Title: The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparison and quality assessment of near-surface sensitive satellite-derived CO2 and CH4 global data sets

Order of Authors: Michael Buchwitz, Dr.; Maximilian Reuter; Oliver Schneising; Hartmut Boesch; Sandrine Guerlet; Bart Dils; Ise Aben; Raymond Armante; Peter Bergamaschi; Thomas Blumenstock; Heinrich Bovensmann; Dominik Brunner; Brigitte Buchmann; John P Burrows; Andre Butz; Alain Chedin; Frederic Chevallier; Cyril D Crevoisier; Nicholas Deutscher; Christian Frankenberg; Otto P Hasekamp; Jens Heymann; Thomas Kaminski; Alexandra Laeng; Günter Lichtenberg; Martine De Maziere; Stefan Noel; Justus Notholt; Johannes Orphal; Christoph Popp; Robert Parker; Marko Scholze; Ralf Sussmann; Gabriele P Stiller; Thorsten Warneke; Claus Zehner; Andrey Bril; David Crisp; David Griffith; Akihiko Kuze; Christopher O'Dell; Sergey Oshchepkov; Vanessa Sherlock; Hiroshi Suto; Paul Wennberg; Debra Wunch; Tatsuya Yokota; Yukio Yoshida
Highlights of manuscript

Buchwitz et al., 2012, The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparison and quality assessment of near-surface sensitive satellite-derived CO₂ and CH₄ global data sets:

- Global satellite data sets of column-averaged CO₂ and CH₄ have been assessed.
- For the first time the quality obtained using different methods has been evaluated.
- CO₂ relative biases are typically approximately 1 ppm relative to TCCON.
- CH₄ relative biases are typically in the 3-13 ppb range relative to TCCON.
- However, also differences have been identified which are not yet well understood.
The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparison and quality assessment of near-surface sensitive satellite-derived CO₂ and CH₄ global data sets


1. Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany
2. University of Leicester, Leicester, United Kingdom.
3. SRON Netherlands Institute for Space Research, Utrecht, Netherlands.
4. Belgian Institute for Space Aeronomy (BIRA), Brussels, Belgium.
5. Deutsches Zentrum für Luft- und Raumfahrt (DLR), Oberpfaffenhofen, Germany.
6. Laboratoire de Météorologie Dynamique (LMD), Palaiseau, France.
7. Karlsruhe Institute of Technology (KIT), Karlsruhe and Garmisch-Partenkirchen, Germany.
8. Swiss Federal Laboratories for Materials Science and Technology (Empa), Dübendorf, Switzerland.
10. European Commission Joint Research Centre (EC-JRC), Institute for Environment and Sustainability (IES), Air and Climate Unit, Ispra, Italy.
11. Jet Propulsion Laboratory (JPL), Pasadena, California, United States of America.
13. University of Bristol, Bristol, United Kingdom.
14. European Space Agency (ESA), ESRIN, Frascati, Italy.
15. National Institute for Environmental Studies (NIES), Tsukuba, Japan.
17. Japan Aerospace Exploration Agency (JAXA), Tsukuba, Japan.
18. Colorado State University (CSU), Fort Collins, Colorado, United States of America.
Abstract

The GHG-CCI project is one of several projects of the European Space Agency’s (ESA) Climate Change Initiative (CCI). The goal of the CCI is to generate and deliver data sets of various satellite-derived Essential Climate Variables (ECVs) in line with GCOS (Global Climate Observing System) requirements. The “ECV Greenhouse Gases” (ECV GHG) is the global distribution of important climate relevant gases – atmospheric CO$_2$ and CH$_4$ - with a quality sufficient to obtain information on regional CO$_2$ and CH$_4$ sources and sinks. Two satellite instruments deliver the main input data for GHG-CCI: SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT. The first order priority goal of GHG-CCI is the further development of retrieval algorithms for near-surface sensitive column-averaged dry air mole fractions of CO$_2$ and CH$_4$, denoted XCO$_2$ and XCH$_4$, to meet the demanding user requirements. GHG-CCI focusses on four core data products: XCO$_2$ from SCIAMACHY and TANSO and XCH$_4$ from the same two sensors. For each of the four core data products at least two candidate retrieval algorithms have been independently further developed and the corresponding data products have been quality assessed and inter-compared. This activity is referred to as “Round Robin” (RR) activity within the CCI. The main goal of the RR was to identify for each of the four core products which algorithm performs best. The algorithms selected will be used to generate the Climate Research Data Package (CRDP), which will essentially be the first version of the ECV GHG. This manuscript gives an overview about the GHG-CCI RR and related activities as conducted during the first two years of this project (September 2010 – August 2012). This activity comprises the establishment of the user requirements, the improvement of the candidate retrieval algorithms and comparisons of the satellite-derived data products with ground-based observations and models. The focus of this manuscript is on the RR approach and results, including the final RR algorithm selection decision and its justification. Comparison with ground-based Total Carbon Column Observing Network (TCCON) data indicates that the “breakthrough” single measurement precision requirement
has been met for SCIAMACHY and GOSAT XCO₂ (< 3 ppm) and GOSAT XCH₄ (< 17 ppb). The achieved relative accuracy for XCH₄ is in the range 3-15 ppb for SCIAMACHY (~3 ppb for 2003-2005 but worse afterwards due to detector degradation) and 2-8 ppb for GOSAT depending on algorithm. Meeting the 0.5 ppm systematic error requirement for XCO₂ remains a challenge: approximately 1 ppm has been achieved at the validation sites but also larger differences have been found in regions remote from TCCON. More research is needed to identify the causes for the observed differences between the various European and non-European satellite XCO₂ data products. In this context GHG-CCI suggests to take advantage of the ensemble of existing data products, for example, via the EnseMble Median Algorithm (EMMA).

Keywords: SCIAMACHY, GOSAT, Greenhouse gases, Carbon dioxide, Methane, Climate Change

1. Introduction

Carbon dioxide (CO₂) is the most important anthropogenic greenhouse gas (GHG) contributing to global warming (Solomon et al., 2007). Despite its importance, our knowledge on the CO₂ sources and sinks has significant gaps (e.g., Stephens et al., 2007, Canadell et al., 2010) and despite efforts to reduce CO₂ emissions atmospheric CO₂ continues to increase at a rate of approximately 2 ppm/year (Figure 1 top panel; see also Schneising et al., 2011, and references given therein). An improved understanding of CO₂ sources and sinks is needed for reliable prediction of the future climate of our planet (Solomon et al., 2007). This is also true for methane (CH₄, Figure 1 bottom panel).

Atmospheric methane levels increased until about the year 2000, were rather stable during ~2000-2006, but started to increase again in recent years (Rigby et al., 2008, Dlugokencky et al., 2009, Schneising et al., 2011, Frankenberg et al., 2011). Unfortunately, it is not well understood why methane was stable in the years before 2007 (e.g., Simpson et al., 2012) nor why it started to increase again (at a rate of approximately 7-8 ppb/year, Schneising et al., 2011).

Global satellite observations sensitive to near-surface CO₂ and CH₄ variations can contribute to a better understanding of the regional sources and sinks of these important greenhouse gases. Information on GHG surface fluxes (emissions and uptake) can be obtained by inverse modeling (e.g.,
Chevallier et al., 2007, Bergamaschi et al., 2009), which compares satellite observations with predictions of GHG distributions from a (chemistry) transport model (e.g., Figure 2) and uses the satellite – model mismatch to improve the surface fluxes used by the model. This or related applications, e.g., use of satellite data for Carbon Cycle Data Assimilation Systems (CCDAS) (e.g., Kaminski et al., 2010, 2012), requires satellite retrievals to meet challenging requirements, as small errors of the satellite-retrieved atmospheric GHG distributions may result in large errors of the inferred GHG surface fluxes (e.g., Meirink et al., 2006, Chevallier et al., 2005) or other model parameters.

The goal of the GHG CCI project is to generate the Essential Climate Variable (ECV) Greenhouse Gases (GHG) as defined by GCOS (Global Climate Observing System): “Product A.9: Global distribution of greenhouse gases, as methane and carbon dioxide, of sufficient quality to estimate regional sources and sinks” (GCOS, 2006). Methane and carbon dioxide measurements with (near-)global coverage are currently only available from Earth orbiting satellites. In order to get information on regional GHG sources and sinks, satellite measurements must be sensitive to near-surface GHG concentration variations. Currently only two satellite instruments deliver (or have delivered until recently) measurements which fulfill this requirement: SCIAMACHY on ENVISAT (March 2002 – April 2012) (Bovensmann et al., 1999) and TANSO-FTS on-board GOSAT (launched in January 2009, and nicknamed “Ibuki”) (Kuze et al., 2009). Both instruments perform (or have performed) nadir observations of reflected solar radiation in the near-infrared/short-wave-infrared (NIR/SWIR) spectral region, covering the relevant absorption bands of CO₂ and CH₄. They also measure the O₂ A-band to obtain the “dry-air column” for computing GHG dry-air column averaged mole fractions and/or to obtain information on clouds and aerosols. These two instruments are therefore the two main sensors used by GHG-CCI. In addition, other sensors or viewing modes are also used (e.g., MIPAS/ENVISAT and SCIAMACHY solar occultation mode for stratospheric CH₄ profiles and IASI/METOP for mid/upper tropospheric CO₂ and CH₄ columns) as they provide additional constraints for atmospheric layers above the planetary boundary layer. The focus of the first two years of the GHG-CCI project (September 2010 – August 2012) was to develop existing retrieval algorithms further, in order to improve the accuracy of the retrieved GHG data products.
Two types of algorithms are distinguished within GHG-CCI: ECV Core Algorithms (ECAs) and Additional Constraints Algorithms (ACAs). ECAs are algorithms to retrieve one or more of the four core GHG-CCI data products, which are the near surface sensitive column-averaged dry air mole fractions of atmospheric CO$_2$ and CH$_4$, denoted XCO$_2$ (in ppm) and XCH$_4$ (in ppb), retrieved from SCIAMACHY and GOSAT NIR/SWIR spectra. ACAs are algorithms to retrieve CO$_2$ and/or CH$_4$ information from satellite data which have no or only little near surface sensitivity but are sensitive to GHG variations in upper layers. All ACAs are listed in Table 3 and further discussed in Section 6.

The focus of GHG-CCI lies on ECAs and their corresponding data products, which is also the focus of this manuscript. Several existing candidate ECAs were selected at the outset of the project for ongoing development, and have been iteratively improved upon through the course of the algorithm inter-comparison and validation activity. This activity is referred to as “Round Robin” (RR) exercise within the CCI. Figure 1 shows results obtained with all GHG-CCI ECAs: the top panel shows northern hemispheric (NH) time series of XCO$_2$ and the bottom panel XCH$_4$ time series. A discussion of the used algorithms and the quality of their corresponding data products is the focus this manuscript. Here only a few general remarks shall be given to explain Figure 1 (a description and discussion of the algorithms is given in the following sections): As can be seen, the various XCO$_2$ time series generated with the different ECAs agree reasonably well. There are however also differences, e.g., a seasonal cycle amplitude differences between the two SCIAMACHY algorithms WFMD (Schneising et al., 2011, Heymann et al., 2012b) and BESD (Reuter et al., 2011). Differences are also due to the different spatial sampling of the various data products. From Figure 1 it can therefore typically not be concluded which data product is the most accurate. This requires, for example, a careful comparison with independent accurate ground-based observations (see Section 5.2). However, one obvious problem can be identified: the SCIAMACHY XCH$_4$ product generated with the IMAP algorithm (Frankenberg et al., 2011) suffers from a significant high bias (relative to several other GOSAT XCH$_4$ data products) during the year 2010 (highlighted by the dotted line). This problem is related to SCIAMACHY detector degradation issues which are not yet properly dealt with by the SCIAMACHY radiometric calibration nor compensated using IMAP (note that the second SCIAMACHY XCH$_4$ data product is not shown here).
algorithm WFMD (Schneising et al., 2011) has not yet been applied to 2010 data; the WFMD time
series “only” covers the years 2003-2009). As will be discussed in more detail below, the most
challenging problems addressed within GHG-CGI are related to achieving the required accuracy: for
XCO₂ this is a challenge because of demanding user requirements and for XCH₄ the most important
challenge was to deal with the proceeding SCIAMACHY detector degradation in the spectral region
needed for methane retrieval which started in October 2005 (see Schneising et al., 2011, and
Frankenberg et al., 2011, for a detailed discussion).

The goal of the RR was to determine which ECA performs best to generate a given GHG-CGI core
data product. The selected ECAs will be used in the third year of this project to generate the Climate
Research Data Package (CRDP), which will essentially be the first version of the ECV GHG. The
description of the RR approach and its results is the focus of this manuscript.

This manuscript is structured as follows: Section 2 presents an overview about the GHG-CGI project
followed by a description of the user requirements in Section 3. In Section 4 the retrieval algorithms
are briefly described. The main part of this manuscript is Section 5 where the RR approach and its
main results are presented and discussed. Section 6 provides a short overview about the Additional
Constraints Algorithms (ACAs) also used within GHG-CGI but not the focus of this manuscript.
Section 7 gives a short overview about the Climate Research Data Package (CRDP) to be generated
using the selected algorithms. A summary and conclusions are given in Section 8.

2. GHG-CGI project overview

The GHG-CGI project covers all aspects needed to generate the ECV GHG and to assess its quality
and the usefulness. This includes the use of appropriate satellite instruments (primarily
SCIAMACHY/ENVISAT and TANSO/GOSAT to generate global XCO₂ and XCH₄ time series),
calibration aspects (“Level 0-1 processing” related aspects, primarily for SCIAMACHY),
development and application of retrieval algorithms to convert the satellite measured spectra into
atmospheric CO₂ and CH₄ information (“Level 1-2 processing”), analysis of the resulting global data
sets including validation and user assessments focusing on inverse modeling of regional surface
fluxes.
Level 1 data (i.e., geolocated and calibrated radiances) are input data for CCI. SCIAMACHY Level 0-1 processing experts are part of the GHG-CCI team in order to provide expertise and to ensure that the findings of the study feed back to improve future Level 1 data products if necessary. Close links have been established with the GOSAT team at JAXA for GOSAT Level 1 data access, expertise and feedback.

The SCIAMACHY and GOSAT Level 1 data products are de-facto used as Fundamental Climate Data Records (FCDRs, see GCOS, 2006) despite the fact that no dedicated inter-calibration or merging efforts are currently foreseen. Consistency between the time series of the two GHG-CCI core satellites is addressed at the level of the Level 2 data products. Ideally, an ECV data product or Thematic Climate Data Record (TCDR) of a given quantity should be a single merged data record obtained from all available appropriate sensors such as SCIAMACHY and GOSAT for satellite derived XCO₂. However, within the present initial stage of this project only first steps in this direction have been carried out (see Section 5).

The ground-based validation of the satellite-derived XCO₂ and XCH₄ data products largely relies on the Total Carbon Column Observing Network (TCCON) (Wunch et al., 2010, 2011) as this network has been designed and developed for this purpose. Methods to also use data from other sources in the future (e.g., NDACC, GAW) are being developed in parallel.

A dedicated GHG-CCI Climate Research Group (CRG) has been set up to represent the users of the satellite-derived CO₂ and CH₄ data products and to provide expertise on inverse modeling of surface fluxes, CCDAS and other user related aspects. A strong link exists between GHG-CCI and the EU FP7 GMES project MACC-II (Monitoring of Atmospheric Composition and Climate - Interim Implementation, http://www.gmes-atmosphere.eu/) that provides feedback on the data quality.

Key activities carried out in the first two years of this project were the establishment of the user requirements (Section 3), the further development of retrieval algorithms (described briefly in Section 4) and data processing and data analysis with the goal to identify which algorithms perform best (“Round Robin” (RR)). The description of these RR activities and their results is the focus of this
manuscript (Section 5). In the third year of this project the selected algorithms will be used to generate
the CRDP (see Section 7), which will subsequently be validated and assessed by users.

3. User requirements

An important initial activity carried out in this project was the establishment of the user requirements. They have been formulated in detail in the GHG-CCI User Requirements Document (URD) (Buchwitz et al., 2011a). The requirements are based on peer-reviewed publications primarily prepared in the context of existing or planned satellite missions and GHG-CCI CRG user expertise and experience with existing satellite data.

Most critical are the requirements on random and systematic errors listed in Table 1. The most challenging requirement is the one on biases for XCO₂. The threshold requirement is 0.5 ppm because even errors of a few tenth of a ppm can result in large errors of the inferred CO₂ surface fluxes when used as input data for inverse modeling schemes (e.g., Chevallier et al., 2005). However, to what extent systematic errors result in biases of the inferred fluxes depends on the spatio-temporal pattern of the systematic errors. A global bias, even if considerably larger than the required 0.5 ppm, would not be critical because it can easily be detected and corrected ad hoc. Most critical are state-dependent systematic errors, which result in regional-scale (~1000 km) biases on medium time scales (~ monthly), because they will likely be missed by bias-correction schemes. As the overall impact of the atmospheric concentration error on the surface flux error depends on the spatio-temporal pattern of the concentration error, the values listed in Table 1 have to be interpreted with care. The requirements reflect what the GHG-CCI users would like to see achieved. The utility of the data can ultimately only be determined by careful analysis. The numbers listed in Table 1 serve to give a rough indication of the required uncertainties but should not be over-interpreted.

The requirements for XCH₄ are also challenging but somewhat less demanding than those for XCO₂. The main reason is that XCH₄ is more variable compared to XCO₂ relative to its background value on the spatio-temporal scales relevant for the satellite retrievals (e.g, Frankenberg et al., 2005, 2011, Meirink et al., 2006, Bergamaschi et al., 2009, Schneising et al., 2011, 2012).
4. Retrieval algorithms

In this section, a brief overview of each retrieval algorithm used for the GHG-CCI RR is given. The reader is referred to peer-reviewed publications for details. All algorithms used within the GHG-CCI RR are also described in the GHG-CCI Algorithm Theoretical Basis Document (ATBD) (Reuter et al., 2012a).

The ECV Core Algorithms (ECAs) generate one or more of the four GHG-CCI core data products, XCO₂ (in ppm) and XCH₄ (in ppb) from SCIAMACHY and GOSAT (each of the four combinations is a separate product). An overview of these algorithms is given in Table 2. Each of these algorithms is briefly described in the following.

4.1 Full Physics (FP) and Proxy (PR) algorithms

Within GHG-CCI, two types of ECAs can be distinguished: The “Full Physics” (FP) algorithms and the light path “Proxy” (PR) algorithms (see also Schepers et al., 2012).

FP algorithms model all relevant physical effects such as scattering by aerosols and clouds and have corresponding elements as part of the to be retrieved state vector. The FP algorithms obtain the dry air column-averaged mole fraction (needed to compute the dry air column-averaged mole fractions of the GHG, i.e., XCO₂ and/or XCH₄) either from the retrieved surface pressure or using meteorological information.

The PR algorithms are based on computing the dry air column-averaged mole fraction using a “reference gas”, which has to be much less variable than the gas of interest on the relevant spatio-temporal scales. The PR method is used for XCH₄ retrieval using CO₂ as a reference gas. The XCH₄ is essentially obtained from computing the ratio of the retrieved CH₄ column and the retrieved CO₂ column. The advantage of this method is that it is potentially very fast, accurate and robust (as several systematic errors cancel in the CH₄/CO₂ column ratio). The disadvantage is that a correction is needed for CO₂ variability, typically based on a global model (see, e.g., Frankenberg et al., 2005, 2011, Parker et al., 2011, Schneising et al., 2009, 2011, Schepers et al., 2012).
4.2 SCIAMACHY XCO₂ algorithms

The Weighting Function Modified (WFM) Differential Optical Absorption Spectroscopy (DOAS) algorithm (WFM-DOAS or WFMD) has been developed to retrieve vertical columns of several atmospheric gases including the GHGs discussed in this manuscript (Buchwitz et al., 2000). During the last decade, this algorithm has been significantly improved and used to generate global multi-year XCO₂ and XCH₄ data sets from SCIAMACHY (Buchwitz et al., 2005, 2007; Schneising et al., 2008, 2009). Within GHG-CCI, WFMD has been further improved and used to generate long-term consistent time series (Schneising et al., 2011, 2012, Heymann et al., 2012a, 2012b). WFMD has been implemented as a fast look-up table (LUT) based retrieval scheme to avoid time consuming radiative transfer (RT) simulations. WFMD is a least-squares method using a single constant atmospheric prior (e.g., single constant CO₂ and CH₄ mixing ratio profiles, a single aerosol scenario, no clouds). WFMD can process one orbit of SCIAMACHY observations in a few minutes on a single workstation.

Aerosols and cirrus clouds are only treated approximately by considering spectrally broad band effects by a low-order polynomial and by post-processing filtering. Overall, this results in small but significant biases, especially for XCO₂ (Heymann et al., 2012a). Recently, an improved version of WFMD has been developed for SCIAMACHY XCO₂ retrieval (Heymann et al., 2012b, see also Figure 2) and the XCO₂ data set generated with this latest version has been used for the GHG-CCI RR.

For SCIAMACHY XCH₄ retrieval, the WFMD version described in Schneising et al., 2011, 2012, has been used (see below).

The Bremen Optimal Estimation DOAS (BESD) FP algorithm was specifically developed for accurate and precise SCIAMACHY XCO₂ retrieval considering aerosols and clouds thereby overcoming limitations of the WFMD algorithm (Reuter et al., 2010, 2011). In contrast to WFMD, BESD is not based on a LUT scheme but uses on-line RT model simulations. BESD is therefore computationally much more demanding. Also, unlike WFMD, BESD is based on Optimal Estimation (OE, Rodgers, 2000) and aerosol and cirrus parameters are state vector elements and retrieved in addition to XCO₂.
4.3 GOSAT XCO$_2$ algorithms

Both GHG-CCI GOSAT XCO$_2$ retrieval algorithms are FP algorithms: the University of Leicester’s (UoL) OCO (Orbiting Carbon Observatory, Crisp et al., 2004) FP (“UoL OCO FP” or OCFP) algorithm (Cogan et al., 2012, Parker et al., 2011) and the RemoteC (or SRON Full Physics (SRFP)) algorithm (Butz et al., 2011). Both algorithms are based on adjusting parameters of a surface-atmosphere state vector and other parameters to the satellite observations, but differ in many details (different RT models, different inversion schemes (OE or Tikhonov-Phillips), different schemes for aerosol modeling and inversion, use of different pre-processing and post-processing steps, etc.) as discussed in Cogan et al., 2012, Parker et al., 2011, and Butz et al., 2011.

4.4 SCIAMACHY XCH$_4$ algorithms

For SCIAMACHY XCH$_4$ retrievals, PR algorithms are used: WFMD (Schneising et al., 2011, see above) and IMAP (Iterative Maximum A Posteriori) DOAS (Frankenberg et al., 2011). These algorithms were already well developed when GHG-CCI started but had essentially only been applied to retrieve XCH$_4$ from the first three years of the ENVISAT mission (e.g., Schneising et al., 2008). Within GHG-CCI, this time series has been significantly extended. The key challenge was (and partly still is, see Figure 1) to deal with the significant detector degradation in the spectral region needed for methane retrievals after 2005 (see Frankenberg et al., 2011, and Schneising et al., 2011, for details).

4.5 GOSAT XCH$_4$ algorithms

To overcome the key limitation of the XCH$_4$ PR algorithms, namely the need to correct the retrieved XCH$_4$ for CO$_2$ variations using a model, FP algorithms are also used within GHG-CCI, but only for GOSAT. GOSAT has higher spectral resolution than SCIAMACHY which is exploited to also retrieve scattering parameters in addition to CH$_4$. Two GOSAT XCH$_4$ FP retrieval algorithms are being used within GHG-CCI, which are also used for GOSAT XCO$_2$ retrieval (see above), OCFP (Parker et al.,
294 2011) and SRFP (Butz et al., 2011), in addition to the two PR algorithms OCPR (Parker et al., 2011) and SRPR (Schepers et al., 2012).

5. Round Robin approach and results

In this section an overview about the GHG-CCI Round Robin (RR) activities is given which have been carried out in the first two years of this project.

5.1 Round Robin approach

The ultimate goal of the GHG-CCI RR was to identify which algorithms and corresponding data products to use for generating the CRDP. This comprised the further development of existing retrieval algorithms with the goal of meeting the challenging user requirements, the application of these algorithms to generate global multi-year XCO$_2$ and XCH$_4$ sets, the comparison with ground-based reference data and inter-comparisons of the data products generated with the competing ECAs.

The selection procedure for ECAs and ACAs is described in the GHG-CCI Round Robin Evaluation Protocol (RREP, Buchwitz et al., 2011b). Initially the plan was to develop a score based selection scheme, i.e., to compute a single number for each algorithm / data product (the higher the number, the better the algorithm), mainly based on satellite – ground-based observation differences. However, this was not pursued because a scientifically sound basis for the classification could not be established. Instead a set of Figures of Merit (FoM), mostly based on differences between satellite and ground-based observations, have been defined (see RREP, Buchwitz et al., 2011b) and evaluated. However, as explained in the RREP and also shown in this manuscript, the comparison with the ground-based observations is only one component for the final selection primarily because of the sparseness of the ground-based network (see Section 5.2). Another major component of the selection procedure was the analysis of (global and regional) maps and time series, including comparisons with global state-of-the-art models, and inter-comparisons of the data products generated with the different candidate algorithms. Some key results of this activity are presented here including a summary of the main RR decision results given in Section 5.6 for ECAs and Section 6 for ACAs.
According to the initial ESA specification of the CCI RR exercise it was required to evaluate “algorithms”. However, complex algorithms such as the ones used within GHG-CCI can hardly be evaluated, especially not in terms of identifying “the best one” in terms of smallest biases. What can be done is to establish whether an algorithm suffers from obvious shortcomings such as underlying unreasonable assumptions. This however has not been found. All XCO\(_2\) algorithms, for example, use different approaches to mitigate biases due to scattering by aerosols and (thin) clouds, but it is virtually impossible to identify \textit{a priori}, e.g., based on a description of the algorithms, which of the approaches will result in the smallest XCO\(_2\) or XCH\(_4\) bias when applied to real data. What can be evaluated are the end products, i.e., the quality of the XCO\(_2\) and XCH\(_4\) data products, assuming appropriate reference data exist for comparison. As shown in this manuscript, even this is not a trivial task primarily due to the sparseness of the TCCON reference data. The GHG-CCI RR decisions are largely based on the analysis of the final data products. However, also other aspects have been investigated, e.g., analysis of retrievals based on simulated measurements (see Buchwitz et al., 2011c, 2012).

It is also important to note that the retrieval teams aimed at producing the best possible end products. This is a very reasonable approach as the quality of the end product is all that matters for the user and only the end product can be fully evaluated with respect to precision and accuracy of the targeted data product. Which input data to use and how to treat them, e.g., in a dedicated pre-processing step, has not been prescribed. Pre-processing steps may be critical for the quality of the end product. This is particularly true if the instrument shows significant degradation as is the case for SCIAMACHY after 2005 especially in the spectral region needed for methane retrieval. To deal with this, quite different approaches have been used by the two algorithms IMAP (Frankenberg et al., 2011) and WFMD (Schneising et al., 2011, 2012). For example, IMAP uses as input data spectra that have been specifically calibrated at SRON and IMAP also uses a single so-called “Dead and Bad detector Pixel Mask” (DBPM), needed to reject detector pixels which are not useful. In contrast, WFMD uses the official standard SCIAMACHY Level 1 data product with standard calibration and several DBPMs,
each optimized for a certain time period, typically covering one or more years (see Schneising et al., 2011, for details).

Finally, it is important to highlight the preliminary nature of the RR. This is due to the fact that all Level 1 input data and retrieval algorithms are continuously being improved. An algorithm currently identified to be the best one will not necessarily be the best one in the future. GHG-CCI therefore needs to be flexible and will aim to consider this in future phases of the CCI.

5.2 Comparison with ground-based observations

The most relevant ground-based observations for the validation of the satellite-derived XCO₂ and XCH₄ data products are the corresponding data products of the TCCON. Detailed validation results will be reported elsewhere (Dils et al., manuscript in preparation), therefore we here give only a short overview highlighting major findings.

Figure 3 shows a comparison of the four GHG-CCI core data products generated with two or more of the candidate algorithms at the 10 TCCON sites listed in Table 4. The results shown in Figure 3 are based on a direct comparison of the co-located satellite and TCCON data products. No correction for different a priori profiles and averaging kernels has been applied. The time period used for comparison depends on the availability of the satellite data products (see Figure 1) and on the TCCON site (see Table 4). Furthermore, the time periods are limited to time periods where the alternative products of the corresponding candidate algorithms are available (SCIAMACHY: XCH₄: 2003-2009, XCO₂: 2006-2009; GOSAT: mid 2009-2010). The results shown in Figure 3 have been generated using a spatio-temporal co-location criterion of 2 hours and 500 km (for alternative co-location criteria see Notholt et al., 2012). In Figure 3 several numerical values are given, which are also listed in Table 5, computed from satellite minus TCCON differences for each single satellite retrieval and the corresponding TCCON mean value. On the left hand side of Figure 3 the mean satellite-TCCON differences are shown for each of the 10 TCCON sites and all four core data products and their corresponding ECAs. For each ECA the standard deviation of the station-to-station bias has been computed (“StdDev”) and the total number of co-located satellite retrievals used for comparison (“N”). The standard deviation of the station-to-station bias is interpreted as a relevant measure of the
systematic error ("relative accuracy" or "relative bias"). The standard deviation is more relevant to characterize systematic errors compared to, for example, the mean difference. Most critical is to achieve high "relative accuracy" (or low "relative bias") not necessarily high "absolute accuracy" (although this would of course be better). For example, a constant offset of the satellite data would not be critical if the data are being used for surface flux inverse modeling (see Section 3) and this is considered by computing the standard deviation. On the right hand side of Figure 3 the standard deviations of the satellite-TCCON differences are shown for each TCCON site. They are a measure of the random error (scatter) of the satellite retrievals. The corresponding mean value over all TCCON sites is used to characterize the mean random error (or "precision") of the corresponding satellite data product. In this context it needs to be pointed out that the estimated systematic and random errors of the satellite retrievals as reported here are upper bounds as the TCCON data used as reference are not free of errors. The systematic and random errors of single TCCON data products are typically 0.4 ppm for XCO\textsubscript{2} (1-sigma) and 4 ppb (1-sigma) for XCH\textsubscript{4} (Notholt et al., 2012, based on Wunch et al., 2010).

The comparison of the two SCIAMACHY XCH\textsubscript{4} retrieval algorithms WFMD and IMAP with TCCON shows the following (Figure 3, first row): Overall, the systematic differences (biases) with respect to TCCON vary from site to site from nearly 0 ppb at Lamont to 20-30 ppb at the southern hemisphere (SH) sites Darwin, Wollongong, and Lauder, but are very similar for WFMD and IMAP. The reason for the large biases at these SH sites have not yet been identified. Agreement is within +/- 10 ppb if these SH sites are excluded. In order to obtain an estimate of the relative biases (i.e., considering that an overall offset is not critical), the standard deviation of the station-to-station biases has been computed: it amounts to 11 ppb for WFMD and 15 ppb for IMAP. The standard deviation of the satellite-TCCON differences (Figure 3, first row, right panel), which is a measure of the single measurement precision (1-sigma), is on average 82 ppb for WFMD and 50 ppb for IMAP. Because nearly all TCCON sites started operation after 2005 (see Table 4), i.e., after loss of important SCIAMACHY methane detector pixels due to detector degradation, the values listed for SCIAMACHY in Figure 3 are not representative for the years 2003-2005. Until the end of 2005 the performance was much better and the corresponding values are listed in curved brackets in Table 5.
possible explanation for the larger scatter (worse precision) of WFMD after 2005 is that WFMD is an unconstrained least-squares algorithm whereas IMAP is based on Optimal Estimation and uses detailed CH$_4$ information (as a function of latitude, altitude and time but not longitude) from a global model as a priori information. This raises the question why the precision of the two data products is similar for 2003-2005. This could be related to the fact that only a single DBPM is used by IMAP whereas WFMD has used a DBPM optimized for 2003-2005. Another possible explanation could be the use of differently calibrated input data. The number of satellite soundings used varies significantly from site to site (~400-20000, not shown), but is very similar for WFMD (N=37628) and IMAP (39489) (at least at TCCON sites, for other locations this may not be true, see Figure 4).

The comparison of the four GOSAT XCH$_4$ retrieval algorithms (OCPR, OCFP, SRPR, SRFP) with TCCON shows the following (Figure 3, second row): The biases depend on the TCCON site but are in the range +/- 15 ppb. The estimated relative bias is best for OCPR (2 ppb) and worst for OCFP (8 ppb). The largest number of data points has OCPR (followed by SRPR). The number of data points is higher for the PR algorithms (OCPR and SRPR) compared to the FP algorithms (OCFP and SRFP).

The FP algorithm with the lowest relative bias is SRFP (3 ppb). The PR algorithm with the lowest relative bias is OCPR (2 ppb). The standard deviation of the satellite – TCCON differences are nearly identical for all four algorithms (Figure 3, second row, right panel).

The comparison of the two SCIAMACHY XCO$_2$ retrieval algorithms WFMD and BESD with TCCON shows the following (Figure 3, third row): BESD has typically lower systematic errors (0.7 ppm) compared to WFMD (1.3 ppm) and also a higher precision (2.3 ppm compared to 5.1 ppm). Ultimately it can be expected that the biases of BESD will be even lower as it has been identified (not shown) that the BESD RR data set suffers from problems related to the SCIAMACHY Level 1 data product used (version 7 consolidation level u, “L1v7u”). This data product was used because it was the latest version available when the final RR data set has to be generated and because it also covers the time period after 2009. The previous Level 1 version 6 (L1v6), used by WFMD, does not suffer from these problems but is only available until end of 2009, where the WFMD data set ends. It has been found that BESD retrievals for selected months using the improved new version L1v7w have much
lower biases especially because the many outliers caused by the L1v7u spectra are not present any more (not shown). It is therefore necessary and planned to reprocess the entire SCIAMACHY data set with BESD using L1v7w, e.g., for the generation of the CRDP. A potentially important pro for WFMD for certain applications is the much larger number of data points.

The comparison of the two GOSAT XCO₂ retrieval algorithms OCFP and SRFP with TCCON shows the following (Figure 3, bottom row): The biases depend on site and are typically in the range +/- 1 ppm. They are very similar for both algorithms. This is also true for the standard deviation of the difference between the GOSAT and TCCON estimates, which is typically in the range 2-3 ppm. The number of co-locations is also nearly identical for both algorithms but varies significantly from site to site (from ~20 at Lauder to ~1000 at Lamont, not shown).

As shown in Table 5, the SCIAMACHY XCH₄ product for 2003-2005 meets the threshold precision requirement (but not for 2006 and later years due to the detector degradation). In contrast, the GOSAT XCH₄ has a much higher precision and even the breakthrough precision requirement is met by all algorithms. All GOSAT XCH₄ algorithms meet the relative accuracy (relative bias) user requirement - some are close to or even better than the goal requirement. For SCIAMACHY this is only true for 2003-2005.

The precision requirement for XCO₂ is also met by all algorithms. WFMD meets the threshold requirement and the other algorithms including BESD even meet the breakthrough requirement. The challenging 0.5 ppm bias requirement has however not yet been met but several algorithms achieve a performance close to the threshold requirement (0.6-0.9 ppm, depending on algorithm).

Concerning the final RR algorithm selection decision, it is important not to over-interpret the numerical values listed in Table 5 due to the sparseness of the TCCON sites. For this and other reasons, the TCCON comparisons presented and discussed in this section are only one key component of the GHG-CCI RR activities. Therefore, more comparisons have been conducted as described in the following sub-sections.
5.3 Inter-comparison of SCIAMACHY XCH$_4$ data products

The multi-year global retrievals obtained from the two SCIAMACHY XCH$_4$ algorithms, WFMD and IMAP, have been compared with one another. Figure 4 shows, as a typical example, a comparison of one month (August 2005) of the global WFMD and IMAP data products (similar figures but for other months are shown in Buchwitz et al., 2012). As can be seen, the monthly XCH$_4$ maps generated with the two algorithms show – depending on region - similar but also significantly different pattern. Both maps show higher methane concentrations over the Northern Hemisphere (NH), where most of the methane sources are located, compared to the Southern Hemisphere (SH). Both data sets agree reasonably well (within typically +/- 10 ppb) over most parts of the SH land areas but over some areas WFMD XCH$_4$ can be up to approximately 20 ppb higher. Over the NH the situation appears to be more complex. Both data sets show elevated methane over large parts of China, south-east Asia and India, but the patterns are not identical, with WFMD being higher over south-east Asia and lower over parts of India compared to IMAP. WFMD and IMAP not only use differently calibrated input data (standard versus non-standard calibration) and different retrieval methods (least squares versus OE), but also different post-processing quality filtering schemes. The latter is reflected by differences in spatial coverage (e.g., WFMD methane is not restricted to land observations only) and number of retrievals over a given region (see right hand side panels of Figure 4). The data density differs significantly depending on region. Typically WFMD has much more data points over Sahara and other areas in the ~10°-40°N latitude range but also over mid/northern Australia and the mid/western part of the US, whereas IMAP has higher data density over South America and mid/high northern latitudes. Large differences between the two data sets are also visible over large parts of northern Africa, where IMAP methane is higher (by approx. 40 ppb) and Greenland, where WFMD methane is higher (by approx. 40 ppb). The reasons for the differences have not yet been identified. It has also not yet been assessed to what extent inferred regional methane fluxes would differ depending on which data set is used as input data for inverse modeling of regional methane fluxes. Significant differences can be expected as the regional differences exceed the bias threshold requirement of less than 10 ppb. The discussion also shows that depending on region the differences can be significantly larger than the
estimated biases listed in Table 5, which are based on the analysis of the satellite data at TCCON sites only. Clearly, more research is needed to understand the differences between the two SCIAMACHY methane data sets discussed in this section.

5.4 Inter-comparison of GOSAT $\text{CH}_4$ data products

Within GHG-CCI, four GOSAT $\text{CH}_4$ retrieval algorithms have been further developed and used to generate global data sets which have been inter-compared and compared with TCCON retrievals and global model data (Buchwitz et al., 2012). The four retrieval algorithms are the FP and PR algorithms developed by SRON (SRFP, SRPF) and Univ. Leicester (UoL; OCFP and OCPR algorithms).

For the PR algorithms, which are based on the retrieval of ratios of the $\text{CH}_4$ to $\text{CO}_2$ columns, followed by a model-based $\text{CO}_2$ correction to compute $\text{XCH}_4$, the column ratios have been compared as well as the final $\text{XCH}_4$ product. As expected, it has been found that the agreement between the ratios is typically somewhat better compared to the $\text{XCH}_4$ products due to differences between the model-based $\text{CO}_2$ correction as used by SRON and UoL (see Buchwitz et al., 2012, for details). Overall and in-line with the discussion presented in Section 5.2, it has been found that the two PR products agree nearly equally well with the TCCON ground-based observations. A direct comparison of the two data products at TCCON sites is also shown in Figure 5 indicating agreement within typically 10 ppb (1-sigma). Nevertheless, inspection of global maps also reveals significant differences, depending on region and time. This is similar to the results found for the SCIAMACHY data sets discussed in the previous section, although the differences shown in Figure 5 for GOSAT are smaller compared to the differences for SCIAMACHY shown in Figure 4. Figure 5 shows a global OCPR-SRPR methane difference map for July 2009. As can be seen, the differences may exceed 10 ppb (approx. 0.5%) over certain extended regions such as parts of North and South America and central Africa. Comparisons between the two FP GOSAT $\text{CH}_4$ data products OCFP and SRFP have also been carried out. Using SRFP, two years of global GOSAT data have been retrieved but the comparison had to be limited to TCCON sites only because of limitations of the OCFP data set which is not yet available globally. It has been found that the inter-station bias is smaller for SRFP ($\approx 4$ ppb) compared to OCFP ($\approx 8$ ppb) and that the scatter of the SRFP data is somewhat smaller compared to the OCFP (14 ppb versus 16...
These findings are consistent with the results presented in Table 5 but have been derived independently (see Buchwitz et al., 2012). It has also been found that the agreement between the two PR algorithms is significantly better than the agreement between the two FP algorithms. This may be due to the fact that PR algorithms are simpler but may also indicate that at the current stage of development the PR algorithms are more mature (note that they also deliver much more data points, see Section 5.2).

5.5 Inter-comparison of XCO₂ data products

Within GHG-CCI two algorithms have been further developed to retrieve XCO₂ from SCIAMACHY, namely WFMD and BESD, and two algorithms to retrieve XCO₂ from GOSAT, namely OCFP and SRFP. In addition, there are three non-European GOSAT algorithms presented and discussed in the peer-reviewed literature whose data products have also been used for comparison: (i) the official operational GOSAT algorithm (v02.xx) developed at the National Institute for Environmental Studies (NIES) in Japan (Yoshida et al., 2011; in the following referred to as “NIES” algorithm), (ii) a scientific algorithm called PPDF (Pathlength Probability Density Function) also developed at NIES (Bril et al., 2007, Oshchepkov et al., 2008, 2009, 2011), and (iii) NASA/JPL’s ACOS (Atmospheric CO₂ Observations from Space) v2.9 algorithm (O’Dell et al., 2012, Crisp et al., 2012).

The global XCO₂ data products from all 7 algorithms have been inter-compared within GHG-CCI (Reuter et al. 2012b, Buchwitz et al., 2012). The analysis revealed the following: The various satellite XCO₂ data products all capture the expected large scale variations of atmospheric CO₂ such as the time dependent north-south gradient (Figure 6) and the CO₂ increase and seasonal cycle (Figure 1) but exhibit differences in the spatio-temporal pattern which – depending on region and time – may exceed the relative bias user requirement of 0.5 ppm.

Typical examples are shown in Figure 6 and Figure 7. Figure 6 shows comparisons of the four GHG-CCI XCO₂ algorithms (BESD, WFMD, SRFP (= RemoteC), OCFP (= UOL-FP)). Figure 7 shows the GHG-CCI algorithms as well as the three non-European algorithms mentioned above (ACOS (v2.9), PPDF (NIES PPDF-D), and NIES (v02.xx)) for the two months September 2009 and May 2010. Also shown is the ensemble data product generated with the EnseMble Median Algorithm (EMMA).
algorithm, discussed below, TCCON XCO₂, and XCO₂ from NOAA’s CO₂ assimilation system CarbonTracker (CT) (Peters et al., 2007). As can be seen, all satellite retrieval algorithms capture the north-south XCO₂ gradient, which is significantly different for the two months shown, in good to reasonable agreement with TCCON and CarbonTracker (Figure 7). As can also be seen, differences between the data products often exceed 0.5 ppm, particularly at locations remote from TCCON sites (e.g., Sahara, South America, Africa). As discussed in Section 5.2, it appears virtually impossible using TCCON, to determine which algorithm performs best at least for GOSAT (for SCIAMACHY it has been shown that BESD outperforms WFMD in terms of single measurement precision and bias not however in terms of number of observations, which is significantly higher for WFMD). It is also likely that “a best algorithm” for all conditions does not exist at present as each algorithm is expected to have its strength and weaknesses. Therefore, which algorithm performs best may depend on the spatio-temporal interval of interest. Clearly, more research is needed to understand the differences between the various XCO₂ data sets shown in Figure 6 and Figure 7.

The situation appears to be similar to that for climate modeling: it is not clear which “model” is the best and (remote from TCCON) there is no truth to compare with. A promising approach to deal with this is to make use of the fact that several state-of-the-art algorithms and corresponding XCO₂ data products are available, i.e., an ensemble of data products, which can be exploited. This is the underlying idea of the EnseMble Median Algorithm (EMMA, Reuter et al., 2012b). The strength of using an ensemble of satellite data products was highlighted at the end of the first year of the GHG-CCI project (Buchwitz et al., 2011c), when biases (0.5%) between Bialystok TCCON XCO₂ and coincident satellite data were identified in the majority of algorithms participating in the GHG-CCI. This bias occurred due to an empirical correction of known magnitude, to account for a laser-sampling bias in the FTS data before September 21, 2009, inadvertently being applied in the wrong direction. A bias in XCH₄ in the early part of the Bialystok time series that occurred due to missing fits in one of the CH₄ micro-windows, was also brought to light by comparisons to the ensemble of satellite retrievals. The identification and quantification of these biases would most likely not have been possible with a
single algorithm / data product, due to difficulty in proving that such relatively small differences are not due to possible retrieval algorithm issues.

A detailed description of EMMA is presented in Reuter at al., 2012b. Therefore here only a short overview is given. The presented version of EMMA (v1.3a) uses the 7 individual satellite XCO₂ products shown in Figure 7 and generates a Level 2 product (i.e., a product containing the XCO₂ of the individual satellite soundings including uncertainty estimate and other information such as averaging kernels) using the median in each 10°x10° monthly grid cell (“voxel”). In short, EMMA works as follows: For each voxel, the mean XCO₂ value is computed for each of the 7 individual data products. The median of the 7 mean values determines which of the individual satellite Level 2 data products is used for the EMMA data product for that voxel (if a certain voxel is not covered by all 7 data products, a smaller number of data products is used). Using the median has several advantages compared to, for example, using the mean value. A key aspect is that the median is robust with respect to outliers. Using the median essentially removes outliers. This is of critical importance as each of the individual data products appears to suffer from outliers but where they appear and when is not known a priori and depends on the algorithm. Of at least equal importance is that the GHG-CCI users need a Level 2 data product (individual soundings) and not a Level 3 data product (e.g., gridded monthly averages). Furthermore, the use of an ensemble of data products possibly permits the generation of more reliable uncertainty estimates, obtained from a combination of the ensemble scatter and the reported uncertainties of the individual algorithms (which are primarily estimates of the random uncertainty). This would in particular be important to get a handle on the systematic error component of the uncertainty, which is very difficult (if not impossible) to reliably quantify for each algorithm individually. For an ensemble, this would strictly speaking require that the median is bias free which is unlikely the case. Nevertheless, the spatio-temporal intervals where the various data products disagree are very likely intervals where the data products need to be used with care. In any case, reliable XCO₂ error estimates of the satellite retrievals are of critical importance for the user of the GHG-CCI atmospheric data products.
Figures such as Figure 7 also permit the determination of which of the data sets agree and which disagree. For example, the EMMA product, but also most of the individual GOSAT products and BESD agree well or at least reasonably with each other as well as with TCCON and CarbonTracker (see green and yellow smileys), whereas this is not always true for the two very fast algorithms WFMD and PPDF (see red smileys). Figure 8 shows pie charts indicating the overall agreement and disagreement of each of the individual algorithms with the median. The results are consistent with the already reported findings, e.g., better performance of BESD compared to WFMD and similar performance of the GOSAT XCO₂ algorithms.

A large number of other comparisons of the individual data products and the EMMA product with TCCON but also with CarbonTracker have been carried out. Figure 9 shows, as an example, a comparison of the amplitude of the XCO₂ seasonal cycle. As can be seen, all satellite data sets suggest that the seasonal cycle is underestimated by CarbonTracker. All algorithms whose data products agree with TCCON within their error bars (i.e., all algorithms except WFMD and PPDF) indicate that this underestimation is ~1.5 +/- 0.5 ppm peak-to-peak. Using only a single data product it would be difficult to “prove” that such a relatively small difference (only ~0.3% of the total column) is significant and not caused by or at least significantly influenced by retrieval issues (see, e.g., the discussion given in Schneising et al., 2011, on this topic). Using an ensemble of data products based on more than one satellite and using several essentially independent algorithms allows drawing more confident conclusions with respect to the interpretation of satellite – model XCO₂ differences than possible using a single data product only. Within GHG-CCI it is therefore planned to continue the efforts on EMMA in addition to further developing the individual algorithms.

5.6 Algorithm selection results

The main goal of the RR exercise was to determine which satellite retrieval algorithms to use to generate the CRDP. Based on the results presented and discussed in the previous sections, the following algorithms have been selected:
5.6.1 SCIAMACHY XCH$_4$

Data products generated with two algorithms have been assessed: WFMD (Schneising et al., 2011, 2012) and IMAP (Frankenberg et al., 2011). Comparison with ground-based TCCON observations revealed that both data products are very similar with respect to biases. This is also true for the estimated single measurement precisions for the time period 2003-2005 where the SCIAMACHY detector in the spectral region needed for methane retrieval did not suffer from major degradation in contrast to later years. After 2005, the WFMD methane shows a larger scatter (~80 ppb) compared to IMAP (~50 ppb). Both data products have to be used with care for the time after 2005 due to potential bias issues related to detector degradation as indicated by the TCCON comparison at southern hemisphere TCCON sites, where both data products show a low bias of 20-30 ppb depending on FTS site. Considering only this analysis, one would conclude that both data products are essentially equivalent and one may therefore select one of them, e.g., IMAP, because of the lower scatter after 2005 and because IMAP is currently the de-facto standard because it has been used for several peer-reviewed publications discussing methane emission related aspects (e.g., Bergamaschi, 2007, 2009, Bloom et al., 2010) and because the IMAP product is also the SCIAMACHY methane product used by the European MACC-II project (http://www.gmes-atmosphere.eu/). Analysis of spatially resolved global methane distributions as generated by the two algorithms however shows significant differences, depending on region and time, which are larger than the required maximum bias of 10 ppb, i.e., are significant for regional-scale methane surface flux inversions. Due to the lack of appropriate reference data such as TCCON, it was not yet possible to determine which of the two data products is the most accurate. Therefore, it has been decided to proceed with both algorithms and to contribute with both alternative data products to the CRDP pointing out the strength and weaknesses of the two approaches. Users will be encouraged to use both data sets, to determine to what extent their findings depend on the data product used, and to report these findings to the GHG-CCI retrieval experts.
5.6.2 GOSAT XCH$_4$

Four algorithms and their corresponding data products have been evaluated: OCFP and OCPR (Parker et al., 2011) and SRFP and SRPR (Butz et al., 2011). All data products show very similar biases and scatter when compared with ground-based TCCON observations. The number of data points is however significantly higher for the “Proxy” (PR) algorithms compared to the “Full Physics” (FP) algorithms and the agreement between the two PR data products is better than for the FP products, indicating a higher level of maturity of the (simpler) PR algorithms. Note that the SCIAMACHY XCH$_4$ algorithms, WFMD and IMAP, are also PR algorithms and that the FP algorithms are relatively new and currently in their early stages of development. Overall, the OCPR algorithm shows a slightly better performance compared to SRPR (primarily in terms of number of data points at TCCON sites). It has therefore been decided to continue with OCPR within GHG-CCI. The PR XCH$_4$ algorithms depend on a CO$_2$ correction using model data. The long-term goal of GHG-CCI is to use a FP algorithm that is independent of a CO$_2$ model. The SRFP FP algorithm shows a somewhat better performance compared to the OCFP algorithm (e.g., lower station-to-station biases at TCCON sites), but also has a lower number of data points compared to OCFP. It has therefore been decided to continue with the SRFP algorithm. In summary, four GOSAT XCH$_4$ algorithms have been evaluated as part of the GHG-CCI RR and two of these algorithms have been selected for further use within GHG-CCI: OCPR and SRFP.

5.6.3 SCIAMACHY and GOSAT XCO$_2$

Within GHG-CCI, two SCIAMACHY and two GOSAT XCO$_2$ algorithms have been further developed and the corresponding data products have been inter-compared. They have also been compared with three other GOSAT XCO$_2$ data products generated outside of this project: with the two GOSAT XCO$_2$ products generated at NIES, Japan, (i.e., the operational GOSAT product (Yoshida et al., 2011) and the scientific PPDF product (Oshchepkov et al., 2011)) and with the NASA ACOS team product (O’Dell et al., 2012, Crisp et al., 2012). Analysis of all seven products indicates that the precision requirement has been met, but not the very demanding bias requirement of less than 0.5 ppm.
(approximately 1 ppm has been achieved at TCCON sites). Clearly, more work on the individual retrieval algorithms is required to achieve this goal and it has been decided to continue with all algorithms. A possible exception is the fast SCIAMACHY XCO$_2$ WFMD algorithm, which shows a reduced data quality in terms of precision and biases compared to the computationally much more demanding BESD algorithm. On the other hand the WFMD product has significantly (3-4 times) more data points compared to BESD and therefore much better coverage compared to any of the other data products including BESD. GHG-CCI aims at taking advantage of the fact that an ensemble of state-of-the-art data products exists which can be exploited. To this end, the EnseMble Median Algorithm (EMMA) has been developed (Reuter et al., 2012b). EMMA generates a Level 2 XCO$_2$ product using the median of the individual data products thereby largely eliminating outliers of the data products generated with the individual algorithms. EMMA may also improve the error characterization using the ensemble scatter. Preliminary analysis indicates that EMMA outperforms each of the individual algorithms. EMMA also permits the identification potential weaknesses of the individual algorithms, which can be used to improve the individual algorithms. Taking this into account, it has been decided to proceed with all satellite XCO$_2$ algorithms and to add the EMMA data product to the GHG-CCI product portfolio.

6. **Additional Constraints Algorithms (ACAs)**

The Additional Constraints Algorithms (ACAs) are algorithms to retrieve CO$_2$ and/or CH$_4$ information for layers above the planetary boundary layer. ACAs are applied to several satellite instruments. An overview about the ACAs used within GHG-CCI is given in Table 3. As the ACA are not the focus of this manuscript the reader is referred to the references listed in Table 3 (including caption) for details on each of these algorithms and corresponding data products.

For ACAs only one algorithm per data product has been considered within GHG-CCI, i.e., ACAs are also being further developed but not in competition and not by covering all aspects (e.g., no dedicated validation). For ACAs a number of criteria have been defined which need to be fulfilled to contribute to the CRDP but detailed user requirements have not been formulated.
Only a limited assessment of the data products generated with ACAs has been conducted during the initial phase of GHG-CCI described in this manuscript because the focus was on ECAs. However, for each of the ACAs listed in Table 3 it has been determined if the selection criteria specified in the Round Robin Evaluation Procedure (RREP, Buchwitz et al., 2011b) have been met. The RREP defines 11 criteria for ACAs which need to be fulfilled for a given ACA to contribute to the CRDP. The criteria are mostly qualitative and refer to a required minimum level of documentation, error analysis and related auxiliary information. All ACA products are potentially useful for GHG-CCI climate applications as they deliver additional information on CO₂ and/or CH₄ thereby providing potentially important constraints when used, for example, within an appropriate inverse modeling framework to derive regional surface fluxes from the satellite observations. However, no detailed user requirements are currently available, no dedicated validation has been performed within GHG-CCI and it has also not been assessed to what extent the existing products are useful or not useful for GHG surface flux inverse modeling. More research is needed to assess the usefulness of these data products for climate relevant applications. It has been identified that all ACAs fulfill the requirements listed in the RREP and that all ACA products can therefore be included in the CRDP.

7. Climate Research Data Package (CRDP)

The goal of the GHG-CCI RR was to decide which algorithms to use to generate the CRDP. The CRDP will be generated during September 2012 to end of February 2013. Table 6 presents an overview of the planned content of the CRDP. The CRDP will be validated during March-May 2013 and subsequently evaluated by the GHG-CCI users (June-August 2013). By the end of August, the CRDP along with the corresponding documentation will be made publicly available via the GHG-CCI website.

8. Summary and conclusions

An overview of the main activities and results achieved during the first two years of the GHG-CCI project of ESA’s Climate Change Initiative (CCI) has been presented, focusing on the CCI “Round Robin” (RR) exercise. The goal of CCI is to generate a number of Essential Climate Variables (ECVs)
in-line with GCOS (Global Climate Observing System) requirements and guidelines using European
Earth observation data and data from ESA Third Party Missions (TPM) such as GOSAT. To achieve
this, several existing state-of-the-art retrieval algorithms for retrieving XCO₂ and XCH₄ from
SCIAMACHY/ENVISAT and TANSO/GOSAT nadir radiance spectra, have been further improved in
order to meet challenging requirements for the targeted regional CO₂ and CH₄ surface flux
(source/sink) application as defined by the GHG-CCI Climate Research Group (CRG). The ultimate
goal of the RR was to identify and select the best algorithms to be used for generating the Climate
Research Data Package (CRDP), which will essentially be the first version of the CCI ECV GHG data
base. In addition, retrieval algorithms for a number of other satellite instruments such as IASI and
MIPAS have also been further developed, but not in competition.

Substantial progress has been made during the first two years (September 2010 – August 2012) of the
GHG-CCI project. For example, longer XCO₂ and XCH₄ time series have been generated from
SCIAMACHY with improved data quality and better error characterization (Reuter et al., 2011,
Frankenberg et al., 2011, Schneising et al., 2011, 2012, Heymann et al., 2012a, 2012b). The same is
ture for GOSAT (Butz et al., 2011, Parker et al., 2011, Schepers et al., 2012, Cogan et al., 2012,
Cogan et al., 2012).

Several retrieval algorithms have been further developed in competition during the GHG-CCI RR and
used to generate global multi-year data sets of XCO₂ and XCH₄ from SCIAMACHY and GOSAT.
The data products have been evaluated by comparison with ground-based TCCON observations, by
inter-comparisons of the data products generated with the different candidate algorithms, and by
comparisons with other data sets including global models. Due to the sparseness of the TCCON
network it was not planned to base the algorithm selection decision only on satellite – TCCON
comparisons. It has been found that nearly all candidate algorithms produce data with very similar
quality at TCCON sites, i.e., show similar satellite – TCCON differences. Significant differences have
however been found remote from TCCON when comparing the global data sets, e.g., when comparing
global maps. Depending on region and time, it has been found that the differences may significantly
exceed the systematic error requirements of less than 0.5 ppm for XCO₂ and 10 ppb for XCH₄. It has
been identified that more research is needed in order to understand the differences between the various data sets. It was not possible for all products to clearly identify which of the candidate algorithms performs best.

For GOSAT XCH₄, two of the four retrieval algorithms have been selected for further development and use within GHG-CCI, namely OCPR and SRFP, each representing a certain type of algorithm ("Proxy" (PR) or "Full Physics" (FP)). Each of the two approaches has different strength and weaknesses. At present the PR algorithms appear to be more mature, e.g., due to the SCIAMACHY heritage, and also deliver more data points but the long-term goal of GHG-CCI is to use a FP algorithm. The reason is that the FP algorithms do not rely on CO₂ model data in contrast to the PR algorithms.

For SCIAMACHY XCH₄, two algorithms have been used and assessed, namely WFMD and IMAP, and it has been decided to continue with both algorithms. The main reason is that it has been found that the data products generated with the two algorithms significantly differ over certain regions and certain time periods but it was not yet possible to identify which algorithm has the lowest biases as the systematic differences at the TCCON validation sites are very similar for both algorithms.

The situation is similar for the XCO₂ algorithms. More work is needed on the individual algorithms in order to meet the challenging relative bias user requirement (< 0.5 ppm) under all conditions. As pointed out when presenting and discussing the user requirements failure to meet this requirement does not imply that the data are useless because the usefulness depends on the spatio-temporal error structure. Nevertheless, GHG-CCI aims at meeting this requirement for as many conditions as possible. Four algorithms have been assessed within GHG-CCI, two for SCIAMACHY and two for GOSAT. In addition, the data products have also been compared with the non-European data products generated at NIES and NASA. Furthermore, the EnseMble Median Algorithm (EMMA) has been developed (Reuter et al., 2012b) based on the statistical properties of the ensemble of products. The EMMA product has important potential advantages compared to the individual products: outliers can be virtually eliminated and the error characterization can be improved by considering the ensemble spread (possibly by providing more reliably error estimates but at least by identifying spatio-temporal
intervals where the various data products disagree indicating intervals where the data products have to be used and interpreted with care). Still bearing in mind the limitations of multi-model ensembles (e.g., Knutti et al. 2010), EMMA shall also be used to identify strengths and weaknesses of the individual products and this information can be used to improve the individual algorithms. It therefore has been decided to add the EMMA XCO₂ product to the GHG-CCI product portfolio. EMMA is in development and will be further improved, e.g., by considering improved products generated with the individual algorithms. The development of EMMA started rather late in this project. Therefore the EMMA product has not been evaluated using the same procedure as used for the other GHG-CCI XCO₂ products. It remains to be shown if a future EMMA product meets the 0.5 ppb XCO₂ bias requirement. EMMA is an intermediate product as the ultimate goal is to select a single algorithm for XCO₂. In this context it needs to be pointed out that single algorithms also have certain potential advantages compared to an ensemble product, e.g., better or at least more straight forward bias correction.

The focus of GHG-CCI are the near-surface sensitive XCO₂ and XCH₄ data products obtained from SCIAMACHY and GOSAT nadir observations, as these data contain most of the information on regional GHG sources and sinks. However, a number of other satellites or satellite instruments (AIRS, IASI, MIPAS, ACE-FTS) and viewing modes (SCIAMACHY solar occultation) have also been considered within GHG-CCI but with lower priority. These algorithms are only shortly discussed in this manuscript. For each of these algorithms (see Table 3), a number of criteria have been defined in order to determine if the corresponding data product should be included in the CRDP or not. As all algorithms meet the criteria listed in the GHG-CCI Round Robin Evaluation Protocol (RREP, Buchwitz et al., 2011b) it has been decided to continue with all these algorithms. However, a detailed quantitative assessment as detailed for the XCO₂ and XCH₄ core products has not been performed for these algorithms and their corresponding data products and also detailed user requirements have not been established.

A summary of the RR findings and conclusions including justification for algorithm selection is given in Section 5.6 “Algorithm selection results”. Section 7 contains a table presenting an overview about
the planned content of the CRDP (Table 6). Due to not yet well understood differences between the
data products generated with the different algorithms it is at present not possible to select a single
“best” algorithm for each data product. More research is needed to achieve this goal.
What is needed are as accurate and precise as possible and as long-term as possible global near surface
sensitive CO₂ and CH₄ data sets fulfilling the needs of the climate and carbon (inverse) modeling user
community. The goal of GHG-CCI is to build up such a time series starting with
SCIAMACHY/ENVISAT (March 2002 – April 2012) and being continued with GOSAT (launch
2009) and future GHG satellite missions such as OCO-2 (Boesch et al., 2011), Sentinel-5-Precurso
(Butz et al., 2012) and possibly CarbonSat (Bovensmann et al., 2010).

9. Acknowledgements

This work was primarily funded by ESA/ESRIN (GHG-CCI) but also received funding from EU FP7
(grant agreement No. 283576, MACC-II), DLR (SADOS), and the State and the University of
Bremen. We thank the members of the GOSAT Project (JAXA, NIES, and Ministry of the
Environment (MoE), Japan) for providing GOSAT Level 1B and Level 2 data products (GOSAT RA1
PI project CONSCIGO). The ACOS v2.9 data were produced by the ACOS/OCO-2 project at the Jet
Propulsion Laboratory, California Institute of Technology, and obtained from the ACOS/OCO-2 data
archive maintained at the NASA Goddard Earth Science Data and Information Services Center. We
thank NOAA for making available the CarbonTracker CO₂ fields. We also thank TCCON and related
funding organizations (NASA grants NNX11AG01G, NAG5-12247, NNG05-GD07G, NASA
Orbiting Carbon Observatory Program, DOE ARM program, the Australian Research Council,
DP0879468 and LP0562346, the EU projects IMECC and GEOmon, the Senate of Bremen).
10. References


of the GHG-CCI project of ESA’s Climate Change Initiative, Technical Report, version 1


GCOS (Global Climate Observing System) (2006), Systematic Observation Requirements for Satellite-based products for Climate – Supplemental Details to the GCOS Implementation Plan, GCOS 107, September 2006.


Schneising, O., Bergamaschi, P., Bovensmann, H., Buchwitz, M., Burrows, J. P., Deutscher, N. M.,

Griffith, D. W. T., Heymann, J., Macatangay, R., Messerschmidt, J., Notholt, J., Rettinger, M.,


Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based

FTS measurements and model results, Atmos. Chem. Phys., 12, 1527-1540.

Simpson, I. J., Sulbaek Andersen, M. P., Meinardi, S., Bruhwiler, L., Blake, N. J., Helmig, D.,

Rowland, F. S., Blake, D. R. (2012), Long-term decline of global atmospheric ethane concentrations

and implications for methane, Nature, www.nature.com/doifinder/10.1038/nature11342

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H.

L. (Eds.): Climate Change 2007: The Physical Science Basis, Contribution of Working Group

I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC),


Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P.,

Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N.,

Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., Steele, L. P., Francey,

R. J., and Denning, A. S. (2007), Weak northern and strong tropical land carbon uptake from

vertical profiles of atmospheric CO₂, Science, 316, 1732–1735.


evon Clarmann, T., Höpfner, M., Kellmann, S., Linden, A., Chauhan, S., Funke, B., Grabowski, U.,

Glatthor, N., Kiefer, M., Schieferdecker, T., Stiller, G., Versick, S. (2009), Retrieval of

temperature, H₂O, O₃, HNO₃, CH₄, N₂O, ClONO₂ and ClO from MIPAS reduced resolution

nominal mode limb emission measurements, Atmos. Meas. Tech., 2, 159-175.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O.,


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., Jimenez, R., Keppel-Aleks, G.,

Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S.,


11. Tables

Table 1: GHG-CCI XCO$_2$ and XCH$_4$ random and systematic uncertainty requirements for measurements over land. Abbreviations: G=Goal requirement (the maximum that needs to be achieved; better performance likely not needed as other errors (e.g., modelling errors) will dominate), B=Breakthrough requirement ("good" performance somewhere between G and T), T=Threshold requirement (the minimum that needs to be achieved for the specified application, here: global regional-scale surface flux inverse modelling). See also main text for a detailed explanation. From GHG-CCI User Requirements Document (URD, Buchwitz et al., 2011a).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Requirement type</th>
<th>Random error</th>
<th>Systematic error</th>
<th>Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XCO$_2$</td>
<td>G</td>
<td>&lt; 1 ppm</td>
<td>&lt; 0.3 ppm</td>
<td>&lt; 0.2 ppm (absolute)</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>&lt; 3 ppm</td>
<td>&lt; 1.0 ppm</td>
<td>&lt; 0.3 ppm (relative)</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>&lt; 8 ppm</td>
<td>&lt; 1.3 ppm</td>
<td>&lt; 0.5 ppm (relative)</td>
</tr>
<tr>
<td>XCH$_4$</td>
<td>G</td>
<td>&lt; 9 ppb</td>
<td>&lt; 3 ppb</td>
<td>&lt; 1 ppb (absolute)</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>&lt; 17 ppb</td>
<td>&lt; 5 ppb</td>
<td>&lt; 5 ppb (relative)</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>&lt; 34 ppb</td>
<td>&lt; 11 ppb</td>
<td>&lt; 10 ppb (relative)</td>
</tr>
</tbody>
</table>

Requirements for regional CO$_2$ and CH$_4$ source/sink determination using SCIAMACHY/ENVISAT and TANSO/GOSAT.
Table 2: Overview GHG-CCI ECV Core Algorithms (ECAs). Details on each of these algorithms are also given in the GHG-CCI ATBD (Reuter et al., 2012a) and in Buchwitz et al., 2012.

<table>
<thead>
<tr>
<th>Algorithm ID</th>
<th>Data product</th>
<th>Sensor</th>
<th>Algorithm short name</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO2_SCI_WFMD</td>
<td>XCO₂</td>
<td>SCIAMACHY/ENVISAT</td>
<td>WFMD (WFM-DOAS)</td>
<td>Schneising et al., 2011, 2012; Heymann et al., 2012b</td>
</tr>
<tr>
<td>CO2_SCI_BESD</td>
<td>XCO₂</td>
<td>SCIAMACHY</td>
<td>BESD</td>
<td>Reuter et al., 2010, 2011</td>
</tr>
<tr>
<td>CO2_GOS_OCFP</td>
<td>XCO₂</td>
<td>TANSO/GOSAT</td>
<td>OCFP (UOL-OCO-FP)</td>
<td>Cogan et al., 2012</td>
</tr>
<tr>
<td>CO2_GOS_SRFP</td>
<td>XCO₂</td>
<td>TANSO/GOSAT</td>
<td>SRFP (&quot;RemoteC&quot;)</td>
<td>Butz et al., 2011</td>
</tr>
<tr>
<td>CH4_SCI_WFMD</td>
<td>XCH₄</td>
<td>SCIAMACHY</td>
<td>WFMD (WFM-DOAS)</td>
<td>Schneising et al., 2010, 2011</td>
</tr>
<tr>
<td>CH4_SCI_IMAP</td>
<td>XCH₄</td>
<td>SCIAMACHY</td>
<td>IMAP</td>
<td>Frankenberg et al., 2011</td>
</tr>
<tr>
<td>CH4_GOS_OCFP</td>
<td>XCH₄</td>
<td>TANSO/GOSAT</td>
<td>OCFP (UOL-OCO-FP)</td>
<td>Parker et al., 2011</td>
</tr>
<tr>
<td>CH4_GOS_OCPR</td>
<td>XCH₄</td>
<td>TANSO/GOSAT</td>
<td>OCPR (UOL-PR)</td>
<td>Parker et al., 2011</td>
</tr>
<tr>
<td>CH4_GOS_SRFP</td>
<td>XCH₄</td>
<td>TANSO/GOSAT</td>
<td>SRFP</td>
<td>Butz et al., 2011</td>
</tr>
<tr>
<td>CH4_GOS_SRPR</td>
<td>XCH₄</td>
<td>TANSO/GOSAT</td>
<td>SRPR</td>
<td>Schepers et al., 2012</td>
</tr>
</tbody>
</table>
Table 3: Overview GHG-CCI Additional Constraints Algorithms (ACAs). (*)Note that CO2_SCI_ONPD is a new algorithm “similar” as the one described in Noël et al., 2011, which has been added in the 2nd year of GHG-CCI. Details on each of these algorithms are also given in the GHG-CCI ATBD (Reuter et al., 2012a) and in Buchwitz et al., 2012.

<table>
<thead>
<tr>
<th>Algorithm ID</th>
<th>Data product</th>
<th>Sensor</th>
<th>Algorithm</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO2_AIR_NLIS</td>
<td>Mid/upper trop. column</td>
<td>AIRS</td>
<td>NLIS</td>
<td>Crevoisier et al., 2004</td>
</tr>
<tr>
<td>CO2_IAS_NLIS</td>
<td>Mid/upper trop. column</td>
<td>IASI</td>
<td>NLIS</td>
<td>Crevoisier et al., 2009a</td>
</tr>
<tr>
<td>CO2_ACE_CLRS</td>
<td>Upper trop. / strat. profile</td>
<td>ACE-FTS</td>
<td>CLRS</td>
<td>Foucher et al., 2009</td>
</tr>
<tr>
<td>CO2_SCI_ONPD</td>
<td>Stratospheric profile</td>
<td>SCIAMACHY</td>
<td>ONPD</td>
<td>(Noël et al., 2011)‘’‘</td>
</tr>
<tr>
<td>CH4_IAS_NLIS</td>
<td>Upper trop. / strat. profile</td>
<td>IASI</td>
<td>NLIS</td>
<td>Crevoisier et al., 2009b</td>
</tr>
<tr>
<td>CH4_MIP_IMK</td>
<td>Upper trop. / strat. profile</td>
<td>MIPAS</td>
<td>KIT/IMK MIPAS</td>
<td>von Clarmann et al., 2009</td>
</tr>
<tr>
<td>CH4_SCI_ONPD</td>
<td>Stratospheric profile</td>
<td>SCIAMACHY</td>
<td>ONPD</td>
<td>Noël et al., 2011</td>
</tr>
</tbody>
</table>
**Table 4:** TCCON sites as used for the validation of the satellite-derived XCH$_4$ and XCO$_2$ Round Robin (RR) data products by the GHG-CCI validation team (from Notholt et al., 2012).

<table>
<thead>
<tr>
<th>Name</th>
<th>ID</th>
<th>Latitude [deg]</th>
<th>Longitude [deg]</th>
<th>Altitude [km]</th>
<th>Time coverage MM/YYYY-MM/YYYY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bialystock</td>
<td>BIA</td>
<td>53.231</td>
<td>23.025</td>
<td>0.183</td>
<td>03/2009 - 03/2011</td>
</tr>
<tr>
<td>Bremen</td>
<td>BRE</td>
<td>53.104</td>
<td>8.850</td>
<td>0.027</td>
<td>01/2009 - 12/2010</td>
</tr>
<tr>
<td>Karlsruhe</td>
<td>KAR</td>
<td>49.102</td>
<td>8.440</td>
<td>0.110</td>
<td>04/2010 - 05/2011</td>
</tr>
<tr>
<td>Orleans</td>
<td>ORL</td>
<td>47.965</td>
<td>2.113</td>
<td>0.132</td>
<td>08/2009 - 11/2010</td>
</tr>
<tr>
<td>Garmisch</td>
<td>GAR</td>
<td>47.476</td>
<td>11.063</td>
<td>0.744</td>
<td>05/2009 - 12/2010</td>
</tr>
<tr>
<td>ParkFalls</td>
<td>PAR</td>
<td>45.945</td>
<td>-90.273</td>
<td>0.442</td>
<td>06/2004 - 04/2011</td>
</tr>
<tr>
<td>Lamont</td>
<td>LAM</td>
<td>36.604</td>
<td>-97.486</td>
<td>0.320</td>
<td>07/2008 - 05/2011</td>
</tr>
<tr>
<td>Darwin</td>
<td>DAR</td>
<td>-12.425</td>
<td>130.891</td>
<td>0.030</td>
<td>08/2005 - 02/2011</td>
</tr>
<tr>
<td>Wollongong</td>
<td>WOL</td>
<td>-34.406</td>
<td>150.879</td>
<td>0.030</td>
<td>06/2008 - 03/2011</td>
</tr>
<tr>
<td>Lauder</td>
<td>LAU</td>
<td>-45.050</td>
<td>169.680</td>
<td>0.370</td>
<td>06/2004 - 06/2011</td>
</tr>
</tbody>
</table>
Table 5: Estimated precision and biases of the satellite XCH$_4$ (top) and XCO$_2$ (bottom) GHG-CCI core data products retrieved with ECAs obtained from comparisons with ground-based TCCON retrievals (see Figure 3 for details). Numbers in curved brackets are for SCIAMACHY methane retrievals during 2003-2005, i.e., before significant detector degradation of the methane channel: values from Buchwitz et al., 2012, are indicated by #) and value from Schneising et al., 2012, is indicated by §). Values in square brackets for SCIAMACHY methane retrieval are from Buchwitz et al., 2012, based on an analysis of all available retrievals (all years) and using a different assessment method. *) The exact version number for BESD is v01.00.01. Also listed are the GHG-CCI user requirements as given the GHG-CCI User Requirements Document (URD (Buchwitz et al., 2011a), see also Table 1, e.g., for the explanation of T, B, G).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sensor</th>
<th>XCH$_4$ [ppb]</th>
<th>Estimated precision single observation</th>
<th>Estimated relative biases</th>
<th>Number of satellite obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WFMD v2.3</td>
<td>SCIAMACHY</td>
<td>82 (~30$^\circ$)</td>
<td>11 (~3$^\circ$) [4-12$^\circ$]</td>
<td>37628</td>
<td></td>
</tr>
<tr>
<td>IMAP v6.0</td>
<td>SCIAMACHY</td>
<td>50 (~30$^\circ$)</td>
<td>15 [4-13$^\circ$]</td>
<td>39489</td>
<td></td>
</tr>
<tr>
<td>OCFP v3.2</td>
<td>GOSAT</td>
<td>16</td>
<td>8</td>
<td>3176</td>
<td></td>
</tr>
<tr>
<td>SRFP v1.1</td>
<td>GOSAT</td>
<td>15</td>
<td>3</td>
<td>2558</td>
<td></td>
</tr>
<tr>
<td>OCFP v3.2</td>
<td>GOSAT</td>
<td>13</td>
<td>2</td>
<td>7323</td>
<td></td>
</tr>
<tr>
<td>SRFP v1.1</td>
<td>GOSAT</td>
<td>14</td>
<td>3</td>
<td>4900</td>
<td></td>
</tr>
<tr>
<td>Required (URD):&lt; 34(T), 17(B), 9(G)</td>
<td>&lt; 10(T), 5(B), 3(G)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sensor</th>
<th>XCO$_2$ [ppm]</th>
<th>Estimated precision single observation</th>
<th>Estimated relative biases</th>
<th>Number of satellite obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>WFMD v2.2</td>
<td>SCIAMACHY</td>
<td>5.1</td>
<td>1.3</td>
<td>30752</td>
<td></td>
</tr>
<tr>
<td>BESD v1 *)</td>
<td>SCIAMACHY</td>
<td>2.3</td>
<td>0.7</td>
<td>9467</td>
<td></td>
</tr>
<tr>
<td>OCFP v3.0</td>
<td>GOSAT</td>
<td>2.7</td>
<td>0.6</td>
<td>2830</td>
<td></td>
</tr>
<tr>
<td>SRFP v1.1</td>
<td>GOSAT</td>
<td>2.8</td>
<td>0.9</td>
<td>2558</td>
<td></td>
</tr>
<tr>
<td>Required (URD):&lt; 8(T), 3(B), 1(G)</td>
<td>&lt; 0.5(T), 0.3(B), 0.2(G)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 6: Overview about the planned content of the GHG-CCI CRDP. §) see Table 2 and Table 3, *) may end later, +) may start earlier, #) mainly high latitudes. Products: (1) mid/upper tropospheric columns, (2) (primarily) stratospheric vertical profiles.

<table>
<thead>
<tr>
<th>Product ID</th>
<th>Product (Level 2, mixing ratios)</th>
<th>Algorithm §)</th>
<th>Coverage</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>XCO₂_SCIA</td>
<td>XCO₂</td>
<td>BESD</td>
<td>Global, land, 2003-2010</td>
<td>-</td>
</tr>
<tr>
<td>XCO₂_GOSAT</td>
<td>XCO₂</td>
<td>OCFP and SRFP</td>
<td>Global, mid 2009-2010</td>
<td>2 alternative products</td>
</tr>
<tr>
<td>XCO₂_EMMA</td>
<td>XCO₂</td>
<td>EMMA</td>
<td>Global, mid 2009-2010</td>
<td>Merged SCIA and GOSAT</td>
</tr>
<tr>
<td>XCH₄_SCIA</td>
<td>XCH₄</td>
<td>IMAP and WFMD</td>
<td>Global, 2003-2010</td>
<td>2 alternative products</td>
</tr>
<tr>
<td>XCH₄_GOSAT</td>
<td>XCH₄</td>
<td>SRFP and OCPR</td>
<td>Global, mid 2009-2010</td>
<td>2 alternative products</td>
</tr>
</tbody>
</table>

Data products generated with Additional Constraints Algorithms (ACAs)

<table>
<thead>
<tr>
<th>Product ID</th>
<th>Product (Level 2, mixing ratios)</th>
<th>Algorithm §)</th>
<th>Coverage</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂_AIRS</td>
<td>CO₂ (1)</td>
<td>NLIS</td>
<td>Tropics, 2003-2007</td>
<td>-</td>
</tr>
<tr>
<td>CO₂_IASI</td>
<td>CO₂ (1)</td>
<td>NLIS</td>
<td>Tropics, 2007-2010</td>
<td>-</td>
</tr>
<tr>
<td>CH₄_IASI</td>
<td>CH₄ (1)</td>
<td>NLIS</td>
<td>Tropics, 2007-2010</td>
<td>-</td>
</tr>
<tr>
<td>CH₄_SCIA_OCC</td>
<td>CH₄ (2)</td>
<td>ONPD</td>
<td>NH mid/high lat., 2003-2010</td>
<td>-</td>
</tr>
<tr>
<td>CO₂_SCI_OCC</td>
<td>CO₂ (2)</td>
<td>ONPD</td>
<td>NH mid/high lat., 2003-2010</td>
<td>-</td>
</tr>
<tr>
<td>CH₄_MIPAS</td>
<td>CH₄ (2)</td>
<td>KIT/IMK MIPAS</td>
<td>Global, 2005-2010</td>
<td>-</td>
</tr>
<tr>
<td>CO₂_ACEFTS</td>
<td>CO₂ (2)</td>
<td>CLRS</td>
<td>Global, 2004-2010</td>
<td>-</td>
</tr>
</tbody>
</table>
12. Figures

**Figure 1**: Top: Northern hemispheric monthly mean XCO$_2$ time series retrieved from SCIAMACHY/ENVISAT (algorithms: WFMD and BESD) and TANSO/GOSAT (algorithms: SRFP and OCFP) satellite data. Shown are monthly mean values for the 0°-60°N latitude range. Clearly visible is the CO$_2$ increase primarily caused by the burning of fossil fuels and the seasonal cycle primarily caused by uptake and release of CO$_2$ by the terrestrial biosphere. Bottom: As top panel but for XCH$_4$ (algorithms: SCIAMACHY: WFMD and IMAP, GOSAT: SRFP, SRPR, OCFP, OCPR).

For a color version of this figure please have a look at the on-line version of this publication.
Figure 2: Global XCO₂ maps from SCIAMACHY (left) and CarbonTracker (right) for two seasons (April-June, top, and July-September, bottom) and two years (2003 and 2009). The CarbonTracker model data have been sampled according to the SCIAMACHY measurements and the SCIAMACHY averaging kernels have been applied to CarbonTracker. Figure adapted from Heymann et al., 2012b. For a color version of this figure please have a look at the on-line version of this publication.
Figure 3: Comparison of the GHG-CCI core ECV data products \(XCH_4\) (top 2 rows, values in ppb) and \(XCO_2\) (bottom 2 rows, values in ppm) from SCIAMACHY/ENVISAT and TANSO/GOSAT generated using all GHG-CCI ECAs with TCCON ground-based observations (see Table 4 for details on the TCCON sites). Shown are the mean difference with respect to TCCON (left) and the standard deviation of the difference (right). A 500 km / 2 hour spatio-temporal co-location criterion has been used to compute the satellite – TCCON differences. The numerical values listed are: Left: “StdDev” is the standard deviation of the mean differences as obtained at the TCCON sites, i.e., a measure of the station-to-station bias and can be interpreted as relative accuracy (relative bias) of the satellite retrievals. “N” is the number of satellite data used for comparison (only those data points are shown where at least 10 satellite observations are available for a given site). Right: “Mean” is the mean value of the standard deviations show by the symbols and is a measure of the achieved overall precision. For a color version of this figure please have a look at the on-line version of this publication.
**Figure 4:** Comparison of two SCIAMACHY XCH₄ data products retrieved using WFMD (top) and IMAP (middle) for August 2005. Global maps of the retrieved XCH₄ are shown on the left and the number of retrievals per 5°x5° grid cell on the right. The WFMD-IMAP difference is shown in the bottom row. Listed in the bottom left are the following parameters: d: mean difference (-2.12 ppb), s: standard deviation of the difference (18.53 ppb), r: linear correlation coefficient (0.75). For a color version of this figure please have a look at the on-line version of this publication.
Figure 5: Comparison of the two GHG-CCI GOSAT XCH₄ PR data products retrieved using the OCPR and SRPR retrieval algorithms. Left: Percentage XCH₄ difference OCPR-SRPR for July 2009. Right: Scatter plot of 6751 co-located OCPR versus SRPR retrievals at TCCON sites. The standard deviation of the difference is 10 ppb (1-sigma) and the linear correlation coefficient is 0.91. For a color version of this figure please have a look at the on-line version of this publication.
**Figure 6:** Maps of monthly mean XCO$_2$ at $10^\circ \times 10^\circ$ resolution as obtained using different GHG-CCI retrieval algorithms: WFMD and BESD for SCIAMACHY, OCFP and SRFP for GOSAT and SCIAMACHY and GOSAT merged using EMMA for September 2009 (left) and May 2012 (right). For a color version of this figure please have a look at the on-line version of this publication.
Figure 7: Comparison matrix of monthly XCO$_2$ maps for September 2009 (top) and May 2010 (bottom) generated using several individual satellite retrieval algorithms: BESD and WFMD for SCIAMACHY and RemoteC (= SRFP), ACOS, UOL-FP (=OCFP), PPDF, NIES for GOSAT. The EMMA data product has been generated from the ensemble of the individual SCIAMACHY and GOSAT XCO$_2$ data products (see main text for details). Also shown is XCO$_2$ from TCCON and NOAA’s CarbonTracker (CT). The diagonal elements show the monthly XCO$_2$ maps (using color bar “mean”). The above diagonal elements show the XCO$_2$ differences for all combinations (color bar “difference”). The below diagonal elements show the numerical values of the Root Mean Square Difference (RMSD) as well as color coded smileys of the RMSD (green: RMSD < 1.2 ppm, red: RMSD > 2.4 ppm, otherwise yellow). For a color version of this figure please have a look at the online version of this publication.
Figure 8: Pie charts showing the agreement (left) and disagreement (right) with the EMMA median obtained using the listed satellite XCO₂ data products. The figure has been obtained using the EMMA Level 3 data product (10°x10°, monthly = 1 voxel). For each voxel the mean XCO₂ value for each algorithm has been computed and the median using all algorithms. The “Agreement with the Median” (left) has been computed as follows: For algorithm i the number of voxels which agree with the median within 0.2 ppm have been counted (= Nᵢ). 100% corresponds to the sum of these numbers (N = Σᵢ Nᵢ). The percentages shown are N/N*100%. The percentages of “Potential Outliers” (right) have been calculated using the same method except that all voxels have been counted where the differences to the median are larger than 2 ppm. As can be seen from the left figure, the data product which agrees best with the median is the ACOS product (v2.9, 21% agreement) followed by the similar UOL-FP (= OCFP) algorithm (19% agreement). The largest number of potential outliers have the data products generated with the two very fast algorithms WFMD (32%) and PPDF (16%). For a color version of this figure please have a look at the on-line version of this publication.
**Figure 9:** Comparison of the XCO₂ seasonal cycle amplitude (peak-to-peak) of the individual XCO₂ algorithms and EMMA with CarbonTracker (v2011) (left) and TCCON (right). As can be seen, all XCO₂ satellite data suggest that the amplitude of the CO₂ seasonal cycle is underestimated by CarbonTracker. The CarbonTracker underestimation is approximately 1.5+/−0.5 ppm peak-to-peak as can be concluded from all satellite data products including EMMA but excluding those data products which do not agree within their error bar with TCCON (i.e., NIES-PPDF-D and WFMD). Figure from: Reuter et al., 2012b. For a color version of this figure please have a look at the on-line version of this publication.
Carbon dioxide (CO₂) column-averaged mole fraction (XCO₂)

Long-term increase (fossil fuel burning, etc.)

SCIAMACHY / WFMD

CarbonTracker / NOAA

Seasonal uptake by growing vegetation (NH)

XCO₂ [ppm]

364  372  380  388  396
Figure 3

Click here to download high resolution image

**Satellite - TCCON differences:**

**SCIAMACHY/ENVISAT XCH₄**

Mean [ppb]

- **BIA**
- **BRE**
- **KAR**
- **ORL**
- **GAR**
- **PAR**
- **LAM**
- **DAR**
- **WOL**
- **LAU**

StdDev: **IMAP**: 15  **WFMD**: 11
N: 39489  37628

Mean: **IMAP**: 50  **WFMD**: 82

**TANSO/GOSAT XCH₄**

Mean [ppb]

- **BIA**
- **BRE**
- **KAR**
- **ORL**
- **GAR**
- **PAR**
- **LAM**
- **DAR**
- **WOL**
- **LAU**

StdDev: **OCFP**: 8  **SRFP**: 3  **OCRPR**: 2  **SRPR**: 3
N: 3176  2558  7323  4900

Mean: **OCFP**: 16  **SRFP**: 15  **OCRPR**: 13  **SRPR**: 14

**SCIAMACHY/ENVISAT XCO₂**

Mean [ppm]

- **BIA**
- **BRE**
- **KAR**
- **ORL**
- **GAR**
- **PAR**
- **LAM**
- **DAR**
- **WOL**
- **LAU**

StdDev: **BESD**: 0.71  **WFMD**: 1.33
N: 9467  30752

Mean: **BESD**: 2.33  **WFMD**: 5.10

**TANSO/GOSAT XCO₂**

Mean [ppm]

- **BIA**
- **BRE**
- **KAR**
- **ORL**
- **GAR**
- **PAR**
- **LAM**
- **DAR**
- **WOL**
- **LAU**

StdDev: **OCFP**: 0.61  **SRFP**: 0.88
N: 2830  2558

Mean: **OCFP**: 2.70  **SRFP**: 2.79
Figure 5
GOSAT XCH$_4$: OCPR versus SRPR

Delta XCH$_4$ (%)

July 2009, OCPR-SRPR

At TCCON sites

1-$\sigma$ = 10 ppb
Correlation = 0.91
6751 common data points
Figure 8

(a) Agreement with Median ($\Delta XCO_2 \leq 0.2$ppm)
- BESD (11%)
- WFMD (7%)
- NIES (15%)
- ACOS (21%)
- PPDF-D (10%)
- RemoteC (17%)
- UOL-FP (19%)

(b) Potential Outliers ($\Delta XCO_2 \geq 2.0$ppm)
- BESD (12%)
- ACOS (9%)
- RemoteC (13%)
- PPDF-D (16%)
- NIES (14%)
- WFMD (32%)
- UOL-FP (5%)
Figure 9: Difference of Seasonal Amplitude

The graph illustrates the difference of seasonal amplitude for various datasets, including ACOS, RemoteC, BESD, UOL-FP, NIES PPDF-D, WFMD, NIES v02.xx, and EMMA. The x-axis represents the datasets CT2011 and TCCON, while the y-axis shows the amplitude in parts per million (ppm) from 0 to 7.