Odin/SMR

Algorithms Theoretical Basis Document - Level 2 processing

Author(s): Patrick Eriksson
Contact: Patrick Eriksson <patrick.eriksson@chalmers.se>
Version: 0.1
Date: 2015-11-16

Chalmers University of Technology
Department of Earth and Space Sciences

Contents

Acronyms

AC	Auto-correlation spectrometer
ARTS	Atmospheric radiative transfer simulator
ATBD	Algorithm theory basis document
AUG	ARTS user guide
ESA	European space agency
FWHM	Full width at half maximum
LM	Levenberg–Marquardt method
LO	Local oscillator
LOS	Line-of-sight
LTE	Local thermodynamic equilibrium
OEM	Optimal estimation method
OSIRIS	Optical spectrograph and infrared imaging system
SMR	Sub-millimetre radiometer
VMR	Volume mixing ratio

Introduction

Aim and scope of this document

Odin/SMR performs passive limb measurements of the atmosphere, mainly at wavelengths around 0.6 mm. The basic output of Odin/SMR is spectra in different frequency bands. After calibration the spectra are denoted as L1b data, and are grouped into limb scans for further processing. The overall aim of this document is to describe the software and algorithms used in this L2 processing, that is, the step of extracting geophysical data from L1b data. The aim includes also to motivate the different decisions involved, such as what instrumental parameters that shall be retrieved in parallel to the geophysical data to provide best possible L2 data product. This document describes the retrieval process and related questions in a general manner, while details and information that vary between observation modes are found elsewhere (see Sec. 1.3).

The document covers only the "operational" Odin/SMR L2 products. These products are based on measurements confined to tangent altitudes above the tropopause, with some margin, to allow that scattering can be neglected in the simulations of atmospheric radiative transfer. Observations that can be affected by scattering (due to tropospheric clouds) are treated separately, and the same applies to data measured when Odin has performed a special scanning sequence (see further Sec. 2.5). The result of these retrievals are formally also L2 products, but inside this document the term "L2" refers only to the operational products.

Background theory and an overview of the Odin/SMR L2 processing is found in Chapter 2. The software tool that handles the tasks directly associated with atmospheric radiative transfer and sensor modelling, the forward model, is described in Chapter 3, while the actual data extraction step is covered by Chapter 4. The L2 data are not complete without a characterisation of spatial resolution and errors, which is the topic of Chapter 5. Finally, Chapter 6 gives a summary, with focus on the most important points to correctly understand the Odin/SMR L2 data. A list of used acronyms is found directly after the table of contents.

Odin/SMR

The Odin satellite

The Odin satellite was launched on the 20th of February 2001, into a sun-synchronous 18:00 hour ascending node orbit, carrying two co-aligned limb sounding instruments: OSIRIS (Optical spectrograph and infrared imaging system) and SMR (Sub-millimetre radiometer). Originally, Odin was used for both atmospheric and astronomical observations, but since 2007 only its aeronomy mission is active. Odin is a Swedish-led project, in cooperation with Canada, France and Finland. Both of Odin's instruments are still functional,

and the present operation of the satellite is partly performed as a ESA third party mission.

The SMR instrument

The Odin/SMR package is highly flexible. In short, the four main receiver chains can be tuned to cover frequencies in the ranges 486–504 GHz and 541–581 GHz, but the maximum total instantaneous bandwidth is only 1.6 GHz. This bandwidth is determined by the two auto-correlation spectrometers (ACs) used for atmospheric observations. The two ACs can be coupled to any of the four front-ends, but only two or three front-ends are used simultaneously. The ACs cover 400 or 800 MHz per front-end, depending on configuration. In the configuration applied for atmospheric sounding, the channels of the ACs have a spacing of 1 MHz, while the frequency resolution is only 1.2 or 2 MHz (depending if a rectangular or Hanning window is applied or not in the conversion from autocorrelations to spectra). To cover all molecular transitions of interest, a number of "observation modes" have been defined. Each observation mode makes use of two or three frequency bands. Single sideband operation is obtained by tunable Martin–Pupplet interferometers. The nominal sideband suppression is better than 19 dB across the image band.

Odin/SMR also has a receiver chain around the 118 GHz oxygen transition, that was heavily used during Odin's astronomy mission. For the atmospheric mission, this front-end was planned to be used for retrieving temperature profiles, but a technical problem (drifting LO frequency) and the fact that the analysis requires treatment of Zeeman splitting have given these data low priority. This ATBD focuses on the processing of the sub-millimetre data but comments on the adoptions required to also handle the 118 GHz data are given

The main reflector of Odin/SMR has a diameter of 1.1 m, giving a vertical resolution at the tangent point of about 2 km. The vertical scanning of the two instruments' line-of-sight is achieved by a rotation of the satellite platform, with a rate matching a vertical speed of the tangent altitude of 750 m/s. Measurements are performed during both upward and downward scanning. The lower end of the scan is typically at about 7 km, the upper end varies between 70 and 110 km, depending on observation mode. In correspondence, the horizontal sampling ranges from 1 scan per 600 km to 1 scan per 1000 km. Measurements are in general performed along the orbit plane, providing a latitude coverage between 82.5°S and 82.5°N. Since the end of 2004 Odin is also pointing off-track during certain periods, e.g. during the austral summer season, allowing the latitudinal coverage to be extended towards the poles.

Main instrumental error sources

The receiver noise temperature of Odin/SMR differs between the front-ends, varying between 2600 and 4500 K. This makes thermal noise an important error source. The pointing of Odin can be reconstructed to an accuracy matching ~ 750 m in tangent altitude, and observed frequencies can be controlled with an accuracy of ~ 1 MHz. The retrieval process includes both a pointing and frequency correction, but there are some remaining errors.

Later revise the number in this paragrapph.

The three error sources discussed above basically follow nominal performance and are well characterised, but there are also features that are not yet understood. First of all, it is clear that the sideband suppression does not reach the nominal value of 19 dB in all frequency bands. Ongoing tests and investigations aim to obtain better knowledge on the actual suppression achieved. Further, the sub-bands of the ACs are not always overlapping perfectly, and some degree of non-linear response cannot be ruled out. Both these issues

are most problematic when measuring spectral features giving a high range of brightness temperatures over the frequency band.

Some other non-nominal features are handled, or are characterised, fairly well. Reflections inside the receiver give rise to "baseline ripple", but these ripples are largely removed as part of the L1b processing (Rydberg et al., 2016). The calibration uncertainty increases linearly with measured brightness temperature, and is presently estimated to ... Add result later.
Odin/SMR applies Dicke switching with respect to "cold sky" but there is some remaining impact of gain variations. These variations give rise to a constant shift of the brightness temperatures across the band (that is, it can be seen as a flat baseline ripple). These shifts have a standard deviation of about 2 K, and uncorrelated between tangent altitudes and front-ends. The retrieval can remove these shifts with marginal impact on the retrievals as long as some part of the measured spectrum corresponds to cosmic background radiation (that is, high tangent altitudes), but this is a main error source for other situations.

Further reading

Murtagh et al. (2002) give an overview of the Odin aeronomy mission, as well as the general technical details of Odin/SMR. Some further technical information is found in Eriksson et al. (2002), and a more detailed description of the observation modes is provided by Merino et al. (2002).

The input to this L2 processing is the L1b data described in Rydberg et al. (2016). The exact settings applied and the general characteristics of the L2 data, that both vary between observation modes, are described in ?. This document does not yet exist

Theory and overview

Terminology and data products

As mentioned in the Introduction, the overall topic of this document is the extraction of geophysical data from observed spectra. This calculation step is normally denoted as the retrieval or the inversion, where mainly the first term is used in this document. The objective of Odin/SMR is to estimate atmospheric quantities, mainly the concentration of different gases, and these variables are accordingly the main target of the retrieval, and, together with auxiliary information, they form the official L2 data. The format of the Odin/SMR L2 data is described in Appendix A. The L2 data are publicly available, visit odin.rss.chalmers.se for registration and downloading instructions.

Besides geophysical variables, the retrieval must consider a number of instrumental aspects to avoid that e.g. pointing errors cause unnecessary large retrieval errors. The result of this part of the retrieval can be seen as diagnostics data. Some of the quality fields of the L2 data are based on the diagnostics data, and the data are also stored internally for further analysis (then denoted as L2I), to e.g. detect instrumental drifts.

The retrieval can also include atmospheric state variables that influence the measurement, but the accuracy of the retrieval is not sufficiently high to justify inclusion in the L2 data. Regarding atmospheric gases, the ones included in the L2 data are denoted as target species, while other retrieved gases are called secondary species.

Radiative transfer basics

There is no ATBD dedicated to radiative transfer, as a general software is used to handle this task (Sec. 2.4.2). Instead the main considerations around radiative transfer are commnted inside this ATBD. For a more detailed discussion, targeting Odin/SMR, the reader is referred to Eriksson et al. (2002).

A strong advantage of using microwaves is that local thermodynamic equilibrium (LTE) can be assumed, even for radiative transfer in the mesosphere and lower thermosphere. Further, as long as the lowest altitude of the line-of-sight, the tangent point, is some kilometres above the tropopause, scattering can safely be neglected as Odin/SMR operates at wavelengths > 0.5 mm. Such a restriction in tangent altitudes is made for the operational Odin/SMR retrievals.

The combination of LTE and no scattering makes the basic simulation task relatively simple. In fact, when it comes to atmospheric radiative transfer, the required simulations boil down to summing up emission along the propagation path, weighted with the transmission between Odin/SMR and the point of emission. That is, the expression for radiative transfer to handle is (see e.g. Chandrasekhar, 1960, Eq. 50):

$$I(\nu) = I_0(\nu)e^{-\int_0^l k(s,\nu)ds} + \int_0^l k(s,\nu)B(T(s),\nu)e^{-\int_0^l k(s',\nu)ds'}ds, \tag{2.1}$$

where I is radiance, I_0 is the radiance along the line-of-sight (LOS) at the top of the atmosphere, where the propagation path starts, l is the length of this propagation path, k is the absorption coefficient, s is the distance along the propagation path, and B is the Planck function giving the emission of a blackbody at temperature T and frequency ν. Eq. 2.1 assumes that the gaseous absorption is unpolarised. This is a valid assumption in the microwave region for all species, except for oxygen which is affected by the Zeeman effect. That is, Earth's magnetic field induces polarised absorption and emission, and a matrix-vector equivalent to Eq. 2.1 is required to handle oxygen (see e.g. Larsson et al., 2014).

Evaluation of Eq. 2.1 gives the monochromatic intensity (radiance) along pencil-beam propagation paths, matching infinite frequency and angular resolution. Accordingly, Eq. 2.1 must be solved for a set of frequencies and propagation paths, to make it possible to incorporate the antenna pattern, channel frequency response and other characteristics of the sensor in the simulations (see e.g. Eriksson et al., 2002; Read et al., 2006).

Theoretical formalism

Forward and inverse models

The retrievals are presented and discussed using the formalism presented in Chapter 3 of Rodgers (2000). The data to be inverted are appended to create the measurement vector, $\mathbf{y}$. This vector is related to other variables as

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \mathbf{b}) + \varepsilon_n, \tag{2.2}$$

where $\mathcal{F}$ is denoted as the forward model, $\mathbf{x}$ as the state vector, $\mathbf{b}$ as the vector of forward model parameters and ε_n represents measurement noise. The distinction between the two arguments of the forward model is that all variables that we want to retrieve form $\mathbf{x}$, while all other quantities required of the forward model are placed in $\mathbf{b}$.

The forward model treated explicitly in this document is the software used to model atmospheric radiative transfer and sensor responses. This is a model operating with discrete quantities, while the "true", hypothetical, forward model, F, representing the actual physical mechanisms in the atmosphere and the instrument, must be seen as a continuous function. If $\mathcal{F}$ is constructed and used carefully, the discrete representation should not cause any fundamental problems, and it is here assumed that all deviations between F and $\mathcal{F}$ can be treated as imperfect values in $\mathbf{b}$. That is, $F = \mathcal{F}(\mathbf{x}, \mathbf{b})$, but the exact values of $\mathbf{b}$ are unknown and all we can do is to use the best possible estimate, $\hat{\mathbf{b}}$. That is:

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \hat{\mathbf{b}}) + \varepsilon_b + \varepsilon_n, \tag{2.3}$$

where

$$\varepsilon_b = F - \mathcal{F}(\mathbf{x}, \hat{\mathbf{b}}). \tag{2.4}$$

This difference (ε_b) is below denoted as the forward model uncertainty.

The retrieval process is formalised in the inverse model, $\mathcal{I}$:

$$\hat{\mathbf{x}} = \mathcal{I}(\mathbf{y}, \mathbf{x}_a, \hat{\mathbf{b}}, \mathbf{c}), \tag{2.5}$$

where $\hat{\mathbf{x}}$ is the retrieved state vector and $\mathbf{c}$ covers all additional variables introduced by the inverse model. The exact nature of the vector $\mathbf{x}_a$ varies between retrieval approaches, but in general it represents an a priori estimate of $\mathbf{x}$.

Linearisation

From this point it is assumed that the retrieval problem is not strongly non-linear, and a local linear analysis is possible. Or expressed differently, that derivatives of the forward and inverse models are approximately valid over a significant range. The Jacobian, or the weighting function matrix, is the partial derivative of $\mathcal{F}$ with respect to $\mathbf{x}$:

$$\mathbf{K_x} = \frac{\partial \mathcal{F}}{\partial \mathbf{x}}. \tag{2.6}$$

In the same manner, we define $\mathbf{K_b}$ as

$$\mathbf{K_b} = \frac{\partial \mathcal{F}}{\partial \mathbf{b}}. \tag{2.7}$$

The contribution function matrix, $\mathbf{G}$, is defined as the partial derivative of the inverse model with respect to the measurement vector:

$$\mathbf{G} = \frac{\partial \mathcal{I}}{\partial \mathbf{y}}. \tag{2.8}$$

Having these partial derivatives at hand, the retrieval error can be related to the fundamental uncertainties. As a first step, the forward model is linearised around $(\mathbf{x}_a, \hat{\mathbf{b}})$:

$$\mathbf{y} = \mathcal{F}(\mathbf{x}_a, \hat{\mathbf{b}}) + \mathbf{K_x}\left(\mathbf{x} - \mathbf{x}_a\right) + \mathbf{K_b}\left(\mathbf{b} - \hat{\mathbf{b}}\right) + \varepsilon_n \tag{2.9}$$

Under these assumptions, we now have a second way to express the forward model uncertainty:

$$\varepsilon_b = \mathbf{K_b}\left(\mathbf{b} - \hat{\mathbf{b}}\right). \tag{2.10}$$

By combining the equations above and rearranging the terms, we can finally derive an expression for the total retrieval error, $\delta = \hat{\mathbf{x}} - \mathbf{x}$:

$$\delta = \left(\mathbf{A} - \mathbf{1}\right)\left(\mathbf{x} - \mathbf{x}_a\right) + \mathbf{G}\mathbf{K_b}\left(\mathbf{b} - \hat{\mathbf{b}}\right) + \mathbf{G}\varepsilon_n, \tag{2.11}$$

where $\mathbf{1}$ is the identity matrix and

$$\mathbf{A} = \mathbf{G}\mathbf{K_x}, \tag{2.12}$$

is the averaging kernel matrix. The terms on the right hand side of Eq. 2.11 are denoted as smoothing error, forward model retrieval error and measurement noise retrieval error, respectively. These error components are discussed further in Chapter 6.

Calculation of retrieval error

Despite that Eq. 2.11 gives an expression for the retrieval error it cannot be used in practice, for the simple reason that $\mathbf{x}$ and $\mathbf{b}$ are not known. The retrieval error can only evaluated in a statistical sense, and for this purpose uncertainties and errors are described by covariance matrices ($\mathbf{S}$). See e.g. en.wikipedia.org/wiki/Covariance_matrix for a description of this type of matrices, as well as the basic calculation rules needed to understand the error propagation. The statistical correspondence to Eq. 2.11 is

$$\mathbf{S}_\delta = \left(\mathbf{A} - \mathbf{1}\right)\mathbf{S_x}\left(\mathbf{A} - \mathbf{1}\right)^T + \mathbf{G}\mathbf{K_b}\mathbf{S_b}\mathbf{K_b}^T\mathbf{G}^T + \mathbf{G}\mathbf{S}_{\varepsilon_n}\mathbf{G}^T. \tag{2.13}$$

The combination of the instrument and the forward model can be seen as the "observation system", then having an uncertainty of

$$\varepsilon_o = \varepsilon_b + \varepsilon_n. \tag{2.14}$$

The covariance matrix of the observation system uncertainty is

$$\mathbf{S}_{\varepsilon_o} = \mathbf{K_b}\mathbf{S_b}\mathbf{K_b}^T + \mathbf{S}_{\varepsilon_n}. \tag{2.15}$$

Using the definition of $\mathbf{S}_{\varepsilon_o}$, Eq. 2.13 can be written in a somewhat more compact form:

$$\mathbf{S}_\delta = (\mathbf{A} - \mathbf{1})\,\mathbf{S_x}\,(\mathbf{A} - \mathbf{1})^T + \mathbf{G}\mathbf{S}_{\varepsilon_o}\mathbf{G}^T. \tag{2.16}$$

Selected set-up

Retrieval method

The need for regularisation

In short, the task of the retrieval method is to derive $\hat{\mathbf{x}}$ from $\mathbf{y}$. It turns out that the mapping from $\mathbf{y}$ to $\hat{\mathbf{x}}$ is not unique. That is, there exists an infinite set of $\mathbf{x}$-vectors that result in a fit with the measurement, considering the measurement noise. Instead, some "optimal" state must be selected, among this set of possible solutions.

If we for a moment assume that the retrieval problem can be treated as fully linear around $(\mathbf{x}_a,\hat{\mathbf{b}})$, we have that $\mathbf{K_x}\,(\mathbf{x} - \mathbf{x}_a) \approx \mathbf{y} - \mathcal{F}(\mathbf{x}_a, \hat{\mathbf{b}})$ and the solution can be written as

$$\hat{\mathbf{x}} = \mathbf{x}_a + \mathbf{G}(\mathbf{y} - \mathcal{F}(\mathbf{x}_a, \hat{\mathbf{b}})). \tag{2.17}$$

It is clear that $\mathbf{G}$, at least roughly, represents an inverse of $\mathbf{K_x}$, but to use the standard inverse of $\mathbf{K_x}$ (only possible if $\mathbf{K_x}$ is square) or to apply the least squares method ($\mathbf{G} = (\mathbf{K_x}^T\mathbf{K_x})^{-1}\mathbf{K_x}$) will result in an extremely high sensitivity to noise. The retrieval problem at hand is said to be ill-posed, which is the standard situation for passive atmospheric sounding.

The optimal estimation method

The normal way to tackle ill-posed problems of this type is "regularisation", where some constrain on the solution is introduced. In passive atmospheric sounding, using statistical information as basis for the regularisation has become the standard approach, and it is also the approach selected for Odin/SMR. The method is probably most widely known as the optimal estimation method (OEM), which is the name used in this document. The method can be seen as an application of Bayes theorem, under certain conditions, and another possible name of the method the maximum a posteriori solution, see further Rodgers (2000).

The aspect of "statistical regularisation" is probably most clearly identified by noting that OEM minimises the following "cost function":

$$C(\mathbf{x}) = (\mathbf{y} - \mathcal{F}(\mathbf{x}, \hat{\mathbf{b}}))^T\mathbf{S}_o^{-1}(\mathbf{y} - \mathcal{F}(\mathbf{x}, \hat{\mathbf{b}})) + (\mathbf{x} - \mathbf{x}_a)^T\mathbf{S}_a^{-1}(\mathbf{x} - \mathbf{x}_a), \tag{2.18}$$

that is,

$$\hat{\mathbf{x}} = \min_{\mathbf{x}}(C(\mathbf{x})). \tag{2.19}$$

The cost function C can be seen as the sum of the two penalty terms. The first term weights how well the solution corresponds to the measurement. If C would be set to only contain this term, the standard least squares solution would be obtained. The regularisation is introduced by the second term, that evaluates the probability of $\mathbf{x}$ based on a priori information.

If the retrieval problem is linear around $(\mathbf{x}_a,\hat{\mathbf{b}})$, $\hat{\mathbf{x}}$ can be calculated using Eq. 2.17, with (Rodgers, 2000, Eq. 4.5)

$$\mathbf{G} = (\mathbf{K}_\mathbf{x}^T\mathbf{S}_o^{-1}\mathbf{K}_\mathbf{x} + \mathbf{S}_a^{-1})^{-1}\mathbf{K}_\mathbf{x}^T\mathbf{S}_o^{-1}. \tag{2.20}$$

However, the inversion of Odin/SMR measurements result in a (moderately) non-linear problem and an iterative process is required to find the minimum of C, see Sec. 4.2.

Considerations around x and b

Formally, $\mathbf{S}_a$ shall be set to $\mathbf{S}_\mathbf{x}$ (Rodgers, 2000), that is, a description of the natural variability (including correlations) of the elements in $\mathbf{x}$, but this rule is not always obeyed. First of all, normally there is no manner in which an exact $\mathbf{S}_\mathbf{x}$ can be derived, and this matrix is set in some parametric way, loosely based on measurements and general knowledge of atmospheric physics. It could also be the case that $\mathbf{S}_a$ is partly used as a "tuning parameter". The most common example is that the variances in $\mathbf{S}_a$ are set with some margin with respect to the true values, to maintain measurement information as far as possible, or expressed differently, not to constrain the solution more than necessary. Further, it is likely that $\mathbf{x}_a$ differs from the true mean value of $\mathbf{x}$, and this systematic deviation should also be represented in $\mathbf{S}_a$ (Eriksson, 2000). For a continued discussion, and comments on the relationship between $\mathbf{S}_a$ and the corresponding regularisation matrix used in "Tikhonov regularisation", see, for example, Eriksson (2000) and Ungermann (2011).

Finally regarding $\mathbf{S}_a$, the point made in Sec. 2.6 of Rodgers (2000) is stressed, for given values along the diagonal, setting $\mathbf{S}_a$ to be purely diagonal means a stronger regularisation than including reasonable off-diagonal elements (to reflect correlations). It is a popular mistake to think that this relationship is reversed. Here it should be noted that using a diagonal $\mathbf{S}_a$ equals the common choice in Tikhonov regularisation to select the "smallest norm" ($\mathbf{L}_0$) solution.

The formally correct choice for $\mathbf{S}_o$ is $\mathbf{S}_{\varepsilon_o}$ (Eriksson, 2000; Rodgers, 2000), but again this theoretical guideline is seldomly followed exactly. The normal choice is to set $\mathbf{S}_o = \mathbf{S}_{\varepsilon_n}$. One important practical consideration is that the inverse of $\mathbf{S}_{\varepsilon_o}$ must be calculated and this is a highly costly operation in the general case, but $\mathbf{S}_{\varepsilon_n}$ is in general a pure diagonal matrix and the inverse can be determined with a minimal calculation cost.

Further, the setting of $\mathbf{S}_o$ is influenced by the exact choice of variables included in $\mathbf{x}$. As discussed in Sec. 4.1.2 of Rodgers (2000) there is no strict division line between $\mathbf{x}$ and $\mathbf{b}$. Obviously, $\mathbf{x}$ must contain a representation of the quantities targeted by the measurements, but for variables of interfering character (e.g. frequency off-set) there is a choice to make. For a perfectly linear retrieval case, exactly the same results will be obtained for the target quantities independent of whether interfering factors are covered by $\mathbf{x}$ or $\mathbf{b}$. However, inclusion in $\mathbf{x}$ is to prefer as this will result in that values describing the interfering effects are obtained, which can provide important diagnostic information (e.g. to detect instrumental changes). In addition, in a non-linear situation a more optimal solution is obtained by placing interfering factors in $\mathbf{x}$, as it is likely that $\mathbf{K}_\mathbf{b}$, and then also $\mathbf{S}_{\varepsilon_o}$ (cf. Eq. 2.15), varies with $\mathbf{x}$. That is, the description of observation uncertainties is likely to deteriorate with the deviation between $\mathbf{x}$ and $\mathbf{x}_a$. If interfering species are part of $\mathbf{x}$, such non-linear effects will be taken care of by the reevaluation of $\mathbf{K}_\mathbf{x}$.

Forward model

Some basic requirements on the forward model were introduced in Sec. 2.2. The simulation problem to be tackled is basically the simplest possible, Eq. 2.1 should be handled by all forward models for passive microwave measurements (but limb sounding requires that refraction and the spherical shape of the Earth is considered when determining the propagation path, which excludes all "plane-parallel" models). On the other hand, the analysis of some Odin/SMR data requires non-standard forward model features: the 118 GHz oxygen measurements give rise to a demand for treatment of Zeeman effects, and Doppler effects due to winds could have a significant impact on mesospheric spectra.

Beside the atmospheric radiative transfer, the second main task of the forward model is to incorporate effects of sensor characteristics. For the Odin/SMR radiometer package the sensor features needing consideration are antenna angular response, mixer, sideband filtering and frequency response of each spectrometer channel. Compensation of the Doppler shift due to the satellite's movement is part of the L1b processing.

A further requirement is that the forward model can deliver the Jacobian ($\mathbf{K_x}$), a demand that should be clear from Eq. 2.20. This task must be handled for all quantities that can be part of the state vector. That is, the Jacobian is needed not only for the relevant atmospheric variables (VMRs, temperature and winds), but also fitting of various sensor imperfections must be handled (pointing off-set, frequency off-set and baseline fit), see further Sec. 5.4.

In contrast to many other satellite missions (e.g. Read et al., 2006), no Odin/SMR specific forward model has been developed. Instead, the general forward model ARTS (Atmospheric radiative transfer simulator) is used for Odin/SMR L2 processing. Earlier operational L2 processing (Urban et al., 2005) made use of the first ARTS version (and also the Moliere forward model (Urban et al., 2004)), while the retrievals described in this document are based on the latest version (version 2.3). At this point it is just noted that ARTS complies fully with the demands listed above, and the presentation of ARTS is continued in Chapter 3.

Non-standard retrievals

As mentioned already in Sec. 1.1, the operational retrievals do not handle all Odin/SMR measurements. Most importantly, scattering is not activated in the forward model simulations and special retrievals have been developed to extract information from spectra containing tropospheric information. This retrieval scheme was introduced by Rydberg et al. (2009) and the latest results are found in Eriksson et al. (2014). However, the method of Rydberg et al. (2009) is only applicable for tangent altitudes below 9 km and latitudes between 30°S and 30°N. This leaves a gap between the operational and the tropical-tropospheric retrievals, and there are still Odin/SMR spectra that are not part of any inversion.

The operational processing is of 1D character (Sec. 3.2), as the standard limb scanning pattern of Odin/SMR does not provide a basis for resolving along-track features. However, Odin has during some shorter periods made measurements dedicated to polar mesospheric clouds. During these measurement campaigns, Odin has just scanned over a narrow altitude region in order to enhance the along-track resolution. Accordingly, to maintain the measurement information of these special observations, 2D (or tomographic) retrievals were needed. These retrievals were also made by combing OEM and ARTS but are presented separately, in Christensen et al. (2015).

The ARTS forward model

General features

Main features

ARTS, the Atmospheric radiative transfer simulator, is a publicly available software that is both developed and maintained as a Free open-source software project. The software was initiated 1999, and the first official version was released some years later (Buehler et al., 2005). The main novelty of this initial version was the fact that ARTS behaves more or less as a scripting language, in contrast to most (or all?) other forward models that are controlled by a fixed set of keywords. The control file was introduced to meet the design goal of high modularity. When defining an ARTS control file, the user can select between a large number of workspace methods and variables. Regarding physical mechanisms covered, ARTS v1 did not stand out in any way, but despite this the software found a number of applications, including operational SMR processing (Sec. 2.4.2).

Already before the first version was released, a development branch was started with the aim of adding more advanced features to ARTS. At the time, treatment of scattering and going beyond horizontally homogeneous atmospheres (1D) were the main aims. This development resulted in a second main version ARTS (Eriksson et al., 2011), and further extensions have been made more recently. ARTS can today handle several observation techniques, as well as simulations for other planets, but only aspects relevant for Odin/SMR are discussed below. Odin/SMR measures thermal emission in a limb sounding geometry and, in this context, the most important features of ARTS are:

- The full polarisation state of radiation can by described, using the Stokes formalism.
- Atmospheric fields can be defined to vary in one dimension (1D; pressure), two dimensions (2D; pressure and angle along the orbit) or three dimensions (3D; pressure, latitude and longitude).
- The planet's overall shape is spherical (also spheroidal allowed for 2D and 3D), there are no assumptions of a "flat Earth".
- Gaseous absorption can be calculated based on several spectroscopic databases and a high number of absorption parameterisations.
- The impact of the magnetic field on oxygen absorption due to the Zeeman effect can be considered (Larsson et al., 2014).
- Refraction can be considered.
- There is extensive support to incorporate sensor characteristics.

- The Jacobian can be obtained for a high number of atmospheric and sensor variables, in general with limited additional calculation time (but only for non-scattering calculations).
- Scattering can be incorporated by two different algorithms, denoted as DOIT (Emde et al., 2004) and MC (Davis et al., 2005).

Access and documentation

The ARTS website is www.radiativetransfer.org. Version 1 has been declared obsolete and is now limited to internal use, while downloading instructions for recent public releases and the development branch are found at www.radiativetransfer.org/getarts. Some external packages are needed, generally or for special features, and these are listed at www.radiativetransfer.org/tools.

To obtain an overview of ARTS as a forward model, the primary reading should be Buehler et al. (2005) and Eriksson et al. (2011). A more practically inclined introduction to ARTS is provided by the "ARTS user guide" (AUG). Some background theory is found in "ARTS theory", and some parts of ARTS are even described by dedicated journal articles (see references throughout this chapter). However, the core documentation of ARTS is the built-in description of workspace methods and variables. This is the main information source for more experienced users and this documentation for the present stable version can be browsed at: www.radiativetransfer.org/docserver-stable. Finally, the best way to get started with practical calculations is to copy and modify some of the test and demonstration control files distributed with the software. For further details and links to the ARTS user guide and other documents, see www.radiativetransfer.org/docs.

Validation and quality assurance

ARTS has been compared to other forward models by Melsheimer et al. (2005), Buehler et al. (2006) and Saunders et al. (2007), as well as in more informal manners. A number of test cases have been defined for ARTS, targeting either a special part of ARTS or a complete calculation of some type. These tests are run at each new commit of code to the software repository, in order to catch bugs and unexpected side-effects of code changes as early as possible. As ARTS is a very flexible tool, it is impossible to design a comprehensive set of tests, but standard applications of ARTS should be fairly well covered by the tests. Some of the tests include comparison to a stored set of reference values, in order to test that the code not only runs, but also that the actual simulation results have not changed beyond some specified tolerance.

The L2 processing chain should have an internal quality check. How?

Configuration used

ARTS version and compilation options

The processing chain contains a check that a specific version of ARTS is used. The version used presently is ARTS 2.3.?. ARTS is written in C++, and a compilation for each platform architecture is needed. The `cmake` build system is used to set up the compilation of ARTS. A minimal version of ARTS is created using the following options:

Add number later.

```
cmake -DCMAKE_BUILD_TYPE=Release -DNO_NETCDF=1 -DNO_DOCSERVER=1 ..
```

Finish this section later. During the development phase we should just use latest version of the development branch. Later we should select a specific version and add check in the code that exactly this version is used. The executable should be minimal ARTS (no DISORT, T-Matrix ...). What level optimisation should be used?

Basic settings

The atmosphere is set to be 1D. That is, atmospheric fields vary only as a function of pressure, and are implicitly treated as constant in the latitude and longitude dimensions. A 1D atmosphere implies further a spherical reference geoid, and the radius of the this geoid is set to 6371 km. The upper limit of the model atmosphere is defind in ARTS by the lowest value in the user defined pressure grid. The range of this grid varies, see Sec. 5.7.1. The lower limit of the atmosphere is determined by the planet's surface, and ARTS demands that a surface is always defined. The altitude of the surface is for these simulations set to 1 km, and regarding radiative properties it is set to behave as a blackbody.

It is possible to use a "dummy" surface for the following reason. Even if data with (bore-sight) tangent altitudes below the tropopause are rejected, the outer parts of the antenna pattern can result in propagation paths that theoretically intersect with the surface, but the net impact of the surface is still zero. This is the case due to the very high optical thickness of the middle and lower troposphere at frequencies around 500 GHz (Ekström et al., 2007). Even for tangent altitudes around 0 km only altitudes above ∼8 km effectively contribute to measured radiance.

The atmosphere is assumed to not cause scattering, which is a valid assumption for the wavelengths of Odin/SMR and tangent altitudes some kilometres above the tropopause.

Gaseous absorption

In lack of scattering and assuming LTE, the local radiative properties of the atmosphere are solely determined by the absorption coefficient. ARTS has two main ways of calculating absorption, "on-the-fly" or by look-up tables, where the latter option is by far the most efficient for simulations of Odin/SMR. That is, molecular absorption cross-sections are pre-calculated and stored in tables. One table is created for each observation mode. The tables contain absorption cross-sections for combinations of pressure and temperature, and final values are obtained by interpolation. It is throughout assumed that the absorption cross-sections do not vary with the mixing ratio of the gas, and that there are no "non-linear" species in the absorption table (this option is required for water vapour in the lower troposphere). ARTS' absorption look-up table system is described in detail by Buehler et al. (2011). The spectroscopic data used to produce the look-up tables are discussed in Sec. 5.7.3).

Radiative transfer

Besides for 118 GHz (see below), the simulations follow Eq. 2.1. Only the first element of the Stokes vector is calculated (remaining three elements are zero for assumed conditions).

Solving equation Eq. 2.1 requires two main steps, to determine the propagation path and to perform the integration along the found path. Refraction must be considered for Odin/SMR (Eriksson et al., 2002), which is achieved by a simple ray tracing scheme (see AUG). The maximum step length of the ray tracing is set to 20 km, which is estimated to give a tangent point accuracy better than 5 m for tangent altitudes above 18 km.

The radiance at the top of the atmosphere, I_0 is set to mimic cosmic background radiation. There is no attempt to model situations when the moon is found in the LOS, since the retrievals reject such spectra. The combination of Odin orbit and keeping the LOS close to the orbit plane results in that there is never a need to consider the Sun.

118 GHz

The standard absorption and radiative transfer settings cannot be used for simulations targeting the 118 GHz oxygen transition, as the Zeeman effect must be considered. First of all, a matrix-vector version Eq. 2.1 must be applied (Larsson et al., 2014), where the full Stokes vector is calculated. Further, absorption must be calculated "on-the-fly" as the look-up table approach does not handle polarised absorption. However, even with an extension of the lookup-tables, the "on-the-fly" option would still be to prefer as in this case only the 118 GHz transition itself must be considered (in contrast to the 0.6 mm frequency bands where hundreds of transitions must be summed up) and a full calculation turns out to be faster than performing a multi-dimensional interpolation.

Sensor characteristics and calibration

The impact of the various sensor responses is incorporated following Eriksson et al. (2006). In this approach calculation time is saved by performing pre-calculations, but, in contrast to the absorption tables, no post-processing is needed. In short, the impact of the different sensor parts is all linear operationsi, and a "response matrix" representing the complete receiver system can be determined. This calculation can be done before the atmospheric radiative transfer has been performed, as long as the calculation grids are set and the description of the sensor responses are at hand. The response matrix obtained, $\mathbf{H}$, is applied as:

$$\mathbf{y}_f = \mathbf{H}\mathbf{i} \tag{3.1}$$

where $\mathbf{y}_f$ is the measurement vector produced by the forward model and $\mathbf{i}$ is all monochromatic pencil beam radiances compiled into a vector. The vector $\mathbf{i}$ must contain a certain number of values to represent the radiation field, as a function of frequency and zenith angle, with sufficient accuracy. The methods setting up $\mathbf{H}$ assumes that both radiances and sensor responses vary linearly between the points were data are available. However, these exists an option to include a polynomial frequency interpolation of monochromatic data, to create a more dense grid before the sensor responses are applied. This option decreases the overall calculation time, but makes the selection of the monochromatic frequency grid more critical. The variables controlling $\mathbf{H}$ and $\mathbf{i}$ are discussed in Sec. 5.7.

The conversion of the Odin/SMR's L1b data to brightness temperatures (T_b, K) must also be mimicked by the forward model. The radiances (I, W/(m^2·Hz·sr)) given by the atmospheric radiative transfer part are converted to follow the Rayleigh–Jeans approximation of the Planck function:

$$T_b = \frac{c^2}{2\nu k_B} I, \tag{3.2}$$

where c is the speed of light and k_B is the Boltzmann constant. This conversion is made before the sensor responses are applied.

Important aspects

Some of the definitions of ARTS have noticeable implications for the L2 processing:

- The basic vertical coordinate in ARTS is pressure. This has the consequence that the retrieval grid for atmospheric quantities is a set of pressures, and that retrieved profiles are functions of pressure.
- With some exceptions of no concern here, ARTS assumes data to vary linearly between grid points. This applies to everything from sensor responses (such as the input antenna pattern file) to how retrieved data shall be interpreted. However, for the pressure dimension the variation is assumed to be linear in the logarithm of the pressure. That is, if a retrieved atmospheric profile shall be interpolated to another set of pressures, it should be done as (using Matlab notation):

  ```
  x_new = interp1( log(p_grid), x, log(p_grid_new) )
  ```

  where `p_grid` is the original pressure grid, `x` is the atmospheric profile to be interpolated, and `p_grid_new` is the pressure grid.
- Should be more to add here

OEM implementation

A general introduction to the optimal estimation method (OEM) is given in Sec. 2.4.1, while details more specific for the Odin/SMR L2 processing are treated in this section.

Software

The Odin/SMR retrievals are performed by a Matlab implementation of OEM that is part of the Atmlab package. This package is available through the ARTS site, at `www.radiativetransfer.org/tools`. The implementation of OEM in Atmlab, `oem.m`, aims at being generic, that is, it shall be possible to couple the function to different forward models. The Atmlab package contains the required functions to interface `oem.m` with ARTS. This functionality has been used to create the second version of the Qpack (Eriksson et al., 2005) inversion system. This version, Qpack2, is used by several groups operating ground-based microwave radiometers (e.g. Tschanz and Kämpfer, 2015). The function `oem.m` was coupled to a lidar forward model by Sica and Haefele (2015).

To be clear, earlier L2 processing made direct use of the Qpack system (version 1), but, in order to allow a higher degree of optimisation with respect to Odin/SMR, this is no longer the case. Instead, the core functionality, on which Qpack2 is based, is used to set up a retrieval system targeting the needs raised by Odin/SMR. The main components taken from Atmlab are `oem.m`, functions to read/write ARTS output/input files, functions to interpolate data of climatology character (denoted as "atmdata" inside Atmlab), and functions to create parametric variance-covariance matrices. All these functions are also used by Qpack2 and, hence, are well tested.

Iteration scheme

The Odin/SMR processing offers a non-linear retrieval problem and an iterative procedure is required to determine the solution (Eq. 2.19). A number of iteration schemes has been suggested. The simplest version is Gauss-Newton (see Rodgers, 2000, Sec. 5.3). This iteration scheme assumes a robust decrease of the cost function (C, Eq. 2.18) during the iterations, but this is seldom the case in practical retrievals and instead Gauss–Newton can result in that C is increasing while iterating. The standard choice among iteration approaches including a manner to enforce a decrease of the cost function at each step, is the Levenberg - Marquardt method (LM), and it is also selected for these retrievals.

LM operates with a parameter γ. With $\gamma = 0$, LM becomes equal to Gauss–Newton iteration. For large values γ, LM instead behaves as a steepest descent method. The later results in a relatively small change in $\mathbf{x}$, and then also a slow convergence rate, but $\mathbf{x}$ is updated in a direction that is close to optimal to ensure a decrease of C. See Rodgers (2000, Sec. 5.7) for a more detailed presentation of LM. There exists different versions of

LM. The version applied here is (Rodgers, 2000, Eq. 5.36)

$$\hat{\mathbf{x}}_c = \hat{\mathbf{x}}_i + \left[(1+\gamma)\mathbf{S}_a^{-1} + \mathbf{K}_i^T\mathbf{S}_o^{-1}\mathbf{K}_i\right]^{-1}\left[\mathbf{K}_i^T\mathbf{S}_o^{-1}(\mathbf{y} - \mathcal{F}(\hat{\mathbf{x}}_i, \hat{\mathbf{b}})) - \mathbf{S}_a^{-1}(\hat{\mathbf{x}}_i - \mathbf{x}_a)\right] \quad (4.1)$$

where $\hat{\mathbf{x}}_c$ is the candidate solution for iteration $i+1$, $\hat{\mathbf{x}}_i$ is the (final) solution after iteration i, and $\mathbf{K}_i$ is the Jacobian ($\mathbf{K}_\mathbf{x}$) using $(\hat{\mathbf{x}}_i, \hat{\mathbf{b}})$ as linearisation point. Other variables are introduced in Sec. 2.3. The iteration is started by setting $\hat{\mathbf{x}}_0 = \mathbf{x}_a$, and $\hat{\mathbf{x}}_c$ becomes $\hat{\mathbf{x}}_{i+1}$ if $C(\hat{\mathbf{x}}_c) < C(\hat{\mathbf{x}}_i)$.

The scheme for updating γ must consider both successful ($C(\hat{\mathbf{x}}_c) < C(\hat{\mathbf{x}}_i)$) and unsuccessful ($C(\hat{\mathbf{x}}_c) \geq C(\hat{\mathbf{x}}_i)$) iterations. It also needs to handle problematic cases by e.g. setting a limit on the number of iterations to perform. The overall iteration scheme applied is:

0 Set start values: $\hat{\mathbf{x}}_0 = \mathbf{x}_a$, $i = 0$ and $\gamma = \gamma_0$.

1 If max iterations reached, $i = i_{\max}$, jump to 5.

2 Calculate $\hat{\mathbf{x}}_c$ by Eq. 4.1.

3 If $C(\hat{\mathbf{x}}_c) < C(\hat{\mathbf{x}}_i)$:

- 3a Set $\hat{\mathbf{x}}_{i+1} = \hat{\mathbf{x}}_c$.
- 3b If convergence is reached (see Sec. 4.3), set $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{i+1}$ and jump to 5.
- 3c Decrease γ with a factor f_s. That is, γ is updated as $\gamma \leftarrow \gamma/f_s$.
- 3d If γ becomes smaller than a threshold value, $\gamma < \gamma_{\min}$, set $\gamma = 0$.
- 3e Continue iterations by jumping to 1.

4 If $C(\hat{\mathbf{x}}_c) \geq C(\hat{\mathbf{x}}_i)$:

- 4a Increase γ. If $\gamma = 0$, set $\gamma = \gamma_{\min}$. Otherwise, updated as $\gamma \leftarrow f_u \cdot \gamma$.
- 4b If γ becomes larger than a threshold value, $\gamma > \gamma_{\max}$, set $\hat{\mathbf{x}} = \hat{\mathbf{x}}_i$ and jump to 5.
- 4c Re-do the iteration with new γ by moving to 2.

5 Perform retrieval characterisation (Chapter 6), using $\mathbf{K}_\mathbf{x} = \mathbf{K}_i$.

The values of γ_0, f_s, f_u, $\gamma_{\min}$, $\gamma_{\max}$ and $i_{\max}$ vary between the retrieval of the different observation modes. Actual values are listed in ?. Where?

Convergence tests and values

Stop criterion

There are different aspects related to convergence. A first question is the criterion for halting the iterations. This issue is discussed carefully in Sec. 5.6.3 of Rodgers (2000) and the details are not repeated here. The stop criterion applied is obtained by combining Eqs. 5.29 and 5.30 in the book section cited:

$$(\hat{\mathbf{x}}_{i+1} - \hat{\mathbf{x}}_i)^T \left(\mathbf{S}_a^{-1} + \mathbf{K}_i^T\mathbf{S}_o^{-1}\mathbf{K}_i\right)(\hat{\mathbf{x}}_{i+1} - \hat{\mathbf{x}}_i) < \Delta\mathbf{x}_{\text{stop}} \cdot n \quad (4.2)$$

where n is the length of the vector $\mathbf{x}$ and $\Delta\mathbf{x}_{\text{stop}}$ is the variable controlling the wanted strictness. The value of $\Delta\mathbf{x}_{\text{stop}}$ =? varies between observation modes, see ?. Where?

If Eq. 4.2 becomes fulfilled, it is considered that basic convergence has been achieved. There are two exceptions, resulting in non-convergence status, that either the iteration number limit ($i_{\max}$) has been reached or that γ has reached its upper threshold ($\gamma_{\max}$).

Normalised cost values

For these Odin/SMR retrievals, the length of $\mathbf{y}$, m, is always much higher than n (length of $\mathbf{x}$). In this situation, the final cost value, defined according to Eq. 2.18, should be approximately m at correct convergence. As m can vary from retrieval to retrieval, even for a single observation mode, a normalised cost, C', is defined to make it simpler to compare cost values from different retrievals. The normalised cost is defined as:

$$C' = C'_y + C'_x, \tag{4.3}$$

$$C'_y = \frac{(\mathbf{y} - \mathcal{F}(\mathbf{x}, \hat{\mathbf{b}}))^T \mathbf{S}_o^{-1} (\mathbf{y} - \mathcal{F}(\mathbf{x}, \hat{\mathbf{b}}))}{m}, \tag{4.4}$$

$$C'_x = \frac{(\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a)}{m}. \tag{4.5}$$

$$\tag{4.6}$$

The L2 data contain C', as well as C'_y to provide information on how the final cost is distributed between the two penalty terms.

Correct convergence

Eq. 4.2 evaluates only if the iteration has reached a stable end point. If the found solution actually corresponds to a global minimum of the cost function is another question, and a problematic one. To ensure a globally optimal solution a more advanced approach than LM is required, but all such methods are computationally very demanding. A simplistic approach is to restart the iteration with other start conditions, where $\hat{\mathbf{x}}_0$ and/or $\gamma = \gamma_0$ are/is given other initial value. Such tests have been performed on a number of cases for each observation mode, to validate the iteration settings, but are not performed on a regular basis.

The final cost value can be used for rough testing of incorrect convergence, but it is hard to give a general critical value. Theoretically, it is expected that $C' \approx 1$, but this is only true on the condition that $\mathbf{S}_o$ gives a correct estimate of the observation uncertainties. For simplicity, we only consider thermal noise, that is $\mathbf{S}_0 = \mathbf{S}_{\varepsilon_n}$. If then the magnitude of thermal noise is over-estimated in $\mathbf{S}_0$, on purpose or by mistake, obtained values of C' will on average be below 1 (assuming good convergence). However, for practical retrievals, C' tend instead to be > 1 due to artefacts in the spectra, simplifications in the retrieval set-up and other reasons resulting in that the residual:

$$\delta\mathbf{y} = \mathbf{y} - \mathcal{F}(\hat{\mathbf{x}}, \hat{\mathbf{b}}), \tag{4.7}$$

contains features besides thermal noise. The so called baseline ripple is a common cause for an excess size of the residual.

For these reasons, a suitable upper limit on C'_y varies between the observation modes. Suggested limits for quality filtering for each mode are found in ?.

Where?

A high final cost can be caused by both bad input data and that the iteration procedure ended up at a local minimum. Without a detailed analysis it is impossible to say which of the two potential problems that is the cause behind a suspiciously high end value of C'.

Some notes on matrix operations

As mentioned above, $m > n$ for these Odin/SMR retrievals. In the terminology of Rodgers (2000), the n-form of equations should then be used for lowest calculation burden, and this guideline is strictly followed.

The matrix algebra is throughout performed by Matlab's internal functions. Explicitly calculating matrix inverses is avoided as far as possible. For example, Eq. 4.1 is evaluated, schematically, as

$$\hat{\mathbf{x}}_c = \hat{\mathbf{x}}_i + \left[(1+\gamma)\mathbf{S}_a^{-1} + \mathbf{K}_i^T\mathbf{S}_o^{-1}\mathbf{K}_i\right] \setminus \left[\mathbf{K}_i^T\mathbf{S}_o^{-1}(\mathbf{y} - \mathcal{F}(\hat{\mathbf{x}}_i, \hat{\mathbf{b}})) - \mathbf{S}_a^{-1}(\hat{\mathbf{x}}_i - \mathbf{x}_a)\right]$$

where $\setminus$ represents Matlab's left matrix divide operator (for details, see Matlab documentation). By using this operator, the inverse of $\left[(1+\gamma)\mathbf{S}_a^{-1} + \mathbf{K}_i^T\mathbf{S}_o^{-1}\mathbf{K}_i\right]$ is not calculated, instead $\hat{\mathbf{x}}_c - \hat{\mathbf{x}}_i$ is found by solving an equation system (by QR decomposition). This option is more numerically efficient, than calculating the actual inverse.

However, a direct calculation of the inverse of $\mathbf{S}_a$ and $\mathbf{S}_o$ is made. These calculations take advantage of the block-diagonal structure (see Sec. 5.6), where each sub-matrix is inverted separately. Further, in some situations the matrix inverse can be obtained in an analytic manner. This is possible for purely diagonal matrices and with exponantially decling correlation[1].

The size of the elements in $\mathbf{x}$ can differ widely. For example, VMR values go down to at least ppb level (10^{-9}), while frequency off-sets can reach 1 MHz. Variances follow the magnitude of elements squared and the span of values in $\mathbf{S}_a$ is accordingly even higher. Without any action, the matrix algebra will be pushed towards, or beyond, the numerical accuracy. To avoid numerical problems (at least due to the reason considered), an internal re-scaling of $\mathbf{x}$ is applied. The re-scaling is done in such way that the variances of $\mathbf{S}_a$, matching the transformed $\mathbf{x}$, all are unity. That is, the elements in $\mathbf{x}$ are scaled with the square root of the corresponding variance in $\mathbf{S}_a$.

[1] Analytical matrix inversion in the case of exponantially decling correlation is discussed in Rodgers (2000, Sec. 10.3.2.2). In contrast to the expression foind in Rodgers (2000), our implementation allows a non-constant standard deviation, but still limited to conmstant correlation length and grid spacing.

Retrieval variables

This chapter reviews the variables that are part of the retrievals, either directly or indirectly. The presentation follows the formalism presented in Sec. 2.3.

Basic variable choices

Odin/SMR performs measurements in several frequency bands simultaneously (Sec. 1.2.2). The set of frequency bands differs between observation modes (Rydberg et al., 2016). The altitude range scanned by Odin/SMR is not fixed, and a Odin/SMR limb "scan" is simply defined as the spectra recorded between two adjacent turning points, in altitude, of the limb scanning (Rydberg et al., 2016). The retrievals covered by this ATBD operate only on one scan, from a single frequency band, at the time. Each combination of scan and frequency band is treated separately, no information is carried between different retrievals.

Following the discussion in Sec. 2.4.1.3, the state vector does not only contain the quantities targeted directly by the Odin/SMR measurements, that is, gas species concentrations and atmospheric temperatures, but also a number of other variables are part of $\mathbf{x}$ (Sec. 5.4). These variables are either instrument correction factors or describe some interfering atmospheric effect. However, a restriction to effects of random nature is made, no fixed forward model variables (such as spectroscopic data) are placed in $\mathbf{x}$. That is, variables causing systematic retrieval errors are not included in $\mathbf{x}$ (and neither in $\mathbf{S}_o$, Sec. 5.3).

Covariance matrices are implemented as sparse matrices, while all other variables are stored as full vectors or matrices. For the retrieval problem at hand, $\mathbf{K_x}$ is relatively dense and the overhead related with a sparse representation results in better performance keeping $\mathbf{K_x}$ as a full matrix.

The measurement vector, y

The measurement vector contains solely Odin/SMR brightness temperature values, with spectra from different tangent altitudes appended. That is,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_s \end{bmatrix}, \tag{5.1}$$

where $\mathbf{y}_1/\mathbf{y}_s$ is the spectrum from the highest/lowest tangent altitude etc.

Each retrieval only contains data from a single scan, as mentioned, but all spectra of the scan are not included. For each mode an upper and lower tangent altitude limit are defined, that determine which spectra of a scan that will be part of $\mathbf{y}$. The altitude limits are defined using geometrical tangent altitudes (that is, refraction neglected) as those are the tangent altitudes found in the L1b data. A minimum scan range is defined in the same manner; no retrieval is made if a limb scan does not cover this minimum range. The tangent altitude masks for each frequency band are found ?.

Where?

Observation uncertainty covariance matrix, $\mathbf{S}_o$

All variables of random nature that have a significant impact on the measurements, including indirect effects through the forward model, are part of the state vector (Secs. 5.1 and 5.4). This means that the term $\mathbf{K_b S_b K_b}^T$ in Eq. 2.15 should give a relatively small contribution (ignoring systematic components) to $\mathbf{S}_{\varepsilon_o}$. Accordingly, $\mathbf{S}_o$ is set to only represent thermal noise:

$$\mathbf{S}_o = \mathbf{S}_{\varepsilon_n}. \tag{5.2}$$

The standard deviations to include in $\mathbf{S}_{\varepsilon_n}$ are set according to Rydberg et al. (2016). The correlation of noise between adjecent channels of each spectrum is considered. The magnitude of this correlation depends on if a Hanning window has been applied or not. No noise correlation between spectra from different altitudes is included.

The state vector, x

The state vector contains a mix of quantities (Sec. 5.1). Besides the 118 GHz band, the quantities always present are: gas species profiles ($\mathbf{v}$), the temperature profile ($\mathbf{t}$), a set of brightness temperature off-sets ($\Delta\mathbf{o}$), a pointing off-set ($\Delta\theta$) and a frequency off-set ($\Delta\nu$). With just these quantities retrieved, and a case involving two gas species, the state vector becomes:

$$\mathbf{x} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{t} \\ \mathbf{\Delta o} \\ \Delta\theta \\ \Delta\nu \end{bmatrix}. \tag{5.3}$$

For 118 GHz the state vector contains no species profiles as only oxygen, having a known VMR, has to be considered for this frequency band.

A species or temperature "profile" is a set of values describing the vertical variation, at a set of specified pressures, that is, the retrieval grid (see also Sec. 3.3). Such retrieval grids are discussed further in Sec. 5.7.1, together with the basic pressure grid of the forward model. Atmospheric temperatures are in Kelvin. Gas species are reported as the VMR in the L2 data, but inside the retrieval process the unit applied for gas species is ....

Do we go for log-retrievals for species? Discuss basic aspects of the log conversion here.

As discussed in Sec. 1.2.3, there is not a perfect compensation for gain variations, causing an off-set in brightness temperatures. This off-set has been found to be flat over the frequency bands, but differ from one tangent altitude to next. The vector $\mathbf{\Delta o}$ contains one brightness temperature off-set for each included spectrum, to compensate for the "baseline shift". That is, a correction for the brightness temperatures is estimated, and corrected spectra can be expressed as

$$\hat{T}_b^{ij} = T_b^{ij} + \Delta\mathbf{o}_i \tag{5.4}$$

where T_b^{ij} is brightness temperature of channel j and tangent altitude i and $\Delta\mathbf{o}_i$ is element i of $\Delta\mathbf{o}$.

The pointing and frequency off-sets are scalar values. This implies that these correction factors are valid for the complete scan. In ARTS the viewing direction, of each LOS, is specified as the angle from zenith. These angles cannot be determined perfectly from attitude data, but the error has been judged to be constant during a scan and a single correction value should be sufficient. Hence, the retrieved zenith angles are

$$\hat{\theta}_i = \theta_i + \Delta\theta, \tag{5.5}$$

where θ_i is the L1b zenith angle of tangent altitude i. A positive $\Delta\theta$ means that the retrieval estimated lower tangent altitudes than obtained by the attitude data. As a rule of thumb, $\Delta\theta = \pm 0.01°$ matches about ∓ 550 m in tangent altitude.

There is a similar uncertainty in the exact frequencies observed. This uncertainty originates in the LO-signal. Already pre-launch tests showed that LO-frequencies vary somewhat with temperature. A correction table was created, but some frequency error remains and there seems also to be some ageing on the LO-chain affecting final frequencies. The resulting frequency uncertainty is totally correlated between the backend channels inside each frequency band. As long as the LO-signal is phase-locked, it is judged that the frequency off-set is constant during a limb scan. However, the phase-locking of the 575 GHz front-end has been failing basically throughout the mission, and the assumption of a constant frequency off-set is more uncertain for this receiver chain. The switch from version 1 to 2 of ARTS enables the retrieval of a non-constant frequency off-set, but this possibility will only be activated later if found necessary. The retrieved frequency of channel i is

$$\hat{\nu_j} = \nu_j + \Delta\nu, \tag{5.6}$$

where ν_j is the frequency found given by L1b data.

The plan is to activate winds for frequency bands containing strong mesospheric lines, but no tests have yet performed. Describe the wind retrieval here, if activated. Otherwise revise text, we have then always the same set of variables in $\mathbf{x}$?

The a priori vector, $\mathbf{x}_a$

Write later. For species we need to decide if we go for the old a priori database, originally from Bordeaux. Or can we get a better, and much newer, alternative? The MIPAS climatology, or from any of all ongoing projects?

For temperature we need to decide if ECMWF-operational or ERA-Interim shall be used as source. And how exactly the transition to MSIS is made.

All sensor correction terms are assumed to have an average of zero. Accordingly, the a priori value of $\Delta\mathbf{o}$, $\Delta\theta$ and $\Delta\nu$ is throughout set to zero.

A priori uncertainty covariance matrix, $\mathbf{S}_a$

The setting of $\mathbf{S}_a$ is the most controversial part of OEM, see Sec. 2.4.1.3 for some comments. This subject is not elaborated further, it is just noticed that for VMR and atmospheric temperature, in general, there exist no direct measurements on which total covariance matrices can be based and ad hoc values are applied to a large extent. In addition, the values selected results in that $\mathbf{S}_a$ in general should over-estimate variability. Accordingly, OEM is not used in a strict sense, but the retrievals can still be classified is regularisation (Ungermann, 2011) and close to optimal result should be obtained as long as $\mathbf{S}_a$ is given "reasonable" values (Eriksson, 2000). On the other hand, the uncertainties in the instrumental variables are known fairly well and the corresponding parts of $\mathbf{S}_a$ should be quite accurate.

Species: Details here depends on retrieval unit selected, and text must be written latter. But likely, we use a ∼50% variability, with a constrain of a minimum variability defined in VMR (converted to relative values).

Temperature: For altitudes up to about 45 km, the standard deviation should be ≈2 K. For higher altitudes, use fixed values, or does CIRA or MSIS contain tables on natural variability that can be used?

The atmosphere exhibits vertical correlation, see Eriksson and Chen (2002) for examples on ozone and temperature correlations. In lack of direct measurements covering all altitudes, vertical correlations are modelled by parametric expressions, using a correlation length l_c. The correlation is modelled to follow a Gaussian function,

$$\rho(z_1, z_2) = exp\left(-4[(z_1 - z_2)/(l_c(z_1) + l_c(z_2))]^2\right), \tag{5.7}$$

or an exponential one,

$$\rho(z_1, z_2) = exp\left(-2\,|z_1 - z_2|\,/(l_c(z_1) + l_c(z_2))\right), \tag{5.8}$$

where $\rho(z_1, z_2)$ is the correlation coefficient between altitudes z_1 and z_2. Note that the mean of the correlation length at z_1 and z_2 is used, that is, $l_c(z_1)/2 + l_c(z_2)/2$.

Where do we report l_c used?

The frequency and pointing off-sets are scalar values, and only a standard deviation (σ) for each quantity must be specified. These standard deviations are set to $\sigma_{\Delta\nu}$ =? MHz and $\sigma_{\Delta\theta}$ =?°, respectively. The brightness temperature off-sets have been found to be uncorrelated between spectra, and the variation to have no variation with altitude. This results in that the covariance matrix for $\mathbf{\Delta o}$ can be written as $\sigma^2_{\mathbf{\Delta o}}\mathbf{1}$. The standard deviation of brightness temperature off-sets is throughout set to ? K.

Set value for all σ.

Any correlations between the different quantities in $\mathbf{x}$ are fully ignored. That is, the

total covariance matrix becomes:

$$\mathbf{S}_a = \begin{bmatrix} \mathbf{S}_{\mathbf{v}_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{S}_{\mathbf{v}_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{S}_{\mathbf{t}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_{\mathbf{\Delta o}}\mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_{\Delta\theta} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2_{\Delta\nu} \end{bmatrix}. \tag{5.9}$$

The inverse of a block-diagonal matrix of this type is also a block-diagonal matrix, consisting of the inverse of the individual matrices, and the inverse of $\mathbf{S}_a$ is calculated in the following manner:

$$\mathbf{S}_a^{-1} = \begin{bmatrix} \mathbf{S}_{\mathbf{v}_1}^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{S}_{\mathbf{v}_2}^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{S}_{\mathbf{t}}^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^{-2}_{\mathbf{\Delta o}}\mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^{-2}_{\Delta\theta} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^{-2}_{\Delta\nu} \end{bmatrix}. \tag{5.10}$$

Forward model parameters, b

Please note that static forward model settings, which are common for all simulations, are presented already in Sec. 3.2.

A first version of everything below exists, but all details should be revised/optimised further. Write each section below, when later looking at each part.

Pressure grids

Describe how pressure grids are set, including retrieval ones. We operate with a fixed set of pressure where integer pressure decades always are included (e.g. 100 and 1 hPa). The spacing is given as number of points per pressure decade, as also done for MLS. Retrieval grids always contain a sub-set of the forward model pressure grid.

Monochromatic frequency grids

Describe the role of these grids, and how they are handled inside ARTS, resulting in a need to have a more dense spacing around transitions. Make an illustrative figure. Describe algorithm to set these grids. Test how the calculation time scales with the number of frequencies. Is it close to linear? If yes, then we should try to improve the existing algorithm to set the grids. The most obvious choice if to use the Ffill option, but that requires careful testing to check that there are no caveats.

Spectroscopic data and absorption continua

The bulk of the data are taken from latest HITRAN. A literature review will be performed to hand-pick values for our most important transitions. The hand-picked data will be put in tables, that can be used to replace the corresponding value in HITRAN by ARTS. The nitrogen and water vapour continua should be set as done for the tropospheric retrievals. For nitrogen this means the expression from MPM93 multiplied with 1.34 following some references in the literature. Details given later.

Absorption look-up tables

Describe how absorption tables are generated, basically just what reference atmosphere that is used, and what pressure and temperature grids that are used. And also state what interpolation orders that are applied.

Radiative transfer

The propagation path used when evaluating Eq. 2.1 is defined by a set of points. A point is added at each crossing of the path with a pressure level (that is, the pressures defining the vertical grid). Additional points are inserted to ensure that the distance along the path does not exceed an user defined threshold. This threshold is denoted $l_{\max}$ in the tables of ?.

Where?

Retrieval characterisation

A full retrieval characterisation is only given for a set of representative cases. This despite that such a characterisation anyhow is largely performed internally for each retrieved scan, but to store the full averaging kernel and error covariance matrices would increase the size of the L2 data files with about two orders of magnitude and such detailed information is anyhow normally not required. Instead, only some all-embracing quantities are included in the L2 data, and full characterisations are provided separately. The background theory of this chapter is found in Sec. 2.3.

Interpretation of errors

The full retrieval error is given by Eq. 2.13, where the total retrieval error is presented as the sum of three terms. The calculation and interpretation of the forward model parameter and the thermal noise retrieval errors are quite straightforward, while the smoothing error term, $(\mathbf{A}-\mathbf{1})\,\mathbf{S_x}\,(\mathbf{A}-\mathbf{1})^T$, is more problematic and frequently causes confusion. The aim of this section is to clarify some aspects of the smoothing error and how this error term is handled in the Odin/SMR L2 data. The averaging kernel matrix has a central role for this discussion and it is defined by Eq. 2.12.

The case of a single species profile

Let us start with a simplified example; that the state vector just contains the vertical profile of a single species. For such a retrieval case, a row of $\mathbf{A}$ shows how changes throughout the true profile contribute to the retrieved value at the altitude of concern. Accordingly the term "averaging kernels" refers to the rows of $\mathbf{A}$. The columns $\mathbf{A}$ give the impact of a delta-type perturbance in $\mathbf{x}$ (Rodgers, 2000).

The characterisation makes no attempt to estimate this smoothing error (the one internal for a retrieval quantity), for the simple reason that a sufficient accurate $\mathbf{S_x}$ cannot be obtained. To replace $\mathbf{S_x}$ with the corresponding matrix used as part of OEM, $\mathbf{S}_a$, is not a viable option as it would give misleading results. The smoothing error should in general be over-estimated, considering how $\mathbf{S_x}$ is constructed (Sec. 5.6). A smoothing error is present when $\mathbf{A} \neq \mathbf{1}$. The averaging kernel matrix deviates from the identity matrix due to the impact regularisation, that roughly can be separated into two main aspects: limited resolution (see Sec. 6.3.2) and non-perfect measurement response (see Sec. 6.3.1).

If the smoothing error is not included in the retrieval error, how should then the two remaining error terms be interpreted? Some guidance to understand this question can be obtained by a rearrangement of Eq. 2.11:

$$\hat{\mathbf{x}} - \mathbf{x}_a - \mathbf{A}\,(\mathbf{x}-\mathbf{x}_a) = \mathbf{G}\mathbf{K_b}\left(\mathbf{b}-\hat{\mathbf{b}}\right) + \mathbf{G}\varepsilon_n. \tag{6.1}$$

If $\mathbf{A}\mathbf{x}_a \approx \mathbf{x}_a$ we get

$$\hat{\mathbf{x}} - \mathbf{A}\mathbf{x} \approx \mathbf{G}\mathbf{K}_\mathbf{b}\left(\mathbf{b} - \hat{\mathbf{b}}\right) + \mathbf{G}\varepsilon_n. \tag{6.2}$$

That is, the forward model parameter and retrieval noise errors describe the difference between the retrieved state and a smoothed version of the true state ($\mathbf{A}\mathbf{x}$). However, as $\mathbf{x}$ is unknown, and then also $\mathbf{A}\mathbf{x}$, a somewhat different view must be used in practice. Retrieved profiles are running averages of the true profiles, and errors refer to the accuracy of these averages. The vertical resolution gives the approximate averaging length. The condition $\mathbf{A}\mathbf{x}_a \approx \mathbf{x}_a$ is roughly fulfilled as long as the measurement response is around 1.

For situations where the measurement provides no information at all, a zero measurement response, the values in both $\mathbf{G}$ and $\mathbf{A}$ become zero. This results in $\hat{\mathbf{x}} - \mathbf{x}_a = 0$. That is, the solution equals the a priori state. A pitfall emerges here, the errors reported also become zero, but this only reflects that neither thermal noise or forward model parameters have had any impact on the solution. The total retrieval error equals the neglected smoothing error.

Multiple retrieval quantities

As a more general example, if the state vector follows Eq. 5.3, the averaging kernel matrix has the structure:

$$\mathbf{A} = \begin{bmatrix}
\mathbf{A}_{\mathbf{v}_1} & \mathbf{A}_{\mathbf{v}_1,\mathbf{v}_2} & \mathbf{A}_{\mathbf{v}_1,\mathbf{t}} & \mathbf{A}_{\mathbf{v}_1,\boldsymbol{\Delta}\mathbf{o}} & \mathbf{A}_{\mathbf{v}_1,\Delta\theta} & \mathbf{A}_{\mathbf{v}_1,\Delta\nu} \\
\mathbf{A}_{\mathbf{v}_2,\mathbf{v}_1} & \mathbf{A}_{\mathbf{v}_2} & \mathbf{A}_{\mathbf{v}_2,\mathbf{t}} & \mathbf{A}_{\mathbf{v}_2,\boldsymbol{\Delta}\mathbf{o}} & \mathbf{A}_{\mathbf{v}_2,\Delta\theta} & \mathbf{A}_{\mathbf{v}_2,\Delta\nu} \\
\mathbf{A}_{\mathbf{t},\mathbf{v}_1} & \mathbf{A}_{\mathbf{t},\mathbf{v}_2} & \mathbf{A}_{\mathbf{t}} & \mathbf{A}_{\mathbf{t},\boldsymbol{\Delta}\mathbf{o}} & \mathbf{A}_{\mathbf{t},\Delta\theta} & \mathbf{A}_{\mathbf{t},\Delta\nu} \\
\mathbf{A}_{\boldsymbol{\Delta}\mathbf{o},\mathbf{v}_1} & \mathbf{A}_{\boldsymbol{\Delta}\mathbf{o},\mathbf{v}_2} & \mathbf{A}_{\boldsymbol{\Delta}\mathbf{o},\mathbf{t}} & \mathbf{A}_{\boldsymbol{\Delta}\mathbf{o}} & \mathbf{A}_{\boldsymbol{\Delta}\mathbf{o},\Delta\theta} & \mathbf{A}_{\boldsymbol{\Delta}\mathbf{o},\Delta\nu} \\
\mathbf{A}_{\Delta\theta,\mathbf{v}_1} & \mathbf{A}_{\Delta\theta,\mathbf{v}_2} & \mathbf{A}_{\Delta\theta,\mathbf{t}} & \mathbf{A}_{\Delta\theta,\boldsymbol{\Delta}\mathbf{o}} & \mathbf{A}_{\Delta\theta} & \mathbf{A}_{\Delta\theta,\Delta\nu} \\
\mathbf{A}_{\Delta\nu,\mathbf{v}_1} & \mathbf{A}_{\Delta\nu,\mathbf{v}_2} & \mathbf{A}_{\Delta\nu,\mathbf{t}} & \mathbf{A}_{\Delta\nu,\boldsymbol{\Delta}\mathbf{o}} & \mathbf{A}_{\Delta\nu,\Delta\theta} & \mathbf{A}_{\Delta\nu}
\end{bmatrix}. \tag{6.3}$$

The sub-matrices along the diagonal, such as $\mathbf{A}_{\mathbf{v}_2}$, can be interpreted exactly as in the simplified example above. In the case of scalar quantities, such as $\mathbf{A}_{\Delta\nu}$, these sub-matrices have size 1 x 1.

The off-diagonal sub-matrices show how the different quantities interfere with each other (Baron et al., 2002). For example, one row of the sub-matrix $\mathbf{A}_{\mathbf{v}_1,\mathbf{t}}$ gives how the corresponding value in gas species profile 1 is affected by variations in the temperature profile, while a row in $\mathbf{A}_{\mathbf{t},\mathbf{v}_1}$ gives how one temperature profile value is influenced by changes in gas species 1.

The retrieval will in general not be able to make a perfect separation between changes in the different retrieval quantities, for example, variations in the atmospheric temperature profile will cause errors in a ozone profile retrieval, even if the temperature profile is part of $\mathbf{x}$. This influence across the retrieval quantities causes an additional smoothing error. To exemplify this, let us continue to use a combined ozone and temperature retrieval:

$$\mathbf{A} = \begin{bmatrix} \mathbf{G}_\mathbf{v} \\ \mathbf{G}_\mathbf{t} \end{bmatrix} \begin{bmatrix} \mathbf{K}_\mathbf{v} & \mathbf{K}_\mathbf{t} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_\mathbf{v} & \mathbf{K}_\mathbf{t} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_\mathbf{v}\mathbf{K}_\mathbf{v} & \mathbf{G}_\mathbf{v}\mathbf{K}_\mathbf{t} \\ \mathbf{G}_\mathbf{t}\mathbf{K}_\mathbf{v} & \mathbf{G}_\mathbf{t}\mathbf{K}_\mathbf{t} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_\mathbf{v} & \mathbf{A}_{\mathbf{v},\mathbf{t}} \\ \mathbf{A}_{\mathbf{t},\mathbf{v}} & \mathbf{A}_\mathbf{t} \end{bmatrix}$$

where the ozone profile is denoted as $\mathbf{v}$. The smoothing error covariance matrix for this

example becomes

$$(\mathbf{A}-\mathbf{1})\begin{bmatrix}\mathbf{S_v} & 0\\ 0 & \mathbf{S_t}\end{bmatrix}(\mathbf{A}-\mathbf{1})^T = \begin{bmatrix}(\mathbf{A_v}-\mathbf{1})\mathbf{S_v}(\mathbf{A_v}-\mathbf{1})^T+\mathbf{A_{v,t}}\mathbf{S_t}\mathbf{A_{v,t}}^T & \dots\\ \dots & (\mathbf{A_t}-\mathbf{1})\mathbf{S_t}(\mathbf{A_t}-\mathbf{1})^T+\mathbf{A_{t,v}}\mathbf{S_v}\mathbf{A_{t,v}}^T\end{bmatrix}$$

where ... indicates matrix elements left out, for brevity reasons. If we focus on the upper left sub-matrix, covering the ozone profile smoothing error, the term $(\mathbf{A_v}-\mathbf{1})\mathbf{S_v}(\mathbf{A_v}-\mathbf{1})^T$ covers the smoothing error that is internal for the ozone profile, while the second term, $\mathbf{A_{v,t}}\mathbf{S_t}\mathbf{A_{v,t}}^T$ is the smoothing error induced by atmospheric temperatures in the retrieved ozone profile. The last term is new compared to an ozone-only retrieval.

Not surprisingly, this new term is identical to the expression to be used if temperature had been a forward model parameter. If $\mathbf{A_{v,t}}$ is expanded, we get that the ozone retrieval error due to temperature interference is (Rodgers and Connor, 2003)

$$\delta_{\mathbf{v}}^{\mathbf{t}} = \mathbf{G_v}\mathbf{K_t}\mathbf{S_t}\mathbf{K_t}^T\mathbf{G_v}^T.$$

This expression gives exactly the same result as treating temperature as a forward model uncertainty and following Eq. 2.15 for the error calculation. This is consistent with the discussion in Sec. 2.4.1.3; that for a linear case exactly the same result is obtained when moving an interfering quantity from $\mathbf{b}$ to $\mathbf{x}$. This implies that the total error remains constant, but the corresponding error term is at the same time moved from being a forward model parameter error to being a smoothing error.

(It should be noted that the last paragraphs make the silent assumption that there is no correlation between the retrieval quantities; off-diagonal sub-matrices of $\mathbf{S_x}$ are zero. This is throughout assumed for the SMR processing and the general case is not discussed.)

The interference between retrieval quantities is included in the error provided. The calculation of these errors also requires $\mathbf{S_x}$, but the accuracy of this matrix is here less important, compared to the estimation of the internal smoothing error. The largest uncertainties of $\mathbf{S_x}$ are found for gas species and temperature, but interference between these quantities is limited and, hence, it does not give rise to any dominating retrieval error. The interference of instrumental quantities on VMR retrievals can be substantial, but $\mathbf{S_x}$ can be set quite accurately for the parts corresponding to instrumental uncertainties.

In summary, the smoothing error internal to each retrieval quantity is not part of the error given. Accordingly, the error refers to the accuracy of the retrieved profile considered as a running average of the true profile. The interference between different retrieval quantities is included. The errors are valid on the constraint that the measurement response is relatively high. If the measurement response is low, there is a systematic bias towards the a priori state, that is not captured by the errors provided, but data with low measurement response shall anyhow not be used for a scientific analysis without special treatment (see Rodgers and Connor (2003) for general guidelines).

Full characterisation

How do we distribute the full characterisations? Take the text below just as some brainstorming around how we shall do this.

Approach and reference cases

The full characterisation does not involve any measurement data. Instead a linear analysis is performed for a set of atmospheres selected from the gas species a climatology. Temperature data are here taken solely from ?. A standard limb scan sequence is assumed. The integration times ?. A retrieval is set-up according to the time and position of the hypothetical measurement, and the characterisation is made for $\mathbf{K_x}$ and $\mathbf{G}$ following Eq. 2.20. That is, a linear analysis around a priori is performed.

MSIS?

One, several, or mix?

How do we define the database of full characteristics? Combinations of months (middle of every second?) and latitudes (-80 to 80 in steps of 20°?). The database should cover all frequency bands. There should also be a discussion on what database case should be selected to match a particular measurement. Should we recommend scaling following the measurement response, where possible?

Averaging kernel matrix, A

The full characterisation reports the complete averaging kernel matrix, together with the positions inside the state vector each retrieval quantity occupies. Each sub-matrix of $\mathbf{A}$ can be extracted if these positions are known. The a priori data for which the averaging kernel matrix are also stored ...

Exemplify how these data can be used.

Error covariance matrices

According to the discussion above, the random retrieval error is defined as the total retrieval error (Eq. 2.13) with the internal smoothing error subtracted. The covariance matrix of this error is schematically calculated as

$$\mathbf{S}_{\delta_r} = (\mathbf{A}-\mathbf{1})\,\mathbf{S_x}\,(\mathbf{A}-\mathbf{1})^T - \begin{bmatrix} \mathbf{S}^1_{\text{is}} & 0 & \dots & 0 \\ 0 & \mathbf{S}^2_{\text{is}} & \dots & 0 \\ 0 & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{S}^q_{\text{is}} \end{bmatrix} + \mathbf{G}\mathbf{S}_{\varepsilon_n}\mathbf{G}^T, \tag{6.4}$$

where q is the number of retrieval quantities and $\mathbf{S}^i_{\text{is}}$ is the internal smoothing error of quantity i, that using the nomenclature of Sec. 6.1 is

$$\mathbf{S}^i_{\text{is}} = (\mathbf{A_i}-\mathbf{1})\mathbf{S_i}(\mathbf{A_i}-\mathbf{1})^T.$$

The $\mathbf{S}_{\delta_r}$ covariance matrix can be divided into sub-matrices exactly as the averaging kernel matrix, where off-diagonal sub-matrices give information on the error correlation between the different retrieval quantities. As any covariance matrix, $\mathbf{S}_{\delta_r}$ is symmetric (which is not the case for $\mathbf{A}$!), and corresponding sub-matrices above and below the diagonal provide redundant information.

For completeness, the covariance matrix of the retrieval error due to thermal noise is also given

$$\mathbf{S}_{\delta_n} = \mathbf{G}\mathbf{S}_{\varepsilon_n}\mathbf{G}^T. \tag{6.5}$$

Do we make use of the symmetry and only store the upper part of the matrices?

Random errors

These errors show how the different sources contribute to the total random error. That is, the error due to thermal noise and all interfering smoothing errors are reported individually. Only standard deviations (σ) are given. The total error $\sigma_{\rm tot}$, is related to the individual error components as

$$\sigma_{\rm tot}^2 = \sigma_{\rm n}^2 + \sum_{i=1}^{q-1} \sigma_i^2 \tag{6.6}$$

where $\sigma_{\rm n}$ is the error due to thermal noise, and σ_i is an interference error. The matching value along the diagonal of $\mathbf{S}_{\delta_o}$ equals $\sigma_{\rm tot}^2$.

Systematic errors

Write this section later as we hopefully can make use of some ongoing ARTS development here. The most important systematic errors should be due to spectroscopic data, such as pressure broadening parameters. These errors are normally determined by perturbing the spectroscopic data, but this is slow and requires a rather intricate set-up. ARTS is right now extended to provide weighting functions for spectroscopic variables, which makes this process both faster and simpler.
Any other systematic error source to consider?

L2 data fields

The L2 data contain only some key characteristics, but these are calculated specifically for each retrieval.

Polish section when L2 format set.

How is **G** set? By pure linear expression, or can we get an effective **G** out of LM?

When things are clearer, consider to add the diagonal sub-matrices of averaging kernel and error covariance matrices. That should be very useful, but would use much more disk space.

Measurement response

The averaging kernel matrix is in the L2 data summarised by three measures, where the measurement response is one. The measurement response is defined as the row sum of the averaging kernel matrix (Baron et al., 2002), but following the treatment of smoothing errors, the summing is here only made inside each sub-matrix along the diagonal. That is, only the measurement response inside each retrieval quantity is considered.

Vertical placement and resolution

Also these measures is based on the diagonal sub-matrices, and not the complete **A**. The vertical resolution is reported as the full width at half maximum (FWHM) of the averaging kernels. An averaging kernel is not necessarily centred around the nominal

retrieval altitude. The vertical placement of the averaging kernel is given as the mean of altitude:

$$z_m = \frac{\int z a(z_0, z)\,\mathrm{d}z}{\int a(z)\,\mathrm{d}z} \tag{6.7}$$

where z is altitude and $a(z)$ is the averaging kernel for the nominal altitude of z_0.

Or give $z_m - z_0$?

FWHM is a quite optimistic value, especially for averaging kernels of the shape given by OEM. Also include spread?

Random retrieval error

The total random error, σ_{tot}, is given following Eq. 6.6.

Summary

Keep text short. Mainly make a "FAQ", as sketched below.

...

Some key issues that should be considered when examining or applying Odin/SMR L2 data:

- The retrievals assume a horizontally homogeneous (1D), non-scattering, atmosphere.
- The main atmospheric vertical coordinate is pressure, i.e. VMR and temperatures are retrieved as a function of pressure. The L2 data also contain an estimate of the geometrical altitudes matching the values of the profile, but this information shall be considered as secondary.

  Correct?
- Retrieved profiles shall be considered as running average estimates of the true profile, with an averaging length roughly given by the vertical resolution.
- Errors refer to the accuracy of running averages, and are only valid under the constraint of high measurement response. Using data with low measurement response requires consideration of smoothing effects.
- The retrievals are best estimates at each pressure level. To obtain values at other pressure levels, apply a linear interpolation using the logarithm of pressure as independent variable.

  Does L2 contain anything else than profiles?
- A single geographical position is assigned to the L2 data, that matches ...

  Geo-position set how?

  However, the data are far from point estimates in the horizontal direction, they represent averages of several 100 km along the viewing direction of Odin/SMR, and not even the centre position of the averages is necessarily found at the stated position.

Bibliography

P. Baron, P. Ricaud, J. de La Noë, P. Eriksson, F. Merino, and D. Murtagh. Studies for the Odin sub-millimetre radiometer: 2. Retrieval methodology. *Can. J. Phys.*, 80: 341–356, 2002.

S. A. Buehler, P. Eriksson, T. Kuhn, A. von Engeln, and C. Verdes. ARTS, the Atmospheric Radiative Transfer Simulator. *J. Quant. Spectrosc. Radiat. Transfer*, 91:65–93, 2005. doi: 10.1016/j.jqsrt.2004.05.051.

S. A. Buehler, N. Courcoux, and V. O. John. Radiative transfer calculations for a passive microwave satellite sensor: Comparing a fast model and a line-by-line model. *J. Geophys. Res.*, 111:D20304, 2006. doi: 10.1029/2005JD006552.

S. A. Buehler, P. Eriksson, and O. Lemke. Absorption lookup tables in the radiative transfer model arts. *J. Quant. Spectrosc. Radiat. Transfer*, 112(10):1559–1567, 2011. doi: 10.1016/j.jqsrt.2011.03.008.

S. Chandrasekhar. *Radiative Transfer*. Dover Publications, New York, 1960. ISBN 0-486-60590-6.

O. M. Christensen, P. Eriksson, J. Urban, D. Murtagh, K. Hultgren, and J. Gumbel. Tomographic retrieval of water vapour and temperature around polar mesospheric clouds using odin-smr. *Atmos. Meas. Tech.*, 8(5):1981–1999, 2015. doi: 10.5194/amt-8-1981-2015.

C. Davis, C. Emde, and R. Harwood. A 3D polarized reversed monte carlo radiative transfer model for mm and sub-mm passive remote sensing in cloudy atmospheres. *IEEE T. Geosci. Remote*, 43(6):1096–1101, May 2005.

M. Ekström, P. Eriksson, B. Rydberg, and D. P. Murtagh. First Odin sub-mm retrievals in the tropical upper troposphere: humidity and cloud ice signals. *Atmos. Chem. Phys.*, 7(2):459–469, 2007. URL www.atmos-chem-phys.net/7/459/2007/.

C. Emde, S. A. Buehler, C. Davis, P. Eriksson, T. R. Sreerekha, and C. Teichmann. A polarized discrete ordinate scattering model for simulations of limb and nadir longwave measurements in 1D/3D spherical atmospheres. *J. Geophys. Res.*, 109(D24):D24207, 2004. doi: 10.1029/2004JD005140.

P. Eriksson. Analysis and comparison of two linear regularization methods for passive atmospheric observations. *J. Geophys. Res.*, 105(D14):18157–18167, 2000.

P. Eriksson and D. Chen. Statistical parameters derived from ozonesonde data of importance for passive remote sensing of ozone. *Int. J. Remote Sensing*, 23:4945–4963, 2002. doi: 10.1080/01431160110111036.

P. Eriksson, F. Merino, D. Murtagh, P. Baron, P. Ricaud, and J. de La Noë. Studies for the Odin sub-millimetre radiometer: 1. Radiative transfer and instrument simulation. *Can. J. Phys.*, 80:321–340, 2002.

P. Eriksson, C. Jiménez, and S. A. Buehler. Qpack, a tool for instrument simulation and retrieval work. *J. Quant. Spectrosc. Radiat. Transfer*, 91:47–64, 2005. doi: 10.1016/j.jqsrt.2004.05.050.

P. Eriksson, M. Ekström, S. Bühler, and C. Melzheimer. Efficient forward modelling by matrix representation of sensor responses. *Int. J. Remote Sensing*, 27:1793–1808, 2006.

P. Eriksson, S. A. Buehler, C. P. Davis, C. Emde, and O. Lemke. ARTS, the atmospheric radiative transfer simulator, version 2. *J. Quant. Spectrosc. Radiat. Transfer*, 112:1551–1558, 2011. doi: 10.1016/j.jqsrt.2011.03.001.

P. Eriksson, B. Rydberg, H. Sagawa, M. S. Johnston, and Y. Kasai. Overview and sample applications of SMILES and Odin-SMR retrievals of upper tropospheric humidity and cloud ice mass. *Atmos. Chem. Phys.*, 14(23):12613–12629, 2014. doi: 10.5194/acp-14-12613-2014.

R. Larsson, S. A. Buehler, P. Eriksson, and J. Mendrok. A treatment of the Zeeman effect using Stokes formalism and its implementation in the Atmospheric Radiative Transfer Simulator ARTS. *J. Quant. Spectrosc. Radiat. Transfer*, 133:445–453, 2014. doi: 10.1016/j.jqsrt.2013.09.006.

C. Melsheimer, C. Verdes, S. A. Buehler, C. Emde, P. Eriksson, D. G. Feist, S. Ichizawa, V. O. John, Y. Kasai, G. Kopp, N. Koulev, T. Kuhn, O. Lemke, S. Ochiai, F. Schreier, T. R. Sreerekha, M. Suzuki, C. Takahashi, S. Tsujimaru, and J. Urban. Intercomparison of general purpose clear sky atmospheric radiative transfer models for the millimeter/submillimeter spectral range. *Radio Sci.*, 40:RS1007, 2005. doi: 10.1029/2004RS003110.

F. Merino, D. Murtagh, P. Eriksson, P. Baron, P. Ricaud, and J. de La Noë. Studies for the Odin sub-millimetre radiometer: 3. Performance simulations. *Can. J. Phys.*, 80: 357–373, 2002.

D. Murtagh, U. Frisk, F. Merino, M. Ridal, A. Jonsson, J. Stegman, G. Witt, P. Eriksson, C. Jiménez, G. Megie, J. de La Noë, P. Ricaud, P. Baron, J. R. Pardo, A. Hauchcorne, E. J. Llewellyn, D. A. Degenstein, R. L. Gattinger, N. D. Lloyd, W. F. J. Evans, I. C. McDade, C.S. Haley, C. Sioris, C. von Savigny, B. H. Solheim, J. C. McConnell, K. Strong, E. H. Richardson, G. W. Leppelmeier, E. Kyrölä, H. Auvinen, and L. Oikarinen. An overview of the Odin atmospheric mission. *Can. J. Phys.*, 80:309–319, 2002.

W. G. Read, Z. Shippony, M. J. Schwartz, N. J Livesey, and W. Van Snyder. The clear-sky unpolarized forward model for the eos aura microwave limb sounder (mls). *IEEE T. Geosci. Remote*, 44(5):1367–1379, 2006.

C. D. Rodgers. *Inverse methods for atmospheric sounding: Theory and practise.* World Scientific Publishing, 1 edition, 2000.

C. D. Rodgers and B J. Connor. Intercomparison of remote sounding instruments. *J. Geophys. Res.*, 108(D3), 2003.

B. Rydberg, P. Eriksson, S. A. Buehler, and D. P. Murtagh. Non-gaussian bayesian retrieval of tropical upper tropospheric cloud ice and water vapour from Odin-SMR measurements. *Atmos. Meas. Tech.*, 2(2):621–637, 2009. ISSN 1867-1381. URL www.atmos-meas-tech.net/2/621/2009/.

B. Rydberg, P. Eriksson, J. Kiviranta, A. Skyman, and D. Murtagh. Odin/SMR: Algorithms Theoretical Basis Document - Level 1b data. Technical report, Department of Earth and Space Sciences, Chalmers University of Technology, 2016.

R. Saunders, P. Rayer, P. Brunel, A. von Engeln, N. Bormann, L. Strow, S. Hannon, S. Heilliette, Xu Liu, F. Miskolczi, Y. Han, G. Masiello, J.-L. Moncet, G. Uymin, V. Sherlock, and D. S. Turner. A comparison of radiative transfer models for simulating Atmospheric Infrared Sounder (AIRS) radiances. *J. Geophys. Res.*, 112:D01S90, 2007. doi: 10.1029/2006JD007088.

RJ Sica and A Haefele. Retrieval of temperature from a multiple-channel rayleigh-scatter lidar using an optimal estimation method. *Appl. Opt.*, 54(8):1872–1889, 2015.

B. Tschanz and N. Kämpfer. Signatures of the 2-day wave and sudden stratospheric warmings in arctic water vapour observed by ground-based microwave radiometry. *Atmos. Chem. Phys.*, 15(9):5099–5108, 2015. doi: 10.5194/acp-15-5099-2015. URL www.atmos-chem-phys.net/15/5099/2015/.

J. Ungermann. *Tomographic reconstruction of atmospheric volumes from infrared limb-imager measurements*, volume 106. Forschungszentrum Jülich, 2011. URL http://hdl.handle.net/2128/4385.

J. Urban, P. Baron, N. Lautié, N. Schneider, K. Dassas, P. Ricaud, and J. de La Noë. Moliere (v5): A versatile forward-and inversion model for the millimeter and sub-millimeter wavelength range. *J. Quant. Spectrosc. Radiat. Transfer*, 83(3):529–554, 2004.

J. Urban, N. Lautié, E. Le Flochmoën, C. Jiménez, P. Eriksson, E. Dupuy, L. El Amraoui, M. Ekström, U. Frisk, D. Murtagh, J. de La Noë, M. Olberg, and P. Ricaud. Odin/SMR limb observations of stratospheric trace gases: level 2 processing of ClO, N_2O, O_3, and HNO_3. *J. Geophys. Res.*, 110:D14307, July 2005. doi: 10.1029/2004JD005741.

Format of L2 data

Who will define the L2 format?