



Final documentation of data delivery system

Deliverable title	Final documentation of data delivery system
Deliverable number	D3.3
Revision	02
Status	Final
Planned delivery date	30/11/2014
Date of issue	28/10/2014
Date of review	7/12/2014
Nature of deliverable	Report
Lead partner	University of Bern
Dissemination level	Public

This work has received research funding from the European Community's Seventh Framework Programme ([FP7/2007-2013]) under grant agreement n°284421.



DOCUMENT PROPERTIES

	FUNCTION	NAME	ORGANISATION	DATE	SIGNATURE
LEAD AUTHOR	Research Associate	K. Hocke	UBern	28/10/2014	
CONTRIBUTING AUTHORS		all scientists of WP-3			
Reviewed by	NORS coordinator	M. De Maziere	BIRA-IASB	7/12/2014	

Table of Contents

Contents

I. INTRODUCTION	4
II. RAPID DATA DELIVERY SYSTEM (RDDS)	6
III. GEOMS HDF DATA FORMAT FOR RDDS	8
IV. ORGANISATION OF THE OPERATIONAL DATA DELIVERY	8
A. MICROWAVE RADIOMETER (MW)	13
B. UV/VIS	18
C. Lidar.....	25
D. FTIR.....	27
E. References	31

I. Introduction

The operational data delivery of NORS is performed by means of the Rapid Data Delivery System (RDDS). The final documentation describes all characteristics of RDDS which are needed for understanding and usage of RDDS. We focus on: 1) the acquisition of the ground-based remote sensing observations obtained by various measurement techniques, 2) characterizing of the available stations and instruments of NORS, 3) the rapid submission of the remote sensing data from the stations to RDDS at NDACC, 4) archiving of the rapid data at NDACC, and 5) rapid delivery of the remote sensing data for the purpose of a fast quality control of satellite measurements and the products of MACC. The final documentation intends to conserve the technical knowledge which we obtained during the NORS project. Technical procedures as the operational generation of HDF GEOMS files or the automatic submission of data files need to be documented, otherwise they would be forgotten within a short time interval. The documentation also serves as a manual for all newcomers, new users and new data providers of RDDS. To get an overview about the interconnections and the role of RDDS, Figure 1 shows a scheme of the Copernicus Earth Observation Programme.

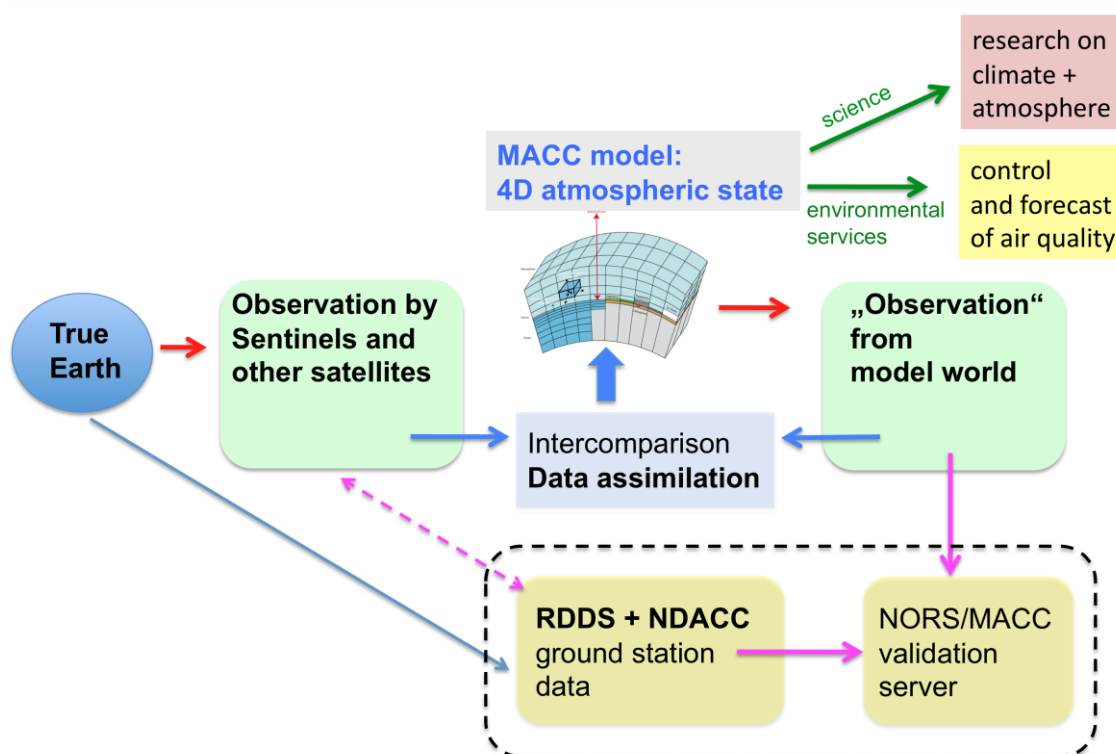


Fig. 1. The Rapid Data Delivery System (RDDS) and the NORS/MACC Validation Server (NVS) are building blocks of the Copernicus Earth Observation programme. NVS performs cross-validation of independent ground-based remote sensing data and colocated MACC products.

The Copernicus Earth Observation programme has in addition other aims such as land survey. Figure 1 is concentrated on the aim to get the most realistic estimate of the atmospheric composition,

dynamics and thermal state as function of space and time. The MACC model interpolates the satellite observations to a regular grid in 4-D by using the laws of physics and chemistry and the data assimilation technique. This approach makes the best usage of the satellite observations and theoretical knowledge included in the model to create the best analysis of the current atmospheric state as well as forecasts. Forecasts of aerosol concentration and reactive gases are of high importance for civilization and health. In addition, the improved knowledge of the 4-D atmospheric state leads to a vital progress in atmospheric sciences and climate change research. As shown in Figure 1, RDDS and NVS are essential for the validation and tuning of the MACC model. RDDS and NDACC data are also suited for a direct intercomparison with coincident satellite observations.

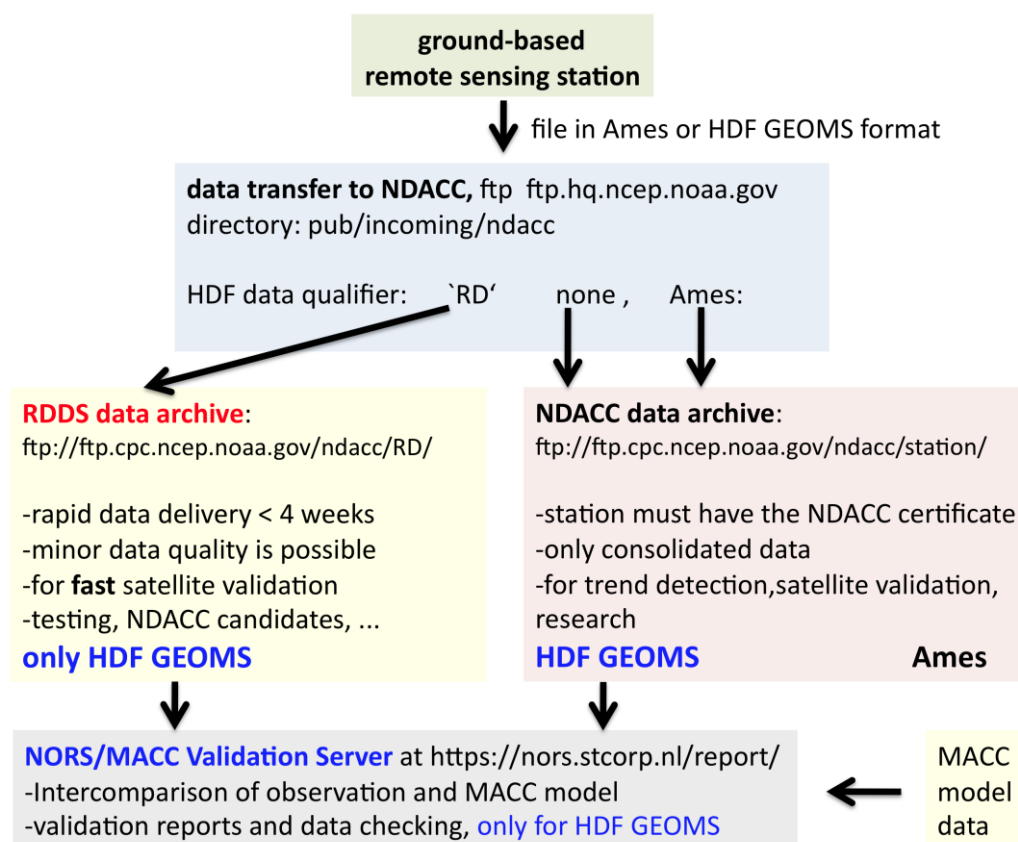


Fig. 2. Flow chart of data files from the station to the data archives of NDACC and RDDS. The NORS/MACC Validation Server automatically looks for the submitted HDF GEOMS files and builds time series of gas species using the latest and most reliable data versions (consolidated data are preferred against the rapid data). At present (November 2014), 15 stations are submitting HDF GEOMS files. The NORS Validation server can be accessed via nors-server.aeronomie.be.

Figure 2 shows a flow chart of the data files from the NORS or NDACC ground-based remote sensing station into the RDDS data archive. RDDS aims essentially at a fast validation of the satellite and the MACC products while NDACC focuses on high-quality data for research, process studies, the detection of long-term trends in atmospheric composition, etc. The detection of long-term trends requires a higher data quality in terms of long-term stability of the instrument and the retrieval

technique. That means, a station has to prove its quality before it can be accepted to become a NDACC station. Complementary to NDACC, RDDS is open to new stations, for example, the test of a NDACC candidate station can be performed within RDDS. The NORS/MACC validation server (NVS) visualizes the ground station data. NVS intercompares the observations with colocated MACC model data. Further, default validation reports are generated automatically by NVS.

II. Rapid Data Delivery System (RDDS)

The RDDS is a backbone of the NORS project since it collects, archives and distributes the ground-based remote sensing data of worldwide-distributed ground stations. RDDS provides the ground-based remote sensing data within four weeks after the measurement. A rapid delivery is important for a rapid control and quality assessment of the upcoming Sentinel satellite series. In addition, the output of the MACC products can be compared to the ground-based remote sensing data of RDDS. The intercomparisons are easily performed by using the NORS Validation Web Server. RDDS is hosted by the NDACC data center and takes advantage of the infrastructure of NDACC. By means of RDDS, the atmospheric community and NDACC have the chance to develop the next-generation network for monitoring of atmospheric composition change.

Another aspect of RDDS is the consolidation of the ground station data. RDDS is the platform and the data source for extensive cross-validation studies carried out by the NORS partners and other scientists. Such studies are permanently required for quality assessment and improvement of instruments and retrieval techniques. The NORS Validation Server already performs cross-validations between stations and the MACC model data. After consolidation or reprocessing of the data sets of RDDS, the data will be transferred into the NDACC data center which has higher standards in terms of accuracy and long-term stability of the measurement series. The transfer from RDDS to NDACC requires a new assignment of the variable "data qualifier" within the HDF GEOMS file and a new version number. The submission process to RDDS and NDACC is described in Figure 2. Generally, the NDACC data are appropriate for long-term trend studies and detailed cross-validations while RDDS allows a fast validation of satellite and MACC model data by means of the independent ground-based remote sensing data.

Index von ftp://ftp.cpc.ncep.noaa.gov/ndacc/RD/

 In den übergeordneten Ordner wechseln

Name

-  bern
-  bremen
-  gallegos
-  izana
-  jungfrau
-  nyalsund
-  ohp
-  reunion_maido
-  reunion_stdenis
-  xianghe

Fig. 3. HDF directories for HDF GEOMS files of Rapid Delivery (RD) data of RDDS in October 2014. The figure shows the path to those NORS stations of the Rapid Data Delivery System (RDDS) which submit RD data (ftp://ftp.cpc.ncep.noaa.gov/ndacc/RD/).

Since the set-up of the RDDS in month 6 of the project (Milestone MS3), some changes to RDDS have been discussed and implemented, essentially to improve clarity for the data user:

- (1) the rapid delivery data are labeled 'RD' instead of 'NRT', in order to be compliant with the common nomenclature for data delivery within one month. This implies that the DATA_QUALITY attribute in the RD data files must have 'RD' as the 2 first characters of the string.
- (2) a separate directory labeled 'RD' has been created in the NDACC database structure, for the submission of these RD data → see ftp://ftp.cpc.ncep.noaa.gov/ndacc/RD/
- (3) A 'RD_readme' file on the NDACC database (ftp-server <ftp://ftp.cpc.ncep.noaa.gov/ndacc/>) and an explanatory paragraph on the NDACC Webpages (www.ndacc.org) explain the contents and usage of this 'RD' directory to the data users. The consolidated data are submitted, as usual, to the station directories.

III. GEOMS HDF Data Format for RDDS

RDDS only accepts the GEOMS HDF data format which supports the intercomparison of coincident atmospheric composition measurements from different stations, satellites and models. For example, the averaging kernel matrices and a priori profiles of the retrieval are provided in the HDF GEOMS data files of ground-based FTIR and microwave radiometers. Another important evolution is related to the Data Format Definitions (see the Data Format Definitions document, D4.1): the development of the NORS Validation Server has highlighted the need for some variable name changes in the HDF templates. These changes have been discussed and agreed among all the NORS partners (and even more generally within the NDACC community) and have been updated accordingly on the AVDC Webpages. It has been agreed that the Rapid Data Delivery will adopt the latest versions of the templates.

IV. Organisation of the Operational Data Delivery

Table 1. Contact addresses of RDDS contributors

Alps/Bern, UBern	microwawe	Klemens Hocke	klemens.hocke@iap.unibe.ch simone.studer@iap.unibe.ch
Alps/Jungfraujoch, BIRA	DOAS	Michel Van Roozendael	michel.vanroozendael@aeronomie.be francois.hendrick@aeronomie.be
Alps/Jungfraujoch, BIRA	MAXDOAS	Michel Van Roozendael	michel.vanroozendael@aeronomie.be francois.hendrick@aeronomie.be
Alps/OHP, CNRS	O3 lidar	Sophie Godin-Beekmann	sophie.godin-beekmann@latmos.ipsl.fr maud.pastel@latmos.ipsl.fr
Alps/OHP/CNRS	DOAS	Andrea Pazmino	andrea.pazmino@latmos.ipsl.fr ariane.bazureau@latmos.ipsl.fr florence.goutail@latmos.ipsl.fr
Alps/Jungfraujoch/ULg	FTIR	Emmanuel Mahieu	Emmanuel.Mahieu@ulg.ac.be
La Reunion, BIRA	FTIR	Martine De Mazière	martine.demaziere@aeronomie.be bavo.langerock@aeronomie.be

La Reunion, CNRS	O3 lidar	Sophie Godin-Beekmann	sophie.godin-beekmann@latmos.ipsl.fr maud.pastel@latmos.ipsl.fr
La Reunion, CNRS	DOAS	Andrea Pazmino	andrea.pazmino@latmos.ipsl.fr ariane.bazureau@latmos.ipsl.fr florence.goutail@latmos.ipsl.fr
Ny Alesund	microwave	Mathias Palm	mathias@iup.physik.uni-bremen.de
Ny Alesund	MAXDOAS	Folkard Wittrock	folkard@iup.physik.uni-bremen.de
Ny Alesund	FTIR	Mathias Palm	mathias@iup.physik.uni-bremen.de
Izana, KIT	FTIR	Thomas Blumenstock	thomas.blumenstock@kit.edu
Izana, INTA	MAXDOAS	Monica Navarro	navarrocm@inta.es
Rio Gallegos	DOAS	Andrea Pazmino, Eduardo Quel	andrea.pazmino@latmos.ipsl.fr equel@citedef.gob.ar
Xianghe	MAXDOAS	Michel Van Roozendael	michel.vanroozendael@aeronomie.be francois.hendrick@aeronomie.be

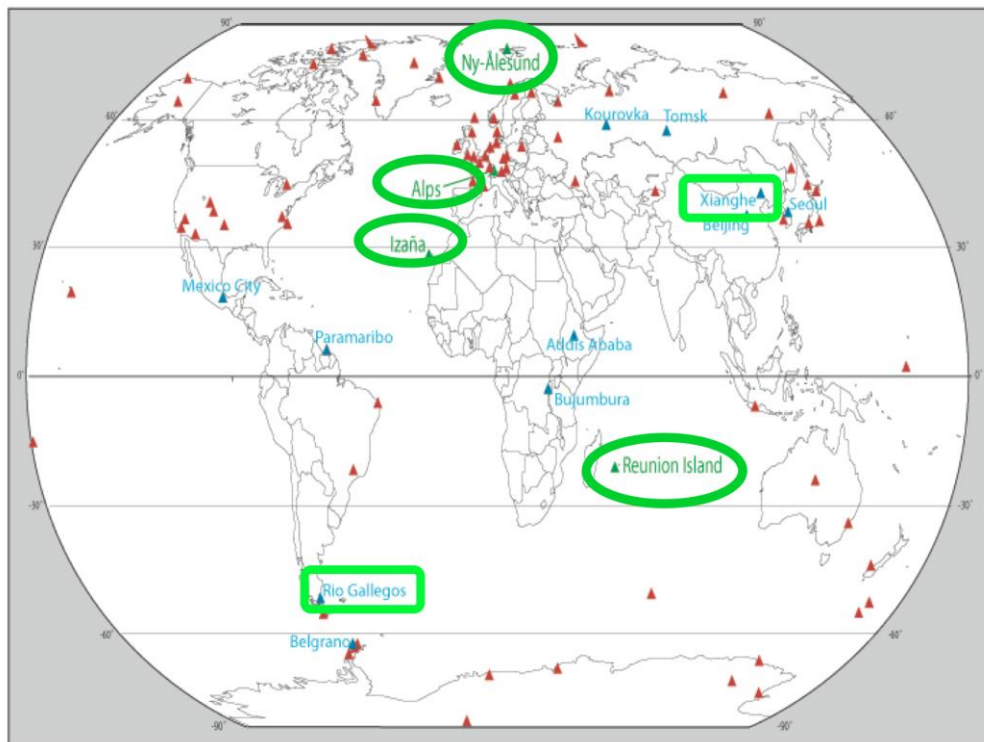


Fig. 4. NDACC network of ground-based remote sensing stations. The four stations with green ellipses are the original NORS network for rapid delivery. Meanwhile two new stations are contributing to the RDDS (green rectangulars)

Table 2. Trace gas and aerosol measurements of NORS/RDDS. The data products are column densities and/or vertical profiles of concentration or volume mixing ratio. Species in brackets mean that the data submission is not regular yet.

	DOAS	MAX DOAS	Lidar	MWR	FTIR
Ny Alesund	(O ₃ , NO ₂)			O ₃	(CH ₄ , CO)
Bremen					O ₃ , CO, CH ₄
Bern (Alps 1)				O ₃ , H ₂ O	
Jungfraujoch (Alps 2)	O ₃ , NO ₂				CH ₄ , CO, NO ₂ , O ₃
OHP (Alps 3)	O ₃ , NO ₂		O ₃		
Izana	O ₃ , NO ₂				CH ₄ , CO, NO ₂ , O ₃
Xianghe		aerosol, NO ₂			
Maido, La Réunion			(O ₃)		CH ₄ , CO, NO ₂ , O ₃ , HCl, HF, HNO ₃
St. Denis, La Réunion	O ₃ NO ₂				
Rio Gallegos	O ₃ , NO ₂				

Guide for Generation of GEOMS HDF-files for RDDS/NORS

1) The IDL program **idlcr8hdf.pro** and templates for meta data of various measurement instruments can be found at Aura Validation Data Center (AVDC):

[GEOMS](http://avdc.gsfc.nasa.gov/index.php?site=1178067684) (<http://avdc.gsfc.nasa.gov/index.php?site=1178067684>)

2) For running of the idlcr8hdf program, you should have access to the image data language IDL. Don't worry if you have no IDL license, the NORS HDF GEOMS files can be generated "manually" by means of IDL VM (virtual machine) which is freely distributed or already part of the newer IDL versions (8 and higher).

3a) Call "IDL VM" for interactive generation of HDF GEOMS files (conveniently with an alias in .bash_profile or .bashrc):

```
alias idlvm="/Applications/itt/idl/idl81/bin/idl -vm"
```

Please select the file "idlcr8hdf.sav" in the graphical user interface and so on. Then mark the "HDF 4" field. You can select a single data file or the data files of a whole month or a year which you should prepare in advance (as ASCII files according to the GEOMS guide lines and the metadata templates).

b) For rapid data delivery, it is better to make a batch or a shell script which calls idlcr8hdf.pro in IDL mode. Further details are provided upon request by Klemens Hocke (klemens.hocke@iap.unibe.ch), Mathias Palm (mathias@iup.physik.uni-bremen.de) or Ian Boyd (iboyd@astro.umass.edu).

4) Quality check. GEOMS HDF sample files can be manually checked at

[Tools/Quality Assurance\(GEOMS, AVDC\)](#)

5) GEOMS files are submitted to RDDS in the same manner as it is done for NDACC:

```
ftp ftp.hq.ncep.noaa.gov
```

```
user: anonymous
```

```
pwd: your email address
```

```
cd pub/incoming/ndacc
```

A. MICROWAVE RADIOMETER (MW)

GROMOS at Bern (Alps)

Technical information about the way the Microwave physics group at IAP, Bern, Switzerland handles operational maintenance, data processing, data archiving, etc. of the ozone radiometer GROMOS.

Labview interface

The labview interface of GROMOS gives us a (qualitative) quick-look on the measurement quality of the observed ozone spectra in real-time. Aside from the spectra, various other variables can be displayed. Tests on the instruments and other incidents are noted down in the Log-Tab.

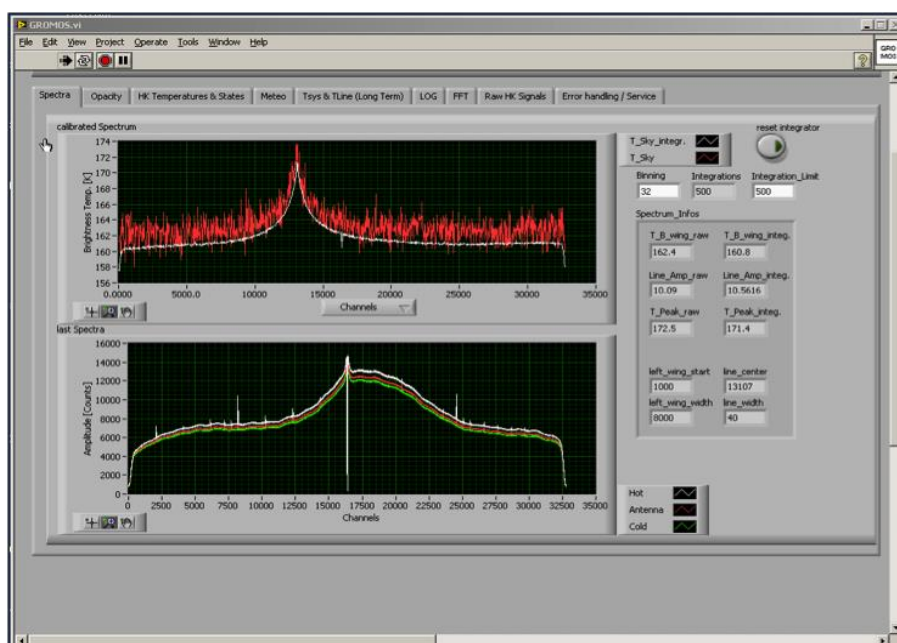


Fig. 5. Top: GROMOS ozone spectra (red: one calibration cycle, white: integrated spectra over approximately 2 h). Bottom: Raw measurements of the sky observation (red) as well as the hot and cold load (white, respectively green).

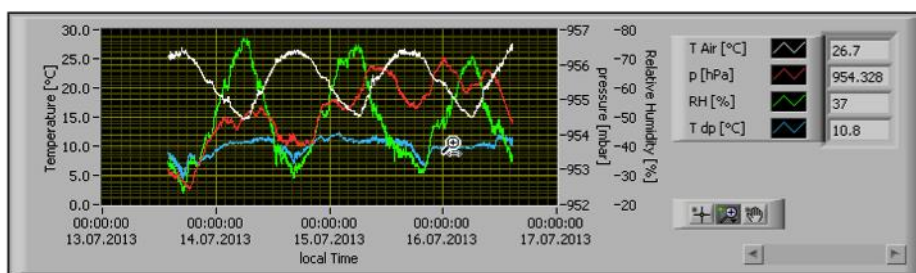


Fig. 6. Temperature, pressure and relative humidity are given in the „Meteo“-Tab. These atmospheric parameters are measured by a Meteostation positioned outside of the GROMOS room.

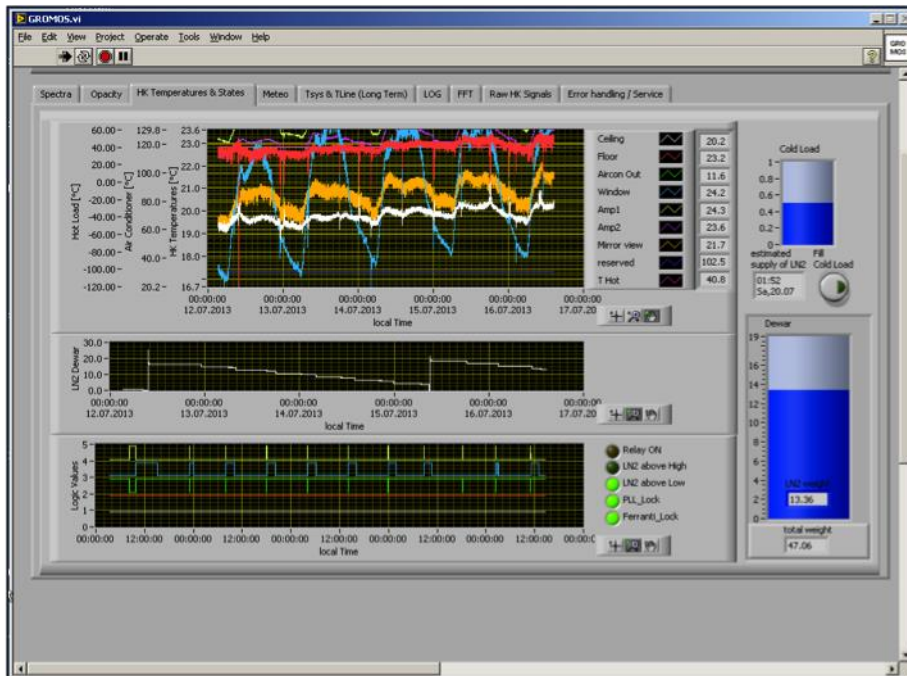


Fig. 7. Additional information on various temperatures of and around the GROMOS instrument as well as the liquid nitrogen level of the cold load is recorded.

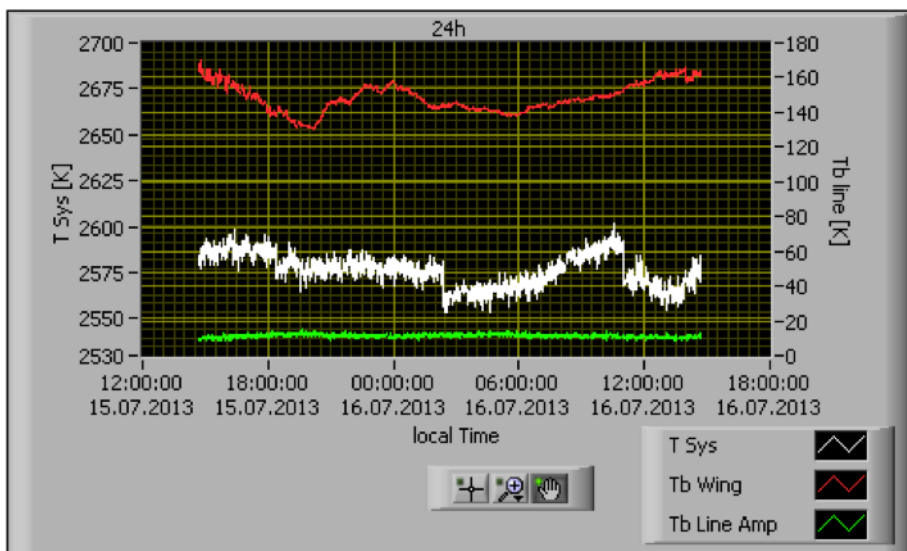


Fig. 8. System temperature T_{sys} of GROMOS is displayed on short- and long-term time scales (given above is the development over the course of 24 hours).

Archiving

Raw data are copied directly from the GROMOS computer to MIRADARA on a daily basis. Data are saved in a binary file, while additional important information such as measurement time, mirror position, etc. are saved in a text file.

MIRADARA consists of two identical, independent RAID 5 systems (QNAP TS-659) with each having (6-1) x 2 TB = 10 TB (Miradara-1 and Miradara-2). Miradara-2 is a copy of Miradara-1 and is regarded as a safety backup. The backup is done weekly.

Data processing

For processing of raw data to level1-files (L1: calibrated spectra) and level2-files (L2: retrieved ozone profiles), we use Matlab (Matlab 2011b). L1 and L2 data are saved in Matlab mat-file format.

Data analysis is also done with Matlab.

MySQL database

Additionally, we save our L2 data together with additional important information (such as integration time, opacity, system temperature, averaging kernels, etc.) in a MySQL database.

MySQL is an open source database management system and is available for Linux, Unix, and Windows platforms. Additionally, it allows a multi-user access. A further advantage is that reading out ozone profiles given a specific criteria (such as e.g. opacity values lower than a threshold) is easily done with a fast performance.

```
mysql> describe level2_header_vK;
```

Field	Type	Null	Key	Default	Extra
profile_ID	int(11)	NO	PRI	NULL	auto_increment
time	datetime	YES	MUL	NULL	
time_min	datetime	YES		NULL	
time_max	datetime	YES		NULL	
integration_time	mediumint(9) unsigned	YES		NULL	
timestamp	timestamp	NO		CURRENT_TIMESTAMP	on update CURRENT_TIMESTAMP
process	enum('yes','no')	YES	MUL	yes	
species	enum('03')	YES		03	
version	float	YES	MUL	NULL	
altitude	smallint(6)	YES		577	
latitude	float	YES		46.9511	
longitude	float	YES		7.4386	
zenith_angle	float	YES		50	
azimuth_angle	float	YES		40	
SZA	float	YES		NULL	
SZA_range	float	YES		NULL	
T_receiver	smallint(6)	YES		NULL	
trop_transmittance	float	YES		NULL	

Fig. 9. MySQL header-table of GROMOS level2 data.

Quality control and consolidation of MW data

The ozone profiles of MW are delivered to RDDS within four weeks. Before consolidation of the ozone data in the NDACC data center, the retrieved time series of ozone profiles are checked and manually controlled by the operator. In rare cases, ozone profiles of poor quality may come through the retrieval chain and into the RDDS. Such runaway data can be easily detected if the time series of a longer time interval, e.g., one year are checked by the operator.

The strategy of the re-analysis is

1) Collection of the ozone data of MW in RDDS for a time interval of about one year (RDDS data) before the consolidation is done

2) Quality control before delivery of the MW data to the consolidated data directory of NDACC:

a) Outlier removal: Remove possible outliers in the ozone series (the operator can mark such profiles with a cursor on the screen, and a flag will be set for invalid ozone profiles in the database. It might be just 3-5 profiles per year. However it is always good to look at the ozone series, - in addition to automatical threshold controls.)

b) Intercomparisons: Perform intercomparisons of the RDDS data and independent satellite data. This gives an estimate of the shape of the MW radiometer and retrieval. Uni Bern has an automatical visualization of the actual ozone series of GROMOS, Aura/MLS, and ECMWF on internet (only accessible for Uni Bern).

Supplementary informations on GROMOS FFTS (vK) data processing:

Step 1.: L0 to L2 => automated: done by cron-job on IAP-server

Run .../home/GROMOS_FFTS_2012/v8/ **direct_level0_level2.m**

⇒ Change the „date_to_start“ and „date_to_stop“ manually

(The routine calls the main subroutines for L0 to L1: level0_level1_f.m
and L1 to L2: level1_level2_K.m)

Step 2.: Insert data into the mysql-tables => automated: done by cron-job on IAP-server

Run .../home/GROMOS_FFTS_2012/v8/ **insert_FFTS_into_mysql_vK.m**

⇒ You will be asked to enter the start and stop date in the format „YYYYMMDD“

⇒

Step 3.: Make daily ascii-files from L2 data

Run .../home/GROMOS_FFTS_2012/v8/ geoms_hdf/ **main_gromos_FFTS_2txt_geoms.m**

⇒ Change the „firstday“ and „lastday“ manually

Step 4.: Make GEOMS-hdf files (with IDL Virtual Machine)*

Run the following commands in terminal (see alias in startup-file)

> geoms

> idlvm

In the interactive window of idlvm that opens :

1. ... click idlcr8hdf.sav & ok (enter)
2. ... choose „Append log output ..“ and „If AVK kernel date ..“
... then click HDF4 & ok (enter)
3. ... click GEOMS-GEN-MWR-O3-Bern.meta & ok (enter)
4. ... ascii files (from Step 3.) are shown: mark all or select days & ok (enter)
5. ... click tableattrvalue_04R004_idl.dat & ok (enter)
6. ... click where hdf-files should be saved: „out“-directory e.g. & ok (enter)

When DONE

- copy data manually from „out“ to „data_vK/YYYY
- and delete „simone_geoms/out/... .hdf-files“ as well as „simone_geoms/... .data“

Step 5.: Transfer GEOMS-hdf to ftp-server of NDACC for rapid data delivery (RD)

Upload data to <ftp://ftp.hq.ncep.noaa.gov/pub/incoming/ndacc/>

(e.g. by connecting (via filezilla) to ndacc_RD_FFTS)

OZORAM at Ny Alesund

The instrument has been described and compared to satellite measurements in (Palm, 2010).

Technical description and processing chain:

The OZORAM is a heterodyne millimeterwave receiver at 142.175 GHz. The frontend consists of a cooled Schottky mixer and a low noise amplifier. The backend is an Aquiris FFT spectrometer (AC240). It has 16384 channels at a width of 61 kHz each.

It has an internal cold load, which can be checked using a blackbody submerged in liquid nitrogen.

The instrument runs automatically and is driven by a computer program written in C++. The raw data are saved locally at Ny Alesund and transferred to the IUP at the Universität Bremen hourly. At the IUP the data are stored on a RAID system and about once a year secured on tape. The local operator writes a protocol once a day and performs a check of the cold load once a week. The results of those checks are kept for future reference.

At Bremen the data are analysed and uploaded to the NORS data base once a day with a quick check of the data operator, the final analysis and correction will be done once a year or on request.

Quality control:

The retrieved data are marked valid using the following conditions:

1. retrieval has converged
2. line center of the O₃ emission is properly fitted, failure to do so points to an instability in the system
3. the atmospheric background is below 200 K
4. the standing waves are below 10 times of the noise level
5. the retrieved water vapor column is positive
6. there are no strong negative overshoots in the retrieved profile

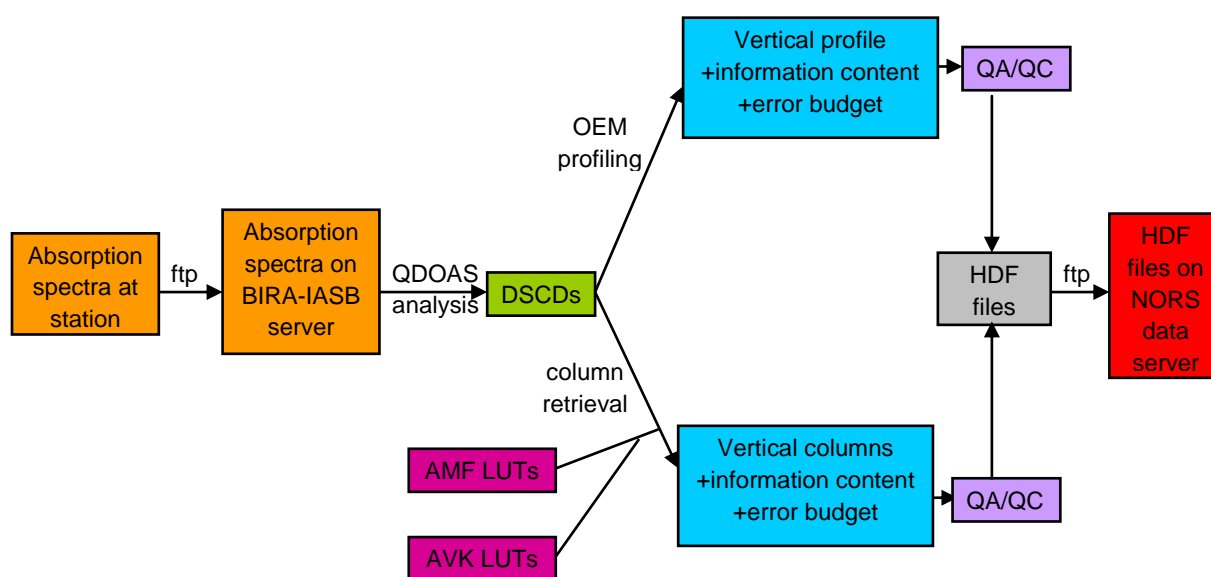
The averaging kernel matrix and the a priori profile are saved with the retrieved profile. The averaging kernel matrix encodes the horizontal averaging has to be used if the OZORAM measurements are compared to independent measurements or model data. The horizontal averaging over the field of view is considered negligible and not taken into account.

Automated processing chain for NORS:

- dedicated computer for analysis of spectra from OZORAM and FTIR for NORS
- get spectra from Ny Alesund
- create PT files from ECMWF
- calibrate and integrate spectra
- analyse spectra using the ARTS/QPACK retrieval software.
- create daily GEOMS-HDF
- plots of recent retrieval will be sent to the operator for quick checking
- submit in GEOMS format to the NDACC database

B. UV/VIS

Alps/Jungfrauoch (DOAS and MAX-DOAS) and Xianghe, China (MAX-DOAS)



The BIRA (MAX-)DOAS rapid data delivery system is described in the flow chart here above. It includes the following steps:

1. Transfer by ftp of the raw absorption spectra from the on-site computers controlling the spectrometers to the BIRA server.
2. The DOAS spectral analysis providing the trace gas differential slant column densities (DSCDs) is performed using the QDOAS software (<http://uv-vis.aeronomie.be/software/QDOAS/>) developed by BIRA and widely used in the DOAS community.
3. At this stage, two different approaches are followed depending of the measurement type:
 - 3.a. MAX-DOAS measurements: tropospheric profiles and columns of trace gases (NO_2 , HCHO) and aerosols are retrieved for each MAX-DOAS scan by applying the bePRO profiling tool developed at BIRA (Cl  mer et al., 2010) to the measured DSCDs. This algorithm, written in FORTRAN, is based on the Optimal Estimation Method (Rodgers, 2000) and the forward model is the LIDORT radiative transfer model (Spurr, 2008). A quality control test is applied to each scan based on the DOFS (degree of freedom for signal) of the retrieval (DOFS should larger than 0.7) and on the RMS of the difference between measured DSCDs and those calculated with the retrieved profiles (the RMS should be lower than a threshold value).
 - 3.b. Zenith-sky DOAS measurements: the retrieval of stratospheric NO_2 and total O_3 vertical columns from measured twilight DSCDs consists of three steps: (1) determination of residual amount in the reference spectrum using the Langley-plot approach, (2) conversion of absolute slant columns into vertical columns using appropriate AMFs, and (3) averaging of the vertical columns over the 86-91   SZA range. The three steps are performed using dedicated MATLAB routines. AMFs are extracted from the NDACC UV-vis look-up tables (<http://www.ndacc.org/>) using a FORTAN routine. A similar approach is used for the extraction of the column averaging kernels to be included in the HDF data files. In case of NO_2 , measurements contaminated by tropospheric pollution events are rejected based on the Langley-plot method (data points not lying on a “smooth” line indicate a tropospheric contamination).
4. The GEOMS HDF files are created with the idlcr8hdf IDL routine. The ascii input data files needed for this step are generated using dedicated MATLAB routines. Data format quality control is made by the idlcr8hdf code itself.

Data are automatically processed and submitted to the NORS/NDACC RD database by ftp on a daily basis using a bash script that calls the different MATLAB, FORTRAN, and IDL components described in steps 2-4. The daily submissions of Jungfraujoch DOAS and Xianghe MAX-DOAS data files are operational since end of March 2013 and beginning of June 2013, respectively. Xianghe is located close to Beijing and it is one of the NORS Capacity Building Stations (see WP10). It should be noted that the MAX-DOAS retrieval at Jungfraujoch is still under development.

The persons responsible for the BIRA (MAX-)DOAS rapid data delivery system are Fran  ois Hendrick and Caroline Fayt (contact: franch@oma.be).

UV/VIS, DOAS at Rio Gallegos

The DOAS at Rio Gallegos provides zenith-sky hdf files (O₃ and NO₂) to RDDS since July 2014.

The contact person for the DOAS instrument at Rio Gallegos is Andrea Pazmino (andrea.pazmino@latmos.ipsl.fr) while the contact person in Argentina is Eduardo Quel (equel@citedef.gob.ar).

The station has the webpage: <http://www.division-lidar.com.ar>

The station is located in south of Patagonia in Argentina

Latitude: -51° 36' 02,03"

Longitude: -69° 19' 09.96"

There are many other instruments. For example, a DIAL ozone lidar which is also part of NDACC.

UV/VIS, DOAS, Alps/OHP and La Reunion

SAOZ team has identical procedure for Alps/OHP (DOAS) and La Reunion (DOAS) zenith sky UV-visible instruments

In the following procedure XX should be replaced by OH for Alps/OHP (DOAS) and RE for La Reunion (DOAS)

Step 1.: L0 to L1 => transfer automated

- a) Level 0 (L0) daily "SXXYYYYMMDDHHmm.efm" files are analysed to obtain Level 1 (L1) daily "SXXYYYYMMDDHHmm.mrs" files
 - S for Ground-Based SAOZ
 - XX corresponds to the station name
 - YYYY year
 - MM Month
 - DD Day of month
 - HH UT Hour
 - mm minutes
 - efm is the extension of the binary file containing the spectra (L0)
 - mrs is the extension of the ASCII file containing the analysis (Slant columns) (L1)
- b) The L0 and L1 are recorded on station on instrument Hard disk
- c) Daily L0 and L1 files are transferred at 12:00 PM to station-server and copied on local station data base
- d) They are transferred every morning by FTP (OHP) or Rsync (La Reunion) at 5:00 AM to LATMOS-server (located in Jussieu, Paris, France).
- e) Then they are transferred every morning at 7:00 AM to NDACC-France server (located at both places in LATMOS-Jussieu, Paris, and LATMOS-Guyancourt, France)
- f) Finally they are copied every morning at 10:00 AM from LATMOS-server to SAOZ data base located on "Espace" computer in LATMOS-Guyancourt, France. This computer is automatically back-up by LATMOS computer team.
- g) The L0 files are located on espace/saoz/sol/0/OHP/YYYY for OHP
The L0 files are located on espace/saoz/sol/0/Reunion/YYYY for La Reunion
- h) The L1 files are located on espace/saoz/sol/1/OHP/YYYY for OHP

The L1 files are located on espace/saoz/sol/1/Reunion/YYYY for La Reunion

Step 2.: L1 to L2 => O3 and NO2 stratospheric columns automated by SAOZ-server

In this step, the retrieval of stratospheric NO₂ and total O₃ vertical columns from measured twilight DSCDs (L1 files) consists of three steps: (1) determination of residual amount in the reference spectrum using the Langley-plot approach, this has been done once, as for SAOZ a single reference spectrum is used for each instrument, (2) conversion of absolute slant columns into vertical columns using appropriate AMFs, and (3) averaging of the vertical columns over the 86-91° SZA range. The three steps are performed using an IGOR routine. AMFs are extracted from the NDACC UV-vis look-up tables (<http://www.ndacc.org/>) using a FORTAN routine. A similar approach is used for the extraction of the column averaging kernels to be included in the HDF data files. O₃ and NO₂ data are filtered to avoid saturated spectra or spectra perturbed by multiple scattering or heavy rain.

- a) Level 1 (L1) (daily *.mrs files) of the year are put together to make an annual file named “**an_en_cours.mrs**” located at the same place on “Espace” computer

This is done by automatic task

Fab_mrs_an_d.cmd at 11:30 AM

- b) Level 2 (L2) are generated every day using the following IGOR software

RETRAITE_all.pxp at 11:50 AM

The annual L2 ascii file containing O₃ and NO₂ stratospheric columns is named

O3_YYYY.XX

The header of the files indicates which AMF have been used, the units and the name of the columns

Climatological AMF used

Units --- O₃ Dobson, NO₂: 1e15 mol/cm²

Year Month Day DoY O₃sr O₃ss dO₃sr dO₃ss NO₂sr NO₂ss dNO₂sr
dNO₂ss

- c) This RETRAITE_all.pxp software generates also the following annual files

I. AMF_O3_YYYY.XX

Daily Climatological AMF at 90SZA calculated

using **o3_amf_interpolation_v2_0_dos.exe**

The AMF is calculated at 510 nm

II. AMF_NO2_YYYY.XX

Daily Climatological AMF at 90SZA calculated

using **no2_amf_interpolation_v1_0.exe**. The AMF is calculated at 470 nm

III. AVK_O3_YYYY.XX

Daily averaging kernel O₃ profile calculated

using **o3_avk_interpolation_v1_0_dos.exe**

IV. AVK_NO2_YYYY.XX

Daily averaging kernel NO₂ profile calculated

using **no2_avk_interpolation_v1_0_dos.exe**

V. Mrs.pxp

This file (which is can be read by IGOR only) contains all level1 results plus, for each measurement, the NO2 diurnal variation calculated using the following software
[NO2_extraction_tool_v2.exe](#)

- d) The L2 files are located on espace/saoz/sol/2/OHP/YYYY for OHP
 The L2 files are located on espace/saoz/sol/2/Reunion/YYYY for La Reunion
- e) There are also a software analysis report and plots which can be used for quality check.
- f) Plot of the level 2 ozone and NO2 stratospheric columns are generated using the following IGOR software
[Graph_all.pxp](#) at 12:30 AM
 The graphs are named [L1_XX_YYYY.png](#)
 The graphs are used for quality control check.
 The ascii file [O3_YYYY.XX](#) and the graph [L1_XX_YYYY.png](#) are transferred every day at 12:55 on a WEB server accessible to all users using the [ftp_web_saoz.cmd](#) .
 The name of the web server is <http://saoz.obs.uvsq.fr/SAOZ-RT.html>

Step 3.: L1 to L2 => NO2 tropospheric columns automated by SAOZ-server

***** *Yet, this is not done automatically* *****

- a) Tropospheric columns are generated every day using the following software
[TRTniv2tropo.pxp](#) at 12:00 AM
 This file contains the tropospheric vertical columns every 2-5 minutes.
 This software needs stratospheric columns O3_YYYY.OH and NO2 diurnal variation output from RETRAITE_all.pxp stored in mrs.pxp .
[NO2troponeu_YYYY.XX](#)
- b) This software generates also daily graphs of the NO2 tropospheric column which are used for quality control check.
 The ascii files and the graph are transferred every day at 13:25 on the same WEB server accessible to all users using the [ftp_tropo_saoz.cmd](#).
 The address of the server is http://saoz.obs.uvsq.fr/SAOZ_tropo_OHP.html for OHP and
http://saoz.obs.uvsq.fr/SAOZ_tropo_Reunion.html for La Reunion

Step 4.: Make GEOMS-hdf files (with IDL Virtual Machine)*

- a) Stratospheric Ozone
 Using the annual ascii files

“O3_YYYY.XX” ,
“AMF_O3_YYYY.XX”
and “AVK_O3_YYYY.XX”
monthly HDF files are generated using the following IDL software
[write_O3saozgb_nors.pro](#)
[idlcr8hdf.pro](#)

b) Stratospheric and tropospheric NO2

From the annual ascii files

“O3_YYYY.XX”
“AMF_NO2_YYYY.XX”
and “AVK_NO2_YYYY.XX”
and “NO2troponew_YYYY.XX”
monthly HDF files are generated using the following software

[write_NO2saozgb_nors.pro](#)
[idlcr8hdf.pro](#)

Step 5.: Transfer GEOMS-hdf to ftp-server of NDACC for rapid data delivery (RD)

- a) Manual delivery:
Upload data to <ftp://ftp.hq.ncep.noaa.gov/pub/incoming/ndacc/>
By connecting manually (via filezilla) to ndacc_RD_UVVis
- b) Automatic transfer to <ftp://ftp.hq.ncep.noaa.gov/pub/incoming/ndacc/>
Using TBD software
This task will be done on the 5th day of each month.

Personal

- a) PI team (Andrea Pazmino, Manuel Pinharanda, Florence Goutail, Jean-Pierre Pommereau) are responsible for data real time analysis and provision of stratospheric and tropospheric columns
- I. Step 1 a), b), c), f)
 - II. Step 2 a), b), c)
 - III. Step 3 a)
- b) Michèle Thetis is responsible for data transfer and L0 and L1 data storage on NDACC-France data base
Step 1 d) and f)

- c) Ariane Bazureau is responsible for HDF files making and transfer on NDACC/RD data base
 - I. Step 4 a) and b)
 - II. Step 5 a) and b)
- d) Maud Pastel is responsible for contact between SAOZ team and NORS project

UV-VIS DOAS and MAXDOAS Izaña Station

The following procedure describes the necessary steps for the Rapid Data Delivery System for DOAS/MAXDOAS instrument at Izaña station

1. Automatic analysis at station PC

At the station, the absorption spectra analysis for the O₃ and NO₂ DSCD using NDACC settings is performed using LANA software after spectra acquisition. LANA software is a DOAS analysis software developed at INTA. In case of ZENITH (DOAS measurements during twilights) measurements, the stratospheric VCD is also obtained using O₃ and NO₂ AMFs extracted from the NDACC UV-Vis look-up tables using 'o3_amf_interpolation_dos.exe' and 'NO2_amf_interpolation_dos.exe' routines. Analysis settings, including reference spectra, are periodically changed on base of the quality control check of the instrument and measurements. Results of analysis are stored in two daily (am and pm) files for ZENITH measurements and two daily (am and pm) files for OFFAXIS (MAXDOAS measurements out of twilights) measurements. These files are in ASCII format.

2. Automatic FTP transfer.

At 9.30 AM every day, the previous day results files are automatically transferred by ftp to a external INTA server and at 10.45 AM every day, to a internal INTA server.

3. Conversion to GEOMS HDF format.

There are two different processes for ZENITH or OFFAXIS data files results:

ZENITH: A software routine converts the stratospheric NO₂ and O₃ VCD data to the averaging VCD at 90° SZA, extract the column averaging kernels from a look-up table using 'no2_avk_interpolation_v2_0.exe' for NO₂ and 'o3_avk_interpolation_v2_0.exe' for ozone and generate the ASCII files with the auxiliary information and the metadata according to the generic metadata template GEOMS-TE-UVVIS-DOAS-ZENITH-GAS-006. IDL IDLCR8HDF software converts these files to the HDF GEOMS format. Two HDF files for each day are generated.

OFFAXIS: A daily filter is applied to the am and pm OFFAXIS files according to the aerosol optical depth obtained from the Izaña CIMEL. Days with AOD > 0.1 are rejected. Conversion to surface O₃ and NO₂ mixing ratio at Izaña level (2370 masl) applying the Gomez et al. (ACP 2014) methodology. A software routine converts those files to generate the ASCII files and metadata according to the generic metadata template GEOMS-TE-UVVIS-DOAS-OFFAXIS-GAS-006. IDL IDLCR8HDF software converts these files to the HDF GEOMS format. Two HDF files for each day are generated.

5. Manual ftp transfer to NDACC database

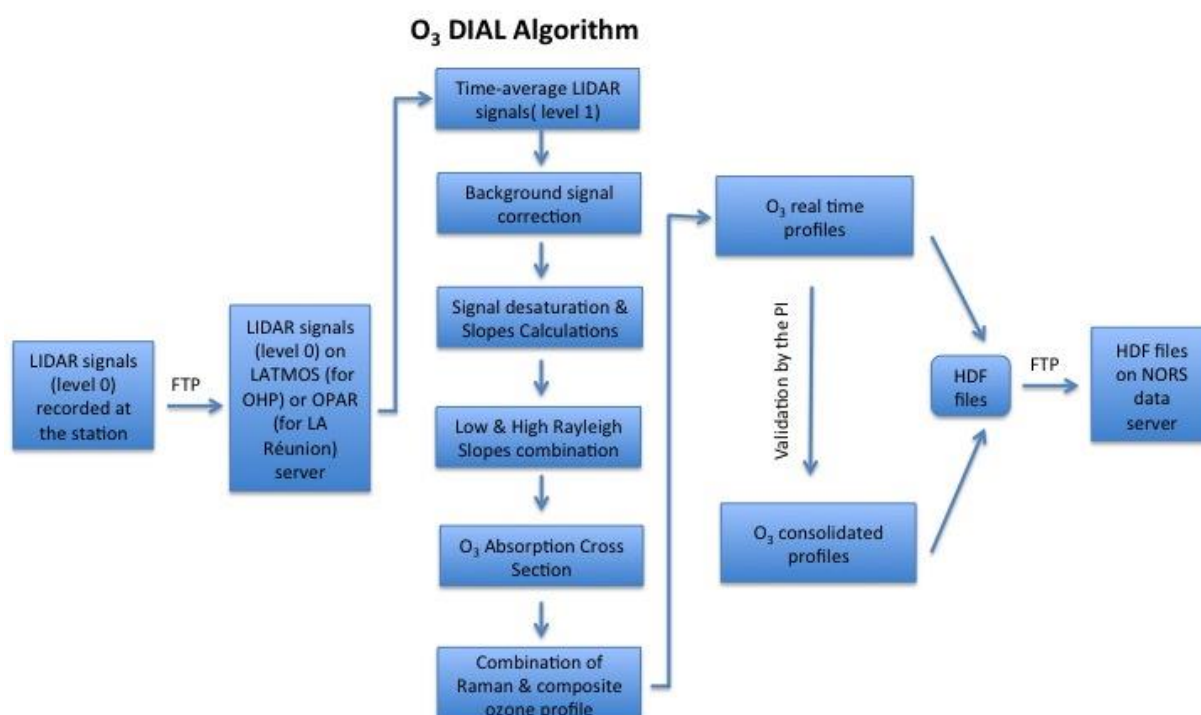
Uploading of the daily HDF files to the NDACC database is done through ftp by hand every four weeks.

The person responsible for the INTA DOAS and MAXDOAS rapid data delivery system is Mónica Navarro Comas (contact: navarrocm@inta.es).

C. Lidar

Alps/OHP, La Réunion and Argentina (LIDAR)

Since June 2013, the LIDAR team has identical data processing and delivery procedure for Alps/OHP and La Reunion LIDAR instruments. A similar procedure is being implemented at the Rio Gallegos in collaboration with the CEILAP team. The procedures are described in the flow chart below:



It includes the following steps:

1. Daily transfer by ftp of the raw lidar signals (level 0) from the on-site computers controlling the LIDAR to the LATMOS or OPAR server connected to the stations.

2. The DIAL algorithm follows basically the theoretical derivation of the ozone number density from the lidar signals. The DIAL software in Matlab has been developed by LATMOS and used by the lidar team at La Réunion. The programme operates every day at 10 pm. O3 profiles are retrieved following 6 steps : 1) Temporal signal averaging over the whole period of measurements (3–4 h in general) in order to increase the signal-to-noise ratio (level 0 to level 1). 2) Correction from : background and dead time in the case of photon counting acquisition, due to the saturation of the photon counting systems with high intensity signals originating from the lower ranges. 3) Combination of the slopes of the low-energy and high-energy Rayleigh signals and Raman signals. 4) The correction of Raman and composite Rayleigh profiles from the Rayleigh extinction using ancillary composite temperature-pressure profile. For example at OHP, Daily P and T from radiosoundings at the nearby Nîmes meteorological station + NCEP + MAP85 is used for the consolidated data. The Nîmes radiosoundings are available to the users only 2 months after the measurements, so for the near-real time data retrieval, only the NCEP data used. They are downloaded every day on the latmos server via FT. 5) Derivation of the ozone absorption cross section profile using the aforementioned temperature profile. 6) Computation of the lidar ozone number density profile, using the 2 pairs of Rayleigh signals and the pair of Raman signals. Summary plots are generated and the retrieval results are visually checked for validation.
3. The GEOMS HDF files are created with the idlcr8hdf IDL routine. The ascii input data files needed for this step are generated using IDL routine. Data format quality control is made by the idlcr8hdf code itself.
4. Transfer GEOMS-hdf to ftp-server of NDACC for rapid data delivery (RD)
Manual delivery:
Upload data to <ftp://ftp.hq.ncep.noaa.gov/pub/incoming/ndacc/>
By connecting manually (via filezilla) to ndacc_RD_UVVis

Automatic transfer to <ftp://ftp.hq.ncep.noaa.gov/pub/incoming/ndacc/> using a bash script that calls the different MATLAB and IDL components described in steps 2 and 3.

This task will be done by the end of each month.

Personnel

Maud Pastel is responsible for the NRT data analysis at OHP and OHP HDF files transfer to the NDACC/RD database (contact: maud.pastel@latmos.ipsl.fr). She is also responsible for the contacts with La Réunion and Argentina lidar teams regarding NORS.

PI team (Sophie Godin-Beekmann for OHP and la Reunion and Thierry Portafaix for la Reunion) are responsible for the final data analysis and provision of O3 stratospheric profiles to the NDACC data base.

Michèle Thetis is responsible for the data Level 0 transfer on NDACC-France database

Guillaume Payen is responsible for la Réunion HDF files making and transfer on NDACC/RD data base (guillaume.payen@univ-reunion.fr)

D. FTIR

Alps/Jungfrau joch

The raw spectra are downloaded from the Jungfrau joch while the NCEP data are retrieved from the NDACC data base. Both operations are performed automatically on a daily basis.

About every 10 days, the retrieval of the RD species are performed (SFIT 2 v3.91) in batch mode and the hdf archives are generated using the idlcr8hdf.sav utility and an IDL VM. The archives are then uploaded manually on the NDACC data base.

In addition, we also perform systematic retrievals of CO₂ and N₂ to appraise the quality and overall consistency of the available spectra (e.g. N₂ fits can reveal misalignment problems, CO₂ outliers issues with the time assignment for a given observation). Apart from these tests, we do not conduct any additional QA/QC evaluations in the RDD chain.

La Reunion

- The ancillary data (6-hourly pressure, temperature and H₂O VMR profiles) is obtained from the NCEP archive using the mod_maker8.pro IDL script
- Measured solar absorption spectra are filtered according to different criteria:
 - o Solar tracking stability
 - o Variation in detector DC level, as an indicator of the stability of the incident light on the detector
 - o Variation of solar intensity, as an indicator of the absence of clouds
- ⇒ These criteria should guarantee that the interferogram has been recorded under stable conditions.
- The target species VMR profiles are retrieved using SFIT2 (in the future this will become SFIT4) from the filtered set of spectra, using the NDACC retrieval strategy. Before submission to NDACC, each retrieval result should pass a quality check, which is based upon the following criteria (thresholds depend on the retrieval strategy and are based on statistics from previous analysis):
 - o The retrieval iterative computation is converging
 - o RMS (difference between measured and calculated spectrum) is below a threshold
 - o The degrees of freedom of the retrieval should be above a threshold.
 - o The error on the target total column should not exceed a threshold
- The GEOMS hdf file is created with idlcr8hdf. The reported pressure, temperature and H₂O profiles in the hdf file originate from the NCEP archive, while the surface pressure is obtained from on site measurements (pressure measurements are gauged to the height of the

interferometer). The reported H₂O profile might also be a retrieved profile, if such a simultaneous retrieval is part of the retrieval strategy for the target molecule.

- Before the data is submitted, summary plots are generated and the retrieval results is visually checked for consistency.
- Data will be processed on a weekly basis.

The automatic routines that call the different components in the RD submission process are a combination of python, matlab and bash scripting.

Izana

The retrieved profiles are uploaded to the NDACC database within four weeks. Before uploading, the measured raw spectra will be calibrated, the profiles retrieved and the output converted into HDF format.

Atmospheric monitoring and quality control

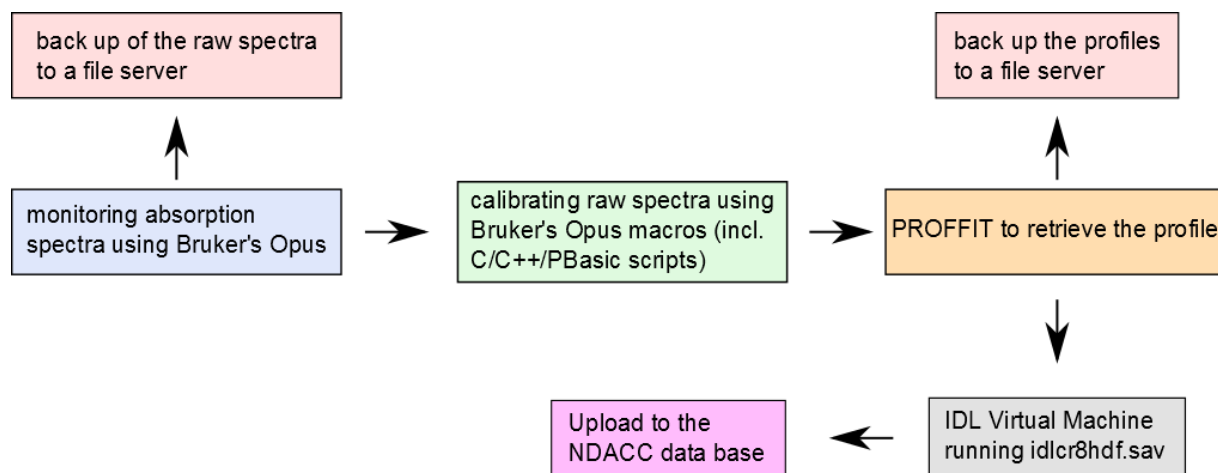
The calibration procedure aims to create spectra that show the absolute radiation intensity falling on the detector. Spectral detector- and spectrometer-sensitivities are corrected, as well as various detector-gain-settings and aperture sizes. To do the absolute radiation calibration, a background (Black Body) measurement is made in regular periods. In order to make sure that the line shape is reproduced correctly by the instrument, cell measurements are made and analysed with LINEFIT (software). Before calibration, the measured interferograms are checked and manually controlled by the operator. Additionally, poor quality spectra are discarded during the calibration routine. The retrieved data is marked valid if the retrieval has converged. The averaging kernel matrix is saved with the retrieved profile. Only in some rare cases, profiles of poor quality may come through the calibration and retrieval chain.

Hardware, Software and data processing chain

Since 2005, a Bruker IFS 125HR is operated on Tenerife Island. For the Operational Data Delivery an Intel Core i7 3770 CPU@3.4GHz machine with 8GB RAM running Win7(64-Bit) with the following software is used:

- Bruker's Opus Software (incl. macros)
- PROFFIT
- IDL Virtual Machine (running idlcr8hdf.sav)
- FileZilla or WS_FTP Software for the upload
- glue language scripts (written in python for the automation)
- diverse C/C++/PBasic scripts

The general processing chain can be displayed in the following way:



Manual delivery

Calibration procedure of raw spectra

The calibration is made by two Bruker's Opus Software macros calA.mtx and calB.mtx. The results are saved in a specific binary format that will be read by the retrieval software PROFFIT.

The calA.mtx macro file performs the black body calibration, DC-correction and baseline correction for each individual spectrum. It leaves the spectra in the Opus file format. It requires:

- callist2.erg file (path to date folder(s))
- cal.inp (path to date folder including pT profiles, path to black body spectra, calibration parameters)

The pT profiles are downloaded every day from '[http://acdb-ext.gsfc.nasa.gov/Data services/met/pro_les](http://acdb-ext.gsfc.nasa.gov/Data%20services/met/pro_les)' using a cron job which also checks if all values are reasonable.

After running calA.mtx, calB.mtx converts the output to the specific binary format for PROFFIT.

Using PROFFIT to retrieve species

PROFFIT is the retrieval tool used at IMK-ASF to analyze the solar absorption spectra. A detailed description would go beyond the scope of this report, so we restrict the description to the minimum.

In general, bat-su96.inp and bat-ex96.inp are used to set up the wanted jobs before running bat-su96.exe and bat-ex96.exe to retrieve the species.

Generating and uploading HDF files to the NDACC database:

To prepare the generation of HDF files, the following input files should be located in \hdf\input_ndacc:

- hdfprp4a.inp (path to date folder, first/last date to be taken into account, Name of Job, etc.)
- hdfprp4b.inp (list of supported keywords, reference file for metadata file)
- hdfprp4c.inp (list of keywords for data file generation, scaling factors)
- *.mdr (in this case, reference meta data file)

Running \ndacc\hdfprep5.exe (incl. Hdfglob5_m.f90, pkunzip (for X86 – 32 Bit) or 7za.exe (for X64 – 64 Bit)) creates meta and data files in folder \ndacc\asc2hdf_ndacc. To generate HDF files using IDL, move or copy meta and data files from \ndacc\asc2hdf_ndacc to \ndacc\idl2hdf. Start idlcr8hdf.sav (IDL binary file) using IDL Virtual Machine and follow the instructions on the screen. Uploading files to the NDACC database is realized by using WS_FTP Client.

Automatic delivery

All steps of the manual delivery can be set up for automatic delivery using a glue language (especially Python) and cron jobs (task planer). For example:

- Python script to download and check pT profiles every single day
- Automatisation of reading the paths to the date folders where spectra are located and automatically creating .inp and .erg files
- Running a cron job that starts Opus Software when new .inp and .erg files were created
- Python script to create .inp files for PROFFIT when spectra calibration finished
- Running a cron job, that checks for new PROFFIT results and creates meta and data files for the HDF generation
- Running IDL Virtual Machine (has to be done manual since IDL needs user input while running)

Personnel

Dr. Thomas Blumenstock, Karlsruhe Institute of Technology, IMK-ASF

Dr. Frank Hase, Karlsruhe Institute of Technology, IMK-ASF

Dr. Matthias Schneider, Karlsruhe Institute of Technology, IMK-ASF

Dr. Omaira Elena García Rodríguez, Centro de Investigación Atmosférica de Izaña, AEMET

Matthaeus Kiel, Karlsruhe Institute of Technology, IMK-ASF

E. References

- Clémer, K., M. Van Roozendaal, C. Fayt, F. Hendrick, C. Hermans, G. Pinardi, R. Spurr, P. Wang, and M. De Mazière, Multiple wavelength retrieval of tropospheric aerosol optical properties from MAXDOAS measurements in Beijing, *Atmospheric Measurement Techniques*, 3, 863-878, 2010
- Palm, M.; Hoffmann, C. G.; Golchert, S. H. W. & Notholt, J. The ground-based MW radiometer OZORAM on Spitsbergen -- description and status of stratospheric and mesospheric O₃-measurements *Atmos. Meas. Tech.*, 2010, 3, 1533 - 1545
- Rodgers, C. D.: *Inverse Methods for Atmospheric Sounding, Theory and Practice*. World Scientific Publishing, Singapore-NewJersey-London-Hong Kong, 2000
- Spurr, R., LIDORT and VLIDORT: Linearized pseudo-spherical scalar and vector discrete ordinate radiative transfer models for use in remote sensing retrieval problems, *Light Scattering Reviews*, Volume 3, ed. A. Kokhanovsky, Springer, 2008
- Studer, S., Hocke, K., Pastel, M., Godin-Beekmann, S., and Kämpfer, N.: Intercomparison of stratospheric ozone profiles for the assessment of the upgraded GROMOS radiometer at Bern, *Atmos. Meas. Tech. Discuss.*, 6, 6097-6146, doi:10.5194/amtd-6-6097-2013, 2013