DDC Report 2018 of WDCC

Document ownership and history	
Owner	WDCC / DKRZ (ipcc.wdc-climate.de)
Location	DDC_report_WDCC_2018.docx
Author team:	M. Stockhause
Version	1.0
Date	2019-02-15

Version history

Date	Version	Comment
2019-01-23	0.1	geographical user distribution missing; numbers of user requests not available
2019-02-13	1.0	numbers reviewed and missing data added

Table of Contents

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 data for the HAPPI-MIP project providing the data for the SR1.5 (Special Report on Global Warming of 1.5°C) were archived during 2018. The data will be added to the IPCC DDC in early 2019.

In 2018 IPCC DDC users downloaded ca. 0.9 PBytes of data in ca. 820 000 individual file downloads, which is about 25 % of the volume and 60 % of the downloaded files in 2017. Neglecting the 05/2017 peak in download volume the downloaded volume in 2018 is with 65 % of that in 2017 consistent with the trend in file downloads. The AR4 data downloads continue to decrease.

612 WDCC users plus an unknown number of ESGF users accessed IPCC DDC data in 2018. This is nearly the same number as for the previous year. As the ESGF users, for which no geographical location is known, are responsible for 40 % of the total downloads, a reliable estimation of the geographical distribution of the users is difficult for 2018. Next year ESGF will provide download statistics. The average number of downloads for an individual IPCC DCC user was ca. 1 300 files, which is similar to the value in 2016 and 60 % of the value in 2017.

The offer to send data for selected areas on storage media was requested by 7 users, 4 of them were located in developing or economy-in-transition countries.

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. This month is responsible for a significantly higher download rate in 2017 with 3.5 PBytes compared to 1 PByte in 2016 and 0.9 PBytes in 2018. Apart from 05/2017, no overall tendency in the monthly downloads is shown over the three years period. Downloads remain on a high level with seasonal variations. The mean monthly download rate for 2018 was ca. 73 TBytes/month.

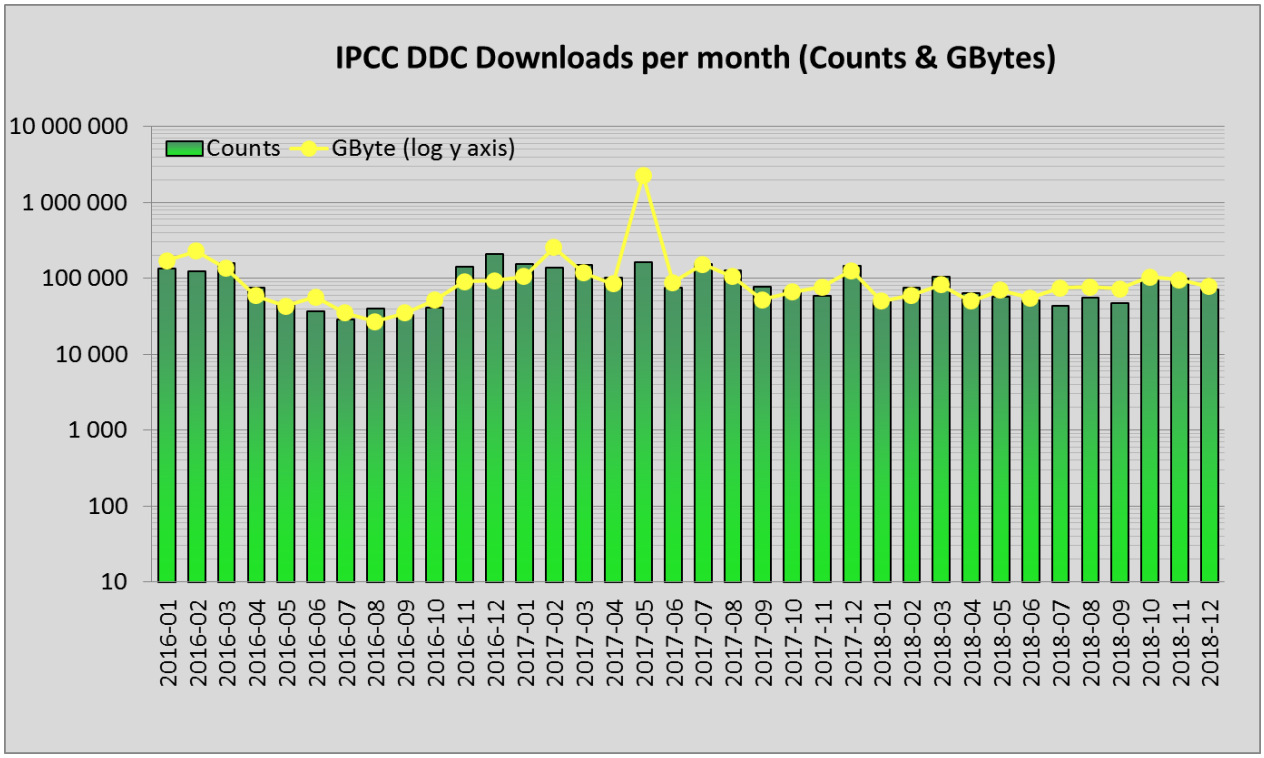

Figure 1: *Total data download counts and volumes per months over the last three years in GBytes from the IPCC DDC reference archive.*

3. Geographical distribution of data access

For the IPCC DDC AR5 data, direct data access at the WDCC and 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 20 % in Europe, similar to the previous years. Accordingly, the share of users located in developing or economy-in-transition countries (Asia, South America, and Africa) was about 63 % (**Figure 2,** bottom), which is again a slight increase from 2017 with 57 %. The number of active users remained nearly the same with 612 in 2018 compared to 620 in 2017. The average number of file downloads per user in 2018 was 1 300, which is similar to that in 2016 and 60 % of that in 2017 (**Figure 2,** top), which is dominated by the highly active European users with a mean file download number per user of ca. 2 500 files. 40 % 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 2019.

Assuming the same user distribution for the ESGF users as the DKRZ users, 62 % of the files were downloaded by European, 27 % by Asian, 11 % by North American and less than 1 % from African, South American, and Australian users. The share of downloads from users from developing and economy-in-transition countries (Asia, Africa, and South America) is with 27 % the same as the share of the Asian users.

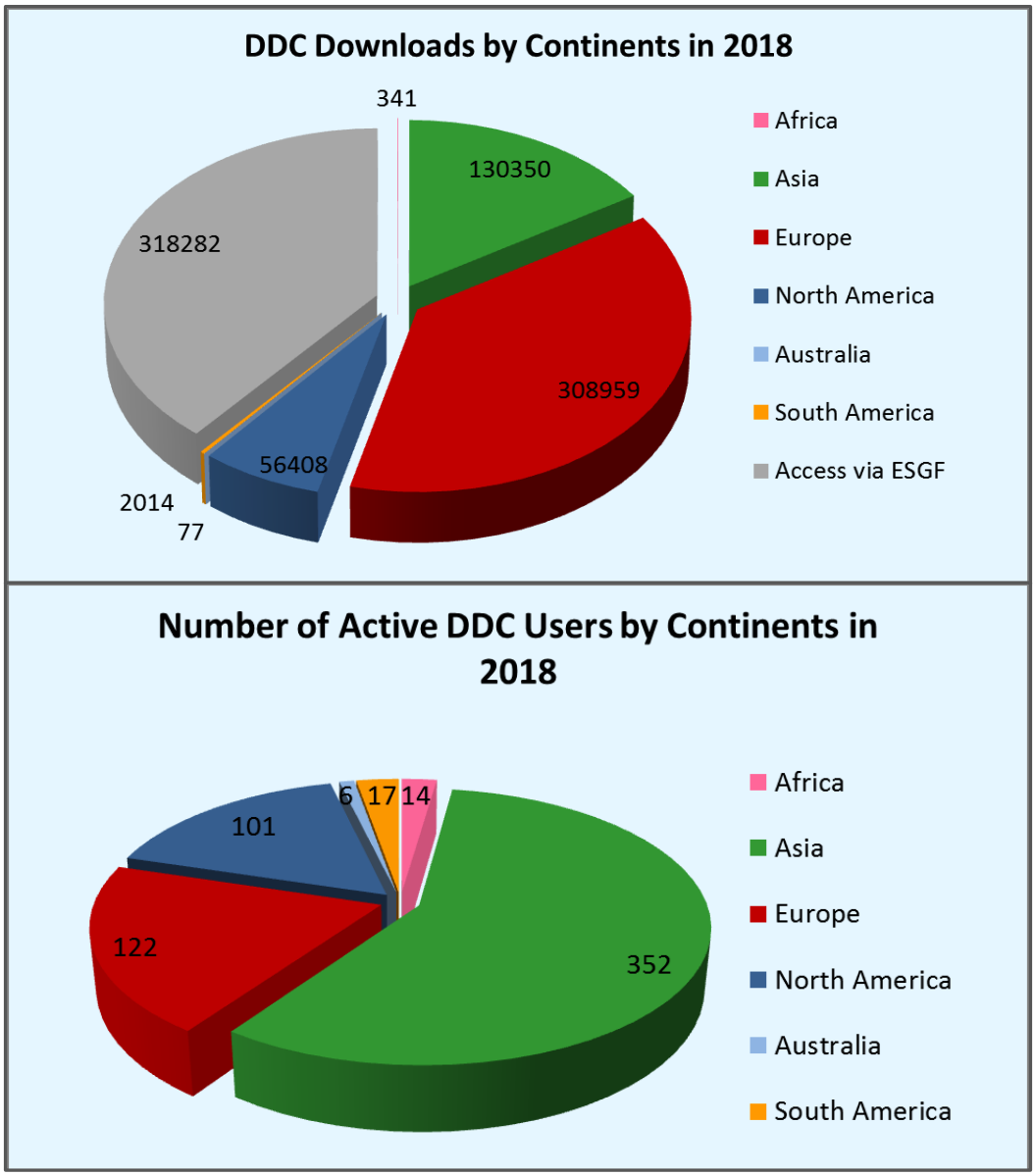

***Figure 2**: Number of active DDC users in 2018 (bottom) and downloads counts of users per continent (top).*

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 from 23 USB devices sent to 13 users in 2017 to 9 USB devices sent to 7 users. Most of the users in 2018 were interested in data from only one geographical area, whereas in 2017 several users were interested in multiple areas. No user was interested in AR4 data on storage media in 2018.

About half of the requestors (4) were located in developing or economy-in-transition countries and the other half (3) in Europe (see **Table 1** and **Table 2**). 2/3 of the USB storage media (6) were sent to requestors in developing or economy-in-transition countries and 1/3 to European users (3).

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.

Table 1: *Number of storage media requests per data area for AR4 and AR5 in 2018.*

Area of Data	No of storage media	Africa	Asia	Australia / Central Pacific	Europe	North America	South America	North Pole	South Pole
AR4	-	-	-	-	-	-	-	-	-
AR5	9	3	2	1	2	-	1	-	-
DDC total	**9**	**3**	**2**	**1**	**2**	**-**	**1**	**-**	**-**

Table 2: *Number of user requests for AR4/AR5 data on storage media per user origin (continent) in 2018.*

User origin	No of users	African users	Asian users	South American users	Australia / Central Pacific	European users
AR4	-	-	-	-	-	-
AR5	7	1	2	1	-	3
DDC total	**7**	**1**	**2**	**1**	**-**	**3**

4. Data access by category AR

The monthly download rates in 2018 from the IPCC DDC reference archive were dominated by AR5 downloads as in the previous years. The tendency is an unchanged high interest in AR5 data but a declining interest in AR4 data (**Figure 3**; online monthly download statistics[1]). AR4 data download deceased by 80 % in number and to about 1/5 in volume between 2017 and 2018, continuing the trend of the previous years. Downloads of SAR and TAR data in 2018 remain in the same order as for the previous years.

In 2018 IPCC DDC users downloaded ca. 0.9 PBytes of data in ca. 820 000 individual file downloads, which is about 25 % of the volume and 60 % of the downloaded files in 2017. Neglecting the 05/2017 peak in download volume the downloaded volume in 2018 is 65 % of that in 2017. The tendency for the download of AR5 data files is a decrease by 42 % compared to 2017 and 24 % compared to 2016. The reason is unknown. For the AR5 data download volume excluding the unreliably high volume for 05/2017, the decrease from 2017 to 2018 is ca. 35 %.

[1] 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

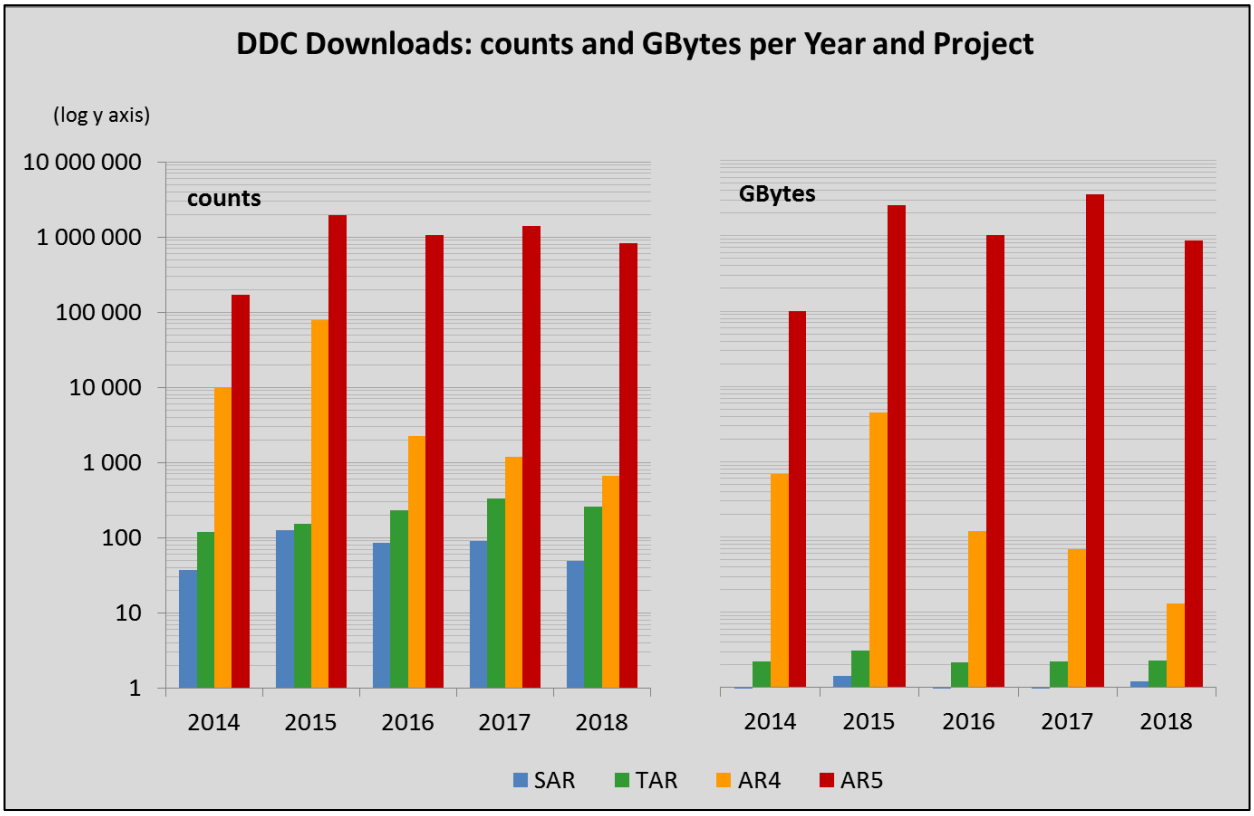

Figure 3: Total annual data download counts (left) and volumes in GBytes (right) over the last five years for the different DDC reference archives (without FAR).

5. Review of user queries

There are no numbers for the handled user requests by WDCC/DKRZ staff available for 2018, same as for 2016 and 2017. A separation of user requests on IPCC DDC issues is not possible.

As the IS-ENES support activity has ended, no ESGF user statistics are available for 2018.

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