CMIP6:

WIP Status Update

A summary of CMIP6 to-date from the WGCM Infrastructure Panel (WIP)

WGCM-24
Thursday 9th December 2021 - Virtual

Paul J. Durack, Matthew Mizielinski
and the WIP membership

Lawrence Livermore
National Laboratory

CMIP: a team sport
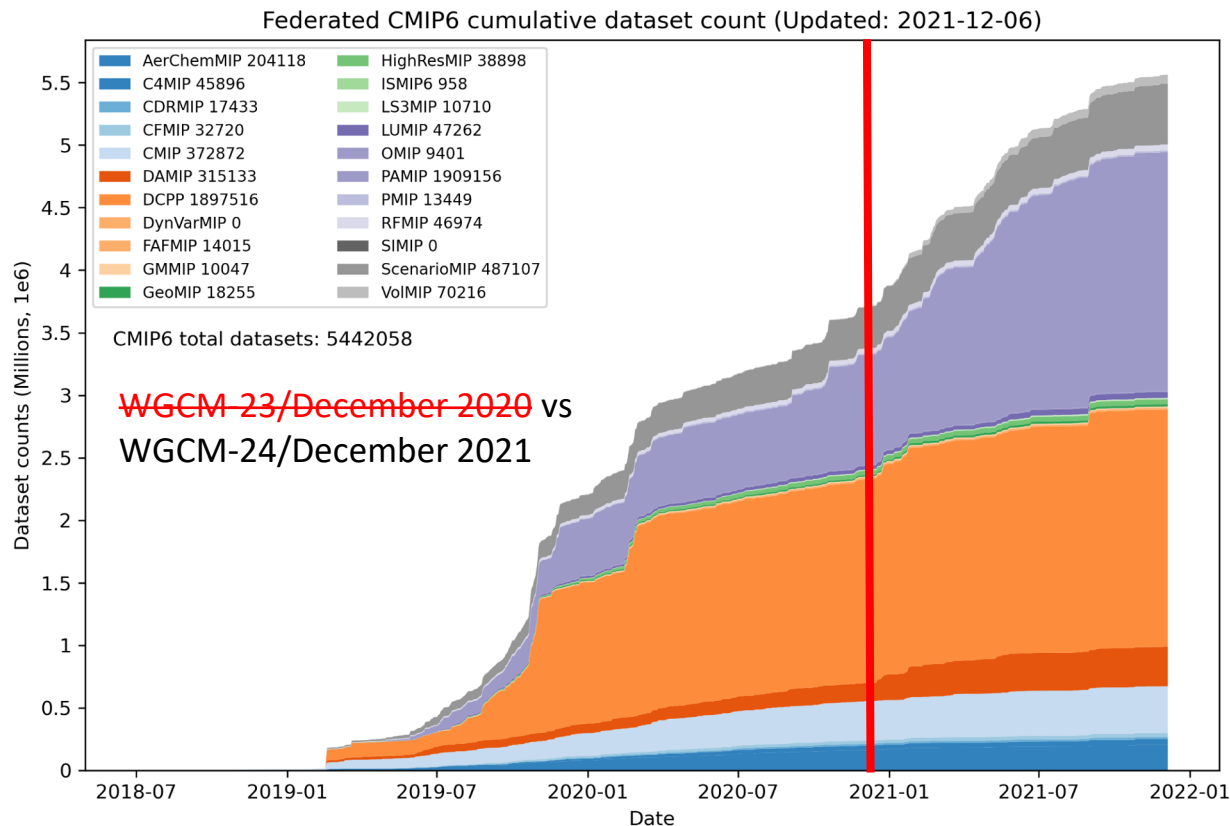
- PCMDI
  - U.S. DOE has provided 32-years of *MIP support
- ESGF
  - Originated by U.S. DOE
  - Major recent contributions from numerous others
- IS-ENES
  - European contribution to ESGF & CMIP infrastructure
- Numerous other projects and institutions, including DKRZ, IPSL, CEDA, ES-DOC, NASA, NOAA, …
- 29+ ESGF nodes
- 140 models, 52 institutions representing 26 countries, and many, many more…
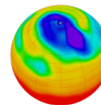
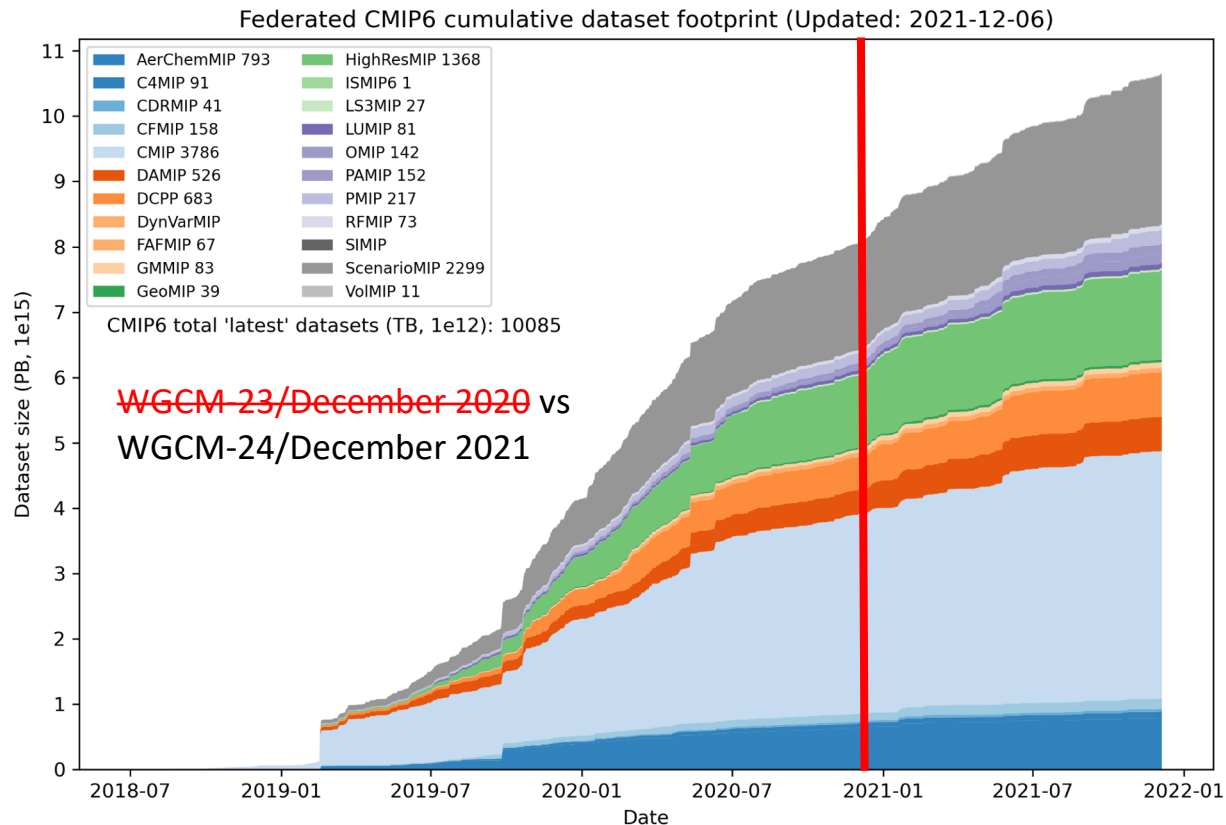\+ **Every modelling group, every forcing dataset provider, …**

ESGF Published data

- Over ~~3.7~~ 5.6 million datasets on ESGF across all CMIP6 activities/MIPs

- Delivery has been seamless – thanks to data challenges and ESGF stability testing

- Datasets - unique variable collections per experiment RIPF

Federated CMIP6 cumulative dataset count (Updated: 2021-12-06)

AerChemMIP 204118	HighResMIP 38898
C4MIP 45896	ISMIP6 958
CDRMIP 17433	LS3MIP 10710
CFMIP 32720	LUMIP 47262
CMIP 372872	OMIP 9401
DAMIP 315133	PAMIP 1909156
DCPP 1897516	PMIP 13449
DynVarMIP 0	RFMIP 46974
FAFMIP 14015	SIMIP 0
GMMIP 10047	ScenarioMIP 487107
GeoMIP 18255	VolMIP 70216

CMIP6 total datasets: 5442058

~~WGCM-23/December 2020~~ vs WGCM-24/December 2021

ESGF Published data

- Over ~~3.7~~ 5.6 million datasets on ESGF across all CMIP6 activities/MIPs

- Delivery has been seamless – thanks to data challenges and ESGF stability testing

- Datasets - unique variable collections per experiment RIPF

- Footprint – storage units in PBs

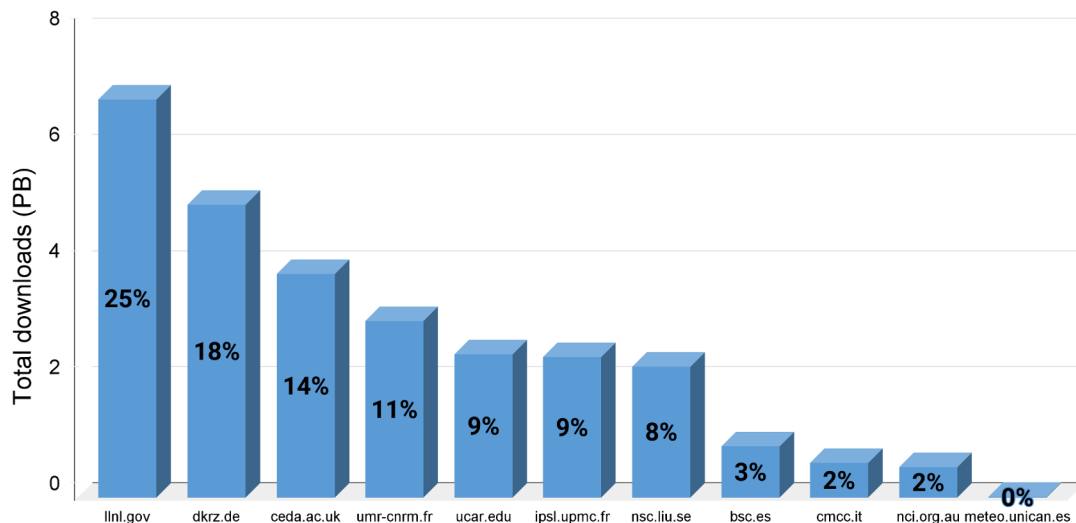Federated CMIP6 cumulative dataset footprint (Updated: 2021-12-06)

Legend:
- AerChemMIP 793
- C4MIP 91
- CDRMIP 41
- CFMIP 158
- CMIP 3786
- DAMIP 526
- DCPP 683
- DynVarMIP
- FAFMIP 67
- GMMIP 83
- GeoMIP 39
- HighResMIP 1368
- ISMIP6 1
- LS3MIP 27
- LUMIP 81
- OMIP 142
- PAMIP 152
- PMIP 217
- RFMIP 73
- SIMIP
- ScenarioMIP 2299
- VolMIP 11

CMIP6 total 'latest' datasets (TB, 1e12): 10085

~~WGCM-23/December 2020~~ vs WGCM-24/December 2021

ESGF publication and replication

- ~~17.7~~ 21.9 PB CMIP6 data available including ~~10~~ ~12 PB of unique and ~~8~~ ~10 PB of replicated data
- ~~16.2~~ 27.9 PB CMIP6 downloads (to November 2021)
- LLNL ~~27~~ 25% downloads to date
- DKRZ ~~20~~ 18%
- CEDA ~~12~~ 14%
- CNRM ~~15~~ 11%
- UCAR 9%
- IPSL ~~13~~ 9%

~~WGCM-23/December 2020~~ vs WGCM-24/December 2021
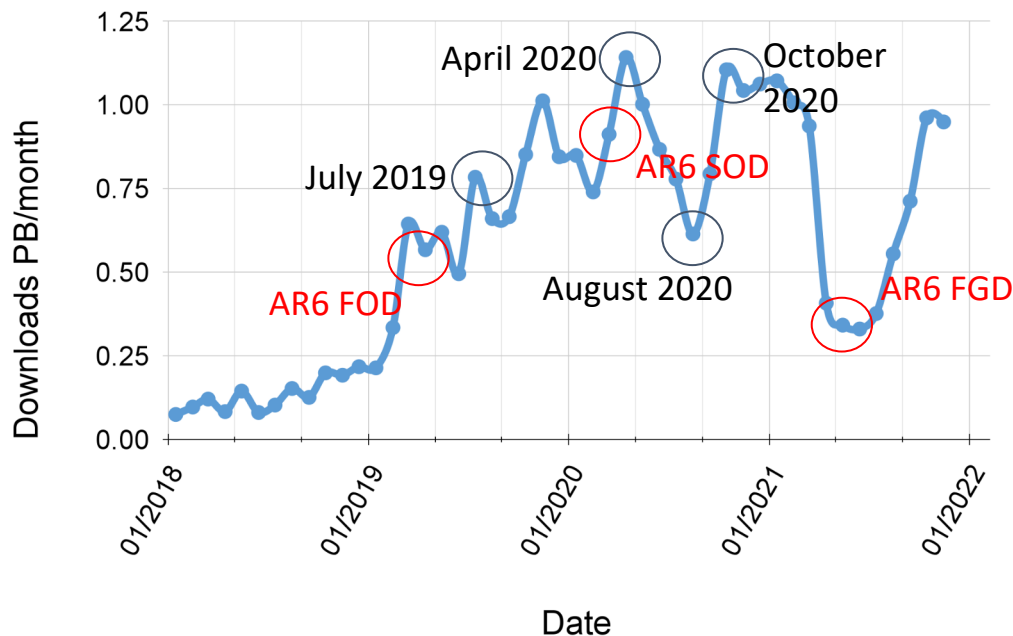
CMIP6 total downloads by ESGF node

Total downloads (PB) vs Distinct ESGF node

- llnl.gov — 25%
- dkrz.de — 18%
- ceda.ac.uk — 14%
- umr-cnrm.fr — 11%
- ucar.edu — 9%
- ipsl.upmc.fr — 9%
- nsc.liu.se — 8%
- bsc.es — 3%
- cmcc.it — 2%
- nci.org.au — 2%
- meteo.unican.es — 0%

http://esgf-ui.cmcc.it/esgf-dashboard-ui/federated-view.html - 29 nodes in total, 11 reporting statistics

Downloaded data

- Over ~28 PB of recorded downloads (to November 2021)

- Use of local (JASMIN/DKRZ/..) or "dark" archives (e.g. Pangeo) means data usage is significantly larger

- Interesting to see COVID/AR6 activity impacts

CMIP6 total downloads by month

CMIP6 controlled vocabulary

- ~~137~~ 140 models registered with CMIP6 CVs
- Each model involved in 6* activities on average
- Experiments grown from ~280 to 322 including six recent "CovidMIP" experiments added to DAMIP
- Added CDRMIP and PAMIP in March 2018
- Added "COVIDMIP" in November 2020

https://github.com/WCRP-CMIP/CMIP6_CVs

CMIP6_CVs registered models - intentions to contribute

Registered models - intentions (y-axis)

MIP/activity id (x-axis): CMIP 107, ScenarioMIP 64, HighResMIP 49, OMIP 47, RFMIP 37, PMIP 31, DAMIP 27, GMMIP 26, C4MIP 26, AerChemMIP 25, CFMIP 24, LUMIP 22, DCPP 22, LS3MIP 21, FAFMIP 19, PAMIP 17, VolMIP 16, GeoMIP 15, CDRMIP 14, ISMIP6 9
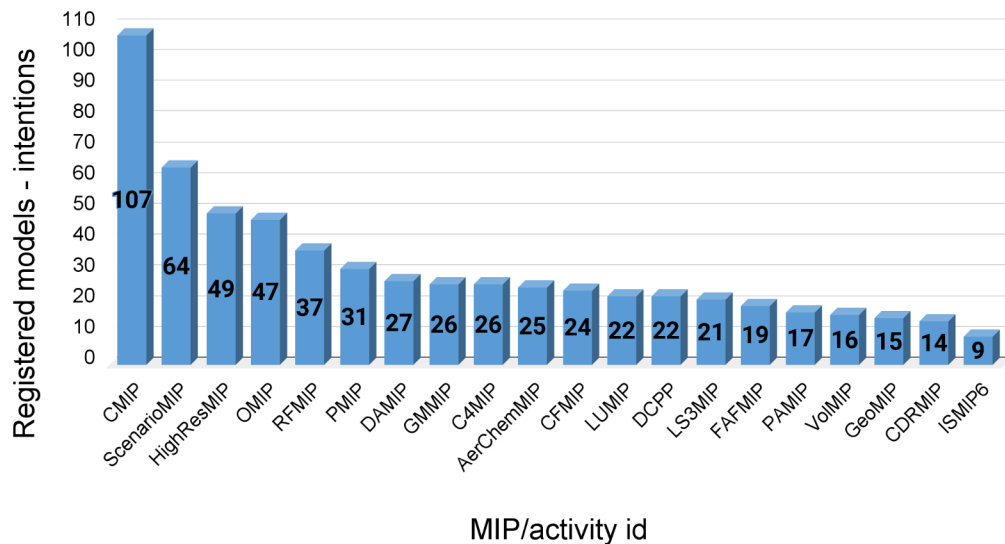
https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_experiment_id.html
https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_institution_id.html
https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_source_id.html

CMIP6 controlled vocabulary

- ~~137~~ 140 models registered with CMIP6 CVs
- Each model involved in 6* activities on average
- Experiments grown from ~280 to 322 including six recent "CovidMIP" experiments added to DAMIP
- Added CDRMIP and PAMIP in March 2018
- Added "COVIDMIP" in November 2020

https://github.com/WCRP-CMIP/CMIP6_CVs

CMIP6_CVs registered models - intentions to contribute

CMIP6 ~68% complete
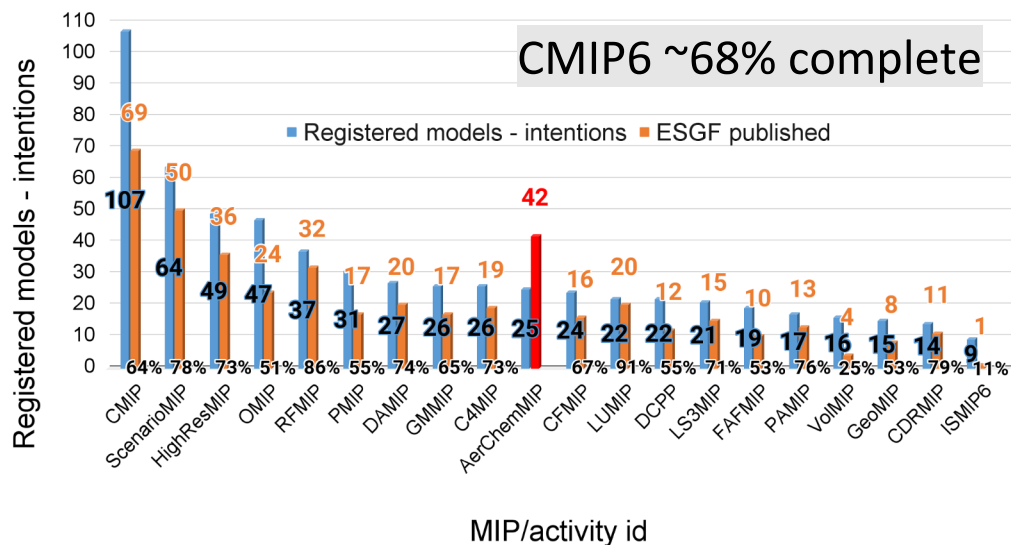
https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_experiment_id.html
https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_institution_id.html
https://wcrp-cmip.github.io/CMIP6_CVs/docs/CMIP6_source_id.html

Paul J. Durack
durack1@llnl.gov

Matthew Mizielinski
matthew.mizielinski@metoffice.gov.uk

Lawrence Livermore National Laboratory