Capabilities of infrared sounder observations for monitoring atmospheric composition and chemistry applications

C. Clerbaux¹,², P.-F. Coheur², S. Turquety¹ and J. Hadji-Lazaro¹

1. Service d’Aéronomie/CNRS, Université Paris 6, BP102, 4 Place Jussieu, 75252 Paris Cedex 05, France

2. Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles, CPI 160/09, 50 Av. F.D. Roosevelt, 1050 Bruxelles, Belgium

ccl@aero.jussieu.fr
pfcoheur@ulb.ac.be
stu@aero.jussieu.fr
jha@aero.jussieu.fr

Final version - August 2003
0. Executive Summary
1. Introduction
  1.1. Motivation of this study
  1.2. Instrument specifications
  1.3. Product requirements
2. Present and planned satellite missions to sound the troposphere
  2.1. Tropospheric sounders
  2.2. Thermal infrared sounders from a polar orbit
    2.2.1. IMG
    2.2.2. MOPITT
    2.2.3. AIRS
    2.2.4. IASI
    2.2.5. TES
  2.3. Improved sounders from a geostationary or drifting orbit
3. Analysis of infrared atmospheric spectra: general approach and methodology
  3.1. Radiance measurement
  3.2. Retrieval of concentration
  3.3. Forward and inverse problem
  3.4. High resolution forward radiative code
  3.5. Neural network code
  3.6. Least square fitting code
  3.7. Sensitivity and accuracy
  3.8. Information content
4. Results
  4.1. Atmospheric variability
  4.2. Absorption spectral range
  4.3. Radiative transfer simulations
  4.4. Accuracies in column retrievals
    4.4.1. Strong infrared absorbers
    4.4.2. Weak infrared absorbers
  4.5. Vertical information content
    4.5.1. Sensitivity of the instrument
    4.5.2. Number of independent information
    4.5.3. Impact of the noise level
    4.5.4. Vertical resolution
  4.6. Global distributions measured by current remote sensors
    4.6.1. CO
    4.6.2. Ozone
    4.6.3. Methane
    4.6.4. SO₂
5. Conclusion
  5.1. Compliance with requirements per species
  5.2. Instrumental improvements
  5.3. Limitations and further works
0 Executive Summary

Following a user consultation process a consolidated list of data products derived from geostationary satellite observations, to support potential operational atmospheric chemistry applications was prepared by Lelieveld (2003). This table (listed here as Table 2) specified the accuracy, temporal and spatial resolution of a selected number of data products. The study did not specify which geostationary satellite observations are needed to obtain the data products, moreover the study did not indicate on how these products are to be derived from the observations.

A geostationary infrared sounder, specified according to needs of numerical weather prediction applications, is being considered as part of EUMETSAT Post-MSG activities. This sounder specifications, valid at time of this study, are listed in Table 1 and covers the region between $660 – 2500 \text{ cm}^{-1}$ with in total 9 bands, and a threshold spectral resolution between $0.625 – 1.25 \text{ cm}^{-1}$ depending on the particular band and a goal spectral resolution of $0.5 \text{ cm}^{-1}$. The $\text{ne\Delta T}$ values are approximately $0.2 – 0.4 \text{ K (G, 0.1 - 0.2 K Threshold)}$.

The present report documents the capabilities of such a geostationary infrared sounder to retrieve the trace gases listed in the consolidated table prepared by Lelieveld. In particular the study considers the tropospheric columnar amount of $\text{O}_3$, $\text{CO}$, $\text{SO}_2$, $\text{H}_2\text{CO}$, $\text{NO}$, $\text{NO}_2$, and PAN. The study aims to:

- Identify spectral absorption range for each product using standard spectroscopic databases
- Perform forward radiative transfer simulations using representative atmospheric cases (e.g. low/high pollution levels)
- Study the impact of noise and spectral resolution on the accuracy of the retrieved products
- Investigate the possibilities of retrieving vertical profiles for selected species and quantify the reachable vertical resolution.

This was done through a combination of detailed radiative transfer simulations of high resolution spectra under clear sky conditions. As input to the radiative transfer calculations the vertical distribution of atmospheric composition was taken from a three dimensional tropospheric chemical transport model (MOZART). The calculated radiances at the top of the atmosphere were input to retrieval routines to analyse the impact of noise and spectral resolution on the accuracy of the retrieved trace gases. Finally to analyse the possibilities to retrieve vertical profiles, the retrieval concept developed by Rodgers (2000) was adopted. Where appropriate IMG/ADEOS data were used to illustrate the outcome of the analysis.

Results of the analysis are

- Retrieval of columnar amount of $\text{CO}$, $\text{O}_3$ and $\text{CH}_4$ with the required accuracy will be possible with the specified threshold values for noise and spectral resolution considered.
- Accuracy of these three above mentioned trace gases, will increase if noise or spectral resolution can be improved. Noise improvements are preferred over spectral resolution. However, only limited vertical information can be extracted from the observations even at goal spectral resolution and noise levels. Significant improvements in especially the goal
spectral resolution would be needed, in order to obtain the goal values for the vertical distribution.

- The goal instrument specification in combination with controlled state of atmosphere (T,q) should allow for limited information on NO₂, SO₂ and possible PAN.

- It is not recommended to expand the spectral domain to cover H₂CO as detection will be difficult even under highly polluted scenarios.

The scenarios analysed here were highly idealised, for instance the radiative transfer simulations were performed assuming cloud free conditions, and for the retrievals it was assumed that the ancilliary data (in particular vertical distribution of temperature and humidity) were perfectly known. To increase confidence in the outcome of the current study, additional analysis are needed which corrects for these limitations.
1 Introduction

1.1 Motivation of this study

The study intends to analyze the capabilities of a Post-MSG GEO infrared sounder to meet the requirements defined for monitoring the atmospheric composition of several key species. It follows previous studies undertaken under the auspices of EUMETSAT: 'The user consultation process' to prepare future geostationary programmes, and the study on ‘Geostationary satellite observations for monitoring atmospheric composition and chemistry applications’ [Lelieveld, 2003]. It relies on the infrared sounder specification provided by EUMETSAT for a Post-MSG instrument and on requirements for a range of data products.

The study aims to:

- Identify spectral absorption range for each product using standard spectroscopic databases
- Perform forward radiative transfer simulations using representative atmospheric cases (e.g. low/high pollution levels)
- Study the impact of noise and spectral resolution on the accuracy of the retrieved products
- Investigate the possibilities of retrieving vertical profiles for selected species and quantify the reachable vertical resolution.

For this study, we rely on forward radiative transfer simulations and retrieval inversion codes developed in the framework of the IASI mission, which were adapted to the present instrument. Processing tools available include analysis of Level 1 data (geo-located radiance spectra) to retrieve Level 2 products (trace gas contents at the measurement locations), and derivation of enhanced Level 3 products (global distributions). Results are illustrated with IMG/ADEOS data or other appropriate sensors when available.

1.2 Instrument specifications

Instrumental specifications were delivered by Eumetsat when available during the course of the study. The following values are extracted from the 'Post-MSG Mission MRD Sounder Discussion' document (June 2003 version). These specifications were defined in priority to meet the needs for NWP global and regional applications.

Spatial resolution

To match the NWP model resolution the sounder must yield spectral radiances suitable for retrieval of temperature and relative humidity with a spatial resolution of $\Delta X = 10 \text{ km}$ (THRESHOLD; GOAL: $\Delta X = 2 \text{ km}$). This resolution of soundings is needed to define the horizontal temperature and moisture gradients. GOAL spatial resolution is preferred because increasing the spatial resolution greatly increases the likelihood of obtaining clear-air soundings, and improves the ability of IRS to obtain soundings adjacent to cloudy regions.
Temporal sampling - Scan area

Sounding measurements are needed with a high vertical, and high spatial resolution and high temporal sampling to enhance the National Meteorological Services’s ability to initialize NWP models with more realistic fields of temperature, moisture and wind. The scan area shall be selectable to offer flexible scan scenarios. The area range from meso-scale areas (1000 by 1000 km) over regional model areas for NWC regional applications up to the full disk for global model applications.

- For global models a temporal sampling of **60 min** over the full disk (THRESHOLD) is required, but some applications may need a higher temporal sampling of **15 min** (GOAL).
- For regional models the high temporal sampling of **10 min** over an area of the size of Europe (THRESHOLD; GOAL: **5 min**) is needed.
- For the support of nowcasting and very short range forecasting a temporal sampling of **2 min** over 1000 by 1000 km areas (THRESHOLD; GOAL: **1 min**) is required.

Spectral coverage

The post-MSG sounder shall cover the spectral range 700 cm\(^{-1}\) to 2500 cm\(^{-1}\) (i.e. 4.0 \(\mu\)m to 14.3 \(\mu\)m) continuously. Reduced performance can be accepted in the range 1250 cm\(^{-1}\) to 1587 cm\(^{-1}\) (i.e. 6.3 \(\mu\)m to 8.0 \(\mu\)m) (TBC). Note: The spectral range shall be extended up to 667 cm\(^{-1}\) (14.99 \(\mu\)m), the most opaque feature of the 15 \(\mu\)m CO\(_2\) band, but NeDT requirement is applicable only up to 700 cm\(^{-1}\). Performance shall be reported up to 667 cm\(^{-1}\) (14.99 \(\mu\)m).

Spectral resolution and noise

The spectral resolution as a function of wavenumber is listed in the table below. The spectral resolution after apodization has been applied. These are apodized values inclusive of any instrumental effect (e.g. lateral shift, etc.) that ultimately is equivalent to spectral resolution degradation.

The radiometric resolution is specified in terms of noise equivalent temperature difference (NeDT) associated to a reference temperature of 280 K at which the NeDT is computed. The specifications given below apply to calibrated spectra, which means that the contribution of the calibration coefficients noise is included. This contribution shall be calculated taking into account the effective temperature of the in-orbit calibrations views.

A secondary objective of the mission is to support user/service need regarding operational chemical applications and air quality monitoring as identified for NWC and chemistry in general. It has therefore been proposed to add a specific chemistry mode, characterized by a spectral resolution of 0.2 cm\(^{-1}\) or better. It is expected that the chemistry mode can be restricted to a sub set of not more than five of the 10 IRS sounder bands defined after. It should be emphasized,
however, that operating the sounder in a very high resolution mode may cause a decrease in
temporal sampling and/or area coverage.

Table 1. Spectral resolution and noise specifications as a function of wavenumber

<table>
<thead>
<tr>
<th>Mission Band</th>
<th>Frequency range cm$^{-1}$</th>
<th>Spectral resolution cm$^{-1}$ (T; G 0.5)</th>
<th>NeDT @ 280 K 1σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRS-0</td>
<td>667-700</td>
<td>0.625 cm$^{-1}$ (T; G 0.5)</td>
<td>TBD</td>
</tr>
<tr>
<td>IRS-1</td>
<td>700-770</td>
<td>0.625 cm$^{-1}$ (T; G 0.5)</td>
<td>0.2 (G: 0.1) TBC</td>
</tr>
<tr>
<td>IRS-2</td>
<td>770-1000</td>
<td>0.625 cm$^{-1}$ (T; G 0.5)</td>
<td>0.24 (G: 0.12) TBC</td>
</tr>
<tr>
<td>IRS-3</td>
<td>1000-1070</td>
<td>1.25 cm$^{-1}$ (T; G 0.5)</td>
<td>0.2 (G: 0.15) TBC</td>
</tr>
<tr>
<td>IRS-4</td>
<td>1070-1210</td>
<td>1.25 cm$^{-1}$ (T; G 0.5)</td>
<td>0.3 (G: 0.2) TBC</td>
</tr>
<tr>
<td>IRS-5</td>
<td>1210-1600</td>
<td>1.25 cm$^{-1}$ (T; G 0.5)</td>
<td>0.2 (G: 0.1) TBC</td>
</tr>
<tr>
<td>IRS-6</td>
<td>1600-2000</td>
<td>1.25 cm$^{-1}$ (T; G 0.5)</td>
<td>0.35 (G: 0.25) TBC</td>
</tr>
<tr>
<td>IRS-7</td>
<td>2000-2250</td>
<td>1.25 cm$^{-1}$ (T; G 0.5)</td>
<td>0.2 (G: 0.15) TBC</td>
</tr>
<tr>
<td>IRS-8</td>
<td>2250-2400</td>
<td>1.25 cm$^{-1}$ (T; G 0.5)</td>
<td>0.3 (G: 0.25) TBC</td>
</tr>
<tr>
<td>IRS-9</td>
<td>2400-2500</td>
<td>1.25 cm$^{-1}$ (T; G 0.5)</td>
<td>0.4 (G: 0.25) TBC</td>
</tr>
</tbody>
</table>

1.3 Product requirements

The product requirements in terms of accuracy, spatial and temporal resolution, are provided in Table 2. The latter was built from the ‘Product Table for operational atmospheric chemistry applications’, based on the study performed by Lelieveld and approved by a consortium of scientists, and from relevant parts of the ‘Consolidated Table of Observational Requirements for Nowcasting and very short range forecasting in 2015-2020, excluding NWP’ from the position paper of Golding et al. [2003] The complete sections extracted from the initial table of the Golding et al. position paper for Chemistry, Dispersion, Chemistry & biology models, Air quality and biomass burning, UV radiation, and Surface pollution, are provided in Annex A.

Table 2 is organized such as to provide the key scientific questions to be addressed for each product. It will be the reference Table for drawing the conclusions at the end of this study. Table 2 is essentially that given in the Lelieveld study, except for the first column where we have added the addressed scientific questions, and the last column, which provides a summary of the requirements. These two columns add important information for the present study. For example ozone was found to be a target product for different fields: it is a climate gas, it is a key species for quantifying UV exposure, it is an important chemical compound for tropospheric oxidation capacity related studies, it has an impact on life quality through respiratory effects, etc. The requirements for all these questions are different and may be fulfilled or not with the proposed instrument. The detailed structure of Table 2 is as follow: Scientific question, Data product, Horizontal resolution, Vertical resolution, Accuracy, Breakthrough level, with:

Scientific question
Scientific question to address:
A= Operational Atmospheric Chemistry Applications [Lelieveld, 2003]
B= Transport: Aviation warning, damage to aircraft, land [Golding et al., 2003]
Data product
Required variable. AOT is aerosol optical thickness (for the fine and course fraction, D<1 μm and D>1 μm, respectively), Aer Reff is aerosol effective radius, SSA is aerosol single scattering albedo, PAN is peroxyacetyl nitrate, VOC stands for volatil organic compounds.

Horizontal resolution (in km)
The spatial resolution that is required before an observation can be used with any benefit (threshold requirement). The spatial resolution beyond which there is no further benefit from improvements is provided in brackets (target requirement).

Vertical resolution (in km)
The vertical resolution that is required before an observation can be used with any benefit (threshold requirement). The vertical resolution beyond which there is no further benefit from improvements is provided in brackets (target requirement). T refers to a total atmospheric column, t is for tropospheric column, bl is for the boundary layer, ft is for the free troposphere, s is for the stratosphere.

Temporal resolution (in hours)
The frequency that is required before an observation can be used with any benefit (threshold requirement). The frequency beyond which there is no further benefit from improvements is provided in brackets (target requirement). Superscript d refers to daytime, n to night.

Accuracy
The accuracy of the variable that is required before an observation can be used with any benefit (threshold requirement). The accuracy beyond which there is no further benefit from improvements is provided in brackets (target requirement). It may be expressed as an RMS error (in %) in the physical variable or as a hit rate for detection of the variable (e.g. in AOT or SSA units).

Breakthrough level
The combination of accuracy/spatial/vertical/temporal resolution that would provide a significant advance in forecasting capability relative to that currently available.
<table>
<thead>
<tr>
<th>Scientific question</th>
<th>Data product</th>
<th>Horizontal resolution in km</th>
<th>Vertical resolution in km</th>
<th>Temporal resolution in hr</th>
<th>Accuracy in %</th>
<th>Breakthrough level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ozone</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A,C1</td>
<td>O$_3$</td>
<td>10 (2)</td>
<td>t (2)</td>
<td>$1^d(n)$ (.5)</td>
<td>10 (5)</td>
<td>5%/10km/tropo/1h</td>
</tr>
<tr>
<td>C2</td>
<td>O$_3$ tot.</td>
<td>100 (10)</td>
<td>T</td>
<td>24 (1)</td>
<td>-</td>
<td>50km/6h</td>
</tr>
<tr>
<td>C2</td>
<td>O$_3$ prof.</td>
<td>50 (10)</td>
<td>2 (.5)</td>
<td>24 (1)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>Pollution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A,E</td>
<td>CO</td>
<td>10 (2)</td>
<td>t (2)</td>
<td>$2^d(n)$ (.5)</td>
<td>10 (5)</td>
<td>10%/10km/tropo/1h</td>
</tr>
<tr>
<td>A,E</td>
<td>SO$_2$</td>
<td>10 (2)</td>
<td>t (2)</td>
<td>$1^d(n)$ (.5)</td>
<td>50 (20)</td>
<td>50%/10km/tropo/1h</td>
</tr>
<tr>
<td>A,E</td>
<td>H$_2$CO</td>
<td>10 (2)</td>
<td>t (2)</td>
<td>$1^d(n)$ (.5)</td>
<td>50 (20)</td>
<td>50%/10km/tropo/1h</td>
</tr>
<tr>
<td>E</td>
<td>NO</td>
<td>10 (2)</td>
<td>t (2)</td>
<td>1 (.5)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>A,E</td>
<td>NO$_2$</td>
<td>10 (2)</td>
<td>t (2)</td>
<td>$1^d(n)$ (.5)</td>
<td>50 (20)</td>
<td>50%/10km/tropo/1h</td>
</tr>
<tr>
<td>A,E</td>
<td>PAN</td>
<td>10 (2)</td>
<td>t (2)</td>
<td>$1^d(n)$ (.5)</td>
<td>50 (20)</td>
<td>50%/10km/tropo/1h</td>
</tr>
<tr>
<td>E</td>
<td>VOC</td>
<td>10 (1)</td>
<td>t (2)</td>
<td>3 (.25)</td>
<td>20 (10)</td>
<td></td>
</tr>
<tr>
<td><strong>UV radiation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>UV-A</td>
<td>10 (2)</td>
<td>surface</td>
<td>$1^d(n)$ (0.5)</td>
<td>20 (5)</td>
<td>regional</td>
</tr>
<tr>
<td>C</td>
<td>UVA</td>
<td>50(2)</td>
<td>surface</td>
<td>3 (.5)</td>
<td>20 (5)</td>
<td>10%/10km/surface/1h</td>
</tr>
<tr>
<td>A</td>
<td>UV-B</td>
<td>10 (2)</td>
<td>surface</td>
<td>$1^d(n)$ (0.5)</td>
<td>20 (5)</td>
<td>regional</td>
</tr>
<tr>
<td>C</td>
<td>UVB</td>
<td>50(2)</td>
<td>surface</td>
<td>3 (.5)</td>
<td>20 (5)</td>
<td>10%/10km/surface/1h</td>
</tr>
<tr>
<td>C2</td>
<td>UV albedo</td>
<td>10 (1)</td>
<td>surface</td>
<td>$&gt;$ 24</td>
<td>20 (10)</td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>UV absorbers (NO$_2$,...)</td>
<td>10 (10)</td>
<td>surface</td>
<td>$&gt;$ 24</td>
<td>20 (10)</td>
<td></td>
</tr>
<tr>
<td><strong>Aerosols</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A,C1</td>
<td>AOTline</td>
<td>5 (0.5)</td>
<td>t (BL+FT)</td>
<td>$1^d(n)$ (0.25)</td>
<td>0.05(0.01)</td>
<td>.05/5km/T/1h</td>
</tr>
<tr>
<td>A,C1</td>
<td>AOTcourse</td>
<td>5 (0.5)</td>
<td>t (BL+FT)</td>
<td>$1^d(n)$ (0.25)</td>
<td>0.05</td>
<td>.05/5km/T/1h</td>
</tr>
<tr>
<td>A,C1</td>
<td>Aer Reff</td>
<td>5 (0.5)</td>
<td>T</td>
<td>$1^d(n)$ (0.25)</td>
<td>30 (10)</td>
<td>30/5km/T/1h</td>
</tr>
<tr>
<td>A,C1</td>
<td>SSA</td>
<td>5 (0.5)</td>
<td>T</td>
<td>$1^d(n)$ (0.25)</td>
<td>0.03</td>
<td>.03/5km/T/1h</td>
</tr>
<tr>
<td>B</td>
<td>volcanic ash conc.</td>
<td>50 (1)</td>
<td>t (.3)</td>
<td>6 (.25)</td>
<td>20 (5)</td>
<td>20%/20km/total/1h</td>
</tr>
<tr>
<td>B</td>
<td>volcanic ash loc.</td>
<td>250 (1)</td>
<td>t (.3)</td>
<td>6 (.25)</td>
<td>20 (5)</td>
<td>20%/20km/total/1h</td>
</tr>
<tr>
<td>B</td>
<td>volcanic ash comp.</td>
<td>100 (1)</td>
<td>t (.3)</td>
<td>1 (.25)</td>
<td>70 (85)</td>
<td>70%/20km/total/1h</td>
</tr>
<tr>
<td>B</td>
<td>sand/dust storm loc.</td>
<td>50 (1)</td>
<td>t</td>
<td>1 (.25)</td>
<td>50 (85)</td>
<td>70%/40km/total/1h</td>
</tr>
<tr>
<td>B</td>
<td>location of volcano</td>
<td>1 (.1)</td>
<td>surface</td>
<td>6 (1)</td>
<td>50 (85)</td>
<td>50%/50km/surface/1h</td>
</tr>
<tr>
<td>C2</td>
<td>total aerosol</td>
<td>50 (10)</td>
<td>T</td>
<td>24 (1)</td>
<td>25 (10)</td>
<td>25/20km/24h</td>
</tr>
<tr>
<td><strong>Meteorology</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>T</td>
<td>10 (2)</td>
<td>2 (1)</td>
<td>1 (.25)</td>
<td>2K (1K)</td>
<td>2K/10km/2km/1h</td>
</tr>
<tr>
<td>A, D</td>
<td>H$_2$O</td>
<td>5 (0.5)</td>
<td>t (BL+FT)</td>
<td>$1^d(n)$ (0.25)</td>
<td>5 (1)</td>
<td>5%/10km/2km/1h</td>
</tr>
<tr>
<td>C2</td>
<td>total cloud water</td>
<td>50 (10)</td>
<td>T</td>
<td>24 (1)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Climate gases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A,D</td>
<td>CO$_2$</td>
<td>50 (10)</td>
<td>t (.5)</td>
<td>$6^d(n)$ (1)</td>
<td>2 (1)</td>
<td>5%/50km/Total/6h</td>
</tr>
<tr>
<td>A,D</td>
<td>CH$_4$</td>
<td>50 (10)</td>
<td>t (1)</td>
<td>$6^d(n)$ (1)</td>
<td>5 (1)</td>
<td>5%/50km/Total/6h</td>
</tr>
<tr>
<td>F</td>
<td>CFC11</td>
<td>50 (10)</td>
<td>S</td>
<td>24 (1)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>CFC12</td>
<td>50 (10)</td>
<td>S</td>
<td>24 (1)</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Present and planned satellite missions to sound the troposphere

2.1 Tropospheric sounders

During the last decade, trace gas abundances in the troposphere were measured from sparsely distributed measurement sites and observations were mostly confined to the surface. The advent of satellite-borne instruments to probe the troposphere has dramatically increased our ability to access the impact of human activities on the chemical composition of the atmosphere and on the climate changes. Tropospheric sounders using either thermal infrared, either UV-visible technology, are currently flying onboard polar-orbiting (LEO) satellites.

Three concepts of passive remote sensors have now demonstrated their ability to provide tropospheric sounding measurements: correlation radiometers (e.g. MOPITT), UV-visible radiometers using the reflected/scattered solar radiation (e.g. GOME and successors SCIAMACHY, OMI, GOME2), and infrared spectrometers using the Earth’s thermal emission as a source (e.g. IMG and follow on AIRS, TES, IASI). The design and technical specifications (spectral range, spectral resolution, footprint on the ground, scanning mode) of each instrument is optimized as a function of the scientific objectives to be achieved during the mission.

All these instruments allow measurements to be made with a high spatial (global Earth) coverage on long time scales for several species at the same time and location. The main drawbacks are on one hand the viewing geometry that provides a weighted integral of atmospheric properties along the optical path and thus limits the vertical resolution, and on the other hand the low revisit time (2 times per day) for a given location, which restricts the usefulness for chemically reactive species. These drawbacks could be overcome by a tropospheric sounder placed on a geostationary (GEO) orbit.

2.2 Thermal infrared sounders from a polar orbit

2.2.1 IMG

The Interferometric Monitor for Greenhouse gases (IMG) [Kobayashi et al., 1999] launched aboard ADEOS in August 1996 was the first high resolution nadir infrared instrument allowing the simultaneous measurement from space of a series of trace gases: H$_2$O, CO$_2$, N$_2$O, CH$_4$, CFCs, O$_3$, and CO. Unfortunately, in June 1997, the ADEOS satellite ceased to collect and transmit data due to a power failure in its solar panel. The IMG instrument was developed by the Japan Resources Observation System Organization (JAROS) for the Ministry of International Trade and Industry (MITI) and was launched as part of the ADEOS payload. ADEOS was a sun-synchronous (equator local crossing time at descending node at 10:30 AM), ground track repeat, polar-orbiting satellite. The instrument performed a global coverage of the Earth, making 14.5 orbits per day with series of six successive measurements separated by 86 km (every 10 sec.) along the track, followed by the observation of deep space and of an internal blackbody for calibration purposes. Due to the huge data flow rate, the operational mode of
IMG was set to 4 days operation/10 days halt alternation, except for one specific period from 1 to 10 April 1997 for which 10 consecutive days are available.

IMG was a nadir-viewing Fourier transform interferometer that recorded the thermal emission of the Earth-atmosphere system between 600 and 3030 cm$^{-1}$, with a maximum optical path difference of 10 cm. The footprint on the ground was 8 km x 8 km, in three spectral bands (band 1 from 2387 to 3030 cm$^{-1}$, band 2 from 2000 to 2500 cm$^{-1}$, and band 3 from 600 to 2000 cm$^{-1}$), corresponding to three different detectors and three geographically adjacent footprints (http://www.eoc.nasda.go.jp/guide/satellite/sendata/img-e.html).

We had access to the IMG Level 1C data through our involvement in the NASDA/MITI Announcement of Opportunity for IMG/ADEOS\(^1\). Results obtained from simulated data analysis in this study have been confronted to retrievals performed using the atmospheric spectra recorded by the IMG instrument. Among the available Level 1C radiance data, we selected four good-quality measurements periods that were cloud-filtered for further analysis of trace gases contents. It represents a set of 20 000 spectra. Most of the illustrations provided in this report use the April period for which 10 consecutive days were available. An example of a radiance spectrum recorded in the thermal infrared by the IMG/ADEOS instrument is provided in Figure 1 with the main infrared atmospheric absorbers identified.

### 2.2.2 MOPITT

The Measurement of Pollution in the Troposphere (MOPITT) instrument, on the NASA EOS Terra satellite, measures the upwelling infrared radiance. Using the technique of correlation spectroscopy whereby a cell of the gas to be measured is used as an optical filter in the infrared to measure the signal from the same gas in the atmosphere, information regarding the distribution of atmospheric CO and CH$\textsubscript{4}$ can be extracted. MOPITT operates by sensing infrared radiation from either the thermal emission/absorption at 4.7 $\mu$m for CO profiles, or reflected sunlight at about 2.2-2.4 $\mu$m for CO and CH$\textsubscript{4}$ column measurements in daylight. The thermal band signals are sensitive to thermal emission from the Earth’s surface as well as atmospheric absorption and emission. The solar band signals are sensitive to atmospheric CO through absorption processes only. By using appropriate data analysis techniques, concentration profiles of CO are obtained on a global basis at a reasonably high horizontal (~22km) and vertical resolution (~3km).

Currently, only clear-sky radiances (i.e. radiances uncontaminated by clouds) and thermal channels are fed to the operational retrieval algorithm [Deeter et al., 2003]. The thermal band signals depend not only on the atmospheric CO distribution but also on various other atmospheric quantities (such as the atmospheric temperature and water vapor mixing ratio profiles) and surface parameters (surface temperature and longwave emissivity). Accurate values for all of these geophysical parameters must be obtained to produce accurate retrievals. Atmospheric temperature and water vapor profiles are obtained by spatially and temporally interpolating reanalysis\(^1\)

\(^1\)Tropospheric and stratospheric studies using ADEOS sensors [PI G. Mége]
Figure 1: IMG infrared spectrum (in radiance units) in the 700-2200 cm$^{-1}$ spectral range. The baseline of the spectra corresponds to the Earth black-body emission at $T=287.6$ K. The main molecular absorption bands are due to rovibrational transitions associated with CO$_2$, O$_3$, CH$_4$, H$_2$O and CO.

Profiles from NCEP to the location and time of each MOPITT pixel. However, sources of geophysical data such as NCEP are unable to provide accurate values of surface temperature and emissivity (both of which are highly variable) at the temporal and spatial resolution demanded by the MOPITT retrievals. Fortunately, information contained in the MOPITT thermal band signals allows retrieval of the surface temperature and emissivity along with the CO profile, and makes external data sources for these quantities necessary only for providing a priori and initial guess values. Thus, rather than assuming fixed values for the surface temperature and emissivity, these two parameters are included in the retrieval state vector (along with the elements of the CO profile). The MOPITT CO "Level 2 Product" therefore consists of retrieved values and estimated uncertainties of the CO profile, CO total column, surface temperature, and surface emissivity. For the CO profile, the retrieved error covariance matrix is also provided.

Figure 2 illustrates the CO total column as retrieved from the MOPITT radiance data for a recent biomass burning event in Central America. Further information on the instrument and retrievals may be found on http://www.eos.ucar.edu/mopitt/.
Figure 2: The widespread and intense biomass burning occurring on the Yucatan Peninsula and other parts of Central America is producing high concentrations of CO in the lower atmosphere, as shown in this image of observations by the MOPITT instrument aboard Terra satellite. This false-color image shows the mixing ratios of CO at an altitude of about 3 km (700 hPa) averaged from April 25 to May 1, 2003. Gray areas indicate where no data are available. Available from http://www.eos.ucar.edu/mopitt/.

2.2.3 AIRS

The Atmospheric Infrared Sounder (AIRS) is a high-resolution infrared sounder selected to fly on the EOS Aqua platform with two operational microwave sounders, AMSU and HSB. It is an high-spectral resolution, grating multispectral infrared sounder (3.74 to 15.4 μm), with a spectral resolving power of 1200, operating in a cross-track-scanning mode. Measurements from the three instruments will be analyzed jointly to filter out the effects of clouds from the IR data in order to derive clear-column air-temperature profiles and surface temperatures with high vertical resolution and accuracy. 1 K temperature retrieval accuracy per 1 km layer in the troposphere is expected. Together, these three instruments constitute an advanced operational sounding system.

An illustration of the capabilities of the AIRS instrument for sounding the surface temperature is provided in Figure 3, which shows an average of surface temperature measurements throughout the globe for April 2003. Further information available on http://aqua.nasa.gov/AIRS3.html.
2.2.4 IASI

The Infrared Atmospheric Sounding Interferometer (IASI) consists of a Fourier Transform Spectrometer with an imaging system, scheduled to be launched on the European METOP platform in 2005. It is based on a Michelson Interferometer and is designed to measure the infrared spectrum emitted by the Earth and the atmosphere, from 645 to 2760 cm\(^{-1}\), at 0.5 cm\(^{-1}\) spectral resolution and high radiometric quality. Data samples are taken at intervals of 25 km along and across track (nadir), each sample having a maximum diameter of about 12 km (see Figure 4). IASI will provide improved infrared soundings of the temperature profiles in the troposphere and lower stratosphere, moisture profiles in the troposphere, as well as some of the chemical components playing a key role in the climate monitoring, global change and
atmospheric chemistry. The Centre National d’Etudes Spatiales (CNES) is leading the IASI programme in association with EUMETSAT. ALCATEL is the instrument Prime Contractor. The IASI will fly on MetOp 1, 2 and 3.

The inversion codes and most of the tools described in this study were developed in the framework of a joint Belgian-French project\(^2\), that was selected in 1995 after he CNES/EUMETSAT IASI Announcement of Opportunity and that is part of the IASI Sounding Science Working Group (ISSWG). More information available on http://smsc.cnes.fr/IASI/index.htm.

### 2.2.5 TES

The Tropospheric Emission Spectrometer (TES) is a high-resolution infrared-imaging Fourier transform spectrometer with spectral coverage of 3.2 to 15.4 μm at a spectral resolution of 0.025 cm\(^{-1}\), thus offering line-width-limited discrimination of essentially all radiative active molecular species in the Earth’s lower atmosphere. TES has the capability to make both limb and nadir observations. In the limb mode, TES has a height resolution of 2.3 km, with coverage from 0 to 34 km. In the downlooking mode, TES has a spatial resolution of 0.53 x 5.3 km with a swath of 5.3 x 8.5 km. TES is a pointable instrument and can access any target within

\(^2\)Chemistry and Climate related studies using IASI [PI C. Clerbaux]
45° of the local vertical, or produce regional transects up to 885-km length without any gaps in coverage. TES employs both the natural thermal emission of the surface and atmosphere and reflected sunlight, thereby providing day-night coverage anywhere on the globe. Observations from TES will further help the understanding of long-term variations in the quantity, distribution, and mixing of minor gases in the troposphere, including sources, sinks, troposphere-stratosphere exchange, and the resulting effects on climate and the biosphere. See http://eos-chem.gsfc.nasa.gov/instruments/tes/introduction.html for further details.

2.3 Improved sounders from a geostationary or drifting orbit

On a longer term, both ESA and NASA are studying new satellite concepts combining several atmospheric sounding instruments to meet the observational requirements for atmospheric chemistry. Innovative platforms combining UV-visible and infrared instruments have recently been submitted to ESA for launch on geostationary orbit (The Geostationary Tropospheric Pollution Explorer mission (GEOTROPE)- PI J. Burrows) or on a drifting orbit (The TROposheric Chemistry and Climate mission (TROC)- PI C.Camy-Peyret). These instruments are expected to overcome the main drawback of polar orbiting satellites, which, with a low revisit time (2 times per day for a given location), have limited usefulness for local or regional pollution studies involving fast chemically reactive species.

3 Analysis of infrared atmospheric spectra: general approach and methodology

3.1 Radiance measurement

Passive remote sensing instruments measure radiance spectra (in W/m².m⁻¹.sr). Atmospheric state variables such as temperature or trace gas concentration may be retrieved from measured radiances using the so called inversion algorithms. The feasibility to retrieve a target species and the accuracy at which it is done, depend on the instrumental design (spectral range and resolution, radiometric noise), on the radiative properties of the target species and on the a priori knowledge of the atmosphere (temperature and moisture profiles etc.). Each instrument is optimized for a given spectral range and is only sensitive to species that combine both spectroscopic absorption lines in this spectral range and large enough abundances to be detected. One important characteristic of current and upcoming sensors is that they combine high spectral resolution with good signal/noise ratio performance, allowing the retrieval of weak absorbers.

The retrieval of accurate concentrations from these spectral radiances relies on the selection of the spectral channels that contain the maximum information on the target gas to retrieve, i.e. which are not contaminated by other absorbers. Also, some channels can not be used as they are saturated and all the initial radiation is fully absorbed. Figure 5 represents the calculated infrared upwelling radiances in the 2000-2250 cm⁻¹ spectral range, where the intense 1-0 vibrational transition of CO occurs (top part of the Figure). In order to illustrate the difficulty
of discriminating between the spectral signature of CO and other absorbing gases in the same spectral range (i.e. \( \text{H}_2\text{O}, \text{CO}_2, \text{O}_3, \) and \( \text{N}_2\text{O} \)) the separated contribution of CO is also provided (bottom part of the Figure).

3.2 Retrieval of concentration

3.2.1 Forward and inverse problem

Let \( \mathbf{y} \) be the measurement vector containing the selected measured radiances, and \( \mathbf{x} \) be the concentration of a given constituent, then the general remote sensing equation can be written as follows [Rodgers, 2000]:

\[
\mathbf{y} = \mathbf{f}(\mathbf{x}, \mathbf{b}) + \mathbf{\varepsilon}
\]

(1)

where \( \mathbf{f} \) represents the forward radiative transfer function, \( \mathbf{b} \) the other parameters having an impact on the measurement, and \( \mathbf{\varepsilon} \) the measurement noise. The vector \( \mathbf{b} \) includes Earth surface radiative features (emissivity and temperature), variables describing the state of the atmosphere (vertical profiles of atmospheric temperature, water vapor and other atmospheric constituents, clouds, aerosols, etc.), and the instrument spectral response function and resolution. The inverse problem consists in retrieving \( \hat{\mathbf{x}} \), an estimate of the true state \( \mathbf{x} \), from the measurement \( \mathbf{y} \), and can be written

\[
\hat{\mathbf{x}} = \mathbf{R}(\mathbf{y}, \hat{\mathbf{b}}) = \mathbf{R}(\mathbf{f}(\mathbf{x}, \mathbf{b}) + \mathbf{\varepsilon} | \hat{\mathbf{b}})
\]

(2)

where \( \hat{\mathbf{b}} \) corresponds to the estimate of the non-retrieved parameters \( \mathbf{b} \), and \( \mathbf{R} \) is the inverse transfer function to be modelized.

For nadir-looking measurements, this problem is known to be non-linear and under-constrained as components of the atmospheric vertical profile do not contribute to the measurements and hence can not be detected. A variety of methods have been developed to retrieve concentration...
profiles from the spectra measured by remote sounders. A large body of literature is available on
the subject, and the most widely used approaches in atmospheric remote sensing are described
in Rodgers [2000]. An inversion method relies on a forward radiative transfer model with a
spectroscopic database like HITRAN [Rothman et al., 2003], and uses a minimization scheme
optimized for the instrument and the target geophysical variable to retrieve. The following
sections describe the forward radiative code we used for the simulations, as well as the two in­
version algorithms currently developed in the framework of the preparation of the IASI mission
( the fast retrieval algorithm based on neural networks, and the least square fitting procedure)
that were also used for this study.

3.2.2 High resolution forward radiative code

The general radiative transfer equation for nadir observation writes as:

\[
f(\nu) = \varepsilon(\nu)B(\nu, T_0) \exp(-\int_0^H \alpha(\nu, s)ds) + \frac{\Omega}{4\pi} \exp(-\int_0^H \alpha(\nu, s)ds) \left[ \int_0^H \alpha(\nu, s)B(\nu, T_s) \exp(-\int_s^H \alpha(\nu, s)ds)ds \right] \]

(3)

where \(\nu\) is the frequency, \(f(\nu)\) is the atmospheric spectral radiance, \(B(\nu, T)\) is the Planck black-body function at the temperature \(T\), \(\varepsilon(\nu)\) is the Earth’s emissivity, \(\Omega\) the instrument field of view, \(s\) the layer thickness in km and \(H\) the distance between the source (the Earth’s thermal infrared radiation in the case of nadir observations) and the detector. \(\alpha(\nu, s)\) is the absorption coefficient at the frequency \(\nu\) and the altitude \(s\). It is obtained by summing over the individual molecular absorption coefficients \(\alpha_n(\nu, s)\), defined for each species and computed using tabulated line parameters and absorption cross-sections.

Line-by-line forward models were used in this work to solve equation 3. For this purpose, the
atmosphere is divided into a number of layers and the monochromatic radiance is calculated
by multiplying the radiances for the different molecules and the different atmospheric layers
according to the average local temperature and pressure conditions. The well-documented and
validated line-by-line radiative transfer model LBLRTM (version 5.1) [Clough et al., 1995]
was used for the present study to generate the synthetic radiance spectra corresponding to the
instrument sounder specification defined by EUMETSAT for a post-MSG instrument (Table 1).
As particular features, LBLRTM includes:

- The molecular line parameters implemented in the latest issue of the HITRAN database
  [Rothman et al., 2003]. A Voigt profile is used for all species at all levels to describe the
  combined effect of pressure and temperature broadenings to the molecular line-shape. An
  algorithm is included to account for the line intensity dependence on temperature.

- The absorption cross-sections for several heavy species, including the halocarbons and
  substitutes

- Line coupling effects for CO₂

- Empirical formulations for the absorption continua of H₂O (CKD2.4), CO₂, O₂ and N₂
  in the infrared spectral region
- Standard atmospheres
- Standard apodization functions to account for the instrumental response

Figure 6 provides a simulation of the nadir radiances to be recorded for the range of spectral resolution values proposed.

![Figure 6: Line-by-line simulations of a nadir upwelling radiance spectrum at various spectral resolutions (0.2 cm\(^{-1}\), 0.5 cm\(^{-1}\), 0.75 cm\(^{-1}\) and 1.25 cm\(^{-1}\)). The 1800-3000 cm\(^{-1}\) part was multiplied by 20 for the sake of clarity.]

3.2.3 Neural network code

Neural network (NN) techniques were chosen to develop an inversion algorithm to retrieve trace gas concentrations from the IASI infrared measurements. A prototype of this code is currently being implemented in the METOP ground-segment by Eumetsat. The NN developed allows to statistically model the transfer function. The latter links the inputs, including the measurements \(\mathbf{y}\) and the estimators \(\mathbf{\hat{b}}\) of some parameters \(\mathbf{b}\), to the output quantities to be retrieved combined in the vector \(\mathbf{\hat{z}}\). The inverse problem described by Equation 2 may be rewritten
\[ \mathbf{x} = F(\mathbf{W}, \mathbf{y}, \mathbf{b}) \]  

(4)

where \( \mathbf{W} \) includes the parameters of the function \( F \) which will be referred as weights and biases of the NN in the following development.

For each species (currently O\(_3\), CH\(_4\), and CO are implemented) the input vector of the neural network module includes the relevant radiance channels [Clerbaux, 1998], the Earth surface temperature, and the atmospheric temperatures on several selected pressure levels. A wide range of radiance channels is used in the input vector because the information redundancy results in increasing the signal to noise ratio. The pressure levels of the atmospheric temperature profile are selected among the levels operationally retrieved during the IASI mission. Hence, for each species, an input is composed of \( N \) Level 1C radiances (measurements vector \( \mathbf{y} \)), and \( M \) Level 2 temperatures (estimators \( \mathbf{b} \) of some non-retrieved parameters \( \mathbf{b} \)), among which \( M-1 \) values correspond to the vertical profile of atmospheric temperature and 1 to the surface temperature, with \( N \) and \( M \) specific to each constituent.

Some of the other parameters \( \mathbf{b} \), not used for the input, are needed to calculate the error budget, and could be added in forthcoming versions of the NN module. The current version of the algorithm provides total column amounts for CO, CH\(_4\), and O\(_3\). For CH\(_4\) and CO, the retrieval of a total column amount gives a good outline of their tropospheric distributions since their vertical concentration distributions are characterized by maximum concentration values in the lower layers of the atmosphere. For O\(_3\), as 90% of the total amount is located in the stratosphere, a total column amount will be mostly influenced by its stratospheric concentration. In order to get a more comprehensive view of its tropospheric distribution, a partial, 0-12 km column amount is calculated along with the total one. The accuracy associated with each retrieval is estimated following the method described hereafter, and is provided with the final retrieved product.

The first step to build an efficient neural network is to find the optimal architecture, which has enough degrees of freedom to solve the problem. The architecture of a multilayer feed-forward NN is defined by the number of layers, the number of neurons on each layer, the topology of their connections, and the elementary transition functions associated with each neuron. An efficient architecture is chosen on the basis of empirical considerations depending on the complexity of the problem to solve. For the retrieval of trace gas concentrations from IASI spectra, our studies have shown that a well-suited architecture is a two layers perceptron: It is composed of an input layer comprising \( N\&M \) neurons (\( N \) radiances \( \mathbf{y} \) and \( M \) temperatures \( \mathbf{b} \)), connected to two hidden layers of \( S_1 \) and \( S_2 \) neurons. The transfer function \( f \) to be modeled being strongly non-linear, non-linear sigmoid transition functions \( f \) have been chosen for the neurons of the hidden layers:

\[ f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \]  

(5)

The output layer is composed of \( P \) neurons with linear transition functions \( g \):

\[ g(x) = \frac{1}{2}x \]  

(6)
For each quantity retrieved, the global transfer function may be written:

\[
\tilde{x}_p = g \left[ \sum_{k=1}^{S_2} w^3_{pk} f \left( \sum_{j=1}^{S_1} w^2_{kj} f \left( \sum_{i=1}^{N} w^1_{ji} y_i + \sum_{i=N+1}^{N+M} w^1_{ji} \tilde{y}_{i-N} + \beta_i^j \right) + \beta_k^j \right) + \beta_p^j \right], \quad p = 1, \ldots, P
\]

(7)

The connection weight matrices consist of \(w^1_{ji}, w^2_{kj}\) and \(w^3_{pk}\), where the index \(i = 1, \ldots, (N+M)\) denotes the elements of the input layer, \(j = 1, S_1\) the neurons of the first hidden layer, \(k = 1, S_2\) the neurons of the second hidden layer, and \(p = 1, \ldots, P\) the neurons of the output layer. The biases associated with the neurons are combined in the vectors \(\beta^1_j, \beta^2_k, \text{ and } \beta^3\).

As already mentioned, NN techniques allow an approximation of the transfer function \(F\), which links the inputs to the outputs of the problem. This approximation, based on statistical theory, requires a comprehensive dataset of known examples, representative of the behavior of the function to be estimated. A part of this dataset, the learning set, is used for the fitting of the NN parameters \(W\) (weights and biases), during the so-called learning phase. This learning set constitutes the a priori information necessary to solve the inverse problem. Examples not included in the learning set can be used to build a validation set to check the generalization capacities of the NN during the learning phase and a test set for the evaluation of the NN performance after the learning phase.

In order to build this comprehensive dataset, the atmospheric concentration profiles provided by the three-dimensional tropospheric chemical-transport model MOZART [Brasseur et al., 1998; Hauglustaine et al., 1998] have been coupled to the high resolution line-by-line radiative transfer code LBLRTM [Clough et al., 1995] in cloud-free and aerosol free conditions with a mean constant emissivity. The temperatures from the European Center for Medium-Range Weather Forecasts (ECMWF) analysis have been used in order to get more realistic simulations, and the simulated spectra have been convoluted with the instrument spectral response function for IASI Level 1C data [Camy-Peyret et al., 2001]. The instrumental noise has been accounted for by adding a random noise to the simulated spectra.

A supervised learning has been chosen for the training of our trace gas algorithm. In this case, the learning phase consists in fitting the NN parameters so that the output \(\tilde{x}\) calculated by the NN agrees with the desired output \(x\) (real state) for the elements of the learning set. A stochastic gradient descent algorithm has been used, based on the calculation of a cost function \(C(W)\) characterizing the quadratic difference between the desired and calculated outputs:

\[
C(W) = ||x - \tilde{x}||^2
\]

(8)

The NN parameters matrix \(W\), first initialized to random values, was modified after each presentation of one example \(n\) of the learning set:

\[
W^n = W^{n-1} - \eta \nabla C^n
\]

(9)
where $\eta$ is the learning rate. At the end of the learning phase, $W$ was fixed and the operational phase could begin. The learning phase requires a long computation time because of the minimization process. Conversely, the operational phase, which only consists of algebraic computations, is very fast (less than 1 second per spectrum).

### 3.2.4 Least square fitting code

In the algorithm used, given the a priori atmosphere (standard or user-defined), a forward radiative model, globally similar to LBLRTM, generates a radiance spectrum by solving equation 3. The state vector $x$ (vertical volume mixing ratio of the molecule studied) that is best consistent with the measurement vector $y$ is then retrieved by minimizing the $\chi^2$ merit function:

$$\chi^2(x) \propto \frac{1}{\sigma^2} \| f(x) - y \|^2$$  \hspace{1cm} (10)

where $f$ is the forward model and $\sigma$ is the variance of the noise vector $\epsilon$. Because of the non-linearity of the problem, the algorithm relies on an efficient minimization algorithm of the Levenberg-Marquardt type \cite{Press1996}, which is defined by the iterative solution:

$$x^{k+1} = x^k + (KK^T + \gamma^k I)^{-1} K^T [y - f(x^k)]$$  \hspace{1cm} (11)

where $K$ is the Jacobian matrix, of elements

$$K_{ij} = \frac{\partial y_i}{\partial x_j}$$  \hspace{1cm} (12)

and where $\gamma^k$ is the regularization parameter ($k$ denotes the iteration step), which, given an initial guess for the set of fitted parameters $x^k$, is updated at each iteration: After $\chi^2(x^k)$ has been computed, the linear equations are solved for $x^{k+1}$ and $\chi^2(x^{k+1})$ is evaluated. If $\chi^2(x^{k+1})$ is larger than $\chi^2(x^k)$, then $\gamma^k$ is increased by a factor of ten and $x^k$ is not updated. In the reverse situation $x^k$ is updated to $x^{k+1}$, and $\gamma^k$ is decreased by a factor of ten. In both cases, the linear equations are solved with the new parameters and the iterative procedure repeated.

In the general least-square algorithm, a global fit method is used, thus allowing retrieving simultaneously the surface temperature and emissivity, as well as the molecular amounts of all species displaying significant absorption in the spectral region investigated. The Levenberg-Marquardt method also provides a straightforward statistical error ($\sigma$) on each determined parameter. It is defined as:

$$\sigma_j = \sqrt{C_{jj}}$$  \hspace{1cm} (13)

where $C$, the covariance matrix, is the inverse of the curvature matrix of elements:

$$\frac{1}{2} \frac{\partial^2 \chi^2}{\partial x_i \partial x_j}$$  \hspace{1cm} (14)
A drawback of the Levenberg-Marquardt method, however, is that it does not take into account a priori information about the smoothness of the solution.

The least-square algorithm has been used in this work to retrieve total column amounts and to estimate the corresponding errors, as a function of instrumental specifications. This means that in the iterative process the shapes of the vertical distributions of the species are kept fixed and only the column are adjusted by globally multiplying the corresponding concentrations.

### 3.3 Sensitivity and accuracy

To characterize the sensitivity of the instrument to the observed species, its weighting functions $\mathbf{K}$, also called Jacobians or sensitivity kernels, are estimated using forward model simulations of the measurements. $\mathbf{K}$ is defined as the partial derivative of the measurement $\mathbf{y}$ with respect to the variable $\mathbf{x}$ observed, so here as the partial derivative of the radiances with respect to the trace gas vertical concentration profile. $\mathbf{K}$ may be computed using a method of perturbation as

$$
\mathbf{K} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \sim \frac{\Delta \mathbf{y}}{\Delta \mathbf{x}}
$$

where $\Delta \mathbf{x}$ is a perturbation applied to the vertical profile $\mathbf{x}$ of the molecule studied. If the measurement $\mathbf{y}$ is a vector of dimension $N$, and $\mathbf{x}$ is defined on $L$ vertical levels, then $\mathbf{K}$ is a $N \times L$ matrix.

Similarly, the sensitivity of the quantities retrieved to the different parameters (temperature profile, emissivity) is investigated through the calculation of gain functions $\mathbf{G}$, also called contribution functions or Jacobians, defined as the partial derivatives of the retrievals with respect to the input data [Rodgers, 2000], i.e. the measurements $\mathbf{y}$ and, if necessary, the other input parameters $\mathbf{b}$. The resulting functions $\mathbf{G_y}$ and $\mathbf{G_b}$ are then defined as

$$
\mathbf{G_y} = \frac{\partial \mathbf{x}}{\partial \mathbf{y}}, \quad \mathbf{G_b} = \frac{\partial \mathbf{x}}{\partial \mathbf{b}}
$$

If the variable $\mathbf{x}$ is of dimension $P$ (e.g. $P$ vertical levels for a profile retrieval or $P$ atmospheric layers for integrated columns retrievals), $\mathbf{G_y}$ is a $P \times N$ matrix, where each element $G_{y,i}^p = \partial \tilde{x}_p / \partial y_i$ corresponds to the contribution of a given input $y_i$ to the retrieval of the output variable $\tilde{x}_p$, with $i = 1, N$ and $p = 1, P$ (similarly, $\mathbf{G_b}$ is a $M \times P$ matrix).

In the case of the NN retrieval, the $G$ functions can be computed analytically by derivation of the NN transfer function. Rodgers [2000] gives several methods for their determination, depending on the complexity of the problem to solve (on its linearity in particular).

The averaging kernels matrix $\mathbf{A}$ results from the product of the $\mathbf{K}$ and $\mathbf{G_y}$ functions and is defined as

$$
\mathbf{A} = \mathbf{G_y} \mathbf{K} = \frac{\partial \mathbf{x}}{\partial \mathbf{x}}
$$
The rows of this matrix, or averaging kernels, correspond to the sensitivity functions characterizing the retrieved variables (one function by retrieval).

The accuracy depends on the systematic errors which may affect the retrievals, and represents the deviation of the observation from the real value [Keith and Anderson, 2001]. Systematic errors may arise from the observation capabilities of the instrument and from the inversion method used. The error associated with the non-ideal sensitivity of the observing system to the real state is called the smoothing error. It depends on both the variability of the trace gas observed and the deviation between the averaging kernel matrix A and the ideal sensitivity matrix I_P. In the case of a profile retrieval, I_P corresponds to the identity matrix: the maximum sensitivity, i.e. sensitivity equal to 1, is reached at the retrieval level, and it is equal to zero for all other levels. In the case of a partial column retrieval, the ideal sensitivity is equal to 1 in the layer considered, and equal to zero for all other layers. The smoothing error covariance matrix is a P × P matrix defined as

\[ S_s = (A - I_P).S_a.(A - I_P)^T \]  

with S_a the covariance matrix representing the variability of the vertical profile of the trace gas studied.

### 3.4 Information content

For the information content evaluation, the linear case is a good approximation provided the forward problem is linear within the error bounds of the retrieval state [Rodgers, 1990]. The forward model is then defined as

\[ y = f(x, b) + \varepsilon \simeq Kx + \varepsilon \]  

where y is the measurement vector of length N, x is the state vector of length L, and K is the N × L matrix of Jacobians.

The result of the optimal estimation algorithm for the Gaussian linear problem [Rodgers, 2000] is

\[ \hat{x} = R(y, \hat{b}) = R(f(x, b) + \varepsilon, \hat{b}) \]  

with the gain functions

\[ G_y = \frac{\partial \hat{x}}{\partial y} = (K^T S_\varepsilon^{-1} K + S_a^{-1})^{-1} K^T S_\varepsilon^{-1} \]  

where x_a is a priori profile (a priori knowledge), S_a is the associated covariance matrix, and S_\varepsilon is the measurement error covariance matrix.

Without measurement error, the number of independent information contained in the measurement is given by the rank of the K matrix characterizing the sensitivity of the instrument. However, if instrumental noise, or any other perturbation, affects the measurements, some information may become unusable, and the number of independent information (the effective
rank of the problem) may become lower than the rank of $K$. In order to estimate this effective rank, the natural variability of the quantity measured is compared to the covariance of the measurement error. Therefore, a new matrix, corresponding to a scaled $K$, is introduced

$$\tilde{K} = S_e^{-\frac{1}{2}}KS_a^{-\frac{1}{2}}$$

(22)

The number of independent observations that can be done with a precision better than the measurement error is equal to the number of singular values $\lambda_i$ of $\tilde{K}$ which are greater or close to unity. A more precise estimation of this quantity is provided by the number of degrees of freedom of the signal.

$$d_s = \sum_{i=1}^{\lambda_i^2} \frac{\lambda_i^2}{1 + \lambda_i^2} = tr(A)$$

(23)

with $\lambda_i$ the singular values of $S_e^{-\frac{1}{2}}KS_a^{-\frac{1}{2}}$.

An estimation of the vertical resolution that can be achieved for a profile retrieval is provided by the shape of the rows of $A$, the averaging kernels. These functions correspond to the sensitivity of the retrievals to the target species. Their width at half maximum gives an estimate of the vertical resolution.

This evaluation of the vertical information content requires the calculation of $K$, $S_a$ and $S_e$. As mentionned in the previous section, $K$ is calculated using a perturbation method. The measurement error has been considered to be only due to radiometric noise. For this study, a non-correlated random noise $\varepsilon$ is simulated by multiplying a specified noise level to a random value (of standard deviation equal to 1), so that $S_e$ is a diagonal matrix whose diagonal elements represent the variance of the radiometric noise $\varepsilon$. The a priori covariance matrix $S_a$ is more difficult to simulate. $S_a$ corresponds to the variability allowed around the $a$ priori state $x_a$ for the retrieval. The number of degrees of freedom of the problem thus strongly depends on this matrix. Figure 7 shows the $S_a$ matrices estimated for O$_3$, CO and CH$_4$ using realistic atmospheric profiles (outputs from the chemistry-transport model MOZART [Hauglustaine et al., 1998] used to constitute the training database of the NN, as described in section 3.2.3). For CO, only the correlations up to 30 km high are represented in order to show the variabilities in the lower atmosphere. Indeed, for this molecule, the largest variability occurs in the upper atmosphere. For the three species, the variance at a given level and the correlations between the levels are strongly height dependent. In order to simplify the study, we have chosen to consider a constant variance, characteristic of the troposphere since it is the region of interest here, and a Gaussian correlation between the atmospheric levels, as shown on Figure 8.
Figure 7: A priori covariance matrices calculated using tropospheric model outputs connected to climatologies, referred to as model $S_a$. 

27
Figure 8: *A priori* covariance matrices calculated using a constant variance, characteristic of the troposphere, (diagonal elements) and Gaussian correlation between the levels, referred to as constant variance $S_a$. For $O_3$, the variance is fixed to 0.025 ppm, to 0.05 ppm for CO and to 0.03 ppm for CH$_4$, and the correlation between the levels is allowed on 10 km.
4 Results

4.1 Atmospheric variability

To be detected from space, a molecule needs first to have a spectral signature within the in­strumental spectral range (see next paragraph), and second to be present with a sufficient at­mospheric concentration. The latter depends on the sources and sinks of each species, and is subject to important variations due to local variability associated with either natural, or anthropogenic contributions. Representative volume mixing ratios for the target products and their associated lifetimes are provided in Table 3. The later may vary as a function of the season and the location. Values representative of an urban polluted area and globally averaged over the globe are provided.

Table 3. Concentration levels at the boundary layer and atmospheric lifetimes for the target gases. Lifetimes representative of urban conditions for the reactive gases (O$_3$, CO, NO$_2$, H$_2$CO and PAN) were calculated for the Ile de France area using the regional Chimere model [Derognat et al., 2003] (Courtesy to Claude Derognat, Service d’Aeronomie).

<table>
<thead>
<tr>
<th>Data</th>
<th>Mixing ratio</th>
<th>Lifetime urban/continental</th>
</tr>
</thead>
<tbody>
<tr>
<td>O$_3$</td>
<td>50-150 ppbv</td>
<td>1 day/6 days</td>
</tr>
<tr>
<td>CO</td>
<td>50-750 ppbv</td>
<td>1 day/1-2months</td>
</tr>
<tr>
<td>SO$_2$</td>
<td>10pptv-1ppbv</td>
<td>1 day</td>
</tr>
<tr>
<td>H$_2$CO</td>
<td>0.1 - 10ppbv</td>
<td>3 hours/10 hours</td>
</tr>
<tr>
<td>NO$_2$</td>
<td>1ppbv-100ppbv</td>
<td>3 hours/1.5 days</td>
</tr>
<tr>
<td>PAN</td>
<td>100pptv-10ppbv</td>
<td>3 hours/3 days</td>
</tr>
<tr>
<td>H$_2$O</td>
<td>1000ppmv</td>
<td>-</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>3.65 ppmv</td>
<td>100 years</td>
</tr>
<tr>
<td>CH$_4$</td>
<td>1.7 ppmv</td>
<td>9 years</td>
</tr>
</tbody>
</table>

As discussed before, the interest is clearly to measure the target products at least over polluted areas. A set of representative profiles (plotted in Figure 9) for moderately polluted conditions was used for the simulations described in the following chapters.

4.2 Absorption spectral range

In this part, for each species listed in Table 2, we identify the main absorption features in the infrared spectral range. The molecular spectroscopic parameters (wavenumber, intensity) from the latest version of the HITRAN database (version 2000 + 2001 updates) [Rothman et al., 2003] are used for this purpose. The instrument spectral range investigated covers the spectral interval provided in Table 4 (700-2500 cm$^{-1}$) and extends towards 600 cm$^{-1}$ and to 3000 cm$^{-1}$ to encompass a larger spectral domain. PAN, CFC-11 and CFC-12 are provided as cross-sections data due to the complexity of their molecular spectra. PAN was missing from the HITRAN data. Cross-section data were measured by Grant Allen/ John Remedios at Leicester
University (UK) and were kindly provided to us for this study. Aerosols also absorb in this spectral range, depending on their characteristics. Table 4 summarizes the infrared absorption bands of the molecules under investigation (target species), as well as for the other molecules that also absorb in this spectral range. Using Table 4 in conjunction with Table 1 that provides the description of the proposed instrument, it can be seen that all the molecules except H$_2$CO have absorption features within the nominal mode. H$_2$CO, which absorbs between 2750 and 3000 cm$^{-1}$, is outside the spectral range.

4.3 Radiative transfer simulations

Radiative transfer simulations using a line-by-line radiative transfer code parameterized for a nadir-viewing instrument were performed at 0.2 cm$^{-1}$ spectral resolution for all the target species. The individual absorption contributions for each molecule (along with the spectroscopic identifications for the main absorption bands) are represented in Figure 10. Figure 11 is a simulation which included different kinds of aerosols. It can be seen that their contribution is weak, usually they absorb as continuum, but that some variability among the different types of aerosols can be seen.
<table>
<thead>
<tr>
<th>Molecule</th>
<th>Center of the band cm(^{-1})</th>
<th>Absorption range cm(^{-1})</th>
<th>Vibrational attribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target molecules for this study</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O(_3)</td>
<td>710</td>
<td>550-900</td>
<td>(\nu_2)</td>
</tr>
<tr>
<td></td>
<td>1043</td>
<td>919-1243</td>
<td>(\nu_3)</td>
</tr>
<tr>
<td></td>
<td>1070</td>
<td>940-1280</td>
<td>(\nu_1)</td>
</tr>
<tr>
<td></td>
<td>2105</td>
<td>1880-2320</td>
<td>(2\nu_1, 2\nu_3, \nu_1+\nu_3)</td>
</tr>
<tr>
<td></td>
<td>2800</td>
<td>2680-2820</td>
<td>(\nu_1+\nu_2+\nu_3)</td>
</tr>
<tr>
<td>CO</td>
<td>2100</td>
<td>2000-2260</td>
<td>1-0</td>
</tr>
<tr>
<td>SO(_2)</td>
<td>1151</td>
<td>1080-1260</td>
<td>(\nu_1)</td>
</tr>
<tr>
<td></td>
<td>1361</td>
<td>1310-1400</td>
<td>(\nu_3)</td>
</tr>
<tr>
<td></td>
<td>2499</td>
<td>2440-2530</td>
<td>(\nu_1+\nu_3)</td>
</tr>
<tr>
<td>H(_2)CO</td>
<td>2780-2874</td>
<td>2700-&gt;3000</td>
<td>(\nu_1, \nu_5)</td>
</tr>
<tr>
<td>NO(_2)</td>
<td>648</td>
<td>650-880</td>
<td>(\nu_2)</td>
</tr>
<tr>
<td></td>
<td>1621</td>
<td>1550-1760</td>
<td>(\nu_3)</td>
</tr>
<tr>
<td></td>
<td>2910</td>
<td>2850-2940</td>
<td>(\nu_3+\nu_1)</td>
</tr>
<tr>
<td>PAN</td>
<td>-</td>
<td>750-1900</td>
<td>Xsections</td>
</tr>
<tr>
<td>H(_2)O</td>
<td>1595</td>
<td>(\leq600-&gt;3000)</td>
<td>(\nu_2)</td>
</tr>
<tr>
<td>CO(_2)</td>
<td>618.1</td>
<td>(\leq600-850)</td>
<td>(2\nu_2(0))</td>
</tr>
<tr>
<td></td>
<td>667.3</td>
<td>(\leq600-850)</td>
<td>(\nu_2(1))</td>
</tr>
<tr>
<td></td>
<td>720.5</td>
<td></td>
<td>(\nu_1)</td>
</tr>
<tr>
<td></td>
<td>1886</td>
<td>1870-1990</td>
<td>(4\nu_2(0)/2)</td>
</tr>
<tr>
<td></td>
<td>2094</td>
<td>2000-2150</td>
<td>(\nu_1+2\nu_2(1/2))</td>
</tr>
<tr>
<td></td>
<td>2137</td>
<td>2000-2150</td>
<td>(2\nu_1)</td>
</tr>
<tr>
<td></td>
<td>2349</td>
<td>2000-2700</td>
<td>(\nu_3)</td>
</tr>
<tr>
<td>CH(_4)</td>
<td>1306.2</td>
<td>900-1970</td>
<td>(\nu_4)</td>
</tr>
<tr>
<td></td>
<td>3020.3</td>
<td>2000-&gt;3000</td>
<td>(\nu_3)</td>
</tr>
<tr>
<td>CFC-11</td>
<td>850</td>
<td>810-880</td>
<td>(\nu_4)</td>
</tr>
<tr>
<td></td>
<td>1082</td>
<td>1050-1120</td>
<td>(\nu_1)</td>
</tr>
<tr>
<td>CFC-12</td>
<td>922</td>
<td>850-950</td>
<td>(\nu_6)</td>
</tr>
<tr>
<td></td>
<td>1160</td>
<td>1050-1200</td>
<td>(\nu_8)</td>
</tr>
<tr>
<td><strong>Other absorbing molecules</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N(_2)O</td>
<td>1285</td>
<td>1210-1340</td>
<td>(\nu_1)</td>
</tr>
<tr>
<td></td>
<td>2222</td>
<td>2120-2270</td>
<td>(\nu_3)</td>
</tr>
<tr>
<td>HNO(_3)</td>
<td>648.8</td>
<td>615-678</td>
<td>(\nu_6)</td>
</tr>
<tr>
<td></td>
<td>763.2</td>
<td>722-810</td>
<td>(\nu_8)</td>
</tr>
<tr>
<td></td>
<td>879.11</td>
<td>816-960</td>
<td>(\nu_5)</td>
</tr>
<tr>
<td></td>
<td>896.85</td>
<td>816-960</td>
<td>(2\nu_0)</td>
</tr>
<tr>
<td></td>
<td>1205.7</td>
<td>1165-1233</td>
<td>(\nu_8+\nu_0)</td>
</tr>
<tr>
<td></td>
<td>1303.5</td>
<td>1098-1388</td>
<td>(\nu_4)</td>
</tr>
<tr>
<td></td>
<td>1325.7</td>
<td>1098-1388</td>
<td>(\nu_3)</td>
</tr>
<tr>
<td></td>
<td>1709.57</td>
<td>1650-1770</td>
<td>(\nu_2)</td>
</tr>
</tbody>
</table>
Figure 10: Radiative transfer simulations (transmittance units) in the 600-3000 cm\(^{-1}\) spectral range for each target molecule.

Complete line-by-line simulations including all absorbing gases were then performed, both for atmospheric situations representative of a non-polluted area (over Pacific Ocean) and for moderately polluted area (North America). Figure 12 provides the simulated atmospheric spectra (in transmittance units for a polluted atmosphere) along with the detailed absorption contributions of the strong and weak absorbing molecules.

It can be seen that the main absorption features in the 600-3000 cm\(^{-1}\) spectral range are associated with rotational transition lines related to the fundamental vibrational bands of H\(_2\)O, CO\(_2\), O\(_3\), N\(_2\)O, CH\(_4\), and CO. All these species combine intense absorption lines in this spectral range with high levels of atmospheric concentration. Weaker absorption contributions (trans-
Figure 11: Radiative transfer simulations in the 600-3000 cm$^{-1}$ spectral range for 5 types of aerosols (rural, urban, desert, maritime, and volcanic) (top). Relative differences with a simulated spectrum for aerosol-free conditions (bottom). Courtesy to Michel Kruglanski (IASB-BIRA, Belgium).
Figure 12: Atmospheric spectrum (in transmittance units) in the 600-3000 cm$^{-1}$ spectral range simulated for North America conditions. Radiative transfer simulations for absorption contributions due to H$_2$O, strong (middle) and weak (absorption lower than 10%) absorbers are also provided (bottom).
mittance lower than \(10\%\) are due to the presence of CFC11, CFC12, HCFC22, HNO\(_3\), OCS, NO and NO\(_3\) and \(\text{H}_2\text{CO}\). All these molecules either have weak infrared absorptions, or have low atmospheric mixing ratios.

4.4 Accuracies in column retrievals

The accuracy at which the concentration of a product can be retrieved from its spectral signature depends on the intensity of the absorption, the location of the absorption (interferences with other gases), and the number of channels available with useful information. It also depends on the instrumental specifications, through the spectral resolution and noise level. The sensitivity of the column retrievals to the instrumental noise and spectral resolution is tackled by means of the least-square fitting algorithm described earlier, using as reference spectra the line-by-line simulations performed at several spectral resolutions (ranging from \(0.2\) to \(1.25\ \text{cm}^{-1}\)) and \(6\) different radiometric noises. It is worth stressing at this point that the a priori atmosphere taken as input into the fitting algorithm is similar to that used for generating the reference radiance spectra. Also, a single parameter (the abundance of the targeted species) is fitted at a time, all others being kept to their initial and perfectly known values. The possible effect of an imperfect knowledge in the temperature or interfering gas profile, of clouds etc. has thus not been investigated, so that the errors in the retrieval of total columns necessarily originates from the instrumental performances. The estimated accuracy given in the following is the \(1\sigma\) value calculated by equation 13.

4.4.1 Strong infrared absorbers

For several of the main absorbers (O\(_3\), CO and CH\(_3\)), the retrievals were performed on microwindows validated through our IMG-IASI experience. For the other molecules (H\(_2\)O and CO\(_2\)), the microwindows selection was done more arbitrarily, by choosing a region characterized by strong but unsaturated lines, where also few interfering lines of other species are lying. A summary of the impact of the instrumental performances on the accuracy of the retrieved total column is provided in the following plots for the main infrared absorbers, i.e H\(_2\)O (Figure 13), CO\(_2\) (Figure 14), O\(_3\) (Figure 15), CH\(_3\) (Figure 16), and CO (Figure 17). The top of the Figure illustrates the forward and fitted spectra both for threshold (in terms of noise and spectral resolution) and goal conditions, and provides the associated retrieved column accuracy. It is seen that for all molecules, the target accuracy is reached for several combinations of noise level and spectral resolution. In general, the retrievals of columns appear to be more sensitive to the noise level than to the spectral resolution, though a significant effect of the latter is observed for CH\(_3\). This could be explained, in part, by the strong overlap of the CH\(_3\) \(\nu_3\) band with the N\(_2\)O \(\nu_3\) band. As a result of the ‘perfect atmosphere’ hypothesis, the accuracies given for the main absorbers are likely underestimated. This is further illustrated in the bottom part of the same Figures, which shows the result of the least square fitting on an observed IMG spectrum, using the same spectral channels as for the simulation, and the derived accuracy. The parameters that were fitted are also indicated on the plot.
Figure 13: Combined impact of instrumental characteristics on the accuracy of the H₂O retrieved columns, for low resolution/high noise (threshold - see Table 1) and higher resolution/low noise (goal - see Table 1). The bottom plot provides the result of the least square fitting on an observed IMG spectrum, using the same spectral channels.
Figure 14: Combined impact of instrumental characteristics on the accuracy of the CO$_2$ retrieved columns, for low resolution/high noise (threshold -see Table 1) and higher resolution/low noise (goal - see Table 1). The bottom plot provides the result of the least square fitting on an observed IMG spectrum, using the same spectral channels.
Figure 15: Combined impact of instrumental characteristics on the accuracy of the \( \text{O}_3 \) retrieved columns, for low resolution/high noise (threshold - see Table 1) and higher resolution/low noise (goal - see Table 1). The bottom plot provides the result of the least square fitting on an observed IMG spectrum, using the same spectral channels.
Figure 16: Combined impact of instrumental characteristics on the accuracy of the CH₄ retrieved columns, for low resolution/high noise (threshold -see Table 1) and higher resolution/low noise (goal - see Table 1). The bottom plot provides the result of the least square fitting on an observed IMG spectrum, using the same spectral channels.
Figure 17: Combined impact of instrumental characteristics on the accuracy of the CO retrieved columns, for low resolution/high noise (threshold - see Table 1) and higher resolution/low noise (goal - see Table 1). The bottom plot provides the result of the least square fitting on an observed IMG spectrum, using the same spectral channels.
4.4.2 Weak infrared absorbers

Specific simulations were performed for the weak absorbing gases, such as NO$_2$, H$_2$CO, PAN and SO$_2$. Radiative transfer simulations were first performed at very high spectral resolution to select the relevant micro-windows that were used in the retrieval step. The forward simulations show that the absorption bands characteristic of NO$_2$, SO$_2$, H$_2$CO and PAN are too weak to be detected under normal atmospheric conditions, with the instrumental characteristics defined for the post-MSG sounder. Additional radiative transfer simulations were accordingly performed for hypothetical high abundances of these species (a scaling factor was applied to the profile), such as to obtain a signal larger than the specified 0.2 K radiometric noise. The scaling factors ranged from about 5 for the detection of NO$_2$ to about 100 for that of SO$_2$. According to Table 3, these scaling factors are compatible with mixing ratios in the boundary layer representative of polluted locations for NO$_2$ and PAN. For H$_2$CO, at least 3 times the abundance of Table 3 is needed to obtain a signal that distinguishes from the noise level.

An illustration of simulations performed for the weak absorbers in shown in Figure 18 with the example of NO$_2$. From the Figure it can be seen that the retrieval of abundances is possible, provided that the atmosphere is well known. Because of the weakness of the absorption, it is also expected that the highest resolution and low noise level will be needed for the detection and the measurement of these species.

Forward calculations were also performed for SO$_2$ vertical distributions representative of different atmospheric conditions. Even in highly polluted areas, it is seen that the SO$_2$ absorption bands hardly distinguishes from the noise using current specifications. Only in the case of volcanic eruptions, where the SO$_2$ concentration released in the atmosphere reaches unusually high values, becomes the signal significant (Figure 19). In these conditions, both the $\nu_3$ and $\nu_1$ band appear clearly. Because it lies in a much favourable atmospheric window, the $\nu_1$ band at 1151 cm$^{-1}$ is expected to be more adequate for SO$_2$ retrievals.

4.5 Vertical information content

A diagnostic of information content following Rodgers [2000] was performed using the theory described in part 3.4. Three species (CO, O$_3$ and CH$_4$) were found suitable to be retrieve with a higher vertical resolution than total column. The impact of the spectral resolution and instrumental noise characteristics on the independent pieces of information that can be retrieved on the vertical has been accessed.

Four spectral resolutions have been tested, first with a constant noise NeDT equal to 0.2 K, and then with NeDT varying between 0.05 and 0.5 K: the threshold resolution of 1.25 cm$^{-1}$, the goal resolution of 0.5 cm$^{-1}$, plus two additional resolutions: 0.75 cm$^{-1}$ (‘reduced performance’) and 0.2 cm$^{-1}$ (‘improved performance’). The atmospheric profiles depicted on Figure 9, corresponding to a polluted area, have been used for all simulations. For each molecule, the studies are undertaken on spectral channels that were selected for IASI [Clerbaux et al., 1998]. It should here be kept in mind that all the calculations were undertaken using a con-
Figure 18: This Figure shows the sensitivity of the post-MSG instrument to the NO\textsubscript{2} main spectroscopic features. The left panel shows the radiative transfer simulations performed for the standard atmosphere studied here (North America conditions) and for the improved resolution mode (0.2 cm\textsuperscript{-1}), without (black line) and with (blue line) the 0.2 K nominal noise. The middle and left panels are simulations corresponding to a high level of NO\textsubscript{2} (5 times the nominal concentration), at both the improved (0.2 cm\textsuperscript{-1}) and threshold (1.25 cm\textsuperscript{-1}) spectral resolutions. In the three plots, a reference simulation without NO\textsubscript{2} is given by a gray line. The red lines and triangles are the results of the spectral fits made using the least-square algorithm.
4.5.1 Sensitivity of the instrument

The Jacobian matrices $\mathbf{K}$ characterizing the sensitivity of the measurements at the threshold resolution (1.25 cm$^{-1}$) to O$_3$, CO, and CH$_4$ are plotted in Figure 20. The magnitude of the Jacobian depends on the intensity of the absorption and the emission. In the case of nadir-viewing infrared, the sensitivity reaches a maximum in the free troposphere, between 6 and 10 km, and rapidly decreases below 2 km. For O$_3$, there is more information above 25 km than for the other molecules due to strong concentrations in the stratosphere (ozone layer). The instrument has a maximum sensitivity in the free troposphere due to the Lorentz collisional broadening of the spectral lines, which increases with pressure. The reduced sensitivity to the lower layers of the atmosphere (atmospheric boundary layer) is a common problem to all nadir-viewing IR remote sounders, mostly explained by the lack of thermal contrast with the surface.

The rank of $\mathbf{K}$ gives a first estimate of the independent information that can be observed by the instrument in idealized conditions (no error). A singular value decomposition (SVD) highlights the principal modes of variation of the spectra (singular vectors of $\mathbf{K}$) with their associated sensitivity (singular values). They are shown for O$_3$ in Figure 21 at the threshold and improved spectral resolutions. The first mode of variation constitutes the largest contribution to the global variability, with a maximum sensitivity in the free troposphere which rapidly decreases below 2 km high and in the stratosphere. The sensitivity to the second mode of variation is about 25 times lower than the sensitivity to the first one, and has two peaks of sensitivity, located slightly
above and below the maximum of the first mode. The sensitivity to the third is then about 100 times lower. At the improved spectral resolution (0.2 cm$^{-1}$), the first three modes of variation are comparable, with however a larger sensitivity to the secondary modes.

This analysis verifies that nadir IR measurements contain height dependent information about the trace gases, with a maximum sensitivity in the free troposphere.

### 4.5.2 Number of independent information

The measurement error and the variability of the constituent studied are taken into account by considering the scaled $K$ matrix $\tilde{K}$. A SVD is used in order to highlight the main sensitivity patterns and calculate the number of independent information contained in the spectra. The number of independent information corresponds to the number of singular values $\lambda_i$ of $\tilde{K}$ that are greater or close to 1. The SVD of $\tilde{K}$ for $\text{O}_3$ is shown in Figure 22. At the threshold resolution, the three first modes have $\lambda_1$ larger or close to 1, implying that about three independent pieces of information could be deduced from the measurements. The first mode corresponds...
Figure 21: Six first singular vectors of the K matrix calculated for O$_3$. The corresponding singular values are indicated on the top of the figures, as well as their relative importance with respect to the first mode.
Figure 22: Six first singular vectors of the $\bar{K}$ matrix calculated for O$_3$. The corresponding singular vectors are indicated on top of the Figures, as well as their relative importance with respect to the first mode.
to variability in the free troposphere, the second mode to variability in the lower troposphere and in the lower stratosphere, and the third one to variability in the free troposphere and in the stratosphere. The number of degrees of freedom calculated using Equation 23 is equal to 2.46. Improving the spectral resolution increases the sensitivity to the second and third modes. At 0.2 cm$^{-1}$ spectral resolution, the number of degrees of freedom increases by about 0.4 points and is equal to 2.83.

The SVD of $\mathbf{K}$ for CO is shown in Figure 23. At the threshold resolution, only the first mode has a singular value larger than 1, and the second one is similar to 1. As for O$_3$, the first mode corresponds to variability in the free troposphere, and the second mode to variability in the lower troposphere and in the lower stratosphere. The number of degrees of freedom is estimated to 1.38. At the improved resolution, the sensitivities to the first and second modes increase significantly, and the calculated number of degrees of freedom increases to 2.24. The impact of the spectral resolution is more important for CO retrieval than for O$_3$.

The SVD of $\mathbf{K}$ for CH$_4$ is shown in Figure 24. It is relatively close to the SVD calculated for CO. At the threshold resolution, only the first mode has $\lambda_1$ larger than 1, but the second mode is still quite important. The number of degrees of freedom is estimated to 1.2. At the improved resolution, the sensitivities to the first and second modes slightly increase, and the number of degrees of freedom is estimated to 1.35. As for O$_3$, the impact of the spectral resolution seems to be relatively small.

4.5.3 Impact of the noise level

The impact of noise on the number of independent vertical information that can be extracted from the observation is shown for O$_3$, CO and CH$_4$ in Figures 25-27 respectively, for 4 different spectral resolutions. For the three molecules, it is seen that the number of degrees of freedom strongly depends on the noise level. The effect of spectral resolution is on the contrary negligible, especially for O$_3$ and CH$_4$ at noise larger than 0.1 K. For CO, however, the number of degrees of freedom increases as the spectral resolution is improved.

4.5.4 Vertical resolution

The vertical resolution that can be achieved is studied using the averaging kernels calculated in the case of a Gaussian linear retrieval [Rodgers, 2000]. The averaging kernels obtained for O$_3$, CO and CH$_4$ are plotted in Figures 28, 29, and 30 respectively.

As previously mentioned, the ideal averaging kernel for the retrieval at a given level is equal to one at this level and to zero elsewhere (Dirac functions). The kernels obtained here are spread peak functions, with maxima which do not always correspond to the retrieval level, implying strong contributions from the neighbouring levels.

For O$_3$, two peaks are obtained in the troposphere at all spectral resolutions, located around 4 and 11 km high, and a third one is obtained in the stratosphere, around 15 km high at the
Figure 23: Six first singular vectors of the $\mathbf{K}$ matrix calculated for CO. The corresponding singular vectors are indicated on top of the Figures, as well as their relative importance with respect to the first mode.
Figure 24: Six first singular vectors of the $\hat{K}$ matrix calculated for CH$_4$. The corresponding singular vectors are indicated on top of the Figures, as well as their relative importance with respect to the first mode.
Figure 25: Number of degrees of freedom calculated for O$_2$ at different spectral resolutions as a function of the noise level. The threshold (0.2 K) and goal (0.15 K) noise levels are indicated by a red and a green arrow respectively.

Figure 26: Number of degrees of freedom calculated for CO at different spectral resolutions as a function of the noise level. The threshold (0.2 K) and goal (0.15 K) noise levels are indicated by a red and a green arrow respectively.
Figure 27: Number of degrees of freedom calculated for CH$_4$ at different spectral resolutions as a function of the noise level. The threshold (0.2 K) and goal (0.1 K) noise levels are indicated by a red and a green arrow respectively.

improved spectral resolution. As already highlighted by the study of the $\mathbf{K}$ matrix, improving the spectral resolution increases the amount of information, particularly in the stratosphere.

For CO, the averaging kernels all peak in the free troposphere, around 8 km high, at the threshold resolution. Improving the spectral resolution results in a splitting of the single peak, thereby providing information on the lower and upper troposphere.

For CH$_4$, the averaging kernels are characterized by a single peak around 8 km high, even at the improved spectral resolution.

The width at half maximum of the kernels provides an estimate of the vertical resolution that may be achieved for the retrieval. It is estimated to be 8-10 km.
Figure 28: Averaging kernels obtained for a linear retrieval of ozone at different spectral resolutions. The colored dots correspond to the retrieval levels associated with each kernel. The corresponding number of degrees of freedom (at threshold noise level of 0.2 K) are indicated on top of the Figures.
Figure 29: Averaging kernels obtained for a linear retrieval of CO at different spectral resolutions. The colored dots correspond to the retrieval levels associated with each kernel. The corresponding number of degrees of freedom (at threshold noise level of 0.2 K) are indicated on top of the Figures.
Figure 30: Averaging kernels obtained for a linear retrieval of CH₄ at different spectral resolutions. The colored dots correspond to the retrieval levels associated with each kernel. The corresponding number of degrees of freedom (at threshold noise level of 0.2 K) are indicated on top of the Figures.
4.6 Global distributions measured by current remote sensors

Global scale distributions obtained from the analysis of measured infrared data have been reported for CO (IMG, MOPITT), O\(_3\) (TOVS, IMG), CH\(_3\) (IMG) and SO\(_2\) (AIRS). The following section illustrates column distributions for these target species.

4.6.1 CO

Global distributions of carbon monoxide total column amounts have been obtained from the IMG spectra using the IASI retrieval algorithm package \cite{Clerbaux:1999, Hadji-Lazaro:1999, Clerbaux:2001, Hadji-Lazaro:2001}. Figure 31 provides the CO distribution retrieved for April 1997.

In \textit{Clerbaux et al.} \cite{Clerbaux:2002} the performance of different retrieval codes (MOPITT, TES) on a common IMG dataset are discussed and the main sources of errors on the retrieved columns are identified. The errors in the temperature profile used as well as uncertainty on the instrumental response function are shown to be the main contributions to explain the discrepancies observed between the results provided by the different groups. Other possible causes of errors are uncertainties associated with water vapor, cloud and aerosol loading contents. Our current guess for averaged accuracy (including all sources of errors) is estimated to 10%.

Retrievals from IMG were difficult to validate using independent measurements, due to the lack of measurements sites for total columns. Some of the NDSC ground-based network stations (those using solar tracking Fourier transform spectrometer) provide CO as a product. But measurements are sparse, and rarely colocated with IMG measurements. Available NDSC stations for the April period are also represented on Figure 31 (enlarged colored dots).

Since the failure of IMG, MOPITT has provided an extensive dataset for CO \cite{Deeter:2003}. Current available products were obtained using the thermal-band channels and are provided both on altitude levels and as total column products. The spatial coverage is higher for MOPITT due to its ability to scan across track, whereas IMG is measuring along the track only. The MOPITT uses a broader spectral range, centered on the main CO absorption lines, whereas the IMG provides high resolution measurements. Both instruments lack sensitivity to the boundary layer and have their maximum sensitivity in the free troposphere. An example of CO distribution as retrieved from the MOPITT data is provided in Figure 32 (with the same color scale as for the IMG plot), also for April but in the year 2000. Current estimates for accuracies are similar to those of IMG \cite{Emmons:2003}.

4.6.2 Ozone

\(O_3\) total column amount

Global distributions of \(O_3\) total column amounts have been obtained from the IMG spectra using the IASI retrieval algorithm package \cite{Turquety:2002}. Total column distributions obtained for the April 1-10 (1997) IMG period are plotted in Figure 33. We derived a mean
Figure 31: CO total column distribution retrieved from the IMG measurements. Data were averaged onto a regular 5° x 5° grid, in latitude and longitude, and were averaged in time over the period from the 1st to the 10th of April 1997. Total column amounts measured by the NDSC ground-based network are represented on the same plot (enlarged dots).

Figure 32: Mopitt CO column distribution averaged over data recorded in April 2000.
accuracy of 3% on these products (ie. for 0.5 cm$^{-1}$ spectral resolution and IASI noise specifications).

Figure 33: Ozone total column distribution retrieved from the IMG measurements. Data were averaged onto a regular $5^\circ \times 5^\circ$ grid, in latitude and longitude, and were averaged in time over the period from the 1$^{st}$ to the 10$^{th}$ of April 1997.

Thanks to the availability of the Total Ozone Mapping Spectrometer (TOMS) instrument onboard the ADEOS plateform, a direct validation of these values with a large number of independent data can be performed. The co-localized total columns were compared within a ± 0.5$^\circ$ area, and a ± 2 seconds time period. After post-processing of the retrieved IMG columns (cloud filter and consistency check), a very good agreement between the two distributions is reached with a rms error $\sim$ 25 DU and a correlation $\sim$ 0.9. The observed differences are mostly explained by the different observation methods used. Figure 34 illustrates the agreement between the latitudinal O$_3$ total column variability observed by both instruments (TOMS, IMG) with their associated standard deviation.

**O$_3$ tropospheric amount**

Using the same algorithm as for total column retrievals, tropospheric ozone distribution are retrievable. The tropospheric column for ozone derived from the IMG spectra for the April period is represented in Figure 35. A 15% accuracy is estimated on average. It may, however, exceed 20% for extreme cases (e.g. for elevate mxr confined in the planetary boundary layer).
Figure 34: Total O$_3$ derived from colocated IMG/ADEOS and TOMS/ADEOS measurements (with associated standard deviations).

Ozonesonde measurements have been used in order to validate the partial O$_3$ columns.

Figure 35: Ozone tropospheric column distribution retrieved from the IMG measurements. Data were averaged onto a regular 5° x 5° grid, in latitude and longitude, and were averaged in time over the period from the 1st to the 10th of April 1997.
4.6.3 Methane

New results have also recently been obtained for retrieval of CH$_4$ from the IMG spectra for the April period. Total column amounts are represented in Fig. 36. The distribution is characterized by a strong North-South gradient as expected, with some smaller regional variations. Some unexpected values are obtained above deserts, probably because of a wrong emissivity used in the retrievals.

![CH4 - IMG data averaged on the period April 1-10, 1997 - 5°x5°](image)

Figure 36: Methane total column distribution retrieved from the IMG measurements. Data were averaged onto a regular 5° x 5° grid, in latitude and longitude, and were averaged in time over the period from the 1st to the 10th of April 1997.

4.6.4 SO$_2$

Recent measurements of the Mt Etna eruption in October 2002 by the AIRS instrument have proven the feasibility to monitor SO$_2$ from a nadir infrared sounder as shown in Figure 37. It should be noted that in that particular case a two-channel selection method was necessary to discriminate the SO$_2$ signal (measured in the 6.7 μm band) from that of water vapor, because the AIRS instrument lacks the useful channel in the atmospheric window around the SO$_2$ 4.7 μm band region.
5 Conclusion

This report aims to study the capabilities of a Post-MSG infrared sounder flying on a geostationary orbit for monitoring atmospheric composition and for chemistry applications. It relies on:

- **Chemistry requirements** for each product. These were derived from previous studies [Lelieveld et al., 2003; Golding et al., 2003], and are provided in terms of horizontal, vertical, temporal resolutions, and accuracy. It varies as a function of the species and the scientific questions to adress. Lower and upper limits for requirements were defined as threshold (value required before an observation can be used with any benefit) and target (value beyond which there is no further benefit from improvements).

- **Instrumental specifications** for a post-MSG infrared sounder. These were provided by Eumetsat during the study. Instrumental characteristics were provided in terms of spatial and temporal resolution, spectral coverage, noise and spectral resolution. These specifications were defined in priority to meet the needs for NWP global and regional applications, and were not driven by chemistry requirements. Upper and lower limits for allowed instrumental performances are referred as goal and threshold.

This study was performed using forward model calculations (using the above mentioned instrumental characteristics) and inverse retrieval codes (neural network and least-squares fitting) developed in the framework of the preparation of the IASI missions, as well as an information
Infrared absorbing species can be divided in four categories, represented in Figure 38:

1. The species combining low spatial/temporal atmospheric variability (associated with a long atmospheric lifetime) and high infrared absorptions (associated with high current concentration levels and intense absorptions in the infrared). These stable gases, represented on the lower right part of the figure, are often referred to as climate gases or greenhouse gases. For measuring these species, a high accuracy is needed, as a small signal variability has to be retrieved from a larger signal contribution.

2. The species combining low spatial/temporal atmospheric variability (associated with a long atmospheric lifetime) and low infrared absorption (associated with low current concentration levels). These species are represented on the lower left part of the figure. The critical point is the possibility to detect the species, i.e. whether its spectral signature can be detected.

3. The species combining high spatial/temporal atmospheric variability (associated with a short atmospheric lifetime) and high infrared absorptions (associated with high current concentration levels and intense absorptions in the infrared). These molecules are represented on the upper right part of the figure. These species, for which accuracy requirements may be relaxed to the spatial/temporal atmospheric variability, are key species to be measured from space.

4. The species combining high spatial/temporal atmospheric variability (associated with a short atmospheric lifetime) and low infrared absorptions (associated with low current concentration levels and/or limited absorption bands in the infrared). These species are represented on the upper right part of the figure. These molecules can only be detected when high atmospheric abundances occur and for optimized instrumental specifications.

### 5.1 Compliance with requirement per species

**Ozone - Breakthrough requirement: 5%/10km/tropo/1h - 50km/6h**

Ozone has several absorption bands throughout the instrumental spectral range. The most suitable channels for ozone retrieval are located in the atmospheric window in the $\nu_3$ absorption bands around 1050 cm$^{-1}$, and are free from interfering constituents. Most of the radiance signal comes from the free troposphere, although a non-negligible part comes from the stratosphere due to the high levels of concentration. The atmospheric lifetime is highly variable depending on the location of the measurement, ranging from less than one day in the lower layers of the atmosphere to two weeks in the stratosphere. The atmospheric spatial variability reaches 300% and 1500% for total and tropospheric ozone respectively.

Due to its high atmospheric variability combined with high infrared absorption, accurate measurements can be performed both for total columns and tropospheric columns. The feasibility
Figure 38: Infrared absorbing molecules as a function of the intensity of their thermal infrared absorption and their atmospheric variability.
has been demonstrated on real atmospheric data provided by IMG/ADEOS. The goal accuracy is met for total column retrievals. For tropospheric columns, the threshold accuracy may be exceeded for specific situations, e.g. when high levels of O₃ are limited to the planetary boundary layer (errors up to 20%). The number of independent pieces of information varies between 2 (NeDT= 0.2K and spectral resolution= 1.25 cm⁻¹) and 2.5 (NeDT= 0.15K and spectral resolution= 0.2 cm⁻¹). The horizontal target and threshold resolution requirements match with the goal and threshold instrument specifications.

Pollution: CO - Breakthrough requirement: 10%/10km/tropo/1h
Carbon monoxide has a single absorption band (1-0 transition) in the instrumental spectral range, around 2000 cm⁻¹. Channels have to be carefully selected due to interfering absorption features due to H₂O, NO₂ and O₃. Only the tropospheric contribution is seen in the radiance signal, with a maximum that peaks around 6-8 km. The atmospheric lifetime ranges from less than one day to two months in pollution-free areas. The atmospheric variability reaches 300%. The feasibility of retrieving total columns has been demonstrated on real atmospheric data provided by IMG/ADEOS. As for ozone, CO can easily be measured with a useful accuracy (between target and threshold) using the above mentioned instrumental characteristics. The number of independent pieces of information varies between 1.3 (NeDT= 0.2K and spectral resolution= 1.25 cm⁻¹) and 2.2 (NeDT= 0.15K and spectral resolution= 0.2 cm⁻¹). The horizontal target and threshold resolution requirements match with the goal and threshold instrument specifications.

Pollution: SO₂ - Breakthrough requirement: 50%/10km/tropo/1h
SO₂ produces only a very weak signal in the nadir mode (transmittance \( \leq 1\% \)). The strongest absorption band, which is the \( \nu_3 \) band at 1361 cm⁻¹ is further overlapped by the strong H₂O \( \nu_3 \) band, making the detection of the molecule by a nadir sounder difficult. Column retrievals are only possible when high mixing ratios occur, which are possibly not met above polluted areas. Strong signatures associated with volcanic plumes can be detected with the required accuracy as demonstrated by the AIRS Science Team. The horizontal target and threshold resolution requirements match with the goal and threshold instrument specifications.

Pollution: H₂CO - Breakthrough requirement: 50%/10km/tropo/1h
The H₂CO absorption band is located around 2750 cm⁻¹, outside the nominal mode of the present post-MSG GEO nadir sounder. The molecule could therefore only be detected by extending the spectral interval towards larger wavenumbers. This band is overlapped by N₂O, CH₄ and H₂O lines but the atmospheric window remains relatively clear. Furthermore H₂CO makes hardly any contribution to a nadir spectrum recorded over unpolluted areas but contributes slightly if pollution levels exceeding 30 ppbv are reached. For the latter case, total column are retrievable with the required accuracy. The horizontal target and threshold resolution requirements match with the goal and threshold instrument specifications.

NO, NO₂ - Breakthrough requirement: 50%/10km/tropo/1h
The most prominent absorption feature of NO₂ in the infrared region between 600 and 3000 cm⁻¹ is the \( \nu_3 \) band at 1621 cm⁻¹. Among the weakly absorbing species (transmittance \( \leq \)
10%), NO₂ is the strongest absorber. However, it should be stressed that the strong ν₃ band lies in a region dominated by the water vapor lines. Despite of this, the NO₂ line at about 1630 cm⁻¹ produces a very distinctive feature if low noise and high-resolution (0.2 cm⁻¹) are combined. The measurement of NO₂ by a nadir sounder will be strongly dependent on the spectral resolution, the noise level as well as the accurate modelling of the water vapor content. Its atmospheric variability is high, ranging from a few hours to 1.5 days. From our simulations, it appears the concentration should exceed 6 ppbv in the planetary boundary layer to detect NO₂ and provide a column product. The horizontal target and threshold resolution requirements match with the goal and threshold instrument specifications. The NO contribution was found to be too weak to be detected.

**Pollution: PAN - Breakthrough requirement: 50%/10km/tropo/1h**
The PAN absorption bands are located between 600 and 1800 cm⁻¹, with a maximum at 1740 cm⁻¹. A band of medium intensity is also located in the atmospheric window, at about 785 cm⁻¹. Because of the broad character of the PAN absorption features (as for CFC11 and CFC12), it is expected that the latter band at 785 cm⁻¹ will be more suitable for detecting and measuring the species. Its atmospheric variability is high, ranging from a few hours to 3 days. As for the other fast reacting species only total column amounts may be retrieved above polluted areas if low noise and high spectral resolution are combined. Our guess for detection limits is 5 ppbv. The horizontal target and threshold resolution requirements match with the goal and threshold instrument specifications.

**Aerosols - Breakthrough requirement: see Table 2**
No specific simulations were undertaken for aerosols during the course of this study. Sensitivity studies performed in the framework of IASI [Kruglanski and DeMazière, 2002] have shown the relative variations in the radiance spectrum due to the presence of different populations of aerosols (urban, maritime, volcanic, desert sand). Most of the impact consists of broad-band absorption features in the atmospheric window. Sand-dust signal in IMG spectra recorded over Sahara were reported in Gribanov et al. [2001] around 1157 cm⁻¹. As for SO₂, the AIRS Science Team has recently reported about the detection of aerosols particules above the Etna eruption plume.

**Meteorology: H₂O - Breakthrough requirement: 5%/10km/2km/1h**
**Meteorology: T - Breakthrough requirement: 2K/10km/2km/1h**
H₂O and temperature retrievals were not investigated in this study. Several teams have recently published their work on temperature and humidity retrievals either for IASI [Prunet et al., 2001; Aires et al., 2002] or for IMG [Amato et al., 1999; Lubrano et al., 2000]. These studies have demonstrated that accurate profiles (error lower than 10% in the troposphere) can be obtained for H₂O. The temperature profiles are measured from the main vibrational bands of CO₂ and can be retrieved with accuracies of 1 K (RMS) at vertical resolution of 1 km, at least in the lower troposphere.

**Climate gases: CO₂ - Breakthrough requirement: 5%/50km/Total/6h**
No specific simulations were undertaken for CO₂ during the course of this study. Chedin et al.
have shown that annual and seasonal variations can be measured for CO\textsubscript{2} and N\textsubscript{2}O from the integrated radiance channels measured by HIRS on the meteorological TIROS-N satellites. Improved high resolution infrared sounders as AIRS and IASI should provide useful information for CO\textsubscript{2} if spatial and temporal averaging are performed in order to meet the required precision. For example, Chedin et al. [2002] showed that 1% uncertainty is achievable for 500kmx500km/15 days spatial and temporal averaging.

**CH4- Breakthrough requirement: 5%/50km/Total/6h**

Methane has one important absorption band ($\nu_3$, located around 1400 cm\textsuperscript{-1}) located in the Post-MSG spectral range. Although its atmospheric variability is weak due to its long lifetime, total column are retrievable with an accuracy of the order of the threshold. Even with improved noise and spectral resolution, 2 pieces of independant information are not possible. The averaging of pixels allowed to reach the 50 km horizontal requirement was not investigated here. The feasibility to retrieve methane global maps was demonstrated from the analysis of IMG data.

**CFC11,CFC12**

No specific simulations were undertaken for these species during the course of this study. Detection of CFC12 has been reported in the IMG spectra by two teams: Lubrano et al, [2002] and Coheur et al. [2003]. The latter also showed that large scale averaging allows the detection of CFC11 and HCFC22 for which atmospheric concentration has increased since the control of CFCs emissions. Spatial variabilities can not be seen from their absorption contributions, but need to be taken into account for a precise estimate of the surface temperature and for a correct representation of ozone which absorbs in the same spectral range.

### 5.2 Instrumental improvements

From the study it appears that the measurements of total columns of CO, O\textsubscript{3} and CH\textsubscript{4} will be possible for all the range of noise and spectral resolution investigated. Both accuracy and vertical information content increase for all three compounds if noise and/or spectral resolution is improved. Given the current specifications, noise improvement should be preferred than spectral resolution improvement. For low absorbing species, we do not recommand to extend the spectral range until the H\textsubscript{2}CO absorption range as detection limits are elevate and possibly will not be reached even for strong pollution events. The goal instrumental specification, combined with a well controlled state of the atmosphere in terms of T and H\textsubscript{2}O profiles, should allow to provide a limited information on NO\textsubscript{2}, SO\textsubscript{2} (volcanoeous) and possibly PAN total column contents.

Returning to specifications provided in Table 1, improved noise and/or spectral resolution mode should be recommended for IRS2 (low priority: atmospheric window + aerosols), IRS3 (high priority:ozone), IRS4 (low priority: SO\textsubscript{2}), IRS5 (low priority: CH\textsubscript{4}), IRS6 (low priority: NO\textsubscript{2}) and IRS7 (high priority:CO).
5.3 Limitations and further works

The following studies should allow to refine some of the results provided here:

1. Improve and optimize the micro-windows selection for the weakly absorbing gases
2. Study the impact of ancillary parameters (temperature profile, other gases contents) on the accuracy
3. Study the impact of geometry (off-nadir soundings), and spatial/temporal averaging
4. Train a neural network algorithm using the instrumental characteristics provided here
5. Perform simulations for cloudy conditions and high aerosol contents
6. Improve determination of detection limits for weak absorbing species using regional CTM models
7. Study the possibility to use HNO$_3$ as a proxy for NO$_3$ measurements

Acknowledgments The authors are grateful to the IMGDIS (Japan) for providing the Level 1 IMG data, to Tony Clough (AER, US) for LBLRTM and to Daniel Hurtmans (ULB, Belgium) for least square fitting retrieval code. The following people have provided useful inputs for this report: Michel Kruglanski and Martine De Mazière (BIRA/IASB, Belgium) for aerosols contribution, Grant Allen and John Remedios (Leicester University, UK) for PAN cross-section data, Didier Hauglustaine (LSCE, France) and Olivier Boucher (LOA, France) for trace gases and SO$_2$ distributions, Claude Derognat (SA, France) for trace gas lifetimes, and Larabee Strow (UMBC, US) for AIRS aerosol measurements.
References


Acronyms

ADEOS ADvanced Earth Observing Satellite
AIRS Atmospheric InfraRed Sounder
CNES Centre National d’Etudes Spatiales
ESA European Space Agency
EUMETSAT European Organization for the Exploitation of Meteorological Satellites
GOME Global Ozone Monitoring Experiment
IASI Infrared Atmospheric sounding Interferometer
IMG Interferometric Monitor for Greenhouse Gases
IMGDIS IMG Data and Information System
ISSWG IASI Sounding Science Working Group
LBLRTM Line-By-Line Radiative Transfer Model
METOP METeorology OPerations
MITI Ministry of International Trade and Industry
MOPITT Measurements Of Pollution In The Troposphere
NASA National Space Agency
NASDA National Space Development Agency of Japan
NDSC Network for Detection of Stratospheric Change
POLDER POLarization and Directionality of the Earth’s Reflectances
TES Tropospheric Emission Spectrometer
TOMS Total Ozone Mapping Spectrometer