospheric Sounding/Environmental Satellite observations, *Appl. Optics*, 45, 716–727, 2006.
- Ceccherini, S., Carli, B., and Raspollini, P.: Equivalence of data fusion and simultaneous retrieval, *Opt. Express*, 23, 8476–8488, 2015.
- Dinelli, B. M., Arnone, E., Brizzi, G., Carlotti, M., Castelli, E., Magnani, L., Papandrea, E., Prevedelli, M., and Ridolfi, M.: The MIPAS2D database of MIPAS/ENVISAT measurements retrieved with a multi-target 2-dimensional tomographic approach, *Atmos. Meas. Tech.*, 3, 355–374, doi:10.5194/amt-3-355-2010, 2010.
- Dudhia, A., Jay, V. L., and Rodgers, C. D.: Microwindow selection for high-spectral-resolution sounders, *Appl. Optics*, 41, 3665–3673, 2002.
- Echle, G., von Clarmann, T., Dudhia, A., Flaud, J.-M., Funke, B., Glatthor, N., Kerridge, B., López-Puertas, M., Martín-Torres,

- F. J., and Stiller, G. P.: Optimized spectral microwindows for data analysis of the Michelson Interferometer for Passive Atmospheric Sounding on the Environmental Satellite, *Appl. Optics*, 39, 5531–5540, 2000.
- Kleinert, A., Aubertin, G., Perron, G., Birk, M., Wagner, G., Hase, F., Nett, H., and Poulin, R.: MIPAS Level 1B algorithms overview: operational processing and characterization, *Atmos. Chem. Phys.*, 7, 1395–1406, doi:10.5194/acp-7-1395-2007, 2007.
- Laeng, A.: Merged 2007–2008 ozone profiles from four MIPAS processors, *Ozone_cci*, available at: <http://www.esa-ozone-cci.org> (last access: 21 April 2017), 2016.
- Laeng, A., Hubert, D., Verhoelst, T., von Clarmann, T., Dinelli, B. M., Dudhia, A., Raspollini, P., Stiller, G., Grabowski, U., Keppens, A., Kiefer, M., Sofieva, V., Froidevaux, L., Walker, K. A., Lambert, J.-C., and Zehner, C.: The Ozone Climate Change Initiative: Comparison of four Level-2 Processors for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), *Remote Sens. Environ.*, 162, 316–343, doi:10.1016/j.rse.2014.12.013, 2015.
- Raspollini, P., Carli, B., Carlotti, M., Ceccherini, S., Dehn, A., Dinelli, B. M., Dudhia, A., Flaud, J.-M., López-Puertas, M., Niro, F., Remedios, J. J., Ridolfi, M., Sembhi, H., Sgheri, L., and von Clarmann, T.: Ten years of MIPAS measurements with ESA Level 2 processor V6 – Part 1: Retrieval algorithm and diagnostics of the products, *Atmos. Meas. Tech.*, 6, 2419–2439, doi:10.5194/amt-6-2419-2013, 2013.
- von Clarmann, T. and Echle, G.: Selection of optimized microwindows for atmospheric spectroscopy, *Appl. Optics*, 37, 7661–7669, 1998.
- von Clarmann, T., Glatthor, N., Grabowski, U., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Mengistu Tsidu, G., Milz, M., Steck, T., Stiller, G. P., Wang, D. Y., Fischer, H., Funke, B., Gil-López, S., and López-Puertas, M.: Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), *J. Geophys. Res.*, 108, 4736, doi:10.1029/2003JD003602, 2003.
- 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, H₂O, O₃, HNO₃, CH₄, N₂O, ClONO₂ 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.