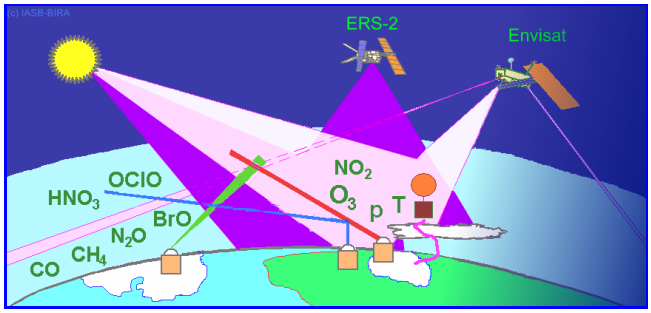
Multi-TASTE Phase F

Technical Assistance to ESA Multi-Mission Validation by Sounders, Spectrometers and Radiometers



Final report

October 2013 – December 2015

authors | auteurs D. Hubert, A. Keppens, J. Granville, F. Hendrick, J.-C. Lambert (BIRA-IASB), J.A.E. van Gijsel, P. Stammes (KNMI)

reference | référence TN-BIRA-IASB-MultiTASTE-Phase-F-FR-Iss2-RevA
date of issue | date d'édition 1 February 2016
issue | édition 2
revision | révision A
ESA contract Nr | contrat ESA No 42000021819/08/I-OL/CCN-1 & CCN-2

title | titre	Multi-TASTE Phase F Final Report / October 2013 – December 2015
reference | référence	TN-BIRA-IASB-MultiTASTE-Phase-F-FR-Iss2-RevA
date of issue | date d'édition	1 February 2016
issue | édition	2
revision | révision	A
status | état	Final draft
document type | type de document	Project Report
ESA contract Nr | contrat ESA No	42000021819/08/I-OL/CCN-1 & CCN-2
prepared by | préparé par	D. Hubert, A. Keppens, J. Granville, F. Hendrick, J.-C. Lambert (BIRA-IASB), J.A.E. van Gijsel, P. Stammes (KNMI)

document change record | historique du document

Issue	Rev.	Date	Section	Description of Change
1	A	12.12.2015	all	Creation of this document
	B	21.12.2015	all	Handling of first ESA feedback
2	A	01.02.2016	all	This issue

Executive summary

This document reports on the activities and results of Phase F extensions of the Multi-TASTE project (ESA Contract No 42000021819/08/I OL/CCN-1 and CCN-2), carried out by teams at BIRA-IASB and KNMI. The main objectives of the project were: (a) to support the full-mission validation of the existing GOMOS, MIPAS and SCIAMACHY Level-2 data sets, (b) to provide current QWG developments with appropriate validation support, (c) to validate future reprocessings, and (d) to consolidate and maintain the correlative data record over the entire Envisat mission lifetime. This executive summary highlights key achievements, while more detailed reports can be found in the main document. Tables which summarize main conclusions and estimates of the quality of Envisat atmospheric data products are provided in Section II.

WP 2 dealt with **data management**. Initially a considerable number of ground-based data records were uploaded and maintained at ESA's Validation data Centre (EVDC). Later on, after activation of a direct mirror between EVDC and the NDACC Data Host Facility (DHF), ESA and the project partners agreed that continuing this task had become irrelevant. The delivery of the Envisat data products by the QWGs to the validation teams experienced significant and unforeseen delays, which propagated several months of delay to the Multi-TASTE Phase F project. In support of current and future delta-validation activities for GOMOS, MIPAS and SCIAMACHY processor upgrades, for each instrument a subset of about 10% of the mission orbits were selected. These new Diagnostic Data Sets (DDS) provide superior spatio-temporal and product coverage than the previously used orbit lists, which led to very robust and representative results in the delta-validation analyses of the MIPAS ML2PP 7.01 and SCIAMACHY SGP 6.00 prototypes.

One of the main focuses of the **correlative analyses** in WP 3 was on the characterisation of and the evolution in the data quality of the existing operational Level-2 data sets of GOMOS (IPF 5.00/5.01 and 6.01), MIPAS (IPF 5.05/5.06 and ML2PP 6.0) and SCIAMACHY (SGP 3.01 and 5.02) for the entire mission. Another focus was to perform the delta-validation of the new Level-2 processors MIPAS (ML2PP 7.01) and SCIAMACHY (SGP 6.00) based on the diagnostics data sets. The KNMI subcontract also included a first look at the full reprocessing by the new MIPAS ML2PP 7.03 processor.

The correlative analysis of the operational **GOMOS** ozone and high resolution temperature profiles (IPF 5.00/5.01 and IPF 6.01) was consolidated during the project. We documented the dependence of the data quality on altitude, latitude and time, and the impact of different screening procedures. Both GOMOS processors are quite similar in terms of data quality. The IPF 6.01 ozone profiles seem a little more precise, with less outliers. There are clear signs that the ozone profile data record has a negative drift below 25 km, which makes the GOMOS data record possibly not suitable for long-term trend assessments in the UT/LS. Our comparison of GOMOS ozone from occultation (operational processor) and bright limb (scientific processor) measurements has furthermore indicated that this instability may not be due to an incomplete correction for the increase in dark charge over the mission.

The ground-based validation of the operational **MIPAS** Level-2 processors IPF 5.05/5.06 and ML2PP 6.0 was consolidated over the reporting period. In general we could not find significant differences in quality between both processors for the vertical profile records of temperature, O3, CH4, HNO3 and N2O. The diagnostic data set for the new MIPAS Level-2 prototype, ML2PP 7.01, was delivered in June 2014. This was followed by an intensive delta-validation effort, that included all products included in the validation contract. The ML2PP 7.01 data quality turned out to be similar as that of IPF 5.05/5.06 and ML2PP 6.0, and perhaps slightly better. Particular attention was given to the temperature profiles, which clearly differed from earlier versions. We found that the quality of V7

temperature data in the 2002-2004 period is worse than for V5 or V6, likely due to changes in the Level-2 set-up (microwindows, continuum, …). Also the 2005-2012 V7 temperature data changed clearly, with a positive trend relative earlier versions, attributed to changes in the Level-1b processor (non-linearity correction). This change, however, seems an improvement in quality in the upper stratosphere and lower mesosphere, though our assessment is inconclusive at lower altitudes. Our other correlative analyses did not find evidence that these changes in temperature would affect the quality of the trace gas products.

We also continued the ground-based assessment of the operational SCIAMACHY Level-2 data products from nadir (O3, NO2, CO, BrO and H2O total columns) and limb observations (O3 and BrO vertical profiles), retrieved by SGP 3.01 and SGP 5.02. Our analyses showed that a few operational products have an unsatisfactory performance (limb O3, limb BrO, nadir CO). We therefore recommend the QWG to consider further developments in the retrieval of these products. In April 2015 the validation teams received a diagnostic data set processed by the new Level-2 prototype, SGP 6.00. An intensive period of delta-validation studies followed, which showed that the SGP 6.00 data is of similar quality, if not slightly better (for a few products, in part of the atmosphere) as SGP 5.02. There was, however, one worrisome result: the appearance at northern mid-latitudes of a negative drift (about 1.5% over the mission lifetime) in the SGP 6.00 O3 column data relative to correlative observations, most likely related to the changes in the Level-1 processor. More detailed investigations of the new processor will be possible once the full mission has been reprocessed.

The validation results were provided to the QWG in the form of oral presentations and written validation reports. In addition we participated in the discussions on how to simplify the use of Envisat data products and contributed to the data product Readme files. The project results were also communicated to the community at numerous international conferences and workshops. A complete list of presentations, written reports and publications can be found in Annexes B and C.

Table of contents

Acronyms and abbreviations

AK	Averaging Kernel
AMF	Air Mass Factor
BIRA-IASB	Koninklijk Belgisch Instituut voor Ruimte-Aeronomie / Institut royal d'Aéronomie spatiale de Belgique (Royal Belgian Institute for Space Aeronomy)
CCN	Contract Change Notice
CF	Cloud fraction (i.e. fractional cloud cover)
CNRS/LATMOS	Centre National de la Recherche Scientifique / Laboratoire "Atmosphères, milieux, observations spatiales"
COT	Cloud Optical Thickness
CTH	Cloud Top Height
DDS	Diagnostic Data Set
DHF	NDACC Data Host Facility
DIAL	Differential Absorption Lidar
DOAS	Differential Optical Absorption Spectroscopy
Envisat	ESA's Environmental Satellite
ESA	European Space Agency / Agence spatiale européenne
EVDC	ESA's Validation Data Centre
FMI	Finnish Meteorological Institute
FR	Full Resolution period MIPAS (2002-2004)
FTIR	Fourier Transform Infra-Red spectroscopy/spectrometer
GAW	WMO's Global Atmosphere Watch
GBL	GOMOS Bright Limb
GOME	Global Ozone Monitoring Experiment
GOMOS	Global Ozone Monitoring by Occultation of Stars
HDF	Hierarchical Data Format
IPF	Instrument Processing Facility
IUP	Institut für Umweltphysik, University of Bremen
KNMI	Koninklijk Nederlands Meteorologisch Instituut
Lidar	Light Detection and Ranging
LS	Lower Stratosphere
LM	Lower Mesosphere
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
ML2PP	MIPAS Level-2 Prototype Processor
MS	Middle Stratosphere
Multi-TASTE	Technical Assistance to ESA Multi-Mission Validation by Sounders, Spectrometers and Radiometers
MWR	Millimetre Wave Radiometer
NDACC	Network for the Detection of Atmospheric Composition Change (formerly NDSC)
NH	Northern Hemisphere
NILU	Norwegian Institute for Air Research
OCRA	Optical Cloud Recognition Algorithm
OR	Optimized Resolution period MIPAS (2005-2012)
QWG	Quality Working Group
RH	Relative Humidity
SACURA	Semi-Analytical Cloud Retrieval Algorithm
SAOZ	Système d'Analyse par Observation Zénithale
SCD	Slant Column Density
SCIAMACHY	Scanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SCIAVALIG	SCIAMACHY VALidation and Interpretation Group
SGP	SCIAMACHY Ground Processor
SH	Southern Hemisphere

SHADOZ	Southern Hemisphere ADditional Ozonesondes
SOM	Self Organising Maps
SZA	Solar Zenith Angle
T	Temperature
TASTE	Technical Assistance to ESA Validation by Sounders, Spectrometers and Radiometers
US	Upper Stratosphere
UT	Upper Troposphere
UT/LS	Upper Troposphere / Upper Stratosphere
UV	Ultra Violet
UV-vis	Ultra Violet - Visible
VCD	Vertical Column Density
VMR	Volume Mixing Ratio
WMO	World Meteorological Organization

I Introduction

The focus of the Phase F extensions of the Multi-TASTE project, ESA Contract No 42000021819/08/I-OL/CCN-1 (Oct 2013 – Dec 2015) and CCN-2 (Dec 2014 – Dec 2015), was (a) on supporting the full-mission validation of the existing GOMOS, MIPAS and SCIAMACHY Level-2 data sets, (b) on providing current QWG developments with appropriate validation support, (c) on validating future reprocessings, and (d) on consolidating and maintaining the validation data record over the entire Envisat mission lifetime. This Final Report summarises the activities during the project, from October 2013 to December 2015.

We start this report with our (tabular) summary of the quality of the operational Envisat Level-2 atmospheric data products, based on the correlative analyses carried out during the project. Then, the document follows the Work Breakdown structure, where each section represents a work package: project management (WP 1, Section III), data management (WP 2, Section IV), correlative analyses (WP 3, Section V) and reporting and valorisation (WP 4, Section VI). The Annexes provide additional graphics (A) and detailed lists of our scientific output (B & C).

II Summary of Envisat Level-2 data product quality

Here we provide quantitative summaries of the quality of the latest version of the operational Envisat Level-2 data products. We report bias and spread estimates from comparisons to correlative data by a suite of ground-based instruments. Whenever relevant, the results are shown as a function of geophysical parameters (e.g. latitude, altitude, …). In all tables below, positive values for bias means that the Envisat product overestimates relative to ground-based measurements.

II.1 Correlative data and validation best practices

Trace gas data product	Correlative instruments	Validation best practices
O3 total column	Dobson, Brewer, DOAS/SAOZ	Lambert et al. (1999, 2000), Balis et al. (2007)
NO2 total column	UV-vis DOAS (incl. SAOZ)	Lambert et al. (2007), Celarier et al. (2008)
BrO total column	UV-vis DOAS	Hendrick et al. (2009)
CO total column	FTIR	de Laat et al. (2010)
H2O total column	radiosonde (integrated profile)	du Piesanie et al. (2013)
T vertical profile	radiosonde, T lidar	Similar to Hubert et al. (2015)
O3 vertical profile	ozonesonde, O3 lidar (DIAL), microwave radiometer	van Gijsel et al. (2010), Keppens et al. (2015), Hubert et al. (2015)
BrO vertical profile	UV-vis DOAS	Hendrick et al. (2009)
CH4 vertical profile	FTIR	Payan et al. (2009)
N2O vertical profile	FTIR	Vigouroux et al. (2007), Payan et al. (2009)
HNO3 vertical profile	FTIR	Vigouroux et al. (2007)

II.2 GOMOS

II.2.1 High resolution temperature vertical profile

Table 1: Absolute differences between GOMOS IPF 6.01 HRTP (full mission, full dark profiles) and radiosonde / T lidar. Positive bias values imply that GOMOS temperatures are warmer than correlative data.

IPF 6.01	Polar (60°-90°)	Mid-latitudes (30°-60°)	Tropics (<30°)
Median bias (K)			
15-20 km	-1	-1	0
20-25 km	0	+2	+1
25-30 km	-1	+2	+1
30-35 km	-2	0	-1
35-40 km		-7	
Comparison spread (K)			
15-20 km	7	6	9
20-25 km	11	6	6
25-30 km	10	7	5
30-35 km	9	9	7
35-40 km		6	
Comments			
• GOMOS screening procedure as described in IPF 6.01 Readme file; only full dark occultations. • All stars & obliqueness classes are taken together. • The altitude coverage is related to the star magnitude and occultation obliquity; profiles from dim-oblique occultations tend to reach less far downward.			

II.2.2 O3 vertical profile

Table 2: Relative differences between GOMOS IPF 6.01 O3 profile (full mission) and O3 sonde / O3 lidar. Positive bias values imply that GOMOS ozone is larger than correlative measurements.

IPF 6.01	60N-90N	30N-60N	30N-30S	30S-60S	60S-90S
Median bias (%)					
< 15 km	0	±10		> +15	>+15
15-20 km	-3	-2	> +15	+9	+7
20-25 km	-7	-3	+4	+2	-2
25-30 km	-6	-2	+1	-1	-1
30-35 km	-4	-1	+1	+1	0
35-40 km	-6	-3	±2	+1	+3
40-45 km	0	-5	-3	-2	< -15
Comparison spread (%)					
< 15 km	>25	>40		>40	>35
15-20 km	15	>15	>25	>15	27
20-25 km	10	11	13	10	16
25-30 km	12	8	7	8	15
30-35 km	12	7	6	6	13
35-40 km	18	9	5	7	21
40-45 km	>25	16	7	13	>25
Comments					
• GOMOS screening procedure as described in IPF 6.01 Readme file. • Below 25 km, the GOMOS data exhibit a negative drift (-5%/decade at 20km, and increasing with decreasing altitude). • At mid-latitudes the summertime ozone is on average 5% smaller than in winter. • No clear dependence of data quality on star class (bright/dim and hot/cool), with changes in bias <5%.					

II.3 MIPAS

II.3.1 Temperature vertical profile

Table 3: Absolute differences between MIPAS ML2PP 6.0 temperature profile (full mission) and radiosonde / T lidar. The statistics are separated for Full Resolution (2002-2004; left) and Optimized Resolution (2005-2012; right) periods. Positive bias values imply that MIPAS temperature is larger than correlative measurements.

ML2PP 6.0	Full Resolution (2002-2004)					Optimized Resolution (2005-2012)				
	60N-90N	30N-60N	30N-30S	30S-60S	60S-90S	60N-90N	30N-60N	30N-30S	30S-60S	60S-90S
Median bias (K)										
> 200 hPa	+0.5	+0.7		+1.1	0	-1.4	-2.0			-2.5
100-200 hPa	0	-0.4	-0.5	-0.5	-0.6	-0.7	-0.6	-0.7	+0.3	-1.2
50-100 hPa	+0.1	-0.1	-0.3	-1.0	-0.5	-0.6	-0.4	-0.1	+0.1	-1.0
20-50 hPa	+0.2	+0.2	-0.5	-1.7	-0.3	-0.9	-0.6	-0.4	-0.7	-1.2
10-20 hPa	+0.1	+0.6	-1.0	-3.0	+0.3	-1.0	-0.5	-0.7	-1.6	-1.1
5-10 hPa		+0.6	-0.6				+0.5	-0.7		
2-5 hPa	(-0.5)	-0.3	-0.9			(-1.6)	+0.2	-1.3		
1-2 hPa	(-0.6)	-1.2	-1.7			(-2.3)	+0.2	-2.0		
0.5-1 hPa	(-0.7)	-1.5	-1.5			(-2.5)	-0.1	-2.0		
0.2-0.5 hPa	(-1.1)	-1.2	-0.8			(-3.0)	-0.8	-1.8		
0.1-0.2 hPa	(-2.5)					(< -3)				
Comparison spread (K)										
> 200 hPa	2.2	2.7		2.6	2.5	2.2	2.4			2.3
100-200 hPa	1.9	2.2	1.7	2.0	2.5	1.5	1.8	1.6	1.8	1.6
50-100 hPa	1.8	1.7	1.3	1.8	2.3	1.3	1.4	1.2	1.7	1.8
20-50 hPa	1.6	1.7	1.2	1.6	2.0	1.3	1.4	1.1	1.6	2.0
10-20 hPa	1.5	2.2	1.6	2.0	2.2	1.5	1.8	1.6	1.8	2.3
5-10 hPa		2.5	1.7				2.2	1.7		
2-5 hPa	(> 4)	3.1	1.8			(> 4)	2.2	1.8		
1-2 hPa	(3.8)	3.6	2.1			(> 4)	2.5	2.1		
0.5-1 hPa	(3.4)	3.5	2.5			(> 4)	3.0	2.7		
0.2-0.5 hPa	(3.4)	3.5	2.8			(> 4)	3.7	3.4		
0.1-0.2 hPa	(3.5)					(> 4)				

Comments

- All MIPAS measurement modes included; screening procedure as described in ML2PP 6.0 Readme file.
- Values between (brackets) are potentially subject to larger sampling uncertainties. These should be considered with care.
- The data quality (bias, spread) has a clear seasonal dependence, especially at mid-latitudes and to a lesser extent in the polar regions. The bias changes by up to 1-2 K over the year, with an absolute minimum in June (day-time profiles) or December (night-time).

II.3.2 O3 vertical profile

Table 4: Relative differences between MIPAS ML2PP 6.0 ozone profile (full mission) and O3 sonde / O3 lidar. The statistics are separated for Full Resolution (2002-2004; left) and Optimized Resolution (2005-2012; right) periods. Positive bias values imply that MIPAS ozone is larger than correlative measurements.

ML2PP 6.0	Full Resolution (2002-2004)					Optimized Resolution (2005-2012)				
	60N-90N	30N-60N	30N-30S	30S-60S	60S-90S	60N-90N	30N-60N	30N-30S	30S-60S	60S-90S
Median bias (%)										
> 200 hPa	+8	> +15	+12	+12	> +15	> +10	>+15	+10	> +15	> +10
100-200 hPa	+4	+12	+15	+6	+10	+6	+8	-2	+8	+10
50-100 hPa	+2	+10	> +20	+5	+7	+2	+6	+7	+4	+5
20-50 hPa	-1	+4	+12	+3	+2	0	+4	+5	+4	+3
10-20 hPa	-3	+2	+3	+1	-2	-1	+5	+4	+4	+2
5-10 hPa	-4	+2	+3	+3		-1	+5	+8	+5	+3
2-5 hPa	-3	+3	+3	+2		0	+4	+10	+5	+6
1-2 hPa			+3			-8	-2	+10	+5	-10
Comparison spread (%)										
> 200 hPa	> 20	> 35	30	> 25	> 25	> 20	> 40	30	> 35	> 25
100-200 hPa	12	24	> 40	18	24	13	26	> 40	25	23
50-100 hPa	7	11	35	10	18	8	12	35	11	18
20-50 hPa	6	6	9	5	13	7	6	8	5	12
10-20 hPa	6	5	5	5	5	8	5	5	5	5
5-10 hPa	8	5	3	4		8	5	4	5	10
2-5 hPa	13	7	3	4		11	8	4	10	18
1-2 hPa			6			18	17	6	16	> 25
Comments										

- All MIPAS measurement modes included; screening procedure as described in ML2PP 6.0 Readme file.
- The comparison spread has a modest seasonal cycle, mainly at high latitudes. Variability is smallest in local summer and largest in local winter.

II.3.3 CH4 vertical profile

Table 5: Summary of MIPAS CH4 profile validation outcome divided into 5 latitude bands. Subsequent columns provide the number of stations, the number of full mission comparisons, the relative bias and spread for ML2PP 6.0, the number of delta-validation comparisons, and the relative bias and spread for ML2PP 7.01 (DDS).

Latitude band	# stations	# V6 comps.	V6 bias (%)	V6 spread (%)	# V7 comps.	V7 bias (%)	V7 spread [%]
Arctic (60N-90N)	4	2307	-5 to +3	2-5	318	-3 to +4	2-5
Mid-north (30N-60N)	3	1295	+1 to +4	2-5	554	+2 to +7	2-5
Tropics (30N-30S)	3	749	-9 to +1	1-4	273	-9 to +5	1-5
Mid-south (30S-60S)	1	560	0 to +4	2-4	15	+3 to +7	2-4
Antarctic (60S-90S)	0	0			0		

II.3.4 HNO3 vertical profile

Table 6: Summary of MIPAS HNO3 profile validation outcome divided into 5 latitude bands. Subsequent columns provide the number of stations, the number of full mission comparisons, the relative bias and spread for ML2PP 6.0, the number of delta-validation comparisons, and the relative bias and spread for ML2PP 7.01 (DDS).

Latitude band	# stations	# V6 comps.	V6 bias (%)	V6 spread (%)	# V7 comps.	V7 bias (%)	V7 spread [%]
Arctic (60N-90N)	4	2955	-17 to +25	6-40	337	-17 to +25	6-40
Mid-north (30N-60N)	4	428	-20 to +20	5-40	141	-7 to +20	5-40
Tropics (30N-30S)	3	380	-17 to +7	5-20	134	-17 to +5	5-20
Mid-south (30S-60S)	0	0			0		
Antarctic (60S-90S)	0	0			0		

II.3.5 N2O vertical profile

Table 7: Summary of MIPAS N2O profile validation outcome divided into 5 latitude bands. Subsequent columns provide the number of stations, the number of full mission comparisons, the relative bias and spread for ML2PP 6.0, the number of delta-validation comparisons, and the relative bias and spread for ML2PP 7.01 (DDS).

Latitude band	# stations	# V6 comps.	V6 bias (%)	V6 spread (%)	# V7 comps.	V7 bias (%)	V7 spread [%]
Arctic (60N-90N)	3	2285	-3 to +2	2-5	170	-3 to +3	2-7
Mid-north (30N-60N)	1	233	-2 to 0	2-3	108	0 to +2	3-4
Tropics (30N-30S)	3	322	-6 to -1	2-4	141	-4 to +4	2-4
Mid-south (30S-60S)	1	589	-4 to +3	1-9	13	0 to +4	1-10
Antarctic (60S-90S)	0	0			0		

II.4 SCIAMACHY

II.4.1 Nadir O3 total column

Table 8: Summary of the validation of SCIAMACHY SGP 5.02 nadir O3 total column using co-located ground-based observations by Dobson, Brewer and UV-visible instruments.

SGP 5.02	
Bias	General positive bias of +1-2%, without latitudinal structure. Furthermore depends on • SZA: negative bias of up to -4% for SZA>80° • cloud cover: weak dependence of up to at most 4%, with best agreement at small cloud cover fractions
Comparison spread	The spread in the comparisons ranges between 3-10%, which is dominated by a combination of measurement uncertainties and atmospheric noise (mismatch uncertainties). At low and middle latitudes, the spread increases for small values of cloud optical depth.
Stability	There is a transitory negative drift in the first years of the mission, but the bias remains stable after 2004.

II.4.2 Nadir NO2 total column

Table 9: Summary of the validation of SCIAMACHY SGP 5.02 nadir NO2 total column using co-located ground-based observations by UV-visible instruments.

SGP 5.02	60N-90N	30N-60N	30N-30S	30S-60S	60S-90S
Bias (molec. cm^{-2})	±3 x 10^{14}	+4 x 10^{14}	±3 x 10^{14}	-5 x 10^{14}	±3 x 10^{14}
Spread (molec. cm^{-2})	5 x 10^{14}	5-8 x 10^{14}	3 x 10^{14}	4 x 10^{14}	5 x 10^{14}
Comments	Apparent bias between NH and SH data, maybe due to difference in sensitivity to tropospheric pollution and/or to residual diurnal cycle effects between SCIAMACHY and NDACC/UV-visible twilight data.				

II.4.3 Nadir BrO total column

Table 10: Summary of the comparison of SCIAMACHY SGP 5.02 nadir BrO total columns and the ground-based UV-visible zenith-sky observations at Harestua (60°N, 11°E).

SGP 5.02 at Harestua (60°N, 11°E)	Relative difference (%)	Absolute difference (molec. cm^{-2})
Bias	-12.8	-6.4 x 10^{12}
Standard deviation	37.4	19.5 x 10^{12}
Comments	• SCIAMACHY vertical columns obtained from SCIAMACHY SCDs and total column AMF from ground-based retrievals. • Using the stratospheric AMF provided with the product increases the negative bias by about 5%. • The annual cycle is well reproduced. • Some outliers visible during 2003-2004 and 2007.	

II.4.4 Nadir CO total column

Table 11: Summary of SCIAMACHY nadir CO total column validation outcome divided into 5 latitude bands. Subsequent columns provide the number of stations, the number of monthly means, and the SGP 5.02 relative median bias and comparison spread in each band.

Latitude band	# stations	# monthly means	Median bias (%)	Spread (%)
Arctic (60N-90N)	5	191	+13	10-70
Mid-north (30N-60N)	4	278	+48	10-50
Tropics (30N-30S)	1	60	+55	20-30
Mid-south (30S-60S)	2	112	+40	10-90
Antarctic (60S-90S)	1	36	+15	20-60

II.4.5 Nadir H2O total column

Table 12: Absolute differences between SCIAMACHY SGP 5.02 nadir water vapour total column (full mission) and radiosonde. Positive bias values imply that SCIAMACHY water vapour is larger than correlative measurements.

SGP 5.02	Bias		Comparison spread	
	(g cm^{-2})	(%)	(g cm^{-2})	(%)
All pixels	-0.04	-6	0.32	30
Land, cloud free	+0.23	+20	0.31	25
Land, cloudy	-0.06	-9	0.33	28
Ocean, cloud free	-0.06	-6	0.28	17
Ocean, cloudy	-0.07	-12	0.29	31

Comments

- SCIAMACHY data screened according to prescription in SGP 5.02 Readme file.
- Co-location criteria: all SCIA pixels within 50 km and 1 h from location and time of sonde launch.
- Data quality degrades with increasing cloud top height under very cloudy conditions.
- At low AMF correction factor the bias becomes increasingly negative, especially for cloudy pixels.
- Bias and spread vary over the course of a year. They are smallest in local spring and largest in local summer
- At small solar zenith angle the bias becomes positive and the variability increases, for all pixel classes.

II.4.6 Limb O3 vertical profile

Table 13: Relative differences between SCIAMACHY SGP 5.02 limb ozone profile (full mission) and O3 sonde / O3 lidar. Positive bias values imply that SCIAMACHY ozone is larger than correlative measurements.

SGP 5.02	60N-90N	30N-60N	30N-30S	30S-60S	60S-90S
			Median bias (%)		
15-20 km	< ±7 / ±10	+4	> +10	-5	< ±15
20-25 km	+5 / > +20	+2	+7	+6	+2
25-30 km	-4 / +12	+1	+12	+12	+3
30-35 km	> +10	+2	> +15	+16	+7
35-40 km		+3	+12	+14	+15
40-45 km		+12	+16	+15	+13
			Comparison spread (%)		
15-20 km	21 / 27	22	> 40	22	> 40
20-25 km	27 / > 35	10	12	11	27 / 18
25-30 km	17 / > 40	10	10	11	17 / 12
30-35 km	> 40	12	12	12	14
35-40 km		17	16	20	24
40-45 km		25	18	38	> 40
			Comments		

The SCIAMACHY data exhibit several important quality issues

- We discourage the use of SCIAMACHY averaging kernels. Smoothing higher resolution profiles using the AKs led to vertical oscillations in the bias (up to 5%) and comparison spread profiles.
- The SCIAMACHY bias has very a pronounced dependences on altitude and latitude, and to a lesser extent on solar zenith angle (or month) and scan angle.
- Data quality is strongly degraded in the Arctic, especially during winter. Use Arctic data with caution.
- Above 30 km there is a negative drift of up to 8%/decade. We discourage the use SCIAMACHY's US data for trend analyses.
- Bias changes of up to 5% should be expected when the provided auxiliary data are used to convert from SCIAMACHY's native (O3 number density at fixed altitude levels) to other profile representations.

II.4.7 Limb BrO vertical profile

Table 14: Summary of the comparison of SCIAMACHY SGP 5.02 limb BrO vertical profiles and the ground-based UV-visible zenith-sky observations at Harestua (60°N, 11°E).

SGP 5.02	Bias (%)		Comparison spread (%)	
at Harestua	15 Feb - 30 Apr	1 May - 30 Nov	15 Feb - 30 Apr	1 May - 30 Nov
15-18 km	+6	+32	50	50
18-21 km	-12	+12	30	30
21-24 km	-32	-12	25	25
24-27 km	-45	-30	25	25
Comments	• No seasonal cycle in the 15-27 km partial column data. • Scientific BrO limb profile product better than SGP in terms of bias and annual cycle.			

III Project management (WP 1)

The Multi-TASTE Phase F CCN-1 started 1st October 2013 and was foreseen to end 30th September 2014. This was then extended by ESA (1st December 2014 to 31st March 2015) to ensure the validation of the new SCIAMACHY SGP 6.00 prototype and to include activities by KNMI, as a subcontractor to BIRA-IASB.

The project was faced with unexpected and significant delays in the delivery of some Envisat Level-2 data products by the QWGs, especially SCIAMACHY SGP 6.00 DDS and MIPAS ML2PP 7.03 full mission. Where possible, and in close collaboration with the QWGs, the teams at BIRA-IASB and KNMI prioritized which Level-2 data products had to be processed first to start the validation process. This minimized further delays in the completion of the validation analyses. Nevertheless, these issues led to an accumulated delay of almost 9 months for the project. At end of project, all tasks, work packages and deliverables were completed, and no notable issues (other than those described above) are reported.

IV Data management (WP 2)

IV.1 Ground-based data handing (WP 2.1)

IV.1.1 Status of correlative data sets

Multi-TASTE uses primarily ground-based measurements provided by monitoring and research networks working under the auspices of WMO's Global Atmosphere Watch, and in particular Global Ozone Observing System (GO_3OS), Network for the Detection of Atmospheric Composition Change (NDACC), and Southern Hemisphere Additional Ozonesondes (SHADOZ). The database manager at BIRA-IASB synchronizes the local copy of the WOUDC, NDACC DHF and NILU archives on a weekly basis. This includes data from various ground-based instruments (Dobson, Brewer, DOAS UV-VIS spectrometer, FTIR spectrometer, ozonesonde, ozone/temperature lidar and microwave radiometer). The automated Correlative2 in-house software system then processes most raw data in preparation for correlative analysis. A separate, semi-automated system was developed that ingests the data not yet supported by the Correlative2 system (FTIR HDF5 files, microwave radiometer files).

IV.1.2 Correlative data updates to EVDC

During the project BIRA-IASB performed regular updates to ESA's Validation Data Centre (EVDC) hosted at NILU for the stations and the ground-based instruments that we are either responsible for or for which we have an agreement with the instrument PIs. We contacted our former Multi-TASTE sub-contractors (contract period 2008-2012) and additional NDACC instrument PIs and asked them to update their data set in the EVDC archive, and to upgrade already archived data sets with newer versions in case of a significant reprocessing. Some partners committed to complete and/or upgrade their data records. Other groups had to recognize their inability to perform such activities themselves as a consequence of collapsing personnel and financial resources. However, in the second half of 2015, NILU activated a direct mirror between the NDACC DHF and EVDC. This allows EVDC users to access the public ground-based data records available in the NDACC archive. The spatial and temporal coverage of the trace gas records at EVDC hereby increased considerably, surpassing the period of Envisat operations and for a broader set of measurement locations. This direct link also ensures that, at any future point in time, the EVDC users will have the latest version of the data at their disposal.

IV.2 Satellite data handling (WP 2.2)

IV.2.1 Status of satellite data sets

The current and previous versions of the operational GOMOS (IPF 5.00/5.01 and IPF 6.01), MIPAS (IPF 5.05/5.06 and ML2PP 6.0) and SCIAMACHY (SGP 3.01 and 5.02) Level-2 data sets were downloaded, processed and made available for correlative analysis. An overview of the availability of the satellite data sets can be found in Table 15. During the project several new Level-2 diagnostic and reprocessed data records were released to the QWG and the validation teams, unfortunately with considerable delays. It concerned the MIPAS ML2PP 7.01 (DDS) and 7.03 (full mission), and SCIAMACHY SGP 6.00 (DDS) products. These satellite data products were processed for subsequent correlative analysis.

Table 15: History of the delivery of Envisat Level-2 data products, the presentation of ground-based validation results and the public release.

Instrument	Level-2 Processor	Full - DDS	Delivery date	First validation results	Consolidated validation results	Public release
GOMOS	IPF 5.00/P, IPF 5.01/P, GOPR 6.0b+f/Q	Full	early 2006		2 Tech Notes	8 Aug 2006
	IPF 6.01/R	Full	8 Aug 2012	QWG#27	QWG#N4	18 Dec 2012
MIPAS	ORM 1.0 (IPF 5.04)	DDS	14 Apr 2010	QWG#23		
	IPF 5.05/R, 5.06/R	Full	Jan-Feb 2011	QWG#25		21 Jun 2010
	ORM 2.0 (ML2PP 6)	DDS	30 Jul 2010	QWG#27	QWG#28	
	ML2PP 6.0/U	Full	26 Jan 2012	QWG#30	QWG#32,#33	7 Jun 2012
	ML2PP 7.01/W	DDS	24 Jun 2014	QWG#36	QWG#37,#38	
	ML2PP 7.03/W	Full	31 Aug 2015	QWG#40		Spring 2016?
SCIAMACHY	SGP 3.01/R	Full	Mar 2008		SQWG#22	Mar 2008
	SGP 5.01/U	DDS	14 Apr 2010	SQWG#11	SCIAVALIG QL	
	SGP 5.02/W	Full	spring 2012		SQWG#22	5 Jun 2012
	SGP 6.00/Y	DDS	28 Apr 2015	SQWG3#3	SQWG3#4 + Tech Note	

IV.2.2 Definition of new Diagnostic Data Sets

ESA asked the project partners to revisit the orbit list for delta-validation exercises. This set of orbits should fulfil following requirements:

- selection of ~10% of the full mission,
- allow correlative delta-validation analyses of all trace gas products included in the validation projects,
- provide results for all correlative instrument types and, where applicable, for different measurement modes of the satellite instrument,
- optimal coverage in space (latitude, altitude) and time (year, season).

We developed a novel optimization algorithm to identify orbit lists satisfying above criteria based on the co-location metadata of the analyses of the latest full mission operational data products. The validation teams presented several orbit lists to the QWGs and ESA. After discussion within the QWG the orbit list recommended by the Multi-TASTE partners was selected to produce all future Diagnostic Data Sets (DDS). The delta-validation exercises during the project (MIPAS ML2PP 7.01 and SCIAMACHY SGP 6.00) were based on these DDS. Validation results for the MIPAS ML2PP 7.01 (DDS) and ML2PP 7.03 (full mission) products were almost identical, illustrating the excellent representativeness of the DDS.

GOMOS DDS

- 4387 orbits;
- Optimized for O3 and HRTP comparisons versus sonde and lidar;
- Delivered as GOMOS_V7_DDS.v2.xlsx to ESA on 3 Feb 2015.

MIPAS DDS

- 4180 orbits;
- Optimized for O3, T, HNO3, CH4 and N2O profile comparisons versus sonde, lidar and FTIR (4059 orbits) and for many more species provided by MIPAS-B flights (121 orbits);
- Presented in Progress Report TN-BIRA-IASB-MultiTASTE-Phase-F-PR1 (Hubert et al., 2014);
- The 4059 orbits were delivered as 20140401__MIPAS_V7_TDS__BIRA-IASB_v2.dat to ESA on 1 Apr 2014.

SCIAMACHY DDS

- 5011 orbits;
- Optimized for O3, NO2, BrO, CO nadir products and O3, BrO limb products versus Dobson, Brewer, UV-vis spectrometers, sonde, lidar and FTIR;
- Presented at SCIAMACHY QWG meeting PM1 (20 October 2014);
- Delivered as SCIAMACHY_DDS_orbits.v4.dat to SCIAMACHY QWG on 25 Nov 2014.

IV.2.3 Assistance to the definition of new Level-2 data formats

The Envisat QWGs are preparing for the preservation of the latest Level-1b and Level-2 data records for the future decades. All data will be stored in netCDF format, in data structures that should be more intuitive and user-friendly than current data records. The validation teams are often the first users of new data records, they have experience with many different satellite data records and they are aware of the requirements and/or habits of data users. As a result, we were able to provide adequate feedback to this process.

IV.3 Quality assurance of ground-based data (WP 2.3)

All contributing instruments are affiliated with measurement networks operating in the framework of WMO's Global Atmosphere Watch (GAW). As such they have to comply with a series of rules and guidance established by international expert bodies to assure good levels of overall quality of the measurements and homogeneity of the network: WMO's Quality Management Framework, GAW guidance documents, NDACC protocols (Measurements Protocol, Data Protocol, Validation Protocol, Instrument Intercomparisons Protocol…) Quality is a concept referring to user requirements. While the overall quality of GAW affiliated instruments fulfils requirements representative for a wide variety of monitoring and analysis applications, stringent requirements imposed by detailed satellite validation studies of subtle effects need sometimes additional data screening, e.g., to ensure detectability of small changes between two satellite data processor versions and early detection of long-term drifts. Therefore the Quality Assurance activities described hereafter have been specifically added by the Multi-TASTE Validation System to address particular issues of satellite validation. Those functions are based on step-by-step and ad hoc experience with data from Envisat, GOME and Third Party Missions.

IV.3.1 Dobson and Brewer UV spectrophotometers

Unless specified, only direct Sun total ozone data acquired by Brewer and Dobson instruments are selected for satellite validation, due to their lower uncertainty. Several stations report Brewer and Dobson data acquired in zenith-sky observation mode as well, useful for monitoring purposes, but

they are not used for detailed satellite validation studies. After this first selection Brewer and Dobson data undergo other quality checks and filtering procedures based on the examination of calibration history reported in WMO Technical Documents (e.g. regular WMO endorsed Dobson and Brewer intercalibration campaigns) and in quality assessments available from WOUDC, where Dobson and Brewer data available at WOUDC are compared in routine to a reference built from a satellite data ensemble (V. Fioletov, personal communication). Instruments or data sets obviously affected by issues such as unexplained bias, drift or noise, are discarded for the period during which the quality of the data is questionable. The next step is an additional filter depending on the instrument type (Figure 1). To avoid unwanted biases due to the known air-mass dependence of instruments equipped with a single monochromator, data acquired by Brewer instruments of the Mark I/II/III type, and Dobson data, are not used if acquired at air-masse higher than 3/3.5, that is, solar zenith angle higher than 75°. Brewer data acquired by Mark-IV instruments are used up to 85° SZA since much less affected by an air-mass dependence. The temperature dependence of the ozone absorption cross-section in the UV, to which Dobson measurements are sensitive, limits the accuracy of Dobson total ozone data by introducing a latitude-dependent bias and a seasonally varying bias. Those biases are well characterised in the literature but no correction has been issued by the Dobson community so far. Nevertheless, thanks to the good documentation, the Dobson data can be used as they are, provided that those temperature induced biases are taken into consideration in the interpretation of validation results. Thanks to their optimized choice of wavelengths, Brewer data are much less sensitive on stratospheric temperatures and can be used as provided.

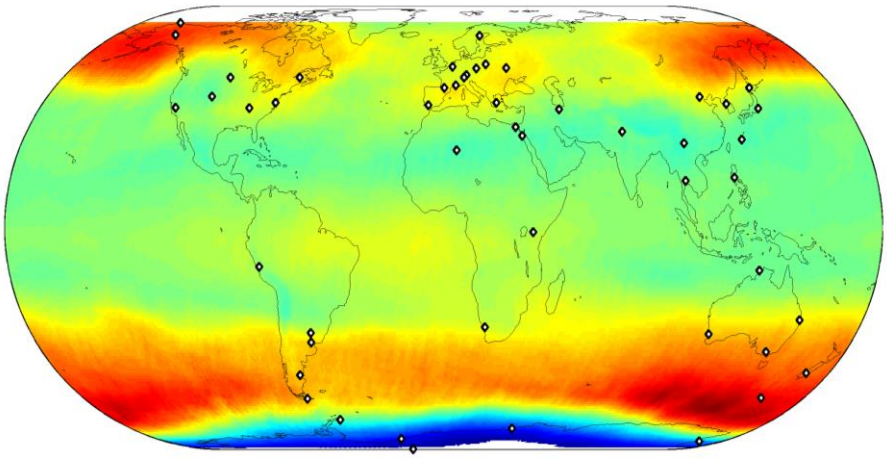


Figure 1: Geographical distribution of Dobson instruments contributing O3 column data to Multi-TASTE studies, on a global map of SCIAMACHY total O3 for October 2010.

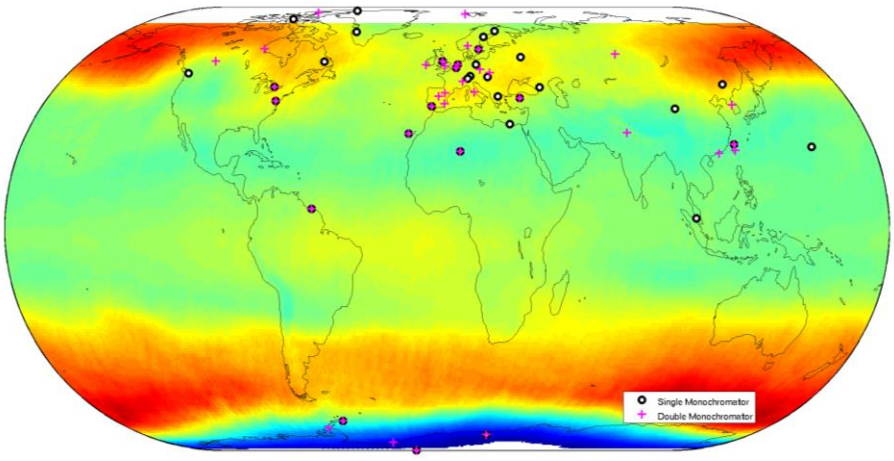


Figure 2: Like Figure 1, but for Brewer instruments, differentiated per type of monochromator.

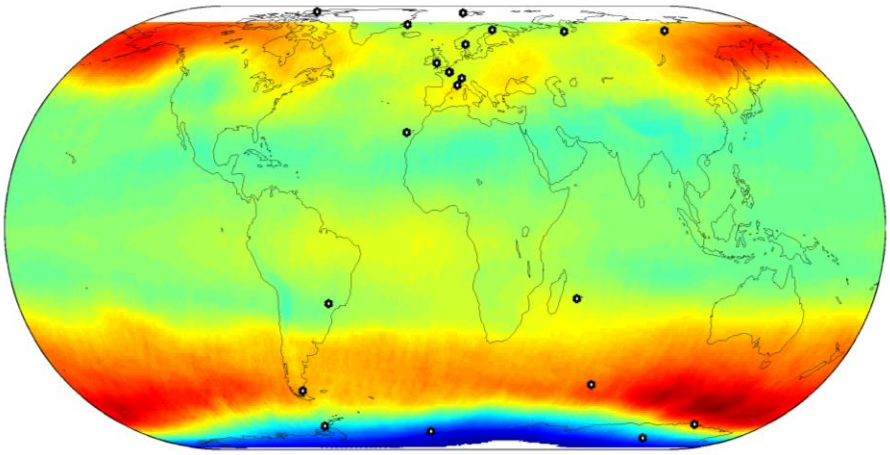


Figure 3: Like Figure 1 but for DOAS UV-visible instruments.

IV.3.2 DOAS UV-visible spectrometer

Correlative ozone and nitrogen dioxide column data undergo first quality checks and filtering procedures, e.g., enhanced NO_2 column values occurring simultaneously with O_4 and H_2O column enhancements are filtered out since most likely affected by uncorrected multiple scattering within dense clouds or snow showers (Pfeilsticker et al., 1999). Most of the NDACC DOAS/UV-visible stations, which retrieve ozone column data with the latest version of the algorithm (Hendrick et al. 2011), provide ozone column data directly applicable to satellite validation. For nitrogen dioxide the situation is more complex, due to the difference in vertical sensitivity between satellite-based nadir UV-visible instruments which are sensitive in both the stratosphere and the troposphere, and ground-based zenith-sky UV-visible instruments which are sensitive mainly to the stratosphere. One solution could have been to select only stations in pristine locations, but then there would have been no station in the Northern middle latitudes. Another solution is to work (1) with a classification of the stations according to the incidence (hereafter, incidence values ranging from 0 for pristine stations through 2 for stations occasionally polluted and up to 5 for stations permanently surrounded by variable tropospheric NO_2) of tropospheric NO_2 on the validation technique, like proposed in Figure 4 and Table 16, and (2) at stations with pollution events (incidence of 3 to 5) include an additional filter based on cloud screening, most of the pollution being below the clouds. Those cloud screening filters are designed empirically according the station and its particular type of cloudiness.

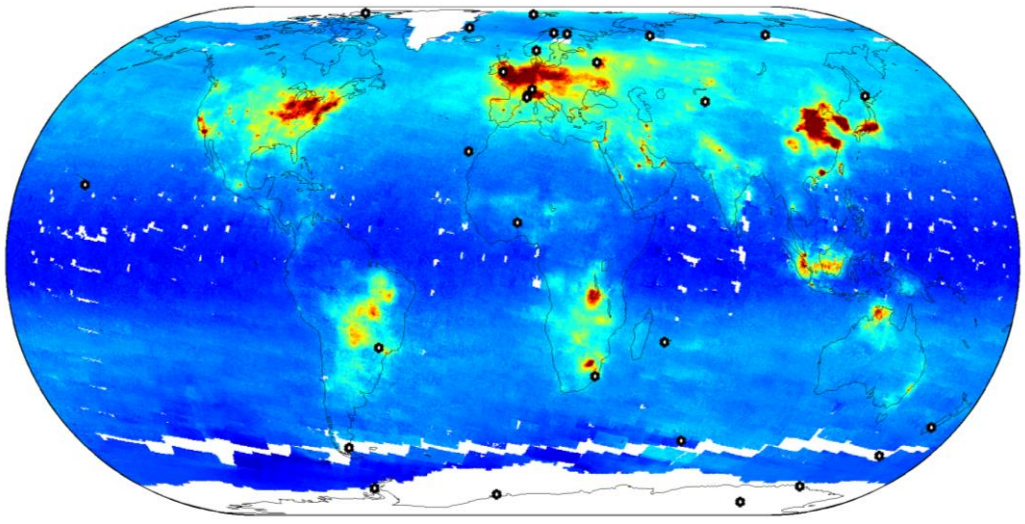


Figure 4: Geographical distribution of NDACC DOAS UV-visible instruments contributing NO2 column data to Multi-TASTE studies, on a global map of SCIAMACHY tropospheric NO2 for September 2010.

Table 16: Classification of NDACC UV-visible instruments contributing data to Multi-TASTE according to the average NO2 tropospheric load. Incidence of tropospheric NO2 is classified with a scale from 0 to 5.

Station	Region	Latitude	Longitude	Altitude	Incidence	Type	Institute
Ny-Ålesund	Spitsbergen	78,91°N	11,88°E	20m	0	SAOZ IFE	NILU IFE/IUP
Thule	Greenland	76,51°N	68,76°W	220m	0	SAOZ	DMI
Scoresbysund	Greenland	70,48°N	21,97°W	10m	0	SAOZ	CNRS/DMI
Sodankylä	Finland	67,37°N	26,67°E	179m	1	SAOZ	CNRS/FMI
Zhigansk	Siberia	66,72°N	123,4°E	50m	0	SAOZ	CNRS/CAO
Harestua	Norway	60,22°N	10,75°E	596m	2	IASB	IASB
Zvenigorod	Russia	55,42°N	36,47°E	220m	3	IAP	IAP-Moscow
Bremen	Germany	53,11°N	8,86°E	27m	5	IFE	IFE/IUP
Aberystwyth	Wales	52,42°N	4,07°W	20m	4	SAOZ	Uni. Manchester
Jungfraujoch	Swiss Alps	46,55°N	7,98°E	3580m	3	SAOZ	IASB-BIRA
Moshiri	Japan	44,4°N	142,3°E	200m	3	NIWA	STEL
Issyk-Kul	Kyrgyzstan	42,63°N	76,98°E	1650m	0	IAP	KSNU
Haute Provence	France	43,94°N	5,71°E	684m	3	SAOZ	CNRS
Rikubetsu	Japan	43,5°N	143,8°E	370m	3	NIWA	STEL
Kiso	Japan	35,8°N	137,6°E		3	NIWA	STEL
Izaña	Tenerife	28,30°N	16,50°W	2367m	1	INTA	INTA
Mauna Loa	Hawaii	19,54°N	155,58°W	3397m	1	NIWA	NIWA
Mérida	Venezuela	8,2°S	71,1°W	4765m	1	IFE	IFE/IUP
Nairobi	Kenya	1,27°S	36,80°E	1795m	2	IFE	IFE/IUP
Saint-Denis	La Réunion	20,85°S	55,47°E	24m	1	SAOZ	CNRS/U.Réunion
Bauru	Brazil	22,35°S	49,03°W	300m	2	SAOZ	CNRS/UNESP
Lauder	New Zealand	45,03°S	169,68°E	370m	0	NIWA	NIWA
Kerguelen	Indian Ocean	49,36°S	70,26°E	10m	0	SAOZ	CNRS
Macquarie	Australia	54,50°S	158,96°E	6m	0	NIWA	NIWA
Marambio	Antarctica (AR)	64,28°S	56,72°W	200m	0	INTA	INTA
Dumont d'Urville	Antarctica (F)	66,67°S	140,01°E	20m	0	SAOZ	CNRS
Rothera	Antarctica (UK)	67,57°S	68,13°W	10m	0	SAOZ	BAS
Syowa	Antarctica (JP)	69,00°S	39,35°E	22m	0	NIWA	STEL
Arrival Heights	Antarctica (NZ)	77,82°S	166,66°E	250m	0	NIWA	NIWA
Belgrano II	Antarctica (AR)	77,87°S	34,63°W	50m	0	INTA	INTA

IV.3.3 FTIR spectrometer

Table 17 provides an overview of the NDACC (http://ww.ndacc.org) ground-based FTIR stations whose column and profile measurements have been used for the comparative validation of the SCIAMACHY CO columns and the MIPAS CH4, HNO3, and N2O profiles, respectively. The locations of these stations are presented on global maps within the main text for column and profile data separately.

All infrared reference data under consideration (should) follow the NDACC FTIR instrument protocols (http://www.ndsc.ncep.noaa.gov/organize/protocols/appendix2) as defined by the NDACC infrared working group (http://www.acd.ucar.edu/irwg). In agreement with published work, no additional filtering has therefore been applied to the CO column data (Dils et al., 2006; de Laat et al., 2010), nor to the CH4, HNO3, or N2O profile data (Vigouroux et al., 2007; Payan et al., 2009; Sepúlveda et al.,

2014) for both the station quality assessment presented here and the comparative validation study outlined in the main text.

The ground-based FTIR station quality assessment combines both own findings as based on data content, co-location, and comparative studies (also see main text), and results presented in the common literature. The outcome of our studies have been indicated in Table 17 and can be summarised as follows:

- Many stations provide adequate FTIR column or profile data on a regular basis for the full Envisat lifetime span 2002-2012 (cells with green background).
- For some measurement series however, part of the Envisat lifetime range is not covered, so that comparative analyses with SCIAMACHY or MIPAS datasets are eschewed towards the limited time period (cells with blue background). Sometimes these station datasets also show statistical constraints: years with less than five ground-based FTIR measurements are removed from yearly averages in the SCIAMACHY CO validation.
- Despite the existence of the FTIR instrument protocols, working with the NDACC data has revealed some unit issues (cells with orange background) that have been reported to the respective station PIs:
    - Unit issue for Bremen and Ny-Ålesund CH4 profile data, both provided by IUP.
    - Unit and averaging kernel issue for the Zugspitze CO column data and N2O profile data in the form of a factor 1000 overestimation.
    - Unit issue for the Harestua CH4 profile data in the form of an overestimation by a factor of one million.
- Finally, some ground-based infrared datasets show comparison results that deviate strongly from the typical bias and spread values at other validation stations, and are therefore considered to provide spurious measurement values (cells with red background). Note that some of these datasets overlap with the previous selection showing unit issues, suggesting that the corresponding FTIR instrument operations or retrievals might be flawed. This is to be further investigated by the station PIs. Note that the comparison results for these datasets are nevertheless included in the main text, partially as a backup for this QA.

According to literature, the systematic and random measurement uncertainties of ground-based FTIR instruments within the NDACC take the following typical values, differentiated per species

- The FTIR CO column random uncertainty is claimed to range between 2% and 5% in general (Dils et al., 2006; de Laat et al., 2010). This range is backed by station-specific spread assessments equalling about 2%, 1.5%, and 1% to 6% at Eureka (Batchelor et al., 2009), Jungfraujoch (Barret et al., 2003), and St. Denis (Réunion Island) (Senten et al., 2008), respectively. The systematic measurement uncertainty at the same three stations takes values within the same 1% to 6% interval.
- Comparisons between NDACC and GAW CH4 profile datasets have revealed a typical bias of about 2.5% and a spread of the order of 1% (Sepúlveda et al., 2014). Detailed studies at Zugspitze (Sussmann et al., 2011) and St. Denis (Senten et al., 2008) have however yielded biases from 5% up to 25 %. The spread at the same two stations, and at Izaña additionally (Schneider et al., 2005), ranges between 1% and 3 %.
- Much less is known about the HNO3 and N2O FTIR profile data from the ground-based NDACC stations. Senten et al. (2008) provide statistics for the St. Denis station, with a systematic uncertainty of 15-34% for HNO3 and 6-10% for N2O, and related random uncertainties of 21-36% and 1-2%, respectively.

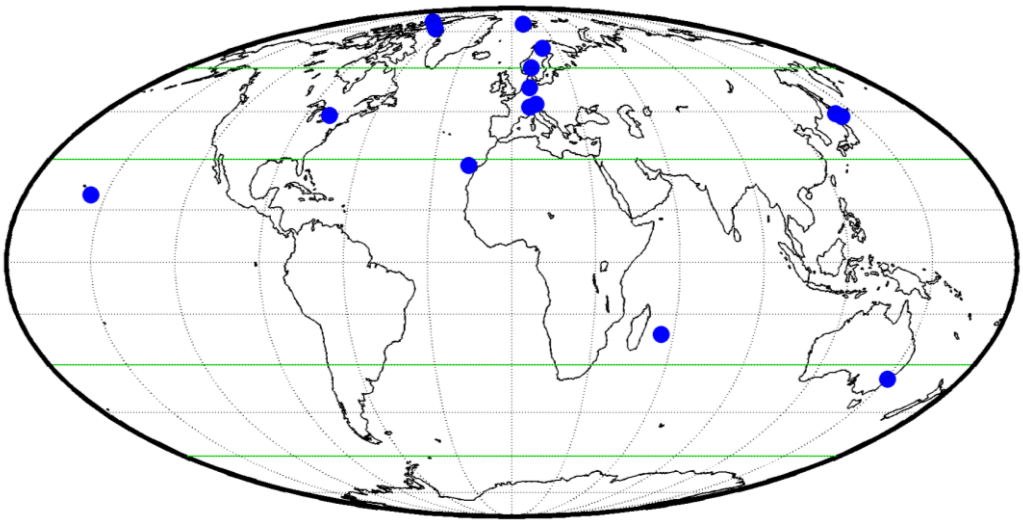


Figure 5: Global distribution of the 15 NDACC FTIR stations providing CH4, HNO3, or N2O vertical profile data. Green lines mark the edges of the latitude bands considered in Table 21.

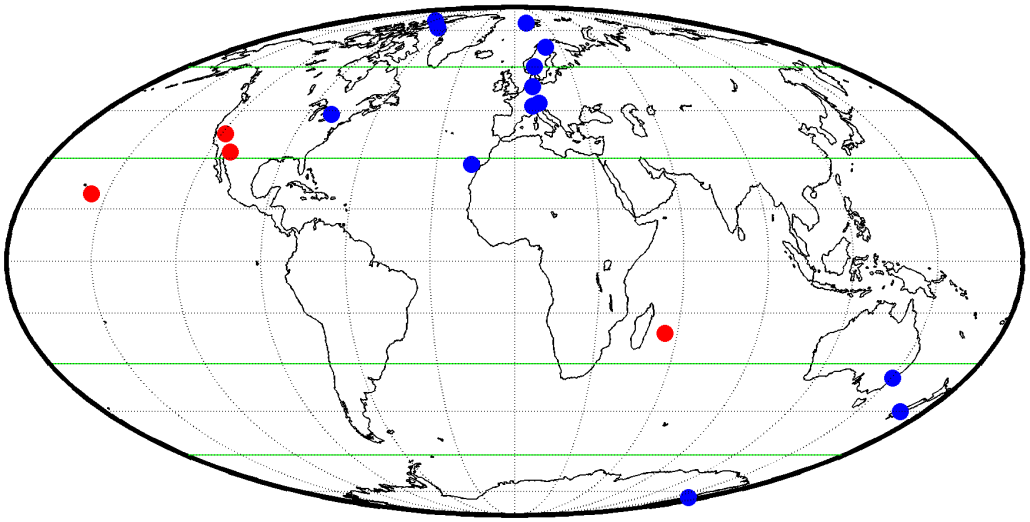


Figure 6: Locations of the 17 ground-based FTIR stations that have provided CO column data for the SCIAMACHY CO column validation, with the 13 stations used in the comparative analysis marked in blue and the four omitted stations marked in red. Green lines mark the edges of the latitude bands considered in Table 22.

Table 17: Overview of the NDACC ground-based FTIR stations (sorted north to south) whose CO column and CH4, HNO3, and N2O profile data have been used for comparative validation. The measurement time span and number of observations is indicated for each species and station. Cell colour codes are clarified in the text.

Station	Lat.	Lon.	CO column	CH4 profile	HNO3 profile	N2O profile
Eureka	79.99	-85.93	2006-2012 (33)	2006-2012 (2652)	2006-2012 (4173)	2006-2012 (2539)
Ny-Ålesund	78.93	11.93	2002-2012 (265)	2002-2012 (955)	2002-2012 (691)	
Thule	76.53	-68.74	2002-2011 (274)	2002-2012 (1469)	2002-2012 (1474)	2002-2012 (2604)
Kiruna	67.84	20.41	2002-2012 (474)	2002-2012 (1930)	2002-2012 (2076)	2002-2012 (1930)
Harestua	60.20	10.80	2002-2012 (327)	2002-2012 (1187)		
Bremen	53.10	8.80	2002-2012 (382)	2004-2012 (1060)		
Zugspitze	47.42	10.98	2002-2011 (1381)	2002-2012 (6685)		2002-2012 (13038)
Jungfraujoch	46.55	7.98	2002-2012 (1240)	2002-2012 (4327)	2004-2012 (4037)	2002-2012 (2665)
Moshiri	44.40	142.30			2002-2007 (714)	
Toronto	43.78	-79.47	2004 (414)	2002-2012 (2013)	2002-2012 (906)	2002-2012 (4041)
Rikubetsu	43.50	143.80			2002-2009 (563)	
Mt. Barcroft	37.58	-118.24	2002 (62)			
Kitt Peak	31.90	-111.60	2002-2005 (501)			
Izaña	28.30	-16.50	2002-2012 (498)	2002-2012 (9089)	2002-2012 (2397)	2002-2012 (4490)
Mauna Loa	19.53	-155.58	2003-2010 (40)	2003-2010 (972)	2003-2010 (557)	2003-2010 (1455)
St. Denis	-20.90	55.50	2004-2011 (91)	2004-2012 (3263)	2004-2012 (1882)	2004-2012 (444)
Wollongong	-34.41	150.88	2002-2008 (550)	2007-2012 (5211)		2002-2012 (6476)
Lauder	-45.04	169.68	2002-2012 (623)			
Arrival Heights	-77.83	166.67	2002-2012 (103)			

IV.3.4 Balloon-borne sonde

O3 sonde measurements are automatically collected every week from the NDACC, SHADOZ and GAW's WOUDC data archives, and then pre-processed by the Correlative2 software to harmonize the format for internal use. In this process, all obvious out-of-range values (latitude, longitude, wrong incorrect magnitude scaling for O3 values, etc…) are fixed or flagged.

In a second step, we remove all measurements with pressure>5 hPa (or altitude>33 km), because of degraded sonde data quality (Smit et al., 2011). We reject measurement levels with clearly unphysical readings (negative O3, pressure<0 hPa or temperature<0 K or >400 K), or during unrealistic jumps in pressure (dp/dt>0 and dz>0.1 km). Entire sonde flights are discarded from further analysis when (a) more than half of the levels are tagged bad, or (b) less than 30 levels are tagged good. In general, this procedure removes <5-10% of the measurement levels and <1% of the profiles.

As a last step, we identified a number of stations that are suitable for our baseline correlative analyses, see Figure 8. This was done using a newly developed method to evaluate the internal consistency of the ozone profile record obtained by the ground-based networks. It is based on the study of the vertical dependence of the bias of ozonesonde relative to five limb/occultation instruments (SAGE II, OSIRIS, GOMOS, MIPAS, Aura MLS). At each station, common features in the sonde bias profiles for relative to the different satellite records allow for an evaluation of the suitability of each station's sonde record. The results for the ozonesonde network were presented at various workshops and conferences (ESA ACVE 2013, EUMETSAT MSC 2015 and NORS/NDACC/GAW Workshop 2015). More details can be found in the respective contributions. We concluded that the homogeneity of the NDACC, GAW and SHADOZ networks is best around 25 km, the spread of the satellite-derived bias across the network is less than 3%. At other altitudes it is better than 5% (Figure 7). These values lie within the estimates of bias and precision for sonde instruments.

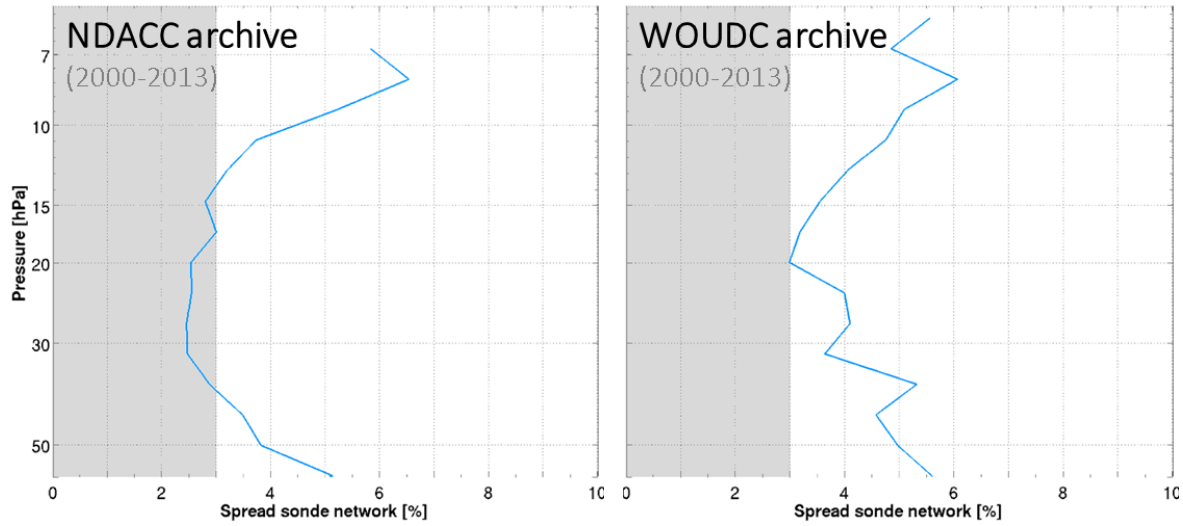


Figure 7: One sigma standard deviation of the (satellite-derived) bias over the NDACC and GAW ozonesonde networks in the middle stratosphere (50 hPa ≈ 20 km, 10 hPa ≈ 30 km). Taken from Hubert et al. (EUMETSAT MSC 2014).

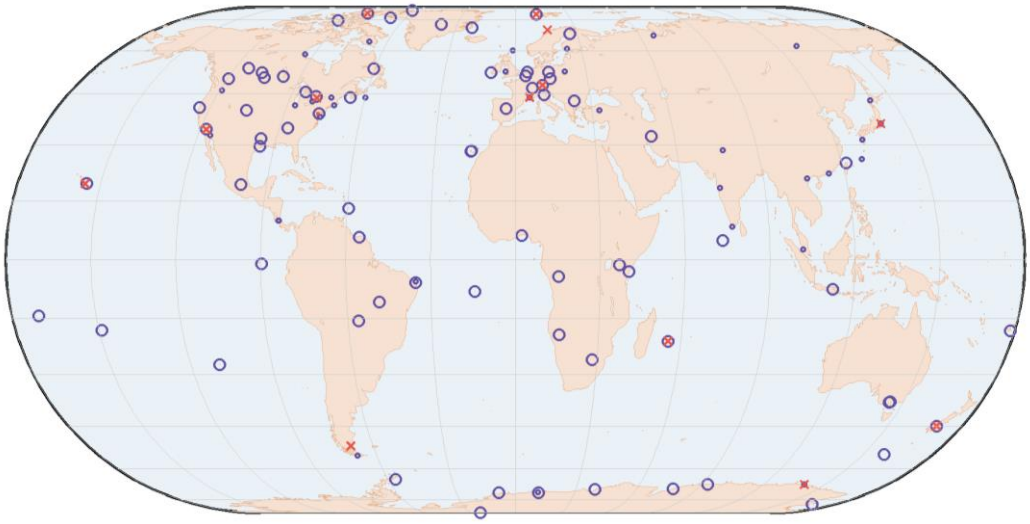


Figure 8: Location of sonde (purple circles) and stratospheric lidar (red crosses) sites used for baseline correlative analyses of O3 and T vertical profiles and H2O total columns. For dedicated analyses with specific needs, other ground-based stations not depicted here are valuable as well.

IV.3.5 Lidar

The QA analysis of stratospheric lidar measurements follows the same procedures described above for ozonesonde data. The main differences are (a) the vertical range of accepted data (O3 lidar: 15-47 km; T lidar: 30-70 km), and (b) the requirement on pressure jumps is irrelevant. Figure 8 shows the lidar sites used in our baseline correlative analyses. The screening typically removes <20% of the measurement levels and <0.5% of the vertical profiles.

IV.3.6 Measurement uncertainties of correlative data

Table 18: Overview of measurement uncertainties for the ground-based trace gas data products provided by and/or used within the Multi-TASTE Phase F projects.

Instrument	Species	Typical range Bias	Typical range Precision	Reference
Brewer	O_3 vert. col.	< ~4%	~1%	Balis et al. (2007)
Dobson	O_3 vert. col.	< ~4%	~1%	Balis et al. (2007)
DOAS/SAOZ	O_3 vert. col.	< ~4%	~1-3%	Hendrick et al. (2011)
	NO_2 vert. col.	Total accuracy <15%		Vandaele et al. (2005), Roscoe et al. (2010)
DOAS	BrO vert. col.	Total accuracy 7-15%		Hendrick et al. (2007)
	BrO profile	11-18%	18-38%	Hendrick et al. (2007)
MAX-DOAS	NO_2 tropo. col.	20-30% depending on pollution		Wittrock et al. (2012)
	BrO tropo. col.	~37%	15%	Theys et al. (2007)
	HCHO tropo. col.	10%	20-25%	Vigouroux et al. (2009)
Ozonesonde	O_3 profile	5-10%	3-5%	Smit et al. (2011)
Radiosonde	T profile	0.2-0.4K	0.2K	Sun et al. (2013)
	H_2O vert. col.	5-15%	2-5%	Miloshevich et al. (2009)
O3 lidar	O_3 profile	2-7.5%	2-7.5%	Keckhut et al. (2004)
T lidar	T profile	1-7.5K	2-7.5K	Keckhut et al. (2004)
Microwave radiometer	O_3 profile	2% (surface) – 30% (100km)	2% (surface) – 50% (100km)	Palm et al. (2010)
FTIR	O_3 vert. col.	~5%	~1%	Senten et al. (2008)
	O_3 profile	4-11%	7-23%	Senten et al. (2008), Schneider et al. (2008)
	CH_4 vert. col.	~20%	~1%	Senten et al. (2008)
	CH_4 profile	5-25%	1-3%	Senten et al. (2008), Sussman et al. (2011), Sepúlveda et al. (2014)
	N_2O vert. col.	~5%	~0.3%	Senten et al. (2008)
	N_2O profile	6-10%	1-2%	Senten et al. (2008)
	CO vert. col.	1-6%	1-6%	Barret et al. (2003), Senten et al. (2008), Batchelor et al. (2009)
	CO profile	3-7%	1-13%	Senten et al. (2008)
	HNO_3 vert. col.	15-34%	21-36%	Senten et al. (2008)
	HCL vert. col.	3-5%	7-11%	Senten et al. (2008)

IV.4 Lidar data standardisation (NDACC ISSI team)

In the previous VALID projects, KNMI participated in an ISSI team to help the NDACC lidar working group standardise the reported temperature and ozone observations in terms of vertical resolution, uncertainties, constants, ancillary datasets and the treatment of gravity. The work carried out by the team lead by Thierry Leblanc is in fact still ongoing, continuing beyond the ISSI project period. The NDACC lidar working group has accepted the final report (which is available from http://www.issibern.ch/teams/ndacc/ISSI_Team_Report.htm). Peer-reviewed publications derived from this report are to be submitted early 2016 (first drafts were circulated at the end of summer 2015). The code for the calculation of the vertical resolution (both FWHM of a finite impulse response and the digital filter cut-off frequency) has been translated into python and C++ (Matlab, IDL and Fortran versions already existed) and some updates have been made to the IDL version. The work is likely to be extended to other observations beyond ozone and temperature (water vapour/aerosols) and possibly even give rise to a common data processor, which would greatly speed up implementation of the standardisation as especially the treatment of uncertainties might require large adjustments in the processing software programs.

In addition, KNMI provided data processing support to the Lauder stratospheric lidar team.

V Correlative analyses (WP 3)

Sections V.1-V.5 describe the validation activities performed at BIRA-IASB (CCN-1), Section V.6 those carried out at KNMI (CCN-2). We provide brief reports and focus on the main results and conclusions. More technical details and results can be found in the references given in the text below.

V.1 O3 and NO2 column (WP 3.1)

V.1.1 SCIAMACHY nadir O3 total column

The operational SGP 3.01 and 5.02 nadir O3 total column data records were compared extensively to correlative data taken by the Dobson, Brewer and SAOZ ground-based networks. The validation method was published by Lambert et al. (1999, 2000), with later updates by Balis et al. (2007). This allows us to estimate the uncertainties of SCIAMACHY data, as well as their dependence on solar zenith angle, season, latitude, and cloud parameters. Extensive studies of these processors were already performed in the preceding project and reported in the Multi-TASTE final report (Hubert et al., TN-BIRA-IASB-MultiTASTE-FR, 2012). Since then, the validation results studies based on correlative Brewer and Dobson data networks remain unchanged. Those based on the SAOZ sub-network of the NDACC UV-VIS network were modified by about 1% at southern mid-latitudes due to a new version (V3) of the SAOZ algorithm, which now includes climatological air mass factors for the conversion of ozone slant column data into ozone vertical column data and several other improvements.

The mean bias of SGP 3.01 and 5.02 relative to the ground-based network data is positive, +1 to +2%, and the standard deviation of the relative differences ranges between 3 and 10%. There is no systematic difference between the processors, the data lie mostly within about 0.6%. The quality of the SGP data does not change over the network, so there is no clear dependence on latitude. A negative drift in SGP 3.01 ozone column values early on in the mission was noticed at numerous but not all stations. It is the result of a transitory decrease in total ozone in the first couple of years, followed by a more stable behaviour after 2004. The introduction in SGP 5.02 of a degradation correction reduced this negative drift in the tropics, but not at mid to high latitudes. The quality of the total ozone column data furthermore depends on solar zenith angle; an underestimation develops of up to 4% for SZA>80°. In addition, there is a relation with cloud parameters. At individual stations the dependence on fractional cloud cover (CF) is usually within 1-4%, with best agreement generally for small CF and variable agreement for large CF. At low and mid-latitudes small cloud optical depth values tend to increase the standard deviation of the differences. Previous findings are summarised in Table 8.

In the second half of 2015 we performed the ground-based assessment of the new Level-2 prototype SGP 6.00, and made a direct comparison of the results to the operational processor. The delta-validation was based on identical SCIAMACHY pixels in the ~5000 orbit diagnostic data set. This showed that the new baseline produces O3 columns with similar quality. There is no change in the standard deviations of the comparisons, and the dependences on SZA and cloud parameters reported for earlier versions were also observed for SGP 6.00 as well (Figure 9). But there are two clear and important differences. Firstly, SGP 6.00 produces systematically 0.2% to 0.6% less ozone than SGP 5.02, which brings it in better agreement with correlative measurements. The bias of SGP 6.00 to the Dobson, Brewer and SAOZ networks is +1 to +1.5%, comparable to the 1% uncertainty attributable to the ground-based observations. The second change is perhaps worrisome: SGP 6.00 is not as stable after 2004 as the current processor. In the northern hemisphere, where the Dobson and Brewer provide best spatio-temporal sampling, SGP 6.00 ozone columns decrease by about 1.5% over

the mission lifetime relative to correlative measurements, while that is not the case for SGP 5.02 (Figure 10). We therefore anticipate important differences between the trend results of both data records. More details can be found in a dedicated Technical Note (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Sep 2015), which was presented to and discussed within the QWG during meeting #4.

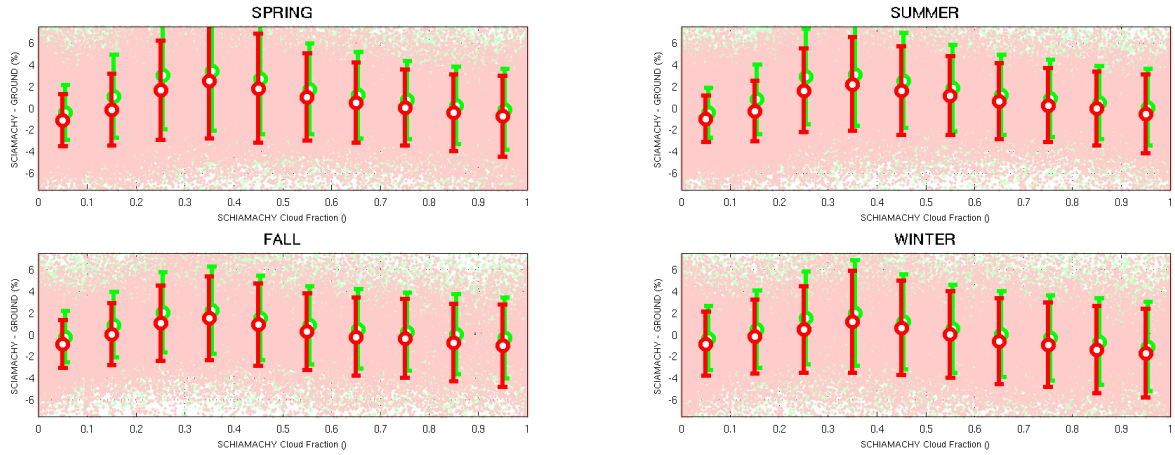


Figure 9: Dependence on fractional cloud cover of the mean and the standard deviation of the relative difference between SCIAMACHY nadir O3 total column and NDACC Brewer instruments at northern mid-latitudes. The delta-validation is based on identical SCIAMACHY pixels retrieved by SGP 5.02 (green) and SGP 6.00 (ref) on the orbits in the diagnostic data set. Each panel shows a different season.

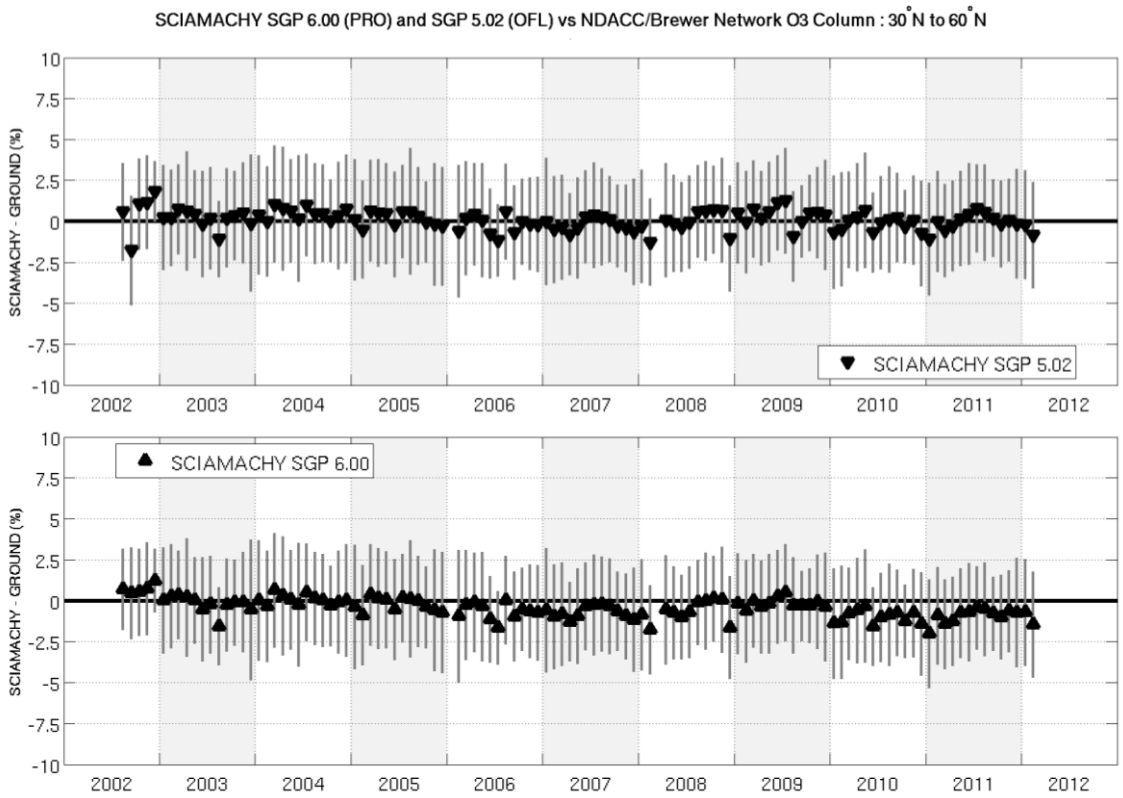


Figure 10: Time series of mean and standard deviation of the relative difference between SCIAMACHY nadir O3 total column and NDACC Brewer instruments at northern mid-latitudes. The delta-validation is based on identical SCIAMACHY pixels retrieved by SGP 5.02 (top) and SGP 6.00 (bottom) on the orbits in the diagnostic data set.

V.1.2 SCIAMACHY nadir NO2 total column

The nadir NO2 total column was compared to correlative data sets collected by ground-based DOAS UV-visible zenith-sky spectrometers performing network operation in the framework of WMO's Global Atmosphere Watch (GAW) contributing network NDACC. The comparison methodology is based on Lambert et al. (2007) and was described extensively in the Multi-TASTE final report (Hubert et al., TN-BIRA-IASB-MultiTASTE-FR, 2012).

The **operational baselines SGP 3.01 and SGP 5.02** generate mutually consistent NO2 column data, which is also consistent with NDACC/UV-visible and GOME GDP 4.1 data records. SGP 5.02 NO2 columns are on average a few 10^{13} to 10^{14} molec. cm^{-2} larger than SGP 3.01, values close to the detection limit of UV-visible spectrometers. In the Northern Hemisphere the direct comparison between SCIAMACHY total column data with NDACC stratospheric column data typically yields apparent positive biases of up to several 10^{15} molec. cm^{-2}. However, the bias is caused by the difference in sensitivity to large concentrations of tropospheric NO2 (e.g. over Europe and Japan). In the Southern Hemisphere, the levels in tropospheric NO2 are lower which enables a more direct comparison between SCIAMACHY and ground-based data. Here SCIAMACHY SGP 3.01/5.02 are biased low with respect to NDACC and GOME GDP 4.1, by about $5x10^{14}$ molec. cm^{-2}. This negative bias exhibits furthermore a seasonal cycle. Our main findings are summarised in Table 9.

In the summer of 2015 we investigated the diagnostic data set reprocessed with **the new prototype processor, SGP 6.00**. Differences between the SGP 5.02 and 6.00 were hardly noticeable and well below the detection limit of the ground-based measurements (Figure 11). At stations without tropospheric pollution and where the diurnal cycle can be accounted for accurately, that is, where direct comparisons between satellite nadir and ground-based zenith-sky measurements provide the most quantitative results, the median agreement ranges within $4x10^{14}$ molec. cm^{-2}. The spread of the absolute difference between SCIAMACHY and NDACC data amounts to a few 10^{14} molec. cm^{-2}, comparable to the combined error bar of the measurements and of the validation method. The enhanced spread at NDACC stations surrounded by pollution sources visible by the satellites (all Northern middle latitude sites, Europe and Japan) and at polar sites where the diurnal cycle is less predictable in spring and winter, is attributable partly to the difference in vertical sensitivity and/or to residual diurnal cycle effects. Our results were documented in a Technical Note (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Sep 2015). which was presented to and discussed within the QWG during meeting #4.

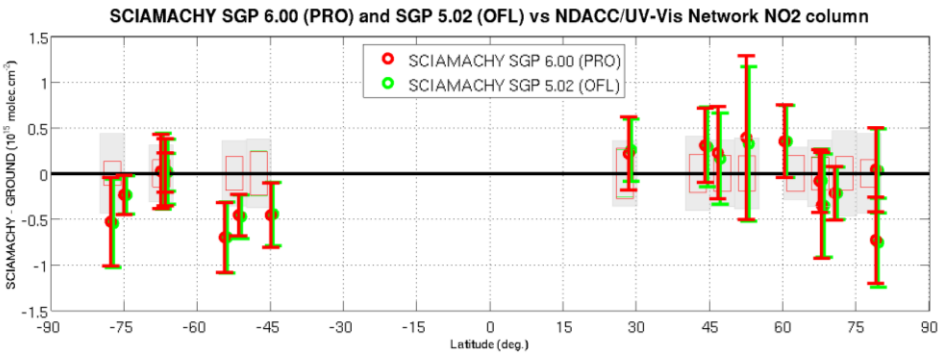


Figure 11: Mean (marker) and 1sigma spread (error bar) of the absolute difference between SCIAMACHY SGP Diagnostic Data Set and NDACC/UV-visible network NO2 column data, as a function of latitude. SGP 5.02 results are depicted in green and SGP 6.00 results in red. The shaded area represents the indicative uncertainty of the NDACC/UV-visible NO2 column measurement. Green and red square contours represent the median value of the uncertainty on SCIAMACHY NO2 column data as reported in SGP 5.02 (green) and 6.00 (red) data files, respectively.

V.2 O3, T profile and H2O (WP 3.2)

All the correlative analyses in this section are based on pseudo-global observations by ozonesonde, stratospheric ozone lidar or stratospheric temperature lidar instruments in the NDACC, GAW and SHADOZ networks. The validation method and processing chain is described in detail in Hubert et al. (2015) for ozone and temperature vertical profiles, and based on du Piesanie et al. (2013) for H2O total columns. The screening, co-location, vertical smoothing and unit conversions are done as specified in the references papers, except when noted explicitly.

V.2.1 GOMOS O3 profile

We performed several GOMOS analyses before project kick-off, before the Multi-TASTE Phase F contracts were signed. We summarise these here, a more detailed account can be found in the Progress Report (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, Apr 2014). In October 2012 we finalized a study on the **influence of GOMOS data screening on IPF 6.01 ozone validation results**. Methods and results are described in a validation report (Hubert et al., TN-BIRA-IASB-GOMOS-IPF6-O3P-SCREENING, Oct 2012), which served as a basis for the discussions on the screening procedure recommended to the users in the GOMOS IPF 6.01 Readme file (GOMOS QWG, ENVI-GSOP-EOGD-QD-12-0117, Dec 2012). We also provided feedback to ESA, in July 2013, on the **quality of IPF 6.01 ozone in the early stages of the mission** (before 11 September 2002). Due to the limited comparison statistics for the period 15 Apr 2002 – 11 Sep 2002 it is very challenging to be conclusive about a change in data quality relative to the rest of the mission. The bias and comparison spread results were not significantly different from those obtained from comparisons over the same period in 2003 or 2004, which made us conclude that if there is a change in data quality in the first few months of the mission it is likely not larger than ~10 %.

In a second stage we verified the data quality of the full mission reprocessing by the **operational processors IPF 5.00/5.01** (and GOPR 6.0c &f) **and IPF 6.01**. The validation method is described in detail in Hubert et al. (AMTD, 2015). The GOMOS data was screened according to the recommendations in the Readme files. The comparison to co-located (500km, 12h) O3 sonde and stratospheric O3 lidar measurements showed that both processors are very similar in terms of bias; overall less than ±5% in the stratosphere, except in the Arctic middle stratosphere (negative by 5-10%), and very positive (>30%) in the UT/LS. The IPF 6.01 comparison data show a spread of 7-10% above 20km and more than 40% in the UT/LS in the relative differences. This is slightly less than the 5.00/5.01 results, indicating that IPF 6.01 profiles are more precise. Our regression analyses showed indications of a long-term time-dependence of the bias in the lower stratosphere, for both GOMOS processors. Ozone values decrease by 5% per decade around 20 km, relative to the sonde and lidar reference. **Caution is therefore needed when using GOMOS data in the UT/LS for long-term studies.** There are no signs of a drift at higher altitudes, between 25-45 km. We also found a time dependence of the bias at smaller scales at mid-latitudes. GOMOS summertime ozone is 2-5% smaller than during winter over the entire stratosphere. There is furthermore no clear dependence on star class of the bias or spread results, see Figure 12, with changes in bias of at most 5% in the middle and upper stratosphere. The results for the IPF 6.01 data set were included in our publication of a comparative assessment of fourteen limb/occultation ozone profile records (Hubert et al., 2015). A summary of the quality of the IPF 6.01 ozone data set can be found in Table 2.

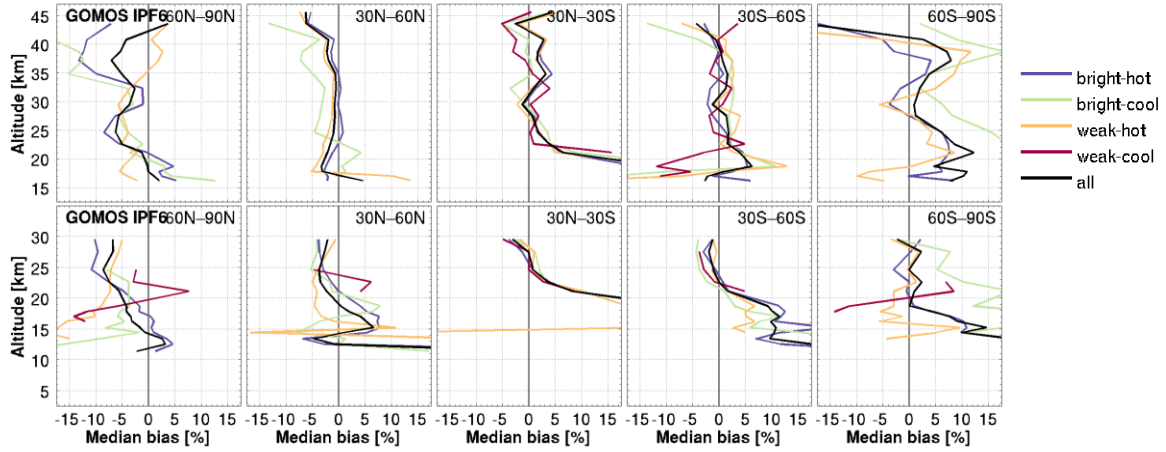


Figure 12: Dependence on star class (bright: magnitude<1.9; weak: magnitude≥1.9; hot: temperature≥7000K, cool: temperature<7000K) of the median bias of GOMOS IPF 6.01 relative to NDACC stratospheric ozone lidar (top row) and NDACC/GAW/SHADOZ ozonesonde (bottom row) instruments in five latitude bands. The black lines shows the complete co-location data set.

GOMOS observed in the stellar occultation and limb geometries. The operational processors retrieve information only from occultations, but FMI developed **a scientific processor (GBL 1.2)** which handles the **bright limb measurements**. This additional data set provides complementary spatio-temporal coverage to the IPF data record. We originally performed ozonesonde and lidar-based correlative analyses of both GOMOS records within ESA's Ozone_cci project. The results were presented to and discussed within the QWG (N4, September 2015). We circulated shortly afterwards a Technical Note around the QWG (Hubert et al., TN-BIRA-IASB-MultiTASTE-PhaseF-GOMOS-O3p-IPF-GBL-1C, October 2015), with graphics of the dependence of bias and comparison spread of both processors on a suite of geophysical and instrument-related variables. Our analysis results were inline with earlier reports by Tukiainen et al. (AMT, 2015) and Hubert et al. (AMTD, 2015), but provided more detailed technical feedback to the QWG. The GBL ozone profiles exhibit a bias less than ±5% in LS and MS, increasing to -10% in the US, and changing sign around the stratopause. There are signs of a dependence of bias on season (or SZA?) and strong indications for a negative drift below 25km. The latter was not expected observation by the QWG, since the negative drift of IPF ozone profiles in the LS was assumed to be caused by an incomplete correction for the (time-dependent) dark charge. The GBL measurements should be less affected by such an incomplete correction, so this indicate that another, common cause may be responsible for the instability in the Level-1b data record.

V.2.2 MIPAS T profile

We consolidated the correlative analyses of temperature profiles retrieved by the **operational Level-2 processors (IPF 5.05/5.06 and ML2PP 6.0)**, both using the IPF 5.05/5.06 Level-1b data as input. The two Level-2 records were compared to NDACC/GAW/SHADOZ sonde and NDACC temperature observations, following the methodology of Hubert et al. (2015). MIPAS data are handled according to the recommendation of the QWG (MIPAS IPF 5 and ML2PP 6 Readme files): prescribed data screening procedure, use pressure as vertical coordinate, smoothen higher-resolution profiles with MIPAS vertical averaging kernels. The temperature bias and comparison spread differs for both MIPAS periods (Full Resolution: 2002-2004, Optimized Resolution: 2005-2012). The FR temperatures are about 0.5-1 K warmer than OR values at most latitudes in the LS and MS (Figure 13). The sign of the FR/OR bias is opposite in LS/MS for 30°S-60°S and in US/LM for 30°N-60°N. In the polar MS/LS and at mid-latitudes the spread in the absolute differences is smaller during the OR period, with a reduction of about 0.2-0.4 K. From Figure 13 we conclude that MIPAS underestimates temperature by 0.5-1 K in most of stratosphere and lower mesosphere. There is a dependence of the bias on latitude and

pressure, and for the observed spread in the differences too. We also found a **very pronounced annual cycle** in the bias for all data versions at mid and high latitudes. The annual cycles for day- and night-time MIPAS measurements have opposite phases, but the phase does not depend on latitude. The effect is largest at mid-latitudes, somewhat smaller in the polar regions, and not seen in the tropics. The sonde-derived bias is 1-2 K smaller (more negative) in June than in December, while the lidar-derived bias is ~2K larger (more positive) in June than in December (Figure 14). Also the spread in the comparisons changes over the year, especially at the poles and at mid latitudes. The spread is minimal in local summer and maximal in local winter, for the sonde and the lidar samples. Table 3 presents a summary of our quality assessment of ML2PP 6.0 temperature, which is virtually identical to that for IPF 5.05/5.06.

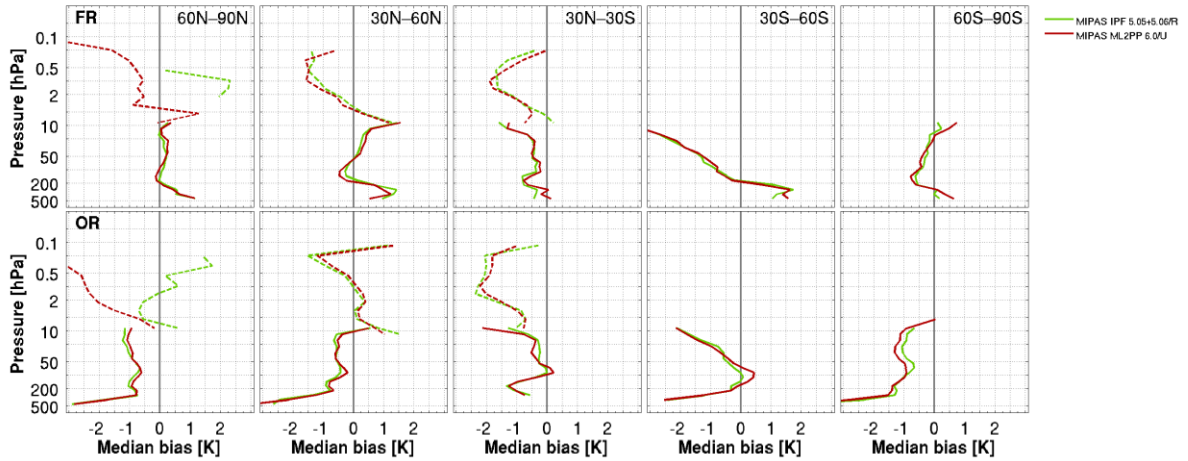


Figure 13: Median absolute difference of MIPAS IPF 5.05/5.06 (green) and ML2PP 6.0 (red) temperature relative to NDACC/GAW/SHADOZ sonde (solid) and lidar (dashed). Results are separated for two periods (FR: 2002-2004; OR: 2005-2012) and for five latitude bands. Negative bias values indicate that MIPAS temperature is cooler than correlative measurements.

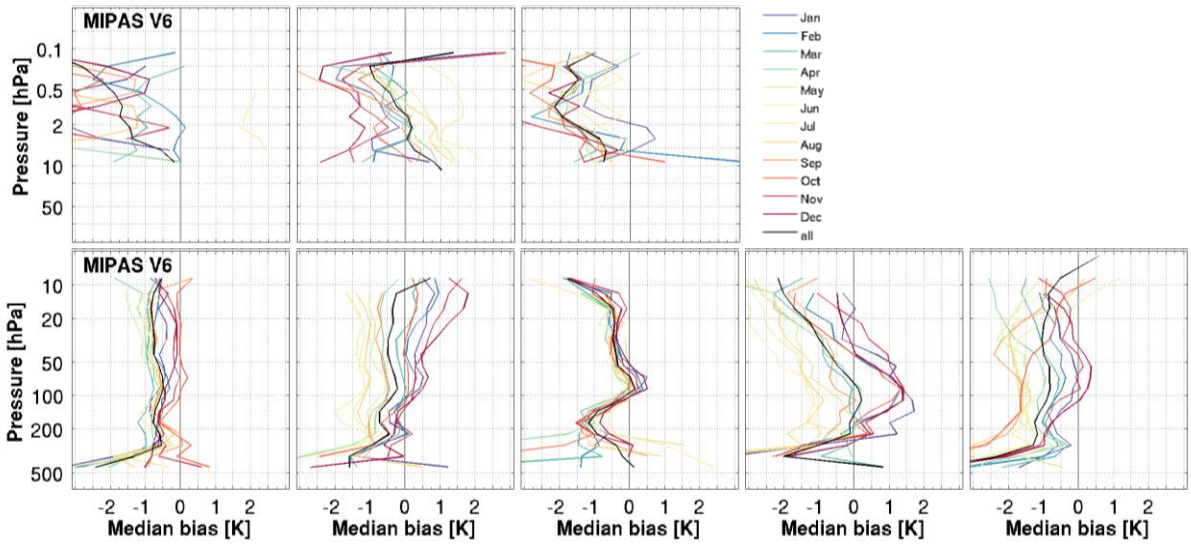


Figure 14: Dependence on month (colours) of the median absolute difference of ML2PP 6.0 temperature relative to NDACC lidar (top) and NDACC/GAW/SHADOZ ozonesonde observations in five latitude bands (from left to right: 60-90°N, 30-60°N, 30°N-30°S, 30-60°S, 60-90°S). Negative bias values indicate that MIPAS temperature is cooler than correlative measurements.

The dependences described above (latitude, pressure, FR/OR, month) are also seen for the **prototype processor ML2PP 7.01** developed during the project. Compared to the previous operational processors, the retrievals are done on new Level-1b data (IPF 7.11) and with a different set up of the Level-2 processor (changed microwindows, continuum, …). The time-dependent non-linearity corrections changed in the V7 Level-1b processor, which were anticipated to impact the temperature retrievals and subsequently, potentially, the trace gas retrievals as well. An extensive analysis of the ~4000 DDS orbits showed that the V7 temperature data are indeed drifting to higher values than the V5 and V6 data during the OR period (Figure 13). The bias increases by about 0.3 K in the MS and by 1 K in the US/LM between 2005 and 2012. This change is an improvement in the US/LM, there V7 temperatures are more stable relative to lidar observations. At lower altitudes the situation is less clear. A clear change in bias was also noticed in the FR period (2002-2004). V7 temperatures are generally 0.5-1 K cooler than V5 or V6, increasing the (negative) bias relative to ground-based observations larger at most pressure levels. The outcome of the delta-validation was discussed during QWG meetings #36 and 37.

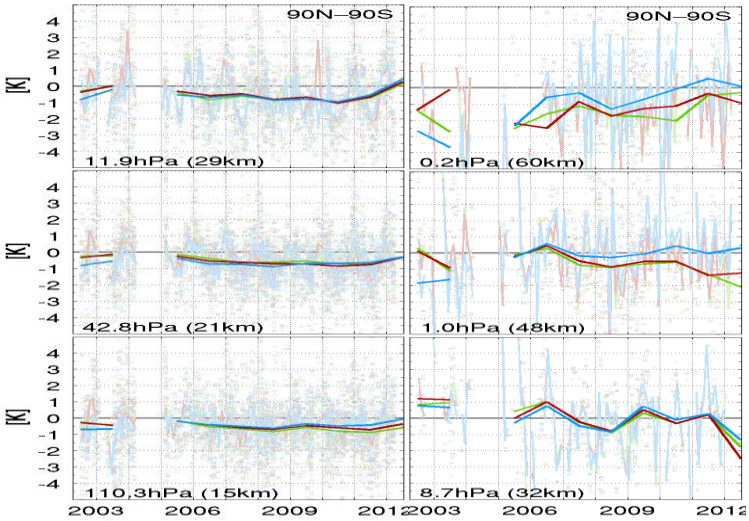


Figure 11: Time series of median bias of MIPAS temperature IPF 5.05/5.06 (green), ML2PP 6.0 (red) and ML2PP 7.01 (blue) relative ozonesonde (left) and lidar (right) measurements. Negative values indicate that MIPAS temperatures are cooler than ground-based values. Results are shown at monthly (light) and annual (dark) time scales at several pressure levels. Only identical MIPAS profiles from DDS orbits are used for the delta-validation.

V.2.3 MIPAS O3 profile

The Progress Report (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, April 2014) contains a study of the **consistency of the MIPAS validation results** in different profile representations (VMR or number density on fixed altitude or pressure levels), for different correlative instruments, and for the validation methodology of BIRA-IASB and KNMI. This dedicated analysis was requested by the MIPAS QWG and the findings were presented to and discussed during QWG meeting #33.

We also consolidated our ozonesonde- and lidar-based correlative analyses of the current (ML2PP 6.0) and previous (IPF 5.05 and 5.06) versions of the operational Level-2 processors. These showed that the **MIPAS V5 and V6 Level-2 ozone profile records are very similar** in terms of bias, spread and their dependence on geophysical (pressure, latitude, month, year, …) and instrument-related parameters (resolution mode, measurement mode). Both V5 and V6 Level-2 records are retrieved from the same Level-1b record (IPF 5.05 and IPF 5.06). The handling of MIPAS data is according to the recommendation of the QWG (MIPAS IPF 5 and ML2PP 6 Readme files): prescribed data screening procedure, use VMR profiles on pressure levels, smoothen higher-resolution profiles with MIPAS vertical averaging kernels. The analysis methodology is that of Hubert et al. (2015). A clear difference in bias is found between the Full Resolution (2002-2004) and Optimized Resolution (2005-2012) periods. For p<20hPa the FR ozone is about 5% smaller than OR ozone; for p>50 hPa FR is 5% larger than OR ozone (Figure 15). As a result, it is mandatory to account for the (altitude-dependent)

FR-OR bias in long-term trend analyses covering the entire mission. MIPAS generally has a positive bias in the MS and US of about 5% relative to ground-based measurements in both data taking periods. In the UT/LS the positive bias is even larger, while in the Arctic it is closer to zero. The spread in the comparisons is not very different for FR and OR. It is minimal between 10-50 hPa (5-7%) and increases towards the stratopause (~10%) and especially the tropopause (>40%). In local summer the spread is smaller than during local winter, this is especially clear at high latitudes. During Antarctic ozone hole conditions the spread at 80hPa increases from ~15% to >40%. The bias changes by up to ~5% with season as well, but no clear pattern could be found. Table 4 gives a quantitative summary for ML2PP 6.0 for the different latitude bands and altitude regions. The V5 and V6 ground-based comparisons differ generally less than 1-2% in bias; the comparison spread is at most ~1% smaller for V6. The ex-ante random uncertainty reported for V6 higher than for V5 in the UT/LS, but not more than 2-3%. The results for the ML2PP 6.0 data set were included in our publication of a comparative assessment of fourteen limb/occultation ozone profile records (Hubert et al., 2015).

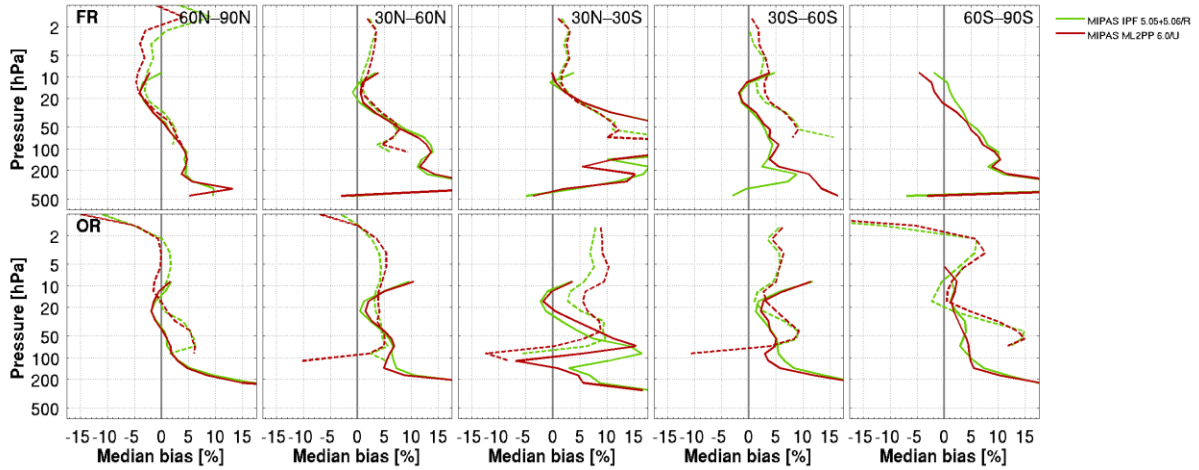


Figure 15: Median relative difference of MIPAS IPF 5.05/5.06 (green) and ML2PP 6.0 (red) ozone relative to NDACC/GAW/SHADOZ ozonesonde (solid) and lidar (dashed). Results are separated for two periods (FR: 2002-2004; OR: 2005-2012) and for five latitude bands. Positive bias values indicate that MIPAS ozone is larger than correlative measurements.

New Level-1b (IPF 7.11) and Level-2 (ML2PP 7.01) processors were developed during the project and were subject to extensive studies by various teams. A delta-validation was carried of the same V5, V6 and V7 profiles for the subset of ~4000 orbits in the newly defined diagnostic data set (Section IV.2.2). The ground-based validation identified minor changes in V7 ozone. The prototype systematically produces 1-2% more ozone than previous versions between 5-50 hPa, and in the Tropics even 3% for p<20hPa (Figure 15). This means that V7 ozone has a higher bias than V5&V6 relative to ground-based measurements. The observed comparison spread, on the other hand, is identical for all versions. The V7 ex-ante random uncertainty is up to 5% larger than the previous processors in the troposphere. Furthermore, all processors have the same dependence on geophysical and instrument-related parameters. The positive drift in V7 temperature (relative to V6) during the OR period (see Section V.2.2) does not have a clear impact (relative to V6) on the stability of V7 ozone. The changes w.r.t. to V6 are less than 1% over the 2005-2012 period. Detailed results were presented during QWG meetings #36 and #37.

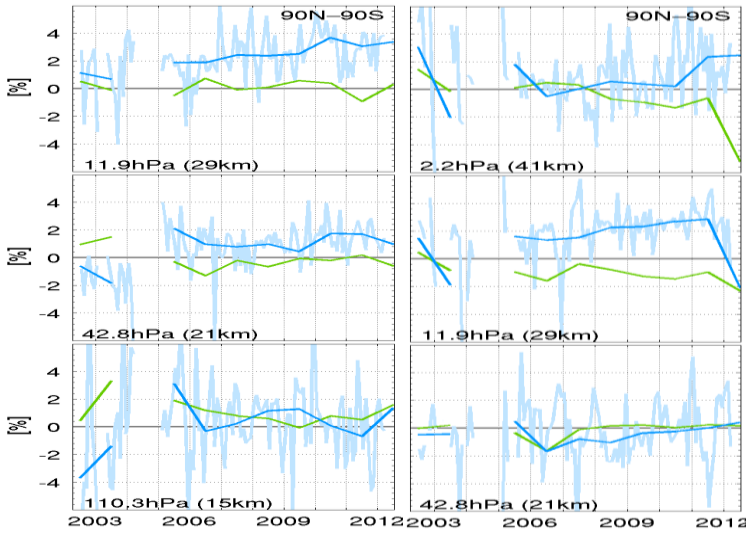


Figure 13: Time series of difference in ground-based median bias of MIPAS ozone IPF 5.05/5.06 (green) and ML2PP 7.01 (blue) relative to that of ML2PP 6.0. Positive values indicate that higher ozone values are found than for V6. Results are shown at monthly (light) and annual (dark) time scales, for the ozonesonde (left) and lidar (right) network at several pressure levels. Only identical MIPAS profiles from DDS orbits are used for the delta-validation.

V.2.4 SCIAMACHY limb O3 profile

SCIAMACHY's **operational processors SGP 3.01 and SGP 5.02** were subjected to a comprehensive analysis using the NDACC/GAW/SHADOZ ozonesonde and NDACC lidar networks. The analysis methodology was recently published by Hubert et al. (2015). It turns out that SCIAMACHY's bias patterns are more pronounced than many other limb/occultation ozone profilers (Hubert et al., 2015). The latitude-altitude structure of the bias is quite intricate, with clear differences between northern and southern hemisphere, and between LS and US (Figure 16). We established that, above 30 km, SCIAMACHY ozone drifts to lower ozone relative to lidar measurements, by up to 8% per decade. In the MS, there are indications of a positive drift of 3% per decade, which is just below the 5% significance threshold. At smaller timescales, the data quality depends on the month of measurement (Figure 17), although SZA may be the underlying driving factor as well. The scan angle seems to impact the bias around 25 km, with East states leading to up to 5% more ozone than West states.

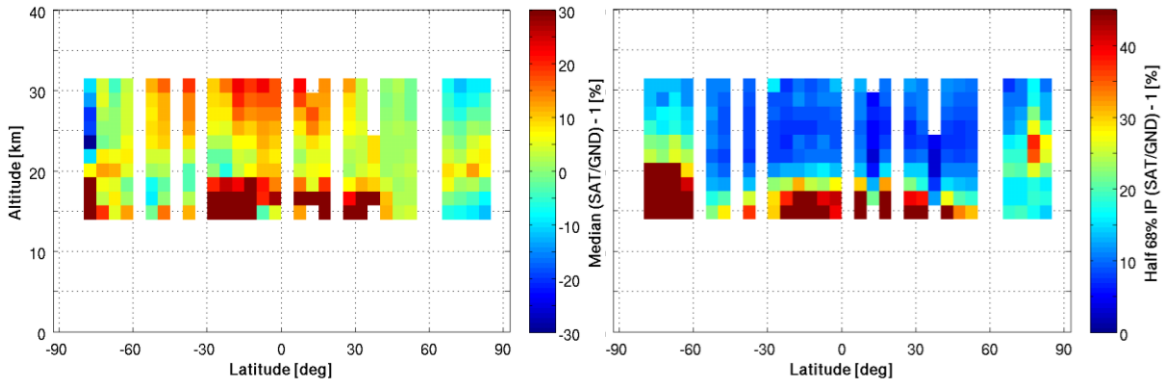


Figure 16: Meridional structure of the median bias (left) and comparison spread (right) of SCIAMACHY SGP 5.02 limb ozone relative to NDACC/GAW/SHADOZ ozonesonde. Only latitude bins with more than 10 comparison pairs are shown. Positive bias values indicate that SCIAMACHY ozone is larger than correlative measurements.

All quality features described above are noticed for both operational data records, but **there are noteworthy differences**. To begin with, the latest processor produces, on average, about 10% more ozone than SGP 3.01. Second, the quality of SGP 5.02 is clearly much worse in the Arctic than for other zonal regions (Figure 17). Around Arctic winter (DJF), its bias and comparison spread reach extreme values between 20-25km. So SGP 5.02 Arctic profiles should probably be avoided or used

with care, while that is not needed for SGP 3.01 data. Third, we identified serious issues in the SGP 5.02 vertical averaging kernels. The vertical behaviour of the relative differences degrades significantly when the correlative profile data are smoothed with the AK. Bias profiles exhibit vertical oscillations between neighbouring grid levels of up to 5% in amplitude, over the entire stratosphere and at all latitudes. The comparison spread has similar vertical oscillations, but solely in the polar middle stratosphere. Such an oscillating behaviour was not seen for AK-smoothed SGP 3.01 comparisons, and therefore likely an artefact of the use of a vertical retrieval grid by SGP 5.02 that is finer than the actual scan locations. The SGP 5.02 AKs are questionable and we therefore discourage their use.

More detailed results can also be found in a technical note (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Sep 2015) and, for SGP 5.02, also in our publication of a comparative assessment of fourteen limb/occultation ozone profile records (Hubert et al., 2015). Table 13 summarises our main conclusions for the operational processors.

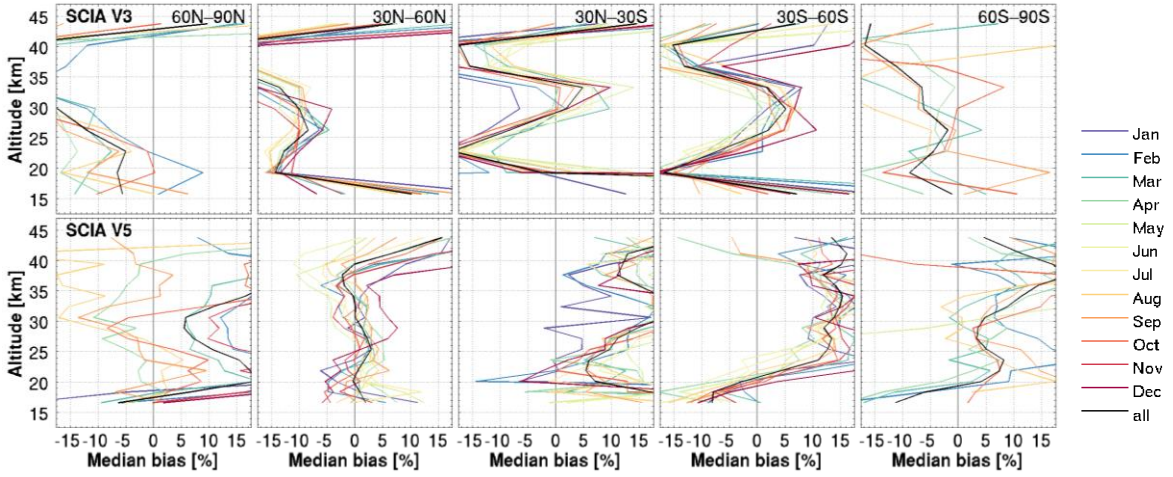


Figure 17: Dependence on month (colours) of the median relative difference of SCIAMACHY SGP 3.01 (top) and SGP 5.02 (bottom) limb ozone relative to NDACC lidar. Positive bias values indicate that SCIAMACHY ozone is larger than correlative measurements.

The QWG developed the new Level-2 prototype processor SGP 6.00 during the project. The diagnostic data set, delivered in April 2015, was used for the ground-based assessment of the prototype and its comparison to identical measurements from the earlier processors SGP 3.01 and SGP 5.02. The most important change in SGP 6.00 is the extension of the O3 profile retrievals to the mesosphere, which necessitated the combination of the radiances from all four measurement states per scan sequence. As a result, the horizontal coverage of the new product is four times smaller while the vertical coverage is larger. Our comparisons to the ozonesonde and lidar networks showed that SGP 6.00 ozone data quality represents only a minor improvement relative to that of SGP 5.02. In some regions of the atmosphere slight improvements were found for the bias, the short-term variability and the estimates of random uncertainty (Figure 18). None of the major quality issues observed for SGP 5.02 (see above) are truly addressed by the new processor, in line with expectations since not the main target during algorithm development. We therefore recommended the QWG to proceed with the reprocessing of limb ozone profiles of the entire mission. The results were presented to and discussed by the QWG during meetings #3 and #4. Exhaustive details of the delta-validation can be found in a Technical Note (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Sep 2015). Initial and consolidated results were presented to and discussed within the QWG during meetings #3 and #4.

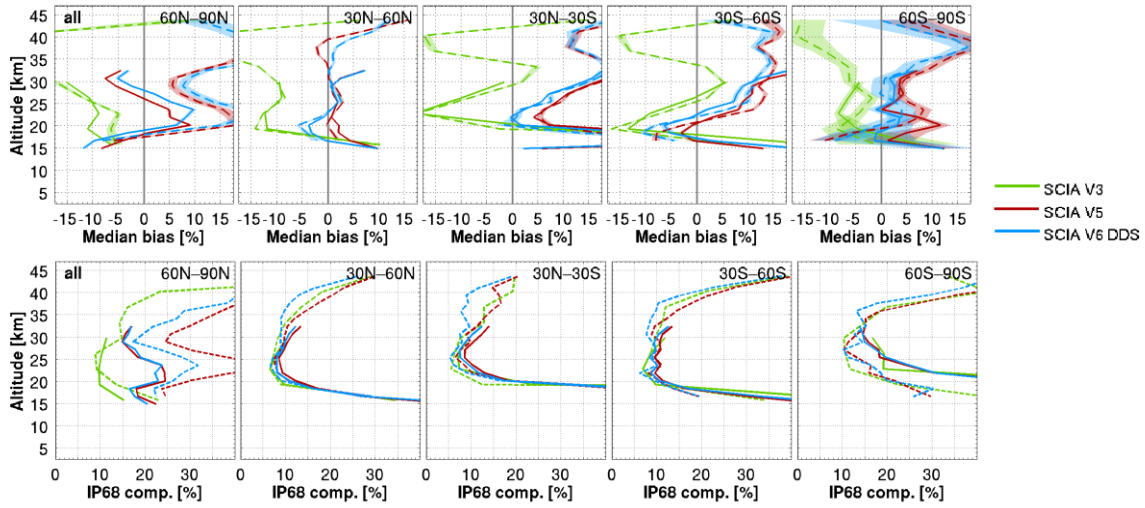


Figure 18: Statistics of the difference of SCIAMACHY SGP 3.01 (green), SGP 5.02 (red) and SGP 6.00 (blue) limb ozone relative to NDACC/GAW/SHADOZ ozonesonde (solid) and lidar (dashed) in five latitude bands. Shown are the median (top row; positive values indicate a high bias of SCIAMACHY relative to ground-based instruments) and the 68% interpercentile (bottom).

V.2.5 SCIAMACHY nadir H2O total column

Total columns of water vapour are retrieved from nadir observations by the operational Level-2 processor since SGP 5.02. The SCIAMACHY water vapour columns were extensively compared to data taken by the radiosonde coupled to NDACC/GAW/SHADOZ ozonesonde flights. The relative humidity (RH) data of the radiosonde is first converted to H2O VMR using the expression of saturated water vapour pressure by Sonntag et al. (1994). VMR is then integrated in the pressure domain to obtain a column of water vapour (in g cm^{-1}). We stop the integration at 10 km since (a) most of the H2O column is in the lower stratosphere, and (b) the accuracy of most humidity sensors degrades once exposed to very cold temperatures (e.g. around the tropopause or in ice clouds). Our validation methodology follows that of du Piesanie et al. (2013).

The full mission data record was available for the validation of the **operational Level-2 processor SGP 5.02**. Hence we could use very tight co-location criteria: all SCIAMACHY pixels falling within 50 km and 1 h of the location and time of a balloon launch. This leads to 10437 co-located pairs, evenly spread over land and ocean (Table 19). Such a stringent window reduces sampling mismatch uncertainties induced by the large spatio-temporal variability of the H2O field, estimated for RH by Sun et al. (2010) to be 3.3% per 3 h and 3.1 % per 100 km of mismatch.

Table 19: Overview of the validation results for SCIAMACHY SGP 5.02 nadir H20 total columns, from comparisons to NDACC/GAW/SHADOZ sonde data. We provide co-location statistics and comparison statistics for the entire data set, and for four disjoint SCIAMACHY pixel classes.

SGP 5.02		# pairs	Fraction	Median bias		Comparison spread	
			(%)	(g cm^{-2})	(%)	(g cm^{-2})	(%)
All	All	10437	100	-0.04	-6	0.32	30
Land	Cloud free	1153	11	+0.23	+20	0.31	25
	Cloud	4403	42	-0.06	-9	0.33	28
Ocean	Cloud free	420	4	-0.06	-6	0.28	17
	Cloudy	4452	43	-0.07	-12	0.29	31

Different SCIAMACHY pixel classes were considered, based on whether representing land or ocean, and whether there are clouds or not (OCRA cloud fraction >0 or =0). We then calculated median and 68% interpercentile of the absolute differences SCIAMACHY minus radiosonde. Co-location and comparison statistics for each SCIAMACHY pixel class can be found in Table 19. SCIAMACHY data is in general too dry, by about 0.06 g cm^{-2} (or ~10%), but not for cloud free pixels over land, where a wet bias is seen of about 0.23 g cm^{-2} (or 20%). The spread in the comparisons is substantial, ranging from 0.28 to 0.33 g cm^{-2} (or 17-31%). In a second phase, we investigated the dependence of data quality with a number of geophysical parameters (cloud fraction, cloud top height, cloud optical thickness, solar zenith angle, AMF correction factor, day of year, time, latitude). The cloud information used here originates from OCRA and SACURA and taken from the SCIAMACHY product. We confirm most conclusions drawn by du Piesanie et al. (2010) whose analysis was based on 18 months of SCIAMACHY data (Feb 2010 to Aug 2011):

- There is a weak relation between bias and cloud cover. Land pixels are too wet compared to correlative data for CF<10%, and too dry in more cloudy conditions, with a maximum in the dry bias at CF=20% (Figure 40). Also the bias over ocean is largest (most negative) for little cloud cover.
- There is no clear relation between bias and cloud optical thickness (Figure 41), but for very cloudy pixels (CF>0.9) the cloud top height plays a role (Figure 42). SCIAMACHY data become increasingly too dry (and more variable) for cloud top heights between 2 and 6 km. The situation is less clear for CTH>7 km.
- There is a seasonal cycle in the comparison spread, with lowest values in local spring (~0.1 g cm^{-2}) and maximal values in local summer (>0.4 g cm^{-2}). Figure 19 shows that the bias changes over the course of a year by at most ~0.4 g cm^{-2} for cloudy ocean pixels and ~0.2 g cm^{-2} for the other pixel classes. SCIAMACHY H2O values are dryer than correlative data in winter and autumn than in other months.
- For AMF correction factors below 1.1, the bias and variability of cloudy pixels increase very clearly down to the 0.8 threshold set by the data provider (Figure 38). Also cloud free ocean pixels tend to develop larger negative biases at small AMF corrections, although not as pronounced as the cloudy pixels.

We noticed that the data quality changes with SZA, whereas du Piesanie et al. (2013) did not report such relation. For high SZA there is almost no dependence of bias with SZA, absolute differences are negative on average. But Figure 20 shows that at small SZA values, around 30°-40°, the bias changes sign and increases to about +0.05 and +0.10 g cm^{-2} at SZA=25°. Also the spread increases with decreasing SZA. Table 12 summarises our conclusions of the SGP 5.02 analysis, Annexe A5.1 contains supplementary figures from the analysis.

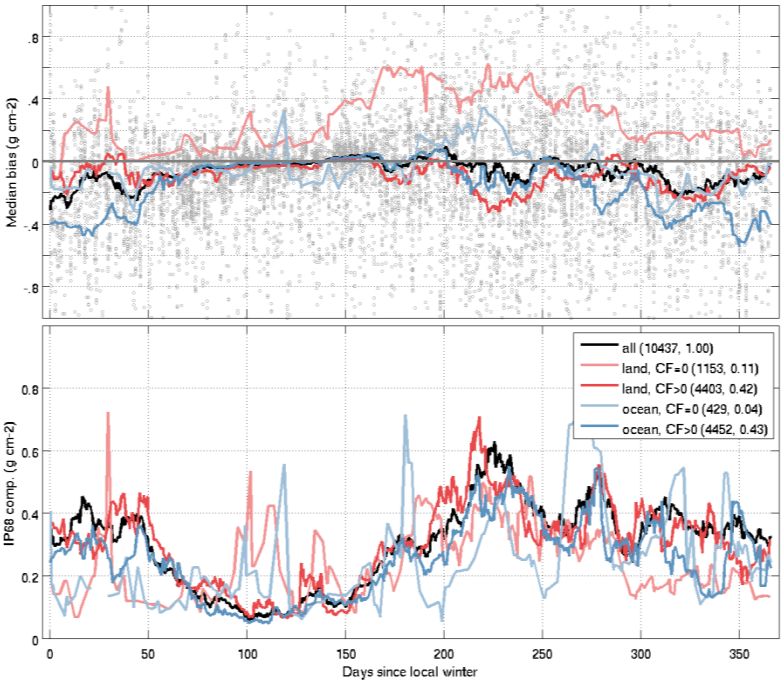


Figure 19: Seasonal dependence of the median (top) and 68% interpercentile (bottom) of the absolute difference distribution of SCIAMACHY SGP 5.02 nadir H2O total column minus radiosonde measurements. Grey markers shows the entire co-location sample, curves represent running median statistics for five SCIAMACHY pixel classes. The x-axis depicts the number of days since the start of local astronomical winter.

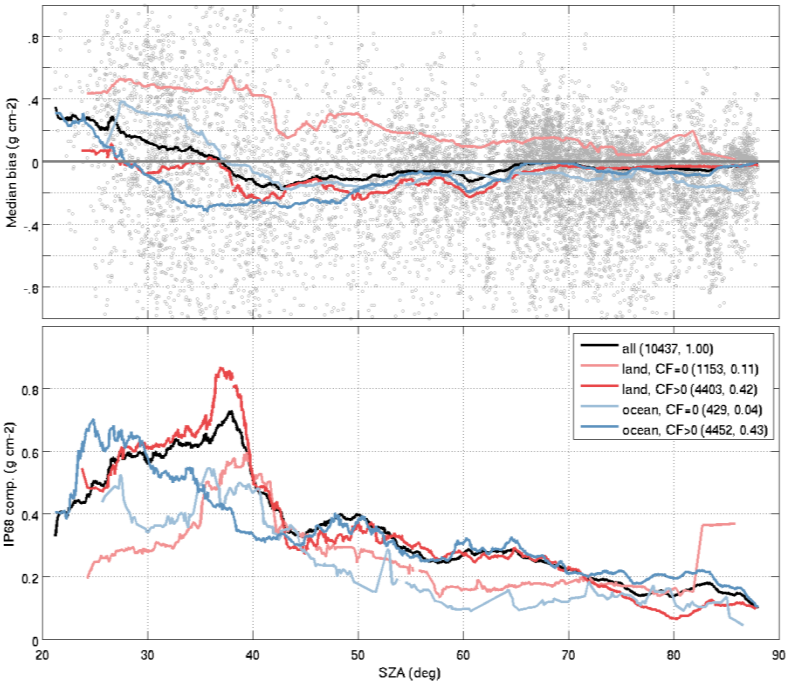


Figure 20: Dependence on SZA of the median (top) and 68% interpercentile (bottom) of the absolute difference distribution of SCIAMACHY SGP 5.02 nadir H2O total column minus radiosonde measurements. Grey markers shows the entire co-location sample, curves represent running median statistics for five SCIAMACHY pixel classes.

The prototype Level-2 processor SGP 6.00 was developed during the project, and a DDS was provided to the validation teams in April 2015. To allow for a sufficient number of co-locations the criteria were loosened with respect to that of the analysis described above. The delta-validation of SGP 6.00 uses only the SCIAMACHY pixels closest to the radiosonde launch location and time, within a radius of 100 km and a period of 3 h. As a result the uncertainty due to co-location mismatch is expected to be a few percent larger than in previous analysis. In total, 2170 pairs were selected, about 1/3 over ocean and 2/3 over land (Table 20). Very little data was available for cloud free conditions, so results may be less representative under this circumstances (bracketed values in the table).

Table 20: Overview of the validation results for SCIAMACHY SGP 6.00 DDS nadir H20 total columns, from comparisons to NDACC/GAW/SHADOZ sonde data. We provide co-location statistics and comparison statistics for the entire data set, and for four disjoint SCIAMACHY pixel classes.

SGP 6.00 DDS		# pairs	Fraction (%)	Median bias ($g\ cm^{-2}$)	Median bias (%)	Comparison spread ($g\ cm^{-2}$)	Comparison spread (%)
All	All	2170	100	-0.04	-6	0.39	30
Land	Cloud free	134	6	+0.22	+17	0.30	20
	Cloud	1239	57	-0.07	-7	0.46	31
Ocean	Cloud free	(29)	(1)	(-0.02)	(-4)	(0.14)	(20)
	Cloudy	768	35	-0.05	-9	0.30	29

There seems to be little difference in the quality of the SGP 6.00 and its predecessor. This is in line with expectations, since no implementation changes were done for the H2O column product in the prototype. The SGP 6.00 bias is positive (+0.22 g cm^{-2} or +17%) for cloud free land pixels, and negative for the other pixel classes (about -0.06 g cm^{-2}, or ~8%). The spread in the comparisons ranges between 0.30 and 0.46 g cm^{-2} (20-30%). Details can be found in Table 20.

All dependences of SGP 6.00 data quality are quite similar to those of SGP 5.02: relation with cloud top height (for very cloudy conditions, Figure 49), with AMF correction factor (Figure 45), with cloud fraction (Figure 47), with season (Figure 50) and with solar zenith angle (Figure 46). We believe that apparent differences are most likely related to the smaller sample. Figure 21 illustrates the dependence on AMF correction factor, all other relevant graphics can be found in Annexe A5.2.

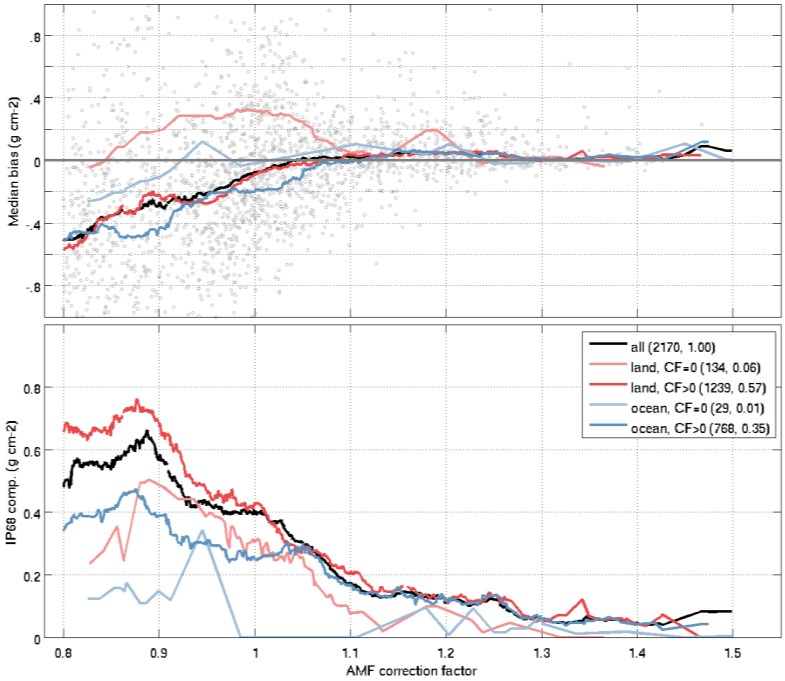


Figure 21: Dependence on AMF correction factor of the median (top) and 68% interpercentile (bottom) of the absolute difference distribution of SCIAMACHY SGP 6.00 DDS nadir H2O total column minus radiosonde measurements. Grey markers shows the entire co-location sample, curves represent running median statistics for five SCIAMACHY pixel classes.

V.3 BrO column and profile (WP 3.3)

V.3.1 SCIAMACHY nadir BrO total column

The retrieval of nadir BrO total columns was added to the operational processing with the SGP 5.01 baseline. We compared the SCIAMACHY BrO columns retrieved by SGP 5.01 and 5.02 to ground-based UV-visible zenith-sky measurements at Harestua, Norway (60°N, 11°E), using the methodology of Hendrick et al. (2009). In order to ensure the photochemical matching between satellite and ground-

based observations, sunrise ground-based columns were photo-chemically converted to the satellite overpass SZAs using a stacked box photochemical model (Hendrick et al., 2007 and 2009). For our final **validation analysis of full mission SGP 5.02 data** we used the total AMFs calculated from retrieved ground-based profiles to compute the vertical columns from SCIAMACHY slant column densities (SCD), and not the stratospheric AMFs included in the SGP product. Doing so improves the bias to ground-based data by about 5%, and leaves the spread essentially unchanged. SGP 5.02 has a negative bias relative to the Harestua UV-visible instrument of -12.8% (or 6.4×10^{12} molec. cm^{-2}) the standard deviation in the comparisons is 37.4% (or 19.5×10^{12} molec. cm^{-2}). The annual BrO cycle is well reproduced. However, there are clearly outlying measurements in the SGP 5.02 data record in 2003-2004 and 2007 (Figure 22).

In a second phase, we investigated the quality of the **new SGP 6.00 prototype**, using the same method as above. SCIAMACHY BrO column data of the diagnostic data set were compared to the UV-visible instrument at Harestua, and these were subsequently compared that those of the full mission validation of SGP 5.02. The analysis is not a delta-validation in the strict sense, and the differences observed between the data versions may be partially explained by differences in sampling. The bias of SGP 6.00 is slightly better than (though not significantly different from) its predecessor, -12.2% (6.1×10^{12} molec. cm^{-2}). The variability in the comparisons for both processors is generally similar, even though the opposite is suggested since the overall spread was nearly halved, to 17.4% (or 8.5×10^{12} molec. cm^{-2}). The latter is due to the absence of severe outliers, which may be a result of the different sampling (Figure 22). Overall, it is quite clear that the data quality of the SGP 6.00 prototype is it least as good as that of the SGP 5.02 processor at the Arctic station of Harestua. Our conclusions are summarised in Table 10 and detailed in a Technical Note (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Sep 2015). The latter was presented to and discussed within the QWG during meeting #4.

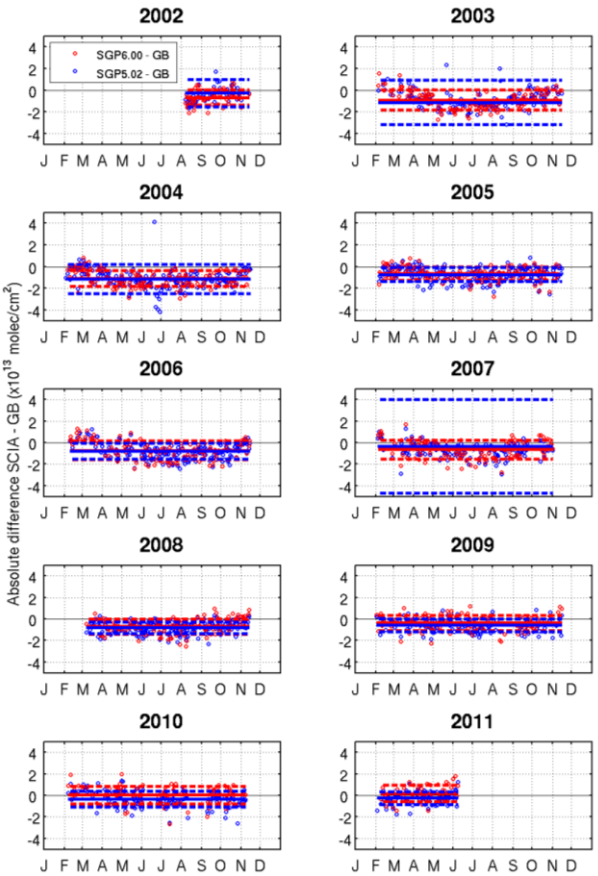


Figure 22: Time series of the absolute difference between SCIAMACHY nadir BrO total columns and ground-based UV-visible zenith-sky observations at Harestua (60°N, 11°E), for the SGP 5.02 processor (full mission) and the SGP 6.00 prototype (DDS).

V.3.2 SCIAMACHY limb BrO vertical profile

Also the limb BrO profiles were compared to the UV-visible zenith-sky measurements at Harestua station in Norway (60°N, 11°E), for the **SGP 5.02 baseline (full mission) and SGP 6.00 prototype (DDS)**. The validation method was published by Hendrick et al. (2009). Ground-based profiles are photo-chemically converted to the solar zenith angle of SCIAMACHY tangent points within 500 km from the station, for days where both instruments provide a measurement. All SCIAMACHY profiles are smoothed with the ground-based vertical averaging kernels, to eliminate the uncertainties due to differences in vertical resolution. The ground-based BrO profiles between December and mid-February were discarded from our analysis, since they have significantly larger uncertainties during this period. Only common coincidences between ground-based and SGP 5.02 or SGP 6.00 observations were selected, representing a total of 2833 morning coincidences from August 2002 to March 2012. This analysis was reported in detail in a Technical Note (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Sep 2015), we repeat our main conclusions in the next paragraph as well as in Table 14.

The SGP 5.02 and SGP 6.00 appear very similar when compared to ground-based profiles (Figure 23). Both show a positive bias at the lower altitude levels (15-18 km for late winter/early fall, and 15-21 km for late spring/summer/early fall), and a negative bias up to -30% (late spring/early fall) or -50% (late spring/summer/early fall) at higher altitude levels. Remarkably, when the SGP profile data are integrated to 15-27 km partial columns they do not exhibit an annual cycle, while it is clearly present in the ground-based data. When compared to the validation results of the scientific processor by IUP Bremen, we find that SGP biases are significantly larger at the same station (+10 to -20%). Also the annual cycle is well captured by the IUP-Bremen scientific product (Hendrick et al., 2009). The operational products seem therefore of lesser quality than the scientific product, at least at an Arctic location such as Harestua.

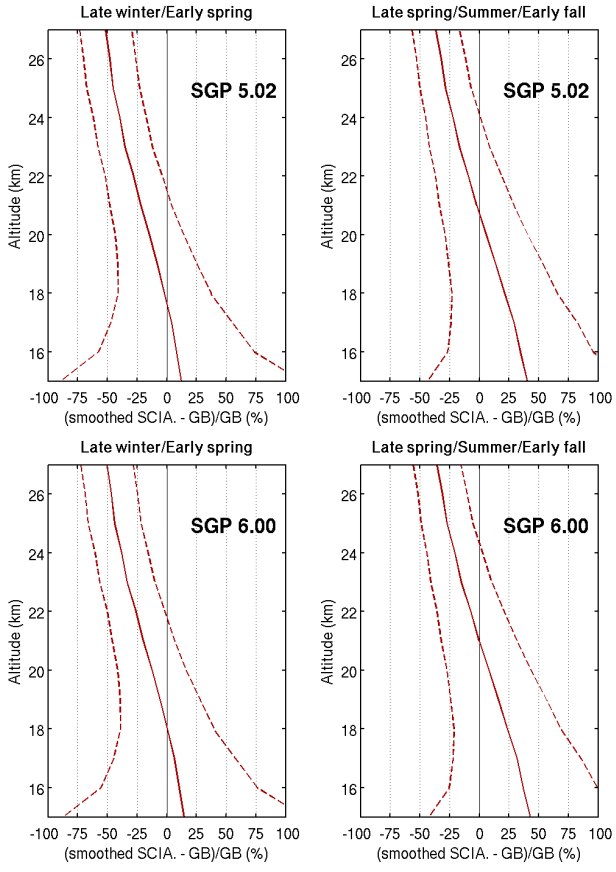


Figure 23: Relative difference between smoothed SCIAMACHY SGP 5.02 (top) or SGP 6.00 (bottom) and ground-based UV-visible profiles at Harestua (60°N, 11°E) for the 2002-2012 period (morning coincidences). The solid/dashed lines represent mean/one-sigma standard deviation.

V.4 CO nadir column and N2O, HNO3, CH4 profile (WP 3.4)

The nadir CH4, HNO3, and N2O profile data selection, screening and processing – including averaging kernel smoothing – for both the Envisat MIPAS satellite instrument dataset and the NDACC ground-based FTIR reference data records have been extensively described in the Multi-TASTE Phase F midterm report (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, Apr 2014) and are therefore not repeated here.

An overview of the 15 NDACC FTIR stations providing CH4, HNO3, or N2O reference profile data is shown by their global distribution in Figure 5 and by listing in Table 21. This table moreover contains each station's operational time span and the number of spatial and temporal co-locations (hence comparisons) with Envisat MIPAS instrument observations that have been obtained within this period, both for the full mission validation of the MLP2PP 6.0 processor and the delta-validation between the IPF 5.05/5.06, ML2PP 6.0, and ML2PP 7.01. The full mission and delta-validation exercises are discussed jointly below, yet for each species separately.

Table 21: List of 15 NDACC FTIR stations (sorted north to south) providing CH4, HNO3, or N2O reference profile data. The number of co-locations (and hence comparisons) for the MIPAS retrieval datasets has been provided for each station and species, for the full mission ML2PP 6.0 processing ("F") and for the delta-validation exercise ("Δ"), respectively.

Station	Lat.	Lon.	Period	CH4 F	CH4 Δ	HNO3 F	HNO3 Δ	N2O F	N2O Δ
Eureka	79.99	-85.93	2007-2009	1452	44	2136	82	1385	32
Ny-Ålesund	78.93	11.93	2002-2012	244	127	223	89	/	/
Thule	76.53	-68.74	2002-2011	369	26	335	26	658	32
Kiruna	67.84	20.41	2002-2012	242	121	261	140	242	106
Harestua	60.20	10.80	2002-2012	181	130	/	/	/	/
Bremen	53.10	8.80	2004-2012	161	75	/	/	/	/
Zugspitze	47.42	10.98	2002-2012	783	384	/	/	1875	905
Jungfraujoch	46.55	7.98	2002-2012	299	157	265	140	233	108
Moshiri	44.40	142.30	2002-2007	/	/	39	0	/	/
Toronto	43.78	-79.47	2006-2011	213	13	92	1	459	21
Rikubetsu	43.50	143.80	2002-2009	/	/	32	0	/	/
Izaña	28.30	-16.50	2002-2012	486	258	191	115	243	136
Mauna Loa	19.53	-155.58	2003-2010	34	8	15	3	54	4
St Denis	-20.90	55.50	2004-2010	229	7	174	16	25	1
Wollongong	-34.41	150.88	2002-2008	560	15	/	/	589	13

V.4.1 MIPAS CH4 profile

MIPAS CH4 profile comparison plots that are vertically resolved between 12 and 30 km are collected in Annexe A1. Both the relative bias and spread are assessed in each station-specific graph, with a seasonal separation for the full ML2PP 6.0 validation (left-hand plots) and a processor version distinction for the delta-validation exercise (right-hand plots). The bias is thereby calculated as the median relative difference, while the spread equals the 68 % interpercentile of the set of relative difference profiles. Note that figures are also provided for the Harestua and Bremen stations; although they have been omitted from the following MIPAS CH4 profile validation analysis (see ground-based data quality assessment in Section IV.2.3).

Table 5 summarises the MIPAS CH4 quality assessment in terms of relative bias and spread within five latitude bands, first for the ML2PP 6.0 full mission data processing and then for the ML2PP 7.01 diagnostic dataset. Only significant biases, i.e. larger than the random error on the median, are thereby considered. Taking into account the (sometimes strong) reduction in the number of ground-

based co-locations (and hence comparisons) between these two validation exercises, the following observations can be made:

- Within the 2 to 5 % overall comparison spread, a globally consistent (ignoring the Antarctic) bias of a few percent negative to positive is detected for the ML2PP 6.0 full mission dataset.
- The ML2PP 7.01 systematic uncertainty falls within the same few percent negative to positive value range, but nonetheless seems to be slightly more positive (of the order of 1 to 3 %) than the (V5 and) V6 DDS relative bias. This increase is however covered by the version-consistent DDS random spread of 2 to 5 %.
- It should be remarked that in contrast with the previous the scarce MIPAS CH4 profile comparisons from the full resolution spectra at the beginning of the Envisat mission show large bias values of up to 30% for all processors (not shown here). The abundance of optimised resolution data usually hides this unsatisfactory performance in the overall comparison statistics.
- Both the ML2PP 6.0 full mission dataset and the three diagnostic datasets show a small vertical dependence of the bias, mostly by having a constant bias above 18 to 20 km, while going to lower (sometimes more negative) values below this altitude. Some stations however somewhat deviate from this tendency. In agreement with the second bullet, integrated subcolumn comparisons (not of primary focus here but presented at QWG meetings) moreover reveal that the ML2PP 7.01 bias slightly increases with respect to previous processings below about 60 km altitude, while slightly decreasing above 60 km.
- With the exception of some stations showing significant seasonal CH4 profile bias changes (e.g. Toronto), in general these changes tend to remain rather limited and no overall seasonal dependence can be detected. It is however observed – but not shown in the plots – that the small seasonal differences of the bias nevertheless seem to decrease with increasing product version.

Time series plots of relative systematic uncertainties (not included) show that no (significant) drifts can be observed.

V.4.2 MIPAS HNO3 profile

MIPAS HNO3 profile comparison plots that are vertically resolved between 12 and 30 km are collected in Annexe A2. Both the relative bias and spread are assessed in each station-specific graph, with a seasonal separation for the full ML2PP 6.0 validation (left-hand plots) and a processor version distinction for the delta-validation exercise (right-hand plots). The bias is thereby calculated as the median relative difference, while the spread equals the 68 % interpercentile of the set of relative difference profiles. After ground-based HNO3 data quality assessment (see Section IV.2.3), no stations have been omitted from the validation analysis.

Table 6 summarises the MIPAS HNO3 quality assessment in terms of relative bias and spread within five latitude bands, first for the ML2PP 6.0 full mission data processing and then for the ML2PP 7.01 diagnostic dataset. Only significant biases, i.e. larger than the random error on the median, are thereby considered. Taking into account the (sometimes strong) reduction in the number of ground-based co-locations (and hence comparisons) between these two validation exercises, the following observations can be made:

- MIPAS HNO3 profiles from the full ML2PP 6.0 dataset show large relative biases and spreads at all stations from the Arctic to the Tropics: Above 25 to 35 km, the bias is strongly positive (up to 25 %), decreasing below to a minimum of the order of -20 % at roughly 22 km altitude. From this minimum towards the ground, the bias again increases to the same high values as for the stratosphere. Although these bias statistics are usually significant, the corresponding spreads range between as low as 5 and as high as 40 %, typically decreasing towards

increasing height from the surface, yet reaching a minimum around the altitude of the most negative bias (22 km).

- Neglecting some smaller deviations, the systematic uncertainty of the ML2PP 7.01 DDS typically follows the same behaviour as the ML2PP 6.0 processor, but nonetheless seems to be slightly more positive (of the order of a few percent) than the (V5 and) V6 DDS relative bias. This increase is however largely covered by the version-consistent DDS random spread.
- The above results are solely valid for the optimised resolution dataset, as the full resolution part contains insufficient data for sound statistics.
- The MIPAS HNO3 profiles show significant seasonal bias dependences, with more negative bias values in local winter times around the bias minimum (around 22 km). Towards higher altitudes, seasonal differences typically become smaller, while they grow and often invert towards the ground. The latter also holds for the seasonal spreads. It is however observed – but not shown in the plots – that the seasonal differences of the bias and spread nevertheless seem to decrease with increasing product version.
- Time series plots of relative systematic uncertainties (not included here) show that no (significant) drifts can be observed for the 12 to 30 km profiles. Integrated subcolumn comparisons (not of primary focus here but presented at QWG meetings) however reveal a possible few percent negative drift above 30 km.

V.4.3 MIPAS N2O profile

MIPAS N2O profile comparison plots that are vertically resolved between 12 and 30 km are collected in Annexe A3. Both the relative bias and spread are assessed in each station-specific graph, with a seasonal separation for the full ML2PP 6.0 validation (left-hand plots) and a processor version distinction for the delta-validation exercise (right-hand plots). The bias is thereby calculated as the median relative difference, while the spread equals the 68 % interpercentile of the set of relative difference profiles. Note that figures are also provided for the Zugspitze and Toronto stations, although they have been omitted from the following MIPAS N2O profile validation analysis (see ground-based data quality assessment in Section IV.2.3).

Table 7 summarises the MIPAS N2O profile quality assessment in terms of relative bias and spread within five latitude bands, first for the ML2PP 6.0 full mission data processing and then for the ML2PP 7.01 diagnostic dataset. Only significant biases, i.e. larger than the random error on the median, are thereby considered. Taking into account the (sometimes strong) reduction in the number of ground-based co-locations (and hence comparisons) between these two validation exercises, the following observations can be made:

- Within the 2 to 10 % overall comparison spread, a globally consistent (ignoring the Antarctic) bias of a few percent negative to positive is detected for the ML2PP 6.0 full mission dataset.
- The ML2PP 7.01 systematic uncertainty falls within the same few percent negative to positive value range, but nonetheless seems to be slightly more positive (of the order of 1 to 3 %) than the (V5 and) V6 DDS relative bias. This increase is however covered by the version-consistent DDS random spread of 2 to 10 %.
- It should be remarked that in contrast with the previous the scarce MIPAS CH4 profile comparisons from the full resolution spectra at the beginning of the Envisat mission show large bias values of up to 50 % for all processors (not shown here). The abundance of optimised resolution data usually hides this unsatisfactory performance in the overall comparison statistics.
- Both the ML2PP 6.0 full mission dataset and the three diagnostic datasets show some vertical dependences of the bias (order of a few %, see first and second bullet), but no consistent view emerges on a zonal or global scale.

- Seasonal changes of the bias and spread in general remain rather limited and no overall seasonal dependence can be detected. It is however observed – but not shown in the plots – that the small seasonal differences of the bias nevertheless seem to decrease with increasing product version.
- Time series plots of relative systematic uncertainties (not included) show that no (significant) drifts can be observed.

V.4.4 SCIAMACHY nadir CO total column

FULL MISSION VALIDATION OF SGP 5.02

The CO nadir column data selection and screening for both the SCIAMACHY satellite instrument and the ground-based FTIR reference data have been extensively described in the Multi-TASTE Phase F midterm report (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, Apr 2014) and are therefore only briefly summarised here: Due to the unrealistically large variability of SCIAMACHY CO column data, users are advised to average SCIAMACHY data at least on monthly scales (whereby negative averages are omitted from further analysis). However, a prerequisite for the meaningful use of monthly means is that the data subsets (SCIAMACHY CO data SGP 5.02 and NDACC data) offer a similar sample of the atmosphere (comparison sampling errors will strongly add to the targeted discrepancies between measurements if this is not the case). The full mission validation analysis is therefore based on the relative difference of monthly means for satellite and ground-based measurement pairs that are both spatially (within 300 km) and temporally (3 hours) co-located. Note that this is not in agreement with the SCIAMACHY CO delta-validation that required additional subsetting (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, Apr 2014).

Table 22: Overview of NDACC FTIR stations (sorted north to south) providing CO total column data, divided into 5 latitude bands (alternating white-blue background). Subsequent columns provide the geolocation in latitude and longitude, the operational time span, and the number of monthly means. The four stations that are omitted from the comparative analysis by lack of statistics are marked in red.

Station	Lat.	Lon.	Period	# coll. (monthly means)
Eureka	79.99	-85.93	2006-2012	22
Ny-Ålesund	78.93	11.93	2002-2012	44
Thule	76.53	-68.74	2002-2011	24
Kiruna	67.84	20.41	2002-2012	50
Harestua	60.20	10.80	2002-2012	51
Bremen	53.10	8.80	2002-2012	59
Zugspitze	47.42	10.98	2002-2011	77
Jungfraujoch	46.55	7.98	2002-2012	77
Toronto	43.78	-79.47	2004	65
Mt. Barcroft	37.58	-118.24	2002	0
Kitt Peak	31.90	-111.60	2002-2005	0
Izaña	28.30	-16.50	2002-2012	60
Mauna Loa	19.53	-155.58	2003-2010	9
St. Denis	-20.90	55.50	2004-2011	9
Wollongong	-34.41	150.88	2002-2008	46
Lauder	-45.04	169.68	2002-2012	66
Arrival Heights	-77.83	166.67	2002-2012	36

When constructing monthly weighted averages of the co-located ground-based FTIR measurements and the SCIAMACHY CO SGP 5.02 dataset, the amount of monthly observations is strongly reduced. It has been decided to exclude stations with less than 20 overlapping monthly means from further

analysis, which is additionally motivated by the later need for yearly statistics (i.e. at least two monthly means per year; also see plots in Annexe A4). Out of the 17 FTIR stations that regularly provide the NDACC DHF with CO column data, as listed in Table 22, the 13 stations that are eventually used for the comparative analysis are indicated by blue dots in Figure 6. This distribution reveals that there is a relative over-concentration of measurements in the Arctic and Northern middle latitudes (30-60°N), which has to be taken into account when evaluating the comparison statistics.

The methodology for determining CO column comparison statistics on both monthly and yearly scales has also been outlined in the Multi-TASTE Phase F midterm report (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, Apr 2014). Yearly mean differences and spreads for SCIAMACHY SGP 5.02 are shown in Figure 37 in Annexe A4 for all 13 FTIR stations under consideration (see above). The comparison results in terms of relative bias and spread are summarised in Table 11 differentiated over the five predefined latitude bands. The major observations are the following:

- SCIAMACHY SGP 5.02 CO column data show a large amount of both positive and negative outliers, even on monthly scales. Negative monthly means are omitted from the comparative analysis.
- The yearly (and monthly) averaged SCIAMACHY CO columns are typically significantly positively biased, showing an apparent increase in median systematic uncertainty towards the equator, going from around 15% at the poles to roughly 55% in the tropics. This meridian dependence however is fully covered by the relative comparison spread, ranging approximately between 10 and 80% on a global scale.
- Due to the large variability of the SCIAMACHY CO column data on monthly scales no seasonal cycle in bias or spread can be observed.
- Neglecting some excessively high positive CO column biases in the limited full resolution dataset, the SGP 5.02 systematic uncertainty typically significantly increases from 2007 onwards, often exceeding the 100 % level. No physical or technical cause is currently known for this deviating conduct.
- Overall, the SCIAMACHY SGP CO product is inadequate in both precision and accuracy, especially for the full resolution measurement range and for post-2006 observations.

DELTA-VALIDATION BETWEEN SGP 5.02 AND SGP 6.00

The SCIAMACHY nadir CO total column delta-validation between SGP 5.02 and SGP 6.00 has been presented in Technical Note "Delta-validation of SCIAMACHY SGP upgrade from V5.02 to V6.00" (Hubert et al., TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Sep 2015). The results of this delta-validation exercise had been summarised in Table 4 (page 19, repeated as Table 23 below) and Table 1 (page 3) in the Technical Note. Initial and consolidated results were presented to and discussed within the QWG during meetings #3 and #4.

Table 23: Summary of the SCIAMACHY CO column delta-validation outcome divided into 5 latitude bands. Subsequent columns provide the number of stations, the number of monthly means from the diagnostic dataset, the DDS median bias shift from SGP 5.02 to SGP 6.00, and the related median comparison spread (which is similar for 5.02 and 6.00) in each band.

Latitude band	# stations	# monthly means	5.02 bias (%)	6.00 bias (%)	SGP 5.02/6.00 spread (%)
Arctic (60N-90N)	4	111	+2	-6	19
Mid-north (30N-60N)	4	176	+43	+36	24
Tropics (30N-30S)	1	42	+25	+55	20
Mid-south (30S-60S)	2	72	+45	+29	37
Antarctic (60S-90S)	1	12	+65	+2	35

V.5 Other validation-related activities

We wrote a Technical Note on how to convert vertical averaging kernels between different unit systems (Keppens et al., TN_BIRA-IASB_MIPAS-AKM-conversions_Keppens, October 2014). This is relevant when the satellite averaging kernels are expressed in different units than the vertical profiles by correlative instruments. The method was proposed to, discussed within and approved by the MIPAS QWG, but is generally applicable to other satellite instruments as well. This work was especially relevant for the validation of MIPAS O3 profiles (VMR on pressure levels) using O3 lidar measurements (number density on altitude levels).

V.6 KNMI activities and results (WP 3a)

This section describes the correlative analyses carried out by the KNMI, during the period of the Multi-TASTE Phase F CCN-2 contract (Dec 2014 – Dec 2015).

V.6.1 Validation approach

ENVISAT data have been obtained and using BEAT/CODA, relevant fields have been extracted from these datasets. The obtained information is then collocated with matching ground/sonde-based observations. Unless specified otherwise in the coming chapters, for temperature, we have used a maximum difference of 300 km and 5 hours and for ozone a maximum difference between observations of 800 km and 20 hours (5 hours above 50 km). Those criteria have been found to be a reasonable trade-off between having sound statistics and not having too different atmospheric conditions being sampled. In some occasions, we have applied a further sub-sampling using the difference in equivalent latitude to avoid sampling very distinct air masses.

As our main focus is on comparisons with lidar, we present our analyses in number density and on an altitude grid, which do not require any transformation of the lidar data, or use of external data. Please note that for MIPAS it is recommended to use a pressure grid.

In most analyses done in the framework of the VALID activities, convolution using the satellite instrument's averaging kernels was not done. In this report we do therefore not report any comparisons with convolved data, which allows for comparisons with analyses of previous versions.
The analyses themselves present (a selection of) the collocation data using a variety of statistics. Presented are means, standard deviations, medians, several percentiles of the differences (2.5, 16, 50, 84 and 97.5) and the spread (difference between the 84 and 16 percentiles). Those indicators are presented as a function of altitude, together with the number of collocations used to calculate those statistics at that altitude. Further details are presented in the text and figures.

V.6.2 GOMOS O3 profile

There has been no release of new/reprocessed data for GOMOS. Validation results of the current version (IPF 6.01) were summarised by KNMI on a poster presented at the ATMOS conference. We have also been involved in the review process of a paper on a comparison of GOME-2A ozone profiles with GOMOS, OSIRIS and MLS. For reference, we combined the GOMOS IPF 6.01 ozone validation results with those obtained by BIRA-IASB. These can be found in Table 2.

V.6.3 GOMOS high resolution T profile

We have done more comparisons of GOMOS temperature profiles with lidar and sonde observations. In general, agreement with the observational data is very good, perhaps with the exception of the upper-most altitudes where the influence of the tie-on (ECMWF a priori) for the GOMOS retrieval is large, and at the lowest altitudes, although the number of collocations is substantially reduced there.

We have looked at various characteristics of the GOMOS temperature profile to see how these affect the comparison results. In one study, special focus was made to the obliqueness in combination with the stellar magnitude following the classes and collocation criteria requested by Viktoria Sofieva.

Table 24. Number of collocations of GOMOS IPF 6.01 high resolution temperature profiles with sonde and lidar for different obliquity ranges and stellar magnitude classes. Collocation criteria used for both: 500 km and 4 hours.

		Obliquity range [°]			
		Vertical [0, 5]		Oblique [5, 90]	
	Collocated instrument	Sonde	Lidar	Sonde	Lidar
Stellar magnitude	Bright [-2, 1.5]	11	18	297	299
	Medium bright [1.5, 2.5]	4	15	360	407
	Dim [2.5, 5]	2	-	162	24

Table 24 above shows that the number of collocated observations under vertical conditions is very limited. For the few comparison cases that are found, we see that the altitude coverage tends to increase with stellar magnitude. For sonde observations, the category of collocations with "oblique, dim stars" is having a smaller altitude coverage (reaching less far downward) than collocations with (medium) bright stars for oblique angles. For the collocations with lidar this limited altitude coverage is also seen.

In Figure 24 we show a comparison for the collocated sonde data introduced above, splitting the results by stellar temperature. The data have been interpolated onto a 200 m grid. We can see that most collocations are with hot (>7000 K) stars and that the altitude coverage is slightly larger for this category. For the comparison with cold (<7000 K) stars, there seems to be a trend of the bias increasing slightly with altitude. This is less visible in the hot stars category results as at some altitudes the bias is stable or reverses sign. The bottom (lowest altitudes) of the difference profiles shows the largest deviations, but the number of collocations also decreases drastically there. Overall, agreement with the sonde data is mostly within 1 K.

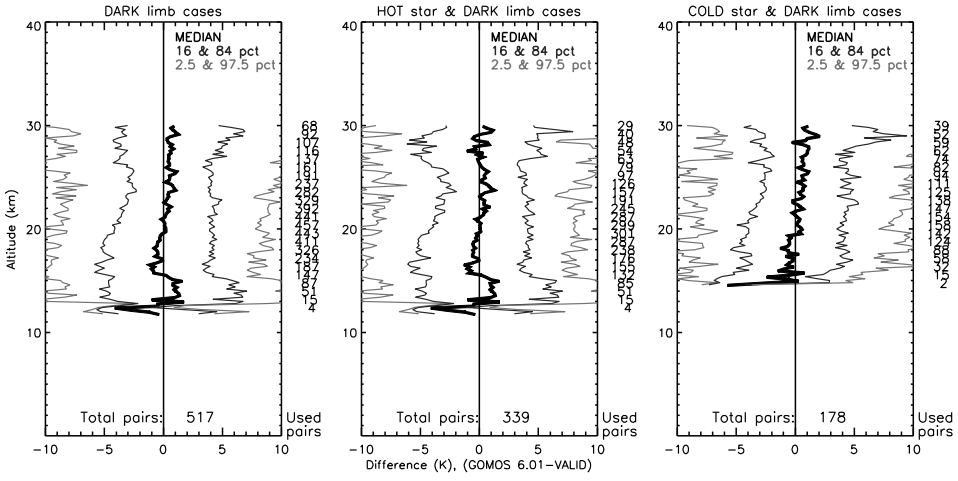


Figure 24: Comparison of GOMOS IPF 6.01 HRTP with temperature profiles from sonde collocated at 400 km and within 5 hours. Left: all cases together, middle: hot stars (stellar T > 7000 K), right: cold stars (stellar T < 7000 K). Shown for each subplot are absolute differences in K with respect to the validation instrument as a function of altitude, for the following percentiles: 2.5, 16, 50 (=median), 84 and 97.5%. Along the right axes the number of collocations at the corresponding altitude is written, and the total number of collocations is listed at the bottom of each subplot.

The same distinction between hot and cold stars, but now in comparison to lidar profiles, is shown in Figure 25. Here we see that the trend with altitude for cold stars does not hold until 30 km as for the sonde analysis. For hot stars, GOMOS temperatures tend to be a little on the warm side below 33 km. Deviations from the lidar temperatures are mostly within 2 K and usually better. In both cases we see some very cold outlier temperatures for GOMOS. It would be useful to add an extensive trend analysis, but this is complicated given the relatively low number of collocations. Table 1 summarises the bias and spread of the comparison between GOMOS IPF 6.01 and lidar, for the three main latitude zones.

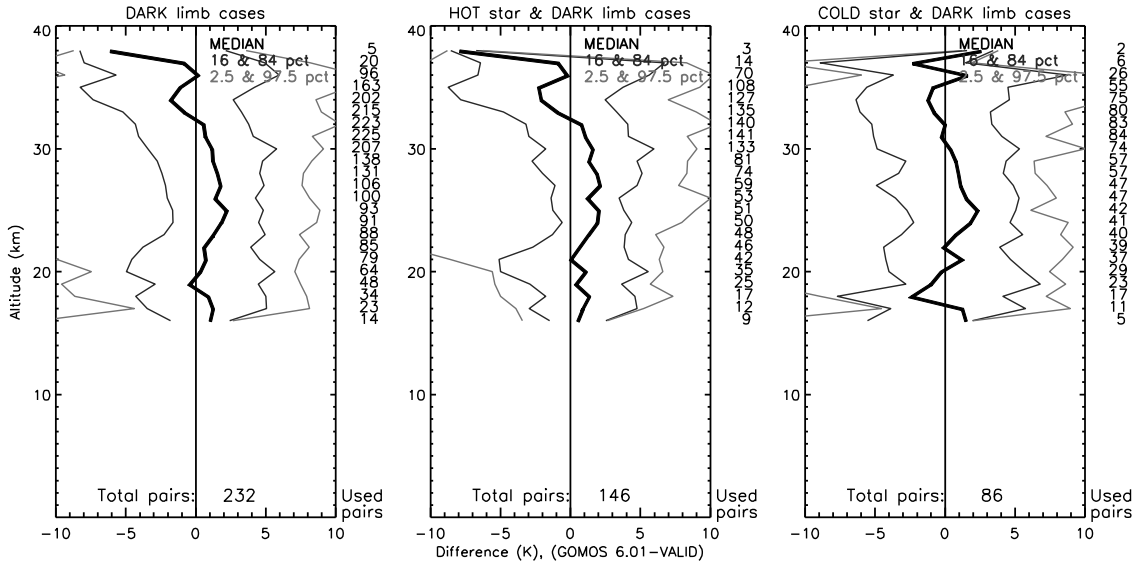


Figure 25: Comparison of GOMOS v6.01 HRTP with temperature profiles from lidar collocated at 300 km and within 5 hours. Left: all cases together, middle: hot stars (stellar T > 7000 K), right: cold stars (stellar T < 7000 K). Shown for each subplot are absolute differences in K with respect to the validation instrument as a function of altitude, for the following percentiles: 2.5, 16, 50 (=median), 84 and 97.5%. Along the right axes the number of collocations at the corresponding altitude is written, and the total number of collocations is listed at the bottom of each subplot.

V.6.4 MIPAS O3 profile

We have carried out various comparisons for MIPAS ozone and temperature data ML2PP versions 7.01 and 7.03. Here we will present results for version 7.03 compared to sonde, lidar and microwave radiometer. Please note that the vertical grid recommended by the MIPAS QWG is the pressure axis, but that we use the lidar's native altitude axis. Also, the lidar data used in the comparisons presented here are not convolved with MIPAS' averaging kernels.

For sonde, the number of collocations with night-time observations is higher than with day-time observations. Little difference can be seen in the median bias for night-time and day-time, but the mean differences are clearly distinct, with a higher mean bias for day-time observations between 21 and 29 km. Figure 26 shows the comparisons with sonde grouped by the three main latitude regions. Agreement is better for higher altitudes and closer to the poles. MIPAS shows strongly overestimated ozone number densities at the lowest altitudes. In the tropics, there is also a peak in the difference profile maximising around 15 km which is seen at multiple sites which is possibly due to (sub-visual) cirrus.

Using the same microwave radiometer site as for the analysis in Figure 32 (Lauder), we have compared its ozone profiles with that of MIPAS version 7.03 in Figure 27. A distinction is made between day-time (orange) and night-time observations, and a further split into full resolution (middle panel) and optimized resolution (right panel) is presented alongside. Oscillations are seen in all percentiles and are of a larger amplitude for day-time conditions. The daytime observations also show a reduced agreement with the microwave radiometer above 60 km in comparison to the night-time observations. Thirdly, the full resolution period is overall more positively biased between 40 and 60 km.

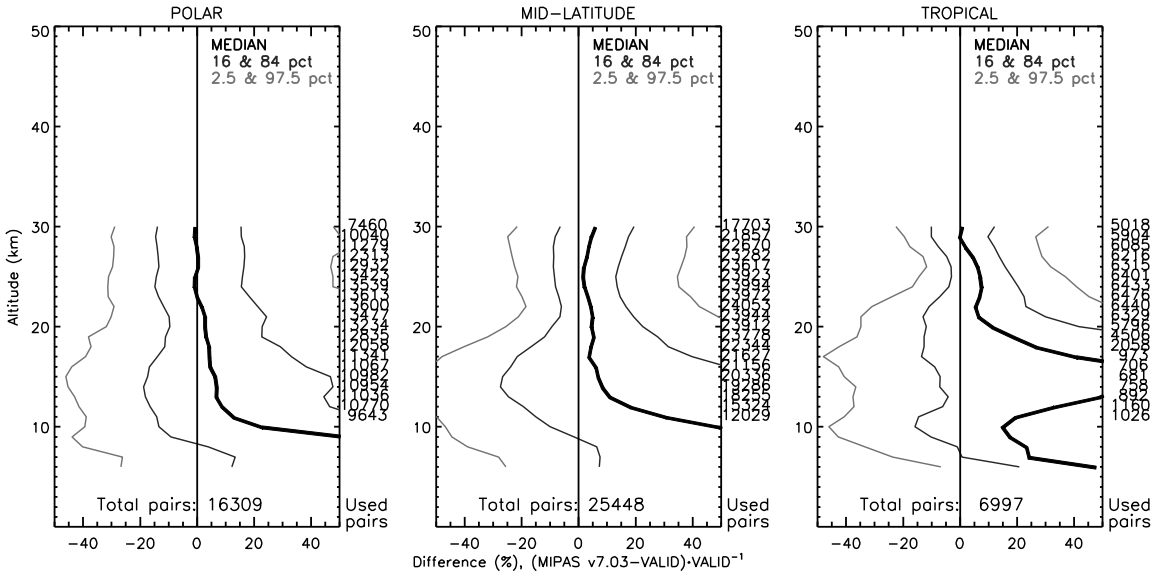


Figure 26: Comparison of MIPAS ML2PP 7.03 ozone profiles with sonde observations for the three main latitude regions.

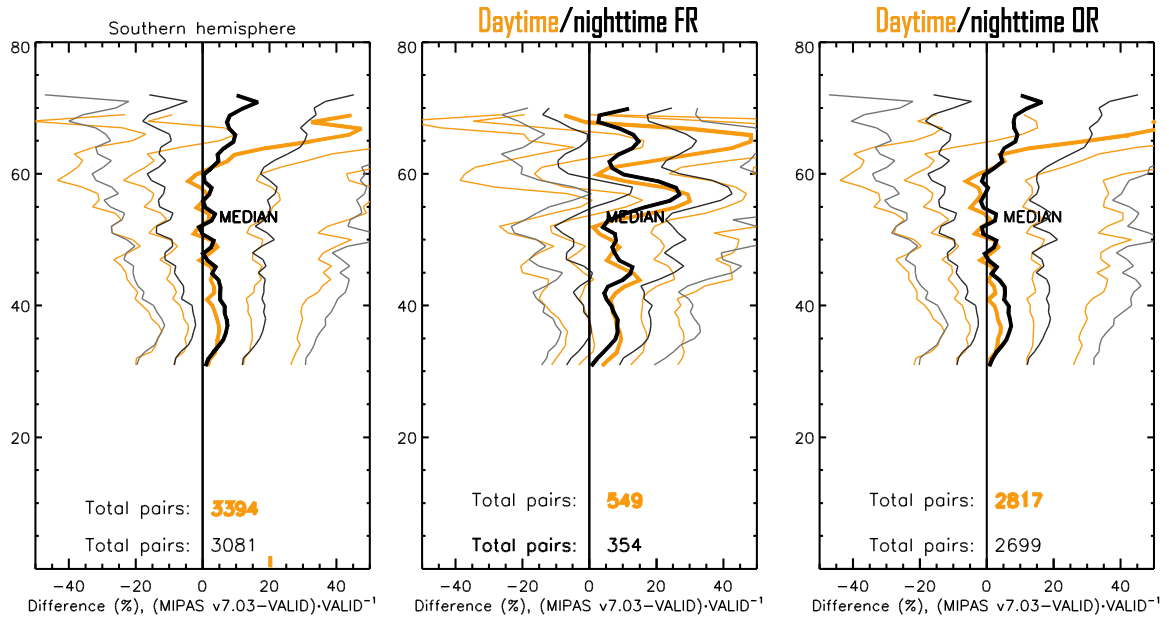


Figure 27: Comparison of MIPAS ML2PP 7.03 ozone profiles with microwave radiometer observations at Lauder (New Zealand), splitting the MIPAS full resolution (FR) and optimised resolution (OR) periods and day-time/night-time (orange/black) observations.

A difference between the two periods and between collocations with night-time/day-time MIPAS observations can also sometimes be seen in comparison to lidar. Figure 28 shows the four combinations of these two characteristics for the three main latitude regions. As the smallest difference in time from the lidar observations is with the night-time observations, these are used for the summary in Table 4. The agreement between the four possible combinations is very good in the mid-latitudes between 25 and 38 km. Outside this range and these latitudes there are differences between the combinations, but the best agreeing combination differs per latitude region and altitude. The 2.5 percentile difference lines are usually quite close to each other. In the mid-latitudes, there is a clear increase of high outliers for the optimised resolution period (blue and green 97.5 percentile lines).

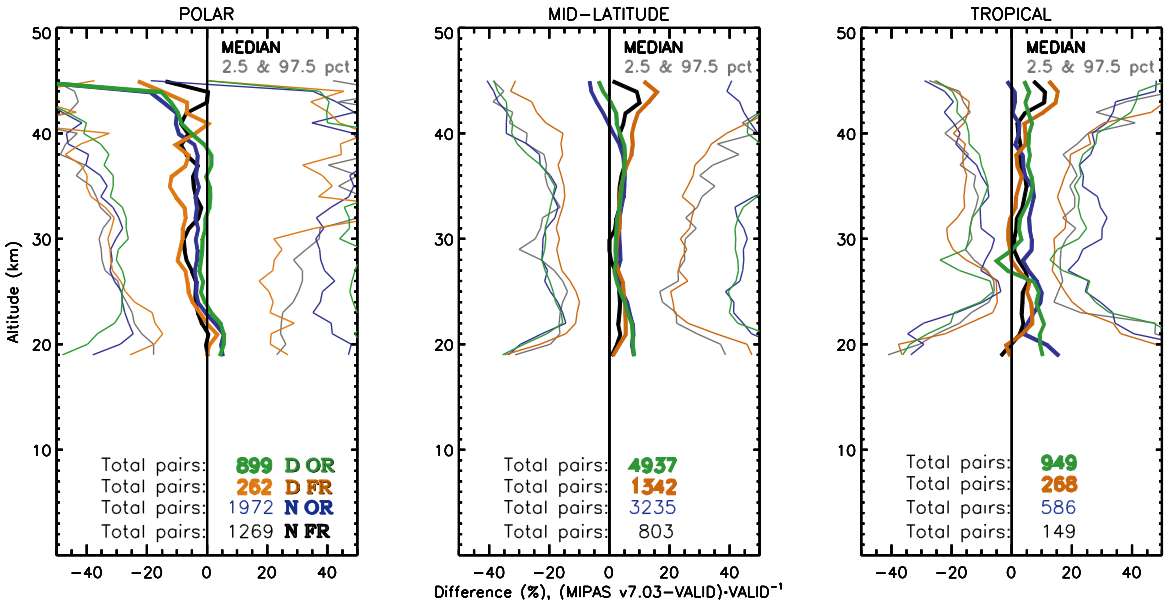


Figure 28: Comparison of MIPAS v7.03 ozone profiles with lidar observations for the three main latitude regions further distinguishing between daytime/night-time and full resolution or optimised resolution.

V.6.5 MIPAS T profile

We have compared MIPAS ML2PP 7.03 temperature profiles with lidar observations. As the criterion for maximum time difference is 5 hours, nearly all of the collocations are with MIPAS night-time observations. Figure 29 below shows the comparison results for the three main latitude regions, with the full resolution (FR) period shown by black lines and the optimised resolution (OR) period shown by blue lines. The two periods only show a general agreement in the tropics. For the OR period, there is always a positive bias at the bottom of the profile and a negative bias at the top of the profile. This is not the case for the comparison at the polar regions during the FR period where the trend with altitude is of a reversed sign. In general, the agreement with lidar is not very good. Table 3 numerically summarises the results obtained by KNMI and BIRA-IASB.

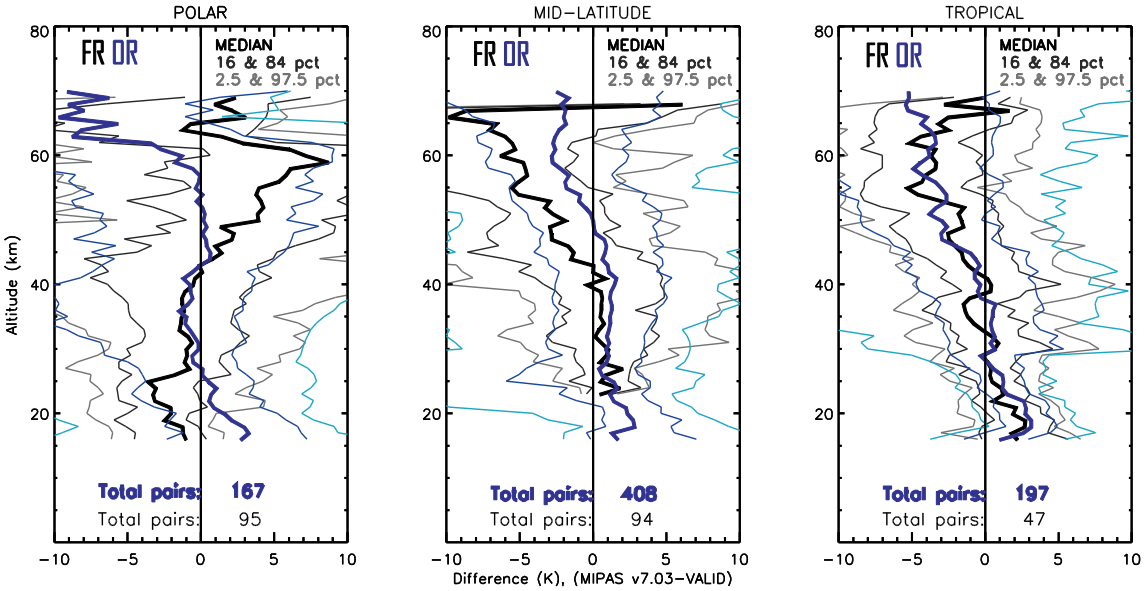


Figure 29: Comparison of MIPAS ML2PP 7.03 temperature profiles with lidar for the three main latitude regions, distinguishing between full resolution and optimised resolution time periods.

V.6.6 SCIAMACHY limb O3 profile

Comparisons of SCIAMACHY SGP 6.00 DDS limb ozone profiles have been done with sonde, lidar and microwave radiometer. Collocations were sought within 20 hours and 800 km. Figure 30 shows the comparisons with sonde, splitting the results to the three main latitude regions. Agreement with the sonde is within about 5% between 20 and 30 km in the polar regions and in the mid-latitudes, but in the tropics the bias increases from 0 around 20 km to more than 10% around 30 km. For the upper half of the altitudes compared here, the spread is largest in the polar regions. For the lower half, we see that the spread increases rapidly when descending downward. Overall, it is more frequent for SCIAMACHY outliers to be too high.

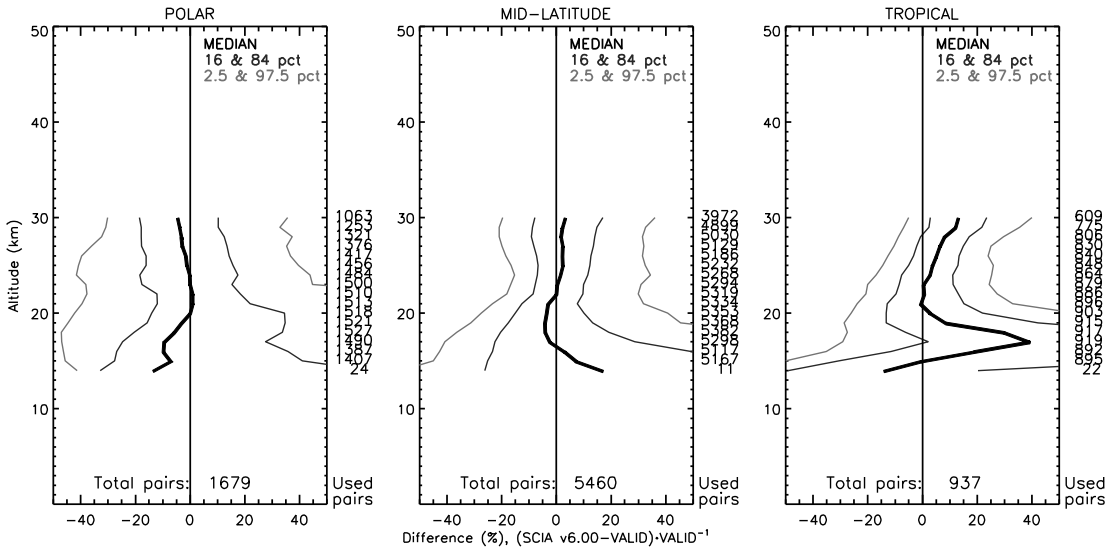


Figure 30: Relative differences between ozone profiles from SCIAMACHY SGP 6.00 DDS and sonde observations for the three main latitude zones) Shown are the 2.5, 16, 50 (=median), 84 and 97.5 percentiles of the differences relative to the sonde. The three panels show from left to right: polar regions, mid-latitudes and the tropics.

Figure 31 shows the comparisons with lidar for the mid-latitudes, adding also the distinction of on which hemisphere the observations were made. The largest part of the comparisons is located on the northern hemisphere, where most lidar sites are. We can also see that there are clear differences in the bias as well as in the 97.5 percentiles. For the northern mid-latitudes, the bias moves a few percent around 0 up to 30 km and then steadily increases to about 8% at the top. The southern mid-latitudes exhibit a positive bias of around 10% between 25 and 40 km with a change towards a more than 10% negative bias above 40 km and below 25 km. The high outliers are more extreme than seen on the northern hemisphere.

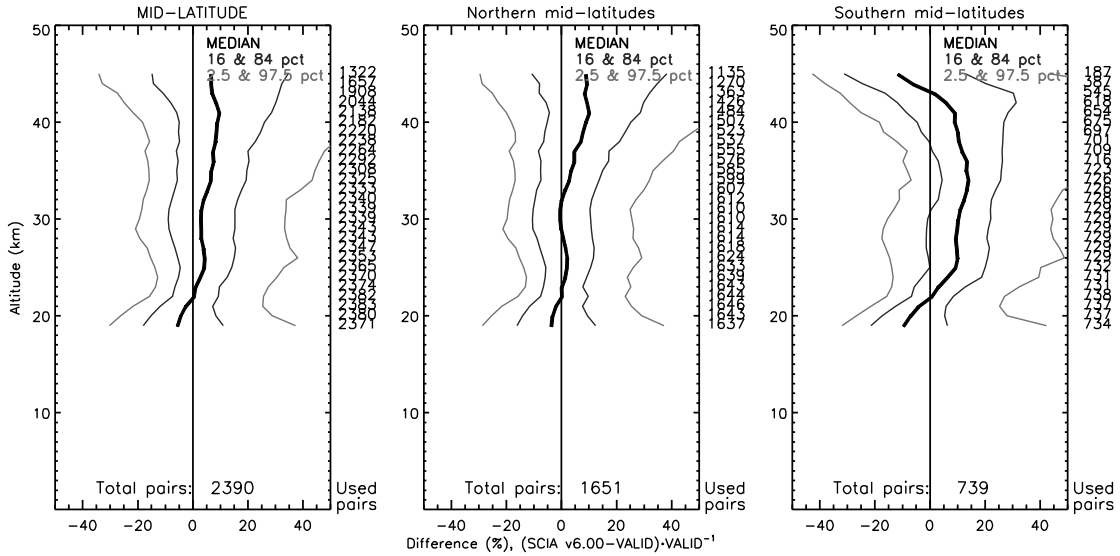


Figure 31: Relative differences in ozone between SCIAMACHY SGP 6.00 DDS and lidar in the mid-latitudes. The left panel groups both hemispheres together, the middle panel only shows the northern mid-latitudes whereas the right panel only shows the southern mid-latitudes.

In version 6.00, the ozone profile retrievals has extended upwards (before most of the data reported above 40 km was coming from the a-priori). We have carried out an analysis using microwave radiometer data to see how the mesospheric ozone profile from SCIAMACHY compares to it. Figure 32 shows results for the site of Lauder, where we have double-collocated the SCIAMACHY observation with a lidar and with a microwave radiometer observation. It can be seen that over the part where we have both lidar and microwave radiometer, the agreement on the differences with SCIAMACHY is very good, up to above ~42 km, where the lidar data become less reliable. As already seen in Figure 31 (Lauder being one of the two sites in the southern mid-latitudes), SCIAMACHY is positively biased. Between 45 km and 60 km, the bias varies between +5% and +15% and rapidly increases above 60 km. A similar large increase in the bias above 60 km is seen for other sites. Table 13 summarises the bias and spread of the comparison between SCIAMACHY SGP 6.00 DDS limb ozone profiles and lidar, for the main latitude zones.

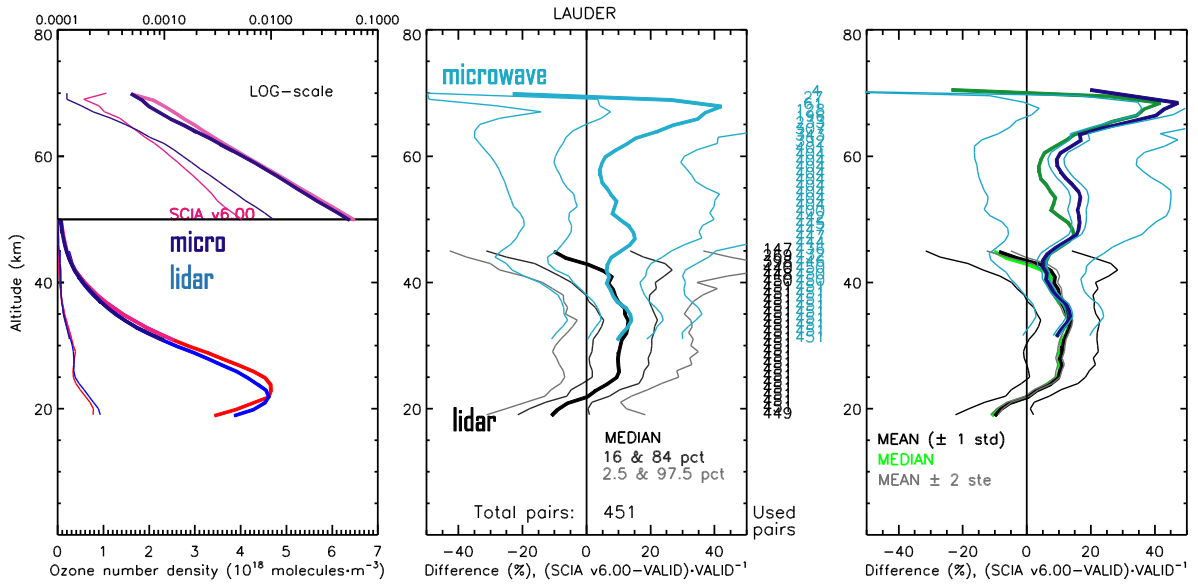


Figure 32: Comparison results for SCIAMACHY v6.00 ozone profiles at Lauder, New Zealand, with respect to the lidar (up to 45 km; lower part in the panels, black lines in middle and right panels) and the microwave radiometer (from 30 km upward; upper part in the panels, blue lines in the middle and right panel with the median in the right panel in dark green).

V.6.7 Identifying possible relations between differences seen in validation with observational variables

During the project, we also published our research work that looks at differences between two datasets with self-organising maps (SOMs) and looks for explanatory links without making a priori assumptions on groups/classes. The technique is illustrated with SCIAMACHY SGP 5.02 limb ozone profiles compared to lidar observations (van Gijsel et al., 2015).

We can see in Figure 33 that obviously some variables are closely related, but that there is an altitude dependence for which variables have the largest influence on the observed differences between SCIAMACHY and the lidar ozone profiles, and that the relations observed can be locally different (in terms of SOM-space). Also after accounting for the correlation with latitude and longitude, we see remaining influences of variables related to those (e.g. solar zenith/azimuth angles). This type of analysis does not always directly pinpoint where the deviations come from, but can offer the algorithm developers new insights and look for possible interactions/issues in the highlighted areas.

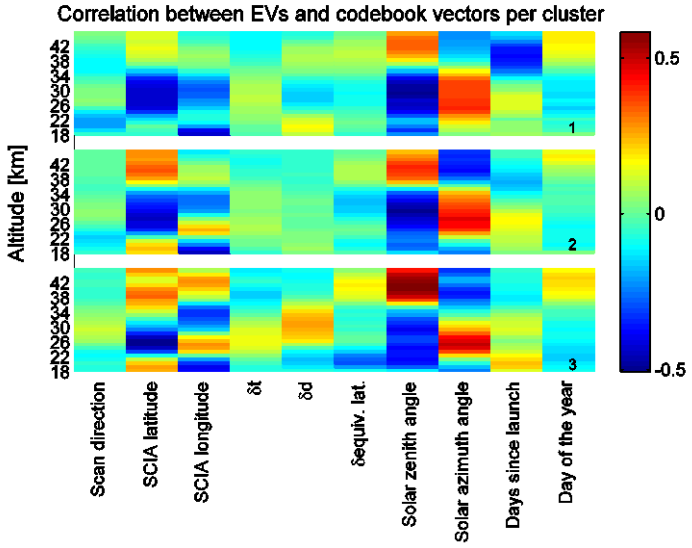


Figure 33: Example from publication: the correlation between explanatory variables (EVs) on the x-axis versus codebook vectors (altitude y-axis) for data grouped into three clusters (sorted top-down)).

VI Reporting and valorisation (WP 4)

Preliminary and consolidated validation results were presented at numerous Envisat QWG meetings. The methodology and final conclusions of our analyses were also reported in written validation reports and in a progress report. Furthermore, we contributed to the discussions and content of the Product Readme files which accompany the publicly released data sets. Multi-TASTE Phase F results were also communicated to the research community at numerous international conferences, workshops and symposia (ACVE, ATMOS, Living Planet, …). In addition, we have provided support to users of Envisat operational Level-2 data products, which led to co-authorship on several peer-reviewed publications.

A complete list of oral and poster presentations, written reports, conference proceedings and peer-reviewed publications can be found in Annexe B. The list of meetings and conferences we attended is given in Annexe C.

VII Bibliography

Balis, D., et al.: *Ten years of GOME/ERS2 total ozone data—The new GOME data processor (GDP) version 4: 2. Ground-based validation and comparisons with TOMS V7/V8*, J. Geophys. Res., 112, D07307, doi:10.1029/2005JD006376, 2007.

Barret, B., et al.: *Ground-based FTIR measurements of CO from the Jungfraujoch: characterisation and comparison with in situ surface and MOPITT data*, Atmos. Chem. Phys., 3, 2217-2223, doi:10.5194/acp-3-2217-2003, 2003.

Batchelor, R. L., et al.*: A New Bruker IFS 125HR FTIR Spectrometer for the Polar Environment Atmospheric Research Laboratory at Eureka, Nunavut, Canada: Measurements and Comparison with the Existing Bomem DA8 Spectrometer*, J. Atmos. Oceanic Technol., 26, 1328–1340, doi:10.1175/2009JTECHA1215.1, 2009.

Celarier, E. A., et al.: *Validation of Ozone Monitoring Instrument nitrogen dioxide columns*, J. Geophys. Res., 113, D15S15, doi:10.1029/2007JD008908, 2008.

de Laat, A. T. J., et al.: *Validation of five years (2003–2007) of SCIAMACHY CO total column measurements using ground-based spectrometer observations*, Atmos. Meas. Tech., 3, 1457-1471, doi:10.5194/amt-3-1457-2010, 2010.

Dils, B., et al.: *Comparisons between SCIAMACHY and ground-based FTIR data for total columns of CO, CH4, CO2 and N2O*, Atmos. Chem. Phys., 6, 1953-1976, doi:10.5194/acp-6-1953-2006, 2006.

du Piesanie, A., et al.: *Validation of two independent retrievals of SCIAMACHY water vapour columns using radiosonde data*, Atmos. Meas. Tech., 6, 2925-2940, doi:10.5194/amt-6-2925-2013, 2013.

Greenblatt, G. D., et al.: *Absorption measurements of oxygen between 330 and 1140 nm*, J. Geophys. Res., 95, 18557-18582, doi:10.1029/JD095iD11p18577, 1990.

Hendrick, F., et al.: *Retrieval of stratospheric and tropospheric BrO profiles and columns using ground-based zenith-sky DOAS observations at Harestua, 60°N*, Atmos. Chem. Phys., 7, 4869-4885, doi:10.5194/acp-7-4869-2007, 2007.

Hendrick, F. et al.: *Multi-year comparison of stratospheric BrO vertical profiles retrieved from SCIAMACHY limb and ground-based UV-visible measurements*, Atmos. Meas. Tech., 1, 273-285, doi:10.5194/amt-2-273-2009, 2009.

Hendrick, F., et al.: *NDACC/SAOZ UV-visible total ozone measurements: improved retrieval and comparison with correlative ground-based and satellite observations*, Atmos. Chem. Phys., 11, 5975-5995, doi:10.5194/acp-11-5975-2011, 2011.

Hubert, D., N. Kalb, J.-C. Lambert et al.: *Multi-TASTE Final Report (October 2008 – October 2011)*, TN-BIRA-IASB-MultiTASTE-FR, http://earth.eo.esa.int/pcs/envisat/calval_res/2012/TN-BIRA-IASB-MultiTASTE-FR-iss1revC-Oct2012.pdf, 2012.

Hubert, D., A. Keppens, J. Granville, F. Hendrick and J.-C. Lambert: *Multi-TASTE Phase F Progress Report #1 / October 2013 – March 2014*, Project Progress Report, TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, 37pp., 17 Apr 2014.

Hubert, D. et al.: *Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records*, Atmos. Meas. Tech. Discuss., 8, 6661-6757, doi:10.5194/amtd-8-6661-2015, 2015.

Keckhut, P., et al. : *Review of ozone and temperature lidar validations performed within the framework of the Network for the Detection of Stratospheric Change*, J. Environ. Monit., 6, 721–733, doi:10.1039/b404256e, 2004.

Lambert, J.-C., et al.: *Investigation of pole-to-pole performances of spaceborne atmospheric chemistry sensors with the NDSC*, Journal of the Atmospheric Sciences, 56, 176-193, doi:10.1175/1520-0469(1999)056<0176:IOPTPP>2.0.CO;2, 1999.

Lambert, J.-C., et al.: *Combined characterisation of GOME and TOMS total ozone measurements from space using ground-based observations from the NDSC*, Advances in Space Research, 26, 1931-1940, 2000.

Lambert, J.-C., et al.: *Initial Validation of GOME-2 Nitrogen Dioxide Columns (GDP 4.2 OTO/NO2 and NTO/NO2): March-June 2007*, IASB/EUMETSAT Technical Note TN-IASB-GOME2-O3MSAF-NO2-01, Issue 1, Revision B, 40 pp., 22 October 2007.

Miloshevich, L. M., et al.: *Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements*, J. Geophys. Res., 114, D11305, doi:10.1029/2008JD011565, 2009.

Noël, S., et al.: *SCIAMACHY Quality Working Group (SQWG-3), Minutes of Status Telecon 20 Nov 2014*, Dec 2014.

Palm, M., et al.: *The ground-based MW radiometer OZORAM on Spitsbergen – description and status of stratospheric and mesospheric O3-measurements*, Atmos. Meas. Tech., 3, 1533-1545, doi:10.5194/amt-3-1533-2010, 2010.

Payan, S., et al.: *Validation of version-4.61 methane and nitrous oxide observed by MIPAS*, Atmos. Chem. Phys., 9, 413-442, doi:10.5194/acp-9-413-2009, 2009.

Pfeilsticker, K. et al.: *Intercomparison of the influence of tropospheric clouds on UV-visible absorptions Detected during the NDSC Intercomparison Campaign at OHP in June 1996*, Geophys. Res. Lett., 26, 1169-1172, doi:10.1029/1999GL900198, 1999.

Pukite, J., et al.: *Extending differential optical absorption spectroscopy for limb measurements in the UV*, Atmos. Meas. Tech., 3, 631-653, doi:10.5194/amt-3-631-2010, 2010.

Roscoe, H. K., et al.: *Intercomparison of slant column measurements of NO2 and O4 by MAX-DOAS and zenith-sky UV and visible spectrometers*, Atmos. Meas. Tech., 3, 1629-1646, doi:10.5194/amt-3-1629-2010, 2010.

SCIAMACHY Quality Working Group: *Readme file for SCIAMACHY Level 2 version 5.02 products*, ENVI-GSOP-EOGD-QD-13-0118, issue 1.2, http://envisat.esa.int/handbooks/availability/disclaimers/SCI_OL__2P_README.pdf, 2013.

Schneider, M., et al.: *Subtropical trace gas profiles determined by ground-based FTIR spectroscopy at Izaña (28° N, 16° W): Five-year record, error analysis, and comparison with 3-D CTMs*, Atmos. Chem. Phys., 5, 153-167, doi:10.5194/acp-5-153-2005, 2005.

Schneider, M., et al.: *Quality assessment of O3 profiles measured by a state-of-the-art ground-based FTIR observing system*, Atmos. Chem. Phys., 8, 5579-5588, doi:10.5194/acp-8-5579-2008, 2008.

Senten, C., et al.: *Technical Note: New ground-based FTIR measurements at Ile de La Réunion: observations, error analysis, and comparisons with independent data*, Atmos. Chem. Phys., 8, 3483-3508, doi:10.5194/acp-8-3483-2008, 2008.

Sepúlveda, E., et al.: *Tropospheric CH4 signals as observed by NDACC FTIR at globally distributed sites and comparison to GAW surface in situ measurements*, Atmos. Meas. Tech., 7, 2337-2360, doi:10.5194/amt-7-2337-2014, 2014.

Smit, H. G. J. and ASOPOS-panel: *Quality Assurance and Quality Control for Ozonesonde Measurements in GAW*, WMO Global Atmosphere Watch report series 201, World Meteorological Organization, http://www.wmo.int/pages/prog/arep/gaw/documents/FINAL_GAW_201_Oct_2014.pdf, 2014.

Sonntag, D.: *Advancements in the field of hygrometry*, Meteorol. Z., N. F., 3, 51-66, 1994.

Sun, B., et al.: *Comparing radiosonde and COSMIC atmospheric profile data to quantify differences among radiosonde types and the effects of imperfect collocation on comparison statistics*, J. Geophys. Res., 115, D23104, doi:10.1029/2010JD014457, 2010.

Sun, B., et al.: *Toward improved corrections for radiation-induced biases in radiosonde temperature observations*, J. Geophys. Res. Atmos., 118, 4231–4243, doi:10.1002/jgrd.50369, 2013.

Sussmann, R., et al.: *Strategy for high-accuracy-and-precision retrieval of atmospheric methane from the mid-infrared FTIR network*, Atmos. Meas. Tech., 4, 1943-1964, doi:10.5194/amt-4-1943-2011, 2011.

Theys, N., et al.: *Retrieval of stratospheric and tropospheric BrO columns from multi-axis DOAS measurements at Reunion Island (21° S, 56° E)*, Atmos. Chem. Phys., 7, 4733-4749, doi:10.5194/acp-7-4733-2007, 2007.

Tukiainen, S., et al.: *GOMOS bright limb ozone data set*, Atmos. Meas. Tech., 8, 3107-3115, doi:10.5194/amt-8-3107-2015, 2015.

Vandaele, A. C., et al.: *An intercomparison campaign of ground-based UV-visible measurements of NO2, BrO, and OClO slant columns: Methods of analysis and results for NO2*, J. Geophys. Res., 110, D08305, doi:10.1029/2004JD005423, 2005.

van Gijsel, J. A. E. et al.: *GOMOS ozone profile validation using ground-based and balloon sonde measurements*, Atmos. Chem. Phys., 10, 10473-10488, doi:10.5194/acp-10-10473-2010, 2010.

van Gijsel, J. A. E. et al.: *Using self-organising maps to explore ozone profile validation results – SCIAMACHY limb compared to ground-based lidar observations*, Atmos. Meas. Tech., 8, 1951-1963, doi:10.5194/amt-8-1951-2015, 2015.

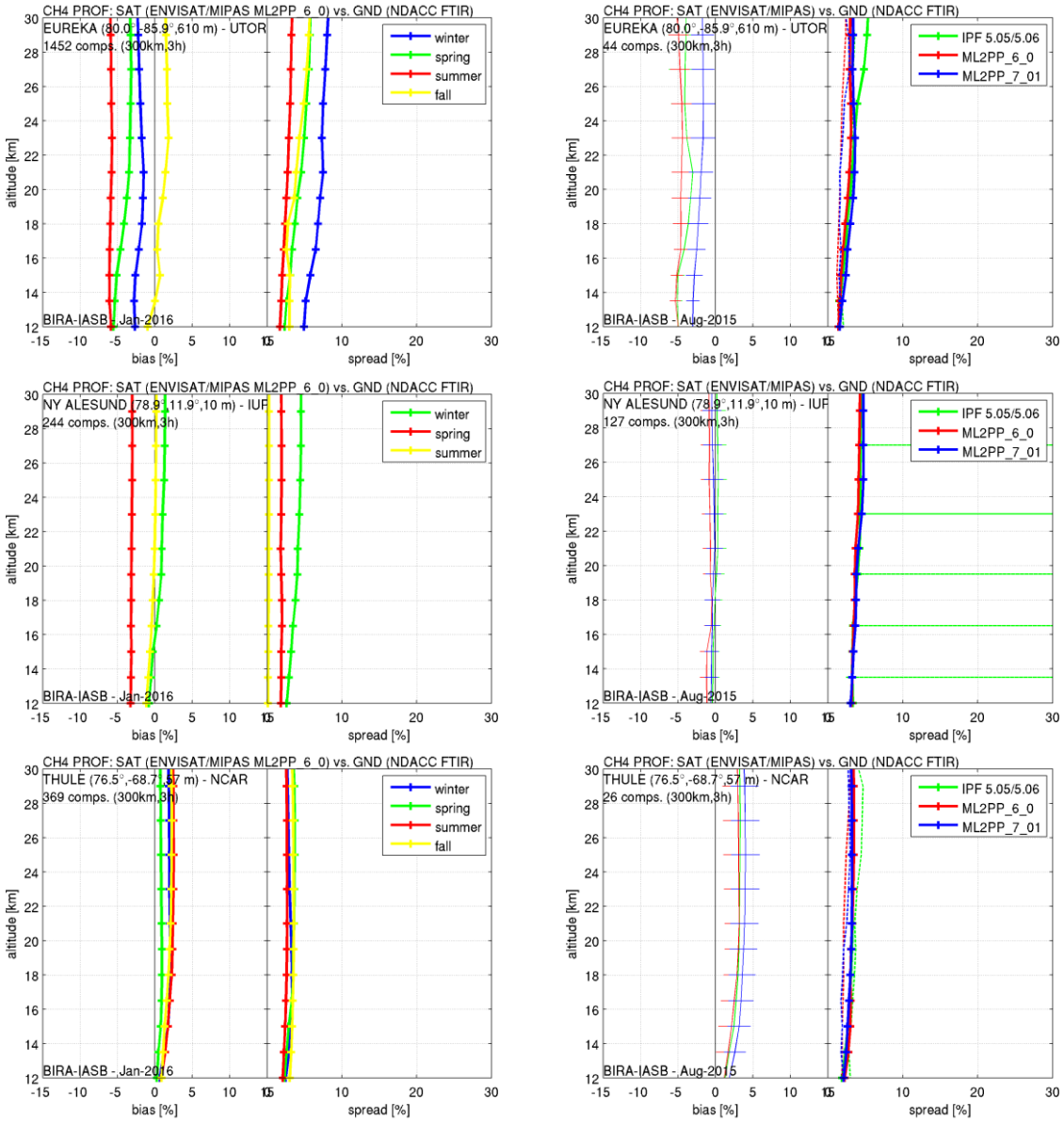
Vigouroux, C., et al.: *Comparisons between ground-based FTIR and MIPAS N2O and HNO3 profiles before and after assimilation in BASCOE*, Atmos. Chem. Phys., 7, 377-396, doi:10.5194/acp-7-377-2007, 2007.

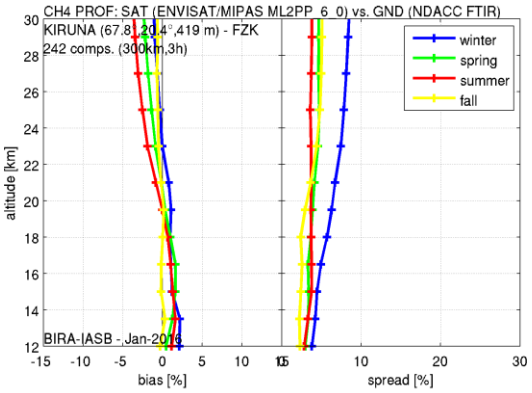
Vigouroux, C., et al.: *Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data*, Atmos. Chem. Phys., 9, 9523-9544, doi:10.5194/acp-9-9523-2009, 2009.

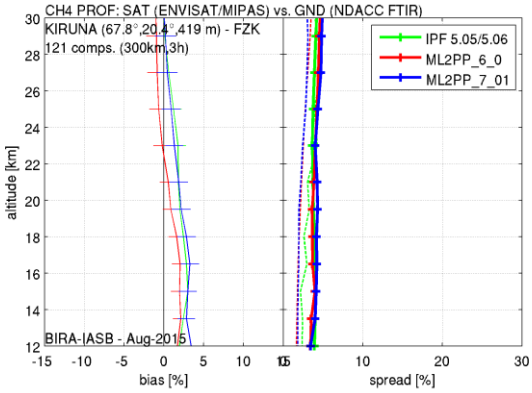
Annexe A : Correlative analyses

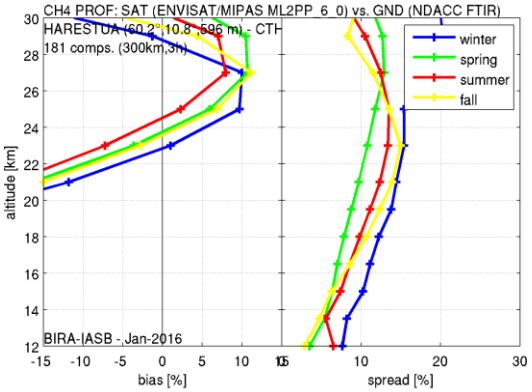
A1 MIPAS CH4 vertical profile

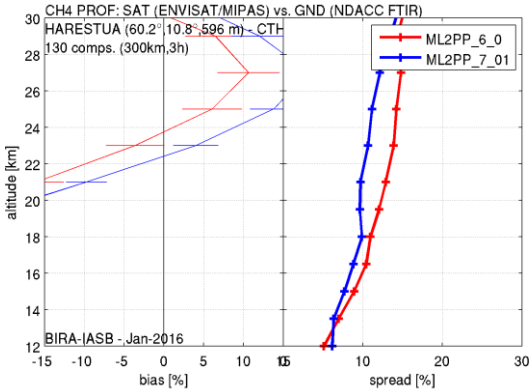
Figure 34: MIPAS CH4 profile comparison plots between 12 and 30 km. Relative bias and spread are assessed in each station-specific graph (rows, sorted north to south), with a seasonal separation for the full ML2PP 6.0 validation (left column) and a processor version distinction for the delta-validation exercise (right column). For the latter also the random error on the mean (as horizontal lines at each altitude) and the combined satellite-ground error (as dashed lines in the spread plots) are provided. The station location and number of co-locations are shown in the upper-left corner of each box.

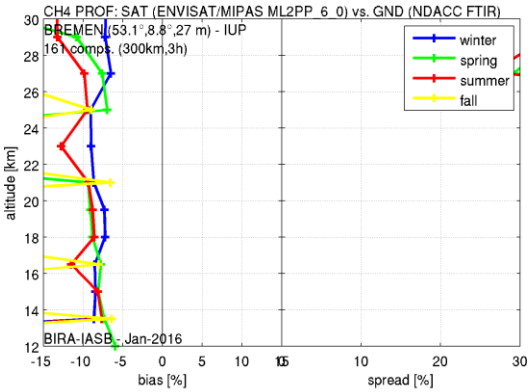


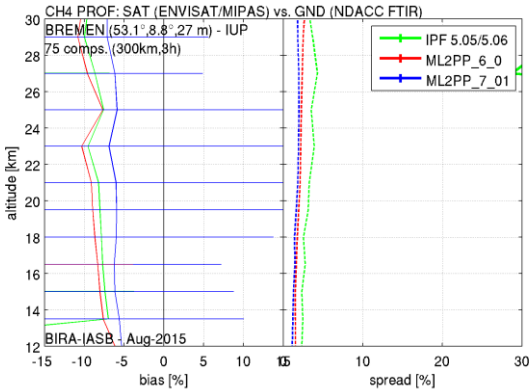

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
KIRUNA (67.8°,20.4°,419 m) - FZK
242 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

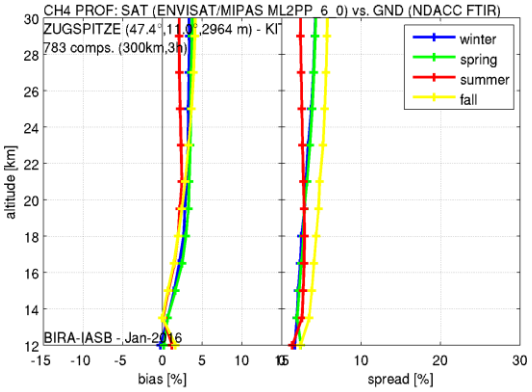

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
KIRUNA (67.8°,20.4°,419 m) - FZK
121 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

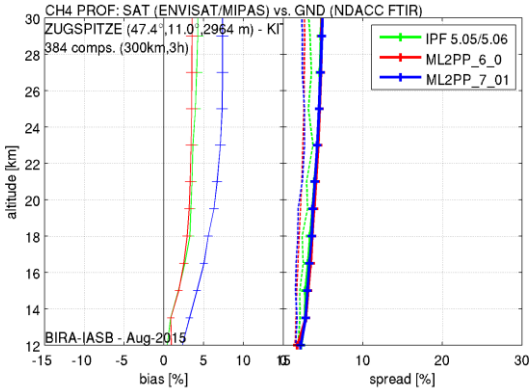

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
HARESTUA (60.2°,10.8°,596 m) - CTH
181 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

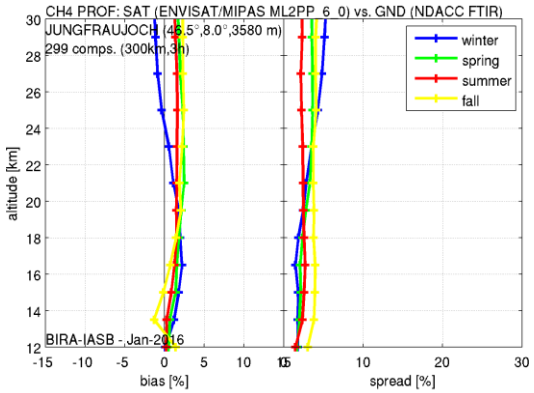

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
HARESTUA (60.2°,10.8°,596 m) - CTH
130 comps. (300km,3h)
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

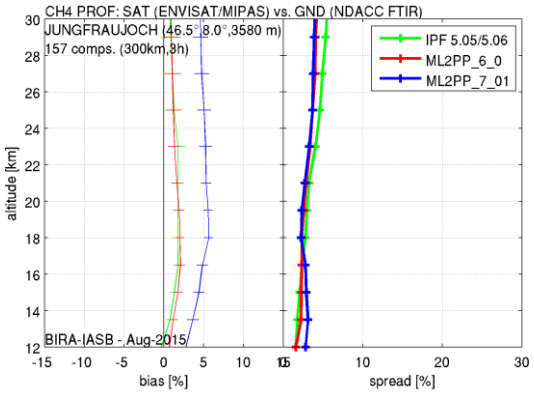

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
BREMEN (53.1°,8.8°,27 m) - IUP
161 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

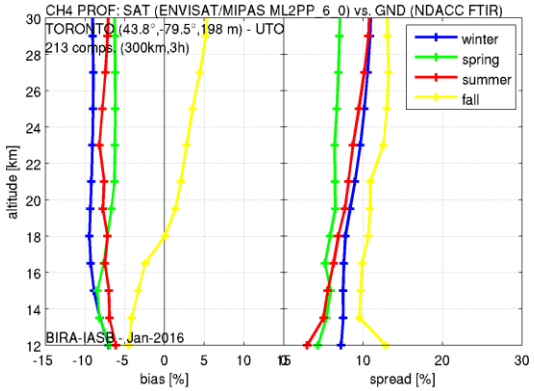

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
BREMEN (53.1°,8.8°,27 m) - IUP
75 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

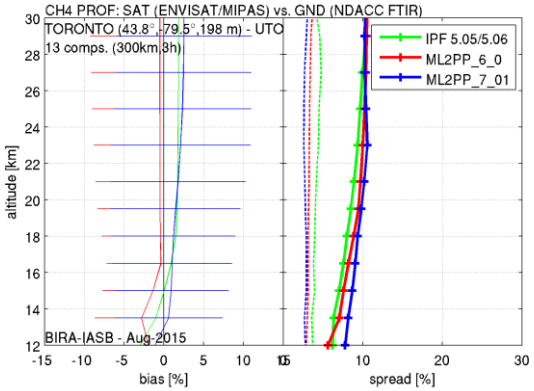

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
ZUGSPITZE (47.4°,11.0°,2964 m) - KI
783 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

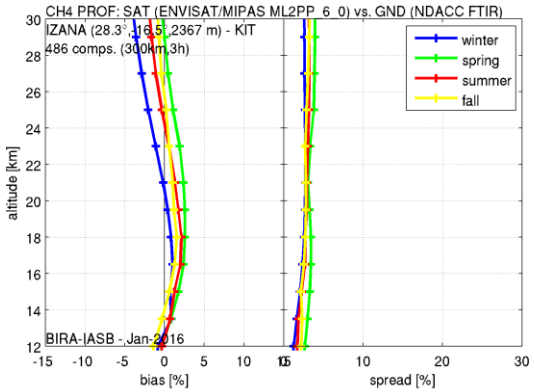

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
ZUGSPITZE (47.4°,11.0°,2964 m) - KI
384 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

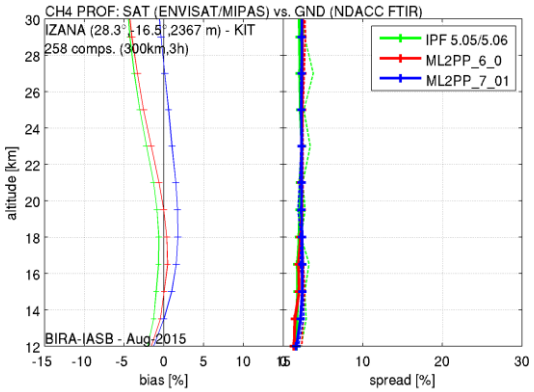

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
JUNGFRAUJOCH (46.5°,8.0°,3580 m)
299 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

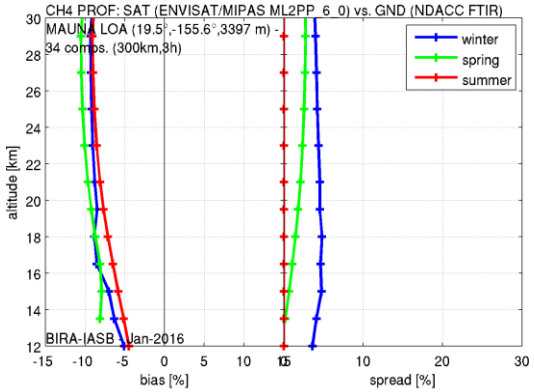

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
JUNGFRAUJOCH (46.5°,8.0°,3580 m)
157 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

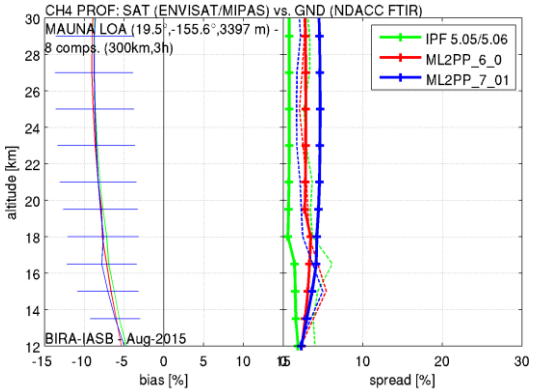

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
TORONTO (43.8°,-79.5°,198 m) - UTO
213 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

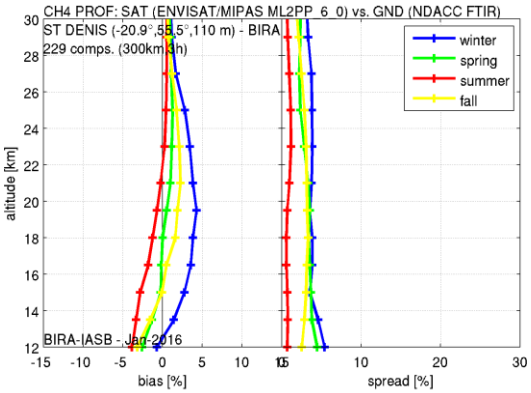

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
TORONTO (43.8°,-79.5°,198 m) - UTO
13 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

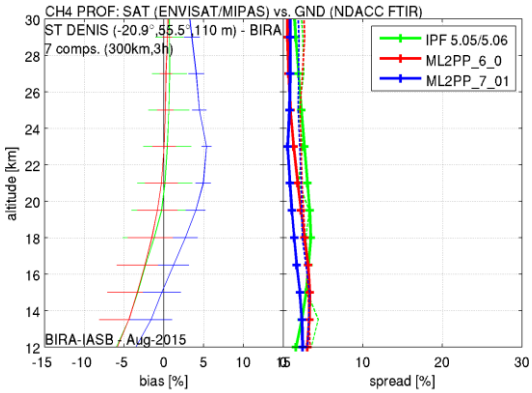

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
IZANA (28.3°,-16.5°,2367 m) - KIT
486 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

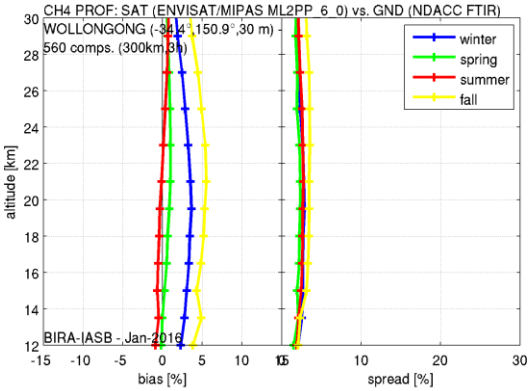

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
IZANA (28.3°,-16.5°,2367 m) - KIT
258 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

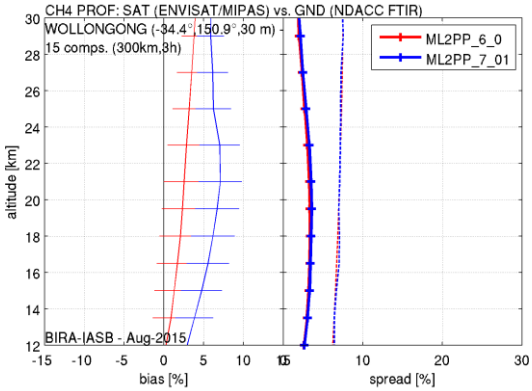

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
MAUNA LOA (19.5°,-155.6°,3397 m) -
34 comps. (300km,3h)
winter
spring
summer
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]


CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
MAUNA LOA (19.5°,-155.6°,3397 m) -
8 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]


CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
ST DENIS (-20.9°,55.5°,110 m) - BIRA
229 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

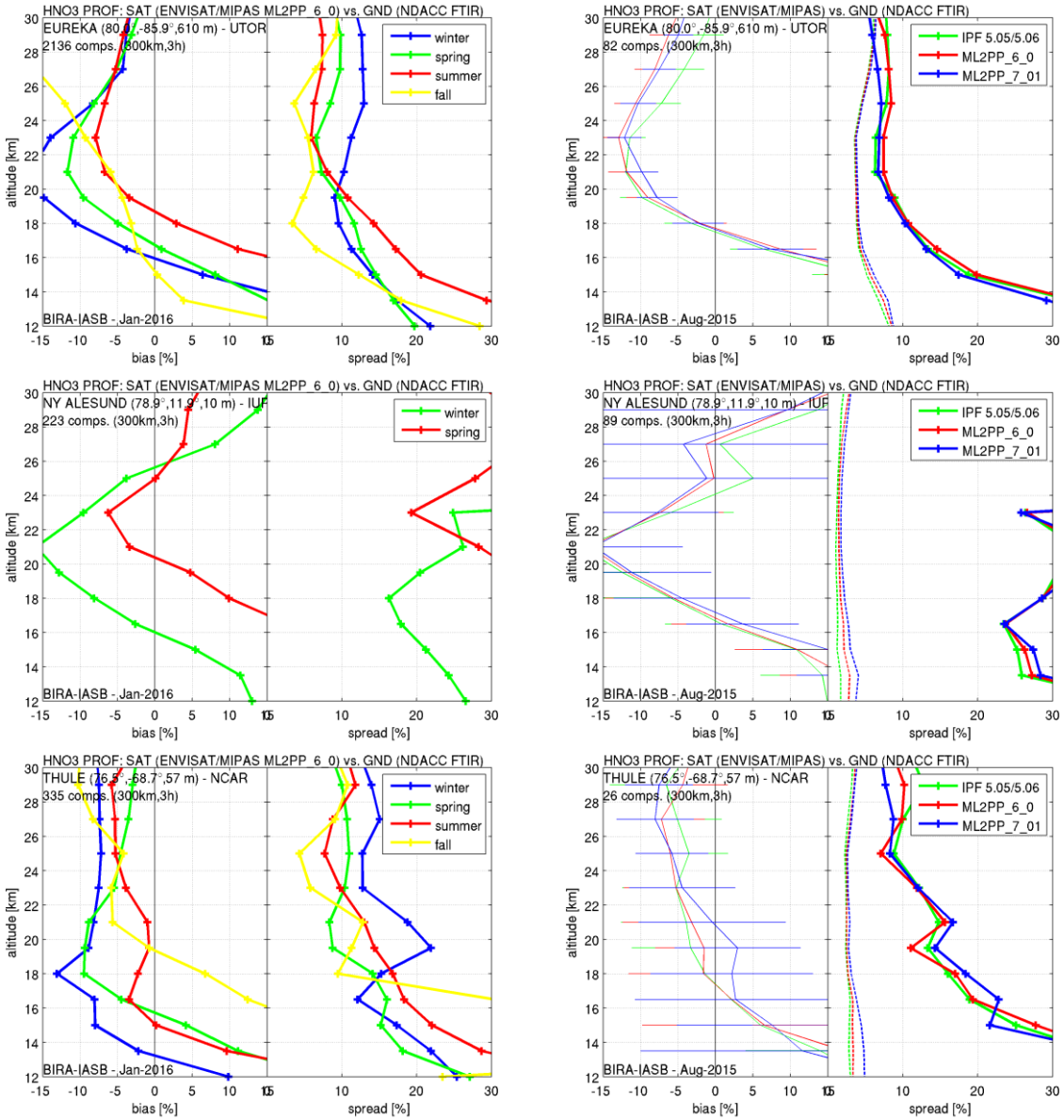

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
ST DENIS (-20.9°,55.5°,110 m) - BIRA
7 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

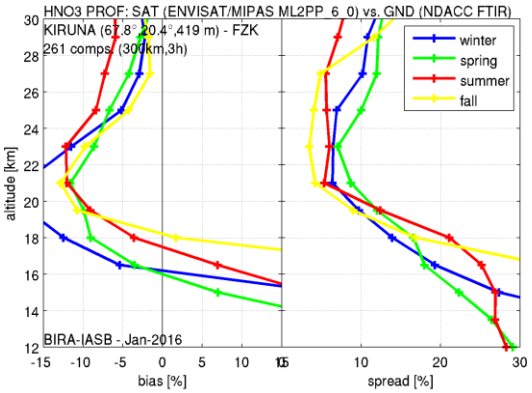

CH4 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
WOLLONGONG (-34.4°,150.9°,30 m) -
560 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

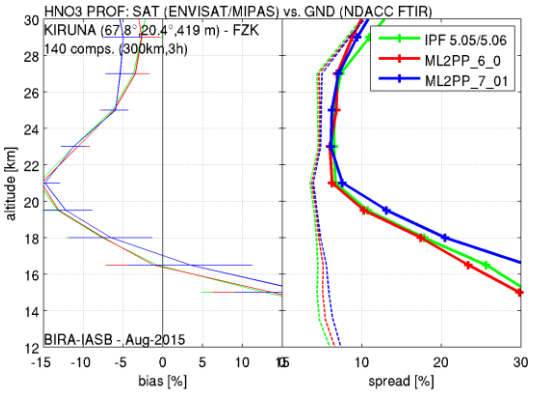

CH4 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
WOLLONGONG (-34.4°,150.9°,30 m) -
15 comps. (300km,3h)
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

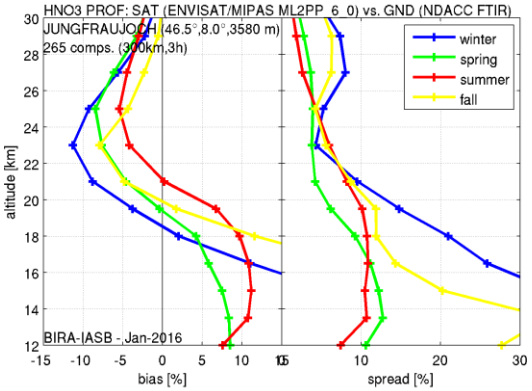
A2 MIPAS HNO3 vertical profile

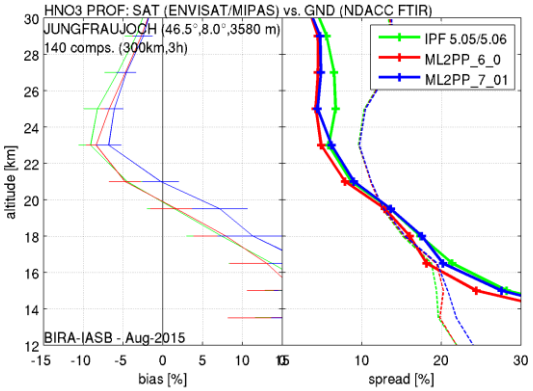
Figure 35: MIPAS HNO3 profile comparison plots between 12 and 30 km. Relative bias and spread are assessed in each station-specific graph (rows, sorted north to south), with a seasonal separation for the full ML2PP 6.0 validation (left column) and a processor version distinction for the delta-validation exercise (right column). For the latter also the random error on the mean (as horizontal lines at each altitude) and the combined satellite-ground error (as dashed lines in the spread plots) are provided. The station location and number of co-locations are shown in the upper-left corner of each box.

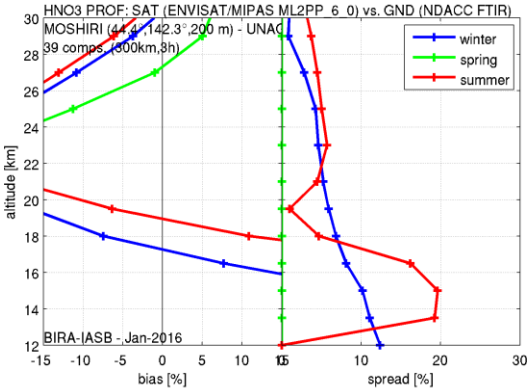


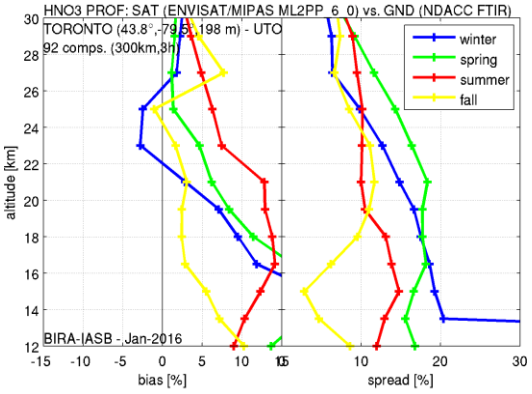

HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
KIRUNA (67.8°,20.4°,419 m) - FZK
261 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

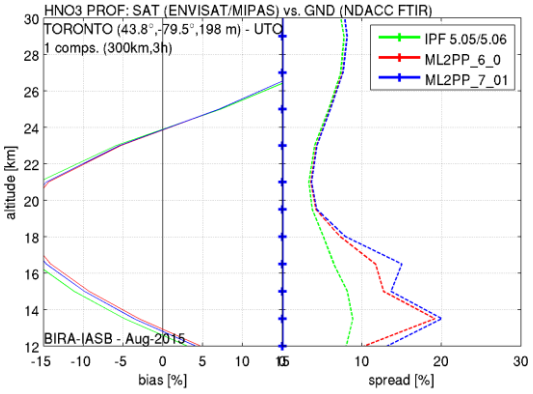

HNO3 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
KIRUNA (67.8°,20.4°,419 m) - FZK
140 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

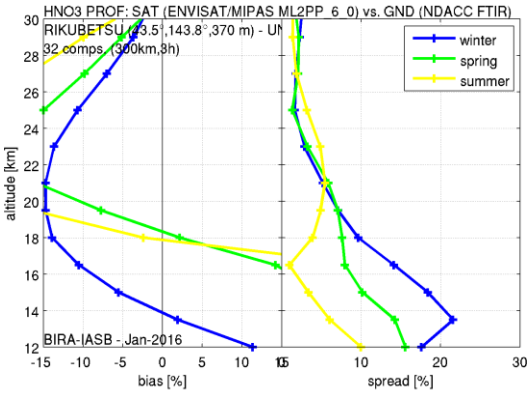

HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
JUNGFRAUJOCH (46.5°,8.0°,3580 m)
265 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

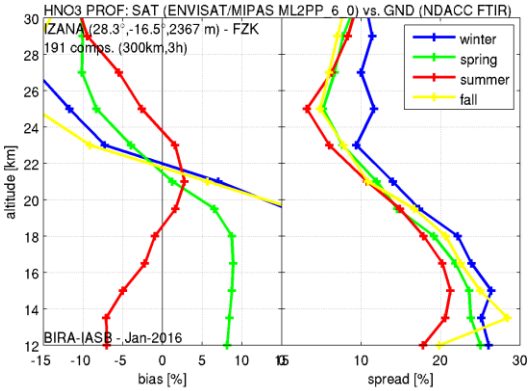

HNO3 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
JUNGFRAUJOCH (46.5°,8.0°,3580 m)
140 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

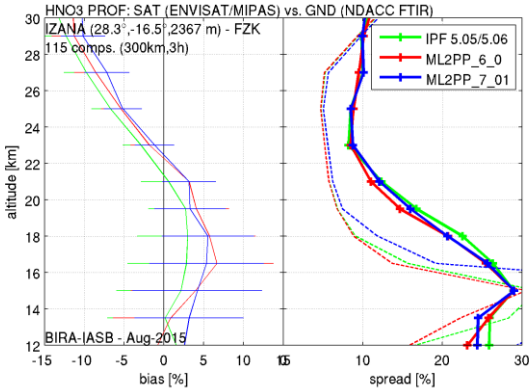

HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
MOSHIRI (44.4°,142.3°,200 m) - UNAC
39 comps. (300km,3h)
winter
spring
summer
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

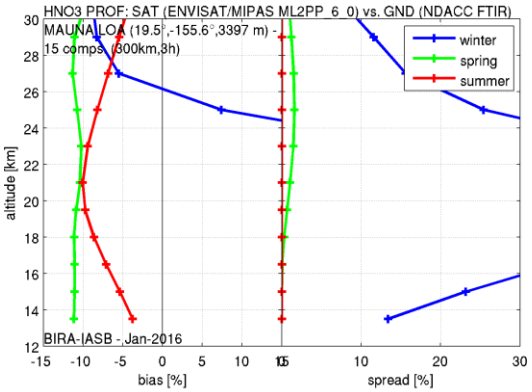

HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
TORONTO (43.8°,-79.5°,198 m) - UTO
92 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

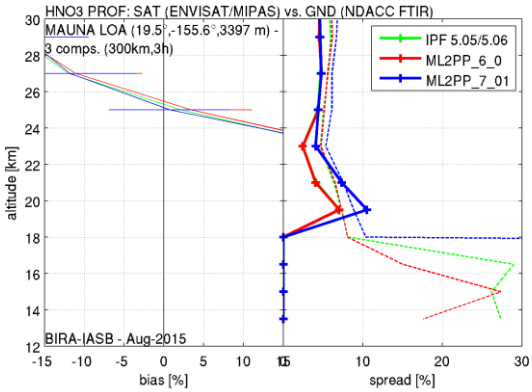

HNO3 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
TORONTO (43.8°,-79.5°,198 m) - UTO
1 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

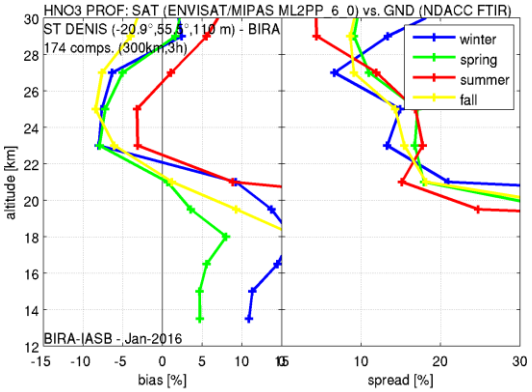

HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
RIKUBETSU (43.5°,143.8°,370 m) - UI
32 comps. (300km,3h)
winter
spring
summer
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

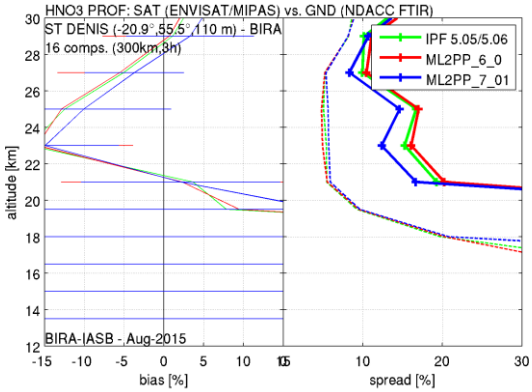

HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
IZANA (28.3°,-16.5°,2367 m) - FZK
191 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]


HNO3 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
IZANA (28.3°,-16.5°,2367 m) - FZK
115 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]


HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
MAUNA LOA (19.5°,-155.6°,3397 m) -
15 comps. (300km,3h)
winter
spring
summer
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

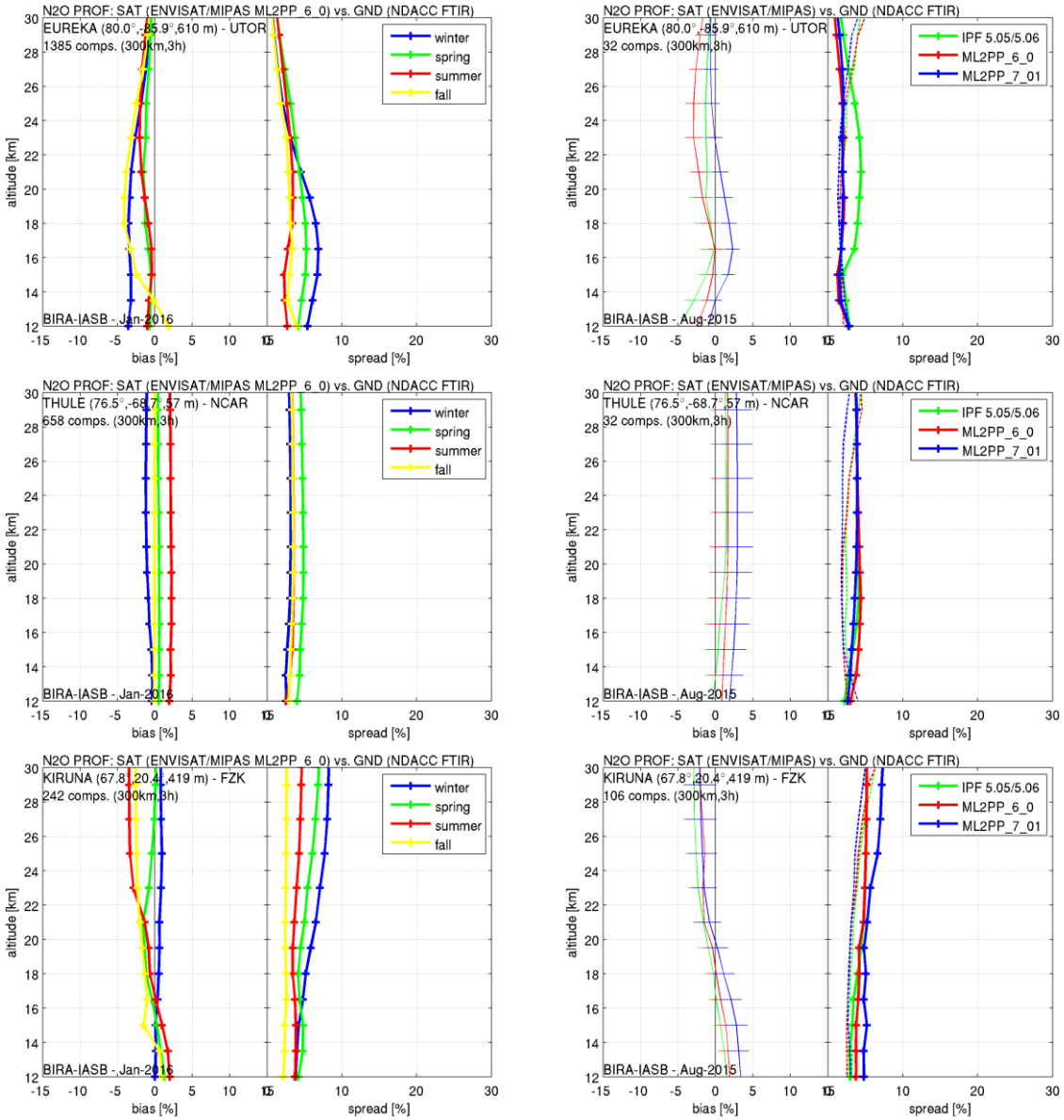

HNO3 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
MAUNA LOA (19.5°,-155.6°,3397 m) -
3 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

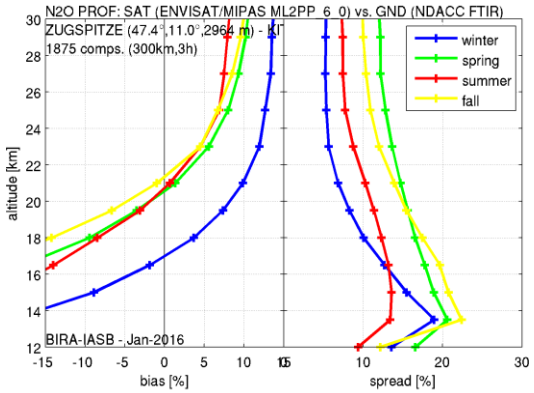

HNO3 PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
ST DENIS (-20.9°,55.5°,110 m) - BIRA
174 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

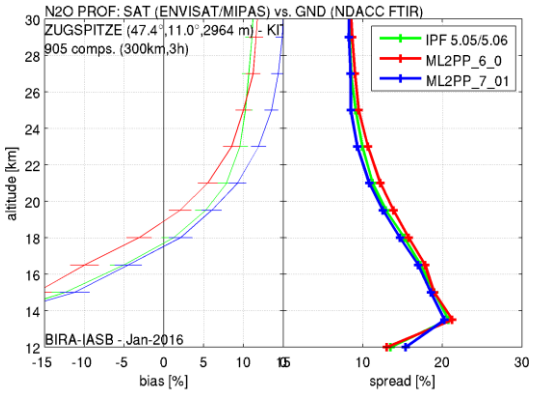

HNO3 PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
ST DENIS (-20.9°,55.5°,110 m) - BIRA
16 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

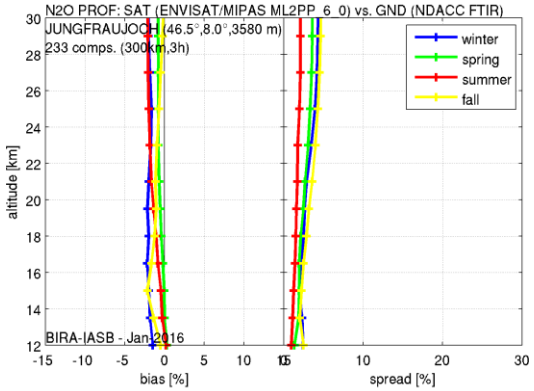
A3 MIPAS N2O vertical profile

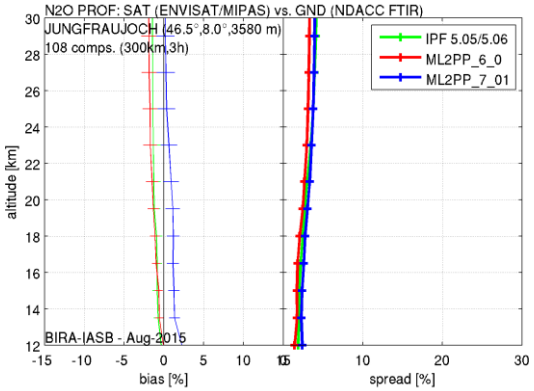
Figure 36: MIPAS N2O profile comparison plots between 12 and 30 km. Relative bias and spread are assessed in each station-specific graph (rows, sorted north to south), with a seasonal separation for the full ML2PP 6.0 validation (left column) and a processor version distinction for the delta-validation exercise (right column). For the latter also the random error on the mean (as horizontal lines at each altitude) and the combined satellite-ground error (as dashed lines in the spread plots) are provided. The station location and number of co-locations are shown in the upper-left corner of each box.

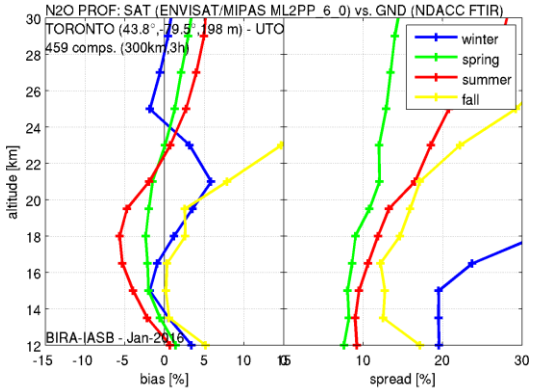


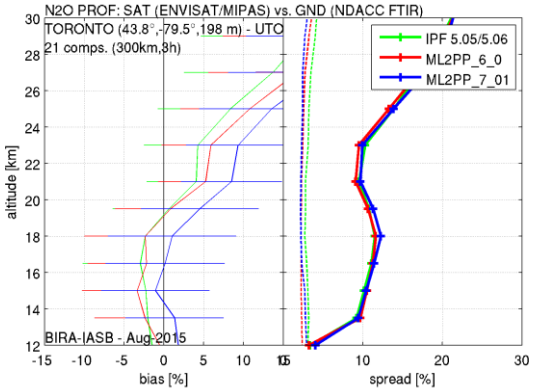

N2O PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
ZUGSPITZE (47.4°,11.0°,2964 m) - KIT
1875 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

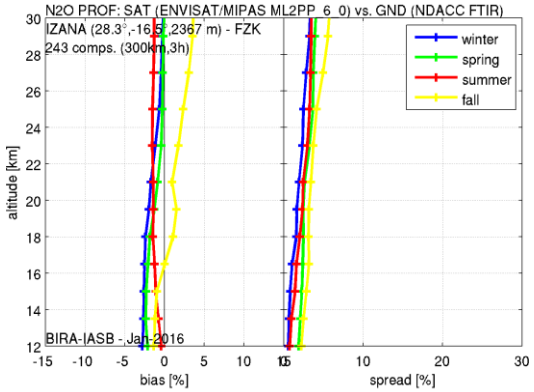

N2O PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
ZUGSPITZE (47.4°,11.0°,2964 m) - KIT
905 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

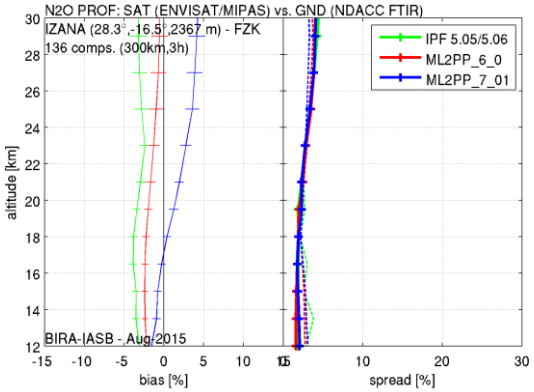

N2O PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
JUNGFRAUJOCH (46.5°,8.0°,3580 m)
233 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

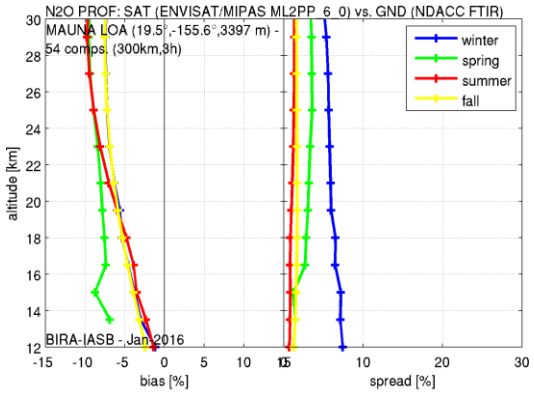

N2O PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
JUNGFRAUJOCH (46.5°,8.0°,3580 m)
108 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

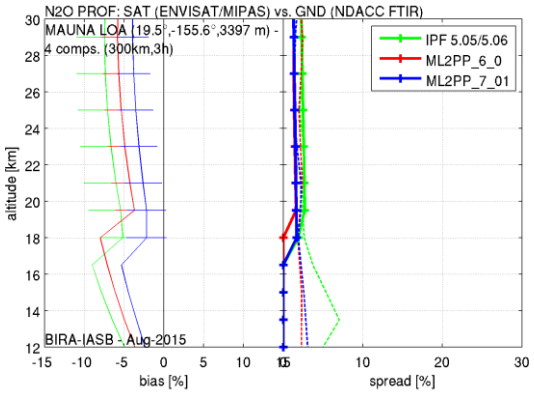

N2O PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
TORONTO (43.8°,-79.5°,198 m) - UTO
459 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

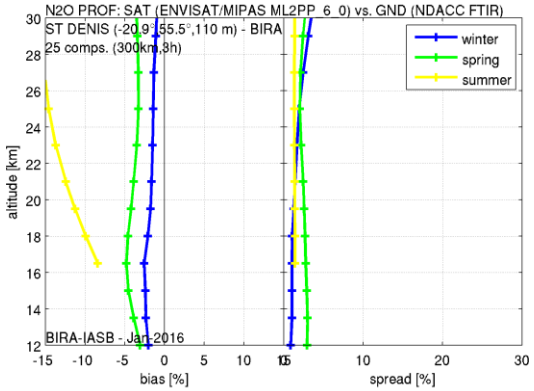

N2O PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
TORONTO (43.8°,-79.5°,198 m) - UTO
21 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

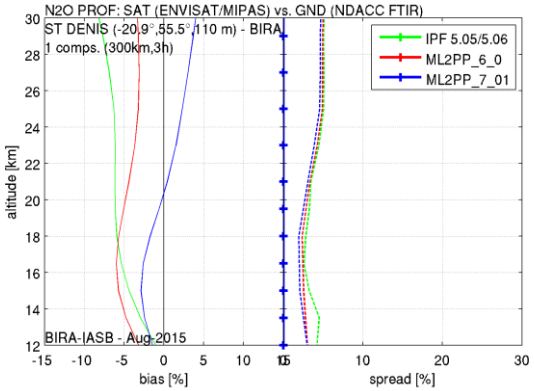

N2O PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
IZANA (28.3°,-16.5°,2367 m) - FZK
243 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

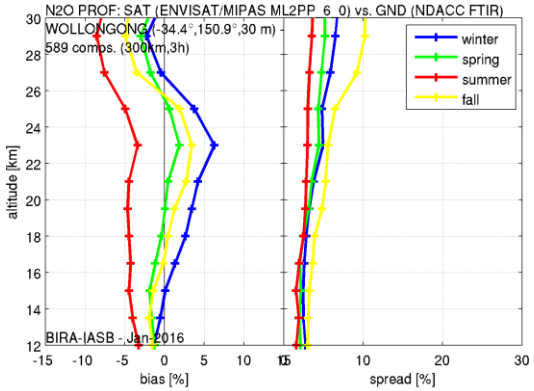

N2O PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
IZANA (28.3°,-16.5°,2367 m) - FZK
136 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

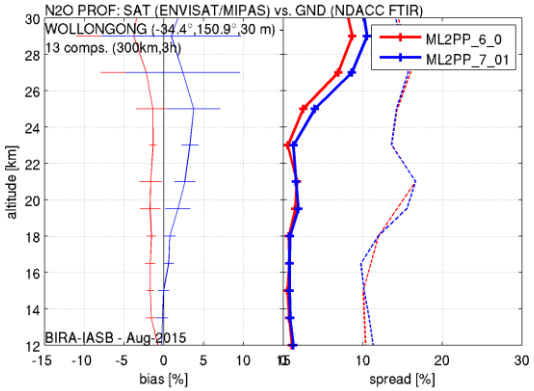

N2O PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
MAUNA LOA (19.5°,-155.6°,3397 m) -
54 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]


N2O PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
MAUNA LOA (19.5°,-155.6°,3397 m) -
4 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]


N2O PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
ST DENIS (-20.9°,55.5°,110 m) - BIRA
25 comps. (300km,3h)
winter
spring
summer
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

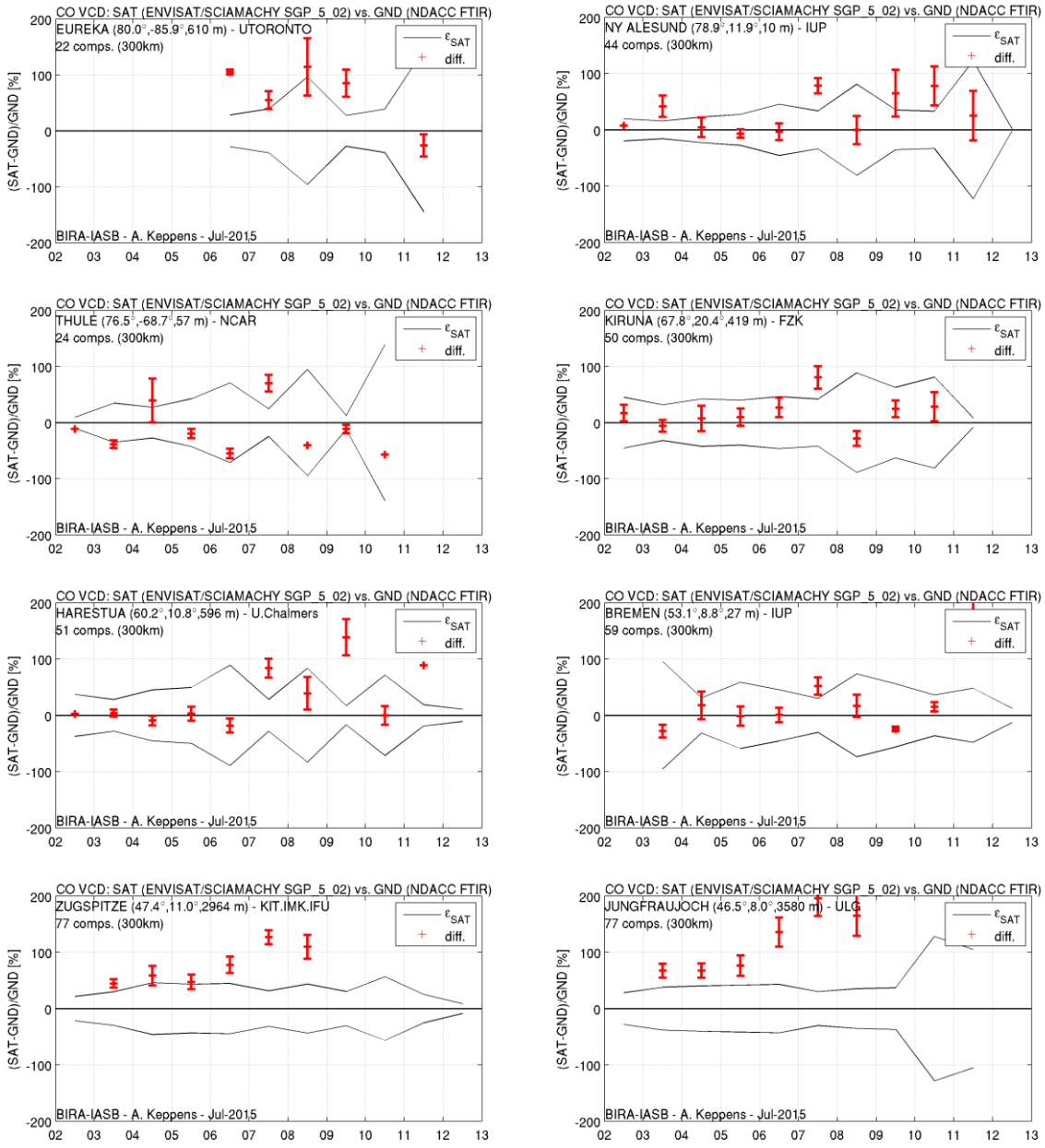

N2O PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
ST DENIS (-20.9°,55.5°,110 m) - BIRA
1 comps. (300km,3h)
IPF 5.05/5.06
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

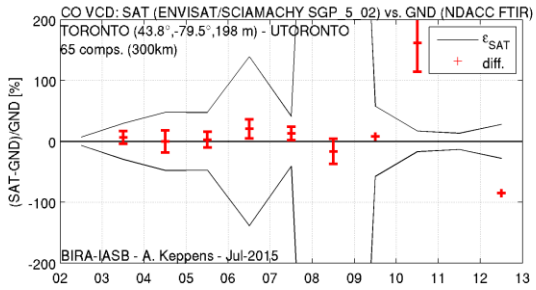

N2O PROF: SAT (ENVISAT/MIPAS ML2PP_6_0) vs. GND (NDACC FTIR)
WOLLONGONG (-34.4°,150.9°,30 m) -
589 comps. (300km,3h)
winter
spring
summer
fall
BIRA-IASB - Jan-2016
altitude [km]
bias [%]
spread [%]

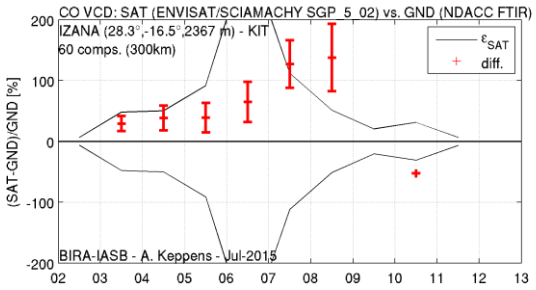

N2O PROF: SAT (ENVISAT/MIPAS) vs. GND (NDACC FTIR)
WOLLONGONG (-34.4°,150.9°,30 m) -
13 comps. (300km,3h)
ML2PP_6_0
ML2PP_7_01
BIRA-IASB - Aug-2015
altitude [km]
bias [%]
spread [%]

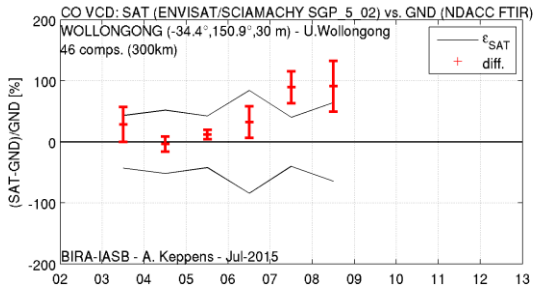
A4 SCIAMACHY nadir CO total column

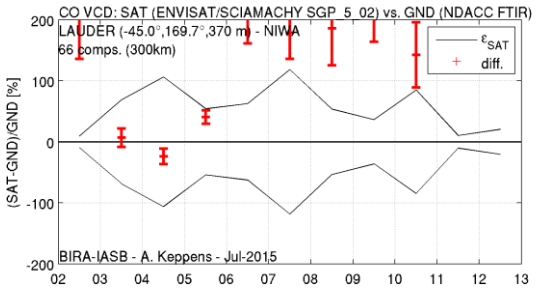
Figure 37: Time series of the median (red crosses) and 68 % interpercentile (black lines) of the relative monthly differences between SCIAMACHY SGP 5.02 nadir CO total column data and NDACC ground-based FTIR data at 13 stations (sorted north to south). Error flags (red vertical lines) equal to the standard errors on the yearly means have been added as well. The station location and number of co-locations (i.e. number of monthly means) are shown in the upper-left corner of each box.

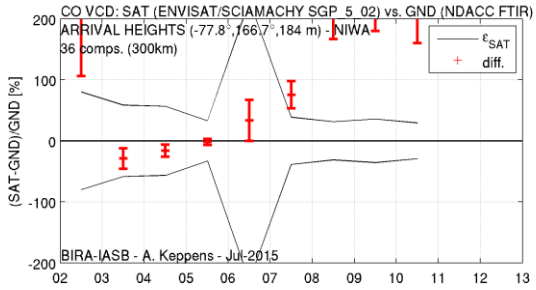



CO VCD: SAT (ENVISAT/SCIAMACHY SGP_5_02) vs. GND (NDACC FTIR)
TORONTO (43.8°,-79.5°,198 m) - UTORONTO
65 comps. (300km)
ε SAT
diff.
(SAT-GND)/GND [%]
BIRA-IASB - A. Keppens - Jul-2015


CO VCD: SAT (ENVISAT/SCIAMACHY SGP_5_02) vs. GND (NDACC FTIR)
IZANA (28.3°,-16.5°,2367 m) - KIT
60 comps. (300km)
ε SAT
diff.
(SAT-GND)/GND [%]
BIRA-IASB - A. Keppens - Jul-2015


CO VCD: SAT (ENVISAT/SCIAMACHY SGP_5_02) vs. GND (NDACC FTIR)
WOLLONGONG (-34.4°,150.9°,30 m) - U.Wollongong
46 comps. (300km)
ε SAT
diff.
(SAT-GND)/GND [%]
BIRA-IASB - A. Keppens - Jul-2015


CO VCD: SAT (ENVISAT/SCIAMACHY SGP_5_02) vs. GND (NDACC FTIR)
LAUDER (-45.0°,169.7°,370 m) - NIWA
66 comps. (300km)
ε SAT
diff.
(SAT-GND)/GND [%]
BIRA-IASB - A. Keppens - Jul-2015


CO VCD: SAT (ENVISAT/SCIAMACHY SGP_5_02) vs. GND (NDACC FTIR)
ARRIVAL HEIGHTS (-77.8°,166.7°,184 m) - NIWA
36 comps. (300km)
ε SAT
diff.
(SAT-GND)/GND [%]
BIRA-IASB - A. Keppens - Jul-2015

A5 SCIAMACHY nadir H2O total column

A5.1 SGP 5.02 full mission

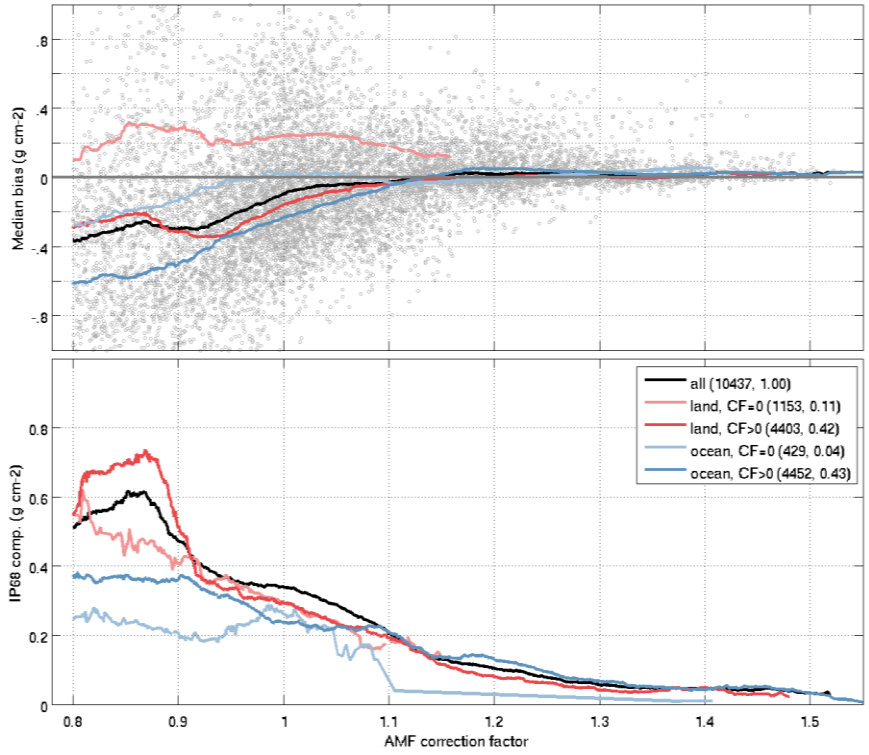


Figure 38: Dependence on AMF correction factor of the median (top) and 68% interpercentile (bottom) of the absolute difference distribution of SCIAMACHY SGP 5.02 nadir H2O total column minus radiosonde measurements. Grey markers shows the entire co-location sample, curves represent running median statistics for five SCIAMACHY pixel classes.

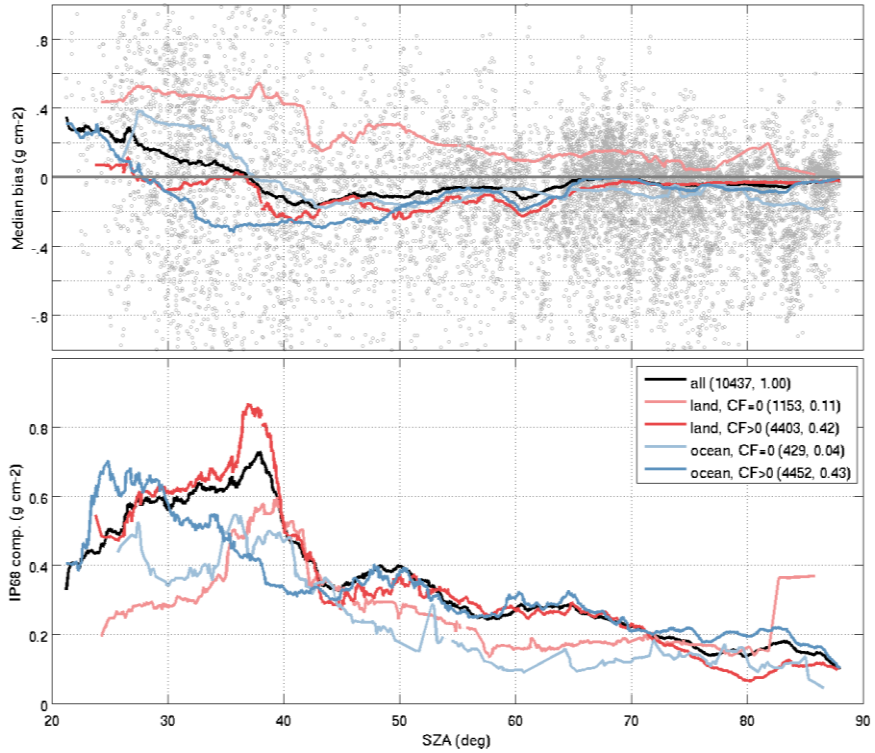


Figure 39: Like Figure 38, but for the dependence on solar zenith angle

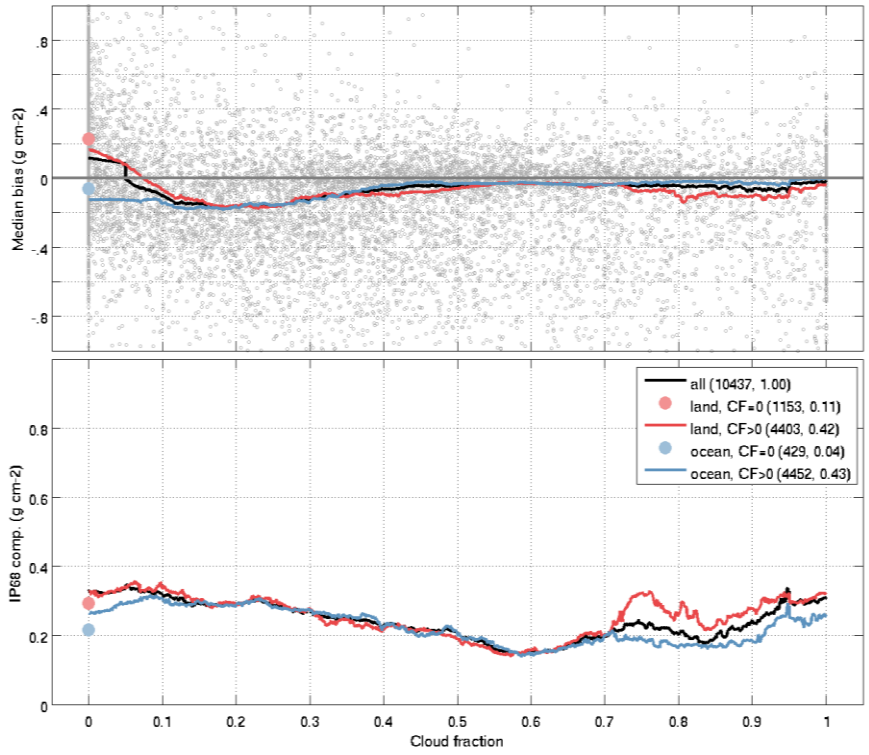


Figure 40: Like Figure 38, but for the dependence on cloud fraction.

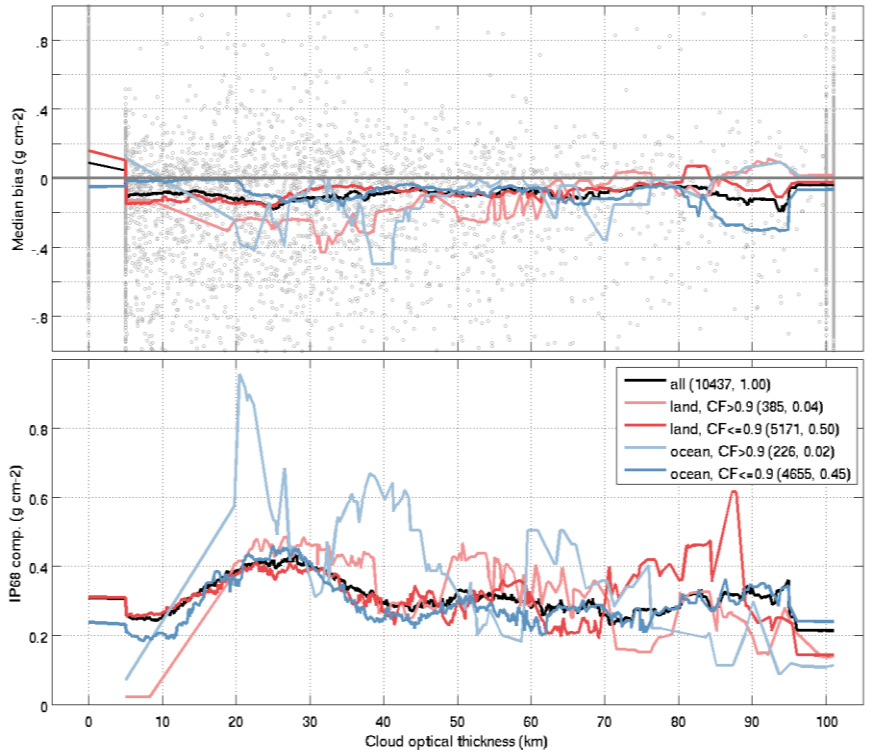


Figure 41: Like Figure 38, but for the dependence on cloud optical thickness.

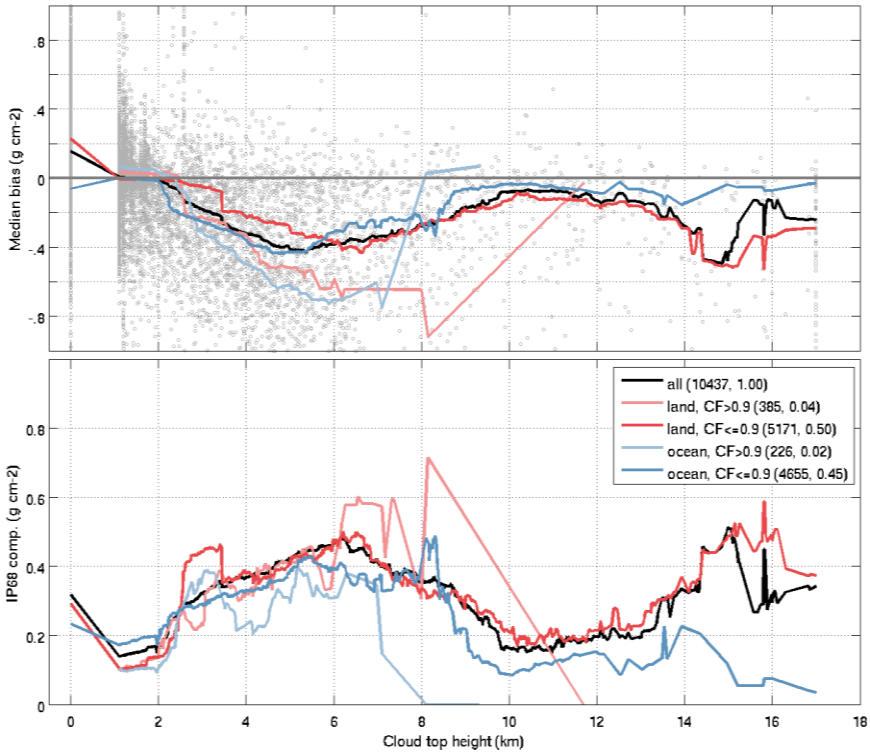


Figure 42: Like Figure 38, but for the dependence on cloud top height under differing cloud conditions. Note: the cloud class definition (see legend) is different from that in the other figures.

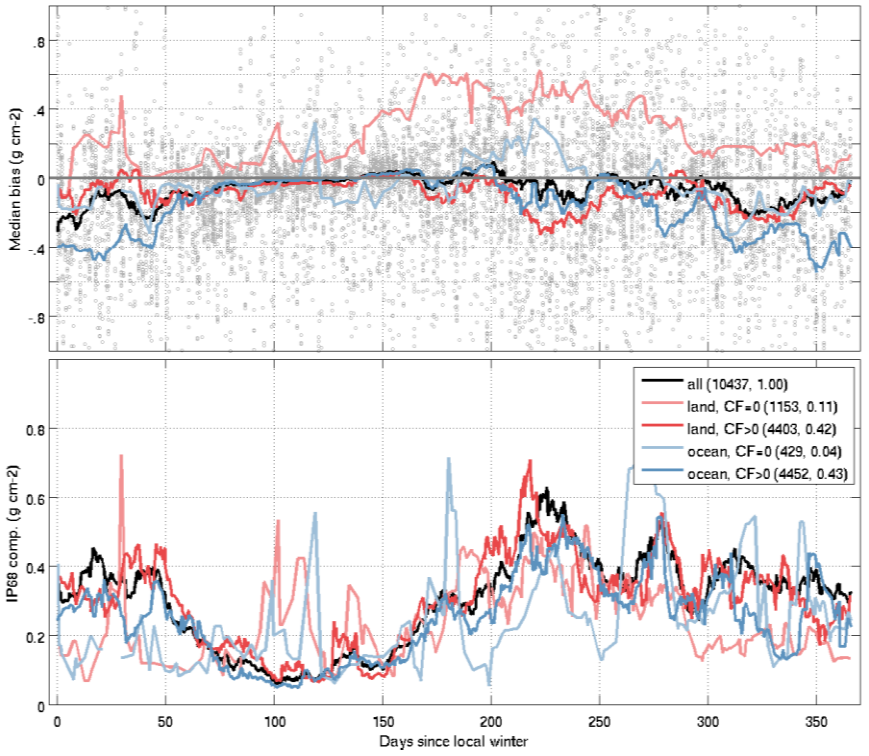


Figure 43: Like Figure 38, but for the dependence on season. The x-axis depicts the number of days since the start of the local astronomical winter.

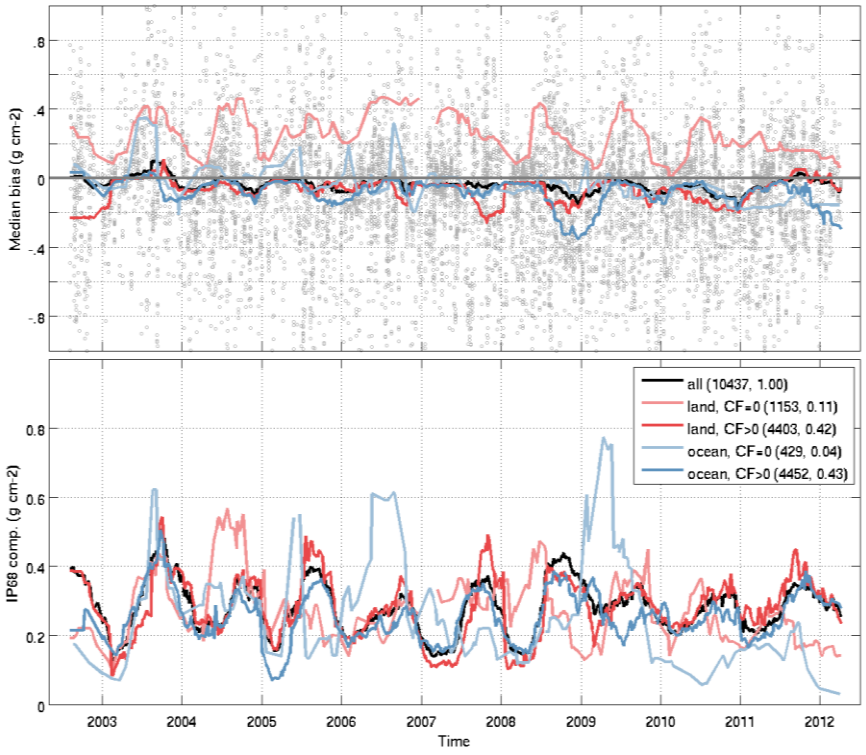


Figure 44: Like Figure 38, but for the dependence on time

A5.2 SGP 6.00 diagnostic data set

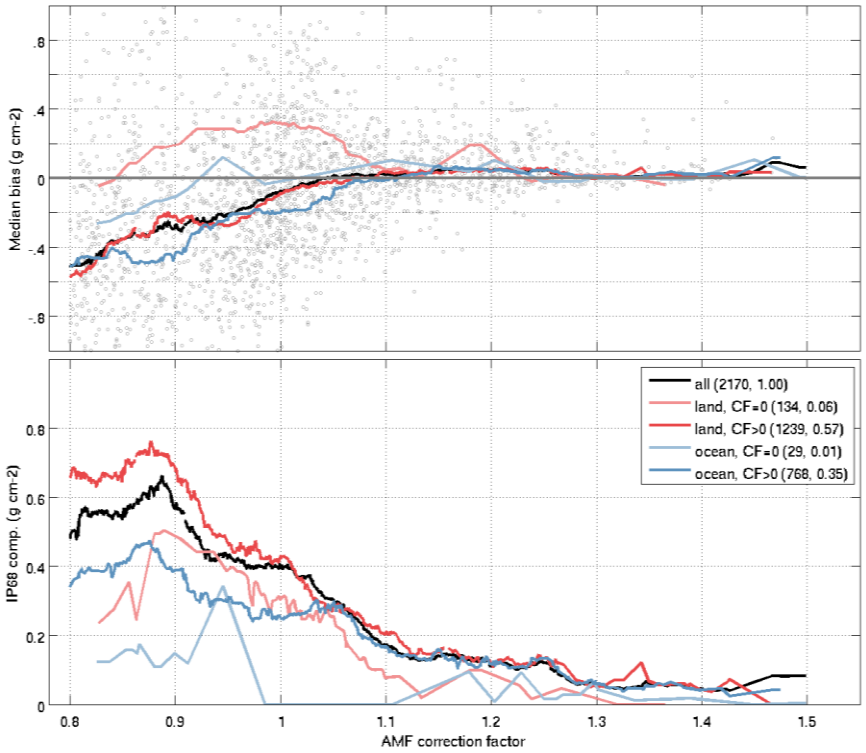


Figure 45: Dependence on AMF correction factor of the median (top) and 68% interpercentile (bottom) of the absolute difference distribution of SCIAMACHY SGP 6.00 DDS nadir H2O total column minus radiosonde measurements. Grey markers shows the entire co-location sample, curves represent running median statistics for five SCIAMACHY pixel classes.

Figure 46: Like Figure 45, but for the dependence on solar zenith angle

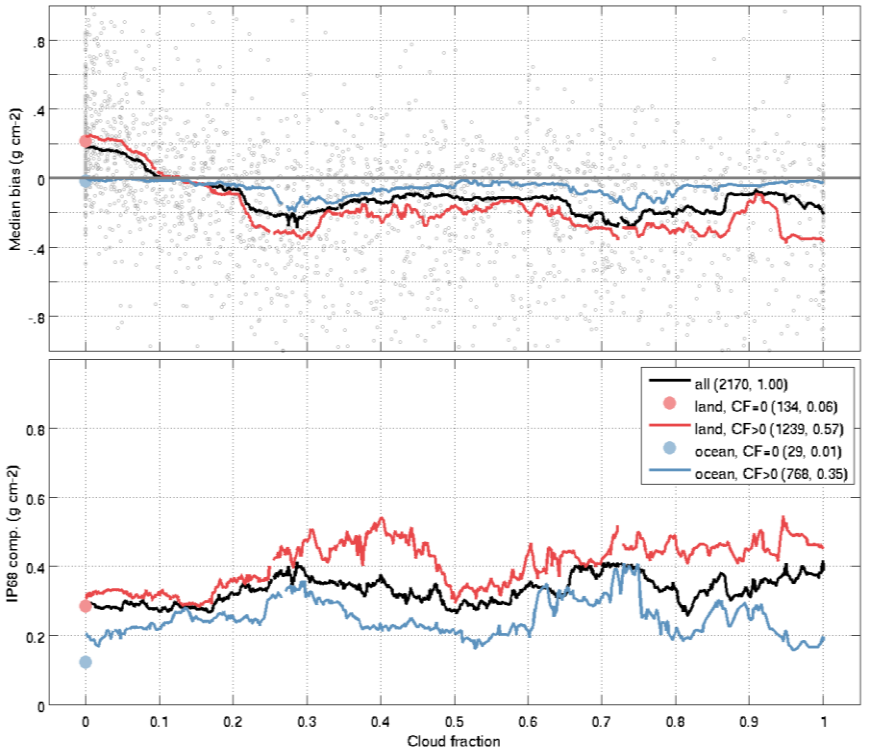


Figure 47: Like Figure 45, but for the dependence on cloud fraction.

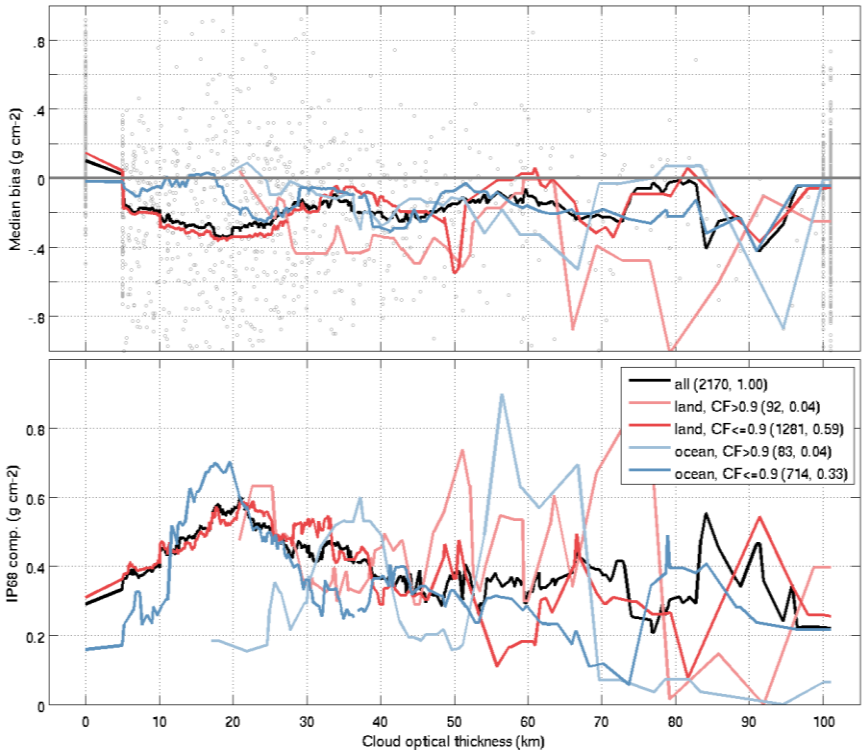


Figure 48: Like Figure 45, but for the dependence on cloud optical thickness.

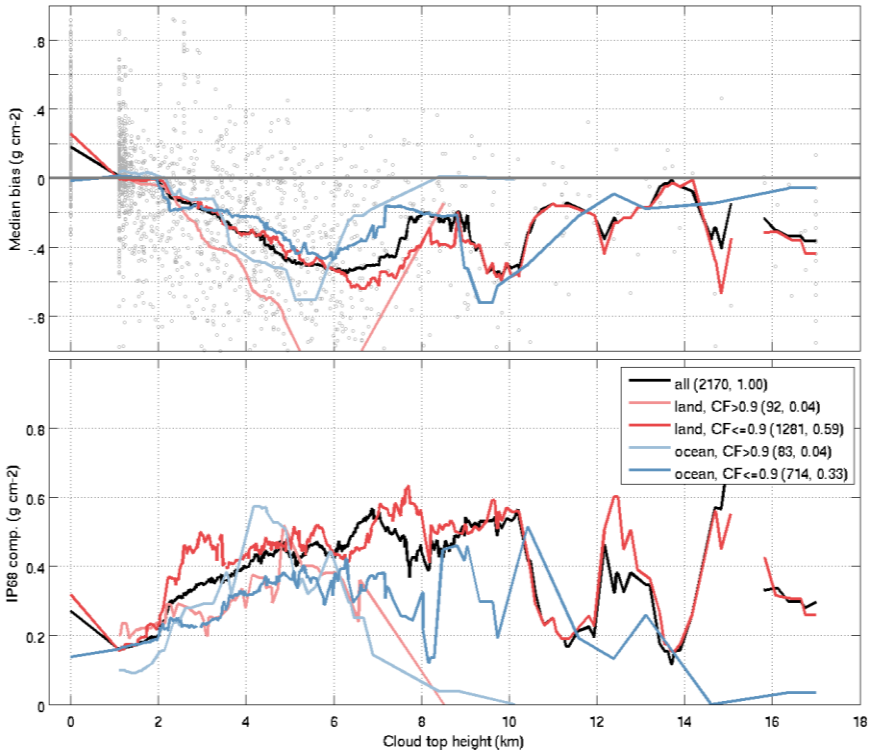


Figure 49: Like Figure 45, but for the dependence on cloud top height under differing cloud conditions. Note: the cloud class definition (see legend) is different from that in the other figures.

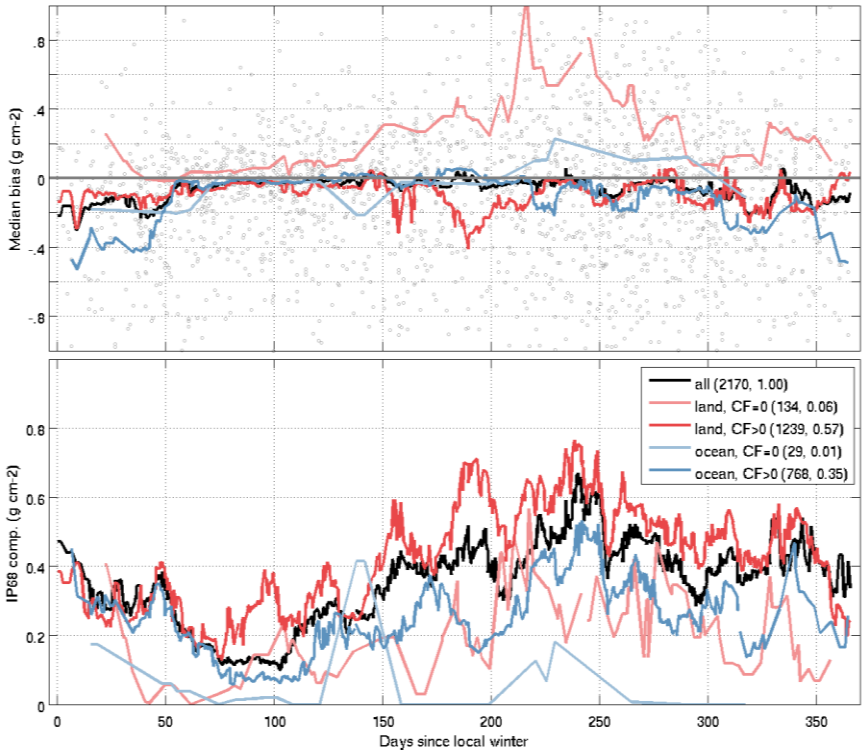


Figure 50: Like Figure 45, but for the dependence on season. The x-axis depicts the number of days since the start of the local astronomical winter.

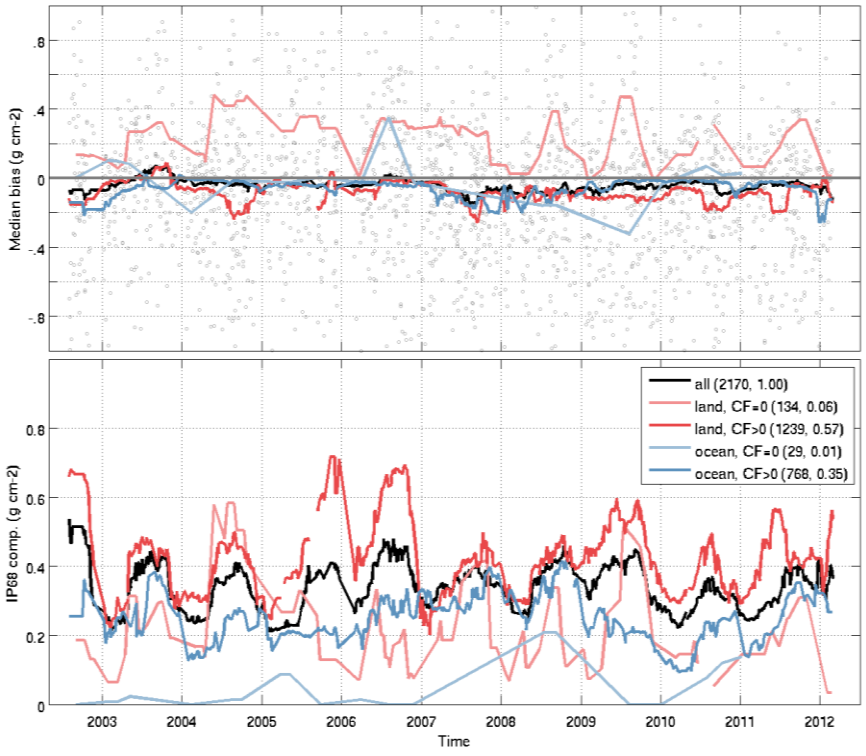


Figure 51: Like Figure 45, but for the dependence on time

Annexe B : Articles and presentations

Peer-reviewed articles

Nair, P. J., Godin-Beekmann, S., Froidevaux, L., Flynn, L. E., Zawodny, J. M., Russell III, J. M., Pazmiño, A., Ancellet, G., Steinbrecht, W., Claude, H., Leblanc, T., McDermid, S., van Gijsel, J. A. E., Johnson, B., Thomas, A., Hubert, D., Lambert, J.-C., Nakane, H., and Swart, D. P. J.: *Relative drifts and stability of satellite and ground-based stratospheric ozone profiles at NDACC lidar stations*, Atmos. Meas. Tech., 5, 1301-1318, doi:10.5194/amt-5-1301-2012, 2012.

Lambert, J.-C., C. De Clercq, and T. von Clarmann, *Comparing and merging water vapour observations: A multi-dimensional perspective on smoothing and sampling issues*, Chapter 9 (p. 177-199) of book "Monitoring Atmospheric Water Vapour: Ground-Based Remote Sensing and In-situ Methods", N. Kämpfer (Ed.), ISSI Scientific Report Series, Vol. 10, Edition 1, 326 p., ISBN: 978-1-4614-3908-0, doi:10.1007/978-1-4614-3909-7_2, © Springer New York 2012.

Richter, A., M. Weber, J. P. Burrows, J.-C. Lambert, and A. van Gijsel, *Validation Strategy for Satellite Observations of Tropospheric Reactive Gases*, Annals of Geophysics, 56, doi:10.4401/AG-6335, 2013.

Adams, C., Bourassa, A. E., Sofieva, V., Froidevaux, L., McLinden, C. A., Hubert, D., Lambert, J.-C., Sioris, C. E., and Degenstein, D. A.: *Assessment of Odin-OSIRIS ozone measurements from 2001 to the present using MLS, GOMOS, and ozonesondes*, Atmos. Meas. Tech., 7, 49-64, doi:10.5194/amt-7-49-2014, 2014.

Hassler, B., Petropavlovskikh, I., Staehelin, J., August, T., Bhartia, P. K., Clerbaux, C., Degenstein, D., De Mazière, M., Dinelli, B. M., Dudhia, A., Dufour, G., Frith, S. M., Froidevaux, L., Godin-Beekmann, S., Granville, J., Harris, N. R. P., Hoppel, K., Hubert, D., Kasai, Y., Kurylo, M. J., Kyrölä, E., Lambert, J.-C., Levelt, P. F., McElroy, C. T., McPeters, R. D., Munro, R., Nakajima, H., Parrish, A., Raspollini, P., Remsberg, E. E., Rosenlof, K. H., Rozanov, A., Sano, T., Sasano, Y., Shiotani, M., Smit, H. G. J., Stiller, G., Tamminen, J., Tarasick, D. W., Urban, J., van der A, R. J., Veefkind, J. P., Vigouroux, C., von Clarmann, T., von Savigny, C., Walker, K. A., Weber, M., Wild, J., and Zawodny, J. M.: *Past changes in the vertical distribution of ozone – Part 1: Measurement techniques, uncertainties and availability*, Atmos. Meas. Tech., 7, 1395-1427, doi:10.5194/amt-7-1395-2014, 2014.

Harris, N. R. P., Hassler, B., Tummon, F., Bodeker, G. E., Hubert, D., Petropavlovskikh, I., Steinbrecht, W., Anderson, J., Bhartia, P. K., Boone, C. D., Bourassa, A., Davis, S. M., Degenstein, D., Delcloo, A., Frith, S. M., Froidevaux, L., Godin-Beekmann, S., Jones, N., Kurylo, M. J., Kyrölä, E., Laine, M., Leblanc, S. T., Lambert, J.-C., Liley, B., Mahieu, E., Maycock, A., de Mazière, M., Parrish, A., Querel, R., Rosenlof, K. H., Roth, C., Sioris, C., Staehelin, J., Stolarski, R. S., Stübi, R., Tamminen, J., Vigouroux, C., Walker, K. A., Wang, H. J., Wild, J., and Zawodny, J. M.: *Past changes in the vertical distribution of ozone – Part 3: Analysis and interpretation of trends*, Atmos. Chem. Phys., 15, 9965-9982, doi:10.5194/acp-15-9965-2015, 2015.

Laeng, A., Hubert, D., Verhoelst, T., von Clarmann, T., Dinelli, B.M., Dudhia, A., Raspollini, P., Stiller, G., Grabowski, U., Keppens, A., Kiefer, M., Sofieva, V., Froidevaux, L., Walker, K.A., Lambert, J.-C., and Zehner, C. : *The ozone climate change initiative: Comparison of four Level-2 processors for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS)*, Remote Sens. Environ., 162, 316-343, doi: 10.1016/j.rse.2014.12.013, 2015.

Hubert, D., Lambert, J.-C., Verhoelst, T., Granville, J., Keppens, A., Baray, J.-L., Cortesi, U., Degenstein, D. A., Froidevaux, L., Godin-Beekmann, S., Hoppel, K. W., Kyrölä, E., Leblanc, T., Lichtenberg, G., McElroy, C. T., Murtagh, D., Nakane, H., Russell III, J. M., Salvador, J., Smit, H. G. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Stübi, R., Swart, D. P. J., Taha, G., Thompson, A. M., Urban, J., van Gijsel, J. A. E., von der Gathen, P., Walker, K. A., Wolfram, E., and Zawodny, J. M.: *Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records*, Atmos. Meas. Tech. Discuss., 8, 6661-6757, doi:10.5194/amtd-8-6661-2015, 2015.

van Gijsel, J. A. E., Zurita-Milla, R., Stammes, P., Godin-Beekmann, S., Leblanc, T., Marchand, M., McDermid, I. S., Stebel, K., Steinbrecht, W., and Swart, D. P. J.: *Using self-organising maps to explore ozone profile validation results – SCIAMACHY limb compared to ground-based lidar observations*, Atmos. Meas. Tech., 8, 1951-1963, doi:10.5194/amt-8-1951-2015, 2015.

Nair, P. J., L. Froidevaux, J. Kuttippurath, J. M. Zawodny, J. M. Russell III, W. Steinbrecht, H. Claude, T. Leblanc, J. A. E. van Gijsel, B. Johnson, D. P. J. Swart, A. Thomas, R. Querel, R. Wang, and J. Anderson : *Subtropical and midlatitude ozone trends in the stratosphere: Implications for recovery*, J. Geophys. Res. Atmos., 120, 7247–7257, doi:10.1002/2014JD022371, 2015.

Technical notes

Hubert, D., J. Granville and J.-C. Lambert, *Influence of Envisat GOMOS data screening on IPF 6.01 ozone validation results*, TN-BIRA-IASB-GOMOS-IPF6-O3P-SCREENING, Validation Report GOMOS QWG, Issue 1.0 / Rev. 0, 18 pp., 31 October 2012.

GOMOS QWG, Contribution to *GOMOS Level 2 processor version GOMOS/6.01 Readme*, ESA, ENVI-GSOP-EOGD-QD-12-0117, 15 pp., 17 December 2012, http://earth.eo.esa.int/pcs/envisat/gomos/documentation/RMF_0117_GOM_NL__2P_Disclaimers.pdf.

Keppens, A., D. Hubert and J.-C. Lambert, *MIPAS Technical Note: Methodology for switching among MIPAS vertical averaging kernel matrix representations*, TN_BIRA-IASB_MIPAS-AKM-conversions_Keppens, Issue 1 / Rev. 2, 7 pp., 16 October 2014.

Hubert, D., A. Keppens, J. Granville, F. Hendrick, J.-C. Lambert, and A. van Gijsel, *Delta-validation of SCIAMACHY SGP upgrade from V5.02 to V6.00*, TN-BIRA-IASB-MultiTASTE-Phase-F-VR1-Iss2-RevA, Multi-TASTE validation report, Issue 2 / Rev. A, Final, 44 pp., 18 September 2015.

Hubert, D., J. Granville, and J.-C. Lambert, *Validation report - GOMOS ozone profiles : IPF V6.01 and Bright Limb V1.2,* TN-BIRA-IASB-MultiTASTE-PhaseF-GOMOS-O3p-IPF-GBL-1C, Multi-TASTE validation report, Issue 1 / Rev. C, Final, 30 pp., 8 October 2015.

Project reports

Hubert, D., A. Keppens, J. Granville, F. Hendrick and J.-C. Lambert, *Multi-TASTE Phase F – extension 2013: Progress Report (October 2013 – March 2014)*, TN-BIRA-IASB-MultiTASTE-Phase-F-PR1, Multi-TASTE progress report, Issue 1.0 / Rev. 0, 37 pp., 17 April 2014.

Oral presentations at Envisat QWG meetings

Hubert, D., A. Keppens, T. Verhoelst, J. Granville and J.-C. Lambert, *GOMOS IPF upgrade from V5 to V6 – Validation of O3 profile*, GOMOS QWG meeting #27, LATMOS, Guyancourt, France, 26-28 September 2012.

Hubert, D., A. Keppens, T. Verhoelst, J. Granville and J.-C. Lambert, *MIPAS ML2PP 6.0 ozone – Bias, spread and stability*, MIPAS QWG #30, ESRIN, Frascati, Italy, 12-14 November 2012.

Hubert, D., T. Verhoelst, A. Keppens, J. Granville and J.-C. Lambert, *Comparison of three SCIA limb ozone products SGP 3.01, SGP 5.02 and IUP V2.5*, SCIAVALIG workshop, KNMI, De Bilt, The Netherlands, 29-30 November 2012.

Hubert, D., A. Keppens, T. Verhoelst, J. Granville, J.-C. Lambert, *NDACC-based validation of SCIA limb ozone data SGP 3.01, SGP 5.02 and IUP V2.5*, SCIAMACHY QWG meeting #22, IUP, Bremen, Germany, 13-14 June 2013.

Hubert, D., A. Keppens, T. Verhoelst, J. Granville, J.-C. Lambert and A. van Gijsel, *MIPAS V6 ozone validation results: consistent (or not?)*, MIPAS QWG meeting #33, ISAC, Bologna, Italy, 4-6 November 2013.

Keppens, A., D. Hubert and J.-C. Lambert, Presentation of *MIPAS Technical Note: Methodology for switching among MIPAS vertical averaging kernel matrix representations*, MIPAS QWG meeting #35, ESRIN, Frascati, Italy, 24-26 March 2014.

Hubert, D., A. Keppens, J. Granville and J.-C. Lambert, *MIPAS V7.01 delta-validation : First results for temperature, ozone and altitude profiles*, MIPAS QWG meeting #36, WebEx conference, 28 July 2014.

Keppens, A., D. Hubert, J. Granville and J.-C. Lambert, *NDACC-based Delta Validation of Envisat MIPAS Data Products: CH4, HNO3, N2O*, MIPAS QWG meeting #36, WebEx conference, 28 July 2014.

Hubert, D., A. Keppens, J. Granville and J.-C. Lambert, *ML2PP V7 delta-validation: Overview of validation results*, MIPAS QWG meeting #37, ESRIN, Frascati, Italy, 7-9 October 2014.

Hubert, D., J. Granville, A. Keppens, F. Hendrick and J.-C. Lambert, *Definition of SCIAMACHY Diagnostic Data Set*, SCIAMACHY QWG meeting PM1, IUP, Bremen, Germany, 20 October 2014.

Hubert, D., J.-C. Lambert, A. Merlaud and M. Van Roozendael, *SCIAMACHY Validation – SCIAVALIG Report to SSAG #45*, 45th SCIAMACHY Science Advisory Group meeting, IUP, Bremen, Germany, 21 October 2014.

Hubert, D., A. Keppens, J. Granville and J.-C. Lambert, *SCIAMACHY Level-2 V6 : First look at Diagnostic Data Set*, SCIAMACHY QWG meeting PM3 (via WebEx), DLR, Oberpfaffenhofen, Germany, 5-6 May 2015.

Hubert, D., A. Keppens, J. Granville, F. Hendrick, J.-C. Lambert and A. van Gijsel, *Upgrade from SGP V5.02 to V6 .00 : Conclusions from delta-validation of Diagnostic Data Set*, SCIAMACHY QWG meeting PM4, WebEx meeting, 22 September 2015.

Hubert, D., J. Granville and J.-C. Lambert, *Validation of GOMOS O3 profile data within Ozone_cci : Bright Limb V1.2 and IPF 6.01*, GOMOS QWG meeting N4, ACRI, Sophia Antipolis, France, 23-24 September 2015.

Hubert, D., J. Granville, and J.-C. Lambert, *First validation of ML2PP V7 full mission : Temperature and altitude data*, MIPAS QWG meeting #40, IFAC, Firenze, Italy, 2-4 November 2015.

Oral presentations at international conferences, symposia and workshops

Hubert, D., T. Verhoelst, A. Keppens, J. Granville, D. Pieroux and J.-C. Lambert, *Towards a merged Essential Climate Variable data record on ozone: stability and consistency of contributing limb profilers*, ESA Atmospheric Science Conference, ATMOS 2012, Bruges, Belgium, 18-22 June 2012.

Hubert, D., T. Verhoelst, A. Keppens, J. Granville, J.-C. Lambert, R. Stübi, P. Nair, S. Godin-Beekmann, NDACC lidar PIs, Ozone Sonde Working Group and Satellite Science and Processing Teams, *Network-based evaluation of fourteen satellite limb/occultation profilers for the next SPARC and WMO ozone trend assessments*, Quadrennial Ozone Symposium (QOS2012), Toronto, Canada, 27-31 August 2012.

Keppens, A., D. Hubert, T. Verhoelst, J. Granville, J.-C. Lambert, A. Delcloo, R. van der A, J. van Peet, R. Siddans, G. Miles, R. Kivi, R. Stübi and S. Godin-Beekmann, *Round-Robin Validation of Nadir Ozone Profile Retrievals Using Ground-Based Reference Data: Methodology and Practice*, ESA Atmospheric Composition Validation and Evolution workshop (ACVE), ESRIN, Frascati, Italy, 13-15 March 2013.

Hubert, D., T. Verhoelst, A. Keppens, J. Granville, J.-C. Lambert, M. Allaart, T. Deshler, B. Johnson, R. Kivi, F. Schmidlin, H. Smit, W. Steinbrecht, R. Stübi, D. Tarasick, A. Thompson, M. Tully, R. Van Malderen and P. von der Gathen, *Assessment of the internal consistency of the NDACC ozonesonde network by comparison with the satellite system of ozone*, ESA Atmospheric Composition Validation and Evolution workshop (ACVE), ESRIN, Frascati, Italy, 13-15 March 2013.

Keppens, A., D. Hubert, J. Granville, J.-C. Lambert, B. Dils, M. De Mazière, C. Vigouroux, and (Multi-)TASTE GHG Team, *(Multi-)TASTE: NDACC-based Validation Of Envisat Greenhouse Gas Data And Their Evolution,* 10th International Workshop on Greenhouse Gas Measurements from Space, ESA-ESTEC, Noordwijk, The Netherlands, 5-7 May 2014.

Hubert, D., T. Verhoelst, A. Keppens, J. Granville, J.-C. Lambert, M. A. F. Allaart, T. Deshler, B. J. Johnson, R. Kivi, F. J. Schmidlin, H. G. J. Smit, W. Steinbrecht, R. Stübi, D. W. Tarasick, A. M. Thompson, M. B. Tully, R. Van Malderen, and P. von der Gathen, *The internal consistency of the ozonesonde network data archive assessed through comparisons with satellite ozone profilers*, 2014 EUMETSAT Meteorological Satellite Conference, Geneva, Switzerland, 22-26 September 2014.

Hubert, D., A. Keppens, J. Granville, and J.-C. Lambert, *NDACC-based validation of Envisat greenhouse gas products and their evolution*, 2014 EUMETSAT Meteorological Satellite Conference, Geneva, Switzerland, 22-26 September 2014.

Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A. Keppens, U. Cortesi, D.A. Degenstein, L. Froidevaux, S. Godin-Beekmann, K.W. Hoppel, E. Kyrölä, T. Leblanc, G. Lichtenberg, I.S. McDermid, C.T. McElroy, H. Nakane, J.M. Russell III, H.G.J. Smit, K. Stebel, W. Steinbrecht, R. Stübi, D.P.J. Swart, G. Taha, A.M. Thompson, J. Urban, J.A.E. van Gijsel, P. von der Gathen, K.A. Walker and J.M. Zawodny, *Ground-based network assessment of the long-term stability and mutual consistency of limb/occultation ozone profile decadal data records*, NORS/NDACC/GAW Workshop, Brussels, Belgium, 5-7 November 2014.

Hubert, D., T. Verhoelst, A. Keppens, J. Granville, and J.-C. Lambert, *O3S-DQA – Internal consistency of the ozonesonde network in the middle stratosphere using satellite data ensembles as reference*, NORS/NDACC/GAW Workshop, Brussels, Belgium, 5-7 November 2014.

Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A., Keppens, U. Cortesi, D.A. Degenstein, L. Froidevaux, S. Godin-Beekmann, K.W. Hoppel, E. Kyrölä, T. Leblanc, G. Lichtenberg, I.S. McDermid, C.T. McElroy, D. Murtagh, H. Nakane, J.R. Russell III, H.G.J. Smit, K. Stebel, W. Steinbrecht, R. Stübi, D.P.J. Swart, G. Taha, A.M. Thompson, J. Urban, A. van Gijsel, P. von der Gathen, K.A. Walker, and J.M. Zawodny, *Uncertainties in recent satellite ozone profile trend assessments (SI2N, WMO 2014) : A network-based assessment of fourteen contributing limb and occultation data records*, Advances in Atmospheric Science and Applications (ATMOS 2015), Heraklion, Greece, 8-12 June 2015.

Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A. Keppens, J.-L. Baray, U. Cortesi, D. A. Degenstein, L. Froidevaux, S. Godin-Beekmann, K. W. Hoppel, E. Kyrölä, T. Leblanc, G. Lichtenberg, C. T. McElroy, D. Murtagh, H. Nakane, J. M. Russell III, J. Salvador, H. G. J. Smit, K. Stebel, W. Steinbrecht, K. B. Strawbridge, R. Stübi, D. P. J. Swart, G. Taha, A. M. Thompson, J. Urban, J. A. E. van Gijsel, P. von der Gathen, K. A. Walker, E. Wolfram, and J. M. Zawodny, *Uncertainties in recent satellite ozone profile trend assessments : A ground network-based assessment of fourteen limb and occultation data records*, 8th Atmospheric Limb Workshop, Gothenburg, Sweden, 15-17 September 2015.

Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A. Keppens, J.-L. Baray, U. Cortesi, D. A. Degenstein, L. Froidevaux, S. Godin-Beekmann, K. W. Hoppel, E. Kyrölä, T. Leblanc, G. Lichtenberg, C. T. McElroy, D. Murtagh, H. Nakane, J. M. Russell III, J. Salvador, H. G. J. Smit, K. Stebel, W. Steinbrecht, K. B. Strawbridge, R. Stübi, D. P. J. Swart, G. Taha, A. M. Thompson, J. Urban, J. A. E. van Gijsel, P. von der Gathen, K. A. Walker, E. Wolfram, and J. M. Zawodny, *How certain are we of the uncertainties in recent ozone profile trend assessments of merged limb/occultation records*, Invited talk at AGU Fall Meeting 2015, San Francisco, California, USA, 14-18 December 2015.

Conference proceedings

Hubert, D., T. Verhoelst, A. Keppens, J. Granville, D. Pieroux and J.-C. Lambert, *Towards a merged Essential Climate Variable data record on ozone: stability and consistency of contributing limb profilers*, in Proc. Advances in Atmospheric Science and Applications (ATMOS 2012), Bruges, Belgium, 18-22 June 2012, ESA Special Publication SP-708, November 2012.

Hubert, D., A. Keppens, J.-C. Lambert, J. Granville, F. Hendrick, and T. Verhoelst, *The Multi-TASTE validation system: Tasting the evolution of reactive and greenhouse gas data products from Envisat and Third Party Missions*, Advances in Atmospheric Science and Applications (ATMOS 2015), Heraklion, Greece, 8-12 June 2015, ESA Special Publication SP-735, November 2015.

Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A. Keppens, U. Cortesi, D.A. Degenstein, L. Froidevaux, S. Godin-Beekmann, K.W. Hoppel, E. Kyrölä, T. Leblanc, G. Lichtenberg, I.S. McDermid, C.T. McElroy, D. Murtagh, H. Nakane, J.R. Russell III, H.G.J. Smit, K. Stebel, W. Steinbrecht, R. Stübi, D.P.J. Swart, G. Taha, A.M. Thompson, J. Urban, A. van Gijsel, P. von der Gathen, K.A. Walker, and J.M. Zawodny, *Uncertainties in recent satellite ozone profile trend assessments (SI2N, WMO 2014) : A network-based assessment of fourteen contributing limb and occultation data records*, Advances in Atmospheric Science and Applications (ATMOS 2015), Heraklion, Greece, 8-12 June 2015, ESA Special Publication SP-735, November 2015.

Posters

Hubert, D., T. Verhoelst, A. Keppens and J.-C. Lambert, *Evaluation of the mutual consistency and long-term stability of eleven ozone profilers using ground-based networks*, 2nd SPARC/IO_3C/IGACO-O3/NDACC workshop on Past Changes in the Vertical Distribution of Ozone (SI^2N initiative), Columbia, Maryland, USA, 16-18 April 2012.

Hubert, D., T. Verhoelst, S. Vandenbussche, J. Granville, D. Pieroux and J.-C. Lambert, *Detectability of past changes in the vertical distribution of ozone with twelve limb sounders*, EGU General Assembly, Vienna, Austria, 22-27 April 2012.

Hubert, D., A. Keppens, T. Verhoelst, J. Granville, J.-C. Lambert, U. Cortesi, D. Degenstein, E. Kyrölä, G. Lichtenberg, T. McElroy, J. Urban, K. Walker, R. Stübi, J.-L. Baray, S. Godin-Beekmann, T. Leblanc, H. Nakane, K. Stebel, W. Steinbrecht, K. Strawbridge, D. Swart, A. Van Gijsel and P. von der Gathen, *Latest updates of ozone profile data from Odin, Envisat and ACE: bias and stability with respect to NDACC/GAW ozonesondes and lidars*, ESA Atmospheric Composition Validation and Evolution workshop (ACVE), ESRIN, Frascati, Italy, 13-15 March 2013.

Verhoelst, T., D. Hubert, A. Keppens, J.-C. Lambert, A. Laeng, T. von Clarmann, G. Stiller, A. Dudhia, B. M. Dinelli, P. Raspollini, K. Walker, P. Bernath, and L. Froidevaux, *A uniform validation of four MIPAS ozone profile retrievals*, ESA Atmospheric Composition Validation and Evolution workshop (ACVE), ESRIN, Frascati, Italy, 13-15 March 2013.

Lambert, J.-C., D. Hubert, J. Granville, A. Keppens, O. Aulamo, R. Kivi and E. Kyrö, T. Blumenstock , F. Hase, G. Kopp and S. Mikuteit, J.P. Burrows, J. Notholt, A. Richter, T. Warneke and F. Wittrock, H. De Backer, C. De Clercq, M. De Mazière, B. Dils, F. Hendrick, S. Vandenbussche, M. Van Roozendael, T. Verhoelst and C. Vigouroux, M. Gil, M. Navarro Comas, O. Puentedura and M. Yela Gonzalez, F. Goutail, A. Pazmiño and J.-P. Pommereau, G. Held, D.V. Ionov and Yu. M. Timofeyev, P.V. Johnston, K. Kreher and S. Wood, E. Mahieu and P. Demoulin, A. Piters, and R. Susmann, *(Multi-)TASTE: Ten Years of Quality Tasting of Envisat Atmospheric Composition Retrievals With the NDACC Network*, ESA Living Planet Symposium, Edinburgh, UK, 9-13 September 2013.

Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A. Keppens, U. Cortesi, D.A. Degenstein, L. Froidevaux, S. Godin-Beekmann, K.W. Hoppel, E. Kyrölä, T. Leblanc, G. Lichtenberg, I.S. McDermid, C.T. McElroy, H. Nakane, J.M. Russell III, H.G.J. Smit, K. Stebel, W. Steinbrecht, R. Stübi, D. Swart, G. Taha, A.M. Thompson, J. Urban, A. van Gijsel, P. von der Gathen, K.A. Walker and J.M. Zawodny, *Ground-based assessment of the bias and long-term stability of fourteen limb and occultation ozone profile data records*, Final SPARC/IO_3C/IGACO-O3/NDACC workshop on Past Changes in the Vertical Distribution of Ozone (SI^2N initiative), FMI, Helsinki, Finland, 18-19 September 2013.

Hubert, D., J.-C. Lambert, T. Verhoelst, J. Granville, A. Keppens, and the Ozonesonde PIs, Lidar PIs, and the Satellite Science and Processing Team, *SI^2N assessment of vertical ozone trends: Stability of limb/occultation data records over 1984-2013 against ground-based networks*, European Geosciences Union, General Assembly 2014, Vienna, Austria, 27 April – 2 May 2014.

Keppens, A., D. Hubert, J.-C. Lambert, M. De Mazière, J. Granville, M. Van Roozendael, and the CINAMON/FTIRval/TASTE/Multi-TASTE Team, *Ten years of NDACC-based support to the maturation of Envisat and TPM atmospheric composition data products*, NORS/NDACC/GAW Workshop, Brussels, Belgium, 5-7 November 2014.

Hubert, D., A. Keppens, J.-C. Lambert, J. Granville, F. Hendrick, and T. Verhoelst, *The Multi-TASTE validation system: Tasting the evolution of reactive and greenhouse gas data products from Envisat and Third Party Missions*, Advances in Atmospheric Science and Applications (ATMOS 2015), Heraklion, Greece, 8-12 June 2015.

Hubert, D., A. Keppens, J.-C. Lambert, J. Granville, F. Hendrick and T. Verhoelst, *The Multi-TASTE satellite validation system: Twenty years of support to the development of reactive and greenhouse gas data products from multiple satellites*, 2015 EUMETSAT Meteorological Satellite Conference, Toulouse, France, 21-25 September 2015.

Annexe C : Participation to project meetings and international conferences, symposia & workshops

Envisat Quality Working Group meetings

- 44th SCIAMACHY Scientific Advisory Committee, KNMI, De Bilt, The Netherlands, 23-24 May, 2012.
- GOMOS QWG meeting #27, LATMOS, Guyancourt, France, 26-28 September 2012.
- MIPAS QWG meeting #30, ESRIN, Frascati, Italy, 12-14 November 2012.
- SCIAVALIG workshop, KNMI, De Bilt, The Netherlands, 29-30 November 2012.
- MIPAS QWG meeting #31 (via WebEx), ESRIN, Frascati, Italy, 18-19 March 2013.
- SCIAMACHY QWG meeting #21, teleconference, 24 April 2013.
- SCIAMACHY QWG meeting #22, IUP, Bremen, Germany, 13-14 June 2013.
- MIPAS QWG meeting #32 (via WebEx), U Oxford, Oxford, United Kingdom, 2-3 July 2013.
- MIPAS QWG meeting #33, ISAC, Bologna, Italy, 4-6 November 2013.
- MIPAS QWG meeting #34, WebEx conference, 25 February 2014.
- MIPAS QWG meeting #35 (via WebEx), ESRIN, Frascati, Italy, 24-26 March 2014.
- WebEx teleconference with ESA on MIPAS ML2PP V7 validation roadmap, 28 March 2014.
- MIPAS QWG meeting #37 (via WebEx), ESRIN, Frascati, Italy, 7-9 October 2014.
- SCIAMACHY QWG meeting PM1, IUP, Bremen, Germany, 20 October 2014.
- 45th SCIAMACHY Science Advisory Group meeting, IUP, Bremen, Germany, 21 October 2014.
- SCIAMACHY Status teleconference, WebEx meeting, 20 November 2014.
- GOMOS QWG meeting #N2 (via WebEx), ESRIN, Frascati, Italy, 26-27 November 2014.
- Multi-TASTE Phase-F CCN2 Kick-Off, Webex meeting, 19 December 2014.
- SCIAMACHY QWG meeting PM2, WebEx meeting, 28 January 2015.
- MIPAS QWG meeting #38 (via WebEx), ESRIN, Frascati, Italy, 18-19 February 2015.
- MIPAS Status teleconference, WebEx meeting, 30 March 2015.
- SCIAMACHY QWG meeting PM3, DLR, Oberpfaffenhofen, Germany, 5-6 May 2015.
- SCIAMACHY QWG meeting PM4, WebEx meeting, 22 September 2015.
- GOMOS QWG meeting N4 (via WebEx), ACRI, Sophia Antipolis, France, 23-24 September 2015.
- MIPAS QWG meeting #40, IFAC, Firenze, Italy, 2-4 November 2015.
- GOMOS QWG meeting N5 (via WebEx), 20 November 2015.
- SCIAMACHY QWG meeting PM5, SRON, Utrecht, The Netherlands, 24-24 November 2015.

International conferences, symposia and workshops

- Second SPARC/IO_3C/IGACO-O3/NDACC workshop on Past Changes in the Vertical Distribution of Ozone (SI^2N initiative), Columbia, Maryland, USA, 16-18 April 2012.
- EGU General Assembly, Vienna, Austria, 22-27 April 2012.
- ESA Atmospheric Science Conference, ATMOS 2012, Bruges, Belgium, 18-22 June 2012.
- Quadrennial Ozone Symposium, Toronto, Canada, August 27-31, 2012.
- NDACC Steering Committee meeting 2012, Garmisch-Partenkirchen, Germany, 15-18 October 2012.
- Atmospheric Composition Validation and Evolution workshop (ACVE), ESRIN, Frascati, Italy, 13-15 March 2013.
- 7th Atmospheric Limb Conference, Bremen, Germany, 17-19 June 2013.

- ESA Living Planet Symposium, Edinburgh, UK, 9-13 September 2013.
- Final SPARC/IO_3C/IGACO-O3/NDACC workshop on Past Changes in the Vertical Distribution of Ozone (SI^2N initiative), FMI, Helsinki, Finland, 18-19 September 2013.
- NDACC Steering Committee meeting 2013, Frascati, Italy, 1-3 October 2013.
- European Geosciences Union, General Assembly 2014, Vienna, Austria, 27 April – 2 May 2014.
- 10th International Workshop on Greenhouse Gas Measurements from Space (IWGGMS-10), ESTEC, Noordwijk, The Netherlands, 5-7 May 2014.
- 2014 EUMETSAT Meteorological Satellite Conference, Geneva, Switzerland, 22-26 September 2014.
- NDACC Steering Committee meeting 2014, Brussels, Belgium, 3-5 November 2014.
- NDACC/NORS/GAW Workshop, Brussels, Belgium, 5-7 November 2014.
- Advances in Atmospheric Science and Applications (ATMOS 2015), Heraklion, Greece, 8-12 June 2015.
- 8th Atmospheric Limb Workshop, Gothenburg, Sweden, 15-17 September 2015.
- NDACC Steering Committee meeting 2015, La Jolla, California, USA, 12-15 October 2015.
- AGU Fall Meeting 2015, San Francisco, California, USA, 14-18 December 2015.