



**open data
portal**
cci

DEL 17 Annual Report on Archive and Metadata Catalogue Operations

Prepared by : Alison Waterfall (STFC-CEDA)
Victoria Bennett (STFC-CEDA)

Approved by : Victoria Bennett (STFC-CEDA)

Customer:	ESA	Author:	STFC-CEDA
Project Name:	CCI Open Data Portal	Contract No.:	4000113814/15/I-NB
Project Ref:	Task 3	Document Ref:	CDAP.REP.024
Issue:	Version 2.2	Issue Date:	7/7/17



Science & Technology
Facilities Council



AMENDMENT RECORD SHEET

The Amendment Record Sheet below records the history and issue status of this document.

ISSUE	DATE	REASON
0.1	11 May 2016	First draft issue for review
1.0	19 May 2016	First formal issue
1.1	16 June 2016	Resolved RIDS
2.0	28 April 2017	Annual report for year 2 issued
2.1	3 rd May 2017	Updated download stats
2.2	7 th July 2017	Resolved RIDS: RID_D.17_v2.1_001 RID_D.17_v2.1_002 RID_D.17_v2.1_003 RID_D.17_v2.1_004 RID_D.17_v2.1_005 RID_D.17_v2.1_006 RID_D.17_v2.1_007 RID_D.17_v2.1_008 RID_D.17_v2.1_009 RID_D.17_v2.1_010 Two appendices are also included with the report providing detailed download statistics for the WMS and WCS protocols: <ul style="list-style-type: none">• Appendix A - wcs_data_201609_201706• Appendix B - wms_data_201609_201706 Section 2.1 updated to encompass these documents in the scope.

DISTRIBUTION LIST

The Amendment Record Sheet below records the history and issue status of this document.

NAME	AFFILIATION
Ed Pechorro	ESA
Simon Pinnock	ESA
Pascal Lecomte	ESA



NAME	AFFILIATION
Cat Downey	ESA
Richard Lowe	TVUK
Kevin Halsall	TUVK
Anna Corlyon	TVUK
Victoria Bennett	STFC
Phil Kershaw	STFC
Anne Chadwick	STFC
Alison Waterfall	STFC
Debbie Clifford	University of Reading
Jon Blower	University of Reading
Jane Lewis	University of Reading
Clive Farquhar	CGI
Alex Wood	CGI
Katy Hurst	CGI
Carsten Brockmann	Brockmann Consult
Tom Block	Brockmann Consult
Norman Fomferra	Brockmann Consult



TABLE OF CONTENTS

1. EXECUTIVE SUMMARY.....	7
2. INTRODUCTION.....	8
2.1 Purpose and Scope	8
2.2 Structure of the Document	8
2.3 Referenced Documents.....	9
3. CCI OPEN DATA PORTAL : CENTRAL ARCHIVE AND METADATA CATALOGUE.....	10
3.1 Interface and server-side tools.....	10
3.1.1 Download methods	10
3.2 Dataset summary	12
3.2.1 Aerosol:.....	15
3.2.2 Cloud.....	16
3.2.3 Fire	17
3.2.4 GHG	18
3.2.5 Glaciers	20
3.2.6 Ice Sheets – Antarctica	21
3.2.7 Ice Sheets – Greenland.....	21
3.2.8 Land cover	24
3.2.9 Ocean Colour	24
3.2.10 Ozone.....	26
3.2.11 Sea ice.....	28
3.2.12 Sea level.....	28
3.2.13 Soil moisture.....	29
3.2.14 SST	30
3.3 Download statistics	32
3.3.1 Data downloads by method (May 2016- April 2017).....	32
3.3.2 Data downloads by dataset.....	33
3.3.3 Historic Data Download Trends.....	33
4. FUTURE ACTIVITIES IN YEAR 3.....	37

DOCUMENT TABLES

No table of figures entries found.

DOCUMENT FIGURES

None



ABBREVIATIONS AND ACRONYMS

The table below shows all abbreviations and acronyms used within the document

AATSR	Advanced Along-Track Scanning Radiometer
AMSR-E	Advanced Microwave Scanning Radiometer - EOS
ALAMO	Aerosol Load and Altitude from MERIS over Ocean
AVHRR	Advanced Very High Resolution Radiometer
CCI	Climate Change Initiative
CEDA	Centre for Environmental Data Archival
CEMS	Climate and Environmental Monitoring from Space
CF	Climate and Forecasts
CFL	Calving Front Location
DOI	Digital Object Identifier
ECV	Essential Climate Variable
EO	Earth Observation
ESA	European Space Agency
ESGF	Earth System Grid Federation
EU	European Union
FTP	File Transfer Protocol
GHR SST	Group for High Resolution Sea Surface Temperature
GIS	Geo-Information System
GLL	Grounding Line Location
GMPE	GHR SST Multi-Product Ensemble
GOMOS	Global Ozone Monitoring by Occultation of Stars
HTTP	Hyper Text Transfer Protocol
IOP	Inherent Optical Properties
IV	Ice Velocity
LC	Land Cover
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
NetCDF	Network Common Data Form



NOAA	National Oceanic and Atmospheric Administration
Obs4MIPS	Observations for Model Intercomparison Projects
OPeNDAP	Open-source Project for a Network Data Access Protocol
OSIRIS	Optical, Spectroscopic, and Infrared Remote Imaging System
RAL	Rutherford Appleton Laboratory
SCIAMACHY	SCanning Imaging Absorption SpectroMeter for Atmospheric CHartography
SEV	Surface Elevation Change
SMR	Submillimeter wave Radiometer
SSM/I	Special Sensor Microwave Imager
SST	Sea Surface Temperature
STFC	Science and Technology Facilities Council
SU	Swansea University
THREDDS	Thematic Real-time Environmental Distributed Data Services
TDS	THREDDS Data Server
TVUK	Telespazio VEGA UK
WFM	Weighting Function Modified



1. EXECUTIVE SUMMARY

In the second year of operations, the focus has been on a) adding new datasets to the archive and updating versions of existing datasets as required b) deploying new download routes and maximising the amount of data served by all data protocols, c) improvements to the CSW to allow refinements to faceted search and to address security issues. Successful discussions also took place with the GEOSS portal with the aim of also making the CCI data visible through their portal, and this has now been implemented.

The current volume of data on the CCI Open Data Portal archive is 99 TB, with approximately 97 separate datasets displayed in the portal dashboard. In addition to this, older versions of the data often remain archived within the metadata catalogue and archive, and for publishing to ESGF, splitting of some datasets has been required which makes the number significantly higher.

In the second year of operations new datasets have been added from 8 of the project teams, and DOI's have been issued to datasets from 5 of the CCI teams. A further set of datasets have recently been released and are scheduled for addition to the portal in the near future.

Going in to year 3, CEDA will continue to lead on the archive and metadata catalogue operations, whilst some of the operational work will be handed over to TVUK and CGI.



2. INTRODUCTION

2.1 Purpose and Scope

This annual operations report on the CCI Central Data Archive and Metadata Catalogue complements the Annual User Assessment Report (DEL-19) and the Archive and Metadata Catalogue Annual Report Annex on Obs4MIPs (DEL-18).

It includes:

- Data availability summary¹
- Information on Interface and server-side tools, with annual totals for usage.
- Data download volume, with annual total broken down per download method
- Cumulative ECV data quality alerts compiled into separate sections per ECV project

The content focuses on updates to the archive in the second year of operations, and summarises for the year the data download statistics presented in each MPR for all datasets. It also describes the progress on making all datasets available via ESGF and WMS.

Two appendices are included with this document providing detailed statistics on the WMS and WCS download protocols:

- Appendix A - wcs_data_201609_201706
- Appendix B - wms_data_201609_201706

2.2 Structure of the Document

The document is structured to have sections as per the bullet points above. Where appropriate, information has been broken down into sections per ECV.

2 Introduction

Current section

3 CCI Open Data Portal: Central Archive and Metadata Catalogue

This section summarises the availability and download statistics for data products on the CCI Central Archive and Metadata Catalogue at the end of year 2, as well as issues encountered in relation to data quality. The section is structured to have sections as per the bullet points above. Where appropriate, information has been broken down into sections per ECV.

4 Future Activities in Year 3

¹ Note that in the PMP we proposed "Data availability per month" which we have replaced with an overall summary of dataset availability during the year. Tracking the addition and replacement and withdrawal of datasets on a monthly basis is not very practical, nor does it present a meaningful metric



This section summarises the activities related to ongoing operations in the Archive and Metadata Catalogue in Year 3.

2.3 Referenced Documents

RD.1 CCI Open Data Portal DEL-09: Technical Specification Document, CDAP.REP.011, Issue 1.0, 30 October 2015

RD.2 CCI Open Data Portal DEL 15: Tech Specification Annex: Archive and Metadata Catalogue Implementation, CDAP.REP.023, Issue 2.1, 30th June 2017



3. CCI OPEN DATA PORTAL : CENTRAL ARCHIVE AND METADATA CATALOGUE

In the second year of operations, the work focused on several strands:

- Updating the datasets in the archive as new data were produced and made available by the CCI teams
- Distributing these datasets via the widest range of download methods. Datasets are being published via the Earth System Grid Federation (ESGF) to provide access to the download services available there, such as http, opendap, WMS and WCS. However, not all datasets were immediately suitable for publication to ESGF and WMS and work was required to add as many datasets as possible.
- Updates to the CSW service to resolve security issues and allow refinement of search facets.
- Interactions with the GEOSS portal resulting in CCI datasets being shown within the GEOSS portal, with regular updates harvested from our CSW.

3.1 Interface and server-side tools

3.1.1 Download methods

In the first phase of operations data was being made available via anonymous ftp, and also via the CEDA archive download routes². In the second year the portal front-end went live, and other data download methods were implemented, specifically http, OPeNDAP, WMS and WCS. Within the portal architecture, these latter methods rely on the publication of datasets to the Earth System Grid Federation (ESGF)³.

The available data access routes are summarised in the Table below:

Download methods	Label	Requires Publication via ESGF
HTTP data download provided by ESGF THREDDS Data Server	CEDA Managed ESGF-HTTP	Yes
OPeNDAP subsetting provided by ESGF THREDDS Data Server	CEDA Managed ESGF – OPeNDAP	Yes

² These CEDA archive download routes are core services provided for all data within the CEDA archive as part of its core remit to the UK's NERC user community, and provide additional access routes over those primary access routes provided in the CCI Open Data portal.

³ As noted in RD.2, the WMS and WCS download routes are provided via a separate THREDDS data server than the standard ESGF data node, as this version of TDS does not fully support WMS and WCS; however, it still makes use of the ESGF publisher to publish the data to this node, and results are still returned in the ESGF search.



Open Geospatial Consortium WMS	CCI ODP Managed – WMS	Yes
Open Geospatial Consortium WCS	CCI ODP Managed – WCS	Yes
Dedicated CCI anonymous FTP server	CEDA Managed – FTP	No
CEDA pydap server	CEDA core – PyDAP	No
CEDA ftp server	CEDA core – FTP	No
CEDA data browser (HTTP access)	CEDA core - HTTP	No

It should be noted that the publication of datasets via ESGF (TDS HTTP, OPeNDAP, WMS and WCS) is only possible for data in suitable formats (NetCDF). All CCI datasets however, can be accessed via the anonymous FTP service, and also via the core CEDA services described above.

3.1.2 Publication to ESGF

Within the portal architecture, as described in RD.2, the http, OpeNDAP, WMS and WCS download methods rely on the publication of the datasets to the Earth System Grid Federation (ESGF). In order to do this an ESGF identifier and search facets have to be identified based on controlled vocabularies, based on the CCI data standards. Where possible this has been done automatically using software that derives the ESGF identifier, assuming compliance with these standards. However, many of the datasets had minor metadata non-compliances that did not allow this. These were identified during the first year of operations and have been fed back to the data providers for incorporation in future datasets. In order to publish existing datasets to ESGF, work has been undertaken on a case by case basis, to manually create the ESGF identifiers and allow publication. It should also be noted that only datasets that are in NetCDF format can be included in ESGF, so some datasets that are in other formats (to meet user requirements for that community), cannot be included for dissemination by the ESGF download methods.

In order to make the data available via the portal WMS and WCS services, the data has to be a) publishable to ESGF (NetCDF format) and b) in a WMS compatible format (e.g. gridded data). Work has been undertaken to identify those in suitable formats, and further work is still ongoing to understand why a few datasets have not published correctly.

An issue was identified with publishing datasets with large numbers of files to ESGF, due to memory limitations in the ESGF publishing software. Various potential solutions to improve the ESGF publisher were investigated, but it has not proved possible to overcome this in the timescale of the project, as ESGF is an external collaboration and not under CEDA control. The best solution to this, in order to enable such datasets to be included in ESGF has been to publish a large dataset in smaller sub-sets e.g. yearly chunks. This has been done for ocean colour and SST data where needed, and has allowed publication in ESGF. This does however have implications for WMS viewing as currently these data can only be viewed in yearly chunks. The success and limitations of this approach are currently being assessed.



3.1.3 Metadata catalogue

As described in RD.2, for all datasets held within the CCI portal, metadata catalogue records have been created. This makes use of the CEDA MOLES catalogue. A Catalogue Service for the Web (CSW) service provides a WAF (Web Accessible Folder) interface which exposes these MOLES records rendered as ISO 19115 records. This enables the CSW to regularly poll MOLES for new ISO records and publish them to the outside. In Year 2, an updated CSW has been deployed.

This CSW interface is queried by the CCI portal front end web presence to provide up to date metadata information to display in the portal dashboard and search. It can also be queried and searched by other services, such as the GEOSS portal or the CCI Toolbox. For instance, the GEOSS portal is now displaying CCI data, which is kept up to date automatically, by regularly harvesting data from the CSW.

The CSW only exports the current version of any CCI dataset within the portal; older data versions may remain available within the archive, but are not exported to be displayed in the portal dashboard and search.

Metadata can also be queried via the ESGF search API (which supports file level metadata). However, as not all datasets can be published via ESGF, and ESGF also retains older versions of the data in many cases, then the CSW should be regarded as the definitive source of all the latest CCI Open Data Portal data.

3.2 Dataset summary

In the second year of operations, many of the CCI portal teams released additional products or new versions of datasets, and these continue to be added to the data archive. Where new versions of a dataset have been released, new catalogue records have been created, and released to the Open Data Portal, via the CSW.

As mentioned in the previous section, in the front end portal dashboard and search interfaces, only the latest version of a data product is shown. However, within the back-end CEDA archive, where older versions of the datasets have previously been displayed in the portal, these are, in general, still made available alongside the newer product. However, they have been completely removed in several cases at the request of the respective CCI teams. The keeping of older versions of the data could be reconsidered on a case by case basis if the space in the archive becomes an issue, in consultation with the relevant CCI team and ESA, as per the process detailed in RD.1 Technical Specification Document, CDAP.REP.011, Issue 1.0.

New data have been added for:

- Landcover – global maps (June 2016)
- Fire – burned area v4.1 product (July 2016)
- Ocean Colour v2 (Nov 2016)
- Greenland Ice Sheets (most Nov 2016, Sentinel-1 Dec/Jan 2017)
- Sea Level (January 2017)
- Sea Ice Concentration (February 2017)
- Cloud (March 2017)
- Obs4MIPS data from the Greenhouse Gas project (August 2016)



DOI's have also been issued for 5 CCI teams:

- Fire burned area v4.1
- One SST product
- Ocean colour v1 and v2
- Sea ice concentration v2
- Greenhouse Gases Obs4MIPS data products

At the end of year 2⁴, the current archive contains 99 TB of data. Approximately 97 datasets are shown via the front-end of the Open Data portal, but further older datasets are also archived in the metadata catalogue and included in the above figure of 99TB.

For a given CCI project, the definition of what constitutes an individual 'dataset' in the portal can be subjective. In general, within the MOLES metadata catalogue and CSW, a dataset follows the same split as is used by the CCI project itself, but in a few cases a different granularity has been required for practical purposes. This is the dataset as is shown in the portal dashboard and search interfaces. However, in publishing to ESGF, the granularity of a dataset has sometimes had to be increased, and well over 100 ESGF datasets have now been published. If MOLES datasets are split within ESGF, then links to all the relevant ESGF records are included in the CSW metadata.

In addition, not all datasets are available via all download methods. All archive datasets are available via the anonymous ftp service and the CEDA core services. However, as mentioned in section 3.1.2, not all of these datasets were suitable for immediate ingestion in the Earth System Grid Federation (ESGF). Only NetCDF formatted data can be published to ESGF, and of those potentially suitable, minor metadata issues have delayed publication in a number of cases as explained in section 3.1.2 and detailed further in subsequent sections. This means that only a subset of the CSW datasets are available in ESGF. Publication via WMS and WCS in the portal is reliant on the ESGF publisher to create the THREDDS catalogue, and also requires data to be suitably gridded. Hence, only a subset of the data published to ESGF is available via the WMS. Effort will continue to increase the proportion of datasets served by all methods.

The status of datasets in the portal (as of 4/5/2017) are summarised in Table 3.1 and described in the subsections below. We define datasets in the portal as those datasets that are displayed in the front end dashboard and search interfaces, which equates to those exposed via the CSW, which should be regarded as the definitive source for the latest CCI portal datasets. Where numbers of datasets are given, these relate to the CSW definition of datasets. For the figures for ESGF and WMS we have reported on the equivalent to a CSW dataset (i.e. if 2 ESGF datasets are required to publish 1 CSW dataset, then this is reported as 1 dataset). Older versions of data are not included in the numbers, as they are not made available through the front end of the CCI portal.

ECV	No of datasets in	No of these potentially	No currently published via	No potentially	No currently published via
-----	-------------------	-------------------------	----------------------------	----------------	----------------------------

⁴ Data volumes are as available on 4/5/2016.



	the portal (available via the CSW) ⁵	suitable for ESGF publication ⁶	ESGF	suitable for WMS publication ⁷	WMS
Aerosol	17	16	8	16	0
Cloud	6	6	6	6	6
Fire	2	1	1	1	1
GHG	12	12	12	0	0
Glaciers	1	0	0	0	0
Greenland Ice Sheets	23	21	0	21 (tbc)	0
Landcover	2	2	2	2	2
Ocean Colour	10	10	10	5	5
Ozone	11	11	2	2	1
Sea ice	2	2	2	2	2
Sea Level	3	3	2	2	1
Soil Moisture	4	4	4	3	3
SST	4	3	2	1	1

Table 3.1: Summary of the number of datasets currently in the CCI portal and available via ESGF and WMS, broken down by ECV. Note that not all datasets are suitable for publication in ESGF or WMS due to the data being in a different non-NetCDF format, or non-gridded data in the case of WMS. Publication to WMS also suitability for publication to ESGF. The numbers of datasets that could potentially be added are shown. Some datasets have not yet been published to ESGF / WMS, due to either metadata or formatting issues, but work is ongoing to include further datasets.

⁵ Corresponding download methods: CEDA managed- FTP, CEDA core- pydap, CEDA core FTP, CEDA core-HTTP (now retried)

⁶ Corresponding download methods: CEDA managed – ESGF HTTP; CEDA managed – ESGF OPeNDAP

⁷ Corresponding download methods: CCI managed – WMS; CCI managed - WCS



3.2.1 Aerosol:

Dataset	Version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	ESGF download available	WMS available
		(CEDA managed – FTP; CEDA core – FTP; CEDA core – OPeNDAP; CEDA core – HTTP)	(CEDA managed-ESGF – HTTP; CEDA managed ESGF- OPeNDAP	(CEDA managed-ESGF – HTTP; CEDA managed ESGF- OPeNDAP
L2 aerosol products from ATSR-2 (SU algorithm)	v4.21	Y	N (see Note 2)	N (see Note 4)
L3 aerosol products from ATSR-2 (SU algorithm)	v4.21	Y	Y	N (see Note 3)
L2 aerosol products from ATSR-2 (ORAC algorithm)	V3.02	Y	N(see Note 2)	N (see Note 4)
L3 aerosol products from ATSR-2 (ORAC algorithm)	V3.02	Y	Y	N (see Note 3)
L2 aerosol products from ATSR-2 (ADV algorithm)	V2.30	Y	N (see Notes 1 and 2)	N (see Note 4)
L3 aerosol products from ATSR-2 (ADV algorithm)	V2.30	Y	N (see Note 1)	N (see Note 1)
Level 3 aerosol products from the Multi-Sensor UVAI algorithm (MS UVAI)	V1.5.7	Y	Y (not daily, see Note 2)	N (reason to be investigated)
Images of Aerosol Absorbing Index produced by the Multi-Sensor UVAI (MS UVAI) algorithm	V1.5.7	Y	N (not NetCDF)	N (not NetCDF)



Level 3 aerosol products from GOMOS (AERGOM algorithm)	v2.19	Y	Y	N (see Note 3)
L2 aerosol products from MERIS (ALAMO algorithm)	V2.2	Y	Y	N (see Note 4)
L3 aerosol products from MERIS (ALAMO algorithm)	V2.2	Y	Y	N (see Note 4)

Current data volume: 992 GB

Note 1: Many of the datasets had minor metadata non-compliances and required manual tagging with the correct vocabulary terms to create the ESGF identifiers before they could be published to ESGF. Feedback has been provided to the aerosol CCI project. Many datasets have now been published, but where the manual tagging was still complex, these are still awaiting publication.

Note 2: Most L2 data products have not yet been published to ESGF whilst a solution for the issue of publishing large numbers of files was awaited. An approach to this has now been adopted and will be applied to these datasets in a subsequent ESGF publishing run. Similarly, the size of the MS-UVAI daily data also proved an issue and will have the same mitigation applied.

Note 3: WMS publication of a number of datasets was attempted but has not yet been successful and the reasons for this are under investigation. It is likely related to the way that the time variable is included in the data file.

Note 4: The L2 datasets are not gridded, so would not be suitable for the WMS.

3.2.2 Cloud

Dataset	Current Version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	ESGF download available	WMS download available
		(CEDA managed -FTP; CEDA core - FTP; CEDA core - OPeNDAP; CEDA core - HTTP)	(CEDA managed-ESGF - HTTP; CEDA managed ESGF-OPeNDAP)	(CEDA managed-ESGF - HTTP; CEDA managed ESGF-OPeNDAP)
AVHRR-AM monthly gridded cloud properties,	V2	Y	Y	Y



v2				
AVHRR-PM monthly gridded cloud properties,v2	V2	Y	Y	Y
MERIS+AATSR monthly gridded cloud properties, v2	V2	Y	Y	Y
ATSR2-AATSR monthly gridded cloud properties, v2	V2	Y	Y	Y
MODIS-AQUA monthly gridded cloud properties, v2	V2	y	Y	Y
MODIS-TERRA monthly gridded cloud properties, v2	V2	y	Y	Y

Current data volume: 231 Gb

This latest version was added to the archive in March 2017, and the previous version removed. In the previous version both L3C and L3U products were held (8 Tb for 3 years). In this version it was decided that the portal should only hold the L3C data directly, due to the size of the L3U data.

3.2.3 Fire

Dataset	Current versions in portal (via CSW)	Available via anonymous ftp and core CEDA methods	ESGF download available	WMS download available
		(CEDA managed –FTP; CEDA core – FTP; CEDA core – OPeNDAP; CEDA core – HTTP)	(CEDA managed-ESGF – HTTP; CEDA managed ESGF- OPeNDAP	(CEDA managed-ESGF – HTTP; CEDA managed ESGF- OPeNDAP



Burned Area Grid Product	v4.1	Y	Y	Y
Burned Area Pixel Product	V3.1,v4.1	Y	Not suitable format	Not suitable format

Current data volume: 42 Gb

DOI's were issued for v4.1 of the data. Only the gridded product is suitable for inclusion in ESGF and the WMS, as the pixel product is not in NetCDF format.

3.2.4 GHG

Dataset	Version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	ESGF download available	WMS download available
		(CEDA managed-FTP; CEDA core - FTP; CEDA core - OPeNDAP; CEDA core - HTTP)	(CEDA managed-ESGF - HTTP; CEDA managed ESGF-OPeNDAP)	(CEDA managed-ESGF - HTTP; CEDA managed ESGF-OPeNDAP)
SCIAMACHY CH4 Level 2 Data Product (CH4_SCI_WFMD), generated with the WFMD algorithm.	CRDP-3 (V3.9)	Y	Y	N (see Note 1)
SCIAMACHY CH4 Level 2 Data Product (CH4_SCI_IMAP) generated with the IMAP-DOAS algorithm.	CRDP-3 (v7.1)	Y	Y	N (see Note 1)
GOSAT CH4 Full Physics Level 2 Data Product (CH4_GOS_SRFP) generated with the SRFP (RemoTeC) algorithm.	CRDP-3 (V2.3.7)	Y	Y	N (see Note 1)
GOSAT CH4 Full Physics Level 2 Data Product, (CH4_GOS_OCFP) generated with the OCFP (UoL-FP)	CRDP-3 (V1.0)	Y	Y	N (see Note 1)



algorithm.				
GOSAT CH4 Proxy Level 2 Data Product (CH4_GOS_SRPR) generated with the SRPR (RemoTeC) algorithm.	CRDP-3 (V2.3.7)	Y	Y	N (see Note 1)
The GOSAT CH4 Proxy Level 2 Data Product, (CH4_GOS_OCPR) generated with the OCPR (UoL-PR) algorithm.	CRDP-3 (V6.0)	Y	Y	N (see Note 1)
The GOSAT CO2 Level 2 Data Product (CO2_GOS_OCFP) generated with the OCFP (UoL-FP) algorithm.	CRDP-3 (V6.0)	Y	Y	N (see Note 1)
The GOSAT CO2 Level 2 Data Product (CO2_GOS_SRFP), generated with the SRFP (RemoTeC) algorithm.	CRDP-3 (V2.3.7)	Y	Y	N (see Note 1)
The SCIAMACHY CO2 Level 2 Data Product (CO2_SCI_BESD) generated with the BESD algorithm	CRDP-3 (V2.01.01)	Y	Y	N(see Note 1)
The SCIAMACHY CO2 Level 2 Data Product (CO2_SCI_WFMD) generated with the WFMD algorithm	CRDP-3 (V3.9)	Y	Y	N (see Note 1)
The merged CH4 Level 2 Data Product (CH4_EMMA) generated with the EMMA algorithm.	CRDP-3 (V1.0)	Y	Y	N(see Note 1)
The merged CO2 Level 2 Data Product (CO2_EMMA)	CRDP-3 (v2.1)	Y	Y	N (see Note 1)



generated with the EMMA algorithm.				
Merged SCIAMACHY and GOSAT Level 3 gridded atmospheric column-average carbon dioxide (XCO2) product in Obs4MIPs format	CRDP-3 (V001) (see Note 2)	Y	Y (but see Note 2)	N (see Note 2)
Merged SCIAMACHY and GOSAT Level 3 gridded atmospheric column-average methane (XCH4) product in Obs4MIPs format	CRDP-3(v001) (see Note 2)	Y	Y (but see Note 2)	N (see Note 2)

Current data volume: 52 GB

Note 1: The GHG products (except for the Obs4MIPs data) are L2 non-gridded products and therefore not suitable for inclusion in the WMS.

Note 2: The Obs4MIPs data is published in ESGF via the Obs4MIPs project rather than that of the CCI portal, and it is not being shown via the portal and CSW. If deemed appropriate this data could be made visible by all download routes.

The portal currently holds only the GHG core products. There are also ‘Additional Constraints Algorithm’ products which are hosted on the GHG server, but which are not considered core products. In consultation with the GHG team, it has currently been decided not to host these directly in the portal archive.

Their final CRDP-4 dataset has recently been released and is scheduled for ingestion into the open data portal archive.

3.2.5 Glaciers

Dataset	Current version in the portal (via the CSW)	Available via anonymous ftp and core CEDA services	Available for ESGF download	Available in WMS
		(CEDA managed-FTP; CEDA core - FTP; CEDA core - OPeNDAP; CEDA core - HTTP)	(CEDA managed-ESGF - HTTP; CEDA managed ESGF-OPeNDAP)	(CEDA managed-ESGF - HTTP; CEDA managed ESGF-OPeNDAP)



Randolph Glacier Inventory gridded data product	V5.0	Y	N (not suitable format)	N (not suitable format)
---	------	---	-------------------------	-------------------------

Current data volume: 8 MB

The current glacier data held in the portal is in ascii format, and as such is not suitable for ESGF. Other data is available and needs to be assessed for inclusion in the portal; this other data are primarily shapefiles.

3.2.6 Ice Sheets – Antarctica

In the last year the first data was released by the Antarctic Ice Sheets team, and is scheduled for ingestion to the open data portal archive.

3.2.7 Ice Sheets – Greenland

Dataset	Current version in the portal (via the CSW)	Available via anonymous ftp and core CEDA services	Available for ESGF download	Available in WMS
		(CEDA managed –FTP; CEDA core – FTP; CEDA core – OPeNDAP; CEDA core – HTTP)	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP
Greenland Calving Front Locations	V2.0	Y	N (not in NetCDF format)	N (not in NetCDF format)
Greenland Surface Elevation Change 1992-2014	V1.2	Y	N (see Note 1)	N (see Note 1)
Greenland Surface Elevation Change from Cryosat	V2.0	Y	N (see Note 1)	N (see Note 1)
Grounding Line Locations	V1.2	Y	N (not in NetCDF format)	N (not in NetCDF format)
Ice Velocity Time Series for the Upernavik region, v1.1	v1.1	Y	N (see Note 1)	N (see Note 1)



Ice Velocity data for the Greenland Margin from the Palsar instrument	V1.1	Y	N (see Note 1)	N (see Note 1)
Ice velocity time series for Jakobshaven Isbrae glacier, 2002-2010	v1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity data for the Greenland margin from ERS2 for winter 1995-1996, v1.1	V1.1	Y	N (see Note 1)	N (see Note 1)
Ice velocity data for the Greenland Northern Drainage basin from ERS-1 for winter 1991-1992	V1.1	Y	N (see Note 1)	N (see Note 1)
Greenland Ice Velocity Map Winter 2013-2014, v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Greenland Ice Velocity Map Winter 2014-2015, v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Greenland Ice Velocity Map Winter 2015-2016, v1.0	v1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series of the Storstrommen Glacier for 2015-2016 from Sentinel-1, v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series of the Helheim Glacier for 2015-2016 from Sentinel-1, v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series	V1.0	Y	N (see Note 1)	N (see Note 1)



of the Zachariae Isstroem for 2014-2016 from Sentinel-1, v1.0			1)	1)
Ice Velocity Time series of the Upernavik Isstroem for 2015-2016 from Sentinel-1,v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series of the Petermann Glacier for 2015-2016 from Sentinel-1,v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series of the 79-Fjord Glacier for 2015-2016 from Sentinel-1,v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series of the Hagen Brae for 2015-2016 from Sentinel-1,v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series of the Jakobshavn Isbrae for 2014-2016 from Sentinel-1,v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Ice Velocity Time Series of the Kangerlussuaq Glacier for 2015-2016 from Sentinel-1, v1.0	V1.0	Y	N (see Note 1)	N (see Note 1)
Greenland Gravimetric Mass Balance from GRACE (DTU_Space)	V1.0	Y	N (see Note 1)	N (see Note 1)
Greenland Gravimetric Mass Balance from GRACE data (TU Dresden)	V1.0	Y	N (see Note 1)	N (see Note 1)

Current data volume: 6.1 GB

Note 1: The Greenland Ice Sheets data has not yet been published to ESGF or WMS due to metadata issues which require manual creation of the ESGF facets. A new release of the data has now been received from the Greenland Ice Sheets CCI team, and is scheduled for inclusion



on the CCI Open Data Portal in the imminent future. Therefore publication to ESGF will also be delayed until this release is available on the portal.

The version 1.1 data now advertised through the portal was added over the course of this year. However, there was some confusion with version numbering between subsequent releases of the ice velocity data, leading to two similarly labelled but different versions. This has been noted on the metadata catalogue pages, and the earlier data remains available in the archive (but not on the dashboard or search interfaces) for reference.

3.2.8 Land cover

Dataset	Current version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	Available for ESGF download	Available in WMS
		(CEDA managed-FTP; CEDA core - FTP; CEDA core - OPeNDAP; CEDA core - HTTP)	(CEDA managed-ESGF - HTTP; CEDA managed ESGF- OPeNDAP	(CEDA managed-ESGF - HTTP; CEDA managed ESGF- OPeNDAP
Water Bodies Map	V4.0	Y	Y	N (reason under investigation)
Land cover maps, v1.6.1	V1.6.1	Y	Y	N (reason under investigation)

Current data volume: 19 Gb

Only a subset of the landcover products are available directly in the Central Data Archive, but metadata catalogue records have been created for the other products, pointing at the external data sources. These are then visible in the portal dashboard and search interfaces.

The data currently on the archive could not be included in the WMS for reasons under investigation. A new version of the land cover maps data is recently been released and is scheduled for inclusion in the Open Data Portal.

3.2.9 Ocean Colour

Dataset	Current version in portal (via	Available via anonymous	Available for ESGF download	Available in WMS



	CSW)	ftp and core CEDA methods		
Global ocean colour data products gridded on a geographic projection (All Products); includes daily, 8-day, 5-day, monthly and climatology products.	2.0	Y	Y (see Note 2)	Y (see Note 2)
Global ocean colour data products gridded on a sinusoidal projection (All Products); includes daily, 8-day, 5-day, monthly and climatology products	2.0	Y	Y (see Note 2)	N (see Note 1)
Global attenuation coefficient for downwelling irradiance (Kd490) gridded on a geographic projection; includes daily, 8-day, 5-day and monthly products	2.0	Y	Y	Y
Global attenuation coefficient for downwelling irradiance (Kd490) gridded on a sinusoidal projection; includes daily, 8-day, 5-day and monthly products	2.0	Y	Y	N (see Note 1)
Global chlorophyll-a data products gridded on a geographic projection; includes daily, 8-day, 5-day and monthly products	v2.0	Y	Y	Y
Global chlorophyll-a data products gridded on a sinusoidal projection; includes daily, 8-day, 5-day and monthly products	v2.0	Y	Y	N (see Note 1)



Global dataset of inherent optical properties (IOP) gridded on a geographic projection; includes daily, 8-day, 5-day and monthly products	v2.0	Y	Y (see Note 2)	Y
Global datasets of inherent optical properties (IOP) gridded on a sinusoidal projection; includes daily, 8-day, 5-day and monthly	V2.0	Y	Y	N (see Note 1)

Current data volume: 93 TB (V1 (30Tb) and v2 (63 Tb)).

V3 of the data has been released, and is scheduled for addition to the open data portal imminently. Earlier versions have been assigned DOI's and will hence be retained in the archive.

Note, although the volume of the data is large, CEDA is committing to archive all versions of the data under our national funding. Hence, older versions will not count towards the 100 Tb of storage assigned to the CCI portal.

Note1: Sinusoidal projection data has not been published via WMS as the sinusoidal grid is not suitable. The equivalent geographic projection data is published via WMS

Note2: For the largest datasets, it was not possible to publish to ESGF as a single dataset, due to memory limitations of the publisher. Hence, data has had to be published in yearly chunks, as previously described. This also affects the WMS and WCS.

3.2.10 Ozone

Dataset	Current version in portal (via CSW)	Available via anonymous FTP and core CEDA methods	Available for ESGF download	Available in WMS
Level 3 Total Ozone Merged Data Product	V01	Y	Y	N (see Note 2)
Level 3 Nadir Ozone Profile Merged Data Product	V2	Y	Y	Y



ACE Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V1	Y	N (see Note 1)	N (see Note 1)
GOMOS Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V1	Y	N (see Note 1)	N (see Note 1)
MIPAS Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V1	Y	N (see Note 1)	N (see Note 1)
ODIN/SMR (544.6 GHz) Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V1	Y	N (see Note 1)	N (see Note 1)
ODIN/SMR Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V1	Y	N (see Note 1)	N (see Note 1)
OSIRIS Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V1	Y	N (see Note 1)	N (see Note 1)
SCIAMACHY Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V1	Y	N (see Note 1)	N (see Note 1)
Merged Level 3 Limb Ozone Monthly Zonal Mean (MZM) Profiles	V2	Y	N (see Note 1)	N (see Note 1)
Merged Level 3 Limb Ozone Semi-Monthly Mean Profiles	V2	Y	N (see Note 1)	N (see Note 1)

Current data volume: 1.5 GB

Note 1: The limb datasets have not been added to ESGF due to minor metadata issues; further work will be undertaken in order to include these in future publication runs.

Note 2: There were problems publishing the total column ozone data to the WMS, and this is being investigated.



Other data products have now been released by the ozone team, and will be added to the portal in the future as appropriate.

3.2.11 Sea ice

Dataset	Current version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	Available for ESGF download	Available in WMS
		(CEDA managed –FTP; CEDA core – FTP; CEDA core – OPeNDAP; CEDA core – HTTP)	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP
Sea ice concentration from the AMSR instruments on a 25km grid	V2.0	Y	Y	Y
Sea ice concentration from the AMSR instruments on a 50 km grid	V2.0	Y	Y	Y

Current data volume: 331GB

V2.0 of the data was added to the archive and DOI's were assigned to this data.

3.2.12 Sea level

Dataset	Current version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	Available for ESGF download	Available in WMS
		(CEDA managed – FTP; CEDA core – FTP; CEDA core – OPeNDAP; CEDA core – HTTP)	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP
Oceanic Indicators of Mean Sea Level Changes	V2.0	Y	Y	Y (see Note 2)
Time series of gridded Sea Level	V1.1	Y	N	N



Anomalies (SLA)					
Sea Level Fundamental Climate Record	Level Data	V2.0	Y	N	Not suitable

Current data volume: 272 GB

Note 1: V2.0 of the sea level data was added to the portal, including an additional FCDR product. This product has not been published to ESGF due to metadata non-compliance and is not suitable for WMS visualisation as it is not a gridded product.

Note 2: Only one of the three indicators was suitable to be published to WMS. One of the other indicators was a time series and not suitable for viewing via WMS. The reason for the unsuitability of the third is under investigation

Note 3: The time series of gridded Sea Level Anomalies could not be published to ESGF (and hence WMS) due to the use of a scalar time variable. This has been reported to the sea level team. Soil moisture

Dataset	Current version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	Available for ESGF download	Available in WMS
		(CEDA managed –FTP; CEDA core – FTP; CEDA core – OPeNDAP; CEDA core – HTTP)	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP	(CEDA managed-ESGF – HTTP; CEDA managed ESGF-OPeNDAP
Soil Moisture 'Active' Dataset	v02.2	Y	Y	Y
Soil Moisture 'Passive' Dataset	v02.2	Y	Y	Y
Soil Moisture 'Combined' Dataset	v02.2	Y	Y	Y
Ancillary datasets used for the "Active", "Passive" and "Combined" products	v02.2	Y	N (see Note 1)	N (see Note 1)

Current data volume: 48 GB



No new data added this year, but a new dataset has been recently released and is scheduled for inclusion in the portal.⁸

Note 1: The Ancillary datasets are a collection of various input data used in the derivation of the other products, and were not considered appropriate for publication via ESGF and WMS.

3.2.13 SST

Dataset	Current version in portal (via CSW)	Available via anonymous ftp and core CEDA methods	Available for ESGF download	Available in WMS
		(CEDA managed –FTP; CEDA core – FTP; CEDA core – OPeNDAP; CEDA core – HTTP)	(CEDA managed-ESGF – HTTP; CEDA managed ESGF- OPeNDAP	(CEDA managed-ESGF – HTTP; CEDA managed ESGF- OPeNDAP
Advanced Very High Resolution Radiometer (AVHRR) level 2 pre-processed (L2P) long term product	V1.0	Y	N (see Note 2)	N (see Note 2)
Along-Track Scanning Radiometer (ATSR) level 3 un-collated data (L3U) long-term product	v1.1	Y	Y (but see Note 2)	N (see Note 2)
Analysis long term product version	v1.1	Y	Y	Y
GHR SST Multi-Product ensemble (GMPE).	-	Y	N (see Note 2)	N (see Note 2)

Current Data Volume: 4.1 TB

No new data were added to the portal in this year. A DOI was issued for v1.1 of the Analysis long term product.

Note 1: There was an issue with the number of files in the ATSR L3U and AVHRR L2P products causing problems with publication to ESGF. A trial solution of publishing the L3U data in yearly

⁸ Since the initial version of this report, this new data version (v3.2) has now been released.



chunks has been implemented, and will be extended to publication of the L2P dataset in future publication runs. The data has not been included in the WMS server yet.

Note 2: The GHRSSST Multi-Product Ensemble (GMPE) dataset contains various types of data from external sources and hence does not follow CCI data standards; these were therefore not appropriate for inclusion in ESGF of WMS



3.3 Download statistics

Monitoring of data downloads via the various download routes has been implemented over the course of the year (prior to those download methods going live). These statistics have been reported in the monthly reports and are summarised below for the different data methods.

It should be noted that ESGF download services only came on line part way through the year, and not all datasets are available for download via this method. The proportion of datasets in ESGF increases through time.

As well as the download methods directly associated with the CCI Open Data Portal, download routes are also available via the CEDA archive, and these are also reported below. The CEDA data browser was the most popular of these routes; however this was switched off at the end of February 2017 and replaced by an OpenDAP service. The download stats for this were not available initially, and hence the stats reported in the monthly report were lacking this figure for a number of months. The download stats have now been reprocessed to include data via this route, and these are reported below and in the attached spread sheet.

WMS and WCS statistics are described separately in section 3.3.4

3.3.1 Data downloads by method (May 2016- April 2017)

Download category	Number of users	Size of data	Number of accesses	Activity Days
CEDA managed – FTP (anon-ftp.ceda.ac.uk)	483	6.15TiB	1296308	1478
CEDA managed – ESGF- HTTP + CEDA-managed ESGF-OPeNDAP	33	0.30 TiB	6639	138
CEDA core – HTTP (2 methods)	507	0.61 TiB	46461	1105
Dbrowser	505	0.46	34723	1089
Dbrowser-mget	15	0.15	11738	16



CEDA core – pydap (2 methods)	262	0.23	48421	525
CEDA core – FTP (2 methods)	19	1.47	308658	64
Total	1170	8.75 TiB	1706487	3310

Table 3.2: Data download by method; Note that in reporting number of methods in stats the CEDA core – Pydap and CEDA core – FTP can consist of two separate methods.

Note that these values include downloads of documentation such as product user guides and readme files. (These are shown in the download statistics reported monthly grouped under /neodc/esacci). The CEDA core –pydap and CEDA core – ftp are reported in the grouped statistics used in data monitoring as 2 methods each. E.g. the pydap server separates out downloads of entire files and subset files in our statistics software, although access is via the same infrastructure and has been collated here to match the download methods described in RD-2. The vast majority of accesses are of complete files. In the case of the CEDA core –ftp, this was due to 2 users still making user of an old historic ftp site instead of our current one.

3.3.2 Data downloads by dataset

Comprehensive data downloads by dataset have been provided in a separate spreadsheet, also broken down by month. As our download statistics reporting has evolved over the year, there may have been differences in how the data was reported over time. All the stats have been reanalysed, and now form a consistent set for the whole year.

Note, as mentioned above the top line of the stats labelled /neodc/esacci refers to miscellaneous downloads that do not correspond to a dataset. This is primarily downloads of readme files and product documentation.

3.3.3 Historic Data Download Trends

The table below shows the data download trends per month.

Please also note that, as already noted, one of the download methods (the CEDA data browser) was removed from operation at the end of February 2017, and the statistics for its replacement are not yet included in this table.

Date	Users	Methods	Datasets	Number of accesses	Size (GiB)	Activity days
201605	105	5	33	211,699	761.516	179
201606	67	4	26	28,623	174.888	106



201607	64	4	71	405,938	695.119	275
201608	84	4	105	268,071	957.341	329
201609	148	5	79	156,426	699.565	358
201610	159	5	57	170,151	676.437	321
201611	124	5	51	92,419	91.953	257
201612	95	3	46	26,667	154.248	188
201701	100	5	55	101,431	1,181.27	227
201702	151	5	61	59,360	98.945	247
201703	214	6	69	129684	515.656	409
201704	164	6	86	56018	1955.63	414
Total:	939	6	149	1,658,054	8.526121 TiB	2,789

Table 3.3 Data download trends per month for all download methods mentioned in the previous section.

3.3.4 WMS and WCS statistics

The Web Mapping Service (WMS) and Web Coverage Service (WCS) statistics are described separately in this section. As the WMS provides online visualisation of the datasets it was not felt to be appropriate to provide like-for-like statistics with the other download protocols; for instance the size of data downloaded will not relate to the same quantity.

The WCS does allow the download of subsets of data, and so in the future should be added to the general download stats; however it is logged within the same download logs as the WMS, and is currently being analysed at the same time. There is little current usage of the WCS, as is discussed below, and these would therefore not affect the overall download stats above.

Log files with basic access monitoring for WMS and WCS data have also been collected for the entire period of operations, and with increased details since April 2017. However, there has been a delay in setting up the tools to report these statistics, and work is still ongoing to provide the comprehensive set of appropriate reporting metrics. Preliminary statistics are provided below, and analysis of the more detailed log files in the future will allow further details to be added for the later period of data.

As a basic reporting metric, the concept of a 'dataset day' has been used, which equates to the number of calendar days in which a given dataset has been requested each month. Total dataset days equate to the sum of the individual dataset days. The data are split into datasets as per the ESGF definition of a dataset, which, as discussed in section Dataset summary, can be at a



finer granularity than the datasets shown within the portal dashboard. Table 3.1 gives a summary of these stats from the period from the 1st September 2016 (when the portal went live), until the end of April 2016. A more detailed breakdown per dataset are provided in external spreadsheets.

Month	WMS (Total dataset days)	WCS (Total dataset days)
201609	283	125
201610	65	7
201611	76	5
201612	52	22
201701	35	2
201702	67	5
201703	7	2
201704	32	0

Table 3.4: Data downloads via WMS and WCS, reporting the total number of dataset days per month.

It should also be noted that the statistics report requests for datasets that have at some point been published via WMS; some datasets, particularly initially, were published, but were not suitable for WMS viewing, resulting in error messages (e.g. aerosols, ghg). These links have either now been removed, or the reasons for the problems with publication are under investigation. These requests are still useful to include in the stats as it gives an idea of user interest.

The usage figures for both the WMS and WCS services are low, and as such, are probably mainly indicative of testing from CCI portal team members. For instance, the high figures in September will be due to activity associated with the 'go-live'. There are potentially several reasons why the use of these services may be currently low, and these will be assessed and addressed where possible. Possible reasons are:

- Users are not aware of these services or how they can be used. For instance, the use of the terms WMS and WCS will only be meaningful to technical users. This can be addressed in the introductory materials for users that will be issued on the website as part of CCN3.
- Users may have a preference for downloading the full dataset, in which case other download routes are more suitable. They may still also find WMS/WCS access useful if more aware of it as per the point above.
- Many of the datasets have not been available via WMS, particularly in earlier time periods. Not all datasets are suitable for publication for WMS, and it has taken a considerable effort to publish many of the other datasets with minor metadata inconsistencies, or due to issues with the ESGF publisher (see section 3.1.2). This has



taken time to resolve, but the number of available datasets is now considerably increased.

- There have been some issues with the performance of the WMS server, which can be slow to respond and sometimes to time out, particularly with larger datasets. We have identified a solution to these time outs, which is currently being assessed for general impact on other ESGF operations, and intend to implement this as soon as possible.



4. FUTURE ACTIVITIES IN YEAR 3

Activities that will be required will include:

- Updates to data products as new versions are released by CCI teams
- Maintenance of metadata catalogue records as additional information and feedback is received from the CCI teams
- Adjustments to the metadata records to support dashboard and search activities.
- Further work to ensure all suitable datasets are included in ESGF and WMS.

In this next phase, CEDA will continue to lead on the archive and metadata catalogue operations, whilst some of the operational work will be handed over to TVUK and CGI. This handover was envisaged for year 2, but has been delayed as the processes are put in place. It is likely that the initial handover will encompass the ingestion of new and updated datasets from a subset of the data providers, and providing updated content for the metadata catalogue. As more processes are automated and made accessible outside CEDA during the course of the project, additional tasks are likely to be handed over.

In year 3 we are proposing to adopt a regular schedule for the publication of new datasets into the archive, to allow for better planning and coordination between the different people involved in the publishing process. This schedule will be communicated to all the CCI teams and regular key users of the portal such as the CCI Toolbox.