TITLE:

SPIN

Product Validation Report final version

Reference: SPIN_PVR_final

Date of issue: 16/04/2014

Distributed to:

This work is supported by the European Space Agency

DOCUMENT PROPERTIES

Title	Product Validation Report final
Reference	SPIN_PVR_final
Issue	05
Revision	00
Status	Final
Date of issue	16/04/2014
Document type	Deliverable

	FUNCTION	NAME	DATE
LEAD AUTHOR	WP 13 science team	Katja Weigel (UB)	
CONTRIBUTING AUTHORS	WP12 manager WP12 science team WP12 science team WP13 manager WP14 manager WP16 manager WP17 manager WP18 manager WP22 manager WP22 science team WP23 manager WP23 science team WP24 manager WP26 manager	Doug Degenstein (US) Landon Rieger (US) Lena Anneke Brinkhoff (UB) Alexei Rozanov (UB) Greg Bodeker (BS) Jo Urban (CUT) Simo Tukiainen (FMI) Andreas Jonsson (UT) Greg Bodeker (BS) Erkki Kyrölä (FMI) Ted Shepherd Charles McLandress (UT) Greg Bodeker (BS) Patricia Liebing (UB)	
REVIEWED BY			
APPROVED BY			
ISSUED BY			

DOCUMENT CHANGE RECORD

Issue	Revision	Date	Modified items	Observations
1	0	19.02.2013	Assembled inputs	
		12.03.2013	Introduction and format-ting	
		19.03.2013	Suggestion for summary and corrections for	

			SCIAMACHY aerosol and water vapour added	
		25.03.2013	Updates of results (Stratospheric Temperature) and Summary	
		26.03.2013	Updates for Stratospheric Temperature	
		19.08.2013	New contributions for WP22 and WP24. Updates and corrections for the whole text.	
		12.09.2013	Updates for WP16	
		11.02.2014	Updates for WP17, WP22, adding WP26	
		12.02.2014	Updates for WP17, WP22, and WP24	
		17.02.2014	Updates for WP16, some corrections and formatting for the whole document	
		20.02.2014	Update for summary (WP22)	
		25.02.2014	Update for WP24	
		18.03.2014	Update for WP18, WP23 added	
		28.03.2014	Updates for WP18 and WP23; several small changes	
		16.04.2014	Updates for WP16	

Table of Contents

Executive summary

The purpose of this document is to detail the validation performed on the data sets developed, matured, and used within SPIN. In the first phase of SPIN these data sets are: climatologies of zonal and monthly mean stratospheric temperatures from ACE-FTS, MIPAS and SMR, stratospheric water vapour from SCIAMACHY, stratospheric aerosols from OSIRIS and SCIAMACHY, stratospheric ozone from GOMOS bright limb measurements, and climatologies of short-lived species from SMR and OSIRIS. In the second phase of SPIN the data sets developed, and their validation detailed in this document, include: two different versions of a merged SAGE II+GOMOS ozone data set, an extended upper stratospheric temperature record, and an improved UT/LS temperature record.

Applicable and reference documents

Anthes, R.A.; Bernhardt, P.A.; Chen, Y.; Cucurull, L.; Dymond, K.F.; Ector, D.; Healy, S.B.; Ho, S.-P.; Hunt, D.C.; Kuo, Y.-H.; Liu, H.; Manning, K.; McCormick, C.; Meehan, T.K.; Randel, W.J.; Rocken, C.; Schreiner, W.S.; Sokolovskiy, S.V.; Syndergaard, S.; Thompson, D.C.; Trenberth, K.E.; Wee, T.-K.; Yen, N.L. and Zeng, Z., The COS-MIC/FORMOSAT-3 Mission: early results, Bull. Amer. Meteor. Soc., 89, 313-333, 2008.

Beyerle, G., T. Schmidt, G. Michalak, S. Heise, J. Wickert, and C. Reigber, GPS radio occultation with GRACE: Atmospheric profiling utilizing the zero difference technique, Geophys. Res. Lett., 32, L13806, doi:10.1029/2005GL023109, 2005.

Beyerle, G., Grunwaldt, L., Heise, S., Köhler, W., König, R., Michalak, G., Rothacher, M., Schmidt, T., Wickert, J., Tapley, B. D., and Giesinger, B., First results from the GPS atmosphere sounding experiment TOR aboard the TerraSAR-X satellite, Atmos. Chem. Phys., 11, 6687–6699, 2011.

Bodeker, G.E.; Hassler, B.; Young, P.J. and Portmann, R.W., A vertically resolved, global, gap-free ozone database for assessing or constraining global climate model simulations, Earth System Science Data, 5, 31-43, 2013.

Borsche, M., Kirchengast, G., and Foelsche, U., Tropical tropopause climatology as observed with radio occultation measurements from CHAMP compared to ECMWF and NCEP analyses, Geophys. Res. Lett., 34, L03702, 2007.

Bourassa, A. E., Rieger, L. A., Lloyd, N. D., andDegenstein, D. A., Odin-OSIRIS stratospheric aerosol data product and SAGE III intercomparison, Atmos. Chem. Phys. Discuss., 11, 25,785-25,811, doi:10.5194/acpd-11-25785-2011, 2011.

Bourassa, A. E., Bodeker, G., Brinkhoff, L. A., Kyrola, E., Liebing, P., Rieger, L., Urban, J., Weigel, K., Algorithm Theoretical Baseline Document, SPIN-UB-ATBD-1.01, ESA SPARC Initiative – SPIN, 2012.

Brohede, S., Haley, C., McLinden, C., Sioris, C., Murtagh, D., Petelina, S., Llewellyn, E., Bazureau, A., Goutail, F., Randall, C., Lumpe, J., Taha, G., Thomasson, L., and Gordley, L., Validation of Odin/OSIRIS stratospheric NO2 profiles, Journal of Geophysical Research 112(D07310), DOI: 10.1029/2006JD007586, 2007.

Brohede, S., McLinden, C., Berthet, G., Haley, C., Murtagh, D., and Sioris, C., A stratospheric NO2 climatology from Odin/OSIRIS limb-scatter measurements, Canadian Journal of Physics 85, 11, 1253-1274, DOI: 10.1139/P07-141, 2007.

Brohede, S., McLinden, C. A., Urban, J., Haley, C. S., Jonsson, A. I., and Murtagh, D., Odin stratospheric proxy NOy measurements and climatology, Atmos. Chem. Phys., 8, 5731-5754, 2008.

Dupuy, E., Walker, K. A., Kar, J., Boone, C. D., McElroy, C. T., Bernath, P. F., Drummond, J. R., Skelton, R., McLeod, S. D., Hughes, R. C., Nowlan, C. R., Dufour, D. G., Zou, J., Nichitiu, F., Strong,

K., Baron, P., Bevilacqua, R. M., Blumenstock, T., Bodeker, G. E., Borsdorff, T., Bourassa, A. E., Bovensmann, H., Boyd, I. S., Bracher, A., Brogniez, C., Burrows, J. P., Catoire, V., Ceccherini, S., Chabrillat, S., Christensen, T., Coffey, M. T., Cortesi, U., Davies, J., De Clercq, C., Degenstein, D. A., De Mazière, M., Demoulin, P., Dodion, J., Firanski, B., Fischer, H., Forbes, G., Froidevaux, L., Fussen, D., Gerard, P., Godin-Beekmann, S., Goutail, F., Granville, J., Griffith, D., Haley, C. S., Hannigan, J. W., Höpfner, M., Jin, J. J., Jones, A., Jones, N. B., Jucks, K., Kagawa, A., Kasai, Y., Kerzenmacher, T. E., Kleinböhl, A., Klekociuk, A. R., Kramer, I., Küllmann, H., Kuttippurath, J., Kyrölä}, E., Lambert, J.-C., Livesey, N. J., Llewellyn, E. J., Lloyd, N. D., Mahieu, E., Manney, G. L., Marshall, B. T., McConnell, J. C., McCormick, M. P., McDermid, I. S., McHugh, M., McLinden, C. A., Mellqvist, J., Mizutani, K., Murayama, Y., Murtagh, D. P., Oelhaf, H., Parrish, A., Petelina, S. V., Piccolo, C., Pommereau, J.-P., Randall, C. E., Robert, C., Roth, C., Schneider, M., Senten, C., Steck, T., Strandberg, A., Strawbridge, K. B., Sussmann, R., Swart, D. P. J., Tarasick, D. W., Taylor, J. R., Tétard, C., Thomason, L. W., Thompson, A. M., Tully, M. B., Urban, J., Vanhellemont, F., Vigouroux, C., von Clarmann, T., von der Gathen, P., von Savigny, C., Waters, J. W., Witte, J. C., Wolff, M., and Zawodny, J. M., Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), Atmos. Chem. Phys., 9, 287--343, doi:10.5194/acp-9-287-2009, 2009.

Eichmann, K.-U., von Savigny, C., Reichl, P., Robert, C., Steinwagner, J., Bovensmann, H., and Burrows, J. P., SCODA: SCIAMACHY Cloud Detection Algorithm from limb radiance measurements, Algorithm Theoretical Baseline Document (ATBD), University of Bremen, Germany, 2009.

Ernst, F., von Savigny, C., Rozanov, A., Rozanov, V., Eichmann, K.-U., Brinkhoff, L. A., Bovensmann, H., and Burrows, J. P., Global stratospheric aerosol extinction profile retrievals from SCIAMACHY limb-scatter observations, Atmos. Meas. Tech. Discussion, 2012.

Free, M.; Seidel, D.J.; Angell, J.K.; Lanzante, J.R.; Durre, I. and Peterson, T.C., Radiosonde Atmospheric Temperature Products for Assessing Climate (RATPAC): A new data set of large-area anomaly time series, J. Geophys. Res., 110, D22101, doi:22110.21029/22005JD006169, 2005.

Fujiwara, M., Voemel, H., Hasebe, F., Shiotani, M., Ogino, S.- Y., Iwasaki, S., Nishi, N., Shibata, T., Shimizu, K., Nishimoto, E., Valverde-Canossa, J. M., Selkirk, H. B., and Oltmans, S. J., Seasonal to decadal variations of water vapor in the tropical lower stratosphere observed with balloon-borne cryogenic frost point hygrometers, J. Geophys. Res., 115, D18304, doi:10.1029/2010JD014179, 2010.

Haley, C., Brohede, S., Sioris, C., Griffioen, E., Murtagh, D., McDade, I., Eriksson, P., Llewellyn, E., Bazureau, A. & Goutail, F., Retrievals of stratospheric O3 and NO2 profiles from Odin Optical Spectrograph and InfraRed Imager System (OSIRIS) limb-scattered sunlight measurements. Journal of Geophysical Research - Atmospheres 109(D16), DOI: 10.1029/2004JD004588, 2004.

Hegglin, M. I., Boone, C. D., Manney, G. L., Shepherd, T. G., Walker, K. A., Bernath, P. F., Daffer, W. H., Hoor, P., and Schiller, C., Validation of ACE-FTS satellite data in the upper troposphere/lower stratosphere (UTLS) using non-coincident measurements, Atmos. Chem. Phys., 8, 1483-1499, doi:10.5194/acp-8-1483-2008, 2008.

Jones, A., Urban, J., Murtagh, D. P., Sanchez, C., Walker, K. A., Livesey, N., Froidevaux, L., Santee, M.L., Analysis of HCl and ClO time series in the upper stratosphere using satellite data sets, Atmos. Chem. Phys. 11, 5321-5333, doi:10.5194/acp-11-5321-2011, 2011.

Khosravi, M., Baron, P., Urban, J., Froidevaux, L., Jonsson, A. I., Kasai, Y., Kuribayashi, K., Mitsuda, C., Murtagh, D. P., Sagawa, H., Santee, M. L., Sato, T. O., Shiotani, M., Suzuki, M., von Clarmann, T., Walker, K. A., and Wang, S.: Diurnal variation of stratospheric and lower mesospheric HOCl, ClO and HO2 at the equator: comparison of 1-D model calculations with measurements by satellite instruments, Atmos. Chem. Phys., 13, 7587-7606, doi:10.5194/acp-13-7587-2013, 2013.

Kyrölä, E., Laine, M., Sofieva, V., Tamminen, J., Päivärinta, S.-M., Tukiainen, S., Zawodny, J., and Thomason, L., Combined SAGE II - GOMOS ozone profile data set for 1984–2011 and trend analysis of the vertical distribution of ozone, Atmospheric Chemistry and Physics, 13(21):10645–10658, 2013.

Lambert, A., Read, W., Livesey, N., Santee, M., Manney, G., Froidevaux, L., Wu, D., Schwartz, M., Pumphrey, H., Jimenez, C., Nedoluha, G., Coeld, R., Cuddy, D., Daffer, W., Drouin, B., Fuller, R., Jarnot, R., Knosp, B., Pickett, H., Perun, V., Snyder, W., Stek, P., Thurstans, R., Wagner, P., Waters, J., Jucks, K., Toon, G., Stachnik, R., Bernath, P., Boone, C., Walker, K., Urban, J., Murtagh, D., Elkins, J. & Atlas, E., Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements. Journal of Geophysical Research - Atmospheres 112 (D24), DOI: 10.1029/2007JD008724, 2007.

Livesey, N., MLS version 3.3 and 3.4 Level 2 data quality and description document, https://mls.jpl.nasa.gov/data/, 2013.

McCormick, M. P., SAGE II: An Overview, Adv. Space Res., 7(3), 219-226, 1987.

McLinden, C. A., Haley, C. S., Lloyd, N. D., Hendrick, F., Rozanov, A., Sinnhuber, B.-M., Goutail, F., Degenstein, D.A., Llewellyn, E. J., Sioris, C. E., Van Roozendael, M., Pommereau, J. P., Lotz, W., and. Burrows, J. P, Odin/OSIRIS observations of stratospheric BrO: Retrieval methodology, climatology, and inferred Bry, J. Geophys. Res., 115, D15308, 2010, doi:10.1029/2009JD012488, 2009.

Müller, R., Grooß, J.-U., Lemmen, C., Heinze, D., Dameris, M., and Bodeker, G.E., Simple measures of ozone depletion in the polar stratosphere, Atmos. Chem. Phys., 8, 251-264, 2008.

Read, W. G., Lambert, A., Bacmeister, J., Cofield, R. E., Christensen, L. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fetzer, E., Froidevaux, L., Fuller, R., Herman, R., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Kelly, K., Knosp, B. W., Kovalenko, L. J., Livesey, N. J., Liu, H.-C., Manney, G. L., Pickett, H. M., Pumphrey, H. C., Rosenlof, K. H., Sabounchi, X., Santee, M. L., Schwartz, M. J., Snyder, W. V., Stek, P. C., Su, H., Takacs, L. L., Thurstans, R. P., Vomel, H., Wagner, P. A., Waters, J. W., Webster, C. R., Weinstock, E. M., and Wu, D. L., Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H2O and relative humidity with respect to ice validation, J. Geophys. Res., 112, D24S35, doi:10.1029/2007JD008752, 2007.

Rozanov, A., Weigel, K., Bovensmann, H., Dhomse, S., Eichmann, K.-U., Kivi, R., Rozanov, V., Vömel, H., Weber, M., and Burrows, J. P., Retrieval of water vapor vertical distributions in the upper troposphere and the lower stratosphere from SCIAMACHY limb measurements, Atmos. Meas. Techn., 4, 933-954, doi:10.5194/amt-4-933-2011, 2011.

Russell, P. B., and M. P. McCormick, SAGE II aerosol data validation and initial data use - An introduction and overview, J. Geophys. Res., 94, 8335-8338, doi:10.1029/JD094iD06p08335, 1989.

Sagawa, H., Sato, T., Baron, P., Dupuy, E., Livesey, N., Urban, J., von Clarmann, T., de Lange, A., Wetzel, G., Connor, B., Kagawa, A., Murtagh, D. & Kasai, Y., Comparison of SMILES ClO profiles with satellite, balloon-borne and ground-based measurements. Atmos. Meas. Tech. 6(12), 3325-3347. DOI: 10.5194/amt-6-3325-2013, 2013.

Schmidt, T., Wickert, J., Marquardt, C., Beyerle, G., Reigber, C., Galas, R., and König, R., GPS radio occultation with CHAMP: an innovative remote sensing method of the atmosphere, Adv. Space Res., 33, 1036–1040, 2004.

Selkirk, H. B., Vömel, H., Valverde Canossa, J. M., Pfister, L., Diaz, J. A., Fernández, W., Amador, J., Stolz, W., and Peng, G., Detailed structure of the tropical upper troposphere and lower stratosphere as revealed by balloon sonde observations of water vapor, ozone, temperature and winds during the NASA TCSP and TC4 campaigns, J. Geophys. Res., 115, D00J19, doi:10.1029/2009JD013209, 2010.

Sofieva, V. F., Rahpoe, N., Tamminen, J., Kyrölä, E., Kalakoski, N., Weber, M., Laeng, A., von Clarmann, T., Stiller, G., Lossow, S., Degenstein, D., Bourassa, A., Adams, C., Roth, C., Lloyd, N., Bernath, P., Hargreaves, R. J., Urban, J., Murtagh, D., Hauchecorne, A., Van Roozendael, M., Kalb, N., and Zehner, C., Harmonized dataset of ozone profiles from satellite limb and occultation measurements. Earth System Science Data, 5:349–363, 2013.

Thomason, L. W., and Poole, L. R., Use of Stratospheric Aerosol Properties as Diagnostics of Antarctic Vortex Processes, J. Geophys. Res., 98, 23,003/23,012, doi:10.1029/93JD02461, 1992.

Thomason, L. W., Poole, L. R., and Deshler, T., A global climatology of stratospheric aerosol surface area density deduced from Stratospheric Aerosol and Gas Experiment II measurements: 1984-1994, J. Geophys. Res., 102, 8967-8976, doi:10.1029/96JD02962, 1997.

Thomason, L. W., and Taha, G., SAGE III aerosol extinction measurements: Initial results, Geophys. Res. Lett., 30, 1631, doi:10.1029/2003GL017317, 2003.

Tegtmeier, S., Hegglin, M. I., Anderson, J., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux, L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Kyrölä, E., Lingenfelser, G., Lumpe, J., Nardi, B., Neu, J., Pendlebury, D., Remsberg, E., Rozanov, A., Smith, L., Toohey, M., Urban, J., von Clarmann, T., Walker, K. A., and Wang, R., A comparison of ozone climatologies from international limb satellite sounders, ournal of Geophysical Research (Atmospheres), 118:12229, November 2013.

Toohey, M., Hegglin, M. I., Tegtmeier, S., Anderson, J., Anel, J. A., Bourassa, A., Brohede, S., Degenstein, D., Froidevaux. L., Fuller, R., Funke, B., Gille, J., Jones, A., Kasai, Y., Kruger, K., Kyrölä, E., Neu, J. L., Rozanov, A., Smith, L., Urban, J., von Clarmann, T., Walker, K. A., and Wang, R. H. J., Characterizing sampling biases in the trace gas climatologies of the SPARC Data Initiative. Journal of Geophysical Research (Atmospheres), 118:11847, October 2013.

van Gijsel, J. A. E., Swart, D. P. J., Baray, J.-L., Bencherif, H., Claude, H., Fehr, T., Godin-Beekmann, S., Hansen, G. H., Keck- hut, P., Leblanc, T., McDermid, I. S., Meijer, Y. J., Nakane, H., Quel, E. J., Stebel, K., Steinbrecht, W., Strawbridge, K. B., Tatarov, B. I., and Wolfram, E. A.: 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.

Velazco, V. A., Toon, G. C., Blavier, J.-F. L., Kleinböhl, A., Manney, G. L., Daffer, W. H., Bernath, P. F., Walker, K. A., and Boone, C. (2011), Validation of the Atmospheric Chemistry Experiment by noncoincident MkIV balloon profiles, J. Geophys. Res., 116, D06306, doi:10.1029/2010JD014928.

Wang, H.-J., Cunnold, D. M., Thomason, L. W., Zawodny, J. M., and Bodeker, G. E.: Assessment of SAGE version 6.1 ozone data quality, J. Geophys. Res., 107, 4691, doi:10.1029/2002JD002418, 2002.

Wurl, D., R. G. Grainger, A. J. McDonald, and T. Deshler, Optimal estimation retrieval of aerosol microphysical properties from SAGE II satellite observations in the volcanically unperturbed lower stratosphere, Atmos. Chem. Phys., 10, 4295-4317, 2010.

Acronyms and abbreviations

ACE-FTS - Atmospheric Chemistry Experiment - Fourier Transform Spectrometer
AMSU - Advanced Microwave Sounding Unit
ATBD - Algorithm Theoretical Baseline Document
AVK - Averaging Kernel
BS - Bodeker Scientific
BSUTLS - Bodeker Scientific Upper Troposphere Lower Stratosphere
CFH - Cryogenic Frost point Hygrometer
CFSR- Climate Forecast System Reanalyses
CHAMP - CHAllenging Mini-satellite Payload
COSMIC - Constellation Observing System for Meteorology, Ionosphere, and Climate
COSPAR - Committee on Space Research
CUSUM - Cumulative Sum
DI - Data Initiative
DU - Dobson Units
ECMWF - European Centre for Medium-Range Weather Forecasts
ENVISAT - Environmental Satellite
EOS - Earth Observing System
ESA - European Space Agency

ESA-TPM - ESA Thrid Party Mission
FMI - Finnish Meteorological Institute
GBL - GOMOS bright limb
GFZ - GeoForschungsZentrum
GOMOS Global Ozone Monitoring by Occultation of Stars
GRACE - Gravity Recovery and Climate Experiment
GPS - Global Positioning System
HALOE - The Halogen Occultation Experiment
IPF - Instrument Processor Facility
JAXA - Japan Aerospace Exploration Agency
LST - Local Solar Time
MIPAS - Michelson Interferometer for Passive Atmospheric Sounding
MLS - Microwave Limb Sounder
MPV - Modified Potential Vorticity
MSU - Microwave Sounding Unit
NCEP - National Centers for Environmental Prediction
NH - Northern Hemisphere
OSIRIS - Optical System for Imaging and low Resolution Integrated Spectroscopy
POAM - Polar Ozone and Aerosol Measurement
PVR - Product Validation Report
PVU - Potential Vorticity Units
RATPAC - Radiosonde Atmospheric Temperature Products for Assessing Climate
RMS - Root Mean Square
RO - Radio Occultation
SAD - Surface Area Density
SAGE - Stratospheric Aerosol and Gas Experiment
SAR Synthetic Aperture Radar
SCIAMACHY - Scanning Imaging Absorption Spectrometer for Atmospheric ChartographY
SCODA - SCIAMACHY Cloud Detection Algorithm
SEM - Standard Error of the Mean
SH - Southern Hemisphere
SMILES - Superconducting Submillimeter-Wave Limb-Emission Sounder
SMR - Sub-Millimetre Radiometer
SPARC - Stratospheric Processes And their Role in Climate
SPIN - ESA SPARC Initiative
SSU - Stratospheric Sounding Unit
SZA - Solar Zenith Angle
TH - Tangent Height
VRO - Vertically Resolved Ozone
WP - Work Package
TSX - TerraSAR-X

1. Introduction

This document describes the validation of the data sets produced or matured within the SPIN project to evaluate their quality. Chapter 2 describes the correlative data sets used for this purpose for stratospheric temperature, water vapour, aerosols and ozone. The co-location criteria are presented in Chapter 3 for data sets which are validated through the comparison of collocated profiles. These data sets are stratospheric water vapour from SCIAMACHY, strat-

ospheric aerosols from OSIRIS and SCIAMACHY, and stratospheric ozone from GOMOS bright limb (GBL) measurements. Due to the different sampling regimens, and different availability of correlative data, the co-location criteria differ between the data sets. For stratospheric temperature, climatologies are validated and therefore co-location criteria are not applicable. The validation results for all of the data sets developed within SPIN are shown in Chapter 4. Chapter 5 describes the concept for the evaluation of SMR and OSIRIS short-lived species climatologies. Results from the evaluation of these climatologies within the SPARC-DI are included. In Chapter 6, the PVR is summarized and inputs for the data product disclaimer are listed.

2. Description of correlative data used for validation/comparison of SPIN products

2.1. Stratospheric temperature

The following radio occultation (RO) data sets were used to validate the stratospheric temperature climatologies developed within WP18 of SPIN:

2.1.1. CHAMP

RO temperature profile measurements were obtained from the CHAMP satellite (Schmidt et al., 2004; Borsche et al., 2007) to generate monthly mean and zonal mean temperature time series on a latitude/pressure grid for the period May 2001 to October 2008. CHAMP provides global coverage of vertical profiles of dry temperature between 0 and 40 km altitude with a vertical resolution of 1.5 km in the stratosphere. The temperature profiles are accurate to approximately 0.5 K between 5-20 km altitude. In the stratosphere, the difference between the dry temperature and the physical temperature is negligible. CHAMP temperature profiles were interpolated in log pressure coordinates from their native pressure levels to the pressure levels defining the satellite-based temperature measurement climatologies i.e. 300, 250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10, 7, 5, 3, 2, 1.5 1, 0.7, 0.5, 0.3, 0.2, 0.15, and 0.1 hPa, although CHAMP temperature profiles at the upper pressure levels (<1 hPa) were not available. Interpolated values were corrected for their representativeness of zonal mean monthly mean temperature using NCEP (National Centers for Environmental Prediction) CFSR (Climate Forecast System Reanalyses) 6 hourly temperature fields on pressure surfaces as:

$$T_{corr} = T_{RO}(\theta, \varphi, P, t) \times \frac{\overline{T_{CFSR}}(5°, P, month)}{T_{CFSR}(\theta, \varphi, P, t)} \quad (3)$$

Where:

T_{corr} is the bias corrected temperature value,

T_{RO} is the RO temperature measurement interpolated onto pressure P at latitude θ, longitude φ at time t,

$\overline{T_{CFSR}}(5°, P, month)$ is the NCEP-CFSR 5° zonal mean monthly mean temperature at pressure P, and

T_{CFSR} is the NCEP-CFSR temperature at the same time and location as T_{RO}.

Applying equation (3) corrects the RO measurements for their sampling bias both in terms of geographical coverage and coverage within the month of interest.

2.1.2. GRACE

This twin satellite configuration is based on CHAMP heritage and was launched on 17 March 2002 (Beyerle et al., 2005). GRACE RO data were processed using the same algorithm as used for CHAMP to create monthly mean 5° zonal mean climatologies on the same pressure

grid as used for CHAMP. GRACE temperature climatologies from January 2006 to December 2011 were made available.

2.1.3. TSX

TerraSAR-X (hereafter referred to as TSX; Beyerle et al., 2011) RO data are routinely processed at GeoForschungsZentrum (GFZ), Potsdam. These data were treated using the same algorithm as used for CHAMP to create monthly mean 5° zonal mean climatologies on the same pressure grid as used for CHAMP. TSX temperature climatologies from July 2008 to March 2012 were made available.

2.1.4. NCEP-CFSR

In addition to the RO temperature climatologies detailed above, climatologies were calculated directly from the NCEP-CFSR reanalyses. These span the period January 1979 to December 2010. The highest level at which the NCEP-CFSR climatologies are available is 1 hPa.

The CHAMP, GRACE and TSX radio occultation products were also used directly in WP24 to create the improved UT/LS temperature record which is then used as the validation standard for the combined MSU4+AMSU9 temperature series.

2.2. Stratospheric Water Vapour

The SCIAMACHY limb water vapour profiles are compared against several other instruments: balloon-borne Cryogenic Frost point Hygrometer (CFH) data, measurements from the Earth Observing System (EOS) Microwave Limb Sounder (MLS) on Aura and the Atmospheric Chemistry Experiment – Fourier Transform Spectrometer (ACE-FTS).
The CFH data are the same as those used by Rozanov et al. (2011). Details on the instrument and measurement campaigns can be found in Vömel et al. (2007), Fujiwara et al. (2010), and Selkirk et al. (2010). The MLS data used were provided by the Jet Propulsion Laboratory, California Institute of Technology under contract with the National Aeronautics and Space Administration. MLS measures microwave radiation in the limb geometry. Here, water vapour from MLS level 2 data version 3.3 (v3.3) is used[1]. A validation of version 2.2, which is similar to v3.3, can be found in Read et al. (2007). ACE-FTS measures solar occultation of thermal infrared radiation. ACE, which flies on the SCISAT platform, is a Canadian-led mission mainly supported by the Canadian Space Agency and the Natural Sciences and Engineering Research Council of Canada. Version 3.0 ACE-FTS water vapour profiles are used as the validation data sets for the SCIAMACHY water vapour product. A validation of version 2.2 ACE-FTS retrievals can be found e.g. in Velazco et al. (2011) and Hegglin et al. (2008).

2.3. Stratospheric Aerosols

2.3.1. OSIRIS

The Stratospheric Aerosol and Gas Experiment (SAGE) II was in operation from 1985 to 2005, providing high quality occultation measurements of aerosol extinction at 385, 420, 525 and 1020 nm (Russell and McCormick, 1989). Although SAGE II does not have a 750 nm channel, data from neighbouring channels can be used to infer the Angstrom coefficient around 750 nm and produce a SAGE II 750 nm extinction product. The Angstrom coefficient can also be compared to the retrieved OSIRIS Angstrom coefficient, although the wavelength

[1] See http://mls.jpl.nasa.gov/data/v3-3_data_quality_document.pdf

discrepancy means comparisons will be qualitative only. Finally, SAGE II was used to retrieve integral properties of the particle size distribution such as the surface area density (Thomason, 1992, 1997). Although OSIRIS measurements are used to retrieve parameters of a log normal distribution, these can be converted to surface area density for comparison with SAGE II.

SAGE III was launched in 2002 and operated until 2005, producing aerosol extinction measurements from 385 to 1545 nm (Thomason, 2003). SAGE III was launched into a sun-synchronous orbit which produced measurements from middle to high latitudes. Although SAGE III has not been used to produce an aerosol particle size product, the wavelengths at 755 and 1545 nm are excellent channels for comparison of the OSIRIS extinction coefficient measurements. Thomason et al. (2010) estimate the accuracy and precision of these channels to be within 10%.

2.3.2. SCIAMACHY

The SCIAMACHY aerosol retrieval approach is validated by comparing the retrieved aerosol extinction profiles with co-located SAGE II measurements. The Stratospheric Aerosol and Gas Experiment II (SAGE II) was aboard the NASA Earth Radiation Budget Satellite (ERBS) and operated from 1984 to 2005. This solar occultation instrument was equipped with a seven-channel sun photometer, which measured, for example, stratospheric aerosols, ozone, water vapour, and nitrogen dioxide (McCormick, 1987). One of the SAGE II data products is stratospheric aerosol extinction profiles at 525 nm wavelength (e.g., Thomason, 1991). The SAGE II stratospheric aerosol data set is generally considered to be one of the stratospheric aerosol data sets with the highest accuracy, and is therefore well suited to serve as a benchmark for comparisons with the SCIAMACHY data set.

2.4. Stratospheric Ozone

To perform initial validation for the GOMOS bright limb ozone profiles, we used reference measurements from the GOMOS night-time occultation (v6 data) and version 3.3 MLS data.
The comparison was done using all available data between 2002 and 2012 with the analyses being done in 20° latitude zones e.g. 50°S-70°S, 30°S-50°S,...50°N-70°N. The difference between the ozone measurements from the two data sets is estimated using the median of the relative individual differences calculated as: (GBL-REFERENCE)/REFERENCE×100 % where the reference is GOMOS (night occultation) or MLS.

3. Description of co-location criteria used to compare SPIN products against selected ground-based and satellite-based observations

3.1. Stratospheric Water Vapour

For the comparisons of the SCIAMACHY water vapour data against MLS, ACE-FTS and CFH measurements, different collocation criteria were chosen because the validation data sets have differing measurement densities. These criteria are:

- CFH: less than 1000 km and 5 hours; data are available from 2002 to 2008, 171 collocations are found for V3.01.
- ACE-FTS: less than 500 km and 6 hours; data are available from 2004 to September 2010, 1888 collocations are found for V3.01.

- MLS: less than 100 km and 6 hours; data are available from 2004 to 2012, 13680 collocations are found for V3.01.

Additionally, a criterion for the modified potential vorticity (MPV) at 475 K is applied. The MPV is calculated from ECMWF data. Profiles are excluded from the comparison if the MPV differs by more than 3PVU or if the profiles are located at the edge of the polar vortex (30-40PVU). The effect of the MPV criterion is not large, which is to be expected since it regards a stratospheric quantity and the SCIAMACHY profiles cover only the upper troposphere and lower stratosphere. Therefore, for future comparisons, a criterion which detects different air masses based on their location with respect to the tropopause should be applied.

3.2. Stratospheric Aerosols

3.2.1. OSIRIS

The global coverage and different wavelengths of SAGE II relative to OSIRIS make large scale zonal comparisons ideal, particularly in the tropics where two volcanic eruptions occurred (Mts Ruang and Reventador and Mt. Manam) over the 4 year overlap of the missions. Although SAGE II provides global coverage, this occurs over relatively long time scales compared to OSIRIS measurements, and so averaging is done in monthly bins to provide continuity of measurements.

The SAGE III orbit does not yield global coverage and takes measurements at well-defined latitudes over the course of the year. While this makes large scale comparisons difficult, the accuracy of the SAGE III 755 nm channel makes comparisons with coincident measurements an excellent test of the OSIRIS 750 nm extinction product for middle to polar latitudes. The coincidence criteria used here are the same as in Bourassa et al. (2011); the reference tangent points of the OSIRIS and SAGE III measurements are within ±6 hours, ±1° latitude, and ±2.5° longitude. These tight criteria should produce comparisons of similar potential vorticity, and because OSIRIS does not make measurements during winter, eliminate comparisons across the steep gradients found at the polar vortex edge.

3.2.2. SCIAMACHY

The SCIAMACHY aerosol extinction is retrieved at altitudes between 12 and 32 km, using the 75°S modified ECSTRA aerosol extinction profile as the *a priori* profile for all latitudes. To determine the ground albedo, the Matthews database (Matthews, 1983) was used, which considers vegetation, land use, and land cover on a 1° × 1° grid. A cloud screening was not applied as previous validation results with SAGE II showed that cloud screening does not affect the comparisons between SCIAMACHY and SAGE II (see Ernst et al., 2012). However, below 20 km the frequent occurrence of tropospheric clouds prevents a robust retrieval at these latitudes. An analysis of the tropospheric cloud detection data set obtained with SCODA (SCIAMACHY Cloud Detection Algorithm) (Eichmann et al., 2009) showed that about 95% of all SCIAMACHY limb measurements are affected by tropospheric clouds.
The overlapping time period between SCIAMACHY and SAGE II used for the validation is from 1 January 2003 to 17 August 2005. All SCIAMACHY limb observations within a spatial distance of 500 km and a temporal difference of 6 hours to a SAGE II measurement were used. SCIAMACHY data with a SZA exceeding 87° were not considered. Moreover, the SCIAMACHY stratospheric aerosol extinction values were converted to 525 nm wavelength for the comparison to SAGE II observations using the assumed spectral dependence of the aerosol extinction coefficient. First, the Angstrom exponent (α) is calculated from the aerosol

extinction coefficients (ε) of the two wavelengths, 470 nm and 750 nm, for each tangent height:

$\alpha = - \log[\varepsilon(750nm)/\varepsilon(470nm)] / \log[750nm/470nm]$

Afterwards, the aerosol extinction coefficient at 525 nm is calculated by

$\varepsilon(525nm) = \varepsilon(750nm)/(750nm/525nm)^{-\alpha}$.

For the comparison, mean profiles and relative differences of mean profiles ((SCIAMACHY-SAGE)/SAGE) of the globally averaged SAGE II and SCIAMACHY stratospheric aerosol extinction profiles at 525 nm wavelength were calculated

3.3. Stratospheric Ozone

To compare the GBL data product developed in SPIN with the GOMOS night occultation data, the following criteria were used: Time difference less than 24 hours and a spatial difference less than 250 km. For the night occultation, measurements derived from stars with Mg>1.7 and T<7000K were excluded from the comparison.

To compare the GBL and MLS ozone profiles, the following criteria were used: Time difference less than 12 hours and a spatial difference less than 200 km. For this comparison we first converted the GBL profiles from a density/altitude grid to a mixing ratio/pressure grid using ECMWF temperature and pressure profiles.

4. Validation results

4.1. Stratospheric temperature

In WP18 monthly and zonal mean temperature climatologies were produced for three ESA or ESA-TPM based instruments, namely ACE-FTS, MIPAS and SMR. The version numbers of the temperature data used here are given in Table 3.2.5.1 of the PSD. To assess the quality of these climatologies they were validated against several radio occultation (RO) climatologies, also produced in WP18. The results of the validation are report in Section 4.1.1. The RO instruments include CHAMP, GRACE, and TSX. Various types of comparisons have been produced, including projecting the climatologies as

- latitude-pressure maps
- latitudinal profiles
- vertical profiles
- seasonal cycles
- monthly mean time series

Some of these comparisons also include reanalysis data, more specifically ERA-Interim and NCEP-CFSR.

Many comparisons were made over the 2004-2008 time period as this is the longest record covered by all three of the ESA-based climatologies.

To avoid the influence of sampling error on the comparisons, annual and multi-year averages are, unless otherwise stated, derived in such a way that a result is only presented if data are available for all time steps over which the averages are calculated (otherwise a no-data-value is applied).

WP24 develops an improved UT/LS temperature record using the ACE-FTS, MIPAS, SMR and radio occultation data sets (CHAMP, GRACE, TSX) derived in WP18. This improved UT/LS record is then used to assess the validity of the MSU4+AMSU9 data splicing. Our improved US/LS temperature record is a monthly mean record in 5° latitude bands on the same pressure grid as used in WP18. The new Bodeker Scientific UTLS temperature data set is validated against:

1. A reanalysis, in particular the NCEP-CFSR reanalysis.
2. COSMIC radio occultation data
3. The Integrated Global Radiosonde Archive (IGRA)

This validation is reported on in Section 4.1.2.

4.1.1. Validation of ACE-FTS, MIPAS and SMR climatologies and SSU-weighted MIPAS times series

Latitude pressure maps

First we compare the reference datasets, and start with the RO climatologies, comparing annual mean temperatures for periods of overlap between the respective missions. The CHAMP climatology extends over 2001-2008, the GRACE climatology extends over 2006-2011 and the TSX climatology extends over 2008-2012. There are, however, data gaps for individual months, so for complete years there is overlap in time between CHAMP and GRACE only in 2007 and between GRACE and TSX only in 2010 and 2011.

Figure 1 shows the annual average temperature in 2007 from GRACE. Figure 2 shows the difference in annually averaged temperature in 2007 between CHAMP and GRACE (CHAMP minus GRACE). GRACE and CHAMP agree to within 0.5 K in the lower stratosphere, around 20 km, but there is a robust bias that increases in magnitude with height, reaching 3-5 K in the upper stratosphere, near 40 km. CHAMP is consistently colder than GRACE in the southern hemisphere and consistently warmer than GRACE in the northern hemisphere.

Figure 3 shows, in a similar fashion, the difference between TSX and GRACE in 2010 and 2011. Similar to the comparison between CHAMP and GRACE, TSX is consistently colder than GRACE in the southern hemisphere and consistently warmer than GRACE in the northern hemisphere. The biases are of equal sign and magnitude (as in the comparison between CHAMP and GRACE) suggesting a hypothesis that a comparison between CHAMP and TSX would show a much smaller bias if a direct comparison was possible.

However, a comparison of TSX with CHAMP annual means is not possible due to the short temporal overlap between their time records. Moreover, comparing different years is not helpful since the magnitude of the interannual variability of monthly zonal mean temperature in the middle atmosphere is comparable to or exceeds the expected biases between the datasets, on the order of a few degrees at most. Figure 4 shows a comparison of CHAMP and TSX over the short overlap between the two datasets, from July to September 2008. The results partly confirm the hypothesis above. In the northern hemisphere CHAMP and TSX

generally agrees to within 1 K or better at all heights. In the southern hemisphere, however, CHAMP is colder by up to 15 K in the upper layers.

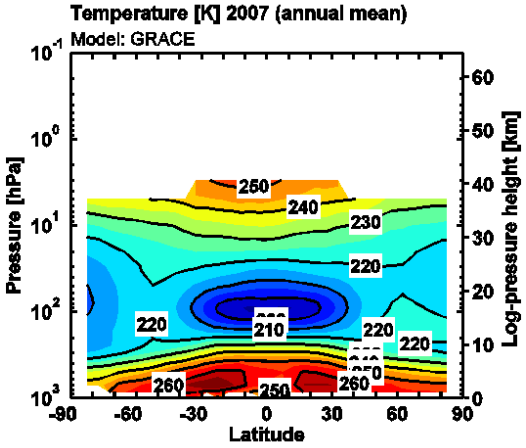

Figure 1: Annual mean temperature in 2007 from GRACE.

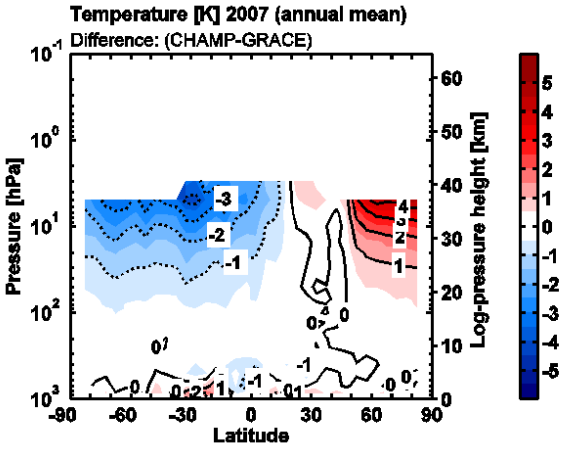

Figure 2: Annual mean temperature difference (CHAMP minus GRACE) for 2007.

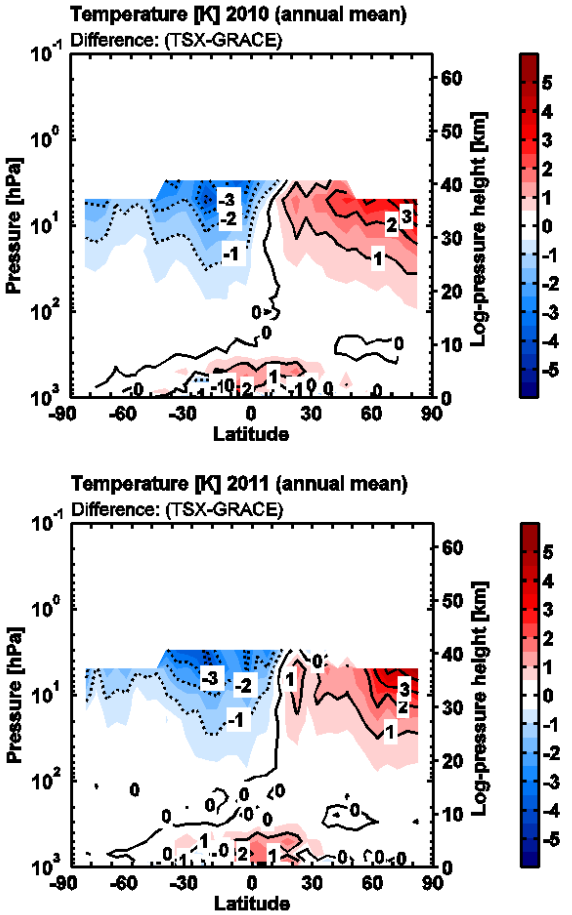

Figure 3: Annual mean temperature difference (TSX minus GRACE) for 2010 (top) and 2011 (bottom).

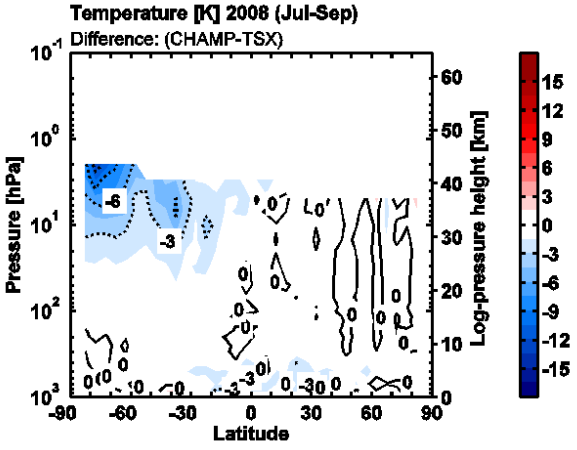

Figure 4: Seasonal mean temperature difference (CHAMP minus TSX) for Jul-Sep 2008.

Until the anomaly in the GRACE climatology is better understood it makes sense to use CHAMP and TSX as reference datasets for the ESA and ESA-TPM based climatologies. There is overlap between the ESA and ESA-TPM climatologies over 2004-2008, and hence CHAMP will effectively be used rather than TSX.

Next we compare the RO climatologies with two reanalysis datasets, ERA-Interim and NCEP-CFSR. Figure 5 shows ERA-Interim minus CHAMP, again annually averaged over 2007. Figure 6 shows NCEP-CFSR minus CHAMP for the same time period. Both ERA-

Interim and NCEP-CFSR are in reasonably good agreement with CHAMP throughout most of the stratosphere, where both reanalyses are generally 0-2 K colder than CHAMP. Although SPIN focuses on the stratospheric temperature record, it can be noted that both reanalyses are much warmer than CHAMP in the troposphere. This bias has not yet been investigated.

Comparisons of NCEP-CFSR and ERA-Interim with TSX in 2010 show similar results (Figure 7).

Figure 8 shows a comparison between NCEP-CFSR and ERA-Interim for 2007. The two datasets are very close, generally within 0.5 K throughout the troposphere and stratosphere below 30 km. Above 30 km NCEP-CFSR is warmer than ERA-Interim by up to 7 K.

It is unclear why both CHAMP and TSX are systematically colder than NCEP-CFSR and ERA-Interim in the stratosphere. The high level of agreement between NCEP-CFSR and ERA-Interim up to 30 km is convincing. To evaluate the ESA and ESA-TPM climatologies we have therefore chosen to show comparisons with ERA-Interim, keeping in mind that both CHAMP and TSX are on average 0-2 K colder than ERA-Interim, and that ERA-Interim temperatures in the upper stratosphere are highly uncertain, as the comparison with NCEP-CFSR shows.

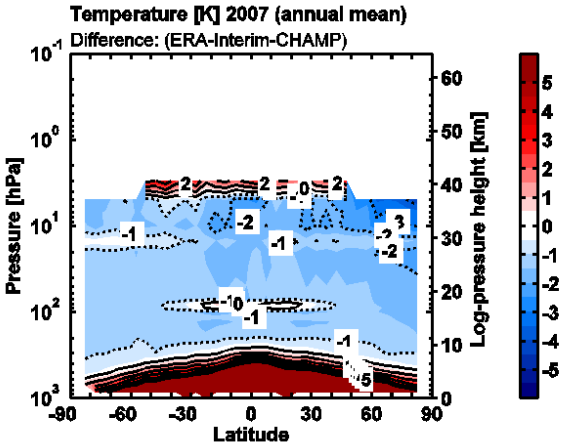

Figure 5: Annual mean temperature difference (ERA-Interim minus CHAMP) for 2007.

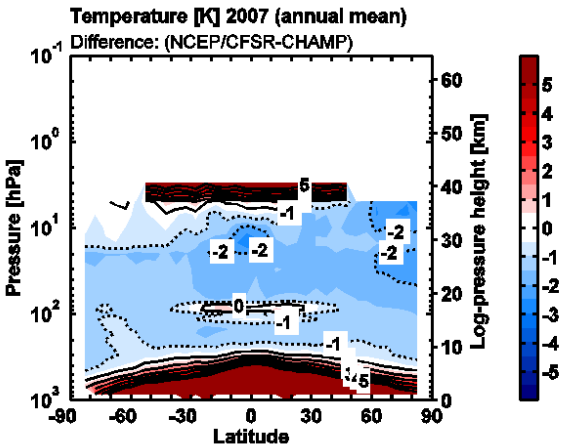

Figure 6: Annual mean temperature difference (NCEP-CFSR minus CHAMP) for 2007.

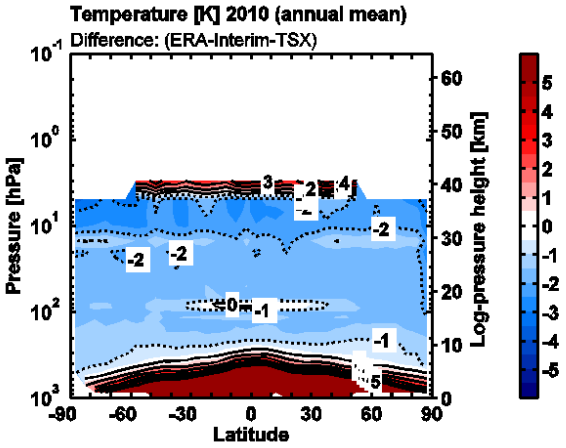

Figure 7: Annual mean temperature difference (ERA-Interim minus TSX) for 2010.

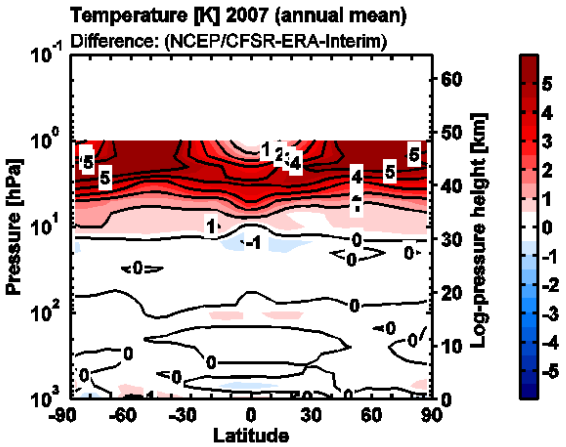

Figure 8: Annual mean temperature difference (NCEP-CFSR minus ERA-Interim) for 2007.

Figure 9 to Figure 13 shows the difference in annual mean temperature between the ESA and ESA-TPM-based climatologies and CHAMP and ERA-Interim. Figure 9 and Figure 10 shows the difference between MIPAS and CHAMP and the difference between SMR and CHAMP, respectively, for 2007. To establish the relative biases between the climatologies also in the upper stratosphere, where the RO climatologies generally don't provide measurements, MIPAS and SMR are also compared to ERA-Interim (Figure 11 and Figure 12).

Note that CHAMP is on average 1-2 K warmer than ERA-Interim in the stratosphere up to about 40 km. This contributes to some of the negative biases displayed by both MIPAS and SMR with respect to CHAMP (Figure 9 and Figure 10). The MIPAS and SMR climatologies both generally compare more favourably with ERA-Interim (Figure 11 and Figure 12).

MIPAS agrees very well with ERA-Interim, generally within 2 K throughout the region of comparison. Specifically the agreement is particularly excellent, within 1 K, throughout the 15-30 km region. SMR is warmer by up to 5 K at 20-25 km, colder by ~5 K at 35-45 km, and colder by ~20 K at 50 km, than ERA-Interim.

The measurement coverage of ACE-FTS is much sparser than that of MIPAS and SMR, and hence latitudinal data gaps are much more prevalent in the ACE-FTS climatology Therefore the ACE-FTS climatology is assessed over the five year period from 2005-2009, over which ACE-FTS provides a continuous record (i.e. for which the climatology provides data in each

month). Note however, that due to the sparsity of the dataset, for a given latitude, it is generally not possible to derive a multi-year average without the introduction of sampling error. This should be kept in mind when interpreting the ACE-FTS results. Figure 13 shows the difference in annual mean temperature between the ACE-FTS and ERA-Interim climatologies.

ACE-FTS is generally within 5 K of ERA-Interim at low and middle latitudes but discrepancies reach up to 25 K at high latitudes. These large biases are mostly due to sampling error, which is expected to be largest at high latitude where the interannual variability in temperature is large. Comparisons of ACE-FTS with ERA-Interim and CHAMP on a month to month basis (not shown) show significantly smaller biases, however such detailed comparisons goes beyond the scope of WP18, which focuses on large scale comparisons that are relevant for developing an extension of the global stratospheric temperature record of SSU in WP23.

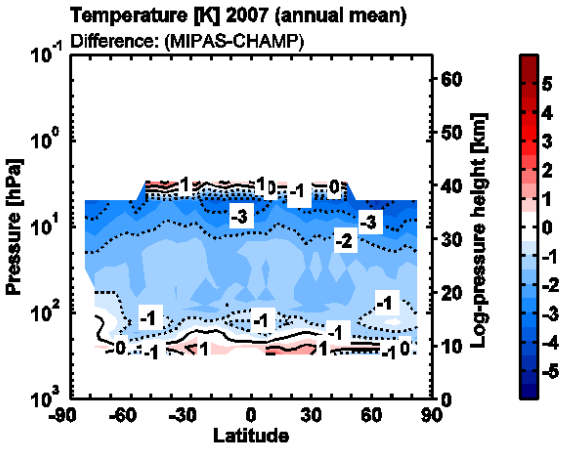

Figure 9: Annual mean temperature difference (MIPAS minus CHAMP) for 2007.

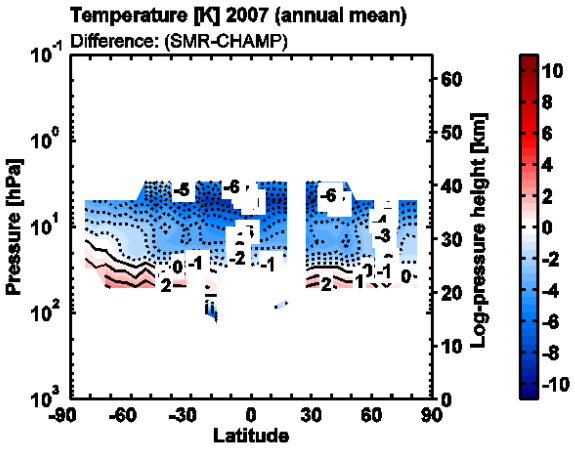

Figure 10: Annual mean temperature difference (SMR minus CHAMP) for 2007.

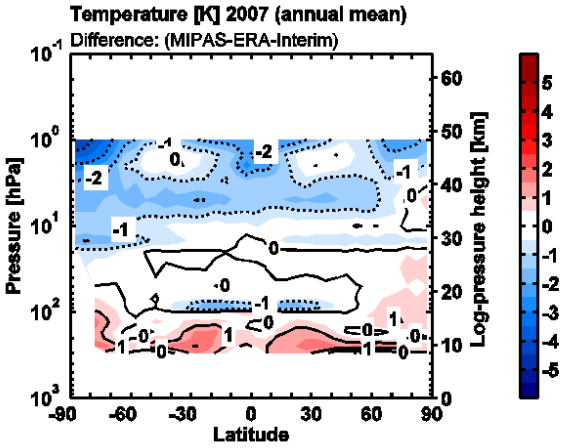

Figure 11: Annual mean temperature difference (MIPAS minus ERA-Interim) for 2007.

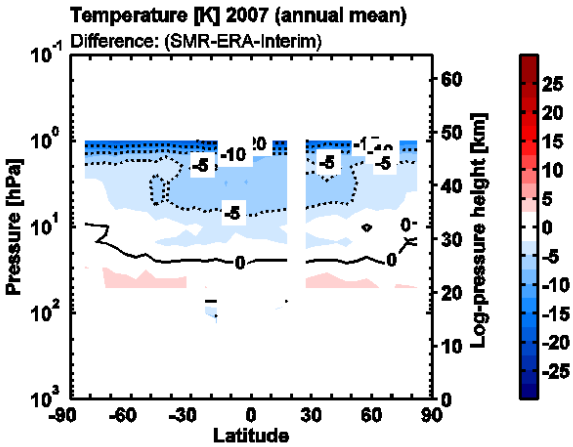

Figure 12: Annual mean temperature difference (SMR minus ERA-Interim) for 2007.

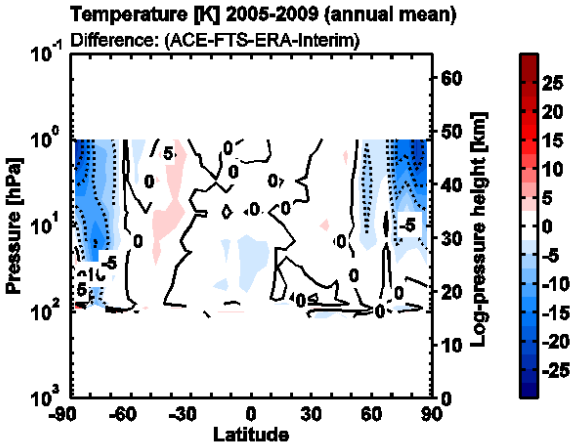

Figure 13: Annual mean temperature difference (ACE-FTS minus ERA-Interim) for 2005-2009.

Latitudinal profiles

Figure 14 shows the temperature climatologies as a function of latitude at 10 hPa in January over the 2004-2008 period.

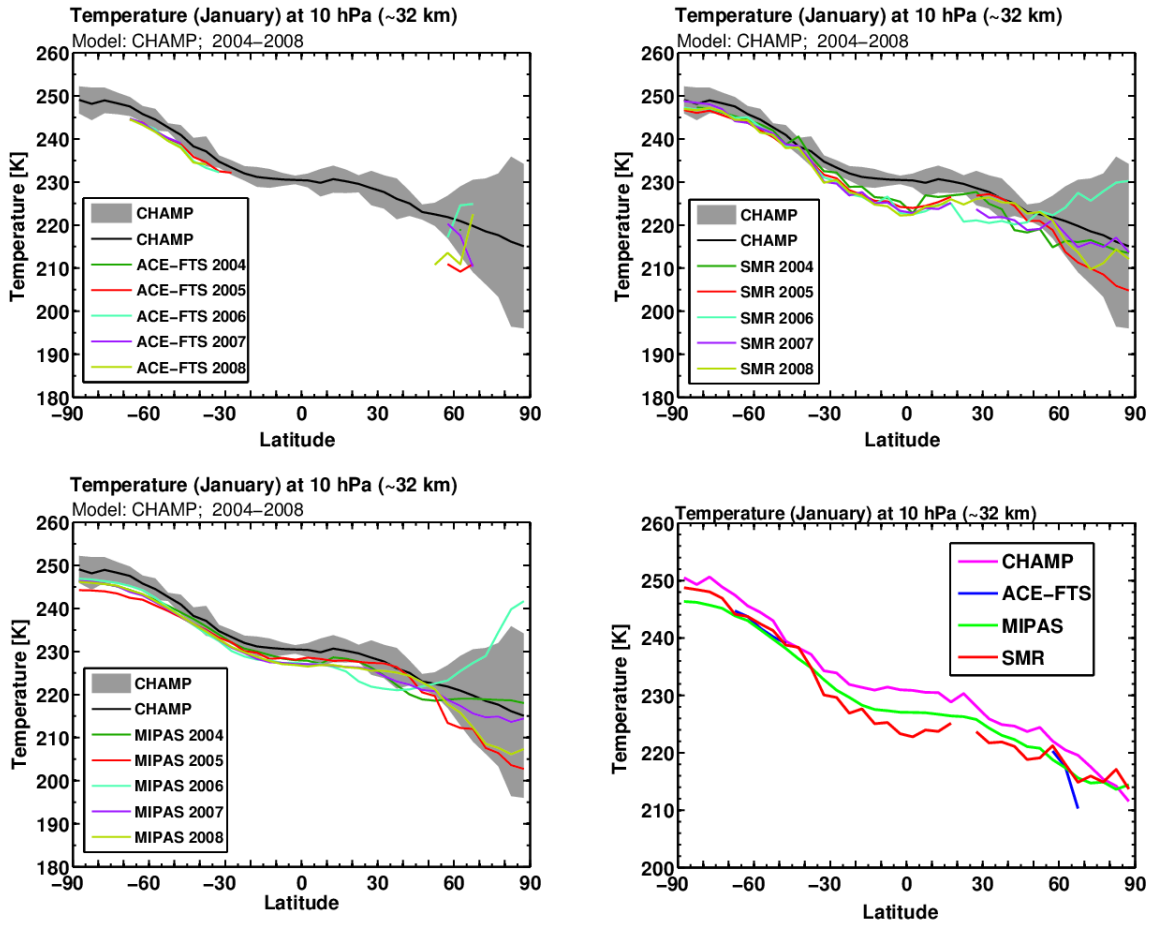

Figure 14: Temperature as a function of latitude at 10 hPa in January for individual years over 2004-2008. The solid black line and the grey shading represent the CHAMP 2004-2008 average and plus minus two standard deviations. The lower right-hand side panel shows the January 2007 averages for different instruments.

Vertical profiles

Figure 15 shows vertical profiles of zonal mean temperature at the equator for January for CHAMP, MIPAS and SMR. The CHAMP data represent all the years 2004 to 2008 (mean and variability), while the MIPAS and SMR data are shown for the individual years. MIPAS is seen to agree better with CHAMP than does SMR, with the latter exhibiting a cold bias of up to 10K in the middle to upper stratosphere. The unphysical temperature drop at 40 km in CHAMP is an artifact; this feature does not occur in the latest version of the climatologies.

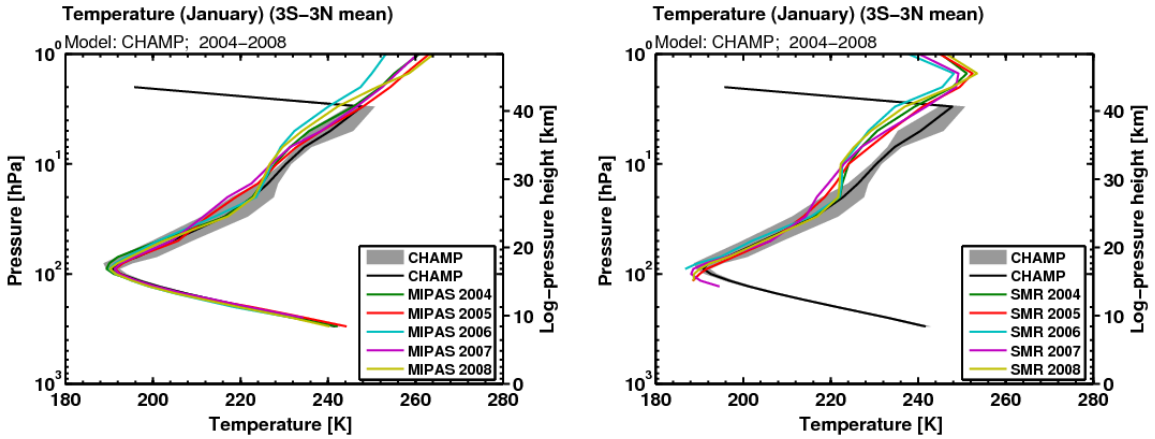

Figure 15: Vertical profiles of temperature at the equator for January: (left) MIPAS versus CHAMP and (right) SMR-STRAT versus CHAMP. CHAMP data is averaged over the years 2004 to 2008; gray shading denotes 2 times the standard deviation. The colours denote the different years from 2004 to 2008

Seasonal cycles

Figure 16 to Figure 18 show the temperature climatologies as a function of month in the tropics at 2, 10 and 50 hPa respectively. All available years, for each climatology, are included in the figures. Figure 19 shows the same type of plot at northern hemisphere middle latitudes at 10 hPa. Note that RO data generally do not extend up to the stratopause, so Figure 16 does not include any RO datasets. Also note that at 50 hPa the SMR climatology does not provide a complete coverage in the tropics, so the seasonal cycle cannot be fully resolved there.

For the tropical comparisons the following can be noted: There is some general offset between the datasets but the seasonal cycle (annual in the lower stratosphere and semi-annual in the upper stratosphere) in terms of magnitude and phase agrees well. There is excellent agreement between MIPAS and ERA-Interim for individual years, e.g. the anomalously warm boreal autumn of 2002 at 2 hPa and the anomalously cold boreal summer of 2010 at the same height. This is to some extent also captured by SMR.

At middle latitudes (Figure 19) there is very little interannual variability in the seasonal cycle. Note, however, that SMR displays significantly more variability than both ERA-Interim and CHAMP, as well as MIPAS.

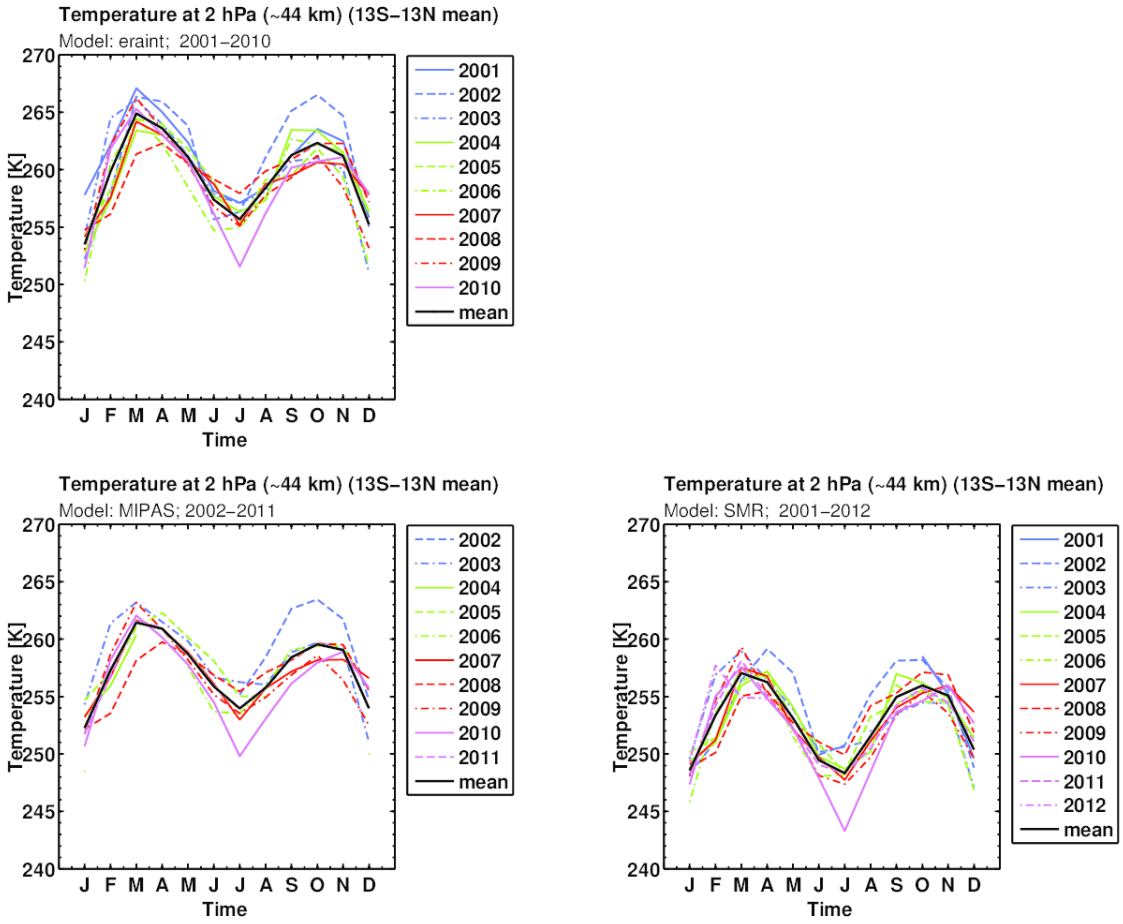

Figure 16: Temperature as a function of month in the tropics at 2 hPa for individual years. Climatologies for ERA Interim, MIPAS and SMR are shown. No RO data is available at this altitude. The data represents a latitudinal average across 15S-15N. The solid black line shows the multi-year mean.

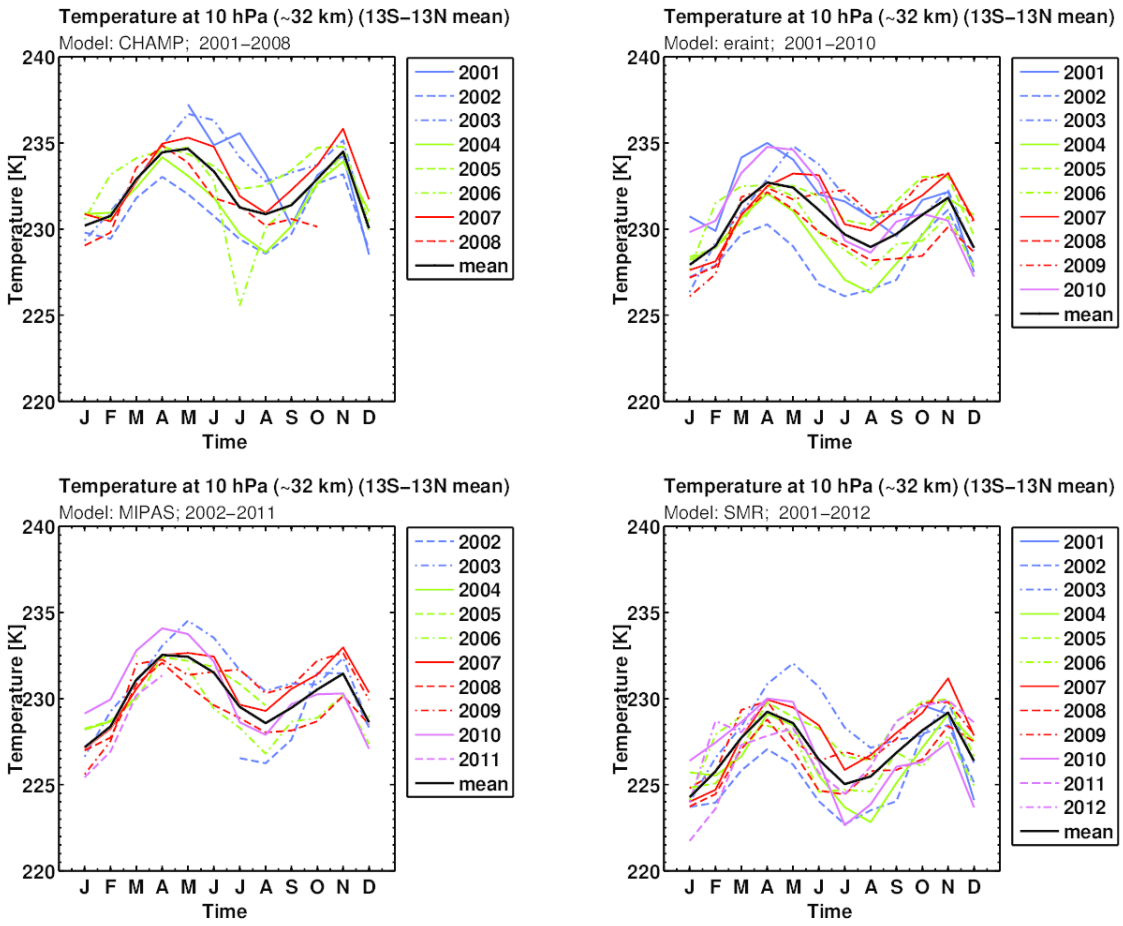

Figure 17: Temperature as a function of month in the tropics at 10 hPa for individual years. Climatologies for CHAMP, ERA Interim, MIPAS and SMR are shown. The data represents a latitudinal average across 15S-15N. The solid black line shows the multi-year mean.

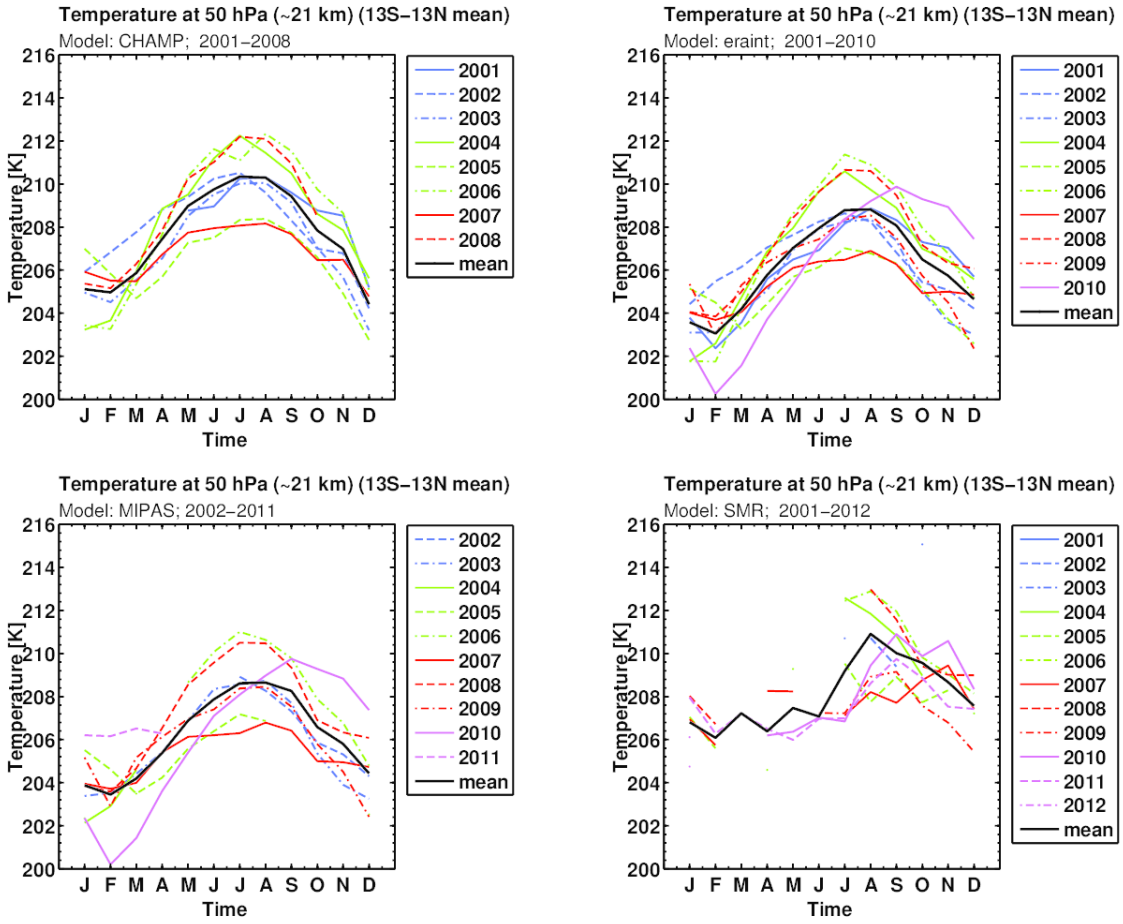

Figure 18: Temperature as a function of month in the tropics at 50 hPa for individual years. Climatologies for CHAMP, ERA Interim, MIPAS and SMR are shown. The data represents a latitudinal average across 15S-15N. The solid black line shows the multi-year mean.

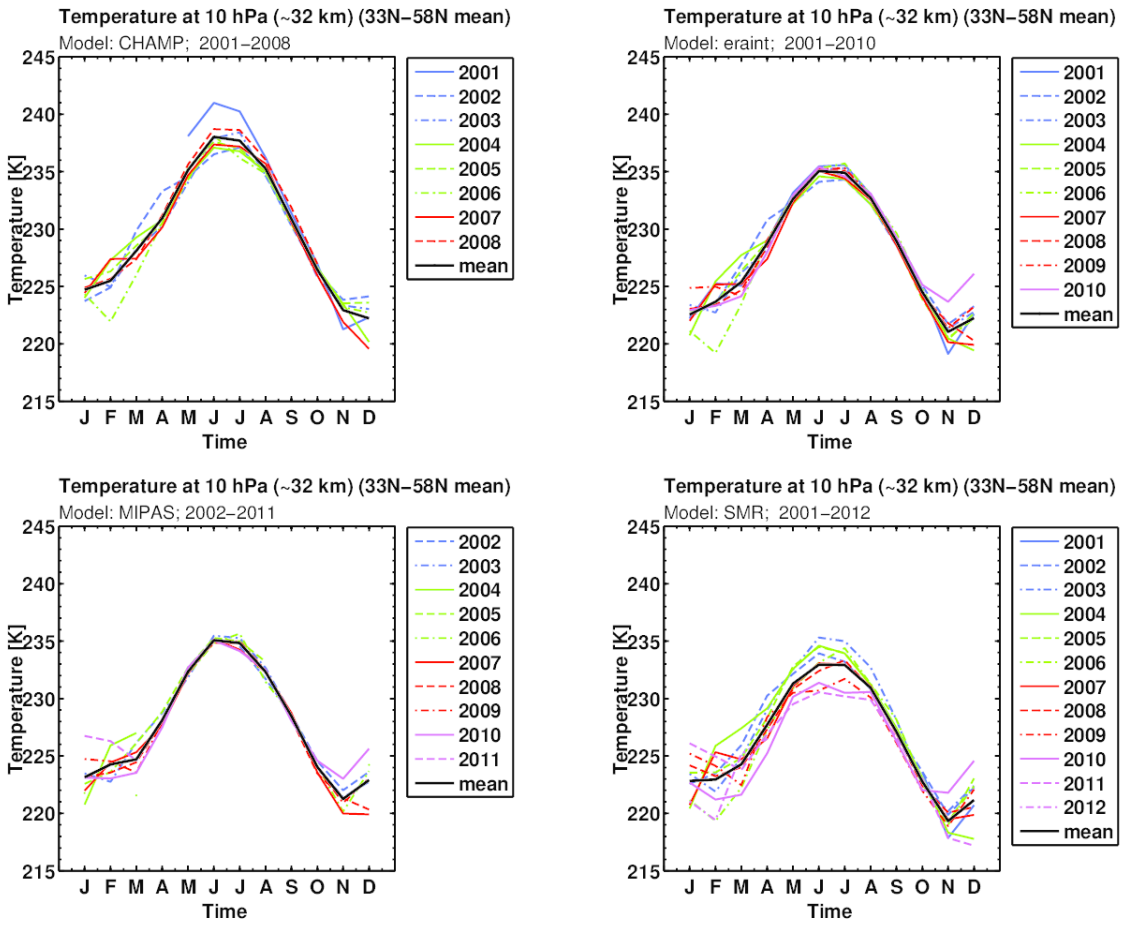

Figure 19: Temperature as a function of month at northern hemisphere middle latitudes at 10 hPa for individual years. Climatologies for CHAMP, ERA Interim, MIPAS and SMR are shown. The data represents a latitudinal average across 30N-60N. The solid black line shows the multi-year mean.

Monthly mean time series

Figure 20 shows time series of deseasonalized monthly and near-global mean (70°S-70°N) temperatures at 2 hPa (~44 km) for ACE-FTS, MIPAS and SMR, as well as for the reanalysis data (ERA-Interim). The much larger variability in ACE-FTS is a result of large spatial sampling errors – there simply is not enough latitudinal coverage to accurately compute a near-global mean. Consequently ACE-FTS will not be used further in the analysis. Regarding the two other ESA instruments, MIPAS and SMR, both exhibit a cooling trend over the ten-year period that is not seen in ERA-Interim. The temporal variability for MIPAS is similar in magnitude to that for ERA-Interim, but is larger for SMR. There is also a ~4 K offset between the two instruments.

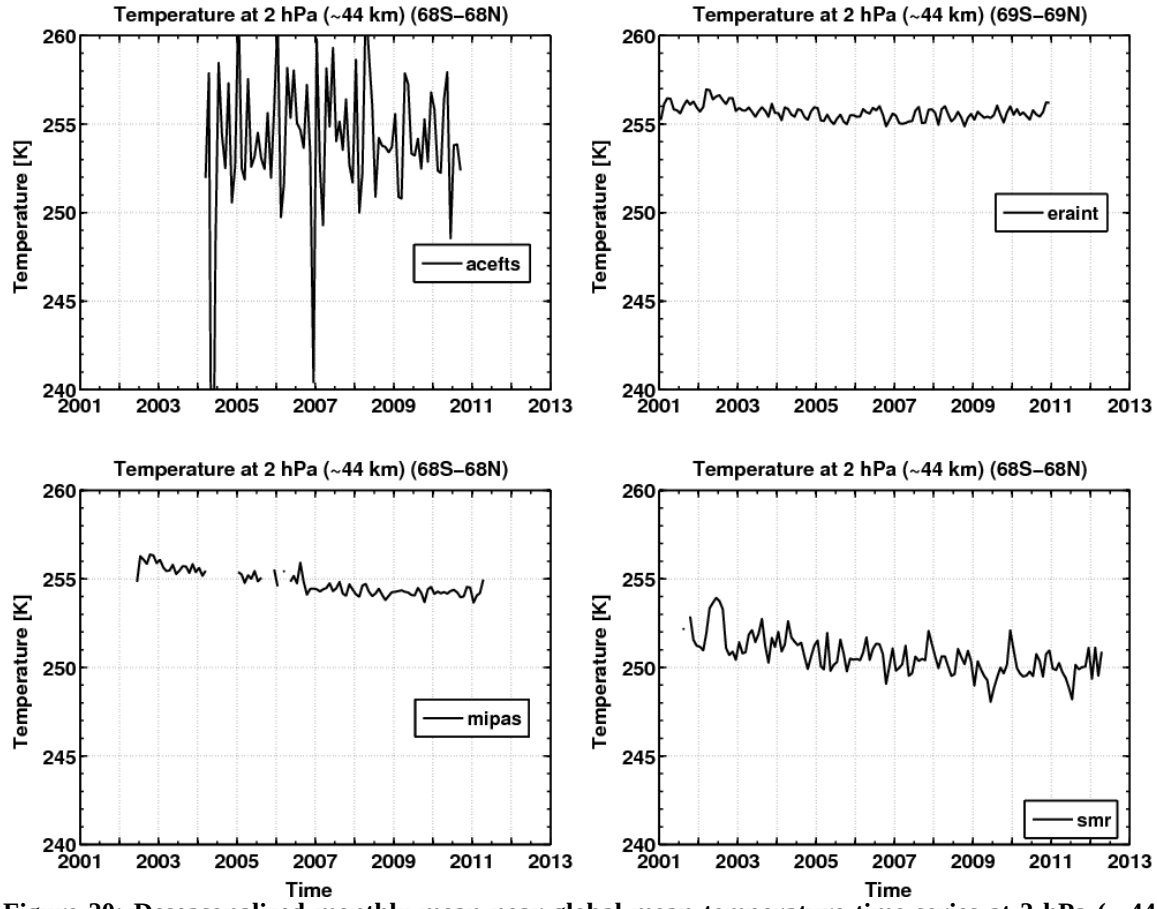

Figure 20: Deseasonalized monthly mean near-global mean temperature time series at 2 hPa (~ 44 km) for: (top left) ACE-FTS, (top right) ERA Interim, (bottom left) MIPAS and (bottom right) SMR.

The next stage of the analysis entails vertically averaging the monthly and zonal mean temperature climatologies of the ESA instruments using the SSU weighting functions. This permits a direct comparison of the ESA climatologies to SSU. Figure 21 shows vertical profiles of the weighting functions for the three channels that are used. The weighting functions (and temperature data presented shortly) are for the new NOAA NESDIS/STAR SSU dataset. Channel 1 (also referred to as SSU25) peaks at 15 hPa, channel 2 (SSU26) at 5 hPa, and channel 3 (SSU27) at 2 hPa.

In order to apply the SSU weighting functions to the ESA data, special attention must be paid to the vertical range of the data since the weighting functions extend from the surface to well into the mesosphere. For MIPAS this is not much of an issue since it has good vertical coverage. However, for SMR and CHAMP this is an issue. The solution arrived at was before applying the weighting functions to fill the regions of missing data below 1 hPa using ERA-Interim temperatures (for the appropriate time period) and above 1 hPa using the CIRA climatology[2].This procedure will be referred to as "filling." In the case where the fill data act to adversely modify the temporal variability of the resulting time series, such profiles are excluded from further analysis (see below for further detail).

[2] The COSPAR International Reference Atmosphere (CIRA-86), NCAS British Atmospheric Data Centre, 2006-03/2012. Available from http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_CIRA .

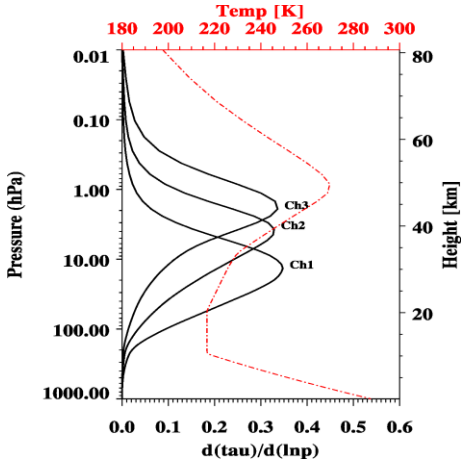

Figure 21: SSU weighting functions for channels 1 to 3 (black) and a typical temperature profile.

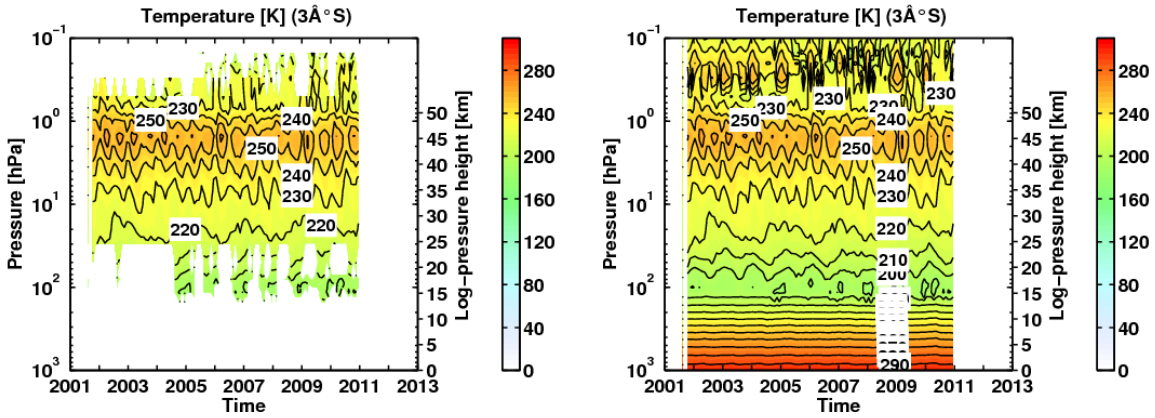

Figure 22: (Left) Deseasonalized monthly mean SMR temperature time series at the equator. (Right) Same but with the missing data below filled with ERA-Interim data and above filled with CIRA data.

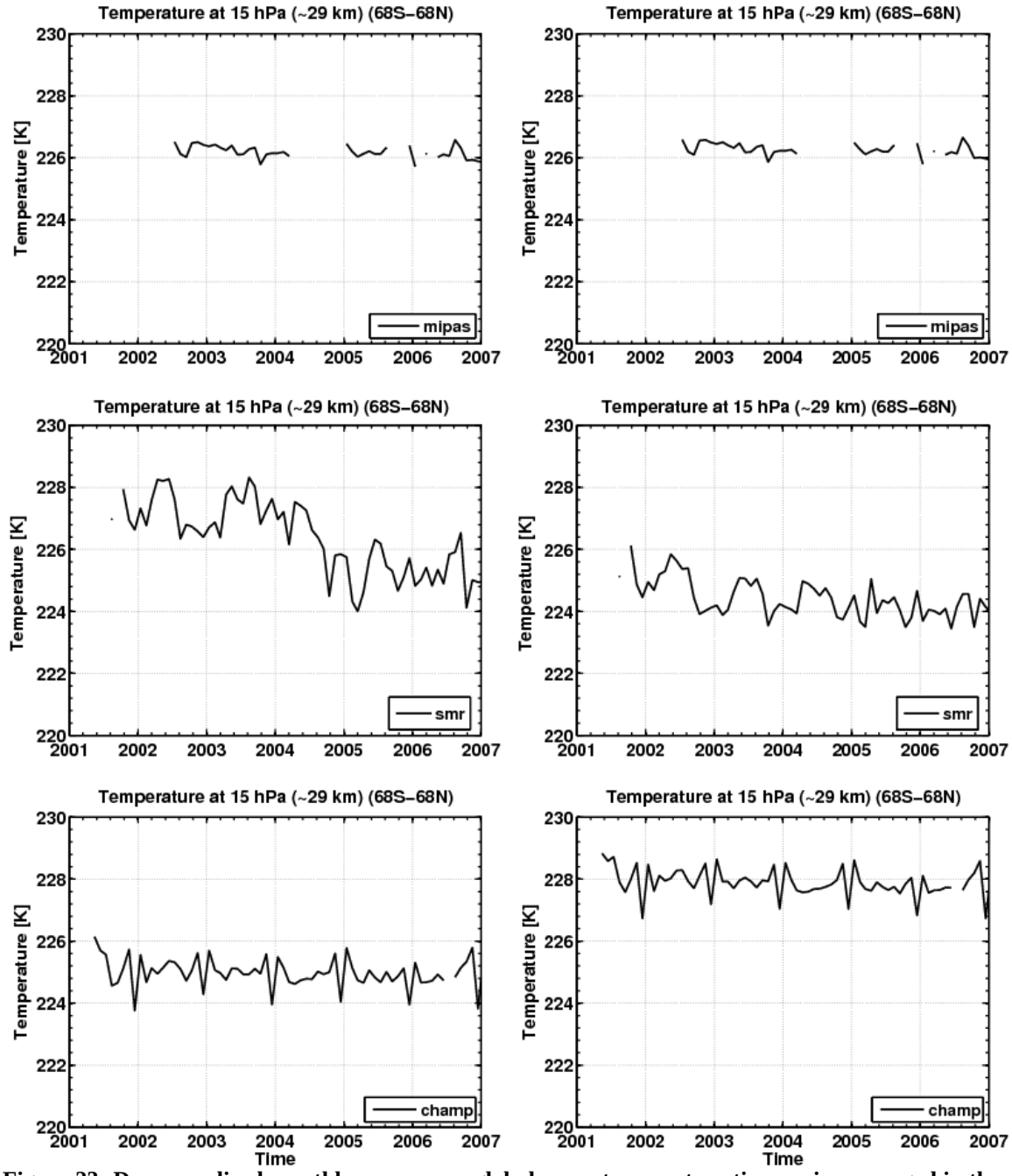

Figure 23: Deseasonalized monthly mean near-global mean temperature time series averaged in the vertical using the SSU Channel 1 weighting function (peak at 15 hPa): (top) MIPAS, (middle) SMR and (bottom) CHAMP; (left) without filling (i.e., the original climatology) and (right) with filling (using ERA Interim below and CIRA above).

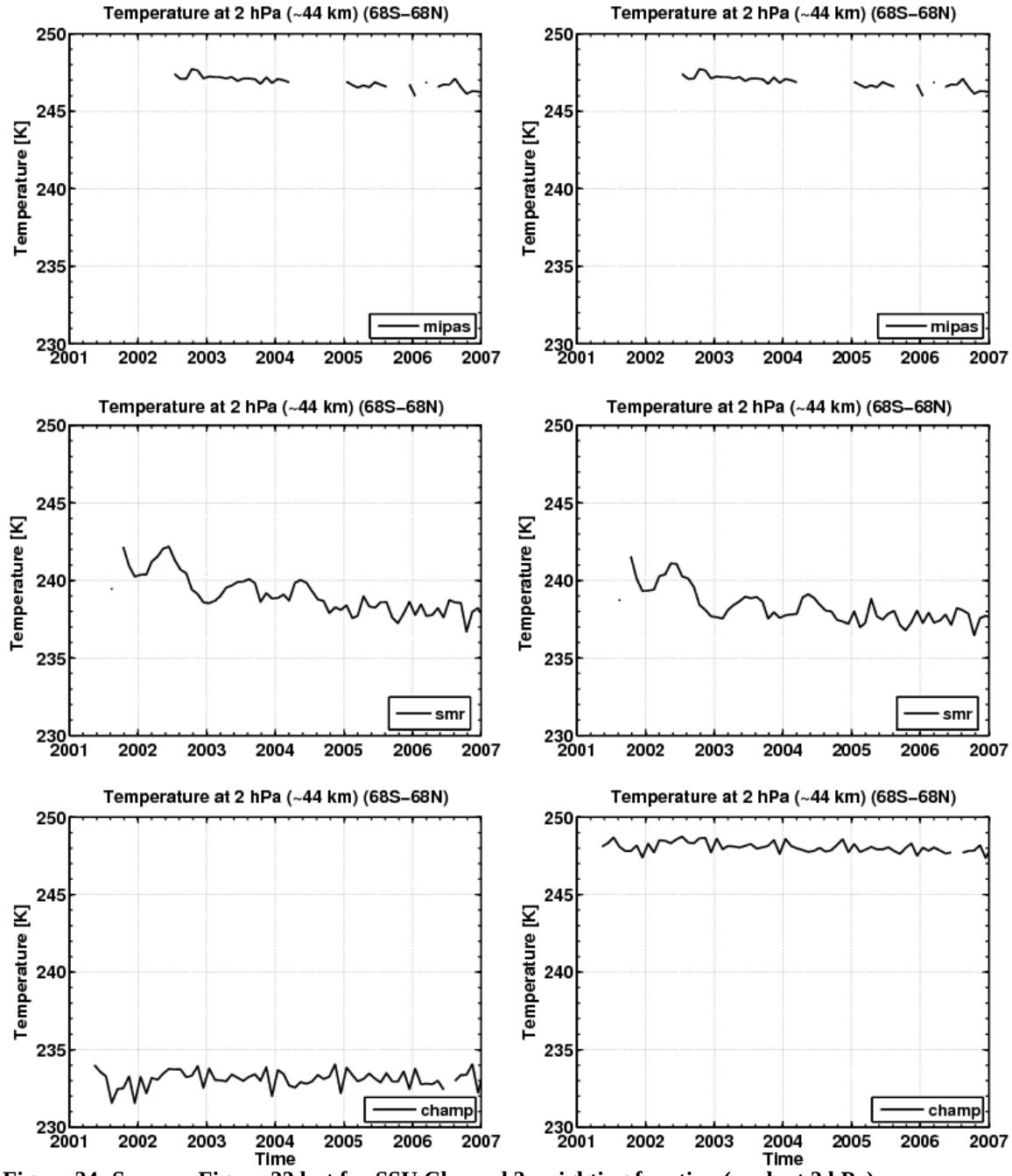

Figure 24: Same as Figure 23 but for SSU Channel 3 weighting function (peak at 2 hPa).

Figure 22 shows an example of the results of filling in the case of SMR temperatures at the equator. The left and right panels show the unfilled (i.e. original) and filled data, respectively. As mentioned above the filled data are then vertically averaged using the SSU weighting functions.

Figure 23 and Figure 24 show the impact of filling on the SSU-weighted MIPAS, SMR and CHAMP temperatures for channels 1 and 3. For the lower stratosphere (Figure 23) comparison of the left and right columns indicates that the filling has little impact on the MIPAS results (top row), which is what is desired. However, for SMR (middle row) the filling has caused a large change. It is in fact the filling in the region below 1 hPa that has caused this change (results not shown). Given that this is the lower stratosphere, the large impact of the fill data on the SSU-weighted SMR temperatures is not surprising (see Figure 22). For CHAMP (bottom row), filling results in an upward shift (i.e., a nearly constant offset) of the temperatures, with the temporal variability largely unaffected. In the upper stratosphere (Figure 24), filling produces almost no change in the MIPAS and SMR results, but a large change, consisting primarily of an upward shift, as well as a change in variability, for CHAMP. The shift is mainly a result of the filling from above (results not shown).

Figure 25 to Figure 27 show the SSU-weighted temperatures for MIPAS, SMR and CHAMP for the three SSU channels, as well as the SSU data themselves. The time series for the three instruments are substantially different. As seen from Figure 23 and Figure 24 a substantial fraction of those differences is a result of filling, which has strongly impacted on SMR in the lower stratosphere (channel 1) and on CHAMP in the upper stratosphere (channel 3). Note also the downward trend and large annual oscillation for SMR for channels 2 and 3, which do not appear in any of the other time series.

It is concluded from the above analysis that the SSU weighting for all three channels can only be applied to the MIPAS data: the SSU-weighted SMR and CHAMP data undergo far too large changes in the lower and upper stratosphere as a result of filling, and in the case of SMR exhibit unrealistically large inter-annual variations. Therefore, SMR data cannot be used to extend the SSU record, while CHAMP data cannot be used to validate the extended SSU record.

Since filling has only a very minor impact on the SSU-weighted MIPAS results (Figure 23 and Figure 24), it will not be used in the subsequent analysis. Instead the SSU weighting functions are only applied over the MIPAS data height range, namely 300 to 0.1 hPa. This was done so as to avoid any possible “contamination” of the SSU-weighted MIPAS time series data by other data sources, namely ERA Interim and CIRA. However, by restricting the height range to these values the vertical integral of the full (i.e. surface to upper mesosphere) SSU weighting functions are slightly underestimated. For channels 1, 2 and 3 the underestimates are ~ 1%, 1% and 3%, respectively. Since the procedure used to do the vertical weighting divides the weighted sum by the sum of the weights, this means that the SSU-weighted time series for channels 1, 2 and 3 are implicitly multiplied by 1.01, 1.01 and 1.03, respectively. This small implicit scaling of the weighted time series could have a slight impact on the results, since, for example, if there was a trend in the filled region but not in the unfilled region, then inflating the filled region to compensate for the incomplete sampling would overestimate the trend. However any such effect is likely to be small.

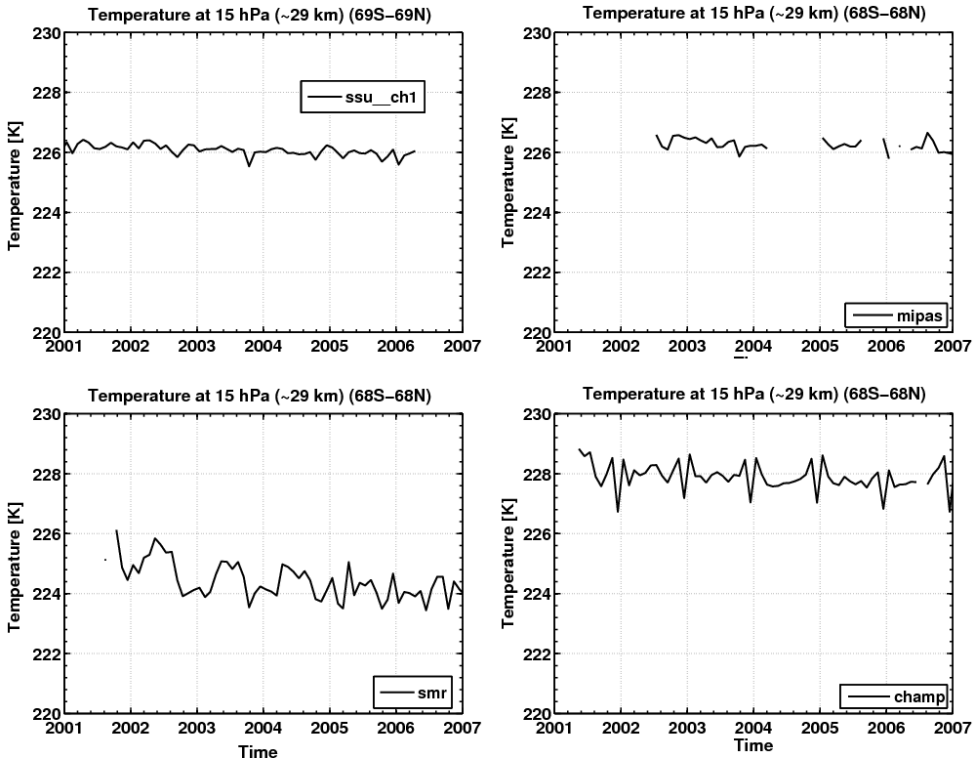

Figure 25: Deseasonalized monthly mean near-global mean temperature time series for SSU Channel 1: (top left) SSU, (top right) SSU-weighted MIPAS, (bottom left) SSU-weighted SMR and (bottom right) SSU-weighted CHAMP.

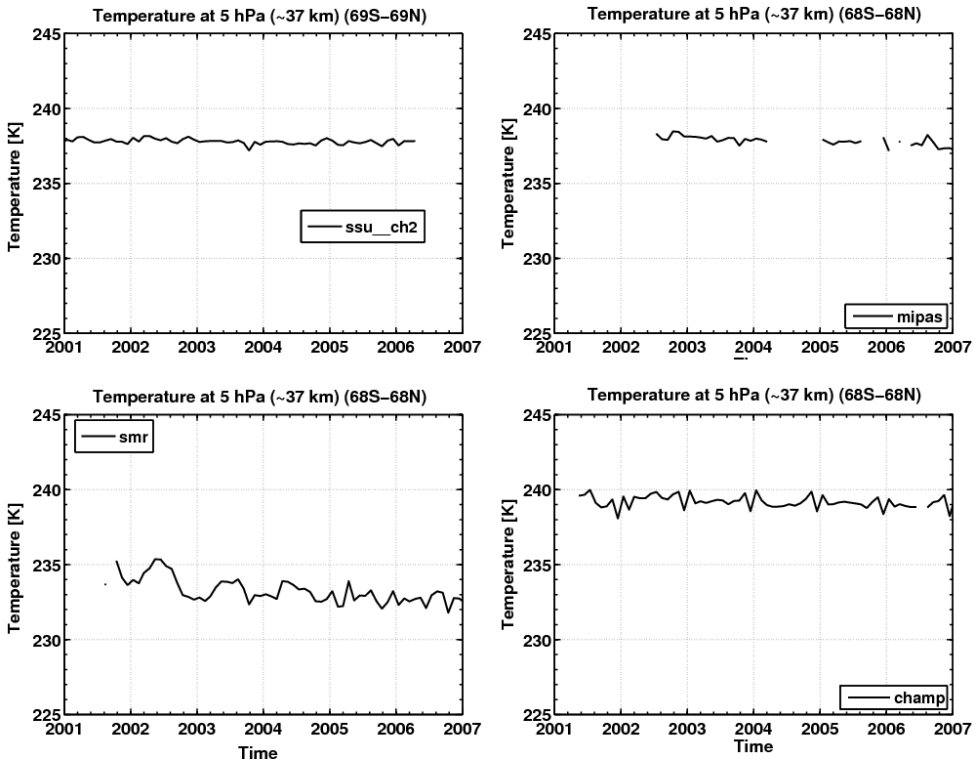

Figure 26: Same as Figure 25 but for SSU Channel 2.

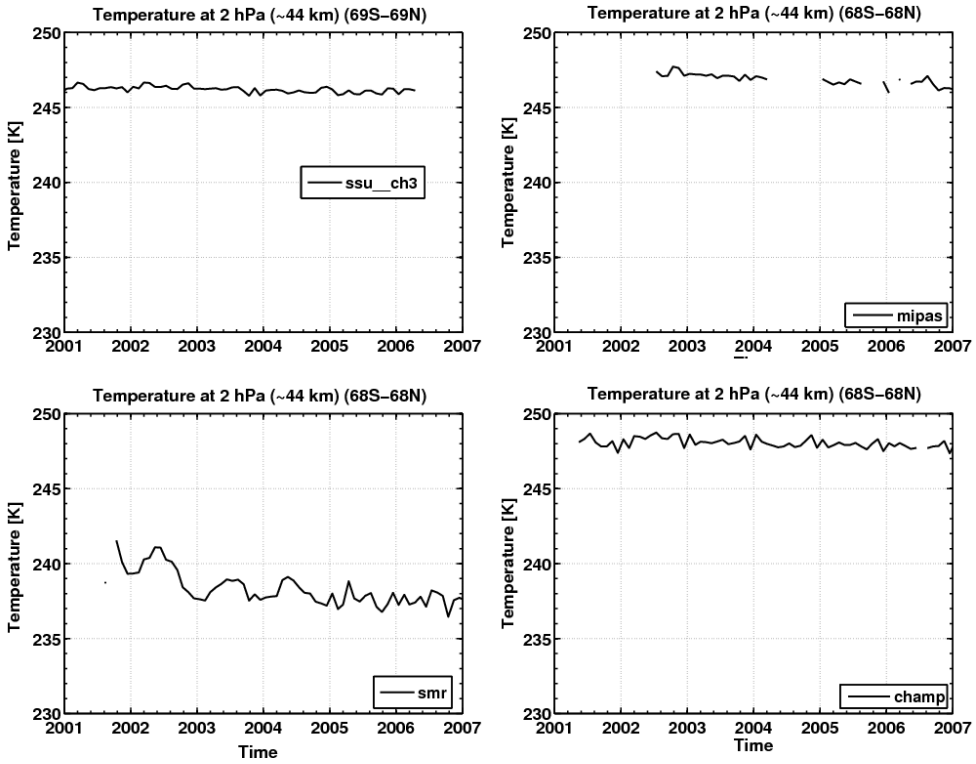

Figure 27: Same as Figure 25 but for SSU Channel 3.

Combined SSU and SSU-weighted MIPAS time series

In this section the MIPAS data are merged with the SSU data to provide a time series extending from 1978 to 2011. To do this differences between the deseasonalized SSU-weighted MIPAS data and the SSU data in the overlap period (2002-2006) must first be characterized and then used to adjust the MIPAS data. This necessitates the use of global (or near-global) mean temperatures, which are to first order unaffected by dynamics and therefore suitable for characterizing the differences. The same property makes them a key quantity for detection and attribution of stratospheric change, hence the strong scientific focus on the near-global time series.

Figure 28 shows deseasonalized near-global mean temperature time series for the three SSU channels, with SSU in red and the SSU-weighted MIPAS data denoted by the *thin* blue line. Comparison of the two curves in the overlap period indicates that MIPAS is consistently warmer than SSU, indicating that to a first approximation constant offsets can be applied to "correct" the MIPAS data. (Note that the SSU data are being treated here as the truth, and so are not modified.) These offsets, which are given by the time average difference between MIPAS and SSU over the overlap period, are ~ 0.18, 0.14 and 0.81 K for channels 1, 2 and 3, respectively. The thick blue lines in Figure 28 show the SSU-weighted MIPAS data with the offsets added on. The thick blue curves now closely follow the SSU data in the overlap region. However, there is a curious "hiccup" in the SSU-weighted MIPAS data in 2006, which is especially prominent in Channels 2 and 3.

Figure 29 is identical to Figure 28 except that it includes the seasonal cycle. Note that the seasonal cycles for MIPAS and SSU are computed separately over their respective periods of time coverage, namely 1979 to 2006 for SSU and 2002 to 2011 for MIPAS. This figure shows

that in the overlap period there is excellent agreement in the seasonal cycles of the SSU and MIPAS data.

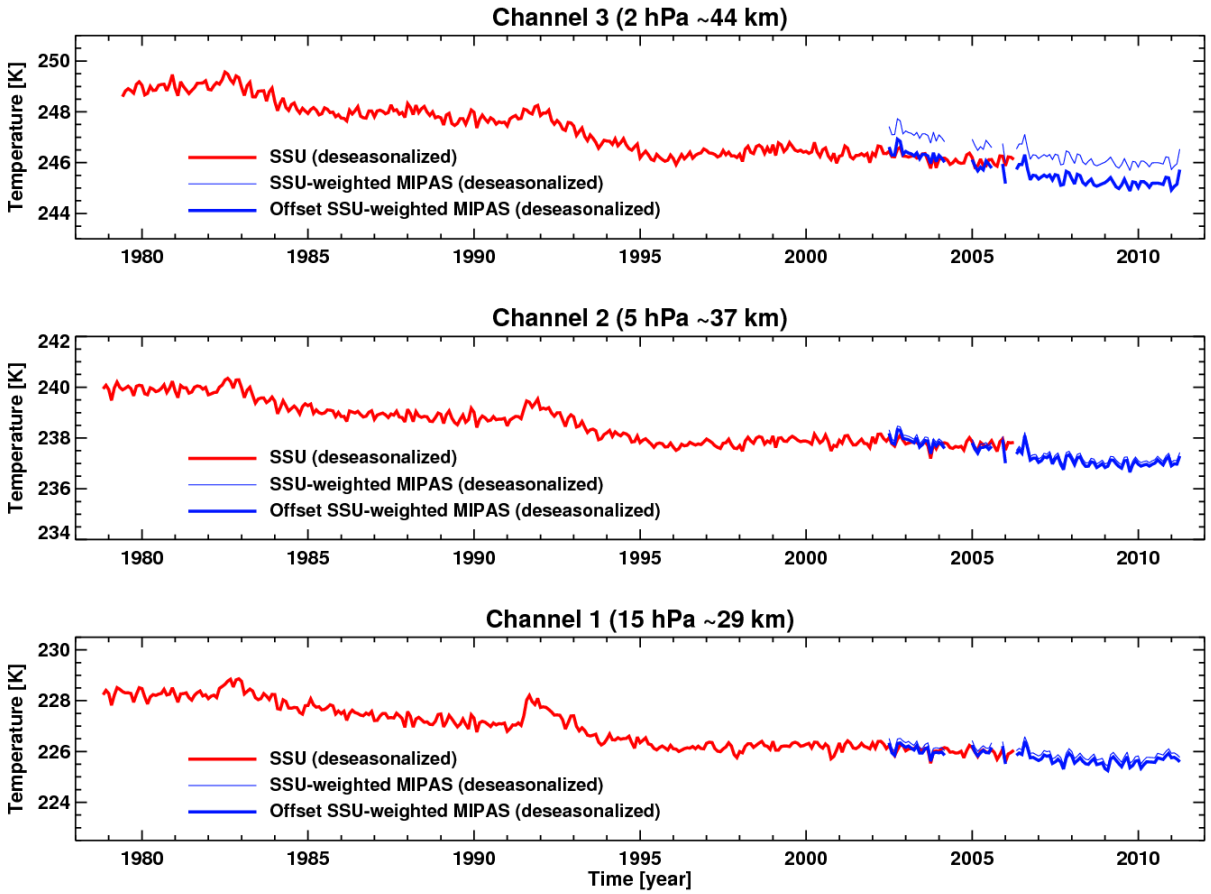

Figure 28: Deseasonalized near-global mean temperature time series for SSU (red) and SSU-weighted MIPAS (thin blue) for the three SSU channels. The thick blue curves denote the deseasonalized SSU-weighted MIPAS data offset by the mean difference between the thick blue and red curves in the overlap period (2002-2006).

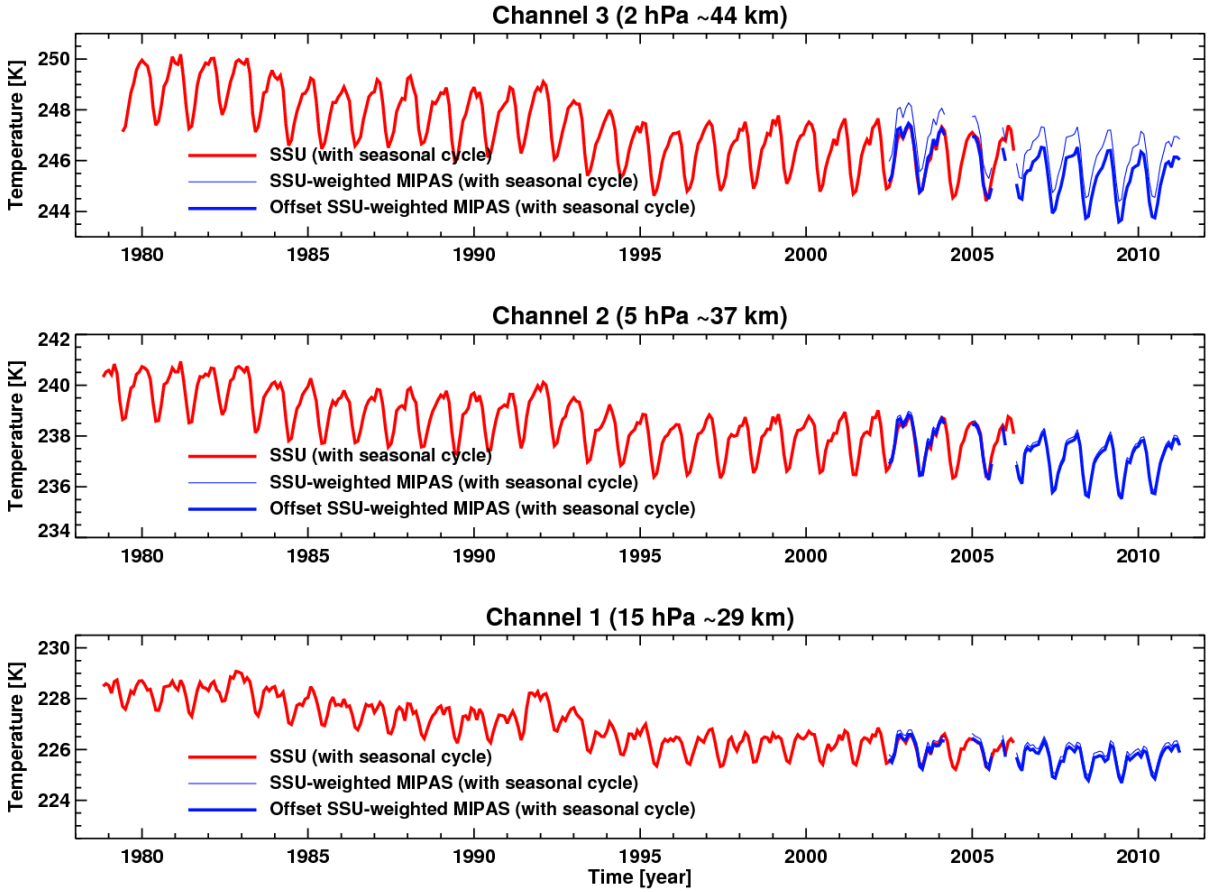

Figure 29: Same as Figure 28 but with seasonal cycle included. Note that the offsets used here were computed from the deseasonalized data.

In a traditional sense, no validation of these adjusted time series is possible since there are no suitable near-global correlative measurements that would be considered reliable climate data records. As has been shown above, CHAMP RO temperature is not suitable because the SSU-weighted CHAMP data is either too sensitive to filling and/or not sufficiently stable in time. These deficiencies likely result from the fact that the weighting functions for all three SSU channels project significantly into the upper stratosphere (see Figure 21), where the quality of the CHAMP temperatures is poor. Reanalyses cannot be used for validation since they assimilate SSU temperature. In-situ data, e.g. from lidars, do not provide global coverage. Therefore, the only way to validate the adjusted time series is to examine the robustness of the adjustment process itself.

To that end, further statistical analysis of the SSU-weighted MIPAS data and the SSU data in the overlap period is shown in Figure 30 and Figure 31. Figure 30 shows scatter plots of the deseasonalized data for the three channels (SSU on the horizontal axis and SSU-weighted MIPAS on the vertical axis). Note that the constant offsets computed above are used to shift the origin of the y-axis in each panel, thus centring the scatter of points in each panel. With the offset applied to the SSU-weighted MIPAS data, the monthly mean near-global data from the two records agree within 0.2 K for nearly all values, and within 0.1 K for most values. Since the random errors associated with each point cannot be easily ascertained, it would not be possible to compute the uncertainties of the two linear fit parameters if the data points were fitted to a straight line. However, the thin lines with a slope of one shown here appear visually

to provide a reasonable fit through the points. And there seems to be little justification from this analysis for performing anything other than a constant offset.

To examine the temporal stability of the offset, Figure 31 shows the Cumulative Sum (CUSUM) of the differences between the offset SSU-weighted MIPAS data and the SSU data for the overlap period. If the differences were random the CUSUM curves would jump around about zero. The fact that they do not means that there are systematic aspects (e.g. temporal drifts) in the differences. (The fact the CUSUM comes down to zero at the end of the record is ensured by the offset, which guarantees a zero time-averaged difference.) In particular, for Channels 2 and 3 CUSUM generally increases in the early part of the overlap period, before the data gap in 2004, and decreases in the later part. This suggests that there may be a systematic difference in MIPAS temperatures before and after the gap in 2004, with the MIPAS temperatures prior to 2004 being systematically higher than those after (assuming that SSU can be used as a transfer function). From the slope of the CUSUM, the relative high bias before the gap is approximately +0.1 K, and that after the gap -0.2 K, implying a relative bias of 0.3 K between the two periods. Although the issue could potentially lie with SSU, one likely candidate to explain this behaviour is the switch in the spectral and vertical resolution of the MIPAS retrieval in January 2005. For Channel 1, there is some temporal structure in the CUSUM prior to the gap but no apparent difference between the pre- and post-2004 portions of the record, suggesting that if there is an impact of the change in retrieval on MIPAS temperatures, it lies in the higher altitudes which contribute less to Channel 1.

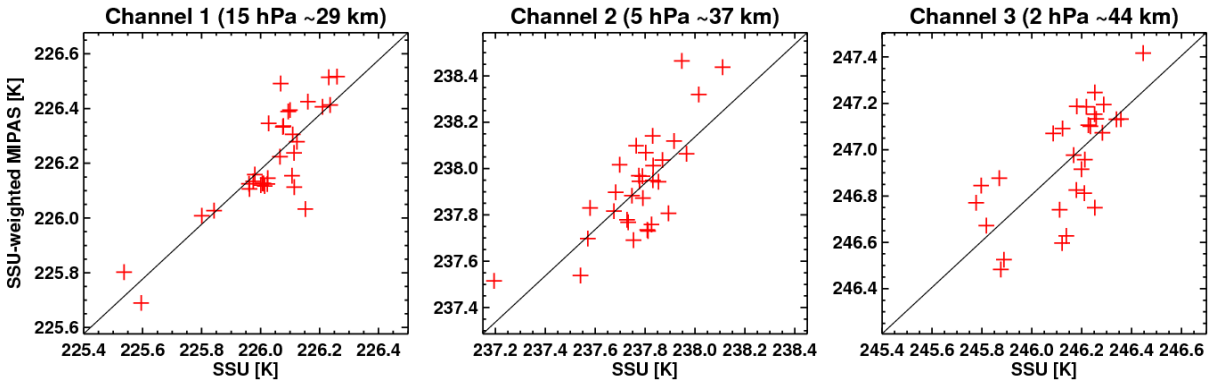

Figure 30: Deseasonalized near-global mean SSU-weighted MIPAS temperatures versus the SSU temperatures for the overlap period (2002-2006) for the three SSU channels. A line with a slope of one is shown in each panel to guide the eye. To center the set of points in each panel, the origin of the y-axes is shifted by the constant offsets computed from the data.

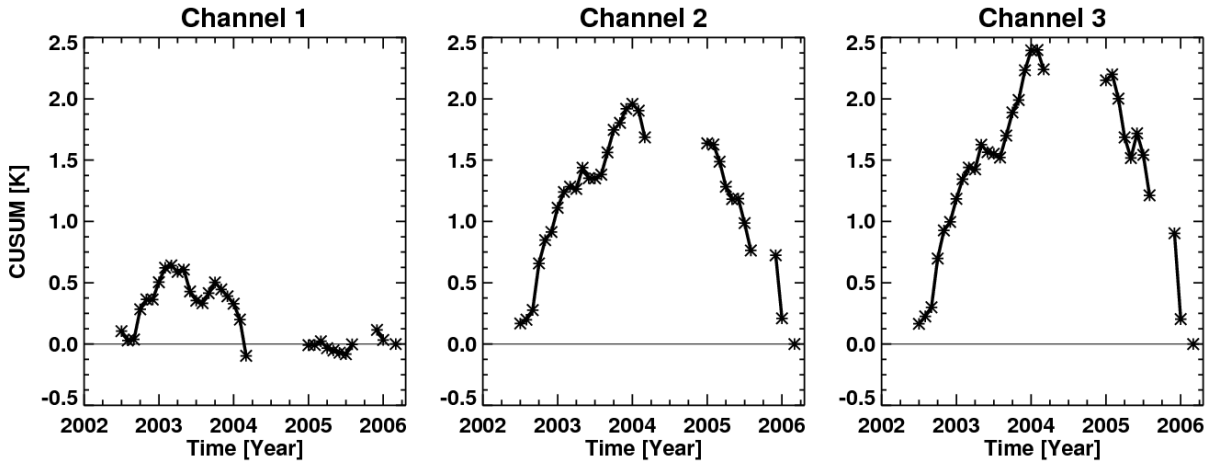

Figure 31: Cumulative sum of the differences between the deseasonalized near-global mean SSU-weighted MIPAS temperatures (with the offsets applied) and SSU temperatures for the overlap period (2002-2006) for the three SSU channels. The gaps at the beginning and middle of the curves denote the times when there were no MIPAS data. The asterisks denote the actual data points.

Figure 32 shows the final merged time series. The SSU results (red) are identical to those shown in Figure 29. For the SSU-weighted MIPAS results with the offsets applied (blue) only the data after the end of the SSU record are shown; with this exception the blue curves shown here are identical to the thick blue curves shown in Figure 29.

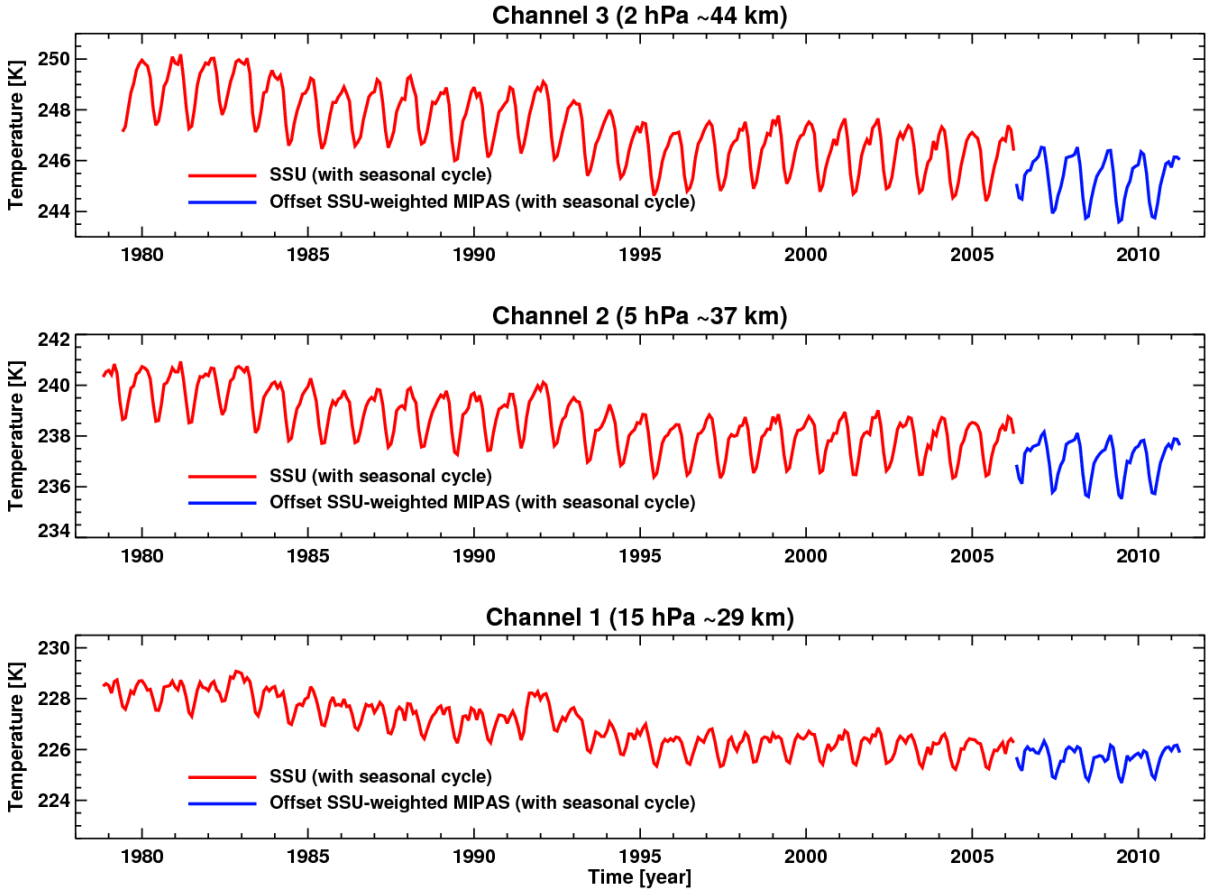

Figure 32: Near-global mean temperature with seasonal cycle included for SSU (red) and the offset SSU-weighted MIPAS (blue) for the three SSU channels. Note that the MIPAS data in the overlap period is not shown.

4.1.2. Validation of BSUTLS temperature record

The Bodeker Scientific Upper Troposphere Lower Stratosphere (BSUTLS) temperature database generated in WP24 was validated in three different ways, viz.:

1. Against the radiosonde-based RATPAC-A (Free et al., 2005) database.
2. Against a COSMIC (Anthes et al., 2008) radio occultation zonal mean monthly mean temperature data set.
3. Against a NCEP CFSR monthly mean zonal mean temperature data set.
4. The comparison of the vertically integrated BSUTLS temperature data set against the merged MSU4+AMSU9 data set itself provides validation of the BSUTLS data set.

Each of these validation exercises is described in separate sub-sections below.

4.1.2.1. Validation of BSUTLS temperature record

The RATPAC-A data set is provided on pressure levels and is aggregated over large geographical regions. Because RATPAC-A data are only provided as anomalies, and because the period over which the anomalies are calculated could not be ascertained from the available RATPAC-A documentation (nor from e.g. Dian Seidel – personal communication), the RATPAC-A anomalies were first subtracted from the BSUTLS time series at each pressure level and then the mean of the resultant time series was subtracted. If the interannual variability between the BSUTLS and RATPAC-A data sets is identical, the resultant time series would be uniformly zero i.e. we compared the ability of the two databases to track interannual variability in the temperature signals. Figure 33 shows the comparison of the interannual variability for the northern and southern hemispheres as a function of pressure and time. At 250 hPa and at 300 hPa differences in interannual variability can exceed 1 K while at lower pressures, differences are generally less than ±0.5 K.

Differences in the interannual variability are smaller on all pressure levels in the tropics (Figure 34). In extra-tropical regions (Figure 35), the anomalies are in general also small at all pressure levels, but can reach up to ±1.5K at low pressure levels of 30 to 50 hPa in 2004/2005 as well as at the high pressure level of 300 hPa in 2011. Similar comparisons are also available for the whole globe and for 20°S to 20°N (not shown here). The ability of the BSUTLS data set to track year-to-year variability over most of the UTLS at the ±0.5 K level indicates that this database includes a valid representative of interannual variability in temperatures over this region.

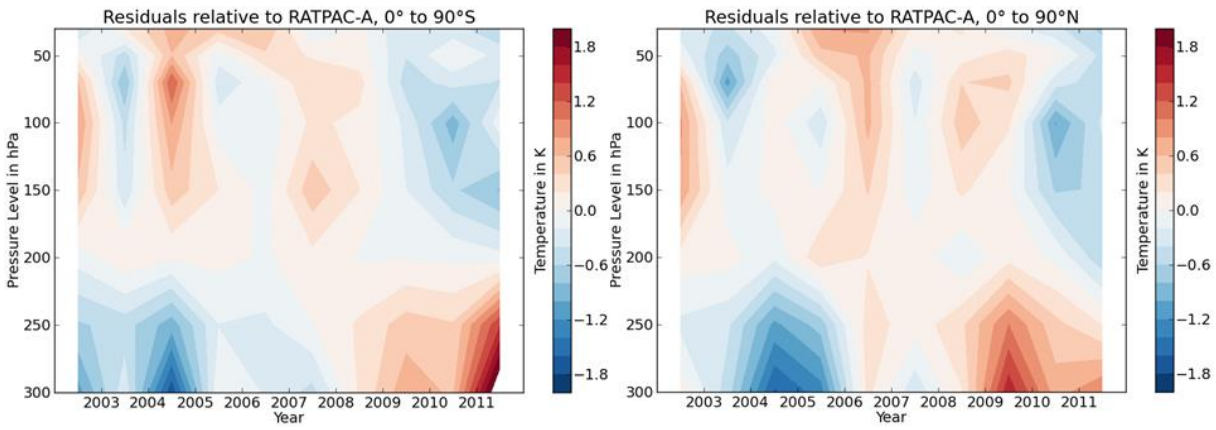

Figure 33: Differences between BSUTLS temperature anomalies and those from the RATPAC-A database. The leftmost panel shows differences for the southern hemisphere while the rightmost panel shows differences for the northern hemisphere.

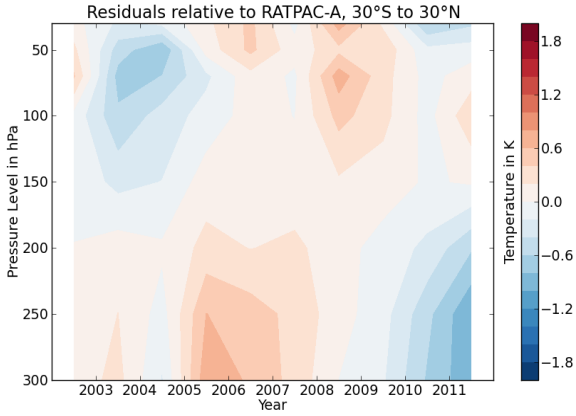

Figure 34: Differences between BSUTLS temperature anomalies and those from the RATPAC-A database in the tropics.

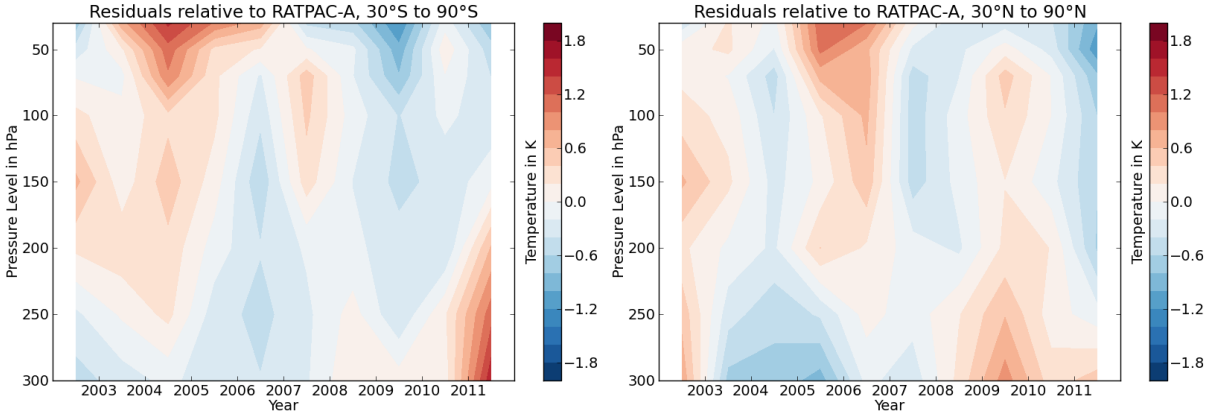

Figure 35: Differences between BSUTLS temperature anomalies and those from the RATPAC-A database in extra-tropics. The leftmost panel shows differences for the southern hemisphere while the rightmost panel shows differences for the northern hemisphere.

4.1.2.2. Validation of the BSUTLS data set against a COSMIC radio occultation data set

The COSMIC radio occultation (RO) data set used for this validation was created in exactly the same way as the RO data sets created in the first part of WP24 and used in the construction of the BSUTLS data set. The COSMIC RO data set was not used in the construction of the BSUTLS data set and was rather retained as a means of validating the BSUTLS data set. Comparisons between the BSUTLS and COSMIC RO data sets were made at all 17 pressure levels for which the databases are available. Examples of comparisons at two different pressure levels are shown in Figure 36. At 10 hPa the BSUTLS temperatures tend to be up to 4 K cooler than the COSMIC temperatures south of ~40°S and up to 4 K warmer than the COSMIC temperatures north of ~40°S. At 100 hPa the differences are much smaller, typically within 1 K, with BSUTLS temperatures being cooler than COSMIC temperatures at higher latitudes and warmer than COSMIC temperatures at lower latitudes.

The latitude zone over the tropics where BSUTLS temperatures are higher than COSMIC diminishes in width with increasing pressure and at around 200 hPa the pattern reverses, i.e.

BSUTLS temperatures in the extra-tropics are warmer than COSMIC and cooler in the tropics (see Figure 37).

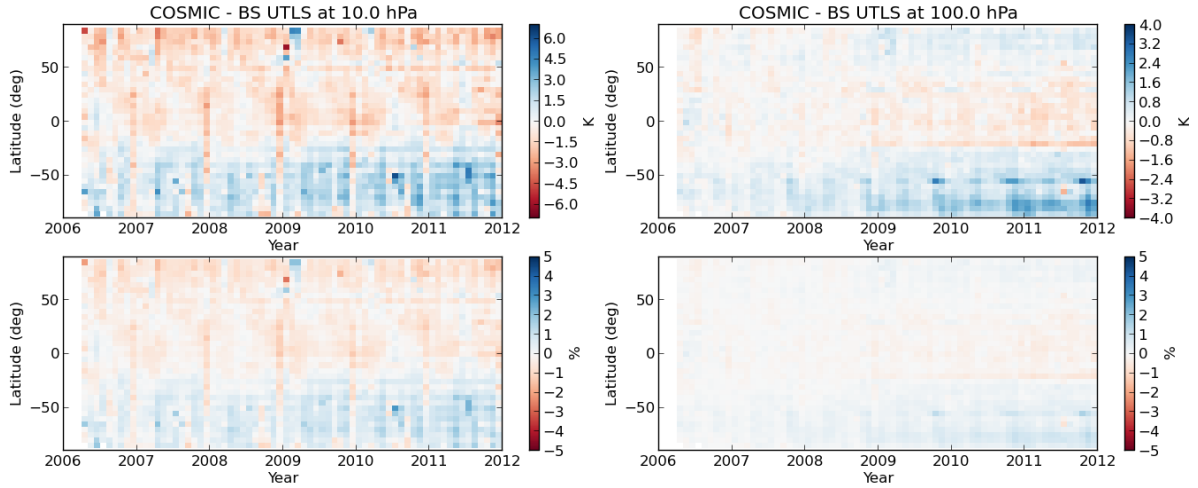

Figure 36: Differences between the BSUTLS and COSMIC RO data sets at 10 hPa (leftmost panels) and at 100 hPa (rightmost panels). Upper panels show the differences in absolute units (K) while the lower panels show the percentage differences.

Such plots are available for 300,250, 200, 170, 150, 130, 115, 100, 90, 80, 70, 50, 30, 20, 15, 10 and 7 hPa.

4.1.2.3. Validation of the BSUTLS data set against NCEP CFSR data

NCEP CFSR data were linearly interpolated onto the log pressure levels of the BSUTLS data set and then area weighted monthly mean zonal means were calculated for all pressure levels.

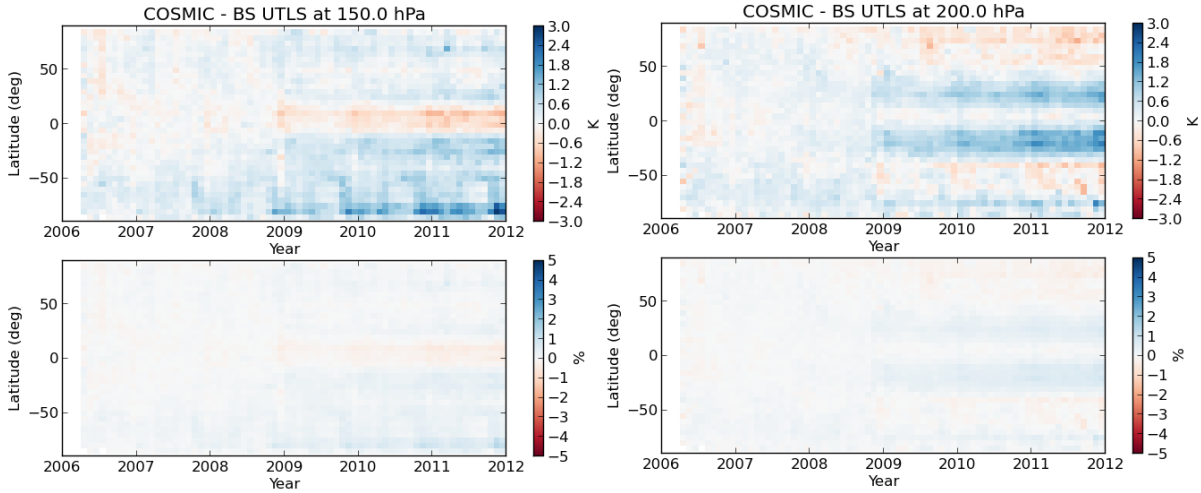

Figure 37: Differences between the BSUTLS and COSMIC RO data sets at 150 hPa (leftmost panels) and at 200 hPa (rightmost panels). Upper panels show the differences in absolute units (K) while the lower panels show the percentage differences.

Figure 38 shows calculated differences between the BSUTLS temperature data set and the NCEP CFSR data set. The BSUTLS data set can be up to 2 K warmer that NCEP CFSR at 10 hPa with a larger bias in the northern hemisphere than in the southern hemisphere. At 90 hPa, differences over the tropics are close to zero but with a clear seasonal cycle. These differences

grow to 1.0-1.5 K over higher latitudes. At 250 hPa the differences vary between 0.0 and 1.5K.

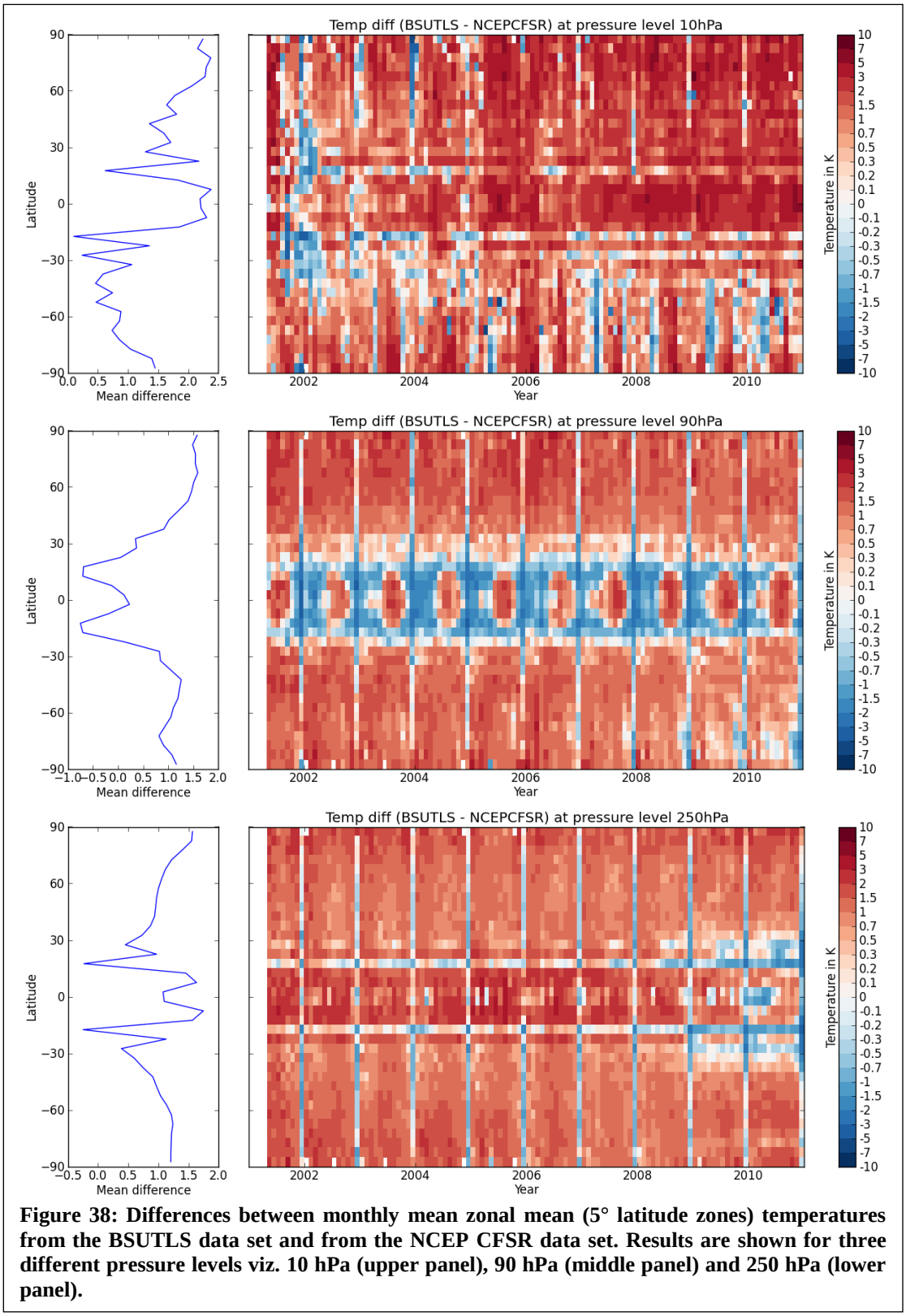

Figure 38: Differences between monthly mean zonal mean (5° latitude zones) temperatures from the BSUTLS data set and from the NCEP CFSR data set. Results are shown for three different pressure levels viz. 10 hPa (upper panel), 90 hPa (middle panel) and 250 hPa (lower panel).

There is no discernible trend in the difference fields. These results suggest that the use of the ESA (TPM) data together with the RO temperature data in the construction of the BSUTLS data set results in a high quality temperature time series that, with sufficient extension, is

likely to be suitable for analyses of temperature trends in the upper troposphere and lower stratosphere.

4.1.2.4. Validation of the BSUTLS data set against merged MSU4 and AMSU9 data

The BSUTLS data set was used to create a proxy for the merged MSU4+AMSU9 data sets. The a weighting function representative of the MSU channel 4 vertical weighting function and the AMSU channel 9 off-nadir weighting function was to calculate weighted means of the BSUTLS data set as described in the ATBD. Values were only considered to be valid if data from all 17 pressure levels are available.

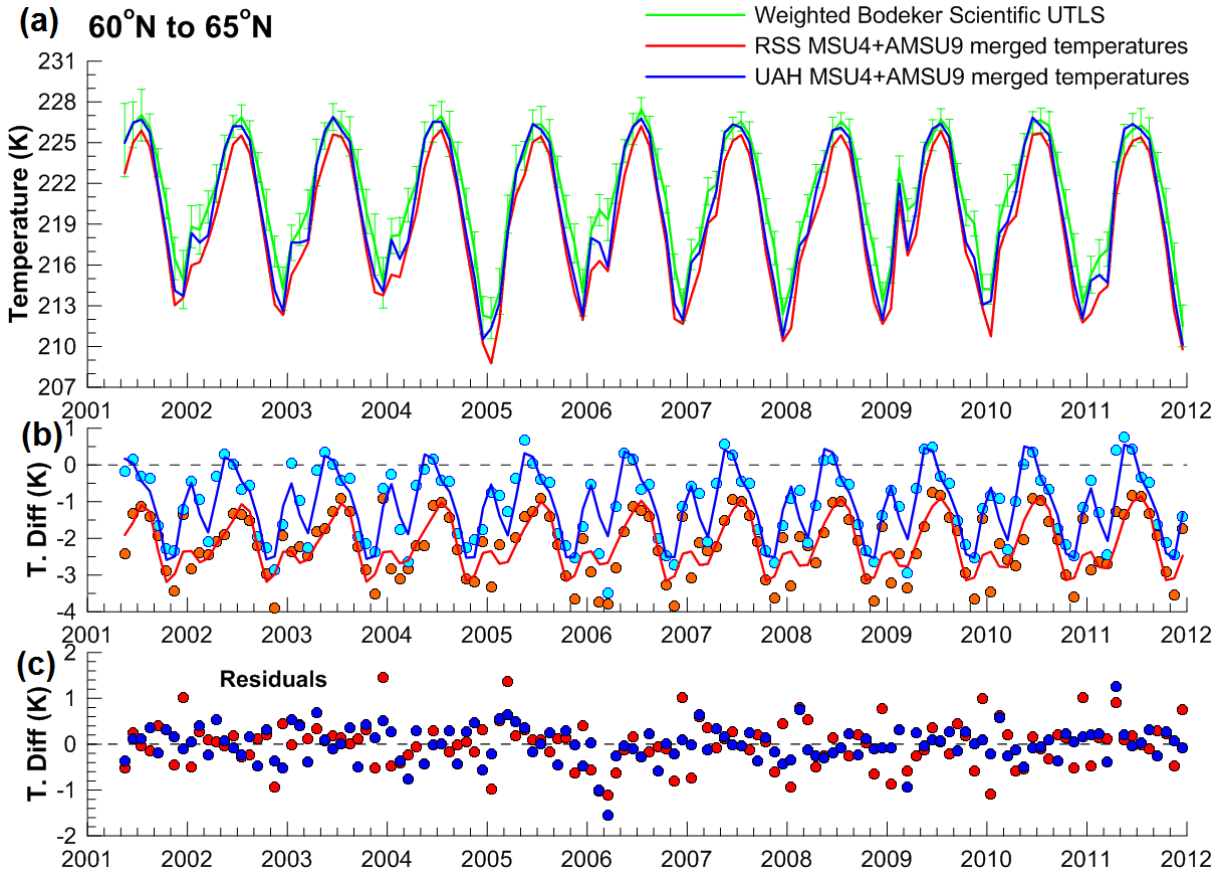

Figure 39: An example of the comparison of the weighted Bodeker Scientific (BS) UTLS data set and the two merged MSU4+AMSU9 data sets. (a) the original raw monthly mean time series, (b) regression model fits (lines) to the differences between the RSS and BS time series (cyan dots and blue line showing the regression model fit) and between the UAH and BS time series (orange dots and red line showing the regression model fit), (c) the residuals from the regression model fits shown in panel (b).

The vertically integrated BS data set was then compared against the two merged MSU4+AMSU9 data sets (available from the UAH and RSS groups) in each 5° latitude zone (see Figure 39 for an example). The regression model fit in panel (b) of Figure 39 is designed to model any systematic biases between the BS vertically integrated database and the merged MSU4+AMSU9 databases. Any steps in the residuals from those regression model fits would be indicative of problems either with the merging of the MSU4 and AMSU9 time series or with the construction of the BSUTLS data set. No such steps are seen in panel (c) of Figure 39 nor in any of the 35 other similar figures for the other latitude zones.

4.2. Stratospheric Water Vapour

The comparisons are done following the method used by Dupuy et al. (2009).
For all comparisons the mean percentage difference between the SCIAMACHY profiles and the other instrument is calculated for all altitudes (z) with:

$$\Delta z = \frac{1}{N(z)} \sum_{i=1}^{N(z)} \delta_i(z)$$

Where $N(z)$ is the number of pairs and $\delta_i(z)$ are all percentage differences $\delta(z)$ between the water vapour profile from SCIAMACHY (x_{SCIA}) and the instrument used for comparison (x_{comp}):

$$\delta_i(z) = 100 * (x_{SCIA}(z) - x_{comp}(z))/x_{ref}(z)$$

The reference x_{ref} is the average of the two profiles:

$$x_{ref}(z) = \left(x_{SCIA}(z) + x_{comp}(z)\right) * 0.5$$

x_{comp} is either interpolated on the retrieval grid of SCIAMACHY (black lines in the following figures) or smoothed with the Averaging Kernel (AVK) of SCIAMACHY (green lines).
The standard deviation of the bias corrected difference is calculated with:

$$\sigma(z) = \frac{1}{N(z) - 1} \sqrt{\sum_{i=1}^{N(z)} (\delta_i(z) - \Delta z)^2}$$

and the standard error of the mean, which is often not visible because it is smaller than the width of the lines with:

$$SEM(z) = \frac{\sigma(z)}{\sqrt{N(z)}}$$

Figure 40 shows the mean percentage differences, the standard deviation of the bias corrected difference and the standard error of the mean (only visible in the right panel) for SCIAMACHY V3.01 and the CFH data. In the right panel the results for all collocations to the processed data set (every 8th day globally, every 2nd day between 45°S and 45°N) are shown the left panel includes all possible collocations. For this purpose the retrieval was run additionally to the processing for all collocations found, the MPV criterion in not applied for these data. The data agree well below about 20km altitude. Here, the difference is smaller than 10% if the CFH data are smoothed with the AVK and up to about 20% for the interpolated version, where the differences are larger at 13-14 km. In this altitude range, the vertical gradient of the water vapour profile is often large, while the vertical sampling of SCIAMACHY is coarse: SCIAMACHY measurements are at about 12.0, 15.3, 18.6, 21.9, and 25.2km, and therefore there is no measurement information at 14km and the values for the retrieval grid at these altitudes are derived from the measurements below and above. They are not purely interpolated but rather a result of the regularization, which takes the profile shape of the *a priori* profile into account. The assumption of the shape is more uncertain for 14 km, than for example for 17 km, because 14 km is often in the troposphere, where the variability of water vapour and its profile shapes is higher than in the stratosphere. This leads to the larger difference between the comparisons based on interpolated or on smoothed CFH values at 14 km.

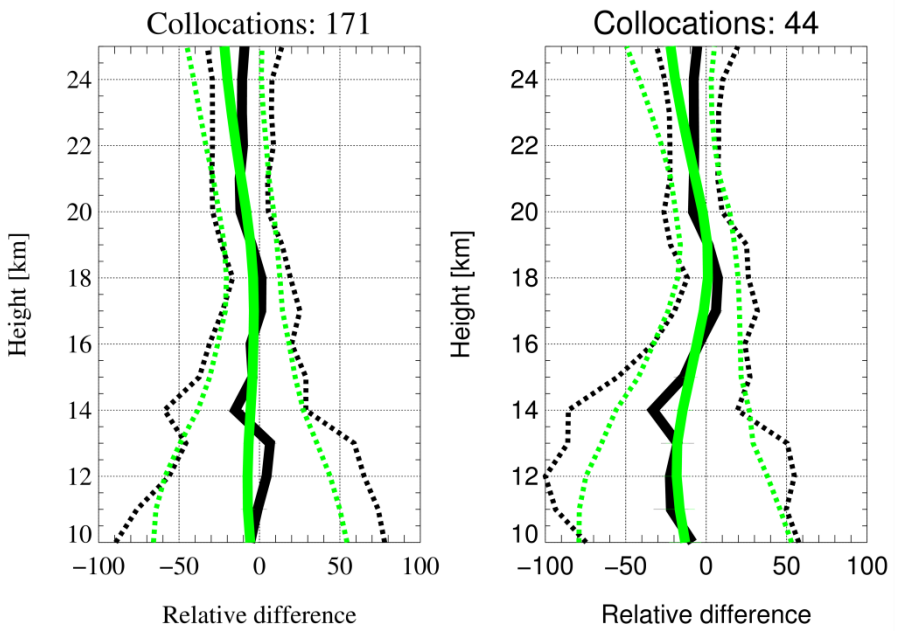

Figure 40: Comparison for SCIAMACHY V3.01 and balloon-borne CFH. Left: all possible collocations, right: collocations to the processed data set (every 8th day globally, every 2nd day between 45°S and 45°N). Black solid lines show the mean percentage difference for CFH data interpolated on the SCIAMACHY retrieval grid, black dotted lines the corresponding standard deviation of the bias corrected difference. Green lines show the mean percentage difference for CFH data smoothed with the SCIAMACHY AVK (solid) and the corresponding standard deviation of the bias corrected difference (dotted). In the right panel the standard error of the mean is indicated as error bars, in the left panel it is smaller than the width of the line and hence not visible.

Above about 19 or 20 km the SCIAMACHY data are 10-20% lower than the CFH data.

For the data, which are part of the already processed V3.01, the differences between SCIAMACHY and CFH are larger than for all possible collocations below 20 km: up to about 20% for the smoothed and 30% for the interpolated CFH data. Hence, the result depends on the availability of collocations and it would be better to regard a larger data set. Therefore, a comparison to two different satellite data sets is also shown. For the satellite comparisons, only data processed for V3.01 (every 8th day globally, every 2nd day between 45°S and 45°N) are used. Due to the small number of collocations, the comparison for the CFH is shown averaged for all available data. The comparisons to the satellite data are split in the latitude band 60°N-90°N, 30°N-60°N, 30°S-30°N, 30°S-60°S, and 60°S-90°S to investigate if the differences vary with latitude.

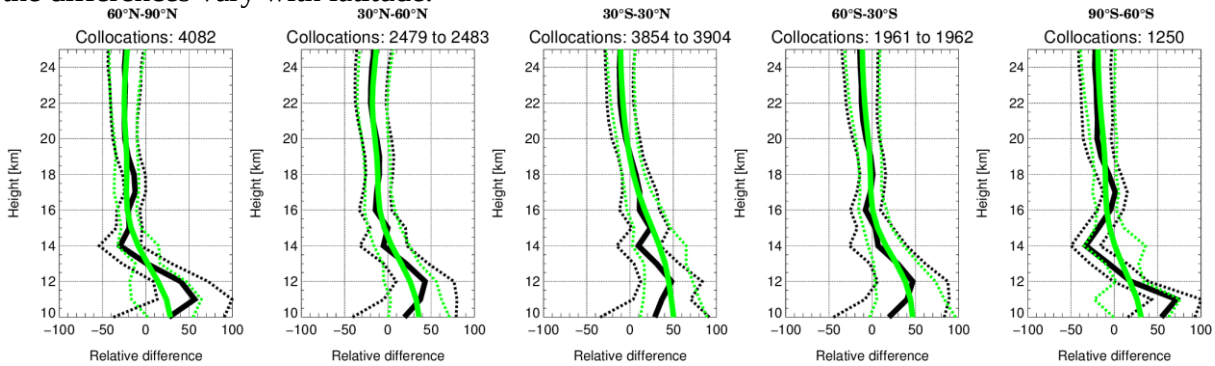

Figure 41: Comparison for SCIAMACHY V3.01 and MLS v3.3 for different latitude bands. Lines as in Figure 40.

Figure 41 shows the comparison between SCIAMACHY V3.01 and MLS v3.3.

Compared to MLS v3.3, for all regions, SCIAMACHY V3.01 is has lower water vapour values in the stratosphere and higher ones in the troposphere. The altitude, where it changes from

negative to positive differences varies with region: 12-14km for polar latitudes, 13-16 for mid-latitudes and 18-19 for the tropics, i.e. always somewhat above the average tropopause for the corresponding region. In the stratosphere, the SCIAMACHY data are lower by 10% (tropics) to 25% (northern hemisphere high latitudes). The difference in the troposphere is up to about 50% (tropics, southern hemisphere mid-latitudes). As for the CFH data, there are differences depending on if the MLS data are smoothed with the SCIAMACHY AVK or interpolated, although the resolution of MLS is comparable to the one of SCIAMACHY. These differences are largest in the troposphere, at altitudes where no SCIAMACHY measurements are available and the profile shape differs from the a priori profile.

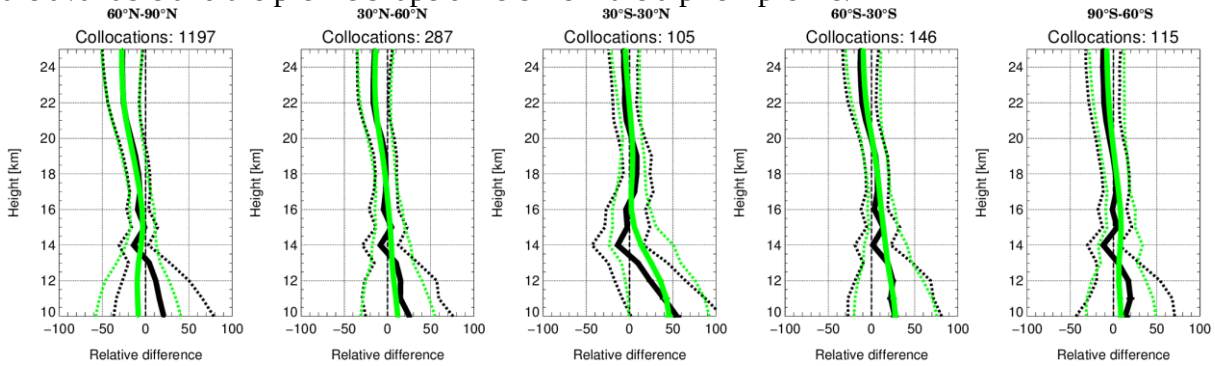

Figure 42: Comparison for SCIAMACHY V3.01 and ACE-FTS V3.0 for different latitude bands. Lines as in Figure 40.

Figure 42 shows the results for the comparison with ACE-FTS. For the stratosphere, the results are comparable to the ones seen for MLS, confirming that there is probably a systematic dry bias in the SCIAMACHY data, which is most pronounced in the Northern Hemisphere high latitudes. For the troposphere, the results are different: In the tropics, the SCIAMACHY data are also higher than ACE-FTS by up to 40% but the agreement is better in the mid-latitudes, where SCIAMACHY data are higher by 10-20% and in the Polar Regions. For all regions the agreement between SCIAMACHY and ACE-FTS is good (differences of 10% or smaller) between about 15 and 19 km altitude. The differences between smoothing the ACE-FTS data with the SCAMACHY AVK and interpolation them onto the retrieval grid confirm the effect of *a priori* profile shape and regularization seen in the other comparisons.

Summarizing the comparisons, SCIAMACHY V3.01 shows a dry bias of about 10-20% in the lower stratosphere, which is most pronounced in the northern hemisphere high latitudes. Predominantly around 14 km altitude the profile shape of the SCIAMACHY water vapour is influenced by the profile shape of the *a priori* profile due to the vertical sampling and the regularization in the retrieval. Compared to MLS, SCIAMACHY shows higher values (up to about 50%) in the troposphere. This is not clearly confirmed by the comparison to ACE-FTS and CFH data but SCIAMACH V3.01 shows also higher values in the troposphere compared to other satellite data (not shown here). Generally, comparisons are challenging in the troposphere due to the high variability of water vapour.

4.3. Stratospheric Aerosols

4.3.1. OSIRIS

4.3.1.1. SAGE II Extinction Comparison

Interpolation of the SAGE II 525 and 1020 nm channels in logarithmic space (log wavelength and log extinction) provides a 750 nm extinction product based on the Angstrom coefficient between 525 and 1020 nm. This derived product is shown in the bottom of Figure 43 for the duration of the SAGE II mission between 20°N and 20°S. The comparative 750 nm OSIRIS

extinction ratio is shown in the top panel. Although exact agreement is not expected since a linear Angstrom coefficient (in log space) is only an approximation, both data sets agree quite well, typically within 20%. The largest discrepancies occur at low altitudes in the volcanic plumes caused by the eruptions of Ruang/Reventador and Mt. Manam. This is likely due to the infrared imager saturating in these conditions which tends to bias the measurements to low aerosol loading conditions. Correction of this is discussed in the recommendations section.

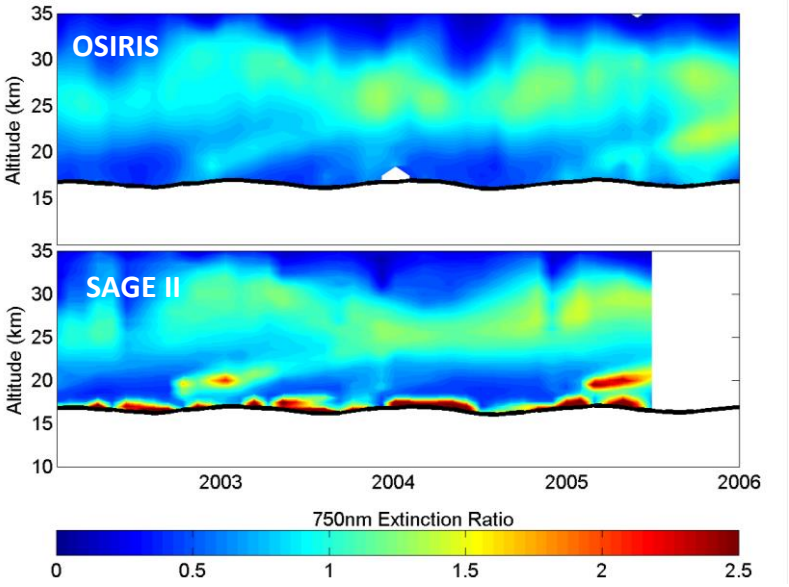

Figure 43: Monthly averaged 750 nm extinction ratio in the tropics. OSIRIS measurements shown in the top panel with SAGE II in the bottom panel.

4.3.1.2. SAGE II Angstrom Comparison

The OSIRIS Angstrom coefficient from 750 to 1530 nm is shown in the top panel of Figure 44. The SAGE II Angstrom coefficient calculated from the 525 and 1020 nm extinction measurements is shown in the bottom panel. In general the OSIRIS Angstrom exponent is larger than the SAGE II measurements, although much of this can be attributed to the wavelength discrepancy of the two instruments. Qualitatively the measurements show many similarities, with a particle size that generally decreases with altitude. In addition, the volcanic eruptions both show a decrease in particle size. The largest difference in the measurements is the six month cycle in the OSIRIS Angstrom data that is not present in the SAGE II measurements. This is likely due to the uncertainty of the 1530 nm albedo used in the OSIRIS retrievals (Bourassa, 2012), which when combined with the Odin orbit produces the 6 month cycle.

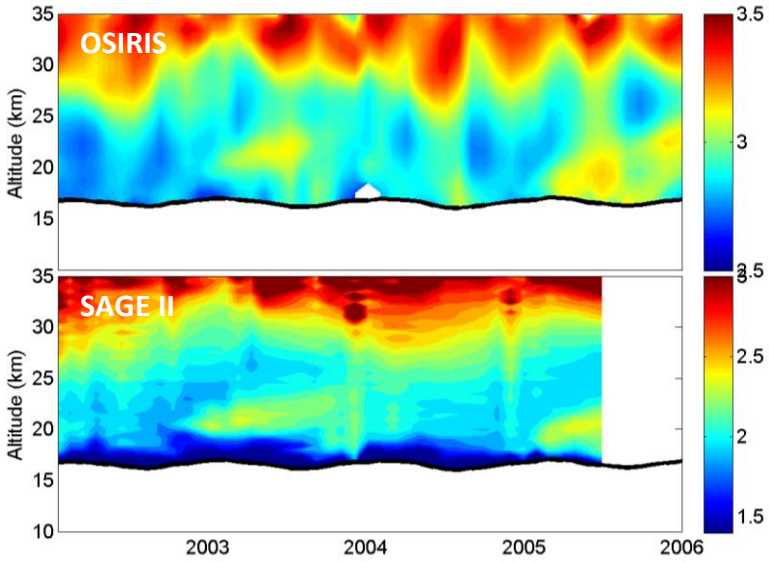

Figure 44: Top panel is the monthly OSIRIS Angstrom coefficient from 750 to 1530 nm from 20°N to 20°S. Bottom panel is the same picture with the SAGE II Angstrom coefficient from 525 to 1020 nm.

4.3.1.3. SAGE II Surface Area Density Comparison

Surface area density can be derived from OSIRIS measurements using the retrieved particle size distribution and extinction. SAGE II also retrieves surface area density, although does so without the assumption of a particle size distribution. These are compared in Figure 45 and Figure 44. The top panel shows the retrieved OSIRIS values, with SAGE II results in the bottom panel. Both data sets show similar features with an increase in surface area density after volcanic eruptions, particularly Mt. Manam. OSIRIS tends to retrieve lower SAD values at higher altitudes. This may suggest OSIRIS is underestimating SAD in low aerosol loading conditions, as SAGE II is expected to underestimate SAD as well, due to the insensitivity to small particles (Wurl, 2010).

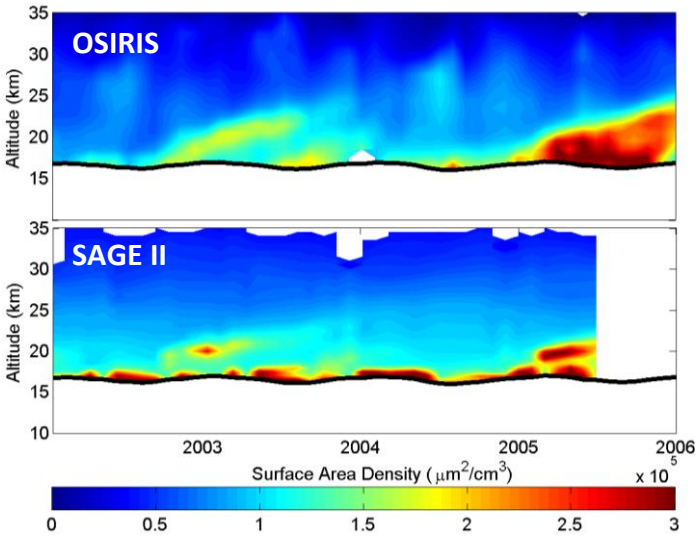

Figure 45: Monthly averaged surface area density from OSIRIS in the top panel and SAGE II in the bottom panel.

4.3.1.4. SAGE III Extinction Comparison

The average per cent difference between the SAGE III and the OSIRIS measurements are shown in Figure 46 for each year from 2002 to 2005. The Version 5.07 results are shown in red and are typically within 10% of the SAGE III measurements, although a systematic bias is evident in 2005 when OSIRIS consistently overestimates the SAGE III measurements by 10-20%. The Version 6.00 retrievals are shown as black lines in Figure 46. These results are comparable at all altitudes, with a slight improvement during 2005, but also slight underestimation in 2003 and 2004 between approximately 23 and 30 km. In general, substantial improvement between comparisons is not expected as the SAGE III measurements are during conditions when the Version 5.07 particle size assumptions are expected to be relatively accurate.

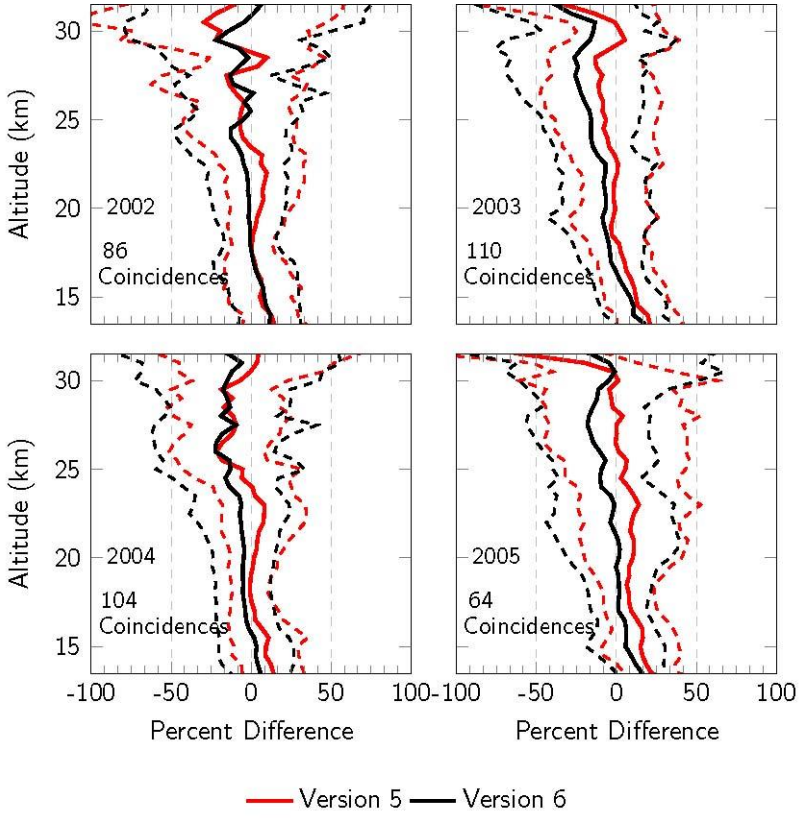

Figure 46: Comparison of SAGE III and OSIRIS aerosol retrievals for yearly coincident measurements. Mean per cent differences are shown as solid lines with standard deviation shown as dashed. Version 5.07 retrievals are shown in red with Version 6.00 shown in black.

4.3.2. SCIAMACHY

Figure 47 to Figure 49 show the validation results of the SAGE II and SCIAMACHY stratospheric aerosol extinction profiles at 525 nm wavelength. The profiles were averaged zonally and over all available co-locations as well as binned into eight 20 latitude bins spanning 80°N and 80°S.

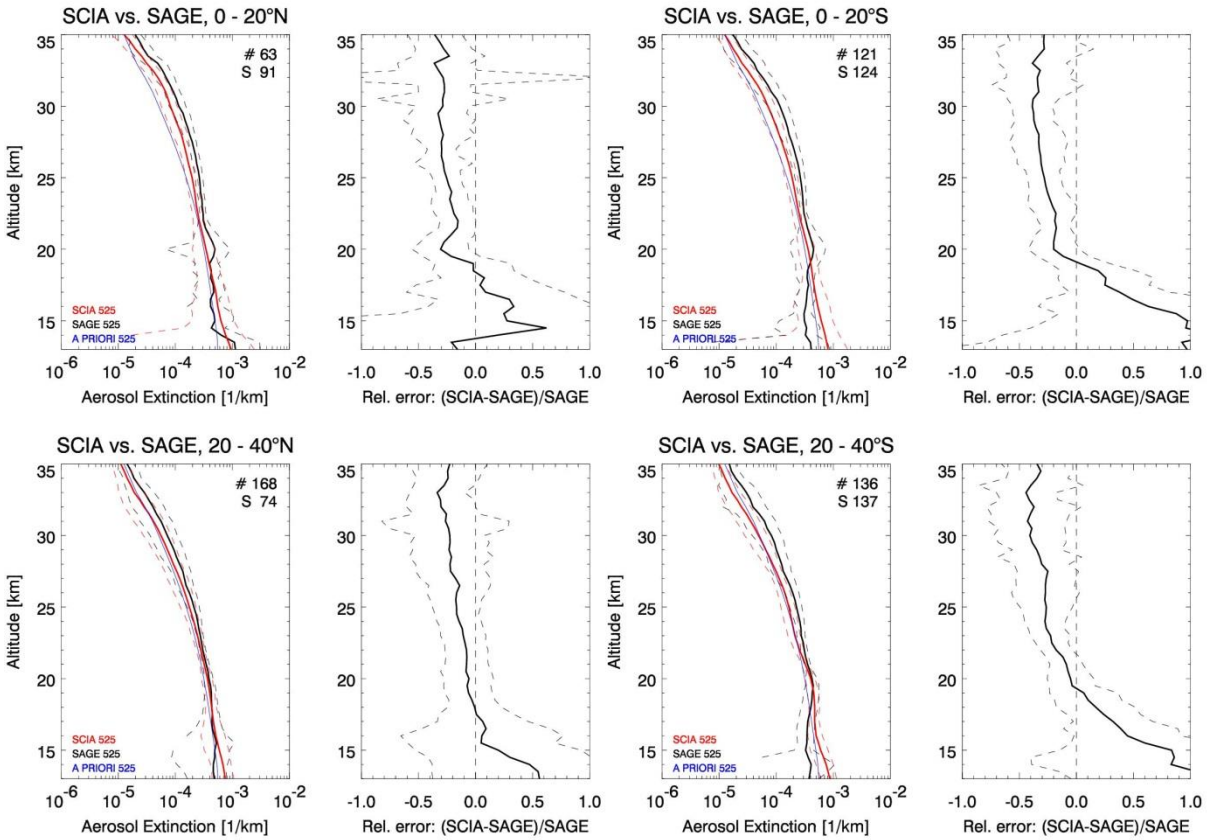

Figure 47: Validation results for the latitude bins: 0-20°N, 0-20°S, 20-40°N and 20-40°S. Left panels: comparison of the retrieved 525 nm aerosol extinction profiles (red) with SAGE II aerosol extinction (black) and a priori profile (blue) in 8 latitude bins with standard deviation (dashed lines). Right panels: mean relative difference between SCIAMACHY and SAGE II aerosol extinction profiles with standard deviation (dashed).

The panels of Figure 47 show the four bins between 40°N and 40°S. For the two northern latitude bins (0-20°N and 20-40°N), the relative difference above the tropopause at ~ 20 km altitude is around -30%. The same appears for the two southern bins (0-20°S and 20-40°S), except for altitudes at 28 km and above. Here the relative difference grows to -40%.
For the latitude bins 40-60°N and 40-60°S (Fig. 2), the difference to SAGE II for altitudes at 15 to 27 km lies between -10 and -20%. Again at higher altitudes, the difference for 40-60°N turns to a maximum value of -40% (at 31 km), while for 40-60°S it stays quite stable until 32 km. Between 60° and 80°N and at 15–23 km, the difference to SAGE II is -15 to +10%, while above the difference increases to a maximum of -50% (Fig. 2). Between 60° and 80°N and at 15–27 km, the difference to SAGE II is smaller or equal to -10%. At 27-32 km it gets worse and grows to +30% at 32 km.

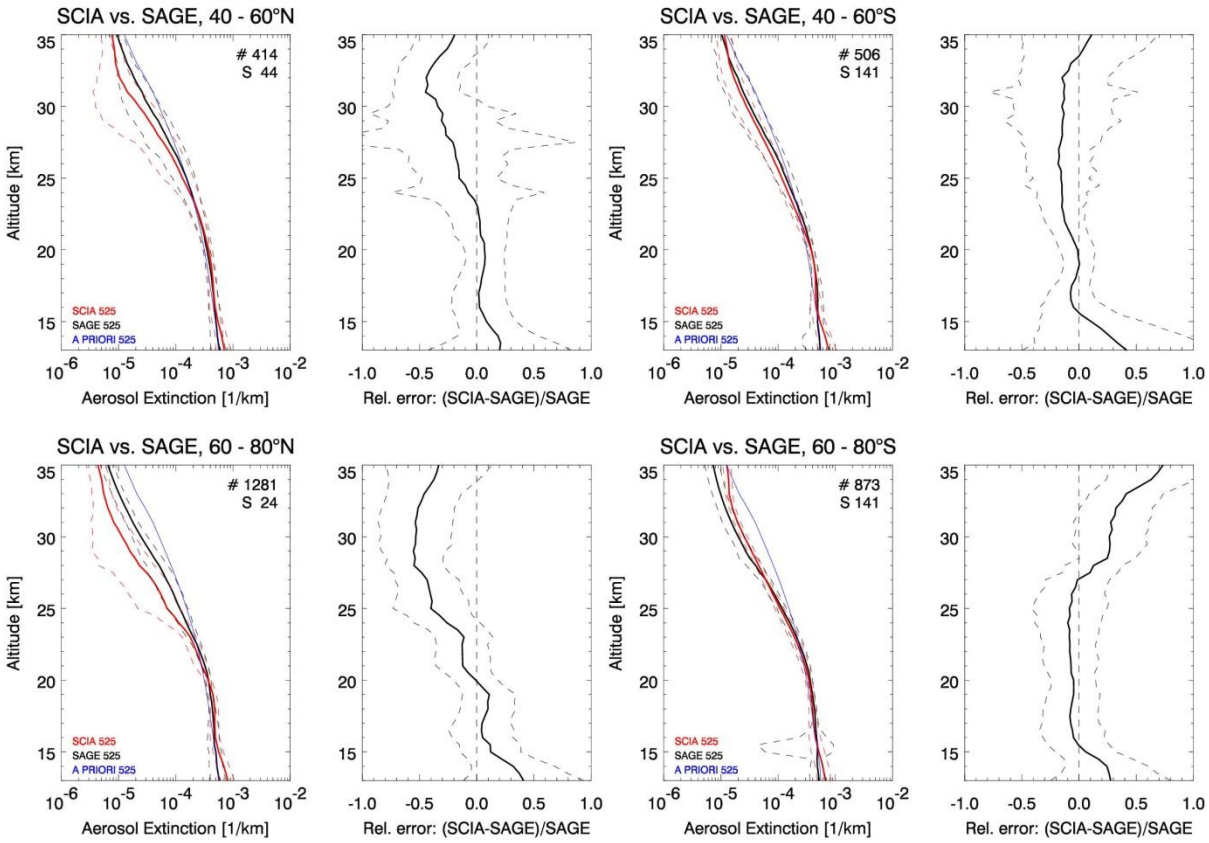

Figure 48: Same as Figure 47, but for the latitude bins: 40-60°N, 40-60°S, 60-80°N and 60-80°S.

To summarize, Figure 49 shows the validation results between SAGE II and SCIAMACHY for all latitudes between 80°N and 80°S. The global average, which contains 3562 co-locations, results in a difference between +10% and -10% at 15-23 km and shows that SCIAMACHY underestimates at 23-32 km with a maximum difference of -30%.

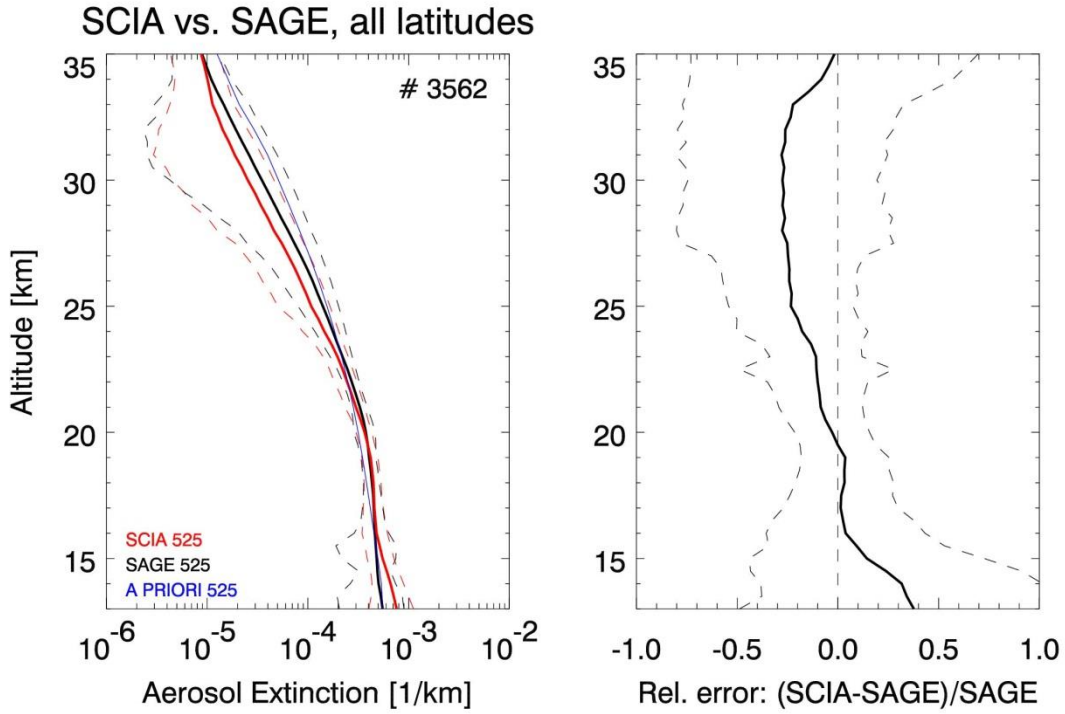

Figure 49: Same as Figure 47, but for all latitudes from 80°N to 80°S.

4.3.3. Statistical comparison between GOMOS, SCIAMACHY and OSIRIS Aerosol Extinction Data

This comparison is based on monthly zonal means. The selection and analysis procedures to obtain the average values of aerosol extinction are described in section 3.3 of the ATBD, as well as the results of a preliminary analysis of each individual data set. The analysis of the individual data sets has shown that the overall variance of the data in the 20-35 km region is typically relatively small when appropriate quality cuts are applied. Furthermore, a large part of the variance is driven by the oscillations generated by the GOMOS and SCIAMACHY retrieval algorithms. The natural variability (at some spatial and temporal distance to major aerosol events) should therefore be even smaller than the observed one. Sampling issues should thus not play a significant role for monthly and zonal averages. An exception may be PSC seasons and latitudes, where the spatial and temporal sampling differences between each instrument become relevant.

The comparison between all three data sets has been performed for the month of June 2003 and the aerosol extinction at 750 nm. GOMOS and SCIAMACHY data only can also be compared at 470 nm. Note that only the GOMOS data hold independent information at this wavelength, because the SCIAMACHY retrieval algorithm uses a fixed relationship between 750 and 470 nm.

Figure 50 shows the extinction profiles at 750 nm in latitude bins of 10 degrees. The envelopes around the profiles indicate the RMS. The GOMOS data (in blue) seem to be too small almost everywhere above 25 km, sometimes even slightly negative. In the tropical region, SCIAMACHY (red) and OSIRIS (green) agree well down to about 20 km, while the differences appear to increase towards Northern latitudes, with SCIAMACHY being higher there than OSIRIS.

Figure 51 shows the comparison of the aerosol extinction profiles between SCIAMACHY and GOMOS at 470 nm. Above 25 km the data are compatible, between 20 and 25 km the SCIAMACHY data are typically a bit lower, especially in the tropics.
Figure 52 and Figure 53 depict the same data as a function of latitude at six different tangent heights between 15 and 31 km. Here, the error bars indicate the error of the mean. Again the growing difference between SCIAMACHY and OSIRIS with higher latitude can be seen as well as the too low values of the GOMOS aerosol extinction at 750 nm at or above 25 km.
The time series of the extinction at 750 nm from SCIAMACHY and OSIRIS were studied as well. Figure 54 shows the time series of monthly zonal (5 degree latitude bins) means and their absolute and relative differences at 25 km as an example. Relative differences w.r.t. OSIRIS can be up to a factor of two, though they typically are within ±30%. The extreme differences occur mostly at high latitudes in winter and are most likely related to different sampling of PSCs. There is a tendency of SCIAMACHY extinctions being higher than OSIRIS at mid to high latitudes, and lower in the tropics. The systematic latitudinal behaviour may in principle be related to the different treatments of the phase function used in the respective retrieval algorithms, combined with the difference and variation of scattering angles over latitude and season. Due to the lower solar zenith angles at tropical and mid-latitudes, the sensitivity of the SCIAMACHY data to the albedo is also larger. In addition, in the tropics there may be effects of spatially or optically thin high clouds, which could not be filtered from the SCIAMACHY data, and which lead to lower extinction values at altitudes above the cloud top.
OSIRIS extinction values seem to be lower than SCIAMACHY's at the beginning of the time series (until about 2005), a behaviour that will eventually lead to quite different long-term trends from the respective data sets.
The analysis shows that, while the GOMOS and SCIAMACHY data at 470 as well as SCIAMACHY and OSIRIS at 750 nm agree on the order of magnitude and qualitative features, systematic differences remain. The possible reasons for these differences are currently under investigation by the respective teams, and possible solutions or mitigations are being studied. Concerning the SCIAMACHY data set it is highly recommended to use an additional cloud filter, a new data set with cloud flags is being released shortly.

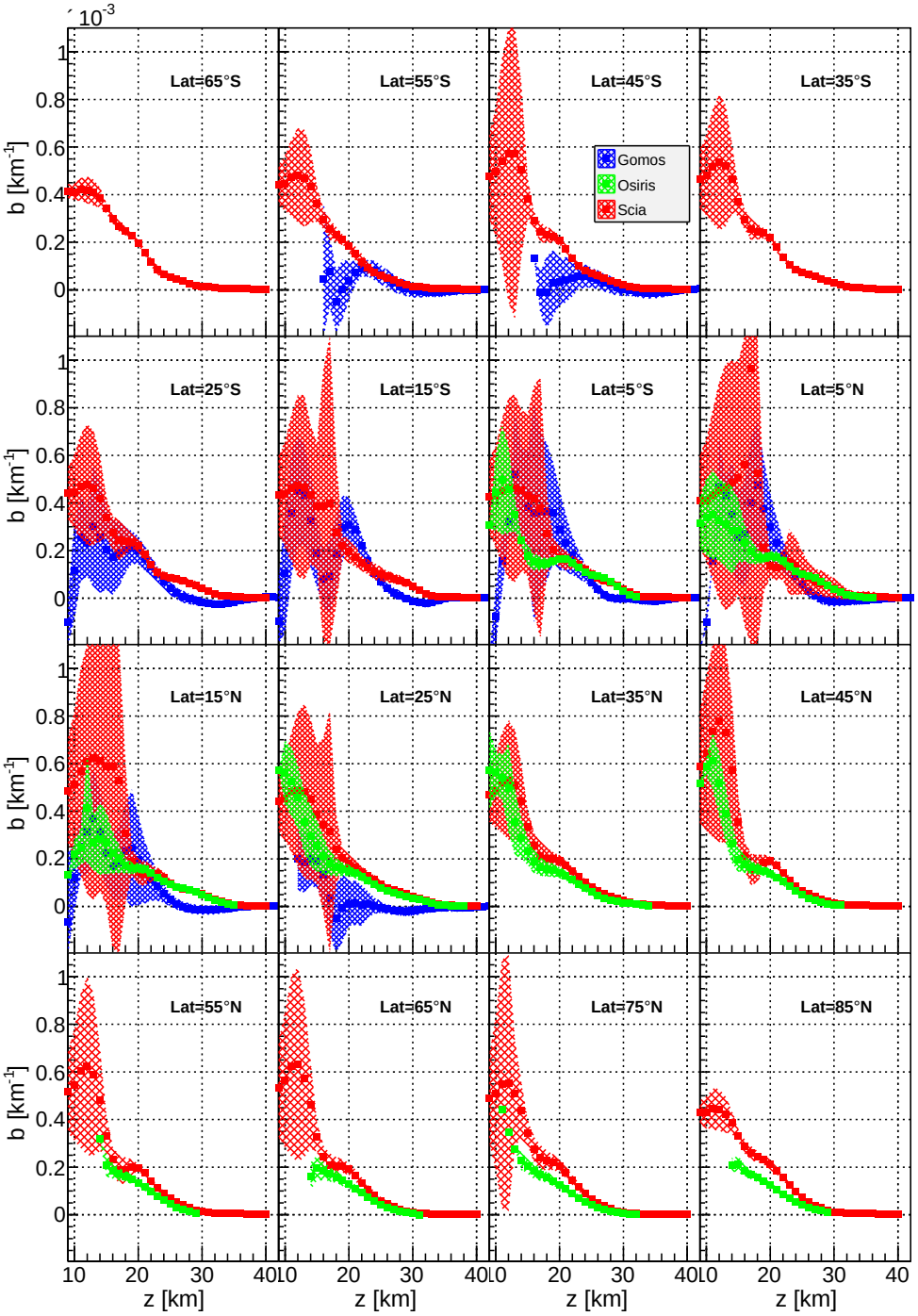

Figure 50: Monthly zonal mean aerosol extinction profiles at 750 nm from SCIAMACHY, GOMOS and OSIRIS for June 2003.

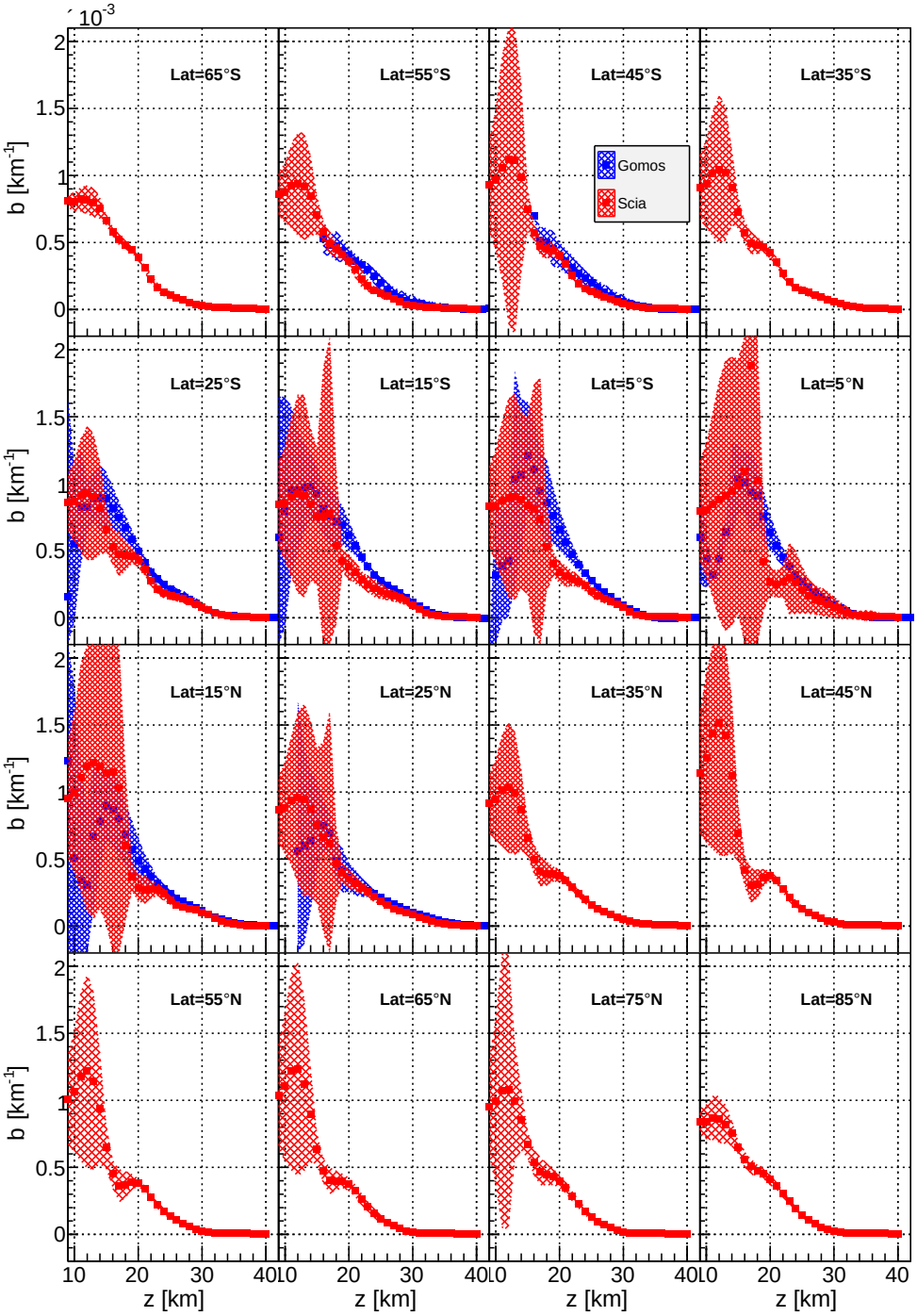

Figure 51: Monthly zonal mean aerosol extinction profiles at 470 nm from SCIAMACHY and GOMOS for June 2003.

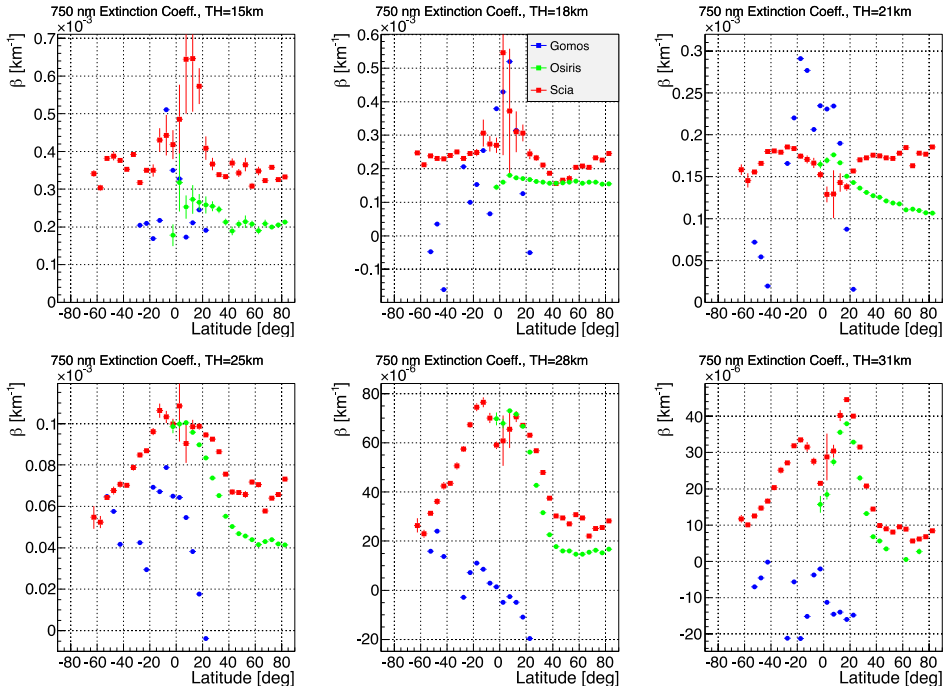

Figure 52: Monthly zonal mean aerosol extinction for June 2003 at 750 nm vs. latitude at 6 selected tangent heights, from SCIAMACHY, GOMOS and OSIRIS.

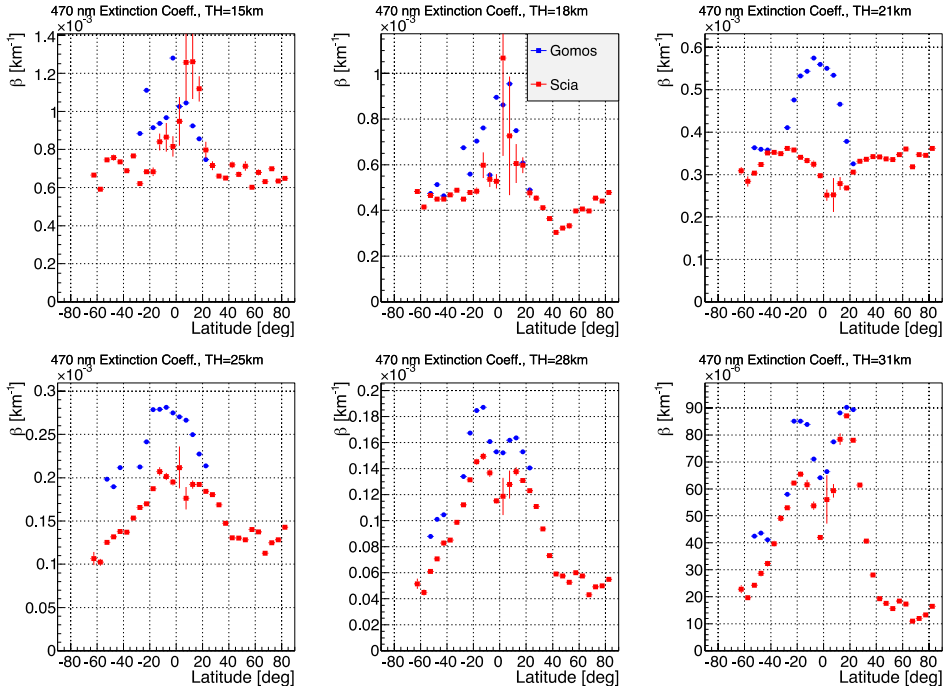

Figure 53: Monthly zonal mean aerosol extinction for June 2003 at 470 nm vs. latitude at 6 selected tangent heights, from SCIAMACHY and GOMOS.

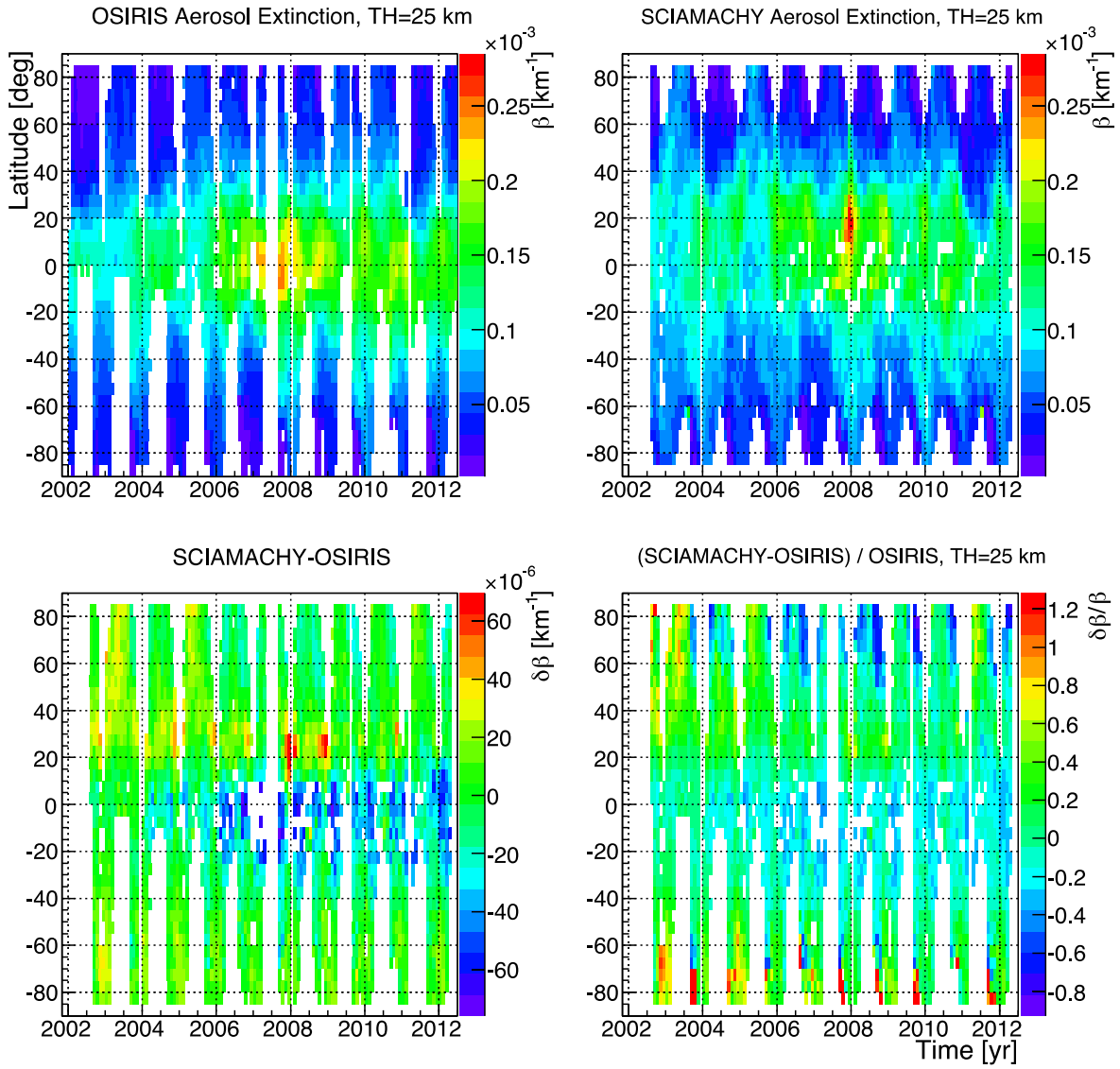

Figure 54: Top: OSIRIS (left) and SCIAMACHY (right) aerosol extinction at 25 km vs. latitude and time. Bottom: Absolute (left) and relative (right) differences of aerosol extinction (SCIA-OSIRIS) at 25 km vs. time.

4.4. Stratospheric Ozone

4.4.1. Validation of the GOMOS Bright Limb (GBL) data set (WP17)

In this section we compare the GOMOS bright limb (GBL) ozone profiles against data from GOMOS night-time occultations, MLS and OSIRIS. The differences are calculated as relative individual differences: (GBL-reference)/reference*100 [%]. Shown are medians of these relative differences.

Figure 55 shows how the GOMOS bright limb ozone profiles compare against the GOMOS night-time profiles (left panel) and against MLS (right panel). The difference is mostly less than 10% except in the 40 km altitude where GBL have 12-18% smaller values. The negative bias above 50 km against the GOMOS night-time data is expected due to diurnal variation of ozone. The comparison with OSIRIS (Figure 56) shows very similar results. The general agreement is good but the 40 km negative bias for all latitudes is clearly visible. Figure 57 and

Figure 59 show the same (co-located) comparisons but plotted as a 2D maps. It should be noted that in Figure 55 to Figure 59 we only used GBL data with the solar zenith angle less than 75 degrees. There is strong positive bias in the GBL data with high solar zenith angles (Figure 60). The reason for this bias is currently unclear.

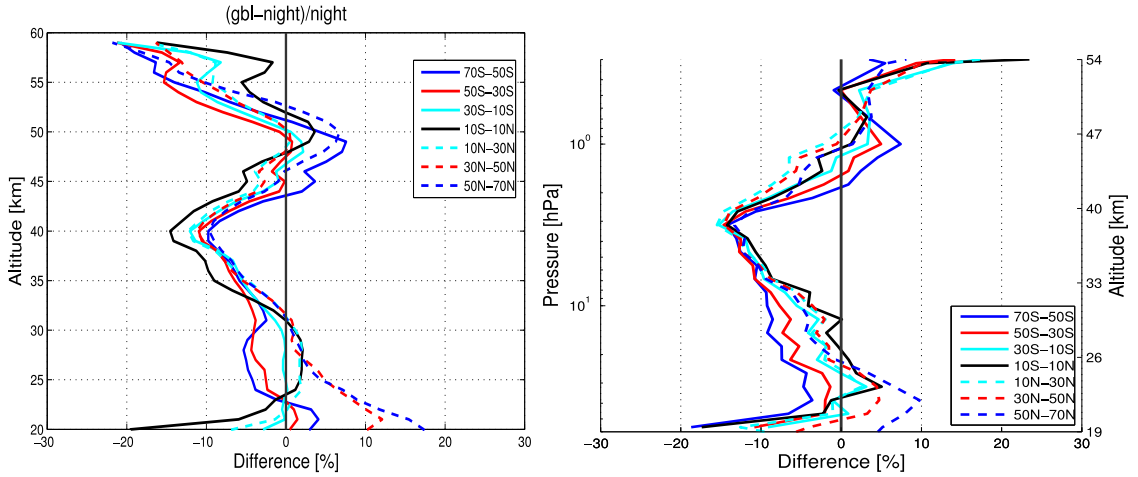

Figure 55: Comparison of GBL against GOMOS night-time occultations (left) and MLS (right).

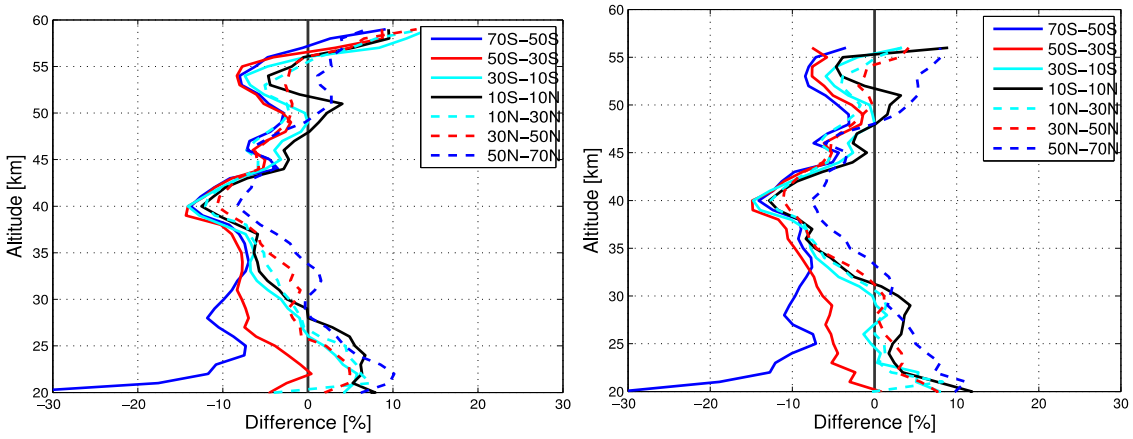

Figure 56: Comparison of GBL against OSIRIS FMI version (left) and OSIRIS Saskatoon version (right).

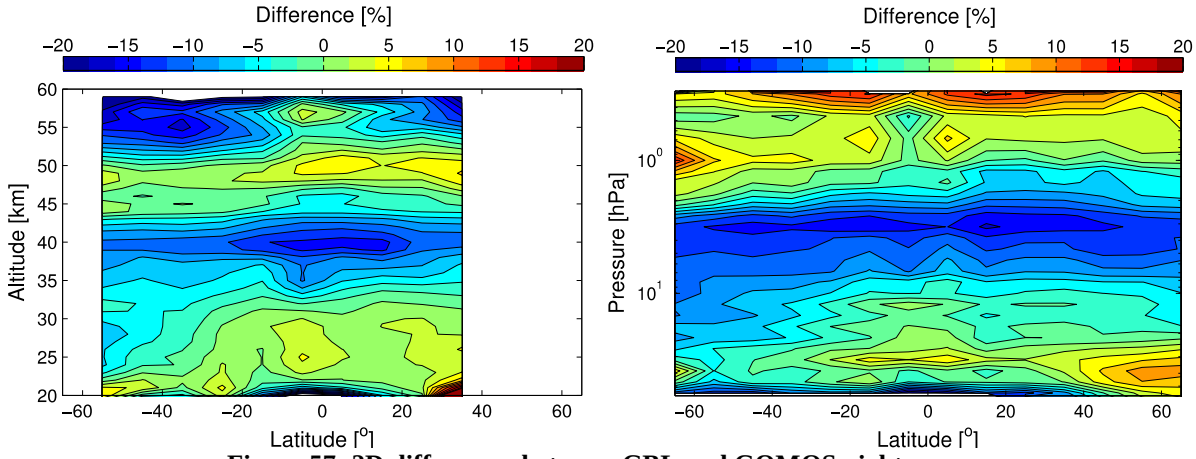

Figure 57: 2D differences between GBL and GOMOS night ozone.

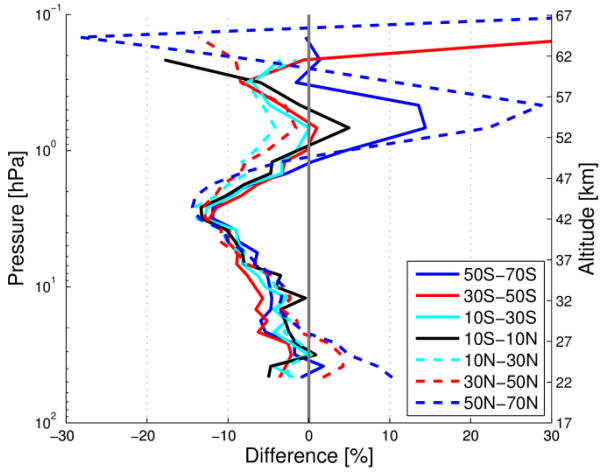

Figure 58: Comparison (left) and between GBL and MLS (right).

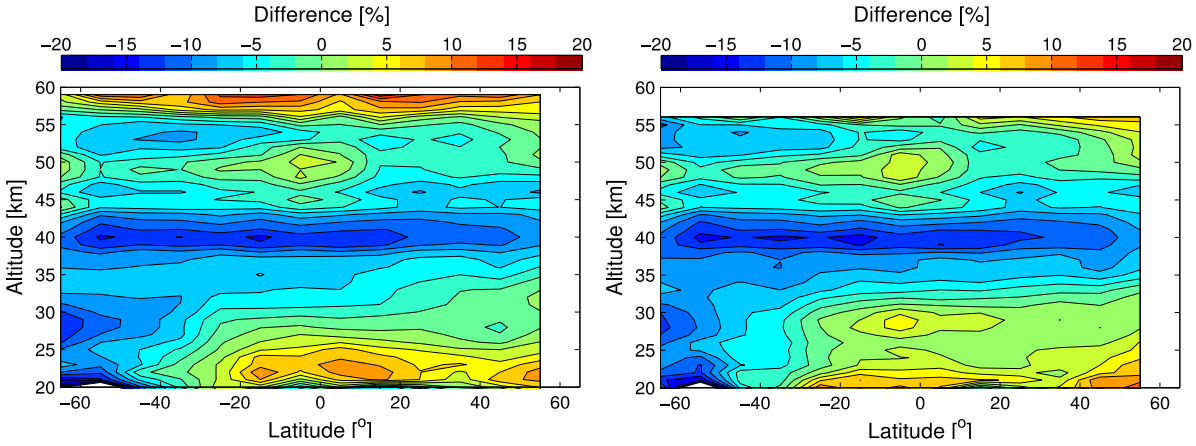

Figure 59: 2D differences between GBL and OSIRIS FMI version (left) and between GBL and OSIRIS Saskatoon version (right).

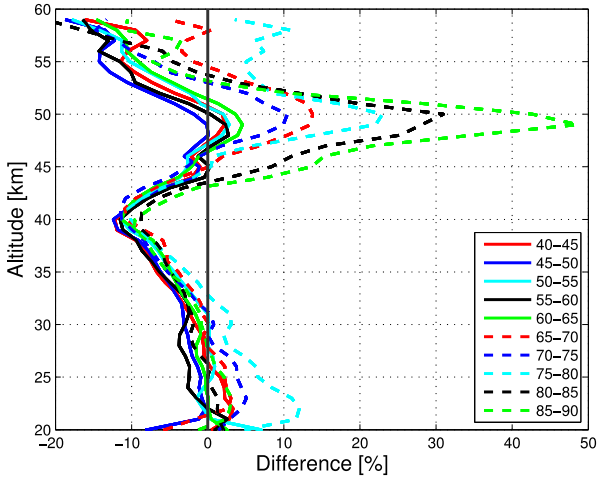

Figure 60: Comparison of GBL and GOMOS night-time profiles with different solar zenith angles.

4.4.2. Validation of the combined SAGE II-GOMOS data set from FMI (WP22)

FMI's combined SAGE II-GOMOS data set (Kyrölä et al., 2013) is built from GOMOS 10 pm measurements (IPF version 6) and SAGE II sunrise/sunset measurements (version 7).

Earlier product versions, GOMOS IPF 5 and SAGE II 6.2, have been successfully validated using ground-based measurements (For GOMOS, see van Gijsel et al., 2010; for SAGE II, see Wang et al., 2002). GOMOS IPF version 6 ozone values differ only slightly from the version 5. SAGE II version 7 ozone values are 1-2% smaller than the ones in version 6.2. The validation results for the new data versions of SAGE II and GOMOS are not yet available. We have compared SAGE II and GOMOS new data sets against each other using collocated measurements in the common period 2002-2005. SAGE II sunrise measurements of ozone were found to be on average 2% smaller than GOMOS measurements whereas sunset measurements were found to be 4% larger than GOMOS measurements. These deviations are largely created diurnal variation of ozone, but some instrumental reasons cannot totally be ignored. In forming the combined data set, these biases were removed by adjusting the SAGE sunrise and sunset ozone values by the observed averaged bias profiles. The new adjusted SAGE II sunrise/sunset data represent now approximately ozone measurements taking place around 10 pm. These artificial data values have not been validated.

FMI's combined SAGE II-GOMOS data is are given in 10-degree latitudinal bands and with monthly time average. As such, it is impossible to validate using ground-based sparse measurement networks. SAGE II and GOMOS monthly data have been compared against to several other satellite data sets in the recent SPARC Data Initiative publication (Tegtmeier et al., 2013). The agreement was found be good, but in some cases the special sampling patterns of occultation measurements were found to be the reason for discrepancies (see Toohey et al., 2013). GOMOS monthly ozone values have also been compared against other satellite measurements in the ESA's Ozone CCI project (Sofieva et al., 2013)

4.4.3. Validation of the merged SAGE II-GOMOS data set from BS (WP22)

The merged SAGE II and GOMOS data set generated by the Bodeker Scientific (BS) group in WP22 has been validated in three different ways, via:

- Comparisons against the FMI merged SAGE II and GOMOS ozone database.
- Comparisons against the Bodeker Scientific vertically resolved ozone database (Bodeker et al., 2012)[3].
- Comparisons of vertically integrated profiles from the merged SAGE II+GOMOS database against the Bodeker Scientific total column ozone database (Müller et al., 2008)[4].

The validation results are reported on in each of the three subsections below.

4.4.3.1. Validation against the FMI merged SAGE II+GOMOS database

The FMI merged database is provided as 10° zonal means and so the BS database, provided in 5° zonal means, was used to calculate area weighted zonal means in 10° latitude zones. Differences between the two databases were then calculated.

[3] Available from http://www.bodekerscientific.com/data/monthly-mean-global-vertically-resolved-ozone.
[4] Available from http://www.bodekerscientific.com/data/total-column-ozone.

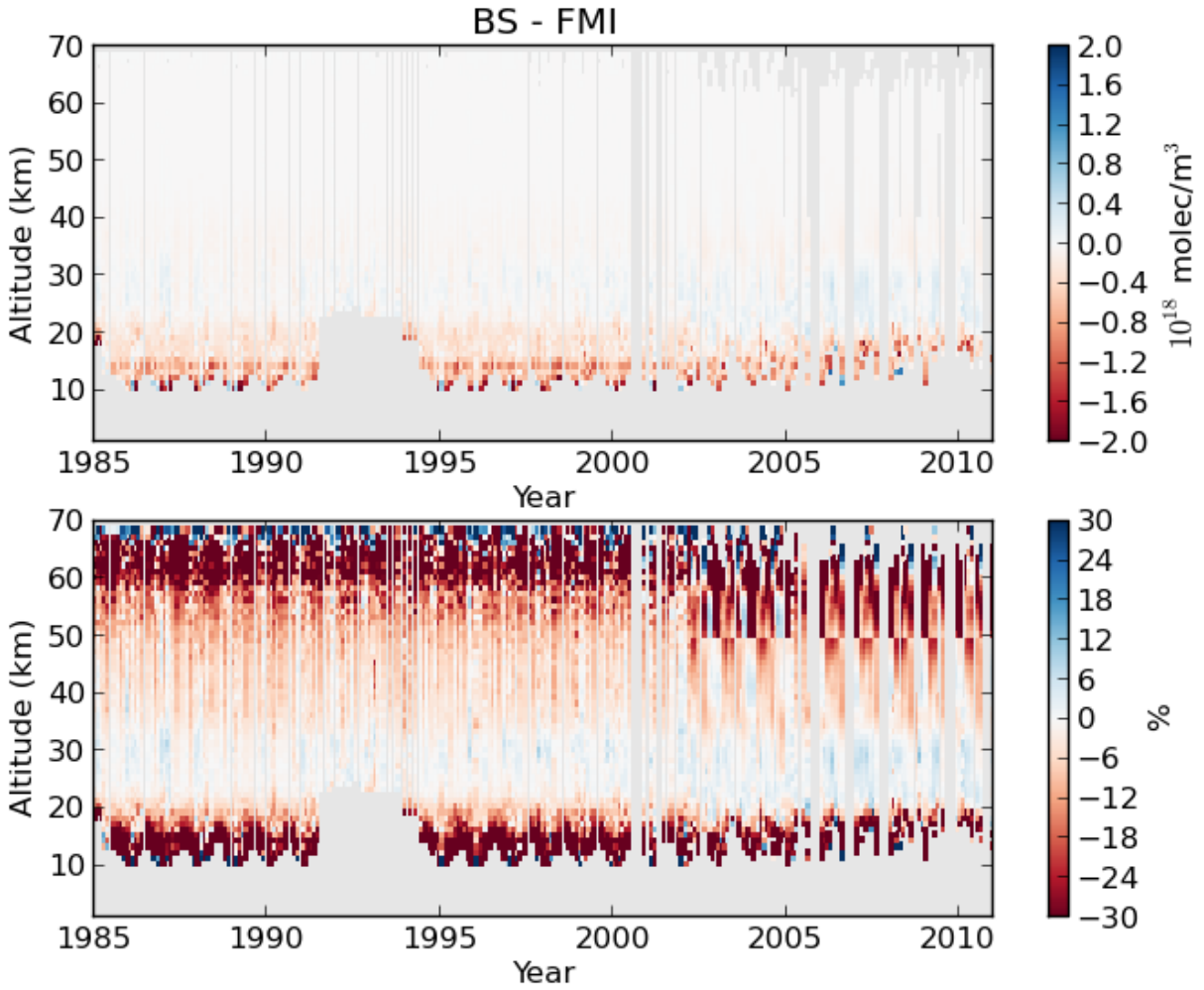

Figure 61: An example of the difference field between the Bodeker Scientific (BS) merged SAGE II and GOMOS ozone profile database and a similar database generated by FMI. The upper panel shows absolute differences in 1018 molecules/m3 while the lower panel shows the percentage differences. Greyed areas show where comparative data are not available. These results are for the region 30°N to 40°N.

One example difference field is shown in Figure 61. Twelve such plots are available since the FMI database does not extend poleward of 60°S and 60°N. In general, the differences between the BS and FMI merged SAGE II and GOMOS ozone profile databases are small i.e. within ±5%. A vertical discontinuity is apparent in the percentage difference field during the GOMOS era, likely due to the change in basis functions used in the calculation of the corrections applied to the GOMOS data.

4.4.3.2. Validation against the Bodeker Scientific vertically resolved ozone (VRO) profile database

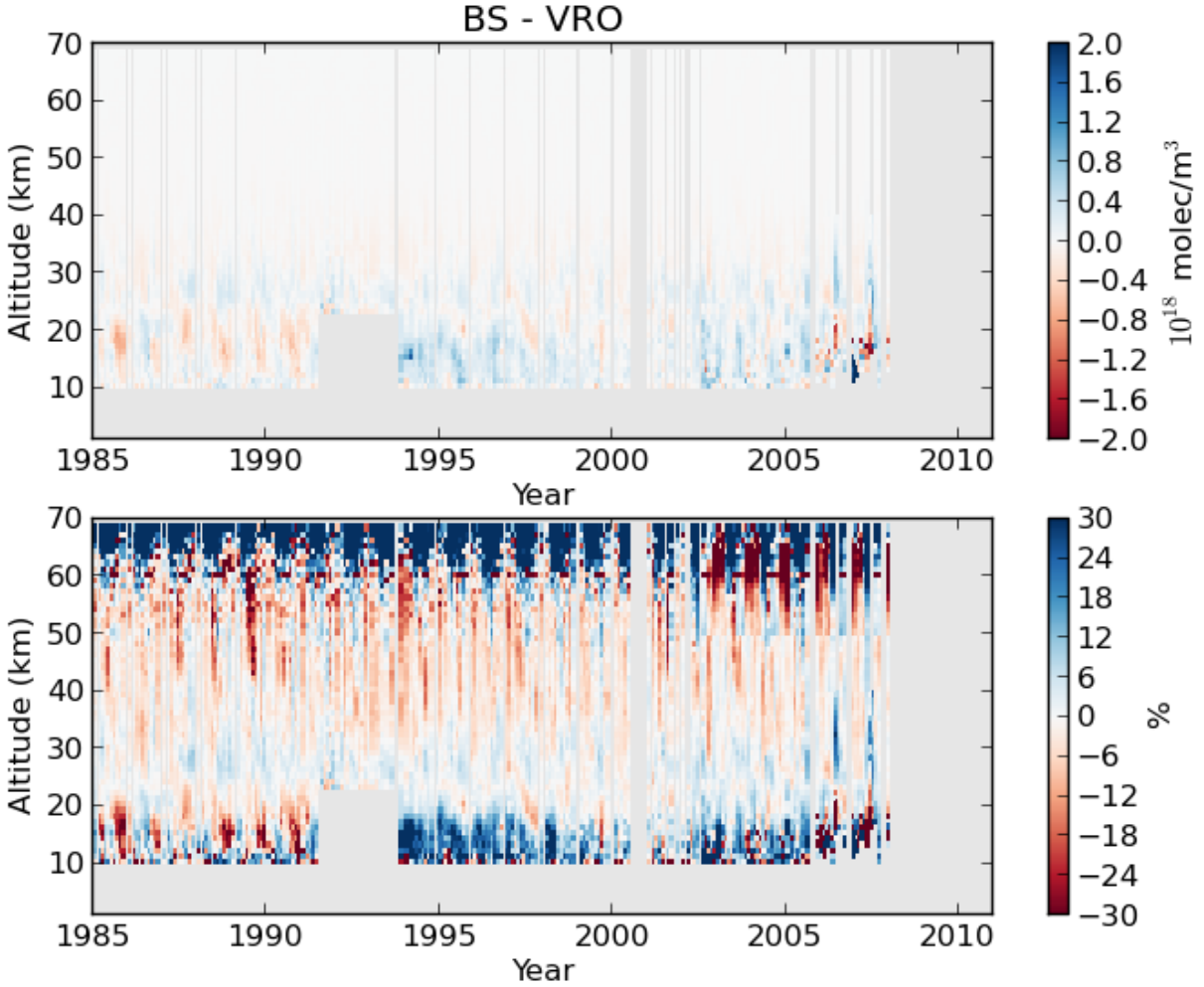

Figure 62: An example difference field between the Bodeker Scientific (BS) merged SAGE II and GOMOS ozone profile database and the Bodeker Scientific vertically resolved ozone database (VRO). The upper panel shows absolute differences in 1018 molec/m3 while the lower panel shows the percentage differences. Greyed areas show where comparative data are not available. These results are for the region 40°S to 45°S.

The BS merged SAGE II and GOMOS ozone database was also validated against the VRO database. Both databases are available in 5° latitude zones and so the differences were calculated at this latitude resolution. An example difference field for the 40°S-45°S zone is shown in Figure 62. The VRO database uses ozone measurements from many sources including satellite-based measurements (including SAGE II but not including GOMOS) and ozonesondes. As in the BS-FMI comparison, the differences through much of the stratosphere are small i.e. typically within ±5%. 35 such difference plots are available. As in the FMI validation, the discontinuity at 50 km is also present in these difference fields.

4.4.3.3. Validation of derived total column ozone

Figure 63 shows a validation of the BS merged SAGE II+GOMOS ozone database against the Bodeker Scientific total column ozone database. Profiles were integrated to obtain total column ozone and where ozone at levels below the lowest levels for which SAGE II and

GOMOS data were available, were obtained from the BS vertically resolved ozone database (primarily ozonesonde data). Differences are generally very small i.e. within ±10 DU (±3%) but appear to be slightly larger in the period following the Mt. Pinatubo eruption and during the period where GOMOS data dominate the database i.e. from the end of 2005 onwards.

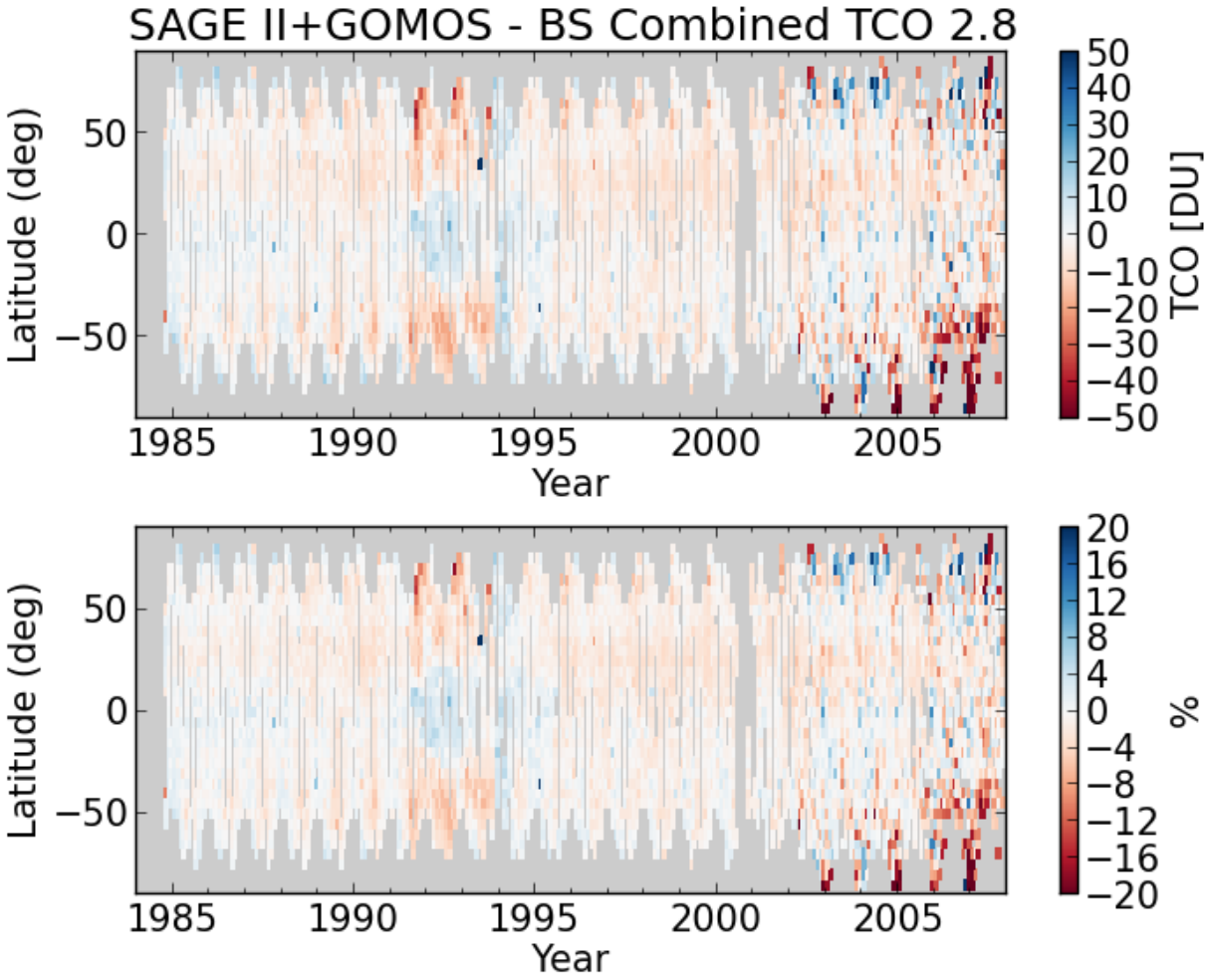

Figure 63: A comparison of the vertically integrated Bodeker Scientific SAGE II + GOMOS database with the Bodeker Scientific total column ozone database. The upper panel shows absolute differences in Dobson Units (DU) while the lower panel shows the percentage differences. Greyed areas show where comparative data are not available.

5. Intercomparison and evaluation of short-lived species climatologies

Short-lived species zonal mean climatologies have been created based on measurements from the two Odin instruments, SMR and OSIRIS. These climatologies, initially created for the period 2002 to 2010 but then extended to cover 2002 to 2013, have been submitted for evaluation to the SPARC data initiative (SPARC-DI). The purpose of the SPARC data initiative is to make a comprehensive assessment of climatologies through inter-comparisons of a large number of satellite climatologies. Short-lived species climatologies from all chemical families (NO_y, Cl_y, Br_y, HO_x) are being assessed in terms of absolute concentrations and spatio-temporal variability. The table below provides a list of Odin short-lived species climatologies

assessed by SPARC-DI relevant for the SPIN project and climatologies used for inter-comparison (from ESA and ESA 3rd party missions as well as from NASA and JAXA missions).

Odin climatology	Instruments for comparison
ClO 2001-2010	SMR, MLS, MIPAS, (SMILES)
BrO 2002-2010	OSIRIS, SCIAMACHY, (SMILES)
NO2 2002-2010	OSIRIS, SCIAMACHY, MIPAS, ACE-FTS, GOMOS, HALOE, SAGE, POAM
NO 2003-2010	SMR, MIPAS, ACE-FTS, HALOE
HNO3 2001-2010	SMR, MIPAS, ACE-FTS
HO2 2003-2004	SMR, MLS, (SMILES)

Short-lived species concentrations depend strongly on the local solar time of the measurement and SPARC-DI climatologies have therefore been created separately for day and night (SMILES) or am and pm (for sun-synchronous and solar occultation sensors) to allow for a more meaningful inter-comparison. For the Odin climatologies additional "scaled" products have been created with the help of scaling factors obtained from a photo-chemical model, in order to facilitate comparisons with other missions. Observations by the Odin instruments are performed at the equator between 6am and 7am (descending node) and between 6pm and 7pm (ascending node) owing to the satellite's sun-synchronous near-terminator orbit. Observations in the winter hemisphere are consequently performed during night, and day-time observations are performed in the summer hemisphere. The Odin orbit is not stabilized and has been slowly drifting in the 12 years since the launch in 2001. As a result, a correction is essential. Odin climatologies of ClO (am, pm) have therefore been scaled to the local time of Aura/MLS observations at the equator (around 1:30am and 1:30pm). This allows for a comparison in the latitude range ~70°S to ~70°N, whilst at higher latitudes the local time of MLS observations changes quickly from day to night or vice versa. Note that with this approach a total of 4 scaled SMR ClO climatologies were produced (SMR am scaled to 1:30am and 1:30pm and SMR pm scaled to 1:30am and 1:30pm), since morning and evening observations (or day and night observations, if more appropriate) from both satellites are handled separately. Note also that scaling has to be performed on individual profiles, i.e. before zonal averaging. As shown earlier by Jones et al. (2011), a consistency check can be performed because individually scaled am and pm observations from Odin should give the same long-term trend when averaged (e.g. monthly zonal means) and plotted against time. Jones et al. (2011) reported trend results only for the equatorial middle to upper stratosphere (35-45km) region, whilst the scaled Odin ClO climatologies created for SPARC-DI and optimized in the SPIN project provide data for trend analyses also at higher latitudes. Scaled climatologies have also been produced for OSIRIS NO_2 and SMR NO following the approach developed and described by Brohede et al. (2008), only that the chosen reference times for SPARC-DI are those of the MIPAS instrument on Envisat which observes at 10am and 10pm. It should also be noted that McLinden et al. (2010) constructed a (scaled) zonal mean BrO climatology from OSIRIS measurements which is also under evaluation in the SPARC data initiative. Note that the SPARC-DI assessment report has not been finalized at the time of writing the final version of the SPIN product validation report (April 2014). Whilst draft chapters for the relevant SPARC-DI short-lived species NO and NO_2 have been prepared, reviewed, and final versions are about to be completed, the intercomparisons for ClO, BrO and HO_2 have not yet been

summarized. The validation of the short-lived species data sets used in the SPIN project for improving on the SPARC-DI climatologies is therefore based on own comparisons.

5.1. Comparisons based on solar-zenith-angle binned level-3 data

Systematic comparisons of short-lived species observations from three microwave sensors (SMR, MLS, SMILES) and two mid-infrared sensors (ACE-FTS, MIPAS) have been performed with help of a 1D photo-chemical model to account for differences in local observation time. These comparisons concern the short-lived species ClO, HOCl, and HO_2 (besides the longer-lived HCl) and aim at comparing the diurnal variation of these species in the tropical mid-stratosphere to the lower mesosphere region. A manuscript describing the results of this study was published in ACP in August 2013 (Khosravi et al., 2013). Examples (for the here relevant species ClO and HO_2) are shown in Figure 64 and Figure 65. Intercomparisons of short-lived species measurements from instruments on sun-synchronous satellites (SMR, MLS, MIPAS) and from solar occultation instruments (ACE-FTS) are challenging since the measurements correspond to different solar zenith angles (or local times). However, using a model as reference, verified by the new SMILES data sets which cover all local times over a period of several months, provides the possibility to indirectly compare the diurnally variable species. The satellite data were averaged for latitudes of 20°S to 20°N for the SMILES observation period from November 2009 to April 2010 and were compared at three altitudes where the observations provide useful data: 35, 45 and 55 km. The study presents the first evaluation of HO_2 Odin/SMR data and the first comparisons of the new SMILES data and the latest version of MLS (version 3.3) data with other satellite observations. It is found that the satellite observations and the carefully initialized model generally agree well in terms of absolute mixing ratios as well as differences between the day and night values. This confirms that gas phase chemistry of these species based on latest recommendations of reaction rate constants is fairly well understood. The generally good agreement of the modelled and observed diurnal variation in terms of shape and amplitude (see Figure 64 for ClO and Figure 65 for HO_2) confirms that such a model can be used to scale observations of short-lived species taken at different local times to a common time, at least in the tropics.

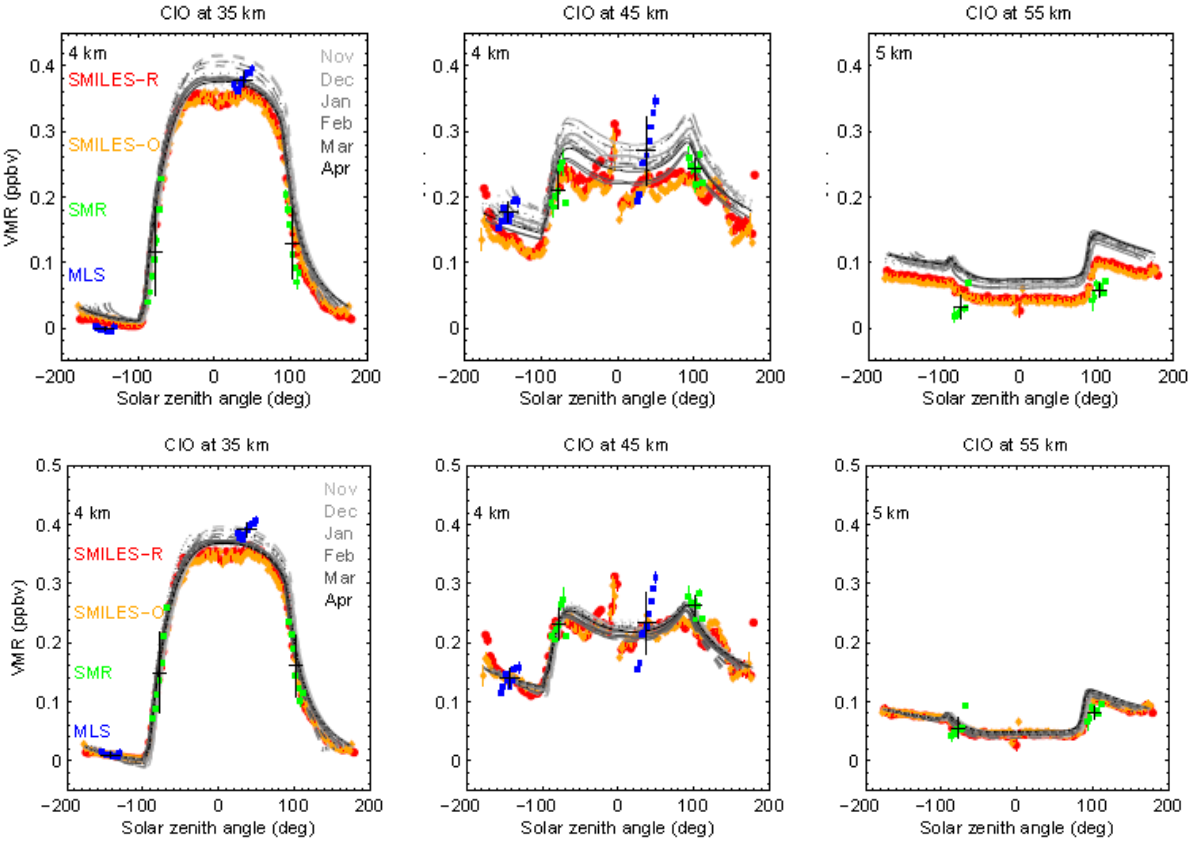

Figure 64: Top row: Modelled diurnal variation of ClO at the altitudes 35, 45 and 55 km (left, middle, right) in the tropics between 20ºS and 20ºN compared to observations made by SMILES, Aura/MLS, and Odin/SMR during the period November 2009 to April 2010. SMILES data are presented by two products; the research product (here called SMILES-R) and the official product (here named SMILES-O). The standard error of the mean is shown in the same colour as the measurements. If not visible, errors are smaller than the symbol size. Model simulations are given for different months (light grey shaded lines: November; dark grey: April). Solid grey lines are for the equator, dotted lines are for 20ºS and dashed lines are for 20ºN. Satellite data have been averaged in bins of 2.5º–3ºdeg in terms of solar zenith angle. The model was run with a vertical resolution of 1 km and the model results have been vertically smoothed using a Gaussian function. The fullwidth-at-half-maximum (in km) is stated in the left upper corner of each plot. MLS and SMR values have been averaged at their ascending/descending nodes. The variability of these measurements in the form of 1-σ standard deviation is shown as vertical error bars in black. The horizontal error bars show the range of solar zenith angles of the respective measurements. Bottom row: same as top row, but offset shifted for high solar zenith angles at night-time or sunrise in order to compare the amplitudes of the SMILES, MLS, and SMR diurnal cycles with the model. The average night-time or sunrise satellite measurements in the overlapping solar zenith angles and the model were subtracted from the values at other solar zenith angles and a correction term respective to SMILES-R for the same solar zenith angle range has been added. Adapted from Khosravi et al. (2013).

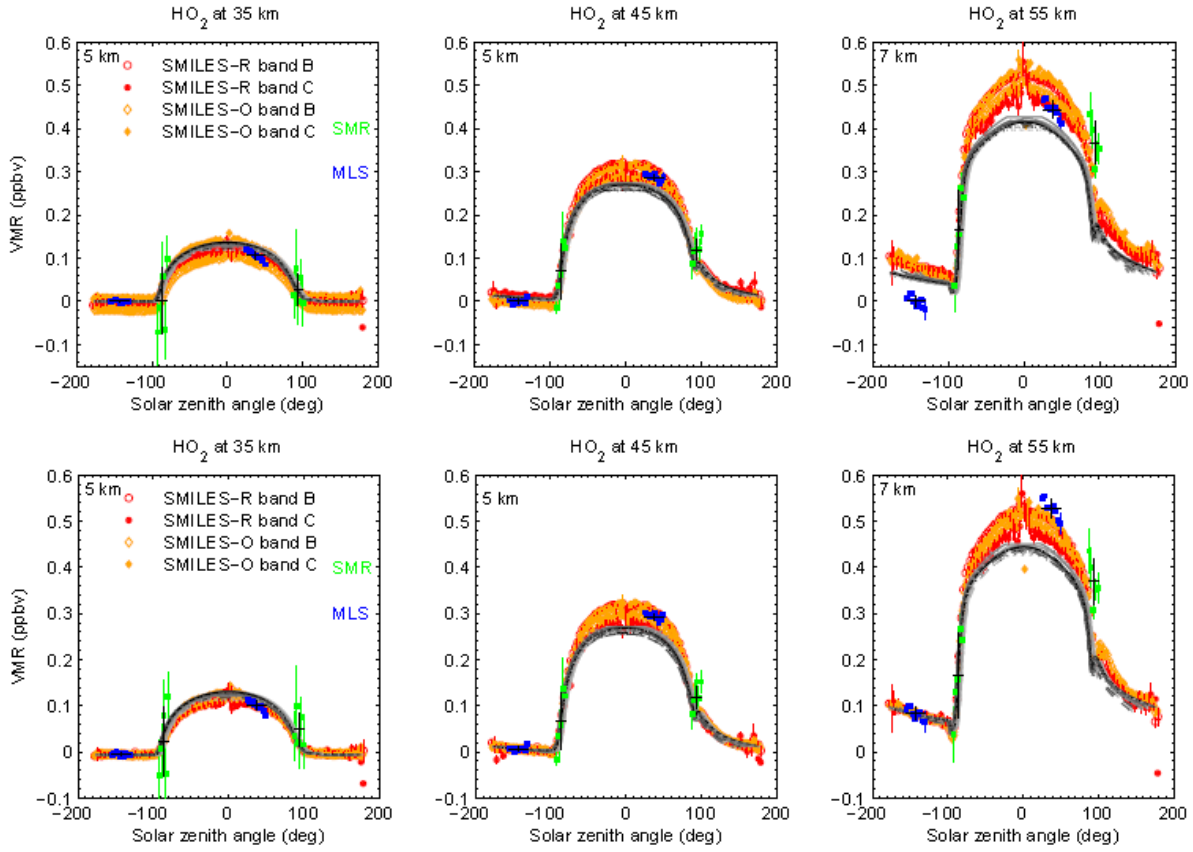

Figure 65: Same as Figure 64, but for HO_2. Top row: HO_2 from the model, Aura/MLS and Odin/SMR. Bottom row: Offset -shifted view for better comparison of diurnal cycle amplitudes. Note that HO_2 from SMR corresponds to the period October 2003 to October 2004 when this species was observed.

5.2. Intercomparisons of local solar time scaled climatologies

Profile retrievals of NO_2 measured by the OSIRIS instrument (Haley et al., 2004) have carefully been validated in the past with correlative measurements from a variety of satellite sensors such as SAGE II, SAGE III, POAM, and HALOE (Brohede et al., JGR-2007). The comparisons with solar occultation sensors, observing by definition at a solar zenith angle of 90°, were performed by scaling the OSIRIS NO_2 concentrations to the solar zenith angles of the occultation sensors. Brohede et al. (CJP-2007) presented then a multi-dimensional NO_2 climatology constructed by scaling the original OSIRIS retrievals to each full hour of a day using the PRATMO 1d photo-chemical model for scaling. A dedicated model run was performed for each OSIRIS measurement based on measured pressure/temperature and ozone data. The SPARC-DI OSIRIS NO_2 climatologies are based on the same approach but are available only for the original measurements and scaled to local solar times of 10am and 10pm.

Concerning BrO, a climatology has been published and thoroughly assessed by McLinden et al. (2010). Owing to the small signal-to-noise ratio of the OSIRIS spectral measurements for BrO, this climatology is based on a dedicated zonal mean retrieval scheme, i.e. individual profiles are not available. A further evaluation of this climatology is part of the SPARC data initiative.

Whilst the study by Khosravi et al. (2013) presents the first evaluation of HO_2 data from Odin/SMR and relatively good agreement with MLS and SMILES, this data set is available only for a short period of time (October 2003 to October 2004) due to an instrumental

problem with the frequency stabilization of the Odin/SMR 576GHz front-end and moreover only for about 1-2 days per month during this period as the relevant measurement mode was used time-shared with many other modes. Production of a local-time scaled climatology, useful for estimating trends, has therefore not been envisaged for this species.
For the abovementioned reason the focus is put within the SPIN project on an improvement of the local solar time scaled SPARC-DI climatologies of ClO, NO, and HNO_3 measured by the SMR instrument on Odin.

5.2.1. Intercomparison of scaled ClO climatologies with external data

In the following we provide results of the comparisons made for ClO. ClO level-2 data from SMR have in the past been evaluated against measurement of the Aura/MLS and SMILES instruments in a number of validation studies, namely those of Lambert et al. (2007), Khosravi et al. (2013) (see Figure 64), and Sagawa et al. (2013). Building on the study of Lambert et al. which was for Aura/MLS level-2 version 2.2, the MLS quality document (Livesey et al., 2013) provides an updated comparison to the most recent MLS data version (v3.3/3.4). All these studies conclude that coincident SMR and MLS ClO profiles show an excellent agreement between 1hPa and 20hPa, whilst MLS data have a low bias with even negative mixing ratios at higher pressure levels. Scaled Odin/SMR ClO climatologies have been produced for daytime (LST 1:30pm) and night-time (1:30am) which corresponds to the MLS equator crossing times.
We have calculated (unscaled) MLS climatologies for day and night observations for comparison with the scaled climatologies produced from SMR observations. Figure 66 shows an example of the climatological profile comparisons for the months of January and July averaged over the period 2005-2012. Results are consistent with earlier validation studies based on coincident profile comparisons: we find generally a good agreement in the pressure range 30-1.5hPa even for the scaled climatologies. Note that the scaled SMR climatologies are based on the assumption that profiles requiring scaling factors larger than 5 to be scaled are discarded. Moreover the model is not valid in the polar winter regions as it doesn't account for heterogeneous chemistry. Exclusion (filtering) of such data leads to a sparser data set in the lower to mid stratosphere during specific periods. The effect of filtering of individual profiles requiring too large scaling factors is also well illustrated in the time-series comparison shown in Figure 67 to Figure 69 for the 3hPa pressure level. The agreement between MLS and the two SMR versions is reasonably good in some of the chosen examples, but one can also see some discrepancies, notably the spikes in SMR ClO scaled to 13:30 (day) in the mid-latitude and Polar Regions during specific periods (autumn) of the year (Figure 69). Tests revealed that those spikes arise from the highest model scaling factors (here up to 5) and can be further suppressed (e.g. necessary for trend analysis) by filtering using the zonal mean scaling factor field provided in the climatologies. As a general rule, one can say that this affects mainly the data scaled to day-time during limited periods of the year and the effect increases with pressure, i.e. disappears at 1hPa. The data scaled to night-time data are not affected and agree better with MLS at 3hPa. Here the opposite effect limits the usefulness for trend analyses, namely low scaling factors leading to very small mixing ratios in the lower to mid-stratosphere limiting the useful vertical range. Finally one can say that in terms of observed variability the best agreement of the scaled SMR time-series with MLS is found at the 1-2hPa levels. However, the agreement in terms of absolute mixing ratios is less good at 1hPa as MLS shows here lower values than the scaled SMR climatologies.

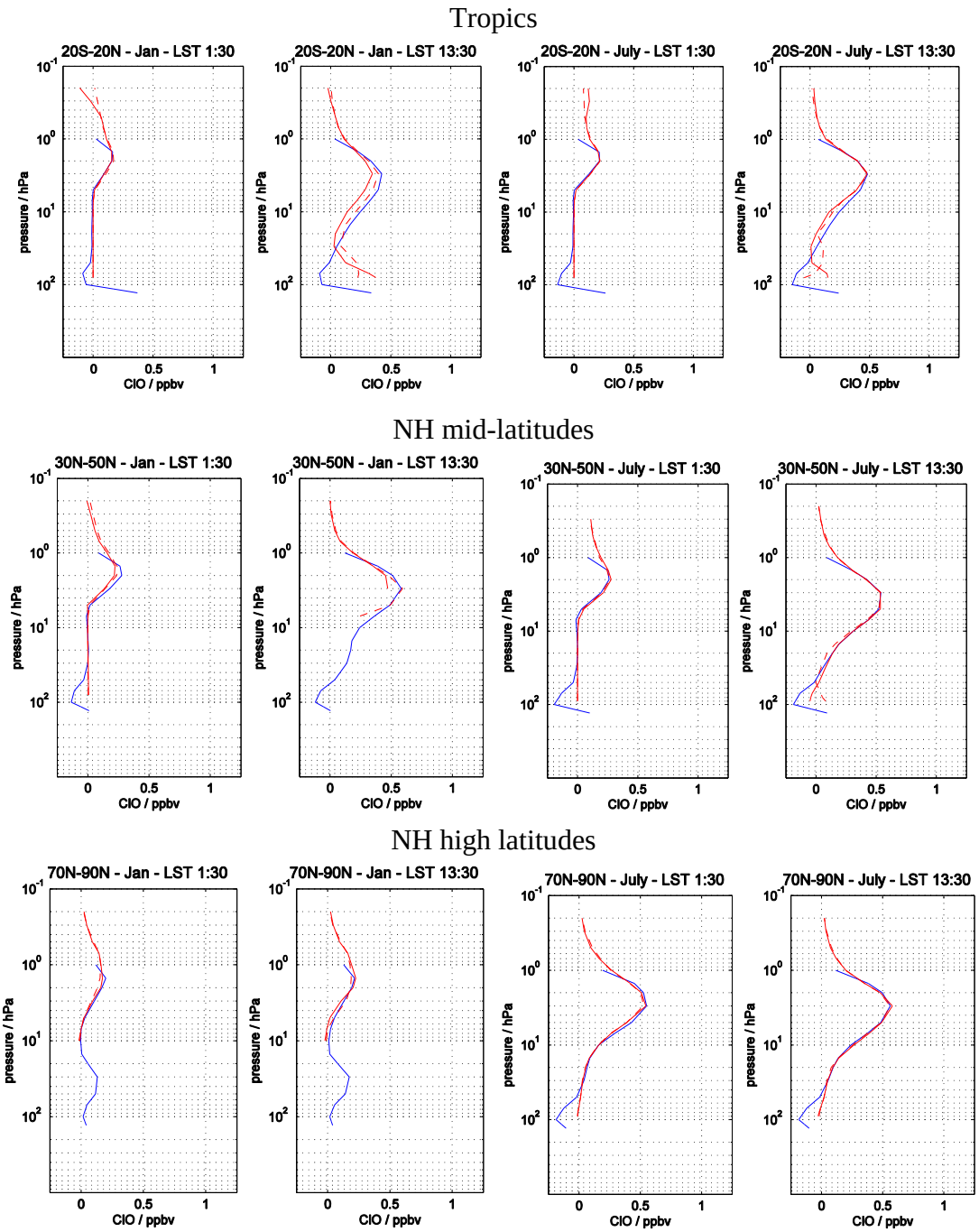

Figure 66: Comparison of local solar time scaled Odin/SMR climatological profiles (red) with unscaled MLS observations (blue) of ClO. Top: tropics; middle: northern mid-latitudes; bottom: northern high latitudes. Examples are shown for January (left) and July (right) and are averaged over the 2005-2012 period. For each month comparisons are shown for the two local times corresponding to day and night. Note that SMR climatologies are further distinguished into climatologies produced from am (morning: solid red line) and pm (evening: dashed red line) measurements.

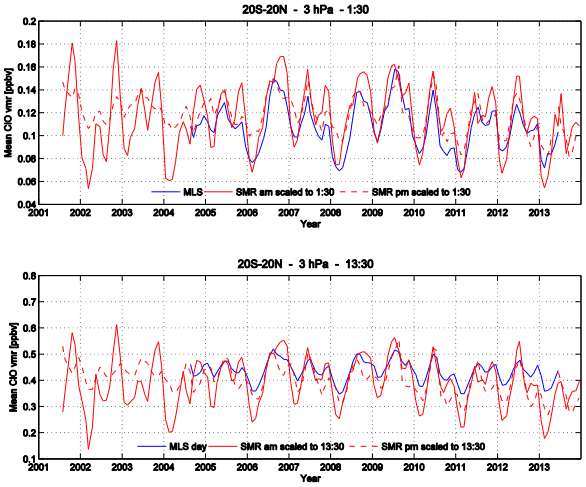

Figure 67: Comparison of local solar time scaled SMR (red) and unscaled MLS (blue) ClO time-series for the 3hPa level and the tropics. Top: night (1:30am). Bottom: day (1:30pm).

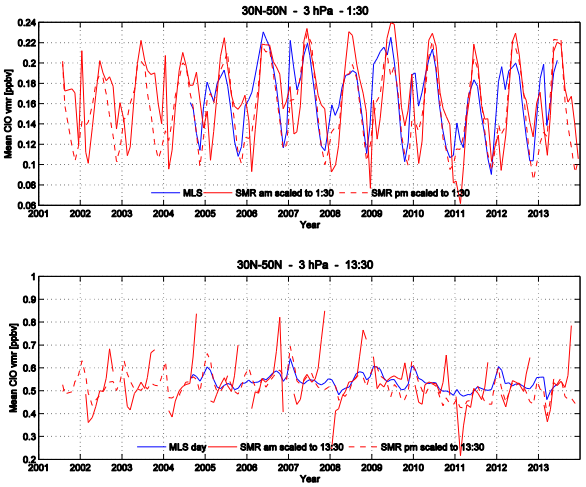

Figure 68: Comparison of local solar time scaled SMR (red) and unscaled MLS (blue) ClO time-series for the 3hPa level and northern mid-latitudes. Top: night (1:30am). Bottom: day (1:30pm).

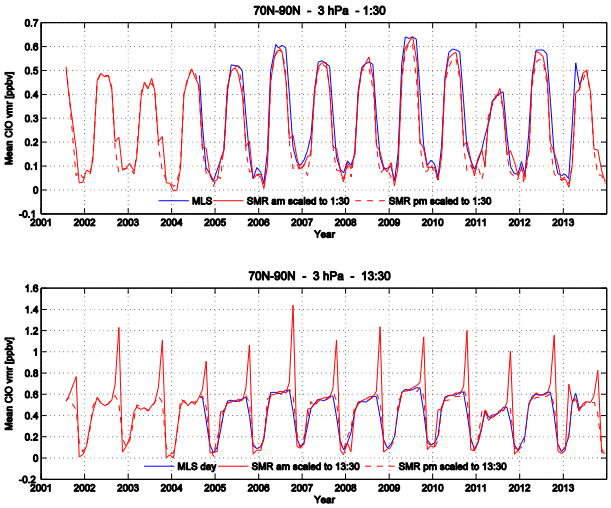

Figure 69: Comparison of scaled SMR (red) and unscaled MLS (blue) ClO time-series for the 3hPa level and northern high latitudes. Top: night (1:30am). Bottom: day (1:30pm).

A trend analysis of the combined SMR and MLS ClO time series would thus require setting a strict criterion for the maximum tolerated model scaling factor. Tests suggested that values of 1.5-2 remove most of the affected data but also lead to a sparser data set during other periods and in other atmospheric regions. Nevertheless it can be said in summary that there is a reasonably good agreement between MLS and SMR scaled ClO climatologies in terms of absolute mixing ratio and temporal variability, as expected from validation studies, with marked differences during autumn in the mid- to lower stratosphere (at pressures larger or equals 3hPa).

5.2.2. Internal consistency of scaled ClO climatologies

The consistency of scaled model time-series obtained separately from descending (am) and ascending (pm) orbit portions can be exploited as an internal quality criterion when it comes to trend analyses for a species under study, as reported earlier by Jones et al. (2011). An internal verification of the scaled ClO climatologies can be done by comparison of scaled am and pm ClO time-series for various altitude and latitude bands. This is illustrated in Figure 70 for the 3hPa level in the tropics.

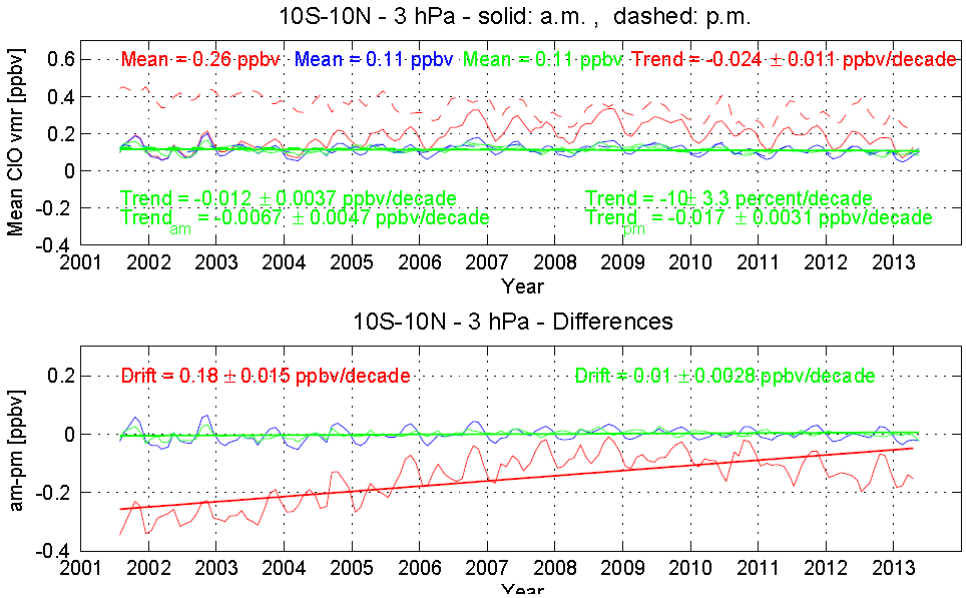

Figure 70: Internal consistency of ClO time-series based on measurements scaled to 1:30am at the 3hPa level in the tropics. Top: unscaled (red), scaled (blue), and deseasonalised scaled (green) time-series of morning (am: solid) and evening (pm: dashed) ClO observations by Odin/SMR. Bottom: Difference of am and pm scaled climatologies.

It can first of all be seen how the scaling improves the long-term stability of the data set (Figure 70, top), as the am and pm based scaled climatologies have only a small drift against each other (Figure 70, bottom). Differences in the time evolution of the am and pm zonal mean climatologies have been accounted for in the scaled am and pm climatologies which show both nearly the same time evolution. Future trend studies for ClO will have to rely on such successfully scaled am and pm zonal mean data sets. The information on drifts can be exploited when estimating the uncertainty or reliability of trends derived from the local solar time scaled climatologies.

.

5.2.1. Intercomparison of scaled NO climatologies

The assessment results for scaled NO climatologies from Odin/SMR created in the SPIN project are based on intercomparisons with SPARC-DI climatologies from MIPAS on Envisat. The Odin observations were scaled to 10am, the local time of MIPAS daytime measurements. Figure 71 shows an intercomparison of climatological profiles from the two satellite instruments for the tropics and southern middle and high latitudes during January (summer) when Odin observations in the southern hemisphere are performed during daytime. Whilst the original Odin NO data agree more or less with MIPAS, the scaled Odin mixing ratios are clearly larger at and below the maximum of the NO profile. Best agreement is achieved at high latitudes where the scaling has only little effect.

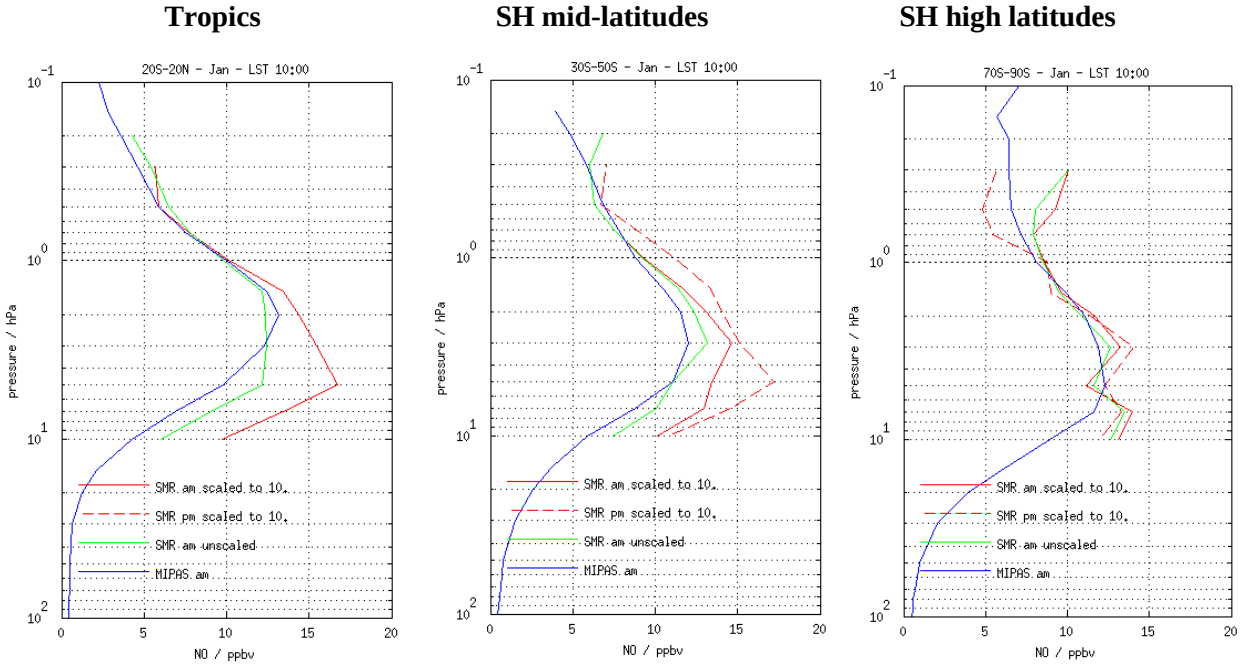

Figure 71: Comparison of unscaled (green) and scaled (red) Odin/SMR climatological NO profiles with unscaled MIPAS (blue) climatological zonal mean data of NO. Left: tropics; middle: southern mid-latitudes; right: southern high latitudes. Examples are shown for January and are averaged over the 2005-2010 period. Comparisons are shown for 10am, the local time of MIPAS daytime observations. Note that SMR climatologies can further be distinguished into climatologies produced from am (morning: solid red line) and pm (evening: dashed red line) measurements.

Time-series comparisons at the 1hPa level are shown in Figure 72 to Figure 74 for southern and northern high latitudes and the tropics. The figures show both, nighttime and daytime data. Note that the model scaling factors make the scaled Odin NO climatology to become virtually zero during night (i.e. in the winter hemispheres), except for the high latitude regions during polar summer where always daytime measurements can be made. Consequently the scaling factors are around one near the summer poles, so that the scaled and unscaled Odin profiles are very similar. At this level a reasonable agreement with MIPAS is found, in terms of absolute mixing ratios and observed temporal variability. Figure 75 shows for comparison the tropical NO time-series at the slightly lower 3hPa level, in order to high-light the positive bias of the scaled Odin/SMR daytime climatologies compared to MIPAS.

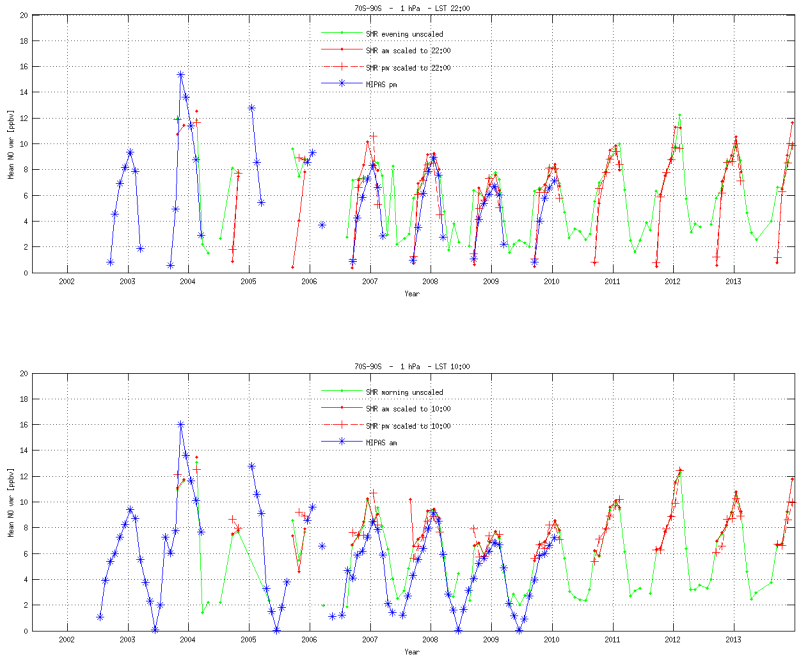

Figure 72: Comparison of local solar time scaled SMR (red) and unscaled SMR (green) and MIPAS (blue) NO time-series for the 1hPa level and SH high latitudes. Top: night (10pm). Bottom: day (10am).

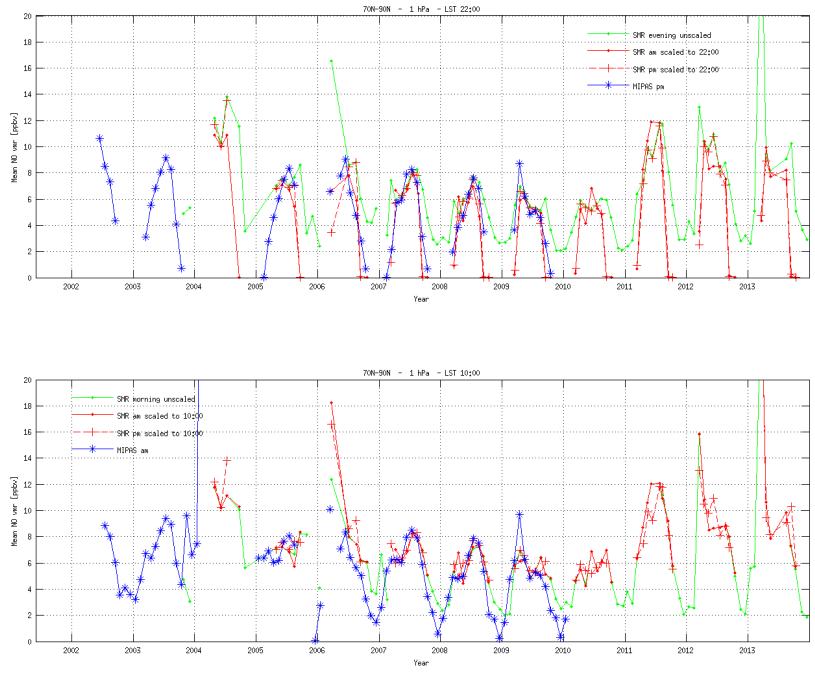

Figure 73: Comparison of scaled SMR (red) and unscaled SMR (green) and MIPAS (blue) NO time-series for the 1hPa level and northern high latitudes. Top: night (10pm). Bottom: day (10am).

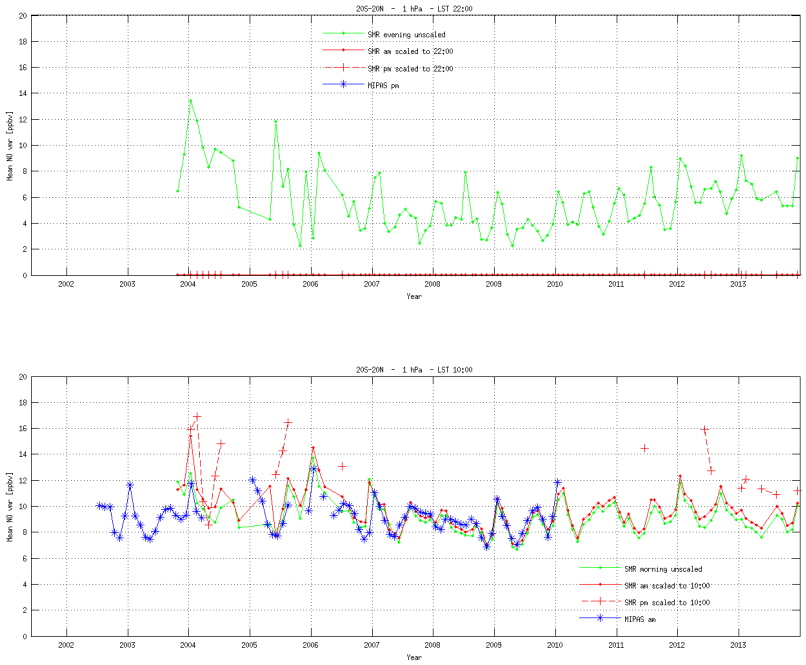

Figure 74: Comparison of local solar time scaled SMR (red) and unscaled SMR (green) and MIPAS (blue) NO time-series for the 1hPa level and the tropics. Top: night (10pm). Bottom: day (10am).

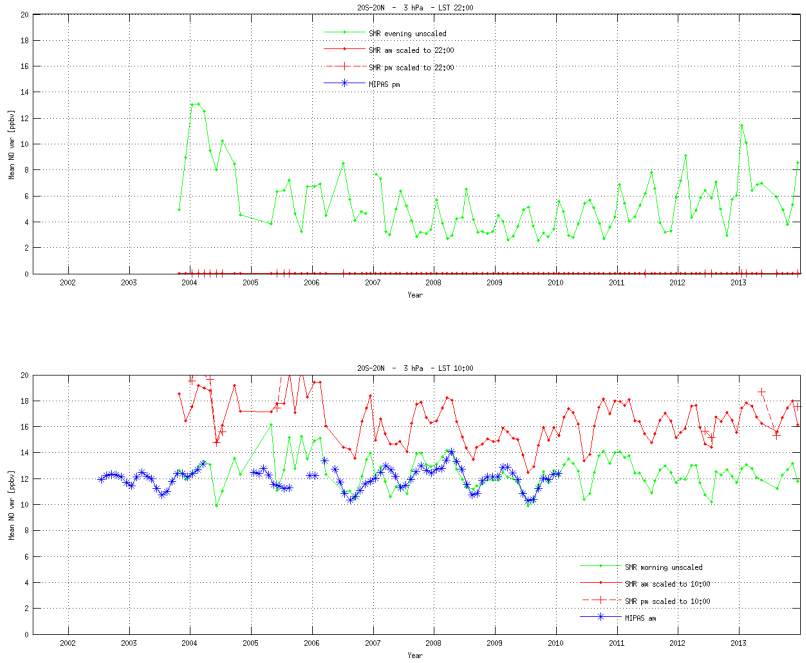

Figure 75: Comparison of scaled SMR (red) and unscaled SMR (green) and MIPAS (blue) NO time-series for the 3hPa level and the tropics. Top: night (10pm). Bottom: day (10am).

5.2.2. Intercomparison of HNO_3 climatologies

This section is dedicated to evaluate the Odin/SMR climatologies of nitric acid against independent observations from Envisat/MIPAS and Aura/MLS. In contrast to the previously discussed shorter-lived species ClO and NO, HNO_3 has a reasonably long lifetime in the stratosphere and model scaling factors are close to one, except above 40km as shown in Figure 76.

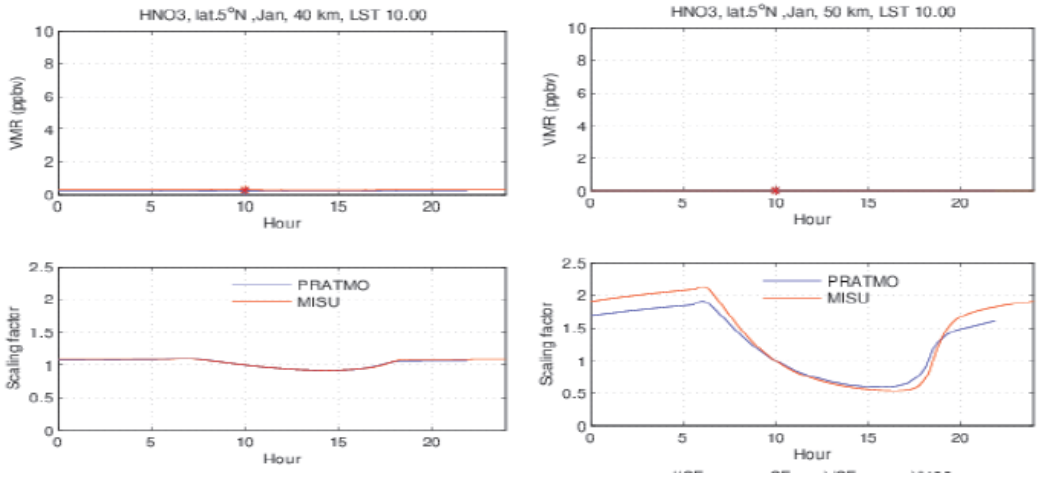

Figure 76: Comparison of model VMR (top) and scaling factors (bottom) for 40km and 50km and the tropics.

Profile comparisons of the SMR, MIPAS, and MLS climatologies of HNO_3 are shown in Figure 77 for low, middle, and high latitudes. Whilst scaling doesn't play a role here, a very good agreement between all the data is generally observed. Exceptions are MIPAS at the highest southern latitudes which observed a different profile shape with higher mixing ratios at low and high altitudes as well as Odin in the 20-30hPa range which is slightly lower. Note that Odin/SMR level-2 HNO_3 data were empirically down-scaled to match average vmr's obtained from MIPAS retrievals (by the Oxford processor) during early years of the Envisat mission in order to correct for a known high bias. Absolute mixing ratios from SMR are therefore not independent of MIPAS.

Time-series comparisons for levels close to the HNO_3 maximum are shown in Figure 79 for the tropics (at 20hPa) and in Figure 79 to Figure 80 at high southern and northern latitudes (at 30hPa). The temporal variability of nitric acid is captured by all satellite instruments which are in very good quantitative agreement in the different latitude ranges. For comparison, Figure 81 shows the comparison at northern high latitudes in the upper stratosphere at the 1.5hPa level. Here the mixing ratios are much lower, but considerable enhancements were observed during the winter seasons. The measurements agree reasonably well.

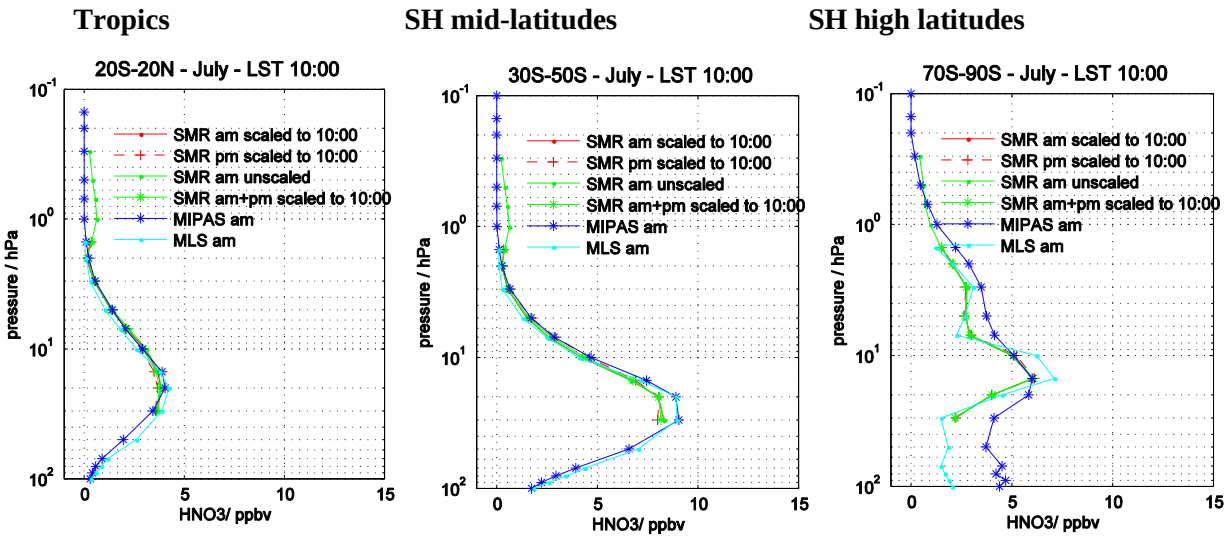

Figure 77: Comparison of unscaled (green) and scaled (red) Odin/SMR climatological HNO_3 profiles with unscaled MIPAS (blue) and MLS (cyan) zonal mean data. Left: tropics; middle: southern mid-latitudes; right: southern high latitudes. Examples are shown for July and are averaged over the 2005-2010 period. Comparisons are shown for 10am, the local time of MIPAS daytime observations. Note that SMR climatologies can further be distinguished into climatologies produced from am (morning: solid red line) and pm (evening: dashed red line) measurements.

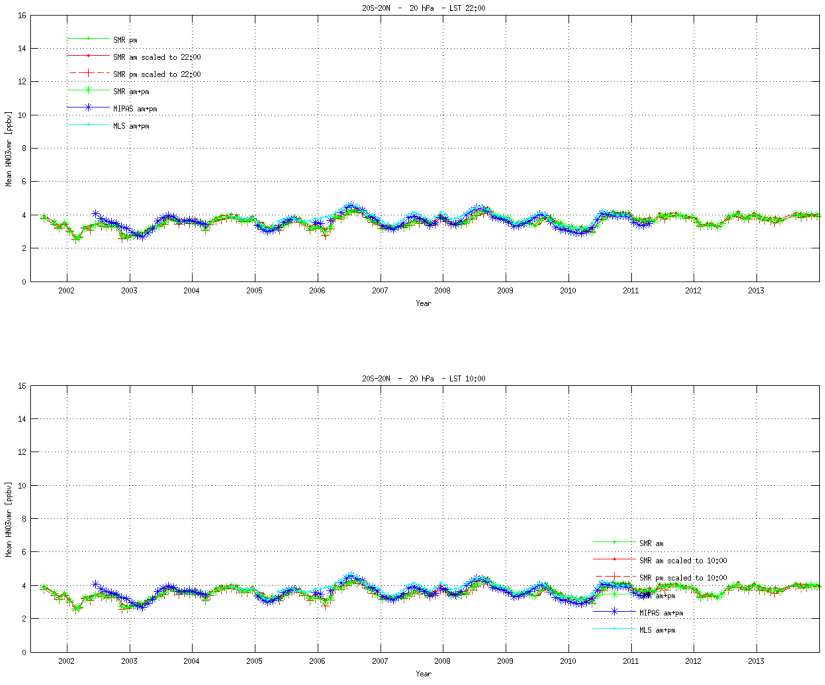

Figure 78: Comparison of scaled SMR (red), unscaled SMR (green), MIPAS (blue), and MLS (cyan) HNO_3 time-series for the 20hPa level in the tropics. Top: night (10pm). Bottom: day (10am).

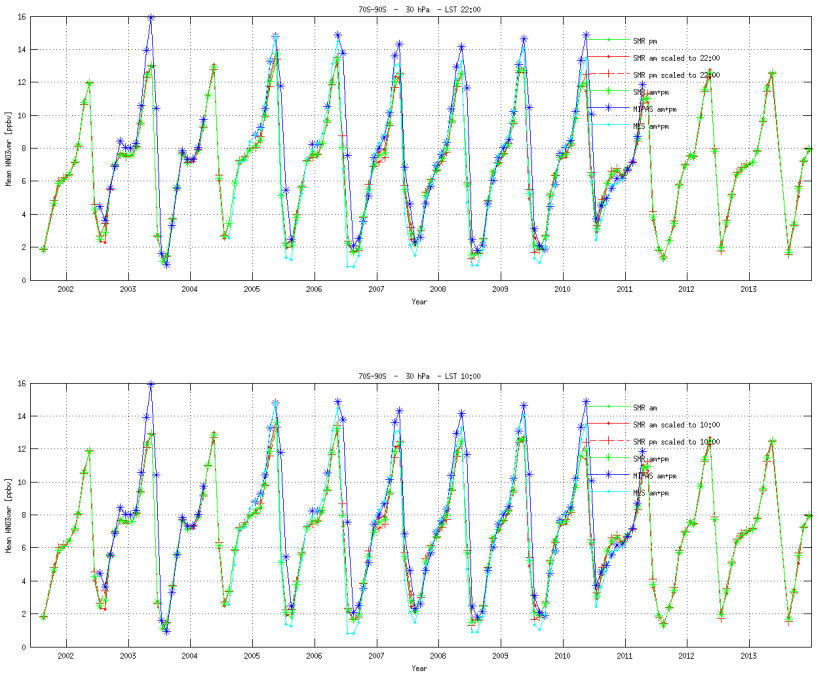

Figure 79: Comparison of scaled SMR (red), unscaled SMR (green), MIPAS (blue), and MLS (cyan) HNO_3 time-series for the 30hPa level and southern high latitudes. Top: night (10pm). Bottom: day (10am).

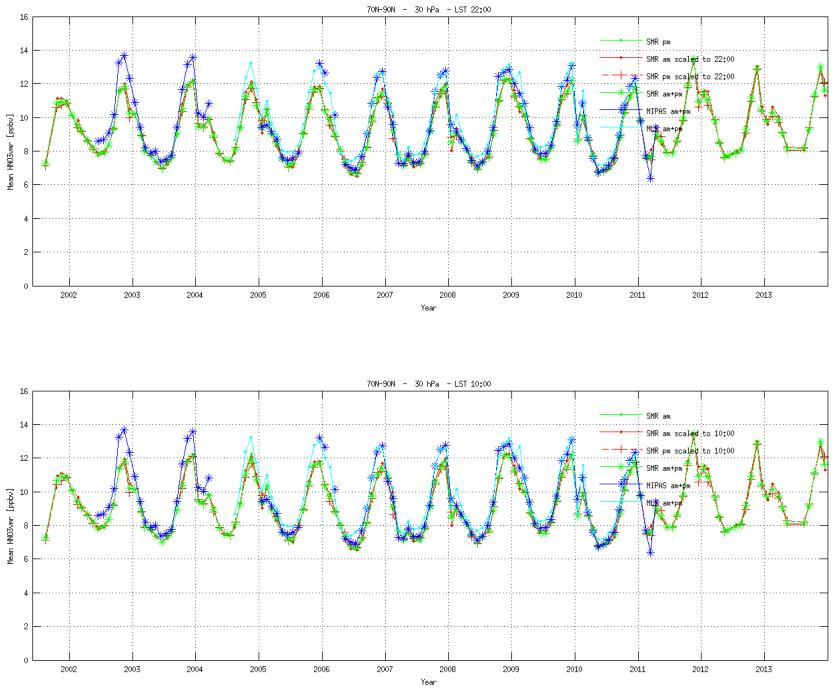

Figure 80: Comparison of scaled SMR (red), unscaled SMR (green), MIPAS (blue), and MLS (cyan) HNO_3 time-series for the 30hPa level and northern high latitudes. Top: night (10pm). Bottom: day (10am).

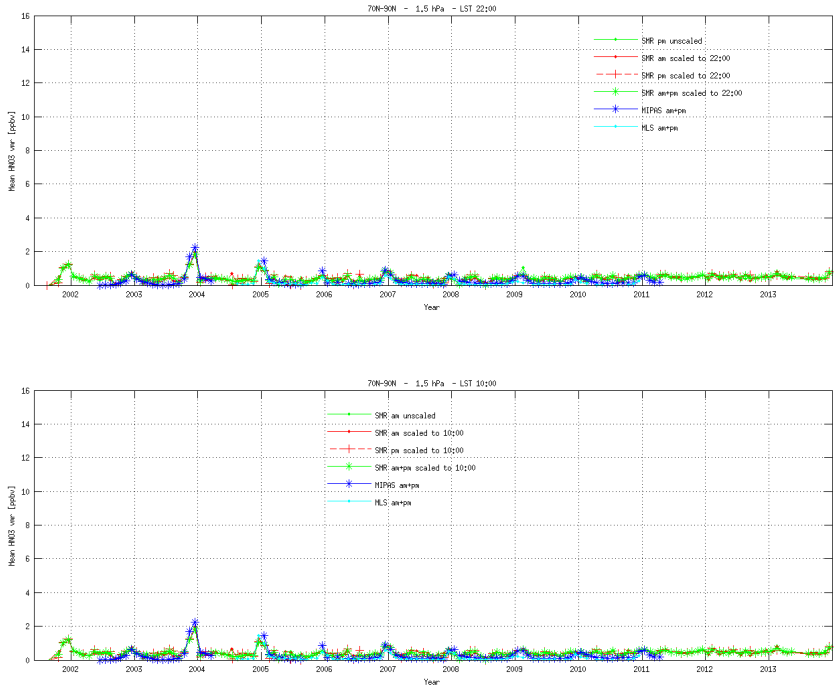

Figure 81: Comparison of scaled SMR (red), unscaled SMR (green), MIPAS (blue), and MLS (cyan) HNO_3 time-series for the 1.5hPa level and northern high latitudes. Top: night (10pm). Bottom: day (10am).

6. Summary and recommendations. Input to data product disclaimer.

The PVR presents the validation of the stratospheric temperature, water vapour, aerosols, ozone, and the evaluation of short lived species climatologies, developed in the SPARC-SA initiative (SPIN). These are the main results and inputs for the data product disclaimer for the data sets of the first phase of SPIN:

6.1. Stratospheric Temperature

Climatologies of zonal and monthly mean stratospheric temperatures from ACE-FTS, MIPAS and SMR have been produced and compared with several RO and reanalysis data sets. Generally, the ESA-based climatologies agree better with ERA-Interim and NCEP-CFSR than with the RO data from CHAMP and TSX, which are on average 1-2K warmer than ERA-Interim in the stratosphere. SMR is warmer by ~5 K at 20-25 km, colder by ~5 K at 35-45 km, and colder by ~20 K at 50 km, than ERA-Interim. Seasonal cycles are well represented. ACE-FTS is not suitable for merging to SSU near-global data because of limited spatial sampling, leading to excessive noise in near-global time series. SMR is not suitable for merging to SSU near-global data because of strange temporal behaviour in near-global time series. MIPAS however looks suitable, with stable temporal behaviour and similar seasonal cycles and inter-annual variability to SSU. CHAMP is not suitable for simulating the SSU channels because of sensitivity to filling and excessive noise in near-global time series, presumably arising from

the upper stratosphere which contributes to all SSU channels but for which RO does not give high quality temperature measurements. The SSU and SSU-weighted MIPAS channels exhibit relative biases during their overlap period 2002-2006; removing these relative biases by adjusting the SSU-weighted MIPAS data leads to an extended SSU near-global temperature record. For Channels 2 and 3 there is some evidence of a temporal inhomogeneity in the MIPAS temperature data between the pre- and post-2004 portions of the record, which would compromise the accuracy of the bias correction and hence of derived long-term changes. Such an inhomogeneity could plausibly have arisen from the change in spectral and vertical resolution in the MIPAS retrievals at that time. There is also a curious "hiccup" in the SSU-weighted MIPAS data in 2006 (outside of the overlap period), which is especially prominent in Channels 2 and 3. Both features would be worthy of further investigation as they currently compromise the value of the MIPAS temperature record for assessing long-term changes in the upper stratosphere.

6.2. Stratospheric Water Vapour

Stratospheric water vapour from SCIAMACHY agrees well with CFH balloon measurements, MLS and ACE-FTS around 15 to 18 km. Above these altitudes a dry bias of about 10-20% is found compared to these data sets, which is largest at the high latitudes in the Northern Hemisphere and smallest in the Tropics. At ~14 km, the comparison differs if the data used are interpolated or smoothed with the AVK of SCIAMACHY. This shows that the vertical sampling of SCIAMACHY and the regularisation used in the retrieval cannot reproduce the strong increase of water vapour in the troposphere at this altitude. Below, the differences do not agree for all regions and instruments. Here, the comparisons is challenging due to the high variability of water vapour in the Troposphere. For the water vapour retrieval, further investigations to improve of the aerosol correction are recommended.

6.3. Stratospheric Aerosols

6.3.1. OSIRIS

The OSIRIS Version 6.00 retrieval produces an aerosol particle size product that compares well with SAGE II measurements and produces 750 nm extinction coefficients that are in good agreement with both SAGE II and SAGE III. However, due primarily to the inability to measure the albedo at 1530 nm, the retrieved particle size has systematic errors related to the viewing geometry of the measurements. The variation of the Angstrom coefficient over a 6 month period is typically 10-15%, suggesting an error of at least this much in the Angstrom coefficient. In addition, the infrared channel is considerably noisier than the optical spectrograph data, resulting in a decrease in precision. Although sufficient to improve the 750 nm extinction results, it is recommended that the retrieved Angstrom coefficient be used as a qualitative measure of particle size rather than an absolute.

In addition, since the noise in the 1530 nm channel affects the precision of the 750 nm retrieval, it is recommended that the retrieved mode radius be used to produce a particle size climatology. This climatology will then be used with only the 750 nm measurements to retrieve the 750 nm extinction coefficient. This should improve the precision of the 750 nm extinction coefficient and reduce the low bias that is evident in the low altitude volcanic plumes as the saturation of the infrared detector will not be as critical.

6.3.2. SCIAMACHY

For the global average, SCIAMACHY and SAGE II agree well between 15 and 23 km. Here, differences up to 10% are found. Above 23 km SCIAMACHY underestimates the aerosol up to 30% in the global average. Larger, varying differences are observed for different 20° latitude bins. Further work is necessary to improve the phase function used in the SCIAMACHY aerosol retrieval.

6.4. Stratospheric Ozone

Stratospheric ozone profiles from GOMOS bright limb measurements agree well with ozone profiles from GOMOS night occultations, MLS and OSIRIS. The difference is less than 10% below 40 km. At 40 km there is a negative bias of 12-18% depending on the latitude. At 50 km there is a positive bias of up to 50% with high solar zenith angles. The negative bias above 50 km against the GOMOS night occultations can be explained by the diurnal cycle of ozone.

While GOMOS ozone profile measurements are more noisy that those from SAGE II, and show systematic biases against the measurements from SAGE II, they can be corrected for these systematic biases and filtered to remove outliers so that they become a valid data source for extending the SAGE II record. Comparisons of a merged SAGE II+GOMOS ozone data set with the Bodeker Scientific vertically resolved ozone profile database (Bodeker et al., 2013) indicate typical differences of ±5% through the stratosphere (20-50 km) with no obvious deterioration in the comparisons through the period that relies exclusively on GOMOS measurements.

6.5. Short-lived species climatologies

Comparisons between climatologies of short-lived species (e.g. BrO, ClO, HNO_3, HO_2, HOCl, NO, and NO_2) from SMR and OSIRIS with other satellite data have been performed as part of the SPARC data initiative. Differences in the observations due to diurnal variation of the regarded short-lived species can be explained by model simulations showing that the gas phase chemistry and reaction rate constants are fairly well understood. Scaled climatologies have been produced for a number of species such as ClO, NO, and NO_2 for the SPARC data initiative. Optimized scaled climatologies of Odin/SMR ClO, NO, and HNO_3 have been produced and were assessed by comparison with climatologies from other satellites such as Aura/MLS for ClO.