

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1



Climate Modelling User Group

Deliverable 3.1 version 2

Technical note on CMUG ECV Quality Assessment Report

Centres providing input: MOHC, MPI-M, ECMWF, MétéoFrance

Version nr.	Date	Status
2.1	1 April 14	Version sent to ESA



METEO FRANCE
Toujours un temps d'avance



Max-Planck-Institut
für Meteorologie

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

**Deliverable 3.1 version 2****Technical note****CMUG ECV Quality Assessment Report****Table of Contents**

1. PURPOSE AND SCOPE OF THE TECHNICAL NOTE	3
2. TERMINOLOGY USED	3
3. METHODOLOGY APPLIED TO ASSESS CLIMATE DATA RECORDS	5
4. ASSESSMENT OF CLIMATE DATA RECORDS FOR CCI ECVS	7
4.1 Sea Surface Temperature	9
4.2 Ocean colour	20
4.3 Sea surface height.....	28
4.4 Clouds.....	39
4.5 Ozone.....	44
4.6 Greenhouse Gases	48
4.7 Aerosols	58
4.8 Land Cover	63
4.9 Fire burnt area	79
5. SUMMARY OF ASSESSMENTS FROM A CLIMATE MODELLING PERSPECTIVE.....	86
6. REFERENCES	89

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

CMUG ECV Quality Assessment Report Version 2

1. Purpose and scope of the Technical note

The purpose of this activity is to assess the quality of the Climate Data Records (CDRs) delivered by the ESA CCI and use them with coupled Earth System Models. To provide added value for climate modelling activities such as initialisation, assimilation, model evaluation and development, trend analysis and monitoring, the CDRs must have 'climate quality' and meet the requirements which have been given in the URDs.

It is important to emphasise this assessment of the CCI CDRs will not be a repeat of the validation performed by the CCI climate research groups but is complimentary to them. CMUG provides an *independent* assessment of the CDRs.

This second version of the report D3.1 (D3.1v2) reports on the assessments of the actual CCI CDRs produced in phase 1 of the CCI project which were available to CMUG on 1 Jan 2014. We recall the first version of this CMUG report (D3.1v1a) (CMUG, 2012) assessed precursor datasets of the CCI ECVs being assessed as a proof of concept.

2. Terminology used

To aid the reader and avoid confusion the definition of the main terms used in this report are given here.

Assessment here is a generic term which refers to a variety of different ways to determine the fidelity of a CDR. The various methods for assessment are given in section 3.

Assimilate here refers to a CDR being used within an atmospheric, ocean or land surface model to adjust the state variables to better fit the observations taking into account the uncertainties of the observations and model first guess.

Climate Data Record (CDR) is a level 2 or 3 dataset for an ECV which has been processed to a standard sufficient for climate monitoring purposes. Level 1 datasets (e.g. top of atmosphere radiances) are referred to as FCDRs (Fundamental Climate Data Record).

CMIP-5 is an exercise to intercompare the current state of art coupled climate models and has provided an ensemble of different experiments for past, present and future climate. This is one of the international Model Intercomparison Projects, CCMVal being another one, for coupled stratospheric chemistry and climate modelling.

Consistency refers to the consistency of related ECVs (e.g. fire and aerosols) in space and time. This is important for relationships between different ECVs and also between different CDRs for the same ECV.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Climate model is a numerical representation of the climate system based on the physical, chemical and biological properties of its components, their interactions and feedback processes, and accounting for some of its known properties.

Earth System Models or ESMs are complex state of the art climate models which represent processes in the Atmosphere, Ocean and Terrestrial domains and the coupling between them.

Ensemble A group of parallel model simulations used for climate projections or predictions. Variation of the results across the ensemble members gives an estimate of uncertainty. Ensembles made with the same model but different initial conditions only characterise the uncertainty associated with internal climate variability, whereas multi-model ensembles including simulations by several models also include the impact of model structural differences.

Essential Climate Variable (ECV) defines a specific variable defining the atmospheric, ocean or land surface state. One ECV can include several different climate data records (e.g. ozone total column and ozone profile). They have been defined by GCOS (2011) for ECVs measured by satellites.

Hindcast is a where a climate or NWP model is run in the past to verify the accuracy of its forecasts with observations.

HyMeX (HYdrological cycle in the Mediterranean EXperiment) aims at a better understanding and quantification of the hydrological cycle and related processes in the Mediterranean, with emphasis on high-impact weather events, inter-annual to decadal variability of the Mediterranean coupled system, and associated trends in the context of global change.

IFS is the ECMWF integrated forecasting system used for NWP, Reanalyses and MACC simulations.

MACC Monitoring Atmospheric Composition and Climate.

Nudging in data assimilation means to add a term to the state vector that is proportional to the difference of the calculated meteorological variable and the observed value. This term "keeps" the calculated state vector closer to the observations.

Pre-cursor refers to a CDR which has similar characteristics to the planned CCI CDRs. It may not be "climate quality". The ESA GlobXXX series datasets are examples of precursors. The main requirement for this purpose is that it can be assessed in a similar way to the CCI CDRs to demonstrate the methodology.

Reanalyses are estimates of historical atmospheric and oceanic temperature, wind, current, and other meteorological and oceanographic quantities, created by processing past meteorological and oceanographic data using fixed state-of-the-art weather forecasting models (atmospheric reanalysis), ocean monitoring and forecasting models (ocean reanalysis) and data assimilation techniques.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Uncertainty refers to a combination of random and systematic (bias) errors for each variable in a CDR. It normally refers to an individual observation but can refer to area and time averaged quantities.

3. Methodology applied to assess climate data records

For climate modelling the four key applications of the CCI datasets are to *Enable Model-Observation Confrontation*, *Provide Boundary Conditions*, *Provide Initial Conditions*, and *Provide Observations capable of assimilation*. Model-Observation Confrontation is the natural first step for a new dataset to be used with NWP reanalyses and climate models and this will be the primary activity performed by CMUG in a number of ways as listed below. Model-Observation Confrontation plays a significant role in the decision process that determines whether a dataset is deemed suitable (from the user's perspective) for the other 3 key applications.

The CMUG assessment will encompass the following aspects for a selection of the CCI climate data records:

Confront

- *consistency* of Global Satellite Data Products in *time* (e.g. stability, uncertainty of bias)
- *consistency with independent observations* (e.g. limb view, in-situ, ground-based remote sensing)
- *consistency with precursor datasets* to understand the differences and assess if the CCI datasets are better representations of the atmospheric/surface state
- *consistency compared to reanalysis fields*
- *consistency across ECVs*
- *ability to capture climate variability and small climate change signals* (e.g. observed trends) for their use in Climate Monitoring and Attribution.

Assimilate and boundary conditions

- *impact* in Model and Data Assimilation Systems (for a few ECVs where appropriate).

There is not a single methodology that can be used universally but several approaches from different science teams and tailored for each ECV are used so only general comments are given here with the details in section 4 for each ECV. In many cases an observation operator is required to compare the measured quantities with the actual model variables although often this operator is fairly trivial. A simple operator would be interpolation from model grid to observation point in space and time. A more complex operator would be a radiative transfer model to compare measured top-of-atmosphere radiances (level 1 data) with model equivalents. If higher level 2 or 3 products are used the operator is usually simpler as the variables are closer to the model variables although the error characteristics of the products can be more complex.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Data used for assessment of CDR	Advantages	Drawbacks
Climate Model (single, ensemble)	Spatially and temporally complete	Model has uncertainties Not all variables available
Re-analyses	Spatially and temporally complete	Analysis has uncertainties Not all variables available
Precursors	Comparing like with like	Some precursors may have large uncertainties
Independent satellite or in situ measurements	Different 'view' of atmosphere/surface	May have much larger uncertainty than CDR, need to include representativity errors
Related observations (surface and TOA fluxes, temperature, water vapour)	Assures consistency with other model variables	May not be spatially or temporally complete

Table 3.1. The various options for assessing CDRs and their advantages and drawbacks

When the products are used in model analyses, the correction of their systematic and random errors may be required. When they are used for direct comparison, the way they are used could be refined and this will be a topic of research in the assessment. In particular the assessment of the uncertainties provided with the data will need to be assessed in an objective manner.

For many atmospheric ECVs a comparison with the ERA-Interim reanalysis is appropriate as one way to assess the overall fidelity of the CDRs, noting the limitations of reanalyses. A comparison with other independent measurements, in situ, ground based remote sensing and other satellite products is also important. Ideally they should exhibit some differences in time sampling or are measurements using a different technique (e.g. limb viewing in infrared or microwave, aircraft sampling). There are also some related products which can be linked to some CCI ECVs (e.g. CO for biomass burning and aerosols, humidity or precipitation for clouds) which should be used in the assessments. Table 3.1 summarises the advantages and disadvantages of the various assessment datasets.

The consistency across ECVs is something that has been specifically identified as being important to the climate modelling community (the CCI project's targeted user community) and the CMUG will look at this aspect of the CCI datasets, drawing attention where necessary to inconsistencies between related ECVs. Increasingly, the climate modelling community approaches consistency from an integrated perspective which includes consistency across ECV product levels, e.g. from Level-1 radiances to Level-2 swath-based geophysical products to Level-3 gridded products, and also extends to ancillary data products such as bias corrections and homogenization terms. It is therefore important that the CCI continues its commitment to open access and traceability, which will entail preserving and making available all such products generated during the project.

An important requirement of an observational dataset for reanalysis is that when assimilated it improves (or at least does not degrade) the short range forecasts of relevant meteorological variables. Assimilation of the CCI products is a longer term goal in the context of reanalysis

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

projects (e.g. ERA-CLIM) and represents a critical test for some CDRs, but given that such tests are expensive to perform they must first be preceded by extensive quality assurance on the observational datasets in order to maximize the prospects for demonstrating beneficial impact.

This report documents some initial assessments of products of the ESA CCI climate data records. Table 3.2 summarises the way the ECVs selected in this report were assessed and it can be noted there are a variety of different methods.

Methodology used for assessment of ECVs	Assessment of ECVs in this report
Comparison with Climate Models (single, ensemble)	Clouds, SSH, Ozone, Land Cover, Fire
Re-analyses	GHG, Aerosol
Precursor datasets	SST, OC, SSH, Cloud, Land Cover, Fire
Independent satellite or in situ measurements	SST, OC, SSH, Clouds, Ozone
Related observations (surface and TOA fluxes, temperature, water vapour)	Land cover
Assimilation	Ocean colour, Ozone

Table 3.2. A summary of the different assessments performed on CCI CDRs in this report.

The following sections describe initial assessments of CCI datasets. Not all CCI ECVs are covered here and soil moisture (V0.1) assessment is reported in CMUG D3.1b (2013a) and the ECMWF assessments of ozone L2 and soil moisture products is in D3.1d CMUG (2013b).

4. Assessment of climate data records for CCI ECVs

A selection of the CCI ECV datasets have been assessed in this report as described below with ECV in a subsection. Table 4.1 lists the version numbers, release dates and any other relevant comments for the CDRs assessed. The assessments given below are only summaries of the work undertaken and for some of them at least it is planned to write up the work for a peer review journal.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

ECV	Version of Dataset	Release Date	Comments
SST	V1	January 2014	L3 (A)ATSR SST
Ocean Colour	V1	December 2013	L3 Chlorophyll conc
Sea Level	V1.1	January 2013	Monthly 1/4°
Clouds	V1 ('prototype data set')	October 2013	L3C/L3S
Ozone	V1	April 2013	Level 3 Merged Total column Level 3 Merged Limb Profile
Ozone	TCO3, fv0100 NPO3, fv1000 LPO3, fv0002	May 2013	Report in D3.1d Merged L3 Total column O ₃ GOME-1 & GOME-2 nadir O ₃ profiles MIPAS limb O ₃ profiles
Greenhouse Gases	CO ₂ BESD, 2.0 CO ₂ OCPF, 4.0 CO ₂ SRFP, 2.1 CH ₄ WFMD3.3 CH ₄ IMAP, 6.0 CH ₄ OCPF, 4.0 CH ₄ SRFP, 2.1	April 2013 April 2013 April 2013 April 2013 April 2013 April 2013 April 2013	Gridded month mean SCIAMACHY GOSAT GOSAT SCIAMACHY SCIAMACHY GOSAT GOSAT
Aerosol	FMI ADV 1.42 SU 4.0 SU 4.1-4.2 ORAC 2.02	May 2013 May 2013 Nov 2013 May 2013	AATSR L3 aerosol
Land cover	V1	Autumn 2013	Not available via CCI LC website, provided by CCI team
Fire	N/A	N/A	Dataset not available but more precursor work presented.
Soil moisture	V0	Summer 2012	Report in D3.1b (MPI) and D3.1d (ECMWF)

Table 4.1. Versions of CCI datasets assessed by CMUG in this report or for the rows in red which other CMUG report the assessment is provided.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

4.1 Sea Surface Temperature

4.1.1 Assessment of the CCI SST from ATSRs and comparison with the ARC dataset

The CCI SST is an effort to produce a complete and homogeneous dataset of SST designed specifically with the climate quality criteria in mind. It covers the period 1991-2010 when data from both ATSRs and AVHRRs are available. The ATSRs produce more accurate SST, due to their well calibrated blackbodies and dual view capability. The advantage of dual view is accompanied by the necessity of narrow swath width (at least in comparison to other satellite cross-track scanning radiometers). For this reason, the CCI SST project attempts to increase the coverage by the use of AVHRRs, which have larger swath width (2900 vs 500 km).

The ARC SST dataset, which was the precursor of the CCI SST project, involved only measurements from the ATSRs. Thus, in order to compare with the SST assessment from ARC (Lean and Saunders, 2013), hereafter the focus is on the validation of CCI ATSR SSTs only. The comparison is not straightforward as there are two main differences between ARC and CCI. Firstly, the method for the skin SST retrieval differs, with ARC using retrieval coefficients calculated from radiative transfer simulations, while CCI uses optimal estimation retrieval for ATSR2 and AATSR (but retrieval coefficients for ATSR1, the same as ARC).

Secondly, the ARC dataset is provided at a spatial resolution of 0.1° , while CCI has a finer resolution of 0.05° . Especially, because of the change in the spatial resolution, CCI and ARC are not directly comparable, as one waits a priori the precision to be worst for the CCI given the fact that it uses smaller number of measurements for the SST retrieval. It should be noted that the quality assessment of the ARC dataset provided previously was based mainly on the comparison of drifting buoy SST with ARC SST at a depth of 1m. As the CCI SST depth is calculated only at 20cm, the respective ARC SST depth (at 20cm) will be used afterwards for consistency. However, according to the results of Lean and Saunders (2013), the comparison statistics are similar for different SST depths, especially during night. The collocation criteria are the same as for ARC but due to the finer resolution of the CCI, the number of match-ups of the CCI with the drifting buoys is greater than for ARC.

The assessment results presented hereafter are based on the same analysis for both CCI and ARC in order to facilitate their comparison. However as mentioned, CCI has more collocations than ARC, due to its finer resolution. In both datasets a common threshold has been applied to eliminate match-ups with difference between buoys and ATSRs greater than $\pm 3K$. This constant threshold has been chosen instead of an approach based on three sigma elimination for two reasons. Firstly, the number of match-ups increases with time following the number of available drifting buoys, while also CCI and ARC do not have the same performance. The choice is not totally arbitrary, as it is based on the fact that the annual standard deviation of the ATSRs-buoys unfiltered differences is relatively close to 1K for both datasets. While the annual percentage of match-ups filtered out by this threshold is in general close to 1% for both datasets either for daytime or night-time retrievals. Nevertheless, there are years when the percentage of filtered match-ups is greater than 2%. For CCI these are 1998-2001 and 2005, while for ARC are 1998 and 2005. The years 1998 and 2005 appear

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

in both datasets, with the percentage of collocations filtered out in 1998 during CCI daytime retrievals approaching 6%.

Firstly, some technical issues found within the version of the CCI SST dataset released at the end of phase 1 are presented. Then, the quality assessment of the dataset is examined alongside a comparison of the ARC assessment. At the end, some recommendations to the CCI SST team are suggested.

4.1.2 Technical issues

The technical issues include regions with no data or with abrupt changes, unphysical values in the files and days without SST retrievals. It is believed that the technical issues are due to glitches/bugs in the retrieval software, but this hypothesis has to be verified by the CCI SST team.

Data gaps and abrupt changes

There are gaps for daytime retrievals over the western Pacific and for night time retrievals over the south-west Atlantic for all instruments (see Figure 4.1.1). These are related to a bug in the software detected by the CCI SST team, so the next version of the CCI SST will be corrected. Also, there is a smaller white gap off the western Indian coast for the daytime retrievals starting on 1992 for ATSR-1 and observed during the whole mission of ATSR-2 (but not observed for AATSR). For ATSR-1, the gap seems to be related to the inclusion of the 1.6 μ m channel in the cloud mask after the failure of the 3.7 μ m channel, as it is not seen at the beginning of the mission and it is not observed during the night time retrieval with the D2 algorithm (after the 3.7 μ m failure).

For the first 7 months of 1993, there are no match-ups at all during night, although there are for ARC (Figure 4.1.3c,d). Also the available number of night retrievals for 1993 is significantly lower in comparison to 1994 or 1995 (thus confirming that the problem is not related to buoys quality or match-up criteria). However, there are no specific issues with the ATSR-1 during this period. Note that this gap could not be detected from the CCI SST dataset, as there is no day/night flag and the D2 retrieval is applied both during day and night after the failure of 3.7 μ m. Thus, an addition of day/night flag in the dataset would be useful (see also below the section 'Recommendations to the CCI SST team').

During night time, there is an abrupt change for all three instruments in the number of retrievals appearing as a straight line at about 2°S and 8°N (e.g. see Figure 4.1.1b for the year 2007). This is not related only to D3 algorithm, as the same issue appears during the ATSR-1 mission after the failure of 3.7 μ m, when the D2 algorithm is used for both day and night retrievals.

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

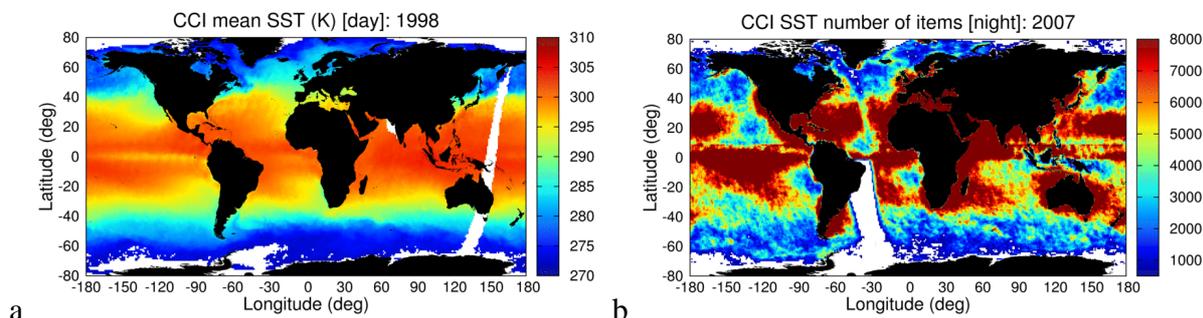


Figure 4.1.1 a) The average SST from the day (D2) retrieval for 1998. b) The number of retrievals realised during night (D3) for 2007. In both panels white areas without retrievals can be observed, during day over the western Pacific and off the west coast of India and during night over the south Atlantic. Also, during night there are abrupt changes in the number of retrievals, which appear as straight line at 2°S (more distinct over the Pacific Ocean) and 8°N.

Fill (unphysical) values

Sometimes either the depth SST or the uncertainties of skin and depth SST take “fill” values despite the fact that only the best quality SST data (Level 5) are used for validation purposes. It should be noted that for these cases all the other values do not seem to be affected, so the quality check applied by the CCI SST team appears not to control at the same time all the parameters of a specific measurement/retrieval.

The depth SST takes the “fill” value -54.53K. The occurrence frequency is of the order of 500 “fill” values per year for the day retrievals and of the order of 10 “fill” values per year for the night retrievals. However, the number of “fill” cases is important during the years 1993 (day ~10 millions), 2000 (day ~12 millions, night ~1 million) and 2002 for AATSR (day ~4 millions, night ~1 million). During these years, the “fill” values for the day retrieval (D2) appear mostly towards the poles (latitude > 60°) with the skin SST being in general lower than 278 K. A limited number is observed when the skin temperature is higher than 278 K very close to coasts at lower latitudes. The same conclusions (for 2000 and 2002) hold for the night retrieval (D3), but now with the majority of “fill” values found close to the Arctic.

The skin SST uncertainty takes the value -0.03 K about 500 times per year for both day and night retrievals for ATSR-2 and AATSR. Then, only for few night retrievals (~10 per year) the depth SST uncertainty becomes -327.68 K. However, the problem is more significant for ATSR-1 with negative skin SST uncertainty occurring from ~2500 (1991-night) up to 160,000 (1995-day) cases. In some other cases, the depth SST uncertainty takes the value 0K. This issue affects ATSR-2 and AATSR, with only 1 case found for ATSR-1. The 0 K depth SST uncertainty happens a few times (generally less than 5 per year) during the day retrieval and ~100 times per year for the night retrieval.

Days without data

There are some days without data at all, when the instrument is not affected by a specific known reason documented in the incident report (e.g. due to out-gassing, blackbody activity, etc.). It is worth highlighting that there is lack of official documentation (except for AATSR) that describes the performance of the instruments during their lifetime and tables/figures



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

containing the orbits with good quality data on a daily basis. The information about the quality of the ATSR archive is based on the following sources:

- i) the mission logs¹ of ATSR-1 and ATSR-2,
- ii) the completeness diagrams² of the TOA_1P product for the version 2.1 (found in the metadata directory),
- iii) the AATSR Instrument Performance, End of Mission Report³ of D. Smith,
- iv) the timelines diagrams⁴ (see the html links) and
- v) the AATSR instrument history⁵.

The days without SST data include the 29th February during the leap years: 1992, 1996, 2000, 2004, 2008 (and probably 2012). Note that the 29th February does not exist in the completeness diagrams of NERC. For ATSR-1 the following missing days have been identified: 26/10/1991, 13/1/1992, 18/1/1992, 8/2/1993, 28/4/1993 and 30/5/1996. For ATSR-2 the following days are missing from the CCI SST archive, but there is no information in the mission logs¹ or the timelines diagrams⁴ before the 3rd June 1998, thus it is not possible to verify the reason for their absence (it seems to be out-gassing periods): 10-13/8/1995, 26-27/11/1996, 15-16/2/1997 and 3/9/1997. For AATSR the following days have been identified as missing: 24/9/2002, 17/11/2005 and 24/7/2007.

4.1.3 Geographical distribution

The quality of the CCI SST is firstly assessed in terms of the geographical distribution of the mean bias shown in Figure 4.1.2. The mean bias is calculated for every 5°x5° grid box with at least 20 match-ups inside it for the period 8/1991-12/2009, thus the white boxes indicate lack of collocations (and the gaps over western Pacific during the day and south Atlantic during night seen in the CCI and discussed in the previous section). However, it should be kept in mind that the number of drifting buoys increased considerably during the recent years (especially after 2003), meaning that the maps reflect mostly the results for AATSR. The same analysis has been repeated for each of the 3 ATSRs independently and the results are similar to Figure 4.1.2a,c for AATSR and ATSR-2 during the day. For ATSR-2 during the night and especially for ATSR-1, it is difficult to conclude as there are far less collocations to give statistically significant results about the bias and a clear indication of the regions that have important bias.

It can be seen that during day the CCI depth SST is significantly warmer than the buoys SST in the tropics, while in general for the other regions either the mean difference is inside the range [-0.1, 0.1] K or slightly warmer than the buoys. Systematic exceptions with differences less than -0.1 K can be observed in the Black Sea and off the coast of Arabian Peninsula. On the other hand during night the CCI SST is found warmer than buoys in the mid-latitudes (mainly in the zone 20-40°), the Indian Ocean and the Western Pacific. Especially, in the case of the north Indian Ocean, there is a sharp transition with the bias being cold off the Arabian Peninsula becoming warm close to the Indian Peninsula. A less significant gradient can be

¹<http://www.atsr.rl.ac.uk/satellite/index.shtml>

²http://neodc.nerc.ac.uk/browse/neodc/aatsr_multimission/

³<http://www.aatsrops.rl.ac.uk/EOMdocs.html>

⁴<http://www.neodc.rl.ac.uk/docs/atsr/timelines/>

⁵http://earth.eo.esa.int/pcs/envisat/aatsr/instrument_history

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



also observed over the eastern part of tropical Atlantic. The cold bias off western Africa and Arabian Peninsula may be related to dust aerosols, which are abundant in these two areas. The same conclusions for both day and night hold for the case of the median bias, but with the maps looking smoother. The location of the day retrievals' bias in the tropics seems to be related with the absorption of radiation by water vapour, which is more abundant in this region of the Earth, and it is more difficult to be taken into account by the 2 channels algorithm, used during the day. The mid-latitude bias during night may be related to systematic errors (prior and nonlinearity) described in the work of Merchant *et. al.* (2006).

The significant warm biases seen in the CCI dataset are absent in ARC both during day and night or at least much less prominent. The regions appearing consistently with warm SST biases with respect to buoys for both CCI and ARC are off the coasts of Indian Peninsula and Southeast Asia and the Gulf of Mexico, with the region around the Indian Peninsula being the more challenging (bias about 0.5K). Regarding the standard deviation of the difference between CCI and buoys (not shown), during the day the larger values are found over the Maritime Continent and Northwest Pacific, while during the night off the eastern coast of Canada and the south-western Atlantic Ocean (between S. America and the white gap). Again the ARC performs better than CCI, in terms of standard deviation. The superiority of ARC SST in comparison to CCI SST is reflected in the area weighted (by the cosine of latitude) mean bias and standard deviation. For CCI the values are 0.09 ± 0.48 K during day and 0.10 ± 0.41 K during night, while for ARC the respective values are 0.07 ± 0.41 K and 0.05 ± 0.36 K.

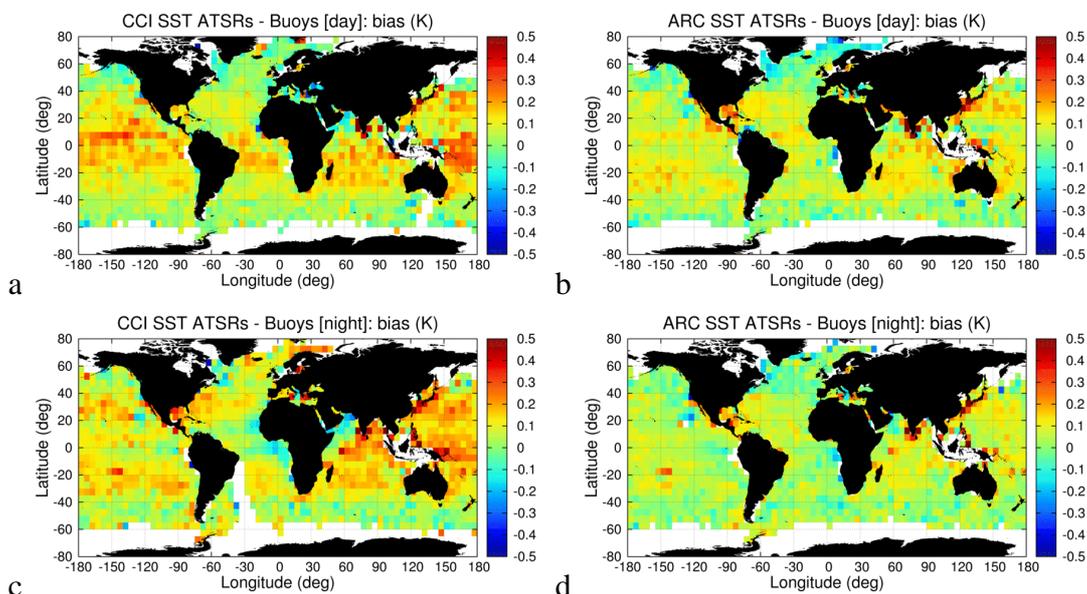


Figure 4.1.2. The mean bias of depth SST from ATSRs minus drifting buoys averaged in $5^\circ \times 5^\circ$ boxes for the period: 8/1991-12/2009. The comparison for daytime observations is given in panels (a), (b) and for night-time observations in panels (c), (d). The left hand panels (a), (c) present the results for CCI and the right-hand (b), (d) panels for ARC.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

4.1.4 Time series

In order to have a clear indication how the bias is evolving with time, as Figure 4.1.2 reflects mostly the results of AATSR, Figure 4.1.3 presents the time series for CCI and ARC. The red symbol is the daily global bias when at least 20 collocations are available. The lack of significant number of match-ups during the ATSR-1 mission can be clearly seen. The blue line is the monthly bias calculated by averaging the match-ups in 5° latitudinal boxes and then by taking the weighted (by the cosine of latitude) mean value. The mean value for a specific month is provided when at least 5 latitudinal boxes have at least 20 match-ups for this month. However, because there are months which do not fulfil these criteria the mean value of all available match-ups is given by the green line in Figure 4.1.3. It can be seen that the two lines are almost identical after 1995. Figure 4.1.3 shows that the variability of the bias in a daily scale is more important than in the monthly average, e.g. with some days exceeding the ± 0.4 K levels (except from ARC night-time SST retrieval). This is an important point to be kept in mind when assessing the quality of SST datasets.

For the CCI SSTs during the day, the first year of ATSR-1 has a warm bias, lower or close to 0.1K, which from July 1992 increases remarkably reaching a maximum in late summer of 1993 and then decreasing close to 0.1 K. This increase probably is related to the failure of the 3.7 μ m channel (on 27th May 1992) and then the inclusion of the 1.6 μ m channel in the cloud mask during day. However, a possible impact of stratospheric aerosols from the Pinatubo eruption on the SST retrieval cannot be excluded. The main reason of the increased bias is difficult to assess at this point due to the lack of night-time retrievals for this period (Figure 4.1.3c). From 1997, the bias becomes negative reaching a minimal value of the whole time series in June 1998 (-0.1 K). In 1999, the bias is below 0.1 K during the first 8 months, then increasing rapidly to 0.2 K (October 1999), staying at this level for 2000 and above 0.1 K during 2001. From 2002 and onwards, there are no abrupt changes and the bias is about 0.1 K, reflecting the stability of AATSR. However, an oscillation can be observed with a warmer bias during the boreal winter months. This oscillation exists also during the periods of ATSR-1 and ATSR-2, but it becomes more obvious after 2002 due to the lack of other significant anomalies.

The night-time CCI SST is more stable than the daytime, as the bias stays close to 0.1 K during the whole period. However, there are two jumps at the beginning of 1997 (decrease from 0.2 K to 0.05 K) and at the end of 2000 (increase from 0.05 K to 0.2 K) with a stable interval between at the level 0.05 K. Once again the AATSR period is more stable in terms of bias, accompanied by a slight decrease after 2001. When switching between ATSR-1 and ATSR-2 an abrupt change of the bias can be observed (for daytime observations a change also exists but it is less significant). This is an expected outcome as ATSR-2 uses also the 3.7 μ m channel in the night-time algorithm.

Concerning the comparison between CCI and ARC, during the period of ATSR-1 there are significant changes of bias, with the mean monthly value being higher about 0.2 K for CCI in comparison to ARC for both daytime and night-time retrievals. Also, the time evolution of the bias is not exactly the same. This fact is surprising as CCI for ATSR-1 used the retrievals with coefficients based on radiative transfer calculations, which is the same SST retrieval method used in ARC. The main difference between the two datasets for this period is the spatial

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1



resolution (finer for CCI), as the cloud screening method is the same (with minor differences concerning the versions of the radiative code and NWP model). This is a point that requires further examination, i.e. how the spatial resolution affects the quality of the SST retrievals and if it is only related to the performance of the cloud mask. But in depth analysis of this point requires two datasets with exactly the same retrieval method of SST and just different resolution, which is not the case for CCI and ARC. For the period of ATSR-2 and AATSR, the two datasets present similar behaviour regarding the daytime retrieval, with the CCI having notable annual oscillations. The biggest difference is observed in 1995, when ARC presents lower bias (mostly negative) than CCI during the first months of ATSR-2. Regarding the night-time retrieval before October 1999 there is important difference between CCI and ARC, with CCI having 0.1 K higher bias than ARC, as seen during the ATSR-1 period. Except that the sign of bias is not the same with CCI presenting a warm bias and ARC having a cold bias. After the end of 1999, the two datasets show the same evolution in terms of bias, with CCI being slightly warmer than ARC, as it is the case for the daytime retrieval.

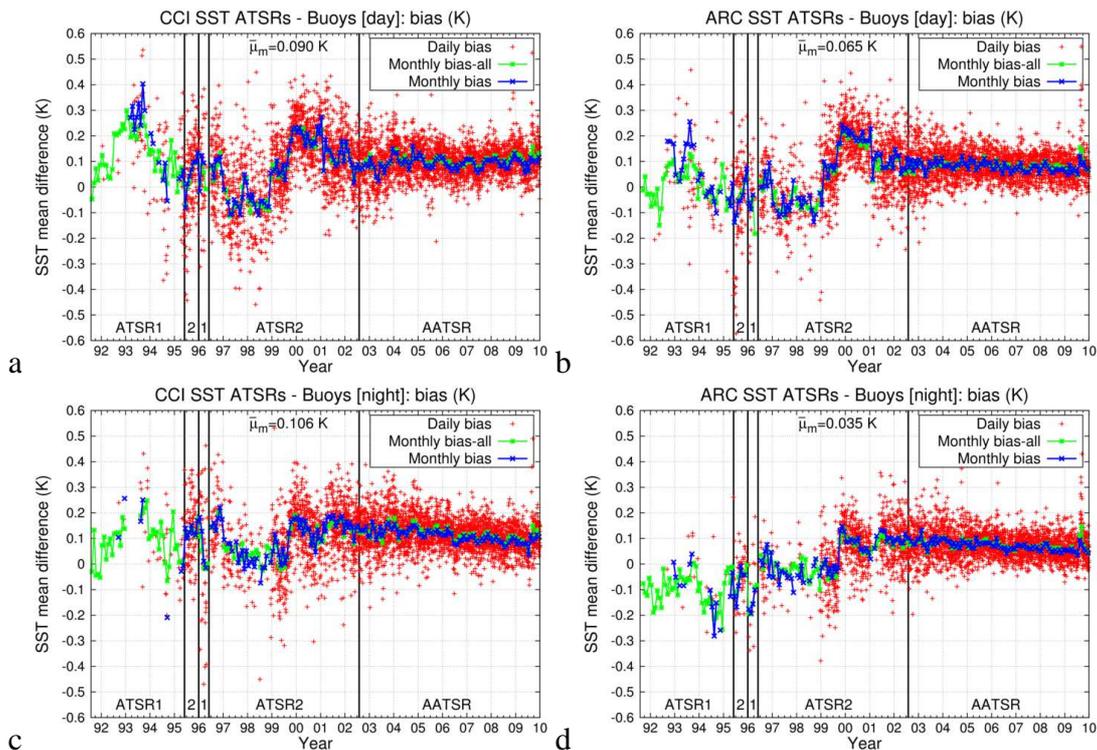


Figure 4.1.3 Time series of the mean bias for the period 8/1991-12/2009: daily (red), monthly for all the available match-ups (green) and monthly weighted from 5° latitudinal boxes (blue). The $\bar{\mu}_m$ is the average value of the mean bias from all the match-ups. The vertical black lines indicate the switch between the ATSRs. The panel order is same as Figure 4.1.2.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

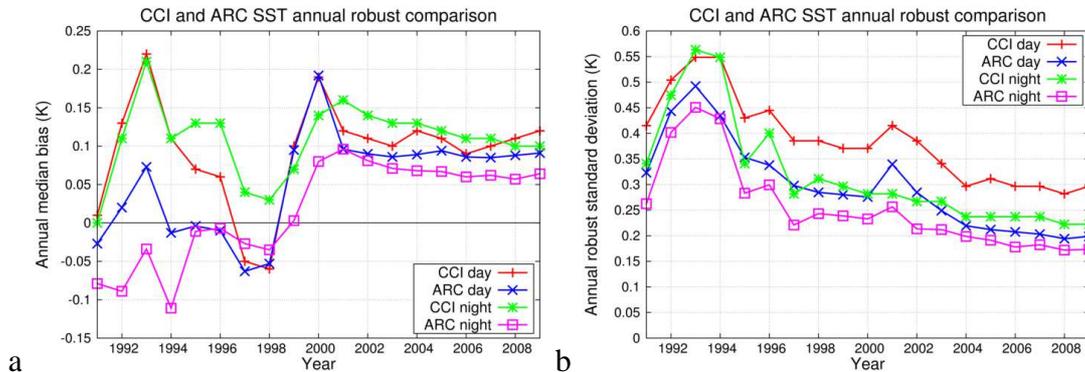


Figure 4.1.4. Annual robust statistics (a) median bias and (b) robust standard deviation calculated from all the available match-ups for each year. Lines colours are: for CCI during daytime (red) and night-time (green) and for ARC during daytime (blue) and night-time (magenta). In panel (a) the horizontal black line indicates the zero bias.

In addition, the time evolution of the monthly standard deviation (not shown) of CCI is alike to ARC. In each dataset the daytime and night-time retrievals show similar patterns of standard deviation, with the daytime taking higher values in general. The standard deviations are increasing with time until 1994 and then decreasing until the switch from ATSR-2 to ATSR-1 in 1996 due to scan mirror failure. After July 1996 the standard deviation stays relatively constant until the end of 1999, and then slightly decreases with time for both datasets. In December 1998–February 1999 and January 2001 (and the next 5 months) there are notably increases of the standard deviation (more than 25%), with the former change being more significant for night-time retrievals and the later for the daytime retrievals. The increase of January 2001 can be attributed to the gyro failure in 15th January 2001, lasted for the next ~5 months. Both increases of standard deviation are accompanied by change in the number of collocations with increase (decrease) for 1998 (2001). The biggest difference between CCI and ARC is the value of the standard deviation, which in the CCI is about 0.1 K higher than ARC. Another difference of the two datasets is that the significant changes in the standard deviation (e.g. the switch between ATSR-1 and ATSR-2) are more abrupt in terms of standard deviation for CCI than for ARC (except from the gyro failure night-time retrievals).

In order to give a more complete picture of the comparison between CCI and ARC, the robust statistics are presented in Figure 4.1.4. The median and the robust standard deviation (which is 1.4826 times the median absolute deviation) are provided because they are less affected by outliers and they fit better the distribution of the differences between ATSRs and buoys. For each year the median and the robust standard deviation (RSD) are calculated for all the available match-ups after the elimination of the match-ups with absolute difference greater than 3 K. Although, for robust statistics this elimination is not necessary (being by definition robust to outliers), this step has been kept in order to retain the same dataset, either for CCI or ARC, as in the previous sections. Figure 4.1.4 clearly shows that ARC is better than CCI both during night and day in terms both of median bias and RSD. Even the ARC daytime retrievals are performing better in general than CCI night-time retrievals. Note that the CCI median bias is always positive, i.e. a warm bias, except in 1997 and 1998 for the daytime retrievals, when the median bias is negative. Also, the median bias of the night-time CCI SST is in general bigger than the corresponding one during daytime. Indeed, the robust statistics for the CCI



CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

and ARC (after the 3 K filter) datasets as a whole (thus biased towards the performance of AATSR) confirm the above conclusions, with values (median \pm RSD) for CCI during daytime 0.10 ± 0.33 K and during night-time 0.11 ± 0.25 K, while for ARC during daytime 0.08 ± 0.24 K and during night-time 0.06 ± 0.20 K. The time evolution of the median bias is consistent with the results of Figure 4.1.3 (note that the ordinate scale is different in Figures 4.1.3 and 4.1.4a). This is also the case for the RSD time series. For example, one can notice the negative impact of switching back to ATSR-1 data during night-time retrieval in 1996, after the scan mirror failure of ATSR-2 (on 22nd December 1995 and the data gap for more than 6 months), which is more important for CCI than ARC. On the other hand, the gyro failure in 15th January 2001 affects more notably the ARC dataset than the CCI, notably for the night-time retrievals. However, the slight increase of the RSD seen during 1998, more obvious in the night-time retrievals, is not related to the increase of the standard deviation in December 1998-February 1999. It most probably reflects the fact of more match-ups presenting larger differences between buoys and ATSR-2, mentioned in Section 4.1.4. Note also, a similar increase in 2005 but only for CCI daytime retrievals, which it can be probably explained by the same reason. The above results indicate the utility of using both statistics in order to assess the quality of a SST dataset, as they are complementary.

4.1.5 Three way error analysis

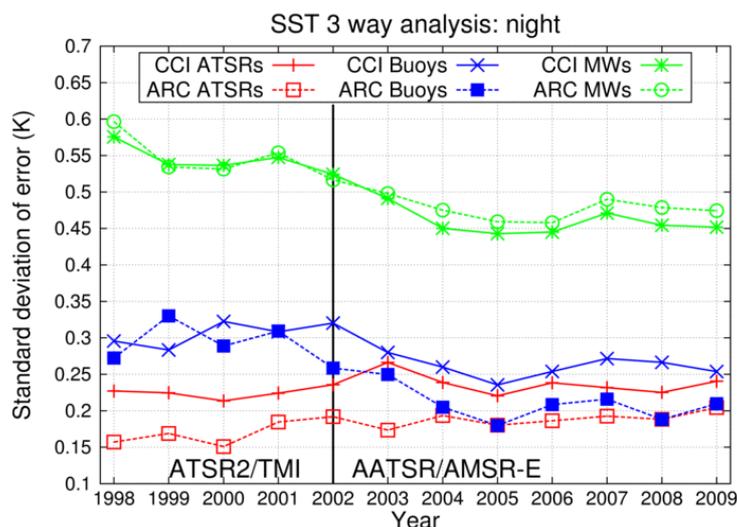


Figure 4.1.5 Annual standard deviation of the error (K) from the three way analysis. The results of ATSR-2/AATSR are shown by red symbols, buoys by blue and TMI/AMSR-E by green. In 2002, the first 7 months are from the combination ATSR-2/TMI and the next 5 months from the combination AATSR/AMSR-E, merged together to give the annual values.

The three way error analysis is described in the study of O'Carroll *et al.* (2008). By using three independent datasets observing the same quantity, the standard deviation of every one can be calculated given the fact that the observations are uncorrelated. Here, the last version of AMSR-E is used (version 7, which differs from the versions used by O'Carroll *et al.* (2008) and by Lean and Saunders (2013)). Also, the results are extended backwards in time by using the SST of TMI (version 4) before August 2002, covering only the latitudes -40 to 40° N. Thus, the couples ATSR-2/TMI are used for the years 1998 to 2001 and AATSR/AMSR-E for

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

the years 2003 to 2009, while for 2002 both couples are used. The match-ups of ATSRs with buoys are collocated with the closest observation of MW SST (provided with resolution of 0.25°) in a time window of 180 min for all three observations. The results are given only during night, in order to minimize the effect of the diurnal cycle. Geographically these collocations are found in the latitude zone -60°N (-40°N for the TMI period) to 40°N , with the number of match-ups increasing with time. As previously, only collocations with absolute differences less than 3 K are used for any couple of the 3 datasets.

Figure 4.1.5 shows that ARC is better than CCI, while both datasets do not show significant evolution with time. The difference between CCI and ARC is not only due to different retrieval method but also due to dissimilar spatial resolution. On the other hand, both buoys and MW SSTs present net amelioration with time. One can notice that the buoys' standard deviation of error is not the same when comparing to CCI or ARC, especially after 2001. This reflects the sensitivity of three way analysis to different datasets (as MW SSTs are the same for both CCI and ARC). It should be mentioned that the threshold used for the elimination of match-ups also affects the results, especially them of buoys, due to the fact that the satellite instruments are more stable with time.

4.1.6 Validation of uncertainty

The characterisation of the uncertainty associated with every SST retrieval is examined. Figure 4.1.6 presents how the mean bias and the standard deviation evolve in terms of uncertainty. Ideally, one expects the bias (red crosses) to be 0 K and the standard deviation (blue squares) to lie on the green lines ($y=x$) for perfect measurements from both buoys and ATSRs. However, the buoys' measurements are characterised by an uncertainty of about 0.2 K (O'Carroll *et. al.*, 2008). Thus, the ideal case of perfect characterisation of the CCI uncertainty means that when the SST uncertainty (from CCI) approaches 0 K the standard deviation should asymptotically approach the value of 0.2 K.

For CCI daytime retrievals, the bias increases slightly with increasing uncertainty. The standard deviation lies very close to 1:1 line for the uncertainty interval [0.4, 0.9] K. Below the uncertainty level of 0.4 K, the standard deviation is almost constant at 0.5 K, which is higher from the expected value of 0.2 K. There are only few observations (light blue bars) with uncertainty above the value of 0.9 K, thus the departure from the 1:1 line is not a major issue. For the validation of the CCI night-time uncertainty, the situation is more complicated as in general the measurements have an uncertainty of either ~ 0.2 K or ~ 0.28 K (Figure 4.1.6c), with only a small number having uncertainties in the interval [0.13, 0.3] K and a very limited number with uncertainty larger than 0.3 K. The bias for the big majority of measurements (having uncertainties of about 0.2 or 0.28 K) is larger than 0.1 K and the standard deviation is about 0.4 K.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

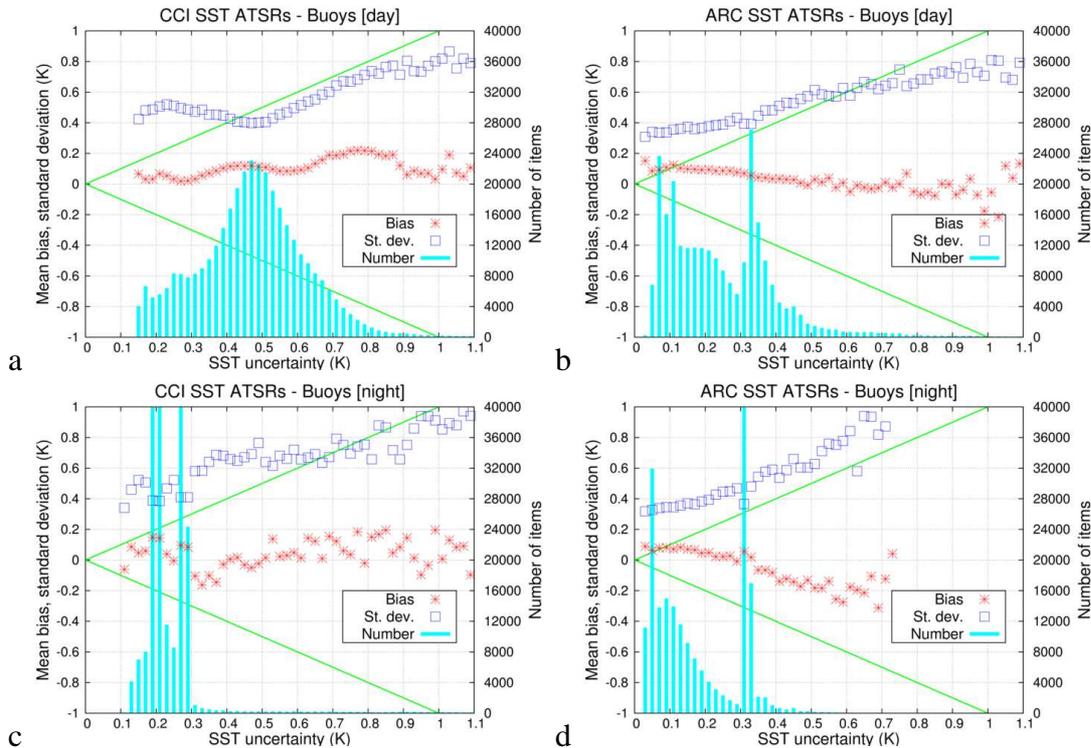


Figure 4.1.6. Mean bias, standard deviation and number of match-ups versus SST uncertainty for all the collocations of the period 8/1991-12/2009. The uncertainty SST bin width is 0.02 K. The 1:1 lines are shown with green colour. The panel order is same as Figure 4.1.2.

One can note that in the ARC dataset, there are uncertainties below 0.1 K with a significant number of observations taking them, which is not the case within the CCI. Figure 4.1.6 shows that the ARC uncertainties approaching better the idealised case for both the bias and the standard deviation than CCI. This is confirmed also by the chi-squared (χ^2) test (by assuming that the uncertainty of buoys is 0.2 K), indicating that when the uncertainty is lower than 0.35 K, the CCI dataset significantly underestimates it both for daytime and night-time retrievals. Of course, the uncertainty assignment of the ARC dataset is far from perfect according to χ^2 -test, but in general is better than the respective of the CCI. However, by examining the percentage of observations of which the ratio bias over uncertainty is lower than 1 and 2 (ideally the percentage should be 68% for ratio lower than 1 and 95% for ratio lower than 2) the CCI dataset seems to be slightly better than ARC. Indeed, for the daytime CCI the percentages are (for ratio lower than 1 and 2, respectively) 76% and 93% and for the night-time are 56% and 83%, while for the daytime ARC the percentages are 58% and 81% and for the night-time are 54% and 76%. In conclusion, the uncertainty assignment of the ARC is in general better than the uncertainty of the CCI, although not for all the uncertainty values or the validation criteria.

4.1.7 Recommendations to the CCI SST team

- It would be useful to provide the time of depth SST also in UTC, rather than in local time (10 am/pm). This will make easier the comparison with other SST observations (e.g. buoys), as the time reported in all databases is the UTC time.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

- Add a day/night flag in the CCI dataset. This information is relevant mostly to the ATSR-1, where the D2 retrieval is used both during night and day, after the failure of the 3.7 μm channel. This discrimination is useful for two reasons during the post 3.7 μm period. Firstly, different cloud masks are used during night and day (use of 1.6 μm after the 3.7 μm failure), thus the cloud filtering is expected to be more efficient for the daytime retrieval (3 channels instead of 2 during night). Secondly, because of the diurnal cycle of SST, especially in regions with low winds and strong solar insolation, the retrieval (D2) may not perform equally well under both day and night conditions. This discrimination will help the better validation of the SST product against buoys during the ATSR-1 period.
- There is no information about the cloud mask of the ATSR-1 retrievals and how it changes with time in the ATBD. This information should be provided.
- A document describing which days have orbits with good quality data and are included in the CCI SST archive should be provided. This will permit to the users to know exactly when the CCI SST data are available, while it will also facilitate the production of the next CCI SST versions. Ideally, the document should also state the reasons for days/orbits with low quality or absence of data.
- It should be noted that there is no official centralised information with easy access (either from ESA or NERC) covering all 3 instruments. This fact can be verified by the multitude of sources needed to check the performance and the data quality of all 3 sensors (see footnotes of the sub-section Technical issues).

4.2 Ocean colour

4.2.1 Introduction

An initial assessment has been made of version 1 (V1) of the ocean colour CCI (OC-CCI) chlorophyll product (<http://www.esa-oceancolour-cci.org>), released in December 2013. Comparison is made here with the precursor ESA GlobColour product (<http://www.globcolour.info>), by assimilating both data sets into a coupled physical-biogeochemical ocean model.

The physical component of the model is the Forecasting Ocean Assimilation Model (FOAM; Storkey et al., 2010), based on the NEMO hydrodynamic model (Madec, 2008) and the CICE sea ice model (<http://climate.lanl.gov/Models/CICE>). A 1° resolution configuration has been used, forced with three-hourly fluxes from the ERA-Interim reanalysis (Dee *et al.*, 2011). There is the option to assimilate physical ocean data, but this has not been included in this study. The biogeochemical component is the Hadley Centre Ocean Carbon Cycle Model (HadOCC; Palmer and Totterdell, 2001). HadOCC is a relatively simple nutrient, phytoplankton, zooplankton and detritus (NPZD) model, which also features a fully coupled carbon cycle and variable carbon to chlorophyll ratio. FOAM-HadOCC has been used in previous studies investigating the assimilation of ocean colour (Ford *et al.*, 2012) and pCO₂ (While *et al.*, 2012) data.

The OC-CCI and GlobColour chlorophyll data are each daily level three products, merging information from the SeaWiFS, MERIS and MODIS-Aqua sensors. The time series begins in

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

September 1997 and data are available until July 2012 for OC-CCI, and present day for GlobColour. The observations are quality controlled, super-obbed and assimilated following the method described in Ford *et al.* (2012). This uses the nitrogen balancing scheme of Hemmings *et al.* (2008), which directly updates all biological and carbon cycle model state variables, aiming to propagate the information from the chlorophyll data as fully and realistically as possible to the unobserved variables.

4.2.2 Experiments

In order to ensure complete independence between the OC-CCI and GlobColour assimilation experiments, the two data sets need to be treated identically, but separately, throughout, including in the calculation of the background fields and error variances used for the quality control and assimilation.

Monthly climatologies for use by the quality control were created for each data set by averaging the observations for the period 1998 to 2011 onto a $1^\circ \times 1^\circ$ grid. A control run with no data assimilation (hereafter CONTROL), was run from January 1989 to July 2012, starting from climatology and rest. The period from January 1989 to August 1997 is treated as spin-up.

Monthly-varying background and observation error variances were calculated using model results for 2003. An initial set of error variances, independent of either data set, were calculated from CONTROL using the quick Canadian method (Polavarapu *et al.*, 2005). These were then used in assimilative runs with each data set. From these results, combined with the observations, a set of error variances were calculated for each data set using the method described by Ford *et al.* (2012). These error variances were used for the main assimilative runs described below, as well as to quality control the observations. At present, the observation uncertainties provided with each data set are used by the quality control, but not the assimilation. However, the system could be developed to make further use of this information in future.

To assess the impact of assimilating the OC-CCI and GlobColour data, assimilative runs with each data set have been performed from 1st September 1997 to 31st July 2012 (hereafter CCI-FULL and GC-FULL), using initial conditions taken from CONTROL. In addition, year-long assimilative runs have been performed for 2003 (hereafter CCI-2003 and GC-2003), using initial conditions from CONTROL on 1st January 2003. This is because time constraints due to the delayed release of the OC-CCI data meant that CCI-FULL could not be completed with sufficient time for a full assessment. Therefore, most of the results presented here are from the one-year runs, with some preliminary results from the full 15-year reanalyses.

4.2.3 ResultsResults – observation processing



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

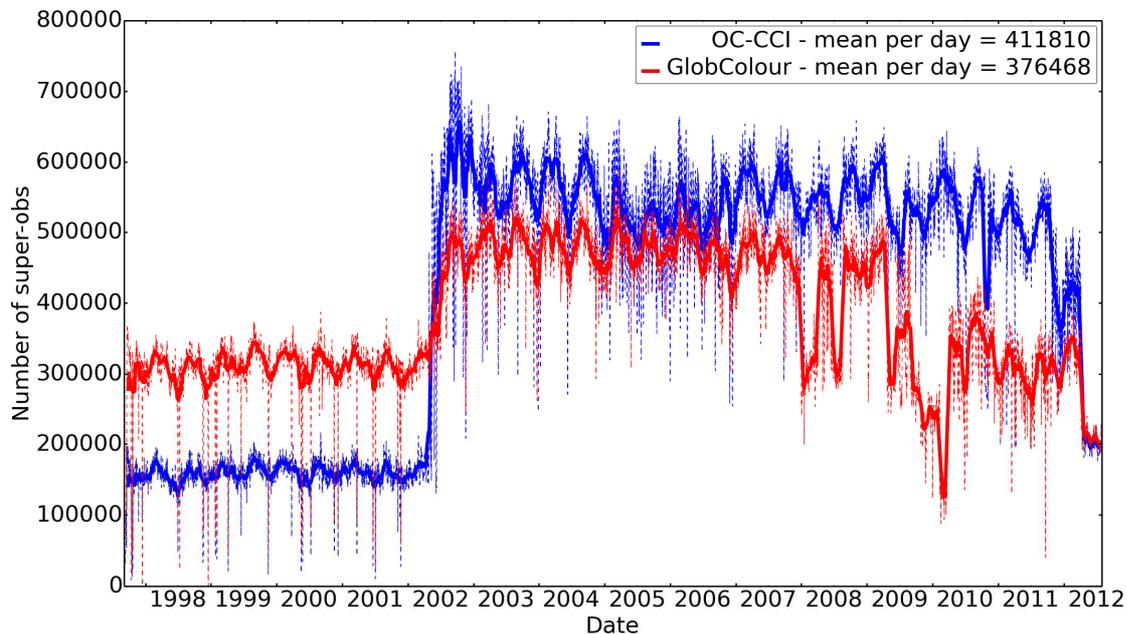


Fig. 4.2.1. Number of super-obs created each day for OC-CCI (blue) and GlobColour (red). Dashed lines show the daily numbers and solid lines the 30 day rolling mean.

A time-series of the number of super-obs created for each product is shown in Fig. 4.2.1. The launch of MERIS and MODIS in 2002 can be clearly seen in both time-series, as can the loss of MERIS data in 2012. During the SeaWiFS-only period, there are fewer OC-CCI observations, but vice-versa during the MERIS era, due to the increased exploitation of MERIS data by the OC-CCI algorithms. The GlobColour data are more sensitive to interruptions in SeaWiFS data in the period leading up to its loss at the end of 2010, and the improved consistency of coverage in the OC-CCI products during this time is expected to be beneficial. The differences seen in Fig. 4.2.1 correspond to differences in the products prior to quality control and super-obbing.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

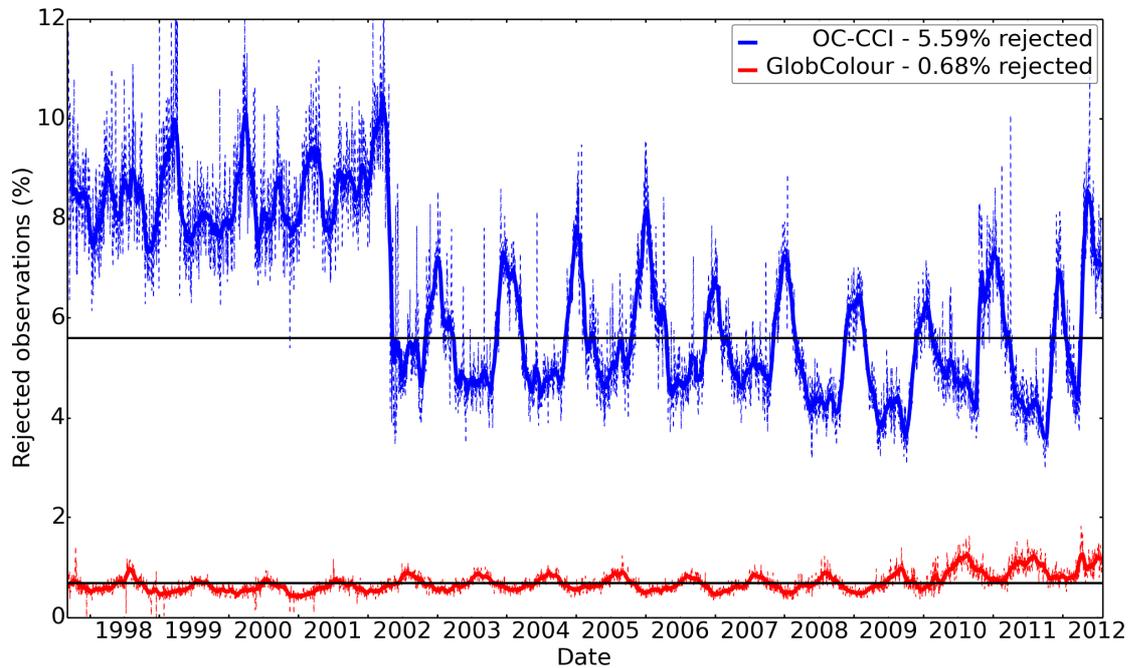


Fig. 4.2.2. Percentage of observations rejected each day by quality control for OC-CCI (blue) and GlobColour (red). The dashed lines show the daily numbers, the solid lines the 30 day rolling mean, and the black lines the mean for each product.

A time-series of the percentage of observations rejected by the quality control is shown in Fig. 4.2.2. For GlobColour the mean rejection rate is 0.68%, remaining reasonably steady throughout the period, although with an increase after 2009, and a clear seasonal cycle. For OC-CCI, a much higher rejection rate is seen, with a mean of 5.59%. This is because a number of OC-CCI observations do not have an associated uncertainty provided, and so are automatically rejected by the quality control. This was originally reported as an issue in the V0 dataset, and is planned to be addressed by the OC-CCI team during phase two of the project. The rejected observations seem fairly evenly distributed, and the impact on the final assimilation results is likely to be small. A substantial decrease in the rejection rate can be seen following the introduction of MERIS data in April 2002, with the time-series remaining reasonably steady before and after this date, albeit it with a large seasonal cycle, out of phase with that seen for GlobColour.

A further issue, also originally reported in the V0 dataset, is that the dates in the OC-CCI files are a day out during British Summer Time. This initially led to all observations being rejected by the quality control, and so the time in the files had to be manually altered.

Results – 2003

The primary aim of the data assimilation is to constrain the model results to better match the assimilated observations, and this is the first thing to check when evaluating an assimilation scheme. As a qualitative demonstration, monthly mean surface chlorophyll for April 2003 from each of the three model runs (CONTROL, CCI-2003 and GC-2003), and each of the observation data sets, is shown in Fig. 4.2.3.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

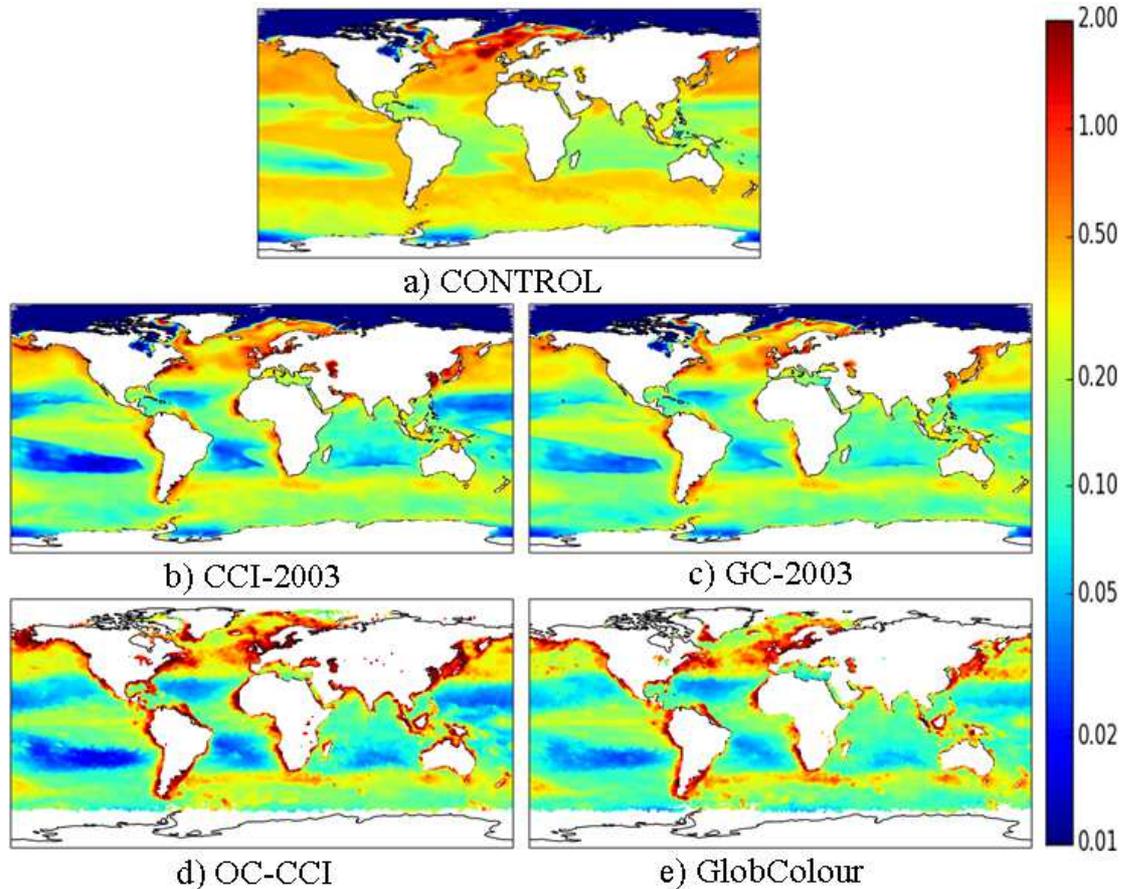


Fig 4.2.3. Mean surface chlorophyll (mg m^{-3}) for April 2003 from each model run and observation data set. The daily OC-CCI and GlobColour products for April 2003 have each been averaged onto a regular 1° grid for this figure. This is because, at the time of writing, OC-CCI monthly composite products are not yet available.

The OC-CCI and GlobColour observations are broadly similar in terms of spatial pattern and magnitudes. Compared to both products, CONTROL has a bias towards high chlorophyll across most of the ocean. General features are reproduced, but there are biases in the more detailed spatial positioning. These biases are considerably reduced in CCI-2003 and GC-2003, with both assimilative runs more closely resembling the two observation products. Overall, CCI-2003 and GC-2003 give similar results, mirroring the similarity between the observation data sets. There are noticeable differences however, and these match differences in the observations. For instance, the Kuroshio region is more clearly distinguished in GC-2003, whereas the Mauritanian upwelling region is more distinct in CCI-2003, which appears to be a result of increased observational coverage in this area. The biggest difference is the chlorophyll concentration in the sub-tropical gyres, particularly in the Pacific, with lower concentrations in the OC-CCI data, and therefore the CCI-2003 model run. A contributing factor is that the GlobColour products have a higher minimum value, but even taking this into account, the sub-tropical gyres are much clearer in the OC-CCI data. In situ observational coverage is poor, so it is unclear which is more realistic.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

It can also be demonstrated statistically that the assimilation improves the fit of the model to the assimilated data, as it is designed to do (not shown, but see Ford *et al.*, 2012). However, the true test comes when comparison is made to independent observations, both of the assimilated and non-assimilated variables. Statistics comparing each model run to in situ chlorophyll and nitrate from the Atlantic Meridional Transect (AMT; <http://www.amt-uk.org>) cruises, and to in situ fugacity of carbon dioxide ($f\text{CO}_2$) from V2 of the Surface Ocean Carbon Atlas (SOCAT) database (Bakker *et al.*, 2013), are given in Table 4.2.1.

	log ₁₀ (chlorophyll)		Nitrate		fCO ₂	
	RMSE	r	RMSE	r	RMSE	r
CONTROL	0.98	0.42	4.42	0.83	37.53	0.60
CCI-2003	0.63	0.77	4.34	0.83	37.52	0.61
GC-2003	0.63	0.77	4.34	0.83	37.65	0.60

Table 4.2.1. RMS error (RMSE) and correlation (r) between each model run and in situ observations of log₁₀(chlorophyll) (log₁₀(mg m⁻³)) and nitrate (mmol.N m⁻³) from AMT, and surface fCO₂ (µatm) from SOCAT. AMT observations taken in the surface 10 m have been used here. fCO₂ observations in water depths less than 150m have been excluded, to remove the influence of coastal regimes the model cannot resolve.

Compared to in situ observations of log₁₀(chlorophyll), taken in the surface 10 m, CCI-2003 and GC-2003 both show a 36% decrease in RMS error compared with CONTROL, and an increase in correlation from 0.42 to 0.77. This is a clear demonstration that the assimilation is improving model chlorophyll, with very similar statistics achieved with both products. A decrease in RMS error and increase in correlation is also seen beneath 10 m (not shown). For nitrate, CCI-2003 and GC-2003 both show a slight decrease in RMS error compared with CONTROL, but the statistics are very similar for all three runs. Nonetheless, this is an important result, as not degrading nutrient concentrations is a major challenge in ocean colour data assimilation. For fCO₂, all three runs compare very similarly against observations in waters deeper than 150 m, as in Table 4.2.1, due in part to fCO₂ being largely physically controlled. However, more localised case studies show that changes in chlorophyll following data assimilation do impact on fCO₂ and air-sea CO₂ flux, and the extent and subtleties of this impact will be explored in more detail in phase 2 of the project. When comparison is made to all available fCO₂ observations (not shown), including those in shallow waters, the RMS errors are greatly increased, because the 1° global model is unable to resolve the coastal processes which dominate in such areas. However, an improvement in both RMS error and correlation is seen in these regions following data assimilation, with a bigger improvement in CCI-2003 than GC-2003. This indicates the assimilation to be beneficial in shelf seas, though this conclusion may not be robust on a global scale, as these are not regions either the model or ocean colour products were designed to properly represent.

Results – 15-year reanalyses

If an observation product is to be useful as a climate data record, as intended with OC-CCI, then it should be consistent and stable through time, with a minimum of bias due to changes in sensors, and no artificially introduced trends due to the data processing. Such issues have previously been reported in the GlobColour data set (Maritorena *et al.*, 2010), and the

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1



consistent processing and bias correction used to generate the OC-CCI products is designed to address these.

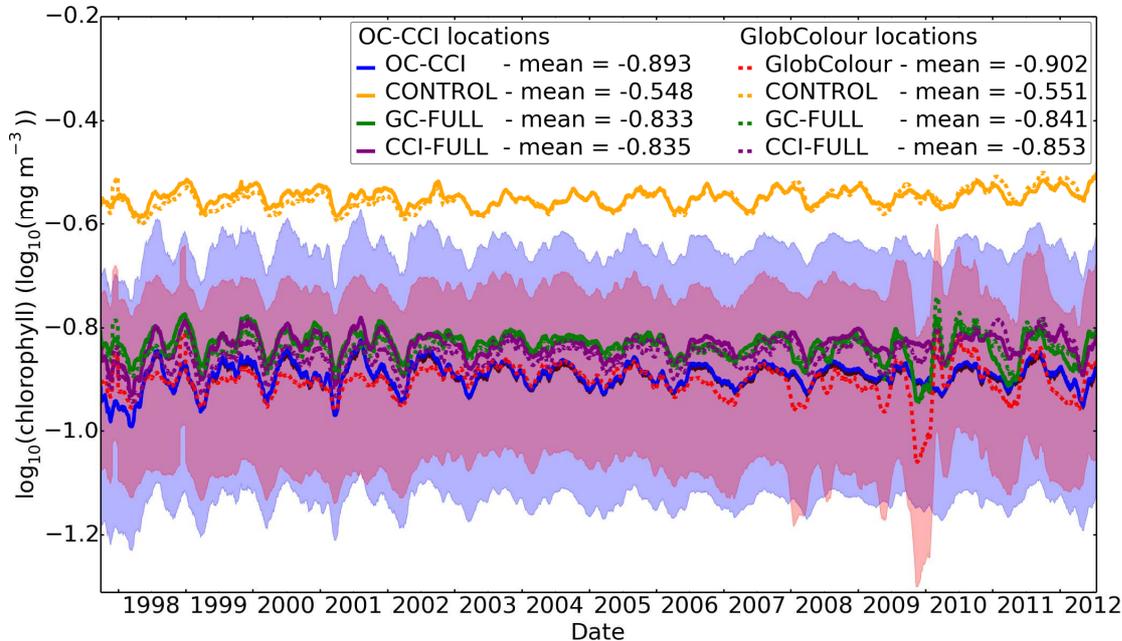


Fig. 4.2.5. 30 day rolling mean global $\log_{10}(\text{chlorophyll})$ from OC-CCI (blue) and GlobColour (red), plus and minus their RMS uncertainties (shaded). Black shading around the OC-CCI line represents the bias uncertainty. Also plotted is global mean $\log_{10}(\text{chlorophyll})$ from CONTROL (orange), CCI-FULL (purple) and GC-FULL (green), computed at both OC-CCI (solid) and GlobColour (dashed) locations.

The OC-CCI and GlobColour observations are each provided with per-pixel uncertainty values. The OC-CCI uncertainties are based on comparisons to in situ data, whereas the GlobColour uncertainties are an output of the processing algorithm. For OC-CCI, an RMS uncertainty and a bias uncertainty are provided in \log_{10} units. For GlobColour, the uncertainty is expressed as a percentage of chlorophyll, and has been converted here to a $\log_{10}(\text{chlorophyll})$ RMS uncertainty, in order to try and obtain a comparable quantity to the OC-CCI RMS uncertainty. However, it should be noted that because of the very different processing methods, such a like-for-like comparison is not straightforward, and interpretation of these quantities should be treated with caution. A time-series of global mean $\log_{10}(\text{chlorophyll})$ for the entire reanalysis period (September 1997 to July 2012) is shown in Fig. 4.2.5, from the three model runs (CONTROL, CCI-FULL, GC-FULL) and the two observation products. Plus and minus the RMS uncertainties have been shaded for the OC-CCI and GlobColour observations (light blue and light red respectively), and the bias uncertainty for OC-CCI is shaded in black. This bias is very close to zero, and therefore difficult to distinguish in Fig. 4.2.5. Plotting plus or minus the RMS uncertainties provides a much larger uncertainty range, with the global mean for each product always lying well within this range for both products. The mean for CCI-FULL and GC-FULL also lies within these ranges, whilst CONTROL is always outside.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Neither the OC-CCI nor the GlobColour data show a clear global trend over the period, however in the latter years the GlobColour products show a much larger amount of variability, which is believed to be largely spurious due to changes in algorithms and sensor characteristics. It is therefore encouraging that this is not seen in the OC-CCI data. Similar differences are seen between CCI-FULL and GC-FULL, demonstrating the impact of this stability on model results. There are also differences in standard deviation between the OC-CCI and GlobColour products (not shown). The GlobColour data typically has a lower standard deviation, but this increases with time, particularly in the final few years. The OC-CCI standard deviation is much steadier, apart from a decrease following the launch of MERIS and MODIS in 2002.

	RMS error (μatm)	Correlation
CONTROL	42.67	0.65
CCI-FULL	41.70	0.66
GC-FULL	42.00	0.66

Table 4.2.2. RMS error and correlation against surface $f\text{CO}_2$ observations from SOCAT. Observations in water depths less than 150 m have been excluded.

Comparison to in situ $f\text{CO}_2$ from SOCAT, covering September 1997 to December 2011, is made in Table 4.2.2. A slight decrease in RMS error is seen due to the assimilation, with the lowest error seen for CCI-FULL, although overall the statistics remain similar for all three runs. Known variability in air-sea CO_2 fluxes is reproduced by the model, such as decreased outgassing in the Tropical Pacific during periods of El Niño, and the anomalous outgassing in the Tropical Atlantic in 2010 recently documented by Lefèvre *et al.* (2013). This latter coincides with an observed slowdown in the Atlantic meridional overturning circulation (McCarthy *et al.*, 2012), and such linkages can be explored further in phase 2 of the project.

4.2.4 Summary

OC-CCI and GlobColour products have been assimilated into a coupled physical-biogeochemical model, and the results compared to a control run. Similar results were seen in each case, with both products improving the model fit to independent in situ observations of chlorophyll, without degrading nutrient or carbon fields. Some differences are seen in model results between the two assimilative runs, both in the ecosystem and carbon cycle, and these will be explored further in phase 2 of the CMUG. At this stage it can be concluded that both data sets are suitable for data assimilation, but neither has yet been definitively shown to be superior to the other, and indeed the different results achieved with each are of scientific interest.

This study used the OC-CCI sinusoidally gridded chlorophyll NetCDF files. Overall, these were convenient to use and fit for purpose. A value for every point on the grid is explicitly stored in the OC-CCI files, whereas with the GlobColour files this information needs to be reconstructed by the user. This makes the OC-CCI files easier to work with and understand, although the trade-off is that data processing takes longer. A couple of technical issues were found regarding the dates during British Summer Time, and the coverage of the uncertainty values, and these have been reported to the OC-CCI helpdesk, who responded quickly and helpfully. Furthermore, the file format was slightly altered between V0 and V1 – this is okay

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

when working with test releases, but for a regular release cycle a fixed format would be required. It would also be desirable for single-sensor along-track products to be available alongside the merged daily-average products. The extra time information in these files, as well as the separate error classification for each sensor, would be of potential benefit for data assimilation. Another requirement is for case II waters products (or ideally a single product designed for use in both case I and case II waters), which could be used for validation of, and assimilation into, the Met Office shelf seas model.

The work reported here represents an initial assessment of the OC-CCI products. In phase 2 this will be extended to analyse in detail the 15-year reanalyses produced. Assessment will also be made of consistency between different marine essential climate variables (ECVs), for instance by assimilating other ECVs and comparing frontal positions of chlorophyll and sea surface temperature, and the alignment of chlorophyll patterns and mesoscale eddies. Other proposed applications of assimilating OC-CCI data include seasonal forecasting, investigating biophysical feedbacks between chlorophyll and temperature due to changes in light attenuation, and assimilating ocean colour data into the Met Office shelf seas biogeochemical model. Further assessment of the OC-CCI uncertainty values could also be made by developing the assimilation scheme to make more use of this information.

4.3 Sea surface height

In a previous version of this document (D3.1, v1, CMUG, 2012), some results of the use of a precursor of the CCI SSH products in the context of the impact of data assimilation and model assessment over the Mediterranean domain were presented by CNRM (Météo-France). The precursor is the SSALTO/DUACS SSH (Dibarboure et al, 2009), combining altimetric data from several satellites (Topex/Poseidon, ERS-1/2, Jason-1, Envisat and OSTM/Jason-2). The impact of data assimilation was first tested through a comparison of averaged SSH over the Mediterranean basin from COMBINE ocean reanalysis assimilating only in-situ observations with the ECMWF assimilation system (Balmaseda et al., 2010), and from GLORYS (1v1 and 2v1) ocean reanalyses performed with the MERCATOR assimilation system that assimilates the precursor dataset (Ferry et al., 2010). A very good agreement of the interannual variability was noted but also a reduction of the annual cycle with the COMBINE dataset compared with the GLORYS one by about 20%. The impact of SSH data assimilation was also tested through simulations of the Nemomed8 model (Sevault et al., 2009), over the 2002-2008 period, constrained with ERA-Interim at the ocean surface and either by COMBINE or GLORYS ocean reanalyses over an Atlantic buffer zone. The agreement is very high as stated by the correlation coefficients between the mean Mediterranean SSH simulated by the model through its free surface equation, and the mean Mediterranean SSH inferred from the reanalyses chosen to constrain the model in the Atlantic buffer zone. The correlation coefficient is slightly higher when the model is constrained by the GLORYS dataset that assimilates the satellite SSH precursor (0.95 compared with 0.84 with COMBINE). These results allow to detect an added value for the reanalysis that includes the assimilation of a satellite derived SSH, and to demonstrate the skill of the Mediterranean sea model at simulating the mean SSH over the domain when it is constrained with an atmospheric reanalysis.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Here we extend these results by including the CCI SSH products in a comparison to the precursor dataset and other datasets over the Mediterranean Sea, and by including simulations performed with a coupled regional climate model only constrained by an ocean reanalysis in an Atlantic buffer zone (and by an atmosphere reanalysis for its atmospheric component).

Besides the precursor SSALTO/DUACS SSH dataset, two other SSH datasets have been considered. The first one (Calafat and Jordà, 2011) is a reconstruction covering the 1950-2008 period using a reduced space optimal interpolation analysis. It combines long-term records from coastal and island tide gauges with spatial covariance structures determined from recent altimetric observations from AVISO. The second reconstruction (Meysignac et al., 2011) covers the period 1970-2006 and is based on a very similar method excepted the fact that the spatial covariance structures are inferred from model simulations (Nemomed8 and PROTHEUS) forced by atmospheric reanalyses. It is also based on tide gauge records and, as the previous, the data is corrected for the atmosphere effect (pressure and wind).

The coupled regional climate model is the so-called RCM4 model (Sevault et al., 2009) developed at CNRM and used for the MedCORDEX international simulation exercise. This is one of the first regional climate system model including a free coupling between a limited area regional atmospheric model (ALADIN-Climat v5, resolution of 50km and 31 vertical levels), a regional Mediterranean sea model (Nemomed8, adapted from the IPSL NEMO v2 ocean model, resolution of $1/8^\circ$ and 43 vertical levels) including a free surface equation, and a river routing component (TRIP, adapted from a model of the University of Tokyo, resolution of 0.5°). The coupling between the atmospheric and oceanic components is done through the OASIS coupler (v3 from CERFACS). The simulation used for the assessment covers the 1981-2010 period. Consistently with the MedCORDEX simulation protocol, the forcing consists in a spectral nudging of the atmospheric component towards ERA-Interim reanalysis and a nudging of the oceanic component towards an ocean reanalysis over an Atlantic buffer zone. To take into account the absence of satellite observation over part of the period, the reanalysis chosen for the ocean boundary layer (temperature, salinity and SSH) is the COMBINE one that only assimilates in-situ observations.

We show in Figure 4.3.1, SSH anomalies obtained from the different observational products and inferred from the model calculations, each being calculated by subtracting the mean value over their corresponding period of data availability (in particular 1980-2012 for the model and 1993-2010 for the satellite-derived products). We can first conclude from this comparison that the SSH inferred from the tide gauge measurements show trends that are slightly lower than the two satellite-derived SSH. However, the differences, lower than 2cm over the period, remain within the range of the regional mean sea level trend error given for the CCI SSH (SLCCI-ErrorReport-30 document, v1.1 dated from 9 April 2013, Table 5), i.e. lower than 5.4cm over the 18-year period ($<3\text{mm/yr}$). The interannual variability is in closer agreement between the two tide gauge-derived products than between one tide gauge-derived product and one satellite-derived product. This might be due to the fact that, for the two datasets based on in-situ observations, the tide gauges are more often located on the northern part of the Mediterranean coasts (this bias being attenuated by the reconstruction of the spatial variability through satellite observations or through model simulations). Figure 4.3.1 also reveals that the CCI SSH is very close to its precursor on average over the basin. We are indeed in a region where the corrections between the two products (in particular for the trends) are not

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



significant, higher latitudes regions being more impacted (see for instance the SLCCI-PVIR-31 document, v1.0 dated from 18 September 2013).

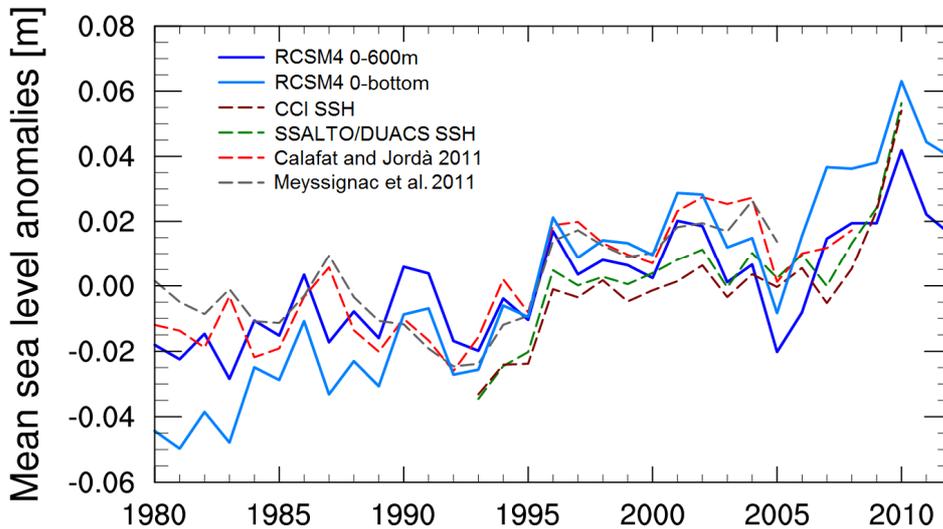


Figure 4.3.1: Averaged Anomalies of SSH over the Mediterranean basin relative to each period of data availability between 1980 and 2012 (in m). For the RCSM4 model (solid lines), there are diagnosed from the free surface model calculation (“dynamic” contribution) adding a mean thermosteric contribution integrated over the Mediterranean Sea surface, and over the upper 600m (dark blue) or the whole water column (light blue). The observations (dashed lines) are derived from two products using only satellite observation, SSALTO/DUACS SSH (green) and CCI SSH (brown), and from two products using tide gauges observations, Calafat and Jordà 2011 (red) and Meyssignac et al. 2011 (grey).

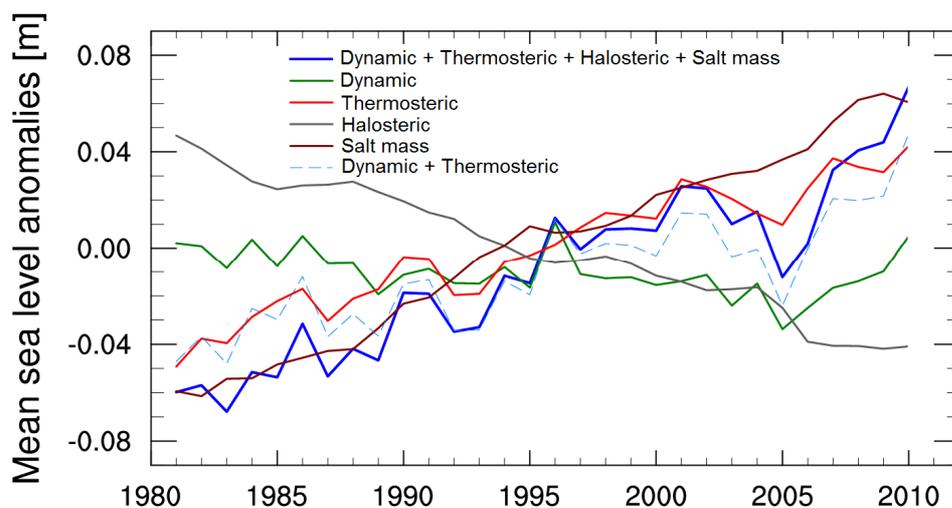


Figure 4.3.2: Contributions to the SSH anomalies relative to the 1981-2010 period, averaged over the Mediterranean domain and inferred from the RCSM4 model outputs (in m). The dynamic contribution (green) is diagnosed from the free surface model calculation, the



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

thermosteric contribution (red) is calculated from the simulated temperature profiles, the halosteric contribution (grey) and the salt mass change contribution (brown) are calculated from the simulated salinity profiles. The sum of the four contributions (dark blue) and the sum of the dynamic and thermosteric contributions (light blue), are also reproduced.

The SSH is calculated in the oceanic component of the model (NEMO model) with a free surface equation that takes into account two terms associated to physical processes explaining its change. The first one is a vertical integration of the convergence of the current (e.g. converging motion makes the level to rise) and is thus a dynamic term. The second results from the fresh water flux at the ocean surface that also implies a mass change and thus a sea level change. Even if these two terms are resulting from mass changes, by commodity we will designate in the following this calculated contribution to SSH changes as the “**Dynamic**” contribution, by reference to the first term of the model equation. An other term of dynamic origin, implying a mass change, is here neglected due to the so-called “**Boussinesq**” approximation that is done in this version of the NEMO model. This contribution comes locally from an advection of water with different density that can be significant when integrated over the Mediterranean basin due to the transport of salty Atlantic water through Gibraltar strait. This transport tends to balance the loss of fresh water at the annual and basin scale, surface evaporation exceeding precipitation at this scale (Jordà and Gomis, 2013). In the following, this mass contribution to sea level change will be called the “**Salt mass**” contribution. Even it is not calculated as part of the free surface equation, it can be diagnosed at the annual and Mediterranean basin scale from the evolution of the salinity profile calculated by the model. It is reproduced in Figure 4.3.2 over the 1981-2010 period where it is confronted with other terms contributing to sea level change at this scale, in particular with the “**Dynamic**” contribution. Two other terms that contribute physically to sea level change and are not included in the model equation are the two steric components of sea level change, the one due to temperature change, or “**Thermosteric**” contribution, and the one due to salinity change, or “**Halosteric**” contribution. These steric contributions result from the change of density implied by the temperature and salinity changes, all over the ocean column. This, keeping the mass constant, ultimately result in a sea level change due to the dilatation or to the contraction of the water volume. They can be diagnosed at the annual and Mediterranean basin scale with the simulated salinity and temperature profiles, and they are also reproduced in Figure 4.3.2.

This Figure helps at evaluating, with of course the limitation of model uncertainties, the relative contributions to sea level changes that are contained in the observational products presented in Figure 1, and in particular the CCI SSH. We can see that the model tends to reproduce opposite effects on mean sea level due to the two steric contributions, consistently with a positive trend in both temperature and salinity. More importantly, the diagnostics from the model outputs also show, in agreement with the finding of Jordà and Gomis (2013) based on in-situ observations, that the “**Salt mass**” contribution is positive and of the same order of magnitude than the “**Halosteric**” component but with an opposite sign. This is one of the reason that motivates the choice commonly done in the Mediterranean regional climate community, to neglect the contribution of salinity changes in the computation of modeled sea level change at the Mediterranean basin scale. An other reason is that the thermosteric

CMUG Deliverable

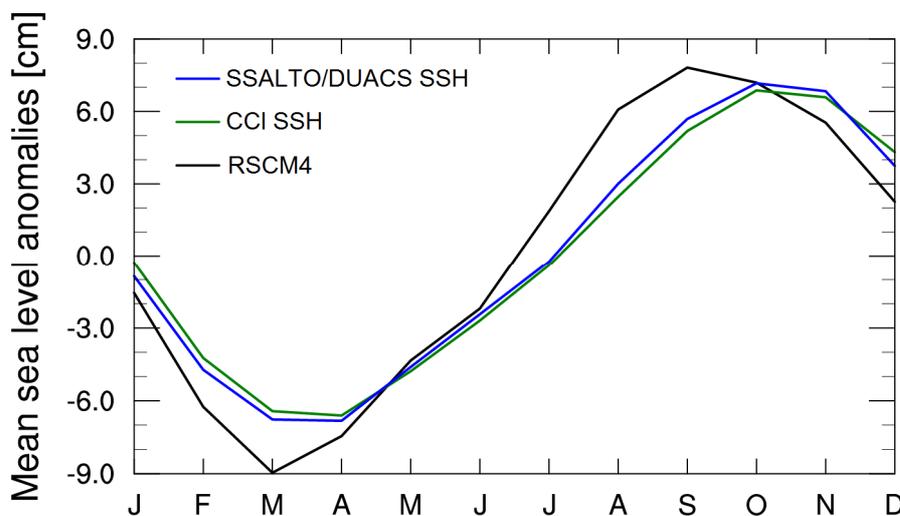
Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



contribution contains the effect of global mean warming that cannot be taken into account through the model equations.

We have reproduced in Figure 4.3.1 (and in Figure 4.3.2; dashed blue line) the diagnostic of SSH anomaly on the basis of this calculation thus obtained only by adding the Thermosteric and the Dynamic contributions calculated with the RSCM4 model outputs. We can see that the simulated anomaly (light blue curve) tends to exaggerate the SSH trend compared with all the observational datasets. This comes from an exaggerated warming that was diagnosed by comparing the model to ocean reanalyses (Sevault et al, 2009). Moreover, this warming mostly results from an unrealistic temperature drift below 600m, while the model is relatively close to observation in the upper part of the Mediterranean Sea. We obtain also a closer agreement between the simulated SSH trend and the observational products when the Thermosteric contribution is calculated only over the upper 600m of the Sea (Dark blue line in Figure 4.3.1). In this case, the SSH trend over the satellite period (1993-2010) is 2.95mm/yr for the model calculation, 2.49mm/yr for the precursor dataset and 2.74mm/yr for the CCI SSH. As far as interannual variability is concerned, as we can see in Figure 4.3.31, the model calculation is not significantly affected by the alternative choice on the Thermosteric component. This variability is not in phase with the two tide gauge products but the agreement seems better with the satellite products. Over the 18-year period, the root mean square difference of SSH anomalies between the model and the precursor is equal to 1.17cm, it is equal to 1.33cm between the model and the CCI SSH anomalies, when it is equal to 0.45cm between the two satellite products.

We can also see in Figure 4.3.3 that the mean seasonal cycle is well reproduced by the model when it is compared with those inferred from the two satellite products. The main discrepancies are a one-month shift in advance for the model and an exaggerated amplitude (by about 20%). This seasonal cycle would have been more importantly underestimated without accounting for the Thermosteric component (by about 50%; not shown).



**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Figure 4.3.3: Averaged seasonal cycle of SSH averaged over the Mediterranean Sea and over the period 1993-2010 (in cm). Results from the RCSM4 model (black) are compared with the two satellite products SSALTO/DUACS SSH (blue) and CCI SSH (green).

Note that the presented model diagnostics suffer from several limitations. First, The contribution of salinity change is not completely negligible as it can be noted in Figure 4.3.2 when comparing the sum of the Dynamic and Thermosteric contributions (dashed blue line) and the sum of the four contributions (dark blue line). In addition, we have not taken into account the impact of the boundary conditions with the nudging of 3D temperatures, 3D salinities and SSH towards the reanalysed fields in the Atlantic buffer zone. This forcing generates density and SSH gradients through Gibraltar Strait that, through the model equations, have also an impact on the simulated SSH at the Mediterranean basin scale and on the individual contributions (Steric and Dynamic contributions). Besides the impact of the global mean warming (and more generally of the global mean steric contribution) to sea level change, an other key contribution is also ignored in this calculation, namely the mass component due to glaciers and ice sheet melting. This contribution is not taken into account by the SSH nudging because, in order to cover a longest period, the MedCORDEX simulations are constrained by reanalyses that are not assimilating satellite SSH. This pleads for dedicated simulations with a nudging towards satellite products like the CCI SSH, over the Atlantic buffer zone, in order to have a better comparison to observed SSH in the inner Mediterranean basin. Some complementary work is ongoing at CNRM in this sense and in order to obtain a diagnostic on the total sea level that should be more directly comparable to observations.

However, most of these shortcomings can be avoided when the model simulated SSH anomalies are compared with the satellite products at the local scale. We have reproduced in Figure 4.3.4 the simulated and observed trends after removing the spatial and temporal averages of SSH over the Mediterranean domain for the 18-year period (1993-2010). The model field only includes the Dynamic contribution but it can be directly compared with observations because the uncertain contributions due to the mass flux at Gibraltar Strait, the global mean steric contribution and the fresh water flux at the sea surface averaged at the basin scale, are subtracted by construction. Their impact on the spatio-temporal variability compared with the mean of course remains at the local scale in the observations, but the comparison show that the agreement between the model and the observational satellite products is fairly good.

We particularly note in the model, consistently with the observations, negative trends in the western part of the basin and in the western part of the Ionian Sea, and positive trends in the eastern part of the Ionian Sea. The amplitudes of the trends however differ from the satellite observations even if they are often in the uncertainty range of the CCI SSH product (<3mm/yr). We can also note in this Figure that, here again, the trends are very similar between the CCI SSH and its precursor. Except at some places in the Aegean Sea, the larger differences are located in the Adriatic Sea where the simulated SSH trend is also closer to that of the CCI SSH product.

We can conclude from this confrontation between the CCI SSH product and its precursor, with tide gauge inferred reconstructions and with model estimates, that it is suitable for

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



regional climate studies over the Mediterranean basin, even at a scale of a few tens of kilometres. The confrontation with the model lets some open questions concerning the way to facilitate the comparison between the models and the observations. This question applies at the global scale since the state-of-the-art climate system models are not directly calculating the mass change contribution to sea level change due to glaciers or ice sheet melting, and to changing river discharge (a contribution that is not negligible according to recent findings). We have seen that the question is even more complicated to solve at the regional scale due to the fact that some terms can no longer be neglected (like those due to salinity changes), in particular in the case of the Mediterranean with the Gibraltar strait mass flux. However, the development of what can be called an “SSH observation simulator”, by analogy with the so-called satellite simulators, seems to be possible either at the global or at the regional scale.

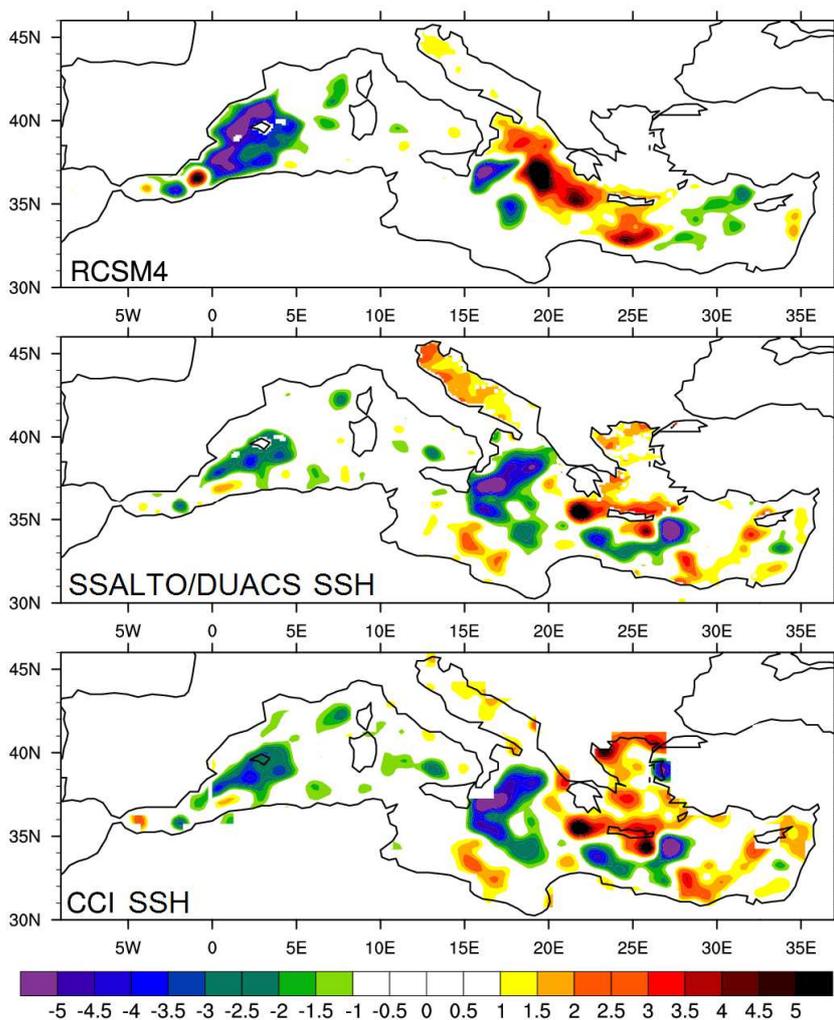


Figure 4.3.4: Trends of SSH anomalies over the 1993-2010 period (in mm/yr). Results from the RCSM4 model (top) are diagnosed from the free surface model calculation. They are compared with the two satellite products SSALTO/DUACS (middle) and SSH CCI SSH (bottom).

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

4.5 Ozone

In a previous version of this document (D3.1, v1, CMUG 2012), some results of the use of precursors of the CCI ozone products in the context of model assessment and data assimilation were presented by CNRM (Météo-France). Here we illustrate some new results obtained with CCI Ozone products consisting first of a comparison to precursor and independent datasets and then a revised model assessment. A wide ensemble of CCI ozone products have been used in the context of a CNES/Météo-France joint work (Prioul, 2013). Here we focus on examples of this work in applications using the merged L3 CCI ozone Total Columns (MERGED_TC) and monthly zonal mean Limb Profiles (MERGED_LP) products.

4.5.1 Confrontation with precursor and independent datasets

One independent dataset has been considered for the comparison. This is the so-called NIWA (National Institute of Water and Atmospheric Research, New Zealand) dataset that includes a total column ozone product consisting in assimilated satellite-based ozone measurements from four Total Ozone Mapping Spectrometer (TOMS) instruments, three different retrievals from the Global Ozone Monitoring Experiment (GOME) instruments, and data from four Solar Backscatter Ultra-Violet (SBUV) instruments (Bodeker et al., 2005). This data is complemented by a global monthly mean vertical ozone profile product combining satellite, balloons and ground-based measurements, spanning the period 1979 to 2007, and available from “Bodeker Scientific” (<http://www.bodekerscientific.com/>). The precursor dataset is the level 2 total column ozone derived from the EUMETSAT Infrared Atmospheric Sounding Interferometer (IASI) flying on board MetOp-A (since 2006) and MetOp-B (2012).

Figure 4.5.1. Histograms of total columns ozone cumulated in October 2008 for the CCI MERGED_TC product (black), IASI Eumetsat (green) and NIWA (violet). The values on the horizontal axis are in Dobson Units and in % on the vertical axis.

The histograms of Figure 4.5.1 for October 2008 show a relatively good agreement between the different total column ozone products. The agreement is however better between the CCI MERGED_TC and the NIWA product with the first being for the most part of the globe lower by about 2% compared with the second, within the range of the CCI product total uncertainty (confirmed for the whole period between 2000 and 2012; not shown). The main discrepancy occurs for the lowest values (near 140 DU) at high austral latitudes (not shown). This result is positive for the CCI product since the NIWA dataset also includes in-situ observations. The disagreement with the IASI dataset concerns mainly near polar latitudes (not shown) and merits further investigation.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

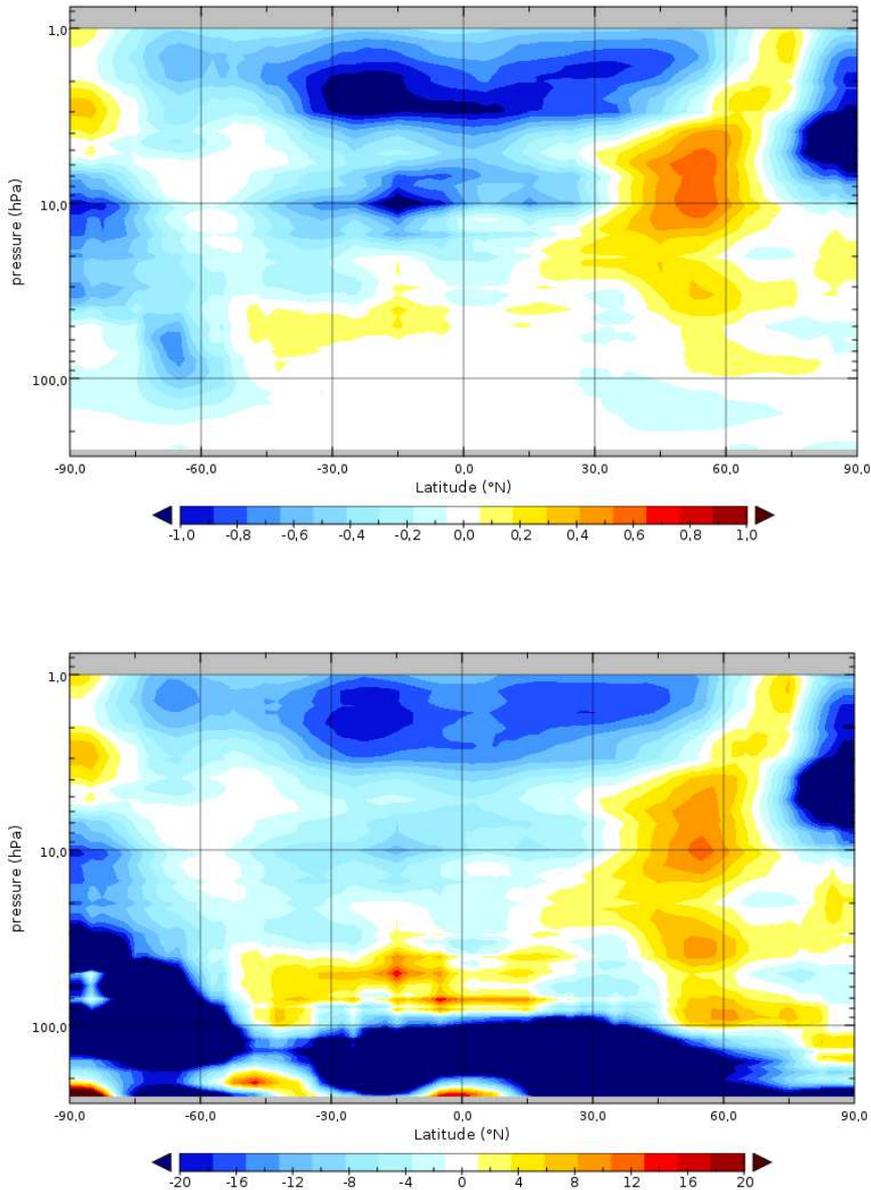


Figure 4.5.2. Vertical profiles of zonal mean differences of ozone mixing ratios in ppmv (top) and in % (bottom), between NIWA and CCI MERGED_TC LP products, averaged over October 2007.

The CCI and NIWA ozone profiles are compared in Figure 4.5.2 for October 2007. This is only one example of comparison that should be more systematically performed over the period of data availability (ending in 2007 for NIWA). The differences are generally below lower than the potential uncertainty of the MERGED_LP product given in the CCI ozone

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

ATBD document (v2 dated 18 March 2014, Table 21). This last should be within the +/-15% range for the seasonal cycle and interannual variability of the ozone mixing ratio. In Figure 4.5.2, when the relative error is greater than this uncertainty estimate, this mainly corresponds to low values of ozone mixing ratio (lower than 1ppmv) in particular below 50hPa and within the sub-polar ozone hole in the southern hemisphere. There are however a few regions with greater ozone content where the differences are higher than the uncertainty range. This concerns the tropical stratosphere above 3hPa and the sub-polar stratosphere above 10hPa in the Northern hemisphere. Attributing the main errors to either one or the other dataset in these regions is outside the scope of this report.

4.5.2 Model assessment

The new version of the CNRM climate model including an on-line representation of stratospheric chemistry processes, the so-called CNRM-CCM model, was assessed with the precursors as presented in D3.1 v1 (CMUG, 2012). Here we illustrate its assessment using the CCI ozone MERGED_TC and MERGED_LP products. The model resolution is of about 2.8° by 2.8° horizontally and the atmosphere is divided in 60 vertical levels. The CNRM-CCM simulation (referred to as CNRM-CM in the following) follows the protocol of the CCMVal project and is the same than the one presented and analysed in Michou et al. (2011) and in D3.1 v1 (CMUG, 2012). In another simulation, called CNRM-CCM nudged, the same model is nudged towards the Era-Interim reanalysis (temperature, humidity and wind components variables). In addition, we also include in the comparison some outputs from two simulations performed with the MOCAGE chemistry-transport model developed at Météo-France at a low-resolution version of 2° by 2° on the horizontal and 60 vertical levels. This version doesn't include the complex stratospheric chemical scheme that is currently used for climate-chemistry interaction studies, but a simple linearized ozone scheme (Cariolle and Tesseydre, 2007) in order to test the impact of ozone data assimilation (Pajot et al., 2012). In the first simulation (MOCAGE), the model is integrated without data assimilation. The assimilation experiment (MOCAGE assimilated) is performed with the Valentina data assimilation system using the incremental 4D-Var method developed at the Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS). The assimilated data consists in the IASI total ozone columns combined with the Microwave Limb Sounder (MLS) ozone profiles (see also D3.1 v1, CMUG (2012)).

The histograms of Figure 4.5.3 for October 2008 show that the two CNRM-CM simulations underestimate the ozone total columns compared with the ozone CCI MERGED_TC product. This is consistent with previous analysis of the CNRM-CM simulation compared with other datasets. The nudging not always results in an improvement of the distribution. This implies that part of the discrepancy comes from the chemical schemes. The distribution given by MOCAGE and the ozone linear scheme is far from the satellite-derived product but here the assimilation is efficient at improving the agreement that is now very good, while the assimilated data are not included in the CCI MERGED_TC product.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

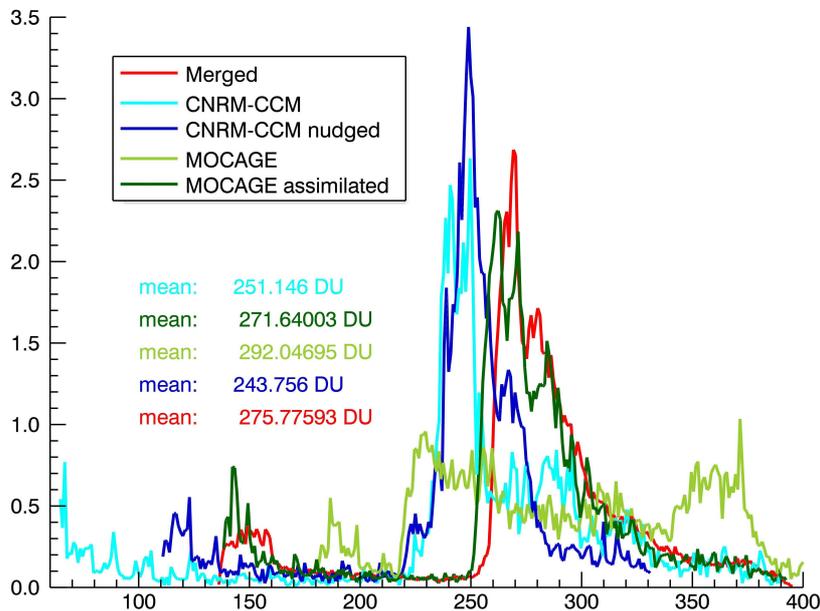


Figure 4.5.3. Histograms of total column ozone accumulated in October 2008 for CNRM-CCM (light blue), CNRM-CCM nudged (dark blue), MOCAGE (light green), MOCAGE assimilated (dark green) and the CCI MERGED_TC product for the 90°S - 84°N latitude band. The values on the horizontal axis are in Dobson Units and in % on the vertical axis.

We have reproduced in Figure 4.5.4 the annual cycle over the year 2008 of monthly mean ozone mixing ratio at 3 hPa, 10 hPa and 50 hPa averaged over different latitudinal bands, calculated from the two CNRM-CCM simulations and from the CCI MERGED_LP product. At 3 hPa, the model simulations reproduce observed variations, although underestimating their amplitudes. These amplitudes are improved by the nudging, showing that part of the discrepancy comes from the representation of the climatic variables at this level. Michou et al. (2011) have shown that CNRM-CCM simulated too low ozone mixing ratios compared with UARS observations (HALOE) and attributed it in part to too warm temperatures. This is consistent with the improved results when the temperatures are nudged towards ERA-Interim even if the ozone mixing ratios remain lower than those of the CCI product. We can however also note that, at this level, in the tropical and at higher northern latitudes, the MERGED_LP mixing ratio are higher than the NIWA product (close to 1ppmv in October 2008 as seen above). This puts in doubt the amplitude of the disagreement and of the error estimate accompanying the CCI dataset reproduced in the figure. At 10 hPa, close to the stratospheric maxima, the amplitude and variation of the annual cycle are fairly well simulated by CNRM-CCM in particular in the nudged simulation. For this last simulation, the highest differences remain in a 15% uncertainty range. At 50 hPa in the high latitude Southern Hemisphere, the ozone mixing ratio is dominated by polar ozone depletion. Figure 4.5.4 shows that the CNRM-CMM model overestimates this depletion due to a temperature bias (Michou et al, 2011) corrected by the nudging. The degradation of the results with the nudging at the tropical latitudes is questionable.

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

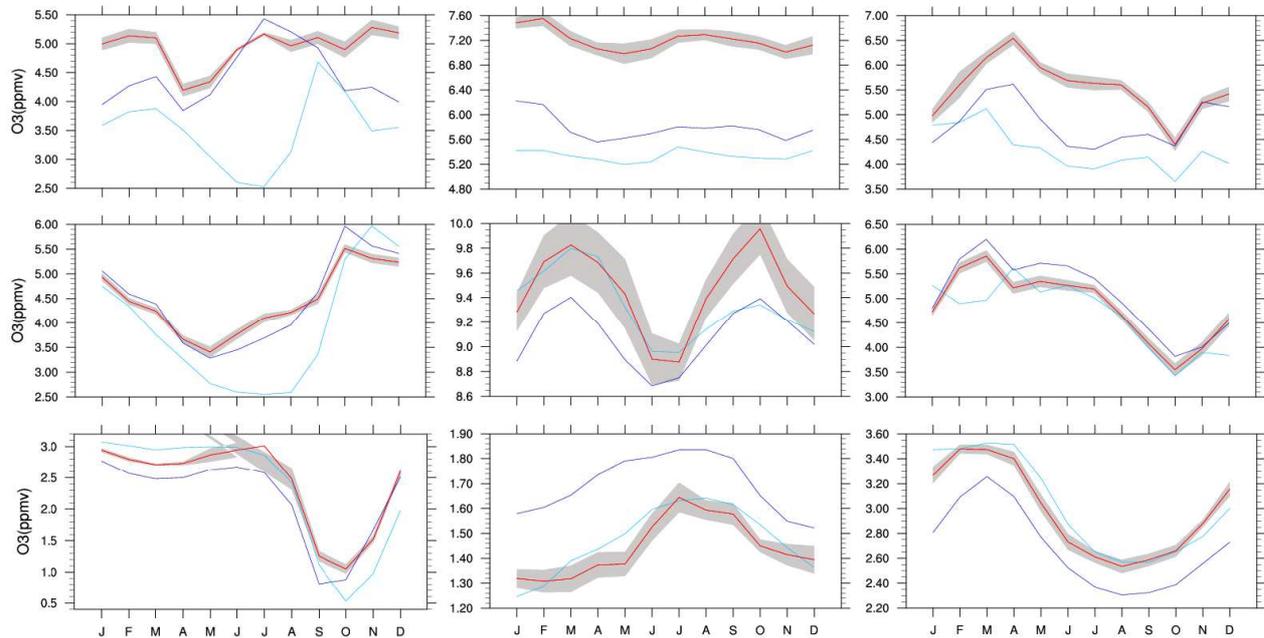


Figure 4.5.4. Annual cycle over the year 2008 of monthly mean ozone mixing ratio (ppmv) at 3 hPa (top line), 10 hPa (middle line) and 50 hPa (bottom line), averaged over the 90S - 60N 60S (first column), 30S - 30N (second column) and 60N - 90N (third column) latitudinal bands. The light blue line corresponds to CNRM-CCM, the dark blue line to CNRM-CCM nudged and the red line to CCI MERGED_LP product associated to an uncertainty estimate (grey shading).

One main lesson from this short analysis is that the ozone CCI-products need a confrontation with all available observational products, in particular those such as the NIWA products that include in-situ observations. The uncertainty range reproduced on the figures seems to be underestimated at least in some regions like the 3hPa level. The confrontation with model simulations, in particular those that are nudged towards, or assimilating other observational products, also confirm that the CCI ozone products are suitable for model intercomparison and data assimilation.

4.4 Clouds

The International Cloud Climatology (ISCCP) data set has been used as a precursor for cloud amount and various other data sets (e.g. MODIS and CloudSat) for other cloud parameters. Here we use the prototype CCI clouds data and show some initial comparisons with the most recent Hadley Centre climate model, HadGEM2.

The CCI cloud products include a merged product – using data from all available sensors – and individual sensor products using AATSR, MODIS and AVHRR. The prototype data set produced in Phase 1 of the CCI covers the three years 2007 – 2009. It follows the framework and philosophy of the GEWEX cloud assessment and has therefore not been specifically

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

designed for climate model evaluation, although there are plans to do more in this regard for Phase 2.

Comparisons with other similar cloud products suggest that the CCI cloud data sets generally underestimate cloud amount over the global oceans at all latitudes. In addition, high-level cloud is underestimated over land over both the tropics and mid-latitudes, while low-level cloud is underestimated over the key areas of marine stratocumulus in the sub-tropics. There is also a severe underestimate of thin cirrus cloud – this suggests that modifications to the cloud detection algorithm may be needed, as in principle the channels available from instruments such as AATSR should allow better cirrus detection compared to ISCCP, for example. These findings have been corroborated by the Cloud CCI climate research group. **Taken together, examination of the current version of the CCI clouds data suggests that it does not at present provide added value compared to ISCCP (or similar precursors) in respect of cloud coverage. The current version of the data sets contains discontinuities in all of the cloud amount products, making comparisons with models unsatisfactory.** For this reason we focus here on other parameters of interest, specifically cloud liquid water and cloud droplet effective radius.

In common with other retrievals using visible and infra-red measurements (ISCCP, MODIS, ATSR) cloud liquid water in the Cloud CCI is not directly retrieved but is instead calculated from the retrieved cloud optical depth and the cloud droplet effective radius. This means that such estimates may be substantially different from retrievals using either passive microwave observations (SSM/I) or space-borne radar (CloudSat).

Figure 4.4.1 compares the CCI cloud liquid water path (LWP) with retrievals from MODIS, ATSR, CloudSat and SSM/I. Note that this is the merged CCI product. Differences between CCI and the two VIS/IR retrievals (MODIS, ATSR) are positive everywhere and the geographical patterns of these differences are very similar. Comparisons with CloudSat and SSM/I are quite different to this, especially in the tropics and sub-tropics, e.g. the underestimate with respect to CloudSat in the main areas of marine stratocumulus cloud. A striking result is the closeness of the CCI cloud LWP to that derived from SSM/I, which is a clear difference with the precursor ISCCP data set (also shown for comparison). Given the very different nature of the sensors it is surprising to find that the CCI data is closer to the microwave retrievals than it is to those obtained from VIS/IR instruments which are components of this merged product.

These differences between the observed estimates of LWP clearly have an impact on their use for climate model evaluation (Figure 4.4.2). The CCI data suggests that HadGEM2 underestimates LWP over land everywhere as well as over the tropical and subtropical oceans. Overestimates are seen over the Southern Ocean and mid-latitude Atlantic and Pacific. Given the comparisons shown in Fig. 4.4.1 this often contrasts with the evaluation using the other LWP products. In general the comparisons are most consistent over the mid-latitude oceans. As expected from Fig. 4.4.1, the comparisons with SSM/I are the most similar over the global oceans.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Figure 4.4.3 shows comparisons of the HadGEM2 simulation of liquid water droplet effective radius (r_{eff}) with retrievals from CCI clouds, together with those from MODIS and ATSR. Note that this quantity has been specifically diagnosed in the model to match the satellite retrievals, i.e. it is not the standard r_{eff} variable. Here the quite different interpretation of the model's simulation according to the data set being used for the evaluation is, if anything, even more apparent. The CCI product indicates far larger drop sizes than either MODIS or ATSR and this is then reflected in the model comparison, which suggests large underestimates of r_{eff} in many regions, over both land and ocean. We can also make use of the CCI's individual sensor products for the evaluation and compare the results to the standard MODIS and ATSR products (Fig. 4.4.4). This suggests that the CCI's large droplet sizes is primarily a function of the retrieval algorithm rather than depending on the sensor used. The final merged value will of course depend on how these sensors are combined – note that the large droplet sizes are also evident with the CCI's AVHRR product (also shown in Fig. 4.4.4), which is also a component of the merged product.

To conclude, it is probably fair to say that we are unable at present to determine the value of the CCI clouds products for climate model evaluation. Clearly more work needs to be done: (a) within the CCI team itself to understand the precise reasons for the differences with other well-established data sets and to improve the CCI products (e.g. the cirrus detection); and (b) in conjunction with the modelling community to ensure the CCI clouds project delivers products that are both useful and add value to those currently being used for model evaluation and development studies.



Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

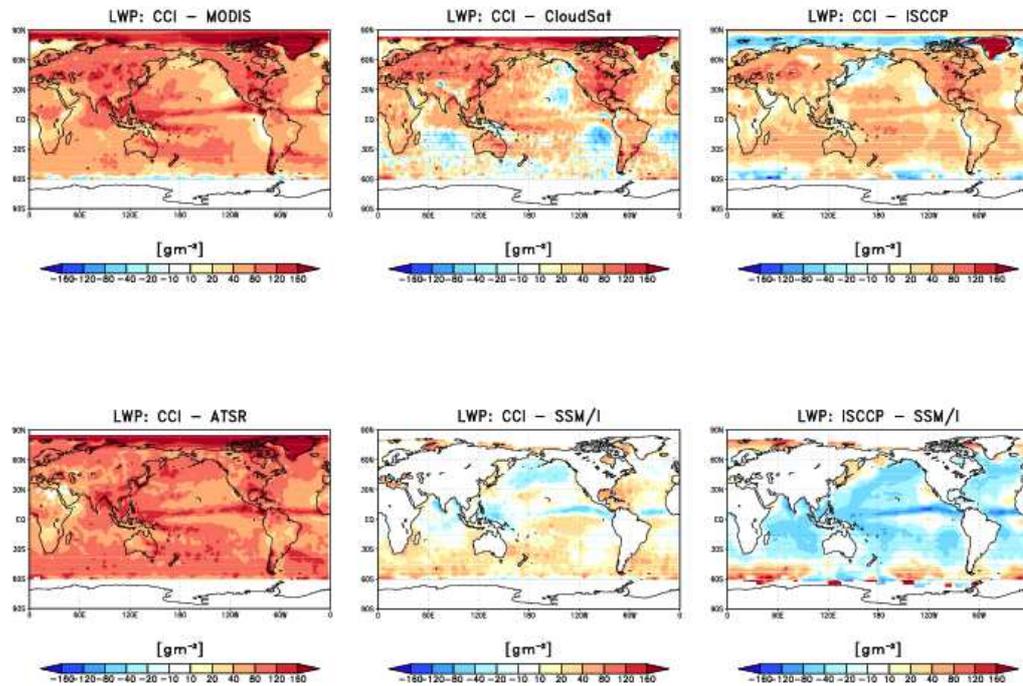


Figure 4.4.1: Comparison of the CCI merged cloud liquid water path product with a range of other data sets. Also shown is the comparison of the ISCCP precursor with the SSM/I product (which is the most similar to the CCI). Data are seasonal averages for June, July, August 2007.

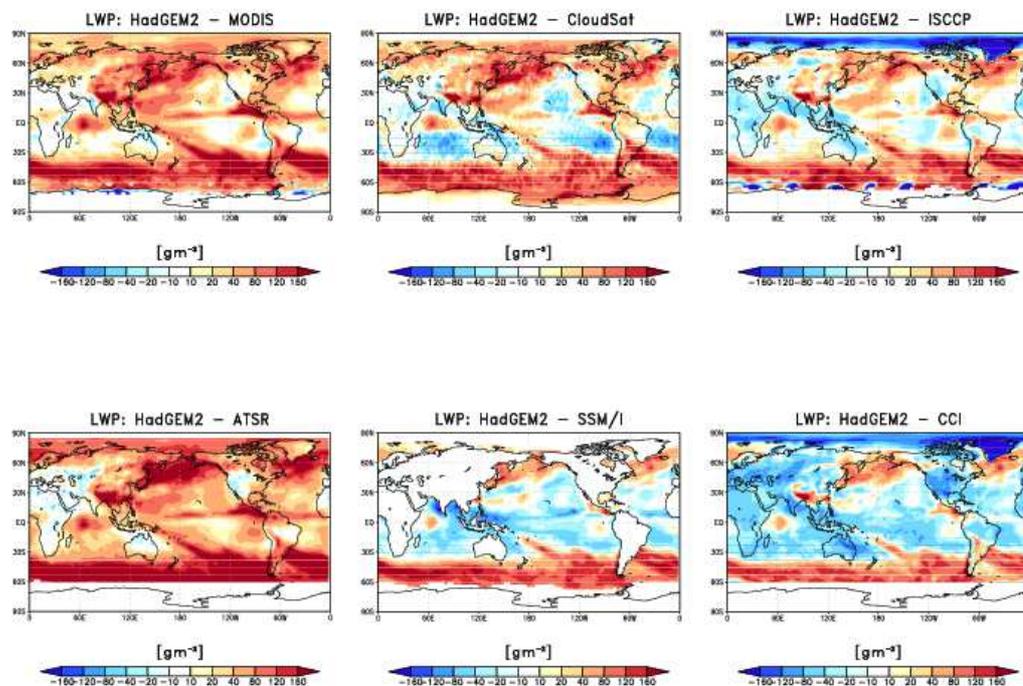


Figure 4.4.2: Evaluation of cloud liquid water path in HadGEM2 for JJA 2007 using a range of data sets including CCI. HadGEM2 values are diagnosed from atmosphere-only simulations driven by observed SSTs.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

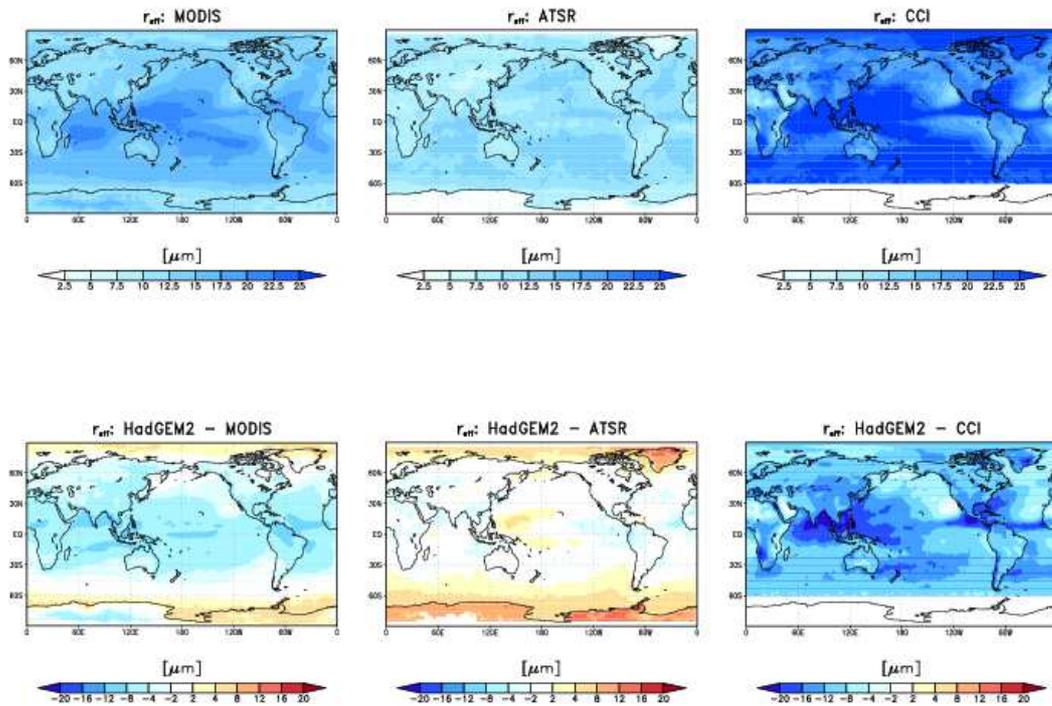


Figure 4.4.3: (Top) Seasonal mean values of cloud droplet effective radius from MODIS, ATSR and CCI for JJA 2007. (Bottom) Comparisons with HadGEM2.

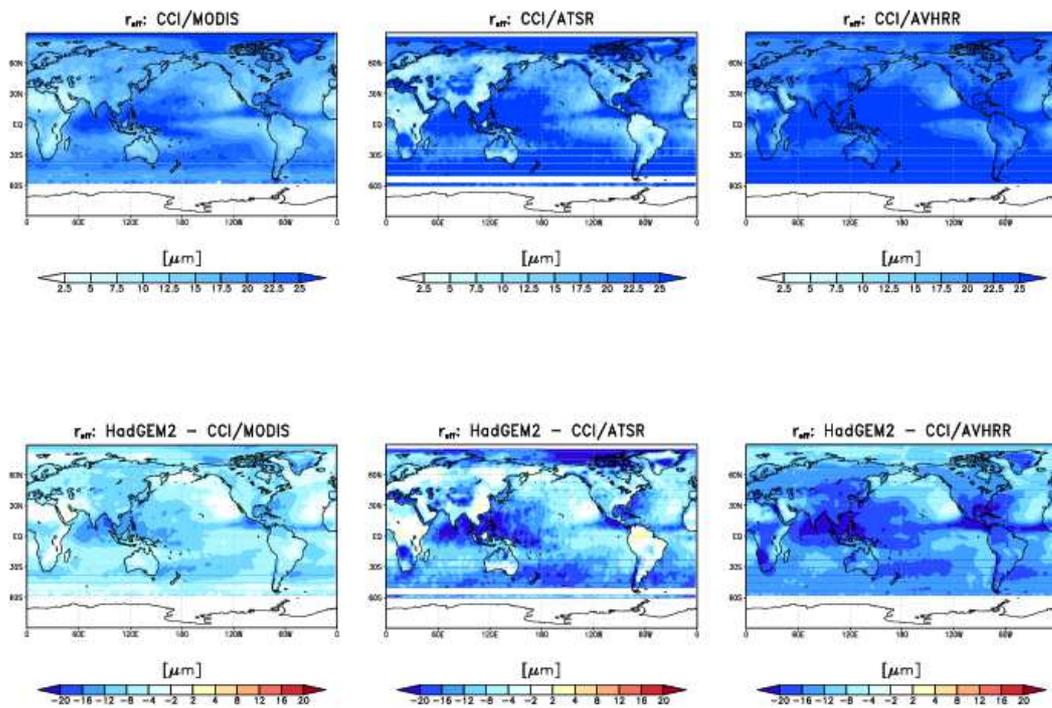


Figure 4.4.4: (Top) CCI individual sensor effective droplet radius products for MODIS, ATSR and AVHRR for JJA 2007. (Bottom) Comparisons with HadGEM2.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

4.5 Ozone

In a previous version of this document (D3.1, v1, CMUG 2012), some results of the use of precursors of the CCI ozone products in the context of model assessment and data assimilation were presented by CNRM (Météo-France). Here we illustrate some new results obtained with CCI Ozone products consisting first of a comparison to precursor and independent datasets and then a revised model assessment. A wide ensemble of CCI ozone products have been used in the context of a CNES/Météo-France joint work (Prioul, 2013). Here we focus on examples of this work in applications using the merged L3 CCI ozone Total Columns (MERGED_TC) and monthly zonal mean Limb Profiles (MERGED_LP) products.

4.5.1 Confrontation with precursor and independent datasets

One independent dataset has been considered for the comparison. This is the so-called NIWA (National Institute of Water and Atmospheric Research, New Zealand) dataset that includes a total column ozone product consisting in assimilated satellite-based ozone measurements from four Total Ozone Mapping Spectrometer (TOMS) instruments, three different retrievals from the Global Ozone Monitoring Experiment (GOME) instruments, and data from four Solar Backscatter Ultra-Violet (SBUV) instruments (Bodeker et al., 2005). This data is complemented by a global monthly mean vertical ozone profile product combining satellite, balloons and ground-based measurements, spanning the period 1979 to 2007, and available from “Bodeker Scientific” (<http://www.bodekerscientific.com/>). The precursor dataset is the level 2 total column ozone derived from the EUMETSAT Infrared Atmospheric Sounding Interferometer (IASI) flying on board MetOp-A (since 2006) and MetOp-B (2012).

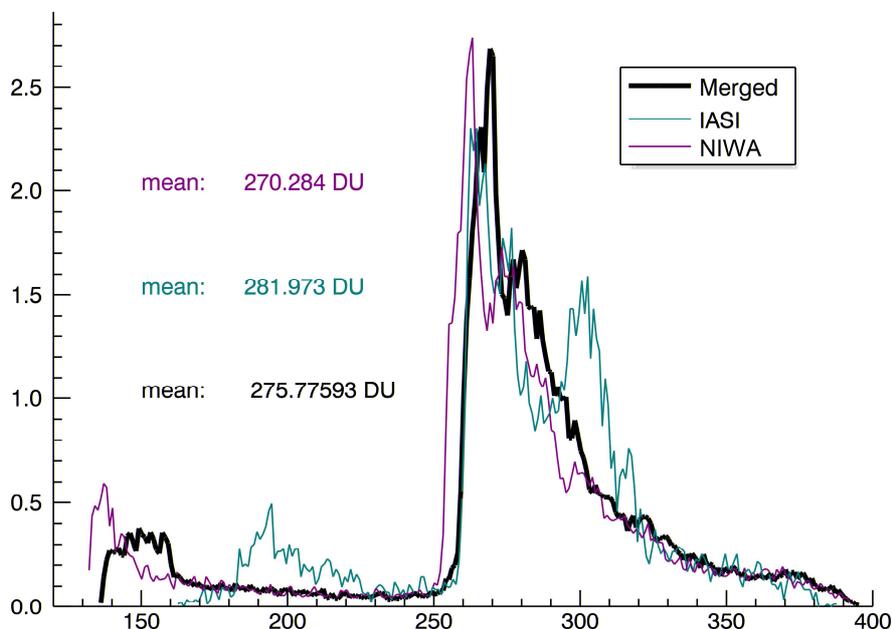


Figure 4.5.1. Histograms of total columns ozone cumulated in October 2008 for the CCI MERGED_TC product (black), IASI Eumetsat (green) and NIWA (violet). The values on the horizontal axis are in Dobson Units and in % on the vertical axis.

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



The histograms of Figure 4.5.1 for October 2008 show a relatively good agreement between the different total column ozone products. The agreement is however better between the CCI MERGED_TC and the NIWA product with the first being for the most part of the globe lower by about 2% compared with the second, within the range of the CCI product total uncertainty (confirmed for the whole period between 2000 and 2012; not shown). The main discrepancy occurs for the lowest values (near 140 DU) at high austral latitudes (not shown). This result is positive for the CCI product since the NIWA dataset also includes in-situ observations. The disagreement with the IASI dataset concerns mainly near polar latitudes (not shown) and merits further investigation.

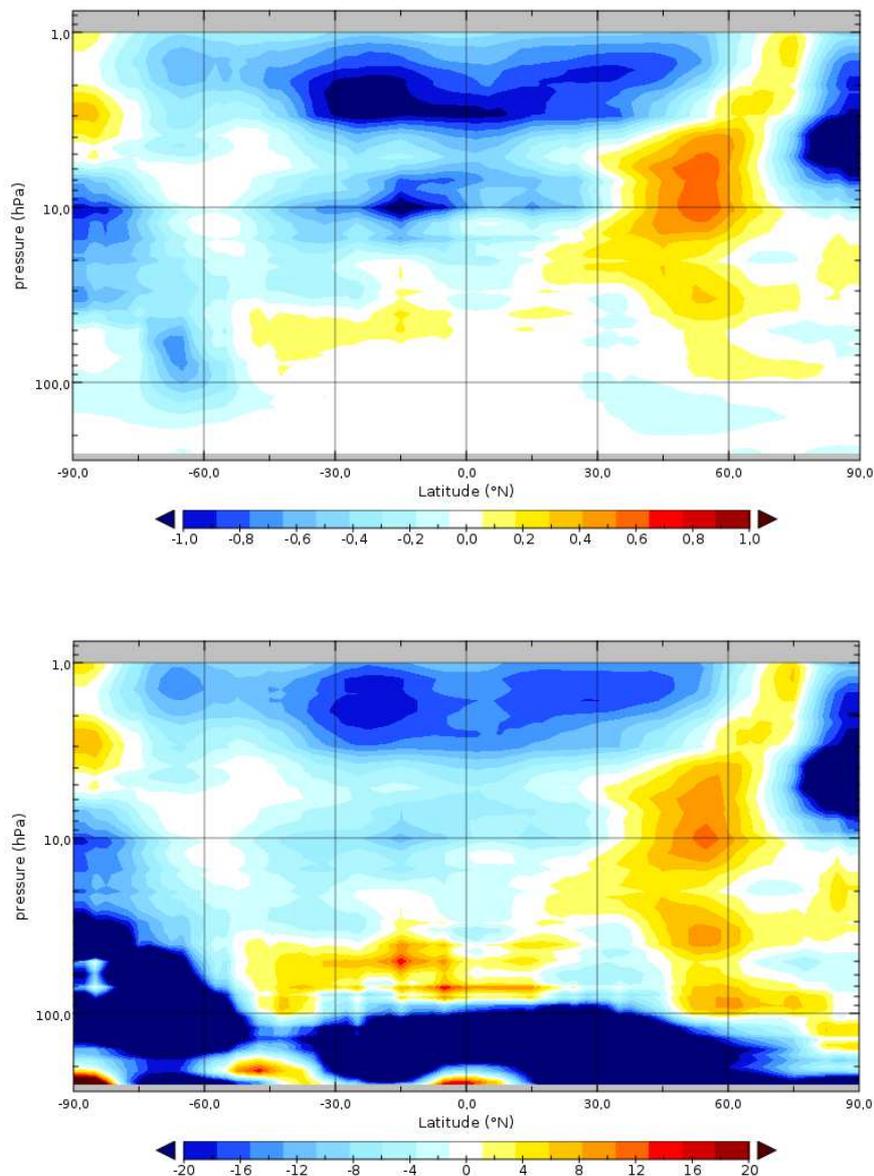


Figure 4.5.2. Vertical profiles of zonal mean differences of ozone mixing ratios in ppmv (top) and in % (bottom), between NIWA and CCI MERGED_TC products, averaged over October 2007.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

The CCI and NIWA ozone profiles are compared in Figure 4.5.2 for October 2007. This is only one example of comparison that should be more systematically performed over the period of data availability (ending in 2007 for NIWA). The differences are generally below than the potential uncertainty of the MERGED_LP product given in the CCI ozone ATBD document (v2 dated 18 March 2014, Table 21). This last should be within the +/-15% range for the seasonal cycle and interannual variability of the ozone mixing ratio. In Figure 4.5.2, when the relative error is greater than this uncertainty estimate, this mainly corresponds to low values of ozone mixing ratio (lower than 1ppmv) in particular below 50hPa and within the sub-polar ozone hole in the southern hemisphere. There are however a few regions with greater ozone content where the differences are higher than the uncertainty range. This concerns the tropical stratosphere above 3hPa and the sub-polar stratosphere above 10hPa in the Northern hemisphere. Attributing the main errors to either one or the other dataset in these regions is outside the scope of this report.

4.5.2 Model assessment

The new version of the CNRM climate model including an on-line representation of stratospheric chemistry processes, the so-called CNRM-CCM model, was assessed with the precursors as presented in D3.1 v1 (CMUG, 2012). Here we illustrate its assessment using the CCI ozone MERGED_TC and MERGED_LP products. The model resolution is of about 2.8° by 2.8° horizontally and the atmosphere is divided in 60 vertical levels. The CNRM-CCM simulation (referred to as CNRM-CM in the following) follows the protocol of the CCMVal project and is the same than the one presented and analysed in Michou et al. (2011) and in D3.1 v1 (CMUG, 2012). In another simulation, called CNRM-CCM nudged, the same model is nudged towards the Era-Interim reanalysis (temperature, humidity and wind components variables). In addition, we also include in the comparison some outputs from two simulations performed with the MOCAGE chemistry-transport model developed at Météo-France at a low-resolution version of 2° by 2° on the horizontal and 60 vertical levels. This version doesn't include the complex stratospheric chemical scheme that is currently used for climate-chemistry interaction studies, but a simple linearized ozone scheme (Cariolle and Tesseydre, 2007) in order to test the impact of ozone data assimilation (Pajot et al., 2012). In the first simulation (MOCAGE), the model is integrated without data assimilation. The assimilation experiment (MOCAGE assimilated) is performed with the Valentina data assimilation system using the incremental 4D-Var method developed at the Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS). The assimilated data consists in the IASI total ozone columns combined with the Microwave Limb Sounder (MLS) ozone profiles (see also D3.1 v1, CMUG (2012)).

The histograms of Figure 4.5.3 for October 2008 show that the two CNRM-CM simulations underestimate the ozone total columns compared with the ozone CCI MERGED_TC product. This is consistent with previous analysis of the CNRM-CM simulation compared with other datasets. The nudging not always results in an improvement of the distribution. This implies that part of the discrepancy comes from the chemical schemes. The distribution given by MOCAGE and the ozone linear scheme is far from the satellite-derived product but here the assimilation is efficient at improving the agreement that is now very good, while the assimilated data are not included in the CCI MERGED_TC product.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

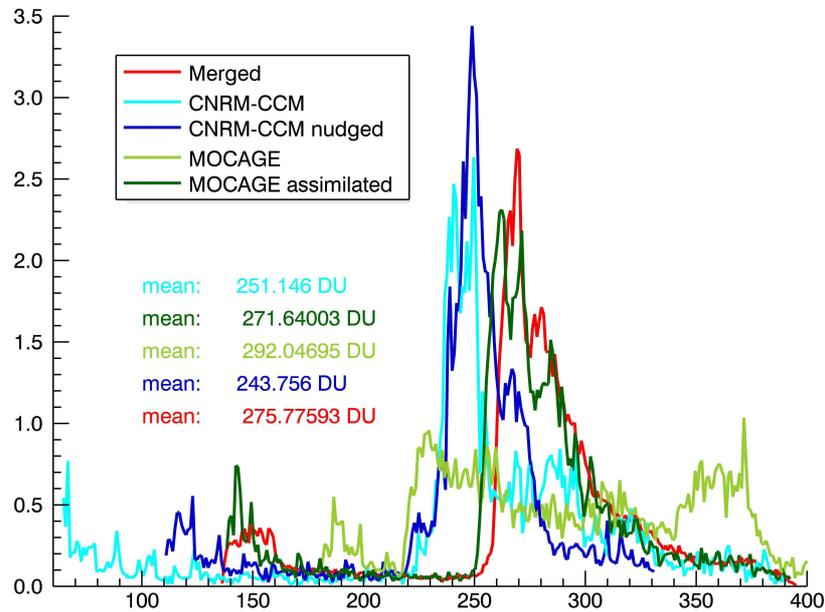


Figure 4.5.3. Histograms of total column ozone accumulated in October 2008 for CNRM-CCM (light blue), CNRM-CCM nudged (dark blue), MOCAGE (light green), MOCAGE assimilated (dark green) and the CCI MERGED_TC product for the 90°S - 84°N latitude band. The values on the horizontal axis are in Dobson Units and in % on the vertical axis.

We have reproduced in Figure 4.5.4 the annual cycle over the year 2008 of monthly mean ozone mixing ratio at 3 hPa, 10 hPa and 50 hPa averaged over different latitudinal bands, calculated from the two CNRM-CCM simulations and from the CCI MERGED_LP product. At 3 hPa, the model simulations reproduce observed variations, although underestimating their amplitudes. These amplitudes are improved by the nudging, showing that part of the discrepancy comes from the representation of the climatic variables at this level. Michou *et al.* (2011) have shown that CNRM-CCM simulated too low ozone mixing ratios compared with UARS observations (HALOE) and attributed it in part to too warm temperatures. This is consistent with the improved results when the temperatures are nudged towards ERA-Interim even if the ozone mixing ratios remain lower than those of the CCI product. We can however also note that, at this level, in the tropical and at higher northern latitudes, the MERGED_LP mixing ratio are higher than the NIWA product (close to 1ppmv in October 2008 as seen above). This puts in doubt the amplitude of the disagreement and of the error estimate accompanying the CCI dataset reproduced in the figure. At 10 hPa, close to the stratospheric maxima, the amplitude and variation of the annual cycle are fairly well simulated by CNRM-CCM in particular in the nudged simulation. For this last simulation, the highest differences remain in a 15% uncertainty range. At 50 hPa in the high latitude Southern Hemisphere, the ozone mixing ratio is dominated by polar ozone depletion. Figure 4.5.4 shows that the CNRM-CMM model overestimates this depletion due to a temperature bias (Michou *et al.*, 2011) corrected by the nudging. The degradation of the results with the nudging at the tropical latitudes is questionable.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

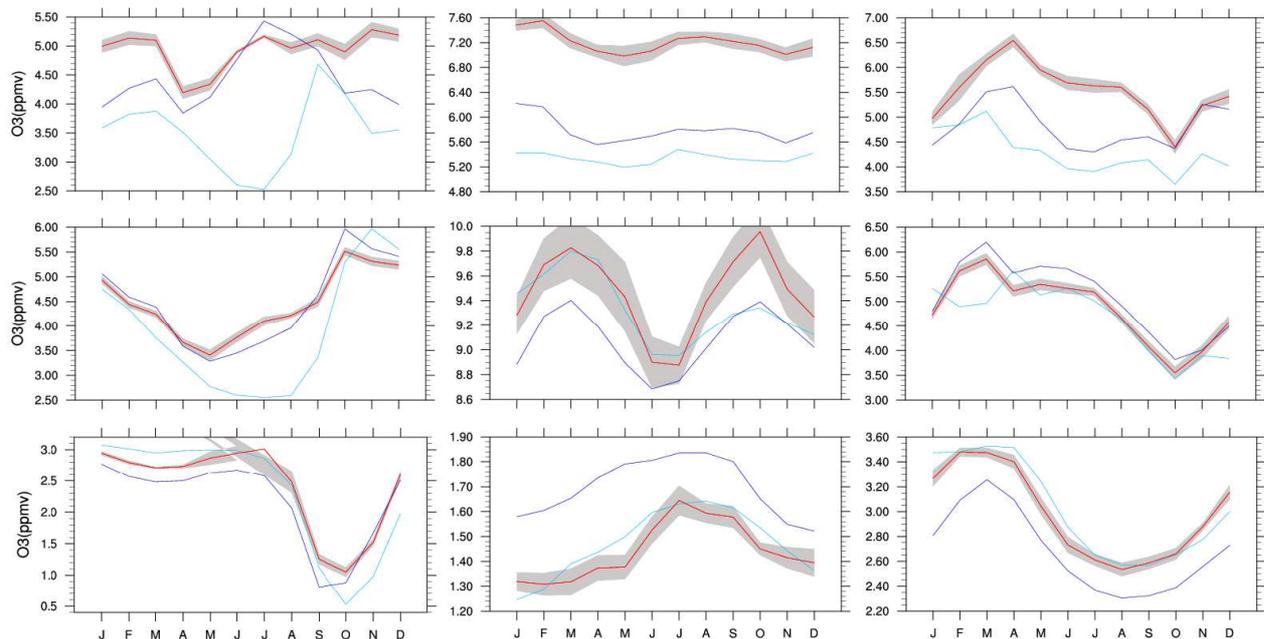


Figure 4.5.4. Annual cycle over the year 2008 of monthly mean ozone mixing ratio (ppmv) at 3 hPa (top line), 10 hPa (middle line) and 50 hPa (bottom line), averaged over the 90S - 60N (first column), 30S - 30N (second column) and 60N - 90N (third column) latitudinal bands. The light blue line corresponds to CNRM-CCM, the dark blue line to CNRM-CCM nudged and the red line to CCI MERGED_LP product associated to an uncertainty estimate (grey shading).

One main lesson from this short analysis is that the ozone CCI-products need a confrontation with all available observational products, in particular those such as the NIWA products that include in-situ observations. The uncertainty range reproduced on the figures seems to be underestimated at least in some regions like the 3hPa level. The confrontation with model simulations, in particular those that are nudged towards, or assimilating other observational products, also confirm that the CCI ozone products are suitable for model intercomparison and data assimilation.

4.6 Greenhouse Gases

GHG_CCI has two lines of production, one for the columnar mole fraction of Carbon Dioxide (X_{CO_2}) and one for the columnar mole fraction of Methane (X_{CH_4}) from two instruments (ENVISAT/SCIAMACHY and TANSO/GOSAT). The two retrievals will be simply referred to as carbon dioxide and methane. For each molecule, several algorithms were considered during Phase 1. A Round Robin (RR) exercise was carried out by GHG_CCI during Phase 1 aiming at selecting the best available algorithms for both variables. Table 4.6.1 lists the algorithms included in the GHG Climate Research Package (CRP) resulting from the RR exercise.

The Climate Monitoring Facility (CMF, see Box A, and CMUG, 2013b, c, d) developed at ECMWF and its database (CMFDb) were used to compare level 3 monthly mean statistics of X_{CO_2} and X_{CH_4} observations retrieved with the algorithms listed in Table 4.6.1 with averages of their model equivalent obtained from the MACC atmospheric reanalysis.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1



Variable	Algorithm (version)	Sensor	Availability	Provider
XCO2	BESD (2.0)	SCIAMACHY	Aug 2002 - Mar 2012	IUP
	OCFP (4.0)	GOSAT	Jun 2009 - Jan 2012	Uni. of Leicester
	SRFP (2.1)	GOSAT	Jun 2009 - Sep 2012	SRON
XCH4	WFMD (3.3)	SCIAMACHY	Jan 2003 - Sep 2012	IUP
	IMAP (6.0)	SCIAMACHY	Jan 2003 - Apr 2012	SRON
	SRFP (2.1)	GOSAT	Jun 2009 - Sep 2012	SRON
	OCPR (4.0)	GOSAT	Jun 2009 - Dec 2011	Uni. of Leicester

Table 4.6.1: List of available algorithms included in the GHG_CCI CRP.

Box A: The Climate Monitoring Facility (CMF)

The CMF is a web-based tool under development at ECMWF that will be put at the disposal of CCI. CMF aims at providing an easy access to a large variety of data and variables to facilitate the assessment of the long-term homogeneity of a given dataset, of the level of agreement of its monthly mean area averaged statistics with equivalent fields available from different sources, and of its consistency with correlated variables. An up-to-date description of the design of the CMF prototype can be found in CMUG (2013c). The CMF database (CMFDb) content was summarised in CMUG (2013b, d).

Like any other tool, the CMF should be used for applications it was designed for (i.e. monitoring and assessing the low-frequency, multi-year variability of regional averages). The comparisons it facilitates are based on pre-calculated statistical regional averages of monthly mean data. As differences may occur in the data coverage of different data streams used to produce those averages, caution should be used when assessing their comparison.

Shortcomings in the CO₂ and CH₄ model and in the observations, as well as lack of a proper anchor for the variational bias correction of these observations led to large biases in the MACC atmospheric reanalyses of CO₂ (A. Agusti-Panareda and S. Massart, personal communication). Thus, CO₂ and CH₄ outputs from two recent research experiments have also been considered in this study. These two experiments are forecast runs (thus not reanalyses) that used optimized fluxes from the flux inversion MACC reanalysis of CO₂ (Chevallier *et al.* 2010) and CH₄ (Bergamaschi *et al.* 2013). The CO₂ fluxes were optimized using only surface observations (no satellite data included), thus the comparisons shown below are completely independent. In contrast, the CH₄ fluxes were obtained using both SCIAMACHY and surface observations, thus the outputs from this research experiment cannot provide completely independent comparisons. It is noted that two different transport models were used in the CO₂ and CH₄ flux inversion (LMDZ and TM5, respectively) and in the forward calculations done by the ECMWF Integrated Forecasting System (IFS) for both CH₄ and CO₂. This is likely to

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

represent a source of error in terms of CO₂ and CH₄ concentrations. It is acknowledged that atmospheric CO₂ and CH₄ fields from the MACC flux inversion reanalysis (which used the same transport model for the flux inversion and forward simulation) also exist (Chevallier *et al.* 2010, Bergamaschi *et al.* 2013). These are more suitable for trend calculations.

When displayed, the MACC atmospheric reanalysis is hereafter simply labelled as MACC. The labels MCO₂ and MCH₄ are instead used in this paper to identify the CO₂ and CH₄ outputs from the two recent MACC forecasts experiments. Both the MACC reanalysis and the MCO₂ experiment are available and used from January 2003 onwards. Experiment MCH₄ was only used for 2008.

A preliminary look at the GHG_CCI data neither highlighted technical issues with the data nor unphysical values. Each dataset includes the minimum required information to make use of it, e.g. data and their uncertainties, as well as quality flags. However, differences exist between the various datasets. For instance, some algorithms provide both the bias corrected and the uncorrected retrievals, in other cases only one of them is provided; the meaning of the quality flags is not used consistently by the various groups. It is recommended that these differences were overcome in future versions and the corresponding Product User Guides updated to reflect the changes.

Whenever available, bias corrected retrievals were used in this study. Also the available quality flags were used to filter out poor quality data.

4.6.1 Carbon dioxide

The MCO₂ outputs make use of optimized CO₂ fluxes computed according to the method described in Chevallier *et al.*, (2010). The latest MACC inversion was found more accurate than the satellite retrievals at the sounding scale (Chevallier, F., and C. W. O'Dell, 2013), though a number of challenges with inverting fluxes from the CO₂ satellite retrievals compared to the current MACC surface-based approach still exist (Chevallier *et al.*, 2014).

Figure 4.6.1 shows the time-series of the anomaly of the global mean monthly averaged carbon dioxide retrieved from the SCIAMACHY and GOSAT measurements using the three algorithms that are included in the GHG_CCI in the Climate Research Package (CRP).

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

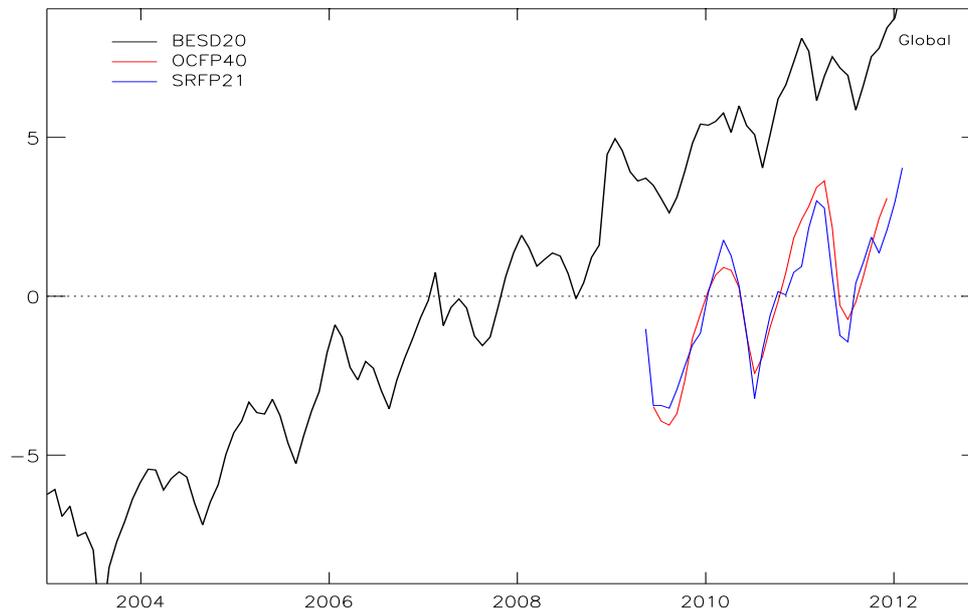


Figure 4.6.1 Time series of the monthly mean globally averaged anomaly (with respect to the mean computed over the whole period of availability) of carbon dioxide retrieved from SCIAMACHY (black line) and GOSAT (blue and red lines) measurements using the three algorithms that are part of the GHG_CCI CRP. Anomaly data are in ppm.

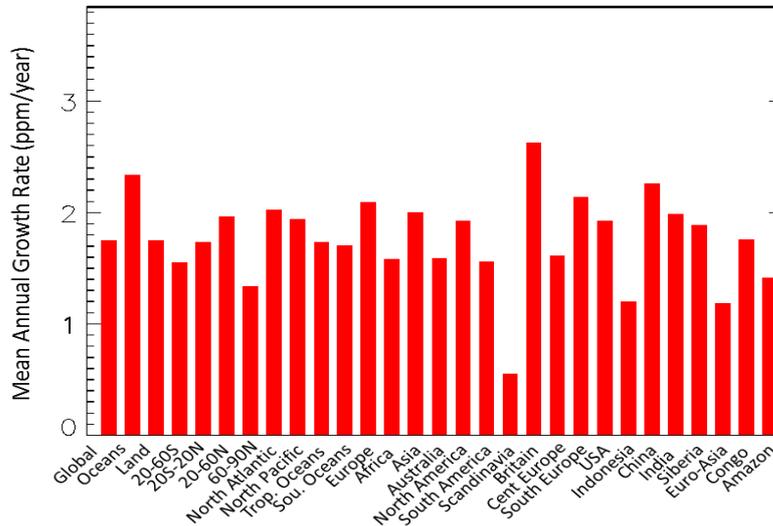
The anomaly for each dataset in figure 4.6.1 is computed as the difference between the monthly mean values and the mean over the period of data availability. Over the period 2003-2011, the (BESD 2.0) SCIAMACHY XCO₂ retrievals are characterized by a mean global annual growth rate of 1.78 ppm/year.

The actual growth rate varies from year to year and depends on the geographical area over which the mean growth rates are computed. Figure 4.6.2 provides the mean annual growth rate of XCO₂ estimated from the SCIAMACHY retrievals during the period 2003-2011. The table on the right hand side of figure 4.6.2 provides the actual global annual change for the years 2004-2011 computed with respect to the 2003 global mean carbon dioxide amount of 375.43ppm. The table also refers to the (BESD 2.0) SCIAMACHY XCO₂ retrievals. For comparison, the mean annual growth calculated from the NOAA ESRL data (Conway *et al.*, 1994) has also been given.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



Global annual XCO ₂ change (ppm)		
	BESD	NOAA
2004	1.52	1.81
2005	2.01	2.03
2006	1.77	2.13
2007	1.66	1.77
2008	1.59	2.08
2009	2.74	1.50
2010	1.61	2.29
2011	1.68	1.92

Figure 4.6.2. Mean Annual CO₂ growth rate in different geographical regions estimated using the (BESD 2.0) SCIAMACHY XCO₂ retrievals for the period Jan 2003 - Dec 2011. The right hand side table gives the actual global mean annual change (with respect to the global annual mean from previous year) obtained from the SCIAMACHY XCO₂ retrievals relative to an estimated annual global mean for 2003 of 375.43 ppm and from the NOAA ESRL data (Conway *et al.*, 1994) that had a 2003 mean value of 374.97ppm.

The three XCO₂ retrievals from SCIAMACHY and GOSAT measurements well compare in the northern hemisphere extra-tropics (figure 4.6.3). In this region, the annual cycle, and the amplitude of the carbon dioxide changes, as well as the observation uncertainties exhibit a high level of agreement. Overall, the MCO₂ output in the northern hemisphere extra-tropics is also in good agreement with the three datasets.

In the tropics and in the southern hemisphere the comparisons show a lower level of agreement. It is noted that the MCO₂ output seems to have a much slower CO₂ growth with time than in the retrievals. Different reasons may have contributed and concurred to this outcome, and more generally explain the differences between observed and modelled fields. Some inter-hemispheric differences exist in the inverse modelling system (Chevallier *et al.*, 2010) as this is constrained by measurements of atmospheric CO₂ that are more sparse in the tropics and in the southern hemisphere thus making the resulting fluxes probably less reliable than those in the northern hemisphere extra-tropics. The errors related to the use of two different transport models in the flux inversion and in the forward simulation are likely also larger in those data sparse regions than in the northern extra-tropics. The comparison method is also prone to errors and, in some cases it can lead to misleading conclusions if the model outputs and observations have substantially different coverage. By design, the SCIAMACHY and GOSAT instruments mostly provide measurements over land. In contrast, models are normally defined at every grid-point. In the southern hemisphere, where the extent of the oceans is greater than that of land, significant differences may be found in the mean values. As CO₂ is lower over the oceans than over land, the model mean CO₂ is expected to exhibit lower values than the mean observations. Differences may also exist in regions characterized by clouds and aerosols, like the tropics, that are a source of uncertainty in deriving CO₂.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Figure 4.6.3 also shows a sudden increase in the MCO₂ analyses between end of 2004 and beginning of 2005 in the southern extra-tropics. This can be explained, at least in part, by an exceptional growth of the atmospheric CO₂ caused by a significant drought in the Amazonian and Central African regions where a large number of trees died, releasing large quantities of CO₂ while contributing in reducing the CO₂ absorbing ability of the forests. The year 2005 was classified as one of the worst droughts in more than a century caused by warmer than usual ocean temperatures.

During the period of SCIAMACHY and GOSAT overlap (Jun 2009 - Apr 2012), the three datasets exhibit a high level of consistency in the extra-tropics. In the tropics, the mean BESD2.0 SCIAMACHY retrievals agree very well with the corresponding mean GOSAT data retrieved from the OCFP4.0 algorithm. In contrast, the SRFP2.1 retrievals show a slower growth that leads to lagged CO₂ annual maximum values of about 2-3 months compared with the other two retrievals. Despite this time lag, the three XCO₂ datasets vary mostly within their corresponding uncertainties.

The observation uncertainties were compared with the residuals between each dataset and their model equivalent and presented in figure 4.6.4 for the same three latitudinal bands shown above. The plots made use of all available data for the three datasets. In general, the SCIAMACHY BESD2.0 uncertainty is of the same magnitude of or larger than the observation minus model departures, particularly in the northern hemisphere extra-tropics. In contrast both GOSAT sets of retrievals exhibit observation uncertainties that are on average smaller than the observation departures from the modelled XCO₂, especially in the tropics and southern hemisphere extra-tropics. The quality of the XCO₂ outputs has clearly a key role, here, as they are used as reference. As mentioned above, the quality of the modelled XCO₂ is potentially hampered by errors or biases in the observations as well as errors associated with the transport model. Thus, it is difficult to anticipate the potential impact of these observations if assimilated and dedicated experiments will need to be run.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

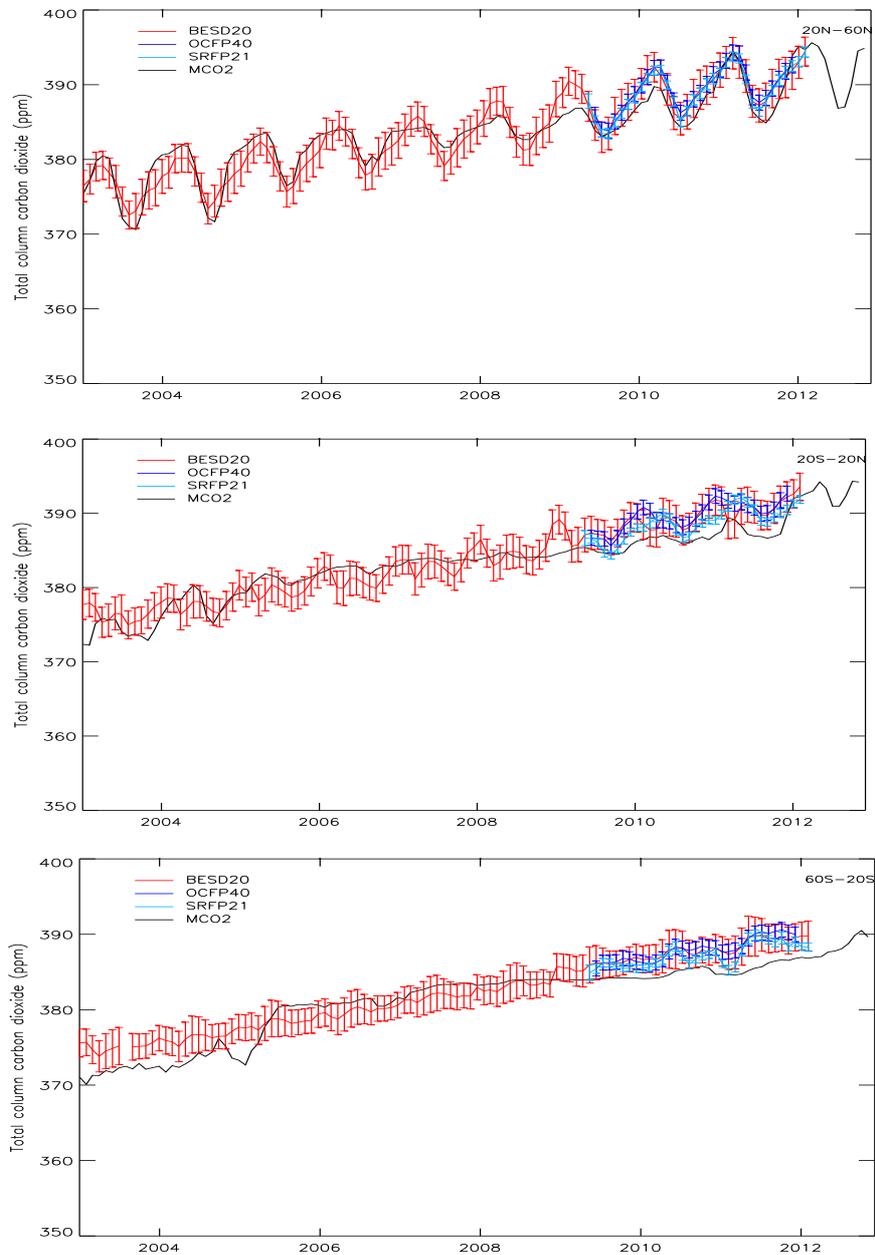


Figure 4.6.3 Time series of the CCI XCO₂ retrievals from SCIAMACHY (BESD 2.0) and GOSAT (OCFP 4.0 and SRFP 2.1) measurements and of the MCO₂ outputs. Data are averaged over three latitudinal bands (from top to bottom): the northern hemisphere extra-tropics, the tropics, and the southern hemisphere extra-tropics. Data are in ppm.


CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

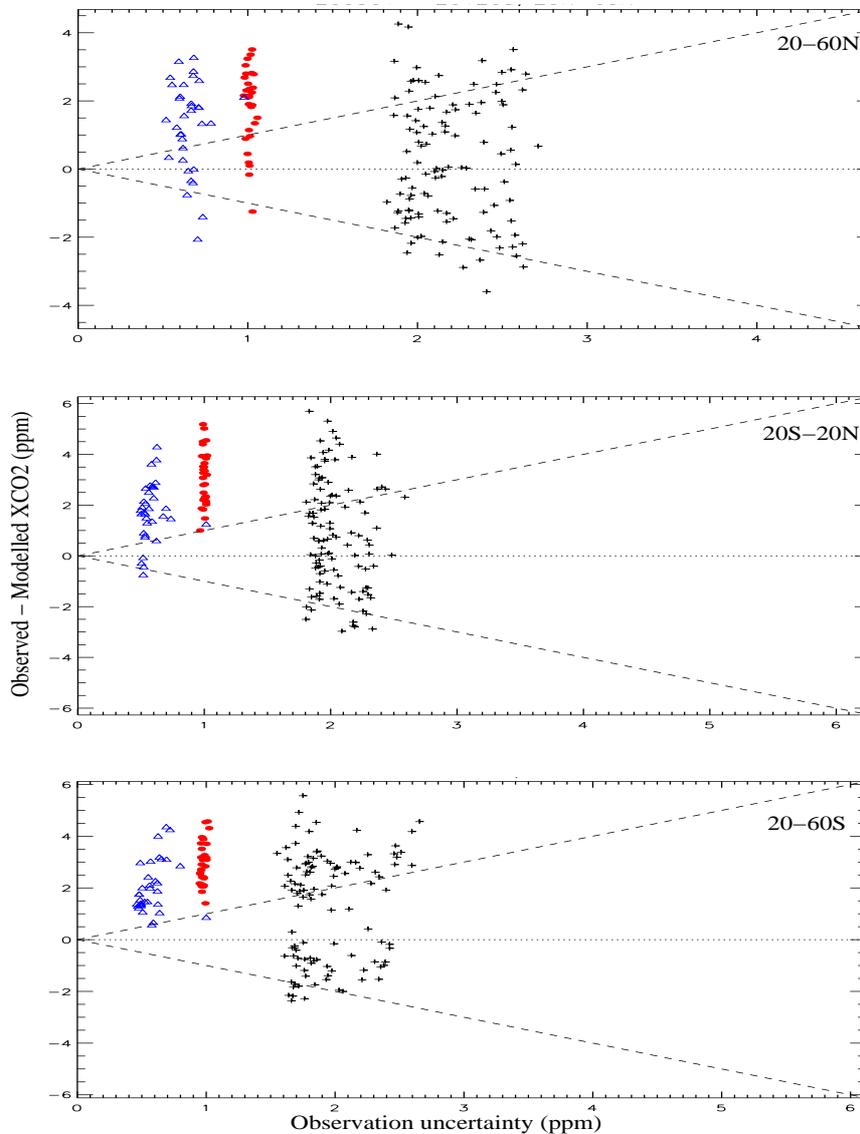


Figure 4.6.4 Scatter plot of the CCI XCO₂ uncertainties of the SCIAMACHY (BESD 2.0, black crosses) and GOSAT (OCFP 4.0, red filled circles, and SRFP 2.1, blue triangles) retrievals and the residuals between each retrieval and the MCO₂ outputs. Data are averaged over three latitudinal bands like in figure 4.6.3. Data are in ppm.

4.6.2 Methane

Monthly mean area averaged statistics of four methane retrievals produced by the GHG_CCI (see table 4.6.1 for details) have been compared with their model equivalent obtained from the MACC atmospheric reanalysis and from an experimental forecast run, MCH₄, only valid for 2008 and that is based on MACC optimized fluxes from Bergamaschi *et al.* (2013). The main difference between the two MACC CH₄ products is that the CH₄ fluxes in the MACC atmospheric reanalysis were derived using uncorrected SCIAMACHY observations. In the most recent simulations, the CH₄ fluxes were instead derived from bias corrected SCIAMACHY data and surface observations.

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



Using the two SCIAMACHY products that are available over a longer period, the mean annual change in XCH₄ was derived for several geographical areas using the 2003-2011 observations (figure 4.6.5). The actual year-by-year global changes are provided in the right hand side table of figure 4.6.5.

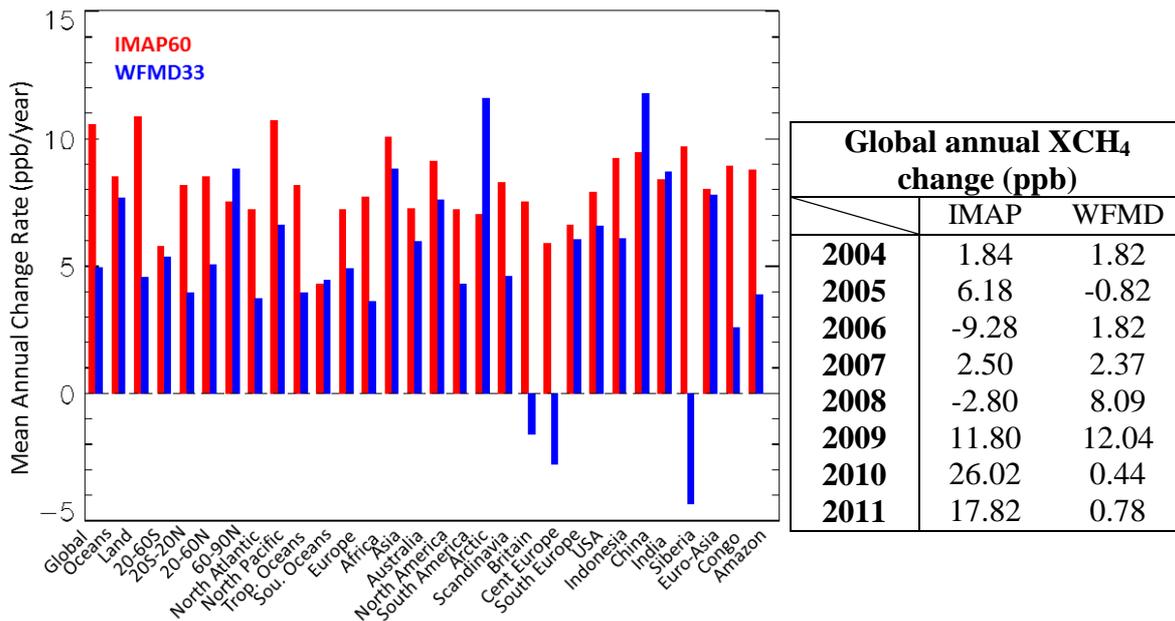


Figure 4.6.5 Mean Annual CH₄ change rate in different geographical regions estimated using the (IMAP 6.0, red, and WMFD 3.3, blue) SCIAMACHY XCH₄ retrievals for the period Jan 2003 - Dec 2011. The right hand side table gives the actual global mean annual change (with respect to the global annual mean from previous year) obtained for the two sets of SCIAMACHY XCH₄ retrievals relative to an estimated annual global mean for 2003 of 1760.24ppb for IMAP 6.0 and 1748.98ppb for WFMD 3.3.

Overall and despite some dependence on the algorithm and geographical areas, the XCH₄ retrievals show a small trend of up to 5-6% per decade. This slow growth rate is also confirmed by figure 4.6.6 that shows the time series of the monthly mean XCH₄ averaged over three latitudinal bands (tropics and northern and southern extra-tropics) for the four GHG_CCI products and the two sets of XCH₄ model outputs. Comparisons were performed over a total of 32 geographical areas with negligible difference in the level of agreement between the averaged observed and modelled methane.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

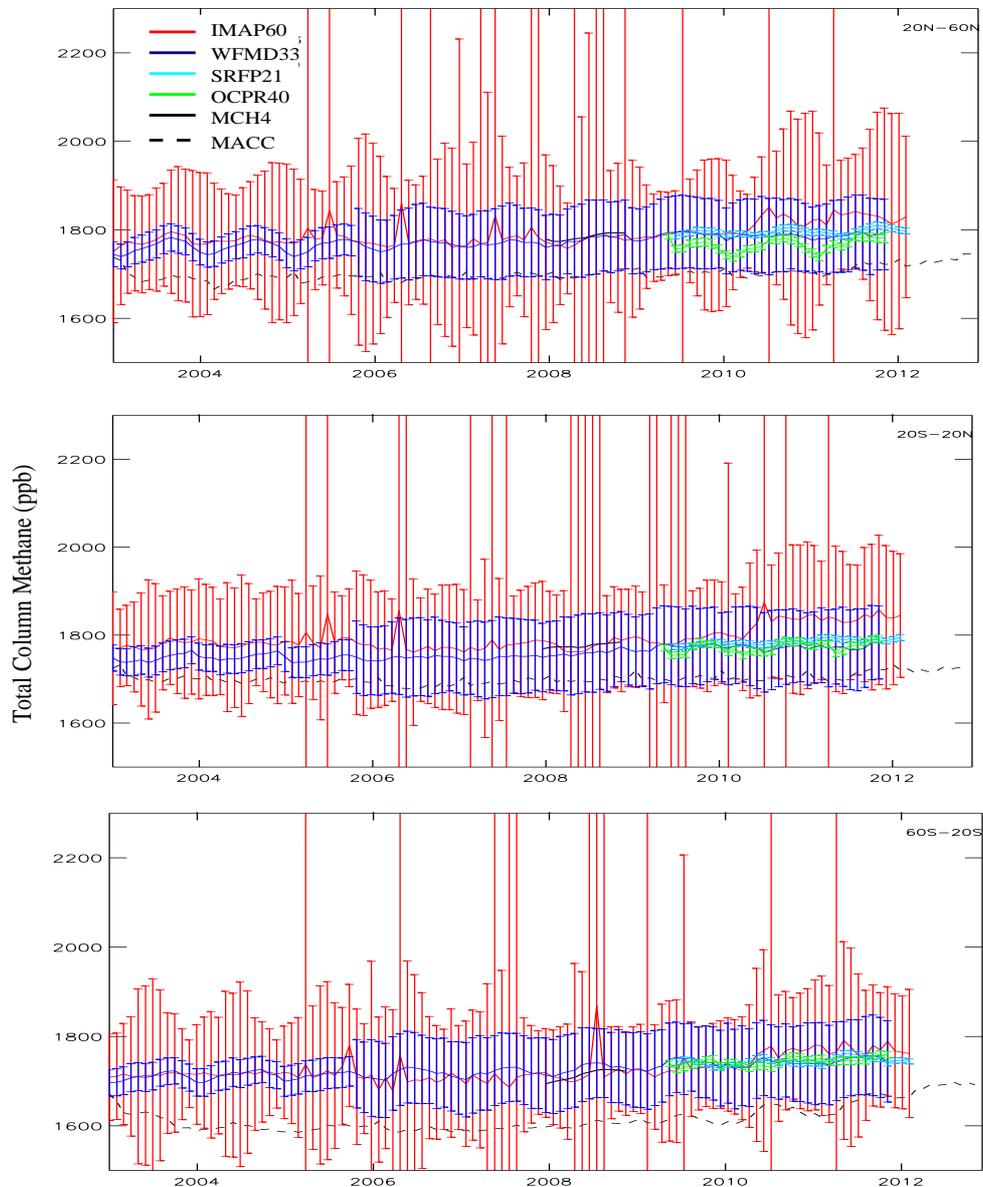


Figure 4.6.6 Time series of the CCI XCH₄ retrievals from SCIAMACHY (IMAP 6.0 and WFMD 3.3) and GOSAT (OCPR 4.0 and SRFP 2.1) measurements and of the MACC (dashed black line) and MCH4 outputs. Data are averaged over three latitudinal bands (from top to bottom): the northern hemisphere extra-tropics, the tropics, and the southern hemisphere extra-tropics. Data are in ppb.

The four GHG_CCI products generally show a good level of agreement and consistency between each other, particularly in the southern hemisphere extra-tropics. A sudden change is noticeable in the IMAP6.0 product (red lines) at the beginning of 2010 in the tropics and in the northern hemisphere extra-tropics. The XCH₄ level reached by the IMAP6.0 retrievals is then maintained for the remaining period of SCIAMACHY availability. Such behaviour is not found in the other three products. This might point to a possible change in the retrieval chain that might be worth investigating further.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

The XCH₄ from the MACC atmospheric reanalyses (dashed black line) show on average a low bias of about 100ppb compared with the GHG_CCI retrievals. About half of such a bias can probably be attributed to the SCIAMACHY measurements that, as mentioned above, were not bias corrected in the flux calculations that produced the fluxes used in that experiment. In contrast, the MCH₄ product (solid black line over 2008) exhibits a much higher level of agreement with the corresponding retrievals. The reader is reminded that the SCIAMACHY data were used in the flux inversion model, thus these comparisons are not independent.

The uncertainty bars over-plotted to the timeseries in figure 4.6.6 show that the two SCIAMACHY retrievals have normally much larger uncertainties compared with the residuals from their MCH₄ model equivalent. It is also clear that for the WFMD 3.3 retrievals the observation uncertainty increased in 2005. This is likely a consequence of degraded measurements that followed the 2005 detector degradation in SCIAMACHY channel 6 that particularly affected the methane retrievals.

Figure 4.6.6 also show that the IMAP 6.0 SCIAMACHY XCH₄ product (red) exhibits sporadically larger uncertainties than usual. Understanding the reason for these large values requires a detailed analysis of the uncertainties on a pixel level (i.e. on the L2 data).

In contrast to the SCIAMACHY retrievals, the uncertainties of the GOSAT observations are much smaller. As the GOSAT instrument was launched in January 2009, with data available from June 2009, there is no overlapping between its retrievals and the MCH₄ analyses, thus it is not possible to report on their level of agreement. However, by extrapolating/projecting the MCH₄ output forward, figure 4.6.6 could offer some indications on the potential level of agreement between analyses and observations that would have likely been good. Using the same argument, the observation uncertainties would have most likely been comparable with the departures between the observations and the modelled CH₄.

As noted in the case of CO₂, there are still issues in the GHG system that will require more work and to be addressed in the future. It is hoped that these observations in combination with surface measurements could have the potential to positively impact future XCH₄ model outputs.

4.7 Aerosols

A preliminary assessment of the Level 3 (L3) Aerosol CCI products was discussed in CMUG (2013b). That assessment was based on observation-model confrontation of pre-calculated monthly mean area averaged statistics using a prototype version of the CMF (see Box A in section 4.6).

The CMUG (2013b) study focussed on the comparison of the L3 Aerosol Optical Depth (AOD) retrieved from the ENVISAT AATSR measurements in 2008 using three Aerosol_CCI algorithms against their MACC reanalysis equivalent. The three algorithms used were the Swansea University algorithm version 4.0 (SU40), the Finnish Meteorological Institute algorithm version 1.42 (ADV142), and the Oxford / Rutherford Appleton Laboratory algorithm version 1.21 (ORAC121).

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Since the CMUG (2013b) study, the ADV142 dataset was extended to span the period from January 2007 to December 2010, while two new algorithm versions (4.1 and 4.2) were released by the Swansea University. In the latter case, the version 4.1 scheme (SU41) was used to retrieve aerosol products for the entire ENVISAT AATSR (Jul 2002 - Apr 2012) mission; while the version 4.2 algorithm (SU42) was used to retrieve two test years, 2008 from the ENVISAT AATSR and 2000 from the ERS-2 ATSR-2 measurements. Compared with SU40, SU41 has a better treatment than SU40 of the cloud-contaminated data over ocean, though compromises in the ocean model led to glint contamination in the AOD (P. North, personal communication). The issue with the SU41 glint contamination was improved in the SU42 ocean model. Over land the two most recent algorithms (SU41 and SU42) are equivalent to SU40 (P. North, personal communication).

Here, an update of the CMUG (2013b) study is presented. The focus is on the inter-comparison of the three SU datasets with the MACC aerosol reanalyses for 2008, and the long-term homogeneity of the SU41 and ADV121 retrievals.

The analysis performed here confirms that negligible to very small differences were found over land between the three SU datasets (not shown); while measurable differences up to 20% compared with SU40 are found over the oceans (figure 4.7.1). At the longest wavelength (1610nm), the higher is the SU version number the higher is the level of agreement with the MACC reanalyses. In contrast, at all other wavelengths the level of agreement with MACC seems degraded in the two newest versions. However, MACC II (2013) discussed some of known issues in the MACC aerosol model. For instance, one shortcoming of the model is related to the representation of sea salt, which seems to be overestimated (e.g. at 550 nm) and leads to a high AOD bias in southern oceanic regions (see also the right panel of figure 18 in CMUG (2013b)). The indication from the MACC II (2013) study is that the 550nm AOD reanalyses could show a bias as large as 40-60% over the oceans. If such a bias was taken into account, a more realistic AOD mean value at 550nm over the oceans could be around 0.1, which would lead to very small residuals from the SU42 dataset.

The SU4.1 product was made available over an extended period of time, which covered the availability of AATSR measurements. In addition, the ADV1.42 product availability was extended from 2008 to cover the period 2007-2010. Long-term level of agreement with the MACC reanalysis was assessed using the CMF and its database.

Figure 4.7.2 shows the global mean AOD anomaly at the four wavelengths for the two products. A very small dependence of the global mean anomaly is found in terms of wavelengths. The two products also show variations with similar amplitude up to about 15% of the global mean values.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

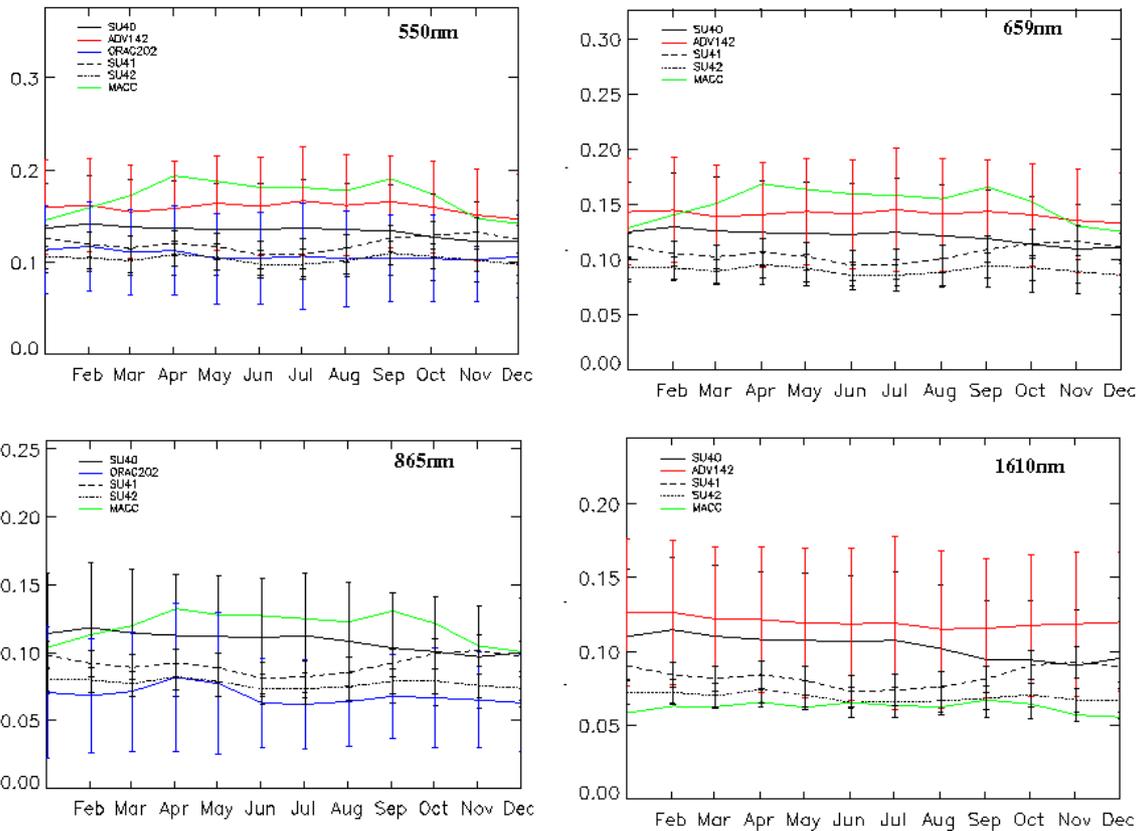


Figure 4.7.1 Mean AOD at 550,659, 865, 1610nm for the Aerosol CCI datasets (SU40, ADV121, ORAC202, SU41, and SU42) retrieved from the 2008 AATSR measurements and their MACC equivalent averaged over the oceans. The vertical bars represent the observation errors of the CCI datasets.

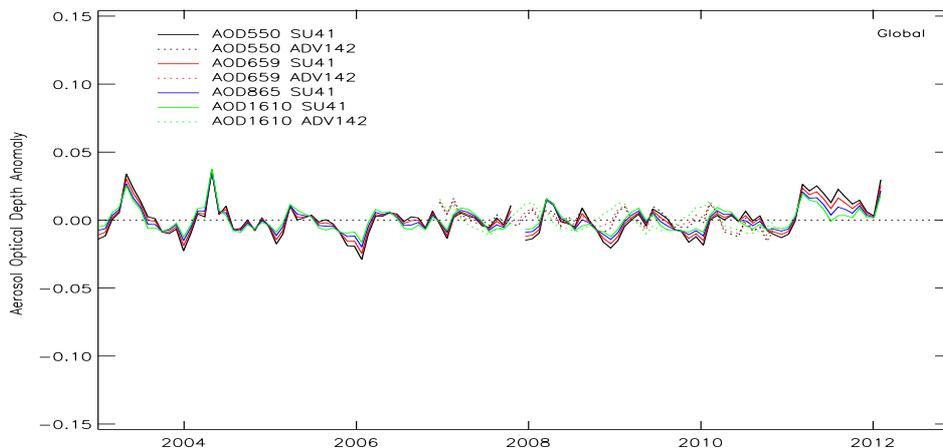


Figure 4.7.2 Global mean anomaly for the SU4.1 retrievals (solid lines) and the ADV1.42 product (dashed lines) computed for the four wavelengths (see legend for details).

In the global mean, the SU4.1 and ADV1.42 show similar levels of agreement at the shortest wavelengths (550 and 659nm) with the MACC reanalyses. Thought, the latter shows a

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



stronger annual variability than the two retrievals (figure 4.7.3). In spite of this, the departures between observations and analyses appear to be smaller than the observation uncertainties in both cases. This is also the case for the comparisons at 865nm between the SU4.1 retrievals and the MACC data.

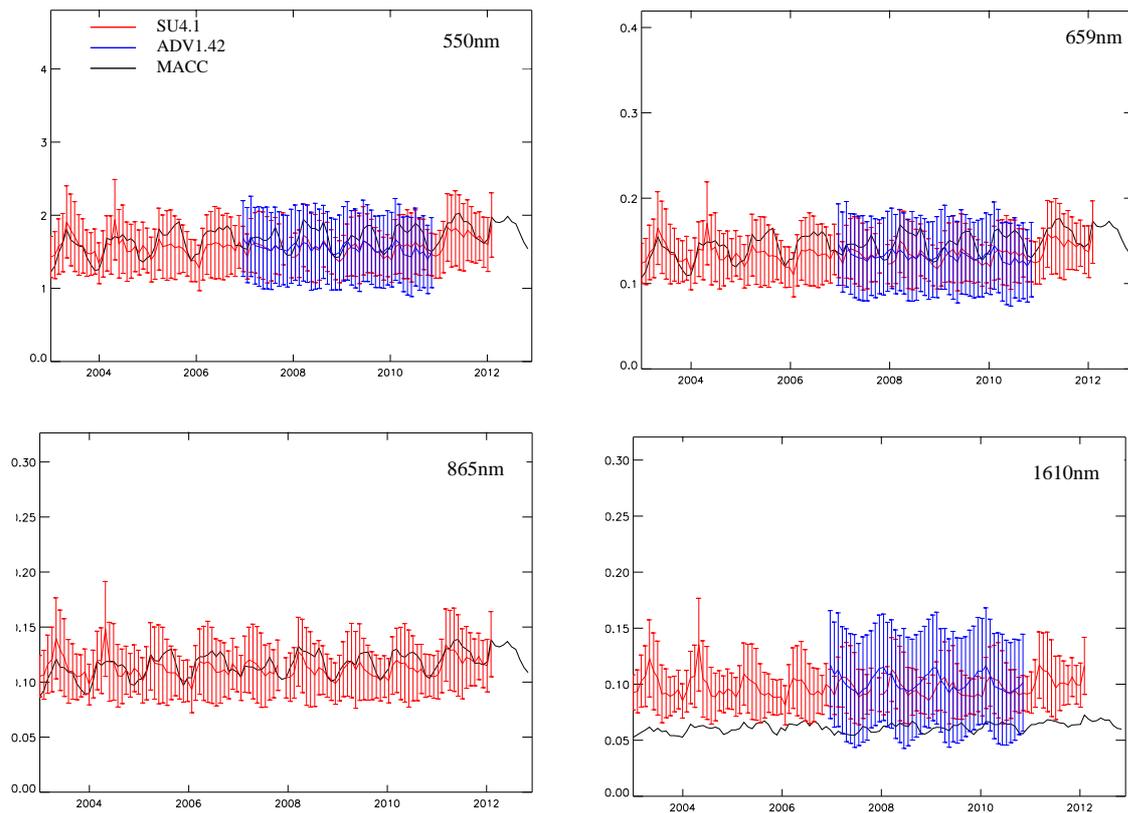


Figure 4.7.3 Like in figure 4.6.2, but for the long-term global mean statistics of SU4.1 (red lines) and ADV1.42 (blue lines) AODs. The black lines refer to the MACC AOD reanalyses.

At the longest wavelength (1610nm), the two Aerosol_CCI products show a good level of agreement in the global mean values, while their departures from the MACC reanalyses are at best as large as their uncertainties.

The good level of agreement found in the global mean at the shortest wavelengths (seen in figure 4.7.3) is not confirmed by the averages over specific geographical regions. Figure 4.7.4 shows, as an example, the comparisons at 550nm produced over land and oceans (right panels) as opposed to the global mean time-series (left panel).



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

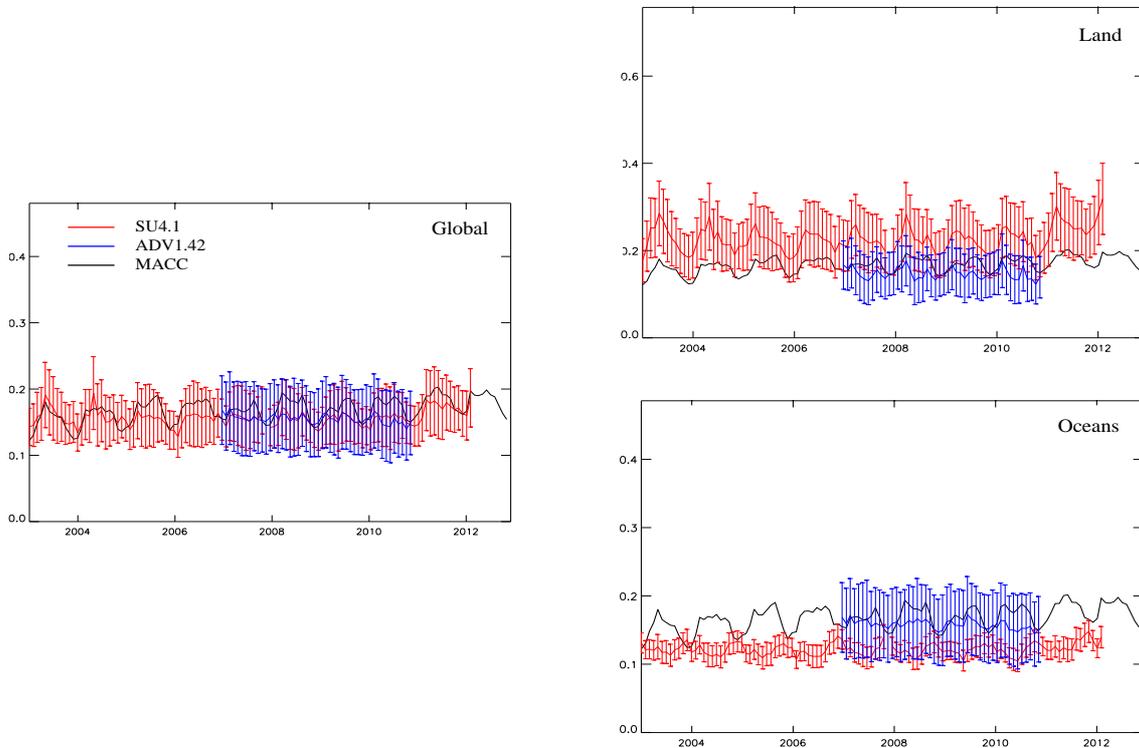


Figure 4.7.4 Time series of the MACC AOD reanalyses and the SU4.1 and ADV1.42 AOD retrievals at 550nm. The statistics were computed for the whole globe (left panel), and over Land (top right panel) and Oceans (bottom right panel).

Over the oceans, the ADV1.42 product shows a higher level of agreement with the MACC reanalysis than the SU4.1 retrievals (bottom right panel of figure 4.7.4). It was already mentioned that a number of shortcomings in the model that produce too high AOD over the oceans, especially the southern oceans, have been identified (e.g. MACC II, 2013) where the MACC reanalyses show an up to 60% positive bias compared to the AERONET measurements. When such a bias is accounted for, the unbiased mean reanalyses would probably be in better agreement with the SU4.1 product than with the ADV1.42 product. Over land, the MACC reanalyses are within the observation uncertainties for both products.

The assessment of the retrieval uncertainties against the observation residuals from the MACC reanalyses has been extended with the latest data (figure 4.7.5). As done in figure 4.7.4, scatter plots were produced for several geographical areas, and shown for the whole globe and then for data averaged over the oceans and land. In the global mean, the observation uncertainties for all datasets are comparable with the corresponding observation residuals from the MACC reanalyses. This is also the case for the ADV1.42 and SU4.0 datasets when the averages over land and oceans are considered, as well as for the SU 4.1 and 4.2 datasets over land. In the case of the two most recent SU datasets, the uncertainties over the oceans appear to be much smaller than the observation residuals. This is also the region where improvements were implemented in the SU algorithm.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

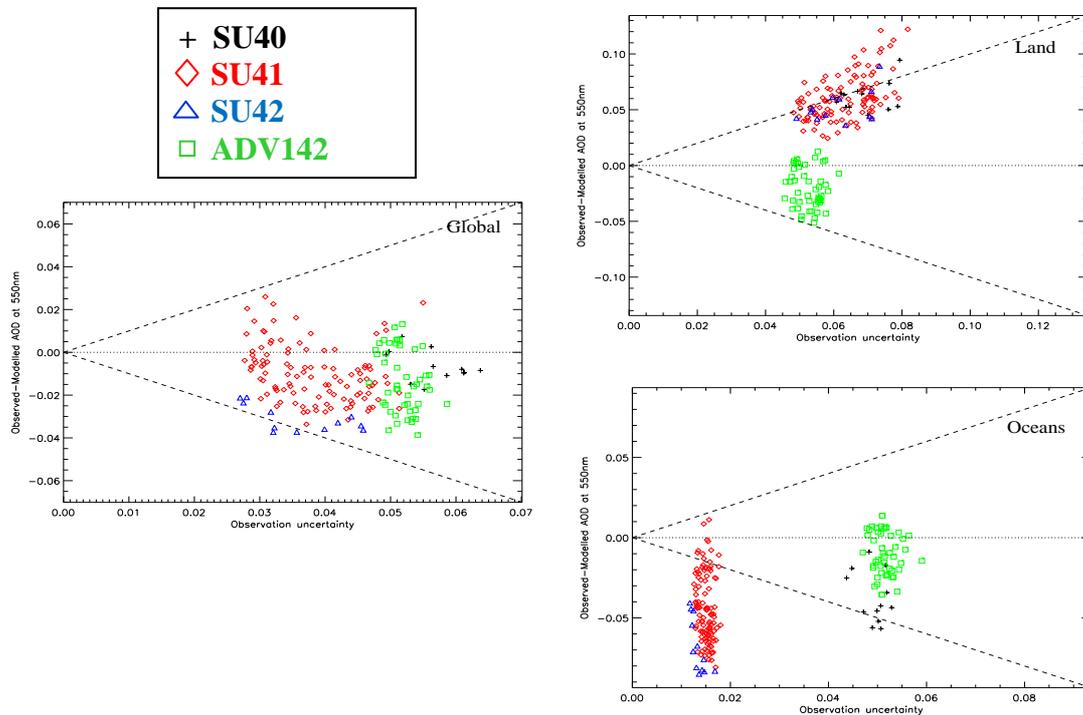


Figure 4.7.5 Scatter plot of the observation departures from the MACC reanalyses at 550nm versus the observation uncertainty. The black crosses, red diamonds and blue triangles refer to the SU 4.0, SU 4.1, and SU 4.2 data, respectively. The green squares refer to the ADV 1.42 AOD data. The left panel shows all observations for the whole globe; the top panel refers to the observations over Land and the bottom right panel to the observations over Oceans.

The scatter plots in figure 4.6.5 strongly rely on the quality of the MACC AOD reanalyses, as they are used as a reference. Additionally, as mentioned above, the MACC AODs are found to have a 60-80% positive bias compared with the AERONET observations over the Southern Oceans. It is argued that by accounting for such a bias over the oceans, the observation uncertainties would likely be comparable with the observation residuals from the bias corrected reanalyses. The two latest SU datasets also appear to be more accurate than the previous one (SU4.0) over the Oceans. Arguably, the MACC AOD analyses would benefit from the assimilation of these observations.

4.8 Land Cover

4.8.1 Introduction

Vegetation interferes with the atmosphere and other components of the climate system at the global scale. The atmosphere-vegetation dynamics has a non linear relationship and any changes in vegetation immediately lead to changes in land surface water, energy and mass fluxes. Modern Earth System Models require land cover information for the initialization of patterns of Plant functional Types (PFT) which constrain many interactive processes. Proper information on land cover or PFT distribution is therefore crucial for a realistic representation of spatiotemporal land surface water, energy and carbon fluxes.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

The CCI LC project provides novel land cover information which differs from the land cover information currently used in climate models. The objectives of the analysis in this study are in particular to answer the following guiding questions

- a. How does an integration of ESA CCI land cover data affect the energy and water fluxes at global scales?
- b. Does the integration of ESA CCI land cover data improve the skill of MPI-ESM in simulating present day climate?
- c. Is the usage of ESA CCI land cover data superior compared to the usage of previously existing land cover products like GlobCover? What is the added value of CCI LC?

The assessments in this study were made using coupled (atmosphere-land) and uncoupled (land only) simulations. Details of the simulation setup are provided in section 4.8.2.

4.8.2 Data and models

4.8.2.1 MPI-ESM and its land surface model (JSBACH)

The model used in the present study is the land surface scheme of the MPI-ESM, JSBACH (Reick *et al.*, 2013). The model is implicitly coupled to the atmospheric component of MPI-ESM (ECHAM6) (Stevens *et al.*, 2013) and simulates all relevant land surface water, energy and carbon fluxes. The present analysis uses version 2.02 of JSBACH which is comparable to the model version which was used for the Coupled Model Intercomparison Project 5 (CMIP5) (Taylor *et al.*, 2012).

4.8.2.2 Land Cover in MPI-ESM

JSBACH uses the concept of plant functional types (PFT) to simulate land surface dynamics. A PFT represents vegetated surfaces which show a similar dynamics and response to external forcing like e.g. radiation and precipitation. Each PFT represents a group of plants with a similar set of functions and which respond similar to climate. Thus, PFT's explicitly take into account the fact that a certain plant type can only occur within certain climatic limits defined by upper and lower limits for e.g. precipitation and temperature. Figure 4.8.1 shows the distribution of PFT types in JSBACH and their frequency distribution. The distribution is originally based on AVHRR land cover information.



CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

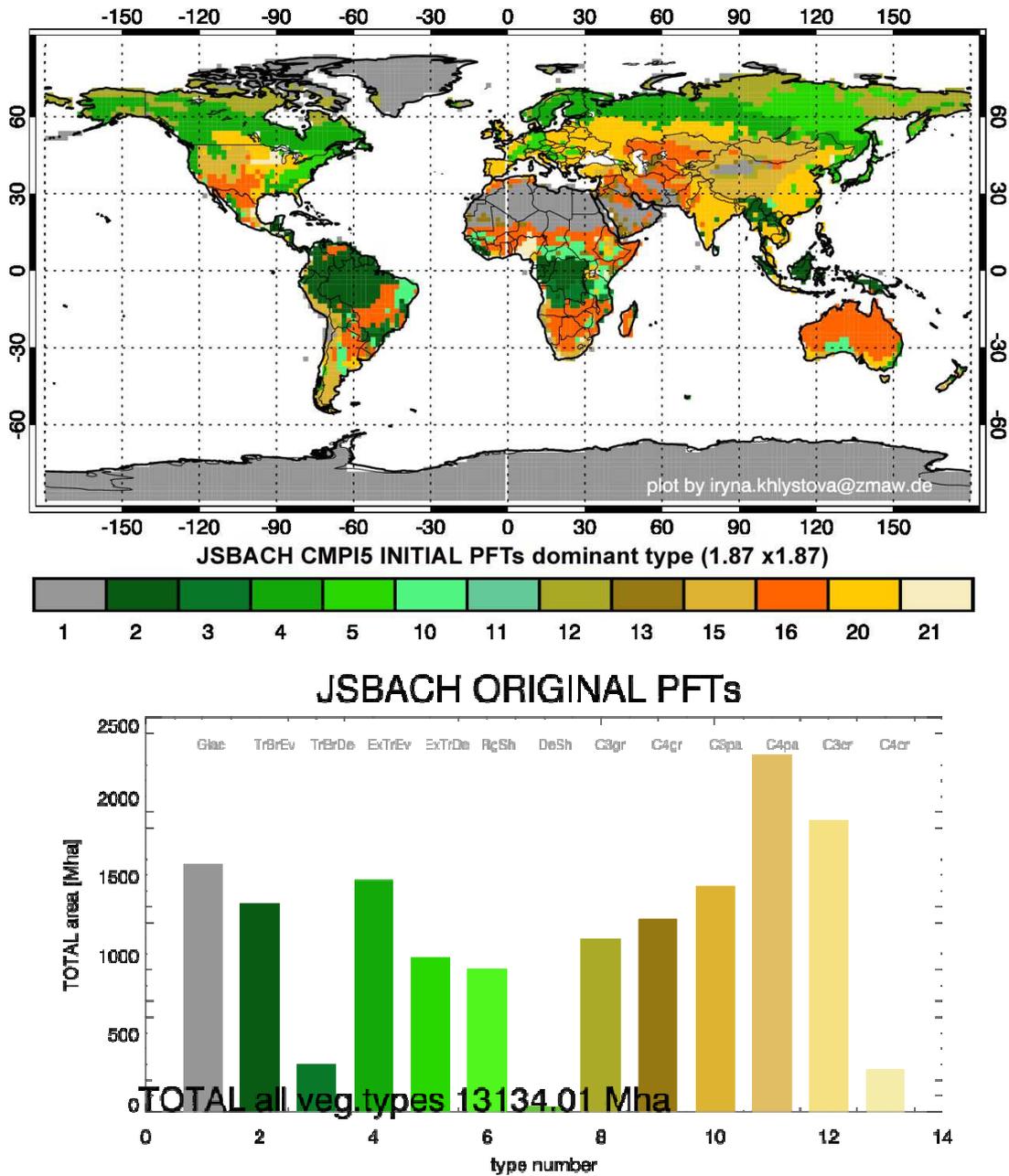


Figure 4.8.1: Majority PFT types in JSBACH and their area covered

4.8.2.3 CCI precursor and CCI land cover data

Within the project, the 3 epochs of the new CCI land cover data and its pre-cursor product from the GlobCover project were used to substitute the standard JSBACH dataset of PFTs. To estimate the PFT fractions from the land cover information provided in either the GlobCover product or the CCI land cover products, the land cover information had to be translated into PFT classes. Details on the method are provided in section 4.8.3. The data used for the present study are summarized in Table 4.8.1.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Source	Period
Globcover	2005
CCI Landcover	epoch 2000
	epoch 2005
	epoch 2010

Table 4.8.1: Summary of land cover datasets used in the study

4.8.2.4 Model offline forcing data

JSBACH simulations can be performed either by forcing the land surface scheme with any kind of meteorological forcing data (e.g. station measurements, reanalysis data) or by coupling JSBACH directly to a Global Circulation Model (GCM), like ECHAM6. For the present study, we use both, coupled and offline simulations. The latter are forced with observed meteorological data. This allows for an assessment of the direct impacts of CCI land cover on the land surface's state, while the coupled model simulations allow to assess the feedback mechanism to the land surface.

Two different meteorological forcing datasets are used for the present study, which allows also to quantify the relative sensitivity of a change in forcing data compared to the change in simulated climate caused by a change in the land cover distribution.

WATCH forcing data (WFD) were created in the framework of the WATCH project (www.eu-watch.org) and are documented in Weedon *et al.* (2011). The data set is available at 0.5° spatial resolution and sub-daily time steps and is based on ERA40 reanalysis data (Uppala *et al.*, 2005). An extensive post-processing was conducted in which the data were interpolated to a rectangular 0.5° resolution grid and elevation corrections as well as undercatch corrections for precipitation and snowfall were applied.

CRU NCEP forcing data: The CRU/NCEP forcing dataset is a combined dataset from the NCEP reanalysis (Kalnay *et al.*, 1996) and the station data based of land surface precipitation and temperature data of the University of East Anglia Climate Research Unit (CRU)⁶.

4.8.3 Methods

4.8.3.1 Integration of ESA CCI data

The JSBACH input PFTs were derived from ESA CCI land cover classes using the general scheme described by Poulter *et al.* (2011). The procedure of Poulter *et al.* (2011) has been adapted to enable conversion of GlobCover and CCI LC information into JSBACH specific PFT fractions. The general preprocessing steps are outlined in Figure 4.8.2 and comprise the regridding of the data, the splitting and reassignment of land cover classes into PFT classes as well as the combination with ancillary climate data to delineate tropical and extratropical PFTs. Figure 4.8.34.8.3 shows the results of the conversion procedure for CCI LC as well as the difference between the resulting PFT distributions.

⁶ <http://www.cru.uea.ac.uk/data>


CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

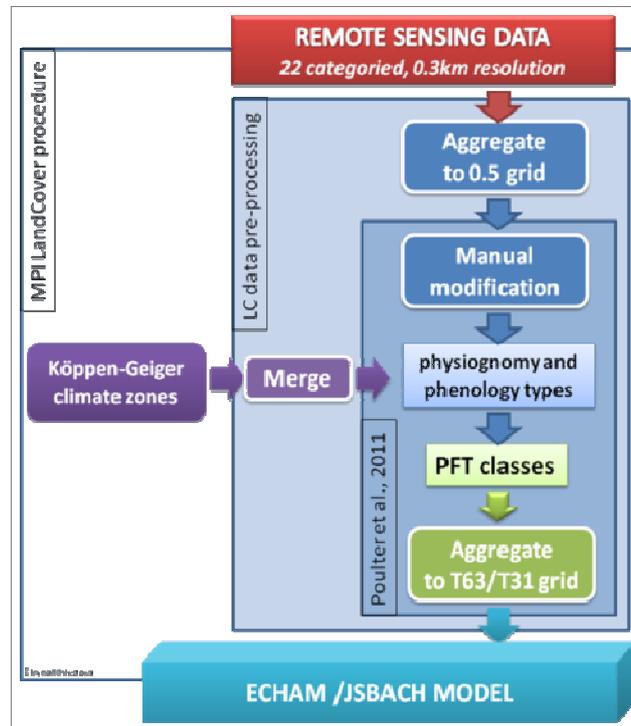


Figure 4.8.2: Schematic representation for the conversion procedure as performed for MPI input generation. The main part of this is adapted from Poulter et al., 2011. Main modifications are made to satisfy JSBACH input specific. File conversion and aggregation: a) Convert from TIFF to netCDF format; b) Tile by type. Reducing resolution to 0.5x0.5 (in order to apply other satellite datasets for further conversion, e.g. KG Biomes classification: i) Reclassification (slightly modified schema by Poulter et al., 2011); ii) reduce to general types (forest, herbac, crop); iii) apply biome mask (Climate classification) scale not used types (e.g. anthrop. water on land); iv) regrid to Model Resolution (T63)



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

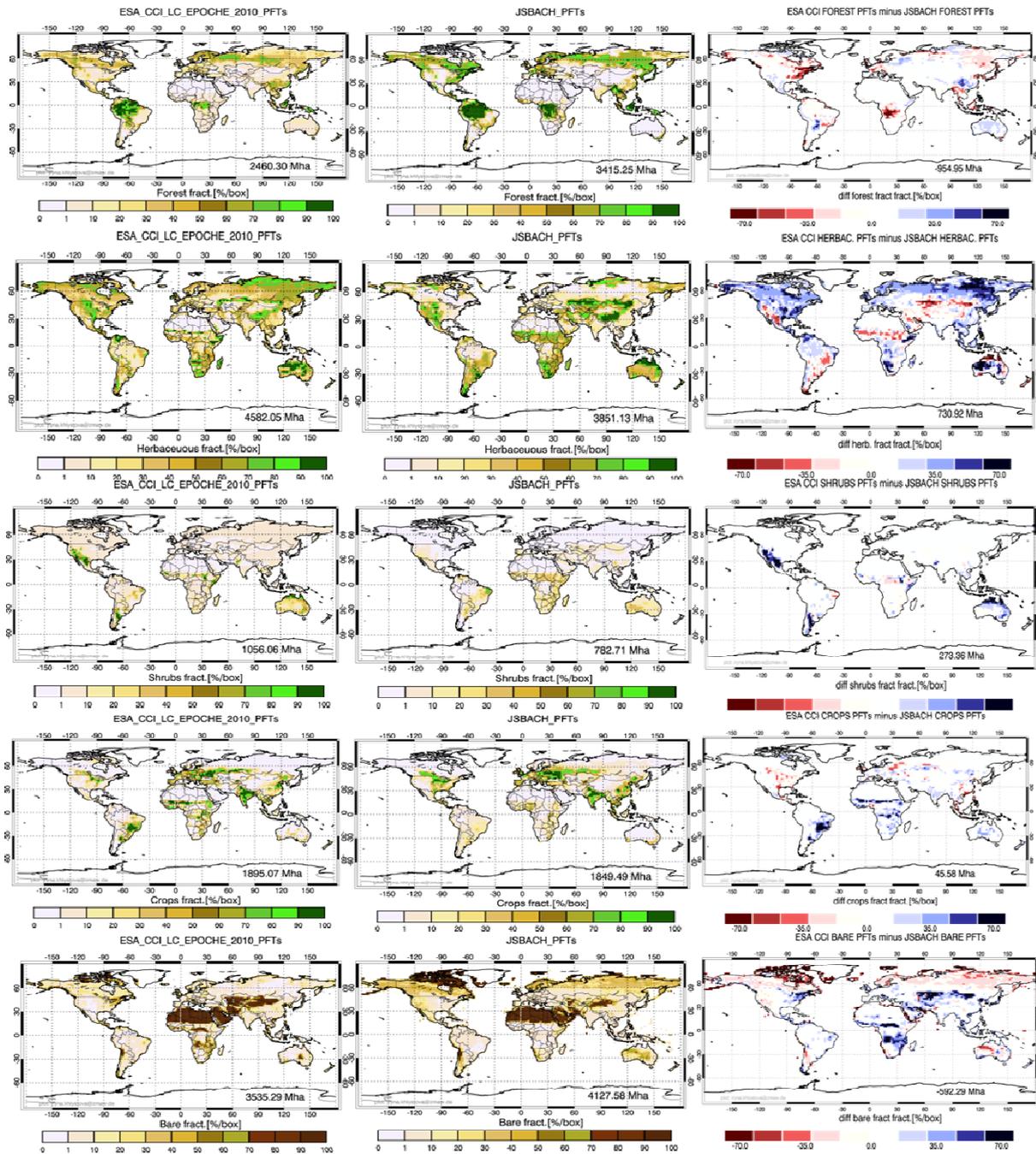


Figure 4.8.3: The differences in the main PFTs – the sum of similar functioning (up to down) forest, grasses, shrubs and bare areas. On the left line the ESA CCI PFTs, on the middle- the JSBACH original, and on the right side – the absolute differences



CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

4.8.3.2 Experiment setup

A multitude of experiments was conducted to assess the effect of the ESA CCI land cover dataset in the MPI-ESM. Coherent simulations were conducted using the original JSBACH PFT maps, GlobCover as a CCI precursor as well as data from all 3 epochs of the ESA CCI LC dataset. Table 4.8.2 summarizes all the experiments. These comprise coupled (land-atmosphere) simulations as well as offline (land-only) simulations with two different forcing datasets. While the coupled simulations quantify the effect of the different land boundary conditions on climate simulations, the offline simulations allow one to assess the effect on biogeophysical variables only without feedback from the atmosphere. Using two different forcing datasets allows further to quantify the relative importance of the forcing data compared to the land-boundary conditions.

Simulation	Climate Forcing	Input PFTs	Simulation years
WFDEI-REF	WFDEI	JSBACH	1979-2000
WFDEI-GCV	WFDEI	GlobCover	1979-2009
CRU/NCEP-GCV	CRU/NCEP	GlobCover	1979-2009
WFDEI-CCI 2000	WFDEI	ESA CCI Epoch 2000	1979-2009
WFDEI-CCI 2005	WFDEI	ESA CCI Epoch 2005	1979-2009
WFDEI-CCI 2010	WFDEI	ESA CCI Epoch 2010	1979-2009
ECHAM6-REF	coupled	Reference	1979-2009
ECHAM6-CCI	coupled	ESA CCI Epoch 2010	1979-2009

Table 4.8.2: Overview of MPI-ESM experiments

4.8.3.3 Model benchmarking

To provide an independent assessment of the quality of the obtained simulation results, an independent model benchmarking is performed by comparing the model simulations against independent observations of surface radiation fluxes as well as near surface temperature data (2m). Table 4.8.3 provides an overview about used observational datasets to evaluate the model performance. As a major model skill metric we use a temporally and spatially weighted root mean square difference (RMSD) statistic based on monthly input data as (Gleckler *et al.*, 2008):

$$E^2 = \frac{1}{W} \sum_t \sum_j \sum_k w_{i,j,k} (y_{i,j,k} - x_{i,j,k})^2 \quad (4.8.1)$$

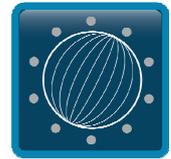
where W is the sum of the weights w which account for the different area sizes of a model grid cell as function of geographic longitude and latitude (i,j dimensions) as well as different length of the months (t dimension). A relative performance metric is defined to enable comparison of model simulations against observations across different variables. For a given variable a typical model error E^i is defined as the median of all RMS error estimates for all models (\mathcal{Y}_k). The median value is used to minimize effects of potential outliers. The relative model error (E'_k) for a particular model is calculated as

$$E'_k = \frac{E_k - \bar{E}}{\bar{E}} \quad (4.8.2)$$

For each observational dataset, one obtains a relative measure to contrast a model against all other models as well as an average model ($\bar{y} = \langle y \rangle$), where brackets indicate the expectation

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



operator. For example, if E^f_k has a value of 0.3, then the RMS error of the model is 30% higher than on average.

Results are illustrated in a portrait diagram, first introduced by Gleckler *et al* (2008) (Figure 4.8.4). Rows in this diagram correspond to different CMIP5 models and columns to different variables. The value of the relative model skill (2) is illustrated in triangles, where the position of each triangle corresponds to a different observational dataset. The advantage of this diagram is that it allows to directly compare the relative performance of the different models, but also to take into account uncertainties from the different observational dataset. If a model is superior to the average model for all observational dataset ($E^f_k < 0$), then it is a robust indicator for a good model performance. In contrast, if a model shows positive as well as negative deviations from the mean model, then it is an indication that the uncertainties in the observations are probably larger than the actual deviation of the model to the observations.

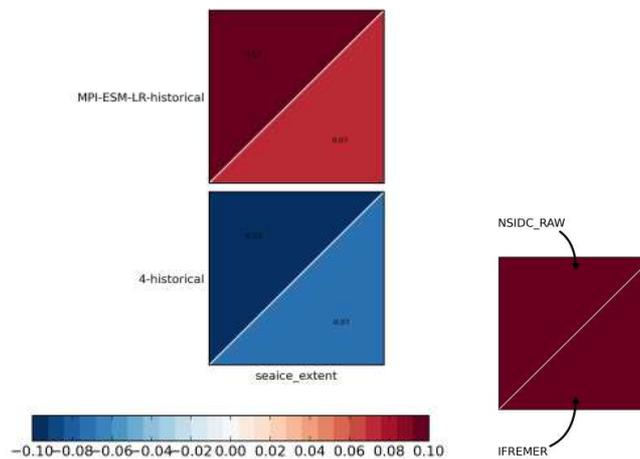


Figure 4.8.4: Example of Portrait diagram (rows=models, columns = variables, triangles=skill score for different observational datasets)

Variable	Observation	Data provider
Surface albedo	MODIS v05	NASA
	CLARA SAL	CMSAF, EUMETSAT
	Globalbedo	ESA
Surface downwelling solar radiation flux	CERES v2.7	NASA
	SRB v3.0	NASA
	ISCCP	NOAA
Surface solar upward flux	CERES v2.7	NASA
	SRB v3.0	NASA
2m temperature	WATCH	EU FP7
	NCEP	NCAR
	CRU 3.0	University of East Anglia

Table 4.8.3: Variables and observational datasets used for independent model evaluation

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

4.8.4 Results

This section provides a very condensed overview about the results obtained and is guided by the major research questions formulated in section 4.8.1.

4.8.4.1 Impact of CCI land cover data on surface water and energy fluxes

How does an integration of ESA CCI Land cover data affect the energy and water fluxes at global scales?

The impact of the integration of the ESA CCI dataset into the MPI-ESM was evaluated by first investigating the impact of a change in the PFT distribution on biogeophysical surface parameters and associated land surface fluxes. This was done using the offline simulation experiments and compared results from the following simulations: WFDEI-REF, WFDEI-GCV, WFDEI-CCI 2005, CRU/NCEP-GCV. The analysis was conducted for 19 different land surface variables and fluxes. The climatological mean fields were analyzed on monthly and seasonal timescales and absolute and relative differences were estimated to investigate in which cases the difference land cover datasets result in different land surface conditions. Figure 4.8.5 shows example maps of relative differences between two simulations for leaf area index and evaporation.

Results for all land surface variables are summarized in Figure 4.8.6 which shows a summary of the global relative differences for all investigated variables. The relative differences are illustrated in Violin plots which illustrate the probability density of the different relative differences as well as the median and inner quartile ranges of the data.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

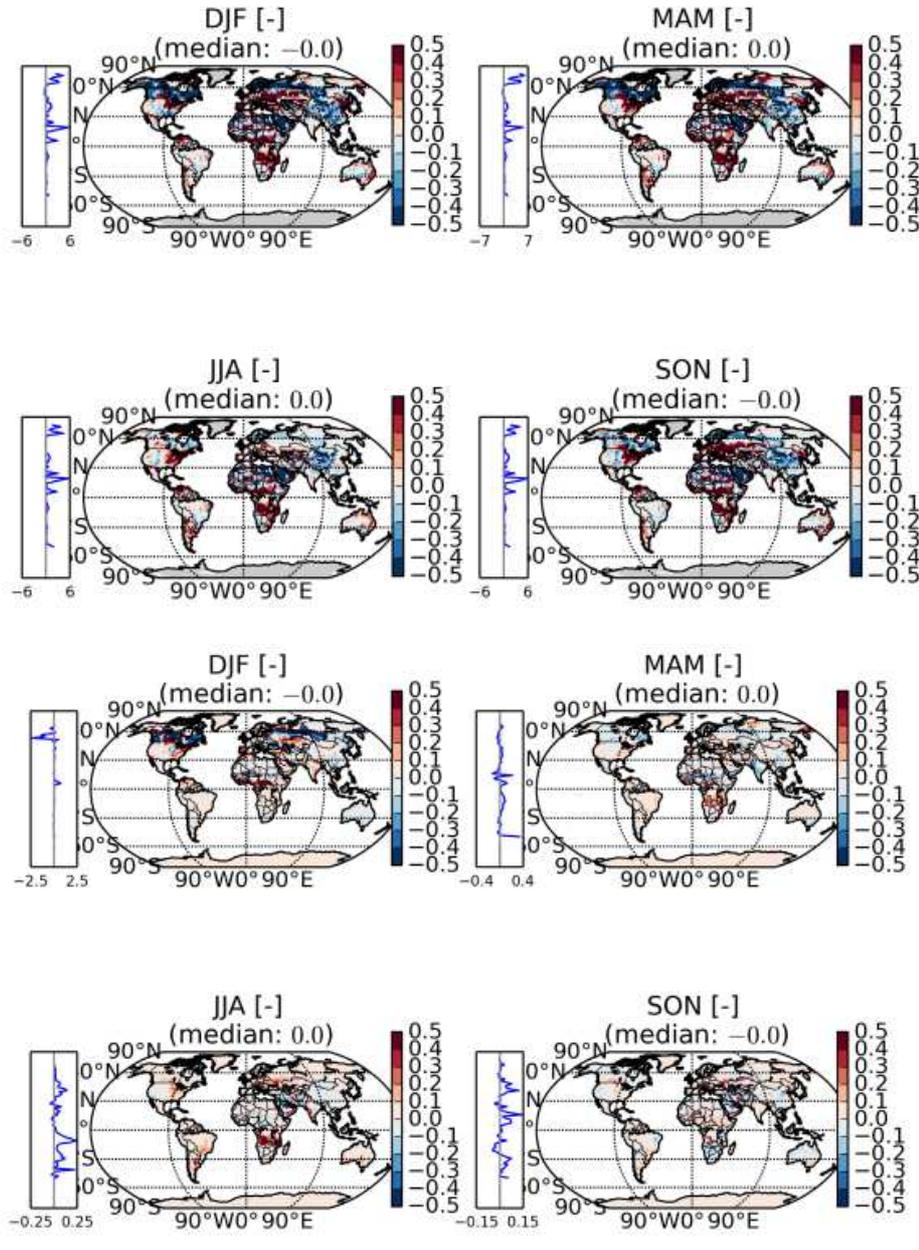
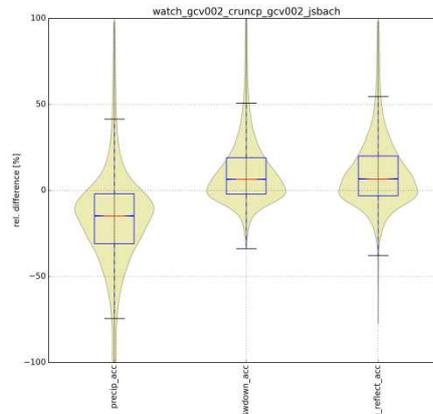
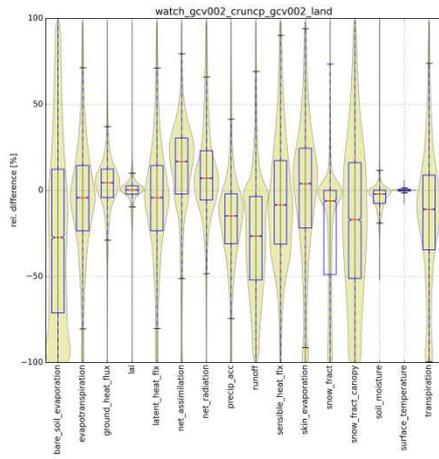


Figure 4.8.5: relative differences between CCI and Globcover for a) leaf area index and b) evaporation

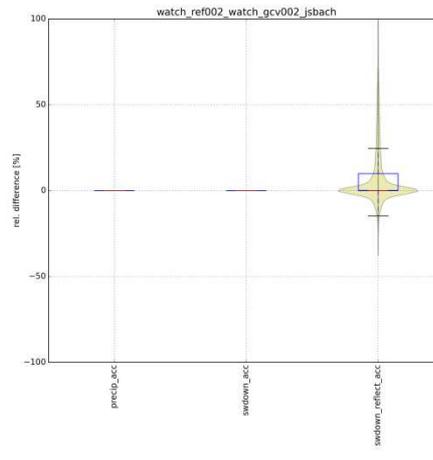
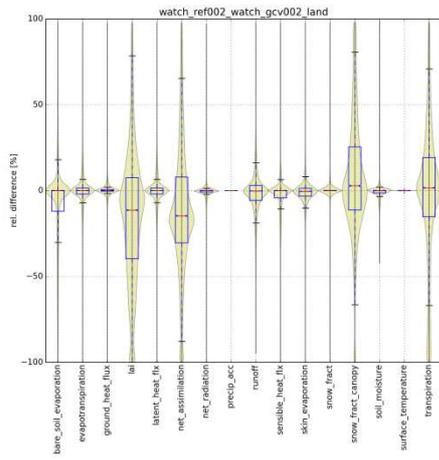


CMUG Deliverable

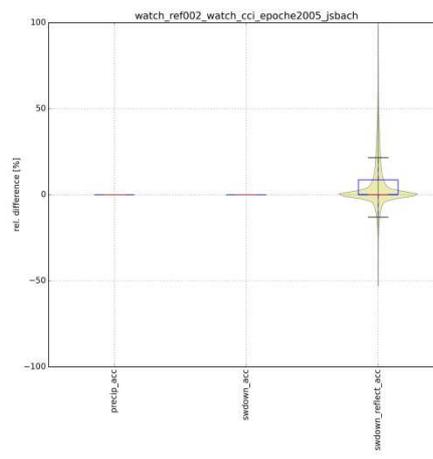
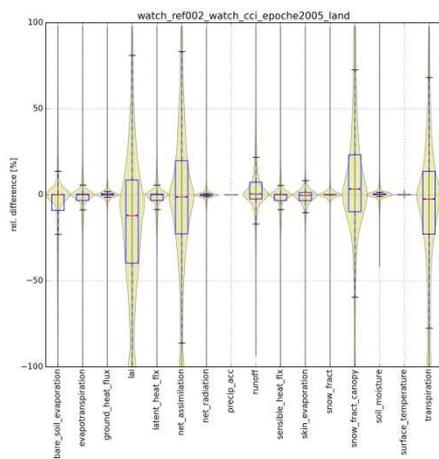
Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



a)



b)

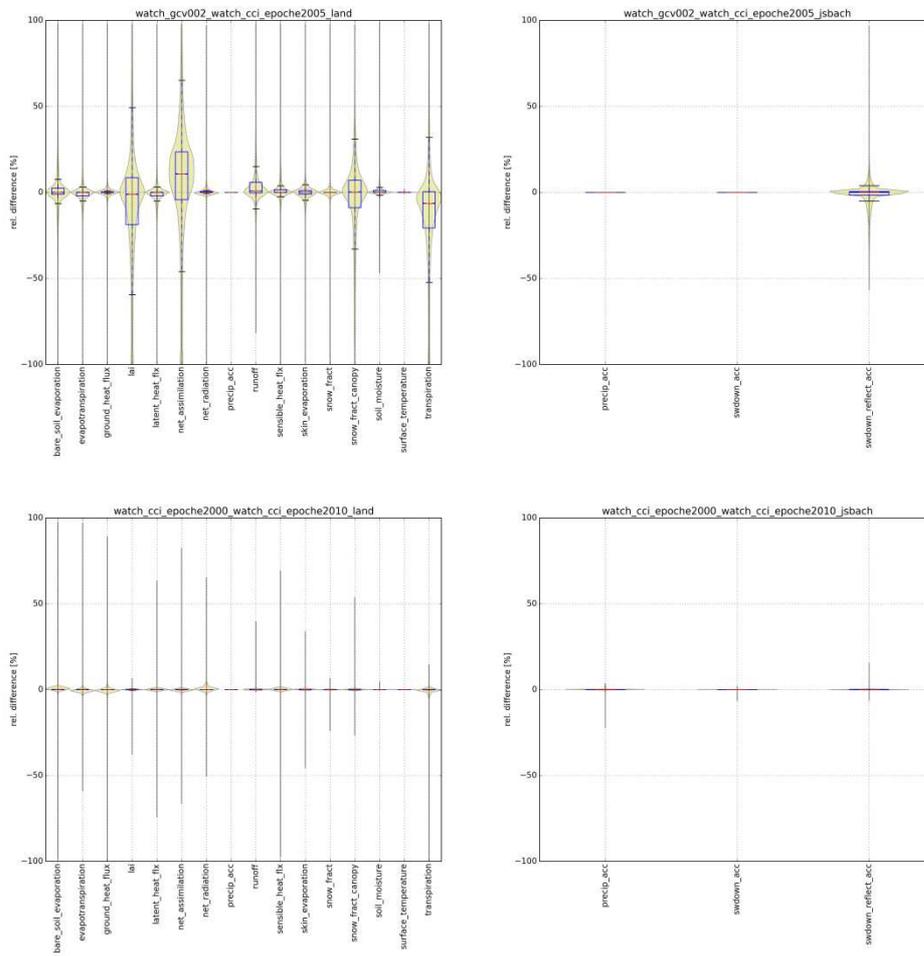


c)



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



d)

e)

Figure 4.8.6: Violin plots of relative differences for investigated parameters between a) WATCH and CRU forcing data (for GCV), b) CTRL and GCV, c) CTRL and CCI, d) GCV and CCI, e) CCI2000 and CCI2010. The Violin plots illustrate the data density for the different relative difference values. Box plots indicate the median value and the inner quartile range of the data.

The Violin plots illustrate clearly where larger differences exist between the different model simulations for all investigated variables. Largest differences are obtained for the comparison of the WATCH forcing and CRU/NCEP forcing based datasets. Major differences for all investigated variables are observed. In contrast changing the land cover information results in much smaller differences. The differences between the CTRL simulations and the simulations using GCV or CCI land cover information are most relevant for leaf area index, snow fraction, transpiration, runoff and net carbon assimilation.

Independent evaluation

An independent evaluation of the model experiments was conducted for surface albedo, surface solar downward radiation flux (SIS) and the (reflected) surface upward solar radiation flux using independent observational data (Table 4.8.3). The relative model skill score for surface albedo using the ESA Globalbedo dataset as a reference is shown in Figure 4.8.7: Relative model skill score for surface albedo compared to ESA Globalbedo. It illustrates that the

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



usage of ESA CCI land cover data results in surface albedo estimates which have an RMSE which is ~2.5% lower than the average RMSE. The relative improvement compared to Globcover is in the order of 7%. Largest RMSE is observed for the simulations using the CRU/NCEP forcing data. However, one needs to emphasize that the global skill score is only a first indicator for model performance and a thorough analysis of the regional and temporal pattern of the differences between models and observations is required.

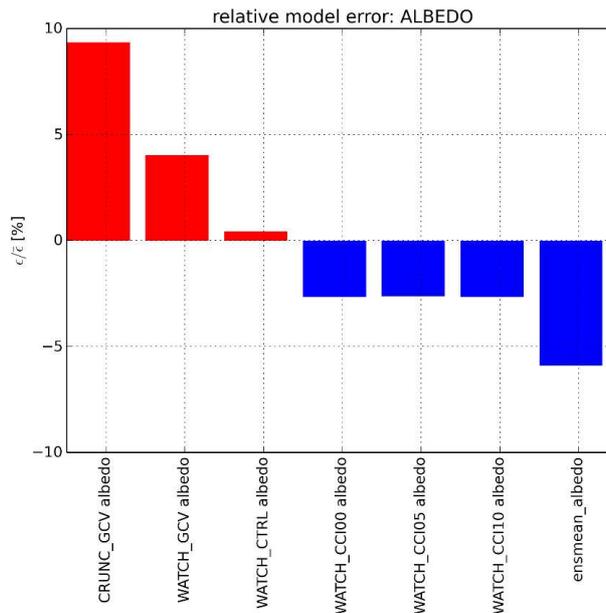


Figure 4.8.7: Relative model skill score for surface albedo compared to ESA Globalbedo

The complete relative model skill score results are illustrated in Figure 4.8.7. These indicate largest errors for the CRU/NCEP based simulations. In general, the usage of CCI dataset results in a slight improvement compared to the usage of Globcover or the reference simulations (ctrl). It needs to be however emphasized that the mean of all experiments (mean-model) outperforms all other experiments and that the differences between the different observational datasets result also in a different model ranking.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

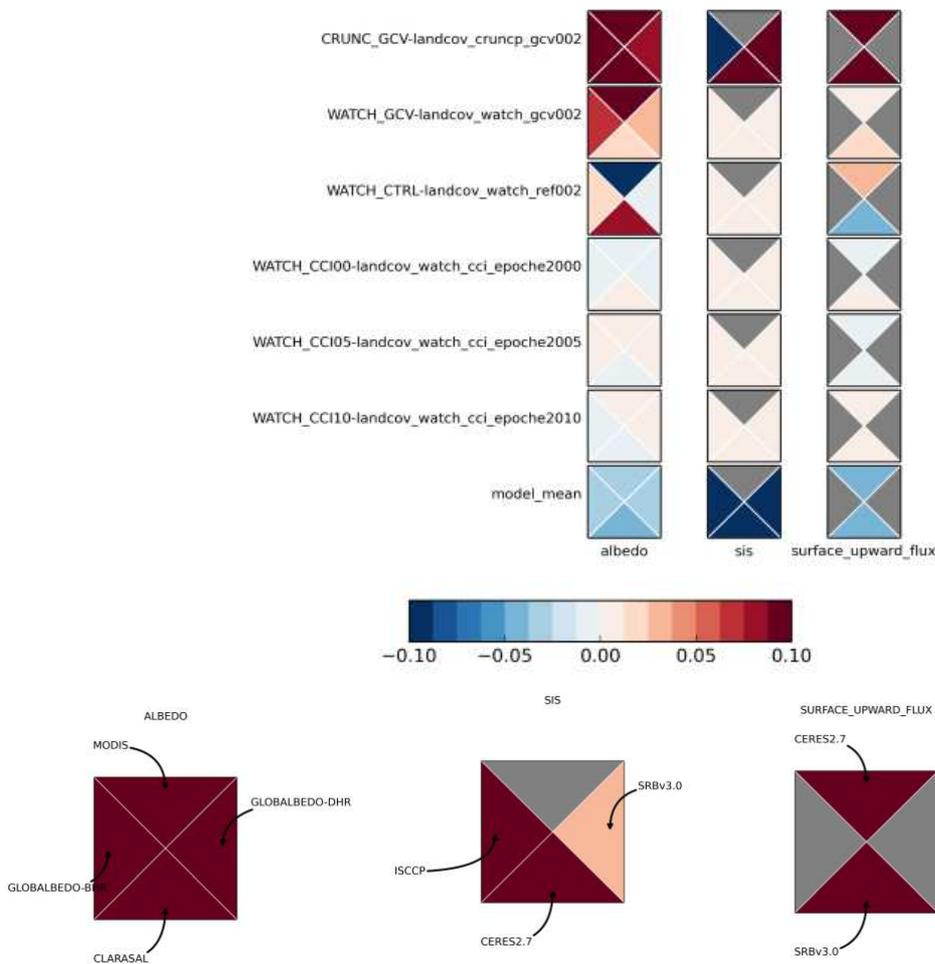


Figure 4.8.8: Relative model skills for the different investigated surface radiation flux components. Relative model performance compared to the median model. A value of 0.1 corresponds e.g. to a 10% worse skill than on average.

4.8.4.2 Impact on coupled climate simulations

Does the integration of ESA CCI land cover data improve the skill of MPI-ESM in simulating present day climate?

To evaluate the impact of CCI land cover and Globcover on the MPI-ESM simulated climate, coupled simulations were performed and evaluated using different observational records. The comparison with global temperature datasets was done for land surfaces only. Figure 4.8.9 shows an example of the absolute and relative temperature differences of the models simulations with/without CCI data as compared to the CRU land surface temperature. Only very minor differences are observed in the temporal mean air temperature fields. However, the differences are changing throughout the year, which is taken into account when calculating the model benchmarking skill score (eq. 4.8.1) and illustrated in Figure 4.8.8.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

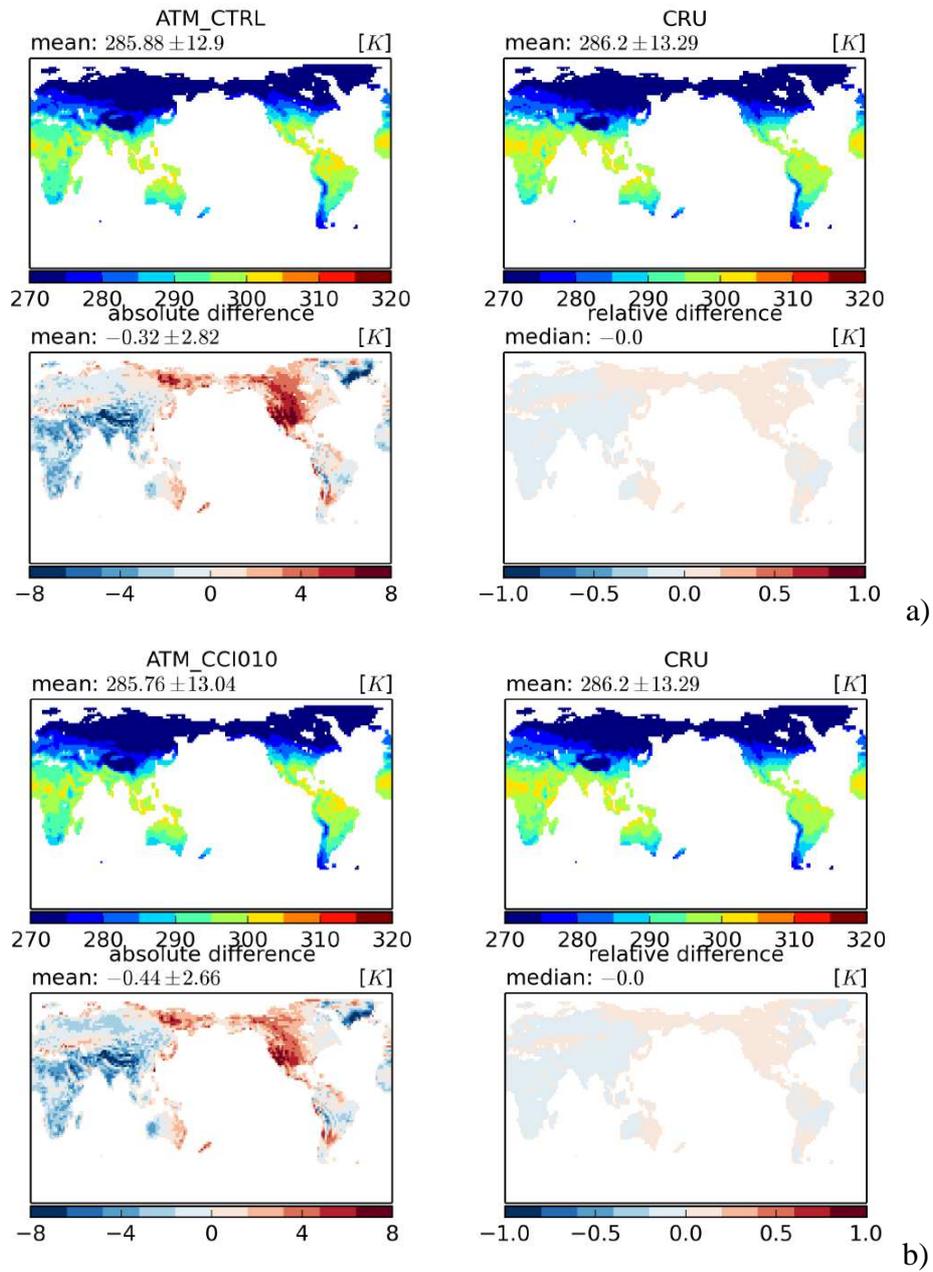


Figure 4.8.9: Climatological mean land surface temperature fields for CTRL simulations (a) and simulations using ESA CCI land cover data (b) as compared against CRU 2m temperature data.

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



Independent evaluation

Figure 4.8.10 shows the relative skill score diagram for the coupled model simulations. In addition to the aforementioned surface radiation fluxes, the surface temperature can be analyzed in coupled simulations. Overall the experiment with the ESA CCI land cover data (ATM_CCI010) shows a slight improvement compared to the Globcover and control simulations. The differences are however in the order of 1.5% ... 3%. It needs to be emphasized that this difference is not expected to be statistically significant as the internal variability within the model is still rather large. The CCI based simulation also shows a coherent improvement in the surface solar radiation flux for all observational datasets. These differences are however also small.

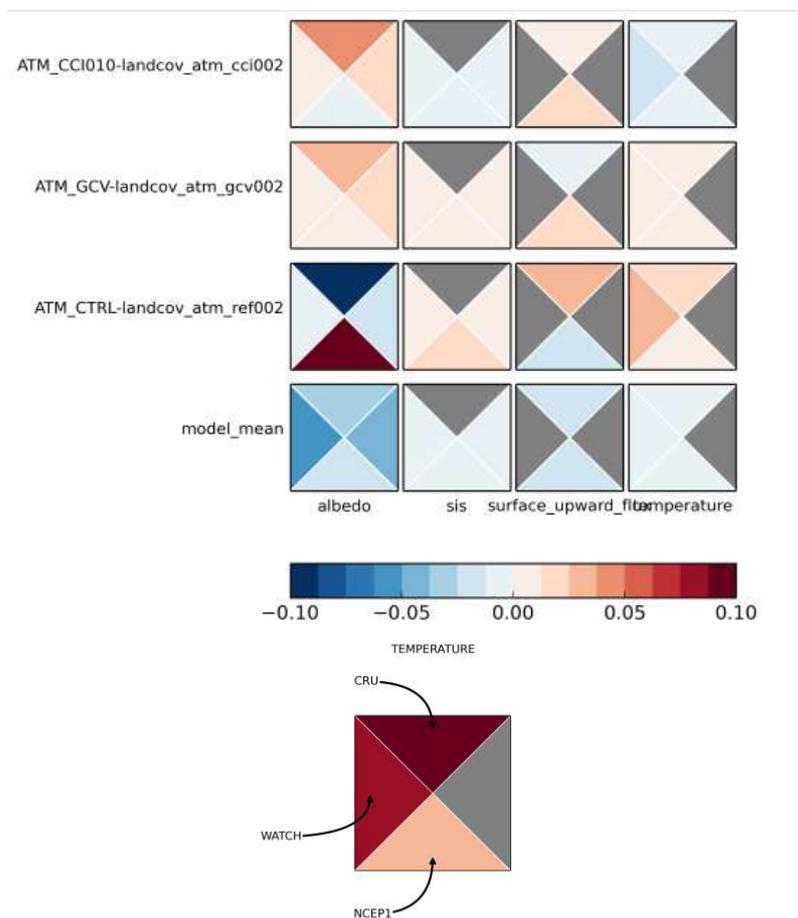


Figure 4.8.10: Relative model skill for coupled model simulations. Legend is the same as for offline simulations (Figure 4.8.), except for temperature.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

4.8.4.3 Overall summary and conclusions

Is the usage of ESA CCI Land cover data superior compared to the usage of previously existing land cover products like e.g. ESA GlobCover? Added value of CCI LC?

The major results of the conducted experiments can be summarized as follows:

- ❖ The usage of different state-of-the-art meteorological forcing datasets (WATCH, CRU/NCEP) results in the largest differences of land surface variables and fluxes. The usage of the forcing data is thus of a much higher relevance than the replacement of the land cover information.
- ❖ The usage of CCI land cover data results in significant changes for important land surface state variables compared to the CTRL simulations.
- ❖ Differences between ESA CCI land cover and its precursor product (Globcover) are smaller than the differences between the CTRL land cover and the ESA CCI land cover.
- ❖ The differences between the different ESA CCI epochs are marginal compared to the impact of changing from the CTRL land cover to the CCI land cover or considering the much higher sensitivity to the forcing data used.
- ❖ Using CCI land cover data results in slightly improved skills in simulating surface albedo as compared against different observational surface albedo data products. The relative improvement is in the order of 2.5% ... 7%.
- ❖ Using CCI land cover data slightly improves the skill in simulating global land surface 2m temperature. The relative improvement is of the order of 1.5% ... 3% as compared against different observational datasets.

It is however emphasized that the coupled model simulations include a large degree of internal variability and that the statistical significance of the temperature simulation skills are marginal compared to the internal model variability.

4.9 Fire burnt area

The burned area fire CCI dataset is assessed by CMUG by prescribing it as a *boundary condition* in a global dynamic vegetation model (JSBACH) as part of the MPI Earth System Model (MPI-ESM). The assessment has been done with a precursor dataset (GFEDv3, Giglio et al. 2010, van der Werf et al, 2010) covering the time period 1997 to 2010 because the CCI dataset was not available at the time this report was written.

4.9.1 Fire in the Earth System

Fire is an important Earth System process, which impacts climate via multiple pathways, including atmospheric chemistry, aerosols, global vegetation patterns, land surface albedo and the carbon and nutrient cycles. At the same time fires are controlled by climate and the frequency of fires is expected to increase with future climate change. As such, fires form a complex feedback cycle in the Earth system which potentially forms an important contribution to the climate sensitivity of the Earth System. The net effect of fires on the climate system remains unclear as depending on the process fires can cool or warm the Earth System (Bowman et al, 2009).



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

4.9.2 Fire in MPI-ESM

Within JSBACH fire is an important perturbation that impacts the vegetation distribution and the carbon cycle. Fires are in the standard setup of JSBACH simulated dynamically with a fire algorithm of intermediate complexity. Long term satellite observations allow for the first time to prescribe burned area as a boundary condition in a global vegetation model. While this improves the representation of fires in the ESM for present day conditions it also allows to use the vegetation model to translate the satellite derived burned area information into fire emissions of trace gases and aerosols.

To simulate the impact of a fire on the carbon cycle the amount of carbon emitted from fires is parameterized as a function of the burned area and the available fuel load simulated in the vegetation model. Together with land cover dependent emission factors, derived from observational data reported in the literature, the simulated carbon emissions can be further related to trace gas and aerosol emission (Figure 4.9.1).

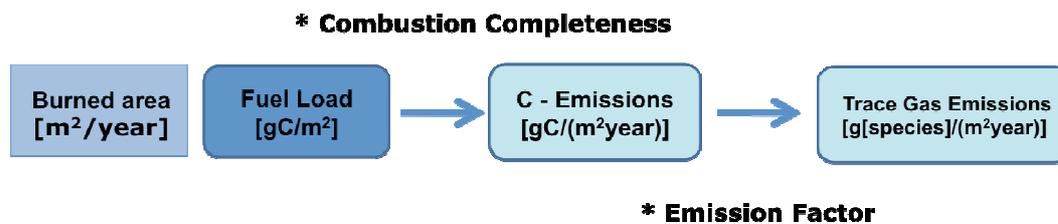


Figure 4.9.1: The conversion of burned area reported in the CCI dataset to the emissions of trace gases in the global vegetation model JSBACH.

4.9.3 Prescribing satellite fire products in MPI-ESM

To assess the impact of the fire CCI dataset on MPI-ESM, an interface was developed which allows burned area to be prescribed from an external data source in the model as a boundary condition, instead of simulating it interactively. The interface is combined with a pre-processing procedure that converts the input data into a format suitable for JSBACH. As the definition of land and ocean boundaries typically differ between models and satellite data products, the interface includes a consistent treatment of the land-sea mask in the input data as well as in the JSBACH model.

As a precursor dataset for the fire CCI the Global Fire Emission Database (GFED, version 3) was chosen (van der Werf *et. al.*, 2010, Giglio *et. al.*, 2010). GFEDv3 reports burned area for the time period 1997 – 2010 on a monthly basis with a spatial resolution of 0.5 x 0.5 deg. For the application in JSBACH the burned area was mapped to the current model standard grid resolution of T63 (~1.8 x 1.8 deg).

Spin-up procedure

As the carbon cycle is characterized by long timescales, such as the carbon uptake in the soil, vegetation model simulations require long spin-up periods before an equilibrium state in the carbon cycle is reached. Equilibrium states are typically achieved for pre-industrial conditions, after which the transient behaviour of the natural system is accounted for, to derive present-day conditions that are comparable to present-day observations.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Long term satellite based burned area products do, however, not represent the timescales necessary to cover observed fire frequencies globally. Some biomes can reach up to one fire occurrence in one thousand years. This information is, however, required to bring the model into an equilibrium state. To overcome this limitation we followed the same approach as in van der Werf et al. (2010) to extrapolate the satellite based observed burned area for the time period 1997 to 2010 further in space using a regionally specific tree cover depending on the remapping procedure, to derive a climatological mean burned area dataset based on present day satellite observations. This climatological mean burned area was used in the spin-up procedure of the vegetation model for climate conditions representative for the 1850s as well as in the following transient simulation up to the year in which satellite based burned area data becomes available (1850 -1996).

Fuel consumption

To parameterize carbon emissions from fires, several assumptions have to be made to relate the simulated fuel load in the vegetation model to the satellite based observed burned area. A vegetation model simulates vegetation biomass in terms of the biomass carbon content. The carbon uptake of plants through photosynthesis is allocated to the biomass, which is distinguished into different strata: wood, green component of living biomass and litter. The vegetation model, thereby distinguishes between 17 plant functional types summarizing different vegetation classes.

Fires lead to combustion of biomass with a carbon release to the atmosphere. Not all biomass affected by a fire is combusted in a fire. Part of the living vegetation is resistant to fire, i.e. it is not affected at all by a fire. Another part gets killed during a fire but is not combusted. The biomass of the killed vegetation is converted from living biomass to dead biomass. In the model this is accounted for by a PFT-dependent combustion completeness factor and a fire mortality. In the present study we tested two fuel consumption approaches (Table 4.9.1): (i) fuel consumption of the standard interactive fire model of JSBACH, (ii) fuel consumption parameterization closely following the approach described in van der Werf et al. (2010).

Burned area distribution among the different PFTs

GFEDv3 provides in addition to the satellite based burned area dataset information on how much of the burned area was observed in tree covered regions and how much in grass covered regions. This information is utilized in the model to distribute the burned area among the different PFTs present in a grid box. The PFTs distribution in a grid box is prescribed in JSBACH using a landcover distribution map as boundary conditions. The information presented in GFEDv3 is based on an overlay of VCT (Hansen et al. 2003) and the observed burned area. Uncertainties in the VCT fields as well as in the land cover information might lead to a mismatch between forest or grass fires reported in GFEDv3 and actual available forest or grass areas presented in JSBACH. In these cases the burned area in the model is attributed first to the vegetation type as reported in GFEDv3 and if the area is not sufficient the remaining area is distributed equally among the vegetation types representing the other class.



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

In the present analysis we further tested, the effect of including pasture in the burned area or not (Table 4.9.1). If pasture areas are excluded from fires the area reported as grass areas in GFEDv3 exceeds in some regions the available grass area in JSBACH. In these cases forested areas are burned instead. If pasture, however, is allowed to burn, the grass plus pasture area in JSBACH is sufficient to account for the grass fires reported in GFEDv3 in almost all regions. In additional experiments we did not use the burn class information in GFEDv3, but distribute the burned area equally among all PFTs present in a grid box, to demonstrate the impact of including this additional satellite based observational record into the overall assessment (Table 4.9.1).

Experiment	Fuel Consumption	Pasture Burning	Burned area distribution	Fire Carbon Emissions [PgC/year]	Burned area Forest [Mha/year]	Burned Area Shrub Land [Mha/year]	Burned Area Grassland [Mha/year]
EXP1	JSBACH	Yes	Equally	3.85	171	83	191
EXP2	JSBACH	No	Equally	3.80	310	186	52
EXP3	GFEDv3	Yes	Equally	2.71	171	83	191
EXP4	GFEDv3	No	Equally	2.04	310	186	52
EXP5	JSBACH	Yes	GFEDv3	3.01	80	41	283
EXP6	JSBACH	No	GFEDv3	2.71	143	72	219
EXP7	GFEDv3	Yes	GFEDv3	2.14	80	41	283
EXP8	GFEDv3	No	GFEDv3	1.54	143	72	219

Table 4.9.1: Different experiments performed for the current analysis. The different experiments make different assumptions about the fuel consumption parameterization: (i) using the standard JSBACH parameterization (ii) using a similar approach as applied in GFEDv3; different assumption on whether pasture is allowed to burn (Yes), or not (No); and on the distribution of the burned area among the PFTs present in one gridbox (equally or prescribed according to the satellite based information provided with GFEDv3). The fire carbon emissions are global averages over the time period 1997 – 2010.

4.9.3 Results

While all experiments use the same climatological mean burned area as boundary condition and the same climatological forcing, the different assumptions on burned area distribution among the PFTs, the different fuel consumption parameterization, as well as the inclusion or exclusion of pasture burning leads to distinct differences in the simulated fire carbon emissions.

Grass and forest fires

The distribution among the different PFTs depends in the JSBACH model on the prescribed land cover map. The land cover map, thereby, changes with time in accordance with land use change. In the present study land cover change is prescribed annually following Hurtt *et al.* (2009). Furthermore, differences in the distribution of the burned area among the PFTs (equally or in accordance with GFEDv3) lead to a different share between grass or forest fires. The resulting burned area in grass covered and forest covered regions for the different experiment is shown in Figure 4.9.2.

CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

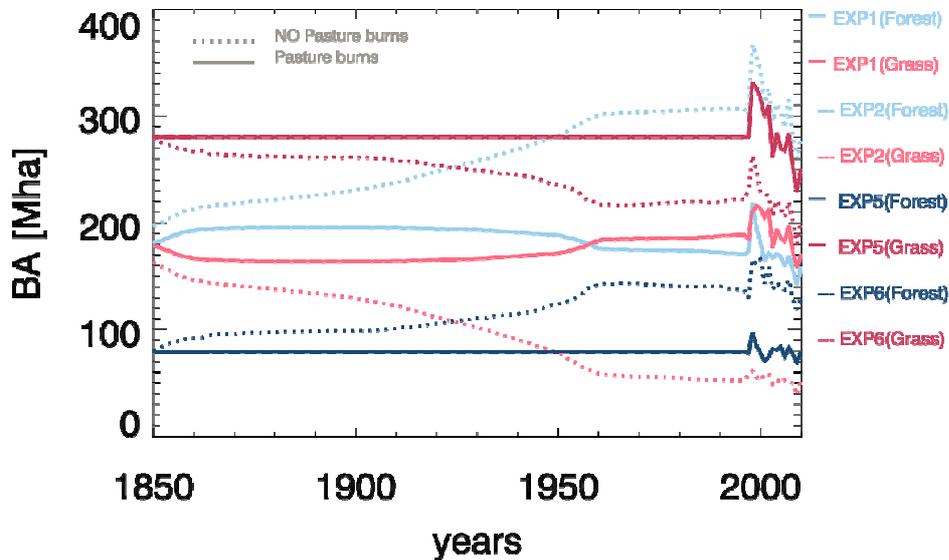


Figure 4.9.2: Global annual burned area for the different experiments (see Table 4.9.1) in grass and forest covered regions. Exp3, Exp4, Exp7, and Exp8 are not shown. For these experiments the burned areas in grass and forest covered regions are identical to Exp1, Exp2, Exp5 and Exp6, respectively. Note, that the total global burned area (grass plus forest) is identical in all Experiments.

In case the burned area is equally distributed among the PFTs present in one grid box (Exp1 and Exp2) the burned area in forested regions and in grass covered regions is very similar on the global annual mean (between 180 and 200 Mha). If pasture is allowed to burn (Exp1) the share between global annual burned area in forest and grass covered regions remains almost constant between 1850 and 1996. If pasture is not allowed to burn (Exp2), however, the burned area in forested regions strongly increase between 1850 and 1996 as with increasing land use change more grass areas are converted into crop or pasture and therefore are not available for burning anymore. The increase in forest fires is thereby accompanied by a decrease in grass fires. If the burned area is distributed among the PFTs following the information of grass versus tree burning provided by GFEDv3 (Exp5, Exp6) the global annual burned area in forested regions is much smaller compared to the experiments in which the distribution is done equally among the PFTs present in a gridbox (~80 Mha compared to ~180 Mha for 1850). This implies that a simple assumption of an equal share of the burned area among the vegetation types present, strongly overestimates the burned area in forested regions and subsequently underestimates the burned area in grass regions. If pasture burning is allowed in the model (Exp5) the burned area in forest and grass covered regions stays constant between 1850 and 1996. If pasture areas are, however, excluded from burning (Exp6) the forest burned area increase between 1850 and 1996 at the expenses of grass fires, which decrease. This is again related to land cover change, which converts mainly grass areas into pasture and crop areas and thereby reduces the area available for grass fires.

In reality land cover change impacts the total burned area and not only the share between forest and grass fires as in our experiments in which we prescribe a constant climatological mean burned area. The change in burned area over the historical period is still very uncertain and can not be derived from observational data alone on a global scale. Available historical



CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

burned area information does, however, suggest that pre-industrial fire occurrence was higher due to less anthropogenic controlled land surfaces (Marlon et al., 2013). For the present study this is not accounted for.

Global annual mean fire carbon emissions

Figure 4.9.3 shows the global annual mean fire carbon emissions as simulated for the different experiments for the time period 1850 to 2010. In addition a zoom into the observational time period 1997 – 2010 is shown and compares the experiments that allow pasture to burn to values reported in the GFEDv3 database.

The experiments using the standard JSBACH fuel consumption parameterisation (Exp1,2,5,6) shows in general higher fire carbon emissions than the one following the GFEDv3 approach (EXP3,4,7,8). If pasture is allowed to burn the carbon emissions are higher than in experiments in which pasture is not allowed to burn (compared Exp 2,4,6,8 with Exp 1,3,5,7 respectively). In case the burned area is distributed equally among the PFTs present in a grid box the fire carbon emissions are higher compared to the distribution following the GFEDv3 approach (compare Exp1,2,3,4 with Exp5,6,7,8, respectively).

For the present day period the simulated carbon emissions for the different experiments fall around the ones reported in the GFEDv3 database. While the experiments using the JSBACH standard fuel consumption parameterization (Exp1 and Exp5) show higher values than the GFEDv3 estimates, the experiments using the GFEDv3 fuel consumption parameterization show lower values. The Exp7 is from the simulation setup the one closest to the GFEDv3 approach using the same burned area, the same distribution of the burned area between forest and grass covered regions and a similar parameterization for the fuel consumption. On average Exp7 experiment results in 2.1 PgC/year whereas GFEDv3 reports 2.0 PgC/year. The interannual variability in the JSBACH experiments and in the carbon emissions reported in GFEDv3 follow each other very closely, as it is largely controlled by the underlying burned area dataset. However, the drop in carbon emissions from the peak in 1997 to a minimum in 2000 is less pronounced in the JSBACH experiments compared to GFEDv3. The spatial variation of the fuel load thereby partly uncouples the global burned area variability from the variability simulated in the global fire carbon emissions.

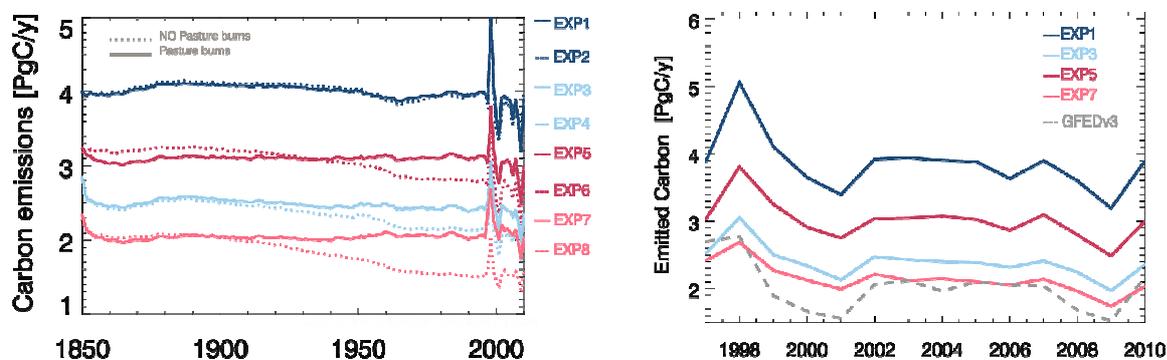


Figure 4.9.3: Global annual mean fire carbon emissions for the different experiments (see Table 4.9.1) for the time period 1850 to 2010 (left) and a zoom for selected experiments into the timeperiod 1997 to 2010 in comparison to the values reported in GFEDv3.



CMUG Deliverable

Number: D3.1
 Due date: March 2014
 Submission date: 1 April 2014
 Version: 2.1

Figure 4.9.4 shows the global distribution of the annual mean fire carbon emissions averaged for the observational period 1997 to 2010 as reported in GFEDv3 compared to the Exp7. While the global totals are relatively similar, the spatial distribution shows distinct differences.

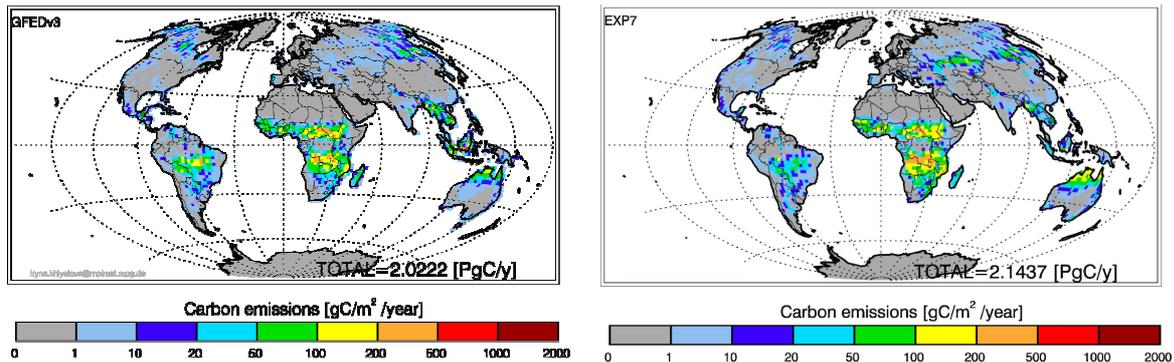


Figure 4.9.4: Annual mean fire carbon emissions averaged over the time period 1997 to 2010 as reported in GFEDv3 (left) compared to Exp7 (right).

Exp7 shows lower values over Southern America, whereas central Asia and Northern Australia show higher values.

4.9.4 Summary of results

In summary, the implementation of the satellite based burned area product allows to relate burned area to the emissions of carbon from fires into the atmosphere. These build valuable boundary conditions for atmospheric chemistry and aerosol studies. So far burned area was used as boundary condition in a global vegetation model only in the GFEDv3 assessment, which applied the CASA model. Here, we use a different model (JSBACH) in a similar approach to analyse the impact of the driving vegetation model on the resulting fire carbon emissions. We performed a range of sensitivity experiments to test different parameterizations for fuel consumption and different assumption on how to distribute the burned area among the vegetation types represented in the vegetation model. Overall we find a large range in fire carbon emission for the time period 1997 to 2010 for the different experiments performed (1.54 to 3.85 PgC/year). The experiment closest to the setting applied in the GFEDv3 assessment leads globally to very similar annual fire carbon emissions (2.0 PgC/year compared to 2.1 PgC/year). Of uttermost importance is thereby the information on the vegetation types that burn, which is provided together with the satellite based burned area product. Without this information the burned area in the model has to be equally distributed among the present vegetation types. As a result the fire carbon emissions are overestimated globally by 28 to 40%.

The modelling framework outlined above has been set up in such a way that the burned area boundary dataset can be easily changed and will be applied to the CCI fire product as soon as it becomes available.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

5. Summary of assessments from a climate modelling perspective

A selection of CDRs from the CCI ECVs have been independently assessed by the CMUG and reported on here. Comments on the integrity of the data are provided as well as a scientific evaluation. The latter has been done in various different ways as listed in Table 3.2. Here we summarise the key points from the assessments given in the previous sections.

An assessment of both ARC and CCI ATSR SSTs are shown by comparison with in-situ observations (drifting buoys) and MW imager data that the ARC SST biases are still slightly superior when compared to the CCI SSTs. However the differences are small and the advantage of having a common processing with the AVHRR SST data to extend the coverage of climate quality SSTs in time and space is an advantage. It is hoped the (A)ATSR CCI SST product will reach the ARC standard in phase 2. It is recommended that the CCI SST be considered for use as input to a climate quality SST analysis (e.g. HadISST) which is then used in climate model runs and atmosphere or ocean reanalyses. The uncertainties provided with the SST product are reasonable but there is scope for improvements in the phase 2 update.

The assessment of the CCI ocean colour dataset is assessed through assimilation over the period 1997-2012 and compared with the GlobColour dataset for the same period. At this stage it can be concluded that both the CCI and GlobColour data sets are suitable for data assimilation in ocean biogeochemical models, but neither has yet been definitively shown to be superior to the other. There are small improvements in the coverage of the CCI data. The observation uncertainties provided with each data set were used by the quality control, but not used in the assimilation itself to weight the observations. However, the system could be developed to make further use of this information in phase 2.

The SSH dataset assessment was for a regional reanalyses using the example shown for the Mediterranean Sea where it is planned to use the data of the HyMeX observing campaign in the next few years to allow more detailed validation of models and satellite datasets. A confrontation with tide gauge-derived mean sea level over the Mediterranean Sea also illustrates that the sampling of in-situ observations might be a key issue when comparing these data to satellite-derived products.

The CCI cloud data sets were compared with other satellite cloud datasets and also the HadGEM2 climate model. Generally the CCI data underestimates cloud amount over the global oceans at all latitudes. High-level cloud is underestimated over land for both the tropics and mid-latitudes, while low-level cloud is underestimated over the key areas of marine stratocumulus in the sub-tropics. A large unrealistic underestimate of thin cirrus cloud was also noted. Hence the current version of the CCI cloud dataset does not provide added value compared to ISCCP (or similar precursors) in respect of cloud coverage. It also contains discontinuities in all of the cloud amount products, making comparisons with models of little value. Other parameters of interest, specifically cloud liquid water and cloud droplet effective radius have been assessed and compared with models and other products. It is interesting to note the cloud LWP values agree quite well with the SSM/I LWP values even though the physical measurement is quite different. The cloud drop effective radius measurements from the CCI dataset provide drop sizes somewhat larger than other satellite estimates. This appears

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

to be due to the retrieval methodology rather than the data itself. At present it is not possible to determine the value of the CCI clouds products for climate model evaluation. Clearly more work needs to be done: (a) within the CCI team itself to understand the precise reasons for the differences with other well-established data sets and to improve the CCI products (e.g. the cirrus detection); and (b) in conjunction with the modelling community to ensure the CCI clouds project delivers products that are both useful and add value to those currently being used for model evaluation and development studies.

As far as the ozone ECV is concerned, the confrontation with model simulations, in particular those that are nudged towards, or assimilating other observational products, confirm that the CCI ozone products are suitable for model intercomparisons and data assimilation applications. This is in agreement with the ECMWF assessments of the ozone CCI products reported in CMUG (2013b). The uncertainty estimate remains a challenge. One way to make progress on this issue could be to generalize the intercomparison of observational products including satellite-derived, in-situ data derived, reanalyses and mixed products, through an international effort equivalent to what is done within the context of Model Intercomparison Projects (MIPs).

Monthly mean XCO₂ and XCH₄ data retrieved with several algorithms from the SCIAMACHY and GOSAT measurements were compared with MACC model outputs using the CMF. It was noted that because of large biases in the MACC atmospheric reanalysis of XCO₂ and XCH₄, model outputs from two experimental forecast runs that used optimised CO₂ and CH₄ fluxes were also used. The XCO₂ retrievals show a mean annual growth that seems to be largely consistent with that obtained from independent sources (e.g. the NOAA ESRL data). The datasets retrieved from SCIAMACHY and GOSAT measurements generally show a good level of agreement in terms of the annual cycle and the amplitude of the carbon dioxide changes, as well as the observation uncertainties, particularly in the northern hemisphere extra-tropics. In the tropics and southern hemisphere extra-tropics, the level of agreement between observed and modelled data is hampered by a number of issues and shortcomings. It was mentioned that inter-hemispheric differences could be associated with less reliable CO₂ fluxes in these regions caused by a much more sparse observing system than in the northern hemisphere. One inconsistency in the MACC system is the fact that two different transport models are currently used in the flux inversion system and in the forward simulation done by the IFS. This error can be exacerbated in regions of sparse data. Finally, comparing pre-calculated averages is not ideal and may lead to misleading results in regions where observations and model exhibit very different coverage or in aerosol and cloud affected regions. The SCIAMACHY XCO₂ uncertainties were normally found as large as or larger than the residuals between the observed and modelled XCO₂. In contrast, the GOSAT uncertainties appear smaller than the observation departures from the model. ####to here##

The GHG_CCI methane retrievals were also compared with the MACC atmospheric reanalysis of CH₄ and with an additional forecast run output that used the MACC optimized CH₄ fluxes. As these CH₄ fluxes were obtained in both cases using the SCIAMACHY data, the comparisons cannot provide an independent assessment for its retrievals. During the period of overlap between SCIAMACHY and GOSAT, a good level of agreement was found among the various products, especially the WFMD SCIAMACHY CH₄ and the two GOSAT

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

datasets. The IMAP retrievals showed larger values than the other three products, particularly in the tropics. The SCIAMACHY uncertainties were found much larger than those in GOSAT, most likely a consequence of instrument characteristics and design. The WFMD product for SCIAMACHY CH₄ is characterized by a sudden increase in the observation uncertainties in 2005 (likely related to the 2005 detector degradation) not seen in the IMAP equivalent. The latter also seem to be sporadically characterized by very large uncertainties. Understanding the reason for these large values requires an assessment of the uncertainties at pixel level.

An assessment of the aerosol products was presented in CMUG (2013b) using the datasets generated for 2008 with the three main AATSR algorithms. Since that report, the ADV1.42 dataset was extended to cover the period 2007-2010; at the same time, two new versions of the Swansea University algorithm (SU) were released with improvements over the oceans. The first version of these two (SU4.1) was also used to produce a decade of AOD data. Here an update of the CMUG (2013b) report has been presented with a focus on these new products. An assessment of the long-term retrieval anomalies at various wavelengths for the ADV1.42 and SU4.1 products did not highlighted particular problems and inconsistencies. The analysis of the 2008 retrievals at the four main wavelengths showed that the two new SU products are characterized, on average, by slightly lower AOD values over the oceans, thus increasing their departures from the MACC AOD reanalyses. This is also confirmed by the comparisons over the two extended periods. Overall, the comparisons would suggest that the ADV product is the one that shows the higher level of agreement with the MACC reanalyses, particularly over the oceans. However large biases were found in the comparisons of the MACC output with the AERONET observations that would suggest a much higher level of agreement between the most recent SU observations and the MACC reanalyses if the bias in the reanalyses was accounted for, and thus an overall improvement of the latest SU products. It is understood that the SU4.2 dataset represents a further improvement of the SU4.1 one. Based on these considerations the SU4.2 data could potentially lead to substantial improvements in future reanalysis productions.

For the surface datasets (e.g. land cover and fire) the assessments are direct comparison between the model and new satellite derived fields, which provide model boundary conditions, and involve exploring the reasons for the differences and impacts on model simulations. The performance of the climate model with the new boundary conditions for example to show the changes in the carbon emissions can be used as a way to assess new surface datasets. It is critical here that all the surface variables are consistent with each other between datasets as the model will struggle to provide consistent surface analyses if not.

For the CCI Land Cover the major results of the climate model experiments are that the meteorological forcing datasets (WATCH, CRU/NCEP) give the largest differences of land surface variables and fluxes which tend to dominate over the use of a different land cover dataset. However use of the CCI land cover data does result in significant changes for important land surface state variables. The differences between ESA CCI land cover and its precursor product (Globcover) are smaller than the differences between the CTRL land cover and the ESA CCI land cover. The differences between different CCI land cover epochs are marginal compared to the impact of changing from the CTRL land cover to the CCI land



CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

cover. Using CCI land cover data results in slightly improved skill in simulating surface albedo as compared with other observational surface albedo data products. The relative improvement is in the order of 2.5% to 7%. Also using CCI land cover data slightly improves the skill in simulating global land surface 2m temperature. The relative improvement is of the order of 1.5% to 3% by comparing with different observed land surface 2m temperature datasets.

For the Fire Burnt Area product only precursor data could be used because a CCI dataset was not available at the time of this report.

A general point is that the models need to be more comparable with satellite-derived products though the development of new diagnostics from modelled fields that are more physically consistent with the ECV geophysical variables. This is illustrated with the SSH ECV, since a total sea level height cannot be directly obtained from a single model output but from the addition of different contributions to be estimated with a combination of modelled variables. This implies to make a careful analyses of the order of magnitude of these contributions in order to select those that need to be calculated. This also raises some specific questions depending on the simulation domain (globe or an oceanic region) and possibly on the modelling protocol (boundary conditions in a regional climate simulation).

Finally it should be clear that what was not attempted in these assessments was to look at cross-ECV consistency which is also an important property for climate modelling applications. It is hoped this aspect can be explored in future studies by the CMUG during phase 2 using the CCI datasets where particular attention will be paid to this aspect. It should be noted that the individual ECV teams through their climate research groups will also be making complementary studies of their datasets but in general focussing less on climate modelling applications.

6. References

Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., Cosca, C., Harasawa, S., Kozyr, A., Nojiri, Y., O'Brien, K. M., Schuster, U., Telszewski, M., Tilbrook, B., Wada, C., Akl, J., Barbero, L., Bates, N. R., Boutin, J., Bozec, Y., Cai, W.-J., Castle, R. D., Chavez, F. P., Chen, L., Chierici, M., Currie, K., De Baar, H. J. W., Evans, W., Feely, R. A., Fransson, A., Gao, Z., Hales, B., Hardman-Mountford, N. J., Hoppema, M., Huang, W.-J., Hunt, C. W., Huss, B., Ichikawa, T., Johannessen, T., Jones, E. M., Jones, S., Jutterstrøm, S., Kitidis, V., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Manke, A. B., Mathis, J. T., Merlivat, L., Metzl, N., Murata, A., Newberger, T., Omar, A. M., Ono, T., Park, G.-H., Paterson, K., Pierrot, D., Ríos, A. F., Sabine, C. L., Saito, S., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K. F., Sun, H., Sutton, A. J., Suzuki, T., Sweeney, C., Takahashi, T., Tjiputra, J., Tsurushima, N., Van Heuven, S. M. A. C., Vandemark, D., Vlahos, P., Wallace, D. W. R., Wanninkhof, R. and Watson, A. J. 2014. An update to the Surface Ocean CO₂ Atlas (SOCAT version 2). *Earth Syst. Sci. Data*, **6**: 69-90.

Balmaseda, A.B., K. Mogensen, F. Molteni, and A. Weaver, 2010: The NEMOVAR-COMBINE ocean re-analysis. *COMBINE Technical Report N°1*, ECMWF, Reading.

Bergamaschi, P., C. Frankenberg, J. F. Meirink, M. Krol, M. G. Villani, S. Houweling, F. Dentener, E. J. Dlugokencky, J. B. Miller, L. V. Gatti, A. Engel, and I. Levin, 2009: Inverse modeling of global and regional CH₄ emissions using SCIAMACHY satellite retrievals, *J. Geophys. Res.*, **114**, doi:10.1029/2009JD012287.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Bergamaschi, P., S. Houweling, A. Segers, M. Krol, C. Frankenberg, R. A. Scheepmaker, E. Dlugokencky, S. C. Wofsy, E. A. Kort, C. Sweeney, T. Schuck, C. Brenninkmeijer, H. Chen, V. Beck and C. Gerbig, 2013: Atmospheric CH₄ in the first decade of the 21st century: Inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, *J. Geophys. Res.*, **118**, 7350–7369, doi:10.1002/jgrd.50480.

Bodeker, G. E., H. Shiona, and H. Eskes, H., 2005 : Indicators of Antarctic ozone depletion, *Atmos. Chem. Phys.*, **5**, 2603–2615, doi:10.5194/acp-5-2603-2005.

Calafat, F.M., and G. Jordà, 2011: A Mediterranean Sea Level reconstruction (1950-2010) with error budget estimates, *Global Planet. Change* **79**, 118-133, doi: 10.1016/j.gloplacha.2011.09.003.

Cariolle, D., and H. Teyssedre, 2007: A revised linear ozone photochemistry parameterization for use in transport and general circulation models : multi-annual simulations. *Atmos. Chem. Phys.*, **7**, 2183-2196, doi:10.5194/acpd-7-1655-2007.

Chevallier F., P. Ciais, T.J. Conway, T. Aalto, B.E. Anderson, P. Bousquet, E.G. Brunke, L. Ciattaglia, Y. Esaki, M. Fröhlich, A. Gomez, A.J. Gomez - Pelaez, L. Haszpra, P.B. Krummel, R.L. Langenfelds, M. Leuenberger, T. Machida, F. Maignan, H. Matsueda, J.A. Morguí, H. Mukai, T. Nakazawa, P. Peylin, M. Ramonet, L. Rivier, Y. Sawa, M. Schmidt, L.P. Steele, S.A. Vay, A.T. Vermeulen, S. Wofsy and D. Worthy, 2010: CO₂ surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, *J. Geophys. Res.*, **115**, doi:10.1029/2010JD013887.

Chevallier, F., and C. W. O'Dell, 2013: Error statistics of Bayesian CO₂ flux inversion schemes as seen from GOSAT. *Geophys. Res. Lett.*, **40**, 1252–1256, doi:10.1002/grl.50228.

Chevallier, F., P. I. Palmer, L. Feng, H. Boesch, C. W. O'Dell, and P. Bousquet, 2014. Toward robust and consistent regional CO₂ flux estimates from in situ and spaceborne measurements of atmospheric CO₂, *Geophys. Res. Lett.*, **41**, 1065–1070, doi:10.1002/2013GL058772.

CMUG 2012 Deliverable 3.1a Technical note on CMUG ECV Quality Assessment *CMUG Report version 1*, June 2012.

CMUG 2013a Deliverable 3.1b CMUG ECV Soil Moisture Assessment *CMUG Report November 2012*

CMUG 2013b Deliverable 3.1d Demonstration of CMF functionality for the assessment of ozone, aerosol and soil moisture Soil Moisture Assessment *CMUG Report September 2013*

CMUG, 2013c. Status of the ECMWF Climate Monitoring Database Facility. *Technical Note/Deliverable 3.5*, May 2013.

CMUG, 2013d. The role of Reanalysis in the production and quality assessment of CDRs. *Technical Note/Deliverable 3.4*, May 2013.

Conway, T. J., Tans, P. P., Waterman, L. S., Thoning, K. W., Kitzis, D. R., Masarie, K. A. and Zhang, N, 1994: Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res.*, **99**(D11), 22,831–22,855, doi:10.1029/94JD01951

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, **137**: 553-597.

CMUG Deliverable

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1



Dibarboure G, Lauret O, Mertz F, Rosmorduc V, Maheu C (2009) SSALTO/DUACS user handbook, (M)SLA and (M)ADT near-real time and delayed time products. *CLS-DOS-NT-06.034. Ramonville-Saint-Agne, France: Aviso Altimetry.*

Ferry, N., L. Parent, G. Garric, B. Barnier, N.C. Jourdain and the Mercator Ocean team, 2010: Mercator global eddy permitting ocean reanalysis GLORYS1V1: Description and results. *Mercator Quarterly Newsletter* **36**, January 2010.

Ford, D. A., Edwards, K. P., Lea, D., Barciela, R. M., Martin, M. J., and Demaria, J. 2012. Assimilating GlobColour ocean colour data into a pre-operational physical-biogeochemical model, *Ocean Sci.*, **8**: 751-771.

Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S. 2010. Assessing variability and long-term trends in burned area by merging multiple satellite fire products, *Biogeosciences*, **7**, 1171-1186, doi:10.5194/bg-7-1171-2010

Gleckler, P.J., Taylor, K.E. & Doutriaux, C., 2008. Performance metrics for climate models. *Journal of Geophysical Research*, **113**(D6), p.D06104.

Hansen, M., DeFries, R.S., Townshend, J.R.G., Carroll, M., Dimiceli, C., Sohlberg, R.A., 2003. Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS vegetation continuous fields algorithm. *Earth Interactions* **7**, 1–15.

Hemmings, J. C. P., Barciela, R. M., and Bell, M. J. 2008. Ocean color data assimilation with material conservation for improving model estimates of air-sea CO₂ flux. *J. Mar. Res.*, **66**: 87-126.

Hurtt, G. C., Chini, L. P., Frohling, S., Betts, R., Feddema, J., Fischer, G., Klein Goldewijk, K., Hibbard, K., Janetos, A., Jones, C., Kindermann, G., Kinoshita, T., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D., and Wang, Y. P.: Harmonisation of global land-use scenarios for the period 1500–2100 for IPCC-AR5, *iLEAPS Newsletter*, **7**, 3 pp., 2009.

Inness A, Baier F, Benedetti A, Bouarar I, Chabrilat S, Clark H, Clerbaux C, Coheur P, Engelen RJ, Errera Q, Flemming J, George M, Granier C, Hadji-Lazaro J, Huijnen V, Hurtmans D, Jones L, Kaiser J, Kapsomenakis J, Lefever M, Leitão J, Razinger M, A, Schultz MG, Simmons AJ, Suttie M, Stein O, Thépaut JN, Thouret V, Vrekoussis M, Zerefos C, the MACC team. 2013. The MACC reanalysis: an 8 yr data set of atmospheric composition. *Atmos. Chem. Phys.*, **13**(8), 4073-4109.

Jordà, G., and D. Gomis, 2013: On the interpretation of the steric and mass components of sea level variability: The case of the Mediterranean basin. *J. Geophys. Res. Oceans*, **118**, 1-11, doi: 10.1002/jgrc.20060.

Kalnay *et al.*, The NCEP/NCAR 40-year reanalysis project, *Bull. Amer. Meteor. Soc.*, **77**, 437-470, 1996.
 Poulter, B. *et al.*, 2011. Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO₂ airborne fraction. *Biogeosciences*, **8**(8), 2027–2036

Lean, K. and Saunders, R. W.: Validation of the ATSR Reprocessing for Climate (ARC) dataset using data from drifting buoys and a three-way error analysis, *Journal of Climate*, **26**, 4758–4772, doi:10.1175/JCLI-D-12-00206.1, 2013.

Lefèvre, N., Caniaux, G., Janicot, S., and Gueye, A. K. 2013. Increased CO₂ outgassing in February-May 2010 in the tropical Atlantic following the 2009 Pacific El Niño, *J. Geophys. Res. Oceans*, **118**: 1-13.

MACC II, 2013. Validation report of the MACC reanalysis of global atmospheric composition. Period 2003-2011. *Technical Note/Deliverable D_83.4, version 1.1, February 2013.*

Madec, G. 2008. NEMO ocean engine, Note du Pole de modélisation, *Institut Pierre-Simon Laplace (IPSL), France, No. 27 ISSN, 1288-1619.*

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Maritorena, S., Fanton d'Andon, O. H., Mangin, A., and Siegel, D. A. 2010. Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues, *Remote Sens. Environ.*, **114**, 1791-1804.

Marlon, J. R., Bartlein, P. J., Daniiau, A.-L., Harrison, S. P., Maezumi, S. Y., Power, M. J., Tinner, W., and Vanniere, B.: Global biomass burning: a synthesis and review of Holocene paleofire records and their controls, *Quaternary Science Reviews*, **65**, 5–25, 2013.

McCarthy, G. D., Frajka-Williams, E., Johns, W. E., Baringer, M. O., Meinen, C. S., Bryden, H. L., Rayner, D., Duchez, A., Roberts, C. D., and Cunningham, S.A. 2012. Observed interannual variability of the Atlantic meridional overturning circulation at 26.5°N, *Geophys. Res. Lett.*, doi:10.1029/2012GL052933 in press.

Merchant, C. J., Horrocks, L. A., Eyre, J. R., and O'Carroll, A. G.: Retrievals of sea surface temperature from infrared imagery: origin and form of systematic errors, *Quarterly Journal of the Royal Meteorological Society*, **132**, 1205–1223, doi: 10.1256/qj.05.143, 2006.

Meysignac, B., F.M. Calafat, S. Somot, V. Rupolo, P. Stocchi, W. Llovel, A. Cazenave, 2011: Two-dimensional reconstruction of the Mediterranean sea level over 1970–2006 from tide gage data and regional ocean circulation model outputs, *Global Planet. Change*, **77**(1-2), 49–61, doi: 10.1016/j.gloplacha.2011.03.002.

Michou, M. et al., 2011: A new version of the CNRM Chemistry-Climate Model, CNRM-CCM: description and improvements from the CCMVal-2 simulations, *Geosci. Model Dev. Discuss.*, **4**, 1129–1183, 2011.

O'Carroll, A. G., Eyre, J. R., and Saunders, R. W.: Three-way error analysis between AATSR, AMSR-E, and in situ sea surface temperature observations, *Journal of Atmospheric and Oceanic Technology*, **25**, 1197–1207, doi: 10.1175/2007JTECH0542.1, 2008.

Pajot, B., S. Massart, D. Cariolle, A. Piacentini, O. Pannekoucke, W. A. Lahoz, C. Clerbaux, P. F. Coheur, and D. Hurtmans, 2012: High resolution assimilation of IASI Ozone data with a global CTM, *Atmos. Chem. Phys. Discuss.*, **11**, 29357-29406, doi:10.5194/acpd-11-29357-2011.

Palmer, J. R., and Totterdell, I. J. 2001. Production and export in a global ocean ecosystem model, *Deep-Sea Res. Pt. I*, **48**: 1169-1198.

Polavarapu, S., Ren, S., Rochon, Y., Sankey, D., Ek, N., Koshyk, J., and Tarasick, D. 2005. Data assimilation with the Canadian Middle Atmosphere Model, *Atmosphere-Ocean*, **43**, 77-100.

Prioul, J.C., 2013: Quality assessment of satellite Ozone products and user assessment of the Météo-France model in the framework of the ESA-CCI program. *CNES/Météo-France internal report, July 2013, 33pp.*

Reick, C., Raddatz, T., Brovkin, V., & Gayler, V. (2013). The representation of natural and anthropogenic land cover change in MPI-ESM. *Journal of Advances in Modeling Earth Systems*. **5**(3). 459-482. doi: 10.1002/jame.20022

Sevault, F., Somot, S., Beuvier, J., 2009. A regional version of the NEMO ocean engine on the Mediterranean Sea: NEMOMED8 user's guide. *Technical note 107, CNRM, Météo-France, Toulouse.*

Stevens, B. et al., 2013. Atmospheric component of the MPI-M Earth System Model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, **5**(2), 146–172

Storkey, D., Blockley, E. W., Furner, R., Guiavarc'h, C., Lea, D., Martin, M. J., Barciela, R. M., Hines, A., Hyder, P., and Siddorn, J. R. 2010. Forecasting the ocean state using NEMO: The new FOAM system. *Journal of Operational Oceanography*, **3**: 3-15.

**CMUG Deliverable**

Number: D3.1
Due date: March 2014
Submission date: 1 April 2014
Version: 2.1

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Sabine, C., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Körtzinger, A., Steinhoff, T., Hoppema, M., de Baar, H. J. W., Wong, C. S., Delille, B., and Bates NR. 2009. Climatological mean and decadal changes in surface ocean pCO₂, and net sea-air CO₂ flux over the global oceans, *Deep-Sea Res. Pt. II*, **56**, 554-577.

Taylor, K.E., Stouffer, R.J. & Meehl, G. A., 2012. An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93**(4), 485-498.

Uppala, S.M. *et al.*, 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131**(612), 2961–3012.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T., 2010, Global fire emissions and the contribution of deforestation, savannah, forest, agricultural, and peat fires (1997–2009), *Atmos. Chem. Phys.*, **10**, 11707-11735, [doi:10.5194/acp-10-11707-2010](https://doi.org/10.5194/acp-10-11707-2010).

Weedon, G. P.; Gomes, S.; Viterbo, P.; Shuttleworth, W. J.; Blyth, E.; Österle, H.; Adam, J. C.; Bellouin, N.; Boucher, O. & Best, M. 2011. Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century *J. Hydrometeorol.*, **12**, 823-848.

While, J., Totterdell, I. J., and Martin, M. J. 2012. Assimilation of pCO₂ data into a global coupled physical-biogeochemical ocean model, *J. Geophys. Res.*, **117**, C03037, [doi:10.1029/2010JC006815](https://doi.org/10.1029/2010JC006815).

End of report