Document Ref.: 5.1

CMUG Phase 2 Deliverable

Number:	**D5.1 v2:** Advanced version of the ESMValTool with ESA CCI datasets and user guide released to CMUG and ESA CCI teams
Due date:	**30 June 2016**
Submission date:	**30 August 2016**
Edition:	**1.0.2**

Climate Modelling User Group

Deliverable 5.1 v2

Advanced version of the ESMValTool with ESA CCI datasets and updated user guide released to CMUG and ESA CCI teams

Centres providing input: DLR, MPI-M (now LMU), SMHI

Edition	Date	Status
1.0.2	30 August 2016	Version for ESA

Deliverable 5.1 v2

Advanced version of the ESMValTool with ESA CCI datasets and updated user guide released to CMUG and ESA CCI teams

Contents

Number: **D5.1 v2:** Advanced version of the ESMValTool with ESA CCI datasets and user guide released to CMUG and ESA CCI teams
Due date: **30 June 2016**
Submission date: **30 August 2016**
Edition: **1.0.2**

Advanced version of the ESMValTool with ESA CCI datasets and updated user guide released to CMUG and ESA CCI teams

1 Summary

Model benchmarking initiatives have become increasingly important to evaluate the quality of coupled Earth System Models (ESMs) and to support the model development process. In the frame of the Coupled Model Intercomparison Project (CMIP), the WGNE/WGCM Climate Model Metrics Panel has been established to define model performance metrics for model-data intercomparison and analyse various aspects of ESM simulations with a multitude of observational datasets. However, ESA data is currently not used in the context of routine model evaluation. There is a strong need in the international community to develop standardized reference datasets, diagnostic and performance metrics tools for evaluation of climate model simulations. At present, basically each research group has established its own approach. Model benchmarking is also important for the Climate Research Groups (CRGs) within the CCI and various CMUG activities. CMUG therefore contributes to the development of a community-wide Earth System Model Evaluation Tool (ESMValTool) that is currently developed by different partners in different projects under DLR lead. CMUG's specific contribution is to include ESA CCI data in the ESMValTool so that they can be routinely used to evaluate models participating in CMIP6.

2 Purpose, scope and content of this report

This deliverable includes a brief overview of an advanced ESMValTool version (v1.0.2), which will both be released together with an updated user's guide to CMUG and the ESA CCI teams in fall 2016. CMUG contributes to this release with technical development of the tool, documentation in form of an overview and an update of the detailed user's guide, and by adding ESA CCI datasets to this release. This advanced version of the ESMValTool is shared for test purposes among CMUG partners in a password restricted area. CMUG partners who wish to work with the advanced versions of the ESMValTool need to register for the password restricted ESMValTool development environment which is hosted at DLR under a subversion controlled repository and a wiki page with documentation. The ESA CCI datasets included in v1.0.2 are ESA CCI aerosol, cloud, sea ice, soil moisture, sea surface temperature, ozone, greenhouse gases, and land cover. Selected new ESA CCI datasets are used in the performance metrics plot as reference dataset to evaluate CMIP5 models. This document will be updated next year to Edition 3 and will contain a summary of the analysis of the performance of the CCI datasets presented in Lauer et al. (in preparation).

3 Brief overview of the ESMValTool

A detailed description of ESMValTool (v1.0) has been published in Eyring et al. (2016a). The ESMValTool's user's and developer's guide is available as supplementary material of Eyring et al.

(2016a) and on the ESMValTool website http://www.esmvaltool.org/. In this section we give a brief overview of ESMValTool (v1.0) which is schematically depicted in Figure 1.

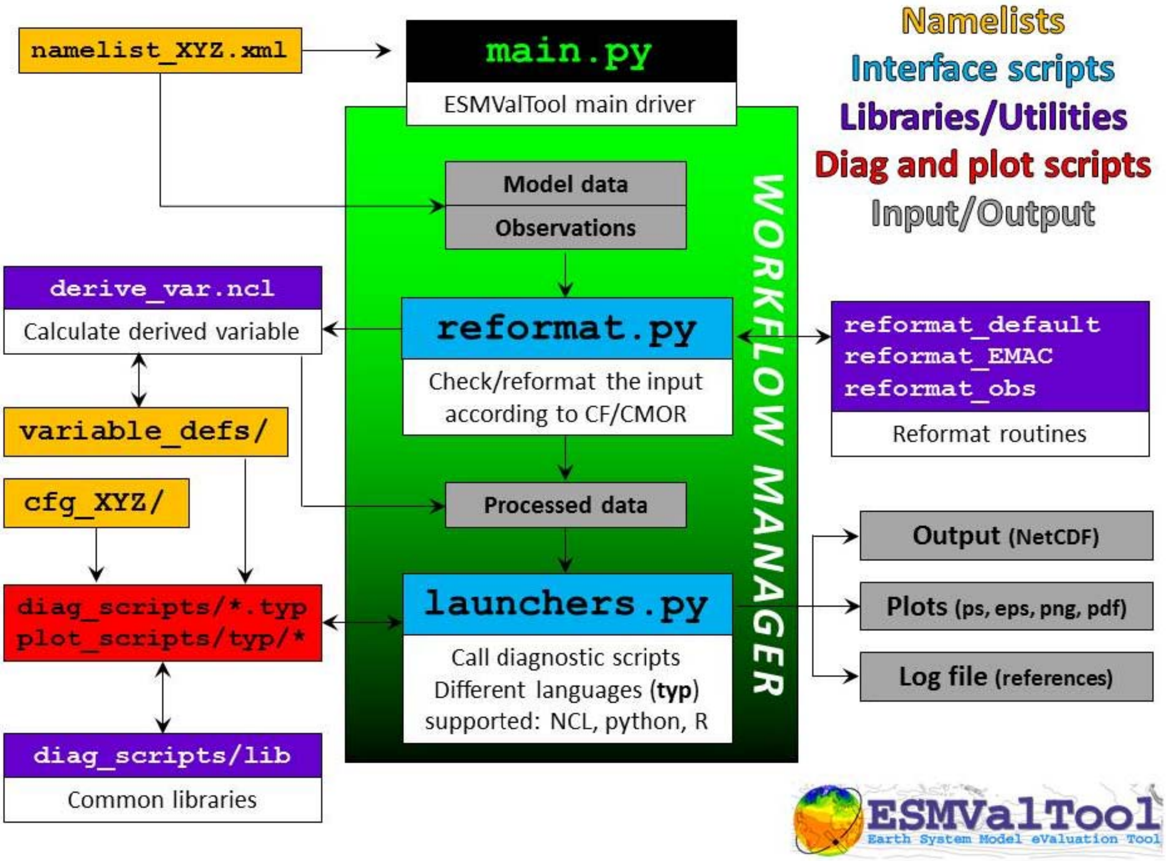

Figure 1: Schematic overview of the ESMValTool structure (from: Eyring et al., 2016a).

The ESMValTool consists of a workflow manager and a number of diagnostic and graphical output scripts as schematically shown in Figure 1. It builds on a previously published diagnostic tool for chemistry-climate model evaluation (CCMVal-Diag Tool, Gettelman et al. (2012)), but is different in its focus. In particular, it extends to ESMs by including diagnostics and performance metrics relevant for the coupled Earth system, and also focuses on benchmarking models with a standard set of diagnostics rather than being mostly flexible as the CCMVal-Diag tool. The workflow manager is written in Python, while a multi-language support is provided in the diagnostic and the graphic routines. The current version supports Python, NCL and R, but it can be extended to other open-source languages. The ESMValTool is executed by invoking the main.py script, which takes a namelist as a single input argument. The namelists are text files written using the XML (eXtensible Markup Language) syntax and define the data to be read (models and observations), the variables to be analysed and the diagnostics to be applied.

Within the workflow, the input data are checked for compliance with the CF and Climate Model Output Rewriter (CMOR, http://pcmdi.github.io/cmor-site/tables.html) standards required by the tool via a set of dedicated reformatting routines, which are also able to fix the most common errors in the input data (e.g., wrong coordinates, undefined or missing values, non-compliant units, etc.). It is additionally possible to define new variables using variable-specific scripts, for example in order to calculate the total column ozone from a 3D ozone field, temperature and surface pressure. The diagnostic and graphic routines are written in a modular and flexible way so that they can be customized by the user via diagnostic-specific settings in the configuration file (cfg-file) and variable-specific settings (in the directory variable_def_dir) without changing the source code. These routines are complemented by a set of libraries, providing general-purpose code for the most common operations (statistical analyses, regridding tools, graphic styles, etc.). The output of the tool can be both netCDF and graphics files in various formats. In addition, a log file is written containing all the information of a specific call of the main script: creation date of running the script, version number, analysed data (models and observations), applied diagnostics and variables, and corresponding references. This helps to increase the traceability and reproducibility of the results.

Besides several small and mostly technical updates and improvements the advanced version of the ESMValTool (v1.0.2) presented here includes the possibility to use variables from the ESA CCI datasets aerosol, cloud, sea ice, soil moisture, sea surface temperature, ozone, greenhouse gases, and land cover for the evaluation of CMIP models. The variables newly implemented are summarized in Table 1.

Table 1 ESA CCI datasets implemented into the advanced CMUG version of the ESMValTool (v1.0.2).

Dataset	Variable(s)	Resolution	Years	Reference(s)
Aerosol	od550aer; od870aer, od550lt1aer, abs550aer	1°x1°	1997-2011	Popp et al. (2016)
Cloud	clt	0.5°x0.5°	1982-2014	Hollmann et al. (2015)
Greenhouse Gases	xco2	5°x5°	2003-2008	Reuter et al. (2011)
Ozone	toz	1°x1°	1997-2010	Van Roozendael et al. (2015)
Land Cover			2000, 2005, 2010	Defourney et al. (2015)
Sea Ice	sic	25 km x 25 km	1992-2008	Sandven et al. (2015)
Sea Surface Temperature	ts	0.05°x0.05°	1991-2013	Merchant et al. (2014a,b)
Soil Moisture	sm	0.25°x0.25°	1978-2010	Liu et al. (2011, 2012), Wagner et al. (2012)

4 ESA CCI data

The following sections give an overview on the ESA CCI data implemented in the ESMValTool v1.0.2 that are now available for comparison with ESM results.

4.1 Aerosol

The ESA Aerosol_cci team produces several aerosol long-term datasets to cover GCOS required aerosol variables such as aerosol optical depth AOD (from radiometers Along-Track Scanning Radiometer ATSR, MEdium Resolution Imaging Spectrometer MERIS, POLarization and Directionality of the Earth's Reflectances POLDER), or stratospheric vertical extinction profile (stellar occultation Global Ozone Monitoring by Occultation of Stars GOMOS). As additional response to the AEROCOM modeling community needs, also information on aerosol composition such as fine-mode AOD (from radiometers) or dust AOD (from Infrared Atmospheric Sounding Interferometer IASI) and absorption AOD (ATSR, POLDER) are derived or estimated by fitting various assumed aerosol mixtures. For various instruments several algorithms are used in parallel since algorithm details lead to varying balances between maximizing coverage and optimizing quality. This can also be used as one means to estimate retrieval uncertainties (condensed into an uncertainty-weighted ensemble dataset). In the case of the ATSR radiometers, three algorithms (ADV of FMI, ORAC of Oxford University and RAL, SU of Swansea University) do perform very similarly, but with regional differences in coverage and accuracy as well as information content. The ESA CCI aerosol product further analysed in this paper consists of a long-term climate record based on data from two similar sensors: the Along-Track Scanning Radiometer (ATSR) on the European Remote Sensing Satellite 2 (ERS2-ATSR2), covering the time period 1995-2003, and the Advanced Along Track Scanning Radiometer on ESA's Environmental Satellite (ENVISAT-AATSR), providing data from 2002 through 2012. Through validation by independent experts with ground-based observations, the Aerosol_cci team identified the data produced by the University of Swansea (SU) algorithm as overall slightly better as described in de Leeuw et al. (2015) and Popp et al. (2016). Into the ESMValTool v1.0.2, version 4.21 of level 3 monthly mean data consider only years with complete data coverage have been implemented so far. Incomplete years from either platform (1995, 1996 and 2003) are not taken into account restricting analyses to the time period 1997-2002. The agreement of the data from the two platforms during the overlapping period 2002-2003 was found to be very good making it easy to combine the two datasets into a single time series. Variables implemented into the ESMValTool so far include total aerosol optical depth (od550aer), fine mode (od550lt1aer), absorption optical depth at 550 nm (abs550aer) and optical depth at 870 nm (od870aer).

4.2 Cloud

The cloud CCI provides data for the following cloud properties: cloud detection (mask/fraction), cloud phase, cloud optical thickness, cloud effective radius, cloud liquid/ice water content and cloud albedo. The retrieved cloud properties are summarized in two global dataset families, the datasets cover the period 1982 through 2014. Pixel based retrievals (L2 products) are further processed to:

- L3U products: sampled into daily global fields (0.05° resolution, 0.02° for MODIS era over Europe) and
- L3C products: aggregated into monthly averages and histograms on a global grid with 0.5° resolution). The monthly averages of LWP and IWP are calculated as in-cloud and all-sky (a.k.a. grid-mean) values.

The cloud CCI data implemented into the ESMValTool v1.0.2 so far consist of the L3C products (AVHRR_NOAA-7-fv2.0) and include the total cloud fraction and the total cloud fraction standard error. The temporal resolution of this product is 1 month.

4.3 Ozone

The ESA CCI total column ozone (toz) dataset consists of combined and harmonized level 3 data covering the time period between 1997 and 2008. Data from the three platforms/instruments, the Global Ozone Monitoring Experiment (GOME) onboard the European Research Satellite 2 (ERS-2/GOME) (1996-2003), ENVISAT/SCIAMACHY (2003-2007), and GOME-2 onboard the Metop satellites (METEOP/GOME-2) (2007-2011) are provided as a merged gridded dataset by the ESA CCI teams. As alternative reference dataset for total ozone columns, we use data from the combined NIWA dataset (Bodeker et al., 2005) (1980-2011).

The ESA CCI limb gridded profile data consist of merged level 3 monthly and zonally averaged data covering the time period 2007-2008 from the six different instruments GOME, the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), SCIAMACHY, the Optical Spectrograph and InfraRed Imaging System (OSIRIS), the Sub-Millimetre Radiometer (SMR), and the Atmospheric Chemistry Experiment (ACE) satellite instruments: the ACE Fourier Transform Spectrometer (ACE-FTS).

4.4 Sea ice

The ESA CCI sea ice dataset provides observational data for sea ice concentration (sic) and sea ice thickness (sit) that are based on satellite retrievals. The sic dataset is (Lavergne and Rinne, 2014):

- daily gridded sic fields based on passive microwave radiometer measurements;
- global maps (both northern hemisphere and southern hemisphere) with 25 km grid spacing;
- both a SSM/I and a AMSR-E dataset, processed and delivered separately;
- daily maps of total standard error (uncertainty) and quality control flags;
- built upon the algorithms and processing software originally developed at the EUMETSAT OSI SAF for their sic dataset (RD-11).

The ESA CCI sea ice datasets implemented into the ESMValTool v1.0.2 consist of the SSM/I data covering the time period 1992 to 2008 and the AMSR-E data covering the time period 2003-2010. Both datasets provide monthly mean sic and sic standard error for the northern and southern hemisphere on equal area grids at a resolution of 25 km. So far, only the variables sea ice concentration and sic standard error from the CCI sea ice dataset have been implemented into the ESMValTool v1.0.2.

4.5 Soil moisture

The ESA CCI soil moisture product is the first ever multi-decadal satellite based soil moisture product and is available for the time period 1978-2010 on a daily basis and at a spatial resolution of 0.25°x0.25°. It has been generated by merging active and passive microwave based soil moisture products from multiple satellite missions (Liu et al., 2011, 2012).

Dorigo et al. (2014) provide a comprehensive validation of the ESA CCI soil moisture using 932 in situ observation sites from 29 different observing networks (Dorigo et al., 2011, 2013). Despite the large difficulties in validating coarse resolution satellite soil moisture products with in situ point like observations (Crow et al., 2012), they conclude that the ESA CCI soil moisture product has an average unbiased RMSE of 0.05 $m^3\ m^{-3}$. Furthermore, it was shown that trends in the CCI observations largely agree with those obtained from various reanalysis products, precipitation, and vegetation vigor (Albergel et al., 2012; Dorigo et al., 2012).

The ESA CCI soil moisture dataset provides a multitude of quality flags and only soil moisture estimates considered reliable are used to create the data product. Snow covered areas and frozen ground are typically masked as well as dense or heterogeneously vegetated areas with high optical depth that are not expected to provide reliable soil moisture estimates (Loew, 2008; Parinussa et al., 2011).

4.6 Land cover

The ESA CCI land cover product (v1.6.1) provides high resolution (300 m) global land cover information for three different years (2000, 2005, 2010). This basic land cover product is complemented by information on land cover conditions at 1 km resolution which comprise climatological information of vegetation state, snow and fire occurrences derived from SPOT vegetation data for the period 1998-2012. In addition a high resolution information layer for surface water bodies at 150 m resolution derived from microwave observations is provided (Bontemps et al, 2012). Information on the accuracy of the ESA CCI land cover product in comparison to other existing global land cover datasets is provided by Tsendbazar et al. (2015).

4.7 Sea surface temperature

The ESA CCI sea surface temperature (SST) dataset (Merchant et al., 2014a,b) provides multi-decadal products of SST derived from infrared brightness temperatures measured from satellites. SST products (Rayner et al., 2015) are generated at full sensor resolution (1 to > 4 km) and are averaged on a regular latitude-longitude grid (0.05°). A gap-filled ('L4 SST analysis') product is currently used with the ESMValTool diagnostics. The L4 SST analysis is a daily optimal interpolation of satellite data with a grid resolution of 0.05°. The interpolation system is the Operational Sea surface Temperature and sea-Ice Analysis (OSTIA) with improved covariance parameterization (Roberts-Jones et al., 2016). The L4 SST analysis has relatively good feature resolution, which is nonetheless lower than the grid resolution, and varies with the density of satellite coverage (Reynolds et al., 2013). Unlike the operational OSTIA products (Donlon et al., 2012) and the older OSTIA-based observational re-processing (Roberts-Jones et al., 2012), no in situ data are used in this CCI product. The product represents the daily value of SST at a nominal depth of 20 cm, representative of the SST measured by drifting buoys and bucket observations. This is possible because the lower-level SST CCI products contain both the skin (radiometric) temperature of the ocean surface at the time of satellite observation estimated based on radiative transfer physics (e.g., Embury et al., 2012a), and a turbulence-model-based adjustment to the 20 cm depth SST at a standardized time of day. The adjusted SST estimate is used as input to the L4 SST analysis. This means that the L4 SST analysis can be treated as independent of in situ data, and useful as a comparison point for the many SST products that are tuned to and/or incorporate in situ data. The standardization of the adjustment with respect to time of day is intended to reduce aliasing of the diurnal cycle into false long-term trends, as satellite overpass times vary (Embury et al., 2012b). All SSTs are provided with estimates of total uncertainty, and L4 SST analysis product includes an operationally produced estimate of sea ice concentration (Good and Rayner, 2014).

Merchant et al. (2014a,b) provide an assessment of the accuracy of this product by comparing more than 2.4 million buoys from different observational networks. A global median difference against drifting buoys of +0.05 K is observed, with a standard deviation (including the ~0.2 K uncertainty in the drifting buoy measurements) of 0.28 K. The comparison with Argo measurements at ~ 5 m depth (only from the latter part of the record) gives +0.04 K and 0.26 K respectively. Systematic regional errors on space scales of ~1000 km range from -0.5 K to +0.5 K, with positive bias of +0.09 K across equatorial regions overall (relative to measurements of the global tropical moored buoy array). Regions persistently affected by mineral atmospheric aerosol, particularly Saharan dust, appear negatively biased.

4.8 Greenhouse gases

The ESA CCI GHG Level 3 (i.e. gridded) data products implemented into the ESMValTool have been specifically created for comparisons with climate model output and are available from obs4MIPs via the Earth System Grid Federation (ESGF) alongside CMIP output in the same format as the CMIP model output. A detailed description of the underlying Level 2 data is available in Buchwitz et al. (2015) and Buchwitz et al. (in preparation).

4.8.1 Column-average dry-air mole fraction carbon dioxide (XCO2)

The ESA CCI GHG XCO2 product is retrieved from measurements of the two satellite instruments SCIAMACHY/ENVISAT (Bovensmann et al., 1999; Burrows et al., 1995) and TANSO-FTS/GOSAT (Kuze et al., 2009). XCO2 is a dimensionless quantity (unit: mol/mol) defined as the vertical column of CO_2 divided by the vertical column of dry air (i.e., all air molecules except water vapor). Details can be found in Buchwitz et al. (2005).

XCO2 is retrieved from radiance spectra in the near-infrared/short-wave infrared (NIR/SWIR) spectral range using (mostly) Optimal Estimation (Rodgers, 2000) retrieval algorithms. For details we refer to the Algorithm Theoretical Baseline Documents (ATBDs) available from the GHG-CCI website for each individual Level 2 data product. Note that a resolution of 5°x5° has been selected (instead of, e.g., 1°x1°) to ensure better noise suppression (note that the underlying individual satellite retrievals are sparse and thus noisy due to very strict quality filtering).

The Level 2 input data used to create the obs4MIPS Level 3 product are all part of the GHG-CCI CRDP3 dataset and are described in Buchwitz et al. (in preparation). From the individual sensor/algorithm Level 2 (L2) XCO2 input data the Level 3 (L3) obs4MIPs product has been generated as follows: to correct for the use of different CO_2 a priori assumptions in the independently retrieved products, all products have been brought to a common a priori using the Simple Empirical CO_2 Model (SECM) described in Reuter et al. (2012). After this a gridded L3 product is generated from each L2 product by averaging all soundings onto a 5°x5° monthly grid. Only those grid cells are further considered having a standard error smaller than 2 ppm. The grid cell uncertainty is computed from the reported L2 uncertainties and a term accounting for potential regional / temporal biases. To avoid potential discontinuities in the obs4MIPs time series, each L3 product has been offset corrected in the overlap period using the mean of the products as reference (conserving the mean value). The offset corrections are typically small and range between -0.4 ppm and +0.6 ppm. The obs4MIPs XCO2 value in a given grid cell is computed as the mean of the individual L3 values and the corresponding total uncertainty is the root-mean-square of the individual L3 uncertainties. Finally a filtering procedure has been applied to remove "unreliable" grid cells considering the overall noise error (1.6 ppm) and total uncertainty (1.8 ppm) of each cell.

An XCO2 uncertainty value (XCO2_stderr) is contained in the obs4MIPs file for each grid cell with valid data (XCO2_nobs > 0). Note that prior to mid-2009 the obs4MIPs product is limited to observations over land due to the solely availability of SCIAMACHY data (GOSAT ocean sun-glint mode observations are only available after mid 2009).

4.8.2 Considerations for model-observation comparisons

A solid comparison between models and the ESA CCI XCO2 product requires sampling the model output at the exact times and location of the satellite observations as well as taking into account the altitude dependent sensitivity of the satellite retrieval. The satellite retrieval is limited to clear sky observations around local noon which also needs to be considered when comparing satellite data with model output. The altitude sensitivity can be considered by applying the satellite averaging kernel to the models' vertical profiles. In this case, all available Level 2 (i.e., individual observations) data products should be used. Note that the XCO2 Level 2 product used for the described obs4MIPs product are the ones from GHG-CCI and that the other products (from NIES and NASA) have been generated using different algorithms.

Due to the gridding / averaging process applied to generate obs4MIPs product, detailed time/location information is not available in the obs4MIPs data product and also averaging kernel is not (yet) part of this product. Typically, however, the satellite XCO2 averaging kernel is close to unity. This is especially the case in the lower troposphere, where the CO_2 variability is typically largest. Therefore applying the averaging kernel typically changes the XCO2 values by less than 1 ppmv (Dils et al., 2014). However, how large this "correction" is depends also on the difference between the respective a priori GHG profile (i.e., CO_2) used for satellite retrieval and the respective model GHG profile. The larger this difference, the larger the averaging kernel correction. If the model profiles are "reasonable" other error sources are likely more relevant for using the obs4MIPs product such as the representativity error.

Another source of uncertainty is the spatial and temporal representativeness of the measurements. A representativity error originates from the fact that the "true" respective GHG field is variable within a given month in a given grid cell but the obs4MIPs values are derived from averaging sparse satellite observations, i.e., are not representative for the "true" monthly mean value of a given grid cell. Note that the agreement with ground-based observations (not considering averaging kernels) for CO_2 is typically within 0.29 +/- 1.2 ppm (1-sigma). These differences therefore include to some extent the representativity error as well as other error sources (e.g., the uncertainty of the reference observations, which is 0.4 ppm (1-sigma) for CO_2. It is recommended to use the reported overall uncertainty range of 0.29 +/- 1.2 ppm (1-sigma) and/or the reported uncertainties for each grid cell as given in the obs4MIPs product file. It can therefore be expected that model minus satellite differences of larger than approximately 2-3 ppmv point to significant differences between the modeled and observed GHG values (differences are significant at the 5% significance level if outside of the 2-sigma (95%) uncertainty range of [-2.1 ppm, 2.7 ppm]).

Note, however, that the obs4MIPs product is new and that not all possible assessments have been carried out yet. The product has been generated by merging an ensemble of underlying Level 2 products. No obvious issues have been identified so far but it cannot be excluded that there are potential issues (depending on data usage / application) introduced by merging the different datasets. For example, the number of observations drops at the beginning of 2012 due to the loss of data from SCIAMACHY/ENVISAT. Another important point to keep in mind when comparing climatological XCO2 values over the ocean to model output is that data are only available for the second half of the ESA CCI record. Averaging over the whole ESA CCI period will therefore result in an overestimation of XCO2 over the ocean. Note also that the satellite observations are obtained around local noon (around 10:00 a.m. local time for SCIAMACHY/ENVISAT and around 1:00 p.m. local time for TANSO-FTS/GOSAT).

5 Examples of using ESA CCI datasets with the ESMValTool

The implementation of the ESA CCI datasets aerosol, cloud, sea ice, soil moisture, sea surface temperature, ozone greenhouse gases, and land cover were the first of the ESA CCI datasets to be implemented within the ESMValTool. Work in the remaining time of CMUG will focus on further scientific interpretation of the results above and on the integration of additional diagnostics and updates of the ESA CCI datasets as outlined in the proposal.

Figure 2 and

Figure 3: caption next page.

Figure 4 are shown exemplary to summarize the main results from Lauer et al. (in preparation). For a more detailed scientific interpretation of the results for the individual CCIs, we refer to Lauer et al. (in preparation).

In order to demonstrate the evaluation of ESMs with the ESA CCI data implemented into the ESMValTool, we use the output from 50 models from the CMIP5 (Taylor et al. 2012). The model data were obtained from the World Climate Research Programme's (WRCP) CMIP5 data archive made available through the Earth System Grid Federation (ESGF). Here, we analyze results of CMIP5 models from the "historical" model runs - twentieth-century simulations for 1850-2005 conducted with the best record of natural and anthropogenic climate forcing. In order to extend the model runs beyond the year 2005, we use results from simulations forced under the Representative Concentration Pathways 4.5 for the years 2006-2013 (Clarke et al., 2007; Smith and Wigley, 2006; Wise et al., 2009). An assessment of the agreement of simulated climatological mean state and seasonal cycle for key variables such as ECVs (GCOS, 2010; Bojinski et al., 2014) with observations is commonly seen as a reasonable starting point for the evaluation of ESMs (e.g., Gleckler et al., 2008; Flato et al., 2013; Hagemann et al., 2013; Eyring et al., 2016b). Following Gleckler et al. (2008) and similar to Fig. 9.7 of Flato et al. (2013), we calculate the relative space-time RMSD of the climatological seasonal cycle from CMIP5 simulations compared with observations for selected variables.

If there are multiple ensemble members available for any given model, we only consider the first ensemble member "r1i1p1" in our analysis. The only exception to this is the analysis of ozone, for which we picked the ensemble members with interactive ozone chemistry as all "r1i1p1" ensemble members use pre-scribed ozone climatologies. All variables except for sea ice coverage (sic) are averaged over the whole globe. Sea ice extent is averaged over the latitude band 60°N to 90°N (Arctic, "NHpolar") and 60°S to 90°S (Antarctic, "SHpolar"). The model results are compared to a reference dataset and where data are available to an alternative observationally based dataset. For the models, results are averaged over the years with observational data available.

Figure 2 provides a synoptic overview of the relative model quality compared with the multi-model mean error. The figure provides a relative metric assessing whether a specific model performs better or worse than the other models. The data have been normalized with the centered median and regridded to the grid of the reference data using a bilinear interpolation method. As such it can be seen as a starting point to investigate the reasons for differences between model and observations. Performance varies across the models and variables, with some models comparing better with observations for one variable and another model performing better for a different variable. The figure includes all variables that have been shown in Flato et al. (2013) and adds variables for ESA CCI data, i.e., soil moisture (sm), sea ice concentration (sic), total ozone columns (toz), aerosol optical depth (AOD) at 550 nm from particles smaller than 1 µm (od550lt1aer), absorption AOD at 550 nm (abs550aer), AOD at 870 nm (od870aer), AOD at 550 nm (od550aer), and sea surface temperature (ts). Except for global average temperatures at 200 hPa (ta_Glob-200), aerosol optical depth of fine particles at 550 nm (od550lt1aer_Glob), and sea ice (sic_NHpolar, sic_SHpolar), the multi-model mean outperforms any individual model.

For calculating the performance metrics for the four aerosol variables od550aer, od550lt1aer, abs550aer, and od870aer shown in Figure 2 the ESA CCI dataset (lower triangles) and the MODIS dataset (upper triangles) are used. Shown are only CMIP5 models with interactive aerosols, models using pre-scribed aerosol climatologies have not been taken into account. Even though most models agree on the basic properties of the AOD distribution (od550aer), the relative spread among the models for absorption AOD (abs550aer) and AOD of fine particles (d < 1 µm, od550lt1aer) is large. It should be noted that the observational uncertainties for these quantities are also larger than for AOD at 550 nm.

For total cloud cover (clt), the choice of the reference dataset can make some difference for the calculated performance of the individual models. A number of models such as, for instance, the GFDL-CM3 and some of the HadGEM models compare worse against the ESA CCI dataset than against the data from MODIS. The ESA CCI cloud data show slightly higher values (10-15%) for total cloud cover in the subtropical stratocumulus regions off the west coasts of North and South America as well as off the coast of Australia. In contrast, cloud amounts in the ESA CCI data are smaller over the tropical Pacific with frequent deep convection (-10 to -20%). These are also regions in which the models typically struggle to reproduce the observations. The average model bias compared with the ESA CCI data is therefore larger than compared with the MODIS data. An exact quantitative assessment, however, requires application of a satellite simulator in the models to take into account satellite overpass times and lower cut-off thresholds, which is beyond the scope of this study. The comparison of total cloud cover done here should therefore only be seen as a starting point for further evaluation of the ESMs.

The performance metric of total column ozone with respect to the ESA CCI (lower triangle) and NIWA (upper triangle) data is shown in Figure 2 only for models with interactive chemistry (CNRM-CM5, GFDL-CM3, GISS-E2-H, GISS-E2-R). The performance of the individual CMIP5 models for total column ozone is quite similar with respect to both observational datasets. This is not surprising as both reference datasets are based on the same satellite observations from GOME-2 and SCIAMACHY (Bodeker et al., 2005; Loyola et al., 2009).Typical model biases include, for instance, an overestimation of total ozone in high northern latitudes (> 60°N) throughout the year and an underestimation of ozone in Antarctica during summer (November to January).

For sea ice extent (sic), the ESA CCI SI SSM/I and the National Snow and Ice Data Center NSIDC-NT (Walsh et al., 2015) observations are used for comparison with the CMIP5 models. Figure 2 shows that the choice of the reference dataset does not impact the results for the model performance in reproducing the observed sea ice concentration significantly. This suggests that the two sea ice datasets are in rather good agreement. Most CMIP5 models show a better performance in reproducing observed sea ice concentration in the Antarctic (SH) than in the Arctic (NH, Figure 2).

The inter-model spread for soil moisture (sm) is large and most models tend to systematically over- or underestimate soil moisture throughout the globe compared with the ESA CCI data. It should be noted, however, that a quantitative comparison is difficult as the observations do not represent exactly the same quantity as simulated by the models. Furthermore, the layer thickness considered is not consistent among the models and the satellite observations (see also Lauer et al., in preparation). Qualitatively, many models such as the FGOALS, GFDL, HadGEM, and MIROC models overestimate the soil moisture particularly in higher latitudes in Asia, as well as Alaska and the northern part of Canada.

Typical biases in the geographical distribution of the simulated SST include a warm bias in the subtropical stratocumulus regions as well as a cold bias in the equatorial Pacific. Individual models performing worse than the multi-model mean (Figure 2) include, for instance, the CSIRO, the

FGOALS, and the MRI models. The reasons for this are rather different, for example the CSIRO model shows a warm bias in the North Pacific whereas the FGOALS model shows a cold bias in the North Atlantic.

A widely used way to summarize overall comparisons of annual mean properties with observations for individual models are Taylor diagrams (Taylor, 2001). The Taylor diagrams shown in

Figure 3: caption next page.

Figure 4 give the standard deviation and linear pattern correlation with observations of the total spatial variability calculated from multi-year annual means. The standard deviations are normalized by the observed values, so the observed climatology is represented in each panel by the filled black dots on the x axis at x = 1. The pattern correlation is given in this polar-stereographic projection by the angular coordinate. The linear distance between the observations and each model is proportional to the root-mean-square error (RMSE) and can be estimated in multiples of the observed standard deviation with the gray circles centered on the observational dots. The multi-model mean values have been calculated over all models with data available (black star). Where available an alternative reference dataset is also shown in Figure 2 (red star). The green circles show estimates of the observational uncertainties (RMSE). The observational uncertainty estimates are provided by the ESA CCI teams and are part of the ESA CCI datasets. The green circles show the multi-year global average uncertainties given as one sigma of the total standard error normalized by the standard deviation of the observations. The RMSE of a given model compared with the observations is therefore smaller than the 1-sigma uncertainty estimate of the observations if the model lies within the green circle.

For total cloud cover, the model show a large spread in pattern correlation between 0.2 and 0.85. Most model are/are not outside of the 1-sigma uncertainty estimate showing that the differences between the models and the observations cannot be solely explained by measurement uncertainties. The integrated aerosol properties aerosol optical depth (AOD) at 550 and 870 nm also show a large inter-model spread. Because of the large observational uncertainties, most models lie within the green circle of the 1-sigma measurement uncertainty making quantitative assessments difficult. This is also supported by the differences between the ESA CCI dataset and the MODIS data for AOD with the MODIS RMSE being close to 1-sigma of the ESA CCI uncertainty estimate. The linear pattern correlation of most models with the ESA CCI data, however, is smaller than that of the ESA CCI data and MODIS (0.8) showing also differences in the geographical distribution of the simulated AOD. Soil moisture can mostly be used for qualitative assessments of the models as the observational uncertainties are larger than the RMSE of many of the individual models. The geographical annual mean patterns of the sea surface temperatures from the models high correlation with the ESA CCI data ranging between 0.94 and 0.98. SST in the subtropical stratocumulus regions as well in the Southern Ocean is overestimated by many models, though. Another typical model bias found in many simulations is an underestimation of the SST in the equatorial Pacific. The correlation coefficients of the modeled total ozone columns with the ESA CCI data are quite high with most models above 0.9 and a ratio of the modeled and the observed spatial standard deviation being close to 1. All models are, however, outside of the 1-sigma uncertainty estimate of the observations. Differences are found, for instance, in the northern high latitudes where the models tend to underestimate the total ozone columns. For the column-averaged CO_2 concentrations, the correlation coefficients of the results from the emission driven simulations with the ESA CCI data are typically quite low and range between 0.2 and 0.6. This is partly caused by the systematical overestimation of XCO2 concentrations by most CMIP5 models and partly by differences in the geographical patterns such as, for example, in northern Europe or Southeast Asia where the models show distinct local maxima that are not clearly visible in the ESA CCI data.

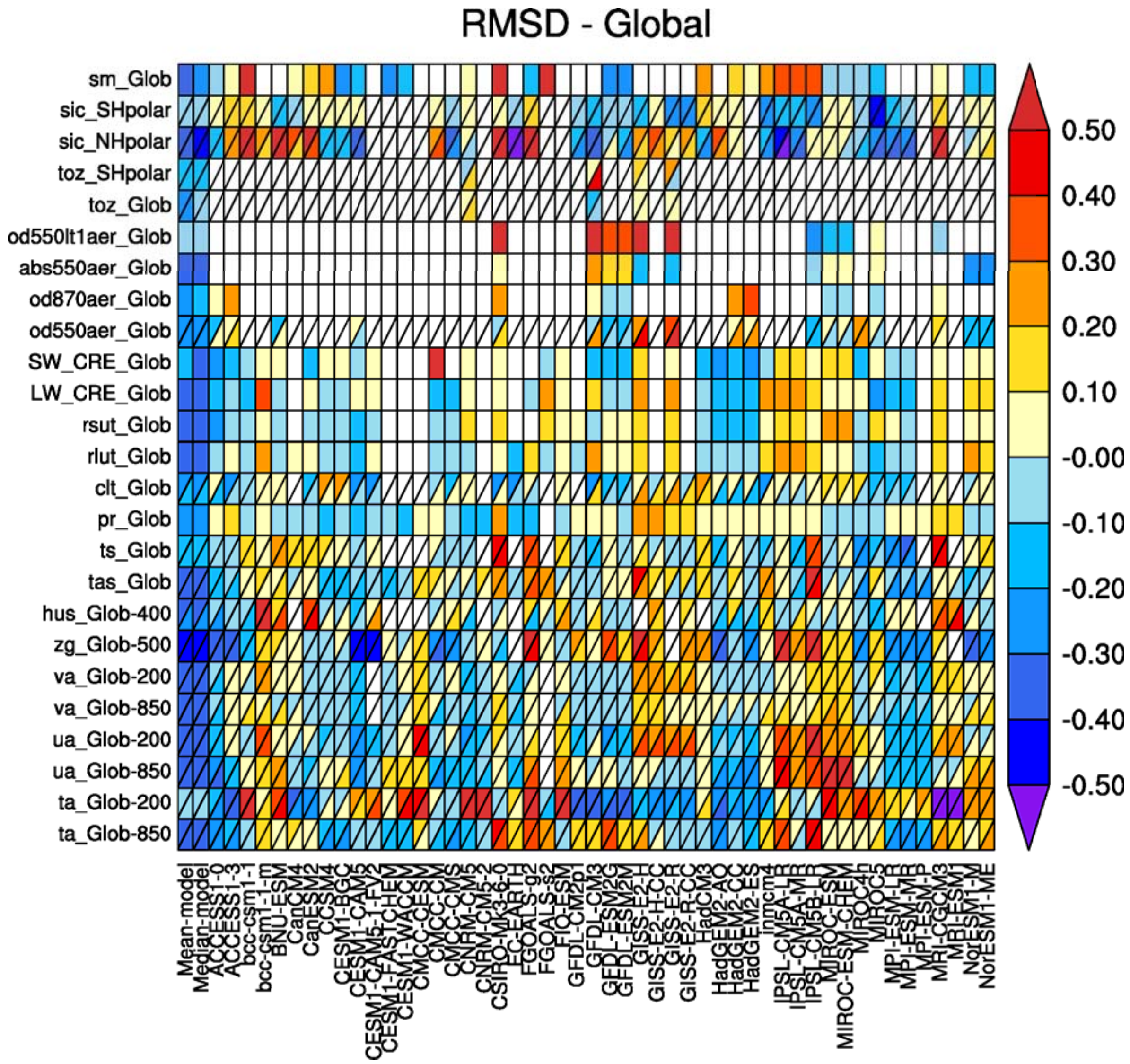

Figure 2: Relative space-time root-mean square deviation (RMSD) calculated from the climatological seasonal cycle of the CMIP5 simulations. The years averaged depend on the years with observational data available and are summarized in Table 1. A relative performance is displayed, with blue shading indicating better and red shading indicating worse performance than the median of all model results. A diagonal split of a grid square shows the relative error with respect to the reference dataset (lower right triangle) and the alternate dataset (upper left triangle). White boxes are used when data are not available for the given model and variable or no alternate dataset has been used. ESA CCI data have been used as reference data (lower right triangle) for the following variables: soil moisture (sm), sea ice concentration (sic), total ozone columns (toz), aerosol optical depth (AOD) at 550 nm from particles smaller than 1 µm (od550lt1aer), absorption AOD at 550 nm (abs550aer), AOD at 870 nm (od870aer), AOD at 550 nm (od550aer), and sea surface temperature (ts) (from: Lauer et al., in preparation).

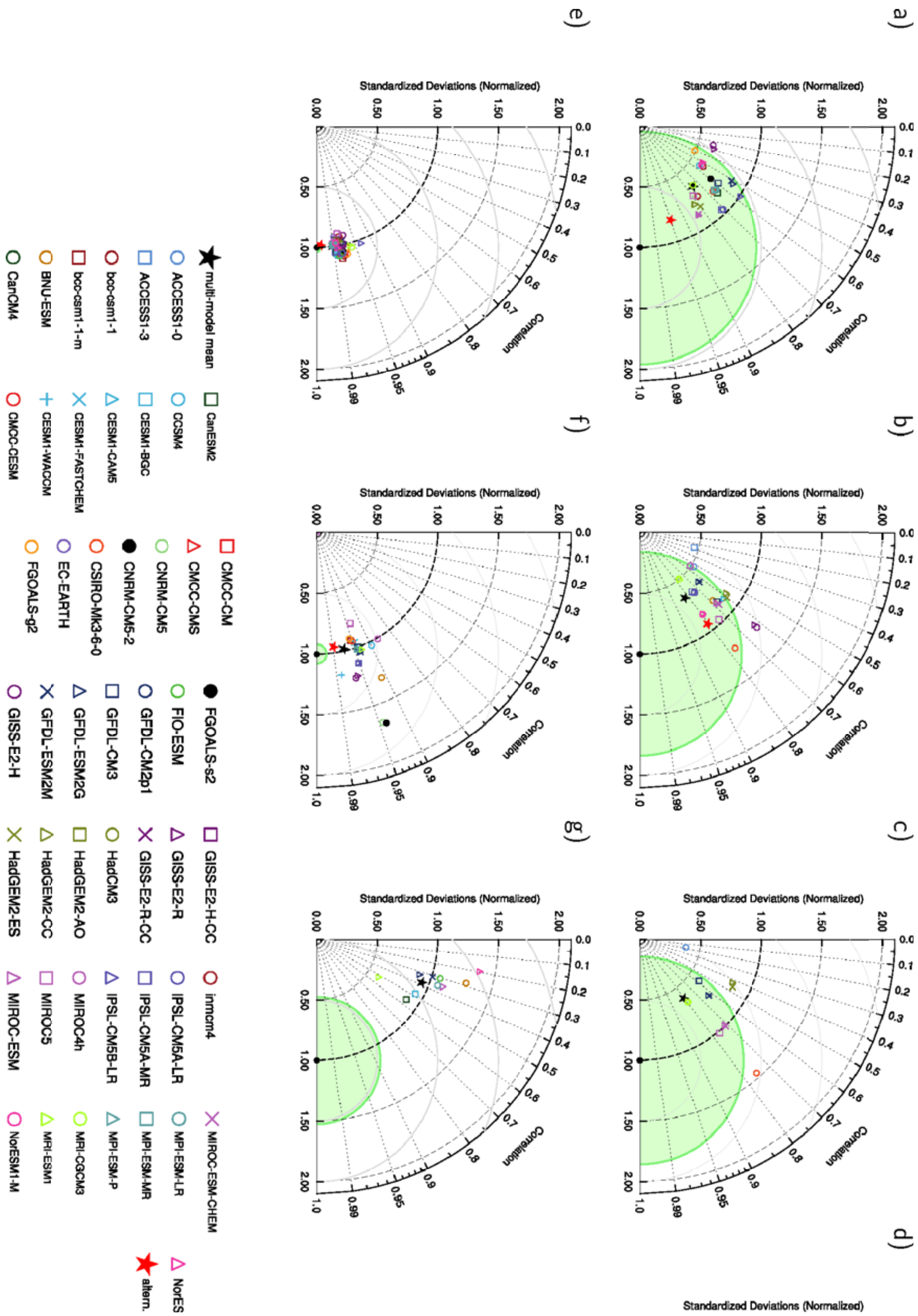

Figure 3: caption next page.

Figure 4: Taylor diagrams showing the multi-year annual average performance of the CMIP5 models in comparison with ESA CCI data for a) total cloud amount, b) aerosol optical depth at 550 nm, and c) aerosol optical depth at 870 nm, d) soil moisture, e) sea surface temperature, f) total column ozone, and g) column averaged CO_2 concentration. Panels a) to e) show CMIP5 historical simulations (extended with RCP4.5), panel f) historical simulations with interactive ozone chemistry, and panel g) emission driven historical simulations (extended with RCP8.5). The multi-model mean values have been calculated over all models with data available (black stars). Where available alternative reference datasets are also shown (red stars). The green circles show estimates of the observational uncertainties (RMSE) (from: Lauer et al., in preparation).

Code Availability

The enhanced version of the ESMValTool with a subset of the ESA CCI Phase 2 data described in this report included is released under the Apache License, VERSION 2.0. This enhanced version and an updated user's guide will be available from the ESMValTool webpage at http://www.esmvaltool.org/ and from github (https://github.com/ESMValTool-Core/ESMValTool) in fall 2016. Users who apply the Software resulting in presentations or papers are kindly asked to cite the ESMValTool documentation paper (Eyring et al., 2016a) alongside with the Software doi (doi:10.17874/ac8548f0315) and version number. The wider climate community is encouraged to contribute to this effort and to join the ESMValTool development team for contribution of additional more in-depth diagnostics for ESM evaluation.

References

Albergel, C., P. de Rosnay, C. Gruhier, J. Muñoz-Sabater, S. Hasenauer, L. Isaksen, Y. Kerr, and W. Wagner, Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sensing Envir., 118, 215-226, 2012.

Bodeker, G. E., H. Shiona, and H. Eskes, Indicators of Antarctic Ozone Depletion, Atmos. Chem. Phys., 5(10), 2603-15, 2005.

Bojinski, S., M. Verstraete, T. C. Peterson, C. Richter, A. Simmons, and M. Zemp, The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy, Bull. Amer. Meteor. Soc., 95, 1431-1443, doi: 10.1175/BAMS-D-13-00047.1, 2014.

Bontemps, S., M. Herold, L. Kooistra, A. van Groenestijn, A. Hartley, O. Arino, I. Moreau, and P. Defourny: Revisiting land cover observation to address the needs of the climate modeling community, Biogeosciences, 9, 2145-2157, doi: 10.5194/bg-9-2145-2012, 2012.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission Objectives and Measurement Modes, J Atmos Sci, 56, 127-150, 1999.

Buchwitz, M., de Beek, R., Burrows, J. P., Bovensmann, H., Warneke, T., Notholt, J., Meirink, J. F., Goede, A. P. H., Bergamaschi, P., Korner, S., Heimann, M., and Schulz, A.: Atmospheric methane and carbon dioxide from SCIAMACHY satellite data: initial comparison with chemistry and transport models, Atmos Chem Phys, 5, 941-962, 2005.

Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Aben, I., Alexe, M., Armante, R., Bergamaschi, P., Bovensmann, H., Brunner, D., Buchmann, B., Burrows, J. P., Butz, A., Chevallier, F., Chédin, A., Crevoisier, C. D., Gonzi, S., De Mazière, M., De Wachter, E., Detmers, R., Dils, B., Frankenberg, C., Hahne, P., Hasekamp, O. P., Hewson, W., Heymann, J., Houweling, S., Hilker, M., Kaminski, T., Kuhlmann, G., Laeng, A., v. Leeuwen, T. T., Lichtenberg, G., Marshall, J., Noël, S., Notholt, J., Palmer, P., Parker, R., Scholze, M., Stiller, G. P., Warneke, T., and Zehner, C.: The greenhouse gas project of ESA's climate change initiative (GHG-CCI): overview, achievements and future plans, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-7/W3, 165-172, 2015.

Buchwitz et al., Remote Sensing of Environment (in preparation).

Burrows, J. P., Holzle, E., Goede, A. P. H., Visser, H., and Fricke, W.: Sciamachy - Scanning Imaging Absorption Spectrometer for Atmospheric Chartography, Acta Astronaut, 35, 445-451, 1995.

Clarke, L. E., J. A. Edmonds, H. D. Jacoby, H. Pitcher, J. M. Reilly, and R. Richels, Scenarios of greenhouse gas emissions and atmospheric concentrations, Sub-report 2.1a of Synthesis and Assessment Product 2.1, Climate Change Science Program and the Subcommittee on Global Change Research, Washington DC, 2007.

Crow, W. T., A. A. Berg, M. H. Cosh, A. Loew, B. P. Mohanty, R. Panciera, P. de Rosnay, D. Ryu, and J. P. Walker, Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products, Rev. Geophys., 50, RG2002, doi: 10.1029/2011RG000372, 2012.

de Leeuw, G., T. Holzer-Popp, S. Bevan, W. Davies, J. Descloitres, R.G. Grainger, J. Griesfeller, A. Heckel, S. Kinne, L. Klüser, P. Kolmonen, P. Litvinov, D. Martynenko, P.J.R. North, B. Ovigneur, N. Pascal, C. Poulsen, D. Ramon, M. Schulz, R.Siddans, L. Sogacheva, D. Tanré, G.E. Thomas, T.H. Virtanen, W. von Hoyningen Huene, M.Vountas, S. Pinnock, Evaluation of seven European aerosol optical depth retrieval algorithms for climate analysis, Remote Sensing of Environment, 162, 295-315, doi: 10.1016/j.rse.04.023, 2015.

Defourney, P., et al.: ESA Land Cover Climate Change Initiative (ESA LC_cci) data: <Product name and Version number> via Centre for Environmental Data Analysis, <date of citation>, 2015.

Dils, B., Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Parker, R., Guerlet, S., Aben, I., Blumenstock, T., Burrows, J. P., Butz, A., Deutscher, N. M., Frankenberg, C., Hase, F., Hasekamp, O. P., Heymann, J., De Maziere, M., Notholt, J., Sussmann, R., Warneke, T., Griffith, D., Sherlock, V., and Wunch, D.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT CO2 and CH4 retrieval algorithm products with measurements from the TCCON, Atmos Meas Tech, 7, 1723-1744, 2014.

Donlon, C. J., M. Martin, J. Starka, J. Roberts-Jones, E. Fiedler, and W. Wimmer, The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sensing of Environment, 116, 140-158, 2012.

Dorigo, W., A. Gruber, R. De Jeu, W. Wagner, T. Stacke, T., A. Loew, C. Albergel, L. Brocca, D. Chung, R. Parinussa, and R. Kidd, Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sensing of Environment, 162, 380-395, 2014.

Dorigo, W., R. de Jeu, D. Chung, R. Parinussa, Y. Liu, W. Wagner, and D. Fernández-Prieto, Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture, Geophys. Res. Lett., 39, L18405, 2012.

Dorigo, W. A., W. Wagner, R. Hohensinn, S. Hahn, C. Paulik, A. Xaver, A. Gruber, M. Drusch, S. Mecklenburg, P. van Oevelen, A. Robock, and T. Jackson, The international soil moisture network: A data hosting facility for global in situ soil moisture measurements, Hydrology and Earth System Sciences, 15, 1675-1698, doi: 10.5194/hess-15-1675-2011, 2011.

Dorigo, W. A., A. Xaver, M. Vreugdenhil, A. Gruber, A. Hegyiová, A. D. Sanchis-Dufau, D. Zamojski, C. Cordes, W. Wagner, and M. Drusch, Global automated quality control of in situ soil moisture data from the international soil moisture network, Vadose Zone Journal, doi: 10.2136/vzj2012.0097, 2013.

Embury, O., C. J. Merchant, and M. J. Filipiak, A reprocessing for climate of sea surface temperature from the Along-Track Scanning Radiometers: basis in radiative transfer, Remote Sensing of Environment, 116, 32-46, doi: 10.1016/j.rse.2010.10.016, 2012a.

Embury, O., C. J. Merchant, and G. K. Corlett, A reprocessing for climate of sea surface temperature from the Along-Track Scanning Radiometers: initial validation, accounting for skin and diurnal variability, Remote Sensing of Environment, 116, 62-78, doi: 10.1016/j.rse.2011.02.028, 2012b.

Eyring, V., M. Righi, A. Lauer, M. Evaldsson, S. Wenzel, C. Jones, A. Anav, O. Andrews, I. Cionni, E. L. Davin, C. Deser, C. Ehbrecht, P. Friedlingstein, P. Gleckler, K.-D. Gottschaldt, S. Hagemann, M. Juckes, S. Kindermann, J. Krasting, D. Kunert, R. Levine, A. Loew, J. Mäkelä, G. Martin, E. Mason, A. S. Phillips, S. Read, C. Rio, R. Roehrig, D. Senftleben, A. Sterl, L. H. van Ulft, J. Walton, S. Wang, and K. D. Williams: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth System Models in CMIP, Geosci. Model Dev., 9, 1747-1802, doi: 10.5194/gmd-9-1747-2016, 2016a.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, doi: 10.5194/gmd-9-1937-2016, 2016b.

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S. C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason, and M. Rummukainen, Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (Ed.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

GCOS, Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC, August 2010, 2010.

Gettelman, A., Eyring, V., Fischer, C., Shiona, H., Cionni, I., Neish, M., Morgenstern, O., Wood, S. W., and Li, Z.: A community diagnostic tool for chemistry climate model validation, Geosci. Model Dev., 5, 1061-1073, 2012.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux, Performance metrics for climate models, J. Geophys. Res., 113, D06104, 2008.

Good, S, and N. Rayner, Product Specification Document, Project Document SST_CCI-PSD-UKMO-201, http://www.esa-sst-cci.org/PUG/documents, 2014 (accessed 13 July 2016).

Hagemann, S., A. Loew, and A. Andersson, Combined evaluation of MPI-ESM land surface water and energy fluxes, J. Adv. Model. Earth Syst., 5, 259-286, doi: 10.1029/2012MS000173, 2013.

Hollmann, R., et al.: ESA Cloud Climate Change Initiative (ESA Cloud_cci) data: <Product name and Version number> via Centre for Environmental Data Analysis, <date of citation>, 2015.

Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, Appl. Opt., 48, 6716-6733, 2009.

Lauer, A., et al.: Benchmarking CMIP5 models with a subset of ESA CCI Phase 2 data using the ESMValTool, Remote Sensing of Environment (in preparation).

Lavergne, T., and E. Rinne, ESA Sea Ice Climate Change Initiative: Phase 1, D3.4 Product User Guide (PUG), Doc Ref: SICCI-PUG-13-07, Version: 2.0, Date: 29 August 2014.

Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. de Jeu, W. Wagner, A. I. J. M. van Dijk, M. F. McCabe, and J. P. Evans: Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, Hydrology and Earth System Sciences, 15, 425-436, doi: 10.5194/hess-15-425-2011, 2011.

Liu, Y. Y., W. A. Dorigo, R. M. Parinussa, R. A. M. de Jeu, W. Wagner, M. F. McCabe, J. P. Evans, and A. I. J. M. van Dijk: Trend-preserving blending of passive and active microwave soil moisture retrievals, Remote Sensing of Environment, 123, 280-297, doi: 10.1016/j.rse.2012.03.014, 2012.

Loyola, D. G., R. M. Coldewey-Egbers, M. Dameris, H. Garny, A. Stenke, M. Van Roozendael, C. Lerot, D. Balis, and M. Koukouli, Global long-term monitoring of the ozone layer - a prerequisite for predictions, International Journal of Remote Sensing, 30, 15, 4295-4318, doi: 10.1080/01431160902825016, 2009.

Merchant, C. J., O. Embury, J. Roberts-Jones, E. Fiedler, C. E. Bulgin, G. K. Corlett, S. Good, A. McLaren, N. Rayner, S. Morak-Bozzo, and C. Donlon: Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI), Geoscience Data Journal, 1, 179-191, doi: 10.1002/gdj3.20, 2014a.

Merchant, C. J., O. Embury, J. Roberts-Jones, E. K. Fiedler, C. E. Bulgin, G. K. Corlett, S. Good, A. McLaren, N. A. Rayner, and C. Donlon: ESA Sea Surface Temperature Climate Change Initiative (ESA SST CCI): Analysis long term product version 1.0, NERC Earth Observation Data Centre, 24th February 2014, doi: 10.5285/878bef44-d32a-40cd-a02d-49b6286f0ea4, 2014b.

Popp, T., G. de Leeuw, C. Bingen, C. Brühl, V. Capelle, A. Chedin, L. Clarisse, O. Dubovik, R. Grainger, J. Griesfeller, A. Heckel, S. Kinne, L. Klüser, M. Kosmale, P. Kolmonen, L. Lelli, P. Litvinov, L. Mei, P. North, S. Pinnock, A. Povey, C. Robert, M. Schulz, L. Sogacheva, K. Stebel, D. Stein Zweers, G. Thomas, L. Gijsbert Tilstra, S. Vandenbussche, P. Veefkind, M. Vountas, and Y. Xue, Development, Production and Evaluation of Aerosol Climate Data Records from European Satellite Observations (Aerosol_cci), Remote Sensing, 8, 421, doi: 10.3390/rs8050421, 2016.

Rayner, N, S. Good, T. Block, SST CCI Product User Guide, Project Document SST_CCI-PUG-UKMO-201, http://www.esa-sst-cci.org/PUG/documents, 2015 (accessed 13 July 2016).

Reuter, M., H. Bovensmann, M. Buchwitz, J. P. Burrows, B. J. Connor, N.M. Deutscher, D. W. T. Griffith, J. Heymann, G. Keppel-Aleks, J. Messerschmidt, J. Notholt, C. Petri, J. Robinson, O. Schneising, V. Sherlock, V. Velazco, T. Warneke, P. O. Wennberg, and D. Wunch: Retrieval of atmospheric CO2 with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results, J. Geophys. Res., doi: 10.1029/2010JD015047, 2011.

Reuter, M., Buchwitz, M., Schneising, O., Hase, F., Heymann, J., Guerlet, S., Cogan, A. J., Bovensmann, H., and Burrows, J. P.: A simple empirical model estimating atmospheric CO2 background concentrations, Atmos Meas Tech, 5, 1349-1357, 2012.

Reynolds, R. W., D. B. Chelton, J. Roberts-Jones, M. J. Martin, D. Menemenlis, and C. J. Merchant, Objective Determination of Feature Resolution in Two Sea Surface Temperature Analyses, J. Climate, 26, 2514-2533, doi: 10.1175/JCLI-D-12-00787.1, 2013.

Roberts-Jones, J., E. K. Fiedler, and M. J. Martin, Daily, Global, High-Resolution SST and Sea Ice Reanalysis for 1985–2007 Using the OSTIA System, J. Climate, 25, 6215-6232, doi: 10.1175/JCLI-D-11-00648.1, 2012.

Robert-Jones, J., K. Bovis, M. J. Martin, and A. McLaren, Estimating background error covariance parameters and assessing their impact in the OSTIA system, Remote Sensing of Environment, 176, 117-138, doi: 10.1016/j.rse.2015.12.006, 2016.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing, 2000.

Sandven, S., et al.: ESA Sea Ice Climate Change Initiative (ESA Seaice_cci) data: ESA CCI SIC v1.11, via Centre for Environmental Data Analysis, 16 Feb 2016, 2015.

Smith, S. J., and T. M. L. Wigley, MultiGas forcing stabilization with minicam, The Energy Journal Special issue, 3, 373-392, 2006.

Taylor, K. E., Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res., 106 (D7), 7183-7192, 2001.

Taylor, K. E., R. Stouffer, and G. A. Meehl, An overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93, 485-498, 2012.

Tsendbazar, N.-E., S. de Bruin, S. Fritz, and M. Herold, Spatial Accuracy Assessment and Integration of Global Land Cover Datasets, Remote Sens, 7, 15804-15821, 2015.

Van Roozendael, M., et al.: ESA Ozone Climate Change Initiative (ESA Ozone_cci) data: <Product name and Version number> via Centre for Environmental Data Analysis, <date of citation>, <DOI>, 2015.

Wagner, W., W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, M. Ertl: Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume I-7, 2012, XXII ISPRS Congress, 25 August - 01 September 2012, Melbourne, Australia, 2012.

Walsh, J. E., W. L. Chapman, and F. Fetterer, Gridded Monthly Sea Ice Extent and Concentration, 1850 Onward, Version 1, Boulder, Colorado, USA, 2015.

Wise, M., K. Calvin, A. Thomson, L. Clarke, B. Bond-Lamberty, R. Sands, S. J. Smith, A. Janetos, and J. Edmonds, Implications of limiting CO_2 concentrations for land use and energy, Science, 324, 1183-1186, 2009.