Università degli Studi di Roma Tor Vergata

Dipartimento di Ingegneria Civile e Ingegneria Informatica
GeoInformation PhD Program

Application of KLIMA algorithm to Carbon Dioxide retrieval from IASI/METOP-A observations and comparison with TANSO-FTS/GOSAT products

Lucia Maria Laurenza

A.A. 2013-2014

Phd Mentor: Prof. Giovanni Schiavon
IFAC-CNR supervisor: Dr. Ugo Cortesi

Table of Contents

Chapter 1

Introduction

1.1 Atmospheric Carbon Dioxide and Global Warming

Carbon Dioxide (CO_2) is the primary Greenhouse gas (GHG) released into the Earth's atmosphere by human activities. It is a well-mixed gas that does not condense or precipitate from the atmosphere, with an atmospheric residence time exceedingly long, being measured in thousands of years. Carbon Dioxide is the key atmospheric gas that exerts principal control over the strength of the so called *Greenhouse effect*, i.e. the trapping of the Sun's warmth in the lower atmosphere, caused by the presence of CO_2 and the other GHGs which absorb the InfraRed (IR) radiation emitted from the Earth's surface. The balance of natural *sources* (any process by which Carbon Dioxide is added to the atmosphere) and *sinks* (any process by which Carbon Dioxide is removed from the atmosphere) of Carbon Dioxide, leading to stable values of about 277 ppm for atmospheric CO_2 concentration in pre-industrial times, has been significantly perturbed by anthropogenic forcing, mostly from fossil fuel burning for energy production, uncontrolled urban development, deforestation and other land use change, since the beginning of the industrial era, in the late Eighteenth Century ([54]).

There are clear hints that Earth's climate is changing. Temperatures are rising, the air temperature near the surface raised with warmer periods and night-time temperatures have risen more than daytime. Regional changes are also evident: the heating is most intense at mid-latitudes and on the continents in winter and spring. Precipitations have increased, mainly over land in high latitude of the Northern Hemisphere, as more extreme climate events, like heavy rainstorms and record high temperatures. Earth's climate is changing in ways that affect weather, oceans, snow, ice, ecosystems and society. Future changes are expected to include a warmer atmosphere, a warmer and more acidic ocean, higher sea levels and larger changes in precipitation patterns. Natural causes alone cannot explain recent changes, and there are evidences that many of the observed changes in the climate can be linked to human activities, principally the release of large amounts of GHGs into the atmosphere, caused largely by people burning fossil fuels to generate energy and power vehicles, that are modifying concentrations and distributions of CO_2 and other heat-trapping gases ([27]).

Current levels of CO_2 in the atmosphere exceed the value of 390 ppm ([33]), with a mean annual global growth rate by decade varying from 0.8 ppm yr^{-1} in 1960-1969, to 1.3 ppm yr^{-1} in 1970-1979, to 1.6 and 1.5 ppm yr^{-1} in 1980-1989 and in 1990-1999 respectively and raising to 1.9 ppm yr^{-1} in 2000-2009, as shown in Figure 1.1 and in Figure 1.2, taken from the studies and model projections of [32], [7] and [65]. Figure 1.3 shows projected GHGs gas concentrations for four different emissions scenarios along with observed and projected changes in global average temperature under three no-policy emissions scenarios. The top three scenarios assume no explicit climate policies. In the emission scenarios and model projections there is a clear raise of CO_2 emissions (left panel of Figure 1.3), especially from fossil fuel and cement, that is been estimated in 2013 as 9.9 ± 0.5 Pg C, an increase of 61% with respect to the last two decades ([33]). The increase in the concentration of GHGs also reduces the efficiency with which the Earth cools to space, that leads an increase of Earth's average temperature, the so-called *Global Warming* effect (right panel of Figure 1.3).

In the period from 1959 to 2008, the fraction of the total anthropogenic emissions of CO_2 that remained in the atmosphere each year, the so called *Airborne Fraction* (AF), has increased from about 40% to 45%. On average, 43% of the total CO_2 emissions each year remained in the atmosphere, even if there is a very large temporal variability (Figure 1.1). Moreover, models suggest that this trend has been caused by a decrease in the capacity of sinks to remove CO_2 from atmosphere, due to the modification of the Carbon sinks in response of the changing in Earth's climate. The rest of the emissions, the *Sink Fraction*, was absorbed partly by oceans and partly by the land biosphere ([**?**]). According to the estimates given by the *Global Carbon Project* for the fate of CO_2 anthropogenic emissions (see, for instance, [41], updated from [57] and [7]), the natural sinks of the ocean and of the biosphere retained respectively 27% and 26% of CO_2 anthropogenic emissions, but their capability for sequestering Carbon Dioxide is strongly variable in space and in time. Over the last 200 years, the oceanic uptake constituted the only true net sink mechanism for anthropogenic CO_2 with the absorption of 118 ± 19 Pg C, whilst the terrestrial biosphere was a net source of CO_2 to the atmosphere in the same period with the release of 39 $\pm$ 28 Pg C [77]. On the other hand, indications exist that the uptake of CO_2 by the ocean has diminished from 34% to about 26%, with the sink strength of the biosphere remaining constant at the level of about 30%.

The knowledge of current state of CO_2 concentration is an essential piece of information to constrain the inverse modelling of surface Carbon fluxes and to identify natural and anthropogenic sources and sinks of CO_2 towards the goal of discriminating human-induced perturbations from background variability.

The first and crucial step to achieve reliable and accurate predictions of future levels of atmospheric CO_2 and to reduce the uncertainties in the projection models is the monitoring of atmospheric emissions, with a measure as precise as possible, provided with a good geographical coverage, of the CO_2 concentration. This can be attained by the employment of satellite missions, which can provide global coverage with short revisiting time and the possibility to perform measurements in remote areas, where it is not possible to use ground stations.

To obtain information about the CO_2 distribution from remote sensing measurements, it is necessary to develop inverse models, which perform retrieval of CO_2 total amount, with the highest possible accuracy. Combining the use of retrieval models with the knowledge of transport mechanisms in the atmosphere, it is also possible to derive another important parameter: the CO_2 fluxes. It is important to obtain an accurate determination of the geographic distribution of surface Carbon fluxes on a variety of spatial scales, from local and regional to continental and global, because they are an essential factor to determine the sources and sinks of natural and anthropogenic Carbon Dioxide ([53]). More accurate information on the distribution and magnitude of sinks and sources, on the partitioning of the total uptake between the ocean and land sinks, on the capacity of sinks to remove CO_2 from atmosphere and on the distinction between natural and anthropogenic sources are relevant not only from a purely scientific point of view but also from a political and economic point of view. An improvement in the understanding of the specific processes responsible for absorption of the CO_2 amounts released into the atmosphere from year to year by natural and anthropogenic sources is a key issue, which dealing with periodical assessments of the efficiency of the adopted Kyoto-like Protocols to reduce GHGs atmospheric emissions and of the effects of measures applied to mitigate the response of Earth's climate to human perturbation of atmospheric Carbon Dioxide concentration.

1.2 The Carbon cycle

Carbon is the fundamental building block of life and an important component of many chemical processes. It is present in the terrestrial atmosphere primarily as Carbon Dioxide but also as other less abundant but climatically significant gases, such as Methane. Carbon is exchanged among Earth's oceans, atmosphere, surface, ecosystem, and exosphere. CO_2 is removed from atmosphere by a number of processes that operate on different time-scales and is subsequently transferred to various reservoirs. Projecting climate into the future and forecasting regional impacts depends on the understanding of this

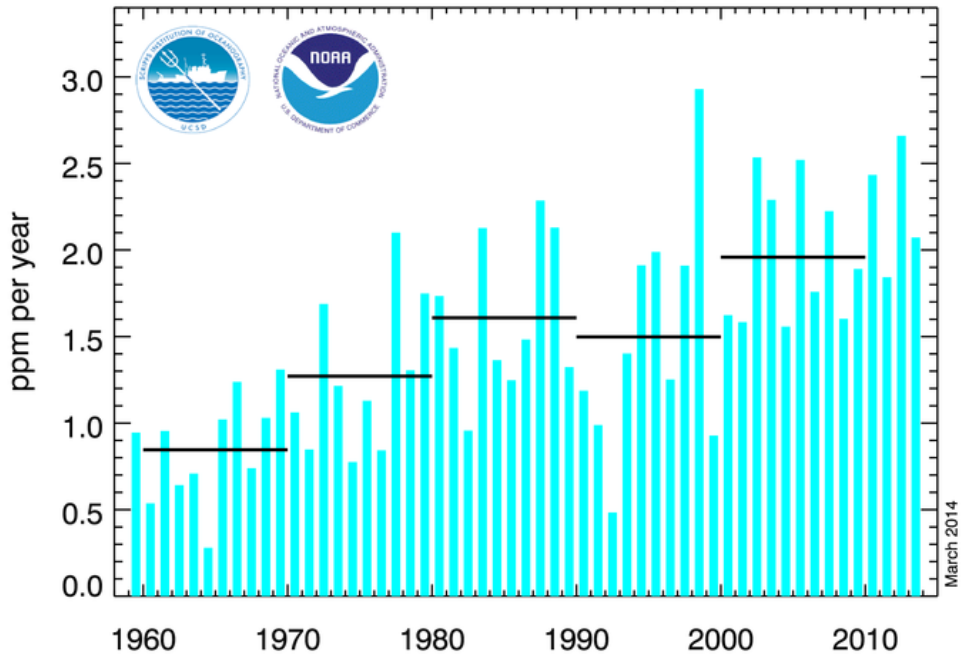

Fig. 1.1: *Annual mean Carbon Dioxide growth rates for Mauna Loa, Hawaii, along with decades averages of the growth rate, plotted as horizontal lines. Based on globally averaged marine surface data, as from values reported Dr. Tans, NOAA/ESRL ([65]) and Dr. Keeling, Scripps Institution of Oceanography ([26])*

cycle.

The Carbon cycle has a slow and a fast component. Through a series of chemical reactions and tectonic activity, Carbon takes between 100-200 million years to move between rocks, soil, ocean, and atmosphere in the slow Carbon cycle. The slow cycle returns Carbon to the atmosphere through volcanoes.

The fast Carbon cycle is largely the movement of Carbon through the biosphere. It is directly connected to plant life and the growing season. In fact, in the Northern Hemisphere winter, when few land plants are growing and many are decaying, atmospheric Carbon Dioxide concentrations climb. During the spring, when plants begin growing again, concentrations drop. This cycle peaks in August. In the fall and winter, as vegetation dies back in the northern hemisphere, decomposition and respiration returns Carbon Dioxide to the atmosphere ([8]).

Anthropogenic emissions of Carbon Dioxide (primarily from the burning of fossil fuels and cement production) have been growing steadily since the onset of the industrial revolution. About half of these emissions are removed by the fast Carbon cycle each year, the rest remains in the atmosphere. Continued emission of Carbon Dioxide to the atmosphere will affect climate and ocean chemistry, subsequently influencing both marine and terrestrial ecosystems. The warming effects of increasing CO_2 and other greenhouse gases have influence on agriculture, natural systems, and a host of environmental variables. Increasing CO_2 in the atmosphere also directly translates to increasing acidity of the oceans.

Atmospheric measurements of the spatial and temporal trends of CO_2 in the atmosphere are essential for reaching quantitative understanding of sources and sinks, and to understanding the potential evolution of global climate. There was unanimous agreement among the models that future climate change will reduce the efficiency of the Earth system to absorb the anthropogenic Carbon perturbation; all models simulated a negative sensitivity for both the land and the ocean Carbon cycle to future climate ([31]). However, there was still a large uncertainty on the magnitude of these sensitivities. To reduce the large uncertainties in climate−Carbon cycle projections, it is important that Carbon cycle models are more completely constrained by more accurate observational data.

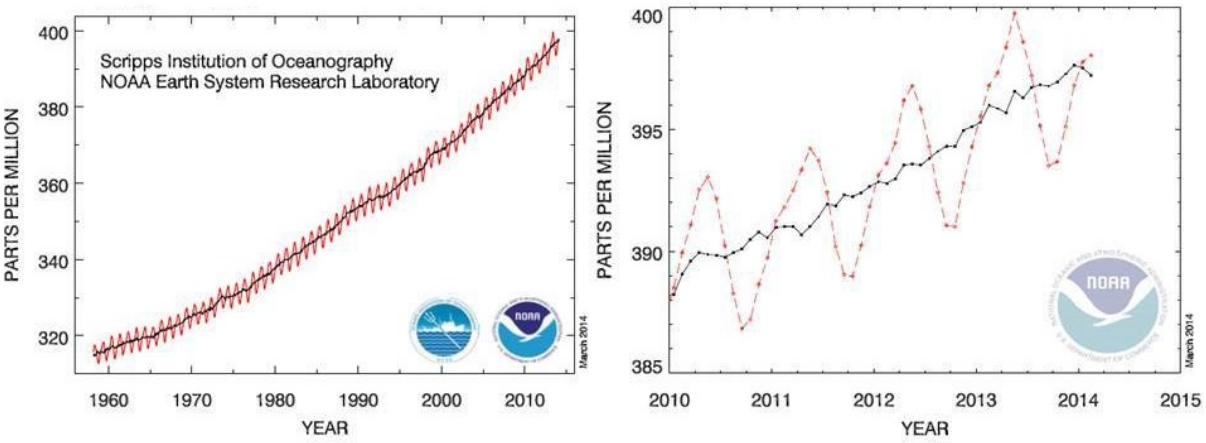

Fig. 1.2: *Left panel: Monthly mean atmospheric Carbon Dioxide at Mauna Loa Observatory, Hawaii. The Carbon Dioxide data (red curve), measured at Mauna Loa constitute the longest record of direct measurements of CO_2 in the atmosphere. The black curve represents the seasonally corrected data. Right panel: recent monthly mean (2010-2014) Carbon Dioxide measured at Mauna Loa Observatory, Hawaii. Red line represents the monthly mean values, the black line represents the seasonally corrected data (Dr. Tans, NOAA/ESRL ([65]) and Dr. Keeling, Scripps Institution of Oceanography ([26])*

1.3 Earth radiation budget and the Greenhouse effect

The Earth's temperature is determined by the balance between energy absorbed from the Sun and emitted into space. The presence or absence of the atmosphere and its thermal structure play a key role in energy balance. The emission of thermal radiation depends on the absorptive properties of the material and its absolute temperature. The rate of heating and cooling is generally not balanced in a given place and at any given time: somehow the temperature of the planet functions as a giant thermostat adjusting the rate of loss of energy to compensate for the excesses of heating and cooling.

The Earth rotation and inclination with respect to the plane of the orbit determines the diurnal and seasonal temperature variations. So, our planet is not at local radiative equilibrium (in which case the rate of heating caused by absorption of radiation at all wavelengths equals that of thermal emission), but is in a global radiative balance hence the energy balance is on long intervals of time and all over the planet. If this radiative balance fails climate changes occur.

If we force upon the radiative equilibrium, we found a surface temperature of 255 K, that is lower than the average temperature of Earth of about 290 K. This temperature represents the average temperature on the ground (which values around 15-18°C) but it is a sort of average temperature at the top of atmosphere (TOA). To obtain a good estimate of the temperature of the ground we need more detailed radiative balance. The soil temperature is greater than the calculated value at the TOA because the radiation emitted does not escape to the ground but it is reabsorbed by the planet and retained by the atmosphere itself.

The IR emission of the Earth is largely absorbed by molecules such as Water Vapour, Carbon Dioxide and Ozone. We expect that this underestimation of the soil temperature is more pronounced as optically thick is the atmosphere.

This difference can be explained by the effect introduced by the atmosphere: the *Greenhouse effect*, which was first discovered by Joseph Fourier in 1824, experimentally verified by John Tyndall in 1863 and quantified by Svante Arrhenius in 1896. Like in a greenhouse, the solar radiation passes through the atmosphere and reaches the surface because the atmosphere is almost transparent to UV and VIS radiation, while the IR radiation emitted from the surface is partly absorbed by the atmosphere and partly is re-emitted toward the surface, due to the presence of GHGs which are radiatively active in the IR. Since part of this re-radiation is back towards the surface and the lower atmosphere, it results in an elevation of the average surface temperature above what it would be in the absence of the gases. The

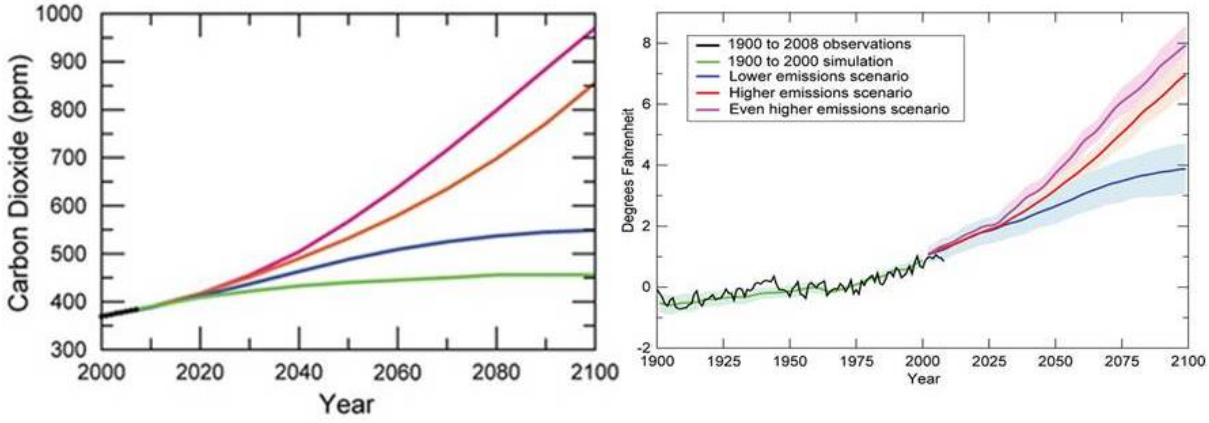

Fig. 1.3: *Left panel: projected greenhouse gas concentrations for four different emissions scenarios. The top three scenarios assume no explicit climate policies. The bottom green line is an illustrative 'stabilization scenario', designed to stabilize atmospheric CO_2 concentration at 450 ppm. Right panel: Observed and projected changes in global average temperature under three no-policy emissions scenarios. The shaded areas show the likely ranges while the lines show the central projections from a set of climate models. Source: USGCRP (2009)([88])*

most abundant GHGs in Earth's atmosphere are:

- **Water Vapour**: gaseous water represents a small but environmentally significant constituent of the atmosphere because it is a potent GHG. Approximately 99% of it is contained in the troposphere;
- **Carbon Dioxide**: CO_2 enters the atmosphere through burning fossil fuels, solid waste, trees and wood products and is removed from the atmosphere by plants absorption;
- **Methane**: CH_4 is emitted during the production and transport of coal, natural gas, oil, by agricultural practices and by the decay of organic waste;
- **Ozone**: O_3 is not emitted directly into the air by specific sources but it is the product of chemical reactions among Ozone precursors in the atmosphere, especially at ground-level. Tropospheric Ozone is a difficult GHG to keep track of, due to its short lifespan and the fact that its concentrations can vary hugely from place to place;
- **Nitrous oxide**: N_2O is emitted during agricultural and industrial activities, as well as during combustion of fossil fuels and solid waste;
- **ChlorofluoroCarbons (CFCs)**: synthetic, powerful GHGs that are emitted from a variety of industrial processes.

Each gas's effect depends on three main factors: how much of these gases are in the atmosphere, how long do they stay in the atmosphere and how strongly do they impact global temperatures.

Water Vapour is the principal contributor which accounts for about 50% of Earth's Greenhouse effect, the contribution of Carbon Dioxide is about 20% and the minor GHGs accounting for the remaining 5%. Because CO_2, Ozone, Methane and CFCs do not condense and precipitate, non-condensing GHGs constitute the key 25% of the radiative forcing, that supports and sustains the entire terrestrial Greenhouse effect. The remaining 75% coming as fast feedback contributions from Water Vapour and clouds ([56]). All of these increases in non-condensing GHG forcer are attributable to human activity. Clouds are the major non-gas contributor to the Earth's Greenhouse effect and also absorb and emit InfraRed radiation, thus having an effect on radiative properties of the atmosphere.

Water Vapour concentrations in the air are controlled by Earth's temperature. Warmer temperatures evaporate more water from the oceans, expand air masses, and lead to higher humidity. Cooling causes Water Vapour to condense and fall out. Carbon Dioxide, on the other hand, remains a gas at a wider

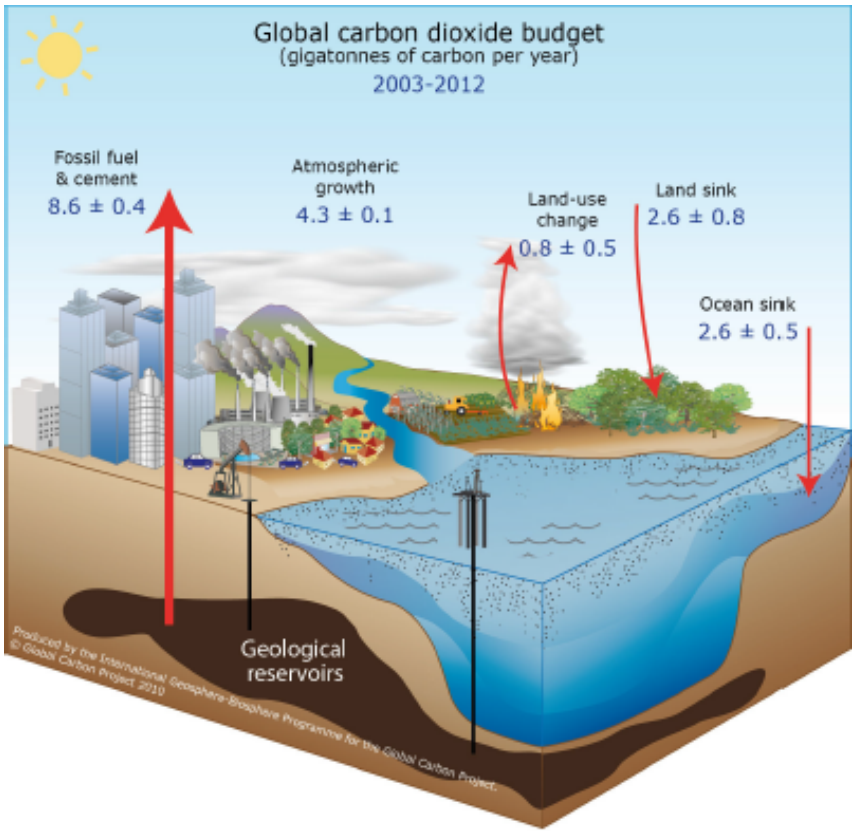

Fig. 1.4: *Schematic representation of the overall perturbation of the global Carbon cycle caused by anthropogenic activities, averaged globally for the decade 2003-2012 (Gtc/yr). The arrows represent emission from fossil fuel burning and cement production; emissions from deforestation and other land-use change; and the Carbon sinks from the atmosphere to the ocean and land reservoirs ([33])*

range of atmospheric temperatures than water. When Carbon Dioxide concentrations drop, Earth cools, some Water Vapour falls out of the atmosphere, and the Greenhouse warming caused by Water Vapour drops. Likewise, when Carbon Dioxide concentrations rise, air temperatures go up, and more Water Vapour evaporates into the atmosphere, which then amplifies greenhouse heating. So, while Carbon Dioxide contributes less to the overall Greenhouse effect than Water Vapour, it is the gas that sets the temperature. Carbon Dioxide controls the amount of Water Vapour in the atmosphere and thus the size of the Greenhouse effect.

1.4 Carbon Dioxide observation from space

A consolidated approach to retrieve quantitative estimates of the strengths and spatial distribution of CO_2 sources and sinks relies on inverse modelling of surface Carbon fluxes. This is based on atmospheric transport models, that include advection, convection and eddy mixing processes and simulate the air-ocean and the air-land exchange, being constrained by observed spatial and time gradients of CO_2 concentration and by the associated uncertainties [39], [74].

First applications of inverse modelling to infer net surface fluxes of Carbon Dioxide were based on best matching in situ measurements from surface flask sampling networks, which despite the high precision of the individual products are too sparse to provide the geographical coverage optimally suited for these inversion studies [79]. The current ground based networks, such as the Total Carbon Column Observing Network (TCCON) described in Chapter 4, do not provide uniform geographical coverage; measurement sites are mainly concentrated on land and in Northern Hemisphere, and they do not provide extensive coverage over oceans and over Poles. Moreover, several stations are not able to work all year

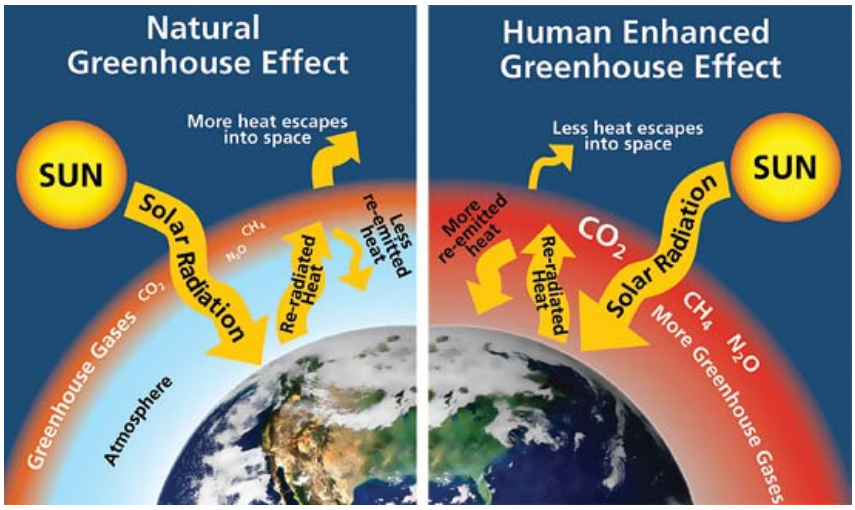

Fig. 1.5: *The Greenhouse effect that occurs in Earth atmosphere. Earth is heated by the absorption of solar radiation and cooled by InfraRed emission but the presence of the greenhouse gases like CO_2 acts like a "blanket" that traps the re-emitted InfraRed radiation in the lower part of the atmosphere, resulting in an elevation of the average surface temperature*

due to weather conditions.

An alternative source of observations, providing the adequate global coverage to characterize the spatial distribution of CO_2 sources and sinks, is now available with measurements of column averaged CO_2 from space-borne remote-sounders, as originally highlighted by [67]. The potential of satellite-based CO_2 observations to drastically improve the performances of the inverse models clearly derives, at first instance, from their better global coverage compared to surface networks, especially over land and from the large amount of data that can be acquired from space platforms. Satellite observations, of sufficient precision and accuracy, can help the classification of regional Carbon surface sources and sinks and help to reduce the flux uncertainties and locate unidentified Carbon sources and sinks, as they offer greater temporal and spatial coverage than the current observing network ([43]). A more accurate estimates of the Carbon fluxes is essential to understanding the response of the Carbon cycle to escalating atmospheric CO_2 levels and for predicting the behaviour of the Earth's future climate. So, a careful validation of satellite observations using ground based measurements and/or chemistry transport models is needed, along with the requirement of consistency between CO_2 retrievals from different space-borne instruments.

Passive remote-sounding of atmospheric CO_2 total columns from space can basically be performed by nadir viewing sensors measuring spectra of the reflected sunlight in the near InfraRed spectral region (CO_2 bands at 1.6 μm and 2.0 μm) using differential absorption optical spectroscopy (DOAS) or observing the radiance emitted by CO_2 bands in the thermal InfraRed (4 μm and 15 μm). Each of the two methods has relative advantages and drawbacks, but observations in the NIR certainly offer a better framework, as they are sensitive down to the lowermost layers of the atmosphere, thus being capable to retrieve the CO_2 content in the entire air column, so that are less affected by uncertainties in vertical transport when applied to inversion models of the Carbon cycle [5]. Moreover, the NIR is less sensitive to temperature and Water Vapour and less affected by possible interferences. Major strengths of CO_2 remote-sounding in the thermal InfraRed, on the other hand, are represented by the better global coverage: TIR measurements can be performed both at day and at night and over land and ocean, whilst NIR observations can only cover the sunlit part of the globe and shall be acquired in the sun-glint mode over the ocean to compensate for the low surface albedo, and by the much stronger absorption bands of Carbon Dioxide.

One of the first attempts to retrieve CO_2 information from space-borne measurements acquired by a nadir-viewing thermal InfraRed sounder and to use the retrieved concentration to infer CO_2 surface fluxes was made with data from the NOAA TIROS-N Operational Vertical Sounder (TOVS) [10], [11]

Mission	Instrument Type	Spectral coverage	Supporting Agency	Launch Status
GOSAT	FT spectrometer	SW-IR, TIR	JAXA, MOE, NIES	2009
GOSAT-2		SW-IR (and TIR)	JAXA, MOE, NIES	2014
OCO-2	Grating spectrometer	SW-IR	NASA	2014
CarbonSat			ESA	Phase-A study for EE-8
MicroCarb			CNES	2016
TanSat			MOST	2015
OCO-3			NASA	Proposed for deployment on ISS
A-SCOPE	Lidar		ESA	Proposed for EE-7, not selected
ASCENDS			NASA (USA)	2016
MERLIN			DLR, CNES	2016

Fig. 1.6: *Dedicated missions for remote-sensing measurements of atmospheric Carbon Dioxide from space*

and demonstrated limited capability to constrain atmospheric transport in the inversion of CO_2 sources and sinks [13].

Current generation of high resolution hyperspectral InfraRed sounders, such as the Atmospheric InfraRed Sounder (AIRS) and the InfraRed Atmospheric Sounder Interferometer (IASI) were shown to provide better results [13], [25], [23], despite their sensitivity to mid upper tropospheric Carbon Dioxide still representing a limiting factor for potential exploitation of InfraRed satellite data in Carbon Dioxide flux inversion. The performances of inverse modelling of the Carbon cycle based on near InfraRed measurements were evaluated by using data from SCIAMACHY [5] and from the Greenhouse gases Observing SATellite (GOSAT) [45]. GOSAT is the first mission specifically designed for monitoring atmospheric CO_2 and CH_4 distribution from space and, at the moment, is the only operative mission dedicated to greenhouse gases observation. Moreover, GOSAT has the unique feature of been equipped with an instrument which performs measurements in the SWIR and in the TIR region. The importance and the interest in monitoring the GHGs is highlighted by a large number of remote-sounders that is under study or already scheduled for launch, as displayed in Figure 1.6. Additional space measurements in the near InfraRed spectral region will become available soon with the replacement of the OCO mission (OCO-2), that is planned to be launched in July 2014 and will provide more precise data on CO_2 concentration, thus making possible a further reduction of the uncertainties associated to the inferred CO_2 surface flux [60], [14].

A comprehensive study, providing a clear overview of current and emerging capabilities for the global monitoring of CO_2 surface fluxes, was conducted by Hungershoefer et al. [42]. They evaluated the potential information content of nine different observing systems, that included the existing network of surface stations for in situ observation of CO_2 concentration, two extended versions of the current network, the nadir sounding in the thermal InfraRed of AIRS, the near-InfraRed measurements of SCIAMACHY, GOSAT and OCO and the active measurements which might be performed by the space-based lidar mission A-SCOPE (Advanced Space Carbon and Climate Observation of Planet Earth) of the European Space Agency [51]. The results of the study confirmed the better performance of space remote-sounding in the NIR with respect to the TIR, and showed a higher information content for the A-SCOPE system in comparison to OCO. Similar performances could be attained by the two extended surface networks, that might be implemented at the same cost of the A-SCOPE satellite mission. Additional insight was gained by testing, along with the nine individual systems mentioned above, different combinations of them.

In summary, a crucial issue to further develop our understanding of the processes controlling the balance between sources and sinks of atmospheric Carbon Dioxide is the assessment and improvement of the quality and spatial and time coverage of observations of CO_2 concentration. This is, in fact, directly linked to the information retrievable from inverse modelling of the Carbon cycle on CO_2 surface fluxes associated to natural and anthropogenic emissions and uptake. And it is mandatory to guide corrective and preventive actions for mitigation of climate change towards most effective measures aimed to reduce CO_2 sources from human activities and to protect and enhance natural Carbon sinks.

The IASI instrument is a Fourier transform spectrometer flying on-board the MetOp-A satellite and operating in the thermal InfraRed range for nadir looking observations of Earth's surface and atmospheric emission. The mission concept is primarily driven by measurement requirements for meteorological sounding and global climate monitoring and the operational retrieval targets are: atmospheric temperature and Water Vapour vertical profiles, O_3, CH_4, N_2O and CO total columns. The broad spectral coverage from 645 to 2760 cm^{-1} and the high spectral resolution of 0.5 cm^{-1} after apodization make also possible to extract from IASI spectra the total amount of other atmospheric minor constituents with spectral signatures in the thermal InfraRed region. In particular, the potential of IASI measurements to provide information on integrated content of Carbon Dioxide was demonstrated by feasibility studies [11], using simulated IASI data to investigate the sensitivity to CO_2 of the shorter and longer wavelength bands of the TIR, at 4.4 μm and 15 μm respectively, and to evaluate the capability of inverse modelling to disentangle CO_2 from the effects of temperature and other interfering species such as Water Vapour and Ozone. Real data from IASI and from AMSU (Advanced Microwave Sounding Unit), both on-board the MetOp-A mission, have also been used to retrieve the upper tropospheric column of CO_2 in the range 11-15 km, based on fourteen channels of IASI in the $\nu 2$ band of CO_2 at 15 μm and collocated measurements of AMSU, not sensitive to Carbon Dioxide as independent information on temperature [23].

The capability of advanced high resolution InfraRed sounders, and specifically of IASI, to retrieve information on CO_2 integrated content in the upper troposphere is definitely relevant both as a direct source of experimental data to constrain inversion models of CO_2 surface fluxes and as an independent measurement for cross-comparison with other existing or planned space-borne sensors such as TANSO-FTS/GOSAT, OCO-2 and CarbonSat. Moreover, IASI has the advantage of providing a very long data series, with the launched missions (MetOp-A 2006, MetOp-B 2012) and the new scheduled mission: MetOp-C (launch scheduled in 2017) and IASI New Generation (IASI-NG), now having reached the accomplishment of design phase-A for launch in the 2020, as part of the EPS-SG (EUMETSAT Polar System-Second Generation, [49]).

1.5 Objectives of this work

This is the general framework in which the KLIMA (Kyoto protocoL Informed Management of the Adaptation) study was collocated. The primary objective of the research activity was to investigate the ultimate performances of the retrieval of CO_2 from IASI observations in the thermal InfraRed, in order to provide additional evidence to the on-going discussion on the measurement of Carbon Dioxide from space. Moreover, since the CO_2 distribution at mid-to-upper tropospheric levels is heavily undersampled, the development of retrieval models in the TIR are particularly important ([86]). As a part of the goals of this thesis work, there is also the achievement of the requirements of TANSO-FTS operational products (0.3%, corresponding to 1 ppm out of 370 ppm, on monthly averages over 1000 km x 1000 km areas) to perform a cross-validation between IASI and TANSO-FTS of atmospheric CO_2 concentrations and to assess the coherence and temporal behaviour of large scale spatial features obtained from the two different sensors.

The study was developed in several steps:

- Adaptation of the KLIMA retrieval algorithm into a non-operational inversion code optimized for fast and accurate retrieval of CO_2 average information from IASI calibrated spectra;
- Integration of the optimized KLIMA CO_2 retrieval code into the ESA Grid Processing On-Demand (G-POD) operational environment [29];
- Inter-comparison of KLIMA CO_2 total column values with EUMETSAT Polar System (EPS) IASI Level 2 (L2) data;
- Inter-comparison of CO_2 products retrieved from collocated IASI and TANSO-FTS measurements for cross-validation of the two instruments;

- Inter-comparison of CO_2 products retrieved from IASI spectra with KLIMA algorithm with the Total Carbon Column Observing Network (TCCON) ground stations products.

From the original MARC code [9] developed by the Earth observation group of IFAC-CNR, a prototype software system, indicated as KLIMA *Reference Retrieval Model (RRM)*, was developed, essentially driven by the requirements for optimal accuracy of CO_2 retrieval product. As a further step, some approximations in the radiative transfer model was introduced to obtain a version of the retrieval code, indicated as KLIMA *Accelerated Retrieval Model (ARM)*, suitable for the integration into the ESA-ESRIN G-POD system, to have access to the computing resources for processing a large dataset of IASI data. The KLIMA *ARM* represents the best trade-off between accuracy and computing time.

The main innovation introduced by the KLIMA *ARM* is the retrieval of CO_2 together with most of the relevant interfering parameters so that interferences contribute directly to the retrieval error and do not pose difficult error assessment problems.

The performance of the *ARM* in terms of total retrieval uncertainty are only slightly degraded with respect to the *RRM* and are fully consistent with the requirement of 0.3% accuracy of CO_2 total column at monthly intervals on regional scales for TANSO-FTS/GOSAT comparison.

1.6 Outline of the Thesis

This Thesis Work consists of seven Chapters, including the Introduction. A general overview of the theory of the Radiative Transfer in the atmosphere and of the KLIMA Forward Model, along with the description of the validation of KLIMA Forward Model are presented in Chapter 2. The Inversion Theory and an overview of the main features of the KLIMA Retrieval Model are presented in Chapter 3. In this Chapter, background information on the characteristics, that the KLIMA prototype software inherited from existing forward and inverse models, is reported along with details of the innovative features that have been implemented in the upgraded version of the code. Moreover, a description of the integration of an optimized version of the KLIMA code on the G-POD for Earth Observation Application System is presented. Chapter 4 reports the options available for the access to IASI, TANSO-FTS and TCCON operational data, in terms of data dissemination means, measurement products types and data processing. The rationale for the use of IASI Level 1 B (L1B), 1 C (L1C) and Level 2 (L2) data obtained from the EUMETSAT archive facility (UMARF) and of TANSO-FTS L2 data downloaded from GOSAT User Interface Gateway (GUIG) is also discussed. In Chapter 5 are presented the results of bulk processing of the IASI measurements, carried out using the G-POD computing resources. In this Chapter is it shown the global distribution of KLIMA L2 products and a description of the errors of KLIMA retrieval. Chapter 6 illustrates a brief overview of the adopted strategy for data analysis and reports the results of inter-comparison between Carbon Dioxide retrieved with the integrated version of the KLIMA algorithm, EUMETSAT L2 operational product (Section 6.3), TANSO-FTS L2 (Section 6.5) and selected TCCON ground stations (Section 6.4). In Chapter 7, final conclusions are presented, along with a possible future works that can be derived from this project.

Finally, in Appendix A it is reported, for completeness, the first inter-comparison between KLIMA and Version 1 (V1) of TANSO-FTS L2 products, that has been updated in the final part of the work to Version 2 (V2).

Chapter 2

The Atmosferic Radiative Transfer: KLIMA Forward Model

The theory of **Atmospheric Radiative Transfer** describes the interaction between electromagnetic energy (solar radiation and terrestrial radiation) and matter (gases, aerosols, clouds) in the Earth's or a planetary atmosphere. In general, it describes the case of loss and gain in intensity due to the interaction of radiation with gas molecules, aerosol particles and clouds associated with different processes of absorption, scattering and emission. This Chapter presents on overview of the theory that lays at the basis of the constructions of a Forward Model for the modelling of the interaction between solar and terrestrial radiation and matter, that is a critical and essential block to infer the physical and chemical characteristics of the atmosphere. It illustrates the general concepts of the processes that take place in the atmosphere, how they are described through the Equation of Radiative Transfer and how all the ingredients for forward modelling are implemented in KLIMA code.

The KLIMA algorithm consists of two distinct modules, that can operate independently or be combined: the Forward Model, discussed in 2.4 and the Retrieval Model, discussed in Chapter 3.

The primary task of this work is the adaptation of the KLIMA algorithm into an inversion code to investigate the ultimate capabilities of the algorithm to retrieve Carbon Dioxide from IASI spectrometer, on-board the METOP-A satellite. The subsequent goals is the integration of an optimized version of KLIMA code into the ESA G-POD operational environment for bulk processing of IASI Level 1 data, and the comparison between KLIMA and TANSO-FTS CO_2 products for cross-validation purposes. To reach these goals a non-operational radiative transfer code has been developed for processing IASI Level 1 data: the Reference Forward Model (RFM), described in Section 2.4.1. KLIMA Reference Forward Model has been validated against a state-of-art Forward Model 2.5. The RFM achieves the best performances for the retrieval of CO_2 columns, with total retrieval error on CO_2 total column of $\sim$1.0%. This configuration, however, does not meet the requirements in terms of program size and running time for integration on G-POD.

In order to achieve a reasonable trade-off between retrieval quality and efficiency and to meet the requirements for integration on G-POD environment a series of approximations was introduced in the RFM. The version implementing these model approximations has been indicated as KLIMA Accelerated Forward Model (AFM), described in Section 2.4.2. The impact of these approximations on the accuracy of the FM was evaluated and the performance of the AFM are only slightly degraded respect to the RFM.

2.1 Theory of the Radiative Transfer in the atmosphere

The most important process responsible for energy transfer in the atmosphere is electromagnetic radiation. Almost all the energy exchange between the Earth and the rest of the Universe is made by radiative transfer. The Earth and the atmosphere absorb continuously energy from the Sun in the form of electromagnetic radiation and re-emit it to space. Sun is the energy source and the radiative exchange of energy between Sun-Earth-Space is the key for energy budget. The most significant spectral regions associated

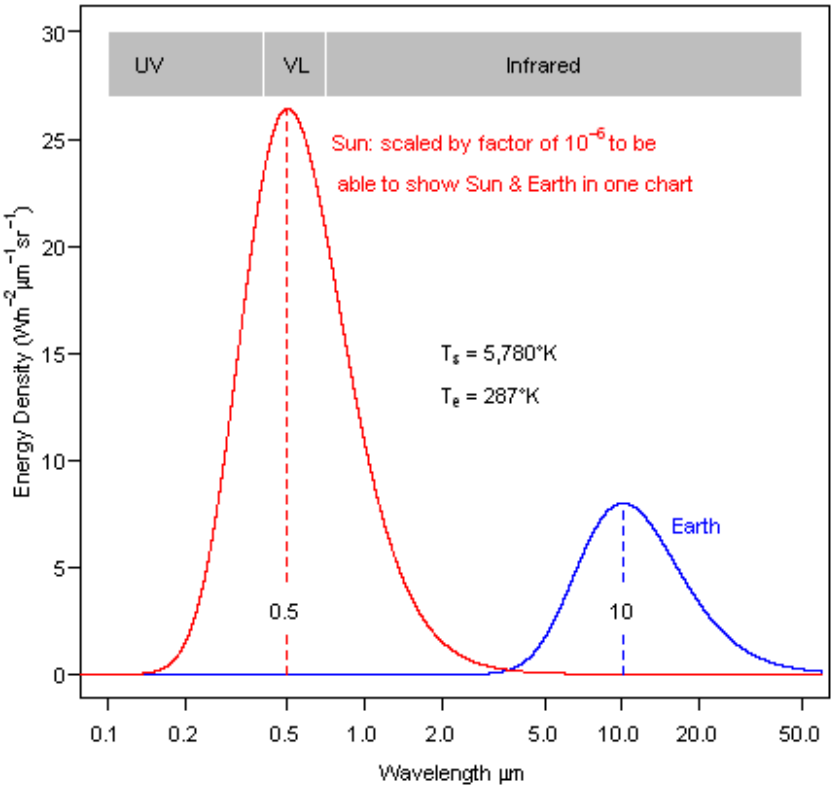

Fig. 2.1: *Comparison of the emission spectra of Sun and Earth*

with radiative energy transfer in planetary atmospheres lie between UltraViolet (UV) and MicroWave (MW). The radiation that propagates in the atmosphere is composed of two main contributions:

- Sun's energy, emitted mainly in the bands of the visible (VIS) and UV, the so-called short-wavelength (SW), characterize by a wavelength $\lambda < 4.0$ μm;
- radiation emitted from surface and atmosphere mostly in the IR region or long-wavelength (LW), with wavelength $\lambda > 4.0$ μm.

The two contributions can be described as the radiation emitted by black bodies at temperature T_{Sun} and T_{Earth} and are significant in spectral regions clearly separated, as showed in Figure 2.1. So, the solar and terrestrial radiation can be treated separately.

2.1.1 Basic radiometric quantities

Flux and intensity are the two measures of the strength of an electromagnetic field that are central to most problems in atmospheric sciences. Remote sensing is based on the use of radiative measurements to probe the atmosphere. In general, it is possible to say that a surface emits a certain irradiance flux or specify for each direction the emitted intensity.

Intensity (or *radiance*) is defined as radiant energy in a given direction per unit time per unit wavelength (or frequency) range per unit solid angle per unit area perpendicular to the given direction. In general, radiance is a function of the coordinates (r), direction (Ω), wavelength (or frequency), time and it is always positive. Radiance, as a function of position and direction, gives a complete description of the electromagnetic field because intensity describes the angular variation of the radiation flux in a certain point. If radiance does not depend on the direction, the electromagnetic field is said to be isotropic. If intensity does not depend on position the field is said to be homogeneous.

Flux (or *irradiance*) is defined as radiant energy in a given direction per unit time per unit wavelength (or frequency) range per unit area perpendicular to the given direction. Monochromatic flux is the

integration of normal component of monochromatic intensity over some solid angle.

2.1.2 Fundamental of extinction and emission

Knowledge on interaction of matter and radiation is necessary for the interpretation of remote sensing results. There are several type of interactions between an electromagnetic radiation field and a medium (e.g., the atmosphere). First, some of the energy contained in the incident beam can be converted into other forms of energy such as heat. This phenomenon is called *absorption*. Second, some of the incident energy can be extracted from the beam and diffused in all directions. This phenomenon is called *scattering*. Scattering can be thought of as absorption of radiant energy followed by re-emission back to the electromagnetic field with negligible conversion of energy. Thus, scattering can remove radiant energy of a light beam travelling in one direction, but can be a source of radiant energy for the light beams travelling in other directions. In addition, if the absolute temperature of the medium is not equal to zero, then it emits radiation in all directions and at all frequencies. This phenomenon is called *thermal emission*.

These processes can be divided into two categories: processes that *remove* the radiant energy from an electromagnetic field and transfers it to other forms of energy (*extinction*) and processes that *increase* the intensity of the electromagnetic field (*emission*).

As a result of absorption and scattering, the energy of the incident beam is reduced by an amount equal to the sum of the absorbed and scattered energy: *Extinction = absorption + scattering*. The fundamental law of extinction is the *Beer-Bouguer-Lambert law*, which states that the extinction process is linear in the intensity of radiation and amount of matter (2.1.3).

On the other hand, the process of emission increases the intensity of radiant energy through the mechanisms of thermal emission and scattering: *Emission = thermal emission + scattering*.

The thermal emission is a process through which each material at a temperature greater than 0 K emits radiation for spontaneous transitions between energy levels of atoms and molecules and follows the Planck's law. At wavelengths longer than 4 μm the contribution of thermal emission in the atmosphere becomes important.

Absorption in a volume is proportional to the initial radiance, the light path, the number density of absorbers and the absorption cross-section. The integrated absorption is called opacity or *optical depth* (OD) of the considered layer of the medium, indicated as τ, which expresses the quantity of light removed from a beam by scattering or absorption during its path through a medium. Optical depth is dimensionless.

Another important quantity, related to the optical depth, is the *transmissivity*, as showed in Figure 2.2, that is the fraction of incident light at a specified wavelength that passes through a sample, described by the relation:

$$T_\lambda(s) = \frac{I_\lambda(s)}{I_\lambda(0)} = exp(\int_0^s \varrho\sigma_{abs,\lambda})\, ds = exp(-\tau_\lambda(s)) \tag{2.1}$$

where ϱ is the density of the material. The total transmissivity of the atmosphere is obtained as the product of transmissivity of the individual components. The sensors that are used for remote sensing of clouds, surface or aerosols must be receptive in the spectral regions where the transmission of radiation through the atmospheric layers is highest: these regions are called *windows*.

2.1.3 The extinction law

Consider a small volume ΔV of infinitesimal length *ds* and area ΔA containing optically active matter: the change of intensity along a path *ds* is proportional to the amount of matter in the path. As the radiation passes through the volume it interacts with particles through processes of absorption and diffusion and a small amount of energy will emerge from the opposite side. The energy that emerges will be given by the initial energy decreased by a contribution. The fundamental law of extinction is the *Beer-Bouguer-Lambert law*, which states that the extinction process is linear in the intensity of radiation

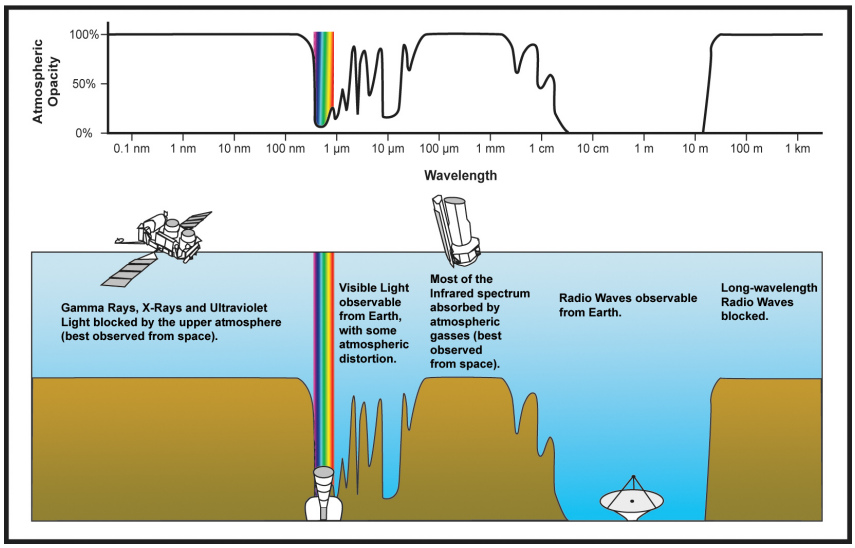

Fig. 2.2: *The atmospheric transmissivity at different wavelengths*

and amount of matter, provided that the physical state (i.e., T, P, composition) is held constant. If the incident intensity at s=0 is $I_\lambda(0)$, then the emergent intensity at a distance s is given by:

$$I_\lambda(s) = I_\lambda(0)exp(-\int_0^s k_\lambda \varrho \, ds) \tag{2.2}$$

The constant of proportionality k_λ is defined as the extinction coefficient and denotes the mass extinction cross section for radiation of wavelength λ. Assuming that the medium is homogeneous, so that k_λ is independent of the distance and defining the path length as:

$$u = \int_0^s \varrho \, ds \tag{2.3}$$

equation 2.2 can be expressed by:

$$I_\lambda(s) = I_\lambda(0)exp(-k_\lambda u) \tag{2.4}$$

This expression is known as *Beer-Bouguer-Lambert law* which states that the decrease in the radiant intensity traversing a homogeneous extinction medium is in accord with the simple exponential function whose argument is the product of the mass extinction cross section and the path length.

Tab. 2.1: *Radiative transfer: A = processes of primary importance, B = processes of secondary importance, C = processes with negligible effects*

	Solar radiation		Earth radiation	
Atmospheric component	Absorption	Scattering	Absorption+Emission	Scattering
Gas	A	B	A	C
Aerosols	B	B	B	C
Clouds	B	A	A	C

2.1.4 Local Thermodynamic Equilibrium

The atmosphere is not strictly in the thermodynamic equilibrium because its temperature and pressure are functions of position; scattering is an example of non-local thermodynamic equilibrium effect. Usually the atmosphere is subdivided into small subsystems, each of which is effectively isothermal and isobaric, referred to as *Local Thermodynamical Equilibrium (LTE)*.

LTE means that the intensive parameters (such as the temperature) are varying in space and time, but are varying so slowly that, for any point, one can assume thermodynamic equilibrium in some neighbourhood about that point. In this situation, the absorbing/emitting medium consists of massive particles which are in equilibrium with each other, and therefore have a definable temperature. For a medium in LTE, the emission coefficient and absorption coefficient are functions of temperature and density only.

The concept of LTE plays a fundamental role in atmospheric studies: e.g., the main radiation laws, which are strictly speaking valid in thermodynamical equilibrium, can be applied to an atmospheric air parcel in LTE. Knowing the temperature profile and the density profile of the medium is sufficient to calculate a solution to the equation of radiative transfer.

2.2 The equation of Radiative Transfer

A *Forward Model* (FM) provides the simulation of the spectra measured by an instrument for given optical path, atmospheric composition and instrumental properties. Thus the FM has to solve the *Radiative Transfer Equation* (RTE) along the Line Of Sight (LOS) taking into account the Instrument Line Shape (ILS) and the Field Of View (FOV) of the instrument.

The RTE is a differential equation for calculating the variations of spectral intensity of the radiation that propagates in the Earth atmosphere. The RTE describes the general case of loss and gain in intensity due to the interaction of radiation with gas molecules, aerosol particles and clouds associated with different processes of absorption, scattering and emission. The main contributions to Radiative Transfer, depicted in Figure 2.3, can be summarized as follows:

- Direct Solar Ray;
- Reflection on the Surface;
- Reflection from Clouds;
- Scattering in the Atmosphere: Rayleigh, Mie and Raman Scattering;
- Absorption in the Atmosphere;
- Emission in the Atmosphere;
- Emission from the Surface;
- Emission from Clouds;

A pencil of radiation, usually represented by its intensity (or radiance) I_λ, traversing a medium will be weakened by interaction with matter through absorption; in particular the reduction in intensity is due to absorption by the material as well as to scattering by the material. At the same time this radiation may be strengthened by thermal emission from the medium plus scattering from all other directions into the pencil under consideration at the same wavelength.

From the Extinction law 2.4, the general equation for radiative transfer in an absorbing and emitting medium can be written in terms of the differential change in the intensity in the form:

$$\frac{dI_\lambda}{k_\lambda \varrho ds} = -I_\lambda + J_\lambda \tag{2.5}$$

where k_λ is the absorption coefficient, ϱ is the density of absorbing gases, s is the slant path and J_λ is the *source function*, which takes into account the increase in intensity due to emission and multiple scattering. This is the general Radiative Transfer Equation without any coordinate system imposed, which is fundamental to the discussion of any radiative transfer process.

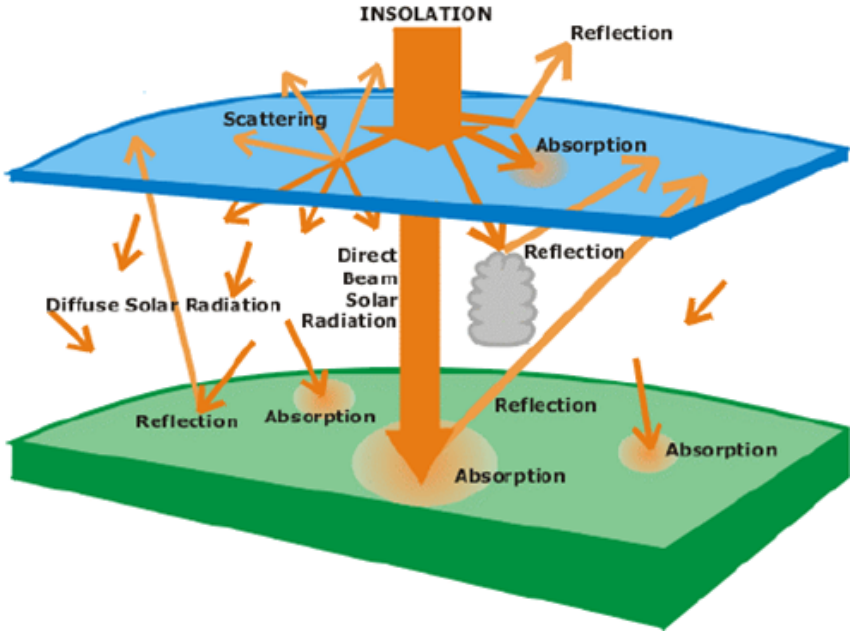

Fig. 2.3: *A schematic description of the main contribution to Radiative Transfer in the atmosphere*

2.2.1 The Equation of Radiative Transfer for plane-parallel atmosphere

In the case for the transfer of thermal InfraRed radiation emitted from the Earth and the atmosphere, it is commonly assumed that the intensity is being independent of time and, in localized portions, the atmosphere is in thermodynamic equilibrium. This implies that the source function is described by the Planck function. Moreover, for many applications, the atmosphere can be approximated by a plane-parallel model to handle the vertical stratification of the atmosphere, so the variations in the intensity and atmospheric parameters (temperature and gaseous profiles) are permitted only in vertical direction (e.g., height or pressure). Plane-parallel atmosphere consists of a certain number of atmospheric layers, each characterized by homogeneous properties (e.g., T, P, optical properties of a given species, etc.) and bordered by the bottom and top infinite plates (called boundaries), as showed in Figure2.4.

Traditionally, the vertical coordinate z is used to measure linear distances in the plane-parallel atmosphere: z = s $\cos(\theta)$ and ds = dz/$\cos(\theta)$, where θ denotes the angle between the upward normal and the direction of propagation of a light beam (or zenith angle) and ϕ is the azimuthal angle.

Under these conditions and considering a non-scattering medium, the intensity is a function of the vertical position and the zenith angle and the basic equation that governs TIR radiation in the height coordinate is (omitting the subscript λ on various radiative quantities):

$$\cos\theta \frac{dI(z;\theta,\phi)}{k\varrho dz} = -I(z;\theta,\phi) + B(z;\theta,\phi) \tag{2.6}$$

where θ denotes the inclination to the upward normal. Introducing the optical depth:

$$\tau = \int_z^\infty k\varrho\, dz' \tag{2.7}$$

the equation 2.6 can be written as:

$$\mu \frac{dI(\tau;\theta,\phi)}{d\tau} = -I(\tau;\theta,\phi) + B(\tau;\theta,\phi) \tag{2.8}$$

where μ = $\cos\,\theta$. Equation 2.8 can be solved to give the upward and downward intensities for a finite atmosphere that is bounded on two sides at τ=0 and τ=τ_* as illustrated in Figure 2.4. In order to solve both upward and downward components for an atmosphere with a total optical depth τ_*, two boundary conditions are required:

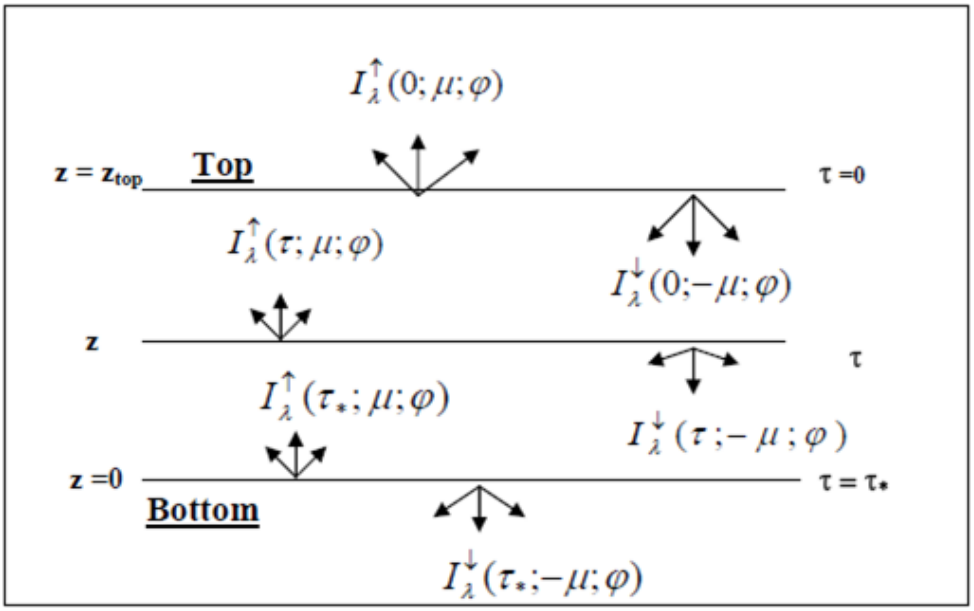

Fig. 2.4: *Upward and downward intensities at a given level τ an at the top (τ=0) and bottom (τ=τ_*)*

- Surface: assumed to be a black body in the IR emitting with the surface temperature T_s;
- Top of the atmosphere no downward thermal emission.

With these boundary conditions the formal solution for upward and downward intensities are given by:

$$I(\tau;\mu,\phi) = B(\tau_*)exp\frac{-(\tau_* - \tau)}{\mu} + \int_{\tau}^{\tau_*} B(\tau')exp\frac{-(\tau' - \tau)}{\mu}\frac{d\tau'}{\mu} \tag{2.9}$$

$$I(\tau;-\mu,\phi) = \int_{0}^{\tau} B(\tau')exp\frac{-(\tau - \tau')}{\mu}\frac{d\tau'}{\mu} \tag{2.10}$$

The upward flux at a given level comes from two sources: the surface emission that is attenuated to that level and the emission contributions from the atmospheric layers characterized by Planck fluxes. Likewise the downward flux at a given level is produced by contributions from the atmospheric layers. Finally, to account for the contributions from all wavenumbers in the TIR spectrum an integration of the monochromatic flux with respect to wavenumbers must be performed. The computation of atmospheric fluxes involves solving the integrations over the wavenumber and along the optical depth.

2.3 Atmospheric Radiative Transfer codes

An Atmospheric Radiative Transfer Model or *Forward Model* (FM) calculates radiative transfer of electromagnetic radiation through a planetary atmosphere, such as the Earth's atmosphere. It provides the synthetic radiances, as they would be measured by the sensor, for a specific state of the atmosphere. FM is an essential and critical part of any retrieval process, it represents the first step where the atmosphere and all the phenomenons link to propagations of radiation are modelled: more accurate is the modelling and more accurate will be the retrieval process.

At the core of a Radiative Transfer Model lies the Radiative Transfer Equation (RTE) that is numerically solved using a solver such as a discrete ordinate method or a Monte Carlo method. The radiative transfer equation is a monochromatic equation to calculate radiance in a single layer of the Earth's atmosphere. To calculate the radiance for a spectral region with a finite width (e.g., to estimate the Earth's energy budget or simulate an instrument response), one has to integrate this over a band of frequencies (or wavelengths).

The correct way to do this is to loop through the frequencies of interest, and for each frequency, calculate the radiance at this frequency. For this, one needs to calculate the contribution of each spectral line for all molecules in the atmospheric layer; this is called a *line-by-line* (LBL) calculation. The method of LBL calculation of the molecular absorption characteristics is widely used in modelling of Radiative Transfer. LBL method is considered to be an exact computation of radiation propagation in the gaseous absorbing/emitting inhomogeneous atmosphere, accounting for all (known) gas absorption lines in the wavenumber range from 0 to about 23000 cm^{-1}. To obtain the signal at the detector is then necessary convolved the result of LBL calculation with the spectral response of the instrument.

2.3.1 The spectroscopic database

For a line-by-line calculation, a FM needs characteristics of the spectral lines, such as the line centre, the line width and the shape; these information are contained in a spectroscopic database, that is a very important ingredient of the radiative transfer calculation. One of the most common and used spectroscopic database is the HIgh-resolution TRANsmission molecular absorption database (HITRAN) developed by the Harvard-Smithsonian Center for Astrophysics ([40]). HITRAN is a compilation of spectroscopic parameters that a variety of computer codes use to predict and simulate the transmission and emission of light in the atmosphere. The database is a long-running project started by the Air Force Cambridge Research Laboratories in the late 1960's in response to the need for detailed knowledge of the InfraRed properties of the atmosphere. This molecular database is updated each time new or modified data are added to the compilation: data are frequently added, modified, or enhanced with respect to the previous versions. The versione implemented in this work, HITRAN2008 described in [71], covers the spectral region from MW to UV and contains 2.713.968 spectral lines for 42 different molecules.

2.4 Forward Modelling of IASI observations

The primary task of this work is the adaptation of the KLIMA algorithm into an inversion code to investigate the ultimate capabilities of the algorithm to retrieve CO_2 total column from nadir sounding observations acquired in the thermal InfraRed by the IASI spectrometer, on-board the METOP-A satellite. The KLIMA algorithm consists of two distinct modules, that can operate independently to produce simulated observations or be combined for inverse processing of real or synthetic data: the Forward Model, discussed in this Chapter, and the Retrieval Model, discussed in Chapter 3. The subsequent goal is the utilization of KLIMA algorithm for the processing of IASI Level 1 data and the creation of a database of CO_2 retrieved data from IASI measurements, used for the comparison with the operational TANSO-FTS products, for a cross-validation of the two instruments. In order to analyse the large amount of IASI data, the KLIMA algorithm has exploited the computing resources of the ESA G-POD system (3.3).

First step to achieve these results has been the development of a non-operational radiative transfer code for processing IASI Level 1 data: the Reference Forward Model (RFM), described in Section 2.4.1. The RFM can simulate wideband nadir radiances, in the complete IASI range from 645 to 2760 cm^{-1}, measured at the top of the atmosphere using line-by-line radiative transfer calculation. The code computes the radiance that reaches the sensor and simulates the instrumental effects. The RFM associated with the retrieval model achieves the best performances for the retrieval of CO_2 columns, with total retrieval error on CO_2 total column of $\sim$1.0%. This configuration, however, does not meet the requirements in terms of program size and running time for integration on G-POD.

In order to achieve a reasonable trade-off between retrieval quality and efficiency, with driving requirements established by the final goal of cross-validation with TANSO-FTS products and by the need for integration on G-POD environment, i.e.: target accuracy of 0.3% (1 ppm out of 370 ppm) on regional scales (1000 km x 1000 km) at monthly intervals and program size not exceeding 1 Gbyte and running time aimed at processing 1 orbit of IASI data in 1 day using G-POD computing resources, a series of approximations was introduced in the RFM. The version implementing these model approximations has

been indicated as KLIMA Accelerated Forward Model (AFM), described in Section 2.4.2.

The impact of these approximations on the accuracy of the FM was evaluated and the results were used for fine-tuning individual choices and implementing an overall procedure optimally suited for the selected target. The performance of the AFM are only slightly degraded respect to the RRM and are fully consistent with the requirement of 0.3% accuracy of CO_2 total column at monthly intervals on regional scales for TANSO-FTS validation.

The major differences between the two versions of the Forward Model (RFM and the AFM) are summarized in Table 2.2.

Tab. 2.2: *Approximations adopted for the KLIMA AFM with respect to the reference configuration (RFM)*

	Reference FM	**Accelerated FM**
Frequency Grid	1/128 of IASI spectral sampling	1/64 of IASI spectral sampling
Line-Mixing	Modelled in the band [645-2760] cm^{-1}	Modelled in the bands [645-825] and [1995-2445]cm^{-1}
Spectroscopic Database	HITRAN (All lines)	HITRAN(skipping all lines smaller than 10^{-5} the largest line in the selected band)
Atmospheric Line Shape	All lines modelled up to ± 25 cm^{-1} from line centre	Lines modelled up to ± 25 cm^{-1} from centre (target species) ± 10 cm^{-1} from centre (other species)
Vertical Grid	90 levels from 1050 to 0.005 hPa	44 levels from 1050 to 0.005 hPa

2.4.1 KLIMA Reference Forward Model

KLIMA Forward Model is a line-by-line radiative transfer model with capability to simulate wide-band spectral radiances acquired by the IASI instrument. It is based on the following key features:

- radiative transfer calculations performed using Curtis-Godson approximation ([72]);
- atmospheric line-shapes modelled with Voigt profile;
- atmospheric continuum model taking into account the main contributions from N_2, O_2, O_3, H_2O, and CO_2.

Scattering effects have not been implement because not necessary under the condition of clear sky and Non Local Thermodynamic Equilibrium (NLTE) effects have not treated because not significant in the tropopause.

The spectroscopic database adopted for data processing is HITRAN2008. Dedicated spectroscopic database and line shape are implemented for CO_2, to take into account the line-mixing effect [63], [64]. The fine frequency grid sampling used to simulate the IASI spectra before the convolution with the instrumental function has been fixed to 0.004 cm^{-1}. Moreover, the correction of the Planck function [18] to take into account the optical depth of the atmospheric layer at the different frequencies has been included.

KLIMA FM has been validated by comparing synthetic IASI measurements generated by KLIMA FM code with those of the FM of the LBLRTM (Line-By-Line Radiative Transfer Model) code ([18]), as reported in Section 2.5. This validated FM, when used with a fine sampling of the functional parameters and a complete accounting of all contributions, provides the *Reference Forward Model* (RFM).

2.4.2 KLIMA Accelerated Forward Model

A series of approximations was introduced in the RFM to improve the efficiency of the code and to meet the requirements on program size and computing time. This Forward Model configuration,

called *Accelerated Forward Model* (AFM), is a trade-off between efficiency (in term of program size and running time) and accuracy of the FM code, taking into account that the goal is the retrieval of the Carbon Dioxide column. For each approximation several tests were performed in order to identify the entity of the approximation that best satisfies the aimed compromise.

The approximations implemented in the AFM with respect to the RFM are here below briefly summarized:

- *Fine Frequency Grid*: a regular grid with steps equal to 1/128 the spectral sampling of the IASI instrument (0.25 cm^{-1}) is adopted for the RFM. A coarser grid, with steps equal to 1/64 the spectral sampling of IASI, was adopted in the AFM.

- *CO_2 line mixing*: in the RFM line-mixing effects are taken into account, by adding to the simulated radiances the contribution of CO_2 line mixing calculated according to the model by Niro et al. [63], [64] over the full IASI spectral range. A significant reduction of the computing time is obtained in the AFM by evaluating the contribution of line-mixing in a sub-set of spectral intervals.

- *Spectroscopic database*: The spectroscopic database adopted for the RFM includes all lines listed in HITRAN2008. In order to limit the total number of spectral features to be considered in the AFM, we evaluated the impact of skipping for each species, all the lines smaller than 10^{-5} the largest line in the selected band.

- *Atmospheric line shape*: In the RFM the atmospheric line shape was modelled up to ± 25 cm^{-1} from the line centre. The impact on computing time and FM accuracy of modelling the atmospheric line shape for non-target species (i.e., all the simulated species except H_2O, CO_2, and O_3) up to ± 10 cm^{-1} from the line centre was evaluated by calculating the difference between the spectrum simulated with and without this approximation.

- *Atmospheric vertical grid*: The vertical grid adopted for modelling the stratification of the atmosphere in the RFM is the same that is used to represent IASI L2 operational products for temperature and Water Vapour, consisting of 90 levels in the pressure range from 1050.0 hPa to 0.005 hPa. In the AFM, the vertical resolution of the pressure grid was degraded: the 21 levels, closest to surface, from 1050.0 hPa to 478.4 hPa (2 km) were the same as in the reference vertical grid. For the levels above, only one every third pressure level was maintained in the new grid. The total number of pressure levels in this approximation is 44.

- *Field of view*: The effects due to the field of view of the instrument neither in the RFM nor in the AFM are taken into account. Assuming uniform surface properties for the observed pixel, for different lines of sight within the field of view the spectral intensity varies by less than 1/10 the nominal IASI Noise Equivalent Spectral Radiance (NESR).

We verified that with the only exception of a few spectral channels, the approximations introduce differences that are significantly smaller than the measurement error.

2.5 Validation of KLIMA Reference Forward Model

A validation of KLIMA RFM was conducted, by comparing synthetic IASI measurements generated by KLIMA FM code with those of the FM of the LBLRTM (Line-By-Line Radiative Transfer Model) code ([18]). The simulations were based on the night-time measurement acquired by IASI over the Southern Great Plains in Oklahoma, USA, on 19 April 2007, during the Joint Airborne IASI Validation Experiment (JAIVEx) [81]. The atmospheric scenario for the simulations was set according to the state retrieved by LBLRTM from this observation. Synthetic IASI measurements were calculated using the KLIMA FM code and the LBLRTM FM and the difference between the two simulations was calculated.

The first results of the comparison between LBLRTM simulated spectrum and the spectrum simulated using KLIMA FM code showed a bias between the two simulations. The reasons of these residuals were

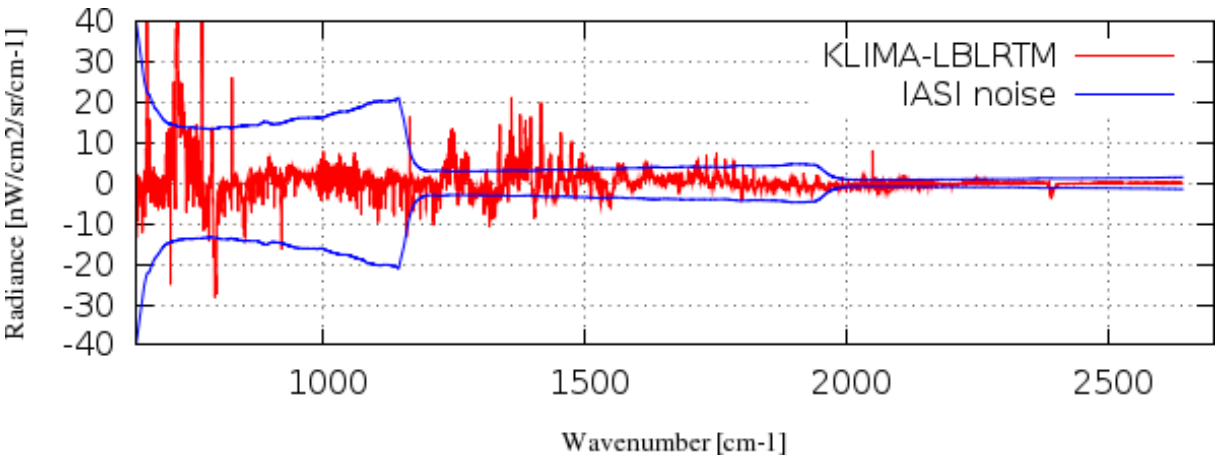

Fig. 2.5: *Residuals between KLIMA simulation and LBLRTM simulation, compared with the nominal values of IASI radiometric noise*

investigated; first of all it was supposed that this residual was due to the different combination of Planck functions adopted to simulate the contribution of each atmospheric layer to the RT. So it was implemented in KLIMA code the correction of the Planck function: the *linear in τ approximation*, as described in [18]; in particular, the *one-term Padé approximant* with $a = 0.278$ was implemented, to take into account the optical depth of the atmospheric layer at different frequencies. The simplest Padé approximant for the effective Planck function is of the form:

$$B(\tau) = [\overline{B} + (a\tau)B_U](1 + a\tau)^{-1} \tag{2.11}$$

which provides the correct result in the optically thick and thin limits. The subsequent step was to replace the instrumental line shape, originally taken from the Unified Meteorological ARchive Facility (UMARF) [87], with a Gaussian with a 1/e point at 0.30 cm^{-1}. Results obtained with these changes are reported in Figure 2.5.

Figure 2.5 reports the result of the comparison, showing the residual difference between KLIMA and LBLRTM synthetic spectra, along with the nominal values of IASI radiometric noise. The IASI spectrum considered in this the comparison is apodized and, accordingly, its radiometric noise is smaller than that of the unapodized spectra that will be considered in subsequent inter-comparisons. Furthermore, the differences shown in this comparison are smaller than the residuals of preliminary tests reported in [22] because of the previous further improvements. The current residuals are, with very few exceptions, significantly smaller than the measurement errors.

In Figure 2.6, we report the IASI measurement used for the validation activity (top) and the residuals between the observations and the simulations performed with the two FMs (bottom - KLIMA red line and LBLRTM blue line). The two codes have comparable performances and reproduce very well the observation. Several more comparisons have been done between synthetic spectra simulated with KLIMA code and LBLRTM on selected atmospheric scenarios and we can conclude that the comparison with a state of the art FM and the good agreement with real observations validate the quality of the KLIMA FM.

2.6 Outline of the Chapter

The first task of this work was the adaptation of the KLIMA algorithm into an inversion code to investigate the ultimate capabilities of the algorithm to retrieve CO_2 total column from observations acquired in the thermal InfraRed by the IASI spectrometer. This leaded to a non-operational, line-by-line radiative transfer code with capability to simulate wide-band spectral radiances acquired by the IASI instrument, the so-called Reference Forward Model. KLIMA Reference Forward Model has been

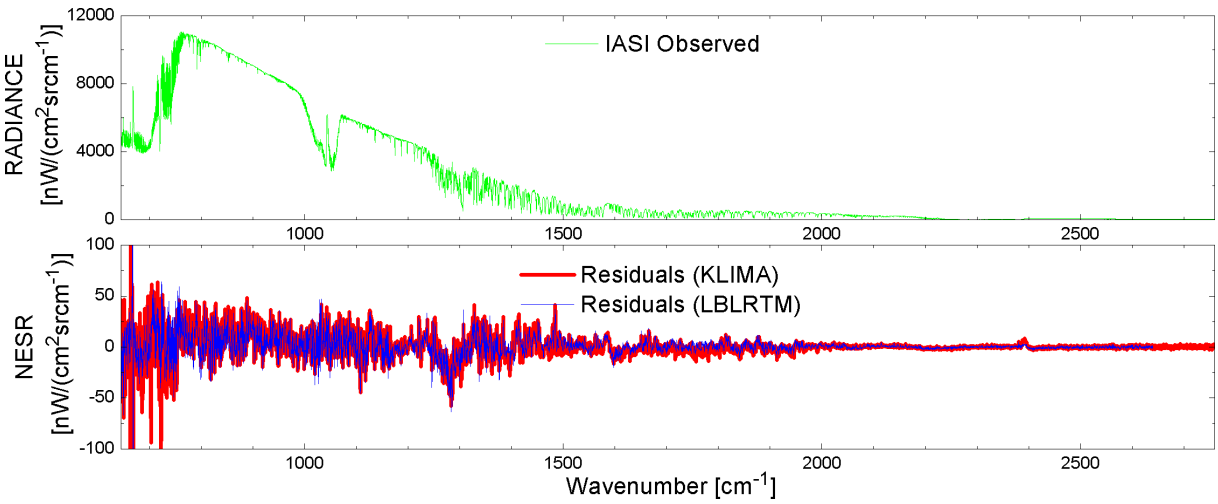

Fig. 2.6: *IASI measurements (top) and residuals wrt KLIMA (red line) and LBLRTM (blue line) simulation (bottom)*

validated by comparing synthetic IASI measurements generated by KLIMA FM code with those of the FM of the LBLRTM code. The residuals of these comparisons are always smaller than the standard deviation of IASI measurements. Subsequently, the quality of the RFM was also confirmed by the absence of significant biases in the average residuals of a large number of retrievals performed on real measurements.

In order to achieve a reasonable trade-off between retrieval quality and efficiency of KLIMA code, and to meet the requirements needed for integration on G-POD environment and the subsequence cross-validation with TANSO-FTS products, a series of approximations has been implemented in the KLIMA Reference Forward Model. It was verified that with the only exception of a few spectral channels, the approximations introduce differences that are significantly smaller than the measurement error. It is possible to conclude that the performance of the AFM are only slightly degraded respect to the RFM and are fully consistent with the requirement of 0.3% accuracy of CO_2 total column at monthly intervals on regional scales for TANSO-FTS validation.

Chapter 3

Inverse Theory: KLIMA Retrieval Model

The forward modelling described in Chapter 2 is generally employed to simulate a measured quantity, e.g., intensity of the radiation, for a predefinite state of the atmosphere. On the contrary, the objective of **inverse modelling** is to retrieve certain characteristics of the atmospheric state, for example concentration vertical profiles of atmospheric trace gases, based on the measured quantities. An intuitive idea of the inverse modelling is given in Figure 3.1: the inverse problem is the question of finding the best representation of the required parameter given the measurements made, together with any appropriate prior information that may be available about the system and the measuring sensor. Associated with the inverse problem there are also questions of understanding and describing the information content of the measurement, the relationship between the true state of the system and that retrieved using inverse method, the error analysis of the overall measuring system and validating results.

This Chapter describes the basic theory and equations that constitute an Inverse Model, which can be used in association with a Forward Model for the retrieval of chemical and physical atmospheric parameters. In particular, it is reported a summary of KLIMA inverse code and the activities carried out for the integration of optimized version of KLIMA code on the ESA G-POD Application System with the purpose of massive processing of IASI Level 1 data.

The KLIMA algorithm consists of two distinct modules, that can operate independently or be combined: the Forward Model, discussed in 2.4 and the Retrieval Model (RM). The retrieval approach, used for the processing of IASI measurements, exploits the wideband of IASI spectra to extract information on CO_2 taking into account the systematic errors due to interfering parameters and making the channel selection no longer necessary. This goal is reached using a multi-target approach: the RM algorithm retrieves the atmospheric CO_2 along with the interfering atmospheric parameters **??**, **??**. The RM uses a constrained Non-linear Least Square Fit (NLSF) approach and the cost function to be minimized takes into account the a priori information (Optimal Estimation Method) and the Marquardt parameter. Exploitation of broadband measurements is made possible by implementing a procedure that reduces the impact of systematic uncertainties. The code implements a multi-target retrieval: more than one species is simultaneously retrieved and systematic errors due to interfering parameters are removed.

Most inversion algorithms developed for both operational and scientific analysis of IASI spectra perform a reduction of the data (3.2); this is not the case of KLIMA code, which was built to include the feasibility of a retrieval approach exploiting all spectral channels of IASI. The *Reference Retrieval Model* (RRM) is obtained with the association of Reference Forward Model (2.4.1) with the RM. This code configuration is used to assess the ultimate retrieval capabilities of IASI for the exploitation of all the spectral channels of IASI, for a broadband and multi-target retrieval of the atmospheric state.

But 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, as described in 3.3.1 and 3.3.2. Using the Accelerate Forward Model (2.4.2) in association with the retrieval code a significant reduction of the computing time is obtained with respect to the use of the RRM. Moreover, no significant loss of accuracy is observed. The ARM version of KLIMA code satisfies the G-POD requirements so it is chosen for integration into the ESA G-POD

system 3.3.

The main elements of innovation introduced by the KLIMA ARM and by its application to the retrieval of Carbon Dioxide from IASI measurements can be recognized in the following aspects: demonstration of the feasibility of a retrieval approach exploiting all spectral channels of IASI to extract information on CO_2 and other target species with no need for channel selection and use of a multi-target retrieval scheme, which removes the systematic errors due to the interfering parameters.

3.1 Inversion Theory

The general inverse problem can be regarded as a question of setting up and solving a set of simultaneous linear or non-linear equations in the presence of experimental error in some of the parameters, and quite possibly in the presence of approximations in the formulation of the equation.

In the Inversion Theory the parameter to be retrieved from the measurements are represented by a *model state vector x*, that contains the number densities of atmospheric constituents defined at discrete altitude levels. Each state vector can be mapped to the measurements space by means of the *forward model operator F* to obtain the corresponding *measurement vector y*, i.e., for each atmospheric state described by a vector *x* an appropriate measured quantity *y* can be simulated using a Radiative Transfer Model as the FM.

Taking into account that measurements are made to a finite accuracy, a measurement error ϵ has to be considered which is generally assumed to be normally distributed with mean zero and known error covariance matrix S_y. The relationship between the model state vector and the measurement vector can be written as:

$$y = F(x) + \epsilon \tag{3.1}$$

In order to solve the inverse problem, this non-linear relationship has to be linearised expanding the forward model operator, *F*, as a Taylor series about a guessed value x_0 of the solution. Ignoring the higher-order terms one obtains:

$$F(x) \approx F(x_0) + \left.\frac{\partial F}{\partial x}\right|_{x_0} (x - x_0) = y_0 + K_0(x - x_0) \tag{3.2}$$

K_0 is the linearised FM operator. In the discrete representation the linearised FM operator is given by the weighting function matrix describing the sensitivity of the measured quantities to the variation of the atmospheric parameters at different altitude levels. This weighting function matrix is calculated with the radiative transfer model.

Atmospheric inversion problems are often "ill-posed" i.e., there is no uniqueness or stability of solutions (with respect to small perturbations in the data). Thus, additional constraints need to be introduced to determine a geophysical solution from the set of mathematically allowed solutions. Most commonly, the methods of statistical regularisation, as described e.g. in [72] are applied, i.e., the maximum likelihood condition, a priori value of the solution (x_0) and its covariance matrix (S_a) are employed to solve the inversion problem. In this case, the solution is found by minimising the following quadratic form:

$$\|(y - y_0) - K_0(x - x_0)\|^2_{S_y^{-1}} + \|x - x_0\|^2_{S_a^{-1}} \rightarrow min \tag{3.3}$$

This results in:

$$x_{n+1} = x_0 + (S_a^{-1} + K_n^T S_y^{-1} K_n)^{-1} K_n^T S_y^{-1}((y - y_n) - K_n(x_0 - x_n)) \tag{3.4}$$

where subscripts *n* and *n+1* denote the number of the iteration. The measurement error covariance matrix (S_y) is usually assumed to be diagonal, i.e. no correlation between measurement errors at different wavelengths is considered.

In the retrieval of the vertical distributions of the atmospheric species the *a priori* covariance matrix (S_a) can be chosen as a block diagonal matrix, i.e. vertical distributions of different atmospheric trace

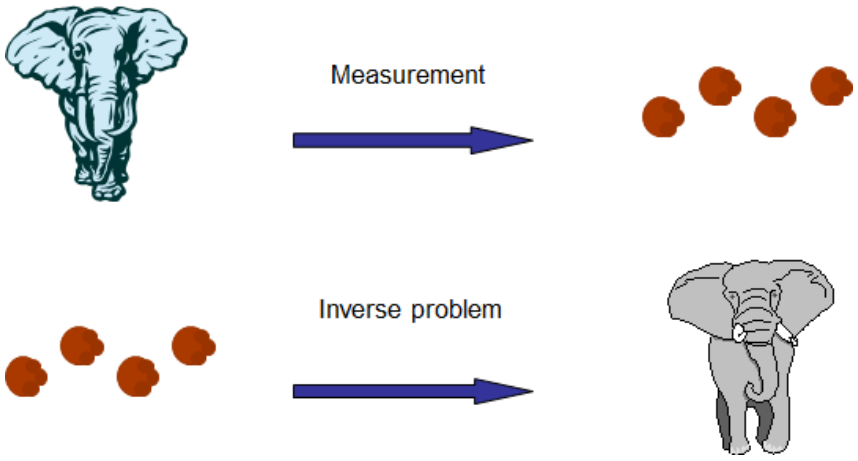

Fig. 3.1: *An intuitive explanation of the idea of inversion theory*

gases are assumed to be uncorrelated. The diagonal elements of S_a represent the variances of the vertical distribution of atmospheric trace gases (σ) which e.g. can be derived from a climatology.

The quality of the obtained solution is characterised by the *a posteriori* covariance matrix:

$$S = (K^T S_y^{-1} K + S_a^{-1})^{-1} \tag{3.5}$$

and by the *Averaging Kernels matrix* (AK):

$$A = \frac{\partial x}{\partial x_{true}} = (K^T S_y^{-1} K + S_a^{-1})^{-1} K^T S_y^{-1} K \tag{3.6}$$

characterising the response of the retrieved solution to the variation of the true atmospheric state. The root square diagonal elements of the *a posteriori* covariance matrix are referenced to as the theoretical precisions.

Employing the averaging kernels, the retrieved solution x can be related to the true solution x_{true} as:

$$x = x_0 + A(x_{true} - x_0) \tag{3.7}$$

i.e., if the model state vector represents a vertical profile of an atmospheric trace gas, the retrieved values at each altitude are expressed as the sum of the *a priori* values at this altitude and of the deviation of the true profile from an a priori profile smoothed with the associated row of the AK matrix. For an ideal observing system, A is a unit matrix. In reality, the rows of the AK matrix are peaked with a finite width, which can be regarded as a measure of the vertical resolution of the retrieved profile.

3.2 KLIMA Retrieval Code

Simultaneous processing of all IASI spectral channels received relatively low attention. The retrieval of atmospheric state parameters from broadband measurements acquired by high spectral resolution sensors generally requires dealing with a prohibitively large number of spectral elements (8461 samples in the case of each IASI observation). Most inversion algorithms developed for both operational and scientific analysis of IASI spectra (a synthetic description of these codes can be find in 4.1.3) perform a reduction of the data (typically based on channel selection, super-channel clustering or Principal Component Analysis (PCA) techniques) in order to handle the high dimensionality of the problem. Previous works on the analysis of IASI Level 2 ([23], [17]) performed a channel selection for the CO_2 retrieval, which associate the highest sensitivity to CO_2. As reported for instance in [23], have been considering channels located in the 15μm band (700 cm^{-1}), with a selection of only 14 channels, from 694.50 cm^{-1} to 704.25 cm^{-1},

This is not the case of KLIMA code, which was built to include the feasibility of a retrieval approach exploiting all spectral channels of IASI. Moreover, the code implements a *multi-target retrieval*: more than one species is simultaneously retrieved to extract information on Water Vapour, temperature and Ozone profiles, as well as on Carbon Dioxide. This multi target retrieval removes the systematic errors due to interfering parameters and makes the channel selection no longer necessary ([24], [22]).

The Retrieval Model (RM) uses a constrained Non-linear Least Square Fit (NLSF) approach and the cost function to be minimized takes into account the a priori information (Optimal Estimation Method - OEM) and the Marquardt parameter. The a priori is used to maintain the retrieved values within physically justified boundaries and only introduces a very weak constraint with negligible bias. On the other hand, the a priori is necessary to model the vertical structure of the profiles (Carbon Dioxide, temperature and Water Vapour) and to account for its induced errors.

The state vector used in the IASI measurements data analysis contains: the profile of temperature and the profile of Water Vapour; the total columns of CO_2, O_3, N_2O, CO, $CH_4$and the surface temperature and emissivity. The profiles are retrieved on a vertical grid of 30 levels from 1050 to 0.005 hPa. For temperature and Water Vapour a priori profiles the operational L2 data provided by UMARF archive have been used.

The challenging computation is made possible by the use of some approximations in the FM, that are tuned for a compromise between speed and accuracy. The FM module is used in the retrieval procedure implement for the analysis of IASI measurements.

The basic features of this retrieval model are:

- Wide band retrieval;
- Multi-target retrieval;
- Marquardt parameters equal to 0.01 for all targets;
- Marquardt and Gauss dumping factor equal to 3;
- Use of the Optimal Estimation.
- Possible use of the VCM that accounts for the FM approximations.

The target parameters of the retrieval are:

- The vertical profile of temperature;
- The vertical profile of Water Vapour;
- The total and partial column of Carbon Dioxide and Ozone;
- The total column of N_2O, CO and CH_4;
- The surface temperature.

The atmospheric scenarios selected for the retrievals and described by the *initial guess* and *a priori* values are taken from the *Initial Guess 2 climatology database* (IG2) [68]. The IG2 database is developed by University of Leicester for the operational analysis of MIPAS/ENVISAT data. The IG2 climatology represents the seasonal average atmosphere through a four season, six latitude band set of states. The four seasons are centred on January, April, July and October. The six latitude bands are $\pm$(90-65°), $\pm$(65-20°) and $\pm$(20-0°). An altitude grid of 0-120 km is employed. The IG2 seasonal climatologies were constructed to cover pressure and temperature, and concentrations of 36 species, including an update for CO_2 due to trends in its concentration. The IG2 profiles have been implemented in the code to retrieve trace gases, temperature and pressure profiles.

The thermal contrast between Earth surface and lowest atmospheric layer (defined as the difference between Earth skin temperature and the air temperature at ground level) was taken equal to zero for the true atmospheric state and equal to +10 K for the a priori state. The assumed a priori uncertainties on the retrieval targets are reported in Table 3.1.

Tab. 3.1: *A priori errors used for the retrieval tests*

Target Parameter	A Priori Error
Temperature	4 K
H_2O	80%
CO_2	10%
O_3, N_2O, CO, CH_4	100%
Surface Temperature	20 K

3.2.1 The Reference Retrieval Model

When the Reference Forward Model (2.4.1) is associated with the retrieval model we have the *Reference Retrieval Model* (RRM). This code configuration is used to assess the ultimate retrieval capabilities of IASI for the exploitation of all the spectral channels of IASI, for a broadband and multi-target retrieval of the atmospheric state. In the lower atmosphere RRM obtains an accuracy of 20% for Water Vapour and of 1 K for temperature. Table 3.2 shows the accuracy of the retrieved parameters. In particular, the total and partial column of CO_2 are retrieved with an accuracy of 1% and 5% respectively.

Tab. 3.2: *Retrieved parameters of retrieved model with Reference Forward Model*

Target Parameter	Accuracy
CO_2 Total Column	1%
CO_2 Boundary Column	5%
O_3 Total Column	0.5%
O_3 Boundary Column	20%
N_2O, CO, CH_4, and Surface Temperature	better than 1%

3.2.2 The Accelerated Retrieval Model

Consistently with the requirements of the GOSAT mission, KLIMA accuracy requirement for CO_2 total column is better than 0.3% (1 ppm out of 370 ppm) on regional scales (1000 km x 1000 km) at monthly intervals. At the same time there are requirements for program size and running time of the retrieval code that are driven by the need to integrate the KLIMA processor into the ESA Grid Processing On-Demand (G-POD) system, to take advantage of the ESA extended calculation capability, necessary for the data processing of IASI. Accordingly with the prerequisite of G-POD system, the required maximum program size is 1 Gb and the running time is that of processing 1 orbit of IASI data within 1 day when using the G-POD computing resources.

Using the Accelerate Forward Model (2.4.2) in association with the retrieval code a significant reduction of the computing time is obtained with respect to the use of the RRM. Moreover, no significant loss of accuracy is observed.

The use of the full IASI spectral band makes the code too computationally demanding, so a reduction of the selected spectra range is necessary. This version of KLIMA code that satisfies the G-POD requirements is the *Accelerated Retrieval Model* (ARM), that is chosen for integration into the ESA G-POD system.

The main elements of innovation introduced by the KLIMA ARM and by its application to the retrieval of CO_2 from IASI can be recognized in the following aspects:

- Demonstration of the feasibility of a retrieval approach exploiting all spectral channels of IASI to extract information on CO_2 and other target species with no need for channel selection or other data reduction techniques to handle the high dimensionality problem posed by IASI data processing;
- Use of a multi-target retrieval scheme which removes the systematic errors due to the interfering

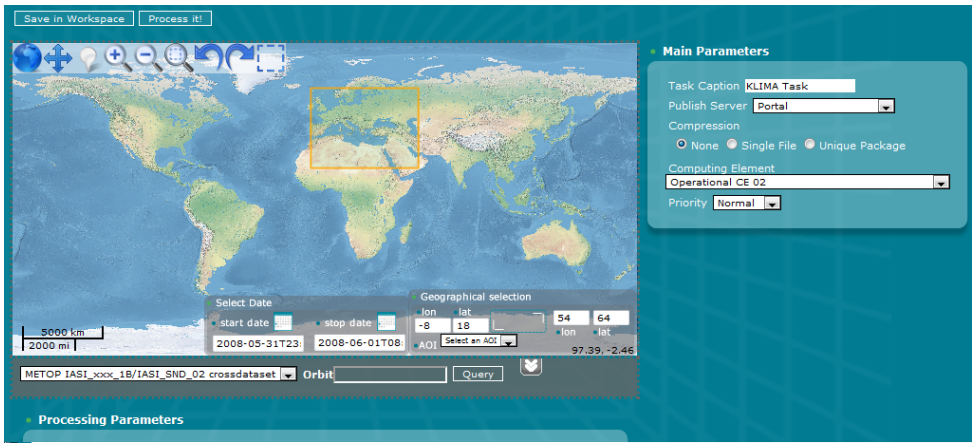

Fig. 3.2: *The ESA G-POD for Earth Observation Application system Configuration Panel*

parameters, in combination with a variance-covariance matrix of the residuals which takes into account the errors on forward model parameters.

3.3 The integration of KLIMA ARM on the G-POD system

With the purpose of massive processing of IASI L1 data, an optimized version of KLIMA code was implemented and integrated on the ESA G-POD Application System, available at ESA-ESRIN. The Earth Observation Grid Processing-on-Demand environment is a generic processing operational environment where specific data handling applications can be plugged into for processing on available Earth observation products. The G-POD user interface is a web-based mechanism for setting up and submitting processing jobs to G-POD clusters. The G-POD environment is accessible using a dedicated Web portal ([37]) allowing the user to select a region of interest, in terms of geographical coordinates and observation time, by using a graphical tool. Selected observations are grouped in a Task and sent in execution on the computer grid. At the end of the retrieval process, the output dataset containing the retrieval results and auxiliary outputs is provided.

The workspace is composed of a Graphical User Interfaces (GUI) that have been developed by the G-POD team for the management of the analysis. Through the configuration panel (Figure3.2) it is possible to select the orbit to be process, to configure the task to submit to the system and to start the analysis. In particular, the configuration panel is composed of a graphical interface dedicated to time and spatial selection, a graphical interface for task configuration, a section dedicated to execute query on orbit database and to the selection of the orbit for data processing and a text interface that allows the user to create or modify the setting files for pre-processor and retrieval.

The KLIMA software suite installed on G-POD, devoted to the analysis of the IASI observations, consists of two independent modules:

- *The KLIMA pre-processor module:* The pre-processor is the interface between the IASI operational products and the retrieval module. It is dedicated to the selection of specific observations from the data orbit files, according to a set of filtering criteria. The pre-processor requires as input: IASI L1B (or L1C) and IASI L2 data files and a settings file containing all input parameters necessary to trigger the pre-processor procedure. From the selected observations the pre-processor extracts the relevant information required by the retrieval (spectral data from L1 file, auxiliary data from L2 files) and arranges them in an input dataset for the analysis using the KLIMA retrieval module.

- *The KLIMA retrieval module:* This module includes the Forward Model and the Retrieval Model that can be independently operated to produce simulated observations or be combined for inverse processing of real or synthetic data from the IASI instrument. The retrieval module reads the input dataset arranged by the pre-processor and performs the retrieval according to the settings provided.

The retrieved quantities and the auxiliary products are exported by the module at the end of the retrieval process and saved in an output dataset.

The analysis starts with the pre-processor that reads the IASI L1 and L2 data files; extracts data related to the observations selected according to the setted filtering criteria and generates the input datasets for the retrieval. The retrieval module requires in input the observation measurements file generated by the pre-processor module. Moreover, the retrieval procedure needs several auxiliary data: the atmospheric climatological data, containing information on the vertical structure of atmospheric pressure, temperature and minor constituents concentration ([69]); the spectroscopic data (HITRAN2008); the line mixing used in CO_2 line-mixing computation ([63],[64]); the error spectra database that models the FM errors and the the irradiance of the solar source at the top of the atmosphere, used to model surface reflection ([30]). At the end of the retrieval process, several output files are produced that contain the simulated spectrum, the original spectrum as extracted from IASI L1 data and the products of the retrieval.

A pre-processing code has been developed in order to perform a screening of the IASI L1 data. The adopted selection criteria allow to exclude the IASI data when a bad quality flag is reported on L1 data (due to the validation of the L1 product) and L2 data (due to the convergence of the L2 iterative retrieval and to the validation of the L2 product) and to exclude the observation in cloudy sky conditions [47], [48]. No selection criterion has been included in IASI data pre-processing in order to exclude latitude range, as performed in [23] due to the atmospheric temperature profile interferences, because the multi-target approach adopted by KLIMA reduces the interferences effects.

Through the use of extensive G-POD computing resources that were made available for KLIMA bulk processing (described more in detail in 5.1), it was decided to process one week per month in the annual range of time selected (from March 1, 2010 to February 28, 2011). The large capacity of G-POD system made possible the acquisition from the UMARF Archive of a total data volume of $\approx$ 5 TB of IASI data.

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. Selection criteria were adopted that limited the analysis to the best measurements, i.e. 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. After the application of the selection criteria and the pre-processing of the selected dataset, and although the processing time per observation was around 2 hours, 5 TB of IASI data corresponding to 2300 orbits, covering the period from 1 March 2010 to 28 February 2011, for a total of approximately 240000 observations have been processed by G-POD during four months, that provided a statistically significant dataset.

3.3.1 Comparison between IASI Level 1B and Level 1C datasets

The original version of the KLIMA code is capable to process either Level 1B (re-sampled, non apodized calibrated spectra and corresponding images) or Level 1C (Level 1B apodized to obtain a nominal Instrument Spectral Response Function (ISRF), 4.1.3) IASI data, but the constraints of the G-POD system required a selection of the Level 1 IASI database to proceed with the bulk processing. A series of tests was conducted to verify that the retrieved XCO_2 products (for a definition of XCO_2 see 5.1) obtained by applying the KLIMA inversion model to L1B and L1C data have negligible differences and the decision was taken to focus on the use of L1C dataset, which consist of geolocated and apodised spectra.

Figure 3.3 reports an example of the results of IASI data processing carried out to test the capability of the KLIMA algorithm to retrieve XCO_2 from IASI L1B and L1C data.

The KLIMA code was used to retrieve XCO_2 in a selected area over North America (between 15°, 40° latitude and -115°, -105° longitude) on 2 December 2010. The results show the linear best fit of the dataset and the good correlation of the two sets of retrieved values, with a correlation factor of 0.92 and a fitted bias of 0.5 ppm. Consistent results were obtained from the two datasets and IASI L1C has been

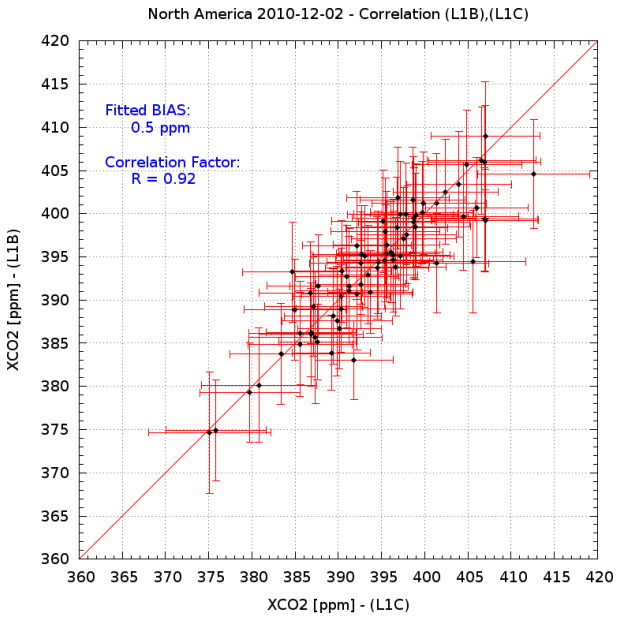

Fig. 3.3: *Scatter plot and relative errors of the two dataset of XCO_2 retrieved using KLIMA code from IASI Level 1B and Level 1C data*

selected to perform the bulk processing of IASI data using the version of KLIMA algorithm integrated on G-POD system.

3.3.2 IASI spectral band selection

The channels more sensitive to Carbon Dioxide features are located in two spectral bands: around 700 cm^{-1} in the $\nu 2$ branch and 2300 cm^{-1} in the $\nu 3$ branch respectively, as depicted in Figure 3.4. They present various sensitivities to Carbon Dioxide and other atmospheric or surface components. However, channels located near 2300 cm^{-1} are less sensitive to other atmospheric components, but the IASI radiometric noise is higher in this spectral region than in the longwave part of the spectrum. Using the multi-target retrieved strategy implemented by KLIMA it is possible to fit the entire spectrum, but the constraints imposed for inverse data processing by the integration and operation on the G-POD system made necessary a selection of the spectral bands.

For the analysis of the CO_2 three bands were considered within the IASI spectral coverage, where CO_2 features are present:

- band1 = 645-800 cm^{-1};
- band2 = 2000-2380 cm^{-1};
- band3 = 2400-2500 cm^{-1}.

The band3 contains less information with respect to band1 and band2 so it was discarded. Several tests were performed to assess the information content that is possible to obtain from a combination of band1 + band2 and from band1 only. Two geographical areas were selected over North America (between 15°, 40° latitude and -115°, -105° longitude) and Asia (between 25°, 50° latitude and 115°, 125° longitude), where a comparison was performed between XCO_2 retrieved with KLIMA algorithm (in the following referred as KLIMA L2) using band1 and using the combination band1 + band2. The results of bands comparison, performed on about 600 observations both for North America and Asia, are shown in Figure 3.5. Figure 3.6 shows, both for selected areas, the distribution of the differences between XCO_2 retrieved using only band1 and using band1+band2. A bias of about -1.5 ppm has been found when only band1 is used in the retrieval; the width of the distribution (about 6 ppm) is in agreement with the XCO_2 retrieval error reported in Table 3.3.

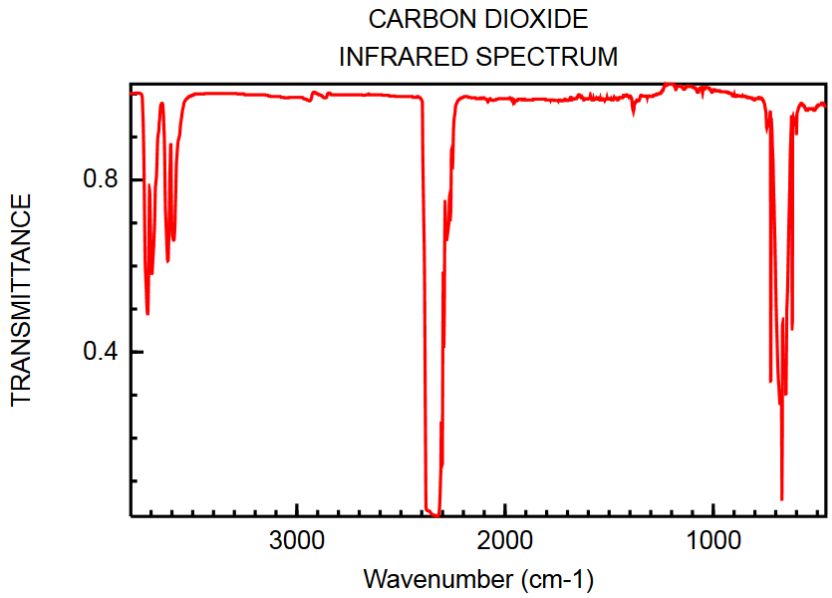

Fig. 3.4: *Carbon Dioxide InfraRed spectrum (Coblentz Society [19])*

The comparison shows that retrievals performed using band1 and the combination of band1 + band2 give comparable results, even if the use of band1 increases the XCO_2 error, as reported in Table 3.3.

Tab. 3.3: *XCO_2 errors using using band1 and the combination of band1 + band2*

XCO_2 errors [ppm]	North America	Asia
band1	5.4	6.7
band1 + band2	3.9	5.2

The analysis was limited to the spectral range 645-800 cm^{-1}, which contains the ν_2 band of CO_2 and provides the best compromise between high sensitivity to Carbon Dioxide and low sensitivity to other atmospheric variables. The smaller information content due to the reduction of the analysed IASI band increases the retrieval error obtained from ARM of about 30 % (from 5.2 ppm to 6.7 ppm on Asia and from 3.9 ppm to 5.4 ppm on North America), but the scatter plots shown in Figure 3.5 give consistent results using a computation time reduced of about 60-70 %. On the basis of these tests, taking into account the reduction of the computation time and the increase of the retrieval error (always within the accuracy requirement needed for the comparison tasks) band1 has been selected for the current analysis to perform the retrieval of Carbon Dioxide and the comparison with TANSO-FTS products.

3.4 Outline of the Chapter

In summary, it is possible to conclude that the Reference Retrieval Model achieves the best performances for the retrieval of Carbon Dioxide columns, with total retrieval error on CO_2 total column of $\approx$ 1.0% and no bias. The RRM, however, does not meet the requirements in terms of program size and running time for integration on G-POD. In order to achieve a reasonable trade-off between retrieval quality and efficiency, it is necessary to use the ARM version of the KLIMA algorithm, which performances, in terms of total retrieval uncertainty, are only slightly degraded respect to the RRM and are fully consistent with the requirement of 0.3% accuracy of CO_2 total column at monthly intervals on regional scales for GOSAT validation.

The use of KLIMA ARM introduces a small bias in the retrieved CO_2 total column when using the full IASI spectral range. This can be however corrected by including the errors due to FM approximations in the VCM of the observations.

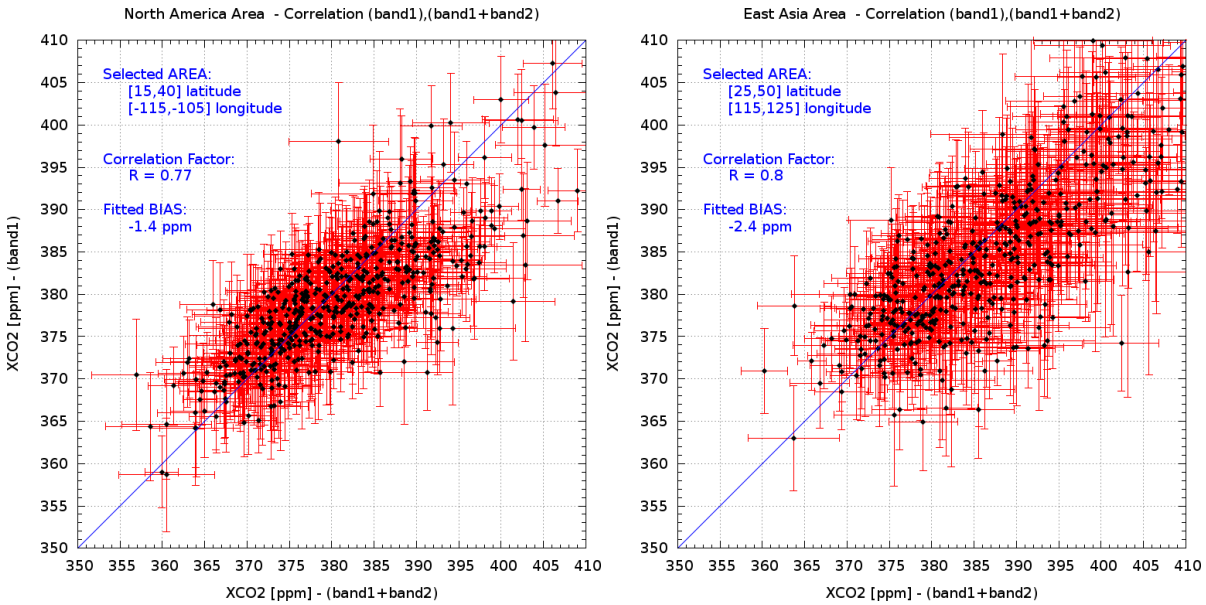

Fig. 3.5: *Scatter plot (and errors) of the XCO_2 retrieved from band1 and band1+band2 over North America (on left panel) and Asia (on the right panel)*

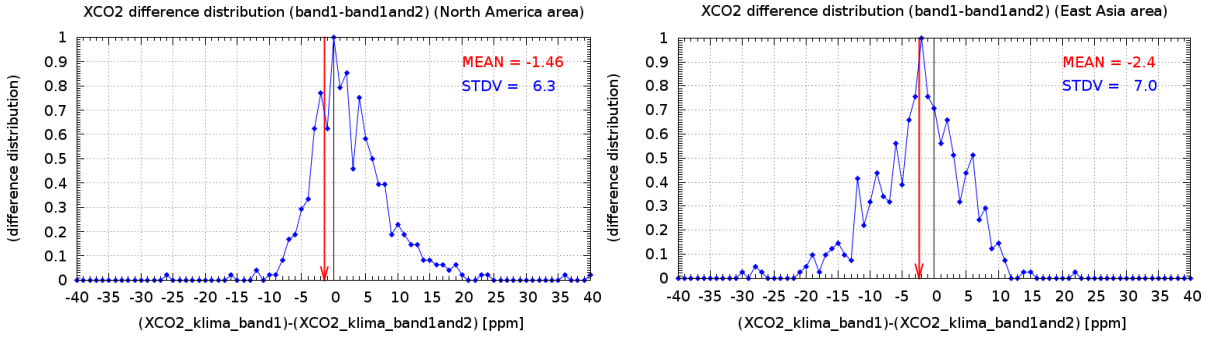

Fig. 3.6: *Distribution of the differences between XCO_2 retrieved using only band1 and using band1+band2 over North America (on left panel) and Asia (on the right panel)*

The original version of the KLIMA code is capable to process both Level 1B (L1B) and Level 1C (L1C) IASI data. A first processing of IASI data was carried out to verify that the retrieval products obtained by applying the KLIMA inversion model to L1B and L1C data have negligible differences 3.3.1. Consistent results were obtained from the two datasets, and IASI L1C, which consist of geolocated and apodised spectra, has been selected to perform the bulk processing of IASI data using the version of KLIMA algorithm integrated on G-POD system.

Moreover, to meet the requirements for the integration on G-POD system, a reduction of the IASI spectral band has been performed. So, the analysis was limited to the spectral range 645-800 cm^{-1}, which provides the best compromise between high sensitivity to Carbon Dioxide, low sensitivity to other atmospheric variables, and the reduction of the computation time of about 60-70 % 3.3.2.

Chapter 4

An overview of the observations: instruments and retrieval codes

In this chapter, we report a brief overview of the sensors examined in the study. We describe the instrument characterization and the different retrieval algorithms applied to the measurement data to obtain information on atmospheric columnar amount of CO_2 from these instruments:

- **InfraRed Atmospheric Sounding Interferometer (IASI)** on-board the MetOp-A platform;
- **Thermal And Near InfraRed Sensor for Carbon Observation-Fourier Transform Spectrometer (TANSO-FTS)** on-board the GOSAT satellite;
- **Total Carbon Column Observing Network (TCCON)** ground stations.

We also summarize the available options for the access to IASI, TANSO-FTS and TCCON operational data and describe the solutions adopted in order to obtain a statistically significant dataset on which to perform the analysis.

4.1 The IASI/MetOp-A mission

4.1.1 IASI instrument description

The InfraRed Atmospheric Sounding Interferometer (IASI)developed by the Centre National dEtudes Spatiales (CNES) in collaboration with the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) consists of a high resolution nadir-viewing Fourier transform spectrometer, associated with an imaging instrument; it is a key payload element of the MetOp series of European meteorological polar-orbit satellites. IASI is designed to measure the spectrum emitted by the Earth's surface and atmosphere in the Thermal InfraRed (TIR) region, with full coverage of the spectral range from 645 cm^{-1} to 2760 cm^{-1} (from 15.5 μm to 3.62 μm). This spectral range was chosen to include the strong absorption features from the CO_2 absorption band around 15 μm, the Ozone band around 9.6 μm, the strong Water Vapour absorption band, and up to the edge of TIR for one of the Methane absorption band.

In the TIR spectral range atmospheric sounding is based on the analysis of atmospheric radiance spectra. The instrument records the radiance signal as a composition of the initial long wave IR emission (the Earth black body emission at the location of the measurement), modified for cloud free scenes by gaseous absorption, and emission processes along the path. Each molecule has a specific signature resulting from rotation-vibration transitions.

Each spectrum is sampled every 0.25 cm^{-1}, with an apodized spectral resolution of 0.5 cm^{-1} (Level 1C spectra), providing a total of 8461 radiance channels [4]. Then concentrations can be retrieved for each molecule absorbing in the atmospheric path using iterative inversion methods.

The sensor is flying on board the Meteorological Operational Satellite (MetOp-A) platform, launched in 19 October 2006 [15], [16], that is part of a series of three satellites, to be launched sequentially every 5 years in order to provide observations for more than 15 years. The second instrument, on-board on the MetOp-B satellite, was launched in September 2012. The third instrument will be mounted on the MetOp-C satellite with launch scheduled for October-November 2016. The satellite flies in a sun-synchronous polar orbit at an altitude of around 817 km and crosses the Equator at two fixed local solar times: 09:30 a.m. (descending) and 09:30 p.m. (ascending). The time to complete one orbit is about 101 min with a total of 14 orbits a day and a repeat cycle of 29 days.

The instrument performs measurements using nadir geometry, scanning the Earth's surface perpendicularly to the satellite's flight track with 15 individual views on each side of the track. Each Instantaneous Field Of View (IFOV) has a size of 50 km x 50 km and consists of 2 x 2 circular pixels; a single pixel has a diameter of 12 km at the nadir point. The maximum scan angle is 48.3° from nadir with a total swath width of 2200 km. A swath with a total of 120 views is achieved in 8 s, allowing global coverage twice a day. Figure 4.1 reports the IASI observation mode.

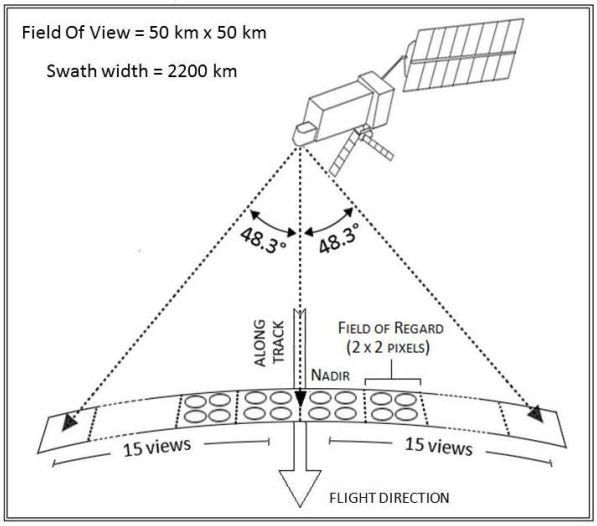

Fig. 4.1: *IASI observing mode: the instrument measures the spectrum of IR radiation emitted by the Earth and the atmosphere. Each Instantaneous Field Of View has a size of 50 km x 50 km and consists of 4 circular pixels of 12 km in diameter*

4.1.2 IASI operational data products

The contribution of IASI lies mainly in weather forecasting, atmosphere composition and climate monitoring, with the main purpose of providing temperature and humidity measurements for weather forecasting. IASI is used to determine ocean and land surface temperatures, temperature and water vapour profiles in the troposphere and lower stratosphere, the fractional cloud cover and cloud top temperature and pressure. IASI also plays a role in monitoring the composition of the atmosphere, providing, as operational products, the total amount of Ozone, CO, CH_4 and N_2O.

The IASI operational products are distributed via EUMETSATs Data Distribution System (EumetCast), the Broadcast System for Environmental Data of EUMETSAT. The CO_2 total columns can be download from the EUMETSAT Unified Meteorological Archive and Retrieval Facility (UMARF) together with the other operational products, even if Carbon Dioxide is provided without an associated retrieval error and its status is still experimental.

The following products are foreseen by the IASI distribution chain:

- Level 0 - Raw IASI measurement data, after on-ground de-multiplexing, include pre-calibrated spectra, the corresponding non calibrated images, calibration images, verification data and auxiliary data necessary for further processing.

- Level 1 A (L1A) - Non apodized calibrated spectra and corresponding images; this step of processing comprises data decoding, radiometric post-calibration, spectral calibration, IASI/AVHRR co-registration via IASI images, geo-location and dating.
- Level 1 B (L1B) - Level 1 A data re-sampled.
- Level 1 C (L1C) - Level 1 B apodized to obtain a nominal Instrument Spectral Response Function (ISRF).
- Level 2 (L2) - Geophysical products derived after co-processing of data from IASI and the MetOp meteorological instruments (profiles of temperature, humidity, surface temperature, trace gas distribution, cloud parameters and other auxiliary data).

The UMARF archive can be accessed through a web interface ([87]), which permits on-line ordering and delivery of the operational products in a variety of formats. In particular, the EPS (EUMETSAT Polar System format) Native format, that is a binary data format organized in sections containing header section providing general data, global data and measurement data (L1A, L1B, L1C) or retrieved data (L2) was selected and implemented in the KLIMA pre-processor (3.3).

The processing of IASI measurements using KLIMA retrieval code requires an input dataset that includes both L1B and L1C data (spectral radiance data and auxiliary data) and L2 operational products (including atmospheric and surface data, total columns of CO_2 and other species).

4.1.3 IASI operational data processing

IASI operational data processing is reported in [1]. The IASI Level 2 processing development targeted the generation of temperature and humidity profile information, the associated surface information and the retrieval of some trace gas species: CO, O_3 , CH_4 , N_2O and CO_2. A large number of scientific studies both internal and external to the IASI Sounding Science Working Group(ISSWG), supported the evolution of the science plan and the processor making in the frame of the EUMETSAT Polar System (EPS) Core Ground Segment development, according to specifications given by EUMETSAT.

The L2 IASI Product Processing Facility (PPF) has a modular structure: a pre-processing step, the cloud detection and cloud-parameters determination, and the geophysical-parameters retrieval.

The pre-processing starts with the acceptance and validation of data. All input data will be checked against valid bounds and invalid or missing data will be flagged. All the flags are then included in L2 product. The cloud processing includes the cloud detection and the determination of cloud parameters. Several method of cloud detection are used due to the absence of single cloud detection method that is able to detect clouds properly in all situations. After identification of a cloud in the IASI IFOV, the cloud-top pressure, the cloud amount and the cloud phase are determined. The number of variables retrieved from the measurements depends on the cloud conditions, which also determine the set-up of the retrieval scheme. Depending on configuration, data availability, surface and cloud conditions different retrieval types are foreseen. The retrieval steps follow with a series of statistical methods to estimate the temperature and humidity profiles, the surface emissivity, the sea and land surface temperature and some trace gas species. A physical retrieval using the optimal estimation method (OEM) is subsequently performed to refine the temperature and Ozone profiles. The first retrieval consists of statistical methods, based on a linear empirical orthogonal function (EOF) regression or a non-linear artificial neural network (ANN), based on the multilayer perceptron neural network with two or three hidden layers. The entire IASI spectrum is used in the linear regression retrieval to determine temperature, Water Vapour profile, Ozone, surface temperature, and surface emissivity. The retrieval of trace gas columns is done using an ANN with a set of selected IASI radiances, sensitive to the particular trace gases, together with the previously derived temperature profile. The final retrieval is a simultaneous iterative retrieval, seeking the maximum a-posterior probability solution for the minimisation of a cost function seeking by the Marquardt-Levenberg method ([72]). The cost function aims at minimising the difference between measured and simulated radiance vectors. Trace gas columnar amounts are retrieved with an accuracy of

5 to 6%. The retrieval of the full state vector is attempted for cloud-free pixels. In case of a partial cloud contamination, the geophysical retrieval is achieved with EOF regression and limited to temperature and humidity.

The synthetic clear-sky radiances are computed with the fast radiative transfer model RTIASI ([58]), modified in RTTOV-10 after the implementation of a new version on October 2011. It is a semi-empirical model that supplies radiances and analytical Jacobians at the top of the atmosphere. Radiation included comes from the up-welling atmospheric emission, the down-welling atmospheric emission reflected at the surface and transmitted to the satellite, the surface emission transmitted to the top of the atmosphere, and the solar radiation transmitted to the surface and reflected to space.

The version 5 (V5) of the processor was released on the operational processing chain on 14 September 2010 ([91]). The release of new processor version could explain the discontinuity, relative to October, that we can observe in the seasonal variation of IASI L2 data, reported in Figure 6.15.

The N_2O , CH_4 , CO_2 products are not operational. These retrievals are still subject to research and development. They benefited from the same algorithm improvements introduced in IASI L2 PPF V5 and are currently produced in an experimental mode.

4.2 The TANSO-FTS/GOSAT mission

4.2.1 TANSO-FTS/GOSAT instrument description

The Greenhouse gases Observing SATellite (GOSAT) (which original Japanese name is IBUKI, that means "breath") is the first satellite mission designed expressively for the study of the greenhouse gases from space. Primary mission goals are the observation of the Short Wave InfraRed (SWIR) and the retrieval of CO_2 and CH_4 column density. The concentrations of GHGs are retrieved only from measurements during the orbital day time. Secondary mission goals are the observations in the TIR with the retrieval of CO_2 and CH_4 altitude profile and column density, possible during orbital night time, other trace gases (O_3 etc.) and other products, such as temperature profile and Earth radiation.

GOSAT is a joint project of Ministry of the Environment (MOE), JAXA (Japan Aerospace Exploration Agency) and NIES (National Institute of Environmental Studies). JAXA is responsible for the development of the instruments, the launch and operations of the spacecraft while NIES develops algorithms for data analysis.

The satellite was launched on January 23, 2009 to the sun-synchronous sub-recurrent orbit (local sun time at crossing the descending node at 13:00) having an approximate altitude of 666 km with an inclination of 98.06° to the equator. The orbital period of GOSAT is approximately 98.1 minutes, corresponding to 14.6 revolutions per day and a revisiting time of 3 days.

GOSAT is equipped with two instruments: the Thermal And Near InfraRed Sensor for Carbon Observation Fourier Transform Spectrometer (TANSO-FTS) and the Cloud and Aerosol Imager (TANSO-CAI) [45]. TANSO-FTS has three narrow bands in the SWIR region (band 1 at 0.76 μm, band 2 at 1.6 μm and band 3 at 2.0 μm) and a wide TIR band (band 4 from 5.5 μm to 14.3 μm) with a spectral resolution of about 0.2 cm^{-1}. The sensor IFOV is a nadir circular footprint of about 10.5 km in diameter. The maximum scan angle is 35° from nadir with a total swath width of 790 km.

TANSO-FTS observes solar light reflected from the Earth surface only during the daytime, as well as the thermal radiance emitted from the atmosphere and the surface during both the daytime and nighttime. The sensor observes scattered sunlight over land using a nadir-viewing observation mode and over ocean using a sun glint observation mode. The SWIR region is observed in bands 1 to 3 in the daytime only and TIR region is captured in band 4 during both the day and the night. The spectral bands of TANSO-FTS and the associated target products are summarized in Table 4.1.

The TANSO-FTS viewing geometry is illustrated in Figure 4.2.

In the normal operations, one interferogram is acquired every four seconds. Orientating mechanism about the along track axis and cross track axis allow the sensor to scan in the cross direction, to point to the sunlight spot and/or predetermined locations, and to keep pointing a place during one whole

Tab. 4.1: *TANSO-FTS Spectral bands*

Spectral band n°	1	2	3	4
Spectral range	VIS	SWIR	SWIR	MWIR/TIR
Coverage (μm)	0.75-0.78	1.56-1.72	1.92-2.08	5.5-14.3
Target of the measurement	O_2, air pressure, cirrus	CO_2, CH_4, H_2O	CO_2, CH_4, H_2O cirrus	CO_2, CH_4, H_2O
Calibration	Solar irradiance, deep space, moon, diode laser			Black body, deep space
Spectral resolution	0.2 cm^{-1} (maximum optical path difference 2.5 cm)			

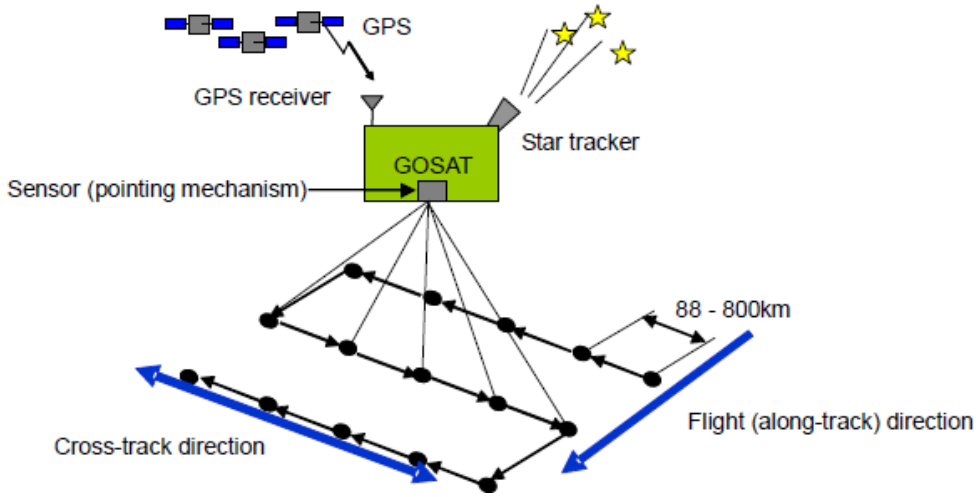

Fig. 4.2: *TANSO-FTS Viewing Geometry ([44])*

interferogram capturing. In routine observation, sensor observers a predetermined number of points scanning 1, 3, 5, 7 or 9 points along the cross track direction and the scan is repeated at an interval along the along track direction, so that the observed points form a grid-like pattern on the Earth's surface. The instrument keeps pointing in the same place during the whole capture of an interferogram. In order to improve Signal to Noise Ratio (SNR), TANSO-FTS may observes a same place repeatedly to sum the observed data up (3 times of repetition, or only once with no repetition).

The SWIR TANSO-FTS L2 data contain the retrieval results of the total columnar amount of CO_2 and the CO_2 in dry air mixing ratio (indicated as XCO_2 , definition of XCO_2 is reported in 5.1). In the L2 data also the error budget (in terms of retrieval, smoothing, interferences and external errors) are provided. The retrieval method (described by Yoshida et al. [89]) is based on the Optimal Estimation approach and on a pre-data screening suitable for the retrieval analyses and a post-quality checking of the retrieved results.

4.2.2 TANSO-FTS operational data products

Several types of GOSAT data products are provided for the users. Data users can search and order L1 data (FTS L1B, CAI L1B, and CAI L1B+ data) and the higher level data products (FTS L2, CAI L2, FTS Level 3, CAI Level 3, Level 4A and Level 4B data products) by accessing the GOSAT User Interface Gateway (GUIG) at the address [38].

The following products are available by the GOSAT spectral data processing and distribution chain:

- FTS Level 1B data: radiance spectra that are obtained by performing the Fourier transformation on the signals detected by FTS.
- FTS SWIR Level 2 data: data products store column abundances of CO_2 and CH_4 retrieved from the radiance spectra in the bands 1 through 3 of FTS. When ordering the L2 data products at GUIG, users can specify the observation areas and periods of their interests.

- FTS TIR Level 2 data: data products are vertical concentration profiles of CO_2 and CH_4 derived from the radiance spectra in band 4 of FTS.

For the goal of performing an inter-comparison of the CO_2 total column retrieved from KLIMA inversion code and TANSO-FTS L2, I downloaded from the GUIG the following data products:

- FTS SWIR L2 V1.20
- FTS SWIR L2 V1.30
- FTS SWIR L2 V1.40

The different versions of the L2 products correspond to different versions of the input TANSO-FTS L1B product (measured spectrum) [89]. The first analysis and comparison was performed using the Version 1 of FTS SWIR data. But, after this first step, a new version of FTS SWIR data (Version 2 (V2)) was released, which have much better quality compared to the initial V1 dataset ([90]). So it was decide to download also the latest version and repeat the analysis with the new dataset:

- FTS SWIR L2 V2.00 (March 2010 - August 2010)
- FTS SWIR L2 V2.10 (September 2010 - December 2010)
- FTS SWIR L2 V2.11 (January 2011 - February 2011)

Access to different releases of this version of TANSO-FTS L2 data (V2.00, V2.10 and V2.11) was necessary, as from the list of FTS SWIR datasets displayed above, in order to cover the entire period of one year identified for the comparison with IASI measurements. The results of the analysis performed using the original V1 dataset are summarized in Appendix A in order to provide a complete view of the activities.

Measurements obtained from IASI and TANSO-FTS SWIR channels present different sensitivities, which can be described by the Averaging Kernel matrix (AK). For a perfect column measurement, the column AK would be 1.0 at all altitudes, but in practice there is a greater sensitivity to some altitudes than others. IASI AKs present a maximum close to 200 hPa and generally reach low values above 900 hPa and below 100 hPa, while the TANSO-FTS AKs are quite constant on the same altitude range, so IASI is more sensitive to mid to upper troposphere while TANSO-FTS is able to detect signal from lower tropospheric layers. SWIR observations are more sensitive to CO_2 near the surface, but less sensitive to the vertical profiles than TIR observations. A synergy use of these bands is expected to complement one another.

For an ideal intercomparison the best condition would have been the use of TANSO-FTS band 4 in the Thermal InfraRed region. In this way both sensors would look at the same spectral region and the information would arrive from the same atmospheric layers. Moreover, with TIR band it is possible to perform measurements both day and night and to collect more data over oceans.

Currently, the column-averaged dry air mole fraction of atmospheric Carbon Dioxide derived from the TANSO-FTS SWIR channels has been largely validated, against ground-based TCCON observation and models, as reported for instance in [61] or [20]. On the contrary, I have not found publications on measurement campaigns carried out to validate the profiles of CO_2 obtained from the TIR band or comparison using ground-based stations. On the dedicated page GOSAT related references ([36]) are present few publications on preliminary retrieval of CO_2 profiles within the upper troposphere. As reported for example in [50], preliminary results show that some systematic biases exist in the TIR spectrum because of the opacity of the mirrors of SWIR bands that obstruct the field of view of band 4 in the TIR. Saitoh at al. [78] is performing a new reprocessing of band 4 data, but data available are still limited and the reprocessing is an ongoing activity. So, due to the non-global coverage in the range of time selected and the fact that the L2 products of the TIR band have not yet been validated, it was decided to focus mainly on the total columnar content obtained from SWIR bands.

Despite the different sensitivity in SWIR and TIR region, it is possible to improve the intercomparison by adopting the strategy described in 6.2, where we take into account the different a priori and AK smoothing contributions.

4.2.3 TANSO-FTS operational data processing

The GOSAT measurements are routinely processed by the GOSAT Data Handling Facility and the operational products can be collected by general users through the GOSAT data product distribution website GOSAT User Interface Gateway (GUIG) ([38]). FTS and CAI data, that the satellite has collected, are received and processed at JAXA Tsukuba Space Center first. Then these data are transferred to GOSAT Data Handling Facility (GOSAT DHF). GOSAT DHF gathers reference data, such as meteorological data necessary for the higher-level data processing, from cooperating institutions on a regular basis. Using these reference data, the observational data from JAXA are processed into column abundances of CO_2 and CH_4, CO_2 sources and sinks, and CO_2 three-dimensional distributions. The data processing is performed in conjunction with other external computing resources. Reference data used for validating the data products are also stored in this facility.

The GOSAT data validation team used high-precision reference data obtained with ground-based and airborne measurements. For validating the L2 column abundances of CO_2 and CH_4, the team uses data from ground-based high-resolution Fourier transform spectrometers and in situ observation instruments installed on aircrafts. The data validation team has carried out a series of data validation activities and compared the L2 data products to the ground-based and airborne reference data. Data obtained in Japan, Europe, Oceania, and North America were used in these validation activities. Also, data collected by the aircrafts of Japan Airlines that participate in CONTRAIL (Comprehensive Observation Network for Trace gases by Airliner) project and the US National Oceanic and Atmospheric Administration's airborne measurement program were employed. The first results of the comparison indicated that the TANSO-FTS SWIR XCO_2 (V1) are biased low by $2.3 \pm 1.2\%$ than the reference values [61]. The latitudinal pattern of the L2 zonal means, however, were broadly consistent with those of the reference values.

4.3 The Total Carbon Column Observing Network (TCCON)

4.3.1 TCCON instrument description

The Total Carbon Column Observing Network (TCCON) is a network of ground-based Fourier transform spectrometers recording direct solar spectra in the near-InfraRed spectral region. From these spectra the column-averaged abundance of several gas species, like CO_2, CH_4, N_2O, HF, CO, H_2O and HDO are retrieved. In particular, CO_2 total amount is retrieved with a precision of 0.25% (less than 1 ppm) for global monthly means ([94]), making it an ideal dataset for validation of retrieved XCO_2 from satellite. The TCCON was established in 2004 with a primary focus of measuring precise and accurate columns of CO_2. Currently, there are 25 sites affiliated with TCCON, 22 of which are operational ([84]). The sites and their locations are shown in Figure 4.3.

TCCON sites span a wide range of atmospheric states and observing conditions, from the Tropics to the Polar regions, continental and maritime, polluted and clean. But there are notable gaps in coverage, particularly in South America, Africa and Asia. To use TCCON for providing a transfer standard between the in situ network and the satellite retrievals of total column CO_2 TCCON retrievals must be placed on the same absolute calibration scale as the in situ network (the World Meteorological Organization (WMO) scale). To address the absolute calibration it was used aircraft profiling above the FTS stations, as reported in [94], [59], [34].

TCCON database contains the retrieval results of the CO_2 in dry air mixing ratio (XCO_2) along with its associated error.

Fig. 4.3: *TCCON site locations. Red circles represent operational sites while blue squared represent the planned future sites*

4.3.2 TCCON operational data products

TCCON data can be downloaded from the TCCON data archive ([85]), maintained by California Institute of Technology, which is also supported by NASA. Through the web page users can select the site, set up the options of interest and order files.

A defined dataset of gases (CO_2, CH_4, N_2O, HF, CO, H_2O, O_2 and HDO) and supporting data, sufficient to allow the calculation of the tropospheric average mixing ratio (e.g. AKs, a priori profiles), are archived in a central TCCON database.

Archived data includes flags to identify which data are deemed of sufficient quality for specific purposes. Data, which pass all defined Quality Control (QC) tests, i.e. data which remain unflagged (flag = 0) at the end of the QC procedure, are deemed suitable for public release. Quality control criteria may evolve in time as our characterisation of measurement errors improves. The QC criteria applied are documented for each public release.

4.3.3 TCCON operational data processing

Individual TCCON sites are responsible for processing their collected interferograms in a standardized procedure to ensure consistency across the network. To this end, a common open-source software package is used network-wide for data processing and analysis. The software package processes the raw interferograms, corrects them for solar intensity variations and phase errors, and computes spectra using a fast Fourier transform algorithm. Subsequently, spectra are fitted with a non-linear least-squares spectral fitting program (GFIT), designed for the analysis of FTS absorption spectra, to retrieve total column amounts and derived output data. GFIT is based on the line-by-line calculations and assumes a Voigt line shape. Thermal emission is negligible in the NIR and scattering is negligible in a direct-sun viewing geometry. GFIT does not do a profile retrieval (i.e. the VMRs at different altitudes are not scaled independently to obtain the best fit to the spectrum). The used spectroscopic database is HITRAN2008 ([95]). The total columns are then corrected for systematic errors and calibrated using comparisons with profiles obtained from aircraft.

Chapter 5

Analisys of KLIMA-IASI retrieved CO_2 total column

5.1 KLIMA-IASI Level 2 processing

After obtained the IASI L1 database from UMARF and with the computing resources of G-POD system, the bulk processing started. In the selected period of one year, from March 1, 2010 to February 28, 2011, one week per month of IASI L1 data were processed on a global geographical coverage, retrieving the CO_2 both on land and on water and both during day and night. Only clear sky IASI observations have been analysed. Moreover, IASI spectrum was analysed between 645 and 800 cm^{-1}, as described in 3.3.2.

On water, because of the moderate CO_2 variability ([62]), a very reduced azimuth coverage has been taken into account (by analysing only one of the pixels closed to nadir in the IASI scan), while on land a week-dependent number of pixels has been imposed and some off-axis azimuth scans are considered, achieving the total amount of 20000 analysed spectra per month.

A screening of the IASI L1 data was performed using a pre-processing filtering. The adopted selection criteria allow to exclude the IASI data when a bad quality flag is reported on L1 data (due to the validation the L1 product) and L2 data (due to the convergence of the L2 iterative retrieval and to the validation the L2 product) and to exclude the observation in cloudy sky conditions. The multi-target approach adopted by KLIMA reduces the interferences effects, so we do not need to include any geographical selection criteria in IASI data pre-processing, as performed for example in the analysis carried out by Crevoisier et al. ([23]), where only data in the Tropic region have been considered, due to exclude the atmospheric temperature profile interferences.

The KLIMA processing campaign was performed on the ESA G-POD facilities in four months, starting from June 2012 until the end of September 2012. A large amount of data was processed, as summarised by the following Table 5.1.

Tab. 5.1: *IASI Data processed on ESA G-POD Earth Observation System*

Processing Time	4 Months
Annual number of orbits	$\approx$ 5110 Orbits
Processed Orbits	$\approx$ 2300 Orbits
Processed Data Volume	$\approx$ 5 TB
Retrieved L2 products	$\approx$ 240000 Observations

After the bulk processing, the retrieved value of the total column amount of CO_2 is used to compute the XCO_2 (column average CO_2 dry air mole fraction) from the total column amount of H_2O and the atmospheric pressure at the surface imposing the hydrostatic equilibrium:

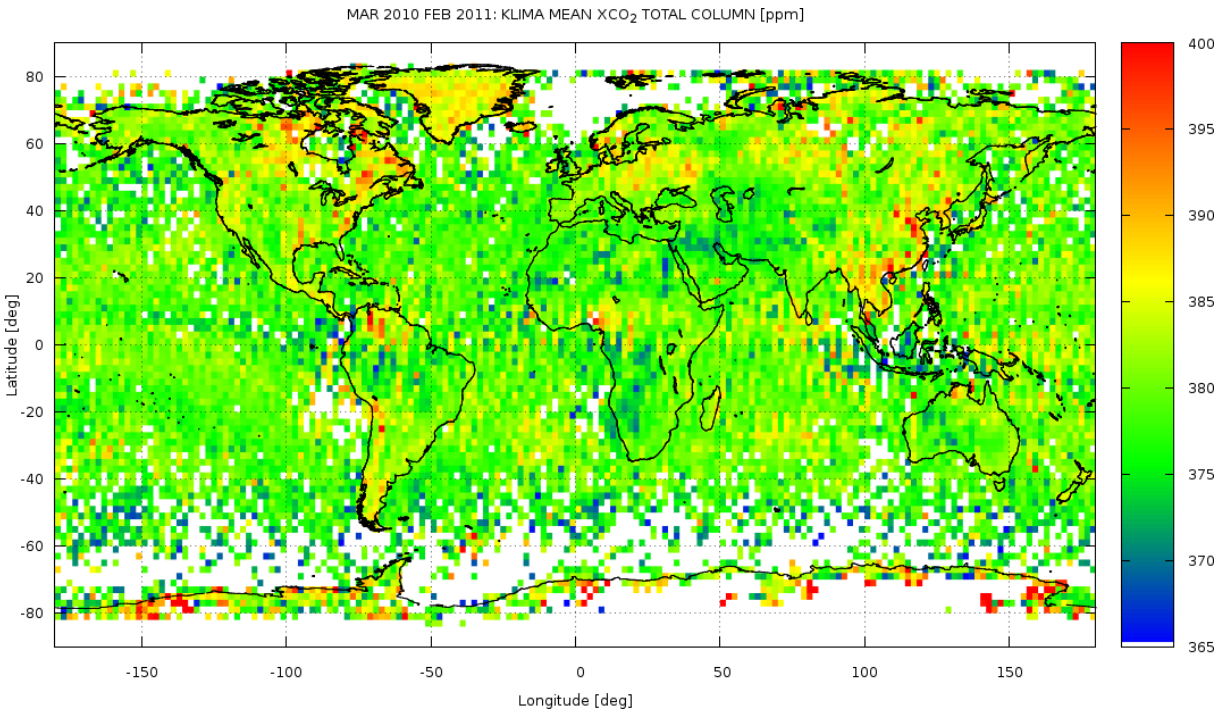

Fig. 5.1: *March 2010 - February 2011: KLIMA L2 XCO_2 total column [ppm] over a grid of 2° x 2°, averaged for one year*

$$XCO_2 = \frac{Column_{CO_2}}{(Column_{Air} - Column_{H_2O})} * 10^6 \tag{5.1}$$

where the $Column_{Air}$ is computed using the L2 IASI value of the surface pressure.

The mole fraction is defined as the number of molecules of CO_2 in any given air parcel divided by the total number of all molecules (except water) in that parcel. For CO_2 it is usually expressed as parts per million. The use of the mole fraction of CO_2 in (dry) air is convenient because it is a conserved quantity: it does not depend on the pressure, temperature, Water Vapour or condensed water content, which are all highly variable. The CO_2 mole fraction is much less variable, and depends only on the history of its sources and sinks, almost all of which are caused by surface processes such as photosynthesis and respiration, and the burning of coal, oil, or natural gas. In practice, this division removes variations due to surface pressure changes, making results from different days or sites more directly comparable.

Equation 5.1 is also used to obtain the XCO_2 value from the L2 IASI EUMETSAT results.

5.2 KLIMA Level 2 products overview

Through the use of G-POD computing resources, one year of IASI L1C spectra was analysed using the optimized version of KLIMA algorithm integrated in the G-POD system. The data selection was performed according to the criteria discussed in Section 3.3. An example of the IASI L2 product provided by KLIMA code is reported in Figure 5.1: the map shows the XCO_2 values [ppm] averaged for one week per month from March 2010 to February 2011 for a global geographical coverage over a grid of 2° x 2°.

With the use of TIR IASI spectral channels, measurements are available both during day and night, even if no significant diurnal variability has been noticed in the dataset observed so far, as it is possible to see in Figures 5.2, that show the KLIMA L2 XCO_2 total column, averaged for one year, for day time case only, night time case only and the difference between the day time global distribution and the night time global distribution. All the maps are plotted over a grid of 2° x 2°. It is possible to see that the difference between day time values of XCO_2 concentration and night time values are around 3-4%, except for some inland areas (over South America, East cost of North America, central Africa and Himalayas). On

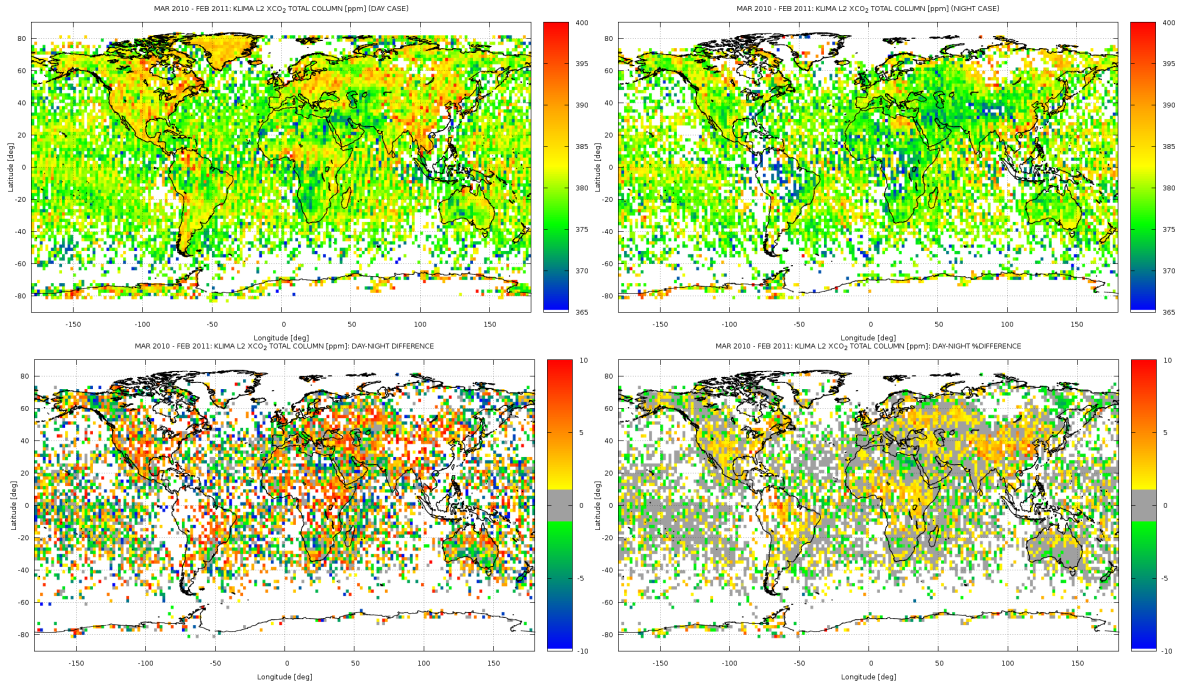

Fig. 5.2: *March 2010 - February 2011: KLIMA L2 XCO_2 total column [ppm] over a grid of 2° x 2°, averaged for one year: daytime case, night time case, difference day-night, % difference day-night*

average, the differences between day and night are smaller over oceans than over land, this reflects the moderate variability of CO_2 over water. Moreover, this suggests that there is not enough sensitivity to catch the CO_2 diurnal cycle. Accordingly, no distinction is made in the plotted results between day and night observations.

In the next pages Panels from Figure 5.3 to Figure 5.14 represent, for each month of the selected year, maps of the global geographical coverage, as a function of latitude and longitude, reported over a grid of 2° x 2°, of the following quantities:

- KLIMA L2 mean XCO_2 total columns [ppm];
- XCO_2 mean retrieval errors [ppm];
- retrieval errors associated with the XCO_2 mean value [ppm] (i.e., the mean error divided the square root on the number of measurements on every grid pixel);
- ratio between the standard deviation of the retrieved values and the XCO_2 mean retrieval error;
- number of observations per pixel.

Mean XCO_2 total columns are calculated for measurements obtained over land and over water for both daytime and night time. For the calculation of these mean errors it is considered the arithmetic average, because the retrieval errors associated with individual products do not present a large monthly variability within the single pixel.

The mean XCO_2 total columns retrieved by KLIMA (top map from Figure 5.3 to Figure 5.14) show a good geographical coverage, data distribution spans from the Polar regions through the Tropics. It is possible to see from the global distribution maps that the longitudinal variations in XCO_2 over the oceans are smaller than those over land. Higher concentration values over oceans are found from October to February, in both Hemispheres. High values of XCO_2 concentration are observed over South America and Central Africa (in winter months), possibly due to the occurrence of atmospheric dust or smoke-like phenomena likely correlated with biomass burning (as reported, for example, in [12]), even if the cores of these events are filtered out by our selection criteria of clear-sky pixels. High values of XCO_2 are

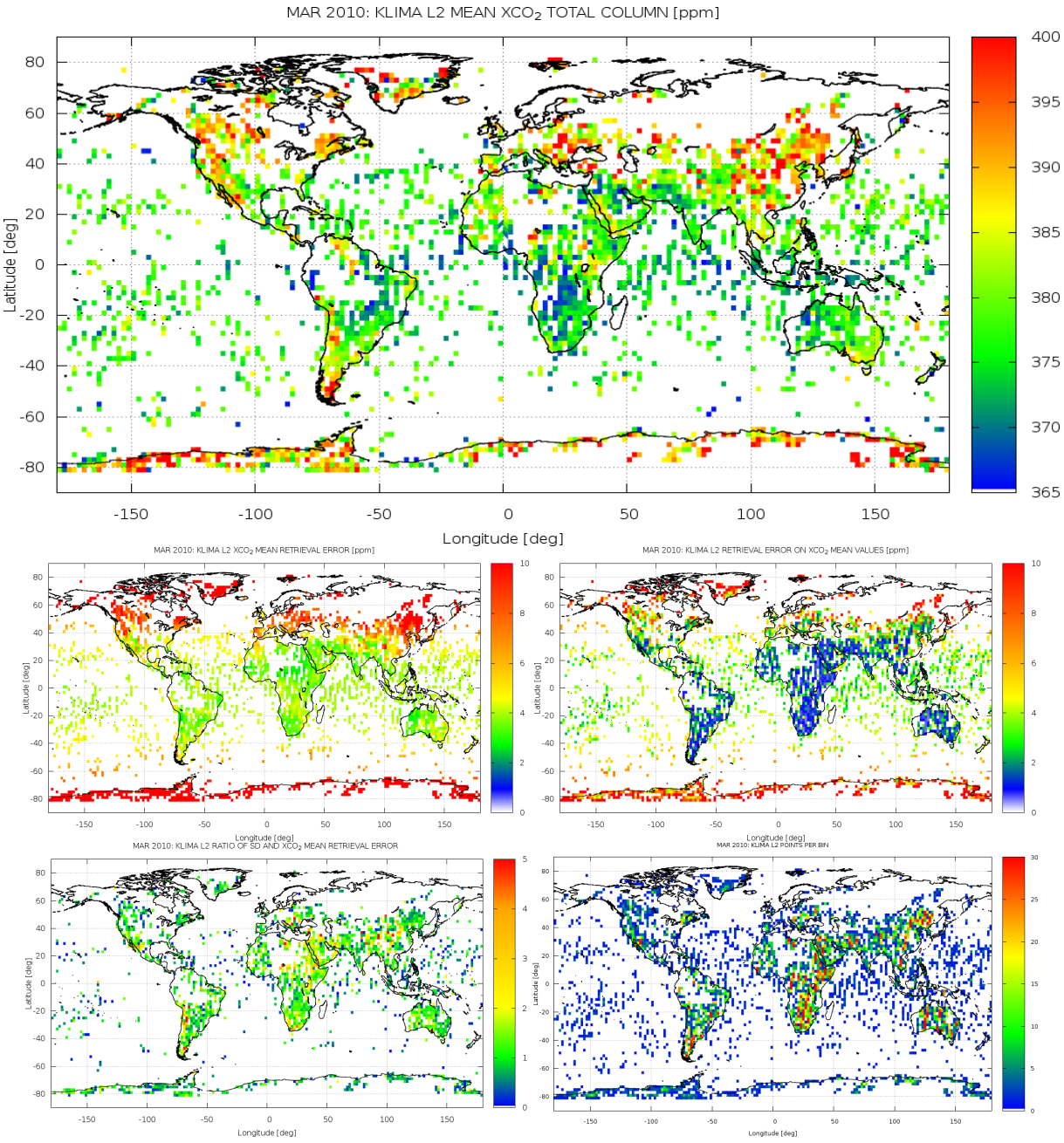

Fig. 5.3: *8-14 MARCH 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

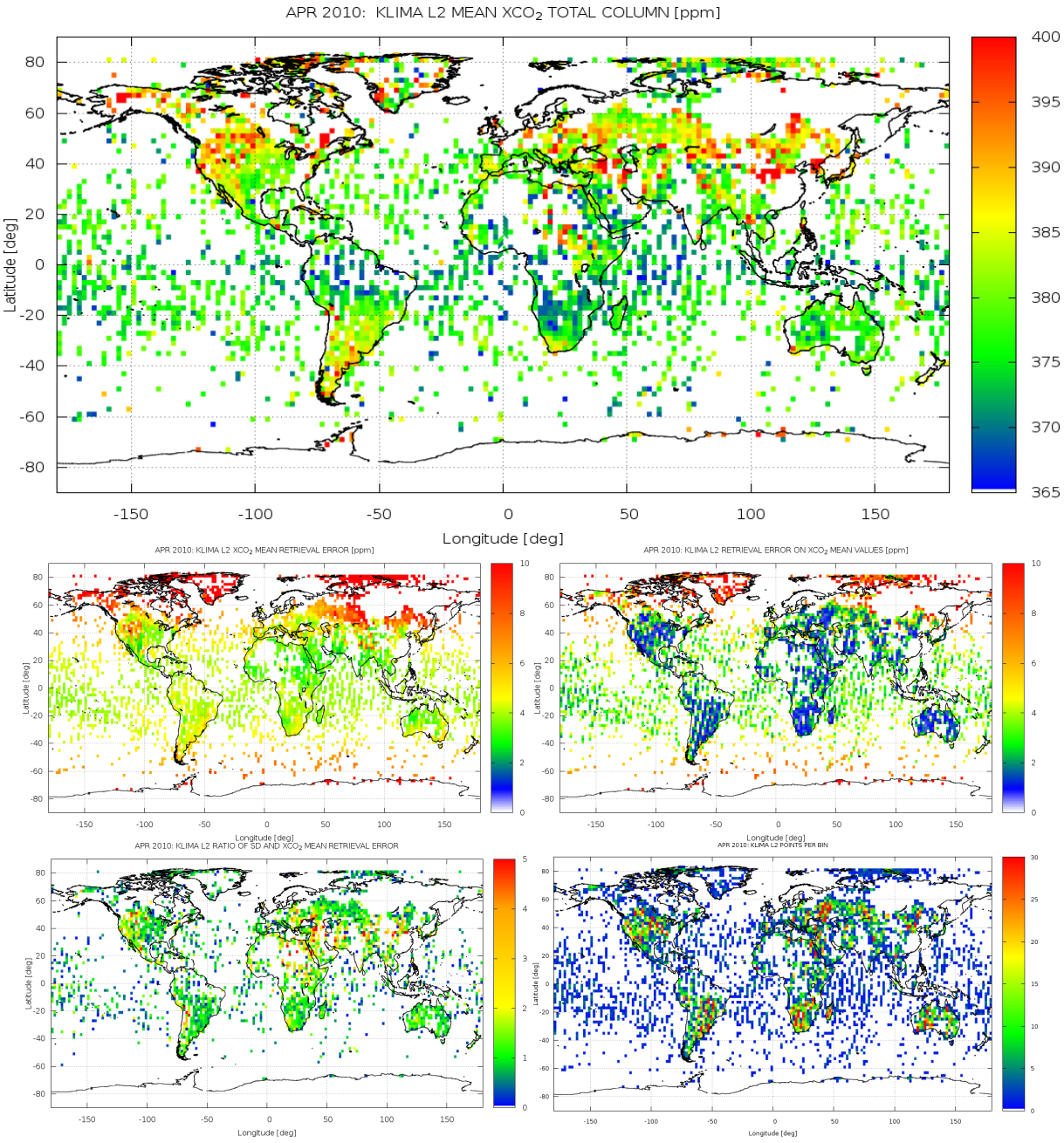

Fig. 5.4: *5-11 APRIL 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

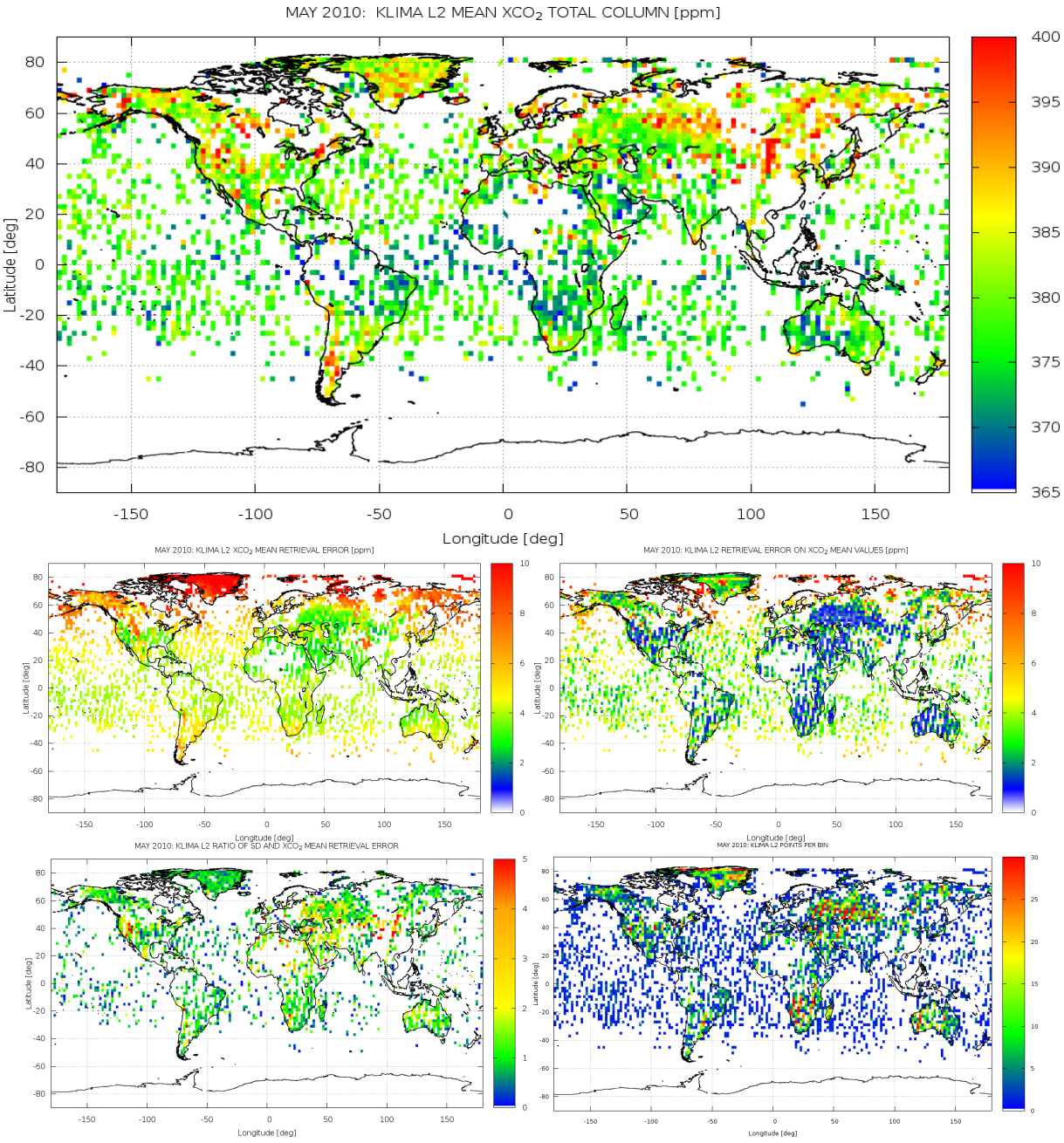

Fig. 5.5: *3-9 MAY 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

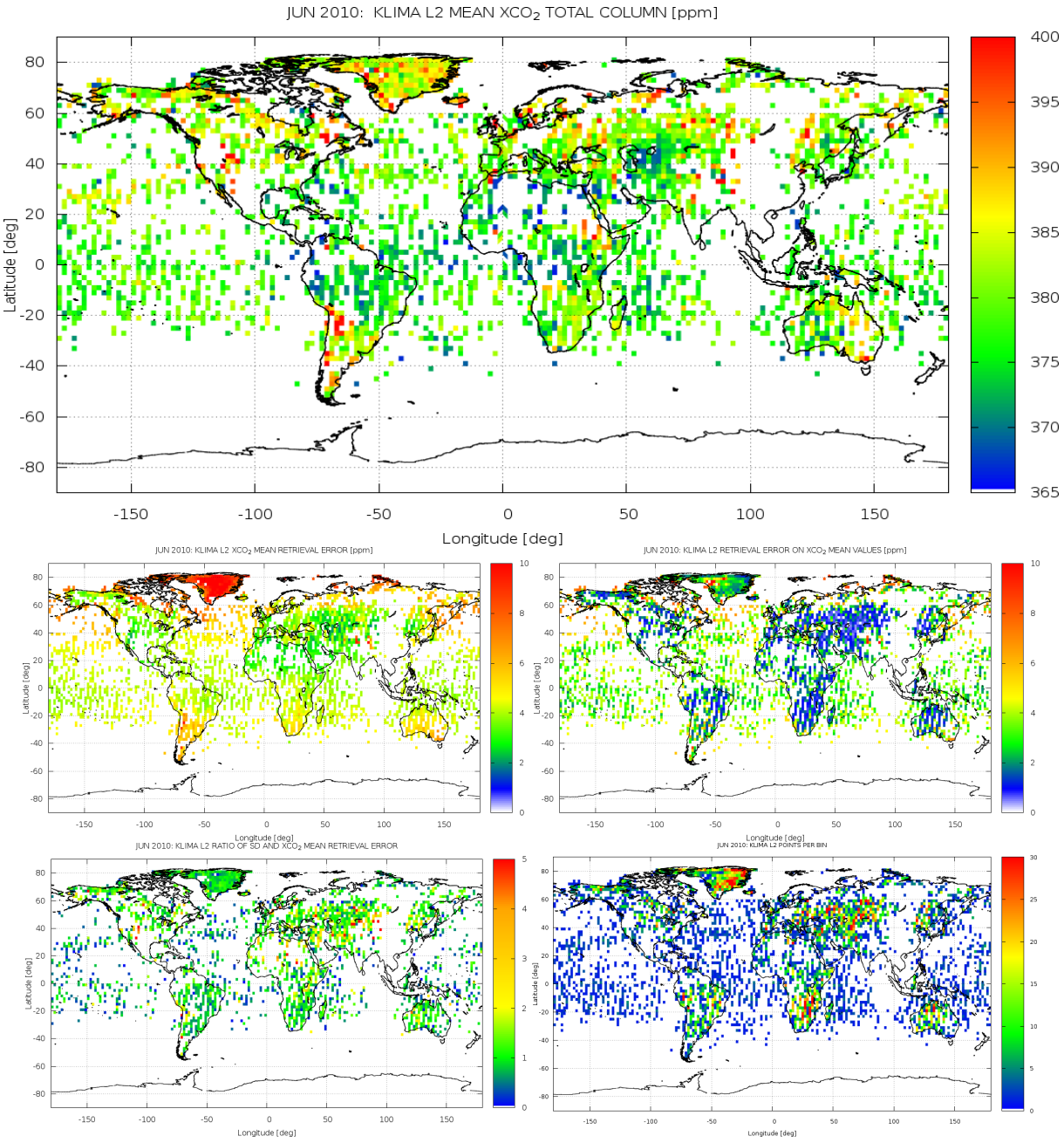

Fig. 5.6: *3-9 JUNE 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

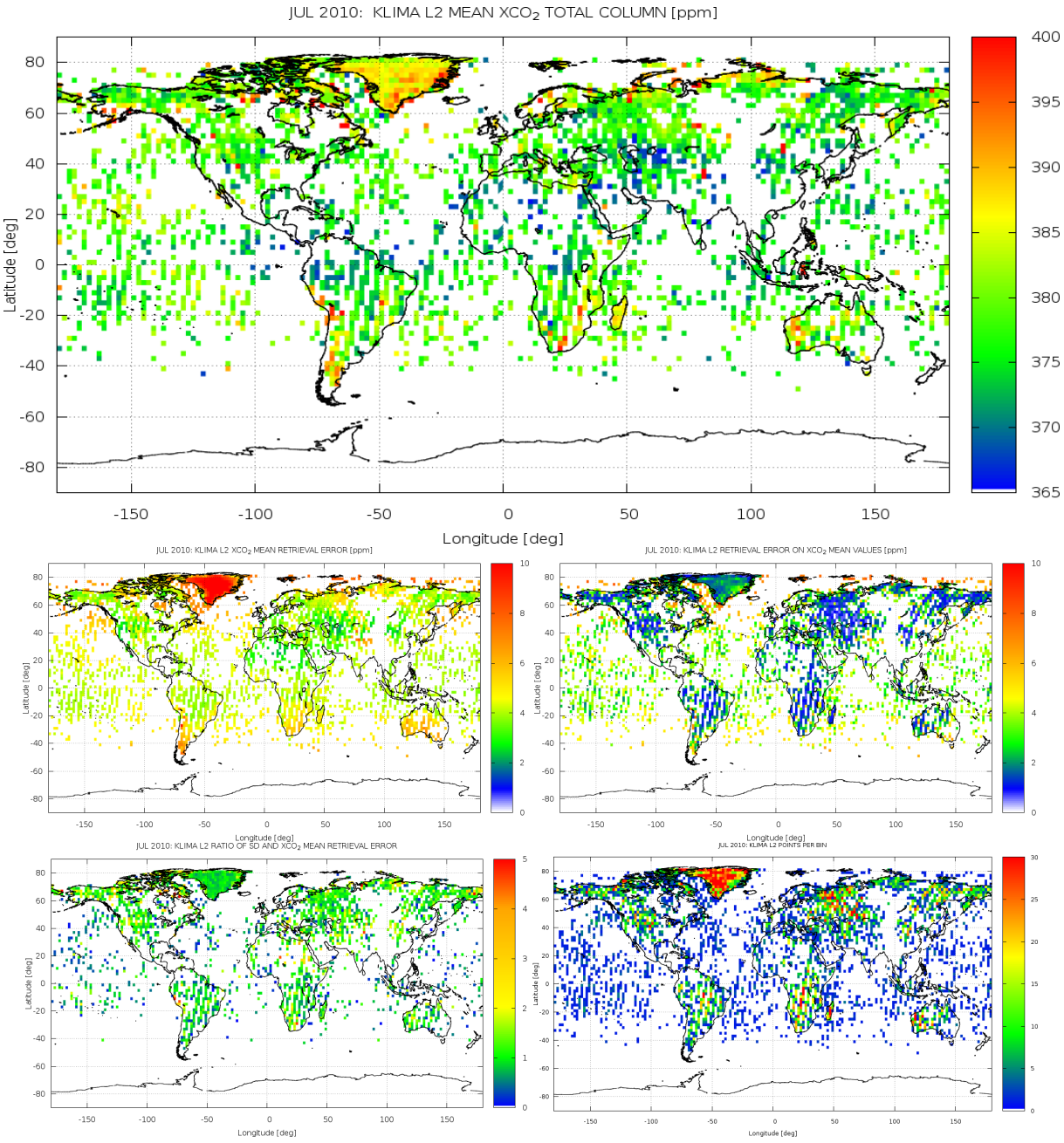

Fig. 5.7: *5-11 JULY 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

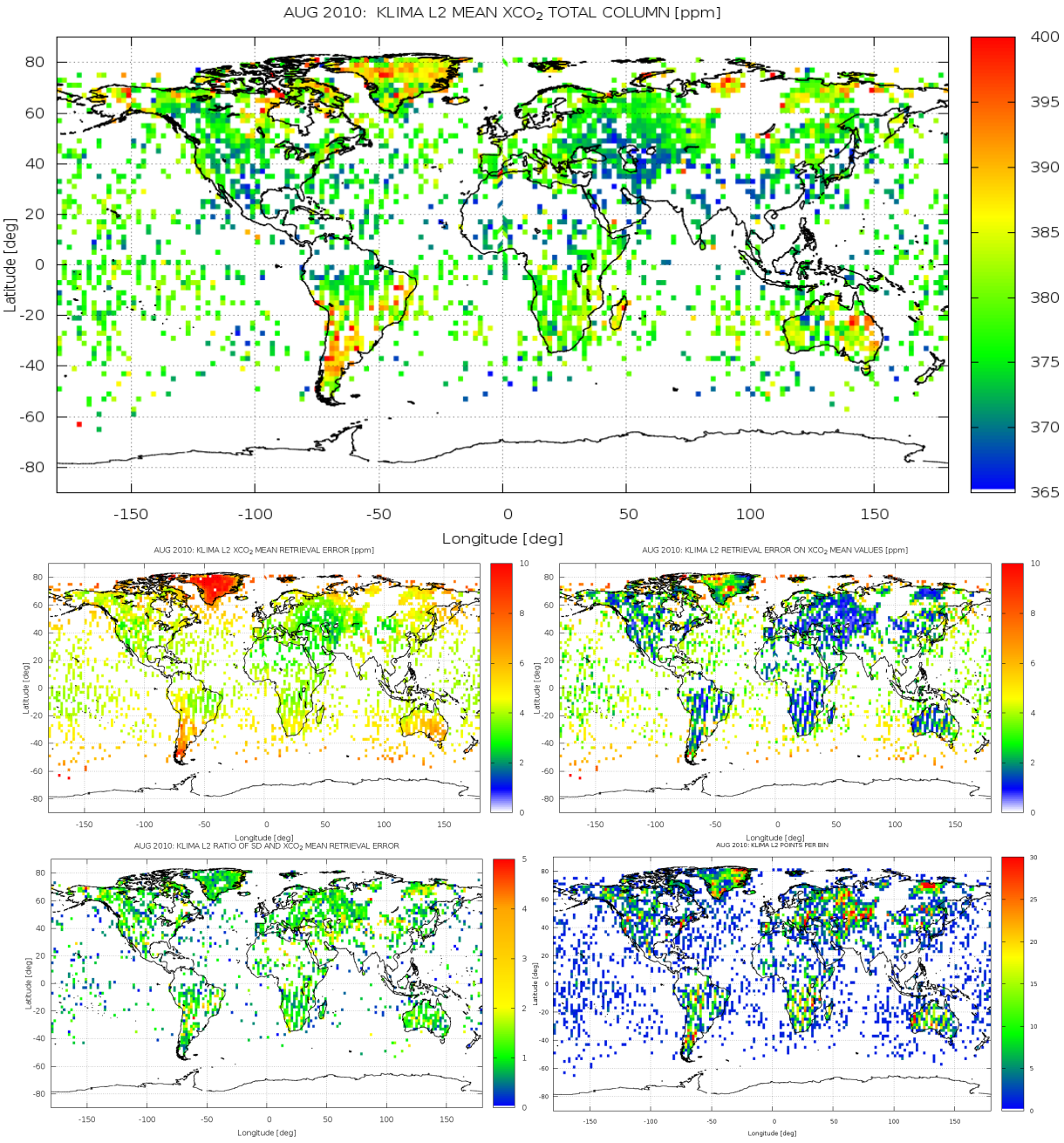

Fig. 5.8: *2-8 AUGUST 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

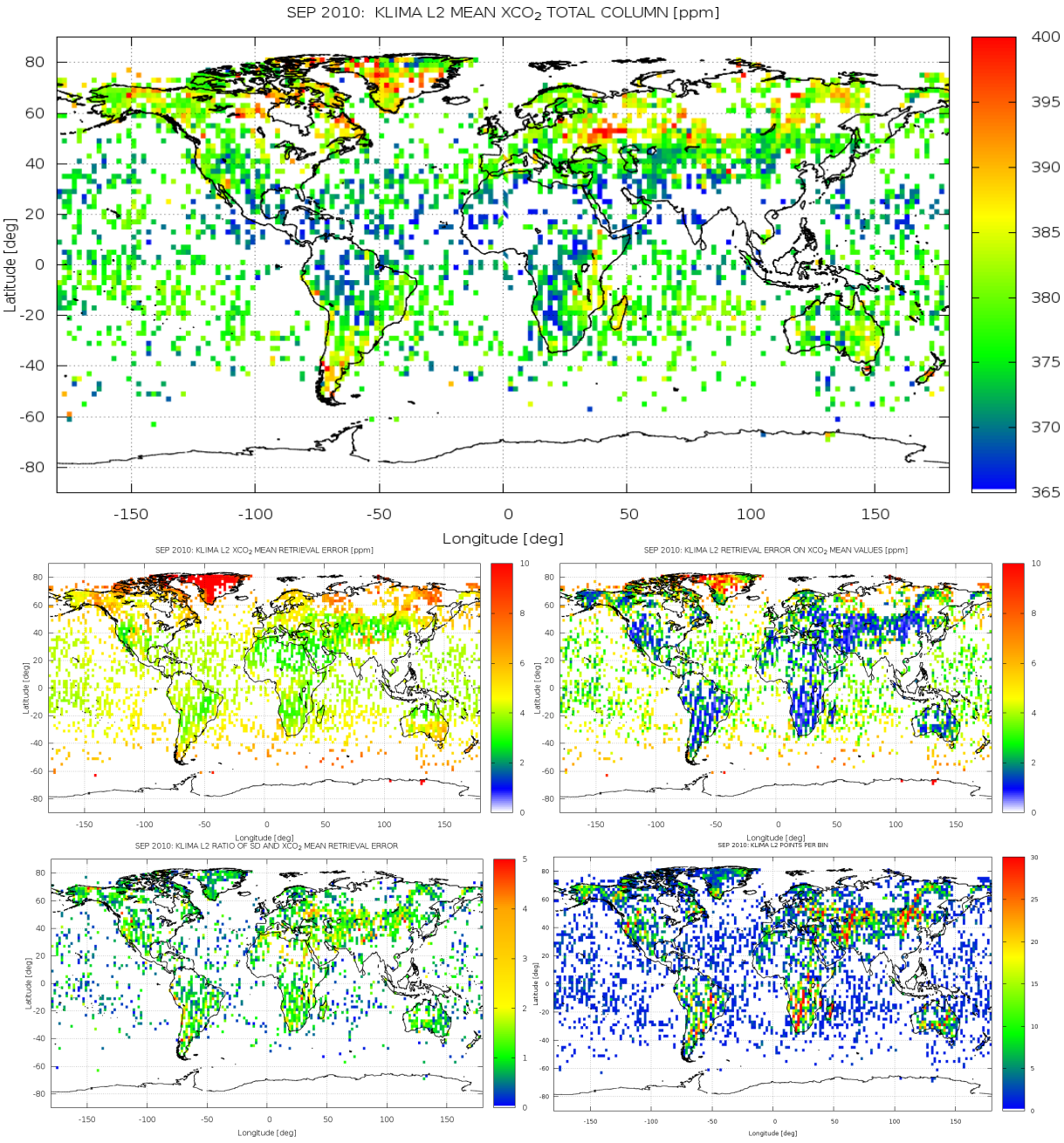

Fig. 5.9: *6-12 SEPTEMBER 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

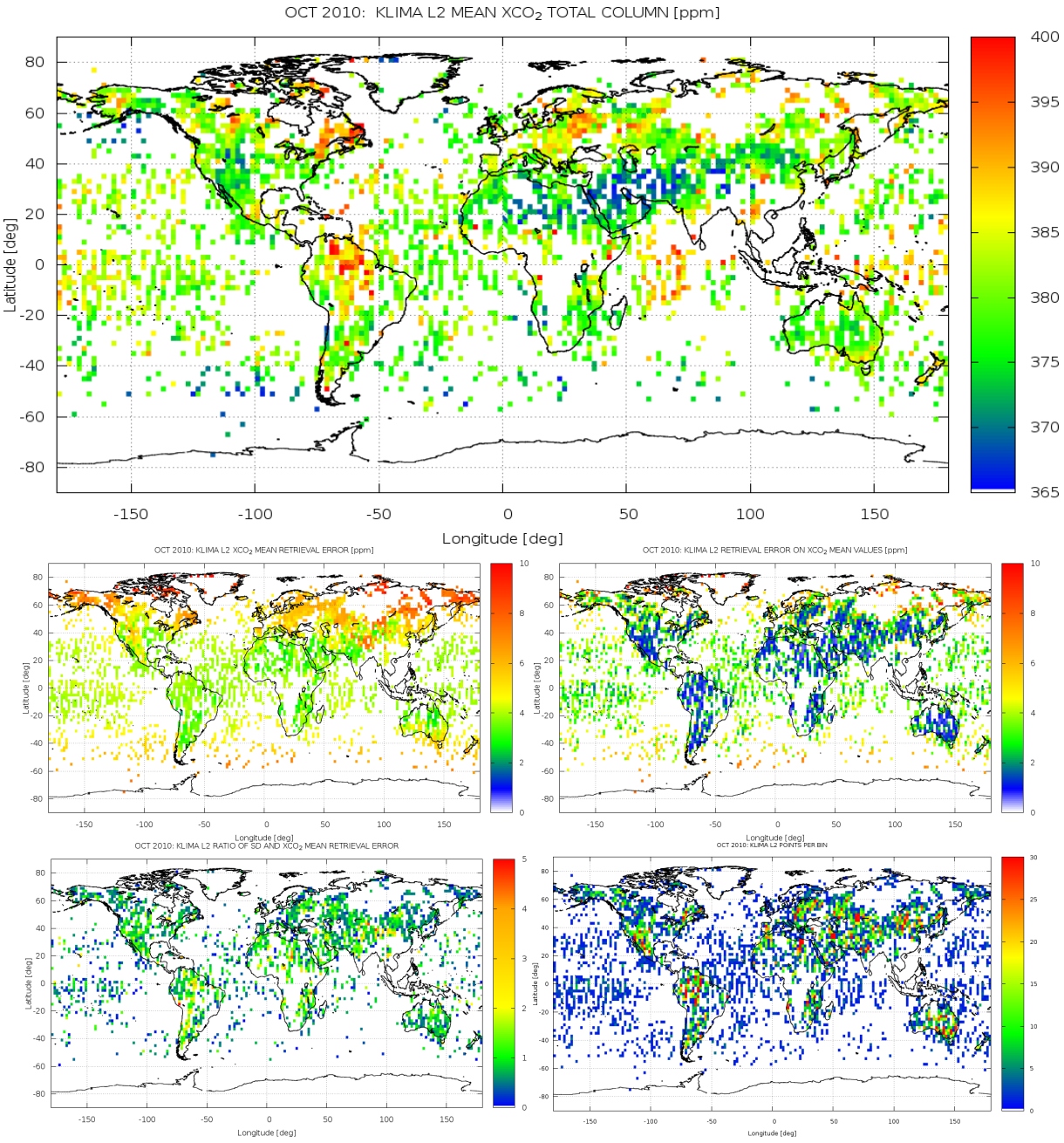

Fig. 5.10: *4-10 OCTOBER 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

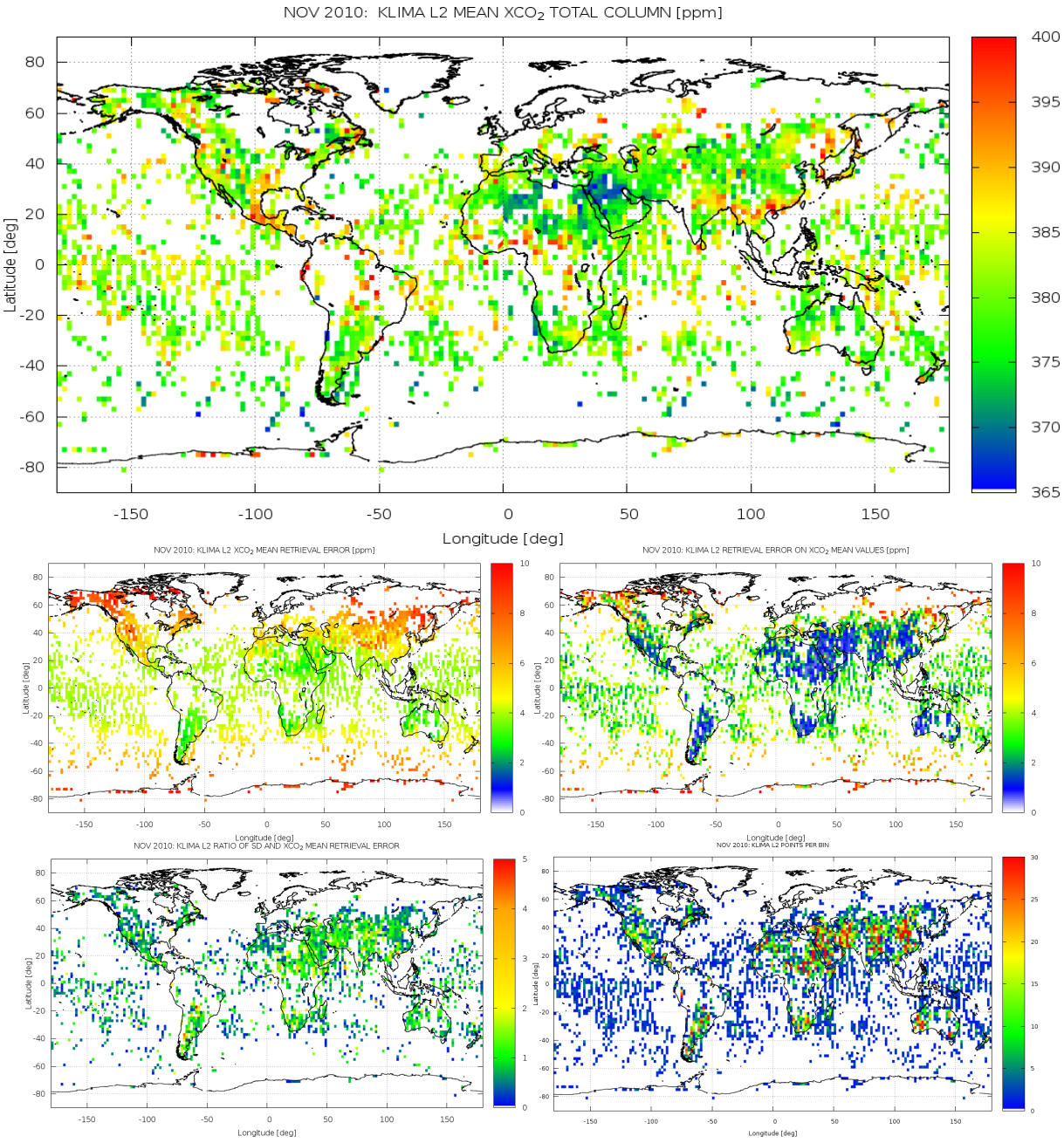

Fig. 5.11: *8-14 NOVEMBER 2010: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

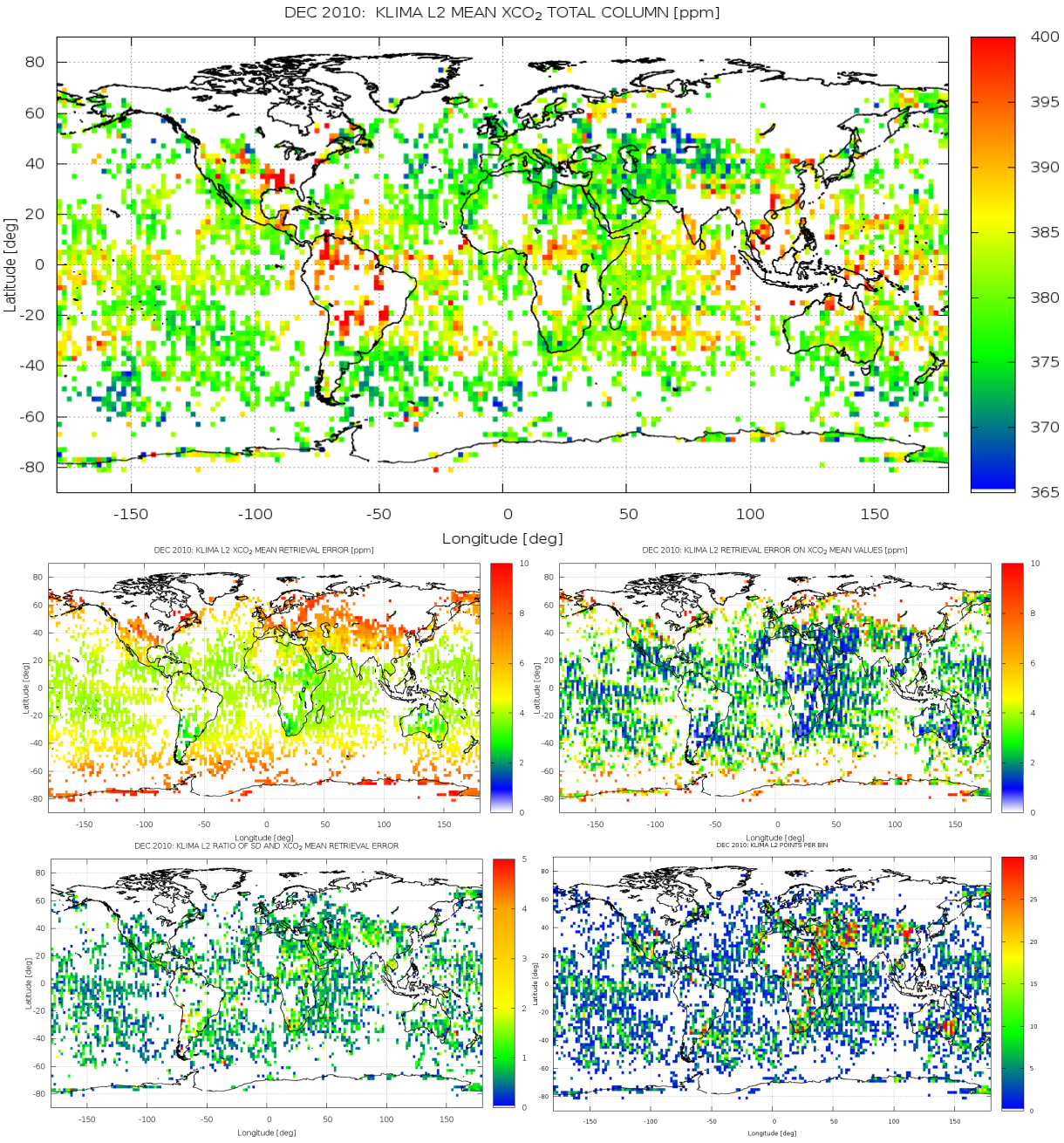

Fig. 5.12: *8-14 DECEMBER: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

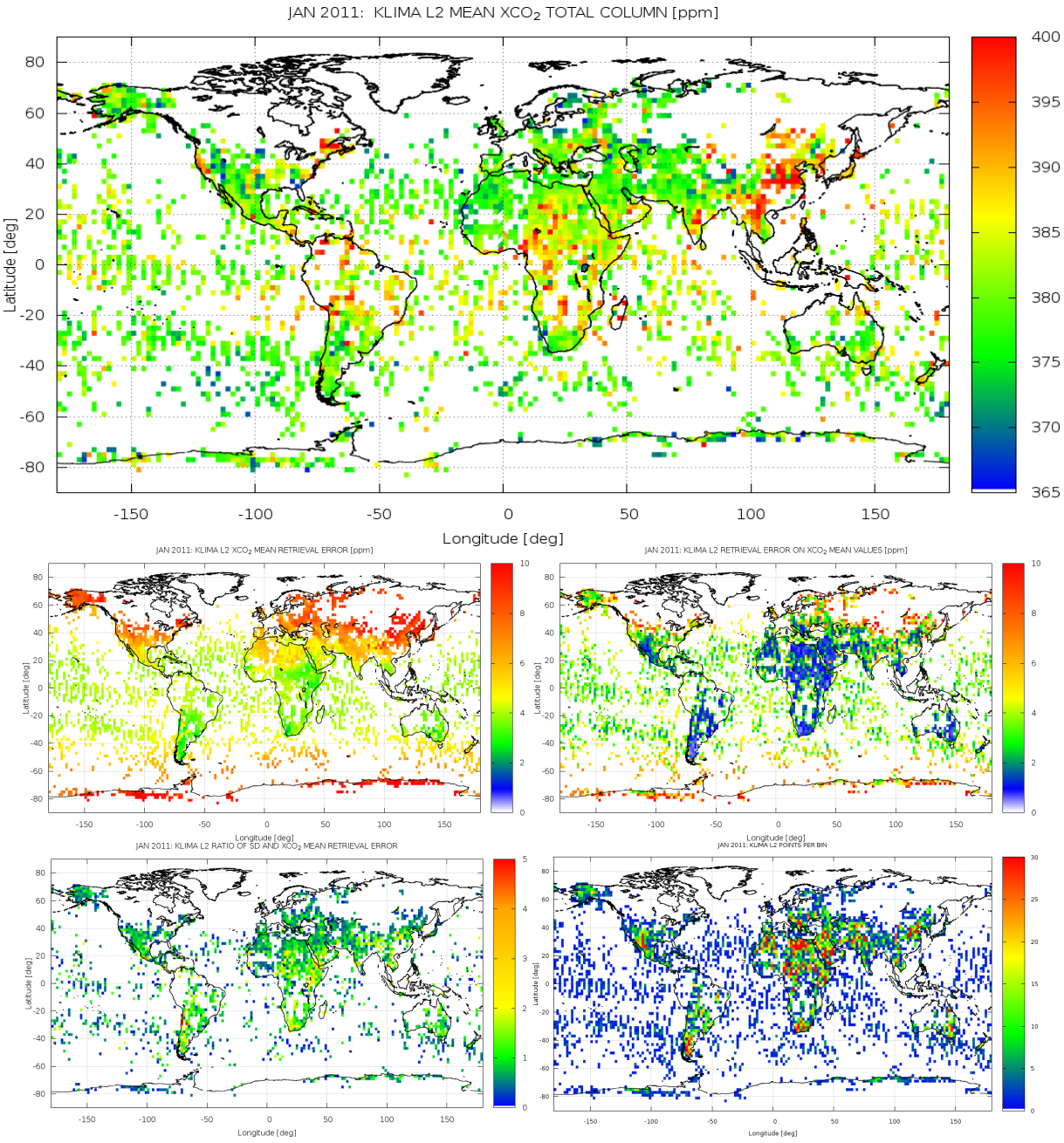

Fig. 5.13: *3-9 JANUARY 2011: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

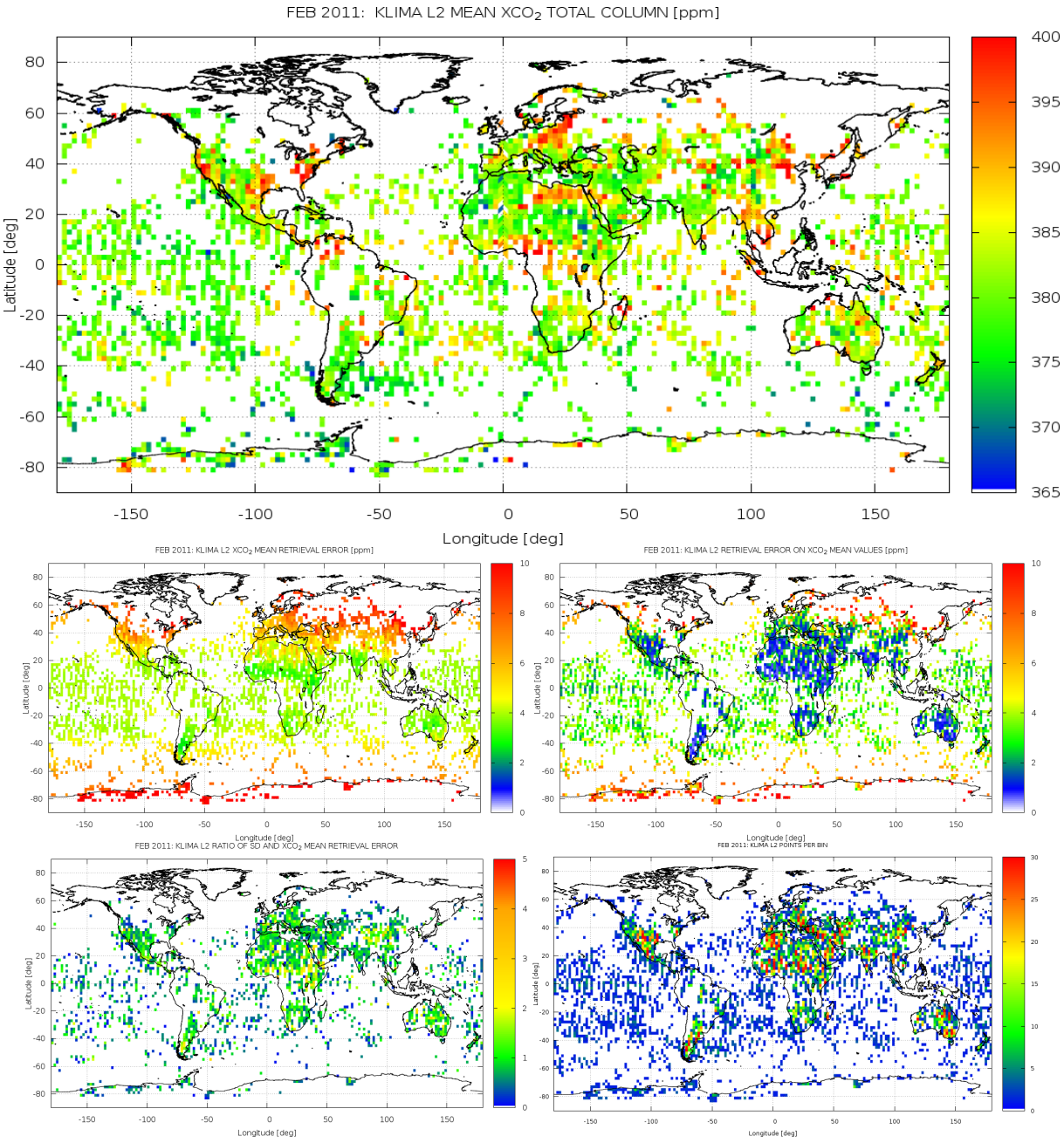

Fig. 5.14: *7-13 FEBRUARY 2011: Global distribution of KLIMA L2 XCO_2 monthly mean; averaged monthly retrieval errors; errors on XCO_2 mean values; geographical and time variability expressed as the ratio between standard deviation and mean error and number of observations per pixel*

also observed in continental China and North America (especially in winter and spring months, from October to May), possibly due to pollution phenomena from anthropogenic sources of CO_2 , such as the combustion of fossil fuels to generate electricity, the combustion of fossil fuels like gasoline and diesel to transportation and industrial processes. When present, data over Greenland and over the South Pole (in the Polar summer period) show considerable values of CO_2 concentration, even if these are the regions affected by the highest averaged error values due to the elevate contrast between surface and the surface nearest atmospheric layer temperature.

In the second map (middle-left) of the figure, from Figure 5.3 to Figure 5.14, I show the geographical distribution of the mean retrieval error on XCO_2 total column values. This is generally in the range 2-4 ppm in the Tropical and Mid Latitude regions, while larger error values are found above 50° (N and S) of latitude. Retrieval errors are larger in winter than in summer in the two Hemispheres and over sea the retrieval error is either larger or smaller than the retrieval error over land depending on the season. These effects are explained by Figure 5.17, that reports the scatter plot of the XCO_2 retrieval error with respect to surface temperature. In this plot, the whole IASI observations dataset has been used, including results obtained over both land and sea and for both day and nigh time. The 5 ppm retrieval error that was observed in the test case discussed in Section 3.2.2, and increased to 6.5 ppm due the further band selection, as described in Section 3.3.2, well represents the average behaviour of the retrieval, but real data show a large variability around this value. The retrieval error monotonically decreases with the increase of the surface temperature. This variation can be explained by the higher intensity of the measured spectrum, when the temperature of the surface is larger. Larger values of the observed radiance correspond to a better precision of the retrieved quantity. Therefore, the observed seasonal and latitude variations of the retrieval error as well as the differences between sea and land mainly depend on the surface temperature conditions. This parameter can be effectively used to determine the cases in which TIR observations provide the best results.

The third map (middle-right) of the figure reports the error on the mean XCO_2 , which is the quantifier of the theoretical performances of KLIMA (1 ppm out of 370 ppm on regional scales (1000 km x 1000 km) at monthly intervals). The results reported in these maps refer to a pixel 2° x 2° at weekly intervals. The blue pixels in the maps meet the requirements even if referring to smaller spatial (the selected pixels are more than one order of magnitude smaller than the require resolution, even if the coverage is not continuous) and time (only one week of observations has been analysed for each month) ranges. The reduced performances observed over the oceans are due to a smaller number of analysed observations. The worst performances are observed in the Polar region, because the precision of KLIMA product depends on the contrast between surface and the surface nearest atmospheric layer temperature (see Figure 5.17).

The fourth map (bottom-left) of the figure shows the ratio between the standard deviation of the retrieved values and their mean retrieval error. This map contains fewer points than the previous error map, because the standard deviation was only calculated when more than two observations are present within the pixel. The plotted ratio, when close to unity, shows the consistency between the estimated retrieval error, determined with error propagation calculations, and the experimental statistical distribution of the values. Values larger than unity are due to either unaccounted errors or geographical and time variability. Most of the pixels are characterized by values of the order of unity and often significantly smaller than unity. Values smaller than unity suggest that a conservative estimate of the retrieval error is made. This is due to the choice of retrieving the vertical profiles of Water Vapour and temperature at high vertical resolution, in order to model the influence of the variability of these profiles on the retrieved value of XCO_2. This choice causes a larger retrieval error for Water Vapour and temperature and through the simultaneous retrieval this error enhancement propagates into the XCO_2 retrieval error. In particular, very small values are observed over all oceans, in Australia, Northern Asia and Europe. Interesting are also the low values observed over Greenland, probably facilitated by the relatively large retrieval errors present in this area. Higher values of the standard deviation are found in the months of AprilMay, when the natural variability of CO_2 is high, and the minimum standard deviation is found in fall, when the natural variability of CO_2 is low. The frequent and homogeneous distribution of values

close to unity provide a strong indication about the correctness of the error assessment and characterizes the corresponding regions as areas of low variability. On the other hand, a high variability is observed in some locations over land, in particular over the Pacific coast of Southern and Northern America, over Southern Africa and over continental Asia. In this case the doubt remains whether we are observing either systematic errors induced by the land background or variability due to the different ecosystems.

Finally, the fifth map (bottom-right) of the figure represents the number of observations per each pixel.

As reported in the Introduction of this thesis work, an accurate retrieval of CO_2 concentration is an essential ingredient for quantifying sources and sinks of CO_2 . The traditional approach is the atmospheric inversion of CO_2 concentrations measured at the earth's surface ([2]). The first and most important problem that arises from using this method is the density of surface station: the sparseness and spatial inhomogeneity of the existing surface network, which suffers the lack of ground station in many areas that are interesting from a Carbon cycle point of view, such as the Amazonian rain forest and the Arctic tundra regions, have limited the ability to understand the quantity and spatio-temporal distribution of CO_2 sources and sinks. Assimilating satellite measurements in atmospheric inversions of CO_2 can bring benefit in the retrieval of sources and sinks. In particular, even if space-based measurements of CO_2 are not as accurate as surface layer measurements, the increased spatial coverage can provide information not available from the sparse surface network ([6]).

5.3 Seasonal variation of KLIMA Level 2 products

Over land, the seasonal variation 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 (April, May and June) than in the Southern Hemisphere, while these higher values decrease during summer months (July, August and September). This behaviour could be explained in terms of plant photosynthesis of the biomass that, for the Northern Hemisphere, is not yet competitive with respiration (primarily due to uptake and release of CO_2).

Figure 5.15 reports KLIMA XCO_2 seasonal variation in Tropics [30°:-30°], Northern Hemisphere mid-latitude [60°:30°] and Southern Hemisphere mid-latitude [-30°:-60°]. For the North mid-latitude case it is possible to see a clear seasonal cycle of about 10 ppm, with a maximum during spring (April and May) and a minimum in summer (July and August). In the Southern Hemisphere the XCO_2 variation is more complex, with a first little maximum in spring and a maximum in summer, and still a minimum in fall. The amplitude of XCO_2 variability in the Southern Hemisphere is also much smaller than in the Northern. A slight decrease of the cycle amplitude with latitudes is also observed. In Tropics the XCO_2 variation is smaller than mid-latitude and the seasonal trend shows only a little minimum in September. Similar results are also reported in [61] and [23].

5.4 Estimate of KLIMA retrieval errors

The monthly maps discussed in the previous section have already provided some useful information about the quality of the XCO_2 values retrieved with the KLIMA code from IASI L1 measurements. However, some further considerations can be made on the products errors.

Figures 5.16 and 5.17 report the scatter plots of the retrieval χ^2 and the XCO_2 retrieval error respectively, with respect to surface temperature. These scatter plots indicate that, while the retrieval error is lower for larger surface temperatures, the primary retrieval quality flag (the χ-test) shows that increasing residuals are observed for larger surface temperatures. A χ^2 greater than unity is usually a sign of unaccounted errors and an indication that the retrieval error must probably be multiplied by the square root of the χ^2 value. The observed increase of χ^2 can be explained by the fact that the residuals increase when the measured radiance is larger.

The errors that contribute to the observed residuals can be: FM errors, such as spectroscopic errors, or to missing model implementations, such as the scattering contribution due to aerosols, but also sys-

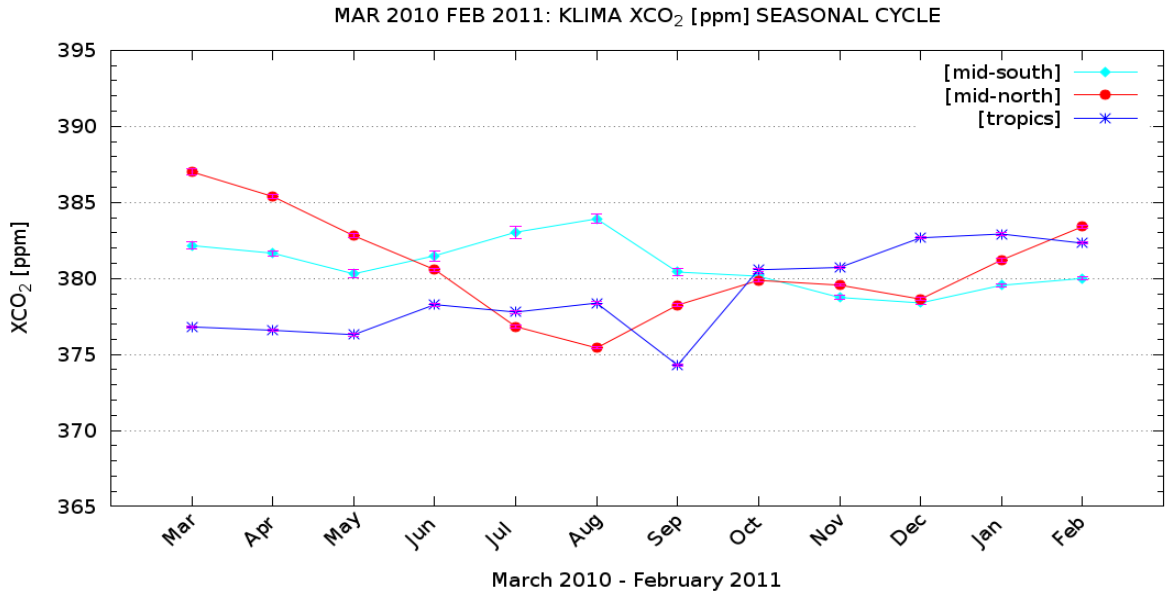

Fig. 5.15: *March 2010 - February 2011: KLIMA L2 XCO_2 seasonal variation in Tropics [30°:-30°], Northern Hemisphere mid-latitude [60°:30°] and Southern Hemisphere mid-latitude [-30°:-60°]*

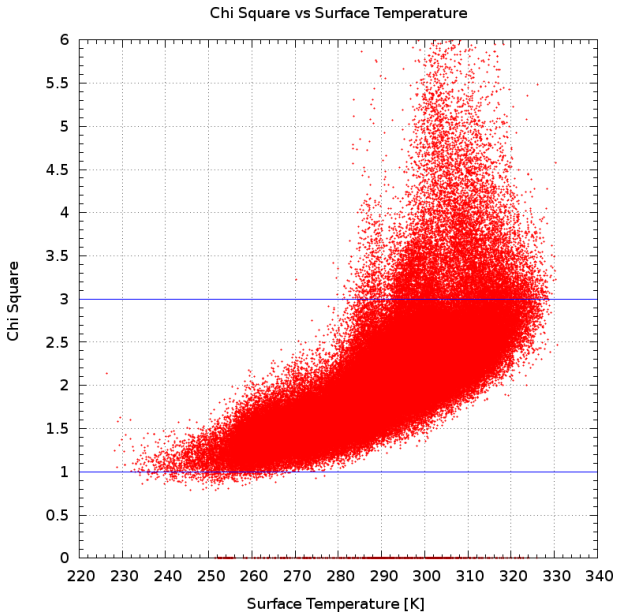

Fig. 5.16: *Scatter plot of the χ^2 retrieval error with respect to the surface temperature. Blue lines indicate the χ^2 range accepted for the comparison activities*

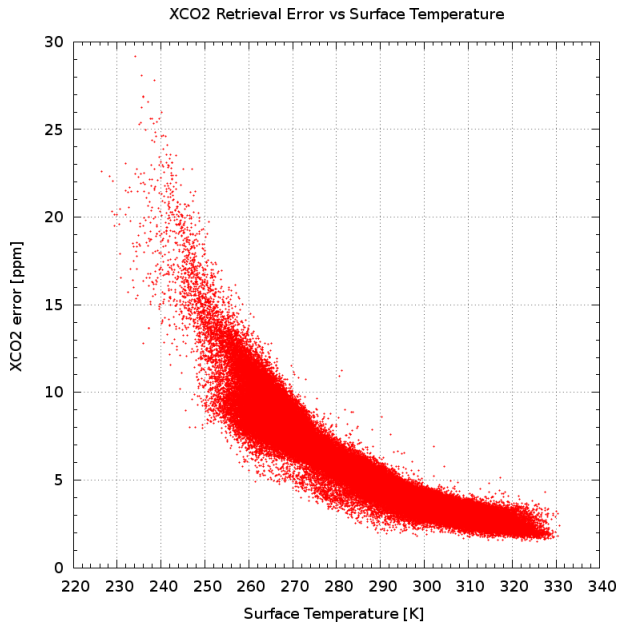

Fig. 5.17: *Scatter plot of the XCO_2 retrieval error with respect to the surface temperature*

tematic errors in the IASI measurements (e.g. introduced in the calibration). These errors, that manifest themselves in the residuals, can also affect the retrieval products increasing the random error of the retrieved product and causing systematic differences (biases) between retrieved and real values. The first effect implies that the precision of the measurements is not fully characterized by the error propagation and by theoretical retrieval error. However, this error enhancement is expected to be very small because in most cases the statistical spread of the observations is consistent with the calculated retrieval error. The variability of the atmosphere (either in time or in space), when present, is the largest cause of precision abatement. The second effect can be contained by reducing the residuals as much as possible, and the result of systematic residual smaller than the IASI noise shows that important progresses have been made in this direction, but the actual entity of the biases can only be assessed with inter comparison and validation exercises.

Some other interesting features can be observed in the scatter plot of Figure 5.16. The events of the scatter plot are contained within a well defined envelope, suggesting that a reproducible law is governing the observed increase of χ^2 with the surface brightness temperature. Indeed a similar behaviour was also observed in the case of retrieval error and surface brightness temperature shown in Figure 5.17. In Figure 5.16, however, next to the main envelope it is possible to notice one or two clouds of events which do not seem to obey to the general rule. These are the events that have high values of χ^2 corresponding to surface temperature of about 287 K and the events that have high values of χ^2 corresponding to surface temperatures between about 300 and 320 K. Figure 5.20 shows the geolocation of the first group of events by selecting retrievals with a χ^2 greater than 3 and a surface temperature less than 292 K. These events only occur in areas where deserts are present. This suggests that the exceptionally large values of χ^2 may be caused by some typical event that can happen in the desert, such as sand storms.

A similar result is obtained by selecting the second group of events. In Figure 5.19 the events with a χ^2 greater than 4,5 have been selected and also in this case areas are identified where sand storms are possible. Given the high surface temperature of the events shown in Figure 5.19, it is not surprising to observe here mainly desert areas, but to be limited to these areas in the case of the events shown in Figure 5.20 strongly suggests that the cause of the high χ^2 is not the desert, but an event occurring in the desert. In order to limit the errors that these non-modelled effects may introduce, it was decided that all the events that have a χ^2 greater than 3 (see the blue lines in Figure 5.16) will not be considered. The geolocation of the IASI observations excluded because the χ^2 is larger than 3 is shown in Figure 5.18. Also in this selection most of the events are in desertic areas.

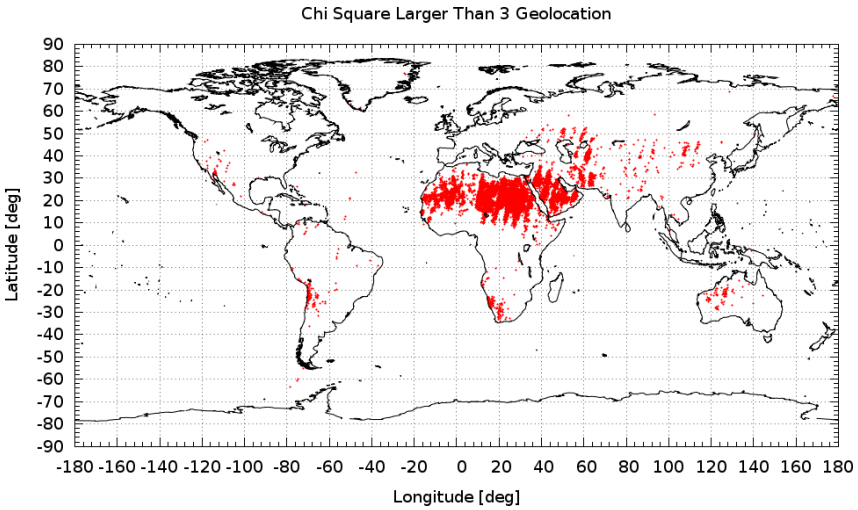

Fig. 5.18: *Geolocation of the analysed KLIMA observations with a χ^2 larger than 3*

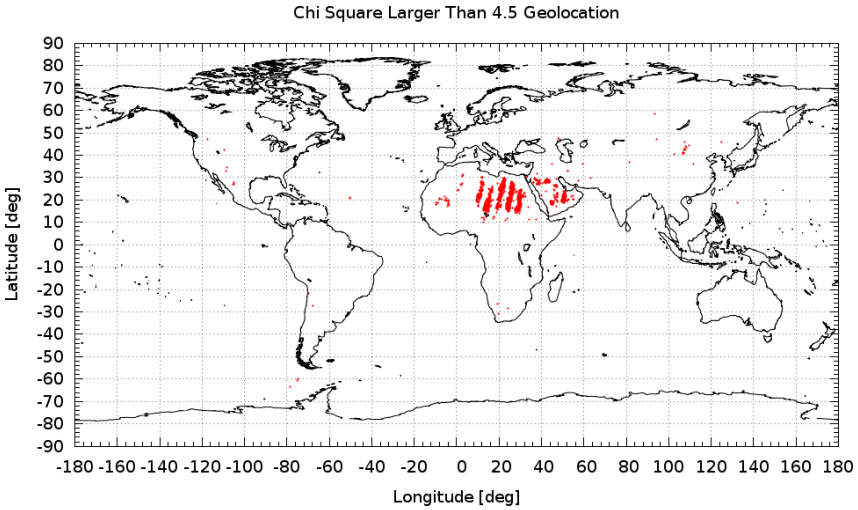

Fig. 5.19: *Geolocation of the analysed KLIMA observations with a χ^2 larger than 4.5*

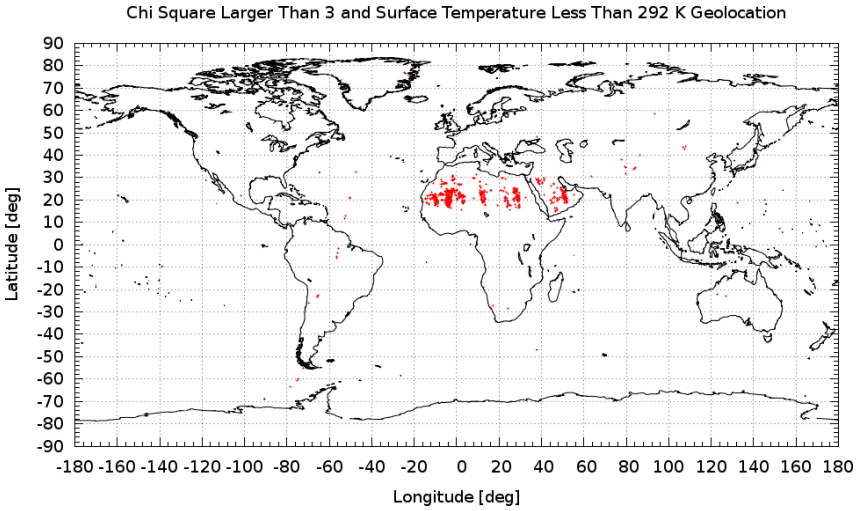

Fig. 5.20: *Geolocation of the analysed KLIMA observations with a χ^2 larger than 3 and the surface temperature less than 292 K*

Chapter 6

Results of Carbon Dioxide comparisons

In general, validation refers to assessing the uncertainty of higher level, satellite sensor derived products, as well as products obtained from other type of sensors, by analytical comparison to reference data, which is presumed to represent the true value of the target. Inter-comparison of data products or model outputs with the reference provides an initial indication of gross differences and possibly insights into the reasons for the differences.

As a part of the activities conducted to evaluate the retrieval quality of KLIMA code, we decided to perform a comparison of the complete CO_2 dataset retrieved using KLIMA with the Level 2 products delivered by EUMETSAT and by TCCON ground stations as a first test of algorithm performances. The final step is the performing of a cross-validation between IASI and TANSO-FTS SWIR atmospheric CO_2 concentrations to assess the coherence and temporal behaviour of large scale spatial features obtained from the two different sensors.

In this Chapter we report the results of the comparisons between CO_2 concentration retrieved with KLIMA algorithm from IASI measurements and:

- standard L2 products delivered by EUMETSAT 6.3;
- operational L2 products collected from TCCON ground stations 6.4;
- operational L2 products obtained from TANSO-FTS/GOSAT SWIR channels 6.5.

6.1 Main strategies adopted for the comparison

To compare the L2 data obtained from KLIMA code with EUMETSAT L2 standard products, with TANSO-FTS/GOSAT SWIR L2 products and with the TCCON ground stations products different strategies have been adopted.

The comparison with EUMETSAT represents a comparison of the performance of two different retrieval algorithms, which refer to the same measurements. This is the simplest case, in which it is possible to make a direct comparison of each XCO_2 value. The comparison has been performed on global geographical scale, evaluating the averaged XCO_2 differences for every weeks in the annual time range selected. The obtained differences are represented in a monthly map on a 2° x 2° pixel grid. Moreover, we and plotted the XCO_2 seasonal cycle and the maps of XCO_2 annual mean values evaluated over a set of selected geographical macro areas. The selected macro areas are shown in Figure 6.1. Results are reported in 6.3.2.

The comparison of KLIMA L2 results with TANSO-FTS operational products has been performed in two ways. The first way consists in considering only the coincident observations. So, we chose the coincidence criteria, in terms of time and geographical distances, as an acceptable compromise between the need to compare quantities retrieved from measurements observing the same scene and a number of samples suitable for a good statistics. The second way consists in comparing the XCO_2 values averaged for each week selected of months from March 2010 to February 2011, for a global geographical coverage.

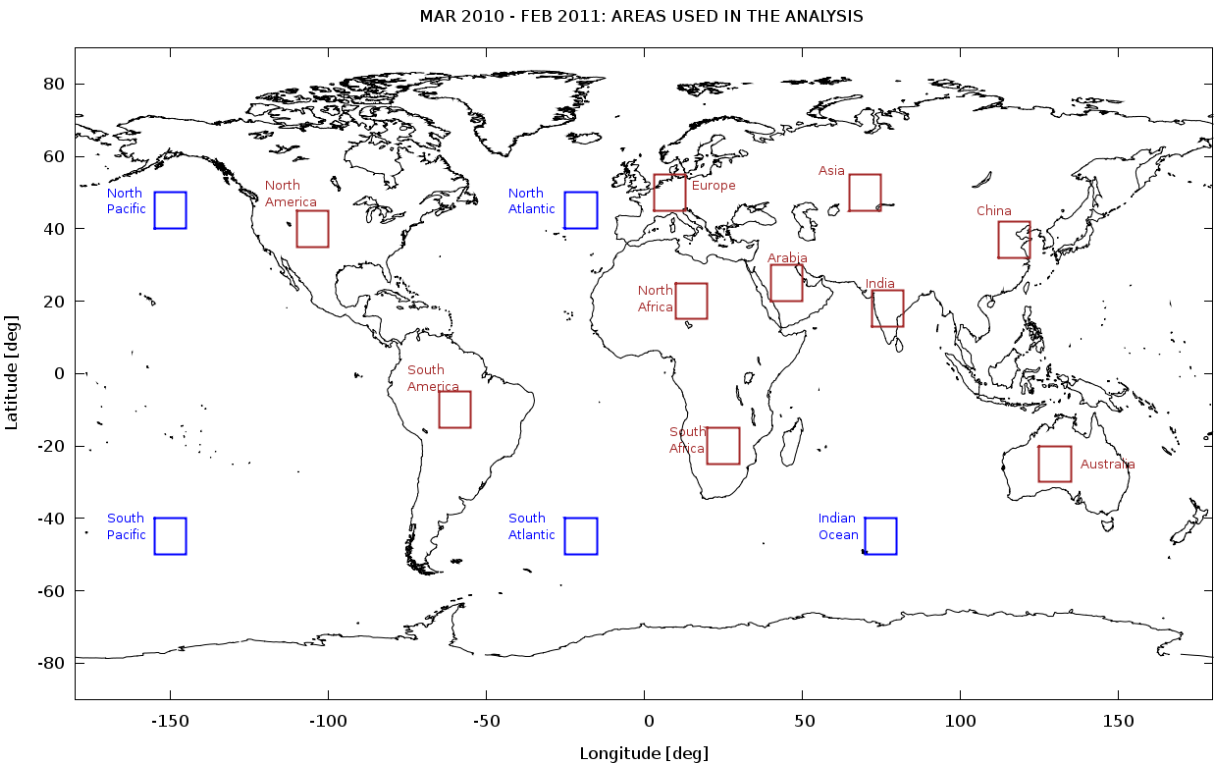

Fig. 6.1: *Macro areas selected for the annual XCO_2 mean values comparisons between KLIMA L2 products, EUMETSAT L2 products and TANSO-FTS L2 products*

The averaged values are then reported on a 2° x 2° and on a 9° x 9° pixel grids. The last grid corresponds to approximately a 1000 km x 1000 km pixel dimension at the Equator. For each month, the averaged values are also reported in a scatter plot. Finally, we evaluated the seasonal cycle and the annual mean over macro areas. Results are reported in 6.5.2.

Due to the limited geographical coverage of ground-based stations, we carried out the comparison between KLIMA L2 results and TCCON products using only the coincident observations. We considered only IASI measurements that satisfy certain criteria of time and spatial distances from TCCON stations. Then we averaged the coincident observations in order to obtain a single value for each day in which the coincidences has been found to evaluate the correlations with KLIMA L2 products. Results are displayed in 6.4.

6.2 Method to compare retrieval products with different Averaging Kernels and different a priori

In the case of comparison of remotely-sensed products obtained from two (or more) different sensors it is necessary to consider the different sensitivity of the instruments to vertical distribution of the target and the use in the retrieval process of some kind of prior constraint in the form of prior information about the observed state. Thus, in order to properly compare indirect measurements obtained from different instruments, the effect of the different smoothing can be described by means of the Averaging Kernels matrix (AK) and the a priori profiles used in the retrieval analysis of both measurements must be taken into account [73]. The AKs describe the altitude-dependent sensitivity of the retrieved profile with respect to the true values. For an ideal measurement, the AK would be 1.0 at all altitudes, but in case of real measurements sensitivity is greater to some altitudes than others and the corresponding values of AK are less than unit. The a priori profiles represent the knowledge of the state before the measurement is made and the external constraint about the observed atmospheric state assumed in the retrieval process to make invertible the problem and determining a geophysical solution in the case of ”ill-posed” problems

3.2.

For the comparison with TANSO-FTS data the a priori profile used in KLIMA data analysis were applied to the retrieved TANSO-FTS SWIR XCO_2 profile, by using the equation:

$$x_{1TANSO} = x_{TANSO} + (A_{TANSO} - I)(x_{aTANSO} - x_{aIASI}) \tag{6.1}$$

The profile is then smoothed by using the IASI AK matrix, according to the equation:

$$x_{2TANSO} = x_{aIASI} + A_{IASI}(x_{1TANSO} - x_{aIASI}) \tag{6.2}$$

where x, x_{1TANSO} and x_{2TANSO} are the TANSO-FTS products adjusted for the a priori and smoothed for the AK respectively, A_{IASI} and A_{TANSO} are the AK matrices, x_{aIASI} and x_{aTANSO} are the a priori profiles used in the retrieval analysis (A_{TANSO} and x_{aTANSO} have been extracted from TANSO-FTS L2 files).

The smoothed profile is successively converted in CO_2 columnar values by means of the operator pressure weighting function, defined as the ratio between the partial dry air column and the total dry air column. The IASI AKs depend on latitude and on season, for this reason the a priory and the IASI AKs, used to modify the TANSO-FTS data, have been evaluated on five latitude ranges (Northern Polar, Northern Mid Latitude, Equatorial, Southern Mid Latitude and Southern Polar) and on four climatological atmospheres (Winner, Spring, Summer and Autumn). The major effect is due to the different smoothing represented by the AKs. In fact, the IASI AKs present a maximum close to 200 hPa and generally reach low values above 900 hPa and below 100 hPa, while the TANSO-FTS AKs are quite constant on the same altitude range.

As shown in many recent works (for example [92], [61], [70], [96] and [20]), also the TCCON data used as a reference in a validation procedure need to take into account the a priori adjustment and the AKs smoothing. In some cases, these effects can be neglected, because the smoothing error is less than the retrieval error (random and systematic); for example, to validate XCO_2 products obtained from TANSO-FTS SWIR measurements on GOSAT satellite, Cogan et al. [20] and Wunch et al. [96] estimated the smoothing error due to the TANSO-FTS SWIR AK in few tenths of ppm, while the retrieval error was about 2-3 ppm so they did not applied AK smoothing.

For the comparison of XCO_2 obtained from IASI measurements using KLIMA code with the TCCON data the effect of the a priori along with the smoothing error, evaluated using Eq. A12 reported in [96], have been estimated of about 5 ppm. In this case, the AK smoothing can not be neglected. However, to perform the AK smoothing, the retrieved XCO_2 profile from TCCON is needed, but it is not directly available from the standard products of the network. For this reason, the comparison between TCCON and KLIMA L2 products can not be considered as a quantitative validation test. On the basis of the previous statements the inter-comparison between KLIMA L2 and TANSO-FTS products adjusted for the a priori profile and smoothed for the AK matrix can be considered a more reliable step in the validation task of the XCO_2 measurement from satellite observations.

6.3 Comparison with IASI EUMETSAT Level 2 products

6.3.1 EUMETSAT products overview

The EUMETSAT CO_2 product has not yet been validated therefore it is not included in the operational IASI L2 products [48]. Despite the experimental status of EUMETSAT IASI L2 CO_2 products, however some advantages to perform this comparison exist, because selected spectral bands are the same for KLIMA and IASI L2 (even if KLIMA algorithm uses a number of channels significantly higher than IASI processor 4.1.3). Moreover, measurements are spatial and time coincident, the observation geometry and IFOV are the same so we do not need to correct for smoothing effects, necessary in the case of different altitude sensitivity.

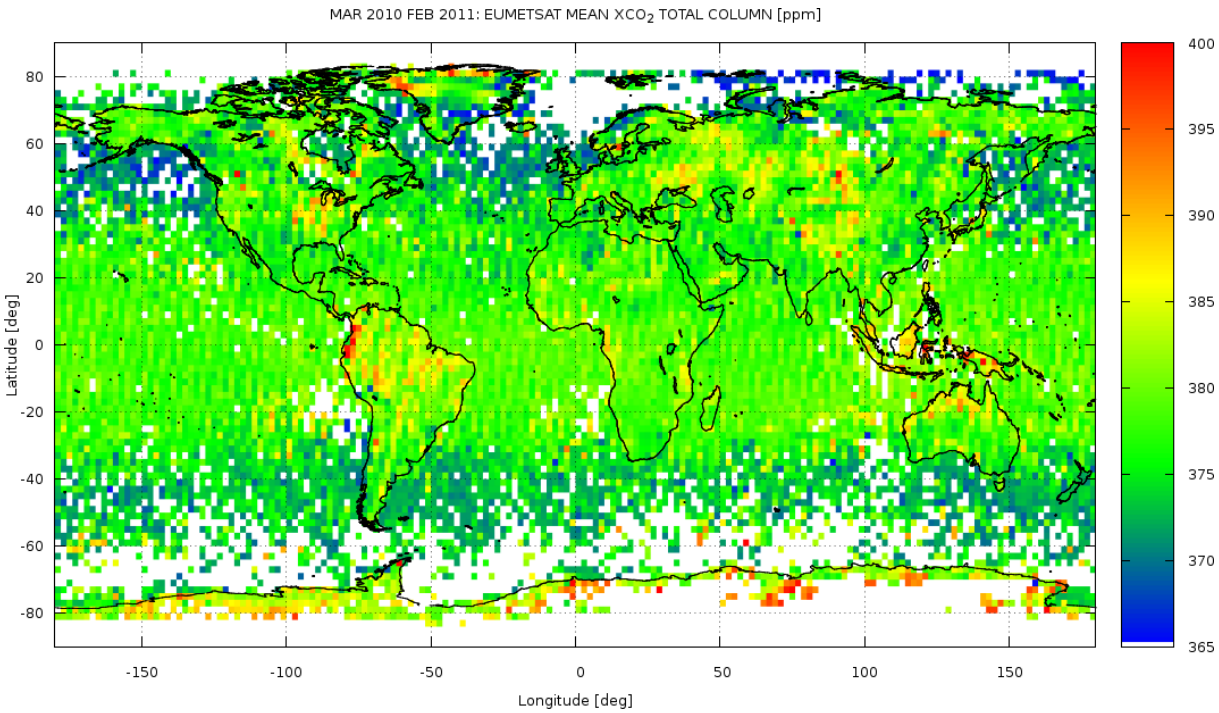

Fig. 6.2: *MARCH 2010-FEBRUARY 2011: EUMETSAT L2 XCO_2 total column [ppm] over a grid of 2° x 2°, averaged for one year*

An example of the IASI L2 product provided by EUMETSAT and used to evaluate the XCO_2 is reported in Figure 6.2: the map shows the XCO_2 values [ppm] averaged for the year from March 2010 to February 2011 for a global geographical coverage over a grid of 2° x 2°. This yearly map is directly comparable with the KLIMA L2 map presented in Figure 5.1 and it is clear that EUMETSAT retrieved a smaller variability than KLIMA. Moreover, this figure highlights larger differences between pixels over land and over ocean compared to KLIMA at latitudes higher than ±45° in the Northern Hemisphere and in the Southern Hemisphere, respectively. In the equatorial belt (at latitudes between -15° and 15°), higher values of XCO_2 (above 385 ppm) are found by KLIMA, whilst EUMETSAT map mostly shows an homogeneous distribution below that threshold in this latitude region, especially over Africa. Similar behaviours is evident over China and Eastern coast of the North America where EUMETSAT values are lower than about 20 - 30 ppm with respect to KLIMA ones.

6.3.2 Results of the comparison

In the Figures from 6.3 to 6.14 are reported, for each month, the maps of the percentage difference between XCO_2 values retrieved by using the KLIMA code and the L2 products delivered by EUMETSAT. The differences are calculated between average values represented in the same 2° x 2° grids used for KLIMA results. The error associated with these differences consists of the KLIMA retrievals error (third map from Figure 5.3 to Figure 5.14) plus the unknown errors of EUMETSAT. From March to September it is possible to see that in Tropical latitudes, both over land and water, EUMETSAT L2 XCO_2 values are higher than KLIMA L2, with a difference of approximately 4-5%; while in Northern and Southern mid and high latitude situation reversed: KLIMA XCO_2 values are higher than EUMETSAT. From October, with the exception of North Africa, Arabia and a small part of continental China, KLIMA values are higher that EUMETSAT, for a maximum difference of 5%. These percent differences, that correspond to about 15-20 ppm, are much larger than the retrieval errors of KLIMA, which typically are of the order of 2 ppm over land, 4 ppm over water and 8 ppm at high latitudes. It is a fact that the global yearly map shows that KLIMA results have a larger dispersion, but EUMETSAT results are likely to have a larger error because based on a subset of spectral points.

In Figure 6.15, it is reported the seasonal cycle of XCO_2 from March 2010 to February 2011, for the Northern and the Southern Hemisphere. In the two figures the green line represents the annual variation of XCO_2 given by EUMETSAT while the red line shows the values retrieved with KLIMA and the green line the values. KLIMA is observing a seasonal periodic variation in the Northern Hemisphere and a small positive trend in the Southern Hemisphere. Quite a different result is observed in the EUMETSAT data where an unrealistic negative trend is present in both the Northern and Southern Hemisphere seasonal cycles. As reported in 4.1.3, on 14 September 2010 a new version of the processor was released and this change in the algorithm could explain the discontinuity, relative to October, that we can observe in the seasonal variation of IASI L2 data.

Figure 6.16 shows the annual average related to the selected macro areas (Figure 6.1), as obtained from KLIMA (red points) and EUMETSAT (green points). The standard deviation of the mean is also shown: the spread of EUMETSAT data is usually smaller than that of KLIMA results. The number of observations for each macro area is also reported, to provide an estimate of the populations of each macro area. This type of plot was made with the intent to show the capability of KLIMA code to highlight the presence of sources and sinks. The macro area of 10° x 10° met the spatial requirement of the project, while the annual average refers to a longer period so as to reduce seasonal fluctuations that have already been shown and discussed in the previous Figures. KLIMA results highlight several of the expected geographical features, such as the higher values over countries characterized by high CO_2 emission (like China, India and North America) and the reduced values of XCO_2 over Oceans in Southern Hemisphere. On the other hand, EUMETSAT results provide nearly constant values over Oceans and over northern and southern continents (for example, there is no change in averaged amounts over North and South America or North and South Africa). EUMETSAT results differ from KLIMA averages by variable quantities, with differences that span from 2-3 to 15-16 ppm. The greater difference of $\approx$ 15 ppm is observed over China, where KLIMA gives a mean value of about 390 ppm while EUMETSAT provides an averaged value of 375 ppm, that is smaller than values found over Australia.

In summary, it is possible to conclude that there is no evident correlation between KLIMA and EUMETSAT results, even if obtained from the same dataset. Despite the fact that, as reported in 6.5, KLIMA results show a larger variability with respect to the spread of the dataset, however, this larger variability is often in the direction of the expected differences and seems to be more realistic than the uniform EUMETSAT results.

The comparison between KLIMA and EUMETSAT dataset is difficult because straightforward correlations do not emerge and the absence of an error estimate for EUMETSAT results prevents a clear conclusions.

6.4 Comparison with TCCON products Level 2 products

The Total Carbon Column Observing Network (TCCON) is a network of ground-based FTS that provides accurate and precise column-averaged abundances of XCO_2 [95]. For this reason, the TCCON products are used as a reference for validation of products from space-borne measurements. To perform the comparison between KLIMA L2 data and TCCON measurements, we selected nine sites covering different latitude and longitude ranges.

In the case of TCCON stations the procedure of smoothing described in 6.2 can not be applied because the needed information are not delivered together with the standard operational products. Nevertheless, some complementary indication about consistency and stability of KLIMA L2 retrieved values can be obtained also from the comparison with TCCON results without smoothing.

6.4.1 TCCON products overview

The TCCON ground stations considered for the comparison are reported in Table 6.1. These stations were chosen for the comparison in order to cover different latitude and longitude range. We also checked for data availability in our selected year. The entire dataset has been downloaded from the web server

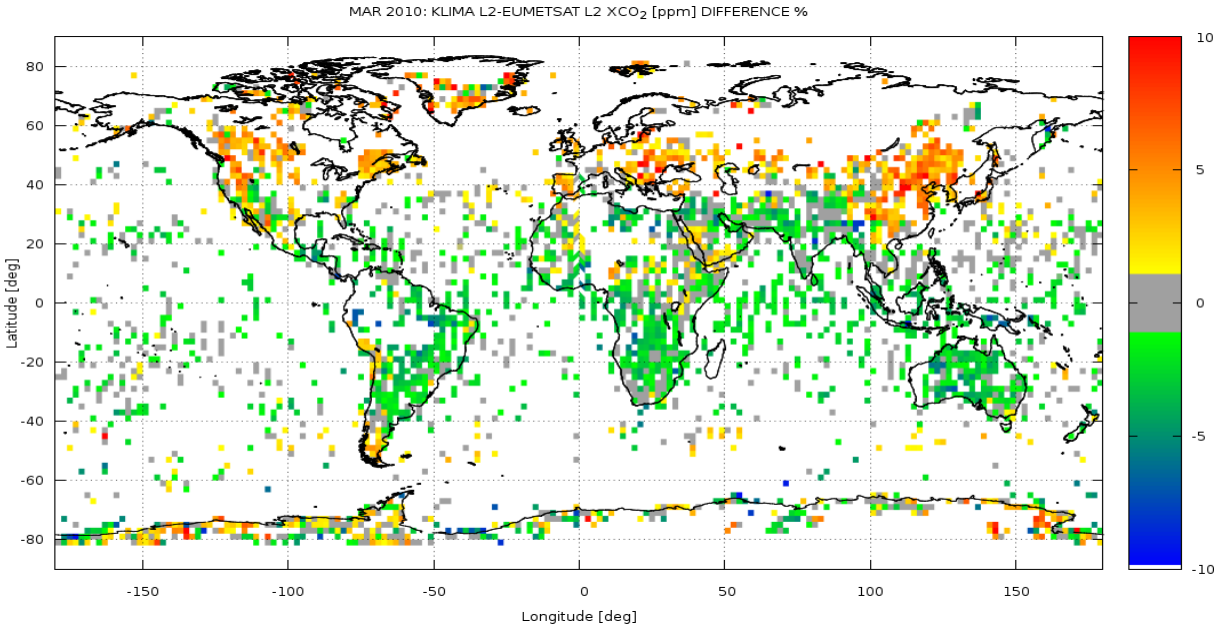

Fig. 6.3: *8-14 MARCH 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

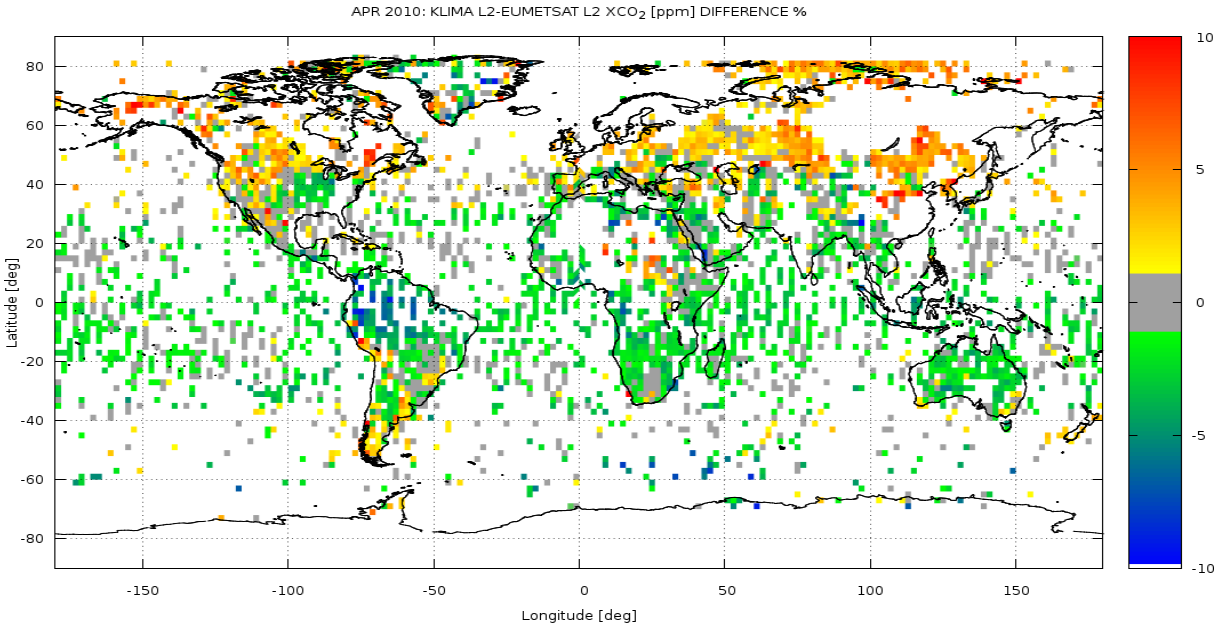

Fig. 6.4: *5-11 APRIL 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

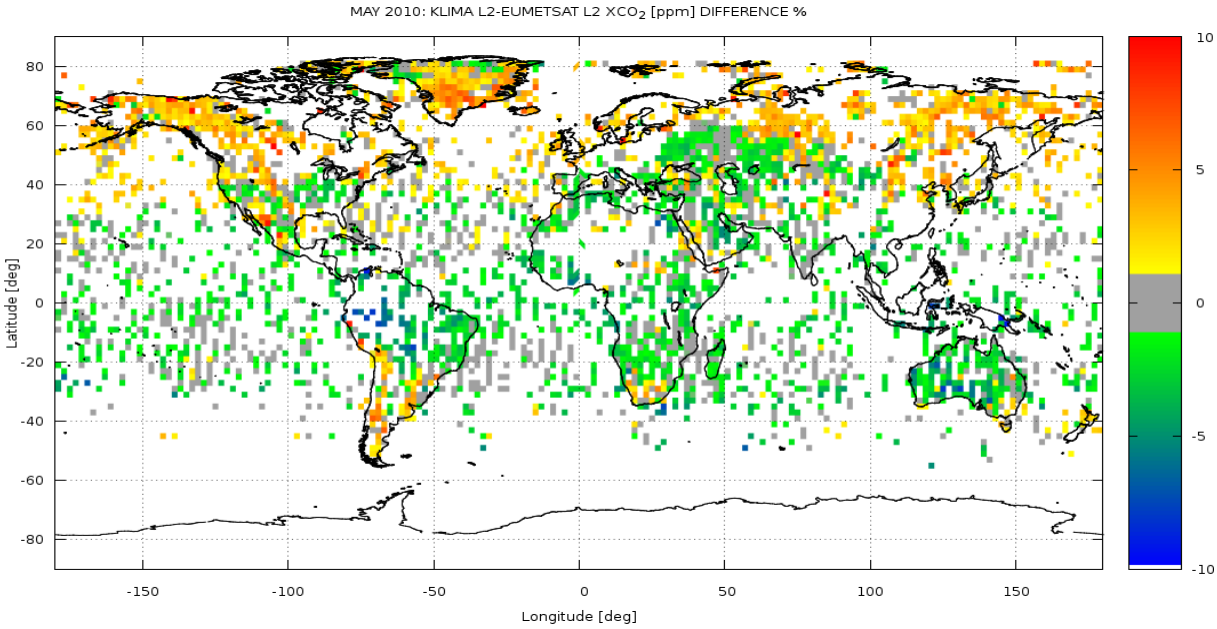

Fig. 6.5: *3-9 MAY 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

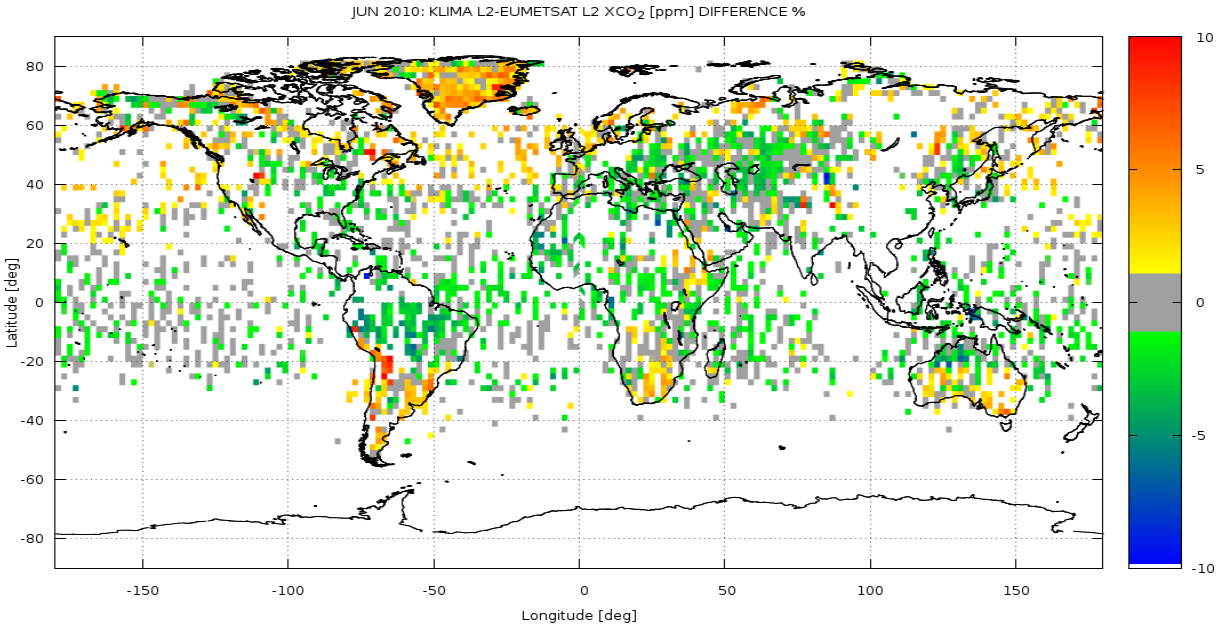

Fig. 6.6: *3-9 JUNE 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

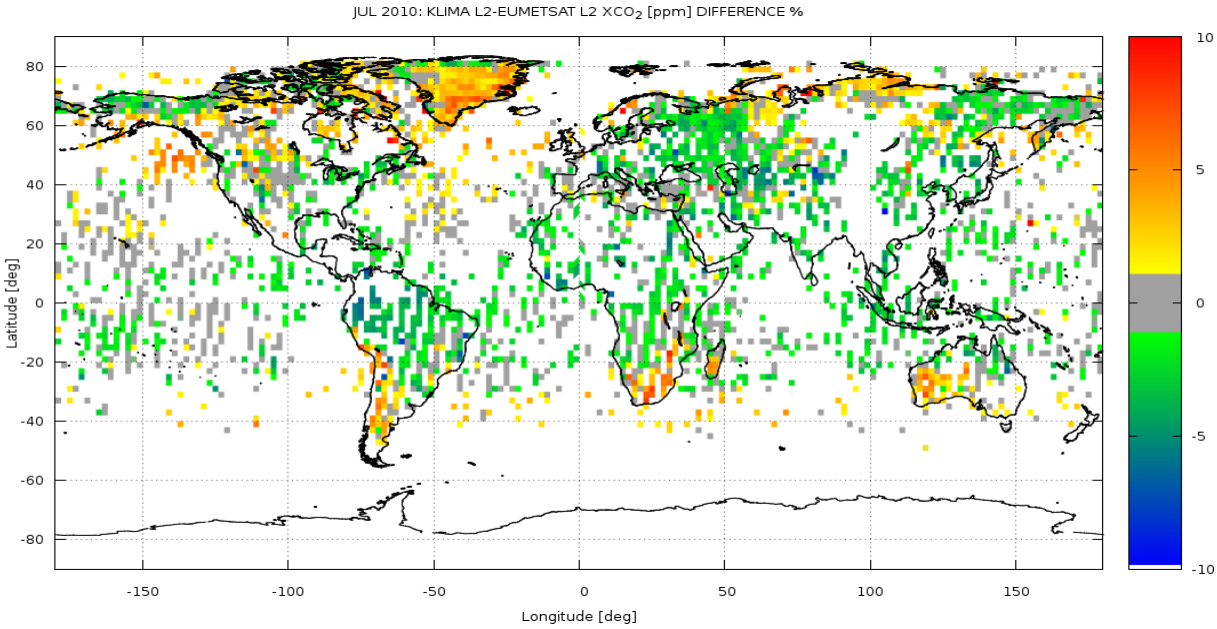

Fig. 6.7: *5-11 JULY 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

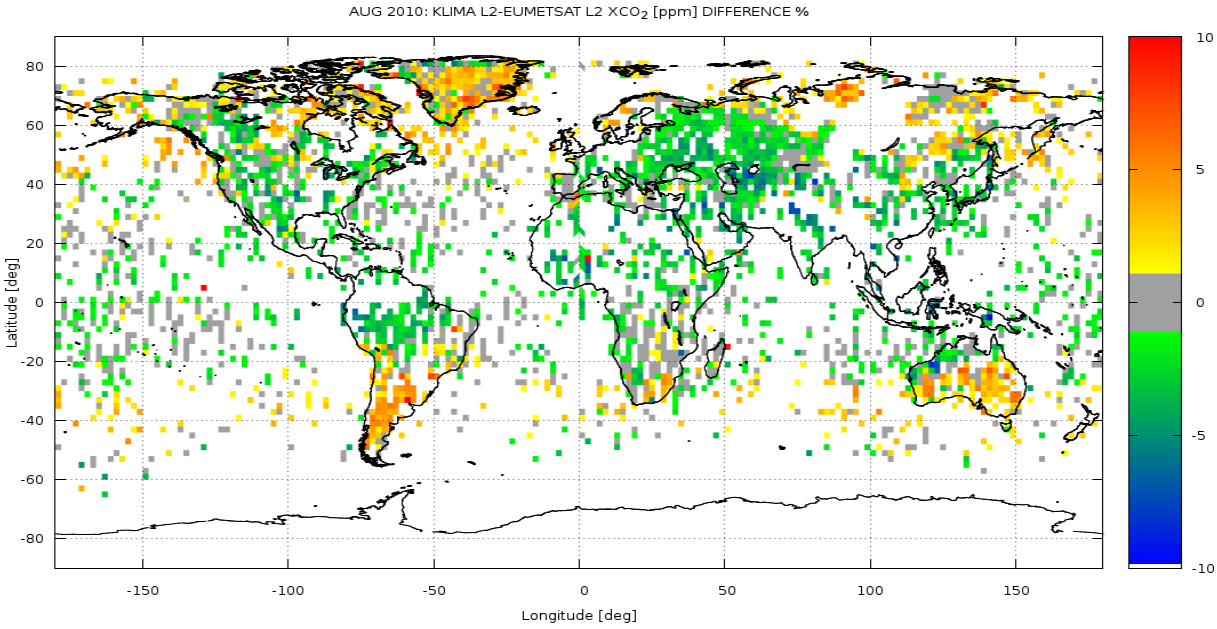

Fig. 6.8: *2-8 AUGUST 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

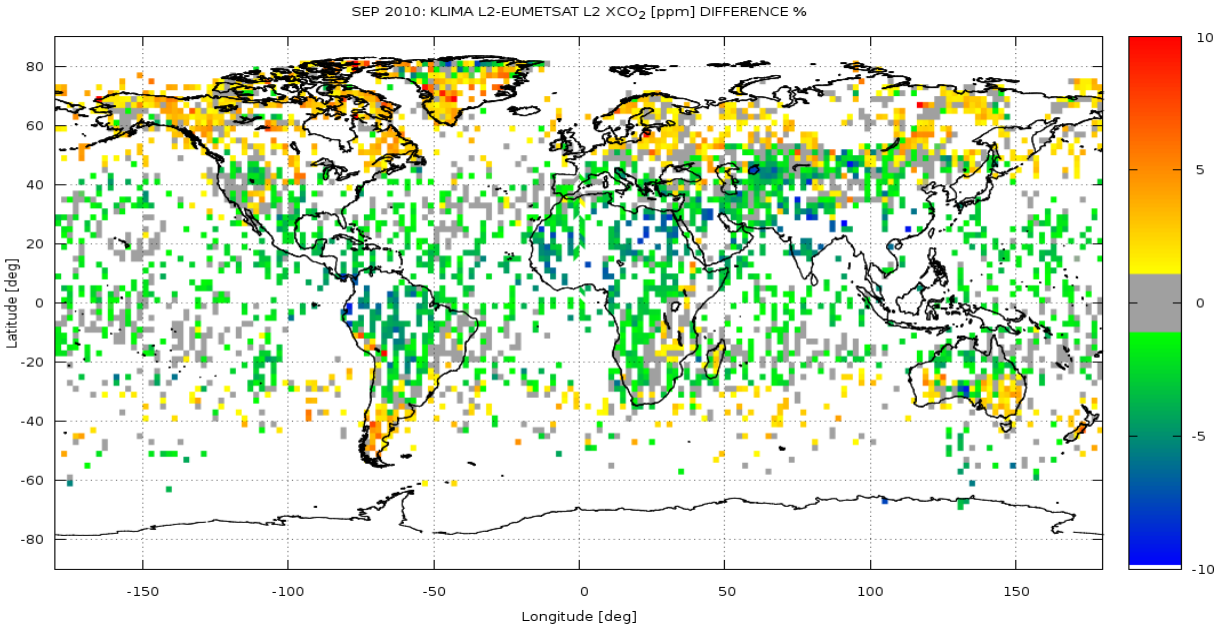

Fig. 6.9: *6-12 SEPTEMBER 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

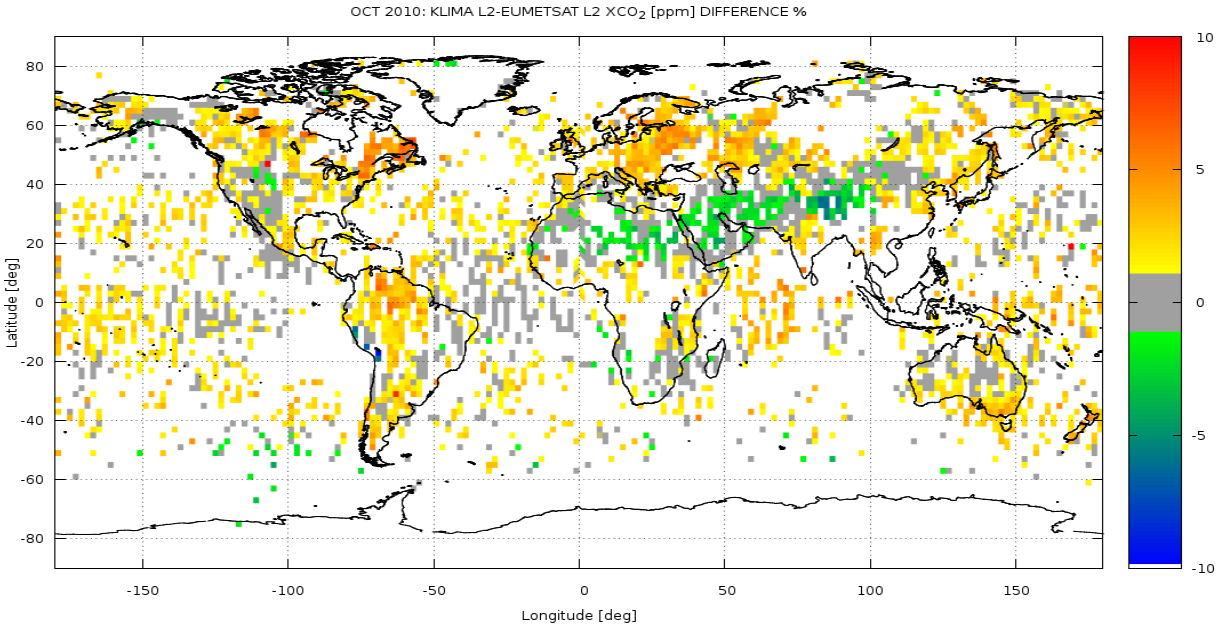

Fig. 6.10: *4-10 OCTOBER 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

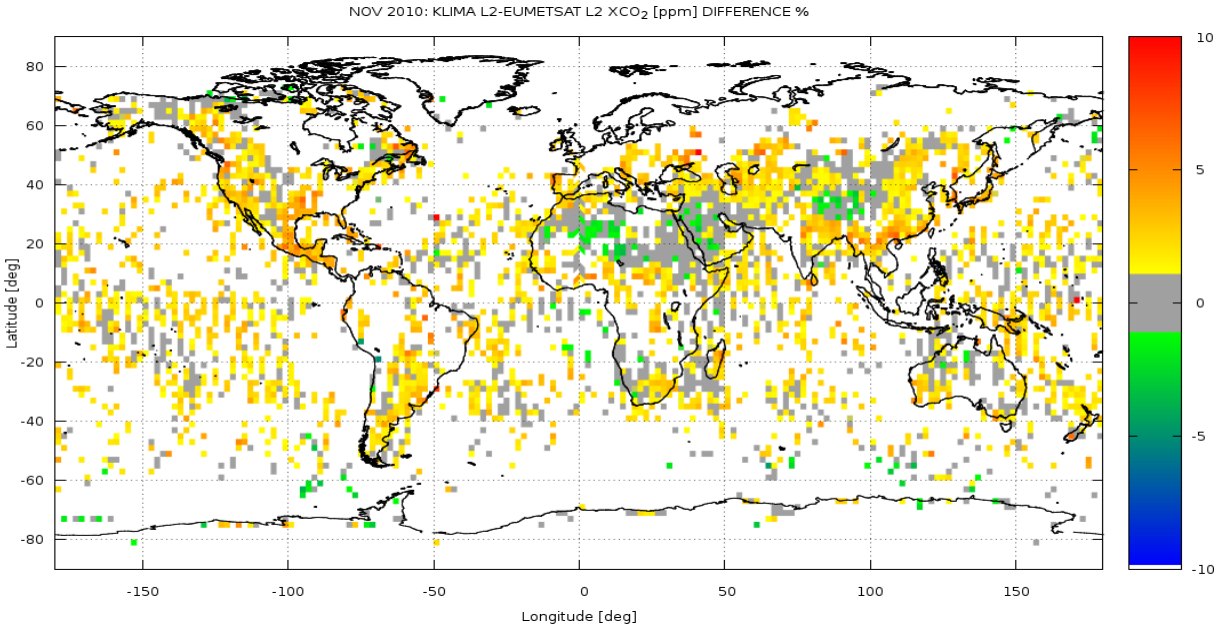

Fig. 6.11: *8-14 NOVEMBER 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

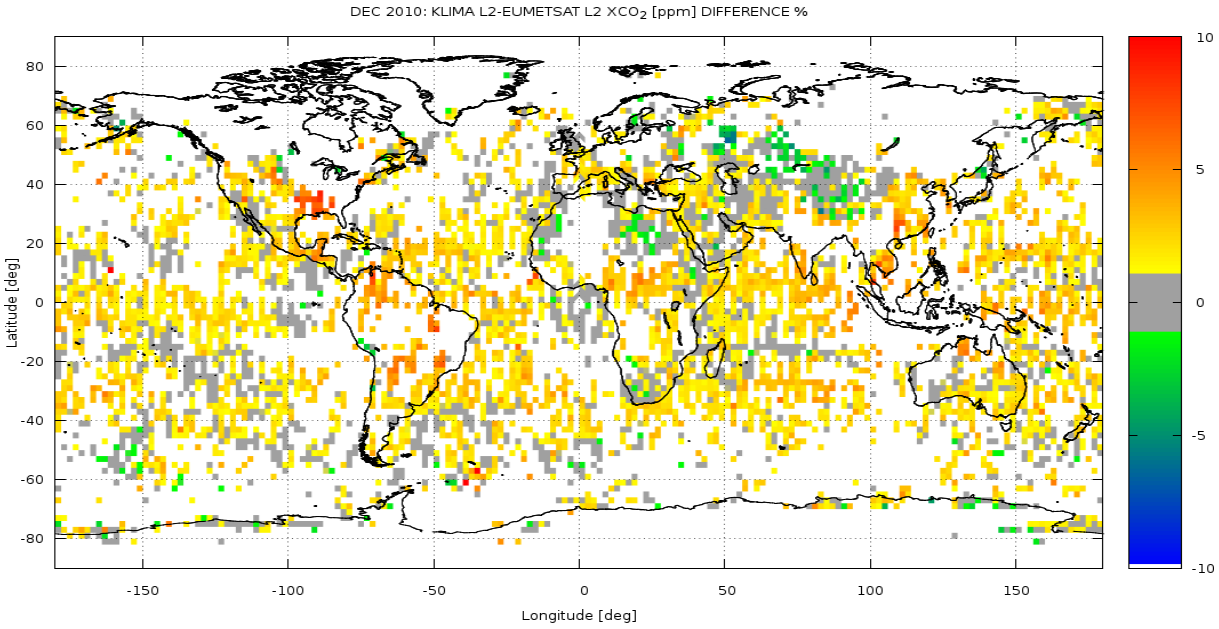

Fig. 6.12: *8-14 DECEMBER 2010: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

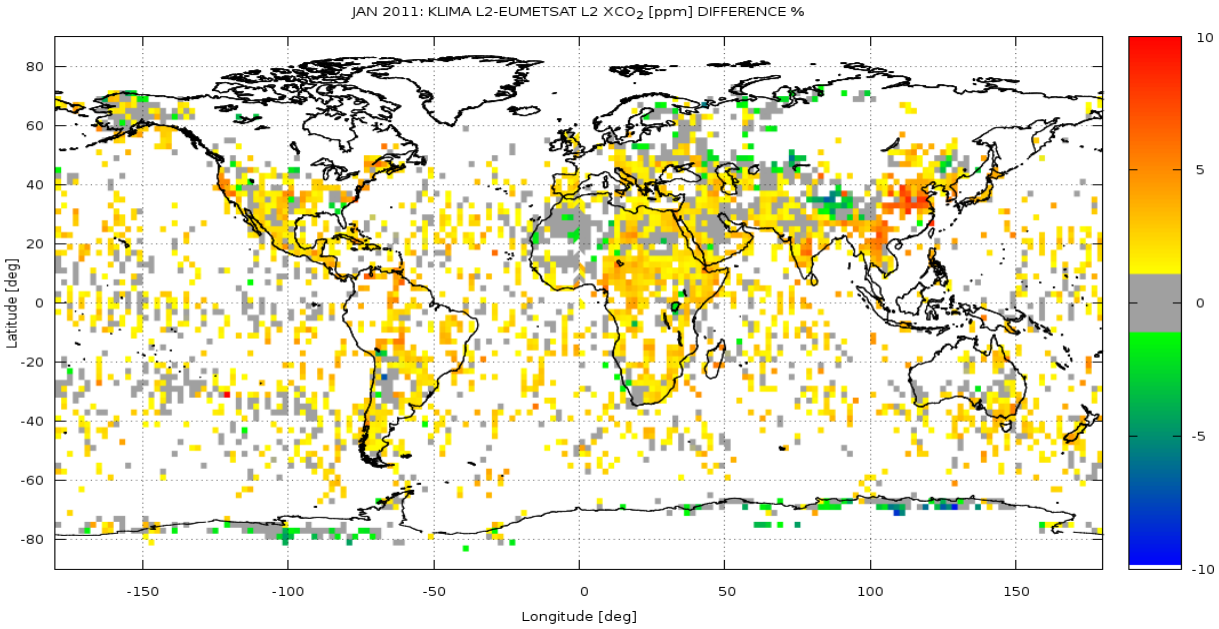

Fig. 6.13: *3-9 JANUARY 2011: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

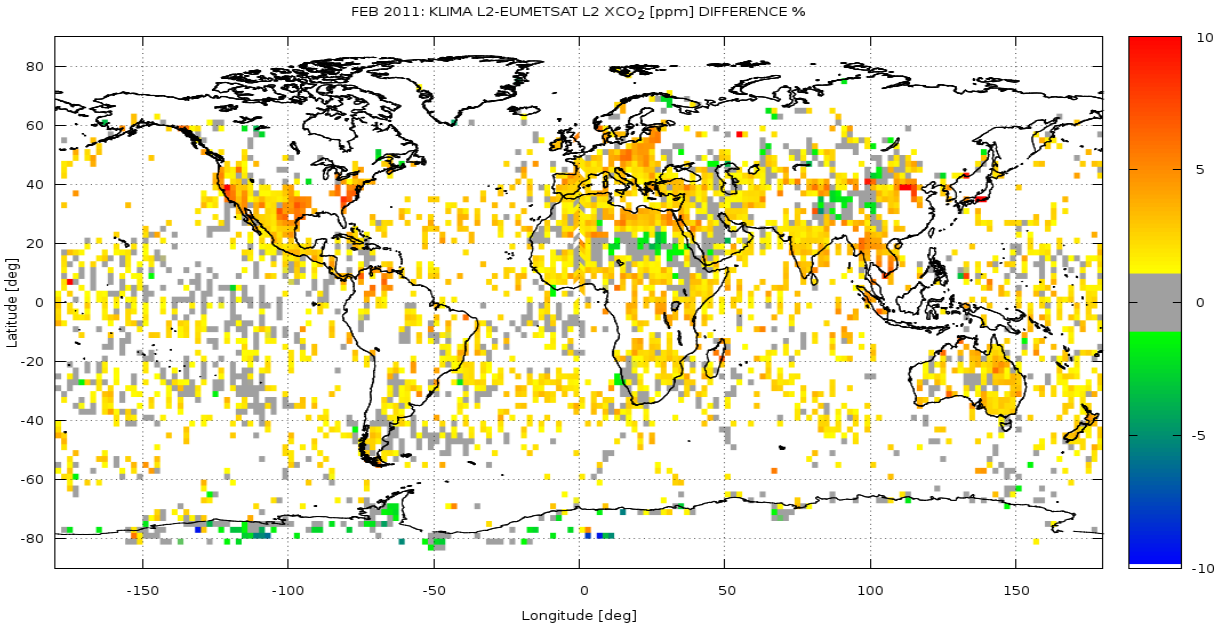

Fig. 6.14: *7-13 FEBRUARY 2011: KLIMA-EUMETSAT XCO_2 comparison: map of % difference*

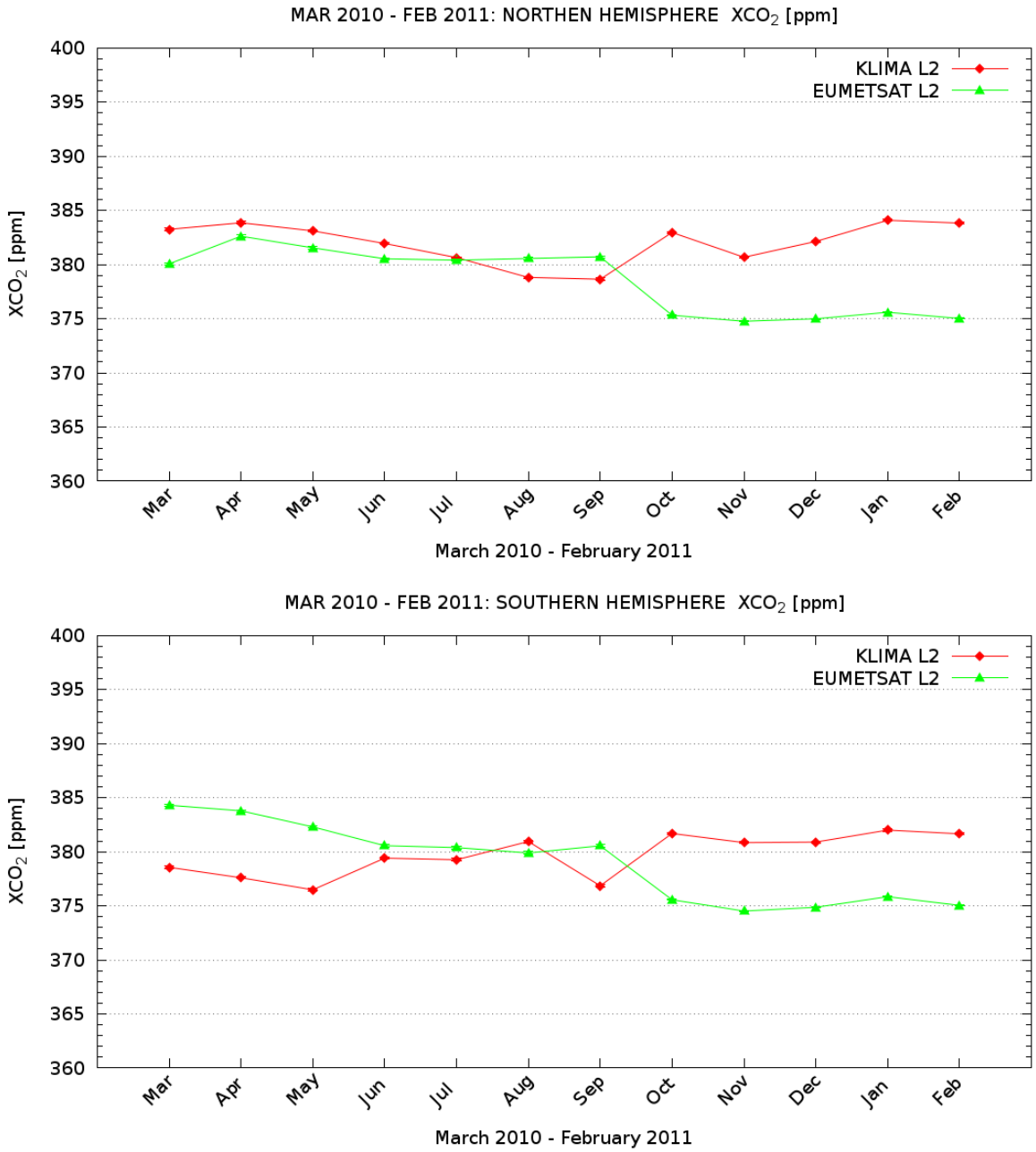

Fig. 6.15: *Seasonal variation of the XCO_2 from March 2010 to February 2011 for Northern Hemisphere (top) and Souther Hemisphere (bottom). The average on the Hemispheres of the XCO_2 retrieved by KLIMA (red) is compared with operational L2 products provided by EUMETSAT (green)*

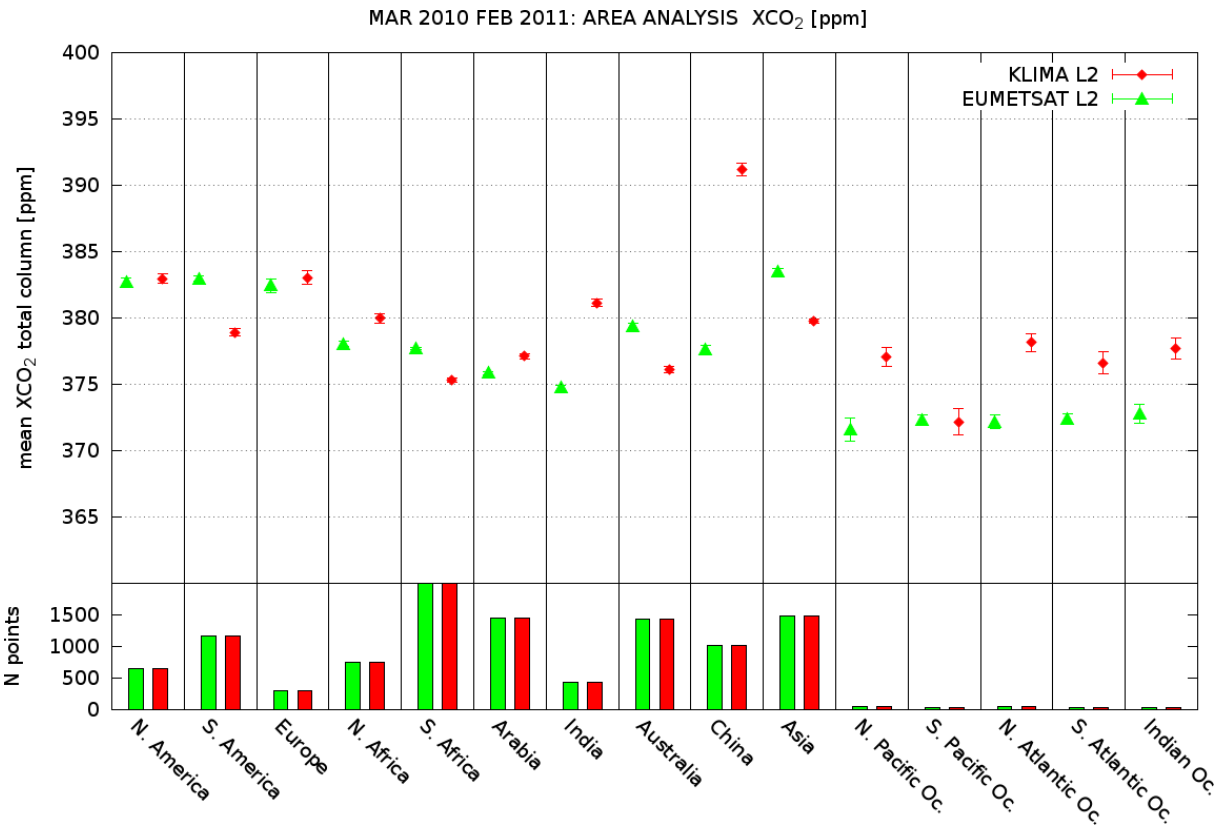

Fig. 6.16: *Top: Average XCO_2 for each selected macro area (Figure 6.1). Red points are KLIMA average XCO_2 green points are EUMETSAT products Bottom: Number of observations for each macro area*

[85]. The TCCON dataset related to the stations used in the comparison in the selected annual range is reported in Figures 6.17, 6.18, and 6.19 where the red symbols represent the TCCON measurements.

6.4.2 Results of the comparison

The comparison with TCCON data has been performed using the following coincidence criteria:

- Distance from TCCON site less then 200 km and measurements acquired within one hour.
- Distance from TCCON site less then 20 km and measurements acquired within one day.
- Distance from TCCON site less then 50 km and measurements acquired within one day.

The first one imposes a very strong constrain in the time distance between the measurements of IASI and TCCON station while the last two criteria are not so strong in time, but impose a strong constrain in term of distance between the footprint of the observations.

For selected TCCON stations a dedicated analysis has been performed using the G-POD resources. In particular for Lamont, Lauder, Park Falls, Sodankyla, and Wollongong the analysis of TCCON measurements are performed on the whole period from March 2010 to February 2011 while for the other stations we considered only the measurements related to the selected weeks of each analysed month.

Both IASI and TCCON measurements that meet the coincidence criteria have been averaged so we obtain a single value for each day in which the coincidences has been found. The time series relative to each TCCON site have been reported in Figures 6.17, 6.18, and 6.19 for each of the different coincident criteria respectively. In these plots the complete time series of TCCON XCO_2 (red points) is compared with the average IASI coincident measurements (blue points).

As a first step we adopted the coincidence criteria of measurements acquired within one hour with a spatial constraints of 200 Km, even if the time constraint of one hour is not so critical to establish a

Tab. 6.1: *Ground-based TCCON sites used for KLIMA L2 product comparison*

Site	Country	Coordinates [Lat., Long.]
Bialystok	Poland	53.23° N, 23.03° E
Garmisch	Germany	47.48° N, 11.06° E
Lamont	USA	36.60° N, 97.49° W
Lauder	New Zealand	45.04° S, 169.68° E
Ny Alesund	Island	78.92° N, 11.90° E
Orleans	France	47.97° N, 2.11° E
Park Falls	USA	45.94° N, 90.27° W
Sodankyla	Finland	67.37° N, 26.63° E
Wollongong	Australia	34.48° S, 150.79° E

coincident crititerium and the cases of one day coincidences have been also investigated. The spatial constraint is critical in this analysis because the statistics is significantly reduced when the pixel dimension used for the comparison decreases, as is underlined in Figure 6.18 when we consider the distance from TCCON site less then 20 km.

TCCON measurements show a small daily variability with respect to the KLIMA XCO_2 products. Moreover, KLIMA data are on average lower than the reference data of the TCCON stations. In the case of the comparison with Wollongong, Lamont and Sodankyla stations within 200 km we have a most populated KLIMA dataset with respect to the other stations and we can see that, especially for the comparison with Wollongong, within the retrieval error, the available KLIMA observations, when corrected for an offset, seem to reproduce well the reference curve. Also in the case of Sodankyla KLIMA dataset seems to be in a discrete agreement with TCCON dataset. The worst agreement are evident in the comparison with Lamont, where we can see a considerable negative bias of KLIMA series.

A fraction of these differences can be explained by mean of the effect of the a priori and of the AKs related to the different analysis.

6.5 Comparison with TANSO-FTS/GOSAT Level 2 products

The inter-comparison of the XCO_2 total column retrieved by using KLIMA code with TANSO-FTS SWIR L2 product is one of the goals of the project. The comparison activity has been performed in the annual range from March 2010 to February 2011 both for land and water observations and for global geographical coverage; however the period of the analysis is restricted to only one week per month since only one week per month of IASI observations has been processed, as described in 5.1. TANSO-FTS L2 operational products have been downloaded by accessing the GOSAT User Interface Gateway (GUIG) for the same period and for the same weeks. The collected TANSO-FTS L2 dataset used for the comparison presents global coverage but it is limited to day observations. Moreover, the points with the ratio between the standard deviation of the retrieved values and their mean retrieval error larger than 2 are not included in order to skip in the comparison the pixels where a large geographical and/or a temporal variability was observed.

As reported in 4.2, in the analysed period TANSO-FTS products show a good global geographical coverage but a poor statistic in terms of number of observations. After applied selection criteria 5.1, KLIMA provides about 240000 XCO_2 observations from March 2010 to February 2011 while in the same period only about 3000 observations from TANSO-FTS SWIR are available. The reduced dataset for TANSO-FTS with respect to KLIMA requires the definition of suitable strategies in order to compare the two products. Three different strategies have been used for this purpose:

- *co-located comparison*: consists in the comparison of the XCO_2 total column retrieved from observations of IASI and TANSO-FTS made in contiguous locations in time and space;

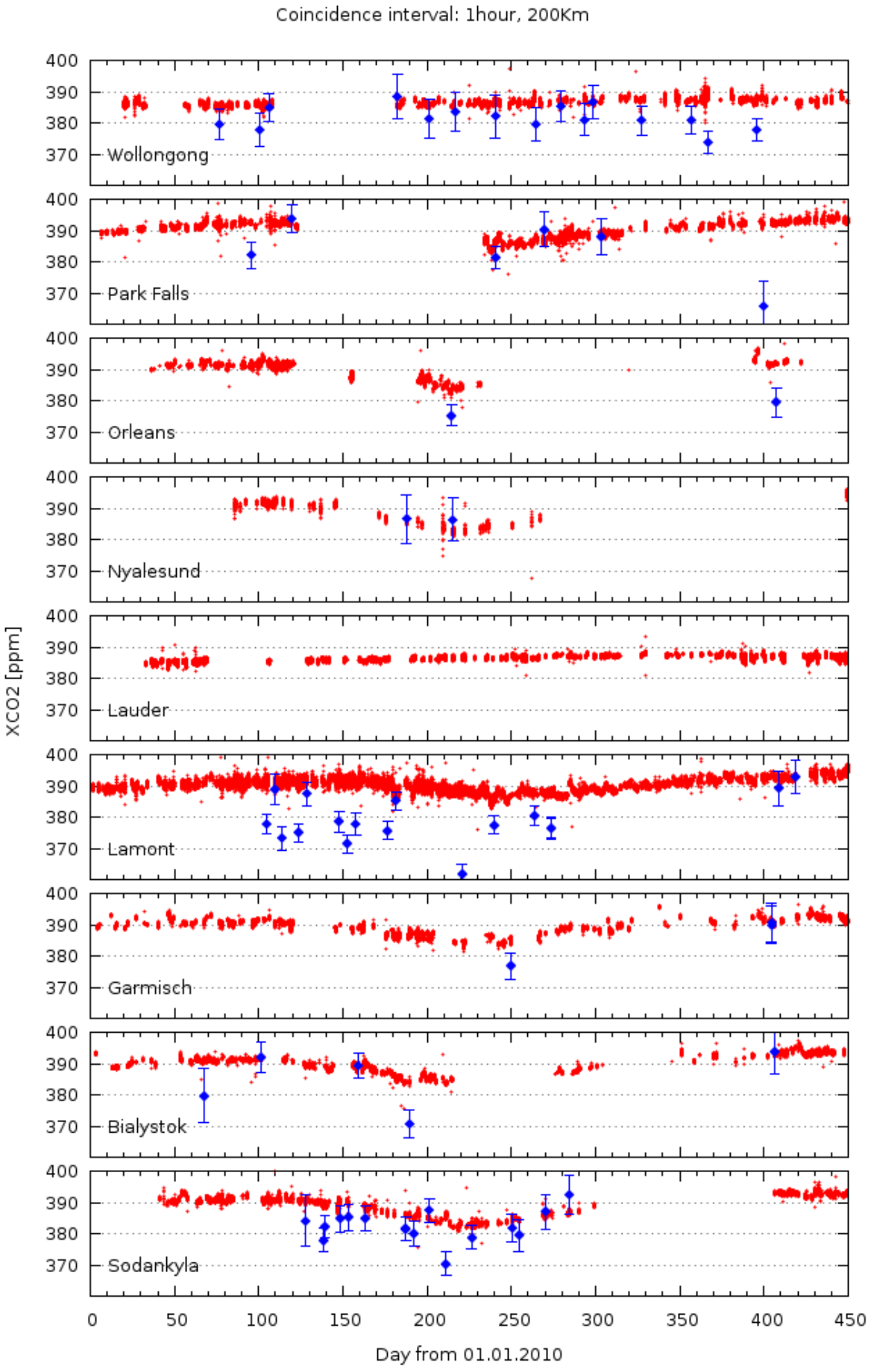

Fig. 6.17: *March 2010 - February 2011: time series of TCCON (red points) and averaged KLIMA IASI L2 (blue points) XCO_2 [ppm] within 200 km and 1 hour*

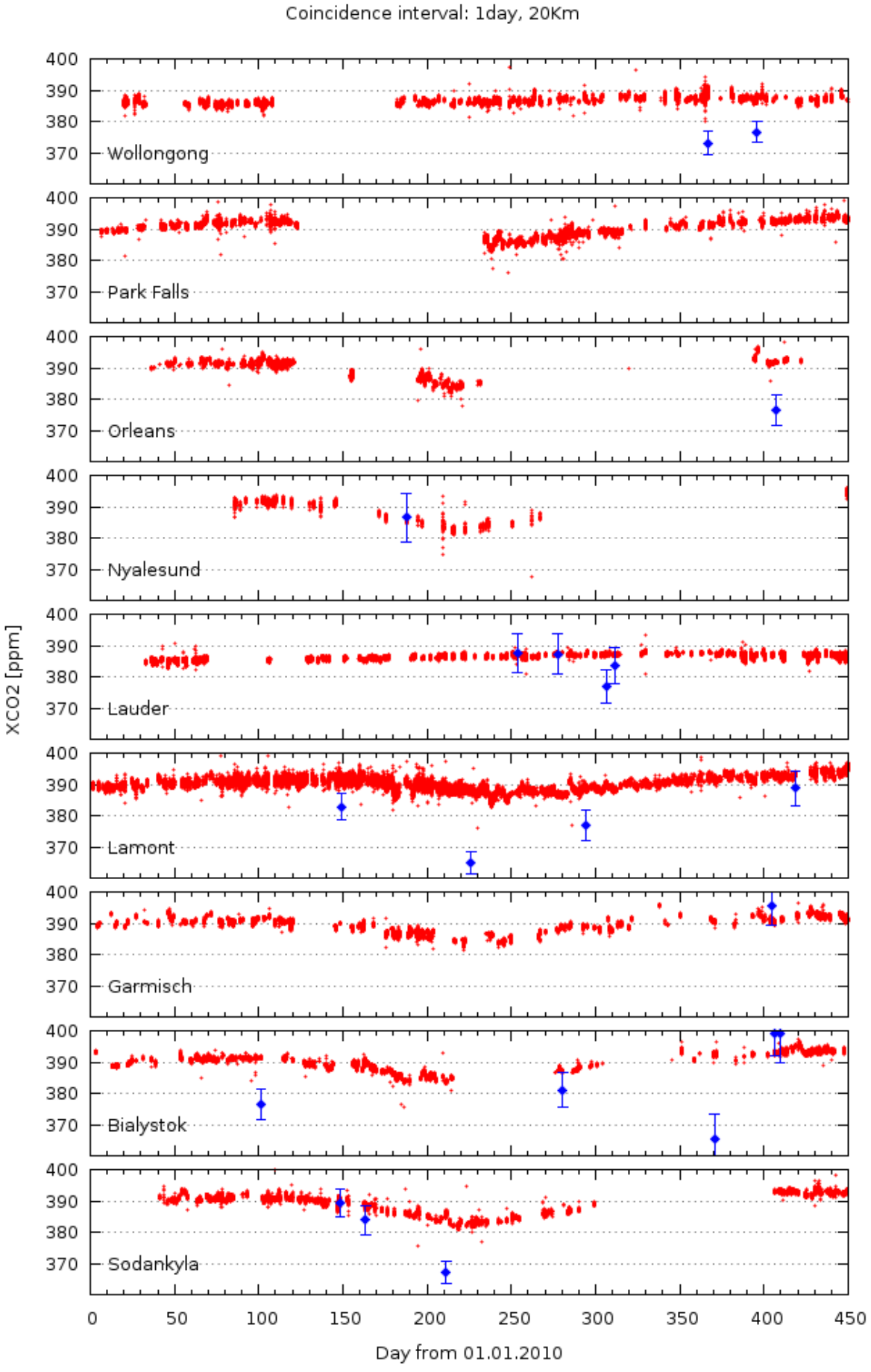

Fig. 6.18: *March 2010 - February 2011: time series of TCCON (red points) and averaged KLIMA IASI L2 (blue points) XCO_2 [ppm] within 20 km and 1 day*

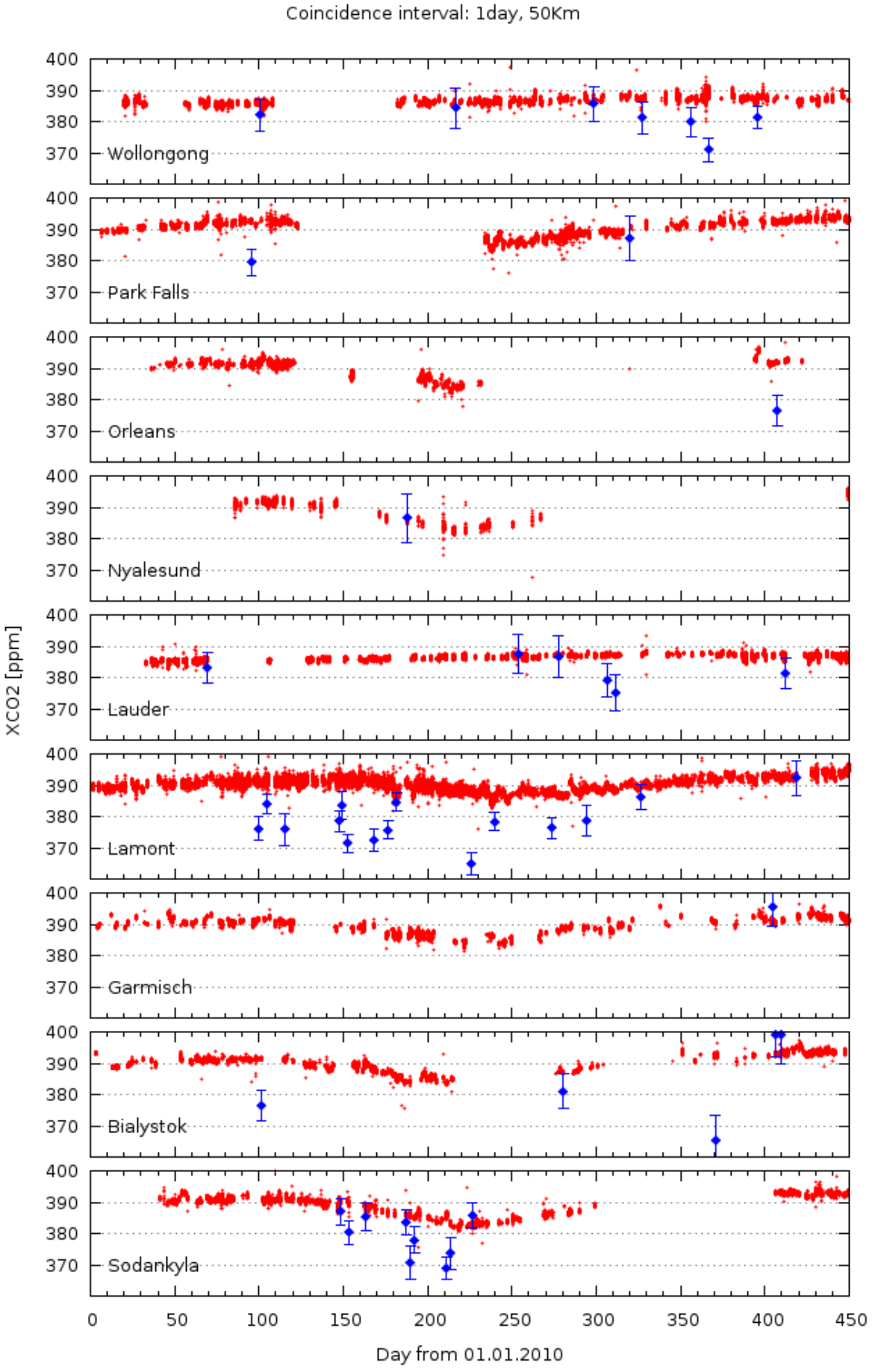

Fig. 6.19: *March 2010 - February 2011: time series of TCCON (red points) and averaged KLIMA IASI L2 (blue points) XCO_2 [ppm] within 50 km and 1 day*

- *averaged comparison*: consists in the comparison of the XCO_2 total column averaged on a suitable spatial and time interval;

- *seasonal-variation comparison*: consists in the comparison of the seasonal variations on macro-areas (Northern/Southern Hemispheres, Pacific Area, some latitude bands).

Furthermore, IASI and TANSO-FTS measurements show a different sensitivity as a function of the vertical profile of the atmosphere. A quantitative comparison of the two XCO_2 products must take into account the different AKs of the two instruments. For this reason, the results presented in the following sections, further than with the direct comparison of IASI and TANSO-FTS retrieved products, report also a cross-comparison after introducing, on TANSO-FTS XCO_2 values, the smoothing effects described in 6.2. This operation of smoothing takes into account the difference of the AKs associated with KLIMA and TANSO-FTS retrievals and the different a priori CO_2 atmospheric profiles. In the case of the direct comparison we are comparing the same climatological product obtained by two different instruments that have different vertical sensitivity. TANSO-FTS is characterized by an almost uniform sensitivity with the altitude and mainly receives information from the lower troposphere while IASI has a maximum of sensitivity in the mid to upper troposphere and poor information in lower troposphere. The smoothing applied to the TANSO-FTS XCO_2 products have the effect to bring back its vertical coverage to the coverage of IASI, allowing a quantitative comparison of the two products.

6.5.1 TANSO-FTS/GOSAT Product Overview

At the beginning of the project, the TANSO-FTS SWIR products version available from GUIG was Version 1 (V1) and first comparisons were made with this first version. However, as reported in [61], the TANSO-FTS SWIR V1 has been found biased low by 8.85 ppm with a standard deviation of 4.75 ppm ([90]) compared with data of the TCCON network. Subsequently, National Institute for Environment Studies (NIES) corrected the TANSO-FTS retrieval algorithm and released a new version of data, Version 2 (V2), which have been corrected for this bias. The improved retrieval algorithm shows much smaller biases and standard deviations, 1.48 and 2.09 ppm for XCO_2 than the version 01.xx ([90]). After this improvement, we decided to use the latest version of TANSO-FTS products and to repeat the comparison with the new dataset:

- FTS SWIR L2 V2.00

- FTS SWIR L2 V2.10

- FTS SWIR L2 V2.11

Only the V2 of the TANSO-FTS/GOSAT products are used for the cross-validation activity. The results of the analysis performed using the original V1 dataset are summarized in Appendix A, in order to provide a complete report of the activity.

An example of the TANSO-FTS/GOSAT products is reported in map of Figure 6.20: map shows the XCO_2 total column for the selected weeks each month in the year from March 2010 to February 2011, averaged over a grid of 2° x 2°. This map is directly comparable with the map obtained by KLIMA and shown in Figure 5.1.

Figure 6.21 shows, for the same coverage considered in the previous Figure, the number of TANSO-FTS XCO_2 measurements that have been averaged. TANSO-FTS has very few measurements at high latitudes (where also the sensitivity of IASI measurements is reduced) and over Oceans, and makes less frequent measurements than IASI. Moreover, measurements are acquired only at day time.

Also in the case of TANSO-FTS products it is possible to observe a variability that is smaller than the one observed by KLIMA. Furthermore, the comparison with GOSAT dataset clearly highlights that, on average, TANSO-FTS provides higher values of XCO_2 total columns than KLIMA. KLIMA product shows high values of XCO_2 in the equatorial belt; this feature is particularly enhanced in TANSO-FTS retrieval products, with XCO_2 annual mean values in the selected period exceeding 390 ppm around the

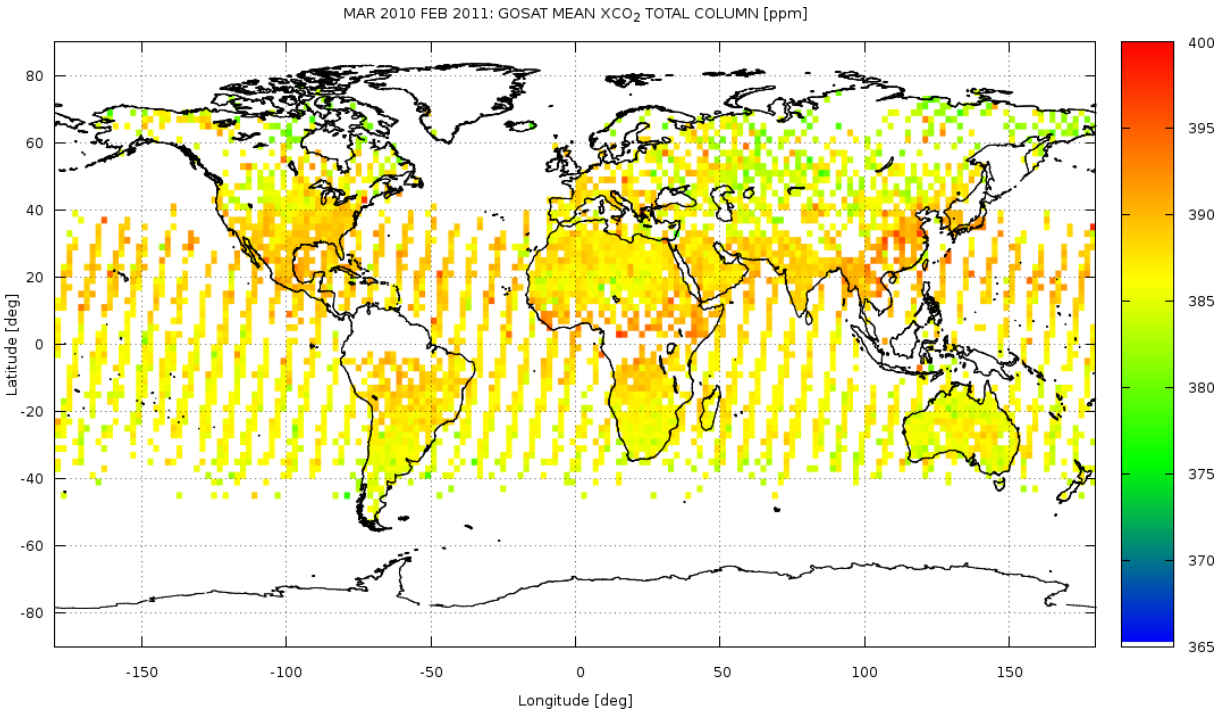

Fig. 6.20: *March 2010 - February 2011: TANSO-FTS L2 XCO_2 total column [ppm] over a grid of 2° x 2°, averaged for one year. TANSO-FTS products are limited to the same selected weeks per month of the KLIMA dataset*

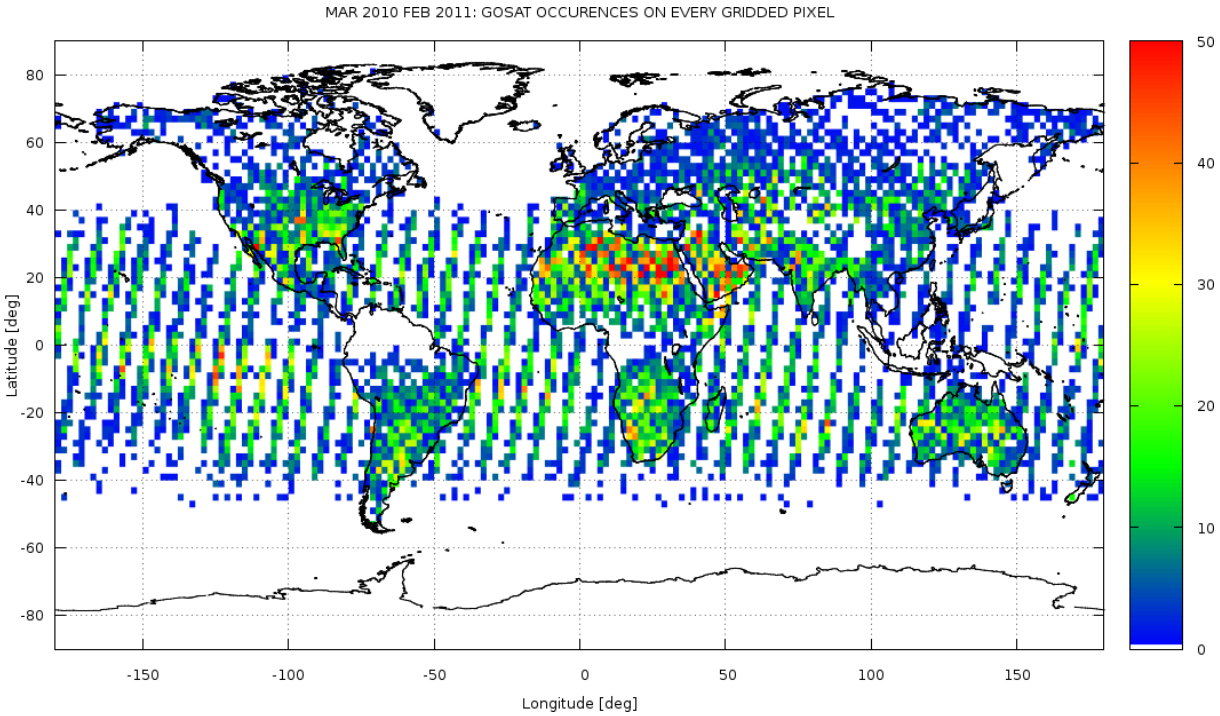

Fig. 6.21: *Map of the TANSO-FTS coverage (in term of number of observations per pixel) for the year from March 2010 to February 2011 over a grid of of 2° x 2°*

Equator. Large values of XCO_2 are also found by TANSO-FTS in South-East Asia and Central America and are consistent with the average distribution obtained from KLIMA.

6.5.2 Results of the comparison

Co-located comparison

The *co-located* strategy consisted in the comparison of the XCO_2 total column retrieved from single measurements of TANSO-FTS and IASI, acquired in the same day and observing the same scene. In particular, the IASI observation has been assumed as *co-located* with the TANSO-FTS when acquired in the same day and if the centre of the IASI ground pixel is located at a distance of less than 20 km from TANSO-FTS one (see Figure 6.22).

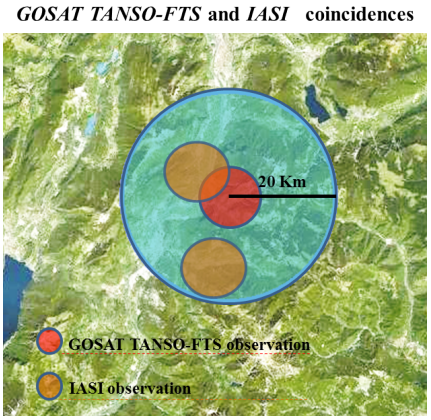

Fig. 6.22: *IASI and TANSO-FTS observations have been assumed as co-located if the centre of the IASI ground pixel is located in a 20 Km circular interval from TANSO-FTS one*

The data selection criterion has been chosen as a trade off between the need to compare quantities retrieved from measurements observing the same scene and the need to have a significant number of samples, suitable for an annual statistics. Measurements executed in the same day (and for the same location) corresponds to the minimum time distance between IASI and TANSO-FTS observations while the spatial interval of 20 Km from the TANSO-FTS centre of pixel can be considered small enough to neglect horizontal XCO_2 variability. The result of the selection is a subset of 519 pairs of XCO_2 total column values for the year from March 2010 to February 2011.

The figures from Figure 6.23 to 6.25 show the correlation between the XCO_2 total column retrieved using KLIMA algorithm and the corresponding value provided by TANSO-FTS V2 operational XCO_2 product, during the period from March 2010 to February 2011. Two different cases are considered: observations over water (displayed in Figure 6.23) and observations over land (displayed in Figure 6.24); the correlation of all co-located observations (both land and water cases) is then shown in Figure 6.25.

Referring to the scatter plots provided in Figures from 6.23 to 6.25, we can see that the XCO_2 total columns retrieved by KLIMA, both in the case of observations acquired over land and over water, are smaller than the corresponding TANSO-FTS values. These differences are larger in the case of measurements over water, even if in this case a limited number of events has been analysed. The range of variability of the datasets is comparable with the relative dispersion of the two observations, leading to an almost uniform distribution of the cloud of points.

It is possible, however, to make some considerations about the distribution and highly populated case of Figure 6.24. For measurements acquired over land the KLIMA retrieval error varies from 2 to 7 ppm while the TANSO-FTS measurements have an error of about 1.5 ppm. The cloud of points has a peak-to-peak spread, mainly caused by the dispersion of IASI values, which seams to be rather large relative to the values of the retrieval errors. In order to try to assess the contribution of the atmospheric variability to points dispersion in the scatter plot, the differences between KLIMA and TANSO-FTS XCO_2 have been

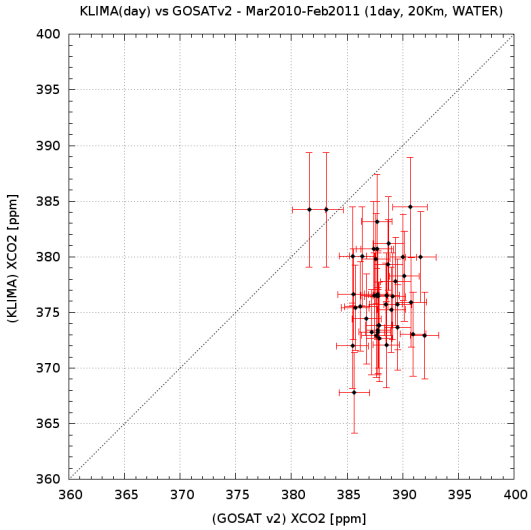

Fig. 6.23: *XCO_2 [ppm] over water for the year from March 2010 to February 2011: correlation between KLIMA retrieved values and TANSO-FTS V2 products*

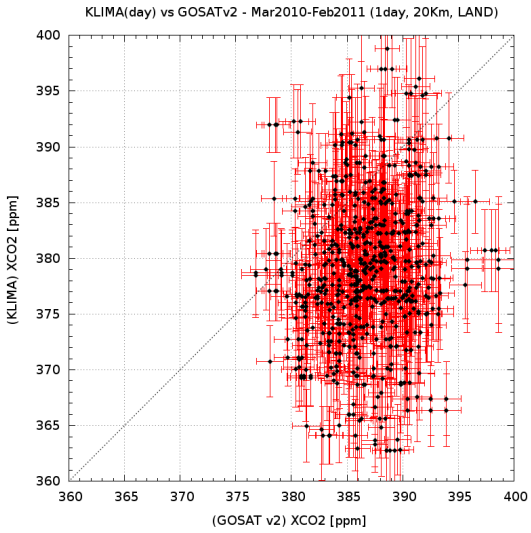

Fig. 6.24: *XCO_2 [ppm] over land for the year from March 2010 to February 2011: correlation between KLIMA retrieved values and TANSO-FTS V2 products*

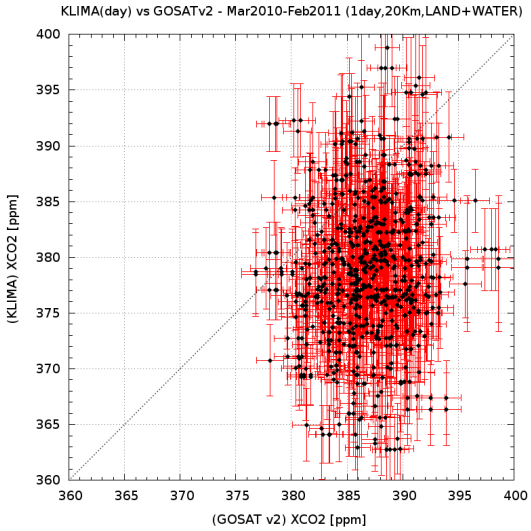

Fig. 6.25: *XCO_2 [ppm] over land and water for the year from March 2010 to February 2011: correlation between KLIMA retrieved values and TANSO-FTS V2 products*

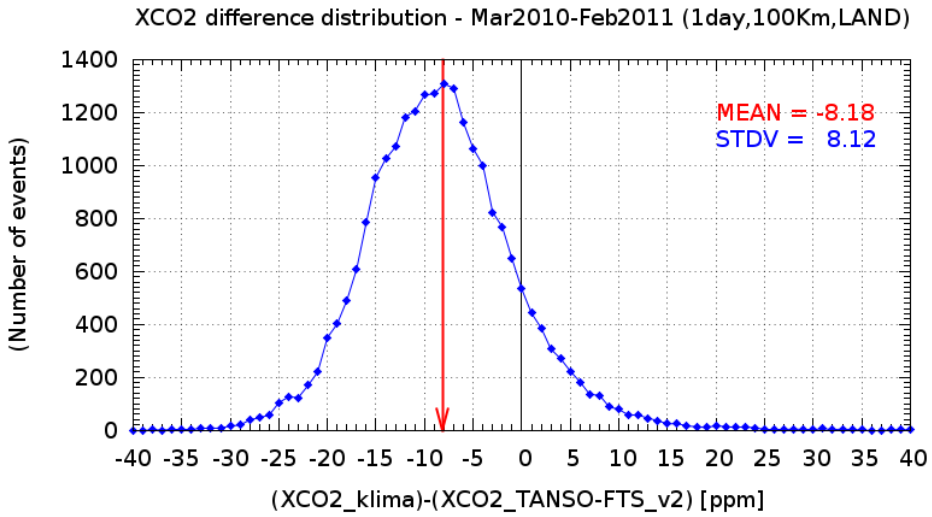

Fig. 6.26: *Distribution of the differences between KLIMA and TANSO-FTS V2 XCO_2 products evaluated with a coincidence criterion equal to 100 km: case observation on land for the period from March 2010 to February 2011*

calculated and plotted in a histogram. To increase the statistics and to better evaluate the data dispersion, the histogram has been built using a larger coincidence criterion with maximum distance equal to 100 km.

In general, KLIMA values show an error and a spread larger than GOSAT. KLIMA retrieved XCO_2 has a spread of about 30 ppm over land while the spread of GOSAT values is limited to about 10 ppm. Moreover KLIMA XCO_2 total column values are, on average, lower than the GOSAT ones. Observations on water have less variability but show the same negative bias.

Figure 6.26 shows the histogram of the differences of the coincident events over land. The bias of IASI relative to TANSO-FTS is about -8 ppm with a standard deviation of 8.2 ppm. As suggested by the spread highlighted in the scatter plots, the standard deviation is larger than expected and its effect is to cover the effective correlation of the two dataset. Additional sources of error must be present. Indeed a difference can be caused by the assumption of different vertical distribution of CO_2 in the two retrieval codes.

In order to take into account for this possible error, the previous comparison has been repeated after applying the smoothing described in 6.2 to TANSO-FTS XCO_2 products. The results of the correlation are reported from Figure 6.27 to Figure 6.29 while the distribution of the differences, evaluated with a coincidence criterion equal to 100 km, is reported in Figure 6.30. Note that when the XCO_2 value is smoothed, also its associated error have to be evaluated accordingly. However in this work the smoothing has been applied only to the XCO_2 values but not to its associated error. As a result, in the correlation graphics the error on TANSO-FTS smoothed values has not been displayed.

The application of KLIMA a priori and AKs to TANSO-FTS retrieved products result in two main effects: over bias and over spread. The negative biases observed in the first comparison is significantly reduced, demonstrating a better agreement between KLIMA and smoothed TANSO-FTS products. On the other hand the spread of TANSO-FTS XCO_2 products distribution increases. These effects are quantified comparing Figure 6.26 (TANSO-FTS) and Figure 6.30 (TANSO-FTS smoothed) where the bias is reduced from -8 to -2 ppm, and the standard deviation of the distribution increases from 8.2 to 9.0 ppm.

Comparison of the averages

Comparison of the *averaged* value strategy consisted in the comparison of the XCO_2 values averaged over a 2° x 2° and a 9° x 9° pixel grid, for the whole year. Figure 6.31 reports for each month the comparison in the form of scatter plot of the two XCO_2 average measurements on the 2° x 2° pixel grid. The points with the ratio between the standard deviation of the retrieved values and their mean retrieval

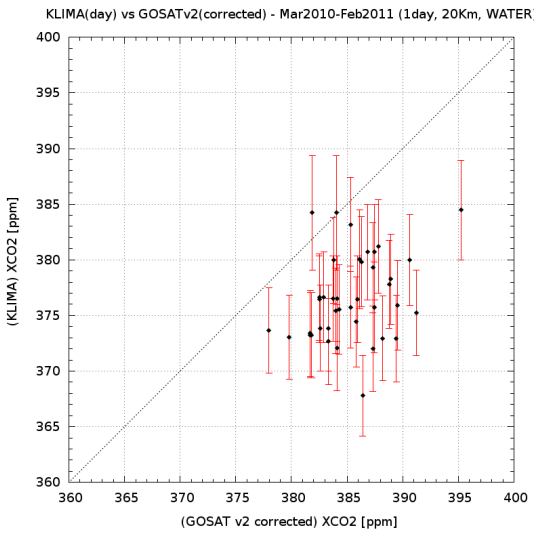

Fig. 6.27: *Distribution of the differences between KLIMA and TANSO-FTS V2 XCO_2 [ppm] smoothed products: case observation over water for the period from March 2010 to February 2011*

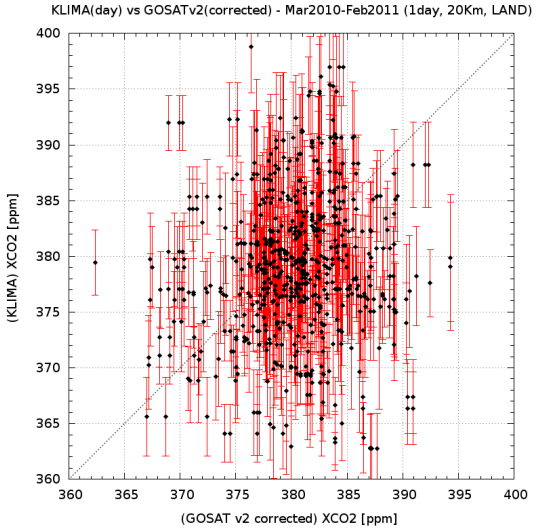

Fig. 6.28: *Distribution of the differences between KLIMA and TANSO-FTS V2 XCO_2 [ppm] smoothed products: case observation over land for the period from March 2010 to February 2011*

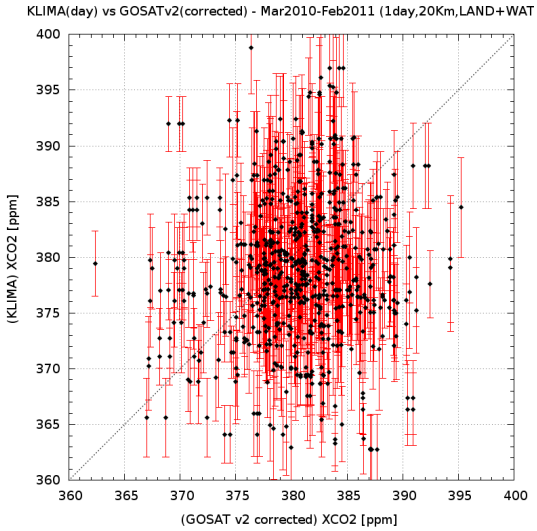

Fig. 6.29: *Distribution of the differences between KLIMA and TANSO-FTS V2 XCO_2 [ppm] smoothed products: case observation over land and water for the period from March 2010 to February 2011*

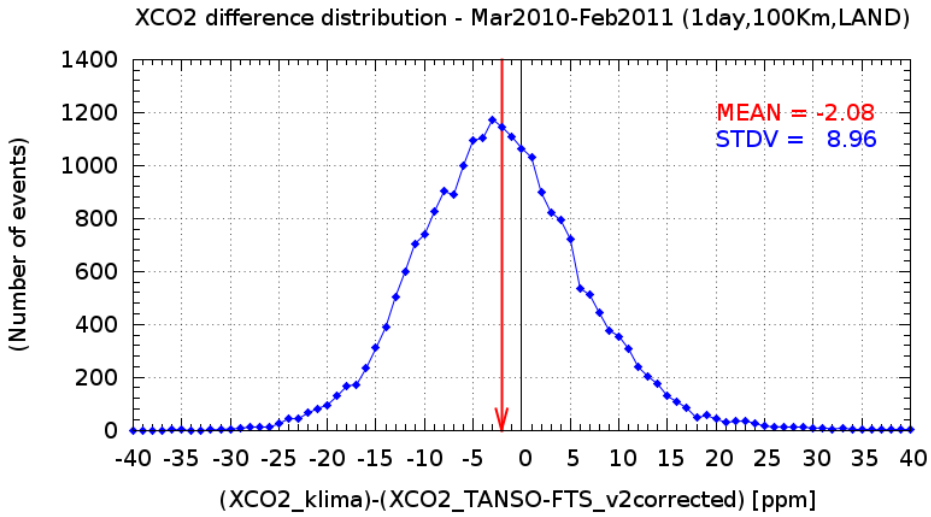

Fig. 6.30: *Distribution of the differences between KLIMA and TANSO-FTS V2 XCO_2 [ppm] smoothed products evaluated with a coincidence criterion equal to 100 km: case observation over land for the period from March 2010 to February 2011*

error larger than 2 are not included in this analisys, in order to skip the pixels where a large geographical and/or a temporal variability could be observed. Each panel shows the scatter plot of KLIMA versus TANSO-FTS (red dots), the TANSO-FTS L2 adjusted for the KLIMA a priori (green dots) and the scatter plot of KLIMA versus TANSO-FTS L2 smoothed (blue dots).

In general, when smoothed values are used, the cloud of points in the scatter plot is better centred relative to 45° line; this indicates a better correlation between the two dataset, even if the negative bias and the spread of KLIMA values still remain.

The scatter plot of Figure 6.32 shows the comparison between the two XCO_2 averaged measurements for each selected week in the time range from MArch 2010 to February 2011 reported on a 9° x 9° pixel grid. The great dispersion of KLIMA values with respect to TANSO-FTS dataset is not reduced using a larger pixel dimension.

Seasonal variations and mean values over macro areas comparisons

The third type of test regards the comparison of *seasonal variations* in the period from March 2010 to February 2011. Figure 6.33 shows the seasonal variation of XCO_2 in the Northern and in the Southern Hemispheres (top and bottom panel, respectively) calculated for the monthly averages of KLIMA and TANSO-FTS L2 SWIR products. The red line indicates the KLIMA measurements while the TANSO-FTS measurements are indicated by a green line in the case of no correction and by a blue line when the smoothing process is applied.

This analisys has been repeated for a subset of latitudinal range. Figures from 6.34 to 6.38 report the results for:

- Polar regions North (latitude from 60° to 90°) and South (latitude from -60° to -90°);
- Mid Latitude North (latitude from 30° to 60°) and South (latitude from -30° to -60°);
- Tropical region (latitude from -30° to 30°).

where KLIMA seasonal trend is plotted in red, TANSO-FTS with no correction is plotted in green and TANSO-FTS in the case of applied smoothing is plotted in blue.

The XCO_2 products obtained from KLIMA and TANSO-FTS SWIR V2 are considerably different when no smoothing is applied; however, when the TANSO-FTS products are smoothed taking into account the KLIMA AKs and the KLIMA a priori a good agreement is achieved both in terms of XCO_2

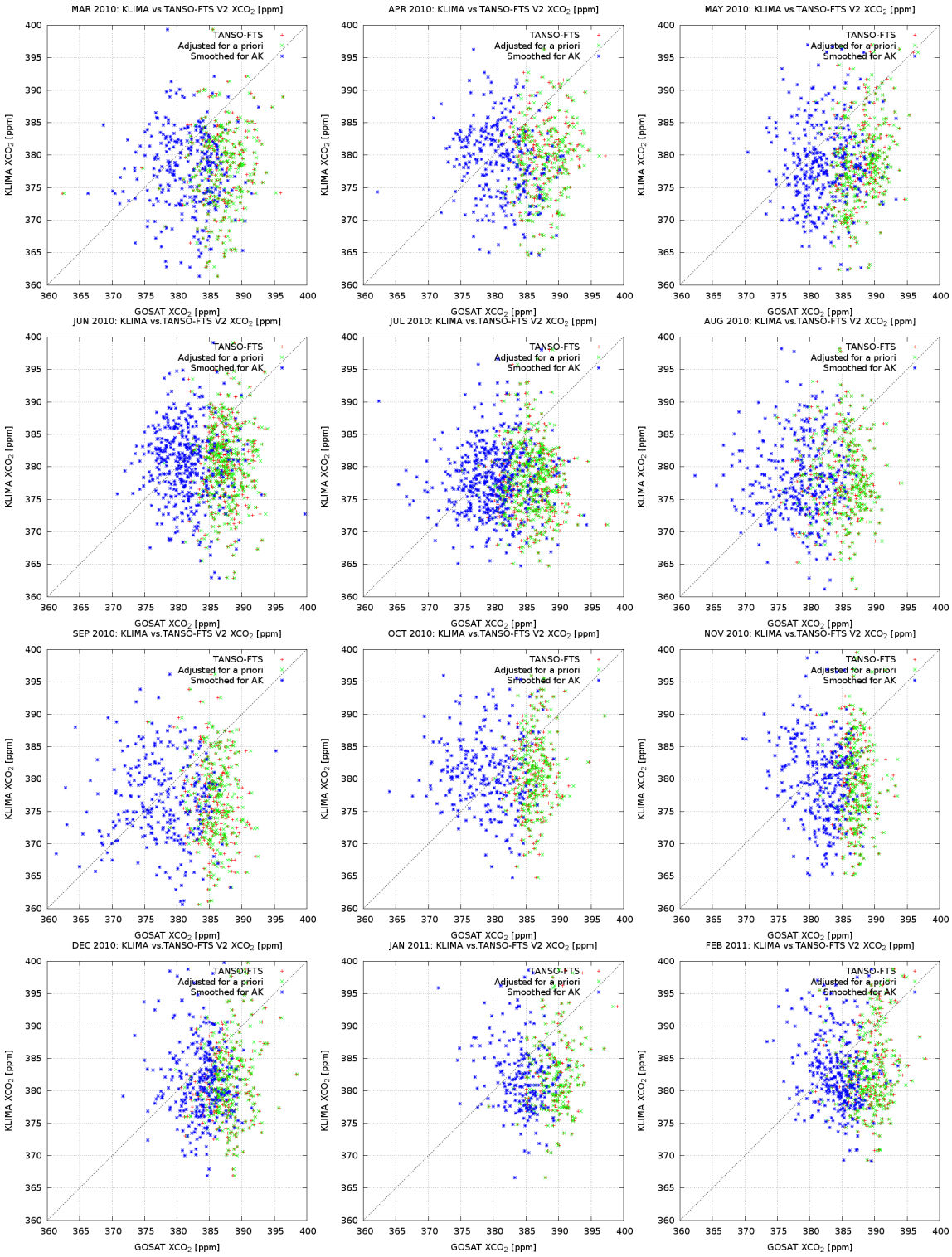

Fig. 6.31: *March 2010 - February 2011: scatter diagrams of KLIMA L2 - TANSO-FTS V2 comparison on a 2° x 2° pixel grid for each month. Red crosses represent the TANSO-FTS SWIR V2 XCO_2 , green marks are TANSO-FTS SWIR adjusted for KLIMA a priori while blue marks are TANSO-FTS SWIR smoothed with KLIMA AKs*

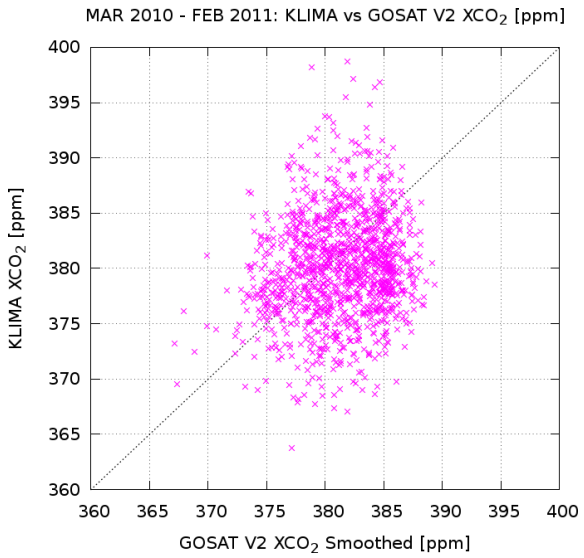

Fig. 6.32: *March 2010 - February 2011: scatter diagram of KLIMA L2 - TANSO-FTS V2 XCO_2 [ppm] comparison on a 9° x 9° pixel grid*

values and in terms of seasonal variations, especially in the Northern Hemisphere. For Southern Hemisphere the agreement is good for the summer period (from June to October) and degrades in winter. In the case of analysis per latitude bands a better agreement is found for Northern Mid Latitude and Polar regions (Figure 6.35 and 6.34) while significant differences remain for Tropical regions (Figure 6.36). The comparison of seasonal trend in Southern Polar region is not possible to perfom due to TANSO-FTS SWIR missing data (Figure 6.38).

In Figure 6.39 is reported the annual average related to selected macro areas (Figure 6.1) calcolated for KLIMA (red points), TANSO-FTS SWIR (green points) and TANSO-FTS SWIR smoothed (blue points). The standard deviation on the mean is also shown. In the bottom of the figure we added the number of observations for each macro area.

In general, over land TANSO-FTS XCO_2 values are reduced after applied the smoothing procedure and the values are closer to the KLIMA ones, but we note an increase of the dispersion of TANSO-FTS results. Over water 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). Despite the small number of samples the behaviour over water is consistent with what has been observed for co-located analysis and KLIMA values are significantly lower than TANSO-FTS ones. Moreover in this case the effect of the smoothing is very small and produces an opposite effect of little increasing of the bias between the two datasets. One of the possible causes of this bias could be the different local time of the two observations: TANSO-FTS has a local time at ascending node of 13:00, while IASI has a local time of 9:30. Apart the South America and areas over the oceans, the geographical variations agree with TANSO-FTS. The observed bias is greater over the regions where systematic errors that affect KLIMA FM are larger (Figure 5.18).

In conclusion we can say that KLIMA L2 and TANSO-FTS SWIR V2 products are on average in a good agreement for land observations, as emerged from the comparison in Northern Hemisphere dominated by land measurements. Differences between the two datasets are more evident for measurement acquired over sea, as also confirmed by the co-located comparison (6.5.2); in this measurement scenario the application of KLIMA AKs and a priori smoothing produces a small change of the TANSO-FTS products.

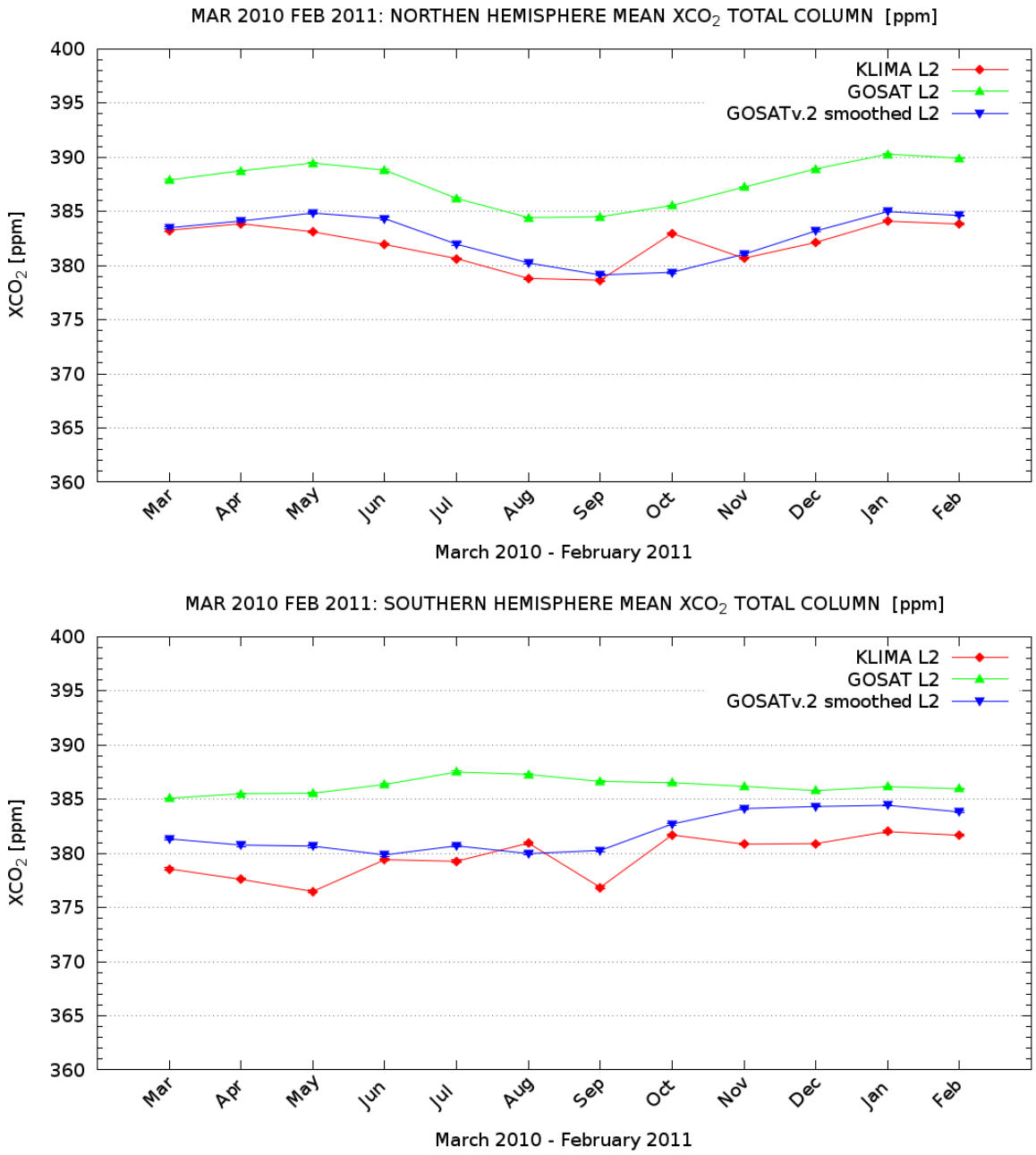

Fig. 6.33: *Seasonal variation of the XCO_2 [ppm] from March 2010 to February 2011. Top: Northern Hemisphere, bottom: Souther Hemisphere. The average on the Hemispheres of the XCO_2 retrieved by KLIMA (red points) is compared with the average of TANSO-FTS SWIR V2 XCO_2 products; green marks refer to TANSO-FTS V2 original products while blue marks refer to V2 smoothed with KLIMA a priori and AKs*

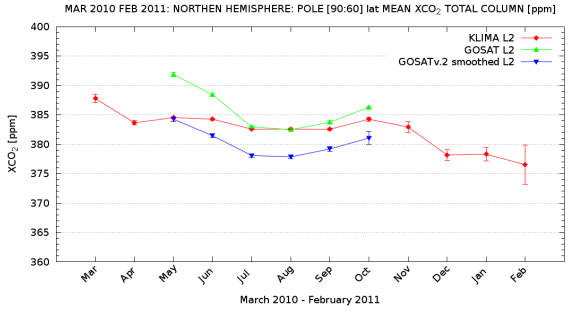

Fig. 6.34: *North Pole: seasonal variation of the XCO_2 [ppm] from March 2010 to February 2011*

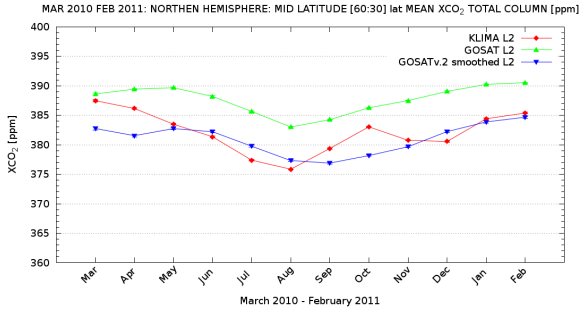

Fig. 6.35: *Mid Latitude North: seasonal variation of the XCO_2 [ppm] from March 2010 to February 2011*

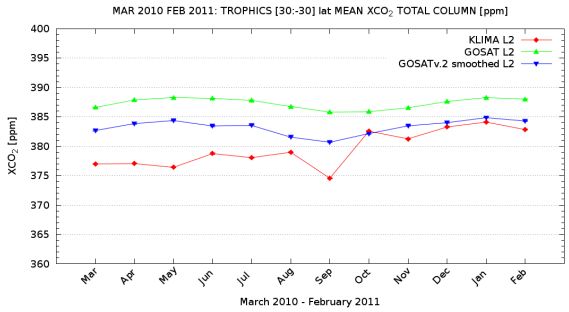

Fig. 6.36: *Tropics: seasonal variation of the XCO_2 [ppm] from March 2010 to February 2011*

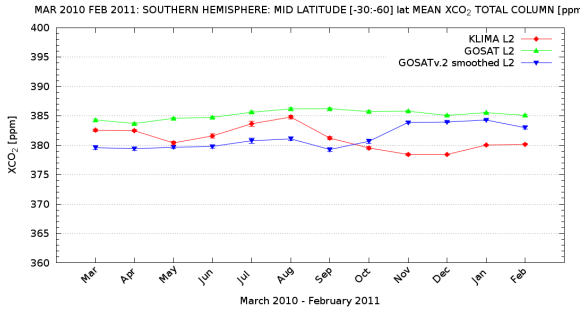

Fig. 6.37: *Mid Latitude South: seasonal variation of the XCO_2 [ppm] from March 2010 to February 2011*

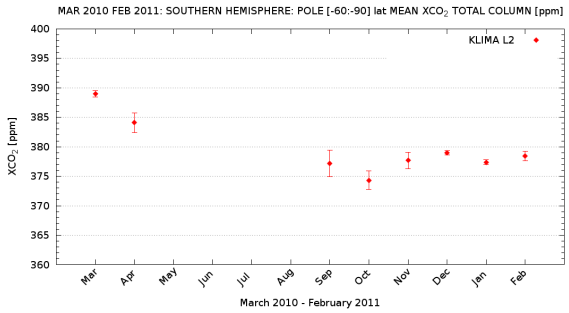

Fig. 6.38: *South Pole: seasonal variation of the XCO_2 from [ppm] March 2010 to February 2011*

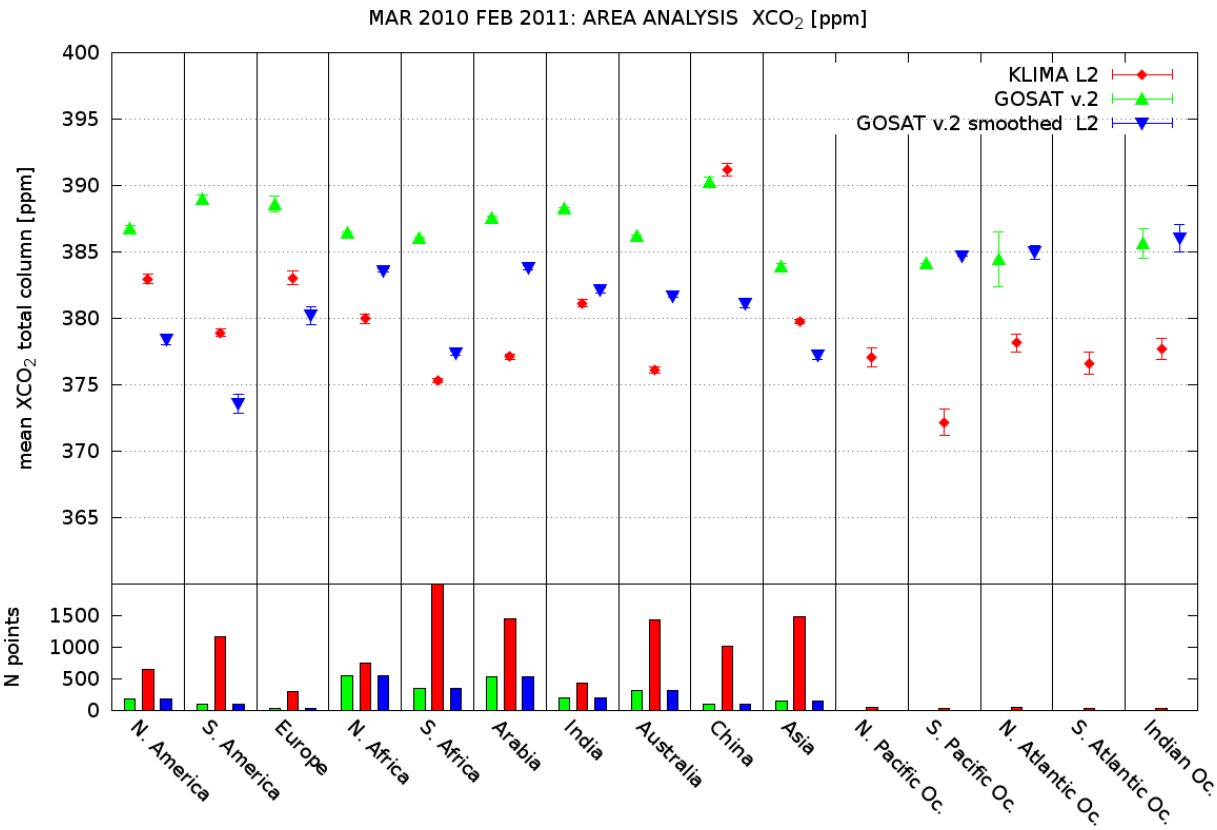

Fig. 6.39: *Top: Average XCO_2 [ppm] for each macro area (reported in Figure 6.1). Red points represent KLIMA average XCO_2 green points represent EUMETSAT products and blue points represent TANSO-FTS SWIR V2 ones. Bottom: Number of observations for each macro area*

Chapter 7

Summary and Conclusions

The objective of the KLIMA study was that of gaining a greater insight in the capabilities of thermal InfraRed passive remote sounding from space to retrieve accurate information on atmospheric Carbon Dioxide. The observations of IASI satellite instrument in this spectral region can provide an important contribution to the increase of coverage and quality of the data necessary to constrain the inverse modelling of surface Carbon fluxes and to determine the natural and anthropogenic sources and sinks of CO_2. To this purpose the potential was analysed of measurements acquired by the IASI high resolution InfraRed sounder on-board the MetOp-A satellite to provide CO_2 total column values with precision and accuracy comparable to those of GOSAT/TANSO-FTS, which is the first satellite-based sensor dedicated to the measurement of Carbon Dioxide. The KLIMA retrieval process, which performs an innovative wide band and multi-target analysis, was used in this study. The combination of these two features makes it possible to exploit for the retrieval all the measured spectral channels and to calculate, in a rigorous way, a retrieval error that also accounts for the errors of interfering parameters.

The primary task of this work was the adaptation of the KLIMA algorithm into an inversion code to investigate the ultimate capabilities of the algorithm to retrieve CO_2 total column from observations acquired in the thermal InfraRed by the IASI spectrometer. This leaded to a non-operational, line-by-line radiative transfer code with capability to simulate wide-band spectral radiances acquired by the IASI instrument, the so-called Reference Forward Model 2.4.1. KLIMA Reference Forward Model has been validated by comparing synthetic IASI measurements generated by KLIMA FM code with those of the FM of the LBLRTM code 2.5. The residuals of these comparisons are almost always smaller than the standard deviation of IASI measurements. Subsequently, the quality of the RFM was also confirmed by the absence of significant biases in the average residuals of a large number of retrievals performed on real measurements.

In order to achieve a reasonable trade-off between retrieval quality and efficiency of KLIMA code and with the purpose of massive processing of IASI Level 1 data, an optimized version of KLIMA code was implemented and integrated on the ESA G-POD Application System. To reach this goal a series of approximations has been implemented in the KLIMA Reference Forward Model chap:RadTran:sect:AccFM that leads to the Accelerated Forward Model. It was verified that with the only exception of a few spectral channels, the approximations introduce differences that are significantly smaller than the measurement error.

When the Reference Forward Model is associated with the retrieval model we have the Reference Retrieval Model 3.2.1. This code configuration is used to assess the ultimate retrieval capabilities of IASI for the exploitation of all the spectral channels of IASI, for a broadband and multi-target retrieval of the atmospheric state. The Reference Retrieval Model achieves the best performances for the retrieval of Carbon Dioxide columns, with total retrieval error on CO_2 total column of $\approx 1.0\%$ and no bias. The Reference Retrieval Model, however, does not meet the requirements in terms of program size and running time for integration on G-POD. Using the Accelerate Forward Model in association with the retrieval code (obtaining the Accelerated Retrieval Model 3.2.2) a significant reduction of the computing time is obtained with respect to the use of the Reference Forward Model. Moreover, the total retrieval

uncertainty are only slightly degraded respect to the Reference Retrieval Model and are fully consistent with the requirement of 0.3% accuracy of CO_2 total column at monthly intervals on regional scales for GOSAT validation.

Using the Accelerated Retrieval Model a significant reduction of the computing time is attained with no deterioration of the retrieval error, however the G-POD requirements are not meet. A reduction of the analysed spectral range, which causes an increase by a factor 1.4 of the retrieval error of the CO_2 column from a single IASI spectrum, is also needed to meet these preliminary computing requirements. So, the analysis was limited to the spectral range 645-800 cm^{-1}, which provides the best compromise between high sensitivity to Carbon Dioxide, low sensitivity to other atmospheric variables, and the reduction of the computation time of about 60-70 % 3.3.2.

The original version of the KLIMA code is capable to process both Level 1B (L1B) and Level 1C (L1C) IASI data. A first processing of IASI data was carried out to verify that the retrieval products obtained by applying the KLIMA inversion model to L1B and L1C data have negligible differences 3.3.1. Consistent results were obtained from the two datasets, and IASI L1C, which consist of geolocated and apodised spectra, has been selected to perform the bulk processing of IASI data using the version of KLIMA algorithm integrated on G-POD system.

After obtained the IASI L1 database from UMARF and with the computing resources of G-POD system, the massive processing started 5.1. In the selected period of one year, from March 1, 2010 to February 28, 2011, one week per month of IASI L1 data were processed on a global geographical coverage, retrieving the CO_2 both on land and on water and both during day and night. Only clear sky IASI observations have been analysed. The G-POD analysis made possible the acquisition of a large amount of data: 2300 processed orbits and a total of 240000 Retrieved Level 2 products, with good geographical and temporal distributions.

With the use of IASI TIR spectral channels, measurements are available both during day and night, even if no significant diurnal variability has been noticed in the dataset observed so far. This suggests that there is not enough sensitivity to catch the CO_2 diurnal cycle.

Data show a geographical and seasonal variability that can be reasonably ascribed to the dynamic of the Carbon cycle together with some scattered points that seem to be outliers (as described in 5.2). 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 5.3. 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).

In the monthly averages, plotted over a 2° x 2° grid, 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. 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 temporal 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 χ^2 indicate that some unaccounted errors are present.

The experimental retrieval error of XCO_2 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. Pixels with a dispersion larger that the estimated error, which indicate either unaccounted errors or variability of the observed atmosphere, are rather uncommon and occur mainly in

specific locations over land.

The average residuals of the retrieval fit are in most spectral channels significantly smaller than the spectral measurement error. However, the χ^2-test 5.4 also shows a correlation of the χ^2 values with the temperature of Earth's surface, with χ^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 of χ^2 greater than 3 occasionally occur, but being always located in desert areas are most probably due to sand storms.

Values of χ^2 greater than unity are a sign of unaccounted errors, which can be:

- Errors due to missing processes in the Forward Model, such as the scattering contribution due to particles and aerosols (i.e. the observed desert-storm effect) and non-LTE diurnal variations (which are expected to be negligible in the troposphere);
- Errors due to external constraints such as the surface emissivity in some areas of the globe that are difficult to characterize (e.g. desert and snow-covered regions);
- Systematic spectroscopic errors;
- Systematic errors in the IASI measurements (e.g. calibration errors).

In conclusion, the diagnostics of the KLIMA retrieved products 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.

As a part of the activities conducted to evaluate the retrieval quality of KLIMA code, comparison of the complete CO_2 dataset retrieved using KLIMA with the Level 2 products delivered by EUMETSAT and by TCCON ground stations was performed, as a test of algorithm performances.

A first comparison was made with the EUMETSAT Level 2 analysis of IASI observations which provides an extensive set of XCO_2 data that can be directly compared with KLIMA results 6.3. These data have not yet been validated and for this product no retrieval errors are given. KLIMA results have a larger dispersion, even if the EUMETSAT results are likely to have a larger error because based on a subset of spectral points. The comparison of these two datasets does not show any straightforward correlations and, since neither of the two has yet been validated, the doubt exists of whether they are observing the same quantity. The two analyses provide different results even when the seasonal variations, in which the monthly averages made in some geographical areas, are compared. Moreover, a new version of the processor was released on the operational processing chain on 14 September 2010 and this changed in the processor could explain the discontinuity, relative to October, that we can observe in the seasonal variation of IASI L2 data.

The comparison between KLIMA IASI and TCCON ground stations 6.4 shows that KLIMA retrievals are low by about 2-3%, even if the number of coincidences is fairly low for a statistical analysis. Moreover, it is not possible to take into account the different a priori profiles and the smoothing effect of the AK because the TCCON vertical profiles are not available from the standard products. The amplitude of this bias is compatible with existing spectroscopic uncertainties. If corrected for this offset, the KLIMA observations reproduce well the variations observed by the ground stations.

The second goal of this work was the comparison of retrieved Level 2 products by KLIMA code with TANSO-FTS/GOSAT Level 2 products to assess the coherence and temporal behaviour of large scale spatial features obtained from the two different sensors. This last comparison was made with the measurements of XCO_2 performed by the TANSO-FTS instrument 6.5 using SWIR channels only due to the lack of TANSO-FTS TIR data. Preliminary results show that some systematic biases exist in the TIR spectrum because of the opacity of the mirrors of SWIR bands that obstruct the field of view of band 4. Saitoh at al. [78] is performing a new reprocessing of band 4 data, but data available are still limited and the reprocessing is an ongoing activity. TANSO-FTS SWIR has very few measurements at high latitudes (where also the sensitivity of IASI measurements is reduced) and over the oceans (where IASI has instead a very good coverage) and it observes a seasonal and monthly averaged variability that

is smaller than the one observed by KLIMA. The AK of the two instruments are quite difference and this make the comparison of the measurements more difficult. Therefore, not only IASI has the potential of providing more information than TANSO-FTS, but it is complementary to this instrument because of important contributions such as the ocean coverage, the availability of night time measurements and the different sensitivity to vertical distribution of the target.

In general a large offset is present between XCO_2 values retrieved from KLIMA and TANSO-FTS SWIR, for which we can give a quantitative estimate in terms of a negative bias of about 8 ppm on the KLIMA-IASI XCO_2 products. The difference is reduced to a negative bias of 1.2 ppm when the comparison is made with TANSO-FTS SWIR measurements smoothed for the effect of the different AKs and a priori profiles. No correlation was found between TANSO-FTS SWIR and KLIMA measurements, when either coincident measurements or monthly averages made in 2° x 2° pixels are considered. However, the two instruments observe similar behaviours when the seasonal variability of XCO_2 is calculated in some latitudinal bands. The agreement is also confirmed by similarities found when comparing the yearly averages of XCO_2 in selected geographical areas of 10° x 10°.

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, that is the first mission dedicated to the measurements of CO_2 and CH_4. Therefore, even if only a fraction of the available information has been extracted by the analysis on the G-POD, the study demonstrates that IASI can provide very extensive and accurate information on XCO_2 total column which is complementary, in terms of altitude and geographical coverage, to what is currently provided by the GOSAT instrument. The KLIMA code can retrieve from IASI stable and internally consistent XCO_2 data even if retrieved values show a negative bias, most probably due to spectroscopic errors. However, the comparisons that are performed in this study, while showing some encouraging agreement in the case of large averages, were not able to constrain the measurements with a stringent validation.

Appendix A

Comparison between KLIMA L2 and TANSO-FTS/GOSAT V1

With the purpose of performing an inter-comparison of the XCO_2 retrieved from KLIMA and TANSO-FTS L2 SWIR in the selected annual range of time (March 1, 2010 to February 28, 2011), we downloaded from the GUIG the following data products:

- FTS SWIR L2 XCO_2 Column amount V1.20
- FTS SWIR L2 XCO_2 Column amount V1.30
- FTS SWIR L2 XCO_2 Column amount V1.40

The different versions of the L2 products correspond to different versions of the input TANSO-FTS L1B product (measured spectrum) [89]. The TANSO-FTS SWIR V1 XCO_2 values were compared with the XCO_2 retrieved with the KLIMA code from IASI L1C. The comparison between TANSO-FTS and KLIMA was performed using the TANSO-FTS L2 data and the KLIMA L2 retrieved values obtained from the bulk processing of one week per month in the annual range of time selected for a global geographical coverage. For each month a complete week has been processed, in order to retrieve the XCO_2 both on land and on water and both during day and night. In the following Figures we report the comparison of the two dataset, as a scatter diagram of mean XCO_2 values, obtained for the global geographical coverage over a grid of 2° x 2°. The Figures report the correlation of KLIMA L2 XCO_2 versus the XCO_2 of TANSO-FTS L2 SWIR V1 (red dots), TANSO-FTS L2 adjusted for the KLIMA a priori (green dots), and the TANSO-FTS L2 smoothed for KLIMA AKs (blue dots).

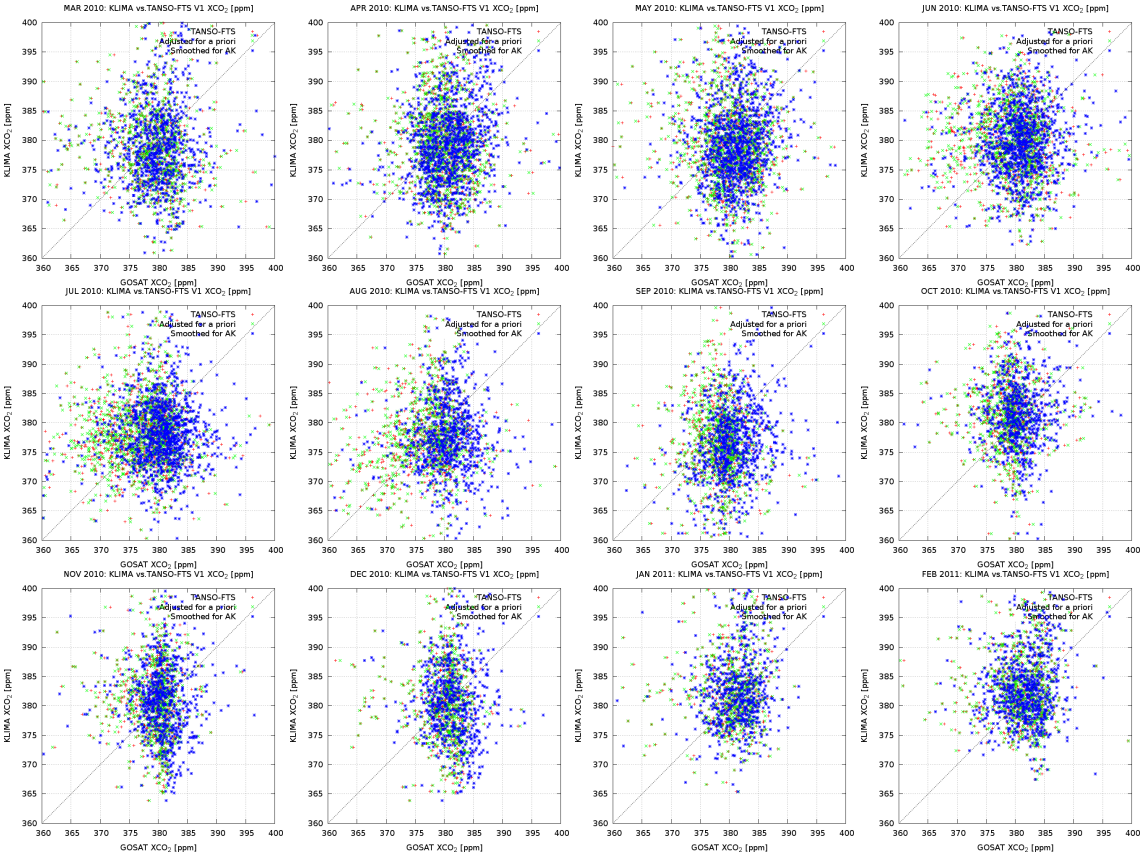

Fig. A.1: *Scatter diagrams of KLIMA L2 - TANSO-FTS V1 comparison for each month. Red points are TANSO-FTS SWIR XCO_2 V1, green points are TANSO-FTS SWIR adjusted for KLIMA a priori, blue points are TANSO-FTS SWIR smoothed with KLIMA AKs*

Acknowledgement

TCCON data were obtained from the TCCON Data Archive, operated by the California Institute of Technology, from the website at http://tccon.ipac.caltech.edu/.

TANSO-FTS L2 data were obtained from the GOSAT User Interface Gateway GUIG, operated by the National Institute for Environmental Studies (NIES), from the website: http://data.gosat.nies.go.jp/GosatUserInterfaceGateway/guig/GuigPage/open.do.

IASI data are received through the EUMETSAT Unified Meteorological Archive and Retrieval Facility (UMARF). The UMARF archive can be accessed through a web interface at http://archive.eumetsat.int/umarf.

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).

The sensor IASI has been developed and built under the responsibility of the Centre National dEtudes Spatiales (CNES, France). It is flown on-board the MetOp-A satellite as part of the EUMETSAT Polar System.

The author wish to thank G-POD Team at ESA/ESRIN for their valuable effort and availability for the bulk processing of IASI data. The author wish to thank the Japan Aerospace Exploration Agency (JAXA), the National Institute for Environmental Studies (NIES) and the Ministry of the Environment (MOE) for the important collaboration and information provided under the Greenhouse gases Observing SATellite (GOSAT) Project, which the project KLIMA is a part (PI: Ugo Cortesi).

Reference documents

RD1:
ESA-ESRIN contract N. 21612/08/I-OL

RD2:
KLIMA Sensitivity Assessment Report

RD3:
Thomas Williams, Colin Kelley et al., Gnuplot 4.6: an interactive plotting program

RD4:
Guido van Rossum et al., The Python Language Refenence, Python Software Foundation

RD5:
Travis E. Oliphant, Guide to NumPy, Trelgol Publishing, 2006

RD6:
John D. Hunter, Matplotlib: A 2D graphics environment, Computing In Science & Engineering

Bibliography

[1] August T., Klaes D., Schlussel P., Hultberg T., Crapeau M., Arriaga A., OCarroll A., Coppens D., Munro R., Calbet X.: IASI on Metop-A: Operational Level 2 retrievals after five years in orbit J. Quant. Spectrosc. Radiat. Transf. 2012, 113, 1340-1371.

[2] S. Basu, S. Guerlet, A. Butz, S. Houweling, O. Hasekamp, I. Aben, P. Krummel, P. Steele, R. Langenfelds, M. Torn, S. Biraud, B. Stephens, A. Andrews, and D. Worthy: Global CO_2 fluxes estimated from GOSAT retrievals of total column CO_2 , Atmos. Chem. Phys., 13, 86958717, 2013

[3] Bianchini, G., Carli, B., Cortesi, U., Del Bianco, S., Gai, M., Palchetti, L.: Test of far-InfraRed atmospheric spectroscopy using wide-band balloon-borne measurements of the upwelling radiance J. Quant. Spectrosc. Radiat. Transf. 2008, 109, 6, 1030-1042.

[4] Blumstein, D., Chalon, G., Carlier, T., Buil, C., Hebert, P., Maciaszek, T., Ponce, G., Phulpin, T., Tournier, B., Simeoni, D., Astruc, P., Clauss, A., Kayal, G., and Jegou, R.: IASI instrument: Technical overview and measured performances, in: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 5543, 169-207, 2004.

[5] Buchwitz, M., de Beek, R., Burrows, J. P., Bovensmann, H., Warneke, T., Notholt, J., Meirink, J.F., Goede, A. P. H., Bergamaschi, P., Korner, S., Heimann, M., and Schulz, A.: Atmospheric Methane and Carbon Dioxide from SCIAMACHY satellite data: initial comparison with chemistry and transport models, Atmos. Chem. Phys., 5, 941962, 2005, 5 `http://www.atmos-chem-phys.net/5/941/2005/`.

[6] Buchwitz, M., Schneising, O., Burrows, J. P., Bovensmann, H., Reuter, M., and Notholt, J.: First direct observation of the atmospheric CO_2 year-to-year increase from space, Atmos. Chem. Phys., 7, 42494256, doi:10.5194/acp-7-4249-2007, 2007.

[7] Canadell J.G., LeQur C., Raupach M.R., Field C.B., Buitenhuis E.T., Ciais P., Conway T.J., Gillett N.P., Houghton R.A., Marland G.: Contributions to accelerating atmospheric CO_2 growth from economic activity, Carbon intensity, and efficiency of natural sinks. Proc. Natl. Acad. Sci. U.S.A. 104, 1886618870, 2007.

[8] `http://earthobservatory.nasa.gov/Features/CarbonCycle/`

[9] Carli, B., Bazzini, G., Castelli, E., Cecchi-Pestellini, C., Del Bianco, S., Dinelli, B.M., Gai, M., Magnani, L., Ridolfi, M. and Santurri, L. : MARC: a code for the retrieval of atmospheric parameters from millimetre-wave limb measurements, JQSRT 105, 476-491 (2007).

[10] Chedin,C., A., Scott, N. A., Crevoisier, C., and Armante, R.: First global measurement of midtropospheric CO_2 from NOAA polar satellites: the tropical zone, J. Geophys. Res., 108, 4581, doi:10.1029/2003JD003439, 2003.

[11] Chedin, A., Saunders, R., Hollingsworth, A., Scott, N. A., Matricardi, M., Etcheto, J., Clerbaux, C., Armante, R., and Crevoisier, C.: The feasibility of monitoring CO_2 from high-resolution InfraRed sounders, J. Geophys. Res., 108(D2), 4064, doi:10.1029/2001JD001443, 2003.

[12] Chedin A.,Scott, N. A., Armante, R., Pierangelo, C., Crevoisier, C., Foss, O., and Ciais, P.: A quantitative link between CO_2 emissions from tropical vegetation fires and the daily tropospheric excess (DTE) of CO_2 seen by NOAA-10 (19871991), J. Geophys. Res., 113, D05302, doi:10.1029/2007JD008576, 2008

[13] Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P.,Breon, F.-M., Chedin, A., and Ciais, P.: Inferring CO_2 sources and sinks from satellite observations: method and application to TOVS data, J. Geophys. Res., 110, D24309, doi:10.1029/2005JD006390, 2005.

[14] Chevallier, F., Bron, F.-M., Rayner, P. J.: The contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework.J. Geophys. Res., 112, D09307, doi:10.1029/2006JD007375, 2007.

[15] Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal InfraRed IASI/ MetOp sounder, Atmos. Chem. Phys., 9, 60416054.d by: Strojnik, M., Vol. 5543, 2004.

[16] Clerbaux, C., Hadji-Lazaro, J., Turquety, S., George, M., Coheur, P. F., Hurtmans, D., Wespes, C., Herbin, H., Blumstein, D., Tournier, B., and Phulpin, T.: The IASI/MetOp I mission: First observations and highlights of its Potential Contribution to GMES, COSPAR inf. Bul., vol. 2007, 1924, 2007.

[17] Clerbaux, C., Boynard, A., Clarisse, L., George, M., Hadji-Lazaro, J., Herbin, H., Hurtmans, D., Pommier, M., Razavi, A., Turquety, S., Wespes, C., and Coheur, P.-F.: Monitoring of atmospheric composition using the thermal InfraRed IASI/MetOp sounder, Atmos. Chem. Phys., 9, 6041-6054, doi:10.5194/acp-9-6041-2009, 2009.

[18] Clough et al.: Line-by-Line Calculations of Atmospheric Fluxes and Cooling Rates. Application to Water Vapour. J. Geo Research, 97, No. D14, 15,761-15,785. 1992.

[19] `http://webbook.nist.gov/`

[20] Cogan, A., Boesch, H., Parker, R., Feng, L., Palmer, P., Blavier, J., Deutscher, N. M., Macatangay, R., Notholt, J., Roehl, C. M., Warneke, T., Wunch, D.: Atmospheric Carbon Dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations. Journal of Geophysical Research D: Atmospheres, 117 (21), D21301., 2012.

[21] Cortesi, U., et al.: KLIMA-IASI Sensitivity Assessment Report, Final Report for Phase 1 of the project Sensitivity Analysis and Application of KLIMA algorithms to GOSAT and OCO validation ESA-ESRIN contract N. 21612/08/I-OL, 2009.

[22] Cortesi U., et al.: Carbon Dioxide retrieval from IASI measurements using the klima inversion algorithm. ESA Living Planet symposium, Bergen. 2010.

[23] Crevoisier C., Chdin A., Scott N. A., Matsueda H., Machida T., and Armante R.: First year of upper tropospheric integrated content of CO_2 from IASI hyperspectral InfraRed observations, Atmos. Chem. Phys., 9, 4797-4810, 2009.

[24] Del Bianco, U. Cortesi, B. Carli.: Utilization of all spectral channels of IASI for the retrieval of the atmospheric state, ESA Living Planet symposium, Bergen. 2010.

[25] Engelen, R. J., Serrar, S., and Chevallier, F.: Four-dimensional data assimilation of atmospheric CO_2 using AIRS observations, J. Geophys. Res., 114, D03303, doi:10.1029/2008JD010739, 2009.

[26] `http://scrippsco2.ucsd.edu/`

[27] http://www.epa.gov/

[28] ESA-ESRIN contract N. 21612/08/I-OL.

[29] Farres, J., Mathot, E. and Pinto, S. G-POD: A Collaborative Environment for Earth Observation at the European Space Agency. Proceedings of the ESA Living Planet Symposium, 2010.

[30] Fiorenza C., Formisano, V.: A solar spectrum for PFS data analysis, Planetary and Space Science 53, 10091016, 2005.

[31] P.Friedlingstein, p.Cox, r.Betts, l.Bopp, w.Von bloh, v.Brovkin, p.Cadule, S.Doney, m.Eby, i.Fung, g.Bala, j.John, c.Jones, f.Joos, t.Kato, m.Kawamiya, W.Knorr, k.Lindsay, h.D.Matthews, t.Raddatz, p.Rayner, c.Reick, e.Roeckner, K.-g.Schnitzler, r.Schnur, k.Strassmann, a.J.Weaver, c.Yoshikawa, and n.Zeng: ClimateCarbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, JOURNAL OF CLIMATE, Volume 19

[32] Friedlingstein, P., Houghton, R.A., Marland, G., Hackler, J., Boden, T.A., Conway, T.J., Canadell, J.G., Raupach, M.R., Ciais, P., Le Qur, C.: Update on CO_2 emissions, Nature Geoscience 3, 811812, doi:10.1038/ngeo1022, Published online, 2010.

[33] http://www.globalcarbonproject.org/

[34] Geibel, M., Messerschmidt, J., Gerbig, C., Blumenstock, T., Chen, H., Hase, F., Kolle, O., Lavric, J. V., Nothol, J., Palm, M., Rettinger, M., Schmidt, M., Sussmann, R., Warneke, T., and Feist, D. G.: Calibration of column-averaged CH_4over European TCCON FTS sites with airborne in situ measurements, Atmos. Chem. Phys., 12, 87638775, doi:10.5194/acp-12-8763-2012.

[35] Gordon I.E., et al.: Current updates of the water-vapor line list in HITRAN: a new "diet" for air-broadened half-widths. J. Quant. Spectrosc. Ra.108 (3):389-402, 2007.

[36] http://www.gosat.nies.go.jp/eng/technology/references.htm

[37] http://gpod.eo.esa.int/

[38] http://data.gosat.nies.go.jp/GosatUserInterfaceGateway/guig/GuigPage/open.do

[39] Gurney, K. R., et al.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626630, 2002.

[40] http://www.cfa.harvard.edu/hitran

[41] Houghton, R. A.: How well do we know the flux of CO_2 from land-use change? Tellus, Series B: Chemical and Physical Meteorology, 62, 337-351, 2010.

[42] Hungershoefer, K., Breon, F.-M., Peylin, P., Chevallier, F., Rayner, P., Klonecki, A., Houweling, S., and Marshall, 5 J.: Evaluation of various observing systems for the global monitoring of CO_2 surface fluxes, Atmos. Chem. Phys., 10, 1050310520, doi:10.5194/acp-10-10503-2010, 2010.

[43] Houweling S., Breon F.-M. , Aben I., Rodenbeck C., Gloor M., Heimann M., and Ciais P.: Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter comparison of measurement techniques and their performance as a function of space and time. Atmos. Chem. Phys., 4, 523538, 2004

[44] http://www.jaxa.jp/press/2009/02/20090209_ibuki_e.html

[45] Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near InfraRed sensor for Carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Optics, 48, 67166733, 2009.

[46] John D. Hunter, Matplotlib: A 2D graphics environment, Computing In Science & Engineering, 9(3), 90–95, 2007

[47] IASI Level 1 Product Guide, EUM/OPS-EPS/MAN/04/0032

[48] IASI Level 2 Product Guide, EUM/OPS-EPS/MAN/04/0033

[49] http://www.eumetsat.int/

[50] Imasu, R. et al.: Upper-atmospheric CO_2 concentration retrieved from thermal InfraRed spectra observed using GOSAT TANSO-FTS (TIR) sensor. SPIE Vol. 7474 74740K-1, 2009.

[51] Ingmann, P.: A-SCOPE, Advanced Space Carbon and Climate Observation of Planet Earth, Report of Assessment, SP-1313/1, ESA Communication Product Office, Noordwijk, The Netherlands, 2009.

[52] Inoue M., Morino I., Uchino O., Miyamoto Y., Yoshida Y., Yokota T., Machida T., Sawa Y., Matsueda H., Sweeney V., Tans P.P., Andrews A.E., and Patra P.K.: Validation of XCO_2 derived from SWIR spectra of GOSAT TANSO-FTS with aircraft measurement data, Atmos. Chem. Phys. Discuss., 13, 3203-3246, 2013

[53] IPCC AR4 WG1, Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M. Averyt, K.B., Tignor, M., and Miller, H.L.: ed. Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2007.

[54] http://www.ipcc.ch/

[55] Jacquinet-Husson, N., Scott, N. A., Chedin, A., Crepeau, L., Armante, R., Capelle, V., and 47 coauthors: The GEISA spectroscopic database: Current and future archive for Earth and planetary atmosphere studies, J. Quant. Spectrosc. Ra., 109, 10431059, 2008.

[56] Andrew A. Lacis, Gavin A. Schmidt, David Rind, Reto A. Ruedy: Atmospheric CO_2 :Principal Control Knob Governing Earth's Temperature, Science 330, 356, DOI: 10.1126/science.1190653, 2010

[57] Canadell J.G., LeQuéré C., Raupach M.R., Field C.B., Buitenhuis E.T., Ciais P., Conway T.J., Gillett N.P., Houghton R.A., Marland G.: Contributions to accelerating atmospheric CO_2 growth from economic activity, Carbon intensity, and efficiency of natural sinks. Proc. Natl. Acad. Sci. U.S.A. 104, 1886618870, 2007.

[58] Matricardi M., Saunders R.: Fast radiative transfer model for simulation of InfraRed atmospheric sounding interferometer radiances. Appl Opt1999;38:567991.

[59] Messerschmidt, J., Geibel, M. C., Blumenstock, T., Chen, H., Deutscher, N. M., Engel, A., Feist, D. G., Gerbig, C., Gisi, M., Hase, F., Katrynski, K., Kolle, O., Lavri, J. V., Notholt, J., Palm, M., Ramonet, M., Rettinger, M., Schmidt, M., Sussmann, R., Toon, G. C., Truong, F., Warneke, T., Wennberg, P. O., Wunch, D., and Xueref-Remy, I.: Calibration of TCCON column-averaged CO_2: the first aircraft campaign over European TCCON sites, Atmos. Chem. Phys., 11, 10765-10777, doi:10.5194/acp-11-10765-2011, 2011.

[60] Miller, C. E., Crisp, D., DeCola, P.L., Olsen, S. C., Randerson, J. T., Michalak, A. M., Alkhaled, Rayner, P., Jacob, D. J., Suntharalingam, P., Jones, D. B. A., Denning, A. S., Nicholls, M. E., Doney, S. C., Pawson, S., Boesh, H., Connor, B. J., Fung, I. Y., OBrien, D., Salawitch, R. J., Sander, S. P., Sen, B., Tans, P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., Law, R.M.: Precision requirements for space-based XCO_2 data, J. Geophys. Res., 112, D10314, doi:10.1029/2006JD007659, 2007.

[61] Morino, I., Uchino, O., Inoue, M., Yoshida, Y., Yokota, T., Wennberg, P. O., Toon, G. C.,Wunch, D., Roehl, C. M., Notholt, J., Warneke, T., Messerschmidt, J., Griffith, D. W. T., Deutscher, N. M., Sherlock, V., Connor, B., Robinson, J., Sussmann, R., and Rettinger, M.: Preliminary validation of column-averaged volume mixing ratios of Carbon Dioxide and Methane retrieved from GOSAT short-wavelength InfraRed spectra, Atmos. Meas. Tech., 4, 1061-1076, doi:10.5194/amt-4-1061-2011, 2011.

[62] Cynthia D. Nevison, Natalie M. Mahowald, Scott C. Doney, Ivan D. Lima, Guido R. van der Werf, James T. Randerson, David F. Baker, Prasad Kasibhatla and Galen A. McKinley: Contribution of ocean, fossil fuel, land biosphere, and biomass burning Carbon fluxes to seasonal and interannual variability in atmospheric CO_2. Journal of Geophysical Research, 113, G01010, doi: 10.1029/2007JG000408.

[63] Niro, F., et al.: Spectra calculations in central and wing regions of CO_2 IR bands. IV: software and database for the computation of atmospheric spectra. J. Quant. Spectrosc. Ra., 95, 469 481. 2005a.

[64] Niro, F., et al.: Spectra calculations in central and wing regions of CO_2 IR bands between 10 and 20 mm. III: atmospheric emission spectra. J. Quant. Spectrosc. Ra., 90, 61 76. 2005b.

[65] `http://www.esrl.noaa.gov/gmd/ccgg/trends/`

[66] Travis E. Oliphant: Guide to NumPy, Trelgol Publishing, 2006, `http://www.tramy.us`

[67] Rayner, P. J., and D. M. OBrien: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Lett., 28, 175 178, 2001.

[68] Remedios JJ.: Extreme atmospheric constituent profiles for MIPAS. In: European symposium on atmospheric measurements from space. vol. 2. Noordwijk, Netherlands: ESTEC; 1999. p. 77983.

[69] Remedios, J. J. et al.: MIPAS reference atmospheres and comparisons to V4.61/V4.62 MIPAS level 2 geophysical datasets, Atmos. Chem. Phys. Discuss., 7, 9973-10017, 2007.

[70] Reuter, M., et al.: Retrieval of atmospheric CO_2 with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results, J. Geophys. Res., 116, D04301, doi:10.1029/2010JD015047.

[71] Rothman, L. S. et al.: The HITRAN2008 molecular spectroscopic database, , JQSRT 110, 533572, 2009.

[72] Rodgers, C.D.: Inverse method for atmospheric sounding: Theory and Practice, 1^{st} ed., World Sci, River Edge, N.J, 2000.

[73] Rodgers, C. and Connor, B.: Inter-comparison of remote sounding instruments, J. Geophys. Res, 108, 41164229, 2003.

[74] Roedenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO_2 flux history 19822001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. Phys., 3, 19191964, 2003.

[75] van Rossum Guido et al.: The Python Language Reference, Python Software Foundation, `http://docs.python.org/2/reference`

[76] Rumelhart, D. E., Hinton, G. E., and Williams, R. J.: Learning internal representations by error propagation, in: Parallel Distributed Processing: Explorations in the Macrostructure of Cognition, vol. 1, edited by: Rumelhart, D. E. and McClelland, J. L., MIT Press, Cambridge, Mass, 318-362, 1986.

[77] Sabine C.L., Feely, R. A., Gruber, N., Key, R.M., Lee, K., Bullister, J.L., Wanninkhof, R., Wong, C.S., Wallace, D.W.R., Tilbrook, B., Millero, F.J., Peng, T-H., Kozyr, A., Ono, T., Rios, A.F.: The Oceanic Sink for Anthropogenic CO_2, Science, 305, 367-371, 2004

[78] Naoko Saitoh, Ryoichi Imasu, Yoshifumi Ota, and Yosuke Niwa: CO_2 retrieval algorithm for the thermal infrared spectra of the Greenhouse Gases Observing Satellite: Potential of retrieving CO_2 vertical profile from high-resolution FTS sensor, JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 114, D17305, doi:10.1029/2008JD011500, 2009

[79] Sarmiento, J. L., and Wofsy, S.C.: A U.S. Carbon Cycle science plan, 1999.

[80] Scott N. A. and Chedin, A.: A fast line-by-line method for atmospheric absorption computations: The Automatized Atmospheric Absorption Atlas, J. Appl. Meteorol., 20, 556564, 1981.

[81] Shephard, M.W., et al.: Performance of the line-by-line radiative transfer model (LBLRTM) for temperature and species retrievals: IASI case studies from JAIVEx. Atm. Chem. Phys., 9, 7397-7417. 2009.

[82] Simeoni, D. Et al.: Design and Development of the IASI instrument, InfraRed Spaceborne Remote Sensing XII. Edited by Strojnik, Marija. Proceedings of the SPIE, Volume 5543, 208-219, 2004.

[83] Schneising, O., P. Bergamaschi, H. Bovensmann, M. Buchwitz, J. P. Burrows, N. M. Deutscher, D. W. T. Griffith, J. Heymann, R. Macatangay, J. Messerschmidt, J. Notholt, M. Rettinger, M. Reuter, R. Sussmann, V. A. Velazco, T. Warneke, P. O. Wennberg, and D. Wunch: Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results, Atmos. Chem. Phys., 12, 1527-1540, 2012.

[84] `https://tccon-wiki.caltech.edu/Sites#TCCON_Site_Map`

[85] `http://tccon.ipac.caltech.edu/`

[86] Yogesh K. Tiwari, Manuel Gloor, Richard J. Engelen, Frederic Chevallier, Christian Rodenbeck, Stefan Korner, Philippe Peylin, Bobby H. Braswell, and Martin Heimann: Comparing CO_2 retrieved from Atmospheric InfraRed Sounder with model predictions: Implications for constraining surface fluxes and lower-to-upper troposphere transport, JGR, VOL. 111, D17106, doi:10.1029/2005JD006681, 2006.

[87] `http://archive.eumetsat.org/umarf/`

[88] Thomas R. Karl, Jerry M. Melillo, and Thomas C. Peterson: Global Climate Change Impacts in the United States, US Global Change Research Program. Cambridge University Press, New York, NY, USA.

[89] Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: Retrieval algorithm for CO_2 and CH_4column abundances from short-wavelength InfraRed spectral observations by the Greenhouse gases observing satellite, Atmos. Meas. Tech., 4, 717734, doi:10.5194/amt-4-717-2011, 2011.

[90] Y. Yoshida, N. Kikuchi, I. Morino, O. Uchino, S. Oshchepkov, A. Bril, T. Saeki, N. Schutgens, G. C. Toon, D. Wunch, C. M. Roehl, P. O. Wennberg, D. W. T. Griffith, N. M. Deutscher, T. Warneke, J. Notholt, J. Robinson, V. Sherlock, B. Connor, M. Rettinger, R. Sussmann, P. Ahonen, P. Heikkinen, E. Kyro, J. Mendonca, K. Strong, F. Hase, S. Dohe, and T. Yokota: Improvement of the retrieval algorithm for GOSAT SWIR XCO2 and XCH4 and their validation using TCCON data, Atmos. Meas. Tech., 6, 1533-1547, 2013

[91] `http://www.eumetsat.int/website/home/Data/ServiceStatus/ProductHistory/index.html`

[92] Washenfelder, R. A., Toon, G. C., Blavier, J.F., Yang, Z., Allen, N. T., Wennberg, P. O., Vay, S. A., Matross, D. M., and Daube,B. C.: Carbon Dioxide column abundances at the Wisconsin tall tower site, J. Geophys. Res., 111, D22305, doi:10.1029/2006JD007154.

[93] Williams, T., Kelley, C., et al.: Gnuplot 4.6: an interactive plotting program, `http://www.gnuplot.info/documentation.html`

[94] Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jimnez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech., 3, 1351-1362, doi:10.5194/amt-3-1351-2010, 2010.

[95] Wunch,D., Toon, G.C., Blavier, J.-F.L., Washenfelder, R.A., Notholt, J., Connor, B.J., Griffith, D.W.T., Sherlock, V., Wennberg, P.O.: The Total Carbon Column Observing Network. Phil. Trans. R. Soc. A (2011) 369, doi:10.1098/rsta.2010.0240.

[96] Wunch, D., et al.: A method for evaluating bias in global measurements of CO_2 total columns from space, Atmos. Chem. Phys., 11,12,31712,337, doi:10.5194/acp-11-12317-2011.

Definitions and acronyms

AF	Airborne Fraction
AFM/ARM	Accelerated Forward/Retrieval Model
AIRS	Atmospheric InfraRed Sounder
AMSU	Advanced microwave Sounding Unit
ANN	Artificial Neural Network
A-SCOPE	Advanced Space Carbon and Climate Observation of Planet Earth
ATOVS	Advanced TIROS (Television and InfraRed Observational Satellite) Operational Vertical Sounder
AVHRR	Advanced Very High Resolution Radiometer
BUFR	Binary Universal Form for the Representation of meteorological data
CNR	Consiglio Nazionale delle Ricerche
CNES	Centre National dEtudes Spatiales
CONTRAIL	Comprehensive Observation Network for Trace gases by Airliner
DHF	Data Handling Facility
DOAS	Differential optical Absorption Spectroscopy
DVB	Digital Video Broadcast
ECMWF	European Centre for Medium-Range Weather Forecasts
ENVISAT	ENVIronmental SATellite
EPS	EUMETSAT Polar System
ESA	European Space Agency
ESRL	Earth System Research Laboratory
EUMETCAST	EUMETSAT's Data Distribution System
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FM	Forward Model
FOV	Field Of View
GOSAT	Greenhouse gases Observing SATtelite
G-POD	Grid Processing On Demand
GTS	Global Telecommunication System
GUI	Graphical User Interfaces
GUIG	GOSAT User Interface Gateway
HDF	Hierarchical Data Format
HITRAN	HIgh-resolution TRANsmission molecular absorption database
IASI	InfraRed Atmospheric Sounding Interferometer
IIS	Integrated Imaging Subsystem
IFAC	Institute for Applied Physics "Nello Carrara"
IFOV	Instantaneous Field Of View
IG2	Initial Guess 2
IPP	Interferogram Processing Program
ISRF	Instrument Spectral Response Function
JAIVEx	Joint Airborne IASI Validation Experiment
JAXA	Japan Aerospace Exploration Agency

KLIMA	Kyoto protocoL Informed Management of the Adaptation
L0	Level 0
L1	Level 1
L1A	Level 1 A
L1B	Level 1 B
L1C	Level 1 C
L2	Level 2
LBLRTM	Line-by-line Radiative Transfer Model
MARC	Millimetre-Wave Atmospheric Retrieval Code
MARSCHALS	Millimetre-wave Airborne Receiver for Spectroscopic CHaracterization of Atmospheric Limb-Sounding
METOP	METeorological SATellite
MHS	Microwave Humidity Sounder
MLP	Multi-Layer Perceptron
MOE	Ministry of the Environment
MSG	Meteosat Second Generation
MSS	Measurement Space Solution
MWIR	MicroWave
MWIR	Mid Wave InfraRed
NASA	National Aeronautic and Space Administration
NLSF	Non-linear Least-Square Fit
NLTE	Non Local Thermodynamic Equilibrium
NSR	Null Space Regularization
NIES	National Institute for Environmental Studies
NIR	Near InfraRed
NESR	Noise Equivalent Spectral Radiance
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
OCO	Orbiting Carbon Observatory
OEM	Optimal Estimation Method
PCA	Principal Component Analysis
QC	Quality Control
REFIR	Radiation Explorer in the Far InfraRed
RM	Retrieval Model
RMSS	Regularized Measurement Space Solution
RT	Radiative Transfer
RFM/RRM	Reference Forward/Retrieval Model
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Chartography/Chemistry
SWIR	Short Wave InfraRed
TANSO-CAI	Thermal And Near-InfraRed Sensor for Carbon Observation- Cloud and Aerosol Imager
TANSO-FTS	Thermal And Near-InfraRed Sensor for Carbon Observation- Fourier Transform Spectrometer
TCCON	Total Carbon Column Observing Network
TIROS	Television and InfraRed Observational Satellite
TIR	Thermal InfraRed
UMARF	Unified Meteorological Archive and Retrieval Facility
UV	Ultra-Violet
V1	Version 1
V2	Version 2

WFM-DOAS Weighting Function Modified DOAS

Tante sono le persone che vorrei ringraziare, che mi hanno sostenuta e guidata nella realizzazione di questa tesi. Vorrei ringraziare prima di tutto il mio relatore, il prof Giovanni Schiavon ed il prof. Solimini, che mi hanno dato la possibilità di entrare in questo programma di Dottorato e di svolgere questo lavoro.

Un sentito ringraziamento va all'Istituto di Fisica Applicata "Nello Carrara" IFAC-CNR di Sesto Fiorentino ed a tutte le persone straordinarie con le quali ho avuto modo di collaborare durante la mia permanenza qui (in rigoroso ordine alfabetico): Flavio Barbara, Bruno Carli, Simone Ceccherini, Samuele Del Bianco, Marco Gai, Piera Raspollini.

Un particolare ringraziamento va anche a: i cioccolatini di Andrea, le torte del Brogio, i caffè di Francesco, le chiacchiere con Chiara, Elisa, Massimo, Giacomo e Gianluca.

Vorrei ringraziare anche la mia famiglia e mio marito Davide, che mi sono stati vicini nei momenti difficili (anche se molti km ci separano, ma è una lontananza solo fisica), che mi hanno sempre sostenuta e spinta ad andare avanti.

Infine un grazie di cuore va al Dr Ugo Cortesi, la persona che forse mi è stata più vicino durante la realizzazione di questo lavoro, che mi ha guidato con i suoi saggi e preziosi consigli e che mi ha sostenuta durante i momenti più difficili, sempre con gentilezza, disponibilità ed intelligenza.