MULTI-TASTE ASSESSMENT OF THE QUALITY AND EVOLUTION OF ENVISAT REACTIVE AND GREENHOUSE GAS DATA PRODUCTS

A. Keppens [(1)], D. Hubert [(1)], J. Granville [(1)], F. Hendrick [(1)], T. Verhoelst [(1)], J.-C. Lambert [(1)]

[(1)] Royal Belgian Institute for Space Aeronomy, Ave. Circulaire 3, 1180 Brussels, Belgium
Email: arno.keppens@aeronomie.be, daan.hubert@aeronomie.be

ABSTRACT

This work reports on the use of the Multi-TASTE QA/validation system in support to the continuous evolution of operational Envisat data products on ozone, greenhouse gases, and temperature. The focus is on the latest upgrades of GOMOS (IPF V5 to IPF V6), MIPAS (IPF V5 to ML2PP V7), and SCIAMACHY (SGP V3 to SGP V6) processors. The studies conclude with altitude and latitude-resolved estimates of bias, spread, and long-term stability of those latest versions.

1. INTRODUCTION

For ten years (2002-2012), ESA's Envisat satellite has provided an important contribution to the global measurement system for atmospheric composition. Its instruments GOMOS, MIPAS, and SCIAMACHY have measured the atmospheric abundance of a variety of trace gases and parameters, including reactive and greenhouse gases. The development of retrieval algorithms used for the derivation of atmospheric abundances from the Envisat-based measured spectra is in permanent evolution and continues during the current post-flight Phase F. This evolution of Envisat data products requires appropriate independent data quality assessments and validations. In particular, the uncertainties and geophysical consistency of the data must be assessed for the wider range of atmospheric states and over the relevant spatial domain, vertical range, and mission lifetime. Every upgrade of the data products and associated processors must be verified through delta-validation studies of the expected improvement. The outcome of delta-validation studies of successive data products furthermore provides highly valuable feedback to the data retrieval teams.

Pioneered in the nineties with the aim to study the mutual consistency between ESA's ERS-2 GOME and NASA's TOMS ozone monitoring instruments, a satellite validation system with multi-mission capacities has been developed at BIRA-IASB over the last two decades. This multi-mission QA/validation system for satellite atmospheric data, which is here referred to as Multi-TASTE system, builds upon state-of-the-art and community-agreed validation protocols and practices, implemented on a generic validation chain. Specific tools and methods include multi-dimensional observation operators for tailored co-location criteria, harmonised unit and representativeness conversions, information content analysers, and a range of statistical techniques to derive robust estimates of bias, spread, and long-term stability of the atmospheric data records. The validation chain comprises comparisons with reference data collected by established measurement networks contributing to WMO's Global Atmosphere Watch (GAW), like the Network for the Detection of Atmospheric Composition Change (NDACC), and NASA's Southern Hemisphere Additional Ozonesondes (SHADOZ) programme. The versatility of data products and retrieval approaches has been a key driver for the development of Multi-TASTE since the beginning. By now, global and long-term validation analyses are possible for both column and vertical profile data products of numerous reactive gases, greenhouse gases and temperature. The system is currently being upgraded with support of the EC QA4ECV and GAIA-CLIM projects in view of upcoming challenges of the Copernicus Sentinel missions.

A brief description of the Multi-TASTE system is first given below, ranging from data content studies to information content analyses and correlative studies using ground-based network observations. The system is then applied for a delta-validation quality assessment of the latest version updates of the trace gas products by the Envisat instruments GOMOS (IPF V5 to IPF V6), MIPAS (IPF V5 to ML2PP V7), and SCIAMACHY (SGP V3 to SGP V6).

2. MULTI-TASTE: A VERSATILE VALIDATION SYSTEM

Keppens et al. provide exhaustive details on the Multi-TASTE validation system and its application to nadir O3 profile retrievals [1-2]. A simplified overview of the system is shown in Fig. 1. The default operating mode is based on community-agreed quality assessment 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.

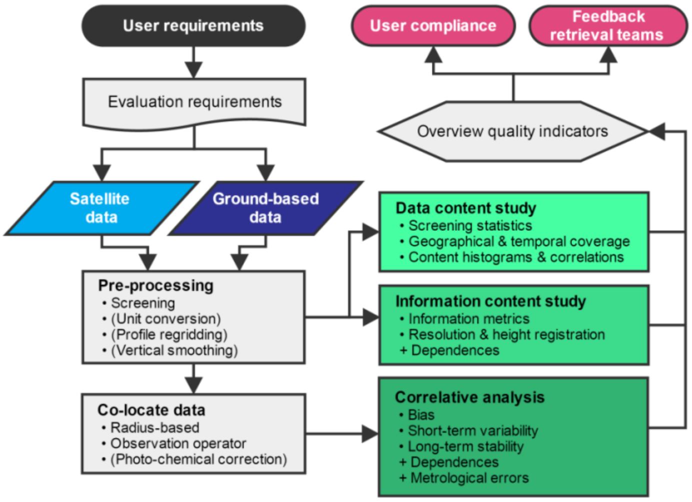

Figure 1. Simplified schematic of the versatile Multi-TASTE validation system. The different components are described in further detail in Keppens et al., Hubert et al., and Verhoelst et al. [1-5].

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 [6]. 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, 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 sampling most atmospheric regimes (Fig. 3, right).

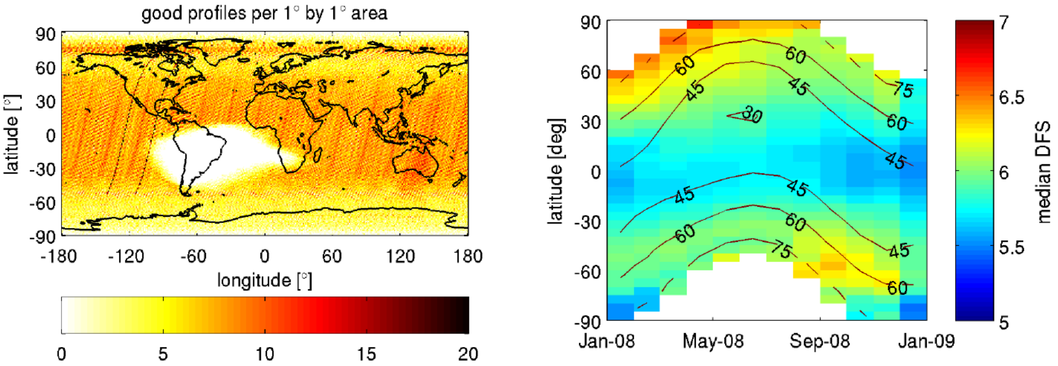

Figure 2. Illustration of information content study for GOME-2 nadir O3 profile data by RAL v2.1 (2008 data). Left: Coverage of screened data. Right: Latitude-time cross section of median DFS. See Keppens et al. [1].

	O3	H2O	T	NO2	BrO	CO	N2O	HNO3	CH4
GOMOS	P	P	P						
MIPAS	P	P	P				P	P	P
SCIAMACHY	C, P	C		C	C, P	C			
Dobson	C								
Brewer	C								
UV-VIS	C			C	C, P				
FTIR						C	P	P	P
Ozonesonde	P	C, P	P						
Lidar	P		P						
MWR	P								

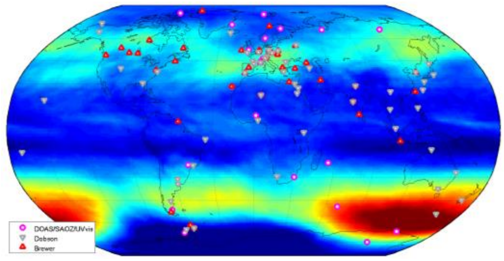

Figure 3. Left: Envisat trace gas products (light blue) validated by the Multi-TASTE system (C: column; P: vertical profile); and ground-based instruments considered for the correlative analysis (dark blue). Right: Geographical distribution of ground-based stations that provided O3 vertical column data 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.

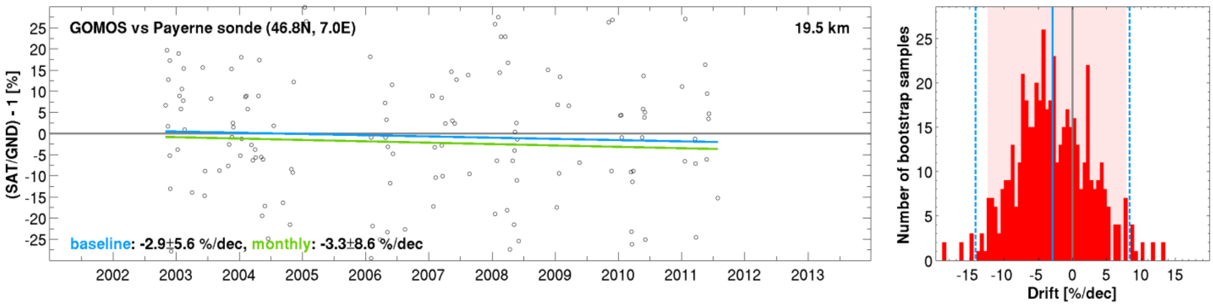

Figure 4. Illustration of a robust time series analysis of the co-located differences of GOMOS IPF 6 O3 profile data and observations by the MeteoSwiss ozonesonde at Payerne, Switzerland. See Hubert et al. [3].

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, …), instrumental, or 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 [3].

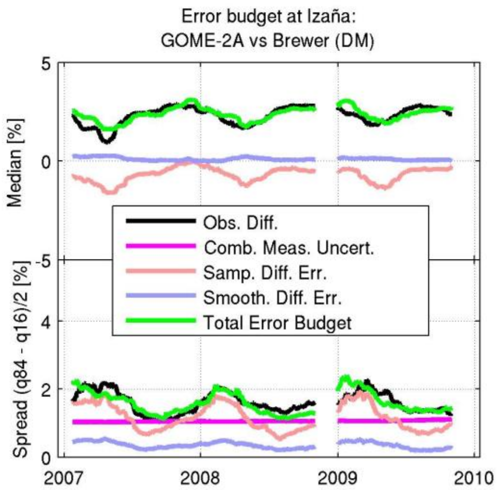

Figure 5. Comparison error budget closure for total O3 column validation. Taking into account comparison errors due to spatio-temporal sampling and smoothing differences between the satellite and the ground-based reference measurements, the median (top) and spread (bottom) of the O3 differences can be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

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).

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. [5].

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.

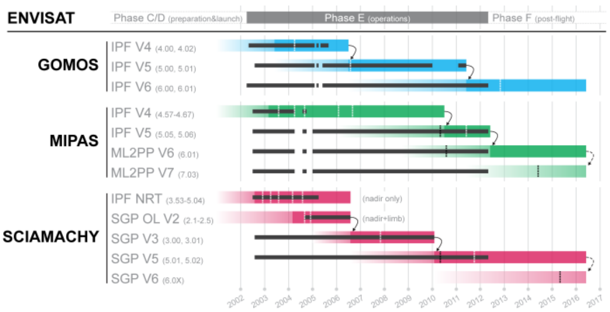

Figure 6. Evolution of the operational offline processors of Envisat's atmospheric composition instruments.

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 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

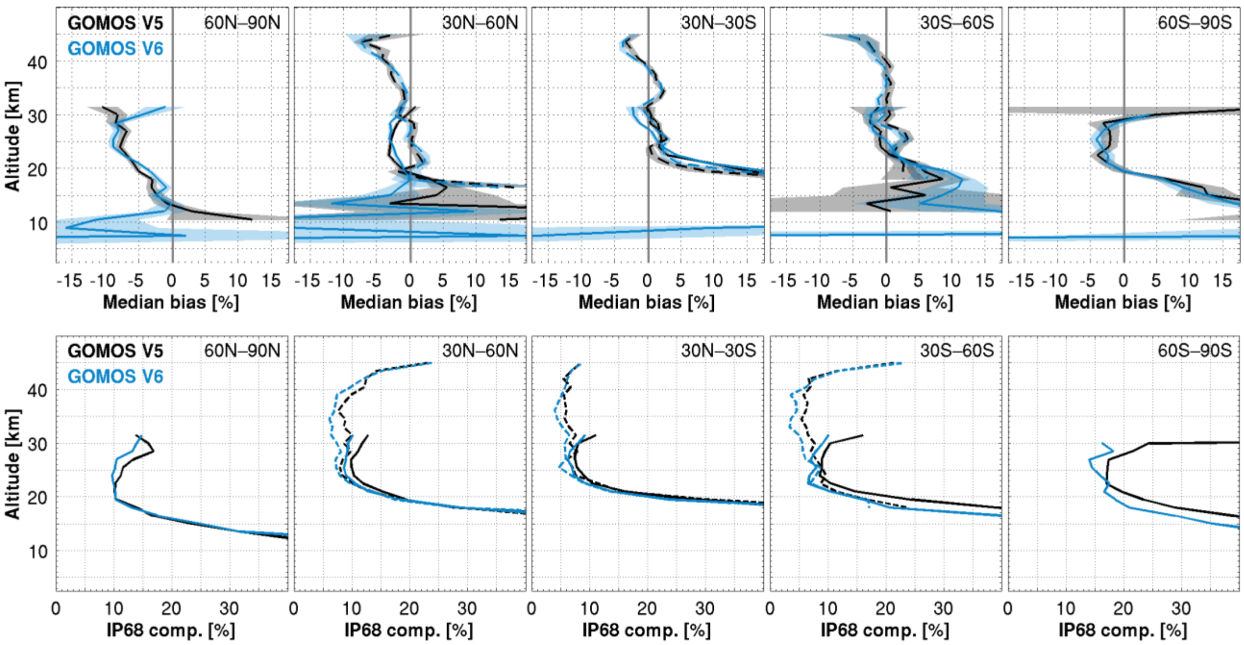

Figure 7. Latest results for GOMOS IPF V5 and V6 O3 profile as derived from comparisons to NDACC/GAW/SHADOZ ozonesonde (solid lines) and NDACC stratospheric lidar networks (dashed lines). Top: median and its 1σ statistical uncertainty; bottom: comparison spread.

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 V5 and V6, MIPAS IPF V5 and ML2PP V6, and SCIAMACHY SGP V3 and V5. In addition, we show the results of the delta-validation analyses on diagnostic datasets from the MIPAS ML2PP V7 and SCIAMACHY SGP V6 processors.

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 previous processor versions can be found in Hubert et al. [7]. Complete results for the full processor versions are collected in the final report of the Multi-TASTE Phase-F project [8].

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 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.

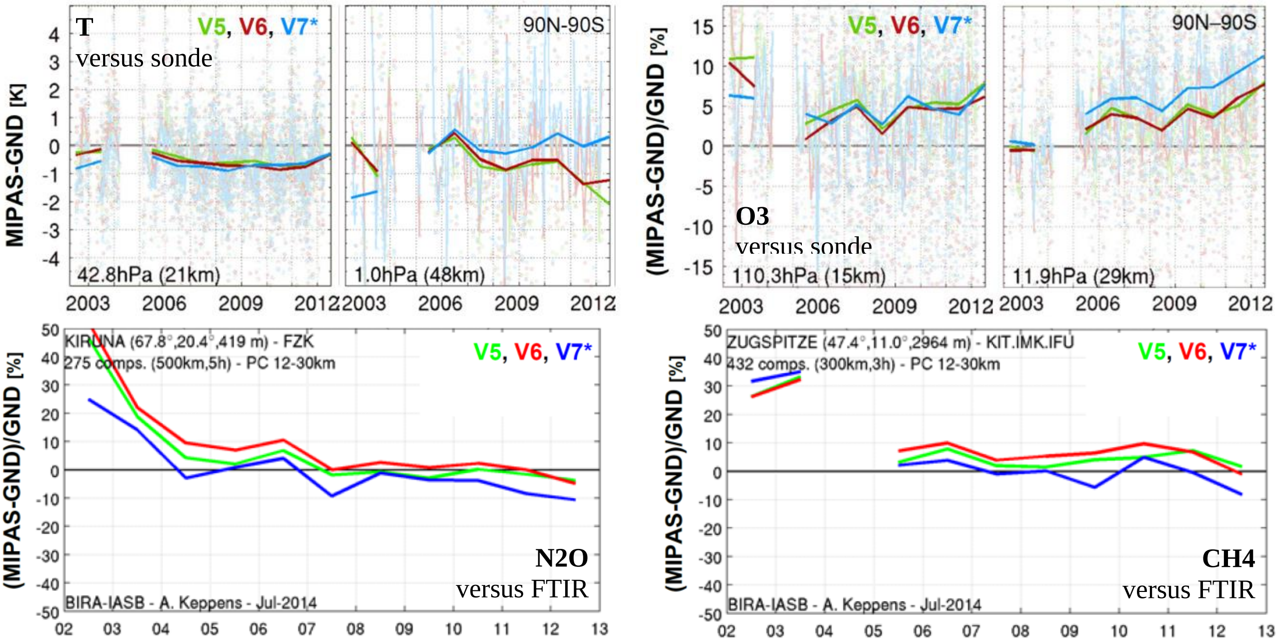

Figure 8. Preliminary delta-validation results for MIPAS IPF 5, ML2PP 6 & 7. Comparison time series at two altitudes for temperature (top left) and O3 (top right) relative to the sonde/lidar networks; and partial columns12-30km of N2O (bottom left) and CH4 (bottom right) relative to two FTIR instruments.

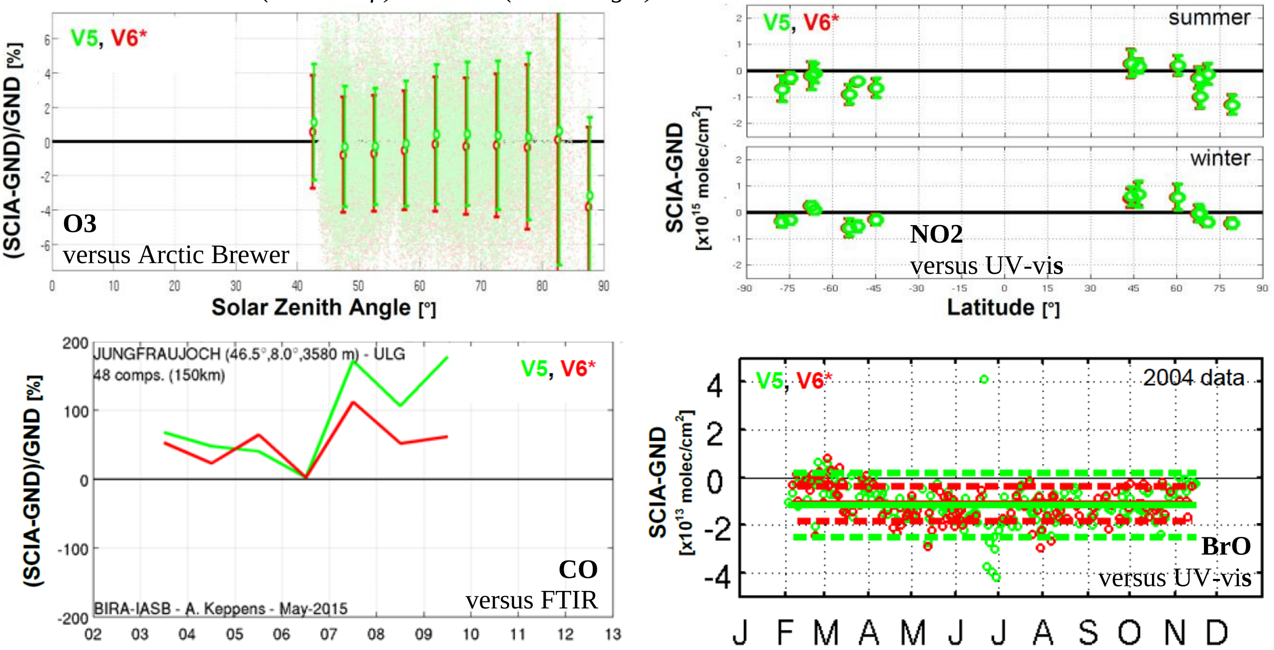

Figure 9. Preliminary delta-validation results for SCIAMACHY SGP 5 & 6. Dependence on SZA of total column comparisons of O3 relative to Arctic Brewers (top left) and on latitude & season of NO2 total columns relative to UV-visible spectrometers (top right); comparison time series of monthly averaged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

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 9-12 km and between 12 and 30 km. 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. 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 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.

6. REFERENCES

[1] 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.

[2] Keppens, A., et al. (2015). Harmonised Validation System for Tropospheric Ozone and Ozone Profile Retrievals from GOME to the Copernicus Sentinels, ESA SP-735(162).

[3] 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.

[4] Hubert, D., et al. (2015) The Multi-TASTE validation system: Tasting the evolution of reactive and greenhouse gas data products from Envisat and Third Party Missions, ESA SP-735(29).

[5] Verhoelst, T., et al. (2015). Metrology of ground-based satellite validation: Co-location mismatch and smoothing issues of total ozone comparisons, *Atmos. Meas. Tech.* **8**, 5039-5062.

[6] 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.

[7] 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.

[8] Hubert, D., et al. (2016). *Multi-TASTE Phase F Final Report / October 2013 – December 2015*, Tech. rep. TN-BIRA-IASB-MultiTASTE-Phase-F-FR-Iss2-RevA, pp87, https://earth.esa.int/documents/700255/702251/TN-BIRA-IASB-MultiTASTE-Phase-F-FR-Iss2-RevA.pdf