

1 **Consistency of satellite climate data records for Earth system monitoring**

2 Thomas Popp¹, Michaela I. Hegglin², Rainer Hollmann³, Fabrice Ardhuin⁴, Annett Bartsch⁵,
3 Ana Bastos⁶, Victoria Bennett⁷, Jacqueline Boutin⁸, Carsten Brockmann⁹, Michael Buchwitz¹⁰,
4 Emilio Chuvieco¹¹, Philippe Ciais¹², Wouter Dorigo¹³, Darren Ghent¹⁴, Richard Jones¹⁵,
5 Thomas Lavergne¹⁶, Christopher J. Merchant^{2,17}, Benoit Meyssignac¹⁸, Frank Paul¹⁹, Shaun
6 Quegan²⁰, Shubha Sathyendranath²¹, Tracy Scanlon¹³, Marc Schröder³, Stefan G. H. Simis²¹,
7 Ulrika Willén²²

8

9 ¹ *German Aerospace Center (DLR), Wessling, Germany*

10 ² *University of Reading, Reading, UK*

11 ³ *Deutscher Wetterdienst (DWD), Offenbach, Germany*

12 ⁴ *Lab. of Ocean Physics and Satellite oceanography, Ifremer, Plouzané, France*

13 ⁵ *b.geos GmbH, Korneuburg, Austria*

14 ⁶ *Dept. Of Geography, Ludwig-Maximilians-Universität, München, Germany*

15 ⁷ *Centre for Environmental Data Analysis, STFC Rutherford Appleton Laboratory, Harwell, UK*
16 *and National Centre for Earth Observation, UK*

17 ⁸ *Sorbonne Université, CNRS, IRD, MNHN, Laboratoire d'Océanographie et du Climat:*
18 *Expérimentations et Approches Numériques (LOCEAN), Paris, France*

19 ⁹ *Brockmann Consult GmbH, Hamburg, Germany*

20 ¹⁰ *Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany*

21 ¹¹ *Environmental Remote Sensing, University of Alcalá, Alcalá de Henares, Spain*

22 ¹² *IPSL – LSCE, Gif sur Yvette, France*

23 ¹³ *TU Wien, Department of Geodesy and Geoinformation, Vienna, Austria*

24 ¹⁴ *National Centre for Earth Observation, Department of Physics & Astronomy, University of*
25 *Leicester, Leicester, UK*

Early Online Release: This preliminary version has been accepted for publication in *Bulletin of the American Meteorological Society*, may be fully cited, and has been assigned DOI 10.1175/BAMS-D-19-0127.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

26 ¹⁵ *Met Office Hadley Centre, Exeter, UK*

27 ¹⁶ *Norwegian Meteorological Institute, Oslo, Norway*

28 ¹⁷ *National Centre for Earth Observation, University of Reading, Reading, UK*

29 ¹⁸ *LEGOS CNES, CNRS, IRD, Université de Toulouse, Toulouse, France*

30 ¹⁹ *Department of Geography, University of Zurich, Zurich, Switzerland*

31 ²⁰ *School of Mathematics and Statistics, University of Sheffield, Sheffield, UK*

32 ²¹ *Plymouth Marine Laboratory, Plymouth, UK*

33 ²² *Swedish Hydrological and Meteorological Institute, Norrköping, Sweden*

34

35 Corresponding author: Thomas Popp, German Aerospace Center (DLR), Münchner Str. 20,
36 82234 Wessling, Germany, Phone: +49 / 8153 28 1382, E-mail: thomas.popp@dlr.de

37

38 **Abstract**

39 Climate Data Records (CDRs) of Essential Climate Variables (ECVs) as defined by the
40 Global Climate Observing System (GCOS) derived from satellite instruments help to
41 characterize the main components of the Earth system, to identify the state and evolution of
42 its processes, and to constrain the budgets of key cycles of water, carbon and energy. The
43 Climate Change Initiative (CCI) of the European Space Agency (ESA) coordinates the
44 derivation of CDRs for 21 GCOS ECVs. The combined use of multiple ECVs for Earth system
45 science applications requires consistency between and across their respective CDRs. As a
46 comprehensive definition for multi-ECV consistency is missing so far, this study proposes
47 defining consistency on three levels: (1) consistency in format and metadata to facilitate
48 their synergetic use (technical level); (2) consistency in assumptions and auxiliary datasets to
49 minimize incompatibilities among datasets (retrieval level); and (3) consistency between
50 combined or multiple CDRs within their estimated uncertainties or physical constraints
51 (scientific level).

52 Analysing consistency between CDRs of multiple quantities is a challenging task and requires
53 coordination between different observational communities, which is facilitated by the CCI
54 program. The inter-dependencies of the satellite-based CDRs derived within the CCI program
55 are analysed to identify where consistency considerations are most important. The study
56 also summarizes measures taken in CCI to ensure consistency on the technical level, and
57 develops a concept for assessing consistency on the retrieval and scientific levels in the light
58 of underlying physical knowledge. Finally, this study presents the current status of
59 consistency between the CCI CDRs and future efforts needed to further improve it.

60 **Capsule**

61 In this study, the ESA Climate Change Initiative (CCI) introduces a three-level
62 definition of consistency between multiple satellite-based Climate Data Records (CDRs) of
63 Essential Climate Variables (ECVs), discusses consistency status and requirements and
64 develops a concept for assessing inter and across ECV consistency.

65 **1. Introduction**

66 The Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report
67 (AR5) and the three Special Reports of the AR6 cycle state that mankind and the biosphere
68 face great threats due to the rapidly changing climate (IPCC, 2013, 2018, 2019a, 2019b). To
69 support political decisions on climate change mitigation and adaptation, and to quantify the
70 implications for economic and non-economic loss and damage, the United Nations
71 Framework Convention on Climate Change (UNFCCC) requires systematic monitoring of the
72 global climate system (e.g. Doherty et al., 2009; UNFCCC Art. 4 and Art. 5, 1992; Paris
73 Agreement 7.7c, Adaptation). In particular, systematic monitoring is important in assessing
74 progress on the aims of the Paris Agreement (e.g. for the global stocktake). The main tools at
75 hand to determine the extent and impacts of climate change on local to global scales and
76 understand its causes are a combination of global and regional climate and Earth system
77 models, reanalysis data, and systematic observations. The latter are indispensable for all
78 Earth system domains (atmospheric, terrestrial, oceanic) to increase the understanding of
79 and quantify processes, budgets and reservoirs within the global Earth cycles (carbon,
80 energy, and water).

81 To promote systematic climate monitoring, the World Meteorological Organization
82 (WMO), Intergovernmental Oceanographic Commission (IOC), United Nations Environment
83 Program (UNEP), and International Science Council (ISC), established in 1992 the Global
84 Climate Observing System (GCOS). GCOS aims at sustained “provision of reliable physical,
85 chemical and bio-chemical observations and data records for the total climate system –
86 across the atmospheric, oceanic and terrestrial domains, including hydrological and carbon
87 cycles and the cryosphere” (GCOS, 2016). GCOS defined a set of currently 54 “Essential
88 Climate Variables” or ECVs (Bojinski et al., 2014) which must be observed in a sustained and

89 **consistent** manner to enable detection of climate trends and provide data suitable for
90 climate model evaluation and climate change attribution.

91 Complementary to relatively sparse airborne and ground-based measurements and
92 inventory data, satellite observations are of ever-growing importance for evaluating,
93 initializing and parameterizing Earth system processes represented in models. This growing
94 importance is due to the increasing satellite global coverage and resolution (in space and
95 time), their improved calibration accuracy and the increasing diversity of relevant
96 observables provided by advances in satellite sensor technologies. Satellite observations can
97 provide a significant contribution for 21 out of the 54 GCOS ECVs. Some of these are
98 exclusively derived from satellite measurements (e.g. the Earth Radiation Budget), whereas
99 for others dedicated spaceborne sensors provide better coverage but lower accuracy or
100 resolution than in situ measurements (e.g. above-ground biomass, column atmospheric
101 concentration of CO₂ and CH₄).

102 Studies of the Earth system require combined analysis of datasets of many variables.
103 Since these are derived from different sources (satellite-, ground-, air- and model-based) and
104 processing systems, one underlying precondition of any such analysis is that the datasets are
105 **consistent**. However, despite the importance of consistency, many open questions remain,
106 ranging from a clear definition of consistency for multiple quantities, to systematically
107 assessing consistency between the many data records used.

108 Possible reasons for inconsistencies include the use of different auxiliary datasets,
109 simplifications in corrections and retrieval algorithms, calibration uncertainties and
110 differences in sampling and gridding. For example, a time series of a single variable built
111 from data records obtained from different sensors may exhibit “jumps” where they are
112 merged with each other, which may hinder any trend analysis. Another example is assigning
113 different land cover classes (e.g. glacier, water, rock or vegetation) to the same pixel by using

114 different glacier masks, which may lead to highly variable budget calculations of related
115 exchange processes.

116 Consistency as an issue in creating satellite-based data records was first met by
117 operational entities like NOAA, EUMETSAT or NASA within their near-real time (NRT)
118 processing chains across different satellite missions. This includes aspects such as common
119 input datasets, gridding methodology, cloud and land/sea masking, aerosol and water
120 vapour corrections, and the land cover map used. The measures taken are typically
121 documented in Algorithm Theoretical Baseline Documents (e.g. consistent OMI-MODIS cloud
122 products: Siddans, 2016 or merged TROPOMI-VIIRS cloud product: NASA, 2014). However,
123 the need for consistency across different variables, domains and processing systems is
124 inherent in climate studies and thus much broader than in the often independent NRT
125 applications.

126 Over the past ten years, space agencies (including ESA, EUMETSAT, NASA, and NOAA)
127 have emphasised the generation and delivery of satellite-based CDRs. Hollmann et al. (2013)
128 describe the efforts of ESA through its Climate Change Initiative (CCI). CCI leverages and
129 harvests the long-term satellite archives available from European and other satellites, and
130 enhances or expands these records with observations from other space agencies to obtain
131 global coverage. In addition, CCI is extending its newly established CDRs with the most
132 recent satellite instruments to guarantee continuation into the future using operational
133 missions (e.g. Sentinel). During its first six years (2011-2017), CCI implemented 14 projects,
134 each targeting one (or two) ECVs; in 2018, CCI was expanded to include nine additional ECVs,
135 as shown in Figure 1. It should be noted that most of the ECVs consist of several quantities,
136 so-called products (detailed information on the products of each CCI ECV for which CDRs
137 have been processed in CCI is available at <http://cci.esa.int>). Of course, products within a
138 particular ECV have to be consistent. A particular element within the CCI program is

139 independent analysis of the quality of its CDRs and particularly their consistency (between
140 different ECVs and products) in a climate modelling context by the CCI Climate Model User
141 Group (CMUG) and several budget closure study projects.

142 Together with the Copernicus Climate Change Service (C3S) and contributions from
143 EUMETSAT through its Satellite Application Facilities (SAFs) such as the Climate Monitoring
144 SAF (Schulz et al., 2009), the NOAA Climate Data Record program
145 (<https://www.ncdc.noaa.gov/cdr>, Bates et al. 2016), and the NASA Measures program
146 (<https://earthdata.nasa.gov/measures>), about 1000 different satellite-based CDRs for GCOS
147 ECV products and further variables are available or will become available in the near future.
148 An overview of these CDRs is given in the ECV inventory
149 (<https://climatemonitoring.info/ecvinventory>), recently established by the joint CEOS-CGMS
150 Working Group on Climate, which conducts regular gap analysis to define future satellite
151 development needs.

152 This study introduces a concept developed in CCI to define and assess consistency
153 between multiple satellite-based ECV products. It is shown that such an assessment allows
154 remaining inconsistencies to be identified and quantified in the light of given CDR
155 uncertainties and relevant physical principles. A key application of assessing and ensuring
156 consistency is in closure studies where multiple CDRs are used together. A selection of topics
157 for such closure studies is briefly discussed in this paper to illustrate the concept.

158 Section 2 discusses different kinds of inconsistencies and develops a definition of
159 consistency, followed by a brief analysis of ECVs covered by CDRs from CCI and consistency
160 needs in Earth system monitoring in Section 3. Section 4 develops a concept for assessing the
161 different levels of consistency and illustrates it with examples from different ECV products in
162 CCI. Section 5 presents a discussion of the main findings and identifies remaining consistency
163 gaps.

164 **2. Consistency in Earth system monitoring**

165 Whilst "consistency" (e.g. between two datasets) is a concept frequently referred to
166 in the observation community, there is, to our knowledge, no comprehensive definition
167 specific to observation datasets of different variables. This may reflect the complexity of
168 relations between the large set of ECVs. This study proposes such a comprehensive
169 definition and an assessment concept for consistency. The focus is on consistency between
170 datasets of different variables, as needed for climate studies, but also single-variable cases
171 are included.

172 According to the common definition of the word "consistency" (Oxford dictionary), it
173 is "the quality of always behaving in the same way or of having the same opinions or
174 standards; the quality of being consistent, i.e., 1/ in agreement with something; not
175 contradicting something, 2/ happening in the same way and continuing for a period of time,
176 3/ consistent with something in agreement with something, not contradicting something, 4/
177 having different parts that all agree with each other". In the observation scientific
178 community, consistency is usually understood as "agreement", "compatibility" or "no
179 contradiction". When considering CDRs, "consistency" goes beyond "agreement" and rather
180 refers to "compatibility". Firstly, agreement per se can only be tested between datasets of
181 the same variable. A mature terminology and a comprehensive set of mathematical tools for
182 this purpose exists, which forms the basis of most calibration, validation and model
183 evaluation activities. Secondly, there can even be cases where two datasets of the same
184 variable agree (their bias is smaller than their combined uncertainties) but are inconsistent
185 (for example if only one of two datasets shows a distinct diurnal or seasonal cycle). In
186 contrast, regionally averaged time series of one variable can disagree (have regional biases
187 larger than the combined uncertainties), but be consistent in their temporal behaviour, as
188 shown for multi-sensor AOD records (Sogacheva, et al., 2020).

189 In a physical sense, consistency can be understood as fulfilling a conservation balance
190 equation (of mass or energy) or exhibiting a correlation in time or space between two data
191 records as expected by a physical theory. In CDR production, also simple category
192 inconsistencies occur (e.g. for one pixel land cover assigns bare soil, while biomass gives a
193 non-zero carbon mass to it).

194 Immler et al. (2010) defined consistency between measurements of the GCOS
195 Reference Upper Air Network (GRUAN) as “when the independent measurements agree to
196 within their individual uncertainties”, which requires knowledge of their (combined)
197 uncertainties. This definition applies to different measurements of the same variable, but in
198 the wider context of Earth system monitoring, a definition of consistency across multiple
199 ECVs is also needed.

200 Several kinds of inconsistency between different data records of the same quantity or
201 of different quantities can be recognised:

- 202 - Inconsistencies due to differences in auxiliary datasets;
- 203 - Temporal inhomogeneities in time series (e.g. due to calibration biases, degradation
204 in data obtained from a sequence of different input data records, or sampling
205 differences in terms of measurement time, frequency, or geographical coverage
206 during gridding);
- 207 - Spatial inhomogeneities due to combining fields from different datasets (e.g. with
208 different observing geometry or different sampling, e.g. all-sky versus clear-sky
209 sampling).

210 Many of these inconsistencies are linked to the statistical properties of the raw data
211 used to create a CDR, when for practical reasons simplifications and aggregations cannot be
212 avoided.

213 To cover the wide range of aspects of consistency, it is convenient to structure it on
214 three complementary levels:

215 (1) Consistency on the technical level: Harmonised data format and metadata
216 description to ease acquisition and combined usage of multiple CDRs;

217 (2) Consistency on the retrieval level: Use of the same auxiliary datasets in retrievals to
218 minimize contradictions in outputs linked to common information (e.g. a water
219 mask);

220 (3) Consistency on the scientific level: Compatibility of the relevant characteristics of two
221 or more CDRs (e.g. patterns, variability, trends, ...) with a reference (represented by a
222 physical equation, a model or a fiducial reference) within their combined
223 uncertainties.

224 While consistency on a **technical level** is easy to define and needs limited scientific
225 insight, it is often a resource-consuming barrier hindering data use. Thus the Earth
226 observation community has focused on this area in recent years (e.g. by adopting common
227 metadata standards). In particular, the CCI program has adopted existing solutions (and
228 when needed developed new ones) that facilitate combined satellite-based CDR use. This
229 includes a harmonized data format (netCDF, with a few exceptions where a different
230 standard is needed for a particular community, e.g. shapefiles for glaciers) and a common
231 metadata convention (CCI data standards: ESA, 2019), which follow the CF convention
232 (<http://cfconventions.org>). It covers additional cross-ECV standardized metadata attributes,
233 using common vocabularies for index terms and harmonized variable names, as well as a
234 harmonized / interoperable data access portal with common catalogue and data services to
235 simplify multi-quantity data search and download within the CCI portfolio
236 (<http://cci.esa.int/data>). This common vocabulary also helps to reduce inconsistent
237 nomenclature, such as labelling slightly different variables as the same retrieved quantity

238 (e.g. due to wavelength dependencies of retrieved information). Furthermore, the
239 underlying documentation of algorithms and datasets in CCI has been harmonized to some
240 extent, as in other initiatives such as the SAF network or NOAA CDR program. This
241 information helps users to quickly understand each dataset and its strengths, weaknesses
242 and limitations. A good example of the benefit of such harmonised climate data records on
243 the technical level is given by the CCI toolbox (<https://climatetoolbox.io>), which can be used
244 for harmonized data pre-processing, analysis and visualisation of the multiple CDRs in a
245 standardized way.

246 On the **retrieval level**, consistency aims at using the same (or a similar) observation
247 strategy (same or similar satellite sensors, frequencies, etc.), and similar auxiliary datasets
248 for the same variable in different retrieval algorithms. Those auxiliary datasets are either
249 categorical datasets, so-called “masks”, or continuous datasets of physical variables. Typical
250 masks used in many retrieval algorithms include, for example, a particular land cover
251 (vegetated areas), land-water, sea ice, snow cover and glacier masks, since many retrieval
252 algorithms behave differently over different surface types. Other masks commonly needed
253 across many variables are cloud masks, since many retrievals in the visible to thermal
254 spectral range need to avoid contamination by clouds. Frequently used continuous auxiliary
255 data fields include meteorological fields (e.g. from reanalysis) and climatologies of
256 atmospheric variables (e.g. water vapour, aerosols, ozone) to conduct atmospheric
257 corrections of visible bands used to retrieve land and ocean ECVs.

258 There is no sharp boundary between retrieval and scientific consistency. Ultimately,
259 **scientific consistency** deals with the compatibility in CDR properties relevant for climate
260 processes. All data records of a single ECV product, if obtained from different sources, need
261 to be consistent within their uncertainties and within sampling differences. One aspect is
262 consistency across borders in space (horizontally and vertically) and in time. Most

263 importantly, systematic biases between datasets need to be avoided as they may lead to
264 errors when evaluating model performance (e.g. Waugh and Eyring, 2008). This applies to
265 different combinations such as one variable based on multiple sensors, one sensor but using
266 multiple algorithms, or combined satellite, model and in situ data. Finally, when several
267 datasets of different variables are included in a physical model or budget equation, multi-
268 variable consistency needs to distinguish uncertainties of calculated closure budgets due to
269 propagated input uncertainties from real physical process imbalances or net effects.

270 3. Consistency needs for CCI Earth System Climate Data Records

271 In this section, linkages on the retrieval and scientific level between the different CCI
272 ECVs (Figure 1) are analysed. This analysis remains at the high level of the GCOS ECVs while it
273 is well understood that most ECVs consist of several different quantities, or so-called
274 products (e.g. the glacier ECV in CCI consists of the three products glacier outlines, elevation
275 change and velocity). In most of the analysis in this study the primary product of an ECV is
276 considered (e.g. aerosol optical depth for aerosol properties) and the most common
277 methodology used to retrieve it. This means that for using a specific CDR of one ECV there
278 may be a need to assess in more detail its linkages if, for example, a new retrieval technique
279 in another spectral range is considered or if another product of this ECV is assessed. Detailed
280 information on the products of each CCI ECV for which CDRs have been processed is
281 available at <http://cci.esa.int>.

282 As a first step, the needs for consistency between ECVs on the **retrieval level** are
283 assessed. Retrievals of Earth system variables from satellite observations aim to produce
284 high quality CDRs by constraining the (often under-determined) inversion equations as good
285 as possible. Typically, the measurements are chosen to have high sensitivity to the target
286 variable, but they are usually subject to perturbations from other variables. In such cases,
287 the inversion needs to either co-retrieve these additional variables or use auxiliary datasets
288 to describe their spatio-temporal distributions. Moreover, different retrieval algorithms are
289 often optimal for use over different surface types as their reflectance or spectral
290 characteristics are highly variable (e.g. over dark water or over bright land). The use of
291 different approaches for obtaining the same variable in different retrieval algorithms is one
292 possible source of inconsistency between CDRs.

293 After processing, all CDRs have to pass validation against external reference datasets
294 (e.g. from ground-based stations) to quantify their accuracy. Furthermore, CCI insists for

295 CDRs to be accompanied by proper uncertainty characterisation (using error propagation or
296 uncertainty characterization during validation) within their data files (Merchant et al., 2017),
297 so that uncertainties can be assessed when using the datasets. However, since reference
298 data can have temporal or spatial representativeness issues and different validation
299 methods also have their inconsistencies, unexplored uncertainties may remain (for the
300 retrieved values themselves and for the estimated uncertainties). Validation and error
301 propagation implicitly quantify inconsistencies from using imperfect auxiliary datasets and
302 retrieval simplifications to within uncertainties. However, proof of consistency needs to
303 explicitly test together the CDRs considered.

304 The part of Table 1 that is above the diagonal summarizes links between ECVs
305 generated and analysed by CCI with regard to their retrieval consistency. A need for retrieval
306 consistency is identified where either one or both retrievals rely on consistent co-retrieved
307 or auxiliary variables of the other ECV (links only within CCI are considered, but there are
308 other products, algorithms or sensors for which these may not apply).

309 The part of Table 1 below the diagonal summarizes the need for consistency on the
310 **scientific level** based on our knowledge of how two variables are linked by Earth system
311 processes or cycles in more detail. For this, the relevance of CCI ECVs for the energy, water
312 and carbon cycles is briefly recalled. Figure 2 lists available or upcoming ECVs for which ESA
313 CCI generates CDRs that contribute to the characterisation of these three main cycles. For
314 simplicity, each ECV is only attributed to the cycle in which it plays the most important role.
315 Practically all ECVs contribute to the energy cycle, either directly through radiation
316 interaction or through mass-attached energy transport in the water or carbon cycle. Studies
317 of sub-elements of these main cycles may also be relevant (e.g. physical processes such as
318 emission, transport, deposition or radiation interactions, chemical transformations; also
319 regional limitations, such as ice-free conditions) which may only require consistency among a

320 reduced set of ECVs. Some further details on the CCI CDRs for the three cycles are provided
321 in the following.

322 **Carbon cycle:** CCI CDRs help quantifying the dynamics of the amount of carbon stored
323 in the atmosphere, oceans and terrestrial biosphere and the fluxes between these reservoirs
324 (see overview about the carbon cycle in Le Quéré et al., 2018). CO₂ in the atmosphere is a
325 key measure of the anthropogenic perturbation to the carbon cycle. The air-sea CO₂ flux is
326 strongly affected by sea-surface temperature (SST) and ocean photosynthetic activity
327 (monitored using ocean colour observations). The CCI CDRs also help constraining carbon
328 fluxes from the land biosphere (e.g. Reuter, et al., 2017) including land use change and
329 biomass burning emissions, together with direct estimates of above-ground biomass and
330 burned area (Chuvieco et al., 2019). Other CCI CDRs of importance to the carbon cycle are
331 snow cover (which affects the duration and start of photosynthetic processes in boreal
332 forests; Pulliainen et al., 2017), similar to the impact of sea ice on marine photosynthesis in
333 high latitudes, soil moisture (which affects land-atmosphere CO₂ fluxes), permafrost (which
334 contains frozen carbon stores with about twice the mass of atmospheric carbon), and sea
335 surface salinity, which, together with SST, determines CO₂ solubility, with important impacts
336 in rainy regions and serves as a proxy for sea water alkalinity (Vinogradova et al., 2019).

337 **Water cycle:** CCI helps quantifying the global water cycle over land and ocean (see
338 overview in e.g. Levizzani and Cattani, 2019) by providing CDRs related to the reservoirs
339 within the water cycle (lake levels, sea level, sea ice, ice sheets, glaciers, soil moisture, and
340 snow), atmospheric water vapour content (water vapour) and clouds. From these, processes
341 such as precipitation and runoff that transfer water between the various reservoirs may be
342 inferred. CCI delivers additional relevant parameters such as sea surface salinity (related to
343 precipitation, evaporation and runoff), SST and LST (determining evaporation), land cover
344 and biomass (both linked to evapotranspiration).

345 **Energy cycle:** CCI also helps constraining the global energy cycle (for an overview see
346 Allan, 2012) by providing CDRs for SST and LST, land and sea ice, as well as snow cover, sea
347 level (which is affected among others by the ocean heat content and land ice melt), sea
348 state, clouds, water vapour, ozone, greenhouse gases and aerosols that help determine the
349 vertical temperature structure of the atmosphere. Finally, the biosphere (biomass) may also
350 be considered a part of the energy cycle since it converts solar energy into chemically-stored
351 energy (organic matter). In the oceans, a significant portion of the organic matter sinks out
352 of the surface layers, exporting the energy to the deep ocean (with the photosynthesis
353 activity being observed indirectly through ocean colour).

354 **4. Concept for assessing consistency on different levels**

355 Due to the complexity of different consistency aspects no single method can be used
356 for assessing consistency of CDRs on various levels. Therefore, a concept employing a range
357 of appropriate methods was developed in CCI, which is summarized here and then illustrated
358 with short examples.

359

360 **4.1 Overview: Methods to assess consistency**

361 All methods for assessing consistency contain several key elements. Firstly, any method
362 needs to be based on physical background knowledge to understand the relevance of any
363 disagreement or incompatibility. Such background knowledge can be a simple principle (e.g.
364 if the land cover is bare soil and the biomass product provides a high biomass value, there is
365 an obvious inconsistency) or knowledge of the sensitivity of a target variable toward an
366 auxiliary dataset, or a more complex physical equation or “model”. Secondly, any
367 assessment needs to select an appropriate characteristic (patterns, time series, masks)
368 tailored towards the relevant process (or cycle) and choose a suitable mathematical tool
369 (metric). Finally, this metric needs to be evaluated against the relevant physical background
370 knowledge while the threshold on the chosen metric for judging consistency depends on the
371 considered process or cycle and the datasets. In order to make any assessment of
372 consistency objective, a study needs to specify the threshold used. This is shown for the
373 following examples for various metrics.

374 In essence, consistency then means that several datasets have been evaluated against the
375 underlying physical background knowledge and were found “fit for purpose” for a specific
376 application domain. This leads to cases where seemingly small values of a chosen metric
377 (compared to its uncertainty) can mean inconsistency, whilst in other cases apparently large
378 deviations mean consistency, as will be shown in the examples of this section. Table 2 lists a

379 variety of related basic principles and methods to assess consistency on different levels used
380 in the following examples.

381

382 4.2 Methods to assess retrieval level consistency

383 As a principle, retrieval level inconsistencies become significant if the difference of
384 the auxiliary data used in two independent processing systems multiplied by the sensitivity
385 of the target variable to the respective auxiliary variable is larger than the target uncertainty.
386 This means that testing retrieval level consistency needs to assess auxiliary dataset
387 differences in the light of target variable sensitivities or incompatibilities.

388

389 **Consistency of categorical auxiliary datasets (“masks”)**

390 A first approach to assess consistency of masks used in independent retrievals lies in
391 visual inspection of combined maps of datasets, as for example, of surface temperature
392 composed from four independent CDRs for land (LST), sea surface (SST), ice (IST) and lake
393 surface water (LSWT) temperatures against required pixel-level agreement of the masks. In
394 CCI the four retrievals use a common land-sea mask (and sea-ice mask), but apply different
395 cloud mask algorithms optimized over land, sea ice and water surfaces. As shown in Figure 3,
396 the reader can visually confirm the absence of any obvious scatter near the land-sea borders,
397 which indicates that the land-sea masks used in the different processing systems are
398 consistent. Additionally, the application of different optimal cloud masking in the retrievals
399 for LST and SST has led to obvious discontinuities in the sampling with temperature
400 observations at the land-sea border, which may be judged as second-order inconsistencies.
401 Such visual inspection of a set of typical scenes can be employed for most ECVs to get an
402 understanding of their physical consistency within one variable across borders of the same
403 mask used in different retrievals. Additionally, Figure 3 shows a case where a contrast in the

404 values in the ECVs between neighbouring pixels (surface temperature of ocean and water)
405 does not mean inconsistency, but reflects physical differences arising from the different heat
406 capacities of water and land.

407 The retrieval of many ECVs needs a cloud mask to avoid cloud contamination. Also
408 cloud properties need a cloud mask to ensure that a pixel truly represents cloud (Poulson, et
409 al., 2012). When, for example, aerosol and cloud property retrievals for the same sensor are
410 implemented as separate algorithms (as is usually the case), individual pixels need to be
411 analysed either as cloud or as aerosol; analysis of the same pixel as aerosol and as cloud
412 under the wrong assumption (cloud-free or aerosol-free) could severely degrade the
413 retrievals and must be minimized (e.g. Sogacheva et al., 2017; Li et al., 2009). To assess if this
414 principle is fulfilled, independent AATSR cloud masks used in the aerosol and cloud products
415 were analysed for four days in September 2008 (covering difficult scenes with high aerosol
416 loads or complicated mixtures of aerosol and clouds). Figure 4 shows a map of different
417 combinations of cloud / no cloud assignment by the two cloud masks and a contingency
418 matrix of those class combinations. The matrix shows, that while 21% of observations are
419 not used for aerosol or cloud retrievals at all (losing sampling coverage but not leading to
420 inconsistency), only 0.3% of them were found to violate the physical principle (i.e., no pixel
421 must be double-analysed as clouds and as aerosols). Even if a very stringent threshold for
422 this fraction of 1% is set (since cloud mask errors lead to very large AOD errors) the two
423 cloud masks are fully consistent. The map also shows that the inconsistent cases (yellow
424 pixels) occur only over land but in all climate zones. Together with the underlying physical
425 principle one can use such a contingency matrix / mapping of class combinations to assess
426 the contingency of masks and to understand where / when inconsistencies mostly occur and
427 need to be corrected.

428 Another typical aspect of multi-quantity spatial consistency is the agreement of

429 locations between the outlines of physically related quantities (different products within one
430 ECV, between different ECVs). For example, glacier outlines are derived from high-resolution
431 satellite imagery or aerial photography using semi-automated mapping techniques or
432 manual on-screen digitization (Paul et al., 2015). Due to their higher spatial resolution, the
433 location of glaciers can be used for land cover as an independent validation source for its
434 “permanent ice and snow” classes. Furthermore, glacier maps serve as an important
435 auxiliary dataset for clouds and LST (to choose the correct retrieval algorithm), and lakes as a
436 reciprocal mask (these can only occur in places not covered by glaciers) for sea ice, ice sheets
437 and permafrost. Again, contingency matrices between glacier or lake location and the other
438 variables can be used to assess consistency in the light of the expected compatible
439 combinations; the threshold for the acceptable fraction of inconsistent pixels needs to be set
440 depending on the potential harm of misclassifications. A limitation for the assessment of
441 categorical auxiliary datasets lies in the fact that mixed cases often exist, in particular for
442 coarser spatial resolutions.

443

444 **Consistency of continuous auxiliary datasets of the same quantity**

445 Often the retrieval of a land / ocean CDR is affected by perturbations in the measured
446 bands due to atmospheric absorption or scattering, so an atmospheric correction needs to
447 be applied. Examples of necessary atmospheric corrections include visible or thermal
448 retrievals impacted by aerosol, water vapour, ozone or other trace gases (e.g. Popp 1995). A
449 first step in algorithm development would be to assess the sensitivities of the measured
450 reflectances to the various absorbing trace gases and to aerosol particles (e.g. Holzer-Popp,
451 et al., 2002). This provides the basis for deciding which corrections can be neglected or made
452 with a simple parameterization, and which need more precise corrections using an auxiliary
453 dataset of distributions of the responsible agents influencing the signal. When the auxiliary

454 datasets come from the same sensor as the target CDR, accurate spatio-temporal matching
455 (pixel colocation) would be possible. However, in cases where the auxiliary data come from
456 different sensors, it may be necessary to deal with spatial and temporal mismatches,
457 introducing a requirement for assessment of the associated additional uncertainties. Figure 5
458 shows a gridded map of differences of aerosol optical depth between the by-products of the
459 ocean-colour atmospheric correction of MERIS data (processed using a NASA algorithm) and
460 the corresponding CCI aerosol ECV product from AATSR, both for 865 nm (both sensors were
461 on-board the same platform ENVISAT and thus exhibit zero time difference). The global
462 average difference of AOD between both products of 0.03 is acceptable for the purpose of
463 aerosol corrections, but the variability is larger for the aerosol ECV product than for the
464 ocean colour product (0.10 ± 0.11 cf. 0.13 ± 0.04 respectively), which has higher AOD values
465 over the open ocean, but lower ones closer to land. Given the importance of AOD in ocean-
466 colour atmospheric correction (IOCCG, 2010), the aerosol-corrected ocean colour ECV can be
467 regarded as consistent with the aerosol ECV in its global average, but not regionally. These
468 results merit further investigations to identify the sources of the discrepancies and to assess
469 the potential to improve the MERIS ocean-colour atmospheric correction algorithm by using
470 concurrent auxiliary information on AOD from the main ECV aerosol product obtained from
471 AATSR.

472

473 4.3 Methods to assess scientific consistency

474 Scientific consistency includes self-consistency within one quantity (when
475 independently retrieved pieces are integrated into a longer time series or a larger map) and
476 mutual consistency between different quantities (different products of one ECV or multiple
477 ECV CDRs) as a consequence of all types of retrieval inconsistencies, limitations of the

478 retrieval algorithms or sensor calibrations, as well as sampling differences between
479 aggregated datasets.

480

481 **Self-consistency of a single quantity**

482 One major problem of satellite-based CDRs is that satellite instruments typically
483 survive in orbit only for a limited time, so that a long-term record needs to be constructed
484 from combining data from a time series of similar sensors. Plotting regional or global data
485 records of the related parts of a time series often allows visual inspection of their
486 consistency, where “jumps” or “breakpoints” are obvious against background knowledge of
487 any known or absent true discontinuities. As example, column-averaged dry-air mole
488 fractions (“vertical columns” X_{CO_2}) of carbon dioxide (Buchwitz et al., 2015) from the
489 greenhouse gas (GHG) ECV are selected. Those CDRs serve as input data for inverse
490 modelling schemes to improve the knowledge on natural and anthropogenic sources and
491 sinks (e.g. Reuter et al., 2017). In creating a multi-sensor CDR covering a longer time period,
492 a merging algorithm (EMMA, Reuter et al., 2013, 2020) corrects potential remaining offsets
493 of individual datasets to avoid jumps in the merged time series. In EMMA, the ensemble
494 members have been bias corrected and brought to common a priori CO_2 profiles before
495 being combined to obtain the merged product. Figure 6 shows at the top the resulting multi-
496 sensor, multi-algorithm monthly mean X_{CO_2} merged record for 2003-2018 for northern mid-
497 latitudes ($30^\circ N$ - $60^\circ N$) with the known nearly linear increase in time and seasonal cycle and
498 no remaining biases, while in the bottom panel differences between individual ensemble
499 members and the merged product before the corrections are shown to be larger than the
500 required X_{CO_2} uncertainties of 0.5 ppm. In this case, this threshold for the target
501 uncertainties of the gap-corrected merged dataset is defined by the user requirement for
502 the application of X_{CO_2} trend analysis.

503 Similarly, spatial inconsistencies in one variable can often be assessed visually by
504 looking at maps combined from independent pieces (different sensors, different overpass
505 times of the same sensor with different observing geometries, different algorithms). In this
506 case, inconsistencies are visible as artificial border lines or gradients that are larger than the
507 noise in the image. Again, physical understanding is needed to decide whether a
508 discontinuity at a physical border is real or erroneous (e.g. surface temperature often shows
509 true differences between land and sea as shown in Fig. 3, while a dust plume should be
510 continuous). Another example for spatial inconsistencies revealed by data overlay are glacier
511 outlines derived from satellite images that have been orthorectified with different digital
512 elevation models (DEMs). In steep and/or high topography geolocation shifts of several
513 pixels (about 30 - 90 m) can occur, making any change assessment (trend analysis) or joint
514 use of sensors nearly impossible (Kääb et al., 2016).

515 Another way of testing the consistency of independently retrieved CDRs for one
516 variable is by comparing estimates of a derived quantity such as their trend with a physical
517 equation. For example, within the GEWEX Water Vapour Assessment (G-VAP, see
518 <http://gewex-vap.org/> for details), inter-comparisons of total column water vapour (TCWV)
519 trend estimates from different CDRs were made and it was concluded that the trend
520 estimates are generally significantly different. It was then shown that several data records
521 disagree with the physical expectation from the Clausius-Clapeyron equation using data over
522 the global ice-free ocean (Schröder et al. 2016, 2019). After homogenisation, a new analysis
523 was applied to the trend estimates and associated results are shown in Figure 7. While the
524 diversity in original trend estimates (-0.15 to +0.12 kg/m²/year) is several times higher than
525 individual uncertainties, it is largely reduced after homogenisation (-0.02 to +0.04
526 kg/m²/year), but still slightly larger than the individual trend uncertainties (up to ± 0.01
527 kg/m²/year). As a consequence, after homogenization there was a significant increase in the

528 fraction of datasets that can be seen as consistent as indicated by agreement of trends
529 within twice their combined uncertainties.

530

531 **Mutual consistency between different quantities**

532 In testing multiple quantity consistency, the role of the underlying background
533 knowledge becomes stronger since the physical processes connecting different ECVs need to
534 be taken into account. One method to test the consistency of two ECVs is by looking at their
535 correlations. For example, in the lower stratosphere, the strong physical dependency of
536 lower stratospheric water vapour on tropical tropopause temperatures can be exploited to
537 test the consistency between climate data records of temperature and stratospheric water
538 vapour as highlighted by Hegglin et al. (2014). This study proposed a new merging method
539 that uses a chemistry-climate model as a transfer function between different satellite
540 instrument records to create a CDR. The methodology allows the bias between instruments
541 to be determined throughout the instrument's lifetime and not only for the overlap period
542 (when old instruments may show first signs of degradation), hence improving
543 characterization of systematic differences (or biases) between datasets. By using the
544 correlation between the newly merged stratospheric water vapour record and the zonal
545 mean temperature from ERA-interim, visual inspection indicated that the new merging
546 method led to physically more consistent results than the traditional one based on bias-
547 correction of instruments during overlap periods. Figure 8 shows the visually well correlated
548 time series of a prototype version of the stratospheric water vapour CDR merged using the
549 new methodology in comparison with zonal mean temperatures at 100 hPa from ERA5 in the
550 tropical region (left panel). We set the threshold for correlations to accept consistency with
551 medium (high) confidence to 0.5 (0.7) since the co-variability of time series of two different
552 variables may also be influenced by other processes which reduce the correlation. In this

553 case (right panel), a correlation of 0.58 suggests that the two variables are physically
554 consistent with medium confidence; if only assessing the last 15 years (not shown) with
555 better data quality, the correlation increases to 0.69 (consistency with high confidence).

556 Alternatively, differences of multi-year trend maps of one variable can be used to
557 assess the consistency of two different ECV CDRs. The example here is the inter-comparison
558 between wave height (measured by satellite altimetry) and sea ice concentrations assessed
559 in Stopa et al. (2016). Daily sea ice concentrations produced from the Special Sensor
560 Microwave Imager (SSM/I) by IFREMER (Ezraty et al., 2007) are used to define open ocean
561 versus sea ice conditions with a 15% concentration threshold at 12.5 km resolution within
562 the Arctic Ocean. For the period 1992-2014, the SSM/I ice concentrations are used along
563 with wind vectors from the Climate Forecast System Reanalysis to reproduce the wave field
564 through the numerical wave model, WAVEWATCH3 (WW3, Tolman et al., 2014), which
565 includes wave-ice interaction through an under-ice parameterization of wave dissipation.
566 Figure 9 shows a comparison of the trends of the significant wave height (H_s) directly
567 measured with altimetry (denoted ALT, Queffeulou and Croize-Fillon, 2015) and from the co-
568 located model data from WW3 (denoted WW3 CoLoc) in which SSM/I ice concentrations
569 have been used. Qualitatively, the regional patterns match between the two datasets,
570 despite stronger trends in the altimeters (of up to 1 cm / y). At present the confidence
571 interval for trends in wave heights is not known. Therefore the quantitative discrepancies
572 between modeled and measured trends here could be due to both systematic time-varying
573 biases in the wave height ECV, which are expected to be only a function of time and sensor,
574 or to a trend error in the surface wind reanalysis used to drive the wave model. However,
575 the wind trends are also constrained by sea level pressure data and sea ice drift (e.g. Spreen
576 et al. 2011). In the future, a wider range of ECVs, combined with in situ data and models,
577 may be used for a quantitative refinement of sea state trends.

578 An example of testing the (anti-)correlation of multiple regional ECV CDRs as
579 predicted by physical theory, is the use of the El Niño Southern Oscillation (ENSO) index for
580 ECV anomalies in the tropical Pacific Ocean Niño3.4 region (5°S-5°N, 190°E-240°E). This
581 natural phenomenon is an ideal candidate for investigating multiple ECV consistency due to
582 its relatively short timescale, large amplitude and multiple ECVs affected by it. This first
583 attempt focusses on the main ENSO signatures at large scale. Physical or biological processes
584 leading to spatio-temporal lags of a few months or a few degrees longitude between some
585 variables have been neglected. This could be refined in future studies. ENSO variability is
586 compared in several ocean (SST, SL, SSS, Chlor_a), atmosphere (CFChigh, TCWV, AOD550)
587 and land (SM, burned area / fire) ECV products - see Table A1 for the acronyms, more
588 detailed information on the datasets and their correlations. All variables were interpolated
589 to a common 1 by 1° grid, de-seasonalised by removing the corresponding monthly mean
590 value and normalised by dividing by the standard deviations for their respective available
591 time period. Figure 10 shows the index variability across the tropical Pacific Ocean for the
592 ECVs in time-longitude anomaly cross-sections. The ocean and most atmosphere ECV time
593 series show consistent spatio-temporal co-variability, as expected. Whereas SST and SL have
594 their largest variability in the Niño3.4 region, CFChigh and TCWV variability peak further west
595 (~180°E), except for the strong El Niño years 1982/83, 1997/98 and 2015/16. Moreover, SSS
596 and Chlor_a are anti-correlated with SST, as expected from a reduced upwelling. For ECVs
597 affected indirectly by El Niño from dry conditions and wild fires over Indonesia (fire, aerosol
598 and soil moisture), the highest correlations occur in their Indonesian time series (10°S-10°N,
599 100°E-150°E). For certain El Niño3.4 years, e.g. 1997, 2007 and 2015 there are clear
600 indicators of co-variability of them and SST (Fig. 10g). Here again the use of a correlation
601 threshold of 0.5 (0.7) for medium (high) confidence on consistency is adopted. In conclusion,
602 by quantifying (anti-) correlations between these nine independently derived satellite ECVs

603 versus the scientific understanding of the ENSO phenomenon, a medium (high) confidence in
604 their consistency can be shown for eight (four) of them.

605

606 4.4 State of consistency assessments for CCI ECVs

607 Several examples of closure / budget studies of partial Earth system cycles
608 demonstrate the usefulness of CCI (and several other) CDRs that are consistent at all three
609 levels. For example, closure of the carbon budget is still an outstanding scientific challenge
610 (Le Quéré et al. 2018) that is impacted by CDR inconsistencies. Different CCI products
611 provide direct and indirect constraints on carbon fluxes that help to improve the consistency
612 of carbon budgets. For example, CCI greenhouse gas products are used to inform
613 atmospheric inversions. Top down inversion results can be complemented by other ECVs to
614 attribute diagnosed fluxes to different components such as biomass carbon changes
615 (biomass CCI product), fire emissions (CCI products on burned area and fire size) and land
616 use change emissions (land cover CCI products).

617 Another example is the regional closure of the water budget. Based on multiple
618 satellite ECVs it has been demonstrated that the water budget can be closed within less than
619 10% uncertainty at a continental annual time scale, while, at monthly time scales, its
620 residuals and uncertainty estimates are larger (about 20%; Rodell et al., 2015). These
621 uncertainties in the water budget closure can be reduced by introducing additional
622 constraints, e.g. by using multiple CDRs with different uncertainties of a single quantity or by
623 additionally forcing closure of the atmosphere and ocean terms. Uncertainties in existing
624 CDRs need to be further reduced and new CDRs of other key variables (most importantly,
625 river discharge and irrigation water use) need to be included or developed to reach the 5%
626 closure error targeted by GCOS (GCOS, 2016).

627 The global mean sea level budget closure has also been assessed within the CCI
628 program by comparing the sum of changes in ocean thermal expansion, land ice melt and
629 liquid water storage on continents with the total observed sea level change. The latter can
630 be estimated globally from satellite altimetry with an accuracy of about 10% on different
631 time scales (e.g. WCRP sea level budget group, 2018). These observations enable closure of
632 the trend in sea level budget with an uncertainty of ± 0.3 mm/year over the last 25 years. The
633 sea level budget involves additional variables from the global water budget (through land ice
634 and liquid water components) and from the global energy budget (through thermal
635 expansion directly related to global ocean heat content; Meyssignac et al. 2017) and thus
636 connects the energy and water budgets. At regional scale, uncertainties in the observed
637 components of the sea level budget are considerably larger (few tens of percent) and need
638 to be further reduced to reach the regional GCOS target.

639 Finally, an assessment of the current state of affairs regarding consistency between
640 the CDRs of the CCI program was made based on the combined scientific expertise of the CCI
641 community; it is not meant to be exhaustive but intended as initial guidance for the use of
642 multiple ECV CDRs or for defining priorities in further consistency analysis. Table 3 provides
643 for each pair of CDRs the consistency status as either: “no evident need to consider
644 consistency”, “further studies needed”, “consistency explicitly ensured by shared processing
645 or co-retrieving”, or “studies already performed”, referenced to Table A2 with the underlying
646 publication or technical report (characterized as “theoretical”, “exemplary / partial” or
647 “comprehensive”). As can be seen from Table 3, quite some work remains to be done where
648 the definition and concept presented in this paper can be applied and further refined.

649 **5. Summary and conclusions**

650 Climate Data Records of Essential Climate Variables derived from satellite
651 instruments provide essential information to monitor the state of the Earth system and its
652 changing climate. A key requirement for these CDRs to be useful for Earth system science
653 applications is that the CDRs are internally and mutually consistent. The ESA CCI program
654 provides a set of CDRs for 21 GCOS ECVs in a common framework, and from the outset has
655 invested heavily in establishing their consistency, as presented in this study. To our
656 knowledge no comprehensive definition of CDR consistency exists. Therefore a three-level
657 definition of consistency applicable to single- and multiple-variable cases is proposed and a
658 concept for assessing if two or more CDRs are consistent with each other and possibly with
659 reference data is presented. On the technical level, straightforward data access and usage,
660 including availability of comprehensive documentation and product user guides, is needed.
661 On the retrieval level, one needs to limit contradictions in the use of auxiliary datasets
662 (masks or continuous fields) of the same variables in separate processing chains. On the
663 scientific level, consistency of multiple ECV CDRs means judging their relevant correlations,
664 patterns, periodicity, trends, etc. (as appropriate for a given variable, process or cycle) in the
665 light of underlying physical background knowledge (e.g., by jointly confronting them with a
666 model). Through this link with background knowledge, “consistency” as defined in this study
667 goes beyond “agreement” and relies rather on “compatibility”. Finding inconsistencies in one
668 or more ECV dataset(s) (i.e. patterns whose disagreements exceed underlying uncertainties,
669 contradict physical principles or a well-founded model) often indicates errors in a dataset or
670 model whose resolution can lead to new scientific understanding.

671 This study also provides an overview of the technical consistency of CCI CDRs
672 (common format and metadata standards, common data portal, harmonized
673 documentation, common uncertainty reporting). An open issue in this regard is

674 harmonization across programs and communities. Here, the CCI program has made an
675 important step by adopting the netCDF format, with the CF and ACDD conventions (the de-
676 facto standard in the modelling community) for its gridded satellite data records. The
677 Climate Data Store (CDS) of the Copernicus Climate Change Service (C3S) is also based largely
678 on CCI standards. Such common standards are a prerequisite for the use of automated data
679 services for accessing multiple data sources with little manual interaction, hence facilitating
680 use of the data in scientific studies across multiple ECVs.

681 The discussion of a concept for assessing consistency and related methods on the
682 retrieval and scientific level shows how consistency with regard to different categorical and
683 continuous auxiliary datasets can be tested and how the assessment of single-variable self-
684 consistency and multiple quantity mutual consistency can be conducted. In all these
685 methods a basic understanding of “the truth” needs to be employed. A relevant
686 characteristic of an ECV and an appropriate metric (e.g. bias, correlation, contingency matrix,
687 ...) for its evaluation need to be chosen. A tabular summary of different methods to assess
688 consistency is given in Table 2. For each of the different metrics, a threshold needs to be
689 defined to judge on consistency of two datasets. This may well differ from commonly applied
690 thresholds for validation purposes since also other processes than consistency may affect
691 the datasets. We suggest as a minimum requirement that each consistency study states the
692 applied thresholds, as is done for the examples in this paper. Whereas the methods used to
693 assess consistency rely on well-established tools for calibration and validation, placing them
694 into the systematic context with relevance to consistency as done here, can serve as a
695 practical guideline to consistency assessment. A brief high-level analysis of the inter-
696 dependencies of CCI ECVs at the retrieval and scientific levels (Table 1) is provided to
697 understand where consistency is needed and thus needs to be checked. Finally a high-level

698 assessment of the current state of affairs regarding consistency assessment between the
699 CDRs of the CCI program (Table 3) is compiled to outline possible further research needs.

700 When discussing consistency, datasets from sources other than satellite data (e.g.
701 Earth system models) are often required to comprehensively study an Earth system cycle,
702 and their uncertainties also need to be considered, together with uncertainties in simplified
703 or estimated budget equations. It is well understood that establishing consistency between
704 two or more variables requires targeted analysis. Within and outside CCI much effort has
705 been spent on quantifying the sensitivities and dependencies of the retrieved quantities.
706 However, a lot more remains to be done in this area.

707 **6. Acknowledgements**

708 This study is based on ongoing work of altogether 30 projects of the ESA Climate
709 Change Initiative (23 ECV projects, the Climate Model User Group project, cross-cutting
710 outreach components on portal, toolbox, visualisation; CCI data standards and system
711 engineering working group). We are grateful to ESA for creating the CCI program which has
712 strengthened the consistency of the many research communities related to developing,
713 processing, qualifying and using satellite CDRs. We are grateful to the several hundred
714 scientists building the CCI community for making a consistent Earth observation based data
715 repository real. The “operational” part of the CCI program has been transferred to the
716 Copernicus Climate Change Service (C3S, (re-)processing to extend the CDRs, associated
717 quality control, user support). We are also thankful for many other datasets from outside CCI
718 and C3S which help cover all relevant ECVs: GOSAT Level 1 data from JAXA, GOSAT Level 2
719 data from NIES and NASA, OCO-2 Level 1 and Level 2 data from NASA, HOAPS data from
720 EUMETSAT CM SAF, water vapour records from the G-VAP data archive, CAMS and ERA-5
721 data from ECMWF / Copernicus Atmosphere and Climate Change Services , SSM/I daily sea
722 ice concentrations from IFREMER, and wind vectors from the Climate Forecast System
723 Reanalysis.

724 **Appendix**
725
726 Tables A1 and A2

727 **7. References**

- 728 Allan, R. P., 2012: The Role of Water Vapour in Earth's Energy Flows. *Surv. Geophys.*, 33, 557
729 – 564, 33:557–564 DOI 10.1007/s10712-011-9157-8.
- 730 Andersson, A., Graw, K., Schröder, M., Fennig, K., Liman, J., Bakan, S., Hollmann, R., Klepp, C.,
731 2017: Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data - HOAPS 4.0.
732 Satellite Application Facility on Climate Monitoring,
733 https://doi.org/10.5676/EUM_SAF_CM/HOAPS/V002.
- 734 Bates, JJ, Privette, JL, Kearns, EJ; Glance, W, Zhao, XP, 2016: Sustained production of
735 multidecadal climate records Lessons from the NOAA Climate Data Record Program. *Bull.*
736 *Amer. Meteor. Soc.*, 97, 1573 – 1581, doi/10.1175/BAMS-D-15-00015.1.
- 737 Bevan, S.L., North, P.R.J., Los, S.O., Grey, W.M.F, 2012: A global dataset of atmospheric
738 aerosol optical depth and surface reflectance from AATSR. *Rem. Sens. Environ.*, 116, 119–
739 210.
- 740 Bojinski, S., M. Verstraete, T.C. Peterson, C. Richter, A. Simmons, and M. Zemp, 2014: The
741 Concept of Essential Climate Variables in Support of Climate Research, Applications, and
742 Policy. *Bull. Amer. Meteor. Soc.*, 95, 1431–1443, [https://doi.org/10.1175/BAMS-D-13-](https://doi.org/10.1175/BAMS-D-13-00047.1)
743 [00047.1](https://doi.org/10.1175/BAMS-D-13-00047.1).
- 744 Boutin, J.; Vergely, J.-L.; Koehler, J.; Rouffi, F.; Reul, N. (2019): ESA Sea Surface Salinity
745 Climate Change Initiative (Sea_Surface_Salinity_cci): Version 1.8 data collection. Centre for
746 Environmental Data Analysis, 2019, doi:10.5285/9ef0ebf847564c2eabe62cac4899ec41.
747 <http://dx.doi.org/10.5285/9ef0ebf847564c2eabe62cac4899ec41>
- 748 Buchwitz, M., M. Reuter, O. Schneising, H. Boesch, S. Guerlet, B. Dils, I. Aben, R. Armante, P.
749 Bergamaschi, T. Blumenstock, H. Bovensmann, D. Brunner, B. Buchmann, J. P. Burrows, A.
750 Butz, A. Chedin, F. Chevallier, C. D. Crevoisier, N. M. Deutscher, C. Frankenberg, F. Hase, O.
751 P. Hasekamp, J. Heymann, T. Kaminski, A. Laeng, G. Lichtenberg, M. De Maziere, S. Noel, J.

752 Notholt, J. Orphal, C. Popp, R. Parker, M. Scholze, R. Sussmann, G. P. Stiller, T. Warneke, C.
753 Zehner, A. Bril, D. Crisp, D. W. T. Griffith, A. Kuze, C. ODell, S. Oshchepkov, V. Sherlock, H.
754 Suto, P. Wennberg, D. Wunch, T. Yokota, Y. Yoshida, 2015: The Greenhouse Gas Climate
755 Change Initiative (GHG-CCI): comparison and quality assessment of near-surface-sensitive
756 satellite-derived CO₂ and CH₄ global data sets. *Rem. Sens. Environ.*, 162, 344-362,
757 doi:10.1016/j.rse.2013.04.024.

758 Chuvieco, E., Mouillot, F., van der Werf, G.R., San Miguel, J., Tanasse, M., Koutsias, N.,
759 García, M., Yebra, M., Padilla, M., Gitas, I., Heil, A., Hawbaker, T.J., & Giglio, L. (2019).
760 Historical background and current developments for mapping burned area from satellite
761 Earth observation. *Remote Sensing of Environment*, 225, 45-64.

762 Doherty, S.J., S. Bojinski, A. Henderson-Sellers, K. Noone, D. Goodrich, N.L. Bindoff, J.A.
763 Church, K.A. Hibbard, T.R. Karl, L. Kajfez-Bogataj, A.H. Lynch, D.E. Parker, I.C. Prentice, V.
764 Ramaswamy, R.W. Saunders, M.S. Smith, K. Steffen, T.F. Stocker, P.W. Thorne, K.E.
765 Trenberth, M.M. Verstraete, and F.W. Zwiers, 2009: Lessons Learned from IPCC AR4:
766 Scientific Developments Needed to Understand, Predict, and Respond to Climate Change.
767 *Bull. Amer. Meteor. Soc.*, 90, 497 – 514, <https://doi.org/10.1175/2008BAMS2643.1>.

768 Dorigo, W.A., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl,
769 M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R.,
770 Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C.,
771 Reimer, C., van der Schalie, R., Seneviratne, S. I. Smolander, T., Lecomte, P., 2017: ESA CCI
772 Soil Moisture for improved Earth system understanding: State-of-the art and future
773 directions. *Rem. Sens. Environ.*, 203, 185 – 215, <https://doi.org/10.1016/j.rse.2017.07.001>.

774 ESA, 2019: CCI Data Standards version 2.0. ESA Climate Office, CCI-PRGM-EOPS-TN-13-0009,
775 http://cci.esa.int/sites/default/files/filedepot/CCIDataStandards_v2-0_CCI-PRGM-EOPS-
776 [TN-13-0009.pdf](http://cci.esa.int/sites/default/files/filedepot/CCIDataStandards_v2-0_CCI-PRGM-EOPS-TN-13-0009.pdf)

777 Ezraty, R., Girard-Ardhuin, F., Piolle, J. F., Kaleschke, L., and Heygster, G., 2007: Arctic and
778 Antarctic sea ice concentration and Arctic sea ice drift estimated from Special Sensors
779 Microwave data. User manual version 2.1. Ifremer/CERSAT,
780 <ftp.ifremer.fr/ifremer/cersat/products/gridded/psi-drift/documentation/ssmi.pdf>.

781 Global Observing System for Climate, 2016: Implementation Needs. WMO, GCOS-200.
782 <http://library.wmo.int>.

783 Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., and Dorigo, W., 2019: Evolution of
784 the ESA CCI Soil Moisture Climate Data Records and their underlying merging
785 methodology. *Earth Syst. Sci. Data*, 11, 717-739, <https://doi.org/10.5194/essd-2019-21>.

786 Hegglin, M. I., D. Plummer, J. Scinocca, T. G. Shepherd, J. Anderson, L. Froidevaux, B. Funke,
787 D. Hurst, A. Rozanov, J. Urban, T. v. Clarmann, K. A. Walker, R. Wang, S. Tegtmeier, K.
788 Weigel, 2014: Variation of stratospheric water vapour trends with altitude from merged
789 satellite data. *Nature Geoscience*, 7, 768 – 776, doi: 10.1038/NGEO2236.

790 Hollmann, R., C.J. Merchant, R. Saunders, C. Downy, M. Buchwitz, A. Cazenave, E. Chuvieco,
791 P. Defourny, G. de Leeuw, R. Forsberg, T. Holzer-Popp, F. Paul, S. Sandven, S.
792 Sathyendranath, M. van Roozendaal, and W. Wagner, 2013: The ESA Climate Change
793 Initiative: Satellite Data Records for Essential Climate Variables. *Bull. Amer. Meteor. Soc.*,
794 94, 1541–1552, <https://doi.org/10.1175/BAMS-D-11-00254.1>.

795 Holzer-Popp, Th., Bittner, M., Borg, E., Dech, St., Erbertseder, Th., Fichtelmann, B.,
796 Schroedter, M., 2002: Das automatische Atmosphärenkorrekturverfahren „DurchBlick“, in:
797 Blaschke, T. (ed.), *Fernerkundung und GIS: Neue Sensoren – innovative Methoden*, H.
798 Wichmann Verlag, Heidelberg, ISBN 3-87907-369-4, S. 78 – 87.

799 Immler, F. J., J. Dykema, T. Gardiner, D. N. Whiteman, P. W. Thorne, and H. Vomel, 2010:
800 Reference Quality Upper-Air Measurements: guidance for developing GRUAN data
801 products. *Atmos. Meas. Tech.*, 3, 1217–1231, doi:10.5194/amt-3-1217-2010.

802 IOCCG, 2010: Atmospheric Correction for Remotely-Sensed Ocean-Colour Products. Wang,
803 M. (ed.), Reports of the International Ocean-Colour Coordinating Group, No.10, IOCCG,
804 Dartmouth, Canada.

805 IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group
806 I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change
807 [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V.
808 Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and
809 New York, NY, USA, 1535 pp, doi:10.1017/CBO9781107415324.

810 IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and
811 III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core
812 Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

813 IPCC, 2018: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global
814 warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission
815 pathways, in the context of strengthening the global response to the threat of climate
816 change, sustainable development, and efforts to eradicate poverty [Masson-Delmotte, V.,
817 P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C.
818 Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T.
819 Maycock, M. Tignor, and T. Waterfield (eds.)]. World Meteorological Organization, Geneva,
820 Switzerland.

821 IPCC, 2019a: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O.
822 Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K.
823 Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. In
824 press.

825 IPCC, 2019b: Climate Change and Land: an IPCC special report on climate change,
826 desertification, land degradation, sustainable land management, food security, and

827 greenhouse gas fluxes in terrestrial ecosystems [P.R. Shukla, J. Skea, E. Calvo Buendia, V.
828 Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van
829 Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P.
830 Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press.

831 Kääb, A.; Winsvold, S.H.; Altena, B.; Nuth, C.; Nagler, T.; Wuite, J., 2016: Glacier Remote
832 Sensing Using Sentinel-2. Part I: Radiometric and Geometric Performance, and Application
833 to Ice Velocity. *Remote Sens.* 8, 598.

834 Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A.,
835 Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneeth, A., Arora, V. K., Barbero, L., Bastos,
836 A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris,
837 I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K.,
838 Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P.,
839 Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S.,
840 Nakaoka, S.-I., Neill, C., Olsen, A., Ono, T., Patra, P., Peregón, A., Peters, W., Peylin, P., Pfeil,
841 B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck,
842 C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P.,
843 Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N.,
844 Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., Zheng, B., 2018: Global Carbon Budget
845 2018. *Earth Syst. Sci. Data*, 10, 2141-2194, <https://doi.org/10.5194/essd-10-2141-2018>.

846 Legeais, J.-F., Ablain, M., Zawadzki, L., Zuo, H., Johannessen, J. A., Scharffenberg, M. G.,
847 Fenoglio-Marc, L., Fernandes, M. J., Andersen, O. B., Rudenko, S., Cipollini, P., Quartly, G.
848 D., Passaro, M., Cazenave, A., and Benveniste, J., 2018: An improved and homogeneous
849 altimeter sea level record from the ESA Climate Change Initiative. *Earth Syst. Sci. Data*, 10,
850 281-301, <https://doi.org/10.5194/essd-10-281-2018>.

851 Levizzani, V. and Cattani, E., 2019: Satellite Remote Sensing of Precipitation and the
852 Terrestrial Water Cycle in a Changing Climate, *Remote Sensing*, 11(19), 2301.

853 Li, Z., Zhao, X., Kahn, R. A., Mishchenko, M., Remer, L., Lee, K.-H., Wang, M., Laszlo, I.,
854 Nakajima, T., and Maring, H. 2009: Uncertainties in satellite remote sensing of aerosols and
855 impact on monitoring its long-term trend: a review and perspective, *Ann. Geophys.* 27,
856 2755–2770.

857 Lizundia-Loiola, J., Otón, G., Ramo, R., & Chuvieco, E. (2020). A spatio-temporal active-fire
858 clustering approach for global burned area mapping at 250 m from MODIS data. *Remote*
859 *Sensing of Environment*, 236, 111493.

860 Merchant, C. J., Paul, F., Popp, T., Ablain, M., Bontemps, S., Defourny, P., Hollmann,
861 R., Lavergne, T., Laeng, A., de Leeuw, G., Mittaz, J., Poulsen, C., Povey, A. C., Reuter,
862 M., Sathyendranath, S., Sandven, S., Soeiva, V. F. and Wagner, W., 2017: Uncertainty
863 information in climate data records from Earth observation. *Earth Syst. Sci. Data*, 9, 511 –
864 527, <https://doi.org/10.5194/essd-9-511-2017>.

865 Merchant, C.J. et al, 2019: Satellite-based time-series of sea-surface temperature since
866 1981 for climate applications, in preparation for Nature Science Data.

867 Meyssignac, B., A.B. Slangen, A. Melet, J.A. Church, X. Fettweis, B. Marzeion, C. Agosta, S.R.
868 Ligtenberg, G. Spada, K. Richter, M.D. Palmer, C.D. Roberts, and N. Champollion, 2017:
869 Evaluating Model Simulations of Twentieth-Century Sea-Level Rise. Part II: Regional Sea-
870 Level Changes. *J. Climate*, 30, 8565–8593, <https://doi.org/10.1175/JCLI-D-17-0112.1>.

871 NASA, 2014: OMI/Aura and MODIS/Aqua Merged Cloud Product:
872 https://cmr.earthdata.nasa.gov/search/concepts/C1265734652-GES_DISC.html.

873 North, P., Briggs, S., Plummer, S. & Settle, J., 1999: Retrieval of land surface bidirectional
874 reflectance and aerosol opacity from ATSR-2 multiangle imagery. *IEEE Trans. Geosci. Rem.*
875 *Sens.*, 37, 526 - 537.

876 Paul, F. and 24 co-authors, 2015: The Glaciers Climate Change Initiative: Algorithms for
877 creating glacier area, elevation change and velocity products. *Rem. Sens. Environ.*, 162, 408
878 - 426.

879 Poulsen, C.A., Siddans, R., Thomas, G.E., Sayer, A.M., Grainger, R.G., Campmany, E., Dead,
880 S.M., Arnold, C.; Watts, P.D., 2012: Cloud retrievals from satellite data using optimal
881 estimation: Evaluation and application to ATSR, *Atmos. Meas. Tech.*, 5, 1889–1910.

882 Popp, T., 1995: Correcting atmospheric masking to retrieve the spectral albedo of land
883 surfaces from satellite measurements, *International Journal of Remote Sensing*, 16, 3483-
884 3508.

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

892 Pulliainen J., M. Aurela, T. Laurila, T. Aalto, M. Takala, M. Salminen, M. Kulmala, A. Barr, M.
893 Heimann, A. Lindroth, A. Laaksonen, C. Derksen, A. Mäkelä, T. Markkanen, J. Lemmetyinen,
894 J. Susiluoto, S. Dengel, I. Mammarella, J.-P. Tuovinen, T. Vesala, 2017: Early snowmelt
895 significantly enhances boreal springtime carbon uptake. *PNAS*, 114, 11081-11086,
896 <https://doi.org/10.1073/pnas.1707889114>.

897 Quartly, G. D., Legeais, J.-F., Ablain, M., Zawadzki, L., Fernandes, M. J., Rudenko, S., Carrère,
898 L., García, P. N., Cipollini, P., Andersen, O. B., Poisson, J.-C., Mbajon Njiche, S., Cazenave, A.,
899 and Benveniste, J., 2017: A new phase in the production of quality-controlled sea level
900 data. *Earth Syst. Sci. Data*, 9, 557-572, <https://doi.org/10.5194/essd-9-557-2017>.

901 Queffelec, P. and Croize-Fillon, D., 2015: Global altimeter SWH data set, Technical Report
902 11.2, IFREMER/CERSAT, [ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/
903 waves/documentation/altimeter_wave_merge__11.2.pdf](ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/waves/documentation/altimeter_wave_merge__11.2.pdf).

904 Reuter, M., H. Boesch, H. Bovensmann, A. Bril, M. Buchwitz, A. Butz, J. P. Burrows, C. W.
905 O'Dell, S. Guerlet, O. Hasekamp, J. Heymann, N. Kikuchi, S. Oshchepkov, R. Parker, S.
906 Pfeifer, O. Schneising, T. Yokota, and Y. Yoshida, 2013: A joint effort to deliver satellite
907 retrieved atmospheric CO₂ concentrations for surface flux inversions: the ensemble median
908 algorithm EMMA. *Atmos. Chem. Phys.*, 13, 1771 - 1780.

909 Reuter, M., M. Buchwitz, M. Hilker, J. Heymann, H. Bovensmann, J. Burrows, S. Houweling, Y.
910 Liu, R. Nassar, F. Chevallier, P. Ciais, J. Marshall, and M. Reichstein, 2017: How much CO₂ is
911 taken up by the European terrestrial biosphere? *Bull. Amer. Meteor. Soc.*, 98, 665 - 671,
912 doi:10.1175/BAMS-D-15-00310.1.

913 Reuter, M., Buchwitz, M., Schneising, O., Noel, S., Bovensmann, H., Burrows, J. P., Boesch, H.,
914 Di Noia, A., Anand, J., Parker, R. J., Somkuti, P., Wu, L., Hasekamp, O. P., Aben, I., Kuze, A.,
915 Suto, H., Shiomi, K., Yoshida, Y., Morino, I., Crisp, D., O'Dell, C., Notholt, J., Petri, C.,
916 Warneke, T., Velasco, V., Deutscher, N. M., Griffith, D. W. T., Kivi, R., Pollard, D., Hase, F.,
917 Sussmann, R., Te, Y. V., Strong, K., Roche, S., Sha, M. K., De Maziere, M., Feist, D. G., Iraci, L.
918 T., Roehl, C., Retscher, C., and Schepers, D., 2020: Ensemble-based satellite-derived carbon
919 dioxide and methane column-averaged dry-air mole fraction data sets (2003-2018) for
920 carbon and climate applications, *Atmos. Meas. Tech.*, 13, 789-819,
921 <https://doi.org/10.5194/amt-13-789-2020>.

922 Rodell, M., H.K. Beaudoin, T.S. L'Ecuyer, W.S. Olson, J.S. Famiglietti, P.R. Houser, R. Adler,
923 M.G. Bosilovich, C.A. Clayson, D. Chambers, E. Clark, E.J. Fetzer, X. Gao, G. Gu, K. Hilburn,
924 G.J. Huffman, D.P. Lettenmaier, W.T. Liu, F.R. Robertson, C.A. Schlosser, J. Sheffield, and

925 E.F. Wood, 2015: The Observed State of the Water Cycle in the Early Twenty-First Century.
926 J. Climate, 28, 8289–8318, <https://doi.org/10.1175/JCLI-D-14-00555.1>.

927 Sathyendranath, S, Brewin, RJW, Müller, D, Brockmann, C, Deschamps, P-Y, Doerffer, R,
928 Fomferra, N, Franz, BA, Grant, MG, Hu C, Krasemann, H, Lee, Z, Maritorea, S, Devred, E,
929 Mélin, F, Peters, M, Smyth, T, Steinmetz, F, Swinton, J, Werdell, J, Regner, P, 2012: Ocean
930 Colour Climate Change Initiative: Approach and Initial Results. IGARSS, 2012, 2024 – 2027,
931 <http://dx.doi.org/10.1109/IGARSS.2012.6350979>.

932 Schröder, M., M. Lockhoff, J. Forsythe, H. Cronk, T. H. Vonder Haar, R. Bennartz, 2016: The
933 GEWEX water vapor assessment (G-VAP) – results from the trend and homogeneity
934 analysis. J. Appl. Meteor. Clim., 55, 1633 - 1649, doi: /10.1175/JAMC-D-15-0304.1.

935 Schröder, M., Lockhoff, M., Fell, F., Forsythe, J., Trent, T., Bennartz, R., Borbas, E., Bosilovich,
936 M. G., Castelli, E., Hersbach, H., Kachi, M., Kobayashi, S., Kursinski, E. R., Loyola, D., Mears,
937 C., Preusker, R., Rossow, W. B., and Saha, S.: The GEWEX Water Vapor Assessment archive
938 of water vapour products from satellite observations and reanalyses, Earth Syst. Sci. Data,
939 10, 1093–1117, <https://doi.org/10.5194/essd-10-1093-2018>, 2018.

940 Schröder, M., M. Lockhoff, L. Shi, T. August, R. Bennartz, H. Brogniez, X. Calbet, F. Fell, J.
941 Forsythe, A. Gambacorta, S.-P. Ho, E. R. Kursinski, A. Reale, T. Trent, Q. Yang, 2019: The
942 GEWEX water vapor assessment of global water vapour and temperature data records
943 from satellites and reanalyses. Rem. Sens., 11, 251, <https://doi.org/10.3390/rs11030251>.

944 Schulz, J., Albert, P., Behr, H.-D., Caprion, D., Deneke, H., Dewitte, S., Dürr, B., Fuchs, P.,
945 Gratzki, A., Hechler, P., Hollmann, R., Johnston, S., Karlsson, K.-G., Manninen, T., Müller, R.,
946 Reuter, M., Riihelä, A., Roebeling, R., Selbach, N., Tetzlaff, A., Thomas, W., Werscheck, M.,
947 Wolters, E., and Zelenka, A., 2009: Operational climate monitoring from space: the
948 EUMETSAT Satellite Application Facility on Climate Monitoring (CM-SAF). Atmos. Chem.
949 Phys., 9, 1687 - 1709, <https://doi.org/10.5194/acp-9-1687-2009>.

950 Siddans, R., 2016: S5P-NPP Cloud Processor ATBD, EUMETSAT Document number S5P-NPPC-
951 RAL-ATBD-0001, available from
952 <https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-NPP-ATBD-NPP-Clouds>.

953 Sogacheva, L., Kolmonen, P., Virtanen, T. H., Rodriguez, E., Saponaro, G., and de Leeuw, G.,
954 2017: Post-processing to remove residual clouds from aerosol optical depth retrieved using
955 the Advanced Along Track Scanning Radiometer, *Atmos. Meas. Tech.*, 10, 491-505,
956 doi:10.5194/amt-10-491-2017.

957 Sogacheva Larisa, Thomas Popp, Andrew M. Sayer, Oleg Dubovik, Michael J. Garay, Andreas
958 Heckel, N. Christina Hsu, Hiren Jethva, Ralph A. Kahn, Pekka Kolmonen, Miriam Kosmale,
959 Gerrit de Leeuw, Robert C. Levy, Pavel Litvinov, Alexei Lyapustin, Peter North, Omar Torres
960 and Antti Arola, Merging regional and global aerosol optical depth records from major
961 available satellite products, *Atmospheric Chemistry and Physics*, 20, 2031 – 2056,
962 <https://doi.org/10.5194/acp-20-2031-2020>, 2020

963 Spreen, G., R. Kwok, and D. Menemenlis (2011), Trends in Arctic sea ice drift and role of wind
964 forcing: 1992–2009, *Geophys. Res. Lett.*, 38, L19501, doi:10.1029/2011GL048970.

965 Stengel M. et al., 2019: Cloud_cci AVHRR-PM dataset version 3: a 35yr spanning climatology
966 of global cloud and radiation properties” in preparation, to be submitted to *Earth Syst. Sci.*
967 *Data*

968 Stopa, J. E., F. Ardhuin, F. Girard-Ardhuin, 2016: Wave climate in the Arctic 1992-2014:
969 seasonality and trends. *The Cryosphere*, 10, 1605 – 1629.

970 Tolman, H. L. and the WAVEWATCH III Development Group, 2014: User Manual and system
971 documentation of WAVEWATCH III version 4.18. NOAA, Technical Note 316,
972 NOAA/NWS/NCEP/MMAB.

973 Vinogradova, N., Lee, T., Boutin, J., Drushka, K., Fournier, S., Sabia, R., Stammer, D., Bayler,
974 E., Reul, N., Gordon, A., Melnichenko, O., Li, L., Hackert, E., Martin, M., Kolodziejczyk, N.,

975 Hasson, A., Brown, S., Misra, S., & Lindstrom, E., 2019: Satellite Salinity Observing System:
976 Recent Discoveries and the Way Forward. *Frontiers in Marine Science*, 6, 23p.
977 <https://doi.org/10.3389/fmars.2019.00243>.

978 Waugh, D. W. and Eyring, V., 2008: Quantitative performance metrics for stratospheric-
979 resolving chemistry-climate models. *Atmos. Chem. Phys.*, 8, 5699 - 5713,
980 <https://doi.org/10.5194/acp-8-5699-2008>.

981 WGClimate ECV Inventory Gap Analysis Report V1.1, 2018: available from
982 http://ceos.org/document_management/Working_Groups/WGClimate/Documents/WGClimate_ECV-Inventory_Gap_Analysis_Report_v1.1.pdf.

984 **Table captions**

985 **Table 1:** Links between ECVs on the retrieval (above the diagonal) and scientific (below the
986 diagonal) level which need to be consistent if used together. Weak linkages are indicated in
987 brackets. Cycles are indicated with the following acronyms: C=carbon cycle, W=water cycle,
988 E=energy cycle. Processes are indicated with the following acronyms: r=radiation interaction,
989 d=deposition, e=emission / evaporation, t=transport, c=chemical transformation,
990 mtf=melting / thawing / freezing, i=ecosystem interaction, a=air sea fluxes of carbon and
991 water, m=mask.

992 **Table 2:** Summary of assessment methods for consistency **on different levels and types**

993 **Table 3:** Consistency analysis status between pairs of CCI ECVs: intrinsically assured (*), study
994 needed (X), study done (c = comprehensive, e = exemplary, t = theoretical) - empty fields
995 indicate that no study is needed, this link cannot be studied (e. g. due to resolution) or the
996 link is considered weak. Numbered references for conducted studies are provided in the
997 appendix (Table A2).

998 **Table A1:** Information on the datasets used for figure 10: versions, DOIs and references. The
999 correlations between the SST Niño3.4 region (averaged 5°S to 5°N, 190°E to 240°E) time
1000 series and the other ECVs's Niño3.4 time series (and for SM, BA and AOD time series with
1001 Indonesia (averaged 10°S to 10°N, 100°E to 150°E) are given in the right column.

1002 **Table A2:** Snapshot of publications or technical reports (available from ESA CCI program)
1003 until the submission of this manuscript behind entries on done consistency studies in Table
1004 3.

1007 **Table 1:** Links between ECVs on the retrieval (above the diagonal) and scientific (below the
 1008 diagonal) level which need to be consistent if used together. Weak linkages are indicated in
 1009 brackets. Cycles are indicated with the following acronyms: C=carbon cycle, W=water cycle,
 1010 E=energy cycle. Processes are indicated with the following acronyms: r=radiation interaction,
 1011 d=deposition, e=emission / evaporation, t=transport, c=chemical transformation,
 1012 mtf=melting / thawing / freezing, i=ecosystem interaction, a=air land/sea fluxes of carbon
 1013 and water, m=mask.

ESA CCI ECVs	Aerosol	Clouds	GHGs	Ozone	Water vapour	Fire	Ice-Sheets	Land cover	Soil moisture	Glaciers	HR land cover	LST	Permafrost	Snow	Biomass	Lakes	Ocean Colour	Sea Ice	Sea Level	SST	Sea State	Sea surface salinity
	Retrieval dependencies																					
Aerosol		x	x	(x)	x	x	x	x				x		x		x	x			x		
Clouds	Wr		x	x	x	x	x	x		x	x	x		x		x	x	x		x		
GHGs	e				x									(x)						(x)		
Ozone		t	c		x									(x)		x	x			(x)		
Water vapour	EW	E	C	c		(x)	x					x		(x)		x	x		x	x		
Fire	CE		Ce	ce				x			x		(x)			x						
Ice-Sheets	d			r	W	d		x	x	x									x			
Land cover	de		Ce			Cie t			x	x	x	x	x	x		(x)						
Soil moisture	e	E	e		We d	i		i		x	x	x		x	x	x	(x)	(x)	(x)	(x)	(x)	(x)
Glaciers	d					d	W	r			x		x	x		x		x				
HR land cover			Ce			Ct		i	m		x		x									
LST	Er	Er		r	EW r	ECe	Wr	r	Wr	m	r		x	x		x		x		x		
Permafrost		Er	Ce		We	Er	m	Er	Er	m	Er	EW r		x		(x)			(x)			
Snow	d	r		r	We	d	W	ri	mtf	Er m	ri	Wt mtf	Er m		(x)	x			(x)			
Biomass			C			Cc		ic	i			C		i								
Lakes	de				W	d	Wt	ti	W	E mtf	t	EW r	WE e	W					x	x		
Ocean colour	de		C	r		d							Cd	m		t		x		x	x	
Sea Ice				r		d						Wr	m				i		x	x	x	(x)
Sea Level					W		W		W	W			W	W		W		W		(x)	x	
SST	Er	Er	r	r	Er	E	mtf					EW t					Er	m	E		(x)	x
Sea State																	i		m			x
Sea surface salinity			C		ea		mtf			mtf				mtf	mtf		CW i	W mtf	WE	Wa	a	

1015 **Table 2:** Summary of assessment methods for consistency **on different levels and types**

	Consistency type	Required background knowledge	Assessment method
Retrieval level			
	Categorical auxiliary data ("masks")	Incompatible mask classes	Visual: combined images Contingency matrix Class combination maps
	Continuous auxiliary data	Target variable sensitivity to auxiliary variable	Visual: homogeneity Difference maps Statistical comparison
Scientific level			
	Self-consistency (single quantity)	Behaviour of one quantity Known record features Known map features Physical equation	Visual: features as expected Quantitative variability Trend analysis
	Mutual consistency (multiple quantities)	Linkage between quantities Physical model Understood Earth system phenomena	Difference maps Trend comparisons Correlations and other measures of co-variability

1016

1017 **Table 3:** Consistency analysis status between pairs of CCI ECVs: intrinsically assured (*), study
 1018 needed (X), study done (c = comprehensive, e = exemplary, t = theoretical) - empty fields
 1019 indicate that no study is needed, this link cannot be studied (e. g. due to resolution) or the
 1020 link is considered weak. Numbered references for conducted studies are provided in the
 1021 appendix (Table A2).

ESA CCI ECVs	Aerosol	Clouds	GHGs	Ozone	Water vapour	Fire	Ice-Sheets	Land cover	Soil moisture	Glaciers	HR land cover	LST	Permafrost	Snow	Biomass	Lakes	Ocean Colour	Sea Ice	Sea Level	SST	Sea State	Sea surface salinity
	Done / needed consistency analysis (retrieval or scientific level)																					
Aerosol	e ¹⁰	*			X	X	X	X	e ²⁶			X		X		X	t ¹¹			X		
Clouds		e ¹⁰	*			X	X	X	X	X	X	X	X	X		e ²⁴	X	X		X		
GHGs			e ¹⁹	*				X	X		X		t ¹		X		X			X		
Ozone				e ¹⁹	*	X	X					X		X		X	X			X		
Water vapour					e ¹⁹	X	X		X			X	t ²	X		X	X		X	X		
Fire						e ¹⁴	e ²³	e ²³	X	X	t ^{16,17}	t ³	e ¹³	e ²⁷	e ¹⁵	X	e ²⁸		e ¹⁸			
Ice-Sheets							e ¹⁴	e ²³	e ²³	X	X	t ^{16,17}	t ³	e ¹³	e ²⁷	e ¹⁵	X	e ²⁸		e ¹⁸		
Land cover								e ²³	e ²³	X	X	t ^{16,17}	t ³	e ¹³	e ²⁷	e ¹⁵	X	e ²⁸		e ¹⁸		
Soil moisture									e ²³	X	X	t ^{16,17}	t ³	e ¹³	e ²⁷	e ¹⁵	X	e ²⁸		e ¹⁸		
Glaciers										X	X	t ^{16,17}	t ³	e ¹³	e ²⁷	e ¹⁵	X	e ²⁸		e ¹⁸		
HR land cover											X	X	t ^{16,17}	t ³	e ¹³	e ²⁷	e ¹⁵	X	e ²⁸		e ¹⁸	
LST												X	t ^{7,8}	X		c ²⁵						
Permafrost													e ⁴	X	X	X		X	X	X		
Snow														e ⁴	X	X		X	X	X		
Biomass															e ⁴	X	X		X	X		
Lakes																e ⁴	X	X				
Ocean colour																	e ⁴	X	X			
Sea Ice																		e ⁴	X	X		
Sea Level																			e ⁴	X	X	
SST																				e ⁴	X	X
Sea State																					e ⁴	X
SSS																						e ⁴

Technical level of consistency achieved
 Common data format
 Common data access portal
 Common metadata standards
 Common documentation standards
 Common visualisation & analysis tools
 (partly) also adopted by C3s Climate Data Store

1022

1023 **Table A1:** Information on the datasets used for figure 10: versions, DOIs and references. The
 1024 correlations between the SST Niño3.4 region (averaged 5°S to 5°N, 190°E to 240°E) time
 1025 series and the other ECVs's Niño3.4 time series (and for SM, BA and AOD time series with
 1026 Indonesia (averaged 10°S to 10°N, 100°E to 150°E) are given in the right column.

ECV	Dataset version, time period used, DOI, references:	Correlation of Niño3.4 SST with
SST	Sea surface temperature ESA SST CCI ATSR and/or AVHRR product version v2.1, 1982-2016 DOI: n/a Merchant et al., 2019	Niño3.4 SST: 1.00
SL	Sea level height SL_cci data v2.0 1993-2015 DOI: 10.5270/esa-sea_level_cci-1993_2015-v_2.0-201612 Legeais et al., 2018 and Quartly et al., 2017	Niño3.4 SL: 0.87
SSS	Sea surface salinity SEASURFACESALINITY_CCI_DATA v1.8 2010-2018 DOI: 10.5285/9ef0ebf847564c2eabe62cac4899ec41 Boutin et al., 2019	Niño3.4 SSS: -0.63
Chlor_a	Chlorophyll-alpha CCI Chlor_a v3.1 (4km_GEO_PML)	Niño3.4Chlor_a: -0.68

	<p>1998-2017</p> <p>DOI: n/a</p> <p>Sathyendranath et al., 2012</p>	
CFChigh	<p>High level cloud fraction</p> <p>Cloud_cci AVHRR-PMv3</p> <p>1982-2016</p> <p>DOI: n/a</p> <p>Stengel et al., 2019</p>	Niño3.4 CFChigh: 0.82
TCWV	<p>Total column water vapour</p> <p>HOAPS 4</p> <p>1988-2015</p> <p>DOI:10.5676/EUM_SAF_CM/HOAPS/V002</p> <p>Andersson et al., 2017, data from 2015 as beta version of HOAPS 4</p>	Niño3.4 TCWV: 0.84
AOD550	<p>Aerosol optical depth at 550 nm</p> <p>CCI ATSR-2/AATSR Swansea v4.1</p> <p>1997-2011</p> <p>https://esgf-node.llnl.gov/search/obs4mips/obs4mips.SU.ATSR2-AATSR.od550aer.mon.v20160922 eridanus.eoc.dlr.de</p> <p>Bevan, S., et al., 2012; North, P., et al., 1999; Popp, et al., 2016</p>	Indonesia AOD550: 0.52
Fire	<p>Burned area</p>	Indonesia Fire: 0.49

	<p>FireCCI51</p> <p>2001-2017</p> <p>DOI:</p> <p>dx.doi.org/10.5285/3628cb2fdb443588155e15dee8e5352</p> <p>Lizundia et al., 2020</p>	
SM	<p>Soil moisture</p> <p>ESA CCI SM merged v04.5</p> <p>1991-2018</p> <p>DOI: n/a</p> <p>Dorigo et al., 2017, Gruber et al., 2019</p>	Indonesia SM: -0.57

1027

1028 **Table A2:** Snapshot of publications or technical reports (available from ESA CCI program)
1029 until the submission of this manuscript behind entries on done consistency studies in Table
1030 3.

- 1031 1. Chadburn, S. E., Krinner, G., Porada, P., Bartsch, A., Beer, C., Belelli Marchesini, L.,
1032 Boike, J., Ekici, A., Elberling, B., Friborg, T., Hugelius, G., Johansson, M., Kuhry, P.,
1033 Kutzbach, L., Langer, M., Lund, M., Parmentier, F.-J. W., Peng, S., Van Huissteden, K.,
1034 Wang, T., Westermann, S., Zhu, D., and Burke, E. J. 2017: Carbon stocks and fluxes in
1035 the high latitudes: using site-level data to evaluate Earth system models.
1036 *Biogeosciences*, **14**, 5143 - 5169, <https://doi.org/10.5194/bg-14-5143-2017>.
- 1037 2. Reuter, M., M. Buchwitz, M. Hilker, J. Heymann, H. Bovensmann, J.P. Burrows, S.
1038 Houweling, Y.Y. Liu, R. Nassar, F. Chevallier, P. Ciais, J. Marshall, and M. Reichstein,
1039 2017: [How Much CO₂ Is Taken Up by the European Terrestrial Biosphere?](#). *Bull. Amer.*
1040 *Meteor. Soc.*, **98**, 665 – 671, <https://doi.org/10.1175/BAMS-D-15-00310.1>
- 1041 3. Carolyn M. Gibson, Laura E. Chasmer, Dan K. Thompson, William L. uinton, Mike D.
1042 Flannigan & [...]David Olefeldt, 2018: Wildfire as a major driver of recent
1043 permafrost thaw in boreal peatlands. *Nature Communications*, **9**, 3041,
1044 <https://doi.org/10.1038/41467-018-05457-1>.
- 1045 4. Westermann, S., Peter, M., Langer, M., Schwamborn, G., Schirrmeister, L.,
1046 Etzelmüller, B., Boike, J., 2017: Transient modeling of the ground thermal conditions
1047 using satellite data in the Lena River Delta, Siberia. *The Cryosphere*, **11**, 1441 - 1463,
1048 [doi:10.5194/tc-11-1441-2017](https://doi.org/10.5194/tc-11-1441-2017).
- 1049 5. Dafflon, B., Oktem, R., Peterson, J., Ulrich, C., Tran, A. P., Romanovsky, V., & Hubbard,
1050 S. S., 2017: Coincident aboveground and belowground autonomous monitoring to
1051 quantify covariability in permafrost, soil, and vegetation properties in Arctic tundra.

- 1052 *Journal of Geophysical Research: Biogeosciences*, **122**, 1321 – 1342,
1053 <https://doi.org/10.1002/2016JG003724>.
- 1054 6. Daiyrov M., C. Narama, T. Yamanokuchi, T. Tadono, A. Kääb, J. Ukita T., 2018:
1055 Regional geomorphological conditions related to recent changes of glacial lakes in the
1056 Issyk-Kul basin, northern Tien Shan. *Geosciences*, **8**, 99.
- 1057 7. Cable, W.L.; Romanovsky, V.E.; Jorgenson, M.T., 2016: Scaling-up permafrost thermal
1058 measurements in western Alaska using an ecotype approach. *Cryosphere*, **10**, 2517 –
1059 2532.
- 1060 8. Zhang, N., Yasunari, T., & Ohta, T., 2011: Dynamics of the larch taiga–permafrost
1061 coupled system in siberia under climate change. *Environmental Research Letters*, **6**,
1062 024–003, <https://doi.org/10.1088/1748-9326/6/2/024003>.
- 1063 9. Nitze, I., G. Grosse, B. M. Jones, V. E. Romanovsky and J. Boike, 2018: Remote sensing
1064 quantifies widespread abundance of permafrost region disturbances across the Arctic
1065 and Subarctic. *Nature Communications*, **9**, 5423.
- 1066 10. Klüser L., S. Stapelberg, Aerosol_cci Cloud_cci cloud mask consistency report v1.1.
1067 DLR / DWD / ESA (briefly summarized in the third example of section 4 in this paper).
- 1068 11. Stebel, Kerstin, et al., 2017: Aerosol_cci2 Technical Note on consistency v1.0. DLR /
1069 ESA.
- 1070 12. Adolf, C., Wunderle, S., Colombaroli, D., Weber, H., Gobet, E., Heiri, O., van Leeuwen,
1071 J. F. N., Bigler, C., Connor, S. E., Gałka, M., La Mantia, T., Makhortykh, S., Svitavská-
1072 Svobodová, H., Vannière, B., and Tinner, W., 2018: The sedimentary and remote-
1073 sensing reflection of biomass burning in Europe. *Global Ecology and Biogeography*
1074 **27**, 199–212. doi: 10.1111/geb.12682.
- 1075 13. Cape, J., Coyle, M., and Dumitrean, P., 2012: The atmospheric lifetime of black
1076 carbon. *Atmos. Environ.*, **59**, 256 – 263, doi: 10.1016/j.atmosenv. 2012.05.030.

- 1077 14. Eichler, A., Tinner, W., Brüttsch, S., Olivier, S., Papina, T., and Schwikowski, M., 2011:
1078 An ice-core based history of Siberian forest fires since AD 1250. *Quaternary Science*
1079 *Reviews*, **30**, 1027 – 1034. doi: 10.1016/j.quascirev.2011.
- 1080 15. Marlon, J. R., Kelly, R., Daniau, A.-L., Vannièrè, B., Power, M. J., Bartlein, P. J., Higuera,
1081 P. E., Blarquez, O., Brewer, S., Brücher, T., Feurdean, A., Romera, G. G., Iglesias, V.,
1082 Maezumi, S. Y., Magi, B., Courtney Mustaphi, C. J., and Zhihai, T., 2016:
1083 Reconstructions of biomass burning from sediment-charcoal records to improve
1084 data–model comparisons. *Biogeosciences*, **13**, 3225 – 3244, doi: 10.5194/bg-13-3225-
1085 2016.
- 1086 16. Pechony, O. and Shindell, D. T., 2010: Driving forces of global wildfires over the past
1087 millennium and the forthcoming century. *Proc. of the National Academy of Sciences*
1088 **107.45**, 19167–19170, doi: 10.1073/pnas.1003669107.
- 1089 17. Chuvieco, E., Mouillot, F., van der Werf, G.R., San Miguel, J., Tanasse, M., Koutsias, N.,
1090 García, M., Yebra, M., Padilla, M., Gitas, I., Heil, A., Hawbaker, T.J., & Giglio, L., 2019:
1091 Historical background and current developments for mapping burned area from
1092 satellite Earth observation. *Rem. Sens. Environ.*, **225**, 45 - 64.
- 1093 18. Chen, Y., Morton, D.C., Andela, N., Giglio, L., & Randerson, J.T., 2016: How much
1094 global burned area can be forecast on seasonal time scales using sea surface
1095 temperatures? *Environmental Research Letters*, **11**, 045001.
- 1096 19. Heymann, J., M. Reuter, M. Buchwitz, O. Schneising, H. Bovensmann, J. P. Burrows, S.
1097 Massart, J. W. Kaiser, D. Crisp, 2017: CO2 emission of Indonesian fires in 2015
1098 estimated from satellite-derived atmospheric CO2 concentrations, *Geophys. Res.*
1099 *Lett.*, **44**, 1537 – 1544, DOI: 10.1002/2016GL072042, pp. 18, 2017.
- 1100 20. Cazenave, A., 2018: Global sea-level budget 1993-present. *Earth System Science Data*,
1101 **10**, 1551 - 1590, <https://doi.org/10.5194/essd-10-1551-2018>.

- 1102 21. Rastner, P., T. Bolch, N. Mölg, H. Machguth, R. Le Bris and F. Paul, 2012: The first
1103 complete inventory of the local glaciers and ice caps on Greenland. *The Cryosphere*,
1104 **6**, 1483 - 1495.
- 1105 22. Huber, J., A. Cook, F. Paul, and M. Zemp, 2017: A complete glacier inventory of the
1106 Antarctic Peninsula based on Landsat7 images from 2000-2002 and other pre-existing
1107 datasets. *Earth Systems Science Data*, **9**, 115 - 131.
- 1108 23. Forkel, M., Dorigo, W., Lasslop, G., Teubner, I., Chuvieco, E., and Thonicke, K. 2017: A
1109 data-driven approach to identify controls on global fire activity from satellite and
1110 climate observations (SOFIA V1). *Geosci. Model Dev.*, **10**, 4443 - 4476,
1111 <https://doi.org/10.5194/gmd-10-4443-2017>.
- 1112 24. Brockmann, C., et al., 2013: Multi-Sensor Cloud Screening and Validation: IdePix and
1113 PixBox. *Proc. of the 2013 European Space Agency Living Planet Symposium*.
1114 <http://livingplanet2013.org/abstracts/850821>.
- 1115 25. Pekel, J.-F., et al., 2016: High-resolution mapping of global surface water and its long-
1116 term changes. *Nature* **540**, 418.
- 1117 26. Klingmüller, K., Pozzer, A., Metzger, S., Stenchikov, G. L., and Lelieveld, J., 2016:
1118 Aerosol optical depth trend over the Middle East, *Atmos. Chem. Phys.*, **16**, 5063 -
1119 5073, <https://doi.org/10.5194/acp-16-5063-2016>.
- 1120 27. Frohling, S., Palace, M.W., Clark, D., Chambers, J.Q., Shugart, H., & Hurtt, G.C., 2009:
1121 Forest disturbance and recovery: A general review in the context of spaceborne
1122 remote sensing of impacts on aboveground biomass and canopy structure. *J.*
1123 *Geophys. Res. Biogeosciences*, **114**.
- 1124 28. Law, K.S., & Stohl, A., 2007: Arctic air pollution: Origins and impacts. *Science*, **315**,
1125 1537 - 1540.

1126 **Figure captions**

1127 **Figure 1:** Temporal coverage of CDRs for ECVs analysed by CCI. Filled bars indicate CDRs
1128 available in 2019, outlined bars CDRs that are planned within the ongoing phase of the CCI
1129 program.

1130 **Figure 2:** The ECVs covered by ESA CCI CDRs, ordered according to the key Earth system cycle
1131 (energy, carbon, water) they help characterise. The cycles are inter-linked, and most water
1132 and carbon cycle ECVs are also relevant to the energy cycle, since energy is stored and
1133 transported in water and matter, at least on transient timescales.

1134 **Figure 3:** Gaps in surface temperature fields (LST and SST from SLSTR on Sentinel-3A on
1135 05/08/2018 at 10:38 UTC) due to masked clouds (grey), showing the absence of scatter at
1136 land-sea borders and sampling discontinuities across some land-sea boundaries due to
1137 different cloud-clearing approaches between LST and SST processing.

1138 **Figure 4:** Consistency overview between Aerosol_cci (Swansea University) and Cloud_cci
1139 (FAME-C) AATSR cloud masks for observations of four selected days in September 2008. No
1140 cloud/no cloud and cloud/cloud situations are solely analysed as aerosol or clouds in
1141 Aerosol_cci and Cloud_cci, respectively. No cloud/cloud situations are wrongly analysed as
1142 aerosols and clouds, while cloud/no cloud situations are not analysed at all.

1143 **Figure 5:** Mean AOD differences at 865 nm between ocean colour MERIS atmospheric
1144 correction by-product and aerosol ECV product from AATSR in May 2003 when both
1145 instruments retrieve AOD.

1146 **Figure 6: Top:** Time series of monthly mean northern mid-latitude XCO₂ (red thick line) based
1147 on merging individual XCO₂ ensemble members (black lines) from GOSAT (since 2009) and
1148 OCO-2 (since 2014). The time series (2003-2018) begins with one XCO₂ product from
1149 SCIAMACHY/ENVISAT. **Bottom:** XCO₂ difference between ensemble members (black lines)

1150 and the multi-sensor / multi-algorithm merged product (red line in top panel). Details see
1151 Reuter et al., 2020.

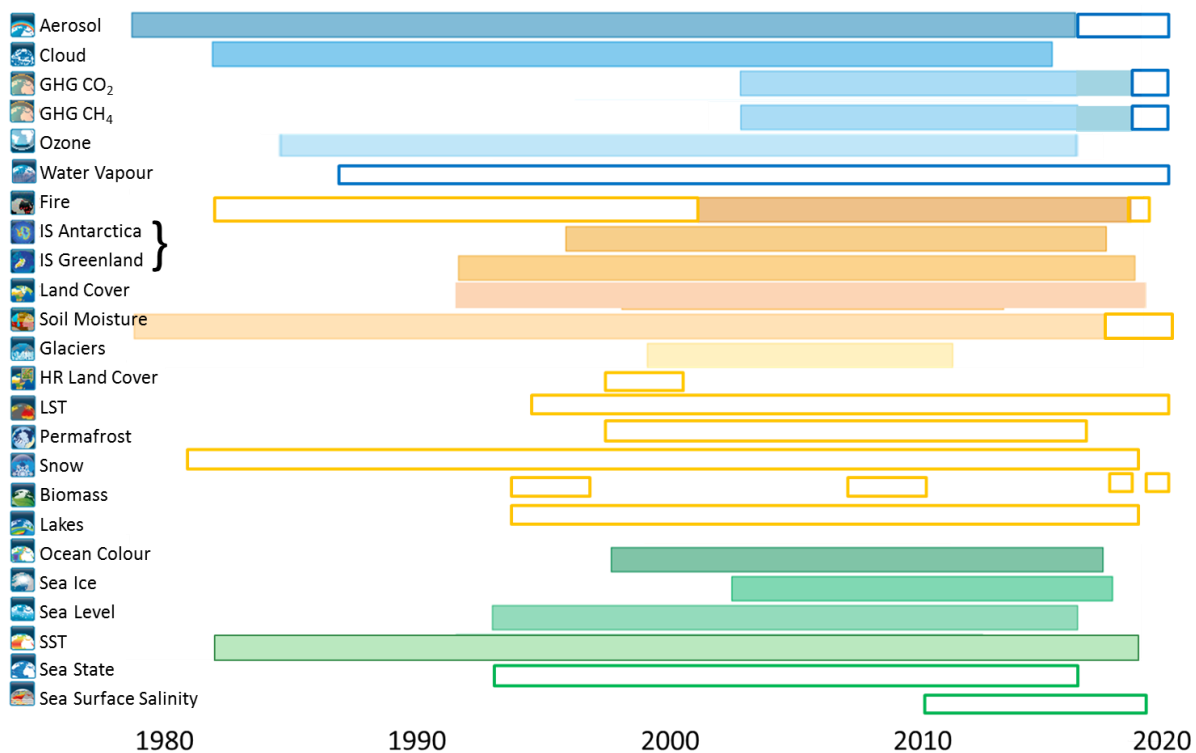
1152 **Figure 7:** Trend estimates computed after (green) and before (black) homogenisation for all
1153 long-term TCWV data records available from the G-VAP data archive (Schröder et al., 2018).
1154 Trend estimates are sorted in ascending order without homogenisation. The grey horizontal
1155 line marks a trend of 0 kg/m²/year (updated from Schröder et al., 2019).

1156 **Figure 8:** The left panel shows the co-variation between a prototype version of the
1157 stratospheric water vapour CDR H₂O (produced within the Water_Vapour_cci) and ERA5
1158 monthly zonal mean temperatures T at 100 hPa. The right panel shows the correlation
1159 between the two datasets.

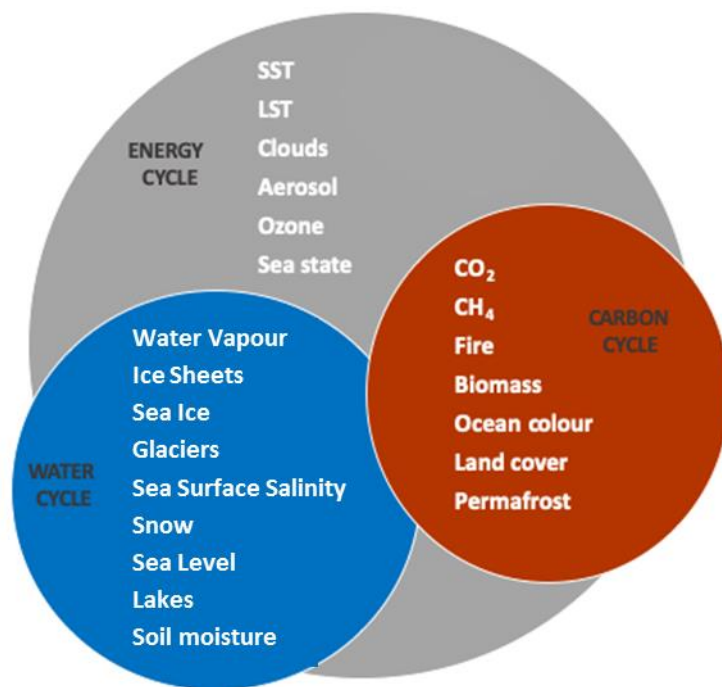
1160 **Figure 9:** Trends of monthly averaged significant wave height H_s data sets with the Mann–
1161 Kendall test (thatched areas) from satellite altimetry (left: ALT), and co-located model WW3
1162 hindcast (right: CoLoc) both given in cm year⁻¹.

1163 **Figure 10:** Zonal month-longitude cross sections (averaged 5°S and 5°N) for 150°E to 280°E
1164 normalized indices of a) sea surface temperature (SST), b) sea level height (SL), c) Sea Surface
1165 Salinity (SSS), d) chlorophyll-alpha (Chlor_a), e) high level cloud fraction (CFChigh), f) total
1166 column water vapour (TCWV). All ECVs are plotted for their respective full year availability.
1167 The black lines in the Hovmöller plots show the Niño3.4 box. g) Time series of Niño3.4 SST
1168 and Indonesia soil moisture (SM), burned area (Fire), and aerosol optical depth at 550 nm
1169 (AOD550). Information on the used datasets is provided in Table A1 in the Appendix.

1170 **Figures**

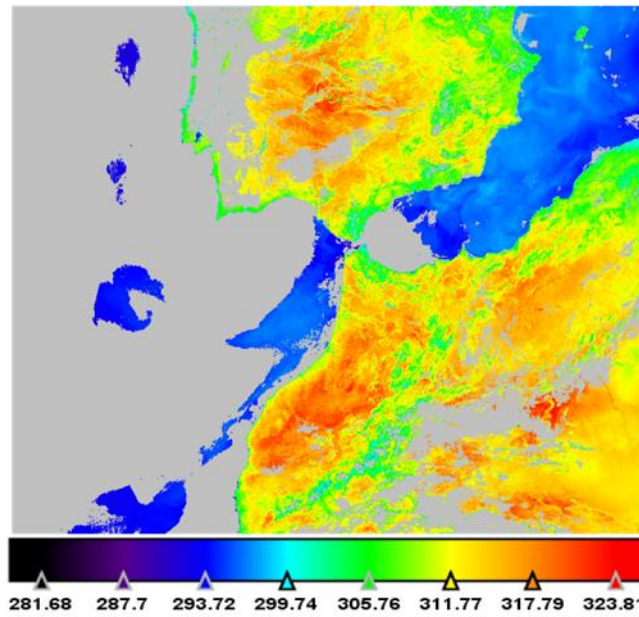


1171
 1172 **Figure 1:** Temporal coverage of CDRs for ECVs analysed by CCI. Filled bars indicate CDRs
 1173 available in 2019, outlined bars CDRs that are planned within the ongoing phase of the CCI
 1174 program.



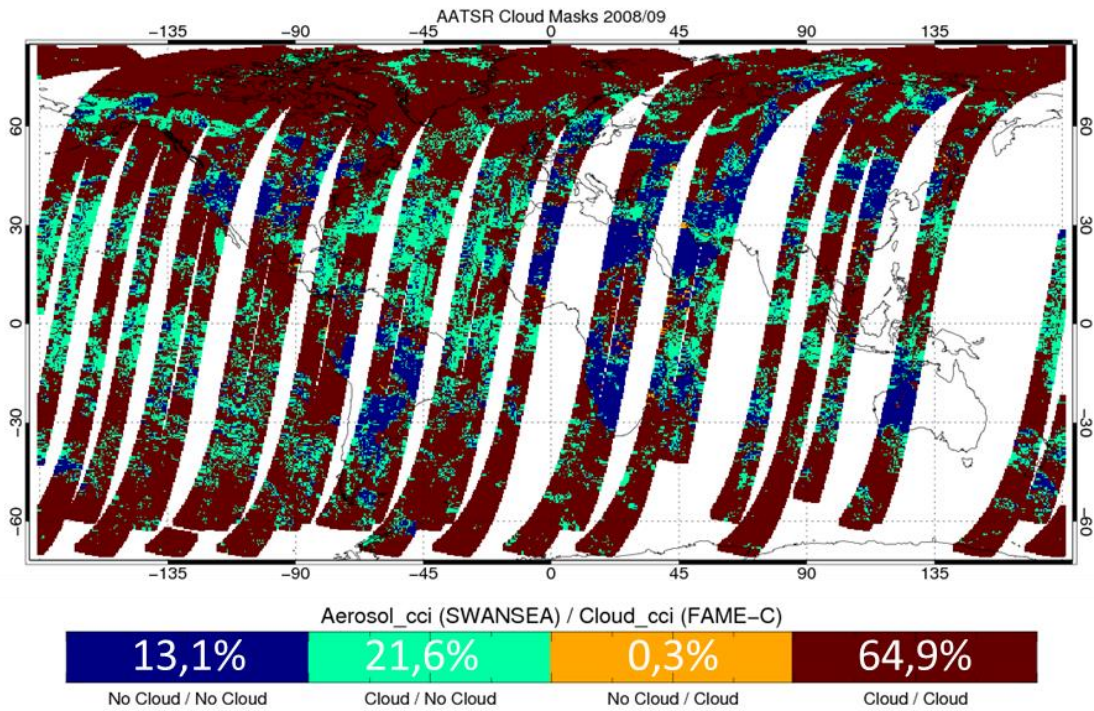
1175

1176 **Figure 2:** The ECVs covered by ESA CCI CDRs, ordered according to the key Earth system cycle
 1177 (energy, carbon, water) they help characterise. The cycles are inter-linked, and most water
 1178 and carbon cycle ECVs are also relevant to the energy cycle, since energy is stored and
 1179 transported in water and matter, at least on transient timescales.



1180

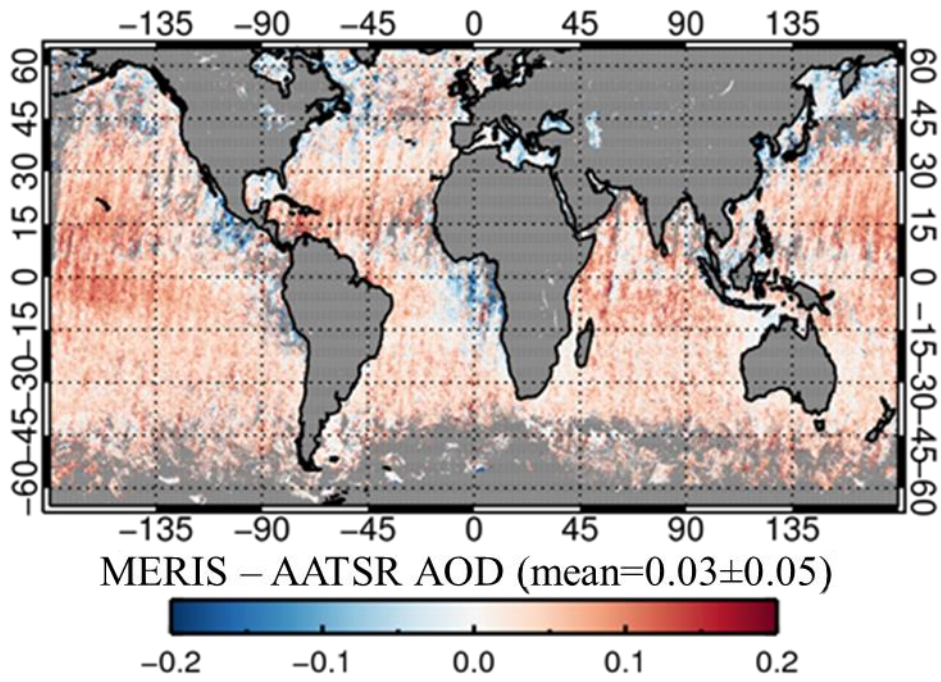
1181 **Figure 3:** Gaps in surface temperature fields (LST and SST from SLSTR on Sentinel-3A on
 1182 05/08/2018 at 10:38 UTC) due to masked clouds (grey), showing the absence of scatter at
 1183 land-sea borders and sampling discontinuities across some land-sea boundaries due to
 1184 different cloud-clearing approaches between LST and SST processing.



1185

1186 **Figure 4:** Consistency overview between Aerosol_cci (Swansea University) and Cloud_cci
 1187 (FAME-C) AATSR cloud masks for observations of four selected days in September 2008. No
 1188 cloud/no cloud and cloud/cloud situations are solely analysed as aerosol or clouds in
 1189 Aerosol_cci and Cloud_cci, respectively. No cloud/cloud situations are wrongly analysed as
 1190 aerosols and clouds, while cloud/no cloud situations are not analysed at all.

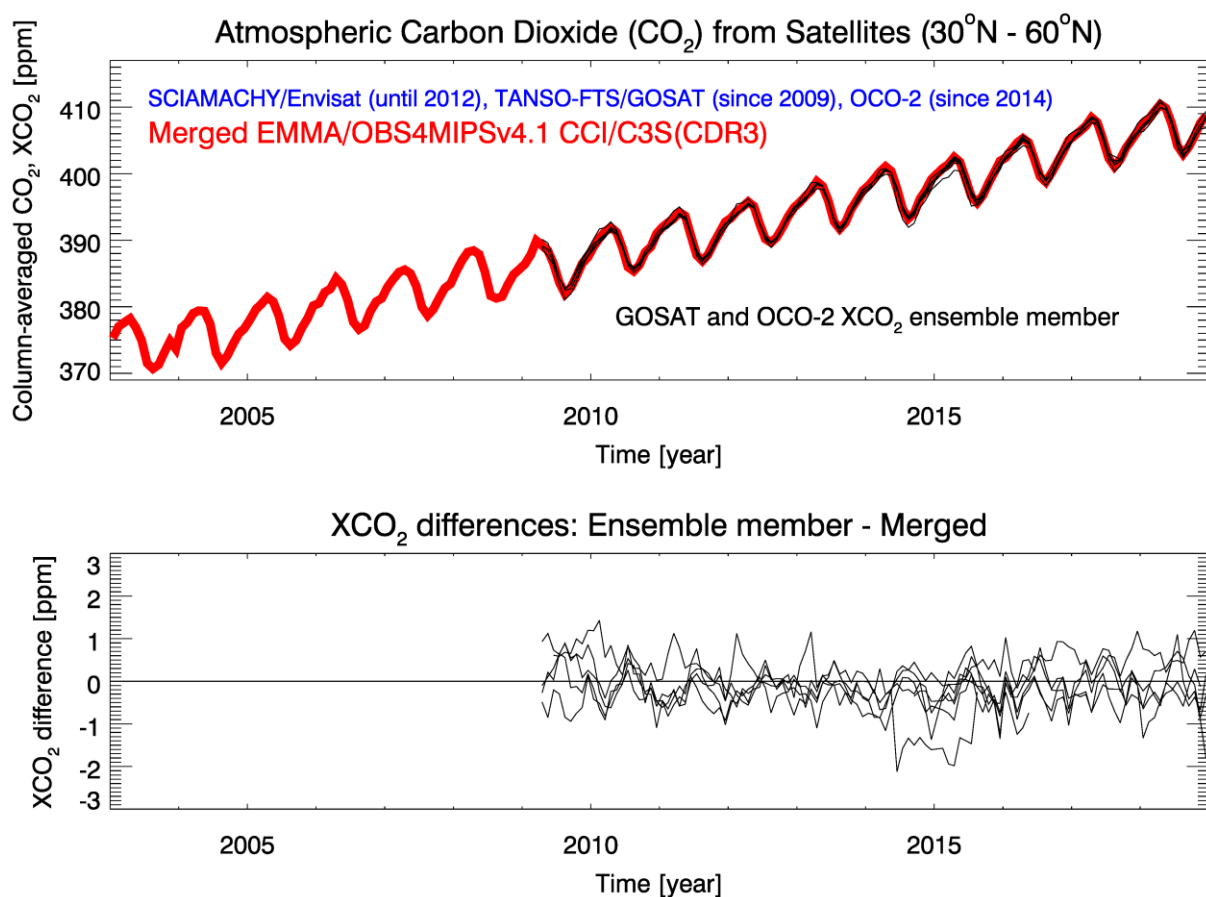
1191



1192

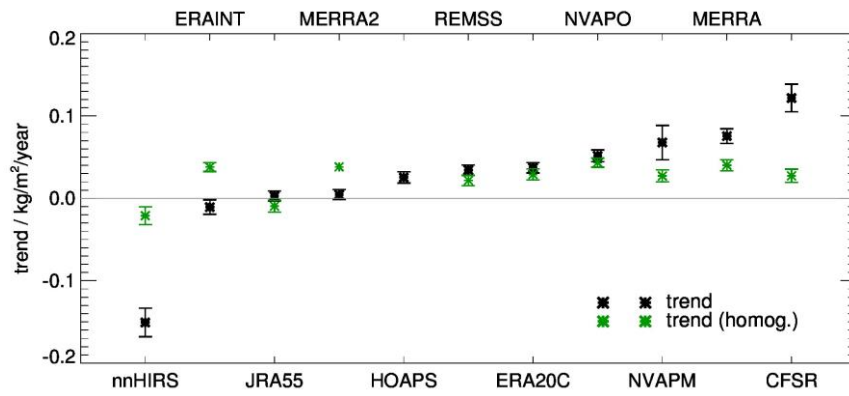
1193 **Figure 5:** Mean AOD differences at 865 nm between ocean colour MERIS atmospheric
 1194 correction by-product and aerosol ECV product from AATSR in May 2003 when both
 1195 instruments retrieve AOD.

1196



1197

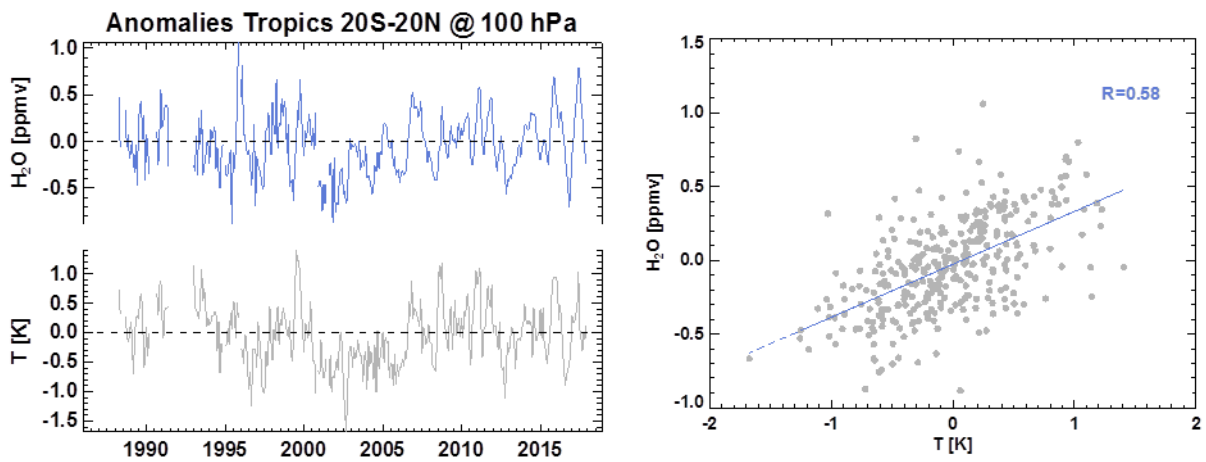
1198 **Figure 6: Top:** Time series of monthly mean northern mid-latitude XCO₂ (red thick line) based
 1199 on merging individual XCO₂ ensemble members (black lines) from GOSAT (since 2009) and
 1200 OCO-2 (since 2014). The time series (2003-2018) begins with one XCO₂ product from
 1201 SCIAMACHY/ENVISAT. **Bottom:** XCO₂ difference between ensemble members (black lines)
 1202 and the multi-sensor / multi-algorithm merged product (red line in top panel). Details see
 1203 Reuter et al., 2020.



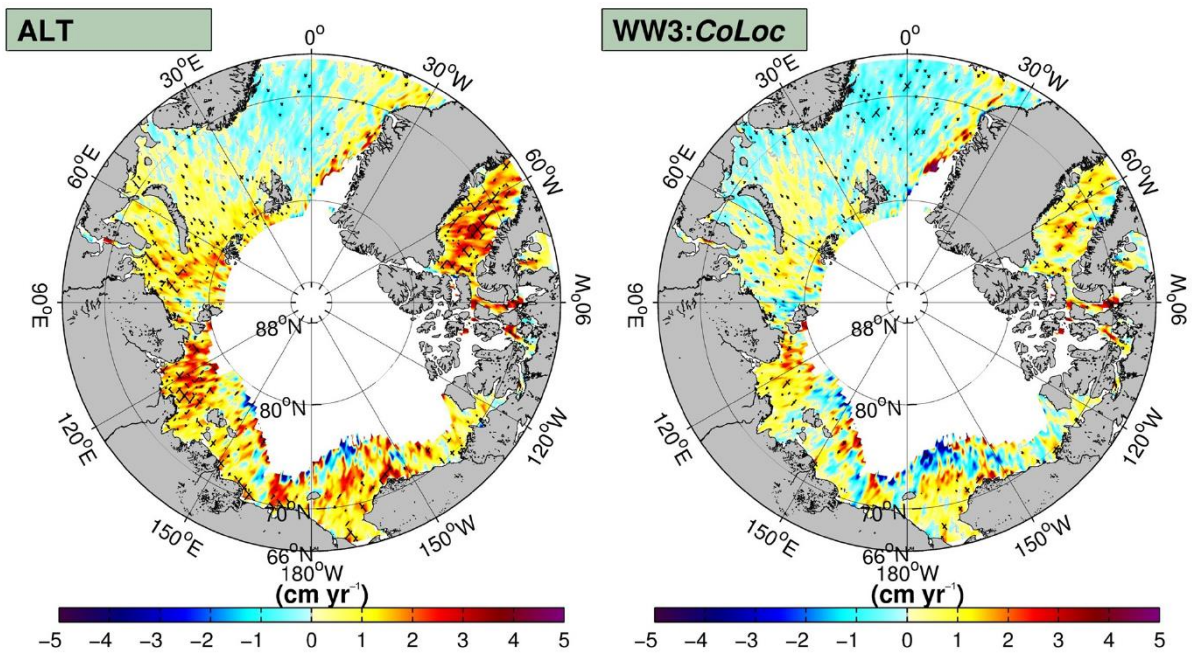
1204

1205 **Figure 7:** Trend estimates computed after (green) and before (black) homogenisation for all
 1206 long-term TCWV data records available from the G-VAP data archive (Schröder et al., 2018).
 1207 Trend estimates are sorted in ascending order without homogenisation. The grey horizontal
 1208 line marks a trend of 0 kg/m²/year (updated from Schröder et al., 2019).

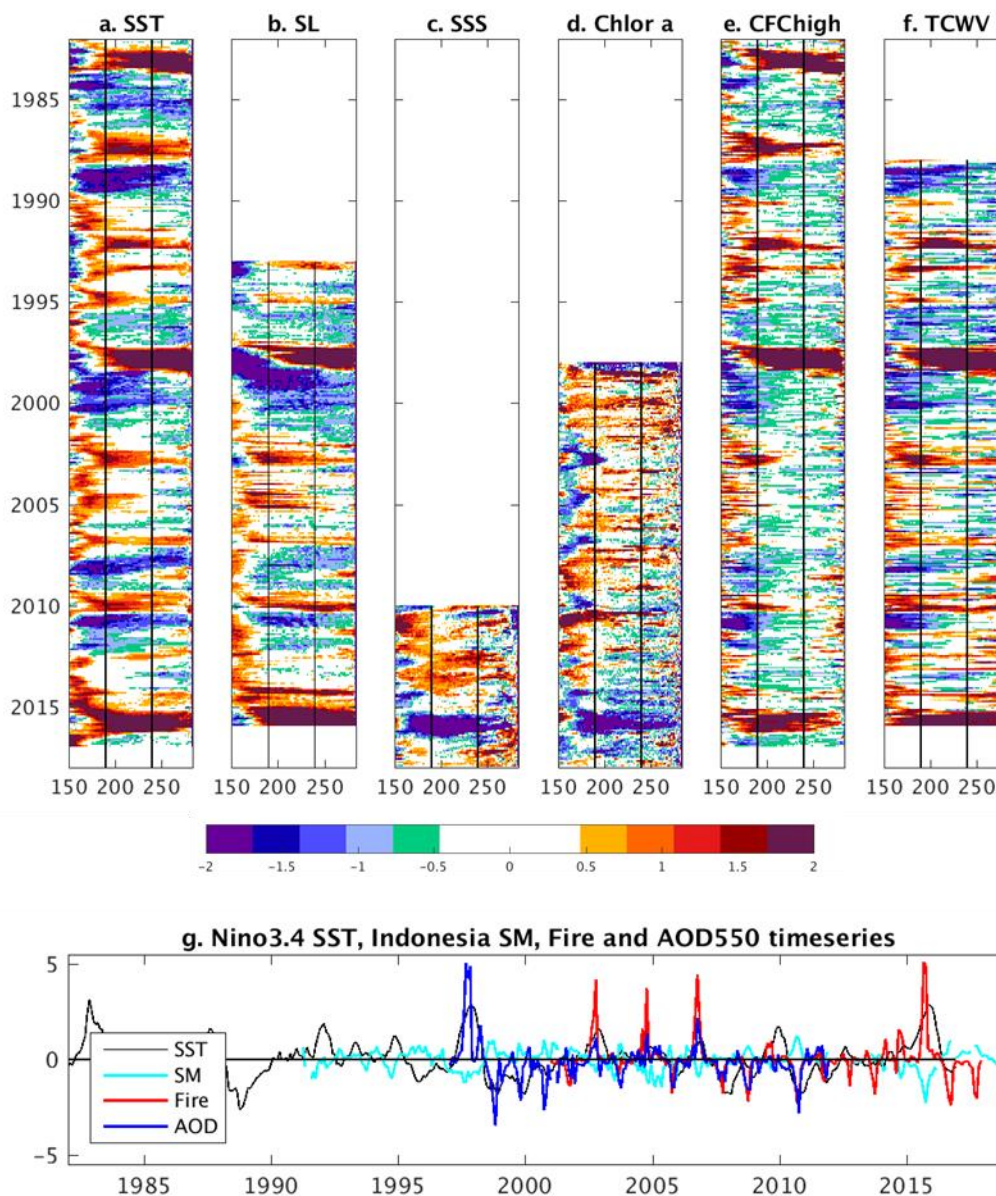
1209



1210
 1211 **Figure 8:** The left panel shows the co-variation between a prototype version of the
 1212 stratospheric water vapour CDR H₂O (produced within the Water_Vapour_cci) and ERA5
 1213 monthly zonal mean temperatures T at 100 hPa. The right panel shows the correlation
 1214 between the two datasets.



1216 **Figure 9:** Trends of monthly averaged significant wave height H_s data sets with the Mann–
 1217 Kendall test (thatched areas) from satellite altimetry (left: ALT), and co-located model WW3
 1218 hindcast (right: CoLoc) both given in cm yr^{-1} .



1219
 1220 **Figure 10:** Zonal month-longitude cross sections (averaged 5°S and 5°N) for 150°E to 280°E
 1221 normalized indices of a) sea surface temperature (SST), b) sea level height (SL), c) Sea Surface
 1222 Salinity (SSS, d) chlorophyll-alpha (Chlor_a), e) high level cloud fraction (CFChigh), f) total
 1223 column water vapour (TCWV). All ECVs are plotted for their respective full year availability.
 1224 The black lines in the Hovmöller plots show the Niño3.4 box. g) Time series of Niño3.4 SST
 1225 and Indonesia soil moisture (SM), burned area (Fire), and aerosol optical depth at 550 nm
 1226 (AOD550). Information on the used datasets is provided in Table A1 in the Appendix.
 1227