

THE MULTI-TASTE VALIDATION SYSTEM: TASTING THE EVOLUTION OF REACTIVE AND GREENHOUSE GAS DATA PRODUCTS FROM ENVISAT AND THIRD PARTY MISSIONS

D. Hubert^(1,*), A. Keppens^(1,*), J.-C. Lambert⁽¹⁾, J. Granville⁽¹⁾, F. Hendrick⁽¹⁾, and T. Verhoelst⁽¹⁾

⁽¹⁾ *Belgian Institute for Space Aeronomy (BIRA-IASB), Belgium*

^(*) *Contacts: daan.hubert@aeronomie.be, arno.keppens@aeronomie.be*

ABSTRACT

Over the past two decades the Multi-TASTE validation system has proven its value in the characterisation and support to the development of atmospheric composition measurements by ESA's GOME, Envisat and Third Party Missions (TPMs). We give an overview of the capabilities and the latest results of this comprehensive, versatile and semi-operational system and address its relevance regarding the recommendations voiced at ATMOS 2012.

1. INTRODUCTION

ESA's Envisat platform is a significant contribution to the global atmosphere observing system. Envisat carries Global Monitoring by Occultation of Stars (GOMOS), Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), and Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY), plus other sensors observing the oceans, land and ice. The satellite was launched in April 2002 on a mid-morning polar orbit with helio-synchronous precession. In April 2012 the communication with the satellite was lost and the end of mission was declared shortly after. GOMOS, MIPAS and SCIAMACHY have enhanced significantly the European capabilities in monitoring atmospheric composition initiated in 1995 with ESA's ERS-2 Global Ozone Monitoring Experiment (GOME), by measuring a variety of parameters from the ground up to the mesosphere, including O₃ and ozone-related species, greenhouse and climate related gases, air quality indicators, and aerosols of various types.

Before being used for scientific and operational applications, the fitness-for-purpose of satellite measurements must be verified carefully by means of geophysical validation studies. Their primary purpose is to derive appropriate quality indicators for a majority of potential users, so that they can decide if the data provided by the satellite experiment respond to the spatial, temporal, and quality requirements specific to their intended application. In particular, uncertainties and the geophysical consistency of these data must be assessed over the whole relevant spatial domain and vertical range, during the entire mission. The outcome

of more dedicated validation studies is also valuable for teams developing the retrieval algorithms and proper feedback must be given to them.

We therefore present the Multi-TASTE system, an adaptable and semi-operational system to assess the data quality of satellite trace gas retrievals. Versatility and traceability have been key drivers for the design of the system, ever since its inception during the commissioning phase of ERS-2 GOME in the mid 1990s. Over the years it has been continuously expanded and improved to identify ever more subtle features in ever more mature satellite data sets. By now, it is capable of providing a suite of global, long-term validation analyses for column and/or vertical profile products of numerous reactive gases, greenhouse gases, temperature and water vapour. The adopted harmonised analysis framework furthermore allows to study the quality of subsequent data releases as well as the mutual consistency of different satellite missions. Multi-TASTE analyses have contributed to the evolution and the characterisation of GOME, Envisat and TPM products in the context of the respective Quality Working Groups and ESA's Climate Change Initiative. The system is currently being prepared for upcoming challenges, e.g. adaptations to support the QA4ECV framework and guidelines, and the operationalization of the data processing flow for the Sentinel-5p TROPOMI mission.

Below, we give a brief description of the different components of the Multi-TASTE system, ranging from data content studies to averaging-kernel based information content analyses to correlative studies using ground-based network observations. Keppens et al. (2015) provide exhaustive details of the system and an application to nadir O₃ profile retrievals from GOME-2 radiance spectra. Here, we illustrate Multi-TASTE's capabilities using the latest versions of the trace gas products by the Envisat instruments.

2. MULTI-TASTE: A VERSATILE VALIDATION SYSTEM

A simplified overview of the Multi-TASTE validation system is shown in Fig. 1. The default operating mode is

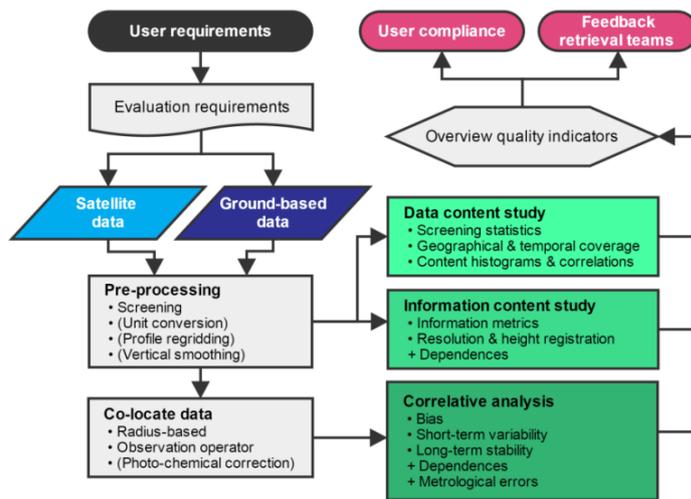


Figure 1. Simplified schematic of the versatile Multi-TASTE validation system. The different components are described in further detail in Keppens et al. (2015), Hubert et al. (2015) and Verhoelst et al. (2015).

based on community-agreed validation protocols and practices. However, the system can easily be tailored for applications with more specific demands since more advanced tools are selectable for most components in the analysis chain. The optimal analysis set-up depends on the evaluation requirements which, in turn, are a result of the inspection and interpretation of the requirements by users, be it retrieval experts or “external” data users. The initial analysis design may then be further refined if unforeseen quality issues emerge from intermediate validation results.

2.1. DATA PRE-PROCESSING

The manipulation of the data under investigation, the satellite records, is ideally kept to an absolute minimum. While the screening of unreliable data is always part of the satellite data flow, that is generally not the case for unit conversions, or for smoothing and regridding operations on the vertical profiles. The latter three pre-processing steps are typically only applied to the ground-based data in the correlative analyses.

Unit conversions are required when the native unit of the (ground-based) reference data differs from that of the satellite record. In that case auxiliary data (pressure, temperature) are exploited to convert e.g. between altitude and pressure or VMR and number density (Keppens et al., 2014). The metrological uncertainty in the comparison results due to differences in vertical smoothing is eliminated by downgrading the higher-resolution record to the lower-resolution record, in most cases the satellite. This is accompanied by a regridding operation on the correlative profiles when the levels of the vertical grid do not coincide with that of the satellite record. Several interpolation methods can be tested.

2.2. DATA CONTENT STUDY

The data content study provides an overview of the general characteristics of a data set. It includes (a) statistics

before and after screening, (b) histograms and time series of the retrieved data and the metadata relevant for the retrieval, and (c) geographical and temporal coverage maps of the satellite products (Fig. 2, left).

2.3. INFORMATION CONTENT STUDY

Retrieved quantities are mixes of a priori constraints and of information that is contributed by the satellite measurement. The vertical averaging kernel matrix characterizes how the retrieval system either smoothes or amplifies departures of the true profile from the prior profile. A study of the algebraic properties of the vertical averaging kernel matrices helps in understanding how the system captures actual atmospheric signals. This type of studies is particularly insightful for nadir profile retrievals.

The Multi-TASTE system computes a variety of diagnostic parameters and quality indicators from the averaging kernels. Several complementary metrics are implemented for information content, ranging from the well known degrees of freedom of signal (DFS), to Fisher and Shannon information content and more recently proposed metrics. Other important indicators include vertical sensitivity, vertical resolution and height registration. The dependence with geophysical parameters such as altitude, latitude, solar zenith angle, etc... is studied for all of these (Fig. 2, right).

2.4. CORRELATIVE ANALYSIS

Correlative analyses are an integral part of most Multi-TASTE validation exercises. These are based on the comparison of the satellite record to independent observations by ground-based networks, such as WMO’s GAW, NDACC and SHADOZ. The networks are composed of numerous stations, scattered around the globe, offering a suite of complementary measurement techniques. The Multi-TASTE system handles data from Dobson and Brewer UV spectrophotometers,

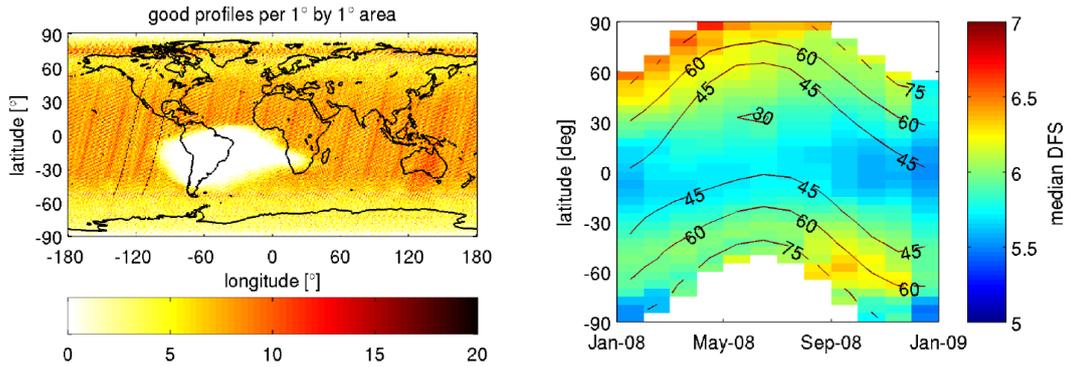


Figure 2. Illustration of information content study for GOME-2 nadir O₃ profile data by RAL v2.1 (2008 data). Left: Coverage of screened data. Right: Latitude-time cross section of median DFS. See Keppens et al. (2015).

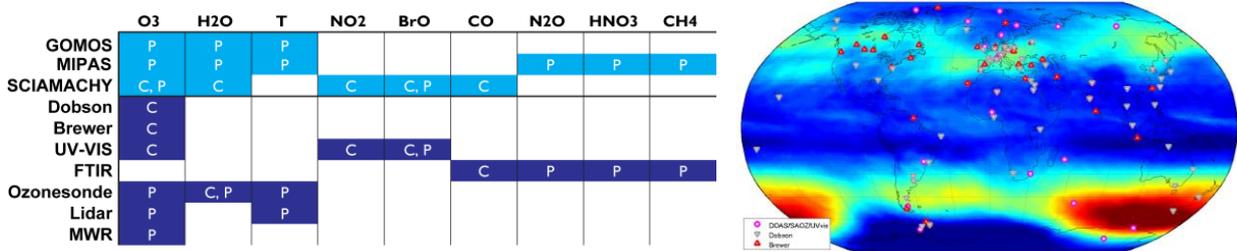


Figure 3. Left: Envisat trace gas products (light blue) validated by the Multi-TASTE system (C: column; P: vertical profile); and ground-based instruments considered for the correlative analysis (dark blue). Right: Geographical distribution of ground-based stations that provided O₃ vertical column data to the NDACC DHF and the WOUDC in overlap with the Envisat mission, on top of the SCIAMACHY mean total O₃ column field of September 2002.

DOAS UV-visible spectrometers, FTIR spectrometers, ozonesondes, stratospheric lidars and microwave radiometers. Together these instruments give access to a whole range of trace gases over most of the atmosphere. Fig. 3 (left) shows the capabilities of the ground-based networks relevant for the validation of a selection of Envisat trace gas products. The good spatial coverage of e.g. the total O₃ column networks allows to sample most atmospheric regimes (Fig. 3, right).

After screening, satellite-ground co-locations are identified. The optimal co-location criteria are a trade-off between mismatch uncertainty and statistical uncertainty, which in turn depend respectively on the natural variability of the species and the amount of data available. For some studies traditional radius-based criteria are sufficient, for other studies more advanced tools are more suitable. E.g. zenith-sky SAOZ spectrometer measurements are sensitive to air masses with a considerable spatial extension at low solar zenith angles. Since this region is typically larger than the pixel footprints of the satellite instrument being validated, the co-location is based on observation operators, which characterize the actual region of sensitivity of the SAOZ instruments.

Other potential sources of co-location mismatch uncertainty are diurnal cycles, e.g. for NO₂, BrO and O₃. To this end, a photo-chemical correction was

developed to correct the measurement time of ground-based NO₂ and BrO data to that of the satellite observation. For the O₃ analysis a correction scheme is currently under development for the comparison studies in the upper stratosphere and mesosphere.

After pre-processing, co-location and (when applicable) correction of the data sets various data quality indicators are derived from the absolute and/or relative difference time series. These include overall bias, comparison spread and long-term stability, which can then all be studied as a function of the relevant geophysical (location, clouds, SZA, ...) or instrumental/retrieval parameters (measurement mode, chi², ...). Each analysis starts off at the level of single stations, and results are aggregated to larger scales when they are sufficiently comparable between stations.

Multi-TASTE relies on robust statistical techniques to limit the influence of outliers on the final conclusions. This is especially important for the estimation of the long-term stability of the records (Fig. 4), since traditional methods are quite sensitive to spurious data. Similarly, inhomogeneity across the ground network is taken into account, e.g. for the limb/occultation O₃ profile analyses. Ad-hoc approaches have been developed to incorporate the uncertainty due to the differences between stations (Hubert et al., 2015).

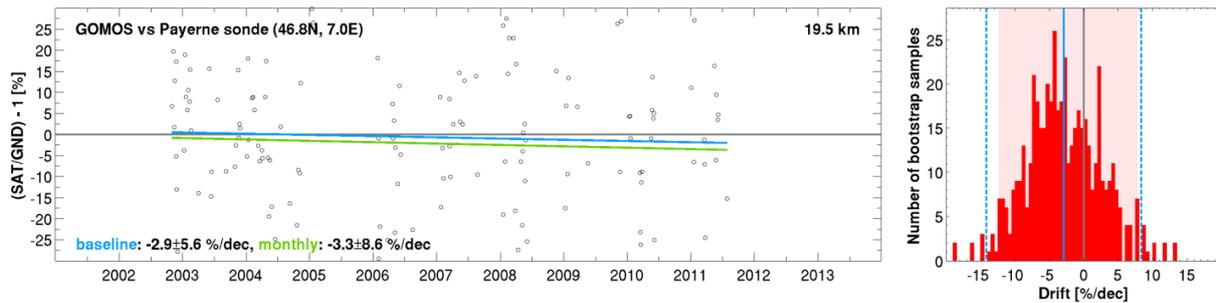


Figure 4. Illustration of a robust time series analysis of the co-located differences of GOMOS IPF 6 O₃ profile data and observations by the MeteoSwiss ozonesonde at Payerne, Switzerland. See Hubert et al. (2015).

2.5. OSSSMOSE

A proper interpretation of validation results requires a corresponding effort to understand the comparison error budget, including uncertainties associated with the comparison metrology: spatial and temporal mismatch in presence of atmospheric gradients and variability, differences in horizontal and vertical smoothing of atmospheric inhomogeneities and structures, and differences in pseudo-global sampling of patterns and cycles. To this end, the Multi-TASTE system includes a simulator of atmospheric remote sensing systems and their metrology, OSSSMOSE (Observing System of Systems Simulator for Multi-mission Synergies Exploration).

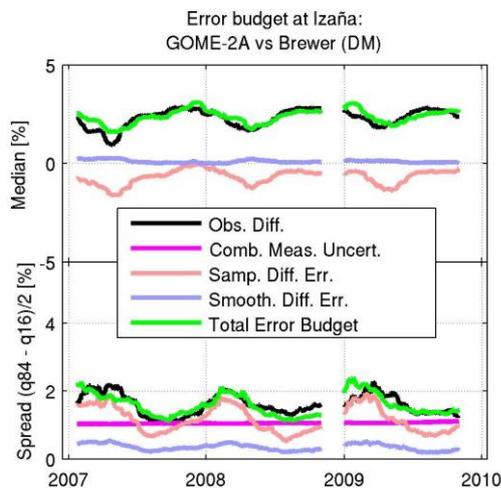


Figure 5. Comparison error budget closure for total O₃ column validation. Taking into account comparison errors due to spatio-temporal sampling and smoothing differences between the satellite and the ground-based reference measurements, the median (top) and spread (bottom) of the O₃ differences can be brought into agreement with the measurement uncertainties.

See Verhoelst et al. (2015).

Its architecture consists in the generation of multi-dimensional observation operators set up by the metadata of existing observing systems, followed by the application of those observation operators onto high-

resolution atmospheric fields. In this way, the system quantifies smoothing and sampling errors associated with a list of remote sensing measurements of atmospheric composition. The system can also model the expected differences between the various measurement types due to differences in sampling and smoothing of atmospheric structures. A successful application on total O₃ column comparisons (Fig. 5) was recently performed by Verhoelst et al. (2015).

2.6. FEEDBACK TO USERS AND RETRIEVAL TEAMS

Once the analysis is finalized, the Multi-TASTE validation reports are tailored to the particular needs of the reader. Retrieval development teams benefit most from e.g. reports of delta-validation exercises, focusing on the differences between incremental processor versions, besides other studies that test e.g. the impact of different screening procedures on final data quality. General users on the other hand are more interested in a detailed characterization of the final data product, with a comprehensive overview of the spatio-temporal coverage of the data set, and of the estimated uncertainties in different parts of the atmosphere and under different measurement conditions. Doing so allows them to easily identify whether or not the data set fits the purpose of their intended application.

3. EVOLUTION OF ENVISAT ATMOSPHERIC TRACE GAS PRODUCTS

Fig. 3 (left) overviews the trace gas products from Envisat's offline operational processors that were validated with the Multi-TASTE system so far. Fig. 6 illustrates the evolution of these processors from the launch of the Envisat platform in 2002 into the post-flight phase of the satellite. About three to four development cycles were carried out over the past years. Coloured gradients depict the phase of prototype development, which is typically followed by the release of a partial diagnostic data set to the validation teams (thin black vertical lines). Subsequently, this (or a slightly updated) version of the prototype is accepted as the operational processor and from then on produces

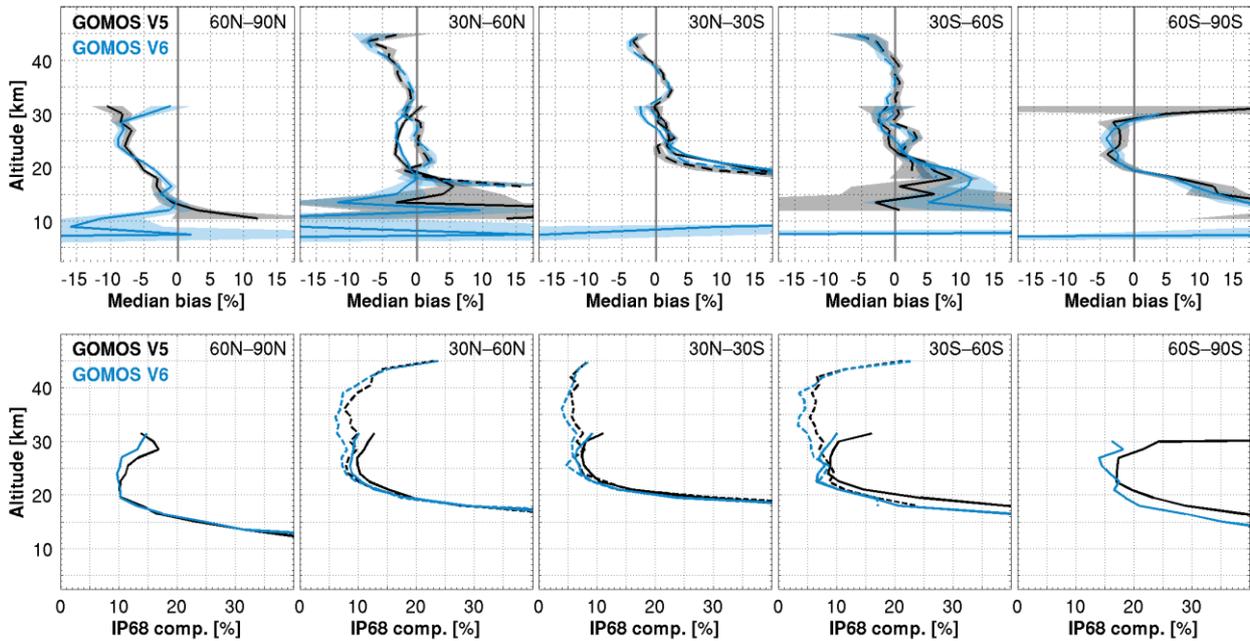


Figure 7. Latest results for GOMOS IPF V5 and V6 O3 profile as derived from comparisons to NDACC/GAW/SHADOZ ozonesonde (solid lines) and NDACC stratospheric lidar networks (dashed lines). Top: median and its 1 σ statistical uncertainty; bottom: comparison spread.

data continuously (solid coloured bars) until the switch to the next version is approved. In some instances, minor processor updates were performed e.g. to overcome data production issues or to incorporate minor quality improvements (white vertical lines). The thinner, horizontal black bars indicate the temporal coverage of each data set.

In the next section we focus on the validation results of GOMOS IPF 5 & 6, MIPAS IPF 5 & ML2PP 6, and SCIAMACHY SGP 3 & 5. In addition, we show the preliminary results of the delta-validation analyses on partial data sets from the forthcoming MIPAS ML2PP 7 and SCIAMACHY SGP 6 processors. The publicly released data set may be produced by an updated version of the current prototype and hence the data quality may differ from what is presented here.

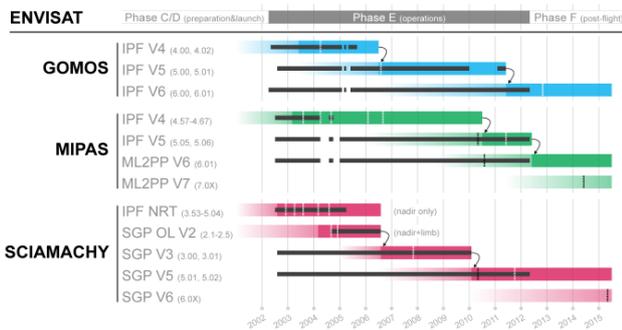


Figure 6. Evolution of the operational offline processors of Envisat's atmospheric composition instruments.

4. SELECTION OF RECENT ENVISAT VALIDATION RESULTS

We summarize a few results of our recent analyses of the operational Envisat products. A more detailed description of the validation of the current and previous processor versions can be found in Hubert et al. (2012). Initial and more complete results for the forthcoming processor versions will appear in the final report of the Multi-TASTE Phase-F project, due fall 2015 (available on ESA's SPPA portal).

4.1. GOMOS

The quality of the current O3 profile product is similar to that of the previous data release. Differences in bias between V5 and V6 remain less than 1-2% (Fig. 7, top), the bias relative to ground-based ozonesonde and lidar is less than 3-5% over the entire stratosphere except in the Arctic (5-10%). The most important change for V6 O3 is the short-term variability, which has reduced by a few percent due to the more refined screening procedure for outliers (Fig. 7, bottom).

4.2. MIPAS

Our delta-validation studies indicated that the prototype V7 temperature data increase over time relative to previous versions V5 and V6 (Fig. 8, top left). The change in temperature trend is most pronounced in the upper stratosphere (above the 5hPa level) and most likely due to the more refined Level-1 calibration scheme for the non-linearity of the detectors. This

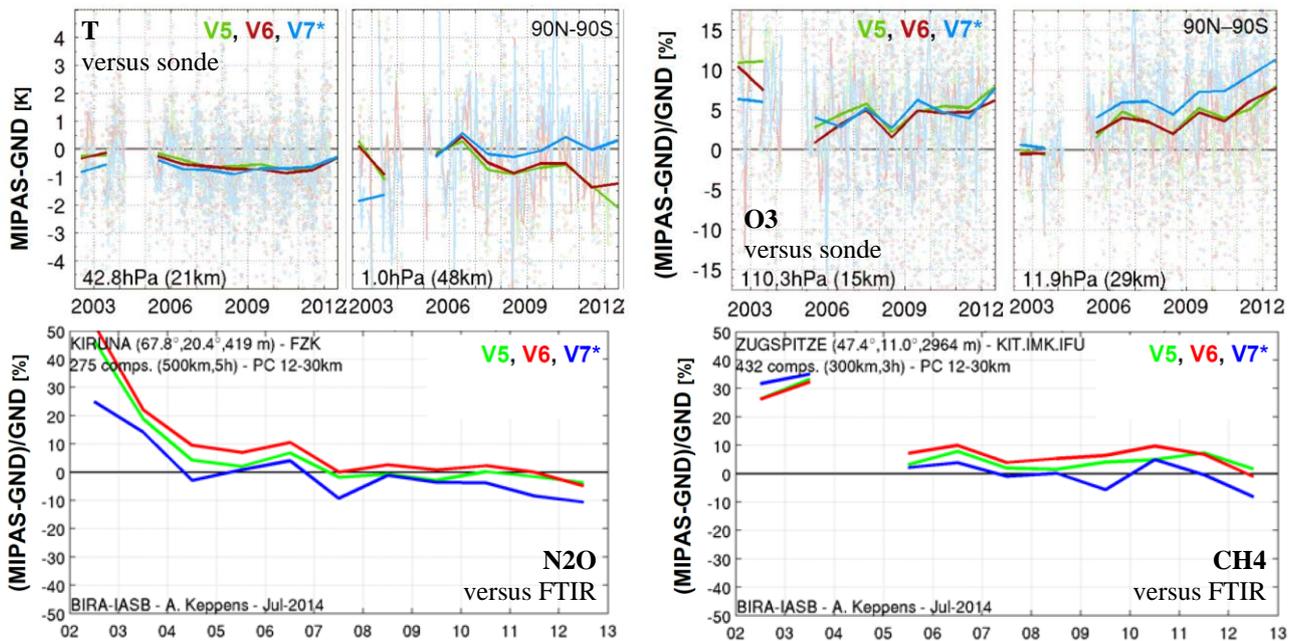


Figure 8. Preliminary delta-validation results for MIPAS IPF 5, ML2PP 6 & 7. Comparison time series at two altitudes for temperature (top left) and O₃ (top right) relative to the sonde/lidar networks; and partial columns 12-30km of N₂O (bottom left) and CH₄ (bottom right) relative to two FTIR instruments.

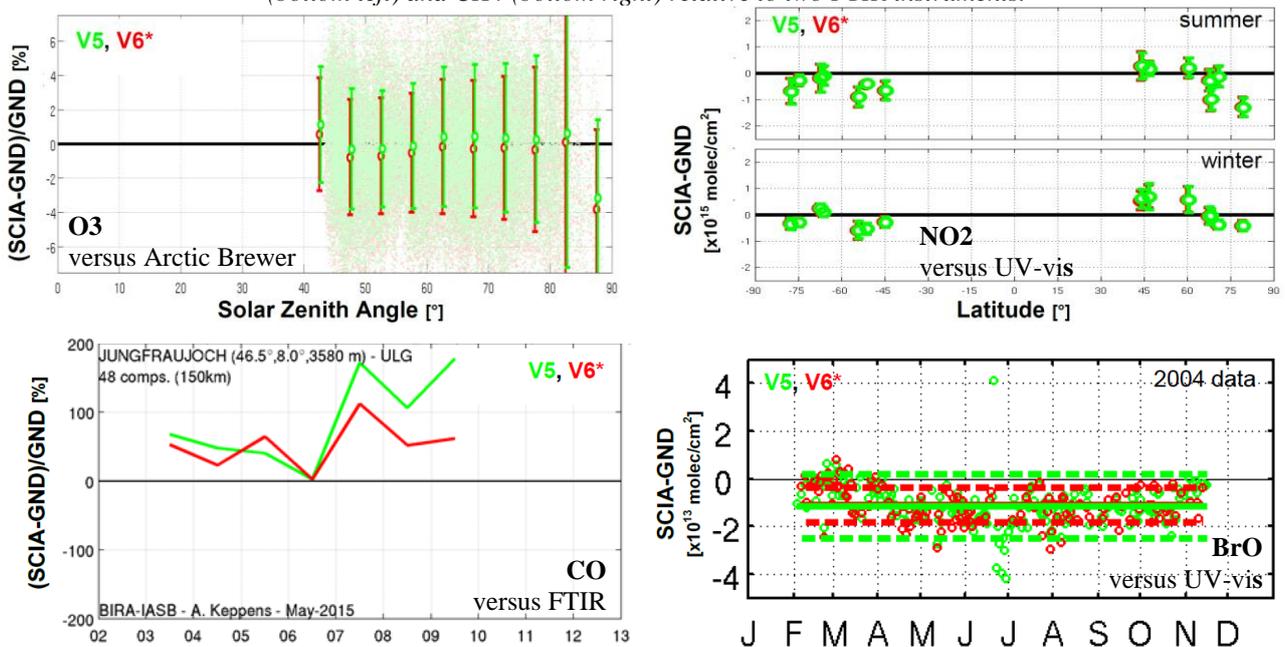


Figure 9. Preliminary delta-validation results for SCIAMACHY SGP 5 & 6. Dependence on SZA of total column comparisons of O₃ relative to Arctic Brewers (top left) and on latitude & season of NO₂ total columns relative to UV-visible spectrometers (top right); comparison time series of monthly averaged CO column data relative to FTIR (bottom left) and BrO column data (2004) relative to the Harestua UV-visible spectrometer.

change may also be the cause of the more negative temperature bias relative to sonde and lidar in the first part of the mission. The V7 O₃ data is 1-2% larger than V5 & V6 in the middle stratosphere, increasing the positive bias relative to ground-based data (Fig. 8, top right). The V7 change in temperature trend may cause a change in O₃ trend as well (towards more positive

values). It is not yet clear to what extent, probably less than 1-3% per decade.

The CH₄, HNO₃ and N₂O profiles have been validated for two fixed sub-columns with approximately unit DFS. These partial columns are obtained by integration of the MIPAS and ground-based FTIR reference

profiles between 9-12km and between 12-30km. Despite the fact that only four FTIR stations provided sufficient statistics for each species, some general observations can be made for the delta-validation of V5-V6-V7. Version 7 CH₄ shows a bias reduction of a few percent with respect to previous versions, but at the cost of an increased spread of the same order (Fig. 8, bottom right). HNO₃ results are comparable for all three product versions, while for N₂O the latest processor shows a slightly reduced bias with an unchanged spread (Fig. 8, bottom left). Remarkably, the seasonal dependence of the bias reduces with increasing processing version for all three species.

4.3. SCIAMACHY

Both V5 and V6 nadir O₃ column data sets are generally consistent with GAW ground-based data records. The V6 O₃ columns are systematically smaller than V5 data, by about 0.5%. This is seen over the entire latitude range covered by the Dobson and Brewer networks, for all seasons and at different solar zenith angle (Fig. 9, top left). The NO₂ nadir column V6 data do not seem to have changed substantially relative to V5, the bias and variability are similar in all seasons (Fig. 9, top right). The large variability of CO nadir column data requires an analysis at least at monthly scales. For the V5-V6 delta-validation analysis monthly means of co-located SCIAMACHY and FTIR data are compared, from which yearly statistics are derived: The V6 bias thereby seems comparable to V5, perhaps even slightly reduced (Fig. 9, bottom left). No significant change is observed for the spread. Preliminary analyses of BrO nadir column data relative to the Harestua UV-visible spectrometer indicate a similar bias of -12%, but especially less outliers in the V6 data set compared to V5 (Fig. 9, bottom right).

A considerable evolution in bias and short-term variability was noted between V3 and V5 limb O₃ profiles, with the latter exhibiting 10-20% more stratospheric O₃ and at least 10% more noise in the Arctic. Initial delta-validation results show less pronounced changes for the new processor prototype. V6 limb O₃ is 2-5% smaller than V5 below 35km, and the noise in most of the upper stratosphere is reduced by about 5%. While this leads to a higher-quality O₃ profile product, a distinctive positive bias remains and Arctic data remains quite noisy.

5. OVERVIEW OF OTHER STUDIES

Various other data sets were assessed with the Multi-TASTE system as well. A prime example is the total O₃ column record (TOC) collected by GOME (on-board ERS-2) and the GOME-2 series (on-board MetOp), see e.g. Koukouli et al. (2015). The GOME-SCIAMACHY-GOME-2 TOC series was also validated extensively

within the scope of ESA's Climate Change Initiative (Ozone_cci). Other major contributions to the assessment of O₃ data products developed within Ozone_cci include nadir profiles (GOME, SCIAMACHY, GOME-2, ...) and limb/occultation profiles (GOMOS, MIPAS, SCIAMACHY, OSIRIS, SMR and ACE-FTS). One of the key advantages of the Multi-TASTE system for this kind of validation applications is its ability to perform analyses of multiple satellite missions in a single analysis and software framework. This allowed us to verify the compliance of each data product with the requirements drafted by climate research and modelling groups in a similar fashion. Results were reported and published, e.g. Lambert et al. (2014), Keppens et al. (2015) and Laeng et al. (2015).

Another advantage of a harmonized analysis approach is the ability to assess the mutual consistency of the data quality of multiple missions. An extensive assessment was done of fourteen limb/occultation O₃ profile records (Hubert et al., 2015), in support of ongoing merging and trend analysis activities within the WMO's 2014 Ozone Assessment and the SPARC/IO3C/IGACO-O3/NDACC initiative (Harris et al., 2015). The availability of an ensemble of satellite records allowed us in turn to evaluate the homogeneity and internal consistency of ground-based networks. Our analyses of the ozonesonde networks (Hubert et al., 2014) contribute directly to the Ozonesonde Data Quality Assessment which is currently revisiting the ozonesonde data record to reduce residual inhomogeneities in space and time (Smit et al., 2012).

Recently, extensive simulations of the OSSSMOSE subsystem showed that the error budget of total O₃ column comparisons can be explained completely (Verhoelst et al., 2015). After taking into account measurement uncertainties, the residual observed small-scale temporal features in bias and spread could be ascribed to differences in sampling and smoothing of the variable atmospheric O₃ field (Fig. 5).

The Multi-TASTE system can be used to optimize a limited list satellite orbits to be processed and prepared for delta-validation exercises. Such *diagnostic data sets* of a few thousand orbits were prepared for GOMOS, MIPAS and SCIAMACHY (Fig. 10). These were designed to allow representative analyses for most geophysical and instrumental states for a maximal number of species.

6. DEVELOPMENTS IN VIEW OF FUTURE MISSIONS

The Multi-TASTE validation system is currently being prepared for upcoming challenges. Adaptations are performed to support the QA4ECV framework and guidelines, which will lead to understandable and fully

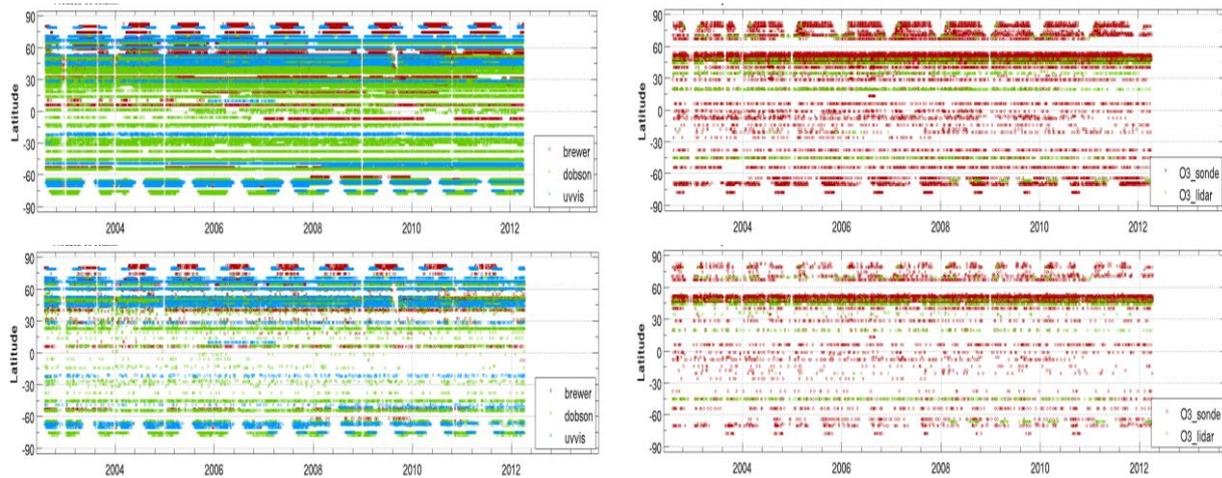


Figure 10. Latitude-time cross section of the co-locations of SCIAMACHY with ground-based data for the complete data set (top row) and an optimized diagnostic data set of 5000 orbits (bottom row). Shown are the nadir total O₃ column (left) and limb O₃ profile (right) and different correlative instrument types (marker colours) operated within the NDACC/GAW/SHADOZ ground-based networks.

traceable quality information for the satellite data used by climate and air quality services. Also, the data processing flow is being operationalized and optimized to handle the increased data volumes of TROPOMI and Sentinels-4 and 5 in order to provide initial feedback in close to near-real time mode.

We also keep track of current developments by ground-based networks such as NDACC. Additional species will be measured in the UV-visible, IR and MW ranges. The validation of the geostationary Sentinel-4 products will benefit from the enhanced ground-based measurement capabilities; e.g. an increased sampling of the diurnal cycle, a higher spatial resolution and an extension to moderate-large SZA for some species. Other aspects that are currently addressed are that of sustainability, long term stability, network homogeneity, and traceability.

7. REFERENCES

- Harris, N. R. P., et al. (2015). Past changes in the vertical distribution of ozone – Part 3: Analysis and interpretation of trends, *Atmos. Chem. Phys. Discuss.* **15**, 8565-8608.
- Hubert, D., et al. (2012). *Multi-TASTE Final Report / October 2008 – October 2011*, Tech. rep. TN-BIRA-IASB-MultiTASTE-FR, pp157, https://earth.esa.int/documents/700255/707912/MultiTASTE_AnnualReport_2_2010.pdf.
- Hubert, D., et al. (2014). *The internal consistency of the ozonesonde network data archive assessed through comparisons with satellite ozone profilers*, presented at 2014 EUMETSAT Meteorological Satellite Conference, Geneva, Switzerland.
- Hubert, D., et al. (2015). Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records, *Atmos. Meas. Tech. Disc.* **8**, 6661–6757.
- Keppens, A., et al. (2014). *MIPAS Technical Note: Methodology for switching among MIPAS vertical averaging kernel matrix representations*, Tech. rep. TN_BIRA-IASB_MIPAS-AKM-conversions_Keppens, Issue 1 / Rev. 2, BIRA-IASB, pp7.
- Keppens, A., et al. (2015). Round-robin evaluation of nadir ozone profile retrievals: methodology and application to MetOp-A GOME-2, *Atmos. Meas. Tech.* **8**, 2093-2120.
- Koukouli, M. E., et al. (2015). Evaluating a new homogeneous total ozone climate data record from GOME/ERS-2, SCIAMACHY/Envisat and GOME-2/MetOp-A, *J. Geophys. Res.*, under review.
- Laeng, A., et al. (2015). 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.
- Lambert, J.-C., et al. (2014). *Ozone_cci - Product Validation and Intercomparison Report (PVIR)*, Tech. rep. 1.0, BIRA-IASB, pp115, http://www.esa-ozone-cci.org/?q=webfm_send/148.
- Smit, H. G. J., et al. (2012). *SI2N/O3S-DQA Activity: Guide Lines for Homogenization of Ozone Sonde Data*, Tech. rep. 2.0, pp48, http://www-das.uwyo.edu/~deshler/NDACC_O3Sondes/O3s_DQ_A/O3S-DQA-GuidelinesHomogenization-V2-19November2012.pdf.
- Verhoelst, T., et al. (2015). Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons, *Atmos. Meas. Tech. Disc.*, accepted.