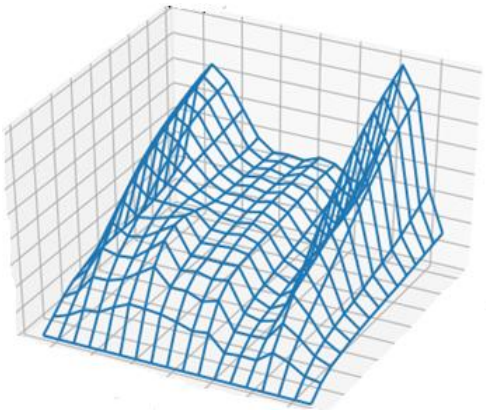
TITLE:

VACUUM'R



Final Report

Reference: VACUUM'_KIT_FR_01_00

Date of issue: 29/06/2021

Distributed to:

Claus Zehner ESA
VACUUM'R team

DOCUMENT PROPERTIES

Title	VACUUM'R Final Report
Reference	VACUUM'_KIT_FR_01_00
Issue	01
Revision	00
Status	Final
Date of issue	29/06/2021
Document type	Report

	FUNCTION	NAME	DATE	SIGNATURE
LEAD AUTHOR	Project Manager	Alexandra Laeng	29/06/2021	
CONTRIBUTING AUTHORS	Science Leader Team Members	Thomas von Clarmann		
REVIEWED BY	Science Leader	Alexandra Laeng		
ISSUED BY	Project Manager	Alexandra Laeng		

DOCUMENT CHANGE RECORD

Issue	Revision	Date	Modified items	Observations
00	00	23/05/2021	-	Creation of the document
00	01	28/06/2021	Comments of co-authors of Regression Function paper were inserted	
01	00	29/06/2021	Comments of team members were incorprated	Finalisation of the document

Chapter 1

Satellite data validation: a parametrization of the natural variability of atmospheric mixing ratios

This project tackles a problem which typically arises when remotely sensed data from different instruments are compared within the framework of validation studies. In quantitative validation, the common approach is to calculate differences of pairs of measurements of the same airmass by the two instruments under comparison. With the aid of χ^2-statistics it is tested if the observed differences can be explained by the estimated error of the differences Rodgers and Connor (2003). The estimated error of the differences includes measurement noise and parameter as well as model errors, as far as they are uncorrelated between the instruments von Clarmann (2006). Also different impact of prior information on the result has to be considered. But still, more often than not, the differences will be too large to be explained by the combined error budget of the measurements under comparison. The reason is that the instruments typically do not sound exactly the same airmass. Spatial and temporal mismatch of the measurements along with natural variability of the state variable measured contributes to the observed differences. This source of differences is quantified in validation papers only in a few exceptional cases (see, e.g., Sheese et al. (2021) for an example where models were used to quantify the related effect of ozone variability).

Instead, natural variability is often used as a universal excuse to defend measurements of which the validation studies suggest that the related retrieval errors are underestimated.

In this study, we present a user-friendly tool to provide quantitative estimates of that component of the differences of observations which can be attributed to the spatial and temporal mismatch and natural variability. The underlying method is based on high-resolved fields of temperature and mixing ratios of trace species as described in Section 1.1. From these model fields the typical variabilities are evaluated as a function of spatial and temporal mismatch

(Section 1.2). In order to avoid unnecessary large data traffic and to reduce as much as possible the impact of model imperfections on the calculated fields, parametrizations of these dependencies, for different trace gases, altitudes, and latitude bins, are developped, by prescribing to the natural vairability function a particular shape, arising from general theory of random functions with stationnary increments (Section 1.3) and confirmed by calculations out of model data. A re-parametrization is developed to improve the validity of the inductive generalization towards other gases and seasons (Section 1.5). The adequacy of our suggested method is critically discussed in Section 1.6 and final recommendations are given in Section 1.7.

1.1 Model data

For this study we have used the data from two models.

BASCOE (Belgian Assimilation System for Chemical ObsErvations) Eulerian chemistry transport model has been developed at BIRA (Belgisch Instituut voor Ruimte-Aeronomie) to be operated in a 4D variational data assimilation context. Model fields used in this study have been produced by the Belgian Assimilation System for Chemical ObsErvations (BASCOE, Errera et al. (2008)). While usually used in the context of stratospheric chemical data assimilation (e.g. Errera et al. (2019)), the Chemistry Transport Model (CTM) of the system has been also used to study of the evolution of the stratospheric composition (Chabrillat et al., 2018; Prignon et al., 2019; Minganti et al., 2020). Here, BASCOE was run for the period 25 Sept-1 Oct 2008 where wind and temperature were taken from the ERA-Interim reanalysis (Dee et al., 2011). The model was run on an 1°x1° horizontal grid, the native 60 vertical levels of ERA-Interim (from the surface to 0.1 hPa) with a time step of 30 minutes. Hourly global fields of 28 relevant trace gases1 were used for this study. For the particular model run used here BASCOE was operated at $1° \times 1°$ horizontal resolution. The altitude grid matched that of the ECMWF (European Centre for Medium-Range Weather Forecasts) ERA-Interim Dee et al. (2011) analysis by which the model was driven. BASCOE provided global fields of all relevant trace gases[1] and temperature for one week at the end of September 2008 at a time resolution of one hour.

WACCM6 (Whole Atmosphere Community Climate Model 6) is the atmospheric component of the Community Earth System Model, Version 2 (CESM2) (Danabasoglu et al., 2019; Emmons et al., 2019; Gettelman et al., 2019; Tilmes et al., 2019) run in UCAR/NCAR/ACOM. The model has a horizontal resolution of 0.9x1.25 degrees with 88 vertical hybrid sigma-pressure levels, and is run using specified dynamics, with nudging of temperature, and U/V winds from the NASA Goddard Earth Observing System, Version 5 (GEOS-5) forecast model.. They contain the fields of 3 species (O_3, H_2O, NO) for 4 weeks in year 2020, one in each season. The data were regridded on the same fixed height grid with 1-km step, using the fields of geopotential heights and temperature, provided with the data.

[1] BrNO, BrO, CCl_4, CFC_{11}, CFC_{12}, CFC_{113}, CH_3Cl, CH_4, ClO, $ClONO_2$, CO, CO_2, H_2O, HBr, HCl, HF, HNO_3, HNO_4, HO_2, HOBr, HOCl, N_2O, N_2O_5, NO, NO_2, NO_3, O_3, OH

1.2 Variabilities

1.2.1 Structure functions

Let X be a random variable defined by the amount of the target trace gase in a given infinitely small air parcel, centered around a point in the atmosphere at a given moment of time, reported in volume mixing ratio (vmr). The amount of the trace gase in any point of the atmosphere at a fixed time (or in any moment of time in a fixed point of the atmosphere) can be viewed as a state of a one-dimensional random process $X(t)$, where t parametrizes distance from an initial point (or time elapsed from an initial moment). This random process is not stationnary, because its statistical characteristics can change with t. The increments $X(t+\tau) - X(t)$ of the process $X(t)$ represent the change of the amount of the trace gase over distance (or over time). In a given sufficiently narrow latitude band, at a fixed altitude, and in a given season, the distribution of the differences $X(t+\tau) - X(t)$ does not depend on t, which means that $X(t)$ is the process with stationnary increments. The basic characteristics of real-valued random process with stationnary increments are the mean value of the increment $E[X(t+\tau) - X(t)]$ and the correlation function of the increment

$$D(\tau) = E|X(t+\tau) - X(t)|^2 \tag{1.1}$$

also called *the structure function* of the process $X(t)$ (Yaglom (1986), ch.23). In our case, as $E[X(t+\tau) - X(t)] = 0$,

$$\begin{aligned} D(\tau) = E|X(t+\tau) - X(t)|^2 = \\ = \sigma^2(|X(t+\tau) - X(t)|) + (E|X(t+\tau) - X(t)|)^2 = \\ = \sigma^2(|X(t+\tau) - X(t)|). \end{aligned} \tag{1.2}$$

The natural variability of a trace gase is the square root of the structure function of the process $X(t)$:

$$\sqrt{D(\tau)} = \sigma(|X(t+\tau) - X(t)|). \tag{1.3}$$

The above provides the formal link between the intuitive definition of the natural variability as the variability of differences and the mathematical machinery of random processes with stationnary increments which is widely used in stydying the processes at smaller spatio-temporal scales, for exemple, in the theory of atmospheric turbulence. This link will allow us to draw conlusions about the nature of the process $X(t)$ basing on the shape of statistics obtained and will justify the choice of the form of the parametrisation of natural variability. Next section explains how the estimation of $\sqrt{D}$ is done out of model fields.

1.2.2 Estimation of variabilities

In order to obtain a statistic of the variability of differences, in a first step the model fields were transformed from their native hybrid altitude grid to a fixed 1-km step geometrical height grid. For this, the geopotential height for each model knot was restored and transformed into geometrical height using the temperature values from the model, which allowed the interpolation of the profiles on a fixed altitude grid. In the second step, the model fields were smoothed

according to the horizontal resolution of the instruments whose precision is to be validated. We have chosen as an example the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS, Fischer et al. (2008)). Its cross track resolution corresponds roughly to an East-West resolution and is defined by the width of the instantaneous field of view at the tangent point, which is 30 km. The along-track smearing, which corresponds roughly the the North-South horizontal resolution is on average roughly 200 km (von Clarmann et al., 2009). This smoothing operation does not influence the shape of the obtained curves, but does reduce the obtained variabilities values by around 0.05%. No vertical smoothing is applied because vertical smoothing typically is considered in validation in explicit manner via the averaging kernels. These smoothed fields are the basis for the statistics of horizontal and temporal variability of the atmospheric state.

Horizontal Variability

We take the model data within a fixed 10-degree latitude bin and at a fixed height; as model dataset used covers only one week, there is no need to fix a season. For all possible pairs of points in the obtained subset, the normalized differences of the volume mixing ratio (vmr) of the target trace species within a predefined radius of 1500 km are calculated:

$$\frac{VMR(location_1, t) - VMR(location_2, t)}{VMR_{mean}}, \tag{1.4}$$

with VMR_{mean} being the mean VMR values of the target trace gas in the chosen latitude band at the chosen height. The constant time index t indicates that only differences are considered where the subtractor and the subtrahend refer to the same time. These differences are binned according to their horizontal separation distance. The following bins were used: 0 to 100 km, 100 to 200 km, etc between the two points. We calculate the standard deviation of the sample of these normalized differences, it provides us with an estimator of natural variability of the target trace gase as function of distance separation. The obtained natural variability of ozone at 35 km altitudes as a function of distance is shown on the left panel of Figure 1.1. The fast growth of the variability for separation distances over 1000 km at high northern latitudes reflects that in many pairs of the corresponding sample, one point lies inside the polar vortex, another lying outside. Note also that the calculated variability values for the distances 0-100 km ist zero for the tropical latitudes (yellow curves at Figure 1.1), and present a peak at subtropical latitudes (clear orange and clear green curves at Figure 1.1). The sample at low latitudes are empty or very small for this distance separation, because the resolution of BASCOE is 1°x1° which is bigger than 100 km at low latitudes. Hence the point 100 km will not be taken into account in the calculation of regression coefficients at these latitudes.

Temporal Variability

Similar as above, for all possible pairs of data points of the entire data set the differences of the volume mixing ratio of the target trace gas within a predefined

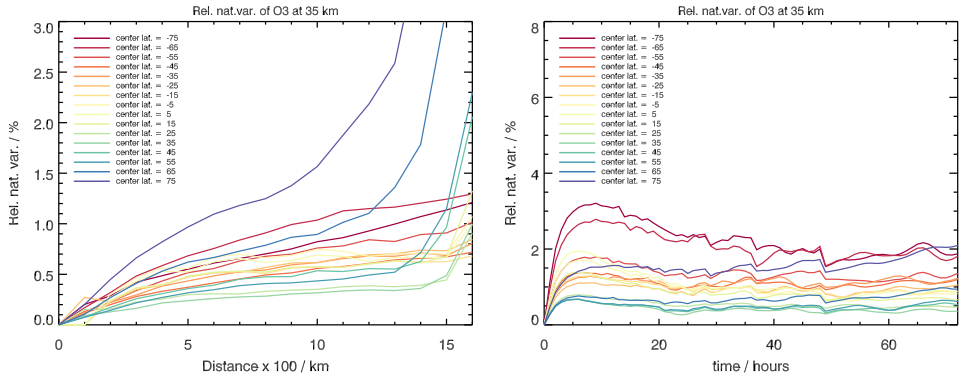


Figure 1.1: Left panel : natural variability of O_3 at 35 km altitude as function of horizontal distance. Right panel : natural variability of O_3 at 35 km altitude as function of time separation.

time period of 72 hours are calculated:

$$\frac{VMR(location, t_1) - VMR(location, t_2)}{VMR_{mean}}. \tag{1.5}$$

The constant location index *location* indicates that only differences are considered where the subtractor and the subtrahend refer to the same location. These differences are sorted according to their time lag. Similar to the horizontal variability, for each time lag, the differences are normalized by the mean vmr within given latitude band at given altitude, then the standard deviation of the sample of normalized differences is calculated. This quantity is the estimator of natural variability of the target species as function of time separation; its values for ozone at 35 km altitude are shown on the right panel of Figure 1.1. As for most of satellite validation exercise, time separation within collocation criteria stays within 5 hours, we made the choice to restrain our analysis to the time separation lag to maximum of 5 hours.

1.2.3 Combination of Horizontal and Temporal Variability

Despite the fact that advection can admittedly cause the correlations between horizontal and temporal components of the variability, at scales considered here, we assume that horizontal and temporal variation of the atmospheric state are uncorrelated. In our analysis we offer independent parametrizations for each of them, which are recommended to be combined by their quadratic sum; we also provide a software performing this summation for a reparametrisation of these quantities on latitudinal gradients. Tests using a statistic of combined horizontal and temporal differences of the type

$$VMR(location_1, t_1) - VMR(location_2, t_2) \tag{1.6}$$

have shown that at scales considered here, the error due to the neglect of correlations is below 0.1% and thus not usually worth an additional effort.

1.3 Parametrization

1.3.1 Motivation

The goal of the present work is to provide the community with information on the natural variability of mixing ratios of trace gases as a function of distance and time separation. This information is meant to be used in the context of validation studies. Instead of providing the entire variability data set, we consider the use of a simple and easy-to-use parametrization as more adequate. The reasons are these. First, the use of parametrizations avoids a considerable amount of data traffic. Second, the fine structure of the fields reflects the actual conditions of the days actually covered by the model run rather than the general behaviour of the atmosphere. And third, our parametrization by continuous regression functions allow easy interpolation.

1.3.2 The regression function

In view of the shape of the curves produced out of model's data (section 1.2.2 and 1.2.2), the natural variability function can be paramerized in the form $D(\tau) = A\tau^{\gamma}$ with $A > 0$ and $0 < \gamma < 1$. An interesting side conclusion that can be made from the obtained shape of the structure function of $X(t)$ is that the process of atmospheric variability (horizontal as well as temporal) of mixing ratios is self-similar: the form of its structure function $D(\tau) = A\tau^{\gamma}$ is invariant under a group of similarity transformations $t \to ht$, $X \to a(h)X$ (Kolmogorov, 1940; Yaglom, 1986); in other words, no characteristic scale can be associated with their structure function. Note, that atmospheric variability as function of distance and time separation could be approximately represented by a two-dimensional random process; this is however out of scope of this paper: our choice is to treat the distance and time mismatch dependences separately, because this is what a typical validation exercice does.

As pointed out in the Section 1.2.2, at high latitudes at the distances over 1000 km the variability grows very rapidely. Also, for low latitudes, the values of variability as function of distance mismatch is meaningless at 100 km : it is calculated on the samples from too small to empty, because of the model horizontal resolution. The choice of 5 hours upper limit for the time dependent variability is driven by the typical values of the time mismatch occuring in satellite validation studies, and the shape of the obtained curves. We calculate the regression coefficients A and γ by minimising the quantity

$$\sum_{i=2}^{10}(y_i - Ax_i{}^{\gamma})^2 \tag{1.7}$$

via Sequential Least SQuares Programming (SLSQP) optimizer (Kraft, 1988) and by giving the value at 100 km as a first constraint for A and 0.5 as a first constraint for γ. The obtained regression curves and the initially calculated model curves for some species are shown in Figure 1.2 for distance separation and Figure 1.3 for time separation. The initial and smoothed (regressed) natural variability surfaces as function of latitude and time separation are shown in Figure 1.4 for ozone at 35 km altitude. In a range between 100 and 1000 km the parametrizations fit the data very well.

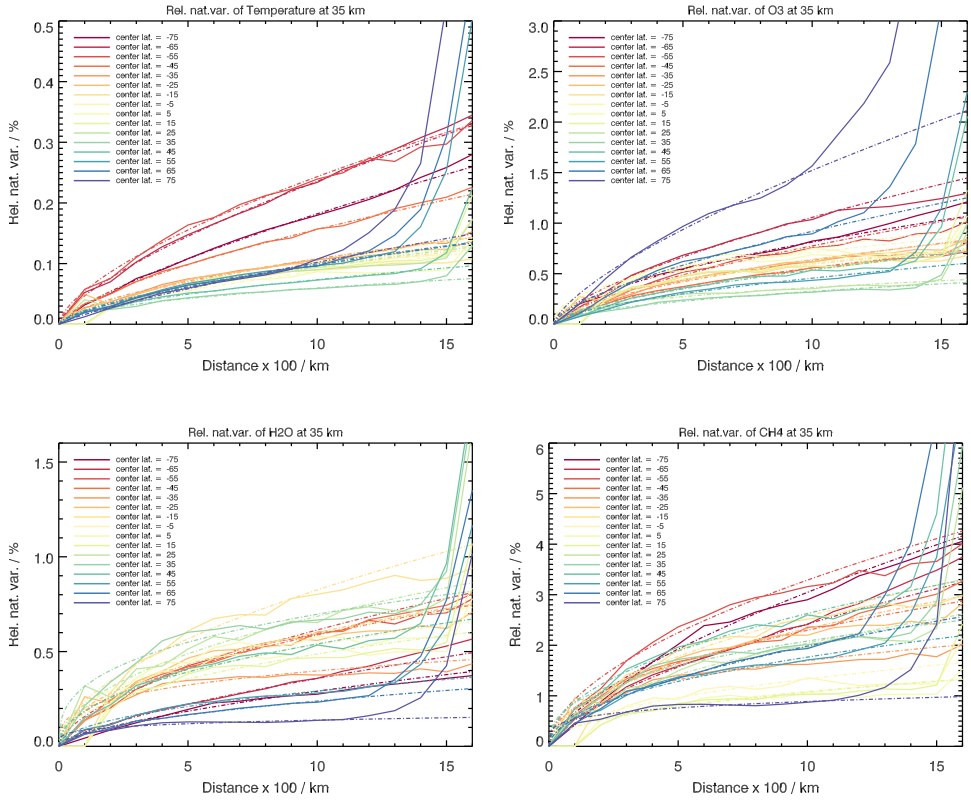


Figure 1.2: Natural variability as function of horizontal distance for temperature, O_3, CH_4, and H_2O (solid lines) and proposed parametrisation (dashed lines) in different latitude bins.

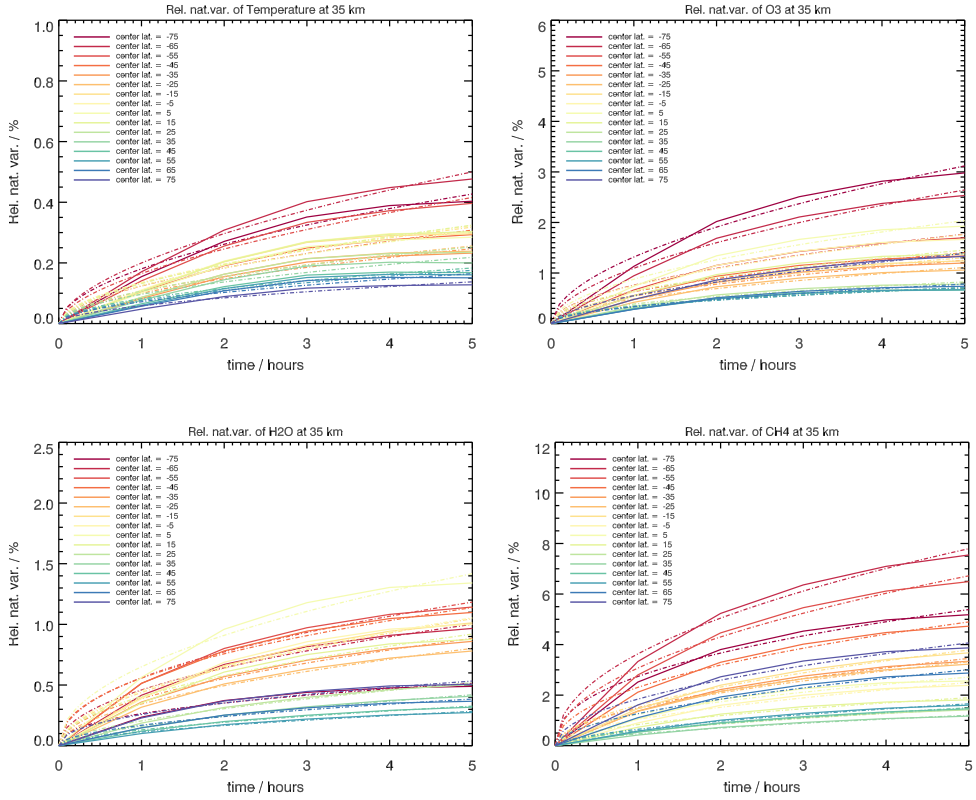


Figure 1.3: Natural variability as function of time separation for temperature, O_3, CH_4, and H_2O (solid lines) and proposed parametrisation (dashed lines).

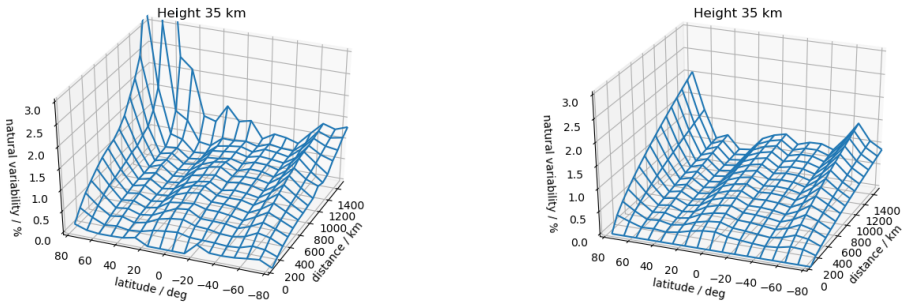


Figure 1.4: Natural variability of O_3 at 35 km altitude as function of horizontal distance separation. Left panel : model fields. Right panel : regressed fields.

1.4 Re-parametrization on latitudinal gradients

An obvious deficiency of our approach is that the variability fields are calculated out of only one week of the data. Operational constraints did not allow to generate a better coverage for these many species at the required resolution. The variability, which we have calculated as function of latitude and distance (time) mismatch, is also season-dependent, because different seasons correspond to different inclinations of the Earth axis and intiuitively all should be schifted in the latitudes while the season changes. There is however a way to encounter the problem: if the user is willing to calculate one additional quantity out of his data, namely the latitudinal gradients of the gas under validation, then the workaround consists in re-parametrisation of the variability on latitudinal gradients of the species. Latitudinal gradients of a gas are defined as

$$\left| \frac{VMR_{l_1} - VMR_{l_2}}{l_1 - l_2} \right|, \tag{1.8}$$

where VMR_{l_1} is the mean VMR of the gas in a latitude band, VMR_{l_2} is the mean VMR in the northwise neighboring latitude band, $l_1 - l_2$ is the width of the latitude band, in our case this is 10 degrees. In relative version of latitudinal gradient, this quantity is normalized with respect to the mean VMR in both bands, and is multiplied by 100. For a fixed distance (or time) mismatch x, in a first approximation, as the variability is calculated as the square root of the variance of the squares of differences, and the latitudinal gradients are calculated as differences, we expect a linear dependence of the variability from the latitudinal gradients.

To test the theoretical considerations above in practice, especially to see if the linear dependence of the natural variability from the latitudinal gradients is the same in different seasons, the link between the natural variability and the latitudinal gradients was tested on the data of a worse resolved model with less species, but in turn, for which the data from different seasons were available, were available, namely data from the Whole Atmosphere Community Climate Model, Version 6 (WACCM6).

The Figure 1.5 shows the ozone distribution in each season of WACCM data, after this manipulation of the data : the distributions are as expected, the latitudinal shift in different season is visible, so the data are suitable for our test.

The Figures 1.6, 1.7 and 1.8 show the natural variability as function of latitudinal gradients for a particular distance separation of 400 km in four seasons of WACCM for H_2O at 30 km, O_3 at 35 km, and NO at 40 km. Each point in these figures corresponds to one 10-degree latitude bands; for the sake of completeness we include the points from all the latitudes. One can observe that in all cases the regression lines are similar, some deviation in summer and winter come from the points corresponding to high (> 70 degrees) latitudes, which is expectable. This hints toward season-independence of linear approximation of the natural variability as function of latitudinal gradients. Hence, the natural variability calculated in just one season and reparametrized on latitudinal gradients of the specie, provides all the information needed for the validation exercice in any season.

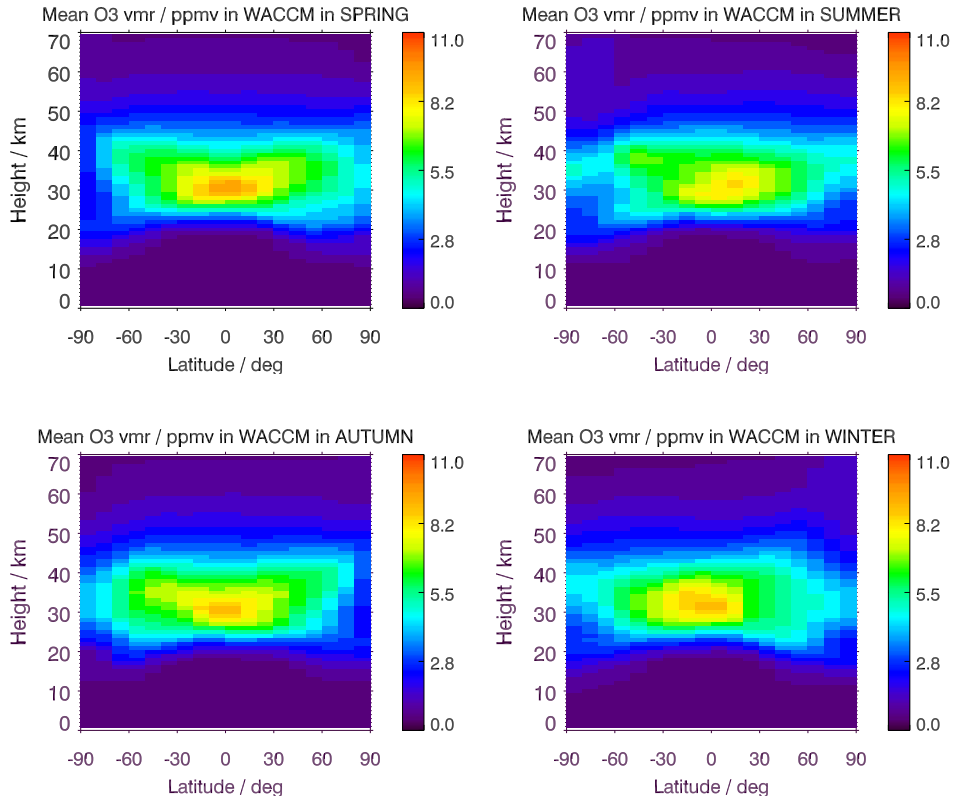


Figure 1.5: Distribution of O_3 in one week of each season in WACCM_2.1.

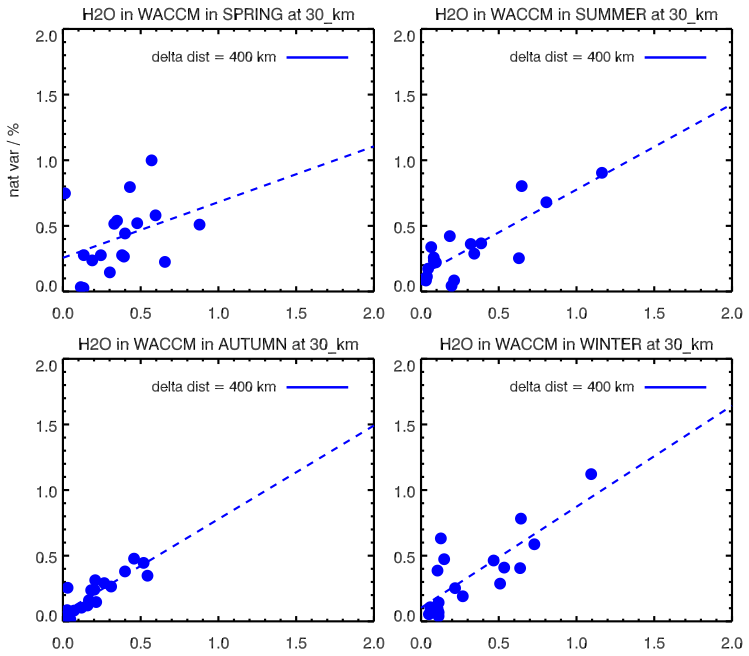


Figure 1.6: Natural variability of H_2O as function of latitudinal gradients and distance separation at 30 km altitude in four seasons in WACCM model.

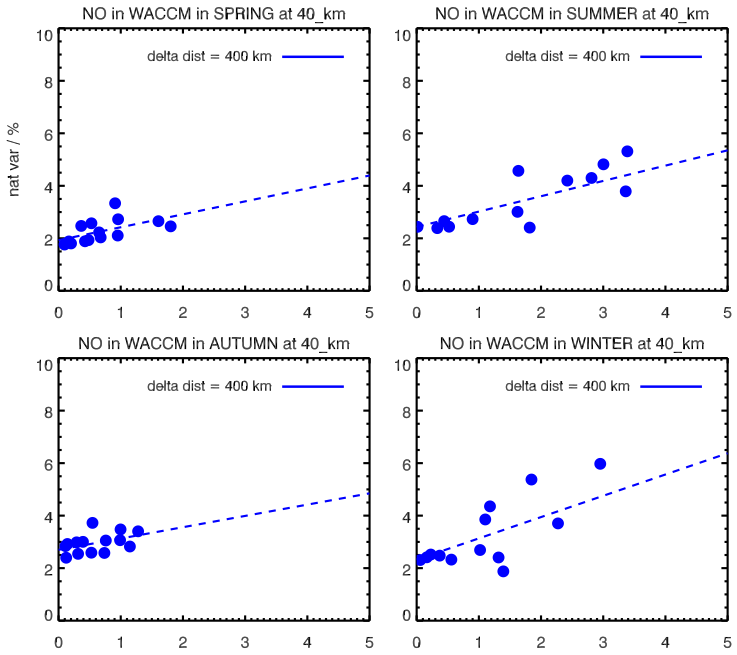


Figure 1.7: Natural variability of NO as function of latitudinal gradients and distance separation at 40 km altitude in four seasons in WACCM model.

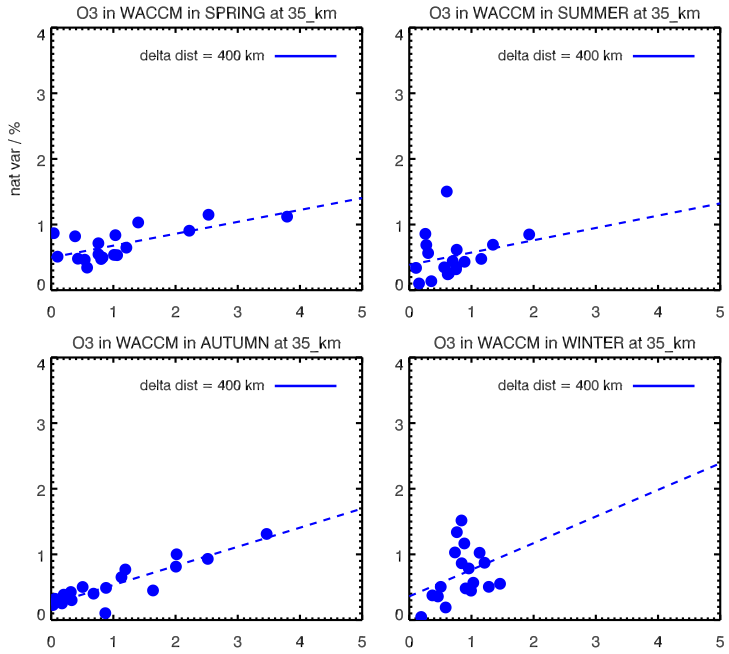


Figure 1.8: Natural variability of O_3 as function of latitudinal gradients and distance separation at 35 km altitude in four seasons in WACCM model.

1.5 What to do in practice: the software

The variabilites as functions of time and distance mismatch are added quadratically and this provides the final variability value for the collocation criteria choosen. In practice, the users will have to calculate out of their data under validation just one additional quantity, namely the latitudinal gradients of the species under validation. This quantity should be calculated on the whole sample, in order to increase the significance of the statistics. Together with regression coefficients values, we provide a software, which takes as input the species name (among 30 available), the distance and time mismatch chosen, the latitude band and the height, and as an output the user obtains a value of the natural variability of the gas, at altitude given, in latitude band given, for mismatch criteria given. If the validation study is perfomed in latitude domain larger then 10 degrees, than the values in correspoding 10-degree bands should be added quadratically.

1.6 Discussion

In the sense of the devil's advocate, we try to raise possible objections against our method and try to rebut them. Since we do not use model data directly but only differences between model data, additive model biases cancel out. Critical minds might plead that there still could be multiplicative biases in the model data, which would affect our statistics of differences. These are, however, not harmful either, if the gradient-related parametrization is used instead of the

latitude-month related parametrization. The reason is roughly this. A model bias affects the horizontal gradients in the same way as the differences used for our statistics. Thus, also the effect of a multiplicative model bias cancels out.

An obvious objection to our method is that the model data used cover only a short time period and might not be inductively generalizable towards other time periods. We agree that, due to the annual cycle, the typical meteorological regimes are shifted in latitude over the year. But again, when the gradient-related parametrization is used, most likely the statistics of the correct meteorological regime is chosen, even if it is found at different latitudes than the validation experiment. The explanation of this is that the natural variability of mixing ratios of most trace species is predominantly driven by the latitudinal gradients. It goes without saying that this does not hold for fast reacting species, and particularly such which are in a photochemical equilibrium.

These parametrizations should not be used where polar vortex dynamics may play a role or for spatial mismatches beyond 1000 km and temporal mismatches beyond 5 hours; but these situations are not the preferred validation scenarios anyway.

1.7 Conclusions

In validation exercice, a universal excuse used to explain the residual discrepancy between the data is the natural atmospheric variability due to the imperfect collocations. This work is the first attempt to quantify this atmospheric variability for big sample of atmospheric constituents and provide the user with a tool to substract from the residual variability the part coming from natural atmospheric variability. The fields of natural atmospheric variability as function of distance and time mismatch were calculated out of high-resolved BASCOE model data. The variability data were described by an easy-to-use regression function and the regression coefficients are provided to the community, together with the software that calculates for given gas, latitudinal gradient, height, and collocation criteria the value of correspoinding natural variability. An independence of linear approximation of the natural variability as function of latitudinal gradients from season was demonstrated on WACCM model data.

Chapter 2

Method to validate the uncertainties of atmospheric composition records

Uncertainties coming with an atmospheric data records are an important characterization of the quality of the data. Uncertainties provided are used for any scientific assessment performed with the data, during the data assimilation, for merging and intercomparison of the datasets. When two datasets are intercompared, the matched profiles do not typically sound the same air mass. This explain why the standard deviation of the differences between two collocated data products is larger than the estimated error of the differences. A part of the observed differences can be attributed to natural variability and less than perfect collocations. Still the standard deviation of the differences of two measurements will almost certainly not agree with the estimated error of the differences, see for instance the Figure 5 in Laeng et al. (2014). Whose error estimate is incorrect? Previous attempts to calibrate precision estimates were going through either getting rid of natural variability or using close measurements or involving external instrument (Laeng et al., 2015; Sofieva et al., 2014, 2021; Fioletov et al., 2006). This chapter describes, implements, and validates a method to validate the uncertainites, by involving an estimation of natural variability from high resolution model calculations.

2.1 Methodology

We assume that every retrieved atmospheric profile $\mathbf{vmr}_{inst}$ can be written as

$$\mathrm{vmr}_{instr} = \mathrm{vmr}_{truth} + \sigma_{syst} + \sigma,$$

where $\mathbf{vmr}_{truth}$ is a true state of the atmosphere, σ_{syst} is the systematic error of the processor, and σ is the random component of the uncertainty of the retrieved profile. This decomposition ignores components that are systematic in one domaine and random in another, but it remains a reasonable assumption. For two instruments, the variance of differences of retrieved collocated profiles

can than be written as

$$\left\langle \mathrm{vmr}_{instr\,1} - \mathrm{vmr}_{instr\,2} \right\rangle = \left\langle \mathrm{vmr}_{truth_1} - \mathrm{vmr}_{truth_2} \right\rangle + \sigma^2_{instr1} + \sigma^2_{instr2}$$

The quantity

$$\left\langle \mathrm{vmr}_{truth_1} - \mathrm{vmr}_{truth_2} \right\rangle$$

is an estimation of the natural variability within given collocation criteria. To alliviate the notations, we will note $\sigma_{instr\,1} = \sigma_1$, $\mathrm{vmr}_{instr\,1} = \mathrm{vmr}_1$ and $\left\langle \mathrm{vmr}_{truth_1} - \mathrm{vmr}_{truth_2} \right\rangle = natvar_{12}$, similar for other indices.

This holds in the ideal situation. Still the standard deviation of the differences of two measurements does not agree with the estimated error of the differences (ploti showing vertical profile of left and right part of the equatio for ACE and MIPAS, MLS ans MIPAS, ACE and MLS). To evaluate whose error estimate is not correct, we can introduce the calibration, or altitude-dependent scaling factors to adjust the uncertainties of partcipating instruments, so that the equation holds. If the error estimation of the instrument is close to reality, than the corresponding scaling factor is equal to 1. To find these scaling factors, a third instrumenti is introduced in the analysis; for three instruments, it gives rise to the system of three equations for three unknowns c_1, c_2, c_3:

$$\left\langle \mathrm{vmr}_1 - \mathrm{vmr}_2 \right\rangle = c_1\sigma_1^2 + c_2\sigma_2^2 + natvar_{1,2}$$

$$\left\langle \mathrm{vmr}_1 - \mathrm{vmr}_3 \right\rangle = c_1\sigma_1^2 + c_3\sigma_3^2 + natvar_{1,3}$$

$$\left\langle \mathrm{vmr}_2 - \mathrm{vmr}_3 \right\rangle = c_2\sigma_2^2 + c_3\sigma_3^2 + natvar_{2,3}$$

Here the σ_i are the precision estimates of the instruments involved, $natvar_{i,j}$ denominates the discrepancies between two sets of data attributable to natural variability, expressed in terms of variance, and c_i are the unknown calibration factors. The variances on the right-hand side of the equations are the discrepancies between the measurements found in validation studies. These are variances of the differences of collocated measurements, which means that biases do not contribute, and in consequence systematic error estimates are not considered. By solving this linear system of equations for c_1, c_2, and c_3, we will obtain the correction coefficient for the precision estimates of each instrument of the triplet.

In order not to overload the notations, we use the same symbol σ_1 in the first and second equations, despite the fact that those are different because they are calculated on the different subsamples of the record of the instrument 1: σ_1 in the first equation is calculated on the subsample coming from the collocations of the instrument 1 with the instrument 2, while σ_1 in the second equation comes from the collocations of the instrument 1 with instrument 3. It could be avoided if the analysis was performed on the triple collocations, but the test have shown that restricting the analysis on triplets significantly reduce the size of the sample giving rise to statistically non-significant result. In matrix form, we are therefore dealing with the linear system

$$\begin{pmatrix} \sigma_1^2 & \sigma_2^2 & 0 \\ \sigma_1^2 & 0 & \sigma_3^2 \\ 0 & \sigma_2^2 & \sigma_3^2 \end{pmatrix} \quad \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} \left\langle \mathrm{vmr}_1 - \mathrm{vmr}_2 \right\rangle - natvar_{12} \\ \left\langle \mathrm{vmr}_1 - \mathrm{vmr}_3 \right\rangle - natvar_{13} \\ \left\langle \mathrm{vmr}_2 - \mathrm{vmr}_3 \right\rangle - natvar_{23} \end{pmatrix}$$

whose determinant $-2\sigma_1^2\sigma_2^2\sigma_3^2 \neq 0$, hence the system always have a solution.

Pair of instruments	Latitude band	Distance mismatch, km	Time mismatch, hours	Sample size
MIPAS/ACE	[50N,70N]	300-400	4-5	231
MIPAS/MLS	[50N,70N]	200-300	2-3	2851
ACE/MLS	[50N,70N]	500-600	6-7	166

Table 2.1: Latitude bands, distance mismatch, time mismatch, and the size of resulting samples for the triplet MIPAS/ACE/MLS.

Pair of instruments	Latitude band	Distance mismatch, km	Time mismatch, hours	Sample size
MIPAS/ACE	[50N,70N]	300-400	4-5	231
MIPAS/OSIRIS	[50N,70N]	200-300	2-3	3696
ACE/OSIRIS	[50N,70N]	300-400	3-4	38

Table 2.2: Latitude bands, distance mismatch, time mismatch, and the size of resulting samples for the triplet MIPAS/ACE/OSIRIS.

The problem is however that we are looking for strictly positive solutions, because c_i is the square of the scaling factor for uncertainty estimate σ_i, and this does not always exist at all the heights.

The difference of this method from the method in Fioletov et al. (2006) is that they use two instruments and estimates the natural variability values together with the variancen of errors from the both instruments using pnly the values of the variance of measured profiles and the variabce of the differences of measured collocated profiles. While in our case the natural variability values are taken from the model.

2.2 Validation of the method

The total procedure was validated by the following self-consistency test. The calibration method is applied to two triples of data sets, which have at least one data set in common but differ in at least one data set. We demonstrate in ot the example of the triplets MIPAS/ACE/MLS and MIPAS/ACE/OSIRIS. We also show the result for the triplet MIPAS/ACE/GOMOS where no conclusion could have been drawn because the solution of the system for c_1 is everywehre negative, hence can not be a square of the adjustment factor.

The latitude bands and the distance and time mismatches for the sample choice were dictated by the sample size : we stayed in the mid-latitudes and have chosen such distance and time mismatch that correspoding sample is the biggest. The Table1 2.1, 2.2, and 2.3 show the choice made for latitude bands, time and distance mismatches, as well as the obtained maximal sample size for three triplets, participating in the validation exercice.

The Figure 2.1, 2.2 and 2.3 show the obtained adjustment factors for three different triplets the triplet of the instruments.

The dashed line corresponds to the ideal uncertainty estimates. The line to the left of the dashed line means that the uncertainty of a processor are

Pair of instruments	Latitude band	Distance mismatch, km	Time mismatch, hours	Sample size
MIPAS/ACE	[50N,70N]	300-400	4-5	231
MIPAS/GOMOS	[50N,70N]	300-400	1-2	359
ACE/GOMOS	[50N,70N]	500-600	4-5	32

Table 2.3: Latitude bands, distance mismatch, time mismatch, and the size of resulting samples for the triplet MIPAS/ACE/GOMOS.

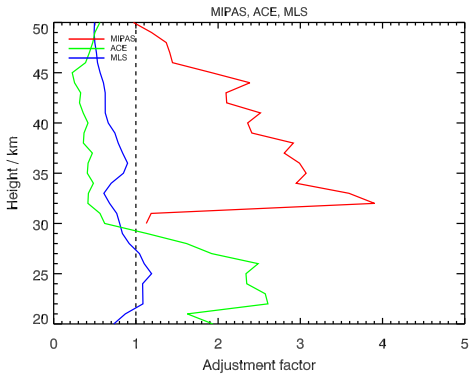


Figure 2.1: Adjustment factors derived from the triplet MIPAS, ACE-FTS and MLS.

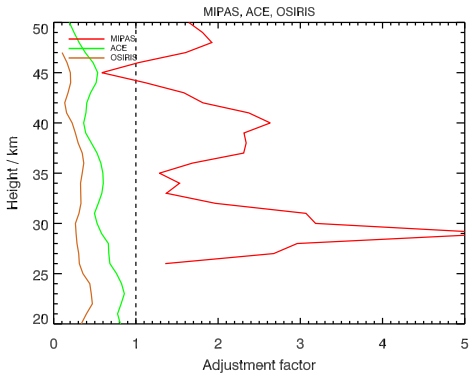


Figure 2.2: Adjustment factors derived from the triplet MIPAS, ACE-FTS and OSIRIS.

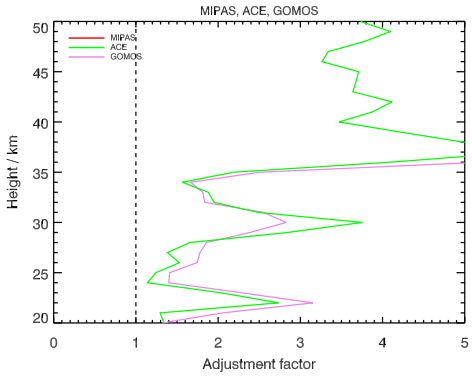


Figure 2.3: Failed attempt to derive the adjustment factors from the triplet MIPAS, ACE-FTS and GOMOS: the solution of the system for c_1^2 are negative at all heights, the factors for the two remaining instruments are not compliant with the analysis of two other triplets..

overestimated and should be scaled by a factor less than 1. The line lying on the right from the dashed line means that the uncertainty of a processor are underestimated and should be scaled by a factor bigger than 1. From both triplets MIPAS/ACE/MLS and MIPAS/ACE/OSIRIS we obtain that ACE uncertainties are overestimated by the mean factor of 2 and MIPAS uncertainties are understimated by the mean factor of 2. Large correction factors for MIPAS do not come as a surprise because for this study measurement noise was the only random error source considered. Other random error components like temperature and pointing uncertainties or uncertainties in the mixing ratios of interfering species were disregarded for this study. It increases the confidence in our method that such issues are discovered. The underestimation (but at a smaller scale) of the error of MIPAS IMK Processor at heights 35-50 km would agree well with the results of (Laeng et al., 2015): we recall here in the Figure 2.4 the third panel of the Figure 19 from(Laeng et al., 2015). The slight overestimation of its error by MLS is also compliant with the result of analysis in (Laeng et al., 2015).

The outlier value in MIPAS adjustment factor at 28 km altitude derived from MIPAS/ACE/OSIRIS triplet comes from particulary big ratio of sample sizes between MIPAS/OSIRIS and ACE/OSIRIS at this height; tests have show that this outlier is dived by 3 if the sample of MIPAS/OSIRIS is diveded by 10. The line for MIPAS adjustment factors stops at 30 km because lower solutions are negative. A general observation, coming from test on different sample sizes and different triplets is that the sample sizes between three pairs participating in the triplet should be of the same order of magnitude, in order to have the statistically significant results and the non-negative solution simulatenously. Thus, the attempt to derive the adjustment factors from the triplet MIPAS/ACE/GOMOS failed because of too smal size of the samples.

It should be mentionned, that the method presents a number of caveats. First, the quantites $\langle \mathrm{vmr}_i - \mathrm{vmr}_j \rangle$ are only sample estimations of the variances, whose the α-level confidence interval depends on the estimated values and size of the sample. This is why we are working with pairwise sample instead of the

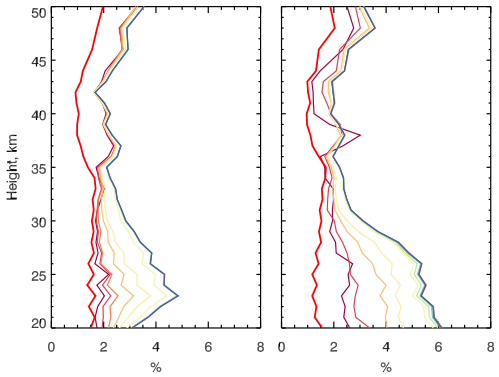


Figure 2.4: The structure function from the error validation analysis for MIPAS KIT Procesor in (Laeng et al., 2015). At the left panel at 35-50 km altitudes, the structure function of MIPAS Processor does not reach the red line of *ex-ante* estimates, which hints toward an underestimation, by a factor approximatively 1.5-1.8 of the error estimation at this height range.

triplets of collocated measurements. Second, the assumption of the independence of σ_i from the *vmr* does not always hold, but this dependence can be quantified (Fioletov et al., 2006).

Despite of this, the calibration constant (adjustment factor) for the MIPAS and ACE datasets, obtained from the application of the method to two triplets, look similar at the height range from 30 to 50 km. This hints toward the viability of this method. Note also that an indirect evidence, that MLS slightly overestimates its uncertainties, is obtained in the analysis of the triplet MIPAS/ACE/MLS; this was already observed in Laeng et al. (2015).

2.3 Conclusion

High-resolution model data are used to estimate typical variabilities of mixing ratios of trace species as a function of spatial and temporal distance. These estimates are then used to explain that part of the differences between observations made with different observing systems that are due to less than perfect collocation of the measurements.

In contrast to the usual way of precision validation where the result more often than not is that one of the two intercompared instruments underestimates its random error budget. Here random error calibration coefficient (adjustment factor) was assigned to each data set. For 3 data sets, 3 pairwise intercomparisons are possible. For each comparison it is postulated that the sum of the random error variances plus the variance due to natural variability was equal the standard deviation of the differences of the data. The latter was calculated from the sample of coincident pairs of measurements. This allowed to set up a system of 3 equations with non-zero determinant. The system was solved to derive the adjustment factors, when it is possible.

The validation of the method was performed on five independent data sets, the method was applied to 3 triplets of data sets, which are different in one data set. The method is validated because for the common data sets the calibration

coefficients resulting from two triplets look similar at the height range where they can be determinated, namely at 30-50 km. The method should be futher validated on the triplets of dense samplers.

2.4 Data availability

The regression coefficients of the parametrisation on latitudes for autumn season and the software for calculating the variability as function of latitudinal gradients are stored in https://publikationen.bibliothek.kit.edu/1000128887

Bibliography

Chabrillat, S., Vigouroux, C., Christophe, Y., Engel, A., Errera, Q., Minganti, D., Monge-Sanz, B. M., Segers, A., and Mahieu, E.: Comparison of mean age of air in five reanalyses using the BASCOE transport model, Atmos. Chem. Phys., 18, 14 715–14 735, https://doi.org/10.5194/acp-18-14715-2018, 2018.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., and et al.: The Community Earth System Model Version 2 (CESM2), JAMES, 12, e2019MS001 916, https://doi.org/10.1029/2019MS001916, 2019.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J., Park, B., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G.and Kinnison, D., Lamarque, J.-F., and et al.: The Chemistry Mechanism in the Community Earth System Model version 2 (CESM2), JAMES, 12, e2019MS001 882, https://doi.org/10.1029/2019MS001882, 2019.

Errera, Q., Daerden, F., Chabrillat, S.and Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D.: 4D-Var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, Atmos. Chem. Phys., 8, 6169–6187, https://doi.org/10.5194/acp-8-6169-2008, 2008.

Errera, Q., Chabrillat, S., Christophe, Y., Debosscher, J., Hubert, D., Lahoz, W., Santee, M. L., Shiotani, M., Skachko, S., von Clarmann, T., and Walker, K.: Technical note: Reanalysis of Aura MLS chemical observations, Atmos. Chem. Phys., 19, 13 647–13 679, https://doi.org/10.5194/acp-19-13647-2019, 2019.

Fioletov, V., Tarasick, D., and Petropavlovskikh, I.: Estimating ozone variability and instrument uncertainties from SBUV(/2), ozonesonde, Umkehr, and SAGE II measurements: Short-term variations, Journal of Geophysical Research Atmospheres., 11, 2993–3002, https://doi.org/10.1029/2005JD006340, 2006.

Fischer, H., Birk, M., Blom, C., Carli, B., Carlotti, M., von Clarmann, T., Delbouille, L., Dudhia, A., Ehhalt, D., Endemann, M., Flaud, J. M., Gessner, R., Kleinert, A., Koopmann, R., Langen, J., López-Puertas, M., Mosner, P., Nett, H., Oelhaf, H., Perron, G., Remedios, J., Ridolfi, M., Stiller, G., and Zander, R.: MIPAS: an instrument for atmospheric and climate research, Atmos. Chem. Phys., 8, 2151–2188, https://doi.org/10.5194/acp-8-2151-2008, 2008.

Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., and et al.: The whole atmosphere community climate model version 6 (WACCM6), Journal of Geophysical Research Atmospheres., 124, 12 380–12 403, https://doi.org/10.1029/2019JD030943, 2019.

Kolmogorov, A. N.: Wiener's spiral and some other interesting curves in Hilbert space, Dokl. Akad. Nauk SSSR, 26, 115–118, 1940.

Kraft, D.: A Software Package for Sequential Quadratic Programming, DFVLR-FB 88-28, 1988.

Laeng, A., Grabowski, U., von Clarmann, T., Stiller, G., Glatthor, N., Höpfner, M., Kellmann, S., Kiefer, M., Linden, A., Lossow, S., Sofieva, V., Petropavlovskikh, I., Hubert, D., Bathgate, T., Bernath, P., Boone, C. D., Clerbaux, C., Coheur, P., Damadeo, R., Degenstein, D., Frith, S., Froideveaux, L., Gille, J., Hoppel, J., McHugh, M., Kasai, Y., Lumpe, J., Rahpoe, N., Toon, G., Sano, T., Suzuki, M., Tamminen, J., Urban, J., Walker, K., Weber, M., and Zawodny, J.: Validation of MIPAS IMK/IAA V5R_O3_224 ozone profiles, Atmos. Meas. Tech., 7, 3971–3987, https://doi.org/10.5194/amt-7-3971-2014, 2014.

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, https://doi.org/10.1016/j.rse.2014.12.013, 2015.

Minganti, D., Chabrillat, S., Christophe, Y., Errera, Q., Abalos, M., Prignon, M., Kinnison, D. E., and Mahieu, E.: Climatological impact of the Brewer–Dobson circulation on the N2O budget in WACCM, a chemical reanalysis and a CTM driven by four dynamical reanalyses, Atmos. Chem. Phys., 20, 12 609–12 631, https://doi.org/10.5194/acp-20-12609-2020, 2020.

Prignon, M., Chabrillat, S., Minganti, D., O'Doherty, S., Servais, C., Stiller, G., Toon, G. C., Vollmer, M. K., and Mahieu, E.: Improved FTIR retrieval strategy for HCFC-22 (CHClF2), comparisons with in situ and satellite datasets with the support of models, and determination of its long-term trend above Jungfraujoch, Atmos. Chem. Phys., 19, 12 309–12 324, https://doi.org/10.5194/acp-19-12309-2019, 2019.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, https://doi.org/10.1029/2002JD002299, 2003.

Sheese, P. E., Walker, K. A., Boone, C. D., Degenstein, D. A., Kolonjari, F., Plummer, D., Kinnison, D. E., Jöckel, P., and von Clarmann, T.: Model estimations of geophysical variability between satellite measurements of ozone profiles, Atmos. Meas. Techn., 14, 1425–1438, https://doi.org/10.5194/amt-14-1425-2021, 2021.

Sofieva, V. F., Tamminen, J., Kyrölä, E., Laeng, A., von Clarmann, T., Dalaudier, F., Hauchecorne, A., Bertaux, J.-L., Barrot, G., Blanot, L., Fussen, D., and Vanhellemont, F.: Validation of GOMOS ozone precision estimates in the stratosphere, Atmos. Meas. Tech., 7, 2147–2158, https://doi.org/10.5194/amt-7-2147-2014, 2014.

Sofieva, V. F., Lee, H. S., Tamminen, J., Lerot, C., Romahn, F., and Loyola, D. G.: A method for random uncertainties validation and probing the natural variability with application to TROPOMI on board Sentinel-5P total ozone measurements, Atmos. Meas. Tech., 14, 2993–3002, https://doi.org/10.5194/amt-14-2993-2021, 2021.

Tilmes, S., Hodzic, A., Emmons, L. K., Mills, M. J., Gettelman, A., Kinnison, D. E., and et al.: Climate forcing and trends of organic aerosols in the Community Earth System Model (CESM2), JAMES, 11, 4323–4351, https://doi.org/10.1029/2019MS001827, 2019.

von Clarmann, T.: Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, Atmos. Chem. Phys., 6, 4311–4320, 2006.

von Clarmann, T., De Clercq, C., Ridolfi, M., Höpfner, M., and Lambert, J.-C.: The horizontal resolution of MIPAS, Atmos. Meas. Techn., 2, 47–54, https://doi.org/10.5194/amt-2-47-2009, 2009.

Yaglom, A.: Correlation Theory of Stationary and Related Random Functions, Vol. II, Springer-Verlag, 1986.