Consolidation of Scientific Baseline for MTG-IRS Processing: Effect of Trace Gases

Final Document, FINAL version
## Document status

### Consolidation of Scientific Baseline for MTG-IRS Processing: Effect of Trace Gases

**Final Document, FINAL version**

<table>
<thead>
<tr>
<th>Issue</th>
<th>Revision</th>
<th>Date</th>
<th>Reason for the revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>02/08/2010</td>
<td>Initial draft version</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>11/10/2010</td>
<td>Final version in response to comments on the DRAFT version by EUMETSAT</td>
</tr>
</tbody>
</table>

## Modification status

<table>
<thead>
<tr>
<th>Issue</th>
<th>Rev</th>
<th>Status *</th>
<th>Modified sections</th>
<th>Reason for the modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>M</td>
<td>1</td>
<td>Precision on the fact that the effect of O3 is considered to be more important</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>M</td>
<td>3.3.1.5</td>
<td>Minor modifications in order to give more precision on the channel selection used</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>M</td>
<td>3.4.1</td>
<td>Minor modifications</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>I</td>
<td>3.4.2</td>
<td>This section was added in order to clearly state advantages and drawbacks of the sequential and combined method</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>M</td>
<td>3.4.3</td>
<td>Important modifications in order to make the recommendations more clear and in order to answer EUMETSAT's comments.</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>M</td>
<td>4.4</td>
<td>Minor modifications and clarifications of the conclusions of PHASE 3</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>I</td>
<td>5</td>
<td>Added this section with final conclusions of this study</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>M</td>
<td>All</td>
<td>Minor editing modifications</td>
</tr>
</tbody>
</table>

* **I** = Inserted  **D** = deleted  **M** = Modified
References

<table>
<thead>
<tr>
<th>Reference</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>[RD1] EUM/MET/SOW/09/0189, issue v2, 7 Mai 2009</td>
<td>Statement of Work for a Study on Consolidation of Scientific Baseline for MTG_IRS L2 processing: effect of trace gases</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>ARPEGE</td>
<td>Action de Recherche Petite Echelle Grande Echelle</td>
</tr>
<tr>
<td>BL</td>
<td>Boundary Layer</td>
</tr>
<tr>
<td>CHIMERE</td>
<td>Modèle de simulation et prévision de la pollution photo oxydante en milieu urbain</td>
</tr>
<tr>
<td>CNES</td>
<td>Centre National d’Etudes Spatiales</td>
</tr>
<tr>
<td>DOFS</td>
<td>Degree of Freedom Signal</td>
</tr>
<tr>
<td>ECMWF</td>
<td>European Centre for Medium-Range Weather Forecasts</td>
</tr>
<tr>
<td>EMEP</td>
<td>European Monitoring and Evaluation Programme</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>EUMETSAT</td>
<td>European Organisation for the Exploitation of Meteorological Satellites</td>
</tr>
<tr>
<td>GCM</td>
<td>General Circulation Model</td>
</tr>
<tr>
<td>GEMS</td>
<td>Global and regional Earth-system (Atmosphere) Monitoring using Satellite and in-situ data</td>
</tr>
<tr>
<td>IASI</td>
<td>Infrared Atmospheric Sounding Interferometer</td>
</tr>
<tr>
<td>IRS</td>
<td>(MTG) Infrared Sounder</td>
</tr>
<tr>
<td>KNMI</td>
<td>Koninklijk Nederlands Meteorologisch Instituut</td>
</tr>
<tr>
<td>L2</td>
<td>Level 2</td>
</tr>
<tr>
<td>LATMOS</td>
<td>Laboratoire Atmosphères, Milieux, Observations Spatiales</td>
</tr>
<tr>
<td>LMD</td>
<td>Laboratoire de Météorologie Dynamique</td>
</tr>
<tr>
<td>MACC</td>
<td>Monitoring Atmospheric Composition and Climate</td>
</tr>
<tr>
<td>MARS</td>
<td>Meteorological Archival and Retrieval System</td>
</tr>
<tr>
<td>MetOffice</td>
<td>Meteorological Office</td>
</tr>
<tr>
<td>MIST</td>
<td>MTG-IRS Science Team</td>
</tr>
<tr>
<td>MOCAGE</td>
<td>Modèle de Chimie Atmosphérique de Grande Echelle</td>
</tr>
<tr>
<td>MODIS</td>
<td>Moderate Resolution Imaging Spectroradiometer</td>
</tr>
<tr>
<td>MTG</td>
<td>Meteosat Third Generation</td>
</tr>
<tr>
<td>MRD</td>
<td>Mission Requirements Document</td>
</tr>
<tr>
<td>OEM</td>
<td>Optimal Estimation Method</td>
</tr>
<tr>
<td>OPD</td>
<td>Optical Path Difference</td>
</tr>
<tr>
<td>SEVIRI</td>
<td>Spinning Enhanced Visible-Infrared Imager</td>
</tr>
<tr>
<td>SRF</td>
<td>Spectral Response Function</td>
</tr>
<tr>
<td>TRAQ</td>
<td>Tropospheric composition and Air Quality</td>
</tr>
<tr>
<td>ULB</td>
<td>Université Libre de Bruxelles</td>
</tr>
<tr>
<td>UTLS</td>
<td>Upper Troposphere/Lower Stratosphere</td>
</tr>
<tr>
<td>VMR</td>
<td>Volume mixing Ratio</td>
</tr>
<tr>
<td>WRF</td>
<td>Weather Research &amp; Forecasting</td>
</tr>
</tbody>
</table>
Table of contents

1. INTRODUCTION .................................................................................................................................................. 9

2. ANALYSIS AND SPECIFICATION OF PROXY DATASET .................................................................................. 10
   2.1. INTRODUCTION ............................................................................................................................................... 10
   2.2. ANALYSIS OF AVAILABLE DATASETS ........................................................................................................... 10
       2.2.1. Existing MTG-IRS proxy data ................................................................................................................. 10
       2.2.2. Resolution of the existing MTG-IRS proxy data .................................................................................. 10
       2.2.3. Existing dataset from chemistry-transport models .............................................................................. 11
   2.3. PROXY DATASETS SPECIFICATION ............................................................................................................ 11
       2.3.1. Discussion of dataset specification ....................................................................................................... 11
       2.3.2. Synthetic presentation of the proposed specifications ......................................................................... 16

3. SENSITIVITY ANALYSIS ..................................................................................................................................... 20
   3.1. INTRODUCTION ............................................................................................................................................... 20
   3.2. SET-UP OF THE INVERSION SIMULATIONS FOR THE SENSITIVITY ANALYSIS PHASE ....................... 20
       3.2.1. Retrieval tool description ...................................................................................................................... 20
       3.2.2. Definition of geophysical states ........................................................................................................... 20
       3.2.3. Observations .......................................................................................................................................... 29
       3.2.4. Presentation of the retrieval approaches used ...................................................................................... 31
   3.3. RESULTS ....................................................................................................................................................... 31
       3.3.1. Reference retrieval .................................................................................................................................. 31
       3.3.2. Improved retrievals .................................................................................................................................. 40
   3.4. CONCLUSION AND RECOMMENDATIONS FROM THE SENSITIVITY ANALYSIS PHASE ..................... 47
       3.4.1. Summary ................................................................................................................................................. 47
       3.4.2. Conclusions on the sequential and combined methods ...................................................................... 48
       3.4.3. Recommendation .................................................................................................................................. 49

4. EVALUATION PHASE ........................................................................................................................................ 50
   4.1. SET-UP OF THE INVERSION SIMULATIONS FOR THE SENSITIVITY ANALYSIS PHASE ....................... 50
       4.1.1. Generation of proxy data ....................................................................................................................... 50
   4.2. METHODOLOGY .......................................................................................................................................... 57
       4.2.1. Testing the impact of O3 on the retrieval of temperature and H2O ...................................................... 57
       4.2.2. Description of scenarios ....................................................................................................................... 58
   4.3. RESULTS ....................................................................................................................................................... 59
       4.3.1. Analysis of the retrieval experiment results for temperature .............................................................. 59
       4.3.2. Analysis of the retrieval experiment results for H2O ........................................................................ 62
       4.3.3. Analysis of the retrieval experiment results for Ozone .................................................................... 65
   4.4. CONCLUSION AND RECOMMENDATIONS FROM THE EVALUATION PHASE ........................................ 70
       4.4.1. Summary and conclusions .................................................................................................................... 70
       4.4.2. Recommendations for future work ....................................................................................................... 70

5. FINAL CONCLUSIONS AND RECOMMENDATIONS .......................................................................................... 71

6. ACKNOWLEDGEMENTS ..................................................................................................................................... 72
1. Introduction

The hyperspectral infrared sounder (IRS) is one of the instruments considered in the phase B MTG activities. The main objective of this instrument is to provide frequent information on the atmospheric temperature, humidity and also wind profiles based on measurements of upwelling radiances at the top of atmosphere in two spectral bands: 700 – 1210 cm\(^{-1}\) and 1600 and 2175 cm\(^{-1}\). A secondary objective that is also explored concerns the utility of the instrument to provide data that will serve for the retrieval of information concerning atmospheric pollution, namely the concentrations of trace gases such as CO and O\(_3\).

In the current phase of activities, a prototype processor is being developed in order to demonstrate with the help of proxy data that the objectives set by the users for the IRS instrument can be met. The objective of the current study carried out by NOVELTIS in collaboration with LATMOS and ULB is to explore the impact of trace gases on the retrieval of H\(_2\)O and temperature and to propose an optimal retrieval strategy.

The work in this project was organised in three separate phases. In phase I, we briefly reviewed the existing proxy datasets and then addressed a number of important issues concerning the specifications of a dataset that can be used to test the impact of trace gases on the retrieval strategy. In phase II we used a database with a small number of representative atmospheric conditions in order to analyse the need to include trace gas retrievals in the main L2 processor to derive state vectors within accuracy required by the users. In addition, a recommendation was made on the strategy of retrieval of T and H\(_2\)O in the presence of trace gases. This strategy was tested in Phase III on proxy data describing a situation with a strong perturbation of O\(_3\) in the upper troposphere caused by a tropopause folding event. In addition, we tested with the help of several scenarios the pertinence of inverting O\(_3\) in the first step of the sequential inversion. Finally we provide recommendations for the retrieval of temperature and H\(_2\)O in the presence of trace gases as well as a description of work recommended for the Optional Phase. The combined impact of CO and O\(_3\) is analysed in Phase II, while in Phase III only the dominant impact of O\(_3\) is considered.

This document is organised as follows:

- section 2 contains the description of the proxy data specifications done in Phase I;
- section 3 contains the results from the Sensitivity Analysis Phase (Phase II);
- section 4 contains the results of the inversions for the selected inversion strategy for an event with tropopause folding event (Phase III);
- section 5 contains the final conclusions and recommendations;
- and finally in section 6 we list acknowledgements.
2. **Analysis and specification of proxy dataset**

2.1. **Introduction**

In this section, we provide specifications (as well as the justifications) for the proxy data that we proposed for the retrieval tests carried out in Phase III of this study. First, we briefly review the existing proxy datasets and then we address a number of important issues concerning the specifications of a dataset that can be used to test the impact of trace gases on the retrieval strategy.

We believe that the data generated according to the specifications that we proposed in this section would lead to a high level of realism in the study of the impact of the interactions between T/H$_2$O and the trace gases. Proxy data that do not follow the specifications can still be used, as it is the case in Phase III (see section 4) to analyse this impact, however at the risk of not capturing all the possible interactions between meteorological and chemical variables.

2.2. **Analysis of available datasets**

2.2.1. **Existing MTG-IRS proxy data**

The document [RD1] contains a description of MTG-IRS proxy datasets available at EUMETSAT. These data consist of MTG-IRS spectra that were simulated for conditions corresponding to the top of the atmosphere with the Optimal Spectral Sampling radiative transfer model [RD2]. The atmospheric and surface conditions used in these simulations were obtained with the help of regional (WRF, MetOffice regional model) and global (ECMWF) meteorological models. These models do not contain chemistry transport modules for simulating trace gases and thus the radiative transfer simulations were done without considering a possible coupling between meteorological and chemical variables.

In order to test the impact of trace gases on the retrieval of T and H$_2$O it is necessary to consider this coupling and have profiles of trace gases that are coherent with variables calculated by the meteorological models. An obvious example of this need is a case with a near surface temperature inversion near a pollution source. Applying a climatological profile of gases such as CO in this case will not take into account a possible accumulation of this pollutant and would not allow testing correctly the impact of this tracer on the retrieval of temperature and humidity.

In order to have profiles of chemical tracers that are coherent with the meteorological variables it is necessary to run a chemistry transport model forced by the meteorological fields. For this reason, in this study in which the impact of the trace gases on the retrieval is tested, we recommend not to use the existing MTG-IRS proxy data in their current state. Instead, we suggest generating missing chemistry fields with a chemical transport model using for forcing either the existing meteorological fields or results from additional meteorological simulations.

2.2.2. **Resolution of the existing MTG-IRS proxy data**

Various horizontal and vertical resolutions were used to generate the existing MTG-IRS proxy data. The ECMWF simulations were conducted with the resolution of 0.5°x0.5°, while the fine zooms of the MetOffice model and of the WRF were run at considerably higher resolutions that allow simulating the impact of heterogeneities in IRS pixels. In this document, we take as a starting point of the specification the fine-scale resolutions considered in the existing proxy data.
2.2.3. Existing dataset from chemistry-transport models

At a global scale, several state-of-the-art chemistry transport models are able to generate output in which the chemical tracers are coupled in a consistent way with the meteorological fields (e.g. LMDz-INCA [RD3], MOZART [RD4], MOCAGE [RD5]). In these simulations the meteorological fields are generally taken from analysis provided by operational weather models (ECMWF, NCEP). The resolution of these global data sets is currently limited due to the computational burden of the chemical calculations (model grids of a hundred or more kilometres). These datasets fail to capture fine scale interactions between meteorological and chemical variables, and are thus not suitable to be used as proxy data in this study. Indeed, anticipating the discussion on the required resolution provided in the next section that a resolution of a few kilometres (meso-scale) is required.

The state-of-the art simulations of chemistry-transport meso-scale models were recently conducted in support to the definition of a pollution satellite mission. These simulations have been conducted in recent international studies supported by ESA for the preparation of SENTINEL 4 and 5 (CAMELOT study, [RD6]) and of TRAQ (ON TRAQ study, [RD7]). In these studies, the regional chemical transport model CHIMERE has been used to generate atmospheric profiles in the lowermost troposphere only. In addition, the meteorological forcing has been done with output from coarse resolution global models (ECMWF) that was interpolated to the fine scale grid of the regional model. In this approach the chemical and dynamical fields are consistent, but the meteorological fields do not contain the fine scale features of meso-scale circulation.

2.3. Proxy datasets specification

This section contains elements for the specification for the proxy data set that we suggest in order to validate the choice of the retrieval strategy for H2O and T in the presence of traces gases (O3 and CO). The specifications consider three stages of the simulation process:

1) the simulation of the geophysical state with surface and atmospheric variables;
2) the simulation of the MTG-IRS spectra with a radiative transfer model;
3) and specification of the statistical errors (in the form of error covariance matrices) of the surface and atmospheric variables as well as the a priori and first guess for the variables to be retrieved.

Point 3 is necessary to have a complete proxy data set for the inversions.

In section 3.1 we discuss the specifications for the three points. Section 3.2 contains a succinct review of the specifications in the form of tables.

2.3.1. Discussion of dataset specification

2.3.1.1. Geophysical database

The objective is to specify a 3-D geophysical database that will be used as input to the radiative transfer simulations of MTG-IRS spectra. This database will include profiles of meteorological variables (temperature and humidity) and chemical species as well as surface parameters (skin temperature, surface emissivity).

Various aspects for the specifications are addressed below: coupling of the chemistry transport models with meteorological models, horizontal and vertical resolution, frequency of output, list of variables, source inventory used in simulations, as well as a size of the region, period and length of the simulations.
2.3.1.1.1 Specification of the modelling system

The data will be generated with the help of a modelling system that includes a meteorological component for calculating the meteorological variables and a chemistry transport component for simulating the transport and chemistry of chemical species. The chemistry transport component must be forced with output from the meteorological component to assure a coherency between the meteorological and chemical variables. The two modules can be run together (online) or separately (offline). In the latter case, the output from the high-resolution meteorological model is saved with specified frequency and is then used to force the chemistry-transport model. The meteorological fields are linearly interpolated in time to obtain forcing between the specified output times.

To obtain simulation at high horizontal resolution it is generally necessary to use several levels of zoom that allow gradually increasing the resolution from the coarse domain to the fine scale domain. Typically, to avoid instability problems, each zoom permits increasing the resolution by a factor of 3 at the maximum, so 2 or 3 levels of zoom are required to go from a grid of 100 km to the high spatial resolution considered here (section 2.3.1.1.2). Zooms can be used both in global models and in regional models. In both cases, the zoom domain uses lateral and upper boundary conditions from the domain with coarser resolution.

2.3.1.1.2 Horizontal resolution of the simulations

One of the key aspects of the simulations concerns the horizontal resolution at which the geophysical database will be generated. Too coarse resolution with respect to the size of the IRS footprint will result in too strong spatial averaging that might impact the conclusions of the study. With the hypothesis of homogenous IRS pixels, a resolution corresponding to the size of the footprint (specified as 4 km at nadir and about 6 km over Europe) is required to resolve the inter pixel gradients. Finally, if heterogeneous pixels are considered, the resolution needs to be increased to capture the spatial scale of the heterogeneities. One can distinguish different mechanisms through which heterogeneities at the scale of an IRS pixel affect the measured spectra, such as:

1) spatial variation of the target variability (i.e. profiles);
2) spatial variation of the surface parameters (i.e. temperature and emissivity);
3) cloud contamination of the scene.

One example of a mechanism that can influence the above points (especially 1 and 3) is convective mixing. In order to resolve this atmospheric process the resolution of 1.5 km was used in the MetOffice regional meteorological model.

As demonstrated by the MetOffice as well as other regional models that were used to generate the existing proxy data, meteorological fields can be obtained at high spatial resolution, given sufficient computer resources.

In principle, chemical transport models can be run at equally high resolution as the meteorological models; however, in practice they are limited by the computational requirements as they use considerably more computer time than meteorological simulations alone. Moreover, the surface emission inventories that are used by these models are not available at the resolution of a sub-grid of an IRS pixel. Fine scale inventories might exist for selected regions only, but at the scale of Europe, the best existing inventory (EMEP) is at 25x25 km resolution and a new inventory with a resolution of 5x5 km is currently under construction by the MACC GMES core service project.

High spatial variability of surface emissions can lead to strong gradients in the mixing ratios of trace gases in the lowest levels of the atmosphere. For this reason it is important to use emissions with resolutions close to the size of the IRS pixel (the newest MACC GMES emissions fill this criterion) and use at least matching horizontal resolution in the chemistry transport and meteorological models. It needs to be determined if such high resolution is needed above the boundary layer. In the absence of local sources (the only in-situ sources are lightning and aircraft sources of NO₃), it is expected that horizontal gradients are weaker as the atmospheric mixing tends to reduce with time the differences in concentrations. However, certain transport driven mechanisms, e.g. convection, mixing of air masses with different origins and tracer concentrations, can also lead to strong local gradients. In the case that it is judged not critical to use high resolution in the

© Noveltis 2010
This document is the property of Noveltis, no part of it shall be reproduced or transmitted without the express prior written authorisation of Noveltis
upper troposphere and the stratosphere, then different models could be used for the boundary layer (high resolution) and above (low resolution). This is important as certain of regional chemistry models were developed for air quality applications and have been applied mostly in the lower atmosphere (standard version of CHIMERE has only 8 levels with the last level below 5km). In this case, to prevent any discontinuities between two models, the upper boundary conditions should be taken from the coarse-resolution model for the corresponding time step.

Another class of regional models (such as the zoom of model MOCAGE, zoom of TM5) can be run with all of the vertical levels used in the original model, however, their horizontal resolution of the zoom can be limited (for example 1°x1° in TM5 in the current version).

In summary, we recommend to use the surface emission inventories with the highest spatial resolution available (about 5x5 km) and use similar resolution for meteorological and chemistry transport simulations in the lowermost atmosphere. It needs to be determined if such high resolution is needed in the upper troposphere and stratosphere.

2.3.1.1.3 Vertical resolution of the simulations

Concerning the vertical resolution, a critical parameter will be the vertical extent of IRS averaging kernels. With the IRS-MTG spectral resolution, the kernels will extend over several kilometres and thus vertical resolution in the troposphere of the order of 1km is sufficient. Higher vertical resolution is required in the boundary layer where:

1) there are strong episodes of pollution with CO values considerably higher than the background level and
2) the exact vertical extent of pollution plumes is important in connection with the radiative temperature contrast with the surface (no temperature contrast indicates that the layer will not be detectable in the spectra).

Due to the importance of this part of the atmosphere, this region is generally well resolved in the chemical and meteorological models.

2.3.1.1.4 Vertical extent of the simulation

It is important that the vertical profiles cover the appropriate vertical extent in order to properly compute the radiance at the top of the atmosphere, i.e. which includes stratosphere as well as surface and troposphere. We propose that the simulations cover a region from the surface to 0.5 mb.

2.3.1.1.5 Frequency of the output

The models using high-resolution zooms use small dynamical and chemical time steps that are adapted to the small size of the pixels. The time between model outputs will thus not be limited by the model but rather the volume of the data that are generated and that need to be stored. To choose appropriate output frequency, one has to consider the revisit time that is of the order of 15 and 60 minutes for local area coverage and full disk coverage modes respectively. Having model output with the frequency of the revisit time for a small region will allow testing if the temporal gradients are well captured by the inversion.

We suggest storing the output at the frequency of 15 minutes for short period episodes in the case the temporal gradients are to be analysed.

2.3.1.1.6 List of surface and atmospheric variables

Variables that need to be provided as input to radiative transfer model include all the surface and atmospheric parameters that impact the thermal infrared spectra at the top of the atmosphere. These parameters include also clouds, but for the purpose of this study, the information on clouds is not required.
A system consisting of a regional meteorological and chemistry transport model can generate most of the necessary atmospheric variables: surface temperature, air temperature and humidity, profiles of: CO, O3, NO2, CO2 and N2O are frequently not included in the standard chemistry transport models but their concentrations can be taken from climatology as these tracers show only little temporal variability. Surface temperature can be taken from the meteorological model, and shall be coherent with the air temperature and tracer profiles. As the surface temperature gradient has strong impact on the sensitivity of the MTG-IRS measurement, surface temperature from meteorological model should be validated against satellite products (e.g. SEVIRI) in order to validate the amplitude and extent of temporal and spatial gradients.

Surface emissivities should be taken from the high spectral resolution dataset from University of Wisconsin (Borbas et al. [RD9]). This database is derived from a combination of high spectral resolution laboratory measurements and surface emissivities for 10 MODIS channels by using principal component analysis regression.

2.3.1.1.7 Choice of region, size of the region and time period for the simulation

One of the criteria for choosing the region for the simulations will be the presence of pollution sources within the region. Such simulations will allow testing the impact of pollution episodes on the retrieval of H2O and T. The region should be sufficiently large to contain;

1) high contrasts in the concentrations of trace gases (polluted vs. clean regions) and
2) different land surface types (in particular land-sea contrast).

We propose a coastal region containing a major industrial area or a mega city with a size of at least 100 x 100 km.

For the choice of the period of simulation, in order to cover a potentially large number of different cases it is recommended to choose several episodes corresponding to different, if possible well documented, periods. The periods should be chosen for different seasons and for different times of the day. Time of the day is important as for example the surface temperature inversions are more likely to form at night.

2.3.1.2. Radiance simulation

The radiance at the top of the atmosphere will be constructed considering an appropriate integration of the atmospheric and surface features by line-by-line radiative transfer through the whole atmospheric path (vertical integration), and by convolution with the Instrument Spectral Response Function (spectral integration).

In the inversions to be carried later in this study, we will not consider Point Spread Function (PSF) integration. For each spectrum, we will retrieve one profile corresponding to the considered pixel. The proxy data can however contain the impact of spatial averaging over the PSF. This approach will result in more realistic spectra (the real instrument will integrate spatially) but it will also generate in our inversions an additional modelling noise term. This noise is not easy to evaluate, but can be considered of second order.

The MTG-IRS Instrument Spectral Response Function (shape and resolution), as well as the spectral sampling, is already specified in the MTG-IRS Science Plan.

2.3.1.3. Error covariance matrices

The Optimal Estimation Method (OEM) (Rodgers [RD8]) will be used in the level 2 processor to retrieve state vectors from simulated spectra. This commonly used inversion method uses observations, a priori information on the state of the atmosphere-surface system as well as a forward model that simulates the measurement. In order to apply the OEM, it is necessary to characterize error covariance matrices on the measurement (observation vector) and on the a priori state vector. The retrieval represents the posterior information provided by the measurement and a constraint given by a a priori information. Thus, the value and the characteristics of the retrieval strongly depend on the a priori information.
The characterisation of the error covariance matrices constitutes a critical element of the OEM method and the choice of the error covariance matrices for the evaluation phase is a major issue addressed in the later stages of this study.

2.3.1.3.1 Measurement error covariance matrix

The main source for the error covariance matrix on the measurement corresponds to the radiometric noise that constitutes the diagonal element of the covariance matrix. We propose that the specified values of the MTG-IRS radiometric noise are used for both spectral bands. A full error covariance matrix that takes into account also cross correlation errors between individual channels (due to spectral calibration for example) could be used as well if this matrix is specified in the framework of MIST.

2.3.1.3.2 State vector a priori error covariance matrix

In order to characterize the a priori error covariance matrix, we first need to define the retrieved state vector parameters as well as the transported state vector parameters.

The transported state vector parameters are the non-retrieved parameters for which the error covariance matrix is transported into the observation vector by use of the jacobian matrix. Thus, the resulting covariance matrix is added to the radiometric noise covariance matrix to constitute the observation vector error covariance matrix.

In order to analyse the impact of trace gases (CO & O₃) on the temperature profile (T) and humidity profile (Q) retrievals from the MTG-IRS measurement, it is necessary to take into account the geophysical parameters for which the MTG-IRS measurement is sensitive and to characterize the appropriate error covariance matrices of each parameter.

The main geophysical parameters of the study are:

- Temperature (T) and humidity (Q) vertical profiles that are the target products;
- CO and O₃ vertical profiles that are the main trace gases that could perturb T and Q retrievals;
- Surface temperature and spectral emissivity which characterize the surface properties;
- CO₂ and N₂O vertical profiles for which the abundance is also needed to be known;

The cloud contamination is also a critical variable to consider, but it is outside of the scope of this study.

The a priori error covariance matrices for each of the above parameters represent the uncertainty on the a priori knowledge of the atmospheric or surface states. They need to be chosen carefully by verifying that their choice is adapted to the questions of the study and is coherent with the chosen a priori state.

The error covariance matrices are usually derived from:

- climatology (mean value and standard deviation) derived from measurements or model simulations;
- estimation of the uncertainty of analysis from meteorological or chemical-transport models.

For these specifications, we propose that the temperature and humidity error covariance matrices applied at EUMETSAT for the inversion of IASI spectra are used in the retrieval of these two variables in this project. The a priori will be the true state corresponding to each spectrum (temperature and humidity profiles simulated by the meteorological model and used as input to the radiative transfer model) with added random noise in accordance with the specified covariance matrix.

For surface temperature that will also be retrieved, we suggest using the following values:

- 2 K over sea and
- 5 K over land.

The true state, perturbed in accordance with this noise levels, will be used as a priori in the inversions.
Surface emissivity will not be retrieved, but the effect of the error due to the incertitude of this parameter will be taken into account through the transport of error. We propose to use errors of 0.01 over sea and 0.02 over land and use the true state that was used to generate the spectra (database from E. Borbas, U. Wisconsin [RD9], see section 2.3.1.1.6).

The error covariance matrices for O₃ and CO as well as corresponding a priori profiles will be taken from recent ESA’s studies (TRAQ, CAMELOT). Depending on the results obtained in the later stages of this study, these two trace gases will be retrieved either at the same time as temperature and humidity (in which case cross correlation ad hoc matrices could be applied), or separately, in which case the error due to the retrieval of CO and O₃ will be transported during the retrieval of the meteorological profiles.

Finally, the error due to the uncertainty in the knowledge of the CO₂ and N₂O profiles will be determined based on variability in results obtained with a global transport model. These tracers will not be retrieved, and true state will be used in the inversions.

Other trace gases, such as NH₃, HNO₃, NO₂ and SO₂ have only a weak signature in the IRS spectrum and we will not take them into account in the retrieval process.

### 2.3.2. Synthetic presentation of the proposed specifications

#### 2.3.2.1 Specification of the modelling system

A modelling system allowing performing high-resolution simulations needs to be chosen. It is possible that the existing proxy data will be used as dynamical forcing for a chemistry transport model.

<table>
<thead>
<tr>
<th>Specification type</th>
<th>model</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meteorological fields + chemical data</td>
<td>Existing proxy data + regional chemistry model</td>
<td>A regional chemistry transport model forced by the meteorological fields needs to be run. Need to provide coherent boundary conditions from a global model. e.g. WRF + CHIMERE or MET OFFICE + CHIMERE</td>
</tr>
<tr>
<td></td>
<td>Other meteorological model + regional chemistry transport model</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Global meteorological model and global chemistry transport model with corresponding zooms</td>
<td>e.g. ARPEGE + MOCAGE +MOCAGE zoom</td>
</tr>
</tbody>
</table>

#### 2.3.2.2 Resolution of the simulations

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Range</th>
<th>comments</th>
</tr>
</thead>
</table>
| Horizontal resolution | About 5 km | 5kmx5km is the resolution of the surface emission inventory generated in the ongoing MACC GMES core service project. There are no inventories at finer scale (except over local areas).
6km is also the size of IRS-MTG footprint over Europe. If necessary (depending on the choice of models), the resolution can be lowered in the upper atmosphere. |
2.3.2.1.3  **Vertical Extent of the Simulations**

**Table 3: Specifications concerning the vertical extent of the model**

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Range</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical Extent</td>
<td>Surface up to 0.5 mb</td>
<td>All of troposphere and stratosphere in order to properly compute the radiance at the top of the atmosphere</td>
</tr>
</tbody>
</table>

2.3.2.1.4  **Frequency of output**

**Table 4: Specifications concerning the frequency of model output**

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Value</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of output</td>
<td>15 minutes</td>
<td>Based on the local area revisit time and in the case the temporal gradients will be analysed.</td>
</tr>
</tbody>
</table>

2.3.2.1.5  **List of variables**

**Table 5: Specifications concerning the list of variables**

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of meteorological atmospheric variables</td>
<td>Profiles of: temperature, humidity</td>
<td></td>
</tr>
<tr>
<td>List of chemical variables</td>
<td>Profiles of: CO, O₃, N₂O, CO₂, NO₂, NH₃</td>
<td>If CO₂ and N₂O are not modelled by the chemical transport model, climatologic or constant profile can be chosen for CO₂. For N₂O climatological profiles can be chosen.</td>
</tr>
<tr>
<td>Surface parameters</td>
<td>Surface temperature</td>
<td>Surface temperature from the meteorological model coherent with the air temperature and tracer profiles</td>
</tr>
<tr>
<td></td>
<td>Surface emissivity</td>
<td>Surface emissivity from the high resolution dataset, e.g. that provided by U. Wisconsin (Eva Borbas, [RD9])</td>
</tr>
</tbody>
</table>
2.3.2.1.6 Surface emission inventories of trace gases

Table 6: Specifications concerning surface emissions

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Surface emission files</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface emission inventories of trace gases to be used in the chemical transport</td>
<td>The emission inventories generated by the MACC project with resolution of 5 km x 5km.</td>
<td>This emission inventory has spatial resolution that matches the size of the IRS pixel.</td>
</tr>
</tbody>
</table>

2.3.2.1.7 Choice of region, size of region, and time period

Table 7: Specifications concerning region and period selection

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice of region</td>
<td>A region with gradients in tracers of chemical pollution</td>
<td>pollution region (megacity, industrial area) + background air</td>
</tr>
<tr>
<td></td>
<td>A region with different surface types</td>
<td>Land + sea</td>
</tr>
<tr>
<td>Size of region</td>
<td>at least 100 km x 100 km</td>
<td>Should be sufficiently large to contain gradients in chemical tracers (clean and polluted regions)</td>
</tr>
<tr>
<td>Period</td>
<td>Periods corresponding to different meteorological conditions and different seasons</td>
<td>Time of the day is also important</td>
</tr>
</tbody>
</table>

2.3.2.2 Radiance simulation

Table 8: Specifications concerning simulations of MTG-IRS spectra with a radiative transfer code

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of radiative transfer model</td>
<td>Line by Line radiative transfer model</td>
<td></td>
</tr>
<tr>
<td>SRF</td>
<td>sinc function defined on +/- 30 cm(^{-1})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.8 cm OPD</td>
<td></td>
</tr>
<tr>
<td>Spectral sampling</td>
<td>0.625 cm(^{-1})</td>
<td></td>
</tr>
<tr>
<td>PSF integration</td>
<td>No integration</td>
<td></td>
</tr>
</tbody>
</table>
### Error covariance matrices

Table 9: Specifications concerning error covariance matrices used in the 1D VAR inversions.

<table>
<thead>
<tr>
<th>Specification type</th>
<th>Error values or sources</th>
<th>A priori state</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error on observations</td>
<td>Radiometric noise that constitutes the diagonal elements of the covariance matrix</td>
<td>Specified values of the MTG-IRS radiometric noise will be used.</td>
<td></td>
</tr>
<tr>
<td>Specification type</td>
<td>Error values or sources</td>
<td>A priori state</td>
<td>Comments</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>T and Q profiles</td>
<td>- Eumetsat level2 <em>a priori</em> error covariance matrix</td>
<td>True state perturbed in a way consistent with the error covariance matrices (Eumetsat)</td>
<td>Retrieved</td>
</tr>
<tr>
<td>T&lt;sub&gt;S&lt;/sub&gt;</td>
<td>2 K (sea) 5 K (land)</td>
<td>True state perturbed in a way consistent with the error values</td>
<td>Retrieved</td>
</tr>
</tbody>
</table>
| ε<sub>S</sub> | - 0.01 (sea) and 0.02 (land) Other choices possible  - Error from Snyder database  - Borbas variability in the 200x200 km² area | True state | Not retrieved  
ε<sub>S</sub> not perturbed, but the effect of the error is taken into account through the transport of error. |
| CO and O<sub>3</sub> profiles | - Results from TRAQ or CAMELOT studies | - Mean profile from TRAQ or CAMELOT studies | Retrieved/not-retrieved? => results from sensitivity analysis |
| CO<sub>2</sub> and N<sub>2</sub>O profiles | - Northern mid-latitude variability from model | True state | Not retrieved (error transport) |
3. Sensitivity analysis

3.1. Introduction
This section contains the results from the Sensitivity Analysis Phase, in which we use a database with a small number of representative atmospheric conditions in order to analyse the need to include trace gas retrievals in the main L2 processor to derive state vectors within accuracy required by the users. The results from this phase are used to formulate recommendations on the retrieval method of temperature and H₂O in the presence of CO and O₃ that is tested in Phase III (section 4).

The section is organised as follows. Section 3.2 contains a description of the set-up of the simulations which includes the description of the inversion tool used, and a description of the geophysical scenarios and error covariance matrices. Section 3.2 also contains a description of the methodology used to address the T and H₂O inversion approach in the presence of trace gases.

In section 3.3 we present the results. In section 3.4 we give the first conclusions and recommendations for consolidation work.

3.2. Set-up of the inversion simulations for the sensitivity analysis phase

3.2.1. Retrieval tool description
The retrieval tool used in this study is NOVELTIS 1DVAR inversion package. The tool is based on the Optimal Estimation Method (OEM) ([RD8]) formulation with Levenberg Marquardt scheme for non linear cases. Based on this retrieval scheme, several tools have been developed. In particular, the 1DVAR package contains:

- an information content analysis package (DOFS, Averaging kernels, error analysis and error transport);
- a channel selection tool based on an iterative method that optimises the information content ([RD8]);
- a grid transformation tool for profiles and error covariance matrices;
- an iterative inversion procedure for retrieving atmospheric and surface properties.

The fast line-by-line radiative transfer model 4AOP is used within the 1DVAR iterative retrieval scheme for the calculations of the spectra and the jacobians.

3.2.2. Definition of geophysical states
In order to study the impact of trace gases (CO and O₃) on temperature and moisture profile (T and H₂O) retrievals, we have considered state vectors containing an ensemble of geophysical parameters that influence temperature and H₂O retrievals. These parameters are discussed in section 3.2.2.1. In order to consider different geophysical conditions, the parameters were selected for 10 different cases which are discussed in section 3.2.2.2.
3.2.2.1. Description of the state vector used in the retrievals

Table 10 lists the geophysical parameters considered in this study to define the state vector. As mentioned in the second column of this table, information on O₃, CO, CH₄, as well as temperature, H₂O and surface pressure and temperature are derived from the output of the TM5 chemistry-transport model runs performed at KNMI, and already used in the framework of a CNES funded study led in 2009. These fields are either calculated by the TM5 model (O₃, CO, CH₄) or are taken from the ECMWF analyses which are interpolated on the horizontal (1°x1° in the zoom region) and vertical (34 levels) grids of the TM5 model. The number of vertical levels of the state vector is 34 and it corresponds to the number of layers in the TM5 model.

Not all of the listed parameters are inverted in our study. As specified in Table 11, besides temperature (air and surface) and H₂O, CO and O₃ are also inverted in several of the tested scenarios (see section 3.2.4 for details).

An important part in the description of the state vector concerns the specification of the error on the prior knowledge for retrieved parameters and for parameters impacting on these retrievals. A brief description of the origin of the various error covariance matrices used in this study is given in the last column of Table 11.

Graphical representation of the correlation matrices together with the corresponding variances are plotted in Figure 1 to Figure 6 for temperature, H₂O, O₃ and CO.

T and H₂O matrices are shown in two configurations (Figure 1 to Figure 4). In the NWP configuration (Figure 1 and Figure 2), the matrices for temperature and H₂O are based on the errors estimated for meteorological forecasts generated by the ARPEGE GCM at METEO FRANCE. The error on temperature profile is in the range of 1 and 1.5 K in the troposphere. For H₂O, the error is specified in the units of specific humidity (g/g). The same error is applied to the different atmospheric conditions (e.g.: winter and summer). A second configuration (Figure 3 and Figure 4) considers climatological variability. For O₃, we have taken the error covariance matrix provided by ULB, and already used during the ESA CAMELOT and ON-TRAQ projects [RD6] and [RD7].

The above matrices are given on their own vertical coordinate systems and we use the grid transformation tool available in the NOVELTIS 1DVAR inversion package, to transform these matrices on the 34 vertical levels at which the inversions are carried out.

For CO we have generated a simple ad hoc matrix as specified in Table 11 as we have encountered numerical difficulties when using the ULB generated matrix on the 34 level discretisation of our state vector.

In the work presented in this report, we generally did not consider cross-correlations between different parameters (all cross correlation terms between different parameters are set to 0).

Table 10: Description of the state vectors considered for the sensitivity analysis

<table>
<thead>
<tr>
<th>Parameters</th>
<th>State vector</th>
<th>Selected cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmospheric temperature profile</td>
<td>ECMWF analysis interpolated on the 1°x1° grid of TM5 and on 34 vertical levels</td>
<td></td>
</tr>
<tr>
<td>Atmospheric H₂O profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmospheric CO profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmospheric O₃ profile</td>
<td>Results from TM5 on 34 vertical layers</td>
<td>10 geophysical cases described in section 3.2.2</td>
</tr>
<tr>
<td>Atmospheric CH₄ profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameters</td>
<td>State vector</td>
<td>Selected cases</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Surface Temperature</td>
<td>ECMWF analysis interpolated on the 1°x1° grid of TM5</td>
<td></td>
</tr>
<tr>
<td>Surface Pressure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface Emissivity</td>
<td>Snyder emissivity ([RD11])</td>
<td>Associated with the 10 geophysical cases: either sea or land (organic bare soil)</td>
</tr>
<tr>
<td>Atmospheric CO₂ profile</td>
<td>380 ppmv (vertically constant)</td>
<td>Climatological profile (global)</td>
</tr>
<tr>
<td>Atmospheric N₂O profile</td>
<td>Climatological profile ([RD10]) with tropospheric values updated to current day values (320 ppbv)</td>
<td>Climatological profile (global)</td>
</tr>
<tr>
<td>Cloud flag</td>
<td>Assumed clear-sky</td>
<td>Clear-sky</td>
</tr>
</tbody>
</table>
### Table 11: Error covariance matrices for state vector parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Inversion</th>
<th>Prior Error Covariance Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmospheric temperature profile</td>
<td>Yes</td>
<td>Based on results from the Arpege Model</td>
</tr>
<tr>
<td>Atmospheric H$_2$O profile</td>
<td>Yes</td>
<td>Based on results from the Arpege Model</td>
</tr>
<tr>
<td>Atmospheric CO profile</td>
<td>No in the reference case</td>
<td>Ad hoc matrix with relative error of 50% and vertical correlations based on the correlation length = $(\exp(-(level(i)-level(j))/10)$, where level is the number corresponding to one of the 34 levels of TM5</td>
</tr>
<tr>
<td>Atmospheric O$_3$ profile</td>
<td>No in the reference case</td>
<td>TRAQ Outputs (European climatologic variability)</td>
</tr>
<tr>
<td>Surface Temperature</td>
<td>Yes (with emissivity error)</td>
<td>2K on sea</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 K on land</td>
</tr>
<tr>
<td>Surface Emissivity</td>
<td>No (error transport on H$_2$O &amp; T retrievals)</td>
<td>Snyder errors (sea/land)</td>
</tr>
<tr>
<td>Atmospheric CO$_2$ profile</td>
<td>No (error transport on H$_2$O &amp; T retrievals)</td>
<td>Northern mid-latitude variability derived from the results of the LMDz-INCA model ([RD3])</td>
</tr>
<tr>
<td>Atmospheric N$_2$O profile</td>
<td>No (error transport on H$_2$O &amp; T retrievals)</td>
<td>Northern mid-latitude variability derived from the results of the LMDz-INCA model ([RD3])</td>
</tr>
<tr>
<td>Atmospheric CH$_4$ profile</td>
<td>No</td>
<td>Not taken into account</td>
</tr>
<tr>
<td>Surface Pressure</td>
<td>No</td>
<td>Not taken into account</td>
</tr>
<tr>
<td>Cloud flag</td>
<td>No</td>
<td>Not taken into account</td>
</tr>
</tbody>
</table>
Consolidation of Scientific Baseline for MTG-IRS Processing: Effect of Trace Gases

Figure 1. Temperature profile error: Standard deviation (left) and correlation matrix (right). Values derived from results obtained with the ARPEGE forecast model.

Figure 2. H₂O profile error: standard deviation (left) and correlation matrix (right). Values derived from results obtained with the ARPEGE forecast model.
Figure 3: Temperature profile error: Standard deviation (left) and correlation matrix (right). Climatological values.

Figure 4: H₂O profile error: standard deviation (left) and correlation matrix (right). Climatological values.
Concerning the surface emissivities, we consider two emissivity spectra provided by Snyder ([RD11]) for two surface categories: sea and organic bare soil. These two emissivity spectra, shown in Figure 7, are used for all ocean and land cases of the geophysical selection, respectively. The prior error on the emissivity is specified in Table 12.
Consolidation of Scientific Baseline for MTG-IRS Processing: Effect of Trace Gases

Figure 7: Surface emissivities for two classes as given by Snyder

<table>
<thead>
<tr>
<th>Spectral band</th>
<th>Sea</th>
<th>Land</th>
</tr>
</thead>
<tbody>
<tr>
<td>700-970 cm⁻¹</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>970-1070 cm⁻¹</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>1070-1160 cm⁻¹</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>1160-1210 cm⁻¹</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>1600-1700 cm⁻¹</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>1700-1800 cm⁻¹</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>1800-1900 cm⁻¹</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>1900-2000 cm⁻¹</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>2000-2100 cm⁻¹</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>2100-2175 cm⁻¹</td>
<td>0.01</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table 12: A priori emissivity error from Snyder database

3.2.2.2. Geophysical cases

To test the impact of CO and O₃ on the retrieval of temperature and H₂O, we have selected 10 atmospheric profiles representing a wide spectrum of atmospheric conditions over north-western Europe. As mentioned in section 3.2.2.1, the chemical profiles were extracted from a database generated by KNMI with the TM5 chemical transport model. The profiles contain (see Table 11) information on CO, H₂O, temperature, O₃, CH₄, as well as surface temperature and pressure. The version used for the simulations contained a zoom over Europe with a spatial resolution of 1°x1°. All selected profiles were extracted from a subset of this zoom containing Northern France, Germany, the Benelux countries and southern England.

The TM5 simulations were forced by meteorological fields (e.g.: winds, air temperature, surface temperature, humidity profiles) taken from the ECMWF analysis interpolated on the TM5’s grid. The 10 vertical profiles of temperature and H₂O were selected for the same model grids as the chemical profiles giving us chemical profiles that are coherent with the meteorological variables.

The version of TM5 contained 34 vertical levels with the centre of the top layer at 0.1 hPa.

A brief description of the main characteristics of the profiles is provided in Table 13. We cover the cases over land and sea, in winter and summer, with and without surface pollution, with and without boundary layer temperature inversions, and with and without CO and O₃ anomalies (presence of air masses with high CO or O₃ content) in the upper troposphere.

© Noveltis 2010
This document is the property of Noveltis, no part of it shall be reproduced or transmitted without the express prior written authorisation of Noveltis
Table 13: A summary description of the 10 selected geophysical cases.

The colours correspond to the colours used in the figures below.

Figure 8 shows the tropospheric profiles of temperature, H2O, CO and O3 for the ten chosen cases. In Figure 8 a and b, the surface temperature is indicated by circles to indicate the presence/absence of vertical temperature contrast. For cases 4, 5 and 7 the temperature of surface is colder than that of the overlying air (night time cases), while for cases 9 and 10 the opposite is true (afternoon cases).
3.2.2.3. Choice of a representative subset of scenarios

Most of the tests will be conducted on the totality of the 10 scenarios. However, for certain tests (see sections 3.3.1.5, 3.3.2.3) it was justified to limit the number of geophysical cases. We chose here a subset of 3 representative cases.

- Case 1: small impact of CO and O₃ errors on the total retrieval (winter case);
- Case 6: non-negligible CO and O₃ impact without temperature inversion;
- Case 7: non-negligible CO and O₃ impact with temperature inversion.

3.2.3. Observations

In order to specify the observations for the purpose of this study it is necessary to provide information on:

1) radiometric noise;
2) spectral sampling;
3) spectral response function (SRF).
We have used the value of 0.625 cm\(^{-1}\) from the MTG-MRD ([RD12]) for the spectral sampling. The SRF used is a sinus cardinal function with OPD of 0.8 cm and defined on +/- 30 cm\(^{-1}\). Figure 9 shows the corresponding instrument function.

The radiometric noise is specified in MTG-MRD and is shown in Figure 10.

![Graphical representation of the instrument spectral response function used.](image1)

![Specified noise level for MTG-IRS.](image2)
3.2.4. **Presentation of the retrieval approaches used**

The results presented in this section 3 are based on the analysis of the information content of the retrieval approach. This method allows calculating the information content given by the DOFS as well as the theoretical error reduction based on:

1) the errors on the prior knowledge of the state vector;
2) the errors of the measurements;
3) the jacobians calculated for the retrieved state vector parameters (T, Ts, H2O, O3, CO) with the 4AOP radiative transfer model.

In order to assess the impact of CO and O3 on the retrieval of temperature and H2O, the information analysis tool has been applied for several retrieval strategies and configurations defined below.

### 3.2.4.1. Reference retrieval

To evaluate the performance of the different inversion approaches tested in this study, it is useful to define a reference retrieval against which other results can be compared. We define this **reference retrieval** as a scenario in which only the surface temperature and profiles of temperature and H2O are inverted. The impact of CO and O3 is calculated only through the transport of *a priori* errors of CO and O3 on the final solution. In this case, the information on CO and O3 present in the measurements (spectra) is not used.

We compare the reference retrieval with results obtained with the information content analysis tool from the case where CO and O3 errors are not transported, i.e., the prior errors on CO and O3 are set to zero (perfect knowledge). The results from this comparison give a first estimation of the impact of CO and O3 errors on the final solution (see section 3.3.1). This impact is referred to as **reference impact** in the remainder of the document.

The reference retrieval is applied to the 10 geophysical scenarios listed in Table 13. No channel selection is considered.

### 3.2.4.2. Improved retrievals

In the next step, we test the theoretical improvement that is achieved when CO and O3 are inverted. The results from two approaches are analysed:

- Sequential retrieval in which CO and O3 are inverted in the first step (considering transport of *a priori* errors of temperature and H2O), surface temperature, H2O and temperature profiles in the second step by using the reduced errors on CO and O3;
- Combined retrieval, when all 4 parameters are retrieved in one step.

The results from these simulations are presented in section 3.3.2.

### 3.3. Results

#### 3.3.1. Reference retrieval

As mentioned in section 3.2.4.1, the reference retrieval and reference impact are used as benchmarks against which the results of the improved simulations are compared in section 3.3.2.

#### 3.3.1. Analysis of geophysical cases

As a first step in presenting the reference scenario, we discuss the DOFS, the weighting functions and the retrieval error on H2O and temperature. The results are shown for the 10 selected geophysical cases (described in section 3.2.2).
3.3.1.1.1 DOFS

Figure 11 shows the performances of the inversion of temperature and H$_2$O from IRS/MTG spectra in terms of DOFS obtained for the reference retrieval. The results, which are shown here as a function of case number, can be summarised as follows:

- DOFS for surface temperature is very close to 1 for all analysed cases;
- With the ARPEGE \textit{a priori} error used, the DOFS for temperature is between 4 and 6 for the analysed cases (mean of 5.5);
- For H$_2$O the mean DOFS is 7.2;
- For the 10 geophysical cases studied, temperature and H$_2$O DOFS are generally anti-correlated;
- Highest DOFS for H$_2$O is obtained for cases with weak H$_2$O content;
- Lowest DOFS for temperature is obtained for winter cases (case 1 and 8);
- Highest total DOFS (sum of DOFS for temperature, H$_2$O and surface temperature) is obtained for cases with boundary layer temperature inversion (4, 5, 7).

![Figure 11: Reference retrieval: information content in terms of Degree Of Freedom of Signal (DOFS) for temperature profile (black), surface temperature (red), H$_2$O profile (green) and total DOFS (blue) as a function of the geophysical case.](image-url)
3.3.1.1.2 T and H₂O vertical profiles

Next, we analyse the retrieval performance for MTG IRS spectra obtained for the 10 geophysical cases. The results obtained for retrieval of temperature and H₂O profiles are presented in terms of weighting functions in Figure 12 and as retrieval errors in Figure 13 and Figure 14. Figure 13 shows the posterior error for temperature and H₂O in absolute units (K and VMR respectively). For H₂O, the results are also shown in relative units in Figure 14. In relative units, for each geophysical case, the H₂O posterior error is divided by the corresponding H₂O profile (left figure) and by the prior error profile (right figure). Results can be summarised as follows:

- Figure 12 indicates that for both temperature and H₂O highest sensitivity is obtained between 850 and 350 hPa. The sensitivity decreases above 350 hPa and below 850 hPa (except for cases with near surface temperature inversions, i.e.: cases 4, 5 and 7);
- Figure 13 shows that the highest temperature error reduction is obtained for summer cases. In the boundary layer, the lowest error is obtained for cases with temperature inversion;
- In the free troposphere, the highest H₂O error reduction in absolute units is obtained for winter cases (i.e.: 1 and 8). In the boundary layer, the highest reduction is for cases with temperature inversion. In relative units, the best performances are obtained for summer cases with about 10% error up to 350 hPa (Figure 14, left).

![Figure 12. Reference retrieval: weighting functions for retrievals of temperature (left) and H₂O (right) profiles as a function of the geophysical case.](image-url)
Consolidation of Scientific Baseline for MTG-IRS Processing: Effect of Trace Gases

Figure 13. Reference retrieval: retrieval errors on temperature (left) and H$_2$O (right) profiles as a function of the geophysical case.

Figure 14. Reference retrieval: retrieval H$_2$O profiles relative errors expressed as a fraction of H$_2$O profile (left figure) and fraction of a priori error (right figure)

3.3.1.2. Impact of trace gases: CO and O$_3$

We make the first evaluation of CO and O$_3$ impact on the temperature and H$_2$O retrieval error budget by comparing the reference case with a case in which CO and O$_3$ are considered as perfectly known. The difference, called here the reference impact, indicates the impact of the transport errors of CO and O$_3$ in the reference case retrieval (see also section 3.2.4.1).

Figure 15 shows the DOFS for temperature and H$_2$O retrievals. In terms of information content, CO and O$_3$ errors have an effect only on temperature for which the DOFS is reduced by around 0.5. For H$_2$O the impact of error transport is negligible.
Figure 15: DOFS for temperature (black) and H₂O (red) retrievals as a function of the geophysical case. Solid lines - retrieval cases with no CO/O₃ impact, dashed lines – retrieval with CO/O₃ a priori error impacts (reference retrieval).

Figure 16 shows the prior errors and the posterior errors for the two retrieval scenarios for the first geophysical case. The results indicate that the impact of transporting errors of CO and O₃ on the total retrieval error is small. Similar results are obtained for the remaining geophysical cases.

To quantify the impact of transporting CO and O₃ errors on temperature and H₂O retrievals, we define:

\[
\text{contribution in absolute units} = \sqrt{E_{ref}^2 - E_s^2}
\]

Eq. 1

and

\[
\text{contribution in relative units} = \frac{\sqrt{E_{ref}^2 - E_s^2}}{E_{ref}}
\]

Eq. 2
where $E_{\text{ref}}$ is the posterior error for the reference case; 
$E_s$ is the posterior error for the case where the CO and O$_3$ error is not transported.

In Figure 17 we show the contribution in absolute units, and in Figure 18 the contribution in relative units.

**Figure 17:** Black solid curve: prior errors for temperature (left figure) and H$_2$O (right figure). Remaining solid curves: total posterior errors (for the case with CO and O$_3$ error transport) for 10 geophysical cases. Dashed curves: Absolute contribution due to the impact of CO and O$_3$ uncertainties on the total retrieved profile errors (Eq.1).

**Figure 18.** Relative contribution due to the impact of CO and O$_3$ uncertainties on the total retrieved profile errors (Eq.2).

The main results can be summarised as follows:

- The impact of transporting the prior error of CO and H$_2$O on the total retrieval error is weak, but non negligible, especially if this error corresponds to a systematic bias;
- The impact of CO/O$_3$ prior errors is about 0.1 K for temperature and less than 0.0001 (VMR) for H$_2$O;
- The impact of CO/O$_3$ prior errors is about 15% (varying between 5% and 20% depending on geophysical conditions and altitude) of the total retrieval error for temperature and about 10% (varying between 1% and 15% depending on geophysical conditions and altitude) for H$_2$O;
• The impact seems to be depending more on thermodynamic conditions than on the CO/O\textsubscript{3} conditions:
  o Generally small impact for winter cases (1, 8)
  o In the middle troposphere: stronger impact of CO/O\textsubscript{3} for cases with good temperature retrieval and weak sensitivity to H\textsubscript{2}O (cases 2, 3, 9: high H\textsubscript{2}O content)
  o In the boundary layer we observe anti-correlation between the impact of CO/O\textsubscript{3} on temperature and the impact on H\textsubscript{2}O:
    ▪ Strong impact on H\textsubscript{2}O and weak impact on temperature for cases with temperature inversion (cases 4, 5, 7).
    ▪ Strong impact on temperature and weak impact on H\textsubscript{2}O for cases with strong H\textsubscript{2}O content in lower troposphere (cases 3, 6, 9, 10) except for case 2.

3.3.1.3. Emissivity impact

Figure 19 shows the impact of transporting errors due to surface emissivity uncertainties on temperature and water vapour retrievals. The contribution of emissivity errors, presented in relative units as defined in Equation 2, is presented, and can be compared to the contribution of O\textsubscript{3}/CO errors shown on Figure 18.

The impact of emissivity error on the T/H\textsubscript{2}O retrieval quality is larger than the impact of trace gases, in particular in the low troposphere.

![Figure 19: Impact of surface emissivity errors: relative contribution from surface emissivity uncertainties to the total retrieved profile errors (Eq.2) in terms of temperature (left panel) and water vapour (right panel).](image)

3.3.1.4. Impact of error covariance matrices

Figure 20 and Figure 21 (zoom) compare the impact of trace gas errors on temperature retrieval (the contribution is shown in absolute units) when two different types of error covariance matrices are considered for temperature and water vapour: Numerical Weather Forecast \textit{a priori} (left panel in Figure 20, same as left panel in Figure 17) and Climatology (right panel in Figure 20). Figure 22 and Figure 23 compare the same impact on water vapour retrieval.

These results indicate that the contribution of trace gas errors remains weak with respect to total error when climatology \textit{a priori} is considered, both for temperature and water vapour retrieval. A larger impact is observed in low troposphere levels, but basically the conclusions on the impact of trace gases should be independent from the \textit{a priori} knowledge on temperature and water vapour.
Figure 20: Solid black line – prior temperature error (in the right figure it is outside of the chosen scale), solid line– posterior temperature errors, dashed line– fraction of the posterior error due to CO and O₃. Left figure: with forecast type error covariance matrices and right figure: with climatological type matrices.

Figure 21: Zoom on the fraction of retrieval Temperature error shown in Figure 20. Left figure: with forecast type error covariance matrices and right figure: with climatological type matrices.

Figure 22: Solid lines– posterior H₂O errors in fraction of profile, dashed lines– fraction of the posterior error due to CO and O₃. Left figure: with forecast type error covariance matrices and right figure: with climatological type matrices.
3.3.1.5. Mitigating the impact by using channel selection

Figure 24 and Figure 25 give an estimate of the impact of channel selection on previous results. A channel selection is applied, which removes a subset of channels sensitive to both O₃ and CO₂ while at the same time preserving maximum information on temperature and water vapour.

Left panels of Figure 24 and Figure 25 give the contribution (in relative units) of CO and O₃ errors when this channel selection is applied (to be compared with Figure 18, which gives the same contribution without channel selection).

Right panels of Figure 24 and Figure 25 give a quantitative estimate of the loss of information due to channel selection.

These results, showing an impact similar to that obtained without channel selection, indicate that the effect of channel selection for reducing the impact of CO and O₃ is negligible (if this selection is designed to preserve the information on temperature and water vapour) in our case study. This conclusion is directly related to the fact that the original effect of CO and O₃ impact is already weak: in these specific conditions, any channel selection cannot significantly reduce this impact.

However, it shall be noticed that a channel selection remains a good way for mitigating any impact of trace gases. On the other hand, channel selection should be used in a sequential strategy as discussed latter in this document. For these general reasons, a channel selection will be considered in our recommendation for further implementation of the retrieval strategy.
3.3.2. Improved retrievals

In this subsection we evaluate the benefit on the temperature and H2O profile retrievals that could result from the inversion of CO and O3 with the MTG IRS measurement, by means of sequential and combined retrievals.

3.3.2.1. CO and O3 retrievals

First, we analyse the performances of CO and O3 retrievals in terms of DOFS (Figure 26) and vertical profile error (Figure 27 and Figure 28) for the ten geophysical cases. These results are shown for the sequential and combined retrievals.

DOFS results

- Both for O3 and CO, the DOFS is almost identical for the two retrieval strategies;
- Mean DOFS about 3.5 for O3 and 1.2 for CO;
- Global sensitivity on CO and O3 are correlated for the 10 geophysical cases;
- Min DOFS for winter cases (i.e.: cases 1 and 8);
- Max DOFS for cases with high water vapour content.
Consolidation of Scientific Baseline for MTG-IRS Processing: Effect of Trace Gases

Figure 26: DOFS obtained for O₃ and CO retrievals as a function of the geophysical state and the retrieval approach (the DOFS for the sequential and combined retrievals are almost identical).

Results on vertical profile errors

- Sensitivity on season of the posterior CO and O₃ errors in the UTLS region: the reduction is clearly smaller for winter cases (1,8);
- Smaller error reduction for both CO and O₃ retrievals in the lower troposphere for cases with temperature inversion and low H₂O content (4,5,7);
- Higher error reduction for CO and O₃ retrievals in the lower troposphere for cases with positive temperature contrast at the surface (and high H₂O content) (3,6,9,10).

Figure 27: Posterior retrieval errors on O₃ (left) and CO (right) profiles as a function of the geophysical case. Results are shown for the sequential retrieval.
3.3.2.2. Sequential retrieval

For the sequential retrievals, we calculate the impact of the CO and O₃ retrieval errors and we compare it to the impact obtained in section 3.3.1 for the reference retrieval (with and without transport of prior CO and O₃ errors).

We begin by presenting in Figure 29 the calculated DOFS for temperature and H₂O. The main conclusions are:

- The impact of CO and O₃ prior errors on H₂O sensitivity is not significant, and the DOFS are very close for the three retrieval approaches (no CO/O₃ error transport, CO/O₃ error transport, CO/O₃ inversion using the sequential method);
- The impact of CO and O₃ prior errors on DOFS for temperature corresponds to a reduction of about 0.5 with respect to the case with no error transport (for both CO/O₃ error transport and the sequential approach);

The reduction of the impact of CO/O₃ on temperature retrieval (shown as the increase of DOFS obtained with the sequential retrieval as compared with the reference retrieval considering simply error transport) is small but significant. The effect of considering sequential retrieval in place of simple error transport is the most significant for winter cases (cases 1 and 8), when the absolute value of DOFS is the smallest.
Next, to compare the error reduction on temperature and H$_2$O for the sequential case with the reference impact we define the **sequential impact** as:

\[
\text{contribution in relative units} = \frac{E_{\text{seq}}^2 - E_s^2}{E_{\text{seq}}},
\]

Eq. 3

Where $E_{\text{seq}}$ is the posterior error on temperature and H$_2$O for the sequential case, and $E_s$ is the posterior error for the case when CO and O$_3$ errors are set to 0.

Figure 30 compares the reference impact with the sequential impact in relative units, giving a measure of the interest of applying a sequential retrieval strategy, with respect to a single T/H$_2$O retrieval considering CO/O$_3$ error transport (reference case).

**Results on vertical profiles**

- The difference between the reference and sequential impacts is highest (highest error reduction) for the winter case (1) with highest reduction in the boundary layer.
- Very small reduction between the reference and sequential impacts in the case with temperature inversion (7).
- For both T and H$_2$O, the mean reduction is of about 2% (relative to 10% CO/O$_3$ impact)
3.3.2.3. Combined retrieval

In the combined retrieval, CO and O₃ as well as temperature (air and surface) and H₂O are retrieved simultaneously using prior errors on these parameters.

The results for DOFS for the combined case are shown in (Figure 31). As was the case for the sequential retrieval, the water vapour DOFS is the same for all three cases. Concerning DOFS for temperature, the difference between the combined and reference scenarios becomes negligible.

Figure 32 shows that for the vertical contributions to the total error, the impact of CO and O₃ errors is similar for the reference and combined retrievals.

![Figure 30: CO/O₃ impact reduction (fraction of the retrieval error) on the temperature (left) and H₂O (right) retrievals as a function of selected geophysical case. Reference impact (solid) and sequential impact (dash).](image)

![Figure 31: Same as Figure 29 but in the case of combined retrievals](image)
At first glance, these results seem surprising because one could expect that combined retrievals improve the performances of temperature and H\textsubscript{2}O retrievals. A possible explanation, that will be further explored in the optional Consolidation Phase, is that in the case of the combined retrievals, surface temperature as well as profiles of temperature, H\textsubscript{2}O, O\textsubscript{3} and CO are retrieved simultaneously. This retrieval represents an inversion of 137 parameters (there are 34 levels in each profile) to retrieve with only about 19 independent pieces of information (DOFS) available for the most favourable conditions (DOFS of 6 for temperature, 1 for surface temperature, 7 for H\textsubscript{2}O, 3.5 for O\textsubscript{3} and 1.5 for CO).

For the combined retrieval experiment, the \textit{a priori} error covariance matrix used in this study, contains the correlation elements on the diagonal blocks given by the successive error covariance matrix of the different parameters of the state vector (T, T\textsubscript{s}, H\textsubscript{2}O, O\textsubscript{3} and CO). But no cross correlation elements were included between these parameters.

Due to this lack of cross correlation elements, the \textit{a priori} error covariance matrix contains a number of independent elements that is much larger than the information content from the measurement (~19). In consequence, the measurement contribution corresponding to the transport of CO/O\textsubscript{3} errors on T/H\textsubscript{2}O retrievals might be negligible as compared with the \textit{a priori} contribution.

To verify the above analysis, two additional tests have been performed, in Sections 3.3.2.3.1 and 3.3.2.3.2 below. First, the impact of vertical resolution of CO and O\textsubscript{3} profile retrievals is tested, as this impacts the number of independent elements to be retrieved, and could modify the above discussion. Second, the effect of cross correlations is illustrated, by analysing the academic case of temperature profile and surface temperature cross correlations.

### 3.3.2.3.1 Impact of reducing vertical resolution of CO/O\textsubscript{3} retrievals

The exercise of reducing the vertical resolution of state vectors to 8 layers for O\textsubscript{3} and 4 layers for CO, instead of 34 original layers, has been done. As shown in Figure 33, no significant impact on T and H\textsubscript{2}O retrievals errors relative to the reference retrieval is found, as the differences between reference and combined retrieval remain negligible.
3.3.2.3.2 Impact of cross-correlation information

It has not been possible to address the impact of cross correlations directly between T/H₂O/Tₛ/CO/O₃, as no realistic statistics on these cross correlations is available, and due to the difficulty to build a numerically consistent set of matrices based on the available statistics on each variable.

In order to validate the discussion and conclusion on this topic, a simpler exercise has been performed, addressing the impact of cross correlations between temperature profile and surface temperature. Realistic cross correlation between vertical levels of temperature profile and surface temperature has been considered. As presented in Figure 34, the impact of cross correlations drastically modifies the results of the combined retrieval. Results when surface temperature error is transported (red curve) and when surface and profile temperatures are retrieved simultaneously (green curve) are similar, and the use of cross correlations in the combined retrieval (pink curve) strongly improves the results on profile temperature, by allowing the full use of data information for both variables. This validates the importance of considering cross correlations in combined retrievals.

![Figure 33: Fraction of total retrieved error for the reference retrieval (solid line) and combined retrieval with CO and O₃ retrieved at 8 vertical levels (dashed line).](image-url)
3.4. Conclusion and recommendations from the Sensitivity Analysis Phase

3.4.1. Summary

A first set of analyses on reference impacts explore the performances of the worst case and of the ideal case, that is, respectively, the T/H2O stand alone retrieval with CO/O3 climatological knowledge (called the reference retrieval), and the T/H2O stand alone retrieval supposing CO/O3 is perfectly known (the difference between the two is called the reference impact). This analysis shows that

- The reference impact remains weak, i.e. of one order of magnitude below the temperature and H2O error retrievals. Moreover, it is negligible for water vapour in terms of information content (DOFS). These results suggest that the impact of CO and O3 on H2O is not direct but is induced by the retrieval error correlation between H2O and temperature.

- The reference impact is analysed for 2 hypotheses on T/H2O prior knowledge: NWP Forecast and climatology: this reference impact is similar for the two a priori hypotheses, as the impact remains weak with climatological a priori.

- The reference impact is analysed with respect to the impact of surface emissivity errors: the trace gas effect is weak in the lower troposphere as compared with the impact of surface emissivity error.

- A simple mitigation strategy of the reference impact is tested, by applying channel selection to minimize CO/O3 effect and preserve T/H2O information. The effect of such a channel selection on our reference results is negligible, due to initial small impact of CO/O3 on T/H2O.
In the next step, the sequential and combined retrieval strategies have been tested. Tests on combined retrieval explore the possible gains and constraints of including trace gas retrieval in the main L2 processor. Combined retrieval information content analysis gives the expected error at convergence. Concerning reference TH2O retrieval, the risk of bias in the retrieval due to erroneous prior CO/O3 knowledge is removed, but the statistical error remains at the same level. The statistical error could be decreased by:

- Explicitly considering the O3/CO error reduction in the retrieval process (sequential retrieval);
- Considering proper cross correlation errors between TH2O and CO/O3 in background covariances

The only gain of the combined retrieval is to remove the risk of a bias, but there is no significant error reduction when cross correlations are not used. On the other hand, proper cross correlations are not available, and retrieval performances will be very sensitive to their definition. Convergence problems will increase the difficulties for practical implementation.

The alternative strategy of two stand alone retrievals is addressed through sequential retrieval information content analysis. Such approach exploits the information in the spectrum, and reduces the statistical error on the TH2O retrieval. This approach offers a good compromise: a way for reducing/removing the impact of CO/O3 on the TH2O retrieval without the need of having proper cross correlations between a priori errors on TH2O and O3/CO. In addition, the sequential approach will provide a retrieval of O3 and CO.

In this phase the combined impact of CO and O3 was considered. In the optional phase, we will perform a qualitative analysis of the useful spectral bands used for T and H2O retrievals and the resulting contaminating effect of CO and O3 (done separately).

3.4.2. Conclusions on the sequential and combined methods

Below we summarise the conclusions on the advantages and drawback of the two retrieval methods. These conclusions are drawn for the case when the current state-of-the-art error covariance matrices for CO/O3 a priori information are used. These matrices are based on climatologies, and do not contain the cross correlations terms.

**Combined retrieval of TH2O/O3:**

- **Advantages**
  - Best method (in theory) to avoid bias linked to trace gases on TH2O retrievals
  - Faster than sequential method (only one iterative retrieval process)
  - Can use a priori error covariances between TH2O and O3 if they are available.

- **Draw backs**
  - Contrary to the sequential method, TH2O stochastic errors due to O3 are driven by a priori O3 error level.
  - Not very robust because of convergence problems (unmatched convergence or bad solution) due to:
    - TH3 retrieval ambiguities because of TH3 retrieval correlations. The larger is the O3 a priori error, the bigger is this ambiguity.
    - Convergence speed discrepancies between TH2O and O3 retrievals. These discrepancies increase for large O3 a priori error.
    - The two above points indicate that it is necessary to have good O3 a priori knowledge, better than climatology, to full exploit the potential of the combined retrieval.

**Sequential retrieval of O3 (first step) and then TH2O (second step):**

- **Advantages**
  - Best method, in terms of convergence, to get robust retrieval of TH2O.
Draw backs

- Slower method than the combined retrieval (need at least 2 retrieval process, first O3 than T/H2O)
- Can induce bias on T/H2O retrievals
  - bias on T a priori profile can induce bias on O3 first step retrieval
  - if bias present on the retrieved O3 , it could in turn impact the T/H2O retrievals
  - need a complete iterative process to avoid bias or need good first guess T/H2O profiles
- Can lead to incestuous behaviour if the same channels are used for both T/H2O and O3 retrievals
  - need a specific channel or band selection for both retrieval steps

3.4.3. Recommendation

The main conclusions from our sensitivity analysis are that the impact of CO and O3 knowledge errors on T/H2O retrieval performances is weak. In particular, the impact is smaller than other effects (surface emissivity, choice of T/H2O a priori), and it could be mitigated by proper channel selections, without significant loss of information on T and H2O.

However, the inclusion of trace gas retrievals, within or before the main retrieval should be considered, to ensure a constant quality of the derived moisture and temperature information especially in cases of large pollution events. If trace gas retrievals were ignored, a bias in the moisture and temperature profiles for heavy polluted areas might be present (see also section 4).

As pointed out above, both of the analysed retrieval strategies that include trace gases have advantages and drawbacks. Although the combined retrieval is in theory a potentially powerful method, given the current state of the a priori information on CO/O3 (based on climatologies, no cross correlation terms) we recommend the sequential approach because:

- it is the most robust method that works well with the available climatological a priori O3 profiles,
- it allows mitigating stochastic errors due to O3,
- it does not require the presence of highly variable, case dependant, information on correlations between T/H2O and O3. The high dependence of these terms on specific meteorological and chemical conditions limits the interest of the simultaneous method with respect to the sequential approach.

For the sequential approach, we recommend to use a first guess retrieval of T/H2O method (e.g. EOF retrieval method) in the first step O3 retrieval to limit the impact of the bias mentioned in the list of drawbacks in section 3.4.2. Moreover, we recommend selecting 2 different spectral bands (or use of channel selections) for O3 and T/H2O retrievals. The impact of channel selection will be demonstrated in the optional Consolidation Phase.

Even though the sequential method is recommended here, we believe it is also necessary to further explore the possibility to characterize and to use appropriate error covariance matrices including cross correlation terms between trace gases and temperature and water vapour. This should be done in order to fully exploit the potential combined approach that could represent a powerful method if better O3 a priori knowledge (or O3 first guess retrieval) is available too.
4. Evaluation Phase

This section contains the results from the Evaluation Phase. The sequential retrieval method that was recommended as a result of the analysis conducted during the previous phase (Sensitivity Analysis Phase) is tested for a high number of IRS/MTG spectra. In the sequential method, the trace gases (CO and O3) are inverted in a first step inversion, and the obtained information is used in the inversion of temperature and H2O. In addition to applying this inversion scheme, we test also with the help of different scenarios the pertinence of inverting O3 in the first step inversion. This section contains the results of the inversions as well as an update of the recommendations for the retrieval of temperature and H2O in the presence of trace gases.

This chapter is organised as follows. Section 4.1 contains a description of the set-up of the simulations which includes the description of the inversion tool used, and a description of the geophysical scenarios and error covariance matrices. Section 4.2 contains a description of the methodology, including a detailed list of scenarios tested, used to address the T and H2O inversion approach in the presence of trace gases.

In section 4.3 we present the results. These results are now consolidated.

In section 4.4 we give the updated recommendations for the inversion of temperature and H2O in the presence of trace gases.

4.1. Set-up of the inversion simulations for the sensitivity analysis phase

4.1.1. Generation of proxy data

In section 2 of the current document NOVELTIS has outlined specifications for the generation of an optimal MTG/IRS proxy dataset for testing in Phase III the impact of trace gases on the retrieval of temperature and H2O.

The specifications for proxy data concerned three major phases:

1. the simulation of the geophysical state with surface and atmospheric variables;
2. the simulation of the MTG-IRS spectra with a radiative transfer model and
3. the specification of the statistical errors (in the form of error covariance matrices) of the surface and atmospheric variables as well as the \textit{a priori} and first guess for the variables to be retrieved.

The generation of the proxy dataset following these specifications was to be undertaken by EUMETSAT. Given that generation of the dataset with a description of the geophysical state would require a substantial modelling effort and also due to the fact that no sufficient interest was manifested by the remaining MIST members, a decision was made not to generate this data at the current stage. In order to provide the input proxy data for Phase III of the study, NOVELTIS has generated the dataset based on existing modelling results. In the generation of the proxy data we followed the recommendations from points 2 and 3 above; however, for point 1, we used an existing data set which does not meet the proposed specifications. This atmospheric dataset is composed of results from two chemistry transport models: TM5 and CHIMERE. The results from TM5 and CHIMERE were provided by KNMI and were generated in the framework of a project funded by the CNES (CNES contract 4500028777/DCT090 from 10.06.2009).
4.1.1.1. Generation of the database describing the surface-air system

The proposed atmospheric-surface database contains two types of variables:

- meteorological variables;
- chemical variables simulated by KNMI (Henk ESKES) with two different chemical transport models.

For the chemical variables, the regional model CHIMERE is proposed for the lowermost troposphere and the global TM5 model for the middle and upper troposphere and the stratosphere. We propose two different models because we believe that it is important to have:

1) high horizontal resolution of a regional model in the boundary layer that allows reproducing strong and localised pollution events near source regions that are of interest in this study;
2) vertical profiles covering the entire troposphere and stratosphere (the CHIMERE model covers only the lowest part of the troposphere).

Both models were forced by the meteorological fields from the ECMWF analysis assuring the consistency between the chemical and meteorological variables. However, as the forcing fields come from a global model, the high resolution dynamics are not simulated (the ECMWF fields are interpolated to the high resolution grid of CHIMERE).

In the section below, we give a brief description of the two chemical models.

4.1.1.1.1 Description of existing datasets

4.1.1.1.1.1 CHIMERE dataset

The results from the standard version of CHIMERE with 8 vertical levels from the surface to 500 hPa were provided by KNMI. This version is commonly used for air-quality applications in which high vertical resolution near the surface is required in order to simulate the impact of surface emissions on air quality. The results used in this project were generated by a version with a zoom that allowed simulating the chemical concentrations at the resolution of:

- 0.25° in longitude which corresponds to approximately \( dx = 17.9 \text{ km} \) (at 50°N);
- 0.125° in latitude which corresponds to \( dy = 13.9 \text{ km} \).

The zoom covered a region between 2°W - 14°E and 46°N – 56°N.

The results available for this study are:

- 2 periods of three days in July 2006;
- 2 periods of three days in January 2008.

The model CHIMERE was forced by ECMWF analysis (MARS) that was interpolated in space and time to CHIMERE’s spatial and temporal grid.

The surface emissions were taken from high resolution emission inventory files with resolutions that are higher than the resolution of the model.

The lateral boundary conditions to the zoom were taken from the coarse resolution simulations of CHIMERE. The top conditions as well as the lateral boundary conditions for the coarse domain were taken from a global climatological simulation.

4.1.1.1.1.2 TM5

The global chemistry-transport model TM5 was used to complete the CHIMERE profiles. The version used contains 34 vertical levels that cover a region from the surface to 0.1 hPa (about 65 km). The results used in this study were extracted from a zoom over Europe that allows simulating the concentrations of chemical tracers at resolution of 1° x 1°.

© Noveltis 2010
This document is the property of Noveltis, no part of it shall be reproduced or transmitted without the express prior written authorisation of Noveltis
As was the case for the CHIMERE model, the TM5 model is an offline model that uses pre-processed meteorological fields from ECMWF analyses interpolated to the TM5 grid. The ECMWF data were read every 6 hours and were linearly interpolated for the periods between two analyses.

### 4.1.1.1.2 Vertical levels in the merged atmospheric database

The atmospheric dataset used in this study represents a merge of results from the two models described above. Table 14 contains a list of the selected levels which includes their approximate half-level pressures and their origin (CHIMERE or TM5).

#### Table 14: List of the vertical levels in the merged dataset

<table>
<thead>
<tr>
<th>Level Number</th>
<th>Model</th>
<th>Pressure (hPa)</th>
<th>Level Number</th>
<th>Model</th>
<th>Pressure (hPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CHIMERE</td>
<td>1008.74</td>
<td>19</td>
<td>TM5</td>
<td>158.67</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1001.85</td>
<td>20</td>
<td></td>
<td>139.12</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>990.06</td>
<td>21</td>
<td></td>
<td>121.79</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>969.94</td>
<td>22</td>
<td></td>
<td>106.33</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>935.64</td>
<td>23</td>
<td></td>
<td>92.36</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>886.30</td>
<td>24</td>
<td></td>
<td>79.56</td>
</tr>
<tr>
<td>7</td>
<td>TMS</td>
<td>824.12</td>
<td>25</td>
<td></td>
<td>67.71</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>750.52</td>
<td>26</td>
<td></td>
<td>54.32</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>669.21</td>
<td>27</td>
<td></td>
<td>40.11</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>583.58</td>
<td>28</td>
<td></td>
<td>28.25</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>498.05</td>
<td>29</td>
<td></td>
<td>18.78</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>418.62</td>
<td>30</td>
<td></td>
<td>11.61</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>349.50</td>
<td>31</td>
<td></td>
<td>6.54</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>299.31</td>
<td>32</td>
<td></td>
<td>3.28</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>264.42</td>
<td>33</td>
<td></td>
<td>1.40</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>233.25</td>
<td>34</td>
<td></td>
<td>0.49</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>205.45</td>
<td>35</td>
<td></td>
<td>0.11</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>180.69</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.1.1.1.3 Variables present in the merged dataset
List of variables describing the atmosphere-surface system is given in Table 15.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Levels 1 6</th>
<th>Levels 7 34</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmospheric temperature profile</td>
<td>ECMWF analysis (MARS, every 3 hours) interpolated in space and time to CHIMERE’s spatial and temporal grid</td>
<td>ECMWF analysis (every 6 hours) interpolated on the 1°x1° grid and on the vertical levels of TM5</td>
</tr>
<tr>
<td>Atmospheric H₂O profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmospheric CO profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmospheric O₃ profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmospheric CH₄ profile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atmospheric CO₂ profile</td>
<td>380 ppmv (vertically constant)</td>
<td></td>
</tr>
<tr>
<td>Atmospheric N₂O profile</td>
<td>Climatological profile ([RD10]) with tropospheric values updated to current day values (320 ppbv)</td>
<td></td>
</tr>
<tr>
<td>Cloud flag</td>
<td>Assumed clear-sky</td>
<td></td>
</tr>
</tbody>
</table>

Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Temperature</td>
<td>ECMWF analysis (MARS) interpolated in space and time to CHIMERE’s spatial and temporal grid</td>
</tr>
<tr>
<td>Surface Pressure</td>
<td></td>
</tr>
<tr>
<td>Surface Emissivity</td>
<td>Snyder emissivity ([RD11])</td>
</tr>
</tbody>
</table>

4.1.1.1.4 Differences between the existing datasets and specifications

4.1.1.1.4.1 Horizontal resolution
One of the main differences between the available dataset and the specifications given in section 2 concerns the horizontal resolution. For the simulation of chemical tracers the resolution of about 15 km is used in the lowermost troposphere (CHIMERE), and of about 100 km for the higher levels (TM5). For the meteorological fields, results from global model are available (ECMWF) which implies that local meteorological effects are not simulated. Furthermore, strong correlations that could be present between the local meteorological features and anomalies in chemical tracer concentrations (e.g.: local temperature inversions) are not present in the proposed database.

4.1.1.1.4.2 Interface between the two chemical models
At the interface of the two chemical models there is often a discontinuity in the concentrations of the chemical tracers between the highest selected layer in CHIMERE and the lowest selected layer in TM5. For the dynamical variables, the discontinuity is much smaller as both models use ECMWF fields. However, due to interpolation problems, a small discontinuity can be present also in the dynamical variables.
The discontinuity in the fields of CO and O$_3$ needs to be corrected as having unrealistic profiles might affect the inversions. Figure 35 shows the typical case of discontinuity in the CO profiles. TM5 CO fields are generally 10 to 20 ppbv higher at the interface between the two models. In highly polluted regions, in the case when pollution is present in the CHIMERE model near the interface (Figure 36), the CO in CHIMERE is higher than the values simulated by TM5. This case represents a realistic profile with highly polluted boundary layer. In order to reduce the discontinuity presented in Figure 35, the merged profile is linearly interpolated for levels 6 (last CHIMERE level) and 7 (first TM5 level used) by using values from levels 5 and 8. For the pollution case presented in Figure 36, the CO profile is not modified as it represents a realistic boundary layer pollution event.

For O$_3$, it is generally CHIMERE that has concentrations that are higher at the interface by about a few ppbv. This feature can correspond to a build of O$_3$ in the boundary layer as a result of pollution. This build-up is also present in the boundary layer of the TM5 but is on average weaker. A particularly strong discontinuity is shown in Figure 37. To avoid such strong discontinuities, O$_3$ profiles are corrected using linear interpolation.

![Figure 35: An example of a discontinuity in the CO profile at the interface of the two models. Black profile: uncorrected data, red profile: after application of the correction.](image)

© Noveltis 2010

This document is the property of Noveltis, no part of it shall be reproduced or transmitted without the express prior written authorisation of Noveltis
Figure 36: An example of a discontinuity in the CO profile at the interface of the two models. The profile is not corrected in this case.

Figure 37: An example of a discontinuity in the O₃ profile at the interface of the two models. Black profile: uncorrected data, red profile: after application of the correction.
4.1.1.1.5 Temporal resolution
The output from the CHIMERE and TM5 models are available with the frequency of one hour. The meteorological variables are also available with the same frequency; however they represent linear interpolations from ECMWF fields stored every 3 or 6 hours (see Table 15).

4.1.1.2. Selection of geophysical states
In order to test the sequential approach for the retrieval of temperature and H₂O, we have selected several atmospheric scenarios with marked departures from the climatological profiles of CO and O₃. The case that was chosen for the inversion concerned a strong perturbation of O₃ in the UTLS region that was due to a pronounced lowering of the tropopause over the North Sea and neighbouring areas on the 7th of January 2008 at 21:00 UTC. The plots showing the horizontal distribution of O₃, H₂O and T in the UTLS are shown in Figure 38. The black line in Figure 38 corresponds to the selected cross section which is located at 4°E and between 46.9375°N and 55.9375°N. The cross section contains 73 modelling profiles. The mixing ratios of O₃ for this cross section, shown in Figure 39, indicate the presence of a strong perturbation that increases considerably the O₃ levels in the UTLS region. Table 16 below summarizes the information on the selected profiles.

Table 16: Summary of the selected profiles

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description of the profiles selected for the evaluation phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of selected profiles</td>
<td>73</td>
</tr>
<tr>
<td>Latitude</td>
<td>Between 46.9375°N and 55.9375°N</td>
</tr>
<tr>
<td>Longitude</td>
<td>4°E</td>
</tr>
<tr>
<td>Date</td>
<td>January 7th, 2008, 21:00 UTC</td>
</tr>
</tbody>
</table>

© Noveltis 2010
This document is the property of Noveltis, no part of it shall be reproduced or transmitted without the express prior written authorisation of Noveltis
4.2. Methodology

4.2.1. Testing the impact of O₃ on the retrieval of temperature and H₂O

In order to test the impact of trace gases on the retrieval of temperature and water vapour, we choose to focus on O₃ for which the impact is stronger than for CO. We evaluate the performance of the sequential method for the retrieval of the profiles of temperature, H₂O and also O₃. Other variables, which are retrieved or not, are set to their true states (the states used in the forward model used to calculate the spectra) in order to isolate the impact of O₃. Nevertheless, realistic errors, described in section 4.2.2 are set on all parameters that are taken into account.

The inversion tests are done for a set of 73 spectra obtained for a geophysical situation described in the previous section. The spectra are calculated with specifications listed in section 3.2.3. This case is characterised by a strong perturbation of ozone in the upper troposphere and lower stratosphere as a result of the folding of the tropopause.
4.2.2. Description of scenarios

In order to quantify the effect of the perturbations in the distribution of O_3 on the retrieval of temperature and water vapour, we have defined several scenarios. These scenarios allow us to answer several questions:

1) what is the precision in the retrieval of T and H_2O in the case the O_3 profiles are known perfectly (first guess is equal to true state and error covariance matrix for O_3 is set to zero)? This scenario, which is called here the "reference", is used as the reference case against the results from the remaining scenarios are evaluated against;

2) what is the precision in the retrieval of T and H_2O in the case we use the a priori O_3 profiles that are different from the true state and we neglect to transport the errors in the O_3 error covariance matrix (or in other words, we set this matrix to zero)? This case, called the "impact" scenario, demonstrates the need to define an appropriate error covariance matrix for ozone.

3) what is the precision in the retrieval of T and H_2O in the case we use the a priori O_3 profiles that are different from the true state and we now transport the O_3 errors as specified in the climatological O_3 error covariance matrix? This case is called the "default" scenario.

4) Finally, in the last case we analyse the impact of using O_3 (a priori and error covariance matrix) not from the climatology but from the first step inversion in the sequential approach. Comparing this case, called "sequential" scenario, with the default case allows us to quantify the benefit of using the first step inversion.

In all of the above retrieval scenarios, temperature and H_2O profiles are retrieved simultaneously. Temperature and H_2O a priori profiles are taken as the mean profiles of the 73 selected atmospheric situations. This is done in order to avoid having biases from the a priori in the statistical analysis presented later in section 4.3. As a priori errors on T and H_2O, we use the same forecast ECMWF error covariances matrices that were used in the sensitivity analysis phase. To justify the choice of the mean profile as a priori profile, we verified that the standard deviations (STD) of the temperature and H_2O profiles of the 73 atmospheric situations were at the same level that the ECMWF forecast errors.

The definition of the four scenarios is summarised in the table below.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>T/H_2O</th>
<th>O_3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A priori / First Guess</td>
<td>Error matrix</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>Mean profile of the time series</td>
<td>ECMWF forecast</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>Mean profile of the time series</td>
<td>ECMWF forecast</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>Mean profile of the time series</td>
<td>ECMWF forecast</td>
</tr>
<tr>
<td>Scenario 4</td>
<td>Mean profile of the time series</td>
<td>ECMWF forecast</td>
</tr>
</tbody>
</table>

In order to isolate the impact of ozone, surface parameters and CO and CO_2 a priori (or first guess) profiles are taken as true state with a climatological error for CO and CO_2 errors of 1K and 1% for surface temperature and emissivity, respectively. These errors correspond to the expected error level from first prior surface parameter retrieval. The table below summarises the chosen parameterisation.
4.3. Results

This section contains the retrieval results obtained for T, H₂O and O₃ for the 4 scenarios defined in section 4.2.2. The retrievals were done for the 73 simulated spectra with MTG/IRS characteristics that characterise the atmospheric emissions for the strong tropopause folding event described in section 4.1.1.2.

4.3.1. Analysis of the retrieval experiment results for temperature

We begin this section by presenting in Figure 40 the vertical cross sections corresponding to the North-South transect (containing 73 profiles) shown as black line in Figure 38. The results shown in the figure were obtained for the default retrieval scenario (scenario 3, see section 4.2.2).

We precise that results obtained for the “sequential” retrieval scenario are very similar and do not present significant differences from the default scenario. A comparison of the results obtained from the different scenarios is shown in Figure 42.

The cross sections on the left side of the Figure 40 represent the true state (upper figure), a priori used (middle figure) and results of the retrieval (lower figure). Rapid analysis of these 3 figures indicates a considerable improvement in the representation of spatial patterns of temperature in the results from the retrieval as compared with the a priori. The figures on the right quantify this improvement. It shows that effectively the biases present in the a priori were considerably reduced and that the retrieval errors below 400 mb are less that 1 K. Larger departures (~ 2-3 K) are present in the upper troposphere indicating that the potential of the MTG/IRS measurement on improvement of T in this region is lower. In addition, this region around the tropopause is characterised by sharp vertical temperature gradients which are not well retrieved. The errors are highest at the most southern and northern latitudes where the departures of the a priori (taken as the mean temperature profile) from the true state were highest.

Statistical results, based on the results obtained from the 73 retrievals for the default retrieval scenario, in terms of mean bias and standard deviation (STD) are shown in Figure 41. These results can be summarised with the following points:

- Mean a priori bias is extremely weak (close to zero). This was expected as the a priori profile was taken as the mean of the 73 true profiles in the database obtained from modelling simulations;
- Mean retrieval bias is negative and small (less than 0.2 K) below 300 hPa;
- The retrieval bias is larger (more negative) in the 150-200 hPa region where it reaches 0.35 K. As mentioned above, this larger bias can be explained by the lower sensitivity of the MTG-IRS spectra in this region and also on the difficulty to retrieve the high vertical temperature gradients;
- STD of the retrievals is less than 0.5 K between 950 and 500 hPa;
- Significant reduction of the STD in the range [900-400 hPa] and no reduction near the surface and above 300hPa. This behaviour is coherent with the theoretical error estimation obtained in phase II, giving the small a priori effective error in the lower and the upper troposphere for that atmospheric case.
Figure 40: Results for temperature profile retrievals (in Kelvins) for the 73 spectra, as a function of the latitude and of the pressure level. On the left side: true state given by the atmospheric models (top), \textit{a priori} knowledge given by climatology (middle) and retrieved profiles (bottom). On the right side: \textit{a priori} departure from true state (top), \textit{a priori} departure from retrieved state (middle) and retrieval departure from true state (bottom).

Results obtained for the default scenario.
Figure 41: Bias (left) and standard deviation (right) of the retrieved temperature profiles (in Kelvins) obtained with scenario 3. A priori (red) and retrieval (black).

Figure 42: Mean retrieval bias (left) and standard deviation (right) of the retrieved temperature profiles (in Kelvins) obtained for all scenarios. A priori (red), Scenario 1 = reference (black), scenario 2 = impact (orange), scenario 3 = default (green) and scenario 4 = sequential (violet).
The results shown in Figure 40 and Figure 41 were obtained for the “default” scenario. An important result of this study is the comparison of the results of the four defined scenarios that allows quantifying the relative performance of the proposed sequential retrieval approach. This intercomparison is done in Figure 42 which shows mean bias and STD for the four scenarios and the a priori.

The left panel of the Figure 42 indicates that, with the exception of the “impact” scenario (scenario 2), results are characterised by small mean bias on the retrieved temperature profiles. For the “impact” scenario, for which bias on O₃ profile was introduced in the a priori without transporting the error, bad convergence and unrealistic retrieval were obtained. This case demonstrates the importance of including O₃ error covariance matrices in the retrieval of temperature. Concerning the mean biases for scenarios 1, 3 and 4, the following remarks can be made:

- Bias for the “reference” scenario (N°1) is negligible in the lower troposphere and ± 1.5 K in the upper troposphere and stratosphere. These results, which were expected to be better, might be improved with optimization of the method, by removing in the observation vectors the O₃ channels that introduce strong correlation between stratosphere and troposphere.
- Bias for the “sequential” scenario (4) are globally not better than for the “default” scenario (3). For instance, bias is improved in higher troposphere but seems to be larger in the lower troposphere.

The STD results for scenarios 1, 3 and 4 shown in the right panel are similar. For the “impact” scenario (2), the error is higher than for the a priori case indicating that in the current set up in which ozone channels are used it is necessary to include the O₃ error transport.

### 4.3.2. Analysis of the retrieval experiment results for H₂O

For water vapour, the analysis of the results is conducted with the help of figures that are similar to those used for temperature. To take into account the large decrease of H₂O with altitude, the results for H₂O are shown in percentage when judged appropriate.

Figure 43 presents the H₂O cross sections with true state, a priori, and retrieved profiles for the 73 selected atmospheric spectra. As was the case in section 4.3.1, the results presented are for the “default” scenario (3). The results are generally similar to those obtained for temperature. In particular:

- The spatial structure of the retrieved profiles is much closer to the true state than the a priori (left panels);
- Below 400hPa, the error of the retrieval of H₂O fields is small with departures of less than 20%.

Above 400hPa the error can be higher. As for temperature, this increase in the retrieval error is due to the decreasing sensitivity of the MTG/IRS measurements. The errors are highest at the most northern latitudes where the departure of the a priori (taken as the mean profile) from the true state is highest.

Results obtained in terms of mean retrieval bias of H₂O and STD are shown in Figure 44. Again the results are similar to those obtained for temperature:

- Mean retrieved H₂O bias is small (less than ± 5% below 300 hPa and about 10% at 200 hPa);
- strong reduction of STD with respect to a priori between 850 and 300 hPa;
- STD is less than 10% of the mean profile in the range [850-400] hPa.
Figure 43: Results for water vapour profile retrievals for the 73 spectra, as a function of the latitude and of the pressure level. On the left side (in VMR): true state given by the model (top), a priori knowledge given by climatology (middle) and retrieved profiles (bottom). On the right side (in percent): A priori departure from true state (top), A priori departure from retrieved state (middle) and retrieval departure from true state (bottom). Case of scenario 3.
Figure 44: Mean H2O retrieval bias (left) and standard deviation (right) of the retrieved H2O profiles (in percent) obtained for scenario 3. A priori (red) and retrieval (black).

Figure 45 shows bias and STD obtained with each retrieval scenarios and compared to a priori departures. Results are globally similar to those obtained for temperature retrievals:

- All scenarios give good results at the exception of the “impact” scenario (N°2);
- Small differences on STD obtained with scenarios 1, 3 and 4;
- Mean relative retrieval bias for scenarios 1, 3 and 4 are very small (< 5%) below 350 hPa but reach 20% in the higher troposphere for scenarios 1 and 4.
- Mean relative retrieval bias for the “sequential” scenario (N°4) generally not better than for the “default” scenario (N°3) below 400 hPa.

Above 400 hPa, scenario 3 shows lower mean relative bias in the higher troposphere. That could be explained by the fact that departure on the H2O retrieval is not a direct impact of O3 bias (not present in scenario 1) but is due to error correlation between temperature and H2O retrievals that should be smaller for the scenario 3.
4.3.3. Analysis of the retrieval experiment results for Ozone

Results on temperature and H$_2$O retrievals show that the sequential retrieval approach does not really improve the performances with respect to the “default” retrieval scenario. This is the case despite the fact that incertitude on O$_3$ profile, which does have an impact on temperature and H$_2$O retrievals, is lower for the sequential approach. To provide first elements to explain these somewhat unexpected results, we analyse in this section the performance of the O$_3$ retrievals.

4.3.3.1. Information content analysis on O$_3$ retrieval

We begin the presentation of the results for O$_3$ by Figure 46 and Figure 47 showing the O$_3$ estimated errors and averaging kernels obtained with the information content analysis tool. This method, already used in Phase II of this study, allows estimating the information content given by the DOFS and the theoretical error reduction based on:

4) the errors on the prior knowledge of the state vector;

5) the errors of the measurements and

6) the calculated jacobians.

The results shown for O$_3$ in this section were obtained by applying this tool to one of the 73 MTG-IRS synthetic spectra:

- The Degree of Freedom for Signal (DOFS) is about 2.8, which means that nearly 3 pieces of independent information are available in the vertical;

- Figure 46, which presents the obtained error budget, indicates that there is a considerable error reduction (less than 20% for the total error) in the upper troposphere and the stratosphere. This result illustrates the good potential of MTG-IRS for O$_3$ retrieval;

- In the lower troposphere, this potential is reduced;

- As indicated by the Averaging kernels shown in Figure 47, measurements introduce high positive and/or negative correlations between different vertical levels and in particular between the troposphere and stratosphere. Such correlations are due to the poor spectral resolution of MTG-IRS with respect to the thickness of the O$_3$ absorption lines. For example, the information in the lower troposphere corresponds mainly to the transport of information from middle troposphere and from the stratosphere.

It could be interesting to apply a channel selection by removing all channels with strong correlation between stratosphere and troposphere, in order to separate as much as possible tropospheric and stratospheric information.
Figure 46: Estimated theoretical errors on O3 retrieved profile obtained for one of the 73 spectra. *A priori* climatological error (black), total retrieved error including smoothing error (red), retrieved error due to the measurement error (green), retrieved error due to model parameter errors (blue).

Figure 47: Averaging kernels obtained for the inversion of one of the 73 spectra.
4.3.3.2. Retrieval of O₃ profiles

The equivalent of Figure 40 for T and Figure 43 for H₂O is shown for O₃ in Figure 48 below. The results are obtained for the “sequential” retrieval scenario (N°4). The retrieved profiles represent an improvement as compared with the a priori; however, retrieval clearly has difficulties with the retrieval of structures present at smaller vertical scales. Nevertheless, the departures from the true state are considerably reduced as shown in more detail by right panels. Departures of the retrieved profiles are smaller than 15% in the middle troposphere (800-400 hPa) and generally smaller than 30% in the higher troposphere and in the boundary layer.

Figure 49 shows the mean retrieval bias, STD and root mean squares (RMS) obtained for O₃ retrievals in comparison to the estimated errors on O₃ retrieved profile (see Figure 46). The following results are obtained:

- For this scenario in which the climatological ozone is used as a priori, there is now a mean bias on the a priori O₃ profiles. This bias is generally reduced with the retrieval. However the vertical structures which are present in the true state in the tropopause region, and which are not present in the climatological a priories, are not retrieved correctly. This induces a mean bias of up to 30% as it can be seen in higher troposphere and lower stratosphere;
- Good reduction of STD in the upper troposphere that decreases from 20% to less than 5%;
- On the other hand, STD increases (up to 10%) in the lower troposphere. Such an effect can be explained by the fact that information in the lower troposphere comes from the transport of information from higher levels (see section 4.3.3.1) accompanied by error amplification.
- Reduction of RMS of about 50% for the whole vertical profile except in the UTLS region where the retrieval bias is still very large due to the poor vertical resolution of O₃ retrieval.
Figure 48: Results for O3 profile retrievals for the 73 spectra, as a function of the latitude and of the pressure level. On the left side (in ppbv): true state given by the model (top), a priori knowledge given by climatology (middle) and retrieved profiles (bottom). On the right side (in percent): A priori departure from true state (top), A priori departure from retrieved state (middle) and retrieval departure from true state (bottom). Case of first step of scenario 4.
Figure 49: Bias (top left), standard deviation (top right) and RMS (bottom) of the retrieved (black) and a priori (red) O3 profiles (in percent). Results from the sequential retrieval scenario.

It is also interesting to compare the obtained effective errors (Figure 49, RMS panel) with the estimated total error shown in Figure 46. This comparison shows that the theoretical O3 errors used as a priori in the second step sequential retrieval of T/H2O generally represent well the effective, statistical O3 errors found in the retrieval. This point can be summarised as follows:

- good agreement in the upper troposphere indicating a good characterization of the a priori;
- overestimation of the theoretical error with respect to the a priori error in the lower troposphere for the atmospheric scenario considered here.

However, biases in the stratosphere and high troposphere O3 retrieval, together with error correlations between this vertical region and the lower troposphere, could be a problem. This could impact the T/H2O retrieval in the second step sequential approach as can be seen in the lower troposphere bias for scenario 4 in Figure 42.
4.4. Conclusion and recommendations from the Evaluation Phase

4.4.1. Summary and conclusions

In this section we have analysed the impact of O₃ on the retrieval of temperature and water vapour for an atmospheric scenario containing a strong O₃ perturbation in the UTLS region caused by a tropopause folding event. We have chosen to focus on the O₃ perturbations as the impact of O₃ is likely to be stronger than that of CO.

To better quantify the impact of the O₃ perturbation on T and H₂O retrievals, we have considered four different retrieval scenarios. Channel selection was not applied in the retrieval of T and H₂O in these scenarios. Scenario 1 (perfect knowledge of O₃ and no bias) is the reference scenario against which other scenarios are compared. Scenario 2 contains no O₃ error covariance matrix during the retrieval of T and H₂O. Scenarios 3 (climatological O₃) and 4 (O₃ from first step sequential inversion) allow evaluating the potential benefit of the sequential approach.

For the scenarios 1, 3 and 4 and below about 300 hPa the mean retrieval bias for temperature and water vapour is small. Above 300 hPa the bias increases and it reflects the lower sensitivity of the MTG-IRS spectra in this region and also the difficulty to retrieve the high vertical gradients present near the tropopause region. The retrieval error is also considerably smaller than the a priori error for both T and H₂O in the 900-400 hPa region. The reduction is much weaker near the surface and above 300hPa. Scenario 2, for which the results are clearly unsatisfactory, shows the importance of including an adapted O₃ error covariance matrix.

The theoretical analysis presented in phase II predicted that the impact of O₃ for scenarios 3 and 4 should be weak; however, the analysis indicated as well that the error reduction should be stronger for the sequential case. The results obtained in the current evaluation phase confirm a small impact of O₃ for both retrieval scenarios (3 and 4), however, contrary to what was expected, without significant improvement for the sequential retrieval. This is the case despite the relatively good performances of the O₃ retrieval.

4.4.2. Recommendations for future work

Even though the results obtained in this study allow us to formulate first conclusions on the impact of trace gases on the retrieval of T/H₂O, as well as to give recommendations for implementation of the level 2 processor for an optimal mitigation of trace gases impact on temperature and water vapour retrievals, some further work is still necessary in order to consolidate these results.

To provide firm conclusions, we propose to analyse the following points:

- Verification of the behaviour of the sequential retrieval process and of its convergence to the solution;
- Analysis of the sensitivity to the error covariance matrix and the impact of vertical correlations. Does the poor vertical resolution of the retrieved O₃ profile, as well as resulting strong posterior vertical correlations, impact the T and H₂O retrievals?
- Analysis of the impact of the use of channel selections.
5. Final conclusions and recommendations

Our results show that if the a priori error on CO and O3 is properly taken into account, the impact of CO and O3 on T/H2O retrieval performance from MTG/IRS spectra is weak. In particular, the impact is smaller than that of other effects such as errors on surface emissivity and choice of T/H2O a priori. However, in case of underestimating of the CO and O3 errors in the retrieval scheme, the impact of O3/CO knowledge error on the quality of the retrieval can be considerable. Appropriate characterisation of the O3/CO a priori errors, and their proper use in the retrieval scheme, is thus a key point for the overall quality of the T/H2O retrieval.

Moreover, the inclusion of trace gas retrievals, within or before the main retrieval should be considered, in order to ensure a constant quality of the derived moisture and temperature information. This is true for the current state-of-the-art estimates of CO/O3 a priori information which are based on climatologies. If trace gas retrievals were not performed, a bias in the moisture and temperature profiles might be present for cases with large departures from climatological tracer profiles, departures that are not well represented by the a priori error statistics.

Two methods are proposed for the retrieval of CO/O3 in addition to T/H2O: the combined and sequential approaches. With the current characterisation of the a priori matrices, we recommend the sequential approach. This method allows mitigating the stochastic errors due to trace gases, while in the combined retrieval, this mitigation is not achieved without a priori information on the cross correlation terms between CO/O3 and T/H2O. The characterisation of these terms is not consolidated in the current state-of-the-art information on the a priori error covariance matrices. Future work on the characterisation of the error covariance matrices, which include specification of the cross correlation terms, should increase the interest of using the combined method.

Results obtained in Phase III with the sequential retrieval for retrievals from synthetic spectra (in contrast to the information content analysis performed in Phase II) do not show expected mitigation of the O3 impact, with respect to the “default” scenario (only transport of the climatological O3 error). We presume that vertical correlations in the O3 retrieval between stratosphere and troposphere, which is due to the poor vertical resolution of O3 retrieval, can impact temperature and H2O retrievals. To verify this, we recommend in future consolidation work, investigating the sensitivity of the T/H2O retrievals to the posterior O3 error covariance matrix as well as to adequate channel selection.
6. Acknowledgements

We would like to thank Henk Eskes from KNMI for providing the output from TM5 model simulations. These data were generated in the framework of a project funded by CNES contract 4500028777/DCT090 from 10.06.2009.