Harmonized ozone profile retrievals from GROMOS and SOMORA

Eric Sauvageat

UNIVERSITÄT BERN

OESCHGER CENTRE
CLIMATE CHANGE RESEARCH

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

MeteoSwiss

Research Report No. 2022-01-MW

June 2022

Institute of Applied Physics
Microwave Physics Division

Sidlerstr. 5
CH-3012 Bern
Switzerland

Tel. : +41 31 631 89 11
Fax : +41 31 631 37 65
E-mail : eric.sauvageat@unibe.ch

Summary

This document contains a description of the new retrieval routines designed in the frame of the GROMORA project. This is a joint project between the Institute of Applied Physics at the University of Berne and the Federal Office of Meteorology and Climatology MeteoSwiss in Payerne which aims at a complete harmonization of the data processing of GROMOS and SOMORA, two passive ground-based microwave radiometers monitoring ozone in the middle-atmosphere.

This document focus on the second part of the processing, namely the retrievals of atmospheric ozone profile from the calibrated and integrated spectra (level 1a and level 1b respectively). It serves as supplementary material to a publication submitted to Atmospheric Measurement Techniques in the frame of a special issue from the Quadriennial Ozone Symposium 2021 (Atmospheric ozone and related species in the early 2020s: latest results and trends).

This document is complemented by a comprehensive documentation of the harmonized time series for the two instruments. For practical reasons, it has been placed in a separate document which can be found together with the data in the BORIS-portal of the University of Bern.

Contents

Chapter 1

Introduction

The GROund-based Millimeter-wave Ozone Spectrometer (GROMOS) and the Stratospheric Ozone MOnitoring RAdiometer (SOMORA) are two Swiss ground-based passive microwave radiometers (MWR) measuring ozone in the middle-atmosphere. GROMOS is operated by the Institute of Applied Physics (IAP) at the University of Berne and SOMORA is operated by the Federal Office of Meteorology and Climatology MeteoSwiss in Payerne. GROMOS and SOMORA are part of the Network for the Detection of Atmospheric Composition Change (NDACC) and they operate continously since 1994 (GROMOS) and 2000 (SOMORA). They both observe the rotational transition line of ozone around 142 GHz and provide continuous, all-weather ozone profiles with hourly time resolution at altitudes between 25 to 65 km. These measurements are important for estimating long-term trends and cross-validating satellites observations.

Both instruments have been designed by the Microwave Physics group at the IAP and, in addition to using very similar technologies, their geographic proximity ($\sim$ 40 km) make them unique in the world of passive microwave observations. This proximity also makes them important for the validation of the technology within the strato-mesospheric ozone community. In November 2019, MeteoSwiss and IAP decided to start a harmonization of the data processing from the two instruments. The goal was to keep two independent ozone time series but to harmonize completely the processing from the raw data (level 0) to the ozone profiles (level 2) of these two instruments.

The first step of the harmonization was concerned with the calibration of the radiometric raw measurements, often refers to as the "level 0 to level 1" processing. This was the topic of a first research report Sauvageat (2021). It is based on the first version (v.1.0) of the GROMORA (**GROMO**S and SO**MORA**) calibration routine. Since then, there have been a few updates to the calibration routine (now in v.2.0), however, the calibration user guide should remain valid in most aspects.

The present report focus the second part of the GROMORA harmonization project: the retrievals of strato-mesospheric ozone profiles from the microwave calibrated and integrated spectra, i.e. "level 1 to level 2". Similar to the previous report, it covers the periods 2009-2022 where both GROMOS and SOMORA use the same spectrometer. This document can be essentially divided into two main parts: the first part describes the harmonized retrieval routine whereas the second part is designed as a user guide for the new retrieval processing routines.

The first part provides a description of the retrievals setup with insight on the theoretical basis of atmospheric spectral inversions and details on the parameters and extra data needed for it. The second part describes the routine organization and guides the users for running, improving or designing new ozone retrievals. A full descriptions of all parameters and functions is not provided in this document but some of them are available within the technical documentation of the GROMORA project provided with the code in the Git repository of the project. Note that all scripts and functions also contains some useful and succinct documentation.

The detailed documentation of GROMOS and SOMORA data series is provided together with the data series for convenience and ease of access to information. It is of great importance for anyone willing to use extensively these data series for new research topics. The data and its documentation can be found in the Bern Open Repository and Information System (BORIS Portal) project dedicated to GROMORA: Harmonized retrievals of middle atmospheric ozone from Swiss microwave radiometers. In addition to the year-to-year description of these series, we also provide a list of key events for both series in the form of text files stored in the Git repository of the project. This should hopefully enable to easily update the documentation in the following years.

In addition, a peer-reviewed publication on this project has been submitted and accepted for publication in "Atmospheric Measurement Techniques" Sauvageat et al. (2022). All the calibration and retrieval codes are freely available on the Git server of the IAP GROMORA-harmo(only with campus account from the University of Bern). For external user, the code can be found either on GitHub or on Zenodo (10.5281/zenodo.6799357).

The document is organized as followed: Chapter 2 describes the retrieval setup. Chapter 3 contains a description of the retrieval routine, its requirements, main principles and present its different outputs. It also describes succinctly how to use the routine with an existing instrument or how to adapt it to a new instrument. Finally, Chapter 4 presents some short conclusions and outlook.

Chapter 2

Retrievals

In the microwave frequency range, the pressure-broadening effect acting on atmospheric emission lines is used to retrieve information on atmospheric constituent vertical profile from the calibrated microwave emission spectra. This so-called retrieval is a well validated technique which has been successfully applied to temperature, wind and many trace gases like O_3, CO or H_2O (Janssen, 1993). Among the different retrieval techniques, we selected the Optimal Estimation Method (OEM) following the formalism described by Rodgers (2000). This statistical method extracts the best estimate of an atmospheric profile from a set of measurements with noise, a priori information and a forward model. In addition, the OEM enables to characterize the error budget of the retrievals and it will be described more in details in section 2.2.

In the case of GROMOS and SOMORA, we use radiometric measurement around the 142.175 GHz ozone rotational transition and compare it against radiative transfer simulations of the atmosphere performed by the Atmospheric Radiative Transfer Simulator 2.4 (ARTS; Buehler et al. (2018)) described in section 2.1. Providing the right set of a priori and instrumental knowledge to the algorithm, it enables us to retrieve hourly ozone profiles at altitude between $\sim$ 20 and 80 km. As a summary, Table 2.1 lists the most important paramaters used in the GROMORA retrievals.

In practice, the retrieval product is a combination of different quantities depending on what the user wants to retrieve. It includes of course the atmospheric profile of interest (e.g. ozone) but can also include further quantities like a water vapor continuum term or some instrumental parameters (section 2.4). The different retrievals quantities for GROMOS and SOMORA retrievals can be found in section 2.5. Note that the addition of more retrieval quantities does not fundamentally change the underlying retrieval theory presented in the next section.

Table 2.1: Retrievals parameters

	source/type
Forward model	ARTS
Species	O_3, H_2O, O_2 and N_2
Spectroscopy	Perrin (JPL & HITRAN)
Atmospheric state	1D ECMWF & CIRA 86
O_3 a priori	WACCM, free-running
H_2O a priori	ECMWF
FM grid	1-112 km, 2 km resolution
Retrieval grid	1-95 km, 2 km, resolution

2.1 Forward model

In the case of ground-based microwave radiometry, the forward model (FM) describes the radiative transfer physics between trace gases emissions and the instrument's receiver. We used the Atmospheric Radiative Transfer Simulator 2.4 (ARTS), an open source software with a special focus on microwave radiative transfer simulations (Buehler et al., 2018).

The goal of the FM is to simulate the radiometric observations (y) from a given atmospheric state (x) and error sources (ϵ) and it can be mathematically described by Eq. 2.1. It basically describes the emission, extinction and transmission of microwave radiation between the ozone molecules and the antenna of the radiometer.

$$y = F(x, b) + \epsilon \tag{2.1}$$

After running the FM, the simulated atmospheric spectrum can be compared with the radiometric observations. Based on a cost function, the state vector x can then be updated and the FM run again to yield an improved simulated spectrum providing a better match with the observations. This iterative process, so-called "inversion" or "retrieval", will be described in the next session.

2.2 Optimal Estimation Method

The inversion of microwave atmospheric emission spectra is ill-posed because the available measurements are usually not able to fully resolve the atmopheric profile that would reproduce the observations. It means that, even in the case of a perfect knowledge of the atmospheric radiative transfer physics and sensor's influence on the observations, it would not be possible to retrieve an atmospheric profile corresponding perfectly to the measurements. Therefore, there is a need to combine the measurements with some additional knowledge, for instance on the state of the atmosphere or the expected shape of the solution.

The Optimal Estimation Method (OEM) retrieval technique enables to inverse atmospheric emission spectra by combining the measurements with these additional infomation (so-called a priori) in an statistically optimal way. It is widely used to solve atmospheric inversion problems and has been applied to trace gases retrievals successfully since many years. It is a common strato-mesospheric ozone retrieval method used in both ground-based and satellites microwave radiometry.

ARTS offers a fully integrated OEM retrieval environment and in this section, we will descibe succinctly the theoretical basis behind this method, including some underlying assumptions. For more detailed information on the OEM or its application to ozone profiling instruments, the reader is redirected to Rodgers (2000), Parrish et al. (1992) or Tsou et al. (1995).

To solve the inverse problem, the OEM relies on Bayes' probability theorem and it tries to find an estimate of the true state by minimizing the following cost function:

$$\chi^2 = [y - F(\hat{x})]^T S_y^{-1} [y - F(\hat{x})] + [\hat{x} - x_a]^T S_a^{-1} [\hat{x} - x_a] \tag{2.2}$$

where $\hat{x}$ is the estimated atmospheric profile of interest and x_a is the a priori profile. S_y and S_a are respectively the measurement and a priori covariance matrix which are defined by the user. These covariance matrices account respectively for the measurement and atmospheric noise and have a significant influence on the retrieval result and its error budget. The definition of both matrices is critical and is described in section 2.4 and section 2.5.1.

The best estimate of the atmospheric profile $\hat{x}$ is computed with an iterative process using (for instance) the Gauss-Newton algorithm and Eq. 2.3:

$$\hat{x}_{i+1} = \hat{x}_i + (S_a^{-1} + K_i^T S_y^{-1} K_i)^{-1} [K_i^T S_y^{-1} (y - F(x_i)) - S_a^{-1} (x_i - x_a)] \tag{2.3}$$

where K is the so-called Jacobian matrix computed by the FM. It corresponds to the derivative of F with respect to x and essentially describes the sensitivity of the FM to changes in the atmospheric profile.

From the covariance matrices and the Jacobian, the sensivity of the retrieved state to the measurement (also called the gain matrix) is computed with:

$$\frac{\partial \hat{x}}{\partial y} = G = (S_a^{-1} + K^T S_y^{-1} K)^{-1} K^T S_y^{-1} \tag{2.4}$$

And the averaging kernel A, which is the sensivity of $\hat{x}$ to changes in the unknown true state (x) is then given by:

$$\frac{\partial \hat{x}}{\partial x} = \frac{\partial \hat{x}}{\partial y}\frac{\partial y}{\partial x} = GK \tag{2.5}$$

2.2.1 Retrieval errors

One advantage of the OEM is that it enables to characterize the error budget of the retrievals. In our retrievals, we assume that the total error is composed of 2 main components: the so-called "smoothing error" and "measurement error" (sometimes called retrieval noise).

The smoothing error describes the error arising from the limited vertical resolution of the instrument. In fact, the radiometers are not able to resolved small scale vertical changes of the true profiles which results in the retrieval of a smoothed version of the true profile. The amplitude of the error can give an impression of the vertical resolution of the observing system and is usually large compard to other error sources.

The measurement error is a consequences of the noise present in the measurements. This error impacts mostly the higher altitudes and sets an upper altitude limit to the retrieval capabilities of the observing system.

More specifically, the total error covariance matrix of the retrieval can be computed as the sum of the smoothing error covariance matrix (S_s) and the retrieval noise covariance matrix (S_o)

$$S_{\hat{x}} = S_s + S_o = (A - I)S_a(A - I) + GS_yG^T \tag{2.6}$$

Note that to provide a realistic error budget to the retrieval, S_y should also take into account any errors caused by the FM or any corrections applied between the simulated and the true observations (e.g. tropospheric or windows correction). It is important to keep in mind that this is usually not the case and that further error sources have to be considered in order to provide a comprehensive error budget.

2.3 Atmosphere and spectroscopy

ARTS offers many possibilities to define the atmospheric state, a priori data and simulation grids. We use one-dimensional pressure, temperature and altitude (PTZ) profiles from the European Centre for Medium-Range Weather Forecasts (ECMWF). More specifically, we use ECMWF operational global analysis dataset. It provides a best estimates of temperature, pressure, water vapor and corresponding altitude at four main synoptics hours 00, 06, 12 and 18 UTC with 1.125° spatial resolution. For our retrieval, we select the closest grid points to either Bern or Payerne and make a temporal average of the main profile on a time period of 7 hours around the measurement time.

The ECMWF dataset is limited to approximately 70 km altitude and therefore, we complete it with a CIRA86 climatology up to 112 km.

Note that for an easier management of the PTZ profiles, the 1D daily profile are extracted before the retrievals and are stored in daily netCDF files for each location. The routine to do so is also located within the GROMORA Git repository.

As atmospheric species, we include ozone, water vapour, oxygen and nitrogen. For ozone, we use the spectroscopic database from Perrin et al. (2005), which is provided with ARTS 2.4 and is derived from the HITRAN and JPL spectroscopic databases. For water vapour, we use the "H2O-PWR98" complete absorption model for water vapour provided by ARTS and based on (Rosenkranz, 1998). For oxygen and nitrogen, we also use some parametrizations provided within ARTS, respecrively "O2-PWR93" and "N2-SelfContStandardType" (for more information, see Buehler et al. (2005) or the ARTS user guide at `https://www.radiativetransfer.org/`).

2.4 Measurements and sensors

To perform meaningful retrievals, it is important to take into account the influence of the sensor on the atmospheric emission spectrum. However, it is a complex task which requires simulating many instrumental components, most of which having unknown effects on the atmospheric spectra.

The influence of the sensor on the retrieval is difficult to take into account because it contains many component that can vary both in time and in the frequency space. The first error souce arising from the instrument is the noise present in every measurement spectrum. The noise originates from thermal emission of the receiver components and is present in any observing system. In microwave radiometry, this noise is reduced through integration of the calibrated spectrum but it remains a key quantity to take into account when solving the inversion problem.

2.4.1 Measurement noise

The measurement noise is an important quantity for OEM retrievals because it defines, together with the a priori covariance, the information that can be extracted from the measurement at each pressure level.

The measurement noise is computed individually for each measurement spectrum based on the noise level observed on the spectrum and is considered to be uncorrelated between the different channels (i.e. S_y is a diagonal matrix). More specifically, we use the variance of differences between neighbouring channels ($\sigma^2_{\Delta y}$) after removal of spurious FFT channels. This variance should be twice the simple variance of the measurement vector ($\sigma^2_{\Delta y} = 2\sigma^2_y$).

Therefore, we define the noise covariance matrix for each retrievals from the integrated spectrum y using:

$$(S_y)_{ii} = \frac{1}{2}\sigma^2_{\Delta y} \tag{2.7}$$

For most of our retrievals, the noise is slightly higher for GROMOS ($\sigma_y \approx 0.7$ K) than SOMORA ($\sigma_y \approx 0.5$ K) because GROMOS has a higher receiver noise temperature and a higher frequency resolution.

2.4.2 Sensor consideration

ARTS has dedicated built-in functions that can model the influence of the most relevant components on the atmospheric observations (Eriksson et al., 2006). For GROMOS and SOMORA, we included the effect of the FFT spectrometer channel response ($|\frac{\sin(x)}{x}|^2$) and the effect of the sideband ratio. The latter is modelled based on measurements of the sideband response from GROMOS and SOMORA and is assumed to follow a sinusodial dependance.

The frequency grids have been defined to cover the range of GROMOS and SOMORA spectrometers with a refined frequency resolution around the ozone line center. It matches approximately the spectrometer resolution at the line center to optimize retrievals at higher altitudes, whereas the spectral resolution is coarser on the line wings to limit computation time.

There are other sensors or external influences which are difficult to estimate and correct during the calibration process or to simulate accurately for each spectrum. This is the case for the instrumental baselines and the tropospheric absorption. The instrumental baselines are modulation of the atmospheric spectrum due to the observing system. They can arise during the mixing process, the sideband filtering or can be due to undesired reflections, typically when observing the calibration targets. Because these quantities can vary and are difficult to simulate or correct, it can be included as retrieval quantities as will be seen in the next section.

2.5 Retrieval quantities

In ARTS, it is possible to add more than one retrieval quantity to the state vector x. It can be some additional atmospheric unknown quantities (e.g. tropospheric continuum) or some unknown sensors influence on the observations. In principle, the inversion problem is then similar to the single quantity retrieval and for each retrieval quantity, it is needed to define a corresponding a priori and covariance matrix. In this section, we will describe the quantities used for the GROMORA retrievals and the reasons justifying their addition.

2.5.1 Ozone

The main retrieval quantity is hourly ozone volume mixing ratio (VMR) vertical profiles. As a priori we use monthly ozone profile extracted from free-running simulations of the Whole Atmosphere Community Climate Model (WACCM). Further, depending on the local solar time, we either use the daytime or nightime a priori ozone profile as shown in Fig. 2.1.

The a priori covariance matrix for ozone varies with altitude and is kept constant for all retrievals. Under 1 hPa ($\sim$ 50 km), it is a fraction of a yearly ozone profile obtained from the WACCM free-running model. Above, it has been adapted in order to optimize the information content from the measurement in the stratosphere and lower mesosphere. To account for the vertical coupling of the atmosphere, we also include some exponentially decreasing covariances between altitude levels that can be expressed as:

$$(S_a)_{i,j} = \sigma_a(z_i)\sigma_a(z_j)\exp\left(-\frac{|z_i - z_j|}{h}\right) \tag{2.8}$$

In our case, the correlation length $h = 1$ km and a cutoff value has been applied to the covariances which suppress any covariance as soon as it is under a certain threshold. The ozone a priori covariance matrix is shown in Fig. 2.2.

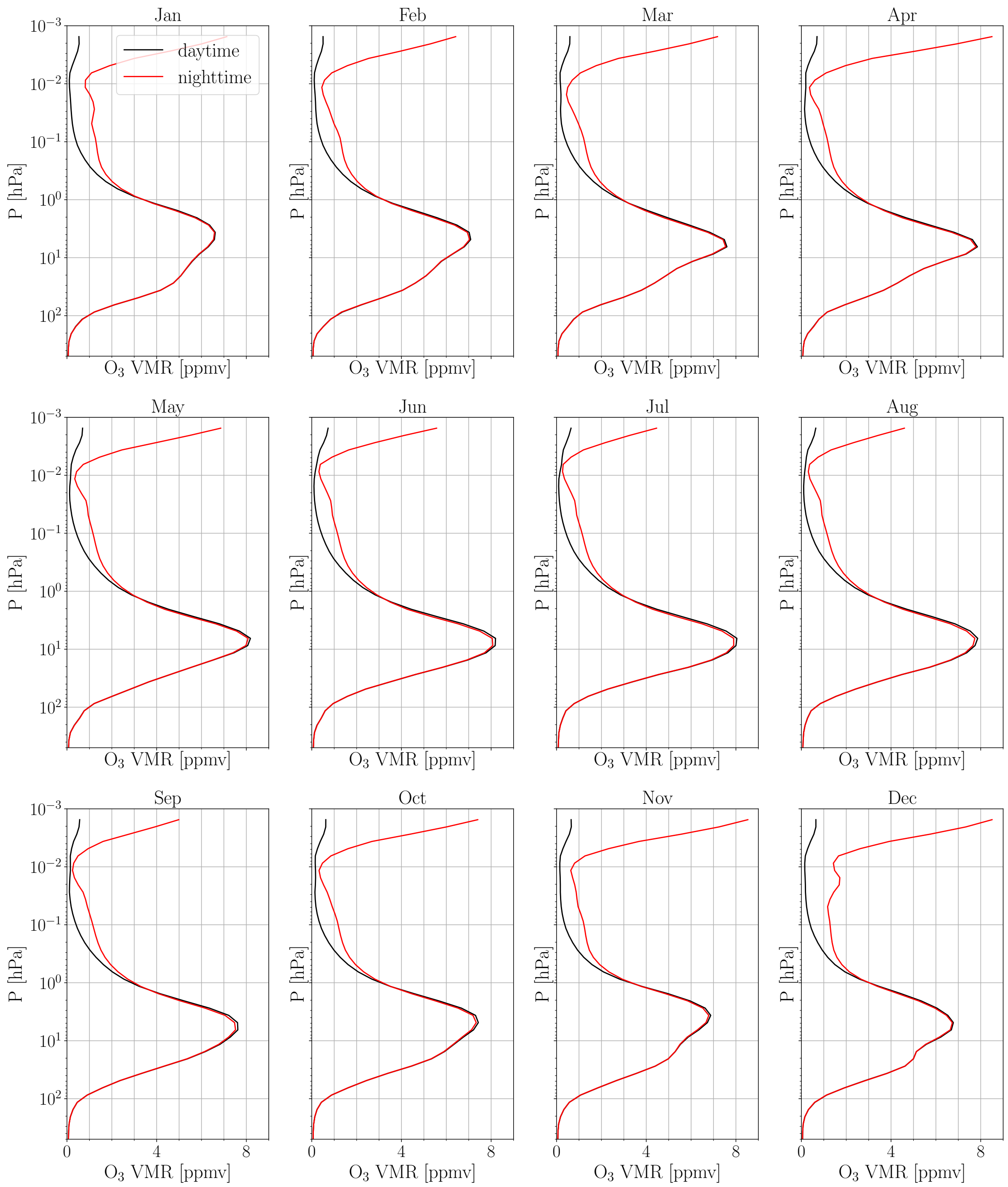

Figure 2.1: Monthly day and nightime apriori profiles used in the GROMORA retrievals.

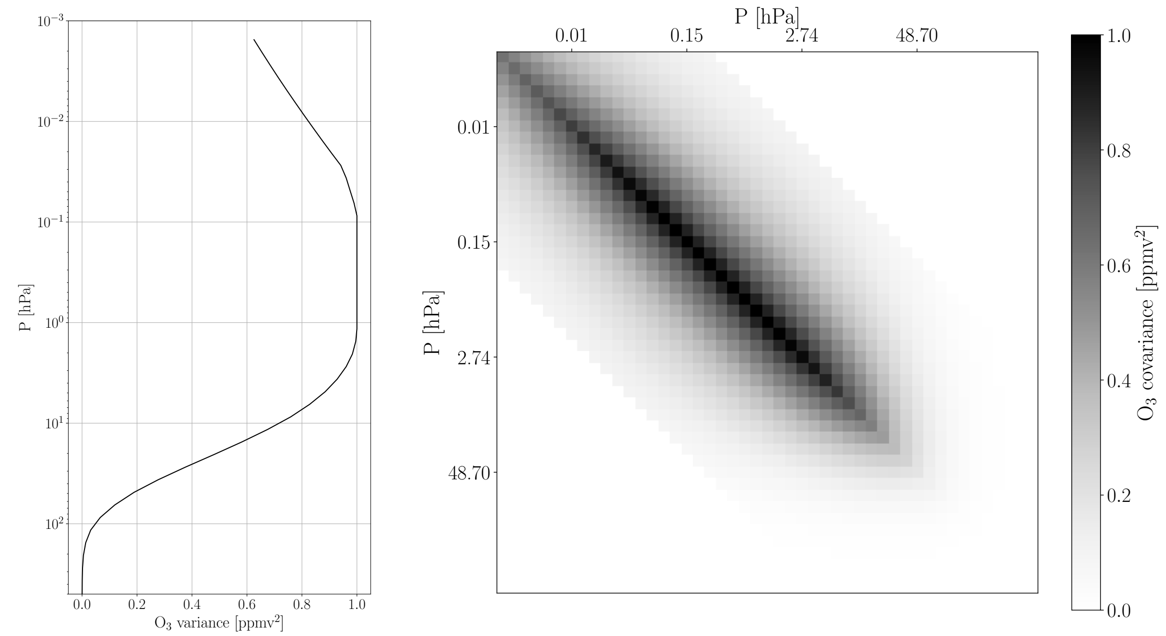

Figure 2.2: Ozone a priori covariance matrix used for the GROMOS and SOMORA retrievals. The left panel shows the diagonal variance profile whereas the right panel shows a 2D representation of the full covariance matrix against the pressure.

2.5.2 Tropospheric contribution

Around the 142 GHz ozone transition line, the tropospheric water continuum contributes significantly to the observed spectra and has to be considered during the inversion process. One simple correction method is the so-called tropospheric correction (Ingold et al., 1998) but it is certainly a better solution - also in view of assessment of the error propagation - to include the tropospheric water vapour as a retrieval quantity within ARTS, as has been done previously for such retrievals (e.g. Palm et al. (2010)).

In practice, we use the ECMWF humidity profile as a priori and we retrieve a tropospheric continuum value by setting a single pressure value as retrieval grid for the water vapour ($p = 500$ hPa in our case). Note that we use relative units and therefore, we only retrieve a scaling of the a priori water vapor profile so that the retrieved value for the continuum have no real physical meaning.

2.5.3 Instrumental contributions

A frequency shift is also retrieved for each spectrum. The reason for including such a retrieval quantity is because the local oscillators of both GROMOS and SOMORA are not perfectly stable and even a slight shift of the reference frequency can have significant bias on the retrievals.

Despite mitigation of baselines using different techniques (e.g., mirror wobbling, non-perpendicular aspect of cold load), it is generally necessary to retrieve some instrumental baselines as well (Palm et al., 2010). In the case of GROMOS and SOMORA, we include a second-order polynomial and different sinusoidal baseline periods. In order to avoid the degradation of the retrievals with the addition of too many sinusoidal baselines, we performed a first processing of the full time series without sine baselines and used the residuals to compute the main sinusoidal baselines for each instrument and period. We observed that the sinusoidal baseline periods remain more less the same on time-scale of month to years so in practice, only a few period changes were applied during the full extent of the time series for each instrument. The details of which baseline periods have been used for which period can be found in the documentation provided with the time series.

Chapter 3

Retrievals routine

In this chapter, we describe the new harmonized retrieval routine written for GROMOS and SOMORA. We provide a short introduction to the routine structure and its main components. The aim is to guide new users of the routine or help for the initial setup of new retrievals.

The routine has been designed in the frame of the GROMORA project and is therefore focused on the ozone retrievals from passive microwave radiometer. Adaptation of the routine to other MW ozone radiometer should be doable without too much effort whereas its extension to other retrievals quantities might me more cumbersome.

To simplify the use and deployement of the routine, all the scripts and functions needed to calibrate and retrieve ozone profiles from GROMOS or SOMORA are located in a single folder or Git repository: GROMORA-harmo. The detailed organization of the folder can be found on different README files located in the repository. These files should complement the explanation given in the present document.

3.1 Requirements and applications

The retrieval routine is written in Python and requires a compiled installation of the Atmospheric Radiative Transfer Simulator (ARTS) version 2.4 Buehler et al. (2018). ARTS is a radiative transfer software written in C++ which runs on Linux systems. All required dependencies and installation instructions from ARTS can be found on the ARTS website or directly on the ARTS Github page.

ARTS is now provided with PyARTS, a Python package providing a comprehensive interface to ARTS original controfiles. In fact, this is the recommended way to access ARTS and run radiative transfer simulations and retrievals. In addition to PyARTS, the GROMORA retrieval routine rely on pyretrievals, a Python library initially written by Jonas Hagen (see original code at pyretrievals). As this library was written before PyARTS and based on a previous ARTS version, we decided to include the updated version directly within the GROMORA retrieval package in order to simplify the deployement of the routines to new users. In addition, it includes some useful scripts to extract ECMWF a priori profiles.

The GROMORA calibration and retrievals scripts are grouped into a single repository available on the Git server of the Microwave Physics group at the University of Bern (GROMORA-harmo). For external user, the code can be found either on GitHub or on Zenodo (10.5281/zenodo.6799357). In case you are willing to contribute to the project, please contact the IAP IT responsible person, which can also grant you access to the original Git repository.

To instal PyARTS and other required Python packages, it is highly recommended to use conda. A specification file can be found in the code repository. Note that you need to activate your environment when compiling ARTS if you want to install correctly PyARTS. For ARTS to work properly, you also need to export the following paths to your shell:

- ARTS_BUILD_PATH
- ARTS_DATA_PATH
- ARTS_INCLUDE_PATH

It can be done either by include it by default in your *.bashrc* file on Linux (recommended) or by explicitely reading at the beginning of your scripts (e.g. using dotenv). For more details on what these paths should contain, please refer to the ARTS documentation.

In terms of computer resources, the retrieval routine is quite computationaly intensive but should run without problem on post-2020 computer.

3.2 Inputs

We have tried to limit to a maximum extent the amount of inputs required for the GROMORA retrievals. Except for the ECMWF daily files, all inputs should be available either from the GROMORA main folder or provided by ARTS. Overall, the operational retrieval routine does require some reading access to the following files:

1. GROMOS and/or SOMORA standard level 1 (calibrated and integrated) files.
2. WACCM ozone climatology: the monthly day/night climatology has been concatenated into a single netCDF file for convenience and can be found in the Git repository of the project (GROMORA-harmo).
3. ECMWF file catalogue for the selected location (see below)
4. CIRA86 monthly climatology, located in ARTS data files.
5. Perrin spectroscopy, located in ARTS data files.

Of course, many of these inputs can be changed for testing purposes and many other options are provided within the scripts so that the retrieval can be adapted to run without the ECMWF daily anaylsis for instance.

3.2.1 ECMWF

For the operational routine to work, a reading access is needed to the ECMWF 6-hourly profiles extracted from the global daily ECMWF analysis. The code for the extraction is located in the *pyretrievals* sub-folder of the git repository.

The so-called file catalogue is simply a folder containing all the 1D PTZ profiles extracted from the global ECMWF files. Note that for the routine to work off the shelf, all files should be located into the same directory.

3.3 Routine description

The routine has been designed to be as generic as possible in order to reduce duplication and harmonize to a maximal extent the retrievals between GROMOS and SOMORA. To apply the principles of generic programming, the GROMORA retrievals routine was designed to be the least instrument dependent as possible, both in terms of parameters and in terms of functions.

To achieve this, we use the concept of abstract base class to define a common Application Program Interface (API) for performing retrievals from GROMOS and SOMORA. In practice, the retrieval setup is defined within this abstract class (named **DataRetrieval** in *gromora_retrievals.py*) using so-called abstract methods and is implemented for different instrument using individual sub-class for each instrument. In practice and to avoid too long files, some functions called from **DataRetrieval** are stored in separate Python modules (i.e. in different files).

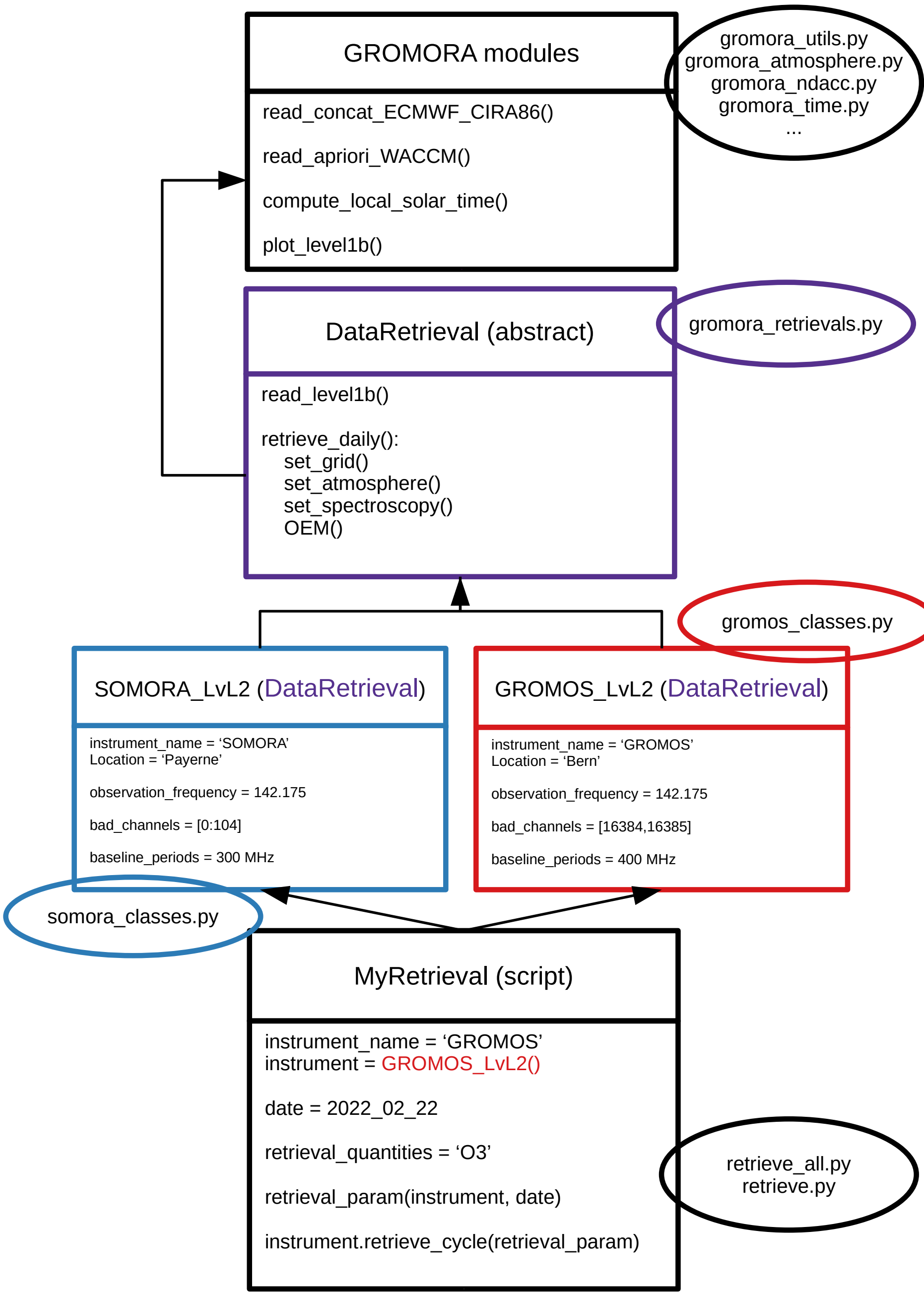

Figure 3.1: Retrieval routine schematic of the GROMORA project.

Similarly to the calibration routine, we also use a dictionary (**retrieval_param**) to store and propagate all relevant parameters for a specific retrieval into the different functions. A main script is used to define the global parameters whereas more specific parameters are defined in **DataRetrieval** or in each instrument's class when they are instrument specific.

In the following sections, we present a brief overview of the functionning of the main script as well as of the main functions used in the operational GROMORA retrievals. For more details, the reader is redirected to the Git repository of the project or to the documentation in the code.

3.3.1 GROMORA retrievals abstract base class

The *gromora_retrievals.py* module contains the core functions of the GROMORA retrievals. It is where the **DataRetrieval** abstract class is defined (see here for more info on abstract base class) which is completely common to both GROMOS and SOMORA and runs the actual retrieval.

The main function of the **DataRetrieval** abstract class is **retrieve_cycle** which performs the actual retrieval for a given day and calibration cycle (or time). This function implements all the steps needed to do the retrievals, setup the forward model and solves the inverse problem by calling the OEM algorithm from ARTS.

It also includes the functions to defines of the main retrieval parameters, grids, covariance matrices, etc... Most of the important functions are located directly within this class whereas some are located in dedicated module and are then called from within the class. An example is the definition of the atmospheric background and a priori profile which are all located in the *gromora_atmosphere* module.

3.3.2 Instrument classes

For each instrument, there is a specific implementation of the GROMORA retrievals abstract class. It consists of individual Python module for every instrument. It is where all instruments dependant parameters can be defined.

In practice, to implement a new retrieval, for instance following a spectrometer change, a new instrument class can be created which will define the updated retrieval parameters. It can then use the same abstract class and most of the existing functions within *gromora_retrievals.py*.

3.3.3 Main script

In the main script, we define all necessary parameters for the calibration and the integration of an MWR on a specific day. For an operational use, there are only a few parameters to modify manually in the main script which includes for instance the date range to retrieve, the retrieval quantities to include or the name of the outputs.

The operational script used to reprocess the data from GROMOS and SOMORA is named *retrieve_all.py* whereas an alternative script *retrieve.py* with further retrieval options can be used for testing or development purposes.

Depending on the instrument defined by the user, the main script will initiate the correct instrument class which will further implement the GROMORA retrieval abstract class and run the retrieval. In principle, the introduction of new retrieval can either be added to the existing main script, or it can be duplicated to avoid confusion.

3.3.4 Other scripts and modules

They are a few other important modules which basically serve as collection of functions for the GROMORA retrievals:

1. *gromora_concatenate_level2*: important module that is used for the concatenation of the daily to the yearly level 2 files. It also contains some plotting functions for a quick look at a short period of retrieved time range.
2. *gromora_atmosphere*: module containing all the function related to the atmosphere definition for the retrievals, including the definition of the a priori profiles.
3. *gromora_time*: module with some useful time function, mostly related to local solar time definition and conversion.
4. *gromora_utils*: utility module with some basic functions.
5. *gromora_ndacc*: module to convert the level 2 data to the NDACC GEOMS data format.
6. *analyse_plot_retrievals*: an old module that contains many plotting functions. Most of them are now moved to a dedicated level 2 analysis repository.
7. *retrieval_module*: an old module containing the retrievals of GROMOS and SOMORA. It has been left in the main folder because it still contains some potentially useful function, like one for the retrieval of ozone from tropospheric corrected spectra.

3.3.5 Using the routine

The retrievals routines of the GROMORA project, while optimized for GROMOS and SOMORA have been written to be as generic as possible. This way of doing provides some flexibility in the application (like adding new instruments) and should help new users to get familiar with the routine pretty quickly as most of the functions and scripts have been harmonized to a maximal extent and are documented as well as possible.

In this section, we will try to give some keys and best practices to use the routine described in this report. While the use of the routine with an existing instrument (i.e. an instrument already included in the GROMORA calibration routine) should be quite straightforward, the adaptation of the routine to a new instrument would require a bit of work and a good knowledge of the routine.

In order to run the retrievals, you first need to activate (or create) your dedicated Python environment and make sure that all the specific paths are setup correctly. Also, make sure that you have access to all inputs and check if the ARTS paths are correctly set. Below, you will find some important concepts and tips to understand how the routine works and can be operated.

Changing the retrieval quantities

The retrieval quantities to include are defined with a string variable called **retrieval_quantities** and stored within the **retrieval_param** dictionary. This might not be the best solution but it limits the amount of variables and should be quite human readable. The different options available are listed in the technical documentation.

Reprocessing the GROMORA dataset

To reprocess the GROMORA ozone time series, please use the *retrieve_all.py* main script with all default parameters. Bascially, you just need to define the time periods and to adapt the input and output folder and filenames.

Update the retrievals

To update the retrieval in itself, you can modify the *gromora_retrievals* base class or add a new retrieval function. It might be needed for instance for the following operations:

- Change of the retrieval method
- Instrumental changes (new sensors considerations (e.g. antenna patterns) or new spectrometer)
- Change to a new atmosphere definition (PTZ) or new a priori data which are not already included in the routine.

3.4 Outputs

3.4.1 Level 2

Due to its widespread utilisation and its self-documentation property, it was agreed to use the netCDF format for storing all levels of our processed data. The outputs of the retrievals routine (i.e. the ozone profiles and retrievals diagnostics quantities) are stored in daily netCDF-4 files. The daily format enable to keep a certain continuity compared to the calibrated and data (which are also saved daily) and results in quite compact files while not producing too many outputs files. Note that it does not mean that the data are daily averages but just that they are grouped together in an output file based on which day they belong.

For a reminder on the netCDF file format and its main functionnalities, the reader is redirected to Sauvageat (2021) of invited to visit unidata: netCDF.

The level 2 of the GROMOA profiles contains not only the ozone profiles, but also all the others retrieved quantities and diagnostics extracted from the OEM retrievals. For instance, it contains the measured and fitted spectra (see Fig. 3.2 and 3.3), the noise content of the spectra as well as all relevant geolocation information. For all retrievals quantities, it also include the apriori, grid and error estimations extracted from the OEM. Typical diagnostics quantities available in the level 2 files can be seen in Fig. 3.4 and Fig. 3.5.

Note that there are 2 types of level 2 files available for the GROMORA retrievals:

1. Daily files, available on request
2. Yearly files which are simply a concatenated version of the daily level 2 with some additional quality checks but without the measured and simulated spectra (due to space limitation). Their goal is to simplify the data sharing and the data analysis. It also introduces the first quality checks of the ozone profiles based on the retrievals diagnostics quantities (Section 3.4.2).

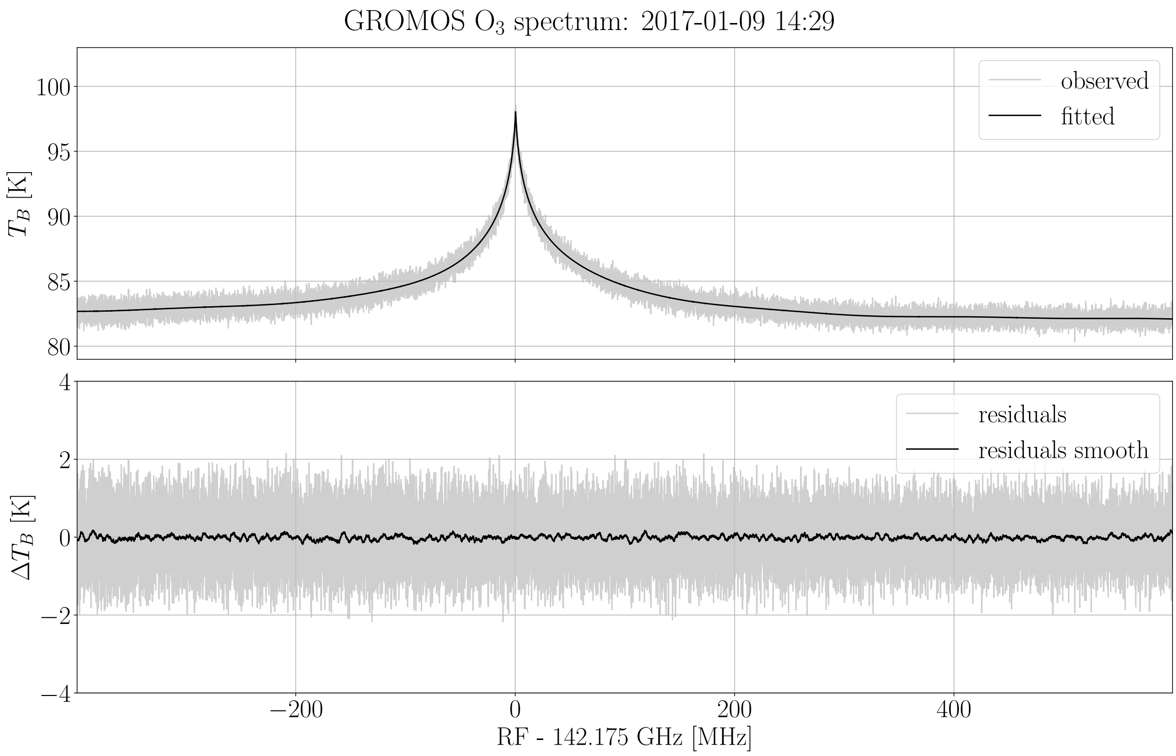

Figure 3.2: Example of measured and fitted ozone emission spectra from GROMOS (daytime)

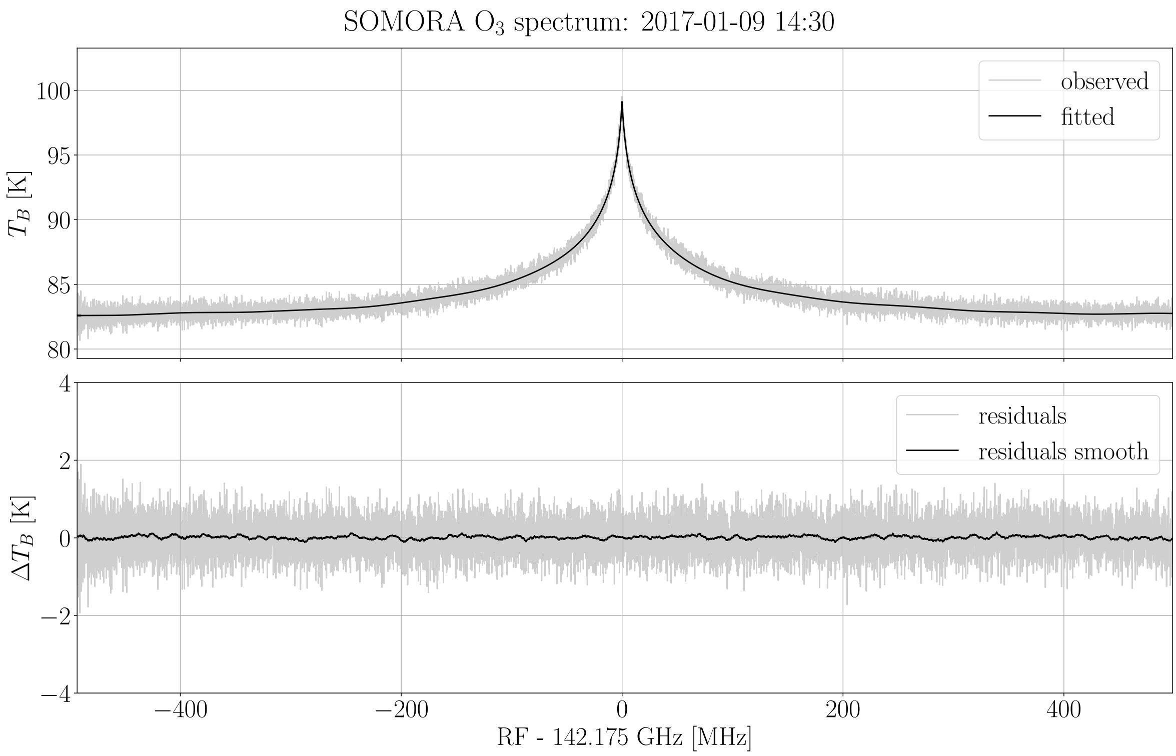

Figure 3.3: Example of measured and fitted ozone emission spectra from SOMORA (daytime)

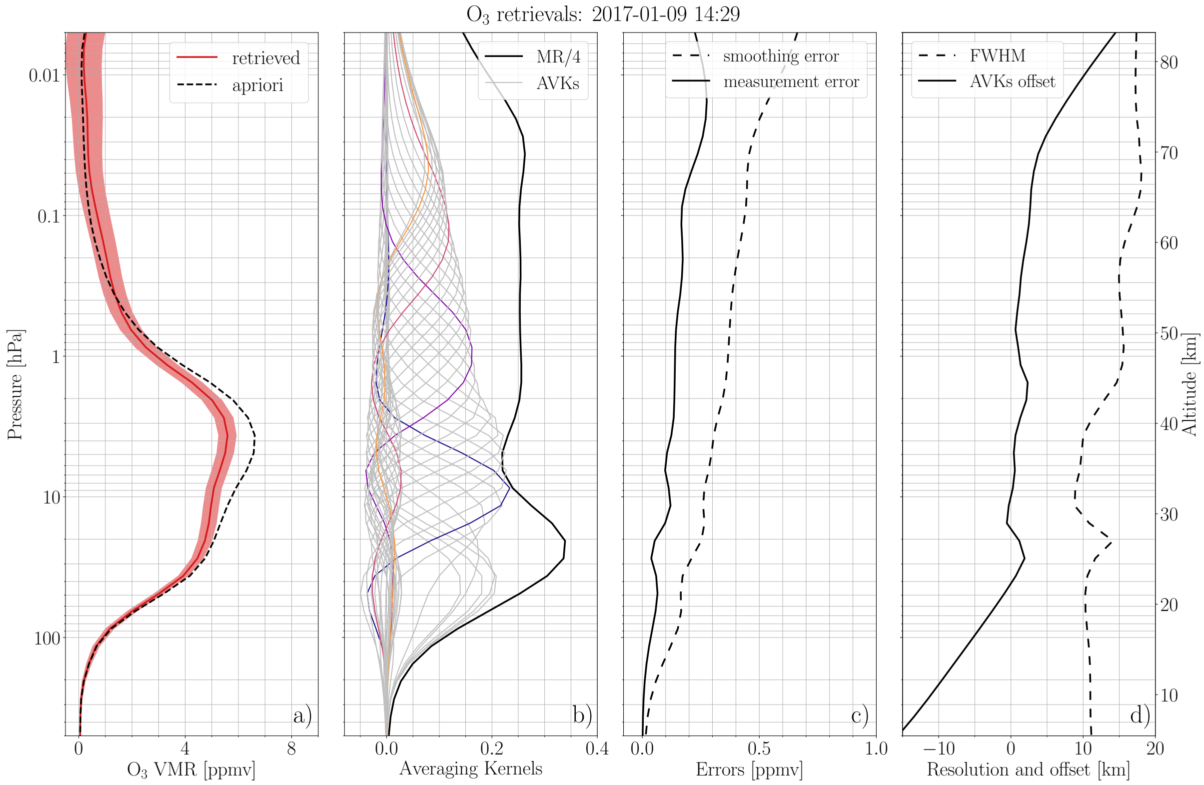

Figure 3.4: Example of ozone profile retrieved from GROMOS (daytime).

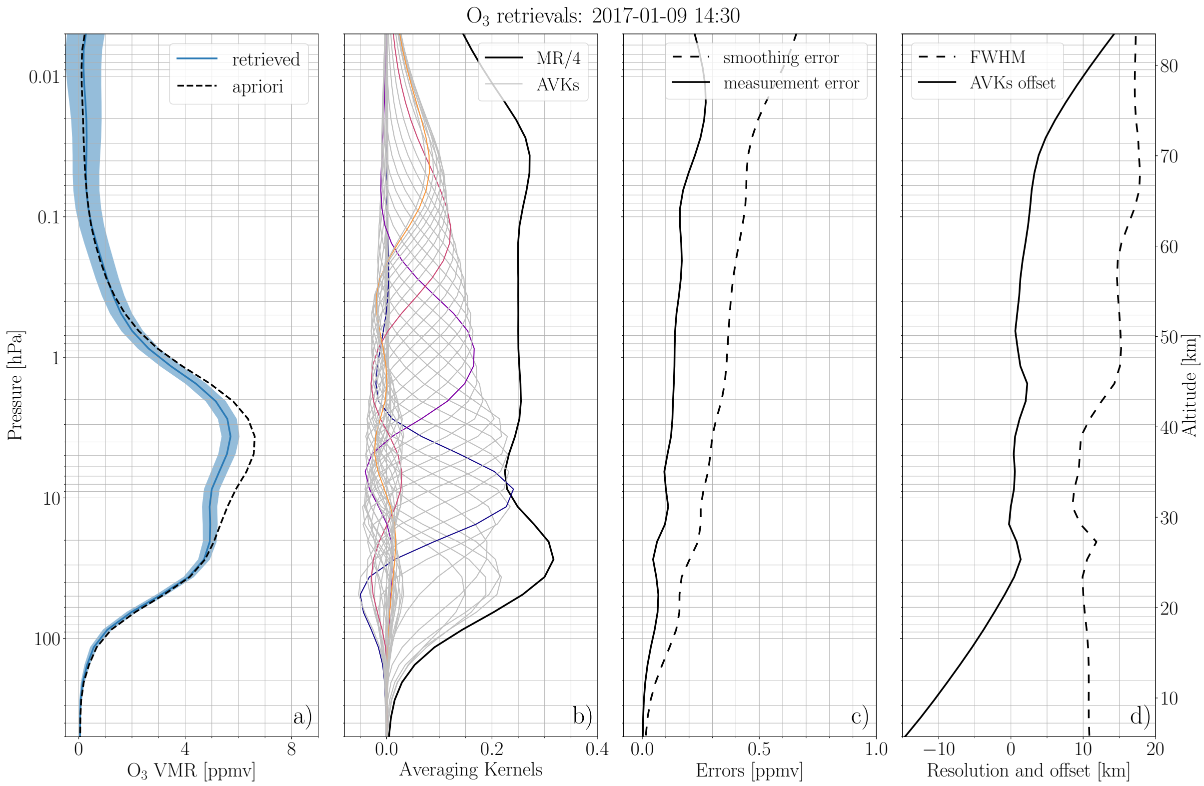

Figure 3.5: Example of ozone profile retrieved from SOMORA (daytime).

3.4.2 Quality checks

The daily level 2 file do not have any quality flags but they contain a set of diagnostics variables that can be use to check the qualtity of the retrieved profiles. In particular, the **oem_diagnostics** variable from ARTS is saved and gives an indication on the convergence status and final cost function of each retrieval (see ARTS documentation). Therefore, the quality checks are made during concatenation of the daily level 2 into yearly files. More specifically, the following flags are introduced:

1. **retrieval_quality**: this flags is based on retrieval diagnostics, more specifically it checks that the retrieval cost and the polyfit constant (degree 0) are within a certain threshold. It is a boolean with the value 1 (True) indicating a good quality of the hourly retrieval.
2. **level2_flag**: this flags identifies the period having good retrieval quality but with known instrumental issues. It is also a boolean with the value 1 indicating a flagged or a bad period.

Note that the flagging of the level 2 is not so conservative so it is generally recommended to remove all flagged data before using the GROMORA data for anaylsis. Depending on the study, it might also be needed to remove further dubious time periods (see time series documentation).

3.4.3 Level 3

The level 3 files are single concatenated files for the whole time series (2009-2021). They are interpolated on regular time grids which can be interesting for certain users and are much faster to read and process than the hourly level 2 data.

Also they are quality controled in the sense that the bad quality retrievals (**retrieval_-quality**) have been removed when the interpolation has taken place. The **level2_flag** are not automatically removed however and still needs to be dealt with depending on the user's needs.

Chapter 4

Conclusion and outlook

A complete new calibration and retrievals routine have been written in the frame of the GROMORA project. The starting point for these new routines was the full harmonization of data processing for GROMOS and SOMORA, two passive ground-based microwave radiometers operated respectively by the IAP and MeteoSwiss.

The routine theoretical basis and operation are described in the present report. The resulting harmonized time series are presented in a separate document which accompanies the data series. The cross-comparisons of the series from GROMOS and SOMORA as well as validation against satellites dataset are the subject of a publication in preparation. All the publications, data and codes related to this project will be grouped in the BORIS Portal in the following project: Harmonized retrievals of middle atmospheric ozone from Swiss microwave radiometers

Overall, the harmonized series from GROMOS and SOMORA show an improved agreement and validate well against external dataset. The documentation help explain most of the spurious periods identified on both time series. The outputs of the routine are also harmonized which makes further studies using both series much easier. In general, it should be noted that the calibration and retrievals of atmospheric constituents from MWRs is a never-ending process. The same applies to GROMOS and SOMORA and whereas the improvements with regards to the previous series suffer no doubts, there are many more work that can be done to improve it further. This work is actually in progress at the time of writing so the reader is definitely encouraged to contact the authors in case of any questions or suggestions.

In particular, the first version of the calibration and retrieval codes were focused on the post-2009 period, where both instruments used the same spectrometer. It is now being extended further to the past and will need to be fully homogeneized at some point. Also, in order for these dataset to be of used for the community, it is needed to upload them to NDACC, which is close to completion.

Bibliography

Buehler, S. A., Eriksson, P., Kuhn, T., von Engeln, A., and Verdes, C. (2005). ARTS, the atmospheric radiative transfer simulator. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91(1):65–93.

Buehler, S. A., Mendrok, J., Eriksson, P., Perrin, A., Larsson, R., and Lemke, O. (2018). ARTS, the Atmospheric Radiative Transfer Simulator - Version 2.2, the planetary toolbox edition. *Geoscientific Model Development*, 11(4):1537–1556.

Eriksson, P., Ekström, M., Melsheimer, C., and Buehler, S. A. (2006). Efficient forward modelling by matrix representation of sensor responses. *International Journal of Remote Sensing*, 27(9):1793–1808.

Ingold, T., Peter, R., and Kämpfer, N. (1998). Weighted mean tropospheric temperature and transmittance determination at millimeter-wave frequencies for ground-based applications. *Radio Science*, 33(4):905–918.

Janssen, M. A. (1993). *Atmospheric remote sensing by microwave radiometry.* Wiley series in remote sensing. Wiley.

Palm, M., Hoffmann, C. G., Golchert, S. H. W., and Notholt, J. (2010). The ground-based MW radiometer OZORAM on Spitsbergen – description and status of stratospheric and mesospheric O_3-measurements. *Atmospheric Measurement Techniques*, 3(6):1533–1545.

Parrish, A., Connor, B. J., Tsou, J. J., McDermid, I. S., and Chu, W. P. (1992). Ground-based microwave monitoring of stratospheric ozone. *Journal of Geophysical Research: Atmospheres*, 97(D2):2541–2546.

Perrin, A., Puzzarini, C., Colmont, J.-M., Verdes, C., Wlodarczak, G., Cazzoli, G., Buehler, S., Flaud, J.-M., and Demaison, J. (2005). Molecular Line Parameters for the "MASTER" (Millimeter Wave Acquisitions for Stratosphere/Troposphere Exchange Research) Database. *Journal of Atmospheric Chemistry*, 51(2):161–205.

Rodgers, C. D. (2000). *Inverse Methods for Atmospheric Sounding: Theory and Practice.* World Scientific. Google-Books-ID: FjxqDQAAQBAJ.

Rosenkranz, P. W. (1998). Water vapor microwave continuum absorption: A comparison of measurements and models. *Radio Science*, 33(4):919–928.

Sauvageat, E. (2021). Calibration routine for ground-based passive microwave radiometer: a user guide.

Sauvageat, E., Maillard Barras, E., Hocke, K., Haefele, A., and Murk, A. (2022). Harmonized retrieval of middle atmospheric ozone from two microwave radiometers in Switzerland. *Atmospheric Measurement Techniques*, 15(21):6395–6417.

Tsou, J. J., Connor, B. J., Parrish, A., McDermid, I. S., and Chu, W. P. (1995). Ground-based microwave monitoring of middle atmosphere ozone: Comparison to lidar and Stratospheric and Gas Experiment II satellite observations. *Journal of Geophysical Research*, 100(D2):3005.

Appendix A

Example of night time retrieved profiles

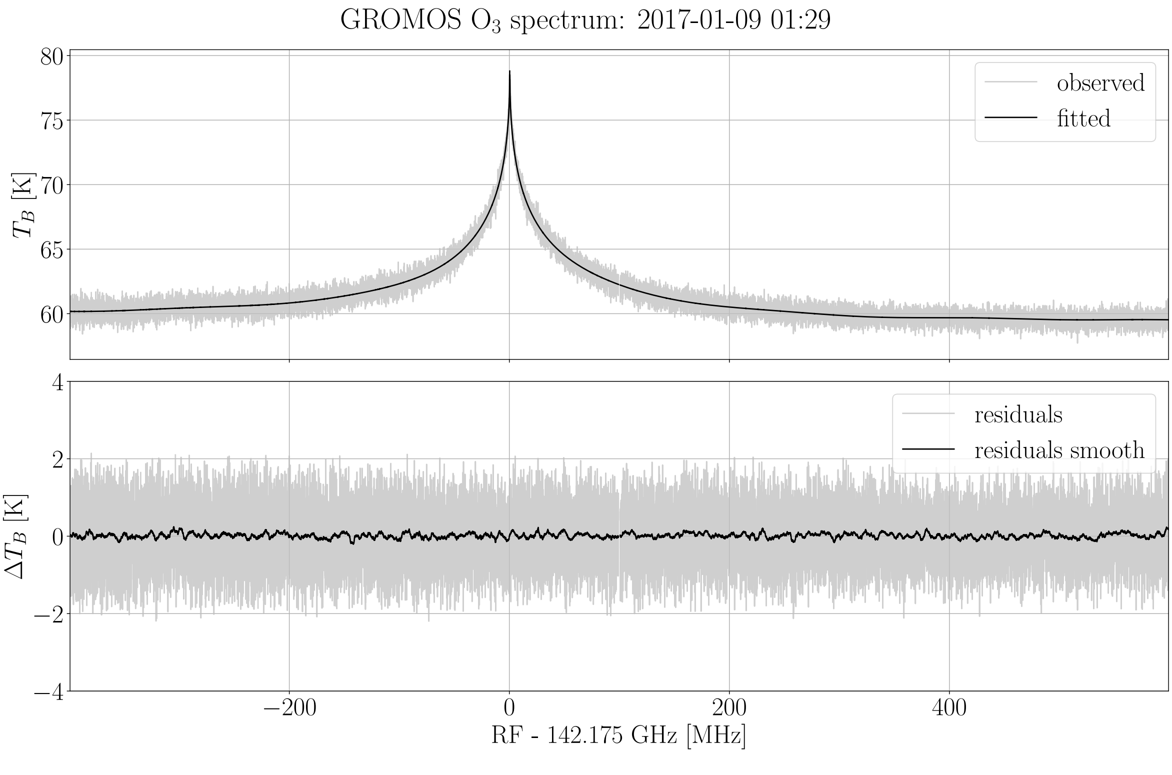

Figure A.1: Example of measured and fitted ozone emission spectra from GROMOS (nighttime)

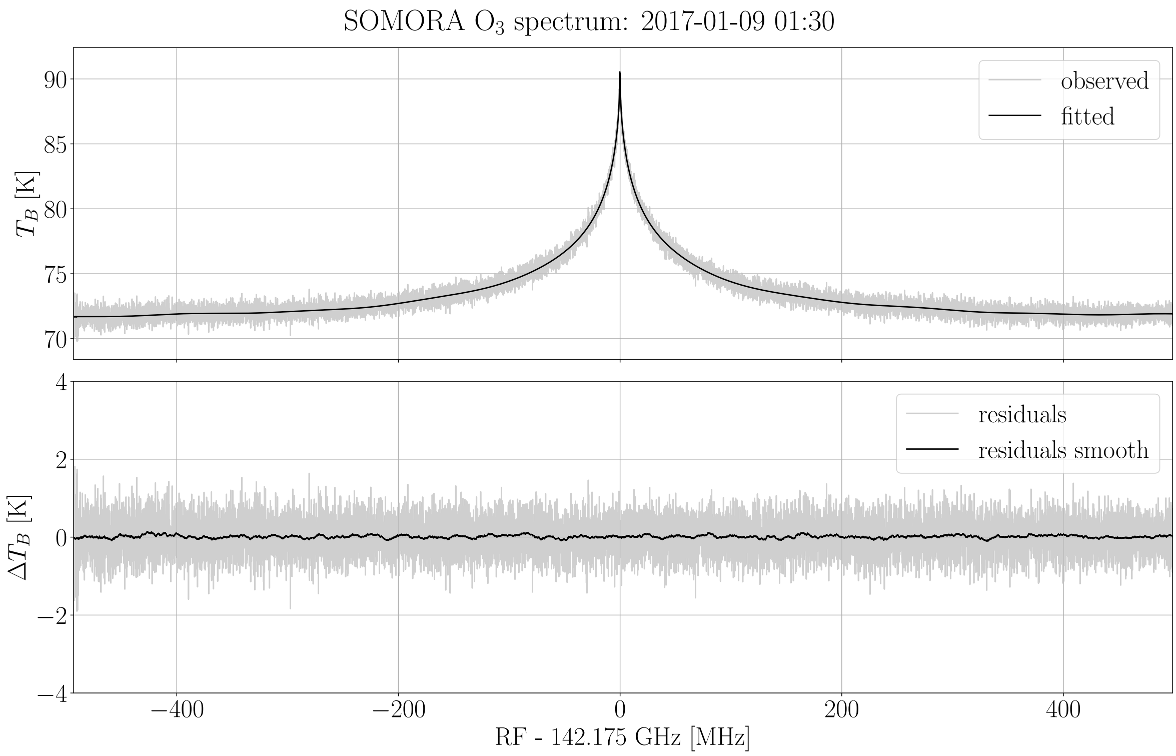

Figure A.2: Example of measured and fitted ozone emission spectra from SOMORA (nighttime)

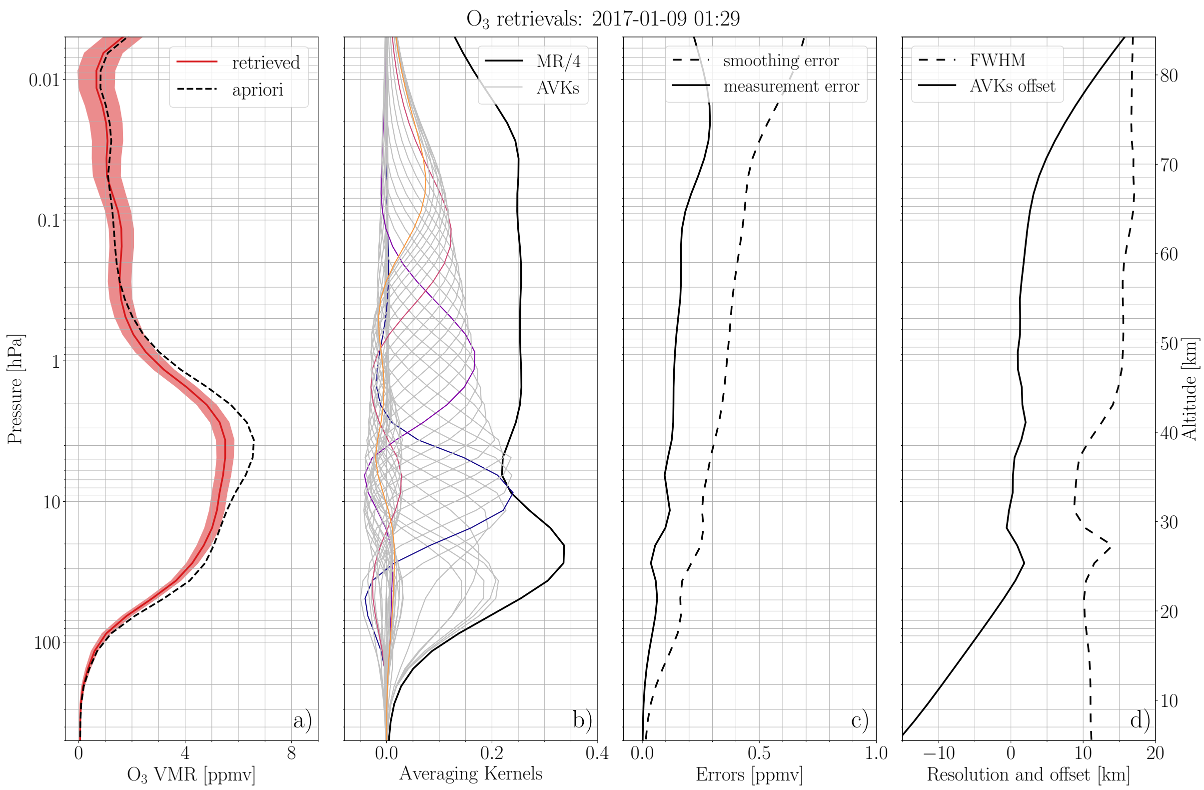

Figure A.3: Example of ozone profile retrieved from GROMOS (nightime).

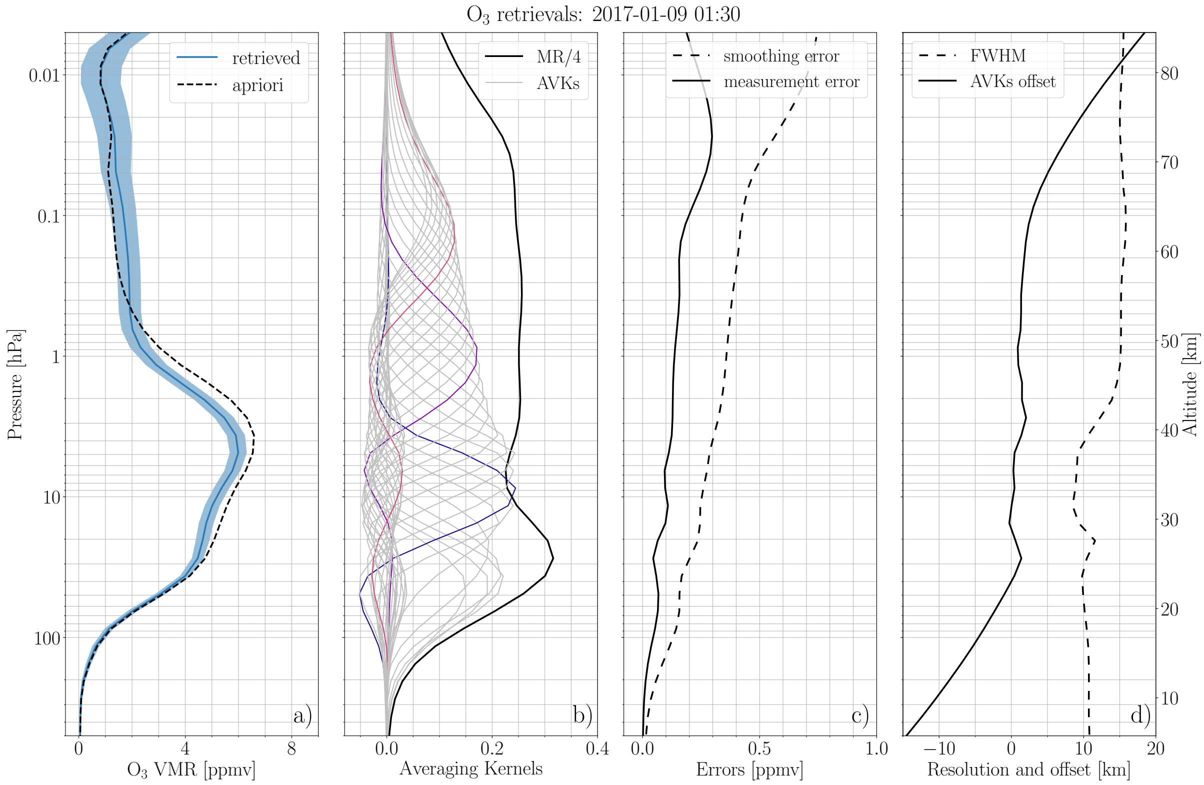

Figure A.4: Example of ozone profile retrieved from SOMORA (nighttime).