On Machines and CPMIP Metrics Underlying CMIP Simulations

Sadie Bartholomew
University of Reading and National Centre for Atmospheric Science
On behalf of the wider ES-DOC team

MS389 Advances in High Performance Computing for Earth Science (Part II of II)
SIAM Conference on Computational Science and Engineering (CSE23)
03.03.23

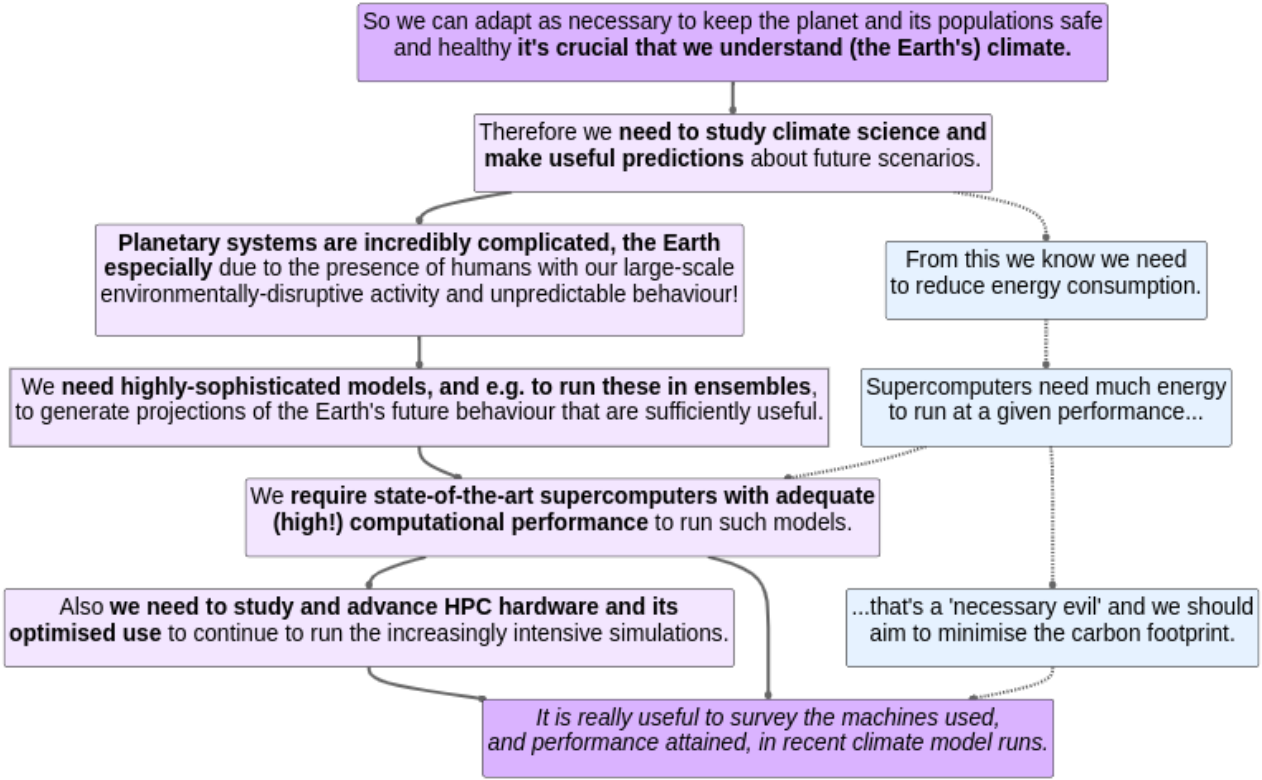
es-doc
Earth System Documentation

is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

University of Reading

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Context: the big picture

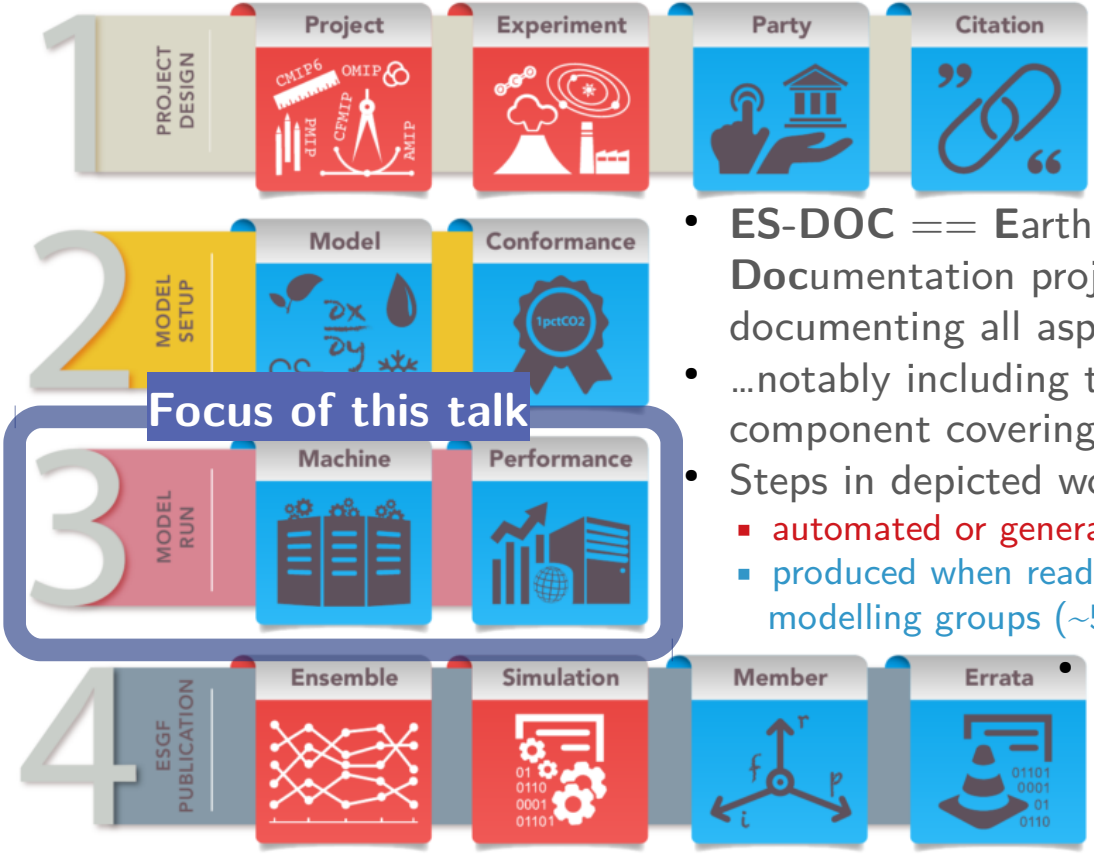


Role of CMIP(6)

- **CMIP** == WCRP (World Climate Research Programme) **C**oupled 'MIP' (**M**odel **I**ntercomparison **P**roject)
- In the words of WCRP (from www.wcrp-climate.org/wgcm-cmip):

> *CMIP provides climate projections that support essential WCRP activities and climate science worldwide, decision and policy-makers communities, in its objective to understand past, present and future climate changes. CMIP and its associated data infrastructure have become essential to the Intergovernmental Panel on Climate Change (IPCC) and other international and national climate assessments.*

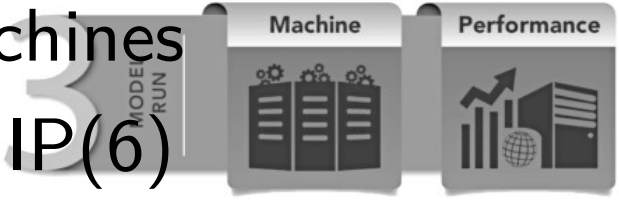
- **CMIP6** is the latest phase of CMIP and is now winding up, with planning for CMIP7 underway.
- Therefore the climate science community is at a critical period to consider lessons learnt from CMIP6 to ensure CMIP7 is as successful as possible.



Role of ES-DOC

- **ES-DOC** == **E**arth **S**ystem **Doc**umentation project, tasked with documenting all aspects of CMIP6...
- ...notably including the 'model run' component covering platform aspects
- Steps in depicted workflow either:
  - automated or generated by ES-DOC
  - produced when ready by the modelling groups (~50)
- ES-DOC builds on previous work of other projects for CMIP ≤5

Characterising HPC machines and performance in CMIP(6)



- Classification according to now-mature **CIM** (== **C**ommon **I**nformation **M**odel) canonical standard, current version 2.x:
  - All open source, stored and developed on GitHub at github.com/ES-DOC/esdoc-cim-v2-schema
  - Machine and performance contained within 'platform' classes: see <link above>/blob/master/platform_classes.py for relevant schema

```
116  def machine():
117      """A computer/system/platform/machine which is used for
118      simulation."""
119      return {
120          "type": "class",
121          "base": "platform.partition",
122          "is_abstract": False,
123          "pstr": ("{}", ("name",)),
124          "is_document": True,
125          "properties": [
126              (
127                  "peak_performance",
128                  "shared.numeric",
129                  "0.1",
130                  "Total peak performance (RPeak in Top500 lingo)",
131              ),
132              (
133                  "linpack_performance",
134                  "shared.numeric",
135                  "0.1",
136                  "Linpack performance (RMax in Top500 lingo)",
137              ),
138          ],
139      }
```

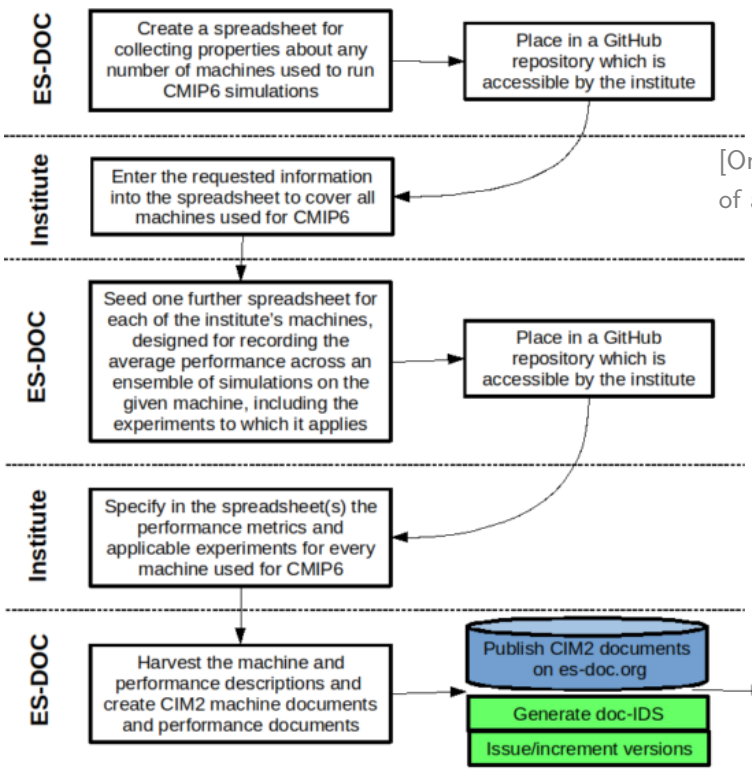
...

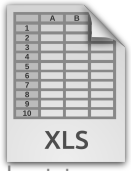
- Platform properties in CIM based on established **CPMIP** (== **C**omputational **P**erformance **MIP**) **metrics**
  - See '*CPMIP: measurements of real computational performance of Earth system models in CMIP6*', V. Balaji et al. (2017), doi.org/10.5194/gmd-10-19-2017
  - These metrics cover aspects such as model e.g. resolution; platform e.g. clock speed; computational cost e.g. Joules per simulated year; and coupling, memory and I/O considerations such as memory bloat and data intensity.

Similar/parallel work to note

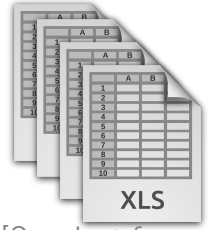
- Also under the project IS-ENES3, Mario Acosta and collaborators evaluated and studied CPMIP metrics for CMIP6 model runs, though for only a subset (~10) of the modelling groups in CMIP6 (~50) that ES-DOC had the potential (with sufficient group engagement) to capture.
  - See: '*IS-ENES3 D4.3: CPMIP performance metrics evaluation for CMIP6 and community advice*', Mario Acosta et al. (2021), doi.org/10.5281/zenodo.6394049

Documentation collection and publication





[One sheet to cover all of a group's machines]



[One sheet for each machine and applicable model combination]

- For these platform aspects, ES-DOC designed and rolled out a two-stage process to gather, via dedicated spreadsheets, first the information on the machines, then the average representative performance of each across the ensembles.
- For further info. see: es-doc.org/cmip6-machine-and-performance/

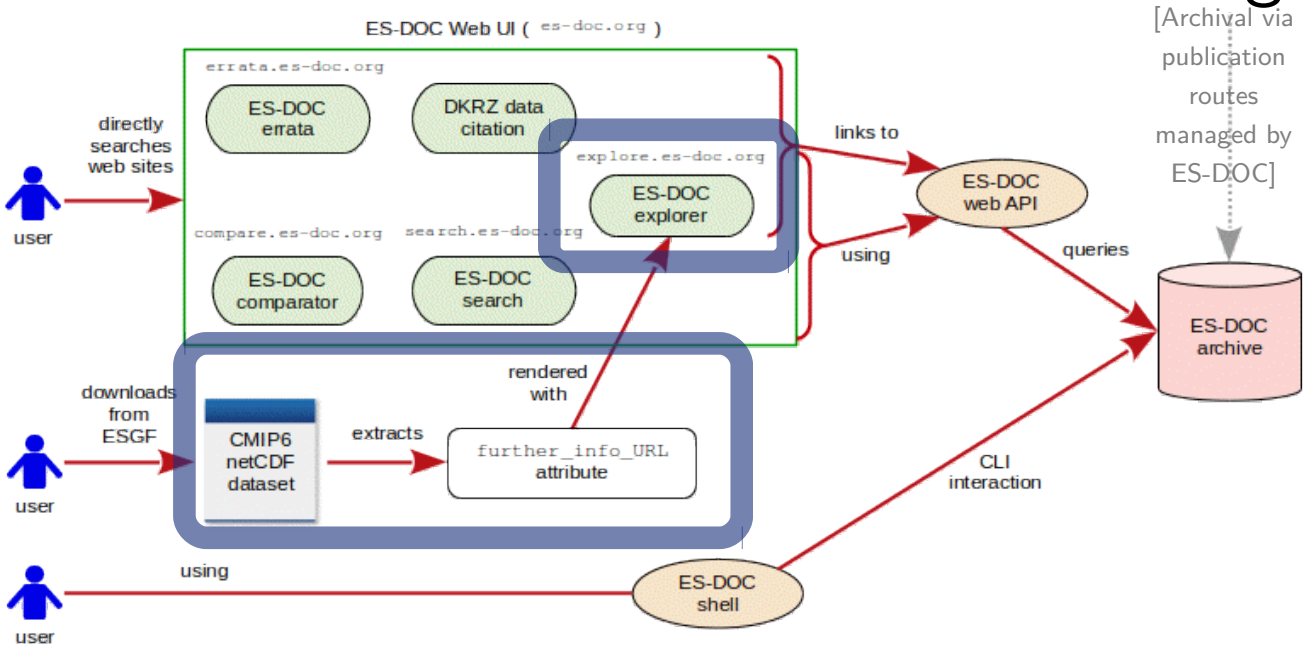
es-doc Earth System Documentation

is-enes INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

University of Reading

National Centre for Atmospheric Science NATURAL ENVIRONMENT RESEARCH COUNCIL

Archived documentation access + linkage



- As well as collecting the information, ES-DOC archives it and makes it openly available as published documentation, notably through a dedicated website (root [es-doc.org](es-doc.org)) e.g. via the 'explorer' (explore.es-doc.org).

Submitted results, 1. machines (hardware)

- All documented machines, but only a subset of the collected machine information for each, shown, where a dot indicates no value was provided:

Modelling group (acronym or identifier)	Group machine(s)	Overall		Example compute pool (first one documented, if multiple) information					Example storage pool (first one documented, if multiple) information		Interconnect	Benchmark performances (all Tflops/s)	
		OS	Vendor	Model number	Number of nodes	Memory per node (GiB)	CPU cores per node	Clock speed (GHz)	File system size (TB)	Storage type	Topology	Peak	LINPACK
EC-EARTH-CONSORTIUM	Beskow	Cray Linux Environment	Cray	.	2060	64	32	2.3	.	.	.	2438.14	1802.51
	Teralith	CentOS 7.7	.	.	1908	96	32	2.1	.	.	.	4335.21	2969.19
	Rhino	Red Hat Enterprise Linux Server 7.3 (Maipo)	.	.	180	.	28	3.07	.	.	.	.	.
	Marenostrum 4	SUSE Linux Enterprise Server 12 SP2	Lenovo	.	3240	96	48	2.1	.	IBM GPFS	.	5765.9	6227.2
CCCR-IITM	INTEL AADITYA	RHEL v6.x	.	.	2384	60	16	.	.	.	.	.	.
CMCC	ATHENA	CentOS Linux release 6.2	IBM	DX 360M4	482	64	16	2.6	300	IBM GPFS	Spine - Leaf	160	.
	ZEUS	CentOS Linux release 7.6	Lenovo	SD530	348	96	36	3	4000	IBM GPFS	Spine - Leaf	1202	.
DKRZ	Mistral	.	.	E5-2680v3 12C	1550	.	2	2.5	54 000	Lustre parallel file system	.	3600	.
INPE	xc50	Cray Linux Environment	Cray XC-50	.	96	192	40	2.4	.	Lustre parallel file system	.	.	.
MPI-M	Mistral	.	.	E5-2680v3 12C	1550	.	2	2.5	54 000	Lustre parallel file system	.	3600	.
CNRM-CERFACS	beaufix2	bullx SCS	.	.	1836	64	40	2.2	.	Lustre parallel file system	.	.	.
DWD	(Unnamed)	Cray Linux Environment	Cray	E5-2680v3	432	128	24	2.5	7204	.	Dragonfly	1459.81	1214.2
MOHC	xce	Cray Linux Environment	Cray	.	12932	192	36	2.1	.	Lustre parallel file system	.	15 400	13 800
	xcf	Cray Linux Environment	Cray	.	12932	192	36	2.1	.	Lustre parallel file system	.	15 400	13 800
	xcs-r	Cray Linux Environment	Cray	.	12932	192	36	2.1	.	Lustre parallel file system	.	15 400	13 800
NERC	Archer	Cray Linux Environment	Cray	.	4920	64	24	2.7	.	Lustre parallel file system	.	2550.53	1642.54
IPSL	Curie	Linux	.	Sandy Bridge	5040	64	8	2.26	<1	Lustre parallel file system	QDR Full Fat Tree	1667.2	1359
	Joliot-Curie	Linux	.	Skylake (Intel Xeon Platinum 8168 CPU)	1656	192	48	2.7	<1	Lustre parallel file system	Multiple	12039.4	6988
MIROC	Earth Simulator (ES3)	SUPER-UX	NEC	.	5120	64	4	1	13 500	Other	2-level Fat-tree	1310.72	n/a
	Earth Simulator (ES4)	CentOS 8	NEC	B401-8 Type 20B	684	128	64	3.4	60 000	Lustre parallel file system	Dragonfly+	13448	9990.7
NOAA-GDFL	Gaea C3	Cray Linux Environment 7	Cray	Cray XC40-LC Haswell	1504	64	32	2.3	32 000	Lustre parallel file system	Dragonfly	1770	1280
	Gaea C4	Cray Linux Environment 7	Cray	Cray XC40-LC Broadwell	2656	64	36	2.1	32 000	Lustre parallel file system	Dragonfly	2982	3213
NASA-GISS	DISCOVER	SuSE Linux Enterprise Server	.	.	1824	128	28	.	10 000	Disk	.	7000	.

… 2. computational performance

- Very sadly, groups are (yet?) to submit performance documentation other than that which was submitted to Mario Acosta and collaborators for their IS-ENES3 work (see slide 6 reference)
- Therefore let's delegate to their work and advertise their analysis! **Figure to right attributed to, and credit to, Mario Acosta et al. (2021)** (full reference on slide 6).

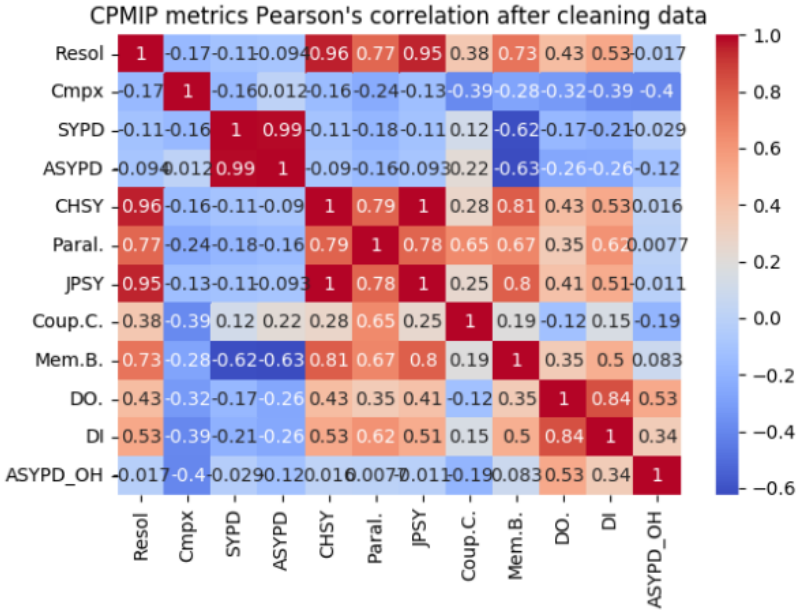


Figure 2. Pearson's correlation among all metrics after cleaning the data.

Challenges and looking ahead to CMIP7

- Core challenge: engagement has been very low, with only 16/49 groups submitting any platform documentation ☹
  - Though strictly modelling groups are obliged to submit full documentation for their CMIP6 contribution, many haven't at all, most not fully. It seems they feel they are too busy otherwise to do this and there isn't adequate reward for doing it or punishment for not doing it that ES-DOC can provide as persuasion. That said, ES-DOC did start the collection late in the CMIP6 timeline (not ideal).
  - In summary: *computers will do what they are asked to do (for better or for worse…), people might not!*
- I fully agree with recommendations for actions concluded by Mario Acosta and team in the 'IS-ENES3 D4.3 …' report (see reference slide 6):

> *Collect CPMIP metrics before or during the CMIP experiment: Spend some resources before the CMIP experiments or at least during the spinup/tuning process… ~ … The development of portable and automatic processes such as the integration with workflow managers could be a solution … ~ … normalize the way to collect some specific metrics …*

Summary

- Ultimately it is really useful to survey the HPC machines used, and performance attained, in recent climate model runs...
- ... so here we consider the hardware and computational performance across CMIP6 (last stage, ending), looking ahead to CMIP7 (next stage).
- Machine and performance were characterised via CIM data model which notably use CPMIP performance metrics.
- ES-DOC collects & publishes documentation for the full CMIP6 workflow:
  - For machine and performance, had a two-stage collection process with modelling groups submitting CIM platform information via spreadsheets processed by ES-DOC infrastructure.
  - Documents connected to relevant ensemble; linked to other aspects via `further_info_URL` concept; and accessible via website (root es-doc.org) such as the 'explorer' (explore.es-doc.org).
- Sadly, low engagement meant limited results submitted so for CMIP7 we also recommend effort to enable (some) auto-generation of such data.
- (But) there is still useful information found in/across the limited results.