XCO$_2$ retrieved from IASI using KLIMA algorithm

Samuele Del Bianco*, Bruno Carli, Marco Gai, Lucia Maria Laurenza and Ugo Cortesi

Istituto di Fisica Applicata 'Nello Carrara' IFAC-CNR, Firenze, Italy
s.delbianco@ifac.cnr.it

I. INTRODUCTION

Carbon dioxide (CO$_2$) is the main greenhouse gas released into the Earth’s atmosphere by human activities. The concentration of CO$_2$ in the atmosphere depends on the balance of natural sources and sinks, which are being perturbed by anthropogenic forcing due to fossil fuel burning, uncontrolled urban development, deforestation and other land use changes. An improvement in our understanding of processes responsible for absorption of CO$_2$ is urgently needed both for a reliable estimate of future CO$_2$ levels, and for the enforcement of effective international agreements for its containment.

A consolidated approach to retrieve quantitative estimates of CO$_2$ sources and sinks relies on inverse modelling of surface fluxes of carbon dioxide based on accurate measurements of its concentration as a function of space and time and simulations of atmospheric transport [Gurney et al., 2002, Roedenbeck et al., 2003]. In this regard, as originally highlighted by [Rayner and O’Brien, 2001], satellite-based observations can considerably improve the performances of the inverse modelling because of their global coverage.

Atmospheric CO$_2$ total columns can be measured from space with nadir looking sensors observing either spectra of the reflected sunlight in the Near InfraRed (NIR) spectral region (CO$_2$ bands at 1.6 $\mu$m and 2.0 $\mu$m) or observing the radiance emitted by CO$_2$ bands in the Thermal InfraRed (TIR) spectral region (CO$_2$ bands at 4 $\mu$m and 15 $\mu$m). NIR observations have the advantage of being sensitive down to the lowermost layers of the atmosphere, which contain information about surface fluxes useful for identification of CO$_2$ sources and sinks; in addition they are less affected by temperature and less contaminated by water vapour lines interferences. In this spectral region, column averaged mixing ratio of CO$_2$ in dry air, denoted as XCO$_2$, has been measured by SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) [Buchwitz et al., 2005] and TANSO-FTS (Thermal And Near-infrared Sensor for carbon Observation- Fourier Transform Spectrometer) [Kuze et al., 2009] and further space measurements will soon become available with the OCO-2 satellite, planned to be launched in July 2014 [Miller et al., 2007, Chevallier et al., 2007].

On the other hand, TIR has the advantage of providing a better coverage with observations over water that do not depend on the sun glint and with continuous measurements during both day and night. Furthermore, a few high quality instruments are already operating in this spectral region. The first retrievals of XCO$_2$ column concentration made with data from TOVS (TIROS Operational Vertical Sounder) [Chédin et al., 2003(a), Chédin et al., 2003(b)] demonstrated limited capability to constrain CO$_2$ sources and sinks [Chevallier et al., 2005], but the new high resolution infrared sounders, such as the AIRS (Atmospheric InfraRed

*Corresponding author: Samuele Del Bianco, Istituto di Fisica Applicata 'Nello Carrara' IFAC-CNR, Firenze, Italy, email: s.delbianco@ifac.cnr.it
Sounder) [Engelen et al., 2009] and IASI (Infrared Atmospheric Sounder Interferometer) [Crevoisier et al., 2009] were shown to attain a substantial gain of information mostly due to their broader spectral coverage and finer resolution.

In particular, IASI provides twice a day global Earth coverage with a very good horizontal resolution (pixel diameter of 12 km for nadir observations) and very broad spectral coverage at high spectral resolution, even if only a small fraction of its spectral channels is currently used because of the complexity of broad band retrievals and of the interference of other atmospheric parameters (in particular water vapour and temperature) in most spectral channels. However, computing resources are now less of a constraint and the problem of interference can be overcome with the simultaneous retrieval of both target and interfering parameters [Carli et al., 2006]. The simultaneous retrieval makes it possible to exploit all the spectral channels that have some information on the target parameter and provides a retrieval error that includes, in a rigorous way, the errors of the interfering parameters. Therefore, with the enhancement of the information and with a better characterisation of the errors the XCO$_2$ retrieved from IASI measurements could be significantly better than what is currently available (2 ppm for a 5° x 5° spatial resolution on a monthly time scale [Crevoisier et al., 2009]).

In this paper, we present the results of a study whose primary objective was to investigate the ultimate performances of the retrieval of XCO$_2$ from IASI observations in the thermal infrared, in order to add insight and additional evidence to the ongoing discussion on the measurement of CO$_2$ from space. As part of this effort, we aimed to achieve the requirements of TANSO-FTS operational products (0.3%, corresponding to 1 ppm out of 370 ppm, on monthly averages over 1000 x 1000 km$^2$ areas) and to perform a cross-validation between the IASI and the TANSO-FTS XCO$_2$ values.

For the retrievals we used the KLIMA (Kyoto protocol Informed Management of the Adaptation) code, obtained by upgrading the algorithm employed for the analysis of REFIR-PAD (Radiation Explorer in the Far InfraRed-Prototype for Applications and Development [Bianchini et al., 2008]) measurements, adapted in turn from the MARC (Millimetre-Wave Atmospheric Retrieval Code) code developed for the MARSHALS (Millimetre-wave Airborne Receiver for Spectroscopic CHaracterization of Atmospheric Limb-Sounding) study [Carli et al., 2007]. Because of the application of this code to the IASI measurements we shall refer to the study as the KLIMA/IASI study.

For more details about the retrieval code, the Forward Model (FM) contained in the retrieval and the analysis set-up (retrieved parameters, a priori information, retrieval grid and so on) see [Laurenza et al., 2014] and the references contained therewith.

In section II to IV of this paper we provide a synthetic overview of the issues covered by our activities for development and optimization of the KLIMA retrieval code, along with a description of the results. In section V the subsequent phase of comparison between KLIMA/IASI and TANSO-FTS products is outlined and the main outcomes are shown. In the conclusions, we provide an assessment of the current status of the work, and highlight the consolidated results of the study along with the open issues that still remain to be investigated.

II. PREPARATORY ACTIVITIES

The preparatory activities included:

- Validation of the FM and of the retrieval code operating wide-band.
- Reduction of the computing time with approximations that induce a systematic error congruent with the required accuracy.

The forward model calculations of the KLIMA code were validated [Laurenza et al., 2014] by comparing its synthetic spectra with those of the LBLRTM (Line-By-Line Radiative Transfer Model) code [Clough et al., 2005]. The atmospheric scenario for the simulations was set according to the atmospheric state retrieved by
LBLRTM from the night-time measurement acquired by IASI over the Southern Great Plains in Oklahoma, USA, on 19th April 2007, during the Joint Airborne IASI Validation Experiment (JAIVEx) [Shephard et al., 2009]. The calculations were made over the full spectral range of IASI from 645 to 2760 cm\(^{-1}\). The difference between the two simulations was, with very few exceptions, significantly smaller than IASI radiometric noise. Subsequently, the quality of the KLIMA FM was also confirmed by the small amplitude of residuals averaged over a large number of retrievals performed on real measurements.

A large number of simulated retrievals with different spectral noise were performed in order to assess the bias caused by the approximations and the accuracy in the averaged retrieved value of XCO\(_2\). The requirement of 0.3% precision and accuracy can be reached by averaging 25 IASI observations, while averaging 100 observations the approximations cause a retrieval bias smaller than 0.04%.

The retrieval code was integrated on the ESA G-POD (European Space Agency Grid Processing On-Demand) system computing resources [Farres et al., 2010] for the bulk processing of IASI measurements. The G-POD computing requirements (program size not exceeding 1 GB and processing of one orbit in less than one day on a single processor) imposed a reduction of the spectral range and of the number of spectra. The analysis was limited to the spectral range 645 - 800 cm\(^{-1}\), which contains the \(v_2\) band of CO\(_2\) and provides the best compromise between high sensitivity to carbon dioxide and low sensitivity to other atmospheric variables. As a consequence, the retrieval error of the performed analysis is about twice the retrieval error that is theoretically possible with the analysis of the full spectral range. Furthermore, only the spectra of the IASI central pixels with respect to nadir have been analysed.

III. KLIMA/IASI ANALYSIS ON G-POD

A total of 240000 IASI spectra, covering the period from March 2010 to February 2011, has been analysed on G-POD. Selection criteria were adopted that limited the analysis to the best measurements (clear sky, small slant angles, flat topography) and to a maximum of 20000 spectra each month. To this purpose, only one week was considered each month and a reduced number of measurements were selected over the oceans. The number of analysed spectra is about 1% of the total number of useful spectra. Nevertheless, the collected information is extensive: indeed the G-POD analysis made possible the acquisition of a large amount of data with good geographical and time distributions, which provides a very important starting point for the assessment of problems and capabilities.

IV. RESULTS OF THE RETRIEVAL

An example of the attained coverage and quality is given in Figure 1 where the global map of XCO\(_2\) averaged over the full period year on a grid of 2\(^\circ\) x 2\(^\circ\) pixels is shown. Data (even if not yet validated) show a geographical variability that can be reasonably ascribed to the dynamic of the carbon cycle together with some scattered points that seem to be outliers. Over land, the seasonal variation (shown in [Cortesi et al., 2014]) of XCO\(_2\) is controlled mainly by photosynthesis in the terrestrial ecosystem. The XCO\(_2\) concentrations are generally higher in the Northern Hemisphere during Spring months than in the Southern Hemisphere, while these higher values decrease during Summer months. This behaviour could be explained in terms of plant photosynthesis that, for the Northern Hemisphere, is not yet competitive with respiration (primarily due to uptake and release of CO\(_2\)) [Morino et al., 2011]. In the monthly averages, plotted over the same grid (shown in [Cortesi et al., 2014]), we observe rather uniform fields over the oceans and a large variability with time and location over
land, even if the monthly variable sampling makes it difficult to identify any meaningful time variation and the reduced statistics highlights the presence of relatively large errors at high latitudes.

In the retrievals, the FM calculations fit well the spectra observed by IASI with residual differences comparable with the spectral noise (the standard deviation of the residuals evaluated on 500 observations reconstructs exactly IASI spectral noise). These residuals, when averaged over a large ensemble of measurements, are usually much smaller than the spectral noise. Only a few isolated atmospheric features show average residuals that are larger than the spectral noise. The same features have been found by [Alvarado et al., 2013]. Further residuals larger than IASI spectral noise (about twice) are close to 649 cm$^{-1}$ and 771 cm$^{-1}$. The first one is due to the approximations introduced in the FM [Laurenza et al., 2014], while the latter has been already observed in the FM validation phase and it requires further investigation.

The $\chi^2$-test [Rodgers, 2000] shows a correlation of the $\chi^2$ values with the temperature of Earth’s surface with $\chi^2$ values close to unit at low temperatures (230 K) and values that monotonically increase up to 3 at the highest temperatures (330 K). This indicates that increasing residuals are observed for larger surface temperatures. Values greater than 3 occasionally occur, but being always located in desert areas are most probably due to sand storms. The KLIMA FM does not include the scattering contribution of aerosol that could affect, in particular, desert measurement scenarios, and it might be a significant source of systematic errors. For this reason, events with $\chi^2$ larger than 3 are filtered out in our analysis.
Figure 2: On the left: Total error (obtained multiplying the retrieval error by the SQRT of the $\chi^2$) of XCO$_2$ retrieved from IASI spectra with the KLIMA code plotted as a function of the surface temperature. On the right: Averaged XCO$_2$ as a function of the surface temperature binned over 1 K.

The retrieval error of XCO$_2$ also varies as a function of the surface temperature with values of 2 ppm at high temperature (330 K) and up to 20 ppm at low temperature (230 K). Indeed high temperatures correspond to large amplitude of the observed spectral radiance and greater signal to noise ratio resulting in a more precise retrieval. This implies that the best precision of IASI observations is obtained in warm seasons and at low latitudes and the worst precision is obtained in the cold seasons and at high latitudes. In comparison with other retrieval exercises it is important to recall that these retrieval errors are obtained from a single spectrum and also account for the errors caused by the uncertainties present in the interfering parameters that are simultaneously retrieved (water vapour profile, temperature profile and surface albedo).

In the monthly averages, where negligible time variability is present, the Statistical Standard Deviation (SSD) of the values retrieved in each pixel was compared with the retrieval errors, determined with error propagation calculations. In most pixels, the SSD is about equal to the retrieval error and often significantly smaller than it. Values larger than the retrieval error are rather uncommon and mainly occur in some locations over land, in particular over the Pacific Coast of South America, over central Africa and over continental Asia, and are probably due to either unaccounted errors or geographical and time variability.

The small values of SSD observed in most of the pixels of the monthly averages suggest that the unaccounted errors are small, but the spectrally localized residuals greater than the measured error and the small deviation from unit of the calculated $\chi^2$ indicate that some unaccounted errors are present. An estimate of the total retrieval error, which takes into account the possible contribution of unaccounted errors, is obtained by multiplying the retrieval error by the SQuare RooT (SQRT) of the $\chi^2$ [Barlow, 1993]. For the analysed data set the total error of retrievals from a single spectrum is shown in Figure 2 to vary from 3 ppm to about 20 ppm as a function of Earth’s surface temperature. In conclusion, the diagnostics of the KLIMA/IASI measurement indicate that the measurements have good coverage. Moreover, the statistical analysis of the residuals and the comparison between the total retrieval error and the SSD of the pixels do not highlight any anomalous behaviour.

The right panel of Figure 2 shows the average of all retrieved XCO$_2$ values as a function of surface temperature in the case of measurements over land (red points) and over water (blue points). We observe a decrease of concentration with increasing temperature over land and with decreasing temperature over wa-
ter which respectively correlates with the CO$_2$ sinks of vegetation, related to plant growth, and of ocean, because a lower water temperature increases the solubility of CO$_2$ into the sea.

V. Comparisons

It is difficult to verify whether the XCO$_2$ values retrieved in the framework of the KLIMA/IASI study capture the variability observed by the other existing measurements, because this variability is often comparable with the measurement error of single observations. On the other hand, averages can be affected by time and space variability that may be present in some locations (especially over land).

Over water, where less variability is expected, the comparison with good quality measurements is only provided by the ground-based in situ sampling made at Mauna Kea. This comparison shows a negative bias of about 12 ppm for KLIMA/IASI, but a rather good agreement in the seasonal variation. However, this is only a qualitative result because of the uncertainties present in a comparison between in situ and column measurements.

The comparison of KLIMA/IASI XCO$_2$ values with the ground-based occultation measurements of the TCCON (Total Carbon Column Observing Network) stations [Wunch et al., 2011] also shows the negative bias of our results. If the bias is corrected, the differences between the two series are consistent with the estimated errors. However, the requirement of 200 km maximum distance, presently adopted in this comparison, provides a data set that is statistically insufficient for a stringent validation. Probably the comparison should be repeated adopting a maximum distance of 500 km, as recently done in the EMMA validation study [Reuter et al., 2013].

The XCO$_2$ retrieved from TANSO-FTS measurement provides a comparable coverage even if, because of the different spectral region of observation, the two instruments have quite different averaging kernels. In this case, a coincidence criterion of one day and 100 km provides a large collection of comparable XCO$_2$ measurements. Moreover, IASI spectral channels are in the TIR region and measurements are available both during day and night; because no significant diurnal variability has been observed in the limited dataset also IASI nighttime measurements have been taken into account in the comparison. The distribution of the differences between the two datasets shows a negative 8 ppm bias of KLIMA/IASI and a standard deviation of 7 ppm. The bias is not a surprise considering the systematic errors that can be caused by the existing uncertainties in the spectroscopic parameters of the HITRAN database used in the FM [Rothman et al., 2009]. The spread is instead much larger than the combined retrieval errors of the two experiments. Taking into account the different averaging kernels the spread does not change significantly, while the negative bias is reduced to 1.2 ppm. Assuming that the two experiments only have positive unaccounted errors (no constraint), this result can only be explained by an unaccounted error of about 6 ppm in KLIMA/IASI. Another comparison between KLIMA/IASI and TANSO-FTS is done by calculating yearly averages in some rather homogeneous geographical areas of $10^\circ \times 10^\circ$, over land and over ocean. The results are summarized in Table 1. The KLIMA/IASI and TANSO-FTS averaged values are listed in the second and third column respectively. The differences between KLIMA/IASI and TANSO-FTS, corrected for the average bias of 8 ppm, are reported in the 4th column. Over ocean, the observations are statistically poorer, especially in the case of TANSO-FTS (in some cases the average values are obtained from a few samples and no TANSO-FTS observations are available in two areas). Over land, the larger differences are located where a stronger geographical variability of CO$_2$ sources is expected (i.e. North America and China). This result could suggest, in addition to a geographical dependent bias, the presence of a constraint in TANSO-FTS measurements. KLIMA/IASI results are quite promising, but further comparisons are needed for a stringent validation.
Table 1: Comparison of the retrieved annual average of XCO$_2$ [ppm] in selected homogeneous macro areas. In red are reported macro areas over land, in blue are reported macro areas over ocean.

<table>
<thead>
<tr>
<th>MACRO AREA</th>
<th>KLIMA/IASI [ppm]</th>
<th>TANSO-FTS [ppm]</th>
<th>(KLIMA/IASI) + bias - (TANSO-FTS) [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>N. America</td>
<td>383.0</td>
<td>386.8</td>
<td>+ 4.2</td>
</tr>
<tr>
<td>S. America</td>
<td>378.9</td>
<td>389.0</td>
<td>+ 2.1</td>
</tr>
<tr>
<td>Europe</td>
<td>383.0</td>
<td>388.6</td>
<td>+ 2.4</td>
</tr>
<tr>
<td>N. Africa</td>
<td>380.0</td>
<td>386.4</td>
<td>+ 1.6</td>
</tr>
<tr>
<td>S. Africa</td>
<td>375.3</td>
<td>386.0</td>
<td>- 2.7</td>
</tr>
<tr>
<td>Arabia</td>
<td>377.1</td>
<td>387.5</td>
<td>- 2.4</td>
</tr>
<tr>
<td>India</td>
<td>381.1</td>
<td>388.2</td>
<td>+ 0.9</td>
</tr>
<tr>
<td>Australia</td>
<td>376.1</td>
<td>386.2</td>
<td>- 2.1</td>
</tr>
<tr>
<td>China</td>
<td>391.2</td>
<td>390.2</td>
<td>+ 9.0</td>
</tr>
<tr>
<td>Asia</td>
<td>379.8</td>
<td>383.9</td>
<td>+ 3.3</td>
</tr>
<tr>
<td>N. Pacific Ocean</td>
<td>377.1</td>
<td>No Data</td>
<td>-</td>
</tr>
<tr>
<td>S. Pacific Ocean</td>
<td>372.2</td>
<td>384.1</td>
<td>- 3.9</td>
</tr>
<tr>
<td>N. Atlantic Ocean</td>
<td>378.1</td>
<td>384.5</td>
<td>+ 1.6</td>
</tr>
<tr>
<td>S. Atlantic Ocean</td>
<td>376.6</td>
<td>No Data</td>
<td>-</td>
</tr>
<tr>
<td>Indian Ocean</td>
<td>377.7</td>
<td>385.6</td>
<td>+ 0.1</td>
</tr>
</tbody>
</table>

VI. Conclusions

The KLIMA code, which performs wide-band and multi-target retrieval, was applied to IASI measurements for the retrieval of XCO$_2$. A total of 240000 IASI spectra, covering the period from March 2010 to February 2011, have been analysed on the G-POD computing resources. This is a preliminary analysis of a small fraction of the information provided by IASI observations. Fits of good quality are obtained and single spectra provide measurements with a precision, which varies from 3 ppm to about 20 ppm as a function of the surface temperature, so that the requirement of 1 ppm on monthly averages over 1000 x 1000 km$^2$ areas can be easily met.

As far as the aim is concerned of developing and testing an inverse model matching the required capability for XCO$_2$ retrieval from IASI spectral radiances, the results made available up to now by the KLIMA/IASI data analysis represent an encouraging step forward. This is also confirmed by the observed geographical and seasonal variability that are often in good agreement with our expectations (see [Cortesi et al., 2014] for a detailed presentation and discussion of these patterns).

Current outcomes of the reported activities - both of KLIMA application to IASI measurements and comparison with TANSO-FTS products - highlight, on the other hand, a series of open issues that recommends for further investigation. Among the most important of these issues is the evidence of a large offset between column-averaged dry air mole fractions of carbon dioxide from KLIMA/IASI and TANSO-FTS retrieval processing, for which we can give a quantitative estimate in terms of a negative bias of about 8 ppm on the IASI XCO$_2$ products. Potential sources of systematic error, to identify as the dominant or combined causes of the observed bias, include the effects of spectroscopic uncertainties and of FM errors, as well as the different vertical sensitivity of the two measurements (the bias is substantially reduced to 1.2 ppm, when we take into account the different averaging kernels). A dedicated effort shall be made to test the different hypothesis and check their impact on the compensation of the bias.

Another topic that emerged from the results of
the comparison is associated to the high value (7 ppm) of the standard deviation of the differences between KLIMA/IASI and TANSO-FTS data corresponding to a spread much larger than the combined retrieval errors of the two experiments. The convolution with the averaging kernels, while reducing the bias value, has negligible impact on the spread. In case we make the reasonable assumption that no strong external constraint is applied in TANSO-FTS retrievals, an unaccounted error of about 6 ppm can be present in KLIMA/IASI XCO$_2$ retrieved values. Tests, which exploit the large available statistics for the assessment of the possible causes of unaccounted error, are in progress.

In summary, we achieved the objective of demonstrating the capability of the KLIMA algorithm to retrieve information on total column of carbon dioxide from IASI measurements with an accuracy comparable to the requirement of the TANSO-FTS instrument. In parallel, our conclusion at the end of the KLIMA/IASI project is that important questions remain to be addressed about the performance of the KLIMA retrieval model. At the present stage, these open questions prevent us from using the results of the comparison for a stringent validation of TANSO-FTS operational products. However, they set a basis for subsequent testing and improvement of the KLIMA code according to a preliminary plan of activities, which is described in detail by [Cortesi et al., 2014].

VII. Acknowledgements

This work was performed under the project ‘Sensitivity Analysis and Application of KLIMA algorithms to GOSAT and OCO validation’ (ESA-ESRIN contract N. 21612/08/I-OL). IASI data are received through the EUMET-SAT Unified Meteorological Archive and Retrieval Facility (UMARF). The authors wish to thank G-POD Team at ESA/ESRIN for their valuable effort and availability for the bulk processing of IASI data. TANSO-FTS L2 data were obtained from the GOSAT User Interface Gateway GUIG, operated by the National Institute for Environmental Studies (NIES). TC-CON data were obtained from the TCCON Data Archive, operated by the California Institute of Technology, from the website at http://tccon.ipac.caltech.edu/

References


