

Dealing with and Contributing to the CMIP5 Data “Tsunami” and Beyond from an Australian Perspective

www.cawcr.gov.au



Mark Collier¹, Stephen Jeffrey², Martin Dix¹, Tony Hirst¹ and Leon Rotstajn¹

¹ CSIRO Marine and Atmospheric Research

² Queensland Climate Change Centre of Excellence

8 April 2011

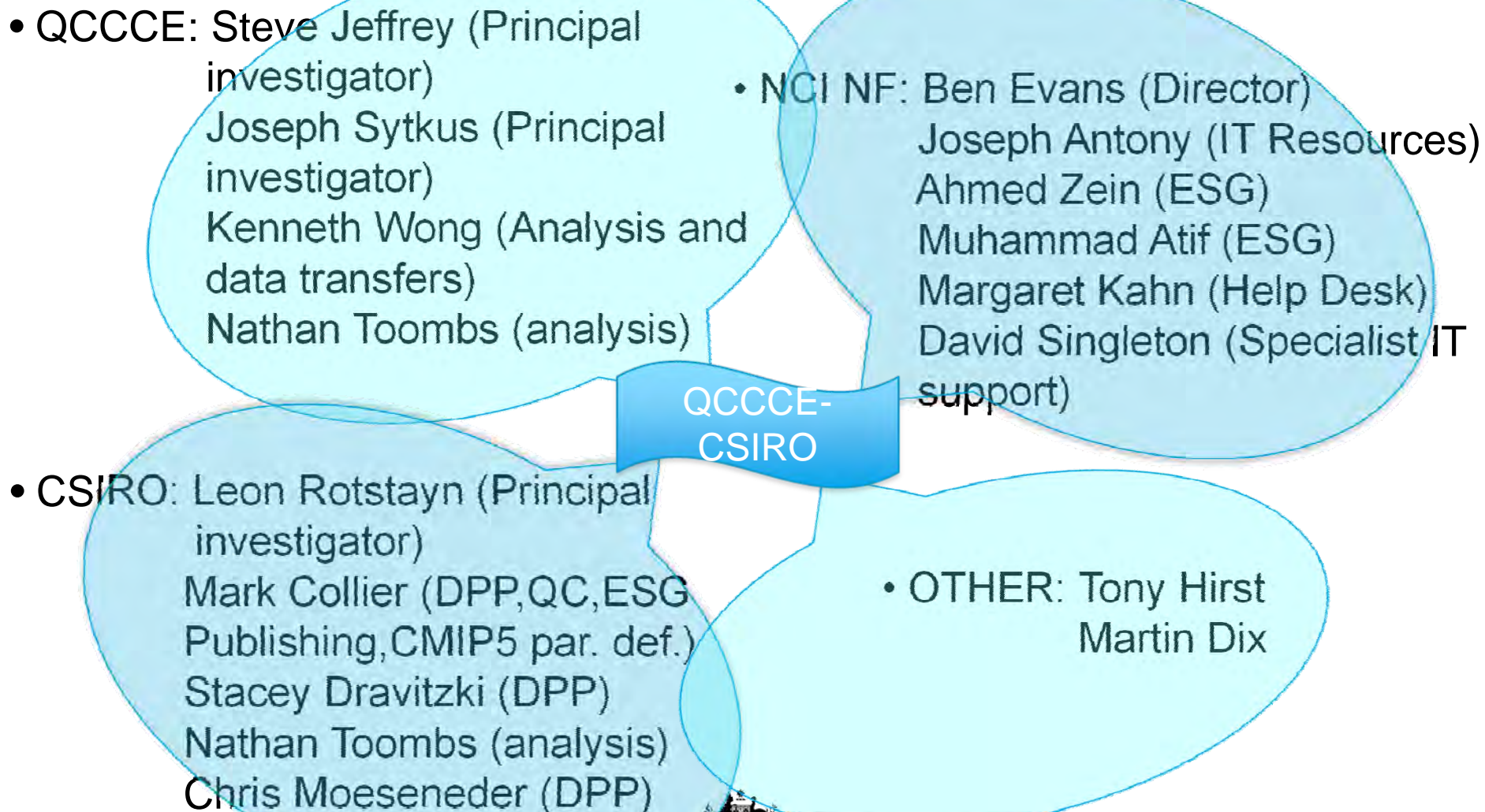


Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology



The CSIRO-Mk3-6-0 Team



CMIP5 CSIRO-Mk3-6-0 Workflow



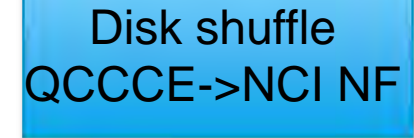
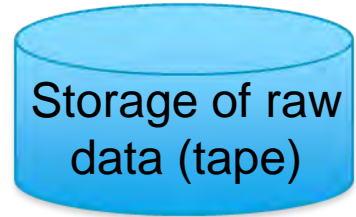
Integrate

Local

(QCCCE)

CSIRO-Mk3-6-0

January 2010



June 2010

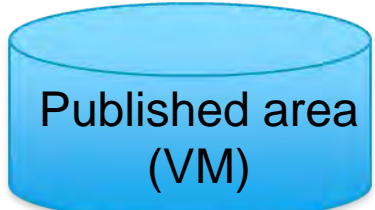


April 2011?



Replication

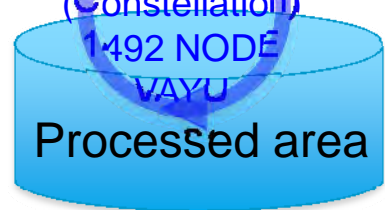
Remote Storage/Staging



Mar./Apr. 2011



March 2011



October 2010



Guiding Documents - CMIP5 “bibles”

There are 3 essential documents that we “religiously” studied:

- A Summary of the CMIP5 Experiment Design – 22 Jan 2011
- Standard Output (Requested Variables) – 23 Mar 2011
- CMIP5 Data Reference Syntax (DRS) and Controlled Vocabulary – 9 Mar 2011

and 3 others of secondary importance

- PCMDI’s CMIP web-site: <http://cmip-pcmdi.llnl.gov/cmip5/>
- CMIP5 Model Output Format and Metadata Requirements – 7 Jan 2011 (guidelines for CMOR)

- CMOR – Climate Model Output Rewriter Software

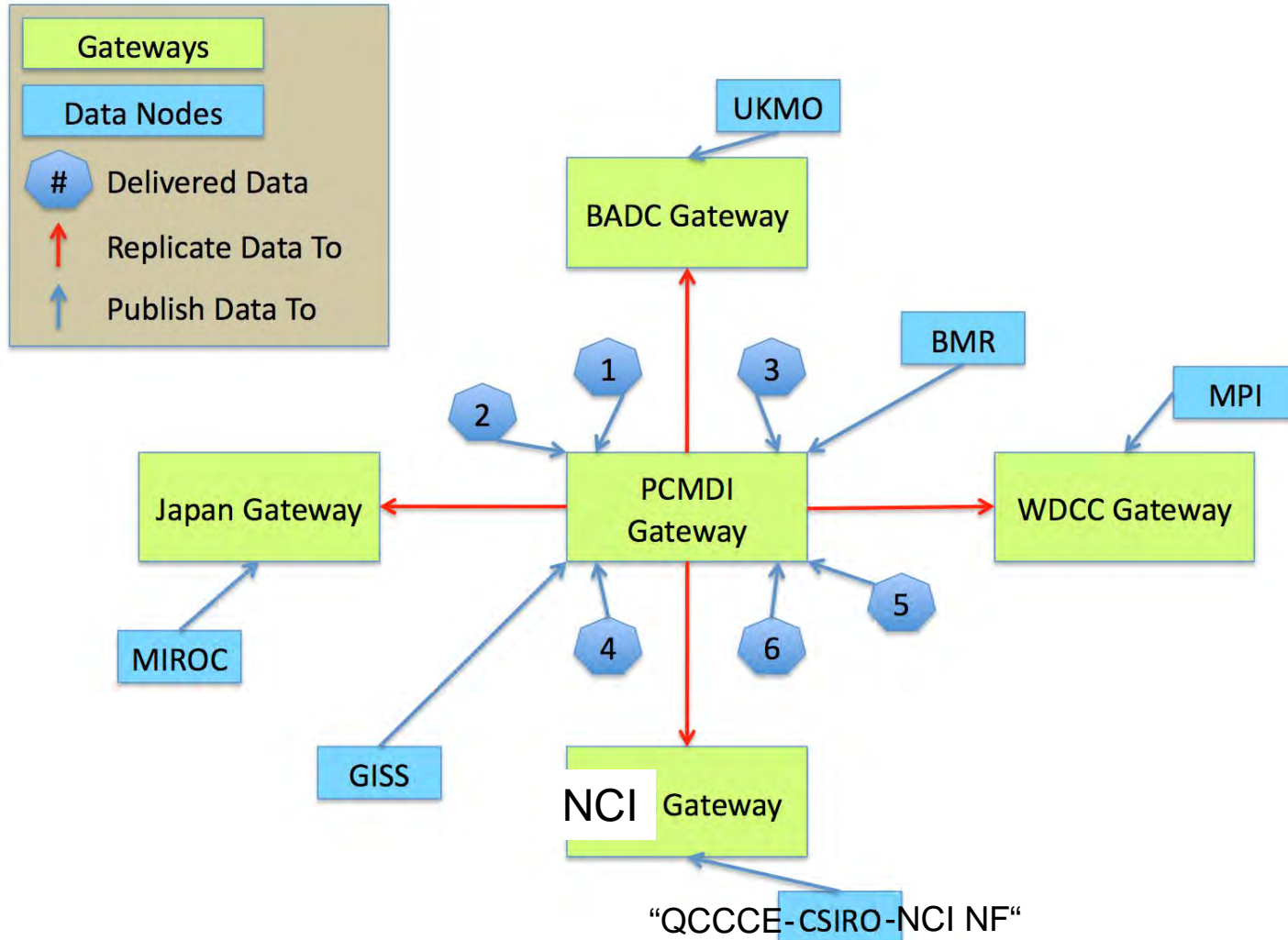
GO-ESSP



- It comprises a group of 40-50 people from climate and weather institutions from around the world, and includes the 2 primary groups Metafor and Earth System Curator
- A collaboration to develop new generation of software infrastructure that will provide distributed access to data, like CMIP5
- Individual software components and a federation of framework is to be continually developed, for example, ESG and the CMIP5 Multi Model Questionnaire
- Major partners are BADC, GFDL, GCCC, LLNL, NCAR and PMEL
- An important part of GO-ESSP is to develop the ESG and metadata conventions (including CF)



ESG Federation





- N
- dis
- T
- BA
- M
- Qu
- the
- T
- cor
- C
- Off
- Mathematics)




ESG Gateway hosted by the Program for Climate Model Diagnosis and Intercomparison

Search: for:

To conduct a search, select a category from the pull down menu and/or enter free text into the the text box.

- ### Search Categories
- Project
 - > CMIP5
 - > ReanalysisCMIP5
 - > test
 - + Institute
 - + Model
 - + Experiment
 - + Frequency
 - + Product
 - + Realm
 - + Variable
 - + Ensemble

Welcome to PCMDI



The Program for Climate Model Diagnosis and Intercomparison (PCMDI) was established in 1989 at the Lawrence Livermore National Laboratory (LLNL), located in the San Francisco Bay area. Our staff includes research scientists, computer scientists, and diverse support personnel.

The PCMDI mission is to develop improved methods and tools for the diagnosis and intercomparison of general circulation models (GCMs) that simulate the global climate. The need for innovative analysis of GCM climate simulations is apparent, as increasingly more complex models are developed, while the disagreements among these simulations and relative to climate observations remain significant and poorly understood. The nature and causes of these disagreements must be accounted for in a systematic fashion in order to confidently use GCMs for simulation of putative global climate change.

- ### Quick Links
- [Getting Started Guide](#)
 - [Create Account](#)
 - [Browse Catalogs](#)
 - [Search for Data](#)
- ### ESG Data Gateways
- [BADC Gateway](#)
 - [DKRZ Gateway](#)
 - [NASA JPL Gateway](#)
 - [NCAR Gateway](#)
 - [NCI Gateway](#)
 - [NERSC Gateway](#)
 - [PCMDI Gateway](#)
 - [ORNL Gateway](#)

Status of the CMIP5 Archive

3/23/2011: HadGEM2-ES (UK Met Office / Hadley Centre) datasets are available from BADC.
3/23/2011: inmcm4 datasets are now available from the PCMDI server.

data
of
ed

CMIP5 Data Reference Syntax (DRS) and Controlled Vocabularies



- A common naming system to be used in files, directories, metadata, and URLs
- To identify datasets wherever they might be located within the distributed CMIP5 Federation
- It defines controlled vocabularies for many of the components comprising the Data Reference Syntax
- Note that there been a disconnect b/w the CMOR (data generators) and ESG (data publishers) communities



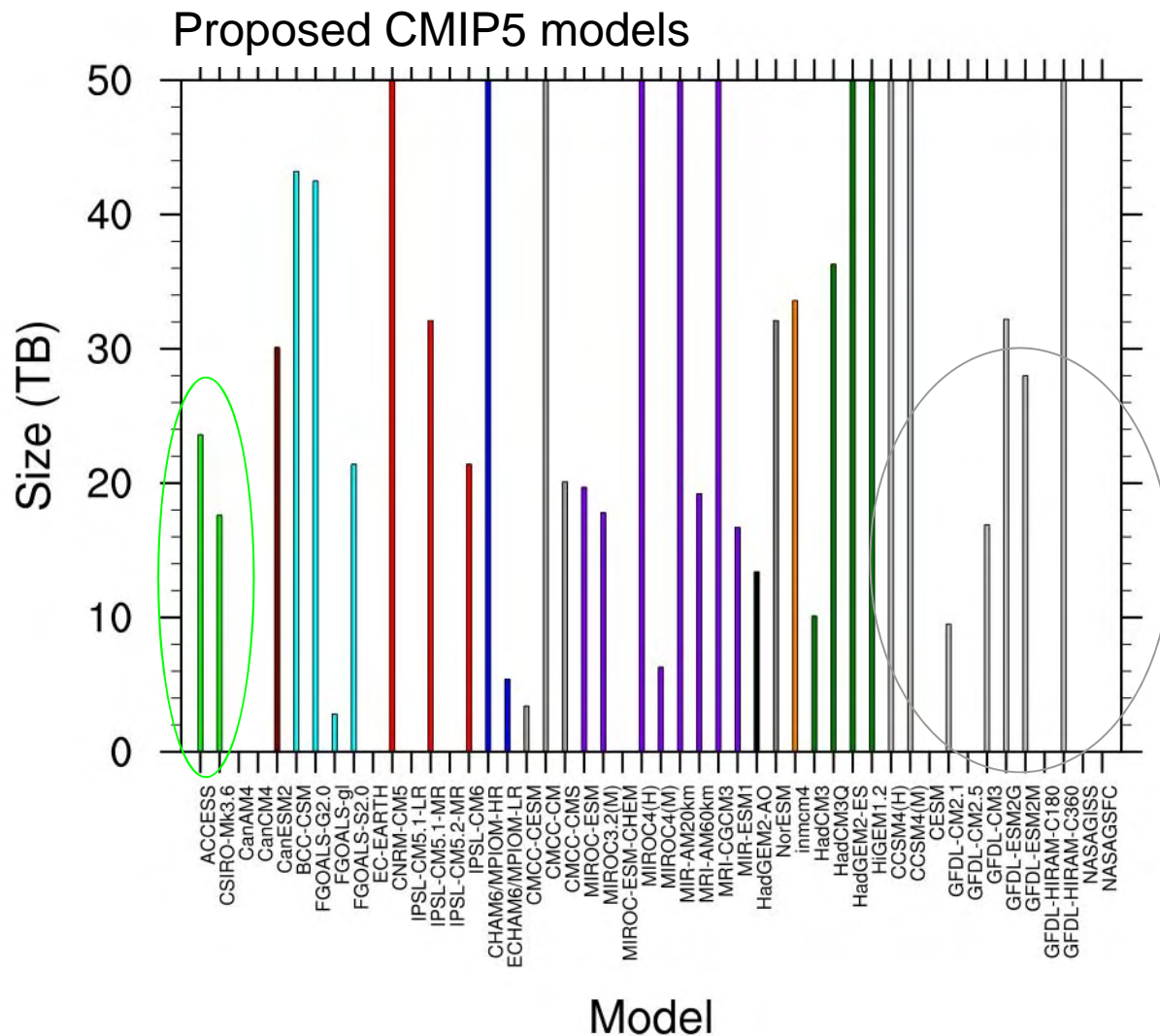
CMIP5 Data Characteristics



- Each file contains a single output variable (along with coordinate/grid variables, attributes and other metadata) from a single model and a single simulation
- There is flexibility in specifying how many time slices are stored in a single file
- Much of the metadata written to the output files is defined in MIP-specific tables of information
- The metadata is constrained by the CF v1.4 convention as specified in the CMIP5 tables
- The output files are written through the NETCDF API following the NETCDF_CLASSIC model and without compression of any kind



CMIP5 Data Totals by Model



48 in total,
compared to 24-25
for CMIP3

WGCM requested
clarification of
submission details
by modelling centers
29 September 2010

Revised Modelling







- A model code error was identified: zero short-wave forcing but non-zero long-wave forcing for volcanic eruptions (hence a warming effect instead of cooling)
- Affected runs: Historic run (with “all forcings”), other historic runs that include volcanic forcing
- Unaffected runs: 1% compounding CO₂, historic runs that don't include volcanic forcing
- More-or-less unaffected: Pre-industrial control
- More-or-less unaffected: RCP projections



Revised Modelling Continued



Experiment	Year range	Mk 3.6.0 Ens.	Mk 3.6.0 + SO ₄ fix Ens.	Mk 3.6.1 Ens.
3.1 Control	500 yr	1		1
3.2 Historical	1850-2005	10	10 45%	10
3.3 AMIP	1979-2008	10	10 100%	10
3.4 Mid-Holocene	100 yr	1		
4.[12] RCPs 4.5, 8.5	2006-2100	10	10 	10
4.[34] RCPs 2.6, 6.0	2006-2100	10	3 	3
4.1-L RCP 4.5	2100-2300	3	3 	3
6.1 1%/yr CO2 to 4x	140	1		
6.2a AGCM + control SSTs	30	1		1
6.2b AGCM + control SSTs + 4x CO2	30	1		1
6.3 4x CO2	150 + 5	1+11		
6.4a AGCM + control SSTs + AA	30	1		1
6.4b AGCM + control SSTs + SA	30	1		1
7.1 Historical (natural)	1850-2011	10	10 36%	10
7.2 Historical (GHGs)	1850-2011	10		10
7.3a Historical (anthropogenic)	1850-2011	10		10
7.3b Historical (all except ozone)	1950-2011	10	10 	10
7.3c Historical (all except AA)	1850-2011	10	10 35%	10
7.3d Historical (AA)	1850-2011	10		10
7.3e Historical (Asian)	1850-2011	10	10 36%	10
Peak number of nodes			61	81



Post-processing Challenge



- Current post-processing challenge: **1.86Tb** 4573 files.

This is for current completed experiments and only up to a maximum of 3 ensembles for some experiments, and for parameters that have the capacity to be defined (subset of the potential)

- Once the volcanic-error affected experiments are re-run, the post-processing challenge will be: **12.15Tb** 27543 files

This is for all experiments up to a maximum of 10 ensembles for some experiments

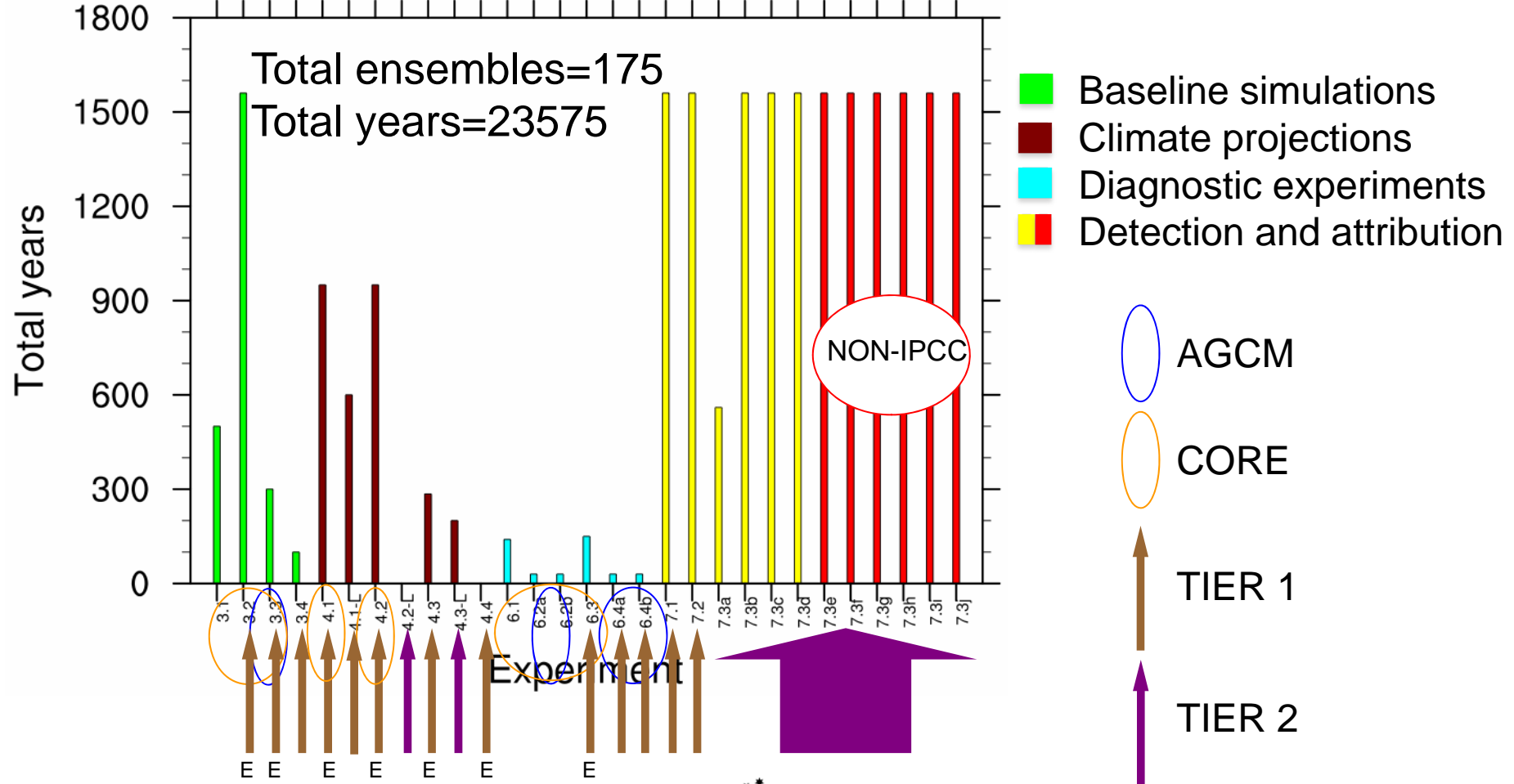
NB. This does not include contributions from daily or 6hourly data for which temporal subsets are requested.



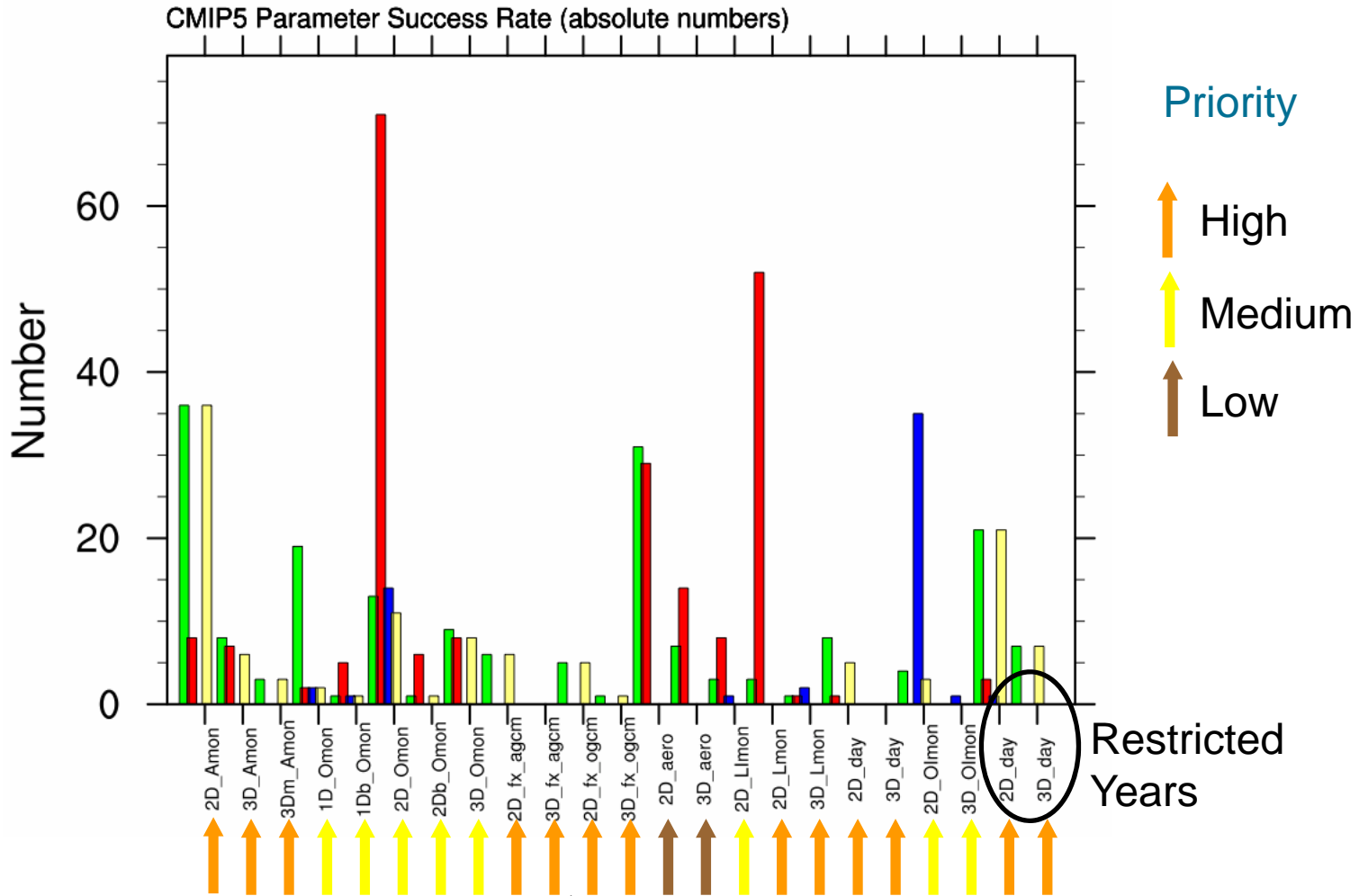
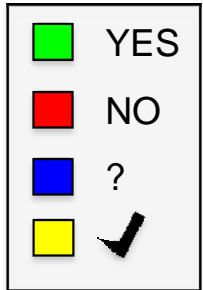
Experiment Success Rate



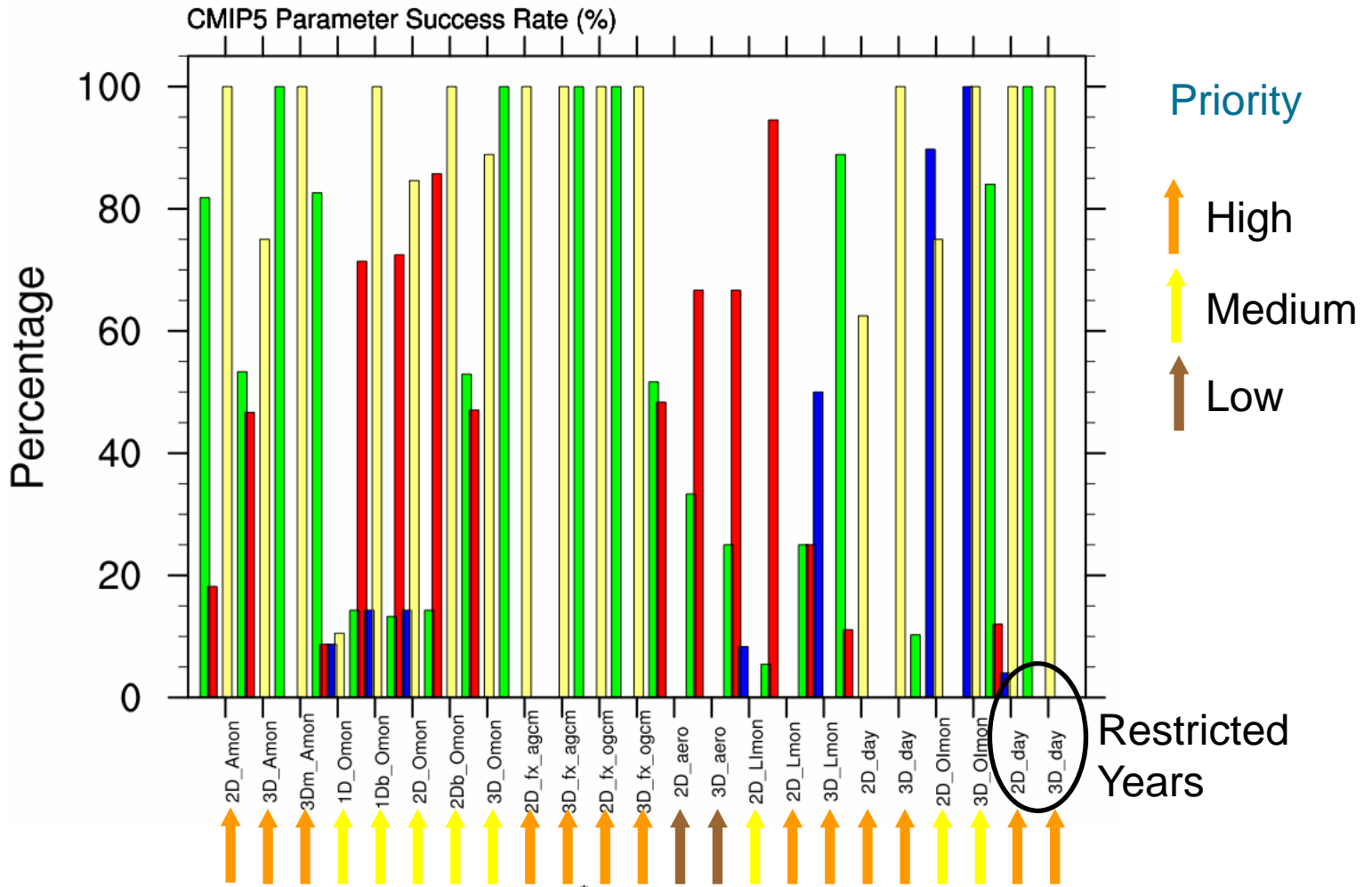
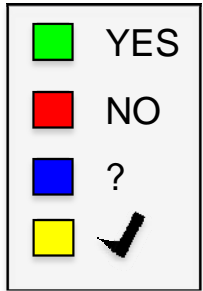
CMIP5 CSIRO-Mk3-6-0 Experiments



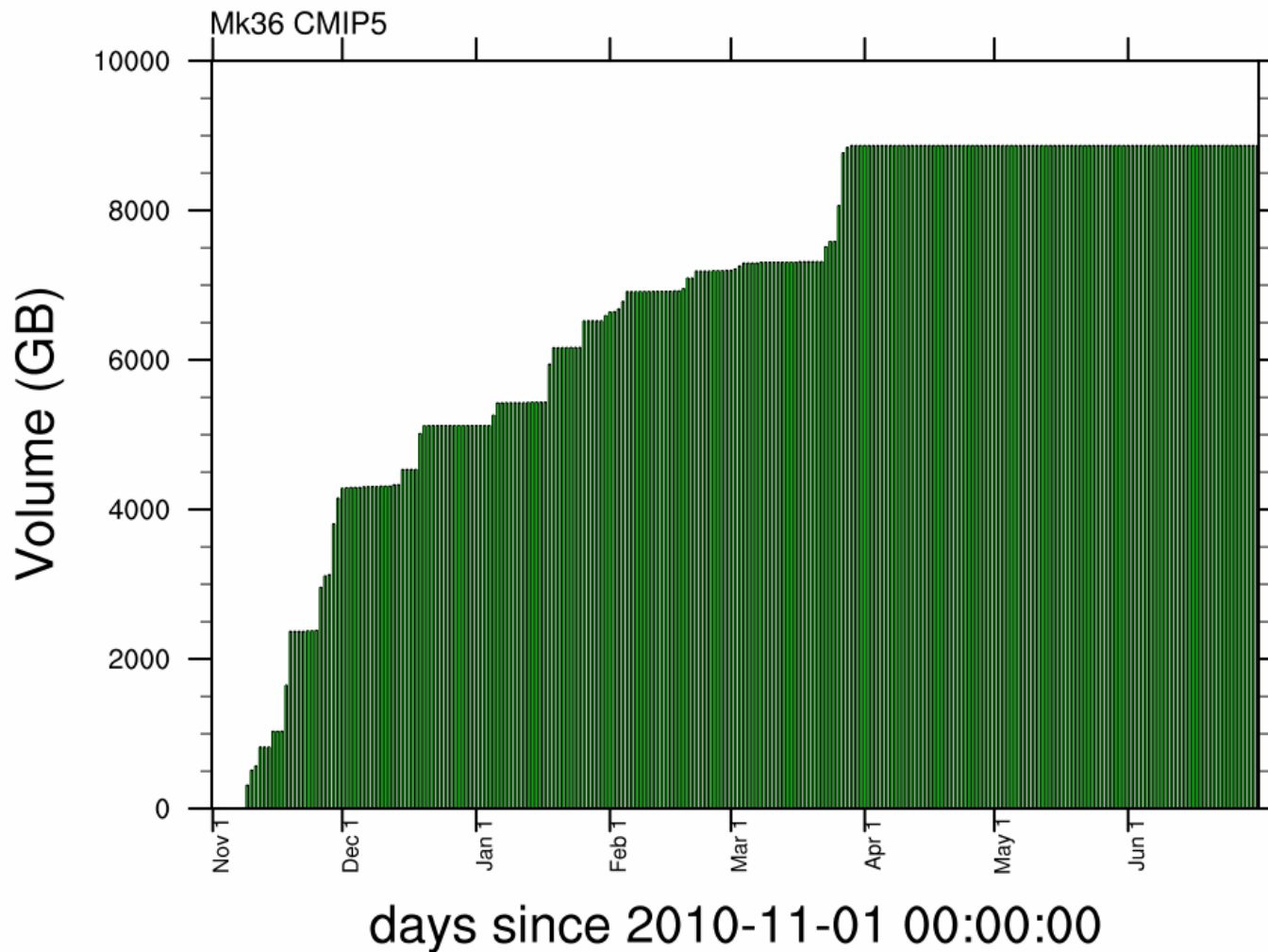
Parameter Success Rate



Parameter Success Rate



Mk36 CMIP5 Data Volume



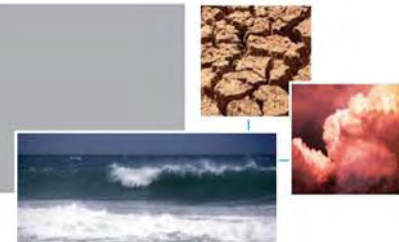
Where is the Mk36 CMIP5 data?



- Currently small quantities of CMIP5 data are being served via ESG
- Our growing collection of data resides on Vayu following DRS awaiting publishing on publicly accessible area on the DataCloud and eventually through the NCI Super-Node/Gateway (DOI)

[I have developed software to process CMIP5 files (“CLIM” tools) into further temporal products (e.g. seasonal anomalies relative to a year range base), these have had limited applications to other model output]

ACCESS Coupled Modelling



Experiment	Year range	Ens.
3.1 Control	500 yr	1
3.2 Historical	1850-2005	1
3.3 AMIP	1979-2008	1
3.4 Mid-Holocene	100 yr	
4.[12] RCPs 4.5, 8.5	2006-2100	1
4.[34] RCPs 2.6, 6.0	2006-2100	
4.1-L RCP 4.5	2100-2300	
6.1 1%/yr CO2 to 4x	140	1
6.2a AGCM + control SSTs	30	1
6.2b AGCM + control SSTs + 4x CO2	30	1
6.3 4x CO2	150 + 5	
6.4a AGCM + control SSTs + AA	30	
6.4b AGCM + control SSTs + SA	30	
7.1 Historical (natural)	1850-2011	
7.2 Historical (GHGs)	1850-2011	
7.3a Historical (anthropogenic)	1850-2011	
7.3b Historical (all except ozone)	1950-2011	
7.3c Historical (all except AA)	1850-2011	
7.3d Historical (AA)	1850-2011	
7.3e Historical (Asian)	1850-2011	



ACCESS CMIP5 submission



- Production post-processing has not begun as

- 1) ACCESS coupled model code not frozen
- 2) and ACCESS coupled model test experiments are still underway

- We are expecting a somewhat smoother post-processing experience with the ACCESS compared to the Mk36 models as

- 1) the basis for the WGCM requests are now clear
- 2) a lot of the model infrastructure is more CMIP5 savvy than the Mk36
- 3) potential for more people to contribute to various aspects and “piggy-back” of Mk36 approach
- 4) no need to transfer data between experiment and post-processing platforms
- 5) better resourced (initial basic submission is 1.5Tb)

Conclusions



- After a “minor” modelling setback we are back on track to publish Mk36 data over the coming weeks and months
- This set of data will constitute the largest ever assembled coupled climate model dataset generated by an Australia coupled climate model, only potentially surpassed by the ACCESS coupled model submission during AR5
- This activity highlights the potential for large and complex projects undertaken by a collaboration of uniquely focused, resourced and skilled research groups
- It is envisaged that several components of the Mk36 post-processing and ESG publishing activity can be adopted in the ACCESS coupled model CMIP5 submission
- Leon and Steve welcome enquiries regarding access to and collaborative projects with the Mk36 model output

Recent Mk36 papers



- Rotstayn et al 2010, “Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment”, IJC
- Rotstayn et al 2011 “Enhancement of ENSO-related natural rainfall variability due to Australian dust”, to be submitted to Atmos. Chem. Phys. (discussion paper awaiting publication)
- Rotstayn et al 2011 “Have anthropogenic aerosols delayed greenhouse gas-induced changes in Indo-Pacific regional circulation rainfall? Submitted to GRL
- Boucher et al 2011 “Climate response to aerosol forcings in CMIP5” Special edition of the CLIVAR newsletter





Australian Government
Bureau of Meteorology

The Centre for Australian Weather and Climate Research
A partnership between CSIRO and the Bureau of Meteorology



Mark Collier
CAWCR and CSIRO Marine and
Atmospheric Research

Phone: 61 3 9239 4422
Email: Mark.Collier@csiro.au
Web: www.cawcr.gov.au

Thank you

www.cawcr.gov.au

