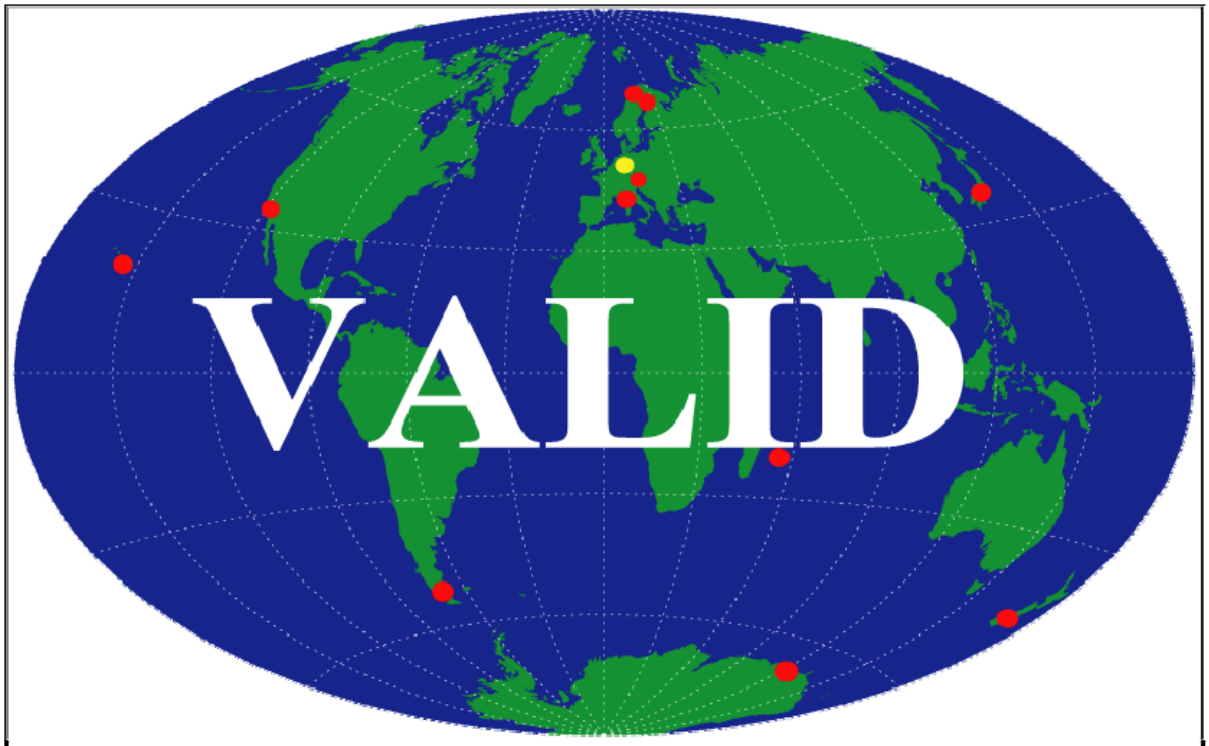



# Satellite validation with lidar

Final report of the VALID-2 project

version November 13th 2013

Authors: Anne van Gijsel, Piet Stammes & Ankie Piters
KNMI, De Bilt, the Netherlands

# 1 Introduction and overview

This report summarises the activities carried out in the VALID-2 (satellite validation with lidar) project. The project has run from 2011 until into 2013. The aim of the project was to validate Envisat ozone and temperature profiles, and to assess the possibility to validate (future) satellite tropospheric $NO_2$ profiles. Besides the "standard' validation work where satellite profiles are compared with ground-based observations, efforts have also been put into assisting the lidar community to homogenise the data processing and reporting (which aids to add consistency to the reference data used for validation purposes), assisting in field campaigns and providing feedback on the GECA website development stages.

This document is organised as follows:

Section 2 gives an overview of the approach taken in all verification/validation studies and complicating factors, shortly describes the reference datasets and gives a summary of the work carried out in the ISSI team on standardisation of the NDACC lidar data, including some more detailed results.

Section 3 introduces the main validation work, providing an overview of the verified/validated ENVISAT datasets, where the main validation results are provided in section 4 for SCIAMACHY, section 5 for MIPAS and section 6 for GOMOS.

Section 7 then presents a comparison of $NO_2$ sonde profiles with $NO_2$ lidar profiles taken during the CINDI and PEGASOS campaigns.

Section 8 finally gives an overview of the documents and presentations prepared during the course of this study.

In the overview table below the ENVISAT sensor versions that have been validated are shown in lavender and the versions that have only been verified in blue.

**Table 1. Validated (lavender) and verified (blue) data versions**

| | | | | |
|---|---|---|---|---|
| SCIAMACHY ozone | v5.01 | v5.02 | v6 (v5.75) | |
| MIPAS ozone + temperature | GMTR 2.1 | GMTR 2.2 | ML2PP v5 | ML2PP v6 |
| GOMOS ozone+HRTP | GBL 1.1 | v5.00 | v6.01 | |

Index

# 2 Validation strategy

Knowledge on the quality of atmospheric profiles is essential for data assimilation, modelling and (trend) studies. To determine a product's strengths and weaknesses, it is usually compared against a reference or standard, or another product with a known quality which is expected to be better than the product under study.

Here we use various quality indicators to describe the differences between the product to be validated or verified, and the chosen standard:

mean: $\mu = \frac{1}{N}\sum_{i=1}^{i=N} x_i$ where N is the total number of observations (x) at a given altitude

standard deviation: $\sigma = \sqrt{\left(\frac{1}{N-1}\sum_{i=1}^{i=N}(x_i-\mu)^2\right)}$

standard error : $s = \frac{\sigma}{\sqrt{N}}$

median: $M = \begin{cases} x_{\frac{n+1}{2}} \longleftarrow rem\left(\frac{n}{2}\right)=1 \\ \left(x_{\frac{n}{2}} + x_{\frac{n}{n}+1}\right) \div 2 \longleftarrow rem\left(\frac{n}{2}\right)=0 \end{cases}$ where $x_1 \le x_2 \le .. \le x_n$ and *rem* is the remainder after division

percentile differences: $x_{n\times p}$ where p = 2.5, 16, 84 & 97.5, $x_1 \le x_2 \le .. \le x_n$ and values are interpolated between bordering indices if n×p is not an exact index. With the 16% and 84% percentiles the 68 interquantile-spread can be calculated. This spread would correspond to that of the range enveloped by the mean ± one standard deviation in the case of a normal distribution. In the following text, it will be referred to as 68-iq spread.

These indicators are computed as a function of altitude to be able to distinguish problematic altitude regions that may differ following other characteristics (e.g. limitations of the reference dataset, regional differences, albedo ranges). These other characteristics are investigated by grouping the data into various classes, provided that the size of the

dataset and subgroups remains sufficiently large to be able to have a statistical significance of any apparent differences.

Various complicating factors can be identified when comparing two datasets:
*Differences in space*: horizontal distance between the two observations as well as the spatial coverages (points versus large pixels).
*Differences in time*: the observed atmospheric states can be different due to non-simultaneity, a measurement's location may not be constant in time (e.g. sonde drifting), the measurement duration can differ substantially (e.g. satellite snapshot versus time-integrated lidar observation of 0.5-2 hours), and discontinued time series may introduce an apparent bias as only a part of the dataset can be compared (abrupt changes in behaviour/trends will be missed).
*Differences in sampled air masses*: even if differences in space and time are relatively small, the profiles to be compared may be very different as a result of sampling different air masses (e.g. near the polar vortex). This may also occur as a result of larger distances or time differences, depending on the observed species.
*Differences in vertical resolutions*: comparing a high vertical resolution profile such as a sonde observation with a low vertical resolution satellite profile will naturally show more deviations. If the satellite retrieval sensitivity is furthermore not limited to the observed altitude layer under comparison but is broader, observed features in the profile may also be displaced and thus indicate deviations.
*Differences in measured quantities*: differences may result as the measurement techniques are not always the same and data may not be reported in the same quantity, thus requiring a conversion (e.g. Dobson units per layer versus number density at a given altitude versus volume mixing ratio at a given pressure) which can introduce uncertainties.
*Differences in used ancillary data*: a conversion such as mentioned above will often require the use of external data, which may not be consistent with the other dataset. Also other external information (e.g. apriori knowledge, cross sections, constants) used by both retrievals can be different.
*Differences in interpretation of variable names/terminology*: it has been noted that when a given data format is prescribed, the interpretation of the name of a required field is not always the same, which will result in different implementations being reported in what appears to be a uniform dataset due to the prescribed data formatting. Also, the metadata coming with a dataset is not always sufficient to fully determine what is being described in the dataset. Using this information in a comparison may thus introduce errors as not the same things are being compared. An example of the first problem (internal variability) has been noticed for reporting the vertical resolution in the case of temperature and ozone lidar data, resulting in the search for a common definition being one of the goals of the ISSI/NDACC standardisation team (more below). An example of

differences in terminology is a definition such as “tropics”, which is not always consistent amongst different comparison studies.
*Differences in the reported uncertainties*: being part of the previous point, it is also common that reported uncertainties do not describe the same thing for different datasets or within a dataset. For sonde data, the information may even be absent. Various groups (e.g. GRUAN, WMO) are working on determination of all uncertainties and homogenisation of the reported data.
*Dependencies (circular validation)*: it is also often seen that for some products or regions, there is no standard reference dataset and products can only be compared on a relative bases (e.g. satellite product B is compared to product A, satellite product C is compared to product B and A is compared to C) or to model output.
*Incomplete/missing/unverified information*: a further complication may come from incomplete knowledge of a dataset. Also data versions are not always mentioned which make a comparison of different studies tricky and complicated.

Unfortunately, it is not feasible to take all of the uncertainties resulting from the above mentioned factors into account. It has to be assumed that all provided information is complete and uniformly formatted/interpreted and that errors from using difference sources of external information are negligible relative to the total measurement uncertainties. Differences resulting from spatial and temporal differences, as well as from sampling different air masses will be studied where possible through different collocation criteria. Differences originating from sampling different scales (horizontal resolution) cannot be assessed here due to the distribution of validation sites, but other international efforts are being taken to study this issue (e.g. The DISCOVER-AQ campaign http://www.nasa.gov/mission_pages/discover-aq/instruments/index.html#.UkloVlPX94E).

## 2.1 Contribution of validation data

Figure 1 shows the locations of the participating partners of the VALID-2 project with blue dots representing the NDACC ozone and/or temperature lidars and the two red dots (shown as one) are the $NO_2$ sonde and the $NO_2$ lidar stations. During the course of the project, the Alomar site and the Tsukuba site were regrettably dropped from the project due to ceasing of their activities. The funding allocated to them in the framework of this project was then used to extend the support to the other sites.

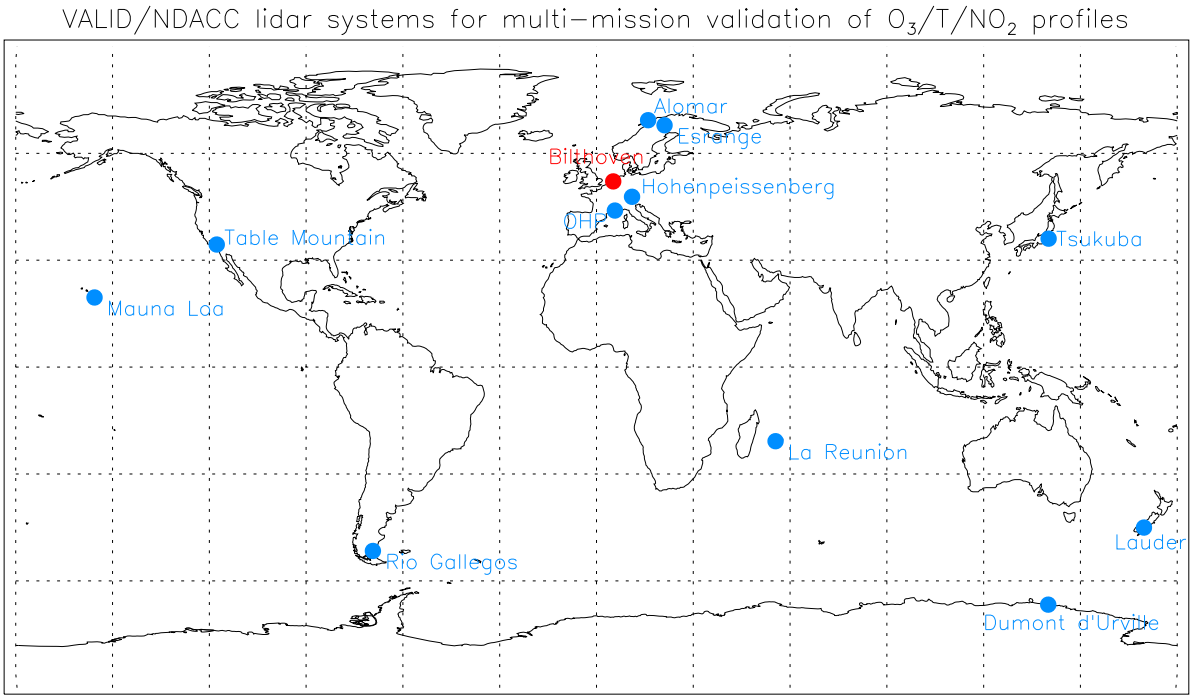



Figure 1. Overview of the planned participating sites in the project. Blue dots indicate temperature/ozone lidars and the red dots represent the $NO_2$ sonde and lidar home sites (Bilthoven/de Bilt).

The lidar team of Lauder (New Zealand) was additionally supported during the two intercomparison campaigns held with the NDACC travelling standard (GSFC mobile lidar) and ozone sonde launches in June/July 2011 and April 2012 respectively.

Below a short summary of the participating instruments is provided.

### 2.1.1 Ozone and temperature lidars

All ozone and temperature lidars, except the one at Esrange, are participating in the NDACC network. Nearly all sites have acquired long time series of measurements and all have undergone intercomparisons with other instruments.

Ozone profiles are derived using the differential absorption technique. The precision of ozone lidar measurements is highest at the lower stratosphere, ranging from about 1-3% up to 30 km, increasing to 2–5% at 40 km, 5-10% at 45 km and to 5–50% at 50 km (Steinbrecht et al, 2009; Nair et al., 2012). We thus limit the upper altitude to 45 km.

Temperature profiles are derived above 25-30 km (where backscattering by aerosol is negligible) using the Rayleigh technique, often combined with retrievals using Raman channels to extend the profiles downwards to approximately 12 km. Above 70 km uncertainties increase due to the tie-on of temperature or pressure and photoncounting uncertainties (Keckhut et al., 2011). For temperature, collocations should when possible be close in time to avoid offsets due to the (semi)diurnal cycle which magnitude depends on location and time of the year (Keckhut et al., 2006). The precision depends strongly on signal strength and can range from 0.3 to 2

K at 40 km, 1-2 K at 50 km, 1-10 K at 70 km (top of reported profiles for some sites) and 3-6 K at 80 km (Funatsu et al., 2011; Steinbrecht et al., 2011, Li et al., 2011). Here we use temperature lidar data in the altitude range 15 to 70 km provided other quality criteria have been fulfilled.
Table 2 shows the sites that have committed to data contribution in the VALID-2 project.

**Table 2. Overview of locations of and provided parameters for ozone/temperature lidars that participated in VALID-2 as data providers.**

| Ground station | Latitude | Longitude | Parameter |
|---|---|---|---|
| Esrange | 67.88 | 21.10 | Temperature |
| Hohenpeissenberg | 47.80 | 11.02 | Ozone, temperature |
| Obs. Haute Provence | 43.94 | 5.71 | Ozone, temperature |
| Table Mountain | 34.40 | -117.70 | Ozone, temperature |
| Mauna Loa | 19.54 | -155.58 | Ozone, temperature |
| La Reunion | -21.80 | 55.50 | Ozone, temperature |
| Lauder | -45.04 | 169.68 | Ozone |
| Rio Gallegos | -51.6 | -69.3 | Ozone |
| Dumont d'Urville | -66.67 | 140.01 | Ozone, temperature |

### 2.1.2 The RIVM $NO_2$ lidar

The RIVM lidar (shown in Figure 2) is operated from a mobile truck and can provide vertical profiles through elevation scanning of various components such as $SO_2$, $NH_3$, Hg, HCL, benzene and $NO_2$ (depending on the used laser dye and resulting emitted wavelengths) with a vertical resolution depending on the observed altitude and integration time (from several meters when close to the surface to two kilometres at four kilometre). The laser beam can be pointed in any direction. The measurement range is typically from 300 m to 2.5 km, sometimes up to 4 km. The lidar uses the DIAL technique as is used by the NDACC lidars for the ozone retrieval. The instrument has participated in various small and large campaigns (e.g. DANDELIONS in 2005/2006, CINDI in 2009, $SO_2$-ship emissions). For $NO_2$ measurements, the laser is alternated between wavelengths of 413.463 nm and 414.112 nm and observations are done for various elevation angles over a period of five minutes or more, combined into one $NO_2$ profile. More information can be found in Brinksma et al. (2008) and Volten et al. (2009).

Figure 2. The RIVM mobile lidar (blue truck).

### 2.1.3 The KNMI $NO_2$-sonde

The $NO_2$-sonde is being developed by the KNMI and is usually flown on small meteorological balloons. It uses the chemiluminescent reaction of $NO_2$ in an aqueous luminol solution, optimised to be specific to $NO_2$ (Sluis et al., 2010). The design has been adapted during the course of development tests and sensitivity as well as stability have substantially improved. The instrument can measure up to three hours, has a vertical resolution of about five meters and measures in the range 1-500 ppbv under all weather conditions. An advantage to the lidar observations is that the sonde reaches higher altitudes, possibly in the future even into the stratosphere. Comparisons have shown that, although the observations are not yet absolute and need to be scaled and corrected for an offset, variations are very well captured giving a high confidence in the observed profile shapes (Sluis et al., 2010; **Figure 3**). The sonde has also been used in a variety of campaigns, such as CINDI (Piters et al., 2012; Vlemmix et al., 2011) and DISCOVER-AQ.

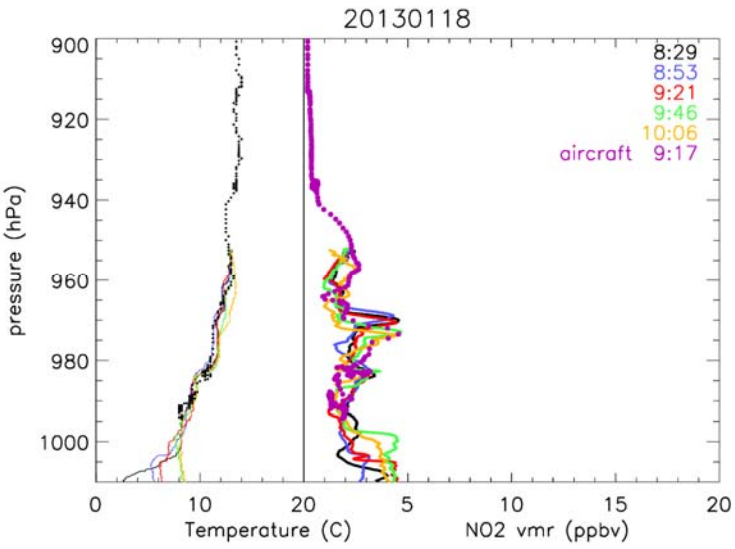


Figure 3. Left: Sonde launch during the PEGASOS campaign in May 2012 in Cabauw. Right: Comparison of various $NO_2$ sonde measurements and nearby aircraft observations of temperature (left side of axis) and $NO_2$ (right side of axis) on January 18th 2013 during the DISCOVER-AQ campaign. Aircraft observations are in black for temperature and in purple for $NO_2$.

Although the $NO_2$ sonde is still experimental (absolute calibration is work in progress), the currently available data are very valuable as the measurements indicate variations on short-time intervals (temporal variation) and over relative short distances (spatial variations). Neglecting such variations will result in large errors for satellite retrievals (when an inappropriate apriori-profile is used) and in model simulations. An example is given in Figure 4.

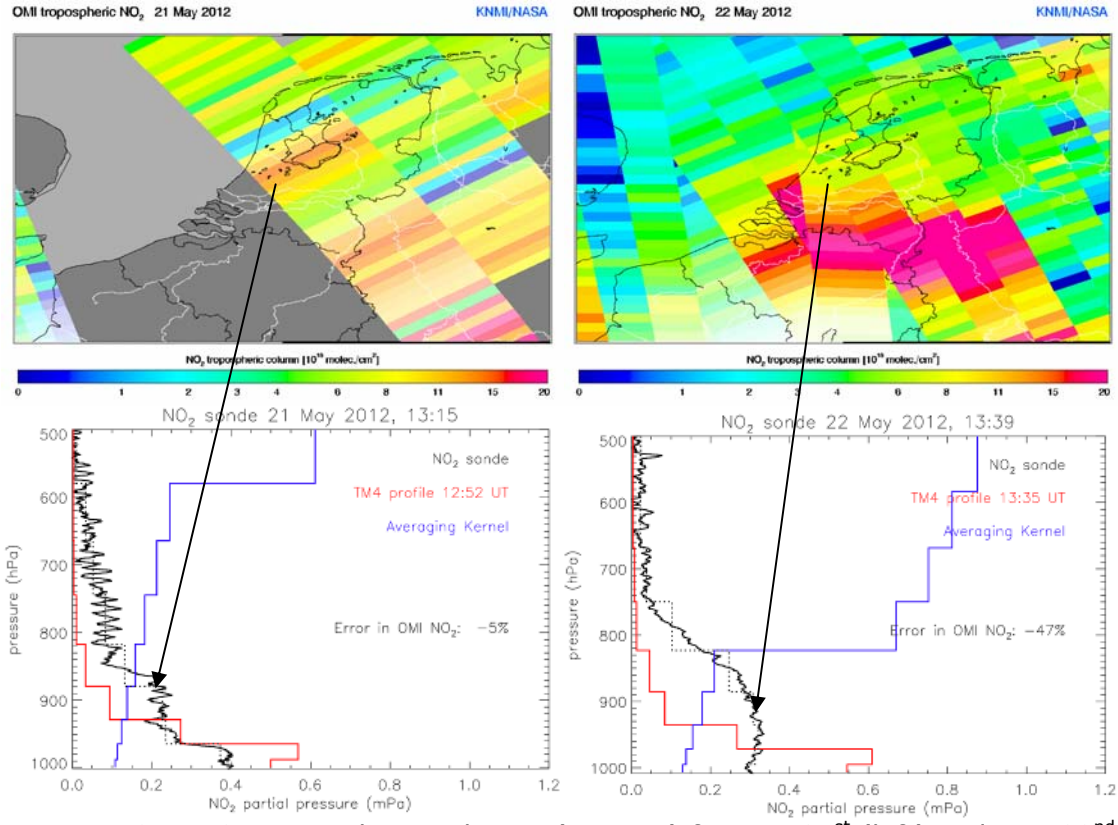



Figure 4. OMI $NO_2$ tropospheric columns (top row) for May 21st (left) and May 22nd (right) in 2012 with $NO_2$ sonde observations (bottom row) for the same dates. The plots on the lower row show, as a function of altitude, the $NO_2$ sonde observations in black as together with the OMI averaging kernel in blue and the TM4 profile that is used as the apriori for the OMI retrieval. The error in the OMI tropospheric $NO_2$ column resulting from assuming a wrong vertical $NO_2$ distribution is 5% for May 21st and 47% for May 22nd, when the differences with the observed $NO_2$ concentrations are much larger.

## 2.2 Data quality and data submission

In the framework of this project, 1386 lidar ozone and temperature data files have been submitted to the calibration/validation database NADIR after checking for erroneous file content and data quality filtering. Data submission is to be completed by a couple of project partners. In comparison to previous years (EQUAL and VALID projects) we can clearly see a reduction in the amount of measurements. This is a regrettable

development as observation sites and scientists are getting less and less (financial) support, which will clearly hinder trend analysis and linking different satellite missions.
$NO_2$ data have not been stored in the NADIR database. $NO_2$ sonde data are not yet absolutely calibrated and it was feared that the data might be used without taking this into consideration as users often do not read the metadata. $NO_2$ lidar data have been reprocessed for the CINDI-campaign for comparison with the $NO_2$ sonde and measurements from the PEGASOS campaign are limited due to a problem with the wavelength calibrator. Currently, there are no ESA satellite $NO_2$ profile data reaching the covered altitudes. A comparison of the sonde and lidar observations is given in section 7.

## 2.3 Data standardisation

As indicated in the section on validation strategy, it is important that the data used for validation is internally consistent in terms of interpretation and reporting. To work towards a higher lidar data quality for validation purposes, as part of this project, we have participated in an ISSI international team, lead by the NDACC lidar working group co-chair Dr. Thierry Leblanc (NASA/JPL/Caltech) on standardisation of the NDACC ozone and temperature lidar data in terms of reported vertical resolution, assessment and propagation of uncertainties, usage of common constants and cross sections and studying the effect of the choice for a certain gravity model on temperature retrievals. The final meeting will take place in September 2013, whereas the team's recommendations will be presented during the bi-annual NDACC lidar working group meeting in November 2013. The outcomes will be presented in an ISSI report and are foreseen to be published as well in a set of journal publications. Where possible, lidar PIs will be assisted in implementing the recommendations. For instance, a plug-in for the computation of vertical resolution for two definitions has been prepared in various languages to facilitate actual implementation ensuring simultaneously that it is correctly implemented. When all recommendations are followed, the NDACC lidar datasets will be more consistent and better documented, and thus of a higher quality for validation and research purposes.
Some information can be found online at http://www.issibern.ch/teams/ndacc/.

### 2.3.1 Vertical resolution

Two definitions have been chosen to be included on a standard base in the NDACC lidar data files. The first is based on the cut-off frequency of a digital filter "dz_cutoff". The second is based on the full-width-at-half-maximum (FWHM) of a response to a perturbation of the impulse-type "dz_fwhm". The two definitions have been put into software plug-ins that will compute the vertical resolution in the PIs retrieval software. The plug-ins are available for Matlab, IDL, fortran and python, which should cover the majority of the PIs. The plug-ins were tested by various team

participants for a number of sites using simulated lidar data to allow checking the correctness of the output (both the vertical resolution reported and the retrieved ozone/temperature profiles). One example is provided in Figure 5 for the digital filter cut-off frequency for simulated lidar signals for the temperature retrieval for Lauder (New Zealand). More examples, also for the vertical resolution tests based on the FWHM, can be found online including additional sites: London (Canada) Mauna Loa (Hawaii), OHP (France) and La Reunion island at http://www.issibern.ch/teams/ndacc/NDACC_Tools_Vertical_Resolution.htm. The tools were found to work correctly and PIs outside the ISSI team will be instructed during the NDACC lidar working group to be held in November 2013 on how to implement the plug-ins.

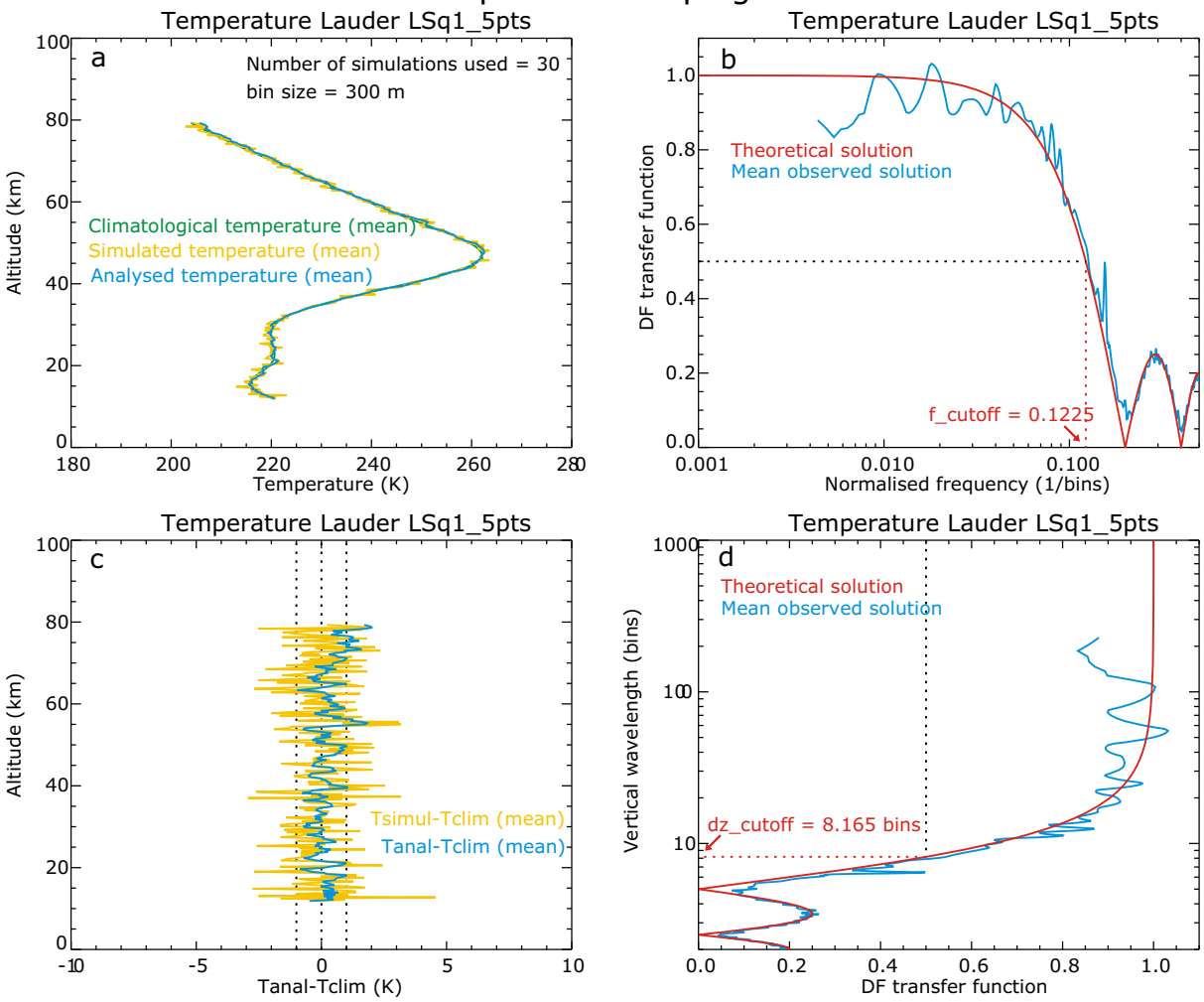



Figure 5. Example of plug-in testing for temperature profiles in Lauder for the digital filter cut-off frequency definition. The test includes a first-order least squares smoothing with 5 altitude bins and shows the mean results for 30 simulations. a) climatological temperature profile (green) and simulated temperature signal (yellow) for Lauder with retrieved profile (blue); b) Digital filter transfer function versus normalised frequency showing the theoretical (red) and observed solution (blue) with the dotted line indicating the cut-off frequency at 0.5; c) observed relative differences with respect to the climatological profile for the simulations (yellow) and the retrieval (blue); d) Vertical wavelength versus transfer function, showing the theoretical (red) and observed solution (blue) with the dotted line indicating the computed vertical resolution (dz_cutoff, resulting in 8.165 altitude bins) corresponding to the cut-off frequency at 0.5.

### 2.3.2 Uncertainties

In the second phase, work has been carried out on identifying all sources of uncertainties in the ozone and temperature retrievals. This has led to the preparation of a long document describing the sources introducing uncertainties, how to quantify these uncertainties or how to provide an estimate of such uncertainty. When possible, a procedure is prescribed and if alternative solutions are implemented, those are penalised. It is intended to assist the PIs in the propagation of the uncertainties throughout the retrieval by means of a toolbox as developed for the computation of the vertical resolution. Although the propagation procedure has been documented now, the uncertainty-propagation toolbox is still to be developed as it is a very complex task given the large variety of procedures in use. A possible alternative would be to go towards a common processor as foreseen for GRUAN and in use with the Earlinet and FTIR communities. Also, where possible, effects of neglecting corrections or of common not fully correct assumptions have been quantified in terms of deviations from the ozone and temperature profiles.

Additionally, a group of frequently used constants were identified. Following an inventory of the used values, it turned out that these constants are not 'constant' as a range of values are in use. The ISSI team has considered the values collected by the CODATA task group on fundamental constants from the BIPM (http://www.bipm.org/extra/codata/) to be the most consistent and recommends to use these values (http://physics.nist.gov/cuu/Constants/Table/allascii.txt) for the constants. To eliminate having to propagate the uncertainty on the value of a constant, when possible, digits are dropped to the digit where the uncertainty is no longer affecting. For example, for the molar gas constant R, the reference value equals 8.3144621 J/mol/K with an uncertainty of 0.0000075 J/mol/K, which leads to the exact value of 8.3145 J/mol/K.

Finally, a study was done on how gravity is being considered in the retrieval software. Also quite a lot of variety was found here, ranging from considering it to be a constant value, to a higher-order degree function depending on altitude and latitude, see Figure 6 and Figure 7. The team has done calculations on the effect of how gravity is treated on temperature profile calculations in a simulation study. Differences up to 6 K in the retrieved temperature profile may theoretically occur if the gravity is assumed to be constant (not latitude- nor altitude-dependent, see Figure 8). The team recommends to use a two-dimensional expression of gravity as a function of latitude and height following the commonly used standard WGS 84 (NIMA-WGS, 1984).

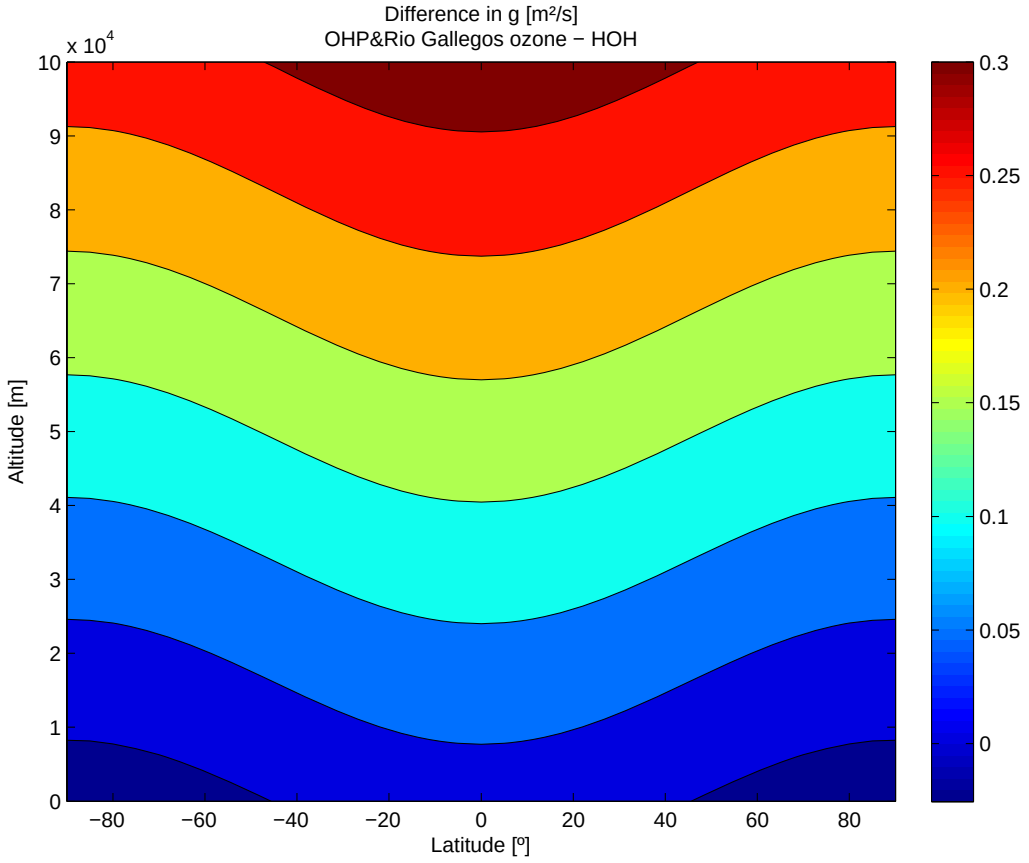



Figure 6. Difference in used gravity following the use of different gravity models. Ozone retrievals for OHP (44°) and Rio Gallegos (-52°) use a gravity that is constant with altitude and latitude, whereas Hohenpeissenberg includes a third-order dependence on altitude as well as a latitudinal dependence. The Hohenpeissenberg model gives very similar results to the WGS-84 model.

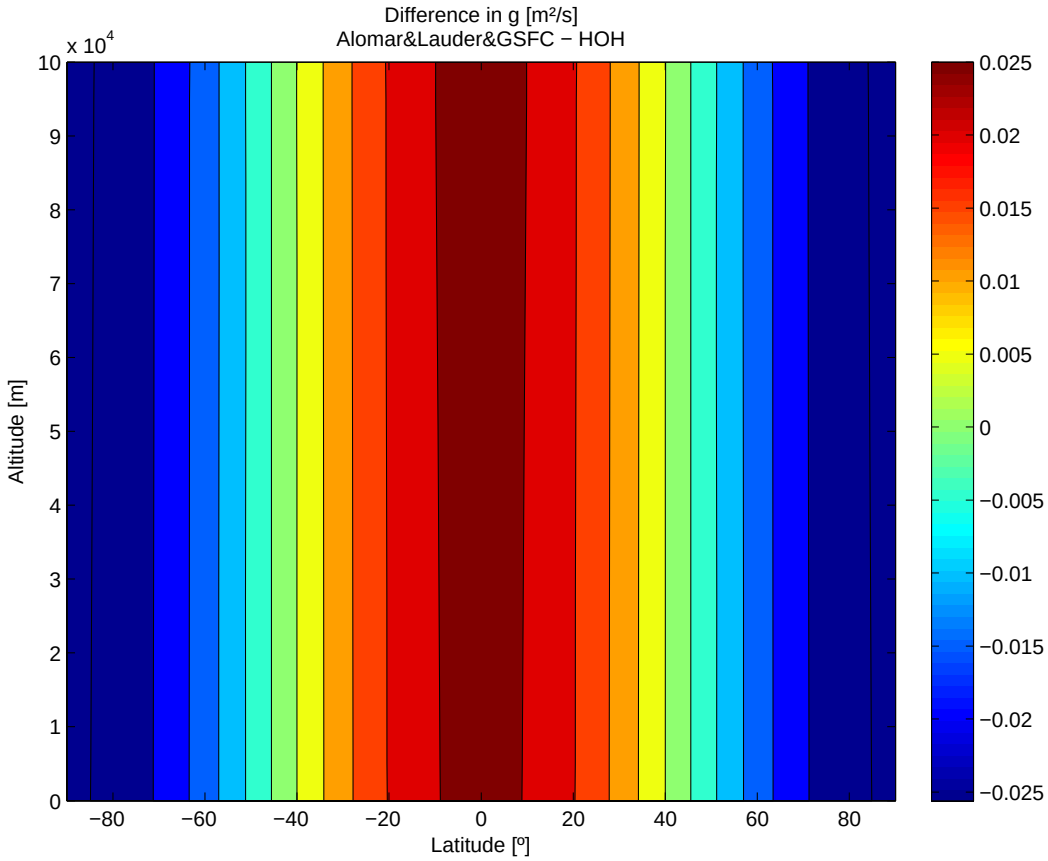



Figure 7. Difference in gravity as a function of altitude and latitude between the implemented gravity computation for Alomar (69°)+Lauder (-45°)+GSFC mobile lidars considering only an second-order altitude-dependence and the Hohenpeissenberg retrievals considering a third-order altitude dependence and a latitude dependence. The Hohenpeissenberg model gives very similar results to the WGS-84 model.

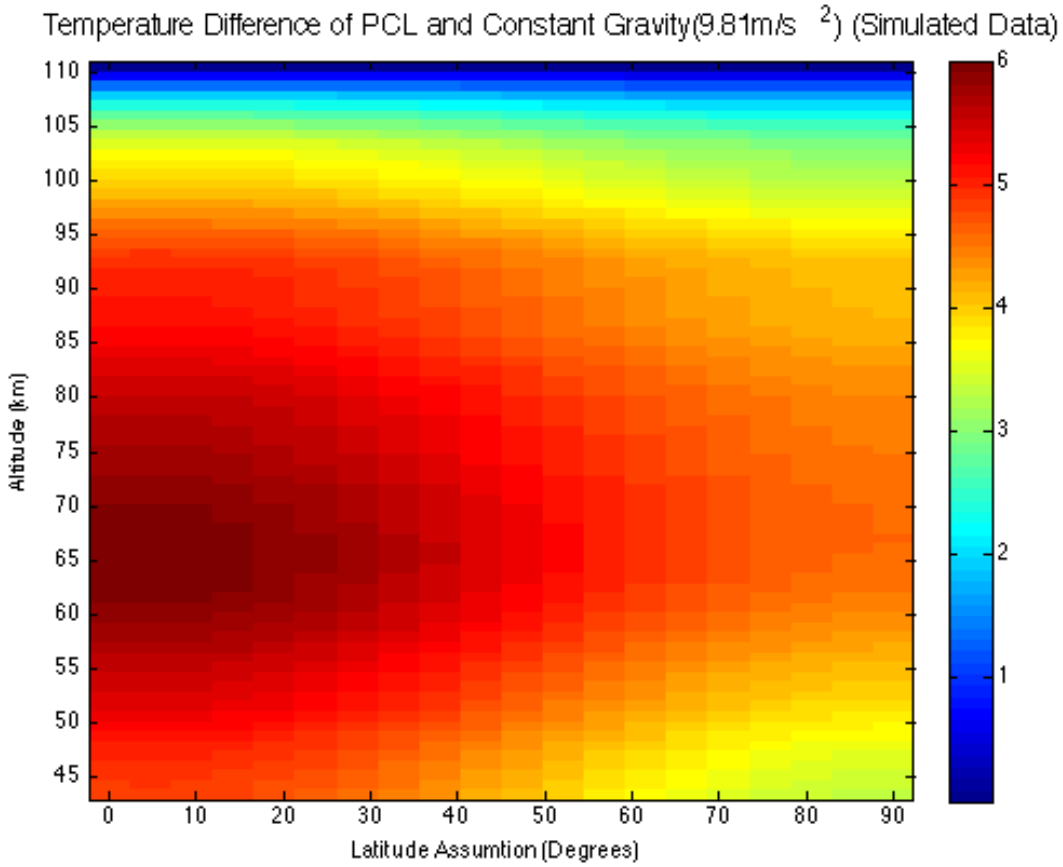



Figure 8. Difference in retrieval temperature (using simulated lidar data) as a function of latitude and altitude when considering gravity to be a constant rather than dependent on altitude and latitude. The reference here is the PCL gravity model, which gives very similar results to the WGS-84 model that is chosen as the NDACC LWG reference.

# 3 Validation and verification results

During the VALID-2 project, comparisons were carried out for data from the three atmospheric profilers on-board ENVISAT: SCIAMACHY, MIPAS and GOMOS. Results and found issues have been discussed with the quality working groups, ESA and/or scientific institutes. In the following chapters, we will summarise the results for the three ENVISAT sensors.
The default collocation criteria used in these studies are as follows:

- for ozone: maximum difference of 20 hours (5 hours above 50 km) and 800 km
- for temperature: maximum difference of 5 hours and 300 km.

Variations of these criteria are also tested, as well as combinations with other requirements (e.g. maximum reported error, illumination conditions, maximum difference in equivalent latitude), which will be reported when these have been applied.

# 4 SCIAMACHY

In the VALID-2 project SCIAMACHY operational level 2 data versions 5.01 and 5.02 ozone profiles have been validated and a verification of version 6 ozone profiles was carried out. Additionally, a set of collocating orbit data for future reprocessings has been compiled for the full time series which has been sent to BIRA to integrate with collocations with other types of instruments. Finally, a study was initiated to attempt to identify factors resulting in similar (good or bad performance) comparisons using clustering by self-organising maps of SCIAMACHY version 5.02 ozone data in comparison to lidar profiles following the technique described in Zurita Milla et al. (2013). This was not successful thus far. A summary of the validation results for version 5.02 is presented below.

SCIAMACHY version 5.02 ozone profiles have been compared with lidar and sonde profiles. All data have been interpolated to a common altitude grid. Lidar and SCIAMACHY data have been filtered to exclude data with an estimated error greater than 30%. The analysis has been done using the validation data 'as is' as well as after convolution with the SCIAMACHY averaging kernels. In the latter case, a value for a given altitude was only considered valid if contributions from altitudes where no validation data were available would be less than 5%.
Analyses have looked at differences between latitude regions, scan angle of the observation, season and year of acquisition, and the cloud flag assigned to the retrieval.
Best agreement with the lidar data is obtained in the polar regions (Figure 9). Overestimation of the ozone concentrations is seen in the mid-latitudes and tropics. The median bias and 68% interquantile spread per altitude bin of 5 km are further summarised in Table 3.

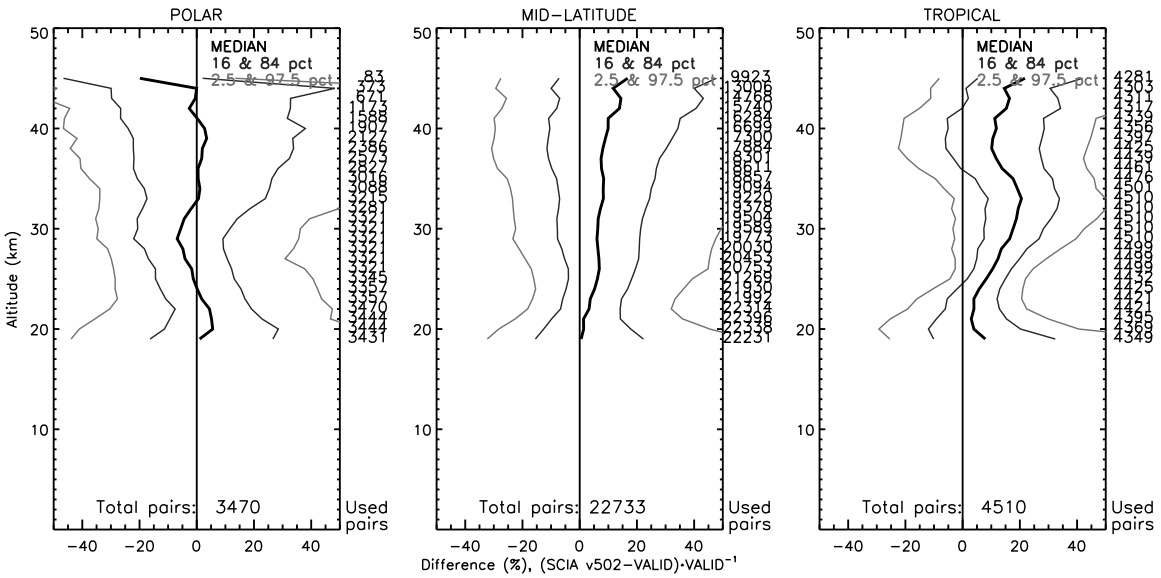



Figure 9. Validation results for SCIAMACHY version 5.02 ozone profiles in comparison to lidar for the polar regions (left), mid-latitudes (middle) and tropics (right panel). Shown are the 2.5, 16, 50 (=median), 84 and 97.5 percentiles of the relative differences as a function of altitude. Along the right axes the number of collocations is listed for the corresponding altitude and the total number of collocations for a latitude group is given at the bottom of each panel.

**Table 3. Validation results for SCIAMACHY version 5.02 ozone profiles in comparison to lidar for three latitude groups. Shown are the median and 68% interquantile spread at the middle of the altitude bin (e.g. 22.5 km for the class 20-25 km). Given numbers are in percent.**

| *Regular* | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | +1 | 43 | +1 | 38 | +8 | 43 |
| 20-25 km | +3 | 27 | +3 | 20 | +4 | 18 |
| 25-30 km | -5 | 27 | +7 | 26 | +13 | 16 |
| 30-35 km | -1 | 39 | +7 | 30 | +20 | 25 |
| 35-40 km | +2 | 54 | +7 | 40 | +11 | 32 |
| 40-45 km | -3 | 62 | +15 | 49 | +16 | 32 |

The differences introduced by the scan angle dependent degradation are most pronounced in the polar region where they can be as large as 15% (median bias). In the mid-latitudes and tropics, the difference observed between different scan angles is within 5% and mostly less (Figure 10). No consistent trend for individual lidar sites was observed when examining results per year and for two altitudes (30 and 40 km).

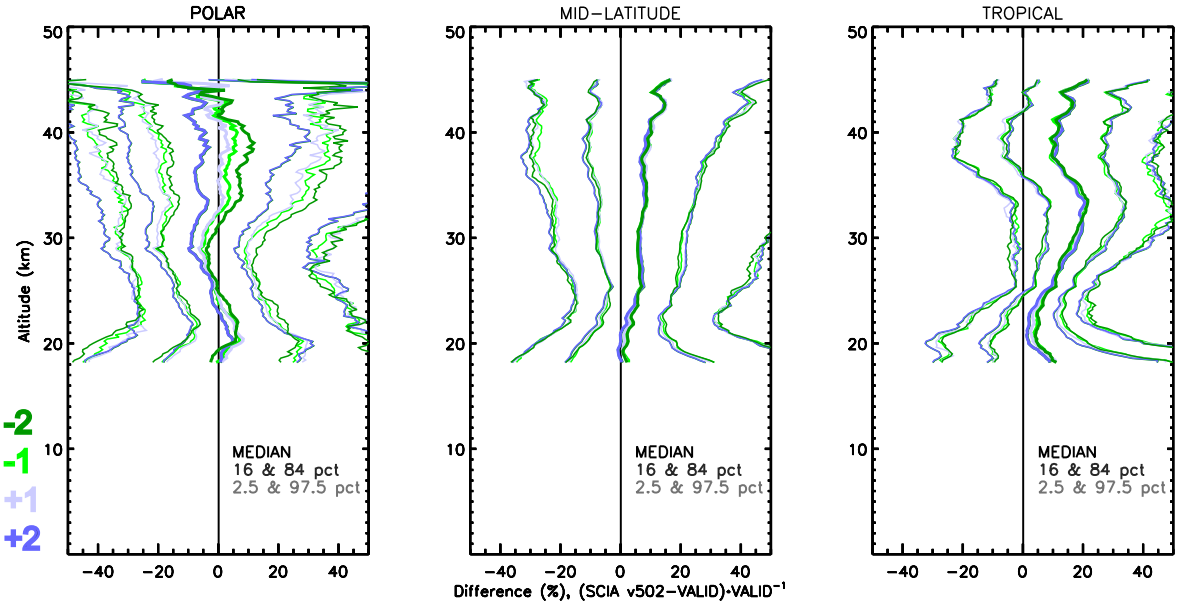



Figure 10. Validation results for SCIAMACHY version 5.02 versus lidar grouped by latitude region. Colours indicate the viewing direction (scan angle): from dark green (-2, left) to blue (+2, right).

Also differences between seasons were found to be largest in the polar regions. The validation results for the autumn and spring seasons are very similar, but the NH summer period (June-August) stands out. One cause could be the reduced number of validation data due to the short nights limiting the lidar observations. In the mid-latitudes a small difference between the summer and winter seasons increasing with altitude for the upper altitudes is seen, possibly related to the large contribution of the apriori information to the SCIAMACHY profiles. In the tropics, the December-February season appears to be different from the rest of the year, which needs further investigation.

# 5 MIPAS

Of the scientific products, MIPAS GMTR 2.1 and 2.2 (ozone and temperature) have been validated. These data are generated by the Bologna team. Of the ESA level two operational products, MIPAS ML2PP v5 & v6 ozone and temperature profiles have been validated. An optimised selection of orbits for delta validation purposes was created in collaboration with the Multi-TASTE team. Shown below are results for ML2PP version 6 and for GMTR version 2.2.

Figure 11 presents the comparison of MIPAS ML2PP version 6 ozone profiles with lidar for day-time (green) and night-time (black) observations. There is obviously an increase of the day-time to night-time differences with increasing altitude, except in the tropics where the reversed behaviour is found. The bias with respect to the lidar observations is latitude-dependent (ML2PP ozone concentrations relatively increase with distance from the poles). More quantitative information is given in Table 4.

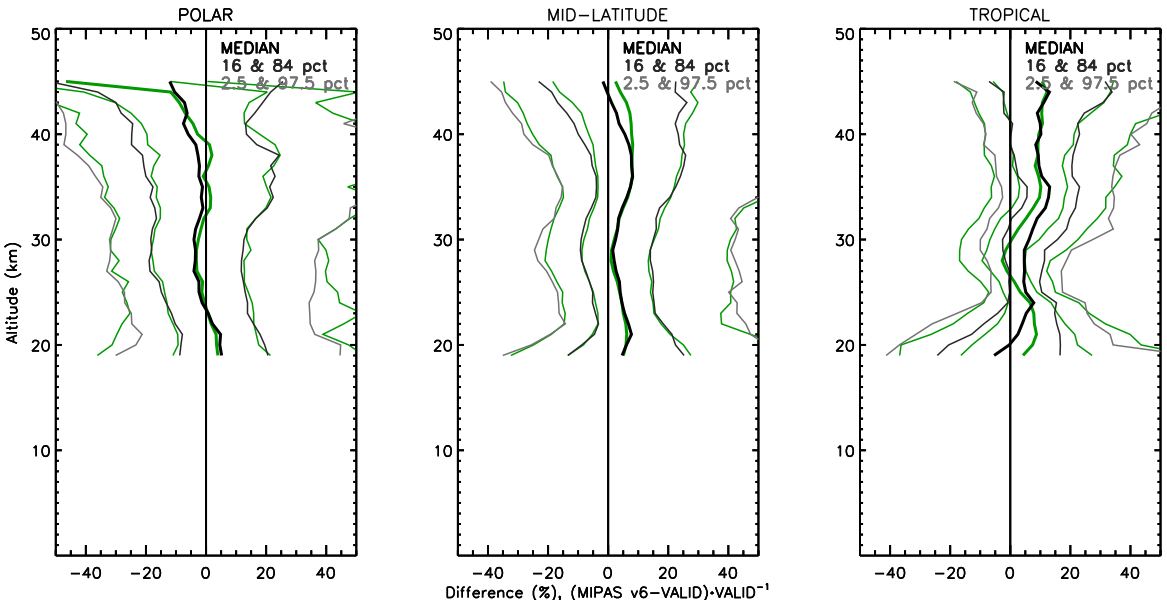



Figure 11. Validation results for ML2PP version 6 ozone profiles in comparison to lidar for the different latitude regions under day time (green) and night time (black) conditions.

**Table 4. Comparison of MIPAS ML2PP version 6 ozone profiles and lidar ozone profiles. Differences are relative to the lidar (MIPAS-lidar/lidar). MIPAS day time observations are listed at the top, whereas night time observations are given in the second half.**

| *Day time* | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | +4 | 33 | +5 | 41 | +4 | 44 |
| 20-25 km | +1 | 28 | +5 | 19 | +8 | 18 |
| 25-30 km | -3 | 30 | +1 | 22 | -3 | 14 |
| 30-35 km | +1 | 35 | +5 | 24 | +8 | 17 |
| 35-40 km | +2 | 40 | +8 | 29 | +9 | 17 |
| 40-45 km | -9 | 45 | +6 | 42 | +11 | 33 |
| | | | | | | |

| *Night time* | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
|---|---|---|---|---|---|---|
| <20 km | +5 | 29 | +5 | 39 | -5 | 41 |
| 20-25 km | +1 | 35 | +6 | 20 | +5 | 19 |
| 25-30 km | -4 | 31 | +2 | 23 | +5 | 13 |
| 30-35 km | -1 | 37 | +5 | 23 | +12 | 18 |
| 35-40 km | -3 | 46 | +8 | 32 | +9 | 20 |
| 40-45 km | -7 | 49 | +1 | 43 | +11 | 33 |

Figure 12 illustrates a comparison between MIPAS GMTR version 2.2 and sonde+lidar+microwave radiometer ozone profiles. The agreement between the GMTR 2.2 data and the validation data is best in the mid-latitude regions. Between 15 and 48 km, agreement in these regions is within -4% to +8% and mostly better (within ±3%). In the polar regions, the profiles show a bias frequently shifting of sign with altitude, but stays within -14% to +7% in the altitude region 12 to 60 km. Above and below stronger deviations are found. The strong positive bias around 10 km is seen on both hemispheres.

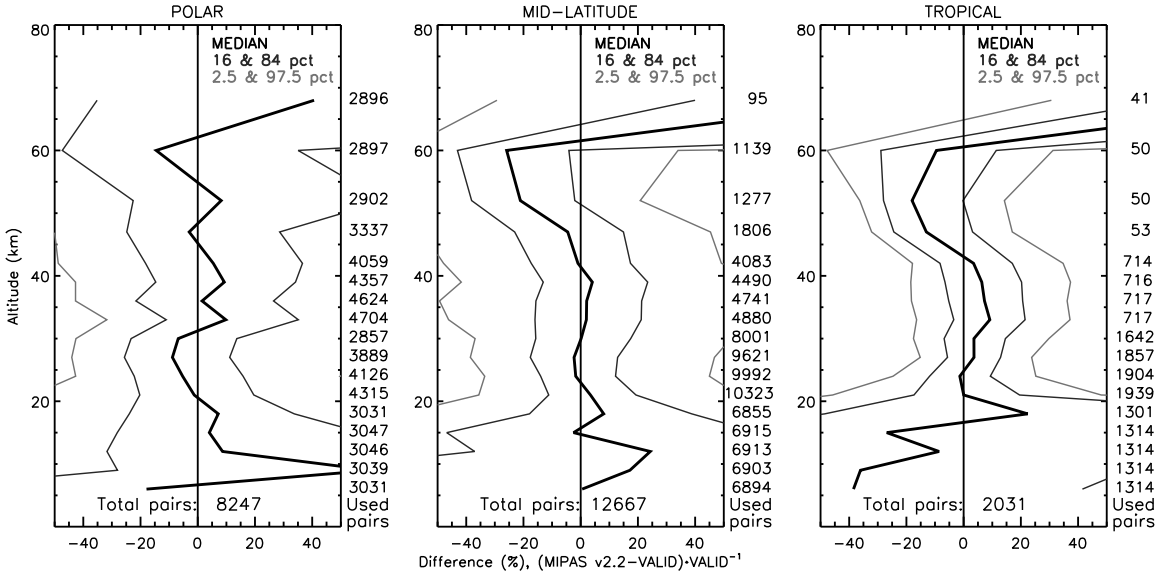



Figure 12. Validation results for GMTR version 2.2 ozone profiles in comparison to sonde, lidar and microwave radiometer data.

For ML2PP version 6 temperature profiles, the validation results in comparison to lidar as shown in Figure 13 and the comparison to sonde in Figure 14. Especially in the lowest 5 km where there is an overlap between the sonde and lidar comparisons (15 to 20 km) a different sign of the bias between these two comparisons is seen, although the differences are not significant. For the higher altitude ranges (comparison with lidar), the bias seems to be trending with altitude in the mid-latitudes and in the tropics where this behaviour is strongest. A large amount of too cold MIPAS temperatures (relative to the lidar) is seen in all regions. In comparison to the sonde, MIPAS temperatures are slightly too cold below 30 km with a strong cold bias at the bottom of the profiles.
This cold bias is also seen for the GMTR version 2.2 temperature profiles (Figure 15). The agreement with the validation data higher up is similar to

that of ML2PP version 6. Part of the oscillations seen may come from the required interpolation.

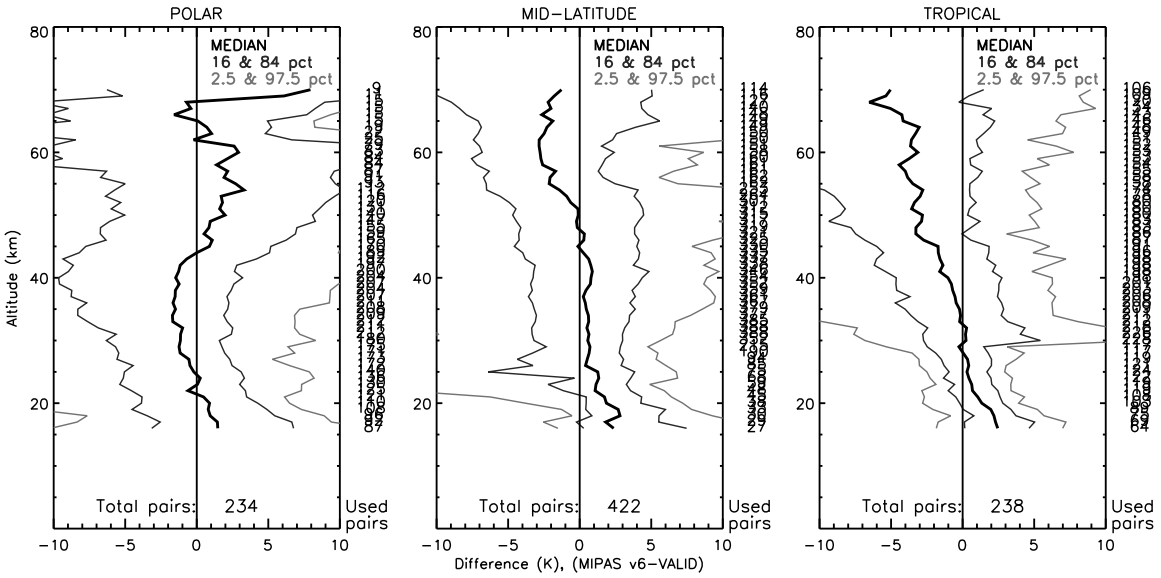



Figure 13. Validation results for ML2PP version 6 temperature profiles in comparison to lidar. Shown are absolute differences for MIPAS night time observations since very few collocations are found for day time as the lidars measure at night.

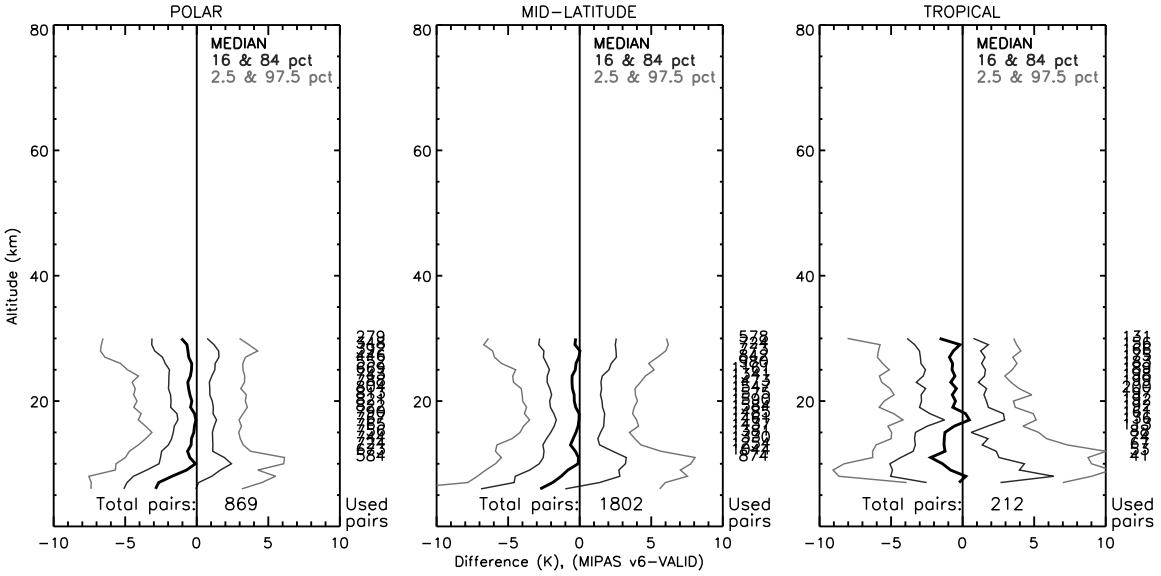



Figure 14. Validation results for ML2PP version 6 temperature profiles in comparison to sonde. Shown are results for MIPAS day time observations.

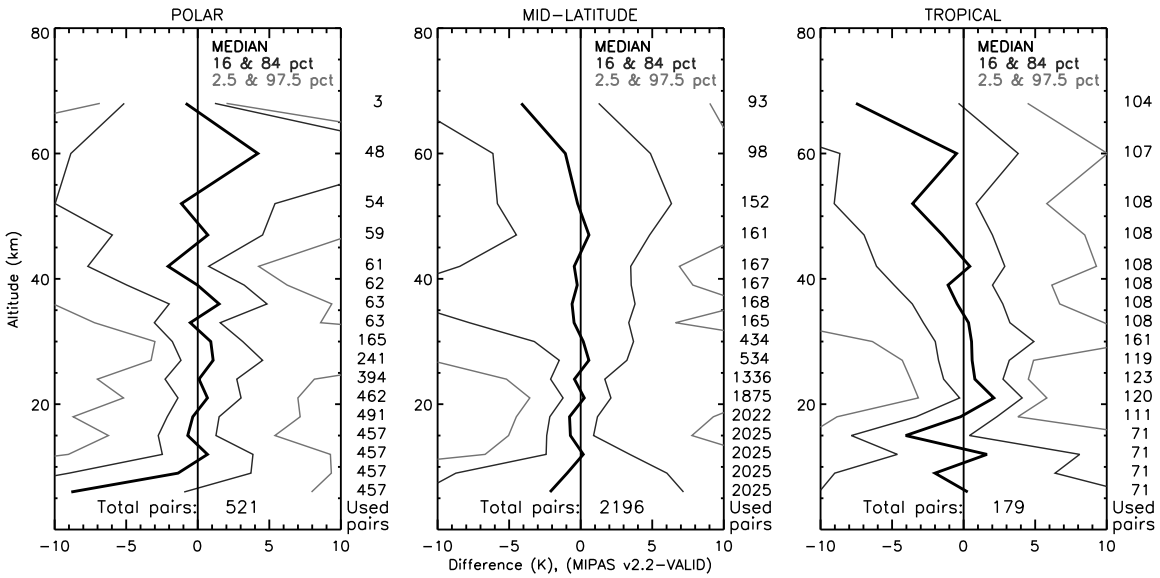



Figure 15. Validation results for GMTR version 2.2 temperature profiles in comparison to sonde and lidar data for the three latitude classes.

# 6 GOMOS

In this project we have validated the ozone and high resolution temperature profiles (HRTP) of the ESA operational level two GOMOS products for v5.00 and v6.01. In addition, comparisons were done for individual years to study stability (version 5.00 and 6.01). Furthermore, also a study was done to check the effect on the data quality when filtering out cool and weak stars. Also, the GOMOS bright limb (GBL) ozone profiles provided by the Finnish Meteorological Institute (FMI) were validated for version 1.1. Finally, a study was done to attempt to verify if there are differences in the data quality of the operational product before and after September 11$^{th}$ 2002. Results from the operational product versions 5.00 and 6.01 and from GBL version 1.1 will be shown here.

Surprisingly, with the upgrade of the operational product from version 5.00 to 6.01, the agreement with lidar had worsened, with version 6.01 presenting lower ozone concentrations, especially for bright and twilight conditions, and to a smaller extent for dark limb conditions and also more pronounced for cool stars (Figure 16). The number of outliers did reduce, in particular below 25 km with dark limb conditions.

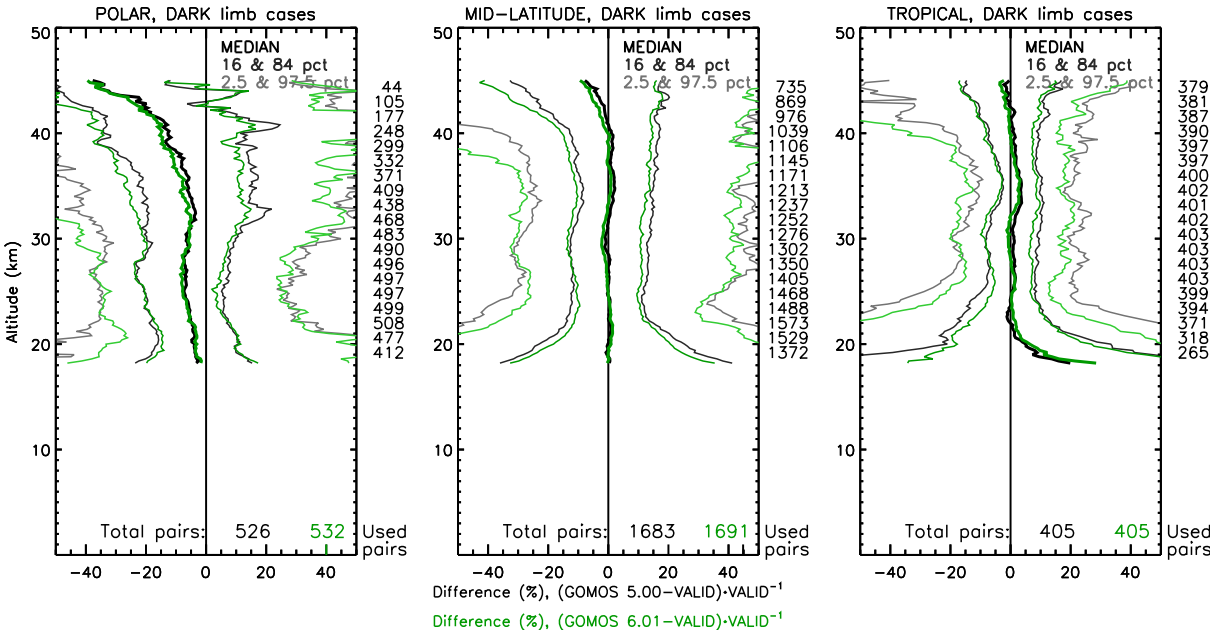



Figure 16. Validation results for GOMOS v5.00 (black) and v6.01 (green) ozone profiles in comparison to lidar for the three latitude regions. Data selection is based upon the solar zenith angle.

For version 6.01, the best agreement with lidar was found for GOMOS observations using weak, hot stars and in the mid-latitudes up to 40 km and in the tropics above the ozone maximum. Differences between using the criterion that the solar zenith angle must be ≥108° or selecting data flagged "dark" were within 1% for the mid-latitudes and tropics (except below 20 km) whereas allowing the use of non-dark-flagged data substantially increases the quantity of useful data (Table 5).

**Table 5. Relative differences between GOMOS version 6.01 ozone profiles and lidar data for three main latitude zones. Given are the median difference and the 68% interquantile spread at a representative altitude (e.g. 22.5 km for the class 20 to 25 km). Top part corresponds to GOMOS data with a solar zenith angle ≥ 108° and bottom part to GOMOS data flagged 'dark'.**

| *Sza ≥ 108°* | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | -3 | 29 | 0 | 51 | +11 | 72 |
| 20-25 km | -5 | 24 | +1 | 23 | +1 | 23 |
| 25-30 km | -8 | 30 | -1 | 23 | 0 | 14 |
| 30-35 km | -6 | 36 | 0 | 24 | +1 | 15 |
| 35-40 km | -11 | 41 | 0 | 27 | +1 | 12 |
| 40-45 km | -21 | 51 | -3 | 44 | -2 | 20 |
| *Flag = dark* | | | | | | |
| <20 km | 14 | 71 | +3 | 51 | +15 | 61 |
| 20-25 km | -2 | 120 | +2 | 22 | +2 | 21 |
| 25-30 km | -9 | 30 | 0 | 26 | +1 | 14 |
| 30-35 km | -12 | 31 | +1 | 25 | +2 | 15 |
| 35-40 km | -10 | 56 | +1 | 29 | +2 | 11 |
| 40-45 km | n/a | n/a | -3 | 43 | -2 | 19 |

In a study comparing the effects of star characteristics, an increased presence of outliers below 25 km was found for weak and cold stars. Mean and standard deviation improved when leaving out observations using weak and cold stars, but the median relative difference showed only a minor improvement. Sofieva et al. (2012) found a strong negative bias for straylight occultations of dim and cool stars in the UTLS and troposphere.

Figure 17 shows a comparison of GOMOS version 6.00 and lidar ozone profiles for different solar zenith angle ranges corresponding to the GOMOS observation. It is clear that in bright limb observations, the retrieved ozone profiles cannot be used. Figure 18 shows the comparison for the scientific retrieval using GOMOS bright limb data, which results in profiles much closer to the lidar data. An overestimation is seen up to 25 km, where above there is an underestimation of the ozone concentration relative to the lidar.

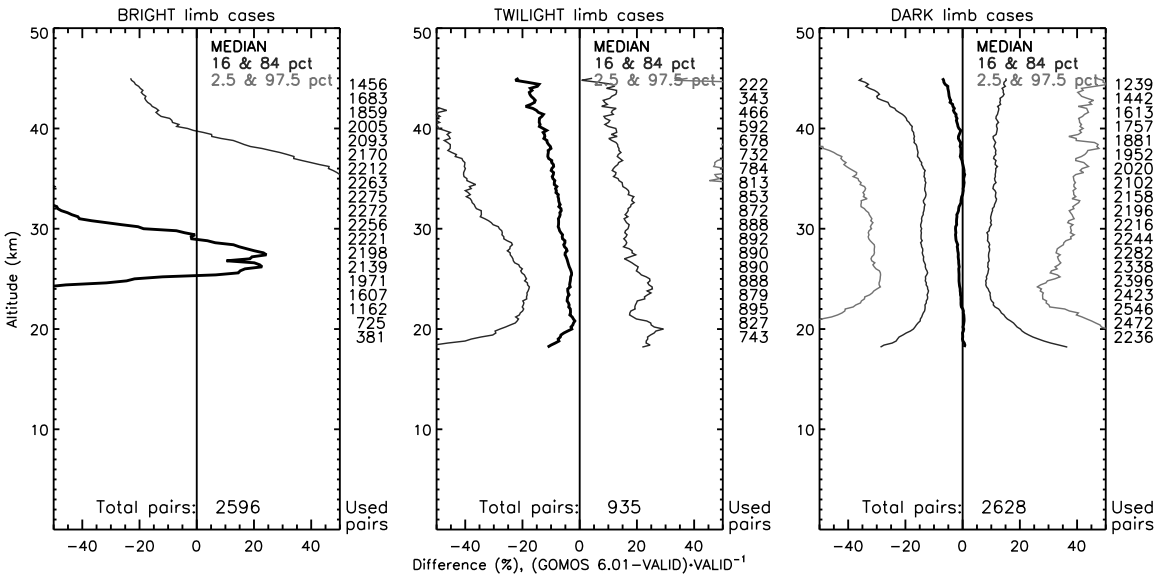



Figure 17. Validation results for the operational GOMOS version 6.01 ozone profiles in comparison to lidar. Left panel: bright limb (solar zenith angle(sza) = [0°,90°]); middle panel: twilight limb (sza = <90°, 108°>); right panel: dark limb (sza = [108°, 180°]).

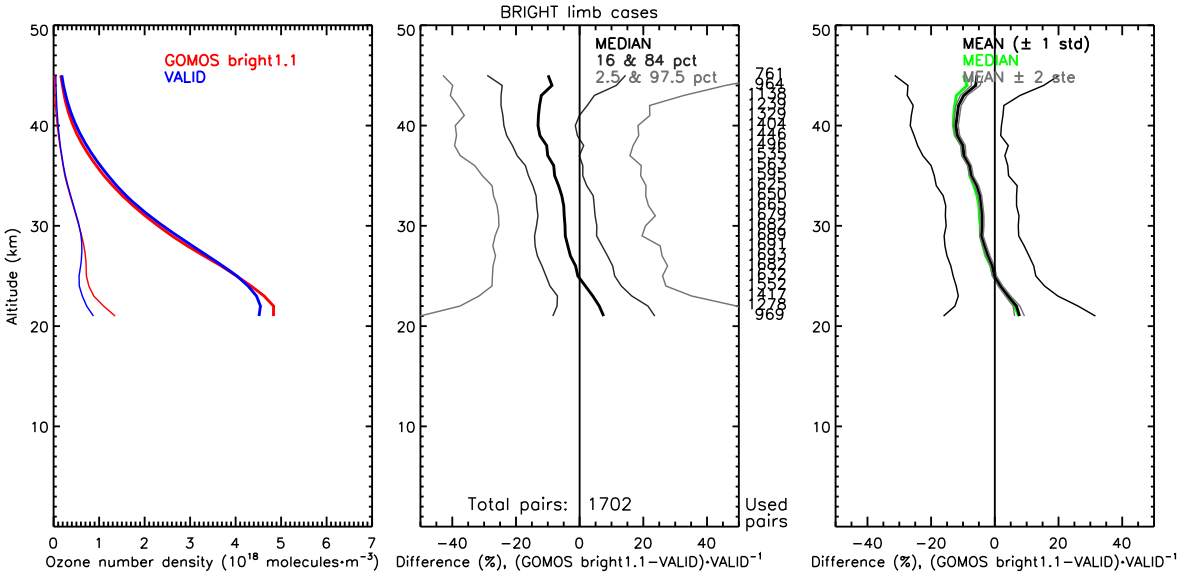



Figure 18. Validation results for the scientific retrieval for GOMOS bright limb ozone data (version 1.1). Shown are the mean GOMOS (red) and lidar ozone profiles (blue) together with the standard deviation (thin lines) in the left panel; middle panel shows the relative differences in percentiles; the right panel shows the median (green) and mean relative differences together with the mean ± one standard deviation (thin black lines) and two standard errors (light grey lines).

The albedo of the observed GOMOS scene relates to the amount of stray light in the GOMOS observation. The stray light removal part of the GBL retrieval algorithm may show sensitivity to large variations in albedo in the observation geometry (Tukiainen et al., 2011). Some trends can be observed: for instance at 30 km the bias becomes increasingly negative with increasing albedo, whereas it becomes more positive at 45 km (Figure 19). When the observed scene is even brighter (albedo > 0.5), the number of outliers increases and the interquantile spread as well.

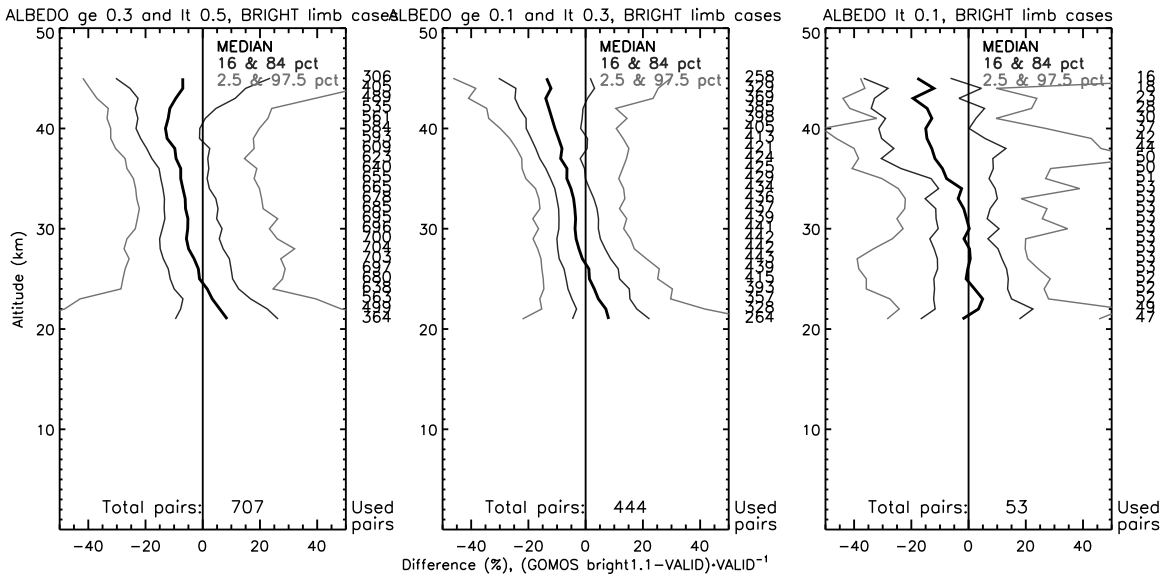



Figure 19. Validation results for GBL version 1.1 ozone profiles in comparison to lidar grouped by albedo of the observed scène. Left panel: albedo ranging between 0.3 and 0.5; middle panel: albedo ranging between 0.1 and 0.3; right panel: albedo lower than 0.1.

With respect to the data quality of the operational GOMOS product before and after September 11th 2002, for temperature and ozone no significant differences could be found. The dataset covering the period before September 11th 2002 was limited in extent and especially for temperature, few collocations were found. One possibility of not finding strong signs of data of reduced quality could be that such measurements might not have made it through all processing steps or be rejected based on the reported uncertainties.

Although the results for ozone of version 6.01 in comparison to version 5.00 were somewhat disappointing, the high resolution temperature profiles did show an improved quality for version 6.01. Overall, the number of collocations had substantially increased (more successful retrievals), the number of outliers had reduced and the agreement with lidar and sonde observations was better. Collocations are mostly with medium bright and bright stars and it was found that the altitude coverage with valid data increases with the observed star's brightness. Cold stars appeared to give somewhat cooler profiles at lower altitudes. Temperatures appear to be somewhat enhanced from version 5.00 to version 6.01, which especially clear in the tropics, where a cold bias was seen with version 5.00 that has turned into a small warm bias with version 6.01. The agreement for version 6.01 HRTP with lidar and sonde is over most altitudes within a few Kelvin (Figure 20) and for comparisons with sonde, differences were not significant.

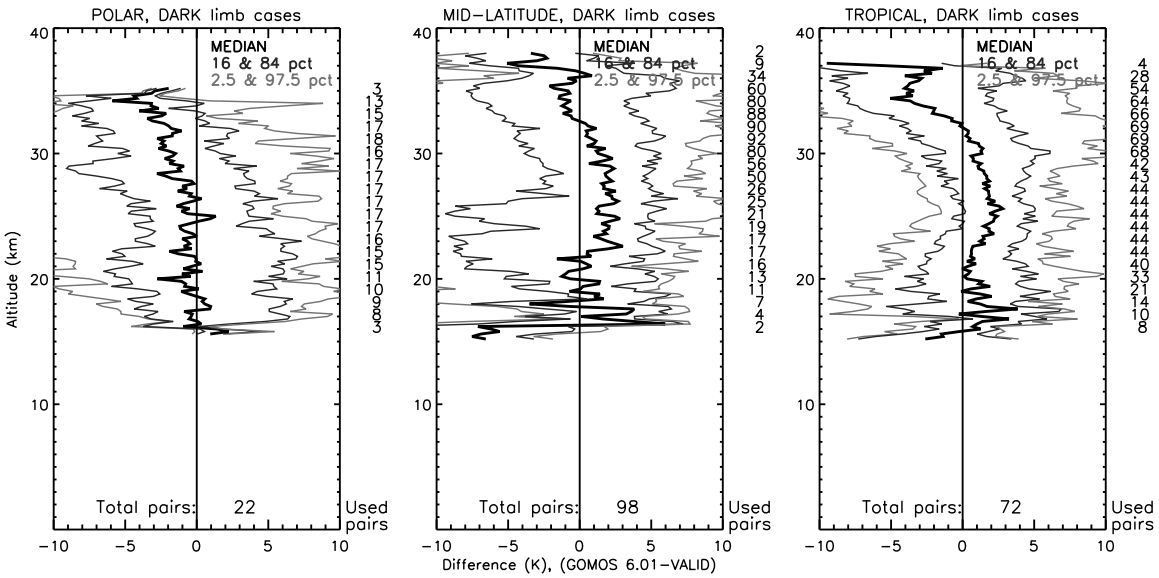



Figure 20. Differences between GOMOS version 6.01 and lidar temperature profiles for the three latitude regions. Shown are the percentile absolute differences.

# 7 Comparison of $NO_2$ data from lidar and sonde

Nitrogen dioxide profiles observed by the lidar and sonde during the CINDI and PEGASOS campaigns are being compared in this section.

The CINDI (Cabauw Intercomparison Campaign of Nitrogen Dioxide measuring Instruments) campaign was held in June-July 2009 at Cabauw (the Netherlands). The main objective of the campaign was to intercompare the $NO_2$ measuring instruments that can be used for validation of tropospheric $NO_2$ from satellites. Apart from $NO_2$, other parameters were measured and intercompared, among which ozone, aerosol, HCHO, CHOCHO, and BrO.

The second campaign dataset is from PEGASOS (Pan-European Gas-AeroSOls-climate interaction Study), as part of which in May 2012 a campaign was held in the Netherlands with a visiting zeppelin. PEGASOS is a project within the European Union 7$^{th}$ framework programme, with major focus on air quality, covering the spatiotemporal scales that connect local surface-air pollution, air quality and meteorological conditions with global atmospheric chemistry and climate. Ground and airborne observations were taken simultaneously.

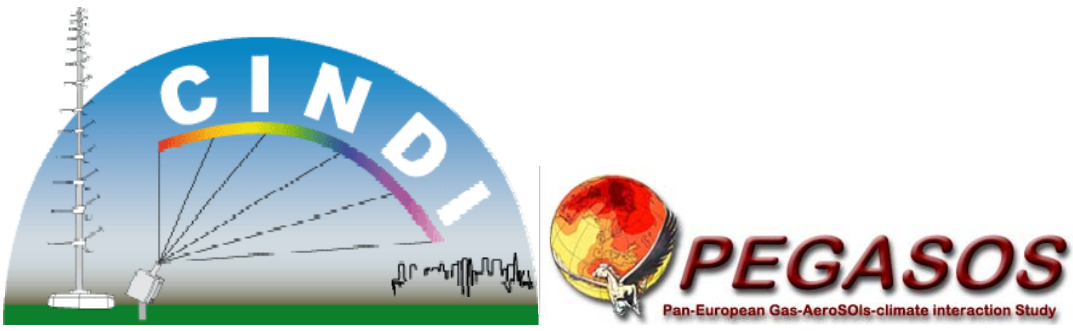



Figure 21. Campaign logos for CINDI (left) and PEGASOS (right).

Seven coincident measurements taken during the CINDI campaign and one collocated measurement during the PEGASOS campaign were selected.

First, measurements taken with $NO_x$-monitors combined with blue light convertors at the same location and two altitudes – 3 and 200 m - are shown in Figure 22 to illustrate variations during the course of the day. Large variations are visible in time, but also differences occur with altitude due to differentiation of the boundary layer-free troposphere). Narrow peaks further indicate that processes can be very local. **Figure 23** then shows how the sonde observations compare to the monitor data. Agreement is very good in the matching period (10:30 until 11:30).

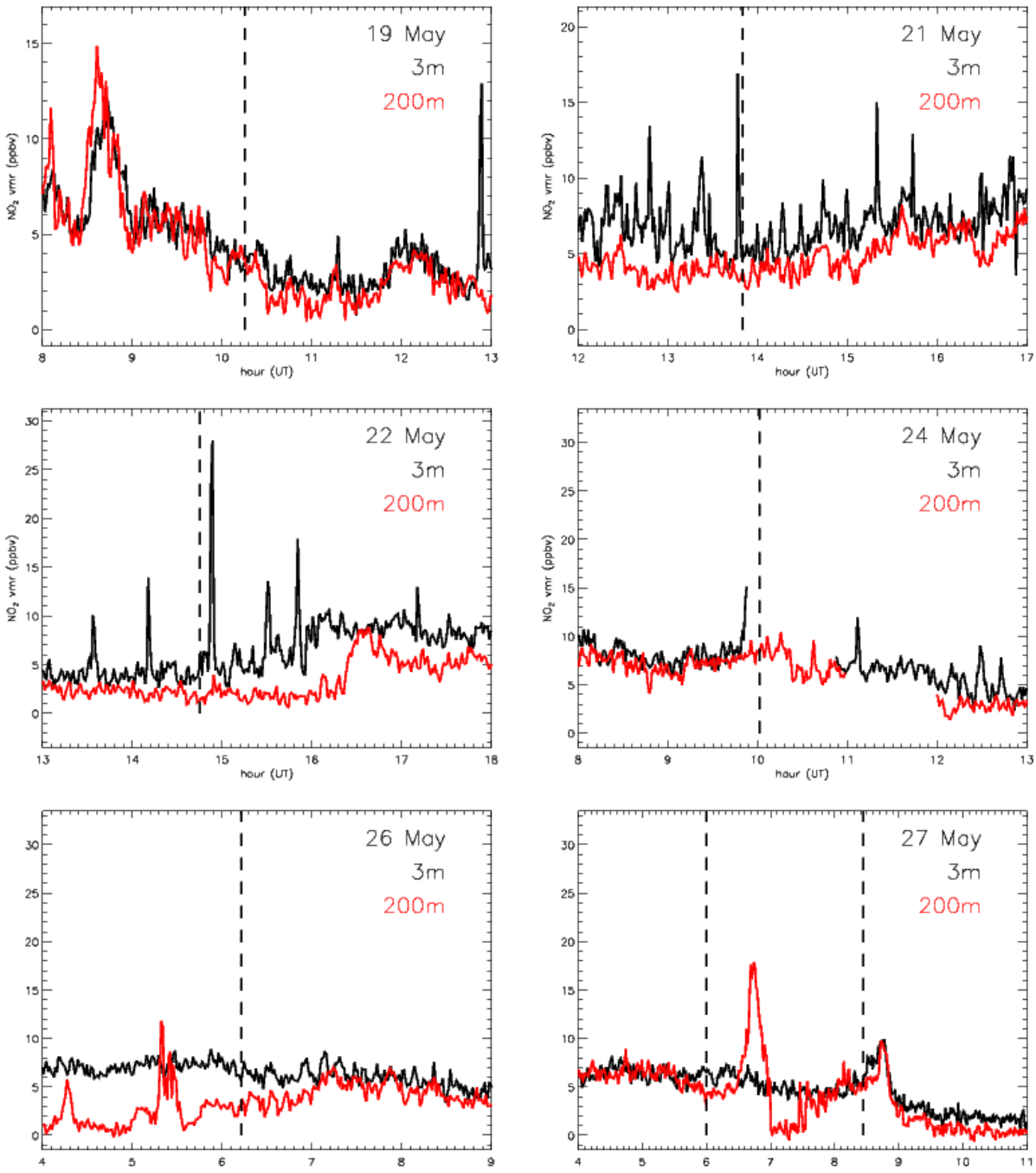



Figure 22. $NO_2$ volume mixing ratio as a function of time measured by the RIVM $NO_x$+blue light convertors-monitors at Cabauw at two altitudes: 3 meter (black) and 200 m (red) on 6 days in May 2012 (PEGASOS campaign period).

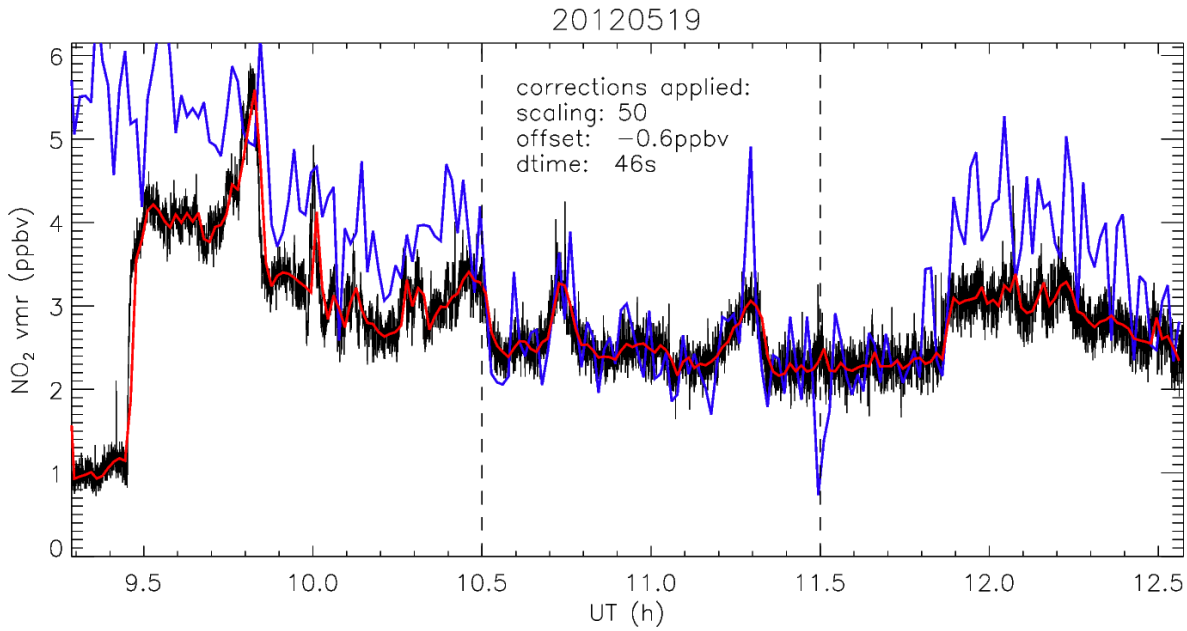



Figure 23. Comparison of $NO_2$ sonde (black, averaged data in red) with $NO_x$ monitor data (blue). Before 10:30 (10.5), measurements are not taken at the same location. Between 10:30 and 11:30, measurements are done at the same location. After 11:30 the sonde data are not corrected for acidification of the sonde solution (which results in a lower sensitivity).

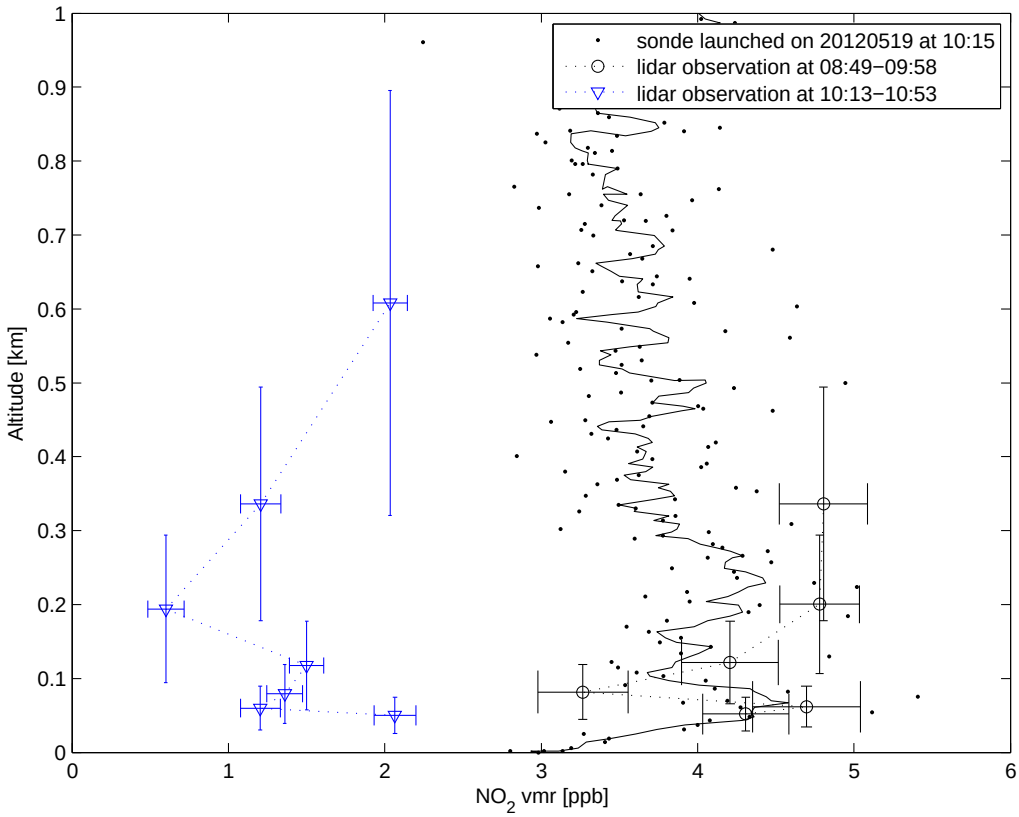



Figure 24. Comparison between $NO_2$-profile derived with the sonde (black dots) and with the lidar (black circles and blue diamonds for consecutive measurements) on May 19th 2012. Note that the lidar observations in this particular campaign could possibly underestimate the $NO_2$ concentration (due to a problem with the wavelength calibration), but the profile shape is said to be reliable.

Figure 24 also gives an example of how fast the $NO_2$ concentration can change. Two consecutive lidar profiles are shown (black circles for the observation between ~9 and 10, blue diamonds for the profile that started 15 minutes later) with differences larger than a factor two. The sonde observation matches in time with the blue lidar profile, but in fact better agrees with the previous (black) one. This can be attributed to the viewing direction of the lidar and the location where the sonde was launched together with the wind direction at that time. The air measured by the lidar before 10 o'clock has moved into the direction of the sonde which was launched at 10:15.

Figure 25 also shows the large variability observed during the course of one day, where high $NO_2$ concentrations observed during the early morning are reduced as the boundary layer height increases and the air mixes. Agreement between the sonde and closest lidar observation (green diamonds) is quite good, with differences at the lowest altitudes probably resulting from the difference in location. Also the reduction in $NO_2$ seen by the sonde above 1.3 km is seen by the lidar observation from 9 to 10 (red squares). More comparisons for the CINDI campaign are given in Appendix I.

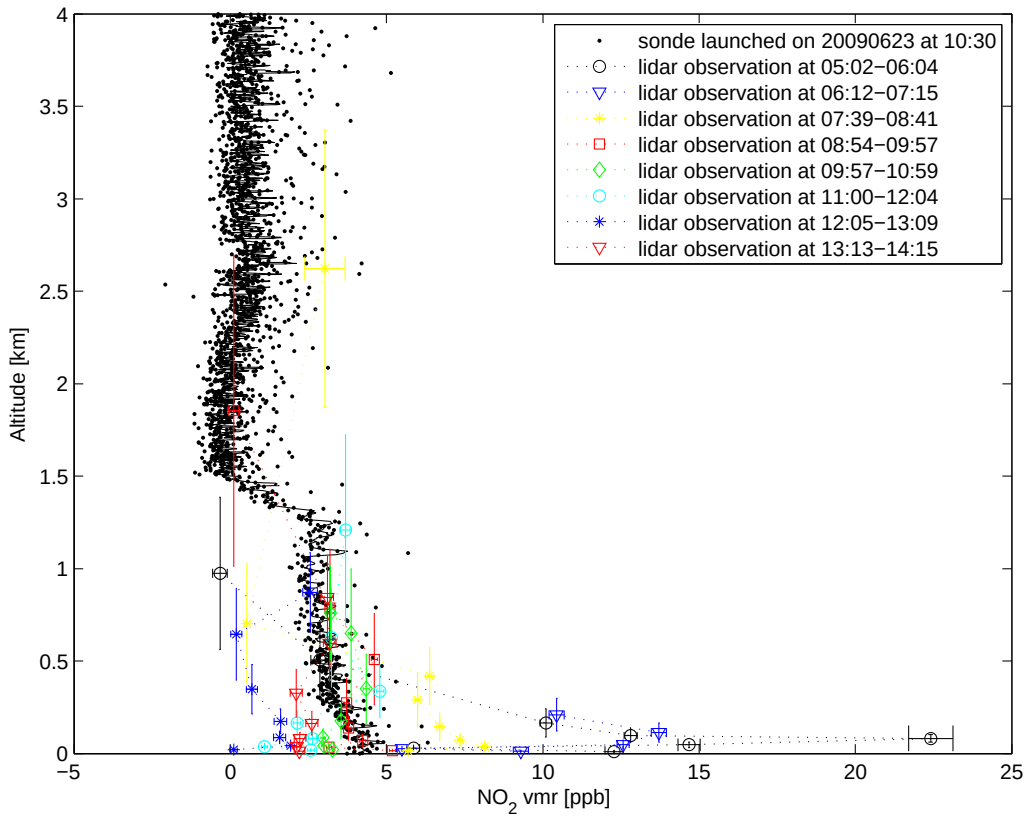



**Figure 25. $NO_2$ profiles measured by sonde (black dots) and lidar (various colours/symbols with dotted lines) on June 23rd 2009.**

We can conclude that variation in $NO_2$ can be very rapidly changing with time and variations can also be very local. When comparing with satellite profiles, this diversity should be taken into account. The current datasets do not permit a comparison with satellite retrievals. However, the information obtained here could be used as input for retrievals and for models, as it was shown in section 2.1.3 that a wrong apriori assumption for a satellite retrieval of a tropospheric $NO_2$ column can lead to substantial (50%) errors. Extending these datasets into a climatology, where - for instance - more seasons are covered, and further development of the $NO_2$ sonde towards absolutely-calibrated profiles is recommended.

# 8 Output

This chapter gives an overview of the presentations and documents prepared and contributed to on the validation activities.

Articles

- Richter, A. et al.: Validation strategy for satellite observations of tropospheric reactive gases (under review at Annals of Geophysics)
- Hubert, D. et al.: Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records (in preparation for AMT)
- Nair, P. et al. (2012): Relative drifts and stability of satellite and ground-based stratospheric ozone profiles at NDACC lidar stations. Atmospheric Measurement Techniques 5, 1301-1318, 2012, www.atmos-meas-tech.net/5/1301/2012/.

Reports

- Validation of SCIAMACHY version 5.02
- MIPAS level 2 version 6 CLOUD processing ozone and temperature validation
- MIPAS 2D GMTR version 2.2 comparison
- Validation of MIPAS ML2PP version 6 ozone and temperature profiles
- On the 20020911 threshold for GOMOS data quality
- Effect of filtering out cool and weak stars
- Validation of GOMOS bright limb version 1.1 ozone profiles
- Validation of GOMOS version 6
- Validation of GOMOS version 6.01 ozone profiles

Conferences, workshops and meetings attended in person

- MIPAS quality working group meeting Florence March 2011
- SSAG De Bilt May 2012
- NDACC lidar working group Potsdam June 2011
- NDACC/ISSI team meeting Bern June 2011
- NDACC symposium La Reunion November 2011
- ATMOS conference Brugge June 2012
- GOMOS quality working group meeting Paris September 2012
- NDACC/ISSI team meeting Bern October 2012
- SCIAVALIG workshop De Bilt November 2012
- GRUAN workshop De Bilt February 2013
- ACVE workshop Frascati March 2013
- SCIAMACHY quality working group meeting Bremen June 2013
- NDACC/ISSI team meeting Bern September 2013

Posters

- Changes and trends of ozone in the southern mid-latitude upper troposphere and lower stratosphere derived from long-term ozonesonde and ozone lidar measurements at Lauder, New Zealand (Guang Zeng et al., IGAC 2012)

- ♠ Validation of atmospheric sensors using lidar – the VALID-2 project, ATMOS 2012
- ♠ Latest updates of ozone profile data from Odin, Envisat and ACE – bias and stability with respect to NDACC/GAW ozonesondes and lidars (Daan Hubert et al., ACVE 2013)
- ♠ Validation of GOMOS v6 ozone and temperature profiles with lidar and sonde data, ACVE 2013
- ♠ Validation of MIPAS ML2PP v6 ozone and temperature profiles with lidar and other data, ACVE 2013

Presentations

- ♠ MIPAS version 5 ozone and temperature validation, MIPAS QWG, March 2011
- ♠ Satellite validation with lidar project, NDACC LWG meeting, June 2011
- ♠ Not so constant constants, NDACC/ISSI team, June 2011
- ♠ MIPAS version 6 ozone and temperature quick validation, MIPAS QWG, October 2011
- ♠ The VALID-2 project, NDACC symposium, November 2011
- ♠ GOMOS version 6 quick validation, GOMOS QWG, March 2012
- ♠ SCIAMACHY ozone profile assessment, SSAG, May 2012
- ♠ GOMOS v5 and v6 comparison with VALID, GOMOS QWG, September 2012
- ♠ MIPAS ozone and temperature VALID comparisons, MIPAS QWG, November 2012
- ♠ SCIAMACHY ozone profiles (VALID results), SCIAVALIG workshop, November 2012
- ♠ Satellite profile validation with lidar (+ sonde & microwave radiometer), GRUAN workshop, February 2013
- ♠ Validation of SCIAMACHY v5.02 ozone profiles with lidar and sonde data (plus MIPAS/GOMOS results), ACVE, March 2013
- ♠ Quick verification of SCIA v575 ozone profiles, SCIAMACHY QWG, June 2013
- ♠ SCIAMACHY version 5.02 ozone profile validation, SCIAMACHY QWG, June 2013
- ♠ Input for Summary ESA Validation Reports Diagnostic dataset, MIPAS QWG, July 2013
- ♠ Digging into GOMOS data, comparisons of individual years – GOMOS vs VALID, communication with FMI, August 2013
- ♠ Contribution to presentations by Arno Keppens and Stefano Casadio on MIPAS comparisons, MIPAS QWG, November 2013

Note: Some presentations (e.g. at ATMOS) included VALID-2 results without informing us – these are not listed here.

# 9 References

Brinksma, E.J., G. Pinardi, H. Volten, R. Braak, A. Richter, A. Schönhardt, M. van Roozendael, C. Fayt, C. Hermans, R.J. Dirksen, T. Vlemmix, A.J.C. Berkhout, D.P.J. Swart, H. Oetjen, F. Wittrock, T. Wagner, O.W. Ibrahim, G. de Leeuw, M. Moerman, R.L. Curier, E.A. Celarier, A. Cede, W.H. Knap, J.P. Veefkind, H.J. Eskes, M. Allaart, R. Rothe, A.J.M. Piters and P.F. Levelt (2008). *The 2005 and 2006 DANDELIONS $NO_2$ and aerosol intercomparison campaigns*. Journal of Geophysical 113, D16S46, doi:10.1029/2007JD008808.

Funatsu, B.M., C. Claud, P. Keckhut, W. Steinbrecht and A. Hauchecorne (2011). *Investigations of stratospheric temperature regional variability with lidar and Advanced Microwave Sounding Unit*. Journal of Geophysical Research 116, D08106, doi:10.1029/2010JD014974.

Keckhut, P., A. Hauchecorne, I.S. Mcdermid, T. Leblanc, H. Bencherif and G. DiDonFrancesco (2006). *Temperature validation using rayleigh-raman lidars*. Proceedings of the 1st EPS/MetOp RAO workshop, 15-17 May 2006, ESRIN, ESA SP-618.

Keckhut, P. et al. (2011). An evaluation of uncertainties in monitoring middle atmosphere temperatures with the ground-based lidar network in support of space observations. Journal of Atmospheric and Solar-Terrestrial Physics 73(5-6), 627-642, doi: 10.1016/j.jastp.2011.01.003.

Li, T., T. Leblanc, I.S. McDermid, P. Keckhut, A. Hauchecorne and X. Dou (2011). *Middle atmosphere temperature trend and solar cycle revealed by long-term Rayleigh lidar observations*. Journal of Geophysical Research 116, D00P05, doi:10.1029/2010JD015275.

NIMA: Department of Defense World Geodetic System 1984, Technical Report, 1984.

Piters, A.J.M., K.F. Boersma, M. Kroon, J.C. Hains, M. Van Roozendael, F. Wittrock, N. Abuhassan, C. Adams, M. Akrami, M.A.F. Allaart, A. Apituley, S. Beirle, J.B. Bergwerff, A.J.C. Berkhout, D. Brunner, A. Cede, J. Chong, K. Clémer, C. Fayt, U. Frieß, L.F.L. Gast, M. Gil-Ojeda, F. Goutail, R. Graves, A. Griesfeller, K. Großmann, G. Hemerijckx, F. Hendrick, B. Henzing, J. Herman, C. Hermans, M. Hoexum, G.R. van der Hoff, H. Irie, P.V. Johnston, Y. Kanaya, Y.J. Kim, H. Klein Baltink, K. Kreher, G. de Leeuw, R. Leigh, A. Merlaud, M.M. Moerman, P.S. Monks, G.H. Mount, M. Navarro-Comas, H. Oetjen, A. Pazmino, M. Perez-Camacho, E. Peters, A. du Piesanie, G. Pinardi, O. Puentedura, A. Richter, H.K. Roscoe, A. Schönhardt, B. Schwarzenbach, R. Shaiganfar, W. Sluis, E. Spinei, A.P. Stolk, K. Strong, D.P.J. Swart, H. Takashima, T. Vlemmix, M. Vrekoussis, T. Wagner, C. Whyte, K.M. Wilson, M. Yela, S. Yilmaz, P. Zieger and Y.

Zhou (2012). *The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results*. Atmospheric Measurement Techniques 5, 457-485, doi:10.5194/amt-5-457-2012.

Tukiainen, S., E. Kyrölä, P.T. Verronen, D. Fussen, L. Blanot, G. Barrot, A. Hauchecorne and N. Lloyd (2011). *Retrieval of ozone profiles from GOMOS limb scattered measurements*. Atmospheric measurement techniques 4, 659–667, doi:10.5194/amt-4-659-2011.

Sluis, W.W., M.A.F. Allaart, A.J.M. Piters and L.F.L. Gast (2010). *The development of a nitrogen dioxide sonde*. Atmospheric Measurement Techniques 3, 1753–1762, doi:10.5194/amt-3-1753-2010.

Steinbrecht, W., T.J. McGee, L.W. Twigg, H. Claude, F. Schönenborn, G.K. Sumnicht and D. Silbert (2009). Intercomparison of stratospheric ozone and temperature profiles during the October 2005 Hohenpeißenberg Ozone Profiling Experiment (HOPE). Atmospheric Measurement Techniques 2, 125–145, doi:10.5194/amt-2-125-2009.

Sofieva, V., S. Tukiainen and E. Kyrölä (October 2012). Technical note: *On GOMOS v.6 ozone data quality*.

Volten, H., E.J. Brinksma, A.J.C. Berkhout, J. Hains, J.B. Bergwerff, G.R. van der Hoff, A. Apituley, R.J. Dirksen, S. Calabretta-Jongen and D.P.J. Swart (2009). *$NO_2$ lidar profile measurements for satellite interpretation and validation*. Journal of Geophysical Research 114, D24301, doi:10.1029/2009JD012441.

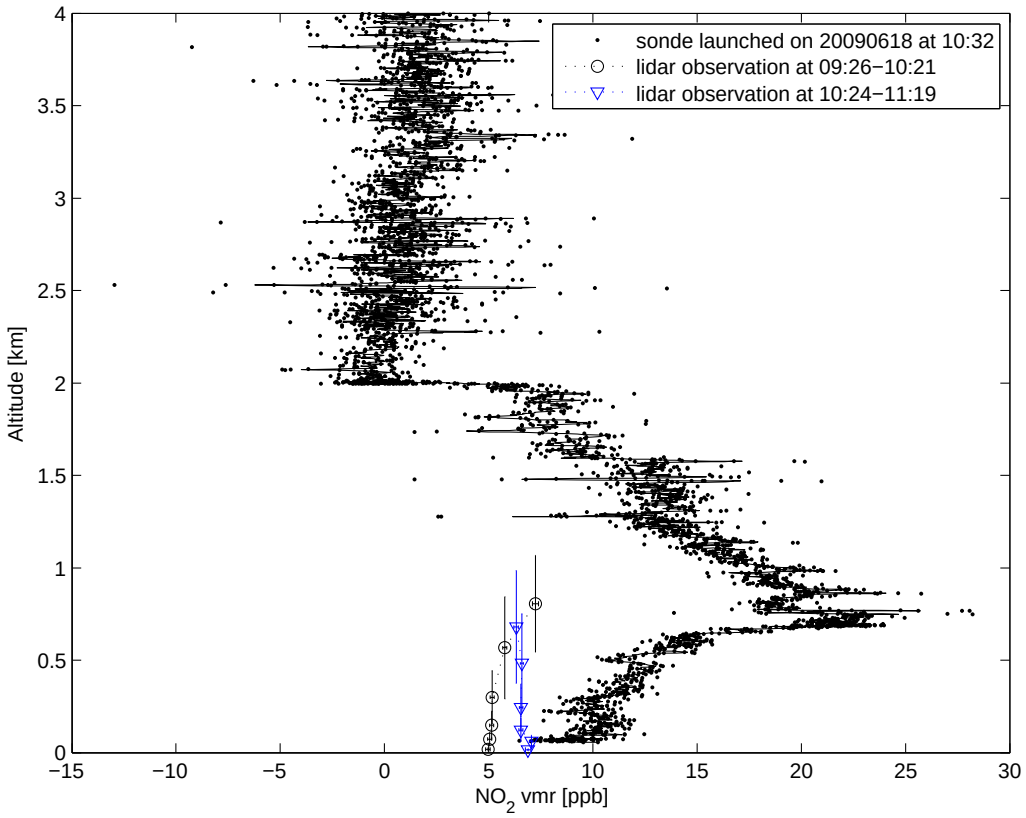
Zurita Milla, R., J.A.E. van Gijsel, N.A.S. Hamm, P.W.M. Augustijn and A. Vrieling (2013). *Exploring spatio-temporal phenological patterns and trajectories using self-organizing maps.* IEEE transactions on geoscience and remote sensing 51(4), 1914-1921, doi:10.1109/TGRS.2012.2223218.

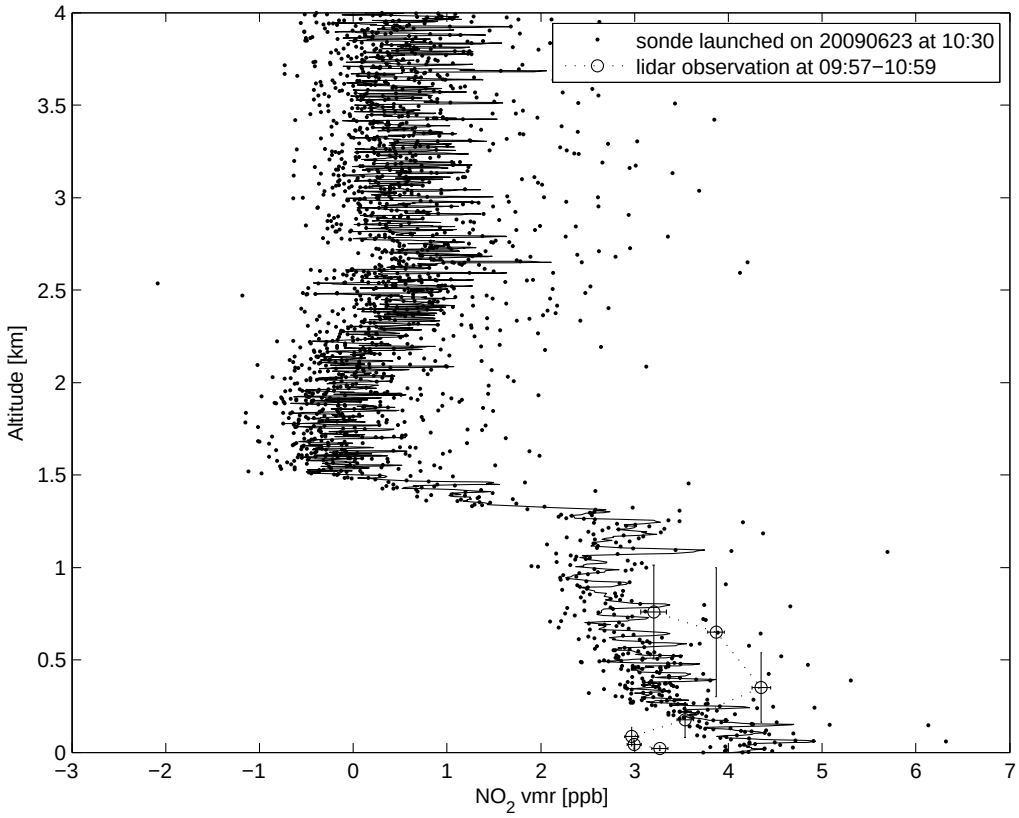
# 10 Acronyms and abbreviations

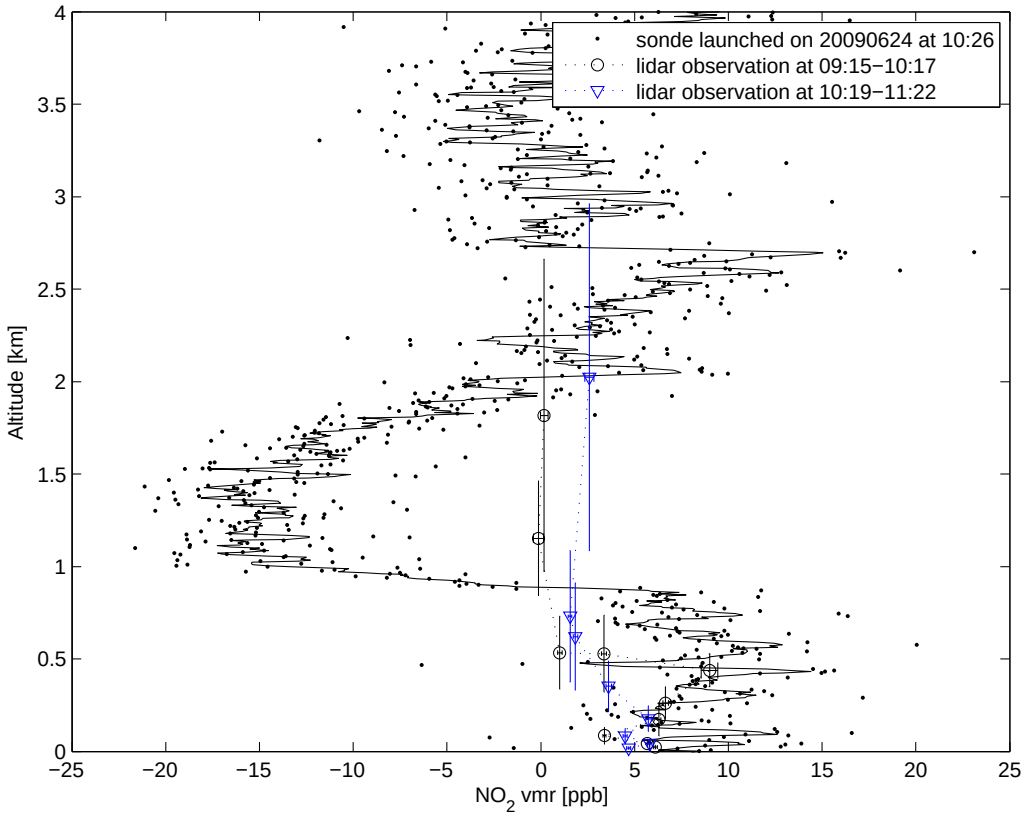
| | |
|---|---|
| ACVE | Atmospheric Composition Validation and Evolution (ESA conference) |
| ATMOS | ESA conference devoted to atmospheric research |
| BIRA | Belgian spatial aeronomy institute |
| CINDI | Cabauw Intercomparison Campaign of Nitrogen Dioxide measuring Instruments (campaign) |
| DANDELIONS | Dutch Aerosol and Nitrogen Dioxide Experiments for the vaLIdation of OMI and SCIAMACHY |
| DIAL | DIfferential Absorption Lidar |
| ENVISAT | ENVIronmental SATellite |
| FMI | Finnish Meteorological Institute |
| GECA | Generic Environment for Calibration/validation Analysis |
| GOMOS | Global Ozone Monitoring by Occultation of Stars |
| GBL | GOMOS Bright Limb (data) |
| GCOS | Global Climate Observing System |
| GMTR | Geofit Multi-Target Retrieval |
| GRUAN | GCOS Reference Upper-Air Network |
| GSFC | Goddard Space Flight Center |
| HRTP | High Resolution Temperature Profile |
| ISSI | International Space Science Institute |
| KNMI | Royal Netherlands Meteorological Institute |
| LWG | Lidar Working Group |
| MIPAS | Michelson Interferometer for Passive Atmospheric Sounding |
| ML2PP | MIPAS Level 2 Processing Prototype |
| NDACC | Network for the Detection of Atmospheric Composition Change |
| QWG | Quality Working Group |
| PEGASOS | Pan-European Gas-AeroSOls-climate interaction Study |
| RIVM | National institute for public health and the environment |
| SCIAMACHY | SCanning Imaging Absorption spectrometer for Atmospheric CHartographY |
| SSAG | SCIAMACHY Science Advisory Group |
| VALID | satellite VAlidation with LIDar (project) |

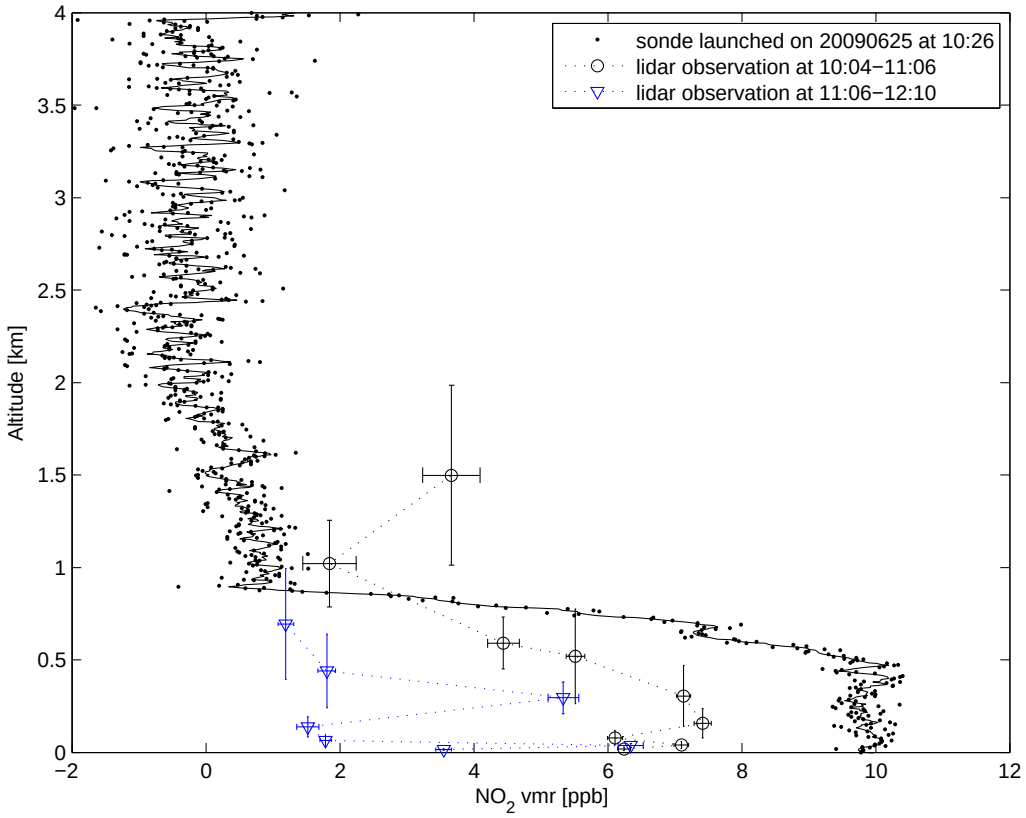
# 11 Appendix I. $NO_2$ concentrations measured by sonde and lidar during the CINDI campaign

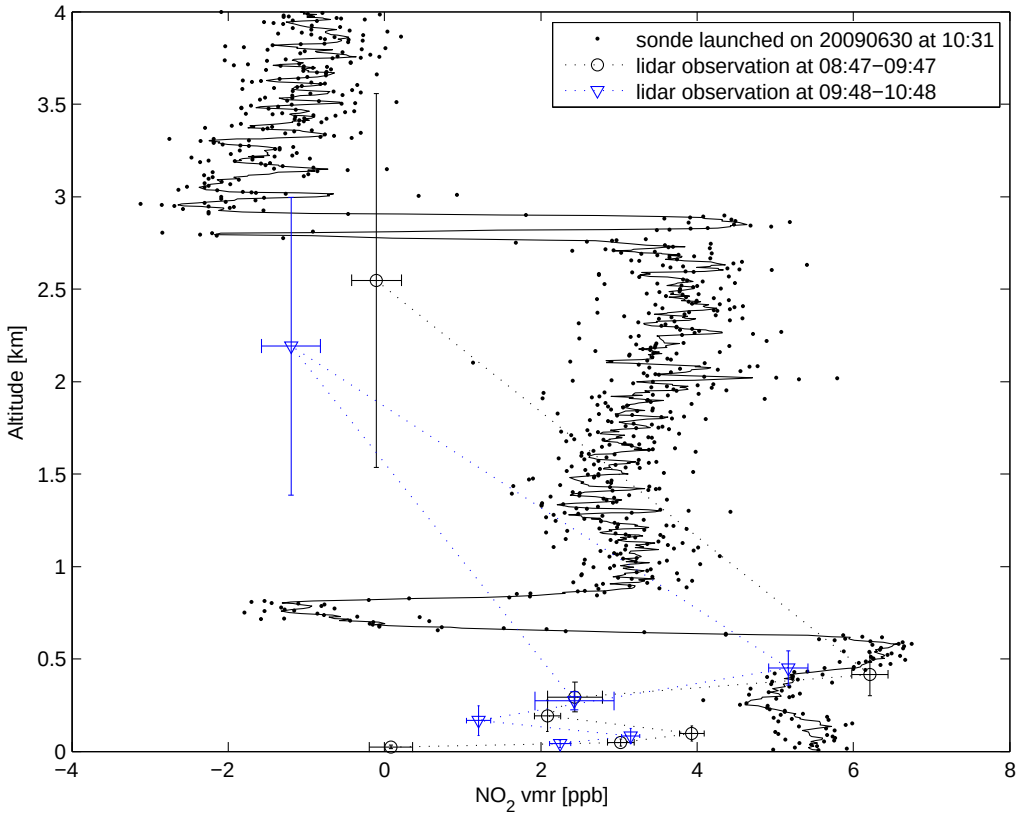
This appendix shows the comparisons between sonde and the nearest lidar observations of $NO_2$ profiles obtained during the CINDI campaign in 2009.

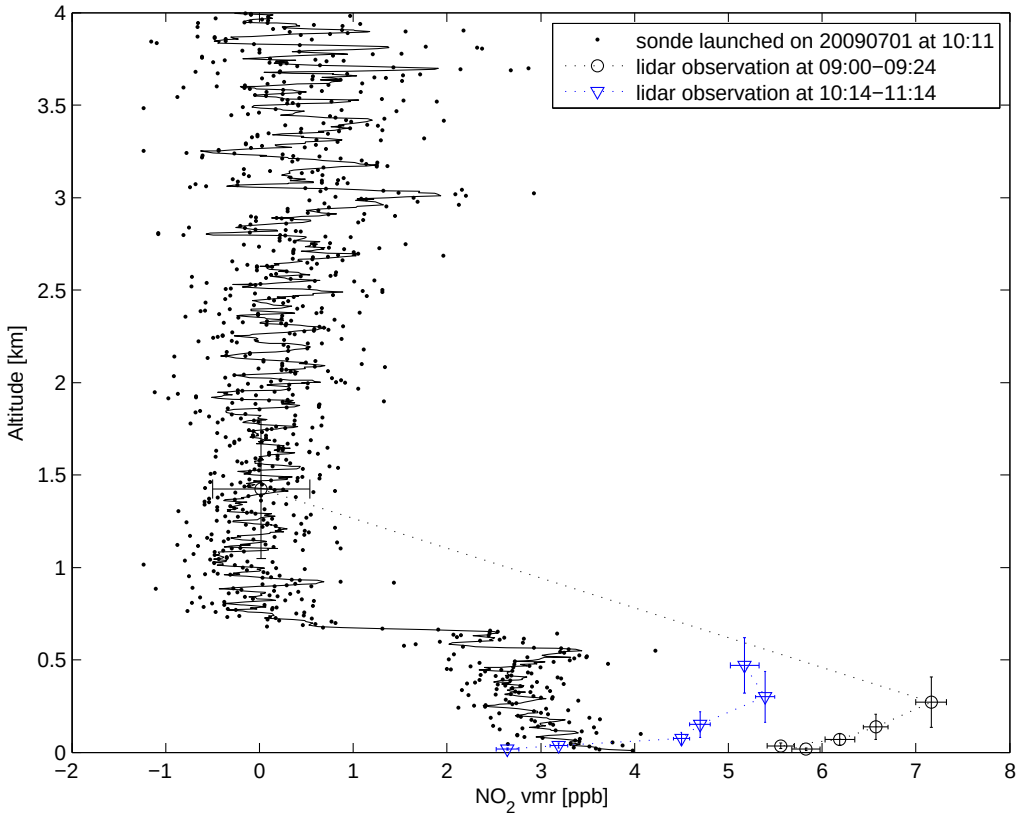


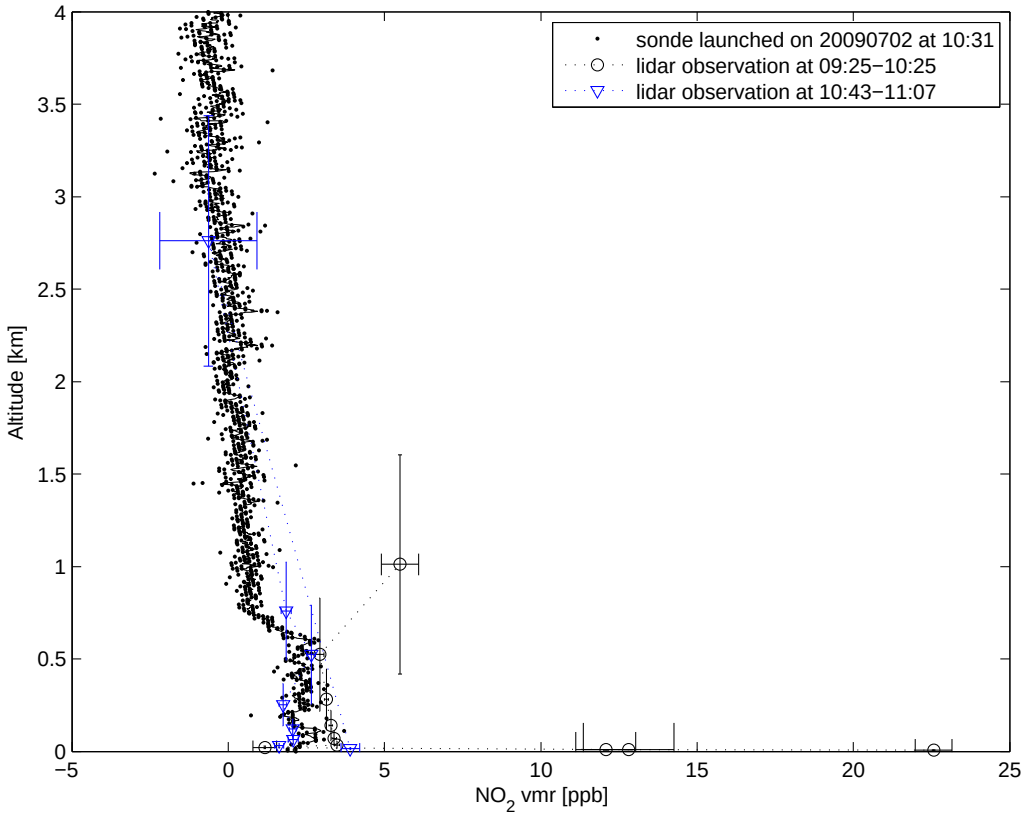

sonde launched on 20090623 at 10:30
lidar observation at 09:57−10:59
Altitude [km]
NO2 vmr [ppb]


sonde launched on 20090624 at 10:26
lidar observation at 09:15−10:17
lidar observation at 10:19−11:22
Altitude [km]
NO2 vmr [ppb]


sonde launched on 20090625 at 10:26
lidar observation at 10:04−11:06
lidar observation at 11:06−12:10
Altitude [km]
NO2 vmr [ppb]
0
0.5
1
1.5
2
2.5
3
3.5
4
−2
0
2
4
6
8
10
12


sonde launched on 20090630 at 10:31
lidar observation at 08:47−09:47
lidar observation at 09:48−10:48
Altitude [km]
NO2 vmr [ppb]
0
0.5
1
1.5
2
2.5
3
3.5
4
−4
−2
0
2
4
6
8


sonde launched on 20090701 at 10:11
lidar observation at 09:00−09:24
lidar observation at 10:14−11:14
Altitude [km]
NO2 vmr [ppb]


sonde launched on 20090702 at 10:31
lidar observation at 09:25−10:25
lidar observation at 10:43−11:07
Altitude [km]
NO2 vmr [ppb]

# 12 Appendix II: Validation summary tables

This appendix presents the validation results in a summarised way for various instruments and data versions. The reported values are the median difference (relative for ozone) together with the 68-interquantile spread for intervals of five kilometre. Nota bene: Please be advised that by summarising, information will be lost. Values cannot be directly compared for different versions, as the input datasets are usually not the same in terms of temporal and spatial coverage. For some cases and regions/altitudes no numbers are provided to avoid confusion/outliers due to the dataset being very limited.

## 12.1 SCIAMACHY ozone

**version 3.01 compared to lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | -6 | 34 | -10 | 29 | -3 | 35 |
| 20-25 km | -6 | 22 | -11 | 15 | -16 | 13 |
| 25-30 km | -14 | 25 | -5 | 22 | -5 | 12 |
| 30-35 km | -23 | 30 | -5 | 27 | 5 | 20 |
| 35-40 km | -35 | 32 | -17 | 29 | -18 | 26 |
| 40-45 km | -16 | 51 | -5 | 45 | -3 | 30 |

**version 5.01 compared to lidar, non-cloudfree cases**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | -4 | 36 | 0 | 35 | 11 | 54 |
| 20-25 km | -5 | 23 | 3 | 20 | 5 | 21 |
| 25-30 km | -8 | 22 | 4 | 25 | 10 | 19 |
| 30-35 km | -13 | 27 | 8 | 26 | 19 | 25 |
| 35-40 km | -9 | 36 | 8 | 34 | 11 | 32 |
| 40-45 km | 0 | 42 | 13 | 40 | 16 | 30 |

**version 5.02 compared to lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 1 | 43 | 1 | 38 | 8 | 43 |
| 20-25 km | 3 | 27 | 3 | 20 | 4 | 18 |
| 25-30 km | -5 | 27 | 7 | 26 | 13 | 16 |
| 30-35 km | -1 | 39 | 7 | 30 | 20 | 25 |
| 35-40 km | 2 | 54 | 7 | 40 | 11 | 32 |
| 40-45 km | -3 | 62 | 15 | 49 | 16 | 32 |

## 12.2 MIPAS ozone

**version ML2PP v5**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 5 | 32 | 4 | 40 | 2 | 46 |
| 20-25 km | 2 | 27 | 5 | 20 | 4 | 23 |
| 25-30 km | -6 | 30 | 1 | 25 | -2 | 18 |
| 30-35 km | -2 | 37 | 5 | 25 | 6 | 20 |
| 35-40 km | -3 | 45 | 10 | 31 | 9 | 20 |
| 40-45 km | -5 | 49 | 6 | 42 | 9 | 30 |

**version ML2PP v6 day time**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 4 | 33 | 5 | 41 | 4 | 44 |
| 20-25 km | 1 | 28 | 5 | 19 | 8 | 18 |
| 25-30 km | -3 | 30 | 1 | 22 | -3 | 14 |
| 30-35 km | 1 | 35 | 5 | 24 | 8 | 17 |
| 35-40 km | 2 | 40 | 8 | 29 | 9 | 17 |
| 40-45 km | -9 | 45 | 6 | 42 | 11 | 33 |

**version ML2PP v6 night time**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 5 | 29 | 5 | 39 | -5 | 41 |
| 20-25 km | 1 | 35 | 6 | 20 | 5 | 19 |
| 25-30 km | -4 | 31 | 2 | 23 | 5 | 13 |
| 30-35 km | -1 | 37 | 5 | 23 | 12 | 18 |
| 35-40 km | -3 | 46 | 8 | 32 | 9 | 20 |
| 40-45 km | -7 | 49 | 1 | 43 | 11 | 33 |

**version ML2PP v6**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 5 | 30 | 5 | 40 | 1 | 43 |
| 20-25 km | 1 | 26 | 5 | 19 | 7 | 18 |
| 25-30 km | -3 | 31 | 1 | 22 | 0 | 16 |
| 30-35 km | -1 | 37 | 5 | 23 | 9 | 18 |
| 35-40 km | -1 | 44 | 8 | 30 | 9 | 18 |
| 40-45 km | -8 | 48 | 4 | 42 | 11 | 33 |

**version ML2PP v6 AVKs applied**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 2 | 27 | 3 | 34 | -1 | 36 |
| 20-25 km | 0 | 17 | 2 | 14 | 4 | 12 |
| 25-30 km | 1 | 16 | 1 | 11 | 0 | 7 |
| 30-35 km | 2 | 21 | 1 | 12 | 1 | 9 |
| 35-40 km | 3 | 36 | 3 | 16 | 1 | 4 |
| 40-45 km | 0 | 100 | 3 | 37 | 2 | 10 |

## 12.3 MIPAS temperature

**version 5, comparison with lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 1 | 10 | 3 | 5 | 3 | 4 |
| 20-25 km | 0 | 9 | 1 | 5 | 1 | 3 |
| 25-30 km | -1 | 9 | 1 | 7 | 1 | 4 |
| 30-35 km | -1 | 11 | 0 | 6 | 0 | 7 |
| 35-40 km | -2 | 11 | 0 | 7 | -1 | 7 |
| 40-45 km | -1 | 13 | 1 | 7 | -2 | 8 |
| 45-50 km | 1 | 15 | -1 | 7 | -3 | 8 |
| 50-55 km | 1 | 15 | -1 | 9 | -3 | 10 |
| 55-60 km | 3 | 22 | -2 | 9 | -4 | 12 |
| 60-65 km | 2 | 18 | -3 | 10 | -4 | 12 |
| 65-70 km | 1 | 21 | -2 | 14 | -6 | 17 |

**version ML2PP 6, comparison with lidar, MIPAS nighttime**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 1 | 9 | 3 | 5 | 2 | 3 |
| 20-25 km | 0 | 8 | 1 | 5 | 1 | 3 |
| 25-30 km | -1 | 7 | 1 | 6 | 0 | 4 |
| 30-35 km | -2 | 10 | 1 | 7 | 0 | 6 |
| 35-40 km | -1 | 11 | 0 | 7 | -1 | 7 |
| 40-45 km | -1 | 14 | 1 | 8 | -2 | 8 |
| 45-50 km | 1 | 14 | 0 | 8 | -3 | 8 |
| 50-55 km | 2 | 16 | -1 | 9 | -3 | 9 |
| 55-60 km | 1 | 22 | -3 | 9 | -4 | 12 |
| 60-65 km | 1 | 17 | -2 | 10 | -3 | 12 |
| 65-70 km | -1 | 20 | -2 | 13 | -7 | 15 |

**version ML2PP 6, comparison with sonde, MIPAS nighttime**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| 5-10 km | -2 | 6 | -2 | 9 | NaN | NaN |
| 10-15 km | -1 | 3 | -1 | 4 | NaN | NaN |
| 15-20 km | 0 | 4 | -1 | 3 | NaN | NaN |
| 20-25 km | -1 | 4 | -1 | 3 | NaN | NaN |
| 25-30 km | -1 | 5 | -1 | 4 | NaN | NaN |

**version ML2PP 6, comparison with sonde, MIPAS daytime**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| 5-10 km | -2 | 5 | -1 | 7 | 0 | 11 |
| 10-15 km | 0 | 3 | -1 | 4 | -1 | 5 |
| 15-20 km | 0 | 3 | 0 | 4 | 0 | 5 |
| 20-25 km | -1 | 3 | -1 | 4 | -1 | 4 |
| 25-30 km | -1 | 4 | 0 | 5 | -1 | 4 |

## 12.4 GOMOS ozone

**version 5.00 solar zenith angle ≥ 108° compared with lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | -3 | 29 | 0 | 59 | 7 | 95 |
| 20-25 km | -6 | 24 | 0 | 26 | -1 | 28 |
| 25-30 km | -7 | 30 | 0 | 25 | 1 | 16 |
| 30-35 km | -3 | 38 | 1 | 25 | 2 | 16 |
| 35-40 km | -8 | 40 | 1 | 26 | 2 | 13 |
| 40-45 km | -19 | 46 | -1 | 42 | 0 | 21 |

**version 5.00 flag = 'dark' compared with lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 27 | 97 | 4 | 71 | 10 | 88 |
| 20-25 km | -10 | 39 | 2 | 31 | -1 | 28 |
| 25-30 km | -14 | 31 | 1 | 30 | 2 | 18 |
| 30-35 km | -7 | 27 | 3 | 29 | 4 | 18 |
| 35-40 km | 1 | 50 | 2 | 28 | 2 | 13 |
| 40-45 km | NaN | NaN | -1 | 40 | -1 | 21 |

**version 6.01 solar zenith angle ≥ 108° compared with lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | -3 | 29 | 0 | 51 | 11 | 72 |
| 20-25 km | -5 | 24 | 1 | 23 | 1 | 23 |
| 25-30 km | -8 | 30 | -1 | 23 | 0 | 14 |
| 30-35 km | -6 | 36 | 0 | 24 | 1 | 15 |
| 35-40 km | -11 | 41 | 0 | 27 | 1 | 12 |
| 40-45 km | -21 | 51 | -3 | 44 | -2 | 20 |

**version 6.01 flag = 'dark' compared with lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | 14 | 71 | 3 | 51 | 15 | 61 |
| 20-25 km | -2 | 120 | 2 | 22 | 2 | 21 |
| 25-30 km | -9 | 30 | 0 | 26 | 1 | 14 |
| 30-35 km | -12 | 31 | 1 | 25 | 2 | 15 |
| 35-40 km | -10 | 56 | 1 | 29 | 2 | 11 |
| 40-45 km | n/a | n/a | -3 | 43 | -2 | 19 |

**version GOMOS bright limb (GBL) 1.1, max chi$^2$ = 3 compared with lidar**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| 20-25 km | 6 | 20 | 6 | 24 | NaN | NaN |
| 25-30 km | -7 | 24 | -3 | 21 | NaN | NaN |
| 30-35 km | -7 | 31 | -5 | 17 | NaN | NaN |
| 35-40 km | -17 | 34 | -9 | 24 | NaN | NaN |
| 40-45 km | -27 | 33 | -11 | 28 | NaN | NaN |

## 12.5 GOMOS high resolution temperature profiles

**version 5, set* lidar 5 K, solar zenith angle ≥ 108°**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| 20-25 km | NaN | NaN | NaN | NaN | -2 | 6 |
| 25-30 km | NaN | NaN | 0 | 7 | -2 | 6 |
| 30-35 km | NaN | NaN | -1 | 8 | -3 | 7 |

**version 5, set sonde 5 K, solar zenith angle ≥ 108°**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | NaN | NaN | -4 | 15 | NaN | NaN |
| 20-25 km | NaN | NaN | -1 | 8 | NaN | NaN |
| 25-30 km | NaN | NaN | 0 | 8 | NaN | NaN |

**version 6, set lidar 5 K, solar zenith angle ≥ 108°**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | NaN | NaN | NaN | NaN | 4 | 11 |
| 20-25 km | NaN | NaN | NaN | NaN | 1 | 4 |
| 25-30 km | NaN | NaN | 2 | 7 | 1 | 5 |
| 30-35 km | NaN | NaN | 0 | 10 | -1 | 7 |

**version 6, set sonde 5 K, solar zenith angle ≥ 108°**

| | Polar | | Mid-latitudes | | Tropics | |
|---|---|---|---|---|---|---|
| Altitude | median | 68-iq spread | median | 68-iq spread | median | 68-iq spread |
| <20 km | NaN | NaN | -2 | 9 | NaN | NaN |
| 20-25 km | NaN | NaN | 1 | 7 | NaN | NaN |
| 25-30 km | NaN | NaN | 0 | 6 | NaN | NaN |

* "set" refers to having the same selection of observations as input for comparison with both versions 5 and 6.

Also note that for most altitudes where NaN is reported here, there are comparisons available, but with a very limited number of collocations.

# 13 Summary

The VALID-2 (satellite validation with lidar) project has run from 2011 until into 2013 and the activities carried out have been described here. The main focus has been on the validation of Envisat ozone and temperature profiles (both scientific and operational products) using lidar data. Additionally, the possibility to validate (future) satellite tropospheric $NO_2$ profiles has been assessed and efforts have been put into assisting the lidar community to homogenise the data processing and reporting (improving consistency to the reference data used for validation purposes) through participation in an ISSI team.

The validation results are summarised in appendix II. In short, for the operational SCIAMACHY v5.02 ozone profiles, best agreement with the lidar data is obtained in the polar regions and an overestimation of the ozone concentrations is seen in the mid-latitudes and tropics.

For the operational MIPAS ML2PP version 6 ozone profiles, in the polar and mid-latitude regions, differences between day-time and night-time collocated data increase with altitude, whereas in the tropics where the reversed behaviour is found. The bias relative to lidar observations is latitude-dependent with ML2PP version 6 ozone concentrations increasing with distance from the poles. The MIPAS temperature profiles show the best agreement with lidar in the polar regions, but the bias seems to be trending with altitude in the mid-latitudes and in the tropics, where this behaviour is strongest.

For the operational GOMOS version 6.01 ozone profiles, surprisingly the agreement with lidar had worsened relative to version 5.00, with version 6.01 presenting lower ozone concentrations, especially for bright and twilight conditions, and to a smaller extent for dark limb conditions and also more pronounced for cool stars. The number of outliers did reduce, in particular below 25 km with dark limb conditions. For the high resolution temperature profiles, an improved quality was observed for version 6.01, with more successful retrievals, a reduction in the number of outliers and a better agreement with lidar and sonde observations.

For $NO_2$ profiles it was shown that variation is naturally very large in both the spatial and temporal domains, and that assuming a wrong profile shape can lead to a very large error in the tropospheric column.