# Report 2019 of the DDC at DKRZ

Document ownership and history					
Owner	IPCC DDC at DKRZ ( <u>ipcc.wdc-climate.de</u> )				
Location	DDC_report_DKRZ_2019.docx				
Author team:	M. Stockhause				
Version	1.0				
Date	2020-02-03				
Version history					
Date	Version	Comment			
2020-02-03	1.0	Report completed incl. internal review			

# **Table of Contents**

1.	Summary	.1
2.	Evolution of data access	.2
3.	Geographical distribution of data access	.2
4.	Data access by category AR	.4
5.	Review of user queries	.5
6.	News and activities	.5

# 1. Summary

The total AR5 data volume provided by IPCC DDC hosted at DKRZ is 1.7 PBytes, 1.6 PBytes in the DDC AR5 Reference Archive and 100 TBytes in the IPCC AR5 WG1 Archive, compared to about 1 TBytes for AR4, and less than 10 GBytes each for the preceding ARs: FAR, SAR, and TAR. 35 TBytes of SR1.5 (Special Report on Global Warming of 1.5°C) data were added to the IPCC DDC web page and of 2019.

In 2019 IPCC DDC users downloaded ca. 0.6 PBytes of data in ca. 600 000 individual file downloads, about 35 % less downloaded volume and 27 % less in the number of downloaded files compared to 2018. The data downloads for the other ARs showed an opposing trend with increases in the numbers of downloaded files between +70 % for AR4 and +400 % for SAR. The offer to send data for selected areas on storage media was requested by 3 users located in developing or economy-in-transition countries in 2019.

705 DDC users plus an unknown number of ESGF users accessed IPCC DDC data in 2019, which is ca. 15 % more than 2018. As the ESGF users, for which no geographical location is known, are responsible for 60 % of the total downloads, a reliable estimation of the geographical distribution of the users is difficult for 2019. Next year ESGF will provide download statistics. The average number of downloads for an individual IPCC DCC user decreased by 60 % from 2018 to 2019.

## 2. Evolution of data access

In the user downloads from the DDC reference archive, the download rates peaked in 05/2017 with a download volume of over 2 PBytes (**Figure 1**). The reason is unknown. For the previous and following years 1 PByte and 0.9 PByte, respectively, were registered. In 2019, the annual download decreased compared to 2018 by about 35 % in volume to a value of 0.6 PBytes and by about 27 % in download counts to a value of ca. 600 000 datasets. The download decrease is observed over the whole year with a slight increase in download volume from 03/2019 to 04/2019 right after the "CMIP6 Model Analysis Workshop" (25 to 28 March 2019 in Barcelona, Spain). The mean monthly download rate in 2019 was with 47 TByte/month about 35 % less the value for 2018.



*Figure 1:* Total data download counts and volumes per months over the last three years in GBytes from the IPCC DDC reference archive.

## 3. Geographical distribution of data access

For the IPCC DDC AR5 data, direct data access at the DKRZ and for AR5 additionally data access via ESGF (Earth System Grid Federation) is supported. For the ESGF data access share, no information about user locations is available.

About 60 % of the registered active users were located in Asia and 18 % in Europe as in the previous years. Accordingly, the share of users located in developing or economy-in-transition countries (Asia, South America, and Africa) was about 66 % (**Figure 2**, bottom), which is again a slight increase from 2017 with 61 %. The number of active users increased by ca. 15 % to 705 in 2019 (**Figure 2**, top). The average number of file downloads per user in 2019 decreased by ca. 60 % compared to 2018, which is higher than the decrease in total download numbers (-27 %) indicating that the average user in 2019 was less active than one in 2018. 60 % of the data was downloaded via the ESGF, for which no user locations have been recorded. Therefore an overall geographical distribution of the IPCC DDC users can only be roughly



estimated. As the ESGF is currently finalizing a user statistic dashboard, this data should be available for 2020.

*Figure 2*: Number of active DDC users in 2019 (bottom) and downloads counts of users per continent (top).

Assuming the same user distribution for the ESGF users as the DKRZ users, 5 % of the files were downloaded by European, 64 % by Asian, 16 % by North American, 14 % from Australian, and less than 1 % from African and South American users each. The share of downloads from users from developing and economy-in-transition countries (Asia, Africa, and South America) is with 65 % double the share in 2018 and is again dominated by the Asian users.

#### 3.1 Data on storage media

The interest in the DDC service to send a data subset for a geographical area on storage devices by mail decreased again from 9 USB devices sent to 7 users in 2018 to only 3 users requesting these service in 2019. 2 of the 3 users were interested in multiple areas resulting in 6 USB devices sent in 2019. No user was interested in AR4 data on storage media in 2019.

All 3 requestors were located in developing or economy-in-transition countries (2 from Africa, 1 from Asia; see **Table 1** and **Table 2**). There might be several different reasons for this decrease of service requests including a higher bandwidth for Asian and South American users.

In addition to these numbers, USB devices are also shared among colleagues. Therefore there might be an unknown number of additional researchers, reusing AR5 data for the regions sent on USB devices to a colleague this year or in previous years.

Area of Data	No of storage media	Africa	Asia	Australia / Central Pacific	Europe	North America	South America	North Pole	South Pole
AR4	-	-	-	-	-	-	-	-	-
AR5	6	2	3	1	2	2	2	-	-
DDC total	6	2	3	1	2	2	2	-	-

Table 1: Number of storage media requests per data area for AR4 and AR5 in 2019.

 Table 2: Number of user requests for AR4/AR5 data on storage media per user origin (continent) in 2019.

User origin	No of users	African users	Asian users	South American users	Australia / Central Pacific	European users
AR4	-	-	-	-	-	-
AR5	3	2	1	-	-	-
DDC total	3	2	1	-	-	-

#### 4. Data access by category AR

The monthly download rates in 2019 from the IPCC DDC reference archive were dominated by AR5 downloads as in the previous years, but the trend in AR5 downloads is opposite to those for data downloads for older ARs (**Figure 3**; online monthly download statistics<sup>1</sup>). Where the download rate for AR5 decreased in 2019 compared to 2018, the interest in the data for the previous ARs (AR4 and older) has increased. The reason for this increase in unknown. A plausible explanation might be the replication of the whole archive by an external repository.

<sup>&</sup>lt;sup>1</sup> Online monthly download statistics are available at:

https://cera-www.dkrz.de/WDCC/ui/cerasearch/statistics?type=downloads\_by\_domain&domain=IPCC-DDC

https://cera-www.dkrz.de/WDCC/ui/cerasearch/statistics?type=downloads\_by\_domain&domain=IPCC-DDC\_AR5

 $https://cera-www.dkrz.de/WDCC/ui/cerasearch/statistics?type=downloads\_by\_domain\&domain=IPCC-DDC\_AR4$ 

AR5 data downloads deceased by 27 % in number and by 35 % in volume between 2018 and 2019, which defines the general data download trend. For AR4, TAR and SAR significant increases in the number of downloaded datasets were registered: +70 % (AR4), +160 % (TAR) and +400 % (SAR).



*Figure 3:* Total annual data download counts (left) and volumes in GBytes (right) over the last five years for the different DDC reference archives (without FAR and SR1.5).

#### 5. Review of user queries

There are no numbers for the handled user requests by WDCC/DKRZ staff available for 2019, same as for the previous years. A separation of user requests on IPCC DDC issues is not possible. For ESGF, no user statistics are available for 2019, either.

In parallel to the regular user support channels, additional requests were directed to individuals at the modelling centers or at the data centers (within ESGF or to WDCC/DKRZ).

#### 6. News and activities

35 TBytes of SR1.5 data were added to the Reference Data Archive following a decision by TG-Data during the TG-Data 1 meeting in 11/2019. The corresponding DDC web page was published in 12/2019.

As a further DDC activity, a Virtual Workspace was offered as collaboration platform in support of the IPCC AR6 authors.