Ensemble-based satellite-derived carbon dioxide and methane column-averaged dry-air mole fraction data sets (2003–2018) for carbon and climate applications

Maximilian Reuter¹, Michael Buchwitz¹, Oliver Schneising¹, Stefan Noël¹, Heinrich Bovensmann¹, John P. Burrows¹, Hartmut Boesch²,³, Antonio Di Noia²,³, Jasdeep Anand²,³, Robert J. Parker²,³, Peter Somkuti²,³, Lianghai Wu⁴, Otto P. Hasekamp¹, Ilse Aben⁴, Akihiko Kuze⁵, Hiroshi Suto⁵, Kei Shiomı⁵, Yukio Yoshida⁶, Isamu Morino⁶, David Crisp, Christopher W. O’Dell⁷, Justus Notholt¹, Christof Petri¹, Thorsten Warneke¹, Voltaire A. Velazco⁹, Nicholas M. Deutscher⁹, David W. T. Griffith⁹, Rigel Kivi¹⁰, David F. Pollard¹¹, Frank Hase¹², Ralf Sussmann¹³, Yao V. Té¹⁴, Kimberly Strong¹⁵, Sébastien Roche¹⁵, Mahesh K. Sha¹⁶, Martine De Mazière¹⁶, Dietrich G. Feist¹⁷,¹⁸,¹⁹, Laura T. Iraci²⁰, Coleen M. Roehl²¹, Christian Retscher²², and Dinand Schepers²³

1 Institute of Environmental Physics (IUP), University of Bremen, 28334 Bremen, Germany
2 Earth Observation Science, University of Leicester, LE1 7RH, Leicester, UK
3 NERC National Centre for Earth Observation, LE1 7RH, Leicester, UK
4 SRON Netherlands Institute for Space Research, 3584 CA Utrecht, the Netherlands
5 Japan Aerospace Exploration Agency (JAXA), 305-8505, Tsukuba, Japan
6 National Institute for Environmental Studies (NIES), 305-8506, Tsukuba, Japan
7 Jet Propulsion Laboratory (JPL), Pasadena, CA 91109, USA
8 Cooperative Institute for Research in the Atmosphere, Colorado State University (CSU), Fort Collins, CO 80523, USA
9 Centre for Atmospheric Chemistry, School of Earth, Atmospheric and Life Sciences, University of Wollongong, NSW, 2522, Australia
10 Space and Earth Observation Centre, Finnish Meteorological Institute (FMI), 99600 Sodankylä, Finland
11 National Institute of Water and Atmospheric Research (NIWA), Lauder, New Zealand
12 Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research (IMK), IMK-ASF, 76021 Karlsruhe, Germany
13 Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research (IMK), IMK-IFU, 82467 Garmisch-Partenkirchen, Germany
14 Laboratoire d’Etudes du Rayonnement et de la Matière en Astrophysique et Atmosphères (LERMA-IPSL), Sorbonne Université, CNRS, Observatoire de Paris, PSL Université, 75005 Paris, France
15 Department of Physics, University of Toronto, Toronto, ON, M5S 1A7, Canada
16 Royal Belgian Institute for Space Aeronomy (BIRA-IASB), 1180 Brussels, Belgium
17 Max Planck Institute for Biogeochemistry, 07745 Jena, Germany
18 Lehrstuhl für Physik der Atmosphäre, Ludwig-Maximilians-Universität München, 80333 Munich, Germany
19 Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt Oberpfaffenhofen, 82234 Welhling, Germany
20 Atmospheric Science Branch, National Aeronautics and Space Administration (NASA), Moffett Field, CA 94035, USA
21 California Institute of Technology, Pasadena, CA 91125, USA
22 Directorate of Earth Observation Programmes, European Space Agency (ESA), ESRIN, 00044 Frascati, Italy
23 European Centre for Medium-Range Weather Forecasts (ECMWF), Reading RG2 9AX, UK

Correspondence: Michael Buchwitz (buchwitz@uni-bremen.de)

Received: 21 October 2019 – Discussion started: 30 October 2019
Revised: 16 January 2020 – Accepted: 20 January 2020 – Published: 19 February 2020

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract. Satellite retrievals of column-averaged dry-air mole fractions of carbon dioxide (CO$_2$) and methane (CH$_4$), denoted XCO$_2$ and XCH$_4$, respectively, have been used in recent years to obtain information on natural and anthropogenic sources and sinks and for other applications such as comparisons with climate models. Here we present new data sets based on merging several individual satellite data products in order to generate consistent long-term climate data records (CDRs) of these two Essential Climate Variables (ECVs). These ECV CDRs, which cover the time period 2003–2018, have been generated using an ensemble of data products from the satellite sensors SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT and (for XCO$_2$) for the first time also including data from the Orbiting Carbon Observatory 2 (OCO-2) satellite. Two types of products have been generated: (i) Level 2 (L2) products generated with the latest version of the ensemble median algorithm (EMMA) and (ii) Level 3 (L3) products obtained by gridding the corresponding L2 EMMA products to obtain a monthly $5^\circ \times 5^\circ$ data product in Obs4MIPS (Observations for Model Inter-comparisons Project) format. The L2 products consist of daily NetCDF (Network Common Data Form) files, which contain in addition to the main parameters, i.e., XCO$_2$ or XCH$_4$, corresponding uncertainty estimates for random and potential systematic uncertainties and the averaging kernel for each single (quality-filtered) satellite observation. We describe the algorithms used to generate these data products and present quality assessment results based on comparisons with Total Carbon Column Observing Network (TCCON) ground-based retrievals. We found that the XCO$_2$ Level 2 data set at the TCCON validation sites can be characterized by the following figures of merit (the corresponding values for the Level 3 product are listed in brackets) – single-observation random error ($\sigma_\text{r}$): 1.29 ppm (monthly: 1.18 ppm); global bias: 0.20 ppm (0.18 ppm); and spatiotemporal bias or relative accuracy ($\sigma_\text{r}$): 0.66 ppm (0.70 ppm). The corresponding values for the XCH$_4$ products are single-observation random error ($\sigma_\text{r}$): 17.4 ppb (monthly: 8.7 ppb); global bias: $-2.0$ ppb ($-2.9$ ppb); and spatiotemporal bias ($\sigma_\text{r}$): $5.0$ ppb ($4.9$ ppb). It has also been found that the data products exhibit very good long-term stability as no significant long-term bias trend has been identified. The new data sets have also been used to derive annual XCO$_2$ and XCH$_4$ growth rates, which are in reasonable to good agreement with growth rates from the National Oceanic and Atmospheric Administration (NOAA) based on marine surface observations. The presented ECV data sets are available (from early 2020 onwards) via the Climate Data Store (CDS, https://cds.climate.copernicus.eu/, last access: 10 January 2020) of the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/, last access: 10 January 2020).

1 Introduction

Carbon dioxide (CO$_2$) and methane (CH$_4$) are important greenhouse gases and increasing atmospheric concentrations result in global warming with adverse consequences such as sea level rise (IPCC, 2013). Because of their importance for climate, these gases have been classified as Essential Climate Variables (ECVs) by the Global Climate Observing System (GCOS) (GCOS-154, 2010; GCOS-200, 2016). The generation of XCO$_2$ and XCH$_4$ satellite-derived ECV data products meeting GCOS requirements using European satellite retrieval algorithms started in 2010 in the framework of the GHG-CCI project (http://www.esa-ghg-cci.org/, last access: 10 January 2020) of the European Space Agency’s (ESA) Climate Change Initiative (CCI) (Hollmann et al., 2013). Since the end of 2016, this activity continues operationally via the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/, last access: 10 January 2020), and the corresponding CO$_2$ and CH$_4$ data products are available via the Copernicus Climate Data Store (CDS, https://cds.climate.copernicus.eu/, last access: 10 January 2020). These ECV data products have been used for a range of applications such as improving our knowledge of CO$_2$ and/or CH$_4$ surface fluxes (e.g., Alexe et al., 2015; Basu et al., 2013; Buchwitz et al., 2017a; Chevallier et al., 2014, 2015; Ganesan et al., 2017; Gaubert et al., 2019; Houweling et al., 2015; Liu et al., 2017; Maasakkers et al., 2019; Miller et al., 2019; Reuter et al., 2014a, b; 2019a; Sheng et al., 2018; Schneising et al., 2014b; Turner et al., 2015, 2019), comparison with climate and other models (e.g., Hayman et al., 2014; Lauer et al., 2017; Schneising et al., 2014a), and for other applications such as computation of CO$_2$ growth rates (e.g., Buchwitz et al., 2018), as well as to better understand changes in the amplitude of the CO$_2$ seasonal cycle (e.g., Yin et al., 2018).

The C3S satellite greenhouse gas (GHG) data set consists of single-sensor satellite data products and of merged (i.e., combined multi-sensor, multi-algorithm) data products. Here we present the latest version, version 4.1, of the merged Level 2 (L2) and merged Level 3 (L3) XCO$_2$ and XCH$_4$ data products, which cover the time period 2003–2018. The L2 products (XCO2_EMMA and XCH4_EMMA) have been compiled with the ensemble median algorithm (EMMA) originally proposed by Reuter et al. (2013) and recent modifications, which are described in Sect. 3.1. These products contain detailed information for each single observation (i.e., footprint or ground pixel) including time, latitude and longitude, the main parameter (i.e., XCO$_2$ or XCH$_4$), its stochastic uncertainty (e.g., due to instrument noise), an estimate of potential systematic uncertainties (e.g., due to spatial or temporal bias patterns), and its averaging kernel and corresponding a priori profile. The L3 products (XCO2_OBS4MIPS and XCH4_OBS4MIPS) are gridded products at monthly time and $5^\circ \times 5^\circ$ spatial resolution in Obs4MIPS (Observ-
Figure 1. Overview of the presented XCO\textsubscript{2} data set. Shown are time series over land for three latitude bands (global, black line: Northern Hemisphere, red; Southern Hemisphere, green) and global maps (half-yearly averages at 1° × 1° obtained by gridding (averaging) the merged Level 2, i.e., EMMA, product). See Sect. 4 for a detailed discussion.

Figure 2. As Fig. 1 but for XCH\textsubscript{4}.


Figure 1 provides an overview of the resulting merged XCO\textsubscript{2} data product in terms of time series for three latitude bands and global maps and the similarly structured Fig. 2 shows the XCH\textsubscript{4} product. As can be seen, XCO\textsubscript{2} and XCH\textsubscript{4} are both increasing with time and exhibit seasonal fluctuations and spatial variations. The spatiotemporal characteristics of the merged data – e.g., the spatial sampling – reflect...
the characteristics of the underlying individual sensor satellite data (described in the data section, Sect. 2). Figures 1 and 2 are discussed in detail in the results section, Sect. 4. How these data products have been generated is described in the methods section, Sect. 3. A summary and conclusions are given in Sect. 5.

2 Data

In this section, we present an overview about the input data used to generate and validate the new XCO\textsubscript{2} and XCH\textsubscript{4} data products.

2.1 Satellite data

The input satellite data used to generate the merged satellite data products are individual satellite sensor Level 2 (L2) data products. Table 1 provides an overview about the satellite XCO\textsubscript{2} input data sets. As can be seen, in total eight XCO\textsubscript{2} L2 data products have been used to generate the merged L2 and Level 3 (L3) XCO\textsubscript{2} data products, each corresponding to a different combination of satellite sensor and retrieval algorithm. An overview about the time coverage of these input data products is presented in Fig. 3. As can be seen, the time period 2003 to March 2009 is only covered by one XCO\textsubscript{2} product, namely XCO\textsubscript{2} retrieved with the Bremen Optimal Estimation DOAS (BESD) algorithm (Reuter et al., 2010, 2011) from the SCIAMACHY/ENVISAT (Bovensmann et al., 1999) instrument. A second SCIAMACHY XCO\textsubscript{2} data product is available, which has been retrieved with the Weighting Function Modified Differential Optical Absorption Spectroscopy (WFM-DOAS or WFMD) algorithm (Schneising et al., 2011), but this product is not used because the merging algorithm EMMA (Reuter et al., 2013, described in Sect. 3.1) requires one or more than two input data products (because the median of a set of elements is, according to our definition which avoids averaging, not defined for two elements). Therefore, one of the two products had to be selected, and the choice was the BESD product for XCO\textsubscript{2} because of somewhat higher data quality compared to the WFMD product (Buchwitz et al., 2017b) (note however that the WFMD product has the advantage of containing a larger number of observations). As can be seen from Table 1 and Fig. 3, several GOSAT input products have been used from April 2009 onwards and two OCO-2 XCO\textsubscript{2} products from September 2014 and May 2015 onwards. Note that additional algorithms/data products are available but have not been used as input, for example the GOSAT BESD XCO\textsubscript{2} product (Heymann et al., 2015) and the OCO-2 RemoTeC XCO\textsubscript{2} product (Wu et al., 2018). These or other additional products may be added in future versions of the merged XCO\textsubscript{2} products. Note also that we always use the bias-corrected version of a data product, if available (some product files contain bias-corrected and uncorrected values).

All listed satellites perform nadir (down-looking) and glint observations and provide radiance spectra covering the relevant CO\textsubscript{2} and CH\textsubscript{4} absorption bands located in the short-wave infrared (SWIR) part of the electromagnetic spectrum (around 1.6 and 2 μm) and also cover the O\textsubscript{2} A-band spectral region in the near-infrared (NIR, around 0.76 μm). All individual sensor input L2 data products have been generated using retrieval algorithms based on minimizing the difference between a modeled radiance spectrum and the observed spectrum by modifying so-called state vector elements (for details we refer to the references listed in Table 1; for additional information see also the Algorithm Theoretical Basis Documents (ATBDs); Buchwitz et al., 2019b, and Reuter et al., 2019b). The exact definition of the state vector depends on the algorithm, but the general approach is based on the optimal estimation (Rodgers, 2000) formalism or similar approaches (see references in Table 1). Among the state vector elements is a representation of the CO\textsubscript{2} vertical profile but also other parameters to consider interfering gases (e.g., water vapor), surface reflection, atmospheric scattering, and other effects and parameters, which have an impact on the (interpretation of the) measured radiance spectrum.

Table 2 and Fig. 4 provide an overview about the satellite XCH\textsubscript{4} L2 input data sets. As for XCO\textsubscript{2}, the time period 2003 to March 2009 is only covered by one SCIAMACHY data product. From April 2009 onwards several GOSAT XCH\textsubscript{4} products are available (see Table 2) and have been used to generate the merged XCH\textsubscript{4} data L2 and L3 data products. For future updates it is also planned to include XCH\textsubscript{4} from the Sentinel-5 Precursor (S5P) satellite (Veefkind et al., 2012), but S5P XCH\textsubscript{4} (Hu et al., 2018; Schneising et al., 2019) has not yet been included as the time period covered by these products is currently quite short (less than 2 years). However, we aim to include S5P XCH\textsubscript{4} for one of the next updates of the merged methane products.
Table 1. Satellite XCO2 Level 2 (L2) data products used as input for the generation of the merged L2 and L3 XCO2 version 4.1 data products. For products which have been generated in the framework of the CCI and C3S projects the corresponding product ID is listed (the other products are external products, which have been obtained from the corresponding websites; see Acknowledgements). Temporal coverage indicates the time coverage of the input data sets.

<table>
<thead>
<tr>
<th>Algorithm/ product acronym</th>
<th>Algorithm/ product version</th>
<th>CCI/C3S Satellite/sensor</th>
<th>Temporal coverage</th>
<th>Comment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BESD</td>
<td>v02.01.02</td>
<td>CO2_SCI_BESD SCIAMACHY</td>
<td>01/2003–03/2012</td>
<td>–</td>
<td>Reuter et al. (2011)</td>
</tr>
<tr>
<td>RemoTeC</td>
<td>v2.3.8</td>
<td>CO2_GOS_SRFP GOSAT</td>
<td>04/2009–12/2018</td>
<td>–</td>
<td>Butz et al. (2011)</td>
</tr>
<tr>
<td>ACOS</td>
<td>v7.3.10a</td>
<td>– GOSAT</td>
<td>04/2009–05/2016</td>
<td>NASA ACOS GOSAT algorithm</td>
<td>O’Dell et al. (2012)</td>
</tr>
<tr>
<td>FOCAL</td>
<td>v08</td>
<td>– OCO-2</td>
<td>01/2015–12/2018</td>
<td>–</td>
<td>Reuter et al. (2017a, b)</td>
</tr>
</tbody>
</table>

Figure 4. As Fig. 3 but XCH4. For details on each product see Table 2.

2.2 Ground-based data

The satellite data products have been validated by comparison with the XCO2 and XCH4 data products of the TCCON (Wunch et al., 2011). TCCON is a network of ground-based Fourier transform spectrometers (FTSs) recording direct solar spectra in the NIR/SWIR spectral region. From these spectra, accurate and precise column-averaged abundances of CO2, CH4, and a number of other species are retrieved. The TCCON data products (version GGG2014) have been obtained via the TCCON data archive (https://tccondata.org/, last access: 15 July 2019). An overview about the used TCCON sites is presented in Table 3.

In Sect. 4.3, we present annual XCO2 and XCH4 growth rates, which have been derived from the new XCO2 and XCH4 OBS4MIPS data products using the method described in Buchwitz et al. (2018). These growth rates are compared with growth rates derived from marine surface CO2 and CH4 observations, which have been obtained from the National Oceanic and Atmospheric Administration (NOAA) (for details including links and last access see Acknowledgements).

3 Methods

3.1 Merging algorithm EMMA

In order to generate the merged L2 products, the ensemble median algorithm is used, which is described in detail in Reuter et al. (2013). Therefore, we limit the description given here to a short overview of the latest version of the EMMA algorithm. To be specific, we initially describe the EMMA XCO2 algorithm and explain differences relevant for XCH4 at the end of this subsection.
Table 2. As Table 1 but for \(\text{XCH}_4\).

<table>
<thead>
<tr>
<th>Algorithm/ product acronym</th>
<th>Algorithm/ product version</th>
<th>CCI/C3S product ID</th>
<th>Satellite/ sensor</th>
<th>Temporal coverage</th>
<th>Comment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>WFMD</td>
<td>v4.0</td>
<td>CH4_SCI_WFMD</td>
<td>SCIAMACHY</td>
<td>01/2003–12/2011</td>
<td>–</td>
<td>Schneising et al. (2011)</td>
</tr>
<tr>
<td>UoL-FP</td>
<td>v7.2</td>
<td>CH4_GOS_OCFP</td>
<td>GOSAT</td>
<td>04/2009–12/2018</td>
<td>Univ. of Leicester full-physics (FP) algorithm</td>
<td>Parker et al. (2011)</td>
</tr>
<tr>
<td>UoL-PR</td>
<td>v7.2</td>
<td>CH4_GOS_OCPR</td>
<td>GOSAT</td>
<td>04/2009–12/2018</td>
<td>Univ. of Leicester proxy (PR) algorithm</td>
<td>Parker et al. (2011)</td>
</tr>
<tr>
<td>RemoTeC-FP</td>
<td>v2.3.8</td>
<td>CH4_GOS_SRFP</td>
<td>GOSAT</td>
<td>04/2009–12/2018</td>
<td>SRON full-physics (FP) algorithm</td>
<td>Butz et al. (2011)</td>
</tr>
<tr>
<td>RemoTeC-PR</td>
<td>v2.3.9</td>
<td>CH4_GOS_SRPR</td>
<td>GOSAT</td>
<td>04/2009–12/2018</td>
<td>SRON proxy (PR) algorithm</td>
<td>Butz et al. (2010)</td>
</tr>
</tbody>
</table>

Table 3. TCCON sites used for the validation of the \(\text{XCO}_2\) and \(\text{XCH}_4\) satellite-derived data products.

<table>
<thead>
<tr>
<th>TCCON site (Acronym)</th>
<th>Latitude (°)</th>
<th>Longitude (°)</th>
<th>Altitude (km)</th>
<th>Start of time series</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eureka, Canada (EUR)</td>
<td>80.05</td>
<td>−86.42</td>
<td>0.61</td>
<td>07.2010</td>
<td>Strong et al. (2019)</td>
</tr>
<tr>
<td>Ny-Ålesund, Spitzbergen (NYL)</td>
<td>78.92</td>
<td>11.92</td>
<td>0.02</td>
<td>04.2014</td>
<td>Notholt et al. (2019a)</td>
</tr>
<tr>
<td>Sodankylä, Finland (SOD)</td>
<td>67.37</td>
<td>26.63</td>
<td>0.19</td>
<td>05.2009</td>
<td>Kivi et al. (2014); Kivi and Heikkinen (2016)</td>
</tr>
<tr>
<td>East Trout Lake, Canada (ETL)</td>
<td>54.35</td>
<td>−104.99</td>
<td>0.50</td>
<td>10.2016</td>
<td>Wunch et al. (2018)</td>
</tr>
<tr>
<td>Bialystok, Poland (BIA)</td>
<td>53.23</td>
<td>23.03</td>
<td>0.19</td>
<td>03.2009</td>
<td>Deutscher et al. (2019)</td>
</tr>
<tr>
<td>Bremen, Germany (BRE)</td>
<td>53.10</td>
<td>8.85</td>
<td>0.03</td>
<td>01.2010</td>
<td>Notholt et al. (2019b)</td>
</tr>
<tr>
<td>Karlsruhe, Germany (KAR)</td>
<td>49.10</td>
<td>8.44</td>
<td>0.11</td>
<td>04.2010</td>
<td>Hase et al. (2015)</td>
</tr>
<tr>
<td>Paris, France (PAR)</td>
<td>48.85</td>
<td>2.36</td>
<td>0.06</td>
<td>09.2014</td>
<td>Té et al. (2014)</td>
</tr>
<tr>
<td>Orléans, France (ORL)</td>
<td>47.97</td>
<td>2.11</td>
<td>0.13</td>
<td>08.2009</td>
<td>Warneke et al. (2019)</td>
</tr>
<tr>
<td>Garmisch, Germany (GAR)</td>
<td>47.48</td>
<td>11.06</td>
<td>0.75</td>
<td>07.2007</td>
<td>Sussmann and Rettinger (2018)</td>
</tr>
<tr>
<td>Park Falls, WI, USA (PFA)</td>
<td>45.94</td>
<td>−90.27</td>
<td>0.44</td>
<td>06.2004</td>
<td>Wennberg et al. (2017)</td>
</tr>
<tr>
<td>Lamont, OK, USA (LAM)</td>
<td>36.60</td>
<td>−97.49</td>
<td>0.32</td>
<td>07.2008</td>
<td>Wennberg et al. (2016)</td>
</tr>
<tr>
<td>Tsukuba, Japan (TSU)</td>
<td>36.05</td>
<td>140.12</td>
<td>0.03</td>
<td>08.2011</td>
<td>Morino et al. (2018a)</td>
</tr>
<tr>
<td>Edwards, CA, USA (EDW)</td>
<td>34.96</td>
<td>−117.88</td>
<td>0.70</td>
<td>07.2013</td>
<td>Ibar et al. (2014)</td>
</tr>
<tr>
<td>Caltech, CA, USA (CAL)</td>
<td>34.14</td>
<td>−118.13</td>
<td>0.24</td>
<td>09.2012</td>
<td>Wennberg et al. (2015)</td>
</tr>
<tr>
<td>Saga, Japan (SAG)</td>
<td>33.24</td>
<td>130.29</td>
<td>0.01</td>
<td>07.2011</td>
<td>Shiomi et al. (2014)</td>
</tr>
<tr>
<td>Burgos, Philippines (BUR)</td>
<td>18.53</td>
<td>120.65</td>
<td>0.04</td>
<td>03.2017</td>
<td>Morino et al. (2018b); Velazco et al. (2017)</td>
</tr>
<tr>
<td>Ascension Island (ASC)</td>
<td>−7.92</td>
<td>−14.33</td>
<td>0.03</td>
<td>10.2018</td>
<td>Feist et al. (2014)</td>
</tr>
<tr>
<td>Darwin, Australia (DAR)</td>
<td>−12.46</td>
<td>130.93</td>
<td>0.04</td>
<td>08.2005</td>
<td>Griffith et al. (2014b)</td>
</tr>
<tr>
<td>Réunion island (REU)</td>
<td>−20.90</td>
<td>55.49</td>
<td>0.09</td>
<td>09.2011</td>
<td>De Mazière et al. (2017)</td>
</tr>
<tr>
<td>Wollongong, Australia (WOL)</td>
<td>−34.41</td>
<td>150.88</td>
<td>0.03</td>
<td>06.2008</td>
<td>Griffith et al. (2014a)</td>
</tr>
<tr>
<td>Lauder, New Zealand (LAU)</td>
<td>−45.04</td>
<td>169.68</td>
<td>0.37</td>
<td>02.2010</td>
<td>Sherlock et al. (2014)</td>
</tr>
</tbody>
</table>
The EMMA XCO₂ data product consists of selected individual L2 soundings from the available individual sensor L2 input products (listed in Table 1). The EMMA L2 product is based on selecting the best soundings (i.e., single ground pixel observations) from the ensemble of individual sensor L2 products. Sounding selection is based on monthly time and 10° × 10° spatial intervals. To decide which individual product is selected for a given month and given grid cell, all input products are first gridded (monthly, 10° × 10°) to consider the fact that the spatiotemporal sampling is different for each individual product (due to different satellite sensors and algorithm-dependent quality-filtering strategies). The selected product is the median in terms of average XCO₂ per month and grid cell (note that in case of an even number of products the product which is closest to the mean is selected). The median is used primarily to remove potential outliers. The advantage of the median is also (in contrast to, for example, the arithmetic mean) that no averaging or other modifications to the input data are required. In order for a grid cell to be assigned a valid value, the following criterion has to be fulfilled: a minimum number of data products having a standard error of the mean (SEOM) of less than 1 ppm has to be available (see grey area in Fig. 3). SEOM is defined by \( \frac{1}{n} \sqrt{\sum_{i=1}^{n} \sigma_i^2} \), with \( \sigma_i \) being the (scaled; see below) XCO₂ uncertainty of the \( i \)th out of \( n \) soundings.

This means that EMMA selects for each month and each 10° × 10° grid cell exactly one product of the available individual L2 input products and then transfers all relevant information (i.e., XCO₂ and its uncertainty, related averaging kernels and a priori profile, etc.) from the selected original L2 file into the corresponding daily EMMA L2 product file. This ensures that most of the original information from the selected individual product is also contained in the merged product.

However, some modifications are applied. In order to remove (or at least to minimize) the impact of different a priori assumptions, all products are converted to common a priori CO₂ vertical profiles (see Reuter et al., 2013, for details). The new a priori profiles are obtained from the simple empirical CO₂ model (SECM, Reuter et al., 2012). SECM is essentially an empirically found function with parameters optimized using a CO₂ model (CT2017; see below). The SECM model used here is referred to as SECM2018 and is an update of the SECM model described in Reuter et al. (2012). The main difference is that SECM2018 is using a recent version of NOAA’s assimilation system CarbonTracker (Peters et al., 2007, with updates documented at: http://carbontracker.noaa.gov/, last access: 10 January 2020), namely CT2017.

SECM2018 is also used to correct for potential offsets between the individual data products by adding or subtracting a global offset (i.e., by using one constant offset value for each individual product applied globally and for the full time series). Time series of the individual data products before and after offset correction are shown in Fig. 5. Note that in Fig. 5 all data are relative to SECM2018, which is a very simple CO₂ model, and therefore all variations and trends seen in Fig. 5 are at least to some extent model errors. As can be seen from Fig. 5, the correction brings the individual data sets typically closer together without changing any of their other characteristics (e.g., their time dependence). But as can also be seen from Fig. 5, better agreement is only achieved on average, not necessarily for all products during the entire time period. For example, the GOSAT RemoTeC product (blue curve) during 2009–2012 exhibits a somewhat larger difference after the offset correction. The approximately 2 ppm (0.5 %) spike at the beginning of the time series is likely due to a positive bias of the underlying BESD data product, which has not been corrected due to the lack of reference data in this time period (see also the discussion of this aspect in Buchwitz et al., 2018). An obvious issue is also the approximately 1.5 ppm (0.4 %) discontinuity in the first half of 2014 of the PPDF-S (photon path length probability density function/simultaneous) product (light-green curve). Depending on application, this may be an issue when this product is used stand-alone, but this is not a problem for EMMA as EMMA identifies and ignores outliers.

Another modification applied to the individual L2 input products is a potential scaling of their reported uncertainty for the individual L2 soundings. The scaling factor has been chosen such that on average the uncertainty of the reported error is consistent with the standard deviation of satellite minus ground-based validation data differences (see Sect. 4.1 for the validation of the reported uncertainties via the uncertainty ratio).

In order to avoid that an individual input product, which has much more observations than the other products (such as OCO-2 compared to GOSAT), entirely dominates the EMMA product, a method has been implemented to prevent overweighting the contributions from individual L2 input data products. The method is based on limiting the number of L2 data points. For each grid cell and month, we perform the following steps: first, we compute SEOM for each algorithm. From these values, we compute the 25th percentile and divide it by \( \sqrt{2} \). The result is used as the minimum SEOM threshold. If SEOM of an individual algorithm is smaller than this threshold, a subset of soundings is randomly chosen such that SEOM becomes just larger than the threshold. If, for example, all \( \sigma_i \) are 1 ppm, then SEOM simply becomes \( 1/\sqrt{n} \).

If in this case, for example, data from four algorithms were available with \( n_1 = 60, n_2 = 80, n_3 = 100, \) and \( n_4 = 1000 \), the SEOM threshold would become \( 1/\sqrt{2n_3} \), which would effectively limit the number of soundings of the fourth algorithm to 200 (chosen randomly).

In addition to the L2 information of the selected data products, EMMA stores the following diagnostic information for each selected sounding: identifier for the selected L2 algorithm and inter-algorithm spread (IAS) within the grid box of the sounding. Within each grid box, IAS is defined as the
algorithm-to-algorithm standard deviation of the grid box averages.

By how much each individual satellite XCO₂ data product contributes to the EMMA XCO₂ product is shown in Fig. 6. Figure 6a shows the relative data weight (RDW), and Fig. 6b shows the number of soundings per month. How the RDW is defined is explained in detail in Reuter et al. (2013). In short, the RDW is defined as the relative number of soundings weighted with the corresponding (square of the inverse) uncertainty. RDW is high if a (relatively) large number of soundings contribute to the EMMA product and if these soundings have (relatively) low uncertainty compared to the other contributing products. The RDW of a product is a measure of how much information on XCO₂ this product con-
tributes to the EMMA product relative to the other contributing products. As can be seen from Fig. 6, the SCIAMACHY BESD product is the only product until early 2009, when the GOSAT time series starts. As can also be seen, OCO-2 dominates in terms of RDW and number of soundings from 2015 onwards. This is because OCO-2 provides much more data with typically better uncertainty compared to the other (GOSAT) product.

The EMMA L2 XCH\textsubscript{4} product has been generated similarly to the EMMA L2 XCO\textsubscript{2} product, i.e., using essentially the same method as described above. A difference is that the offset correction has been done with a CH\textsubscript{4} model instead of SECM2018. This model is the simple CH\textsubscript{4} climatological model (SC4C), and we use the year 2018 update referred to as SC4C2018 in the following. The SC4C2018 model is similar to SECM2018 but for XCH\textsubscript{4}. It is a model-based CH\textsubscript{4} climatology adjusted for the annual growth rate (note that this model has also been used as the climatological training and a calibration data set as described in Schneising et al., 2019). The EMMA algorithm SEOM limit controlling the minimum number of data points per grid box, month, and algorithm has been set to 12 ppb for XCH\textsubscript{4}. The impact of the offset correction for merging the XCH\textsubscript{4} product is shown in Fig. 7. Note that in Fig. 7 all data are relative to SC4C2018, which is a very simple CH\textsubscript{4} model, and therefore all variations and trends seen in Fig. 7 are at least to some extent model errors. As for CO\textsubscript{2} (Fig. 5) the offset correction typically brings the various XCH\textsubscript{4} products closer together but does not change any of their other characteristics. The PPDF-S product suffers from a discontinuity (of 8 ppb or 0.4 %) in the first half of 2014 (see above for a similar problem for PPDF-S XCO\textsubscript{2}).

Figure 8a shows the RDW, and Fig. 8b shows the number of soundings per month for all individual sensor XCH\textsubscript{4} products contributing to the XCH\textsubscript{4} EMMA product. Until early 2009, the SCIAMACHY WFMD product is the only product contributing to the EMMA product. Note that the RDW of the SCIAMACHY products drops at the end of 2005 (in contrast to the absolute number of soundings per month). The reason is the increase in the uncertainty of this product due to detector degradation (see, e.g., Schneising et al., 2011, for details). As can also be seen, the two GOSAT proxy (PR) products (i.e., CH\textsubscript{4}_GOS_OCPR and CH\textsubscript{4}_GOS_SRPR) dominate the XCH\textsubscript{4} EMMA product because they contain more soundings compared to the other (GOSAT) data products.

### 3.2 Algorithm to generate the Level 3 OBS4MIPS products

The version 4.1 L3 XCO\textsubscript{2}_OBS4MIPS and XCH\textsubscript{4}_OBS4MIPS data products have been obtained by gridding (averaging) the version 4.1 L2, i.e., XCO\textsubscript{2}_EMMA and XCH\textsubscript{4}_EMMA, products using monthly time and 5° × 5° spatial resolution. The algorithm for the generation of the OBS4MIPS products is described in Reuter et al. (2019b). Therefore, we here provide only a short overview.

For each individual product, the gridding is based on computing an arithmetic, unweighted average of all soundings falling in a grid box. For each grid box, the standard error of the mean is computed using the uncertainties contained in the corresponding EMMA product files. In order to reduce noise at least two individual observations must be present and the resulting standard error of the mean must be less than 1.6 ppm for XCO\textsubscript{2} and less than 12 ppb for XCH\textsubscript{4}.

Besides XCO\textsubscript{2} or XCH\textsubscript{4}, the final L3 product also includes (per grid box and month) the number of soundings used for averaging; the average column-averaging kernel; the average a priori profile; the standard deviation of the averaged XCO\textsubscript{2} or XCH\textsubscript{4} values; and an estimate for the total uncertainty computed as the root sum square of two values, where one value is SEOM and the other value is IAS as computed by EMMA. For cases including only one algorithm, the second value is replaced by quadratically adding spatial and seasonal accuracy determined from the TCCON validation.

### 3.3 Validation method

The validation of the merged satellite-derived XCO\textsubscript{2} and XCH\textsubscript{4} data products is based on comparisons with ground-based XCO\textsubscript{2} and XCH\textsubscript{4} TCCON observations (using version GGG2014). We present results from two somewhat different validation methods (the EMMA method, Reuter et al., 2013; and the QA/QC method, Buchwitz et al., 2017b; see below), which are similar to other validation methods used in recent years (e.g., Butz et al., 2010; Cogan et al., 2012; Dils et al., 2014; O’Dell et al., 2018; Parker et al., 2011). These methods differ with respect to details such as the chosen colocation criterion, whether the data are brought to a common a priori or not, and if yes which a priori has been used. In the following, we will highlight some of these details as relevant for the two validation methods used for this paper.

Both methods used for the validation of the L2 EMMA products are based on colocating each individual satellite XCO\textsubscript{2} (or XCH\textsubscript{4}) observation with a corresponding value obtained from TCCON using predefined spatial and temporal colocation criteria (see below). The comparisons take into account different a priori assumptions regarding the vertical profiles of CO\textsubscript{2} (or CH\textsubscript{4}) as used for the generation of the L2 input products by converting either the satellite data (QA/QC method) or the TCCON data (EMMA method) to a common a priori. This a priori correction is based on using the satellite averaging kernels and a priori profiles, which are contained (for each single observation) in the EMMA product files. The magnitude of the a priori correction (the explicit formula is shown as Eq. 3 in Dils et al., 2014) depends on the deviation (difference) of the averaging kernel from unity and on the difference of the a priori profiles. Because the averaging kernel profiles are typically close to unity (note that both satellite and the TCCON retrievals correspond to cloud-free...
conditions) and because the a priori profiles are not totally unrealistic, the a priori correction is typically very small (approximately 0.1 ppm for XCO\textsubscript{2} and 1 ppb for XCH\textsubscript{4}).

The first validation method is the EMMA quality assessment method, which is described in Reuter et al. (2013). Note that EMMA is not only a merging method but also a data quality assessment method, as the assessment of the quality of all satellite input data (listed in Tables 1 and 2) is a key aspect of EMMA. The second method is the quality assessment/quality control (QA/QC) method (Buchwitz et al., 2017b), which is applied to all satellite XCO\textsubscript{2} and XCH\textsubscript{4} data products generated for the Copernicus Climate Change Service (C3S), i.e., to the merged products but also to all the individual sensor CCI/C3S L2 input products, which are also available via the Copernicus Climate Data Store (CDS) (see products with CCI/C3S product ID listed in Tables 1 and 2).

Key differences between the QA/QC method and the EMMA method are listed as follows.
Figure 9. April 2011 XCO\textsubscript{2} at 10° × 10° spatial resolution showing (i) the individual sensor/algorithm input data sets (panels in rows 1–4; see Table 1 for details), (ii) EMMA XCO\textsubscript{2} (bottom left), and (iii) the inter-algorithm spread (IAS, 1σ) as computed by EMMA (bottom right; see main text for details). Also shown in the bottom-right panel are the locations of the TCCON sites (pink triangles) and the range of IAS values covered by them (see color bar). Note that the OCO-2 maps (row 4) are empty because this satellite was launched after April 2011 (see Fig. 10 for OCO-2 XCO\textsubscript{2}).

- Colocation criteria: QA/QC used ±2° latitude and ±4° longitude as the spatial colocation criterion, but EMMA used 500 km (both methods use the same temporal colocation criterion of 2 h).

- Filtering criterion surface elevation: EMMA requires a surface elevation difference of less than 250 m between a TCCON site and satellite footprints, whereas the QA/QC does not use this filtering criterion.

- A priori correction: both methods correct for the use of different a priori CO\textsubscript{2} vertical profiles in the various retrieval algorithms, but QA/QC uses the TCCON

a priori as common a priori, whereas EMMA uses the SECM2018 model for CO\textsubscript{2} and the SC4C2018 model for CH\textsubscript{4} (see Sect. 3.1).

- Approach to quantify seasonal bias and linear bias trend: the EMMA method is based on fitting a trend model, which includes an offset term, a slope term, and a sine term for seasonal fluctuations (see Reuter et al., 2019c) and computes the seasonal bias from the standard deviation of the fitted seasonal fluctuation term and obtains the bias trend and its uncertainty from the fitted slope term. The QA/QC method (Buchwitz et al., 2019a) uses (only) a linear fit to obtain the bias trend...
and its uncertainty and computes the seasonal bias from the standard deviation of the seasonal biases (as also done by Dils et al., 2014, for their quantity seasonality).

– Criteria for enough data: both algorithms use several different thresholds for the required minimum number of colocations per TCCON site and minimum length of overlapping TCCON time series.

Despite all these differences, quite similar overall figures of merit have been obtained with both methods (see results section, Sect. 4). This indicates that the overall data quality results do not critically depend on the details of the assessment method (the same conclusion has also been reported for earlier comparisons of results from different assessment methods, e.g., Buchwitz et al., 2015, 2017b).

4 Results and discussion

4.1 Products XCO2_EMMA and XCO2_OBS4MIPS (v4.1)

When generating an EMMA product, a set of standard figures are generated such as Figs. 5 and 6 already discussed but also maps of the EMMA product and of the various input data products for all months of the 2003–2018 time period. Two of these figures are shown here, namely the figures for April 2011 (Fig. 9) and April 2015 (Fig. 10) (note that 2011 is the last full year with data from SCIAMACHY and that 2015 is the first full year with OCO-2 data). The maps in the first four rows of Figs. 9 and 10 show the individual sensor/algorithm L2 input data. As can be seen, the spatial XCO2 patterns are quite similar (e.g., north–south gradient),
Table 4. Overview validation results at TCCON sites for data product XCO2_EMMA (version 4.1).

<table>
<thead>
<tr>
<th>TCCON site</th>
<th>Random error single obs. (1σ) (ppm)</th>
<th>Uncertainty ratio (–)</th>
<th>Overall bias bias satellite – TCCON (ppm)</th>
<th>Seasonal bias bias satellite – TCCON (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>QA/QC EMMA</td>
<td>QA/QC EMMA</td>
<td>QA/QC EMMA</td>
<td>QA/QC EMMA</td>
<td>QA/QC EMMA</td>
</tr>
<tr>
<td>SOD</td>
<td>1.19 1.33</td>
<td>1.16 1.10</td>
<td>0.57 0.18</td>
<td>– 0.22</td>
</tr>
<tr>
<td>BIA</td>
<td>1.11 1.16</td>
<td>1.44 1.37</td>
<td>0.06 0.10</td>
<td>– 0.26</td>
</tr>
<tr>
<td>BRE</td>
<td>1.66 1.30</td>
<td>0.90 1.14</td>
<td>1.09 0.55</td>
<td>– 0.15</td>
</tr>
<tr>
<td>KAR</td>
<td>1.45 1.40</td>
<td>0.96 0.99</td>
<td>1.18 0.52</td>
<td>1.17 0.40</td>
</tr>
<tr>
<td>PAR</td>
<td>1.30 –</td>
<td>0.99 –</td>
<td>–0.49 –</td>
<td>– –</td>
</tr>
<tr>
<td>ORL</td>
<td>1.18 1.40</td>
<td>1.15 1.04</td>
<td>0.30 0.45</td>
<td>0.75 0.39</td>
</tr>
<tr>
<td>GAR</td>
<td>1.48 1.46</td>
<td>0.91 1.04</td>
<td>1.28 0.36</td>
<td>0.83 0.22</td>
</tr>
<tr>
<td>PFA</td>
<td>1.08 1.27</td>
<td>1.31 1.11</td>
<td>0.09 –0.37</td>
<td>0.70 0.18</td>
</tr>
<tr>
<td>LAM</td>
<td>1.26 1.47</td>
<td>1.08 0.95</td>
<td>–0.09 –0.61</td>
<td>0.17 0.38</td>
</tr>
<tr>
<td>TSU</td>
<td>1.54 –</td>
<td>0.95 –</td>
<td>0.54 –</td>
<td>0.61 –</td>
</tr>
<tr>
<td>EDW</td>
<td>1.48 –</td>
<td>0.78 –</td>
<td>1.16 –</td>
<td>0.21 –</td>
</tr>
<tr>
<td>CAL</td>
<td>1.57 –</td>
<td>0.75 –</td>
<td>–0.46 –</td>
<td>0.15 –</td>
</tr>
<tr>
<td>SAG</td>
<td>1.41 –</td>
<td>1.06 –</td>
<td>–0.17 –</td>
<td>0.31 –</td>
</tr>
<tr>
<td>ASC</td>
<td>1.16 –</td>
<td>1.44 –</td>
<td>0.65 –</td>
<td>0.60 –</td>
</tr>
<tr>
<td>DAR</td>
<td>1.06 1.06</td>
<td>1.01 1.02</td>
<td>–0.23 0.52</td>
<td>0.66 0.34</td>
</tr>
<tr>
<td>REU</td>
<td>0.75 –</td>
<td>1.73 –</td>
<td>0.29 –</td>
<td>– –</td>
</tr>
<tr>
<td>WOL</td>
<td>1.21 1.19</td>
<td>1.00 1.00</td>
<td>–0.53 –0.66</td>
<td>0.24 0.17</td>
</tr>
<tr>
<td>LAU</td>
<td>1.13 –</td>
<td>1.03 –</td>
<td>0.14 –</td>
<td>0.10 –</td>
</tr>
<tr>
<td>Mean</td>
<td>1.28 1.30</td>
<td>1.15 1.07</td>
<td>0.30 0.10</td>
<td>0.50 0.27</td>
</tr>
<tr>
<td>SD</td>
<td>0.23 0.14</td>
<td>0.23 0.12</td>
<td>0.60 0.48</td>
<td>0.33 0.10</td>
</tr>
</tbody>
</table>

Figure 11. Average XCO2 inter-algorithm spread (1σ) during 2003–2018. As can be seen, the scatter is typically around 1 ppm except over parts of the tropics (in particular central Africa), the Himalayas, and at high latitudes, where the scatter can be larger. but there are also significant differences, especially with respect to the spatial coverage. The spatial coverage depends on time and is related to the different satellite instruments but also due to algorithm-dependent quality filtering. The largest differences are between the SCIAMACHY BESD product (top left in Fig. 9) compared to the other products, as the SCIAMACHY product is limited to observations over land, whereas the GOSAT and OCO-2 products also have some ocean coverage due to the ocean-glint mode, which permits the acquisition of an adequate signal (and therefore also signal-to-noise ratio) also over the ocean (note that the reflectivity of water is poor outside of sunglint conditions in

www.atmos-meas-tech.net/13/789/2020/ Atmos. Meas. Tech., 13, 789–819, 2020
the used SWIR spectral regions around 1.6 and 2 µm). The EMMA product is shown in the bottom-left panels of Figs. 9 and 10, and in the bottom-right panel IAS is shown, which quantifies the level of agreement (or disagreement) among the various satellite input data sets. The IAS maps also show the location of the TCCON sites (pink triangles) and the IAS values at the TCCON sites (see pink triangles above the color bar). As can be seen, the TCCON sites are typically located outside of regions where the IAS is highest.

The average IAS for the entire time period 2003–2018 is shown in Fig. 11. As can be seen, the scatter is typically in the range 0.6–1.1 ppm with the exception of parts of the tropics, in particular central Africa, the Himalayas, parts of southeast Asia, and high latitudes. High latitudes typically correspond to large solar zenith angles, which is a challenge for accurate satellite XCO₂ retrievals, as this typically corresponds to low signal and therefore low signal-to-noise ratio resulting in enhanced scatter of the retrieved XCO₂. In areas with frequent cloud coverage, such as parts of the tropics, sampling is sparse and this may also contribute to a larger scatter.

Detailed validation results for all individual sensors and the EMMA XCO₂ Level 2 data products are shown in Appendix A (Fig. A1) for all TCCON sites. The validation results are summarized in Table 4 (per site) and Table 5 (overall) together with the corresponding results of the QA/QC assessment method.

Table 4 lists all TCCON sites, which fulfill either the EMMA method or the QA/QC method criteria with respect to a minimum number of colocations and length of time series. Listed are the numerical values (in ppm), which have been computed for several figures of merit. This includes (i) the overall estimation of the single-observation random error computed as the standard deviation of the satellite minus TCCON differences; (ii) the uncertainty ratio, which is the ratio of the mean value of the reported (1σ) uncertainty to the standard deviation of the satellite–TCCON difference (computed to validate the reported uncertainties); (iii) the overall bias computed as the mean value of the satellite–TCCON differences; and (iv) the seasonal bias, computed as the standard deviation of the biases determined for the four seasons. Also shown in the last two rows are the mean value and the standard deviation of the values listed per TCCON site in the rows above. Several of these values have been used to compute the values listed in Table 5, which shows the overall summary of the quality assessment.

Table 5 lists (i) the mean value of the single-observation random error, (ii) the global bias computed as the mean value of the biases at the various TCCON sites, (iii) the regional bias computed as the standard deviation of the biases at the various TCCON sites, (iv) the mean seasonal bias, and (v) the spatiotemporal bias computed as the root sum square of the regional and of the seasonal bias. The spatiotemporal bias is used to quantify the achieved performance for relative accuracy, which characterizes the spatially and temporally varying component of the bias (i.e., neglects a possible global bias (global offset), which is reported separately).

The linear bias trend has also been computed by fitting a line to the satellite–TCCON differences (not shown here). The mean value of the linear trend (slope) and its uncertainty (1σ, obtained from the standard deviation of the slope at the various TCCON sites) are −0.05 ± 0.06 ppm yr⁻¹ for the EMMA method and −0.06 ± 0.09 ppm yr⁻¹ for the QA/QC method. This means that no significant long-term bias trend has been detected; i.e., the satellite product is stable.

As can be seen from Table 5, the values computed independently using the EMMA and the QA/QC assessment methods are quite similar, which gives not only confidence in the overall quality assessment documented in Table 5 but also in the products and the used validation methods.

Note however that the quality of the satellite data (at least at TCCON sites) is very likely better than Table 5 suggests (i) because the TCCON retrievals are not free of errors (the 1σ XCO₂ uncertainty is about 0.4 ppm; Wunch et al., 2010) and (ii) because of the representation error originating from the (real) spatiotemporal variability of XCO₂ around the TCCON sites. The overall error related to this is difficult to quantify, but some indication can potentially be obtained by additional assessment results such as the one shown in Fig. 12. Figure 12 shows the biases as obtained with the EMMA method at the various TCCON sites used for the EMMA method comparisons. Shown are not only the mean satellite–TCCON differences as obtained for the EMMA product but also for all the individual sensor/algorithm input products. The differences are shown as anomalies with respect to the mean; i.e., the sum of the differences in each row is zero. This is equivalent to assuming that for a given satellite product the mean value over all TCCON sites is zero. As can be seen from Fig. 12, the satellite–TCCON differences are dominantly positive (orange and red colors) for higher-latitude TCCON sites and mostly negative (blue colors) for lower-latitude TCCON sites. In order to rule out that this is an artifact of the EMMA assessment method, the overall biases computed with the QA/QC method and biases computed

| Parameter Assessment method Mean |
|----------------------------------|---------------------------|
| QA/QC   | EMMA   | Mean   |
| Random error single observations (1σ) (ppm) | 1.28 | 1.30 | 1.29 |
| Global bias (ppm) | 0.30 | 0.10 | 0.20 |
| Regional bias (1σ) (ppm) | 0.60 | 0.48 | 0.54 |
| Seasonal bias (1σ) (ppm) | 0.50 | 0.27 | 0.39 |
| Spatiotemporal bias (1σ) (ppm) | 0.78 | 0.55 | 0.66 |
Table 6. TCCON XCO$_2$ bias in parts per million (ppm; satellite – TCCON). “–” means that the number of available colocations is less than the threshold required by the corresponding assessment method. Note that this table includes only a subset of the 10 sites shown in Fig. 12, namely only those sites with a mean bias being considerably (more than 1.5 times) larger than the standard deviation of the biases.

<table>
<thead>
<tr>
<th>Satellite product</th>
<th>Assessment method</th>
<th>TCCON site</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SOD</td>
</tr>
<tr>
<td>XCO2_EMMA</td>
<td>QA/QC</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>EMMMA</td>
<td>0.18</td>
</tr>
<tr>
<td>CO2_SCI_BESD</td>
<td>QA/QC</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>EMMMA</td>
<td>0.32</td>
</tr>
<tr>
<td>CO2_GOS_OCFP</td>
<td>QA/QC</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>EMMMA</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>DP$^a$</td>
<td>0.57</td>
</tr>
<tr>
<td>CO2_GOS_SRFP</td>
<td>QA/QC</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>EMMMA</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>DP$^b$</td>
<td>0.89</td>
</tr>
<tr>
<td>GOS NIES</td>
<td>EMMMA</td>
<td>0.29</td>
</tr>
<tr>
<td>GOS NASA</td>
<td>EMMMA</td>
<td>1.04</td>
</tr>
<tr>
<td>OCO-2 FOCAL</td>
<td>EMMMA</td>
<td>0.02</td>
</tr>
<tr>
<td>OCO-2 NASA</td>
<td>EMMMA</td>
<td>0.40</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>0.44</td>
</tr>
<tr>
<td>Standard deviation</td>
<td></td>
<td>0.28</td>
</tr>
</tbody>
</table>

Assessment method DP is the method used by the data provider. For $^a$ see Boesch et al. (2019). For $^b$ see Wu et al. (2019).

The XCO2_OBS4MIPS product has also been directly compared with TCCON using a comparison method based on the comparison of the monthly satellite product with TCCON monthly mean values. The results are shown in Fig. 13. As can be seen, the mean difference (satellite – TCCON) is 0.18 ppm (which is close to the mean value of the global bias of 0.20 ppm listed in Table 5), the standard deviation is 1.18 ppm (as expected, because of the spatiotemporal averaging, which is somewhat smaller than the value of 1.29 ppm obtained for the XCO2_EMMA product listed in Table 5), and the linear correlation coefficient is 0.99. The spatiotemporal bias, computed as the standard deviation of 3-monthly averages at the TCCON sites listed in Fig. 13, is 0.7 ppm.

Figure 1 presents an overview of the XCO$_2$ data product in terms of time series for three latitude bands and global maps. XCO$_2$ is increasing almost linearly during the 16-year time period (for a discussion of the derived annual growth rates see Sect. 4.3). The main reason for this increase is CO$_2$ emission due to burning of fossil fuels (Le Quéré et al., 2018). The seasonal cycle, which is caused primarily by quasi-regular uptake and release of atmospheric CO$_2$ by the terrestrial vegetation due to photosynthesis and respiration (e.g., Kaminski et al., 2017, Yin et al., 2018), is most pronounced over the Northern Hemisphere. The half-yearly maps for 2003 are based on SCIAMACHY on board ENVISAT (Burrows et al., 1995; Bovensmann et al., 1999) satellite data, and the maps for 2018 contain data from the GOSAT (since 2009) (Kuze et al., 2017; Kussul et al., 2013; Osadchuk et al., 2015; Ziegenhagen et al., 2015) satellite data.
al., 2016) and OCO-2 (since 2014) (Crisp et al., 2004) satellites. GOSAT and OCO-2 also provide good-quality XCO\(_2\) retrievals over the oceans due to their sunglint observation mode.

### 4.2 Products XCH4_EMMA and XCH4_OBS4MIPS (v4.1)

As for XCO\(_2\), monthly maps have also been generated for the EMMA XCH\(_4\) data product. Two examples are shown in Fig. 14 for September 2010 and in Fig. 15 for September 2018. The individual sensor XCH\(_4\) input data are shown in the first four rows, and the EMMA XCH\(_4\) product is shown in the bottom-left panel. The bottom-right panel shows the IAS. As can be seen, the spatial patterns of the XCH\(_4\) maps are similar but not identical. The IAS shows a quite large variability. The scatter is larger compared to the corresponding XCO\(_2\) IAS (Figs. 9 and 10, bottom-right panels), and spatially the grid cells with larger spread are more equally distributed over the globe but with largest differences over the southern part of Asia.

Detailed validation results are shown in Appendix A (Fig. A2), and the validation results are summarized in Tables 7 and 8, which have the same structure as the corresponding XCO\(_2\) tables (Tables 4 and 5). These tables also list the results of the QA/QC assessment method, which results in quite similar (within a few ppb) overall quality assessment results (Table 8) as obtained with the EMMA method. The linear bias trend has also been computed by fitting a line to the satellite–TCCON differences (not shown here). The mean value of the linear trend (slope) and its uncertainty (1\(\sigma\), obtained from the standard deviation of the slope at the various TCCON sites) are \(-0.1 \pm 0.4\) ppb yr\(^{-1}\) for the EMMA method and \(0.5 \pm 0.8\) ppb yr\(^{-1}\) for the QA/QC method. As for XCO\(_2\), this means that no significant long-term bias trend has been detected; i.e., the satellite product is stable.

Figure 16 shows the TCCON station XCH\(_4\) bias anomaly as also shown for XCO\(_2\) in Fig. 12; i.e., Fig. 16 shows the biases as obtained with the EMMA method at the various TCCON sites used for the EMMA method comparisons. As for XCO\(_2\) not only the mean satellite–TCCON differences as obtained for the EMMA product are shown but also the differences for all the individual sensor/algorithim input products. The differences are shown as anomalies with respect to the mean; i.e., the sum of the differences in each row is zero. As can be seen from Fig. 16, the pattern of satellite–TCCON XCH\(_4\) differences has some similarity with the XCO\(_2\) difference pattern shown in Fig. 12. For example, the differences are mostly positive at Sodankylä and Garmisch-Partenkirchen and mostly negative at Darwin and Wollongong. But there are also significant differences, for example,
with respect to the sign of the bias (e.g., Park Falls, Bremen, Karlsruhe).

The XCH4_OBS4MIPS product has also been directly compared with TCCON (Fig. 17) using the same method as also used for product XCO2_OBS4MIPS (Fig. 13). As can be seen from Fig. 17, the mean difference (satellite – TCCON) is −2.88 ppb (which is close to the mean value of the global bias of −2.0 ppb of product XCH4_EMMA listed in Table 8), the standard deviation is 8.65 ppb (as expected, because of the averaging, which is somewhat smaller than the value of 17.4 ppb obtained for the XCH4_EMMA product listed in Table 8), and the linear correlation coefficient is 0.97.

Figure 2 presents an overview of the XCH4 data product in terms of time series for three latitude bands and global maps. As can be seen, XCH4 was nearly constant during 2003–2006 (apart from seasonal fluctuations) but has been increasing since 2007 (for a discussion of the trend and annual growth rates see Sect. 4.3). The reason for this is likely a combination of increasing natural (e.g., wetlands) and anthropogenic (e.g., fossil fuel related) emissions and possibly decreasing sinks (hydroxyl, OH, radical), but it does not seem currently possible to be more definitive (e.g., Worden et al., 2017; Nisbet et al., 2019; Turner et al., 2019; Howarth, 2019; Schaefer, 2019).

### Table 8. Validation summary for data product XCH4_EMMA (version 4.1).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Assessment method</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QA/QC</td>
<td>EMMA</td>
</tr>
<tr>
<td>Random error single observations</td>
<td>21.2</td>
<td>13.6</td>
</tr>
<tr>
<td>Global bias (ppb)</td>
<td>−4.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Regional bias (1σ) (ppb)</td>
<td>5.2</td>
<td>3.7</td>
</tr>
<tr>
<td>Seasonal bias (1σ) (ppb)</td>
<td>2.2</td>
<td>2.5</td>
</tr>
<tr>
<td>Spatiotemporal bias (1σ) (ppb)</td>
<td>5.6</td>
<td>4.4</td>
</tr>
</tbody>
</table>

![Figure 12: Average XCO2 differences (satellite – TCCON) for the different satellite XCO2 products at 10 TCCON sites as used by the EMMA assessment method. The differences are shown as anomalies; i.e., the sum of the values corresponding to a given row is zero. Note that here “ACOS” refers to NASA’s ACOS algorithm as applied to GOSAT and that “NASA” refers to NASA’s ACOS algorithm as applied to OCO-2.](image)

![Figure 13: Summary of the comparison of product XCO2_OBS4MIPS with TCCON monthly mean XCO2 (each symbol corresponds to one month and to one TCCON site; each color corresponds to a different TCCON site; TCCON site colors and site IDs (see Table 3) are shown on the right). The comparison is based on 1446 monthly values. The mean difference (satellite – TCCON) is 0.18 ppm and the standard deviation of the difference is 1.18 ppm. The linear correlation coefficient R is 0.99.](image)
Figure 14. September 2010 XCH$_4$ at $10^5 \times 10^5$ spatial resolution showing (i) the individual sensor/algorith input data sets (panels in rows 1–4; see Table 2 for details), (ii) EMMA XCH$_4$ (bottom left), and (iii) the inter-algorithm spread (IAS, 1$\sigma$) as computed by EMMA (bottom right; see main text for details). Also shown in the bottom-right panel are the locations of the TCCON sites (pink triangles) and the range of IAS values covered by them (see color bar).

ministration (NOAA) (shown in blue color in Fig. 18b), which are based on marine surface CO$_2$ observations (ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_gr_gl.txt; last access: 30 July 2019). As can be seen from Fig. 18b, the agreement of the satellite-derived XCO$_2$ growth rates with the NOAA surface-CO$_2$-based growth rates is better from year 2010 onwards compared to the time period before when the EMMA data set consists only of one SCIAMACHY data set instead of the full ensemble. For 2018, the XCO$_2$ growth rate is $2.1 \pm 0.5$ ppm yr$^{-1}$, which is lower than the NOAA surface CO$_2$ growth rate of $2.43 \pm 0.08$ ppm yr$^{-1}$. Note that the 1$\sigma$ uncertainty ranges of the two growth rate estimates overlap, which indicates that the two growth rate estimates are consistent.

The growth rate of atmospheric methane is also an important quantity (e.g., Nisbet et al., 2019). The method of Buchwitz et al. (2018) has now also been used to compute annual XCH$_4$ growth rates from satellite XCH$_4$ retrievals. Figure 19a shows the time series of the globally averaged OBS4MIPS version 4.1 XCH$_4$ data product over land. As shown by the linear fit, the average growth rate is $7.9 \pm 0.2$ ppb yr$^{-1}$ during 2010–2018, i.e., for the time period where an ensemble of GOSAT data has been used. The annual growth rates are shown in Fig. 19b for the satellite-derived XCH$_4$ (red) and for the NOAA growth rates (ftp:
Figure 15. As Fig. 14 but for September 2018. Note that the SCIAMACHY/WFMD map (top left) is empty because this product ended in April 2012 (see Fig. 14 for SCIAMACHY/WFMD XCH₄). For product GOSAT/PPDF (row 4) no data were available for this month (see Fig. 14 for GOSAT/PPDF XCH₄).

//aftp.cmdl.noaa.gov/products/trends/ch4/ch4_gr_gl.txt; last access: 30 July 2019) derived from marine surface CH₄ observations. For 2018, the XCH₄ growth rate is 10 ± 6 ppb yr⁻¹, which is close to the NOAA surface CH₄ growth rate of 9.46 ± 0.56 ppb yr⁻¹.

5 Summary and conclusions

Satellite-derived ensemble XCO₂ and XCH₄ data products have been generated and validated. These data products are the version 4.1 Level 2 (L2) products XCO₂_EMMA and XCH₄_EMMA and the Level 3 (L3) products XCO₂_OBS4MIPS and XCH₄_OBS4MIPS and cover the time period 2003–2018. The data products are freely avail-

able for interested users via the Copernicus Climate Data Store (CDS, https://cds.climate.copernicus.eu/, last access: 10 January 2020), where also earlier versions of these data products are accessible. The L2 products have been generated with an adapted version of the EMMA algorithm (Reuter et al., 2013), and the L3 products have been generated by gridding (averaging) the EMMA L2 product to obtain products at monthly time and 5° × 5° spatial resolution in Obs4MIPS format. The products have been validated by comparisons with TCCON ground-based XCO₂ and XCH₄ retrievals using TCCON version GGG2014.

From January 2003 to March 2009 the products are based on SCIAMACHY/ENVISAT, and from April 2009 onwards the products use an ensemble of one SCIAMACHY (until
any of the individual sensor input data sets) with as-high-as-
possible accuracy including all information needed, e.g., for
surface flux inverse modeling. The median approach helps
to reduce the occurrence of potential outliers and thus reduces
spatial and temporal biases in the generated data products.

Detailed quality assessment results based on comparisons
with TCCON ground-based retrievals have been presented.
We found that the XCO$_2$ Level 2 data set at the TCCON
validation sites can be characterized by the following
figures of merit (the corresponding values for the Level 3
product are listed in brackets) – single-observation ran-
dom error (1σ): 1.29 ppm (monthly: 1.18 ppm); global bias:
0.20 ppm (0.18 ppm); and spatiotemporal bias or relative ac-
curacy (1σ): 0.66 ppm (0.70 ppm). The corresponding val-
dues for the XCH$_4$ products are single-observation random
error (1σ): 17.4 ppb (monthly: 8.7 ppb); global bias: −2.0 ppb
(−2.9 ppb); spatiotemporal bias (1σ): 5.0 ppb (4.9 ppb). It
has also been found that the data products exhibit very good
long-term stability as no significant linear bias trends have
been identified.

The new data sets have also been used to derive annual
XCO$_2$ and XCH$_4$ growth rates, which are in reasonable
to good agreement with growth rates from the National Oceanic
and Atmospheric Administration (NOAA) based on marine
surface observations.

An important application for the EMMA products is to
use them together with inverse modeling to obtain improved
information on regional-scale CO$_2$ (e.g., Houweling et al.,
2015) and CH$_4$ (e.g., Alexe et al., 2015) surface fluxes. Ap-
lications for the corresponding OBS4MIPS products are,
for example, climate model comparisons (e.g., Lauer et al.,
2017) and studies related to annual growth rates (e.g., Buch-
witz et al., 2018). It is however important to note that these
merged products are not necessarily the most optimal prod-
ucts for all applications as they do not contain all data from
a given satellite sensor. For example, users interested primar-
ily in emissions from power plants or other localized CO$_2$
sources will prefer the original OCO-2 Level 2 data product
(e.g., Nassar et al., 2017; Reuter et al., 2019a). Especially for
users interested in only parts of the time series it is recom-
ended to use the individual sensor products in addition to
the merged product as this may significantly increase the ro-
bustness, reliability, and uncertainty characterization of key
findings.
Figure 18. (a) Monthly values of the globally averaged $\text{XCO}_2$ (over land) as computed from the OBS4MIPS version 4.1 $\text{XCO}_2$ data product. The corresponding annual mean $\text{XCO}_2$ values are also listed. The increase during 2010–2018 is $2.28 \pm 0.04$ ppm yr$^{-1}$ as obtained via a linear fit. (b) Annual $\text{XCO}_2$ growth rates (red, with $1\sigma$ uncertainties; the corresponding numerical values are also listed with $1\sigma$ uncertainty in brackets) and $\text{CO}_2$ growth rates from NOAA (shown in blue) obtained from marine surface $\text{CO}_2$ observations.

Figure 19. (a) Monthly values of the globally averaged $\text{XCH}_4$ (over land) as computed from the OBS4MIPS version 4.1 $\text{XCH}_4$ data product. The corresponding annual mean $\text{XCH}_4$ values are also listed. The increase during 2010–2018 is $7.9 \pm 0.2$ ppb yr$^{-1}$ as obtained via a linear fit. (b) Annual $\text{XCH}_4$ growth rates (red, with $1\sigma$ uncertainties; the corresponding numerical values are also listed with $1\sigma$ uncertainty in brackets) and $\text{CH}_4$ growth rates from NOAA (shown in blue) obtained from marine surface $\text{CH}_4$ observations.
Appendix A

In this appendix, detailed validation results are shown for the individual sensor and EMMA XCO$_2$ and XCH$_4$ Level 2 data products.

The comparison of the various XCO$_2$ data products with TCCON XCO$_2$ at 10 TCCON sites is shown in Fig. A1. These 10 TCCON sites fulfill the EMMA criteria in terms of a sufficiently large number of colocations as defined to obtain robust conclusions per site. The individual soundings of the EMMA XCO$_2$ product are shown as white circles with a black border. As can be seen, they are located within (mostly close to the center of) the range of values of the individual sensor/algorithm XCO$_2$ values, which is expected.

Figure A2 shows the comparison of the EMMA XCH$_4$ product (white circles with a black border) and of the individual sensor XCH$_4$ input products with TCCON XCH$_4$ originating from the EMMA assessment method. As for the EMMA XCO$_2$ product (Fig. A1), the EMMA XCH$_4$ is located near the center of the clouds of XCH$_4$ values, as expected.
Figure A1. \( \text{XCO}_2 \) time series at 10 TCCON sites during January 2009–December 2018 as obtained using the EMMA quality assessment method. TCCON GGG2014 \( \text{XCO}_2 \) is shown as thick black dots, the individual satellite L2 input products are shown as colored dots, and the EMMA product is shown as white circles with black borders. The derived numerical values are listed in Table 4.
Figure A2. XCH₄ time series at 10 TCCON sites during April 2010–December 2018 as obtained using the EMMA quality assessment method. TCCON GGG2014 XCH₄ is shown as thick black dots, the individual satellite L2 input products are shown as colored dots, and the EMMA product is shown as a white circles with black borders. The derived numerical values are listed in Table 7.
Data availability. The EMMA and OBS4MIPS XCO2 and XCH4 version 4.1 data products (but also several data sets used as input; see data sets with CCI/C3S product ID in Tables 1 and 2) are available (from early 2020 onwards) via the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/, ECMWF, 2020a) Climate Data Store (CDS, https://cds.climate.copernicus.eu/, ECMWF, 2020f), including documentation such as the product user guides (http://www.iup.uni-bremen.de/carbon_ghg/docs/CCI/CDR3_2003-2018/PGUGS/CCS_D312b_Lot2.3.2.3-v1.0_PUGS-GHG_MAIN_v3.1.pdf, Buchwitz et al., 2019c; http://www.iup.uni-bremen.de/carbon_ghg/docs/CCI/CDR3_2003-2018/PGUGS/CCS_D312b_Lot2.3.2.3-v1.0_PUGS-GHG_ANNEX-D_v3.1.pdf, Reuter et al., 2019d).

Author contributions. MR generated the EMMA and OBS4MIPS XCO2 and XCH4 version 4.1 data sets. MR and MB performed the data analysis. MB wrote the first version of the paper with support of MR. The following authors provided input data or expertise on data sets: MR, MB, OS, SN, HB, JPB, HBoe, ADN, JA, RJP, PS, LW, OPH, IA, AK, HS, KS, YY, IM, DC, CWO’D, JN, CP, TW, VAV, NMD, DWTG, RK, DP, FH, RS, YVT, KS, SR, MKS, DD, LG, DGF, LTI, CMR, CR, and DS. All authors contributed to significantly improve the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The generation of the EMMA Level 2 and OBS4MIPS Level 3 data sets and the corresponding data analysis has been funded primarily by the European Union (EU) via the Copernicus Climate Change Service (C3S, https://climate.copernicus.eu/, last access: 10 January 2020) managed by the European Centre for Medium-Range Weather Forecasts (ECMWF).

The work presented here strongly benefited from additional funding by the European Space Agency (ESA) via ESA’s Climate Change Initiative (CCI, http://www.esa-ghg-cci.org/, last access: 10 January 2020) projects GHG-CCI/1GHG-CCI+.

The further development of the FOCAL retrieval algorithm used to generate the OCO-2/FOCAL XCO2 input data product would not have been possible without cofunding from the European Union (EU) Horizon 2020 (H2020) research and innovation program projects CHE (grant agreement no. 776186) and VERIFY (grant agreement no. 776810). The generation of the XCO2_OBS4MIPS product also benefited from cofunding from EU H2020 project CC4C (grant agreement no. 821003).

We thank several space agencies for making available satellite Level 1 (L1) input data: ESA/DLR for SCIAMACHY L1 data, JAXA for GOSAT Level 1B data, and NASA for the OCO-2 L1 data product. We also thank ESA for making the GOSAT L1 product available via the ESA Third Party Mission (TPM) archive.

We thank NIES for the operational GOSAT XCO2 and XCH4 Level 2 products (obtained from: https://data2.gosat.nies.go.jp/, last access: 4- September 2019) and the NASA team for the GOSAT and OCO-2 ACOS Level 2 XCO2 products (the NASA GOSAT L2 data product was obtained from: https://oco2.gesdisc.eosdis.nasa.gov/data/GOSAT_TANSO_Level2/ACOS_L2_Lite_FP/7.3/.

www.atmos-meas-tech.net/13/789/2020/

www.atmos-meas-tech.net/13/789/2020/ Atmos. Meas. Tech., 13, 789–819, 2020


M. Reuter et al.: Carbon and climate applications


Reuter, M., Buchwitz, M., and Schneising-Weigel, O.: Algorithm Theoretical Basis Document (ATBD) – ANNEX D for
M. Reuter et al.: Carbon and climate applications


