Australia's CMIP5 submission using the CSIRO-Mk3.6 model

Stephen Jeffrey[1], Leon Rotstayn[2], Mark Collier[2], Stacey Dravitzki[2], Carlo Hamalainen[1], Chris Moeseneder[3], Kenneth Wong[1] and Jozef Syktus[1]

[1]Queensland Climate Change Centre of Excellence,
Ecosciences Precinct, Dutton Park, Queensland
[2]The Centre for Australian Weather and Climate Research,
CSIRO Marine and Atmospheric Research, Aspendale, Victoria
[3]CSIRO Marine and Atmospheric Research, Ecosciences Precinct, Dutton Park, Queensland

(Manuscript received July 2012 revised December 2012)

A comprehensive set of climate modelling experiments has been performed to provide input into the Coupled Model Intercomparison Project – phase 5 (CMIP5). The CSIRO-Mk3.6 climate model was used to prepare a joint CMIP5 submission under a partnership between the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the Queensland Climate Change Centre of Excellence (QCCCE). The submission includes data for 163 variables from 22 experiments. The raw model output has been post-processed into CMIP5 format using the Climate Model Output Rewriter and publicly released on the Earth System Grid. The Mk3.6 submission includes data for most of the Core, Tier 1 and Tier 2 CMIP5 longer-term experiments which don't require modelling of the carbon cycle. Data have been provided for nine historical experiments driven by a range of forcings to support detection and attribution studies. Most experiments have been performed as an ensemble of runs, with ensemble sizes exceeding CMIP5 recommendations.

Introduction

The Coupled Model Intercomparison Project[1] (CMIP) is an international effort to assist analysts and model developers in understanding the differences between models and the mechanisms driving those differences. The project commenced in 1995 with modelling groups undertaking a number of 'control' experiments to enable model intercomparison. The most recent component of the project – phase 5 (CMIP5) – continues this tradition with the definition of a coordinated and extensive set of experiments to support:

(i) assessment of model performance (by determining how well models simulate the past);
(ii) climate predictions for the future (projections based on emissions scenarios); and
(iii) intercomparison of model results.

Analysis of the CMIP5 results is expected to contribute most of the new modelling results presented in the Intergovernmental Panel on Climate Change Fifth Assessment Report (AR5), due for release in 2013–2014.

Participating in CMIP5 requires a substantial commitment from modelling groups. To contribute data for a given experiment involves experiment modelling, post-processing the model output to CMIP5 format, and public release of the data. Each stage has its own challenges, which requires participating groups to have extensive scientific and technical expertise and supercomputing resources.

In 2009 the CSIRO and Queensland Climate Change Centre of Excellence (QCCCE) commenced a three-year collaborative project. The goal was to prepare a joint submission to CMIP5 using the CSIRO-Mk3.6 climate model. This paper is the first in a series of papers in this issue of the Australian Meteorological and Oceanographic Journal (AMOJ) describing the CSIRO-QCCCE contribution to CMIP5. It provides a description of the tools, datasets and activities involved in building the CSIRO-QCCCE CMIP5 data archive. Subsequent papers in the series will focus on model performance and projections.

[1]http://cmip-pcmdi.llnl.gov

Corresponding author address: Stephen Jeffrey, email: stephen.jeffrey@qld.gov.au

Model

The CSIRO Mark 3.6 (Mk3.6) global climate model (GCM) was developed from the earlier Mk3.5 version, which was described in detail by Gordon et al. (2002, 2010). It is a coupled atmosphere–ocean model with dynamic sea-ice. It also has a soil–canopy scheme with prescribed vegetation properties. The ocean, sea-ice and soil–canopy models are unchanged between Mk3.5 and Mk3.6. The main differences between Mk3.5 and Mk3.6 are the inclusion of an interactive aerosol treatment and an updated radiation scheme in Mk3.6. Rotstayn et al. (2010) gave an overview of Mk3.6, and also assessed the model's simulation of Australian mean climate and natural rainfall variability associated with ENSO, with generally favourable conclusions.

The atmospheric component is a spectral model, which utilises the flux form of the dynamical equations (Gordon 1981). It has eighteen vertical levels and horizontal resolution of approximately 1.875° × 1.875° (spectral T63). The ocean model is based on version 2.2 of the Modular Ocean Model (MOM2.2; Pacanowski 1996). Every atmospheric grid-box is coupled to two oceanic grid-boxes: enhanced north–south resolution in the ocean model was implemented with the aim of improving the representation of tropical variability. The ocean model thus has resolution of approximately 1.875° × 0.9375° and has 31 vertical levels. The sea-ice model is based on O'Farrell (1998), with revised numerics as described by Gordon et al. (2010).

The model physics were not changed specifically for use in CMIP5; modifications were limited to those required to use CMIP5 forcing datasets. Additional diagnostics were however added so that data could be provided for a number of variables requested by CMIP5 which were not previously available from Mk3.6.

Forcing datasets

The CMIP5 experimental design (Taylor et al. 2012) defines a series of experiments which are driven by various combinations of anthropogenic and natural forcings. Anthropogenic forcing agents in our runs are long-lived greenhouse gases (GHGs), ozone and aerosols. Changes in land use are not included. We prescribed CMIP5-recommended, annual-mean concentrations of long-lived GHGs (carbon dioxide, methane, nitrous oxide and chlorofluorocarbons CFC-11, CFC-12 and HCFC-22), and monthly mean, spatially-varying ozone concentrations. Ozone concentrations are based on the AC&C/SPARC ozone database (Cionni et al. 2011); they are three dimensional in the troposphere, and zonally averaged in the stratosphere. Emissions of anthropogenic aerosols and aerosol precursors also follow CMIP5 recommendations (Lamarque et al. 2010), with modifications as described in Rotstayn et al (2012). In addition to direct effects and indirect aerosol effects on liquid water clouds, the model also includes a simple treatment of the effect of black carbon on snow albedo (Hansen and Nazarenko 2004). The anthropogenic-forcing datasets are prescribed for both the historic period and the four Representative Concentration Pathways (RCPs); see Moss et al. (2010) for an overview of the RCPs. We also included the historic time series of annual-mean total solar irradiance recommended for CMIP5; this includes estimates of both the 11-year solar cycle and changes in background irradiance (Lean 2000, Wang et al. 2005). Experiments requiring a fixed pre-industrial solar irradiance used a 13-year average (1844–1856). Total solar irradiance data for years beyond 2008 were not available at the time we commenced the simulations; an 11-year cycle (1997–2008; 2000 was erroneously omitted) was repeated for years 2009–2300.

CMIP5 does not specifically include a prescribed dataset for volcanic forcing. We prescribed zonally averaged distributions of stratospheric sulfate based on Sato et al. (1993). A recent update[2] which extended the dataset beyond 1999, was not available at the time we commenced the simulations, so we set volcanic forcing to zero from 2000 onwards.

Experiments

CMIP5 defined a broad range of experiments encompassing control, paleoclimatic, sensitivity, hindcast and forecast experiments, as well as providing scope for modelling groups to define their own historical experiments for detection and attribution studies. The experiments are arranged in two focus groups (near-term and longer-term), with each group subdivided into 'Core', 'Tier 1' and 'Tier 2', in order of priority. Modelling groups were encouraged to undertake the 'Core' experiments (in either or both focus groups) first, followed by Tier 1 and Tier 2 experiments if resources were available.

CSIRO-QCCCE addressed all of the experiments in the longer-term focus group with the exception of two paleoclimate and four cloud diagnostic experiments, and those experiments requiring a carbon cycle within the model. The contributed set comprises all the core experiments (except those requiring a carbon cycle) and most of the Tier 1 and Tier 2 experiments in the longer-term group. Experiments in the near-term (decadal) focus group could not be addressed as the CSIRO-Mk3.6 model cannot be initialised with oceanic observations for decadal prediction.

The CSIRO-QCCCE submission consists of data for 22 experiments, including six historical experiments designed for detection and attribution studies. Most experiments have an ensemble size of ten, exceeding the CMIP5 requirement of three member ensembles. Access to ensemble members 6–10 of the 'historicalMisc' experiment is (at the time of publication) restricted to groups collaborating with the Mk3.6 modelling team. The restriction will be removed in the near future. The experiments addressed are shown in Table 1.

[2]http://data.giss.nasa.gov/modelforce/strataer

Table 1. The CSIRO-QCCCE CMIP5 experiments.
Notes: † ensemble members 1–3 are extended to 2300; †† ensemble members 2–12 are five years in length consistent with the CMIP5 specification; ††† experiment commenced in 1950 as ozone changes prior to 1950 were considered negligible; * climatology computed using years 191–290 of the piControl experiment; ** includes the indirect effect of black carbon on snow albedo; AGCM denotes atmospheric-only global climate model; natural forcings include solar irradiance and stratospheric sulfates; anthropogenic forcings include greenhouse gases (GHGs), anthropogenic aerosols and ozone; p1–p6 physics version is used to distinguish between the six different forcing sets used in historicalMisc; and *** Asia is defined as the region from 10°S to 45°N and from 65°E to 150°E. All forcings are time-varying unless specified otherwise. RCP 4.5 forcings were used to extend the historicalGHG and historicalMisc experiments from 2006–2012.

Experiment	*Ens. size*	*Years*	*Summary*	
piControl	1	1-500	All forcings fixed at 1850 levels	
historical	10	1850–2005	All forcings are time varying	
AMIP	10	1979–2009	AGCM experiment using observed SSTs and sea-ice	
midHolocene	1	1–100	All forcings fixed at 1850 levels; mid-Holocene orbital parameters	
rcp45	10	2006–2100†	RCP 4.5	
rcp85	10	2006–2100†	RCP 8.5	
rcp26	10	2006–2100	RCP 2.6	
rcp60	10	2006–2100	RCP 6.0	
1pctCO2	1	1–140	Cumulative increase in CO_2 of 1 per cent/year up to 4 × 1850 concentration; all other forcings fixed at 1850 levels	
sstClim	1	1–30	AGCM experiment with SST and sea-ice climatology* from piControl experiment; all other forcings fixed at 1850 levels	
sstClim4xCO2	1	1–30	AGCM experiment with SST and sea-ice climatology* from piControl experiment; CO2 fixed at 4 × 1850 concentration; all other forcings fixed at 1850 levels	
abrupt4xCO2	12	1–150††	CO_2 fixed at 4 × 1850 concentration; all other forcings fixed at 1850 levels	
sstClimAerosol	1	1–30	AGCM experiment with SST and sea-ice climatology* from piControl experiment; anthropogenic aerosols** fixed at 2000 levels; all other forcings fixed at 1850 levels	
sstClimSulfate	1	1–30	AGCM experiment with SST and sea-ice climatology* from piControl experiment; anthropogenic sulfate aerosols** fixed at 2000 levels; all other forcings fixed at 1850 levels	
			Natural forcings	*Anthropogenic forcings*
historicalNat	10	1850–2012	Time varying	Fixed at 1850 levels
historicalGHG	10	1850–2012	Fixed at 1850 levels	GHGes: time varying All others: fixed at 1850
historicalMisc[p1]	10	1850–2012	Fixed at 1850 levels	Time varying
historicalMisc[p2]	10	1950–2012†††	Time varying	Ozone: fixed at 1949 All others: time varying
historicalMisc[p3]	10	1850–2012	Time varying	Aerosols**: fixed at 1850 All others: time varying
historicalMisc[p4]	10	1850–2012	Fixed at 1850 levels	Aerosols**: time varying All others: fixed at 1850
historicalMisc[p5]	10	1850–2012	Time varying	Aerosols: fixed at 1850, but time varying in Asian region*** All others: time varying
historicalMisc[p6]	10	1850–2012	Volcanic: time varying All others: fixed at 1850	Fixed at 1850 levels

The pre-industrial control experiment was initialised from starting conditions taken from another preindustrial control experiment that showed little evidence of drift; there were some minor differences in aerosol treatments, but the net radiative balance at the top of the atmosphere was similar in the two experiments. A 160-year spin-up period allowed the model to adjust to the new CMIP5 pre-industrial forcings, before data collection commenced. Data throughout the spin-up period were not submitted to CMIP5. The pre-industrial control experiment was allowed to stabilise for a further 80 years before being used to initialise other experiments. Those experiments with a multi-member ensemble were initialised from starting conditions taken from the pre-industrial control at random intervals between ten and 17 years. A longer separation between ensemble members would have been desirable but was not possible due to time constraints. The time intervals were randomised to reduce any correlations between ensemble members due to aliasing onto natural modes of variability, such as the El Niño–Southern Oscillation. It is also expected that any remaining correlations would be much reduced by the early 20th century (i.e. after several decades of integration). Table 2 (Appendix 1) details the simulation date at which each experiment and ensemble member branched from the parent experiment. Table 2 also shows the year range in the pre-industrial control experiment which aligns with the year ranges in all other experiments.

The AMIP and mid-Holocene experiments were both allowed to stabilise before data collection commenced. Spin-up periods of 30 and 300 years were used for those experiments, respectively.

Simulated temperature changes

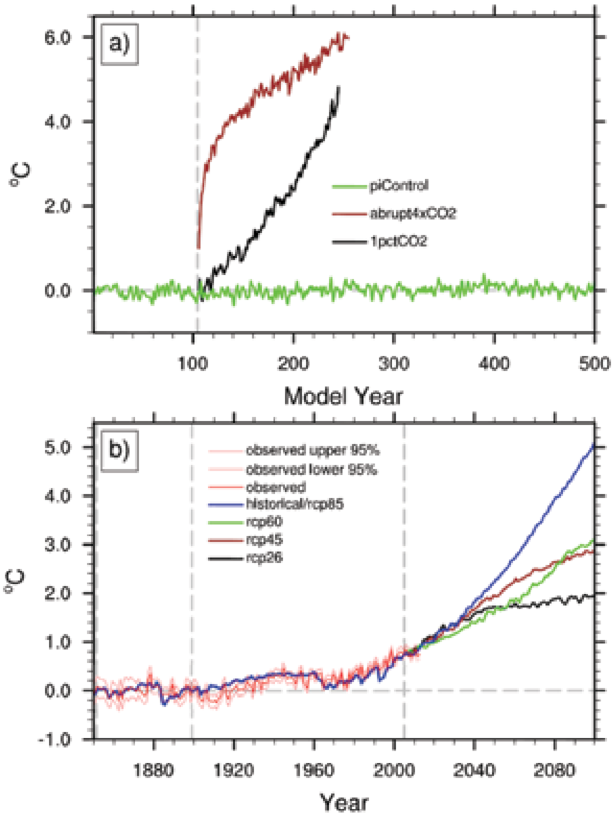
The pre-industrial control experiment can be used to provide a baseline for the other experiments and also to determine if there is an apparent drift in the model. Figure 1 shows the globally averaged surface air temperature (T_s) for the pre-industrial, historical, four projection and two idealised (or sensitivity) experiments. The top panel indicates the Mk3.6 model is stable, in the sense that the drift in T_s is small in the pre-industrial control run; the linear trend of 0.024 K/century compares favourably with earlier models from CMIP3, most of which had drift exceeding 0.05 K/century (Sen Gupta et al. 2012, Fig. 3).

The bottom panel compares the modelled results (for the historical period) with observations (HadCRUT4[3]; Morice et al. 2012). The uncertainty in the HadCRUT4 observations (see bottom panel) is discussed by Morice et al.; after 1950, the total uncertainties in the annual, global-mean temperatures are generally less than 0.2 °C. This includes the effects of coverage limitations, for which we have not attempted to correct the modelled time series in Fig. 1. The modelled historical time series is generally within the uncertainty range of the observations, though there is some indication of an underestimate by the model after about 1990. This may suggest that the model's aerosol forcing is somewhat too strong (Rotstayn et al. 2012). After 2005, the evolution of T_s in each RCP is qualitatively consistent with the radiative forcing (van Vuuren et al. 2011, Fig. 10).

For further analysis of the Mk3.6 datasets, the reader is referred to other papers in this issue of AMOJ.

[3]http://www.metoffice.gov.uk/hadobs/hadcrut4/index.html

Fig. 1. Global-mean surface air temperature anomalies (°C) for: (a) pre-industrial control, abrupt4xCO2 and 1pct-CO2 experiments; and (b) HadCRUT4 observations and historical, rcp26, rcp45, rcp60 and rcp85 experiments. Data are presented as anomalies relative to base periods: years 80–129 (piControl), 1–20 (1pct-CO2 and abrupt4xCO2) and 1850–1899 (all datasets in lower panel). With the exception of the observational data, the solid lines in (b) are ten member ensemble means. The pink lines in (b) show the uncertainty in the observations. The vertical dashed lines represent the model year (year 104) where the idealised experiments branched from the pre-industrial control (a), and the base (1850–1899) and historical (1850–2005) periods in (b).



Variables

The CMIP5 output specification[4] defines the datasets requested for each experiment. In many cases, climate variables are requested for a subset of experiments, ensemble members and simulation years. The CSIRO-QCCCE submission includes data for 163 climate variables and ten time-invariant variables (such as cell volumes). Data were submitted for those variables which were output directly from the Mk3.6 model, and those which could be readily derived from model outputs. The CSIRO-QCCCE submission also includes a number of extensions, providing data for ensemble members and simulation years beyond that requested by CMIP5. The datasets available are summarised in Table 3 (Appendix 2).

[4]http://cmip-pcmdi.llnl.gov/cmip5/docs/standard_output.pdf

Three-dimensional atmospheric datasets have been provided using the vertical coordinate defined by CMIP5 for each variable, which in most cases is either the model's native vertical coordinate or a prescribed set of pressure levels. Vertical interpolation was used to map three-dimensional variables onto the required vertical coordinate. Grid cells were interpolated individually and all variables were interpolated as a function of pressure (*p*). The following interpolation types were used: linear for wind variables and relative humidity; p^3 for specific humidity; and $ln(p)$ for temperature variables and geopotential height. The interpolation method used was selected for consistency with the interpolation method used within the model.

In general, data were not extrapolated beyond the vertical levels supported by model. The Mk3.6 model uses a hybrid σ:pressure vertical coordinate with the lowest level at approximately 995 hPa. CMIP5 requested a number of daily and monthly atmospheric variables on pressure levels up to and including 1000 hPa. Extrapolation was used to generate data at the 1000 hPa level for grid cells having surface pressure greater than 1000 hPa.

All variables were presented on their native horizontal grid. Horizontal interpolation (bilinear) was only used when deriving a new variable from multiple inputs for which the constituent variables were not all on a common grid.

Post-processing

Customised software was developed to post-process the Mk3.6 model output into CMIP5 format[5]. Existing data manipulation tools could not be used as post-processing required the flexibility to load and manipulate raw model outputs which were defined on multiple time domains (six-hourly, daily, monthly and time-invariant datasets), and multiple horizontal and vertical coordinates. For some variables, post-processing involved little or no modification of the raw Mk3.6 data, while others required extensive data manipulation. Common processing tasks included: vertical interpolation from model levels to prescribed pressure levels (or model pressure levels to prescribed pressure levels), averaging over time domains, conversion of units, range checking, and derivation of new variables from one or more existing Mk3.6 parameters.

The Climate Model Output Rewriter library Version 2 (CMOR2)[6] was used to output all datasets. CMOR2 was used as it has a number of benefits: (1) files are written according to the strict CMIP5 formatting requirements; (2) CMOR2 performs some simple quality assurance checks; and (3) the outputs are guaranteed to achieve partial compliance with the CMIP5 quality control standard.

Quality control

CMIP5 datasets must pass a series of quality control[7] (QC) checks before they can be published and formally cited. The respective levels are: QC0 (informal spot checking of data); QC1 (data and metadata compliance checks automatically imposed by CMOR2 and the Earth System Grid data publishing software); QC2 (data consistency checks); and QC3 (extended checking of data and metadata). Datasets prepared using CMOR2 and published using the Earth System Grid software (Williams et al. 2009) are guaranteed to achieve compliance with QC level 1. Datasets must be manually checked using the QC level 2 tools[8] to achieve QC2 compliance. Individual modelling groups are responsible for ensuring their datasets pass the QC1 and QC2 checks. Datasets passing both QC1 and QC2 are then subject to QC3 checking by the World Data Centre for Climate[9]. Datasets meeting QC3 requirements are assigned a Digital Object Identifier[10] (DOI) and can then be formally cited in the scientific literature.

The Mk3.6 CMIP5 datasets have passed the level 1 and 2 quality control checks. At the time of publication, the datasets were undergoing QC3 checks.

Data archive

CMIP5 datasets are hosted on the Earth System Grid (ESG)[11]. The ESG is an international network of data nodes which enables users to transparently access data irrespective of where the data are physically stored. The Mk3.6 datasets are publicly accessible for non-commercial use and are physically stored at the data node hosted by the National Computational Infrastructure National Facility (Canberra, Australia)[12]. The datasets are being replicated at the World Data Centre for Climate for QC3 data checking and to enable users to access the data when the primary host is inaccessible. At the time of publication, the Mk3.6 archive consisted of approximately 90 000 files and occupying about 30 TB of disk space. The data are stored in NetCDF format, conforming to the 'classic' data model.

The catalogue of CMIP5 datasets is mirrored at all ESG nodes. The catalogue is publicly accessible and has search facilities for identifying data of interest. To download data users must register at one of the data nodes and apply for access to the relevant data group. The 'CMIP5 research' group would be appropriate for most research and education purposes. Once the user has logged in to the ESG and selected the dataset(s) of interest, the relevant file(s) can be downloaded using the web browser. Alternatively, users may

[5]http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax.pdf
http://cmip-pcmdi.llnl.gov/cmip5/docs/CMIP5_output_metadata_requirements.pdf
[6]http://www2-pcmdi.llnl.gov/cmor (CMOR Version 2.5, C-interface).

[7]https://redmine.dkrz.de/collaboration/projects/cmip5-qc/wiki
[8]https://redmine.dkrz.de/collaboration/projects/cmip5-qc/wiki/Qc_level_2
[9]http://www.dkrz.de/daten-en/wdcc?set_language=en
[10]http://www.doi.org/
[11]http://www.earthsystemgrid.org/home.htm
[12]http://esg2.nci.org.au/esgf-web-fe/

wish to download a small UNIX shell script that can be used to transfer the files using the File Transfer Protocol (FTP). The shell script method is recommended for downloading large volumes of data, but users will require access to an authentication tool (such as MyProxyClient, MyProxyLogon or Globus) and wget (an FTP client).

Summary

The CSIRO and QCCCE have prepared a comprehensive CMIP5 submission using the Mk3.6 climate model which includes data for 163 variables from 22 experiments. We expect the Mk3.6 submission will be of particular interest to a range of analysts because: (1) it includes a suite of historical experiments tailored to detection and attribution studies; and (2) the ensemble size of most experiments far exceeds the CMIP5 minimum requirement. Furthermore, as most of the data were publicly released on the ESG throughout July–December 2012, we expect the Mk3.6 results will be included in many of the publications being used as input to the IPCC's Fifth Assessment Report.

Acknowledgments

This work was supported by CSIRO Marine and Atmospheric Research and the Queensland Climate Change Centre of Excellence (Queensland Government Department of Science, Information Technology, Innovation and the Arts). QCCCE provided the high-performance computing facilities for the model experimentation. The National Computational Infrastructure (National Facility) provided computing resources for post-processing the raw model datasets, hardware for hosting the Mk3.6 CMIP5 data on the Earth System Grid, and technical support.

The project team would like to acknowledge the assistance provided by Martin Dix, Hal Gordon, Tony Hirst, Didier Monselesan, Siobhan O'Farrell and Eva Kowalczyk. Published datasets for global average thermosteric sea level change and sea surface height above geoid were prepared by Didier Monselesan.

References

Websites

Climate Model Output Rewriter web site: http://www2-pcmdi.llnl.gov/cmor

Coupled Model Intercomparison Project web site: http://cmip-pcmdi.llnl.gov

CMIP5 - Modeling Info - Producing Model Output web site: http://cmip-pcmdi.llnl.gov/cmip5/output_req.html

CMIP5 Quality Control web site: https://redmine.dkrz.de/collaboration/projects/cmip5-qc/wiki

Digital Object Identifier web site: http://www.doi.org/

Earth System Grid web site: http://pcmdi9.llnl.gov/esgf-web-fe/

Forcings in GISS Climate Model web site: http://data.giss.nasa.gov/modelforce/strataer

HadCRUT4 dataset of global historical surface temperature anomalies web site: http://www.metoffice.gov.uk/hadobs/hadcrut4/index.html.

World Data Centre for Climate web site: http://www.dkrz.de/daten-en/wdcc?set_language=en

Other references

Cionni, I., Eyring, V., Lamarque, J.F., Randel, W.J., Stevenson, D.S., Wu, F., Bodeker, G. E., Shepherd, T.G., Shindell, D.T. and Waugh, D.W. 2011. Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing. *Atmos. Chem. Phys., 11,* 11 267–11 292, doi: 10.5194/acp-11-11267. http://www.atmos-chem-phys.net/11/11267/2011/acp-11-11267-2011.html.

Gordon, H.B. 1981. A flux formulation of the spectral atmospheric equations suitable for use in long-term climate modeling. *Mon. Weather Rev., 109,* 56–64.

Gordon, H.B., Rotstayn, L.D., McGregor, J.L., Dix, M.R., Kowalczyk, E.A., O'Farrell, S. P., Waterman, L.J., Hirst, A.C., Wilson, S.G., Collier, M.A., Watterson, I.G. and Elliott, T.I. 2002. The CSIRO-Mk3 Climate System Model, Technical Paper No. 60, CSIRO Atmospheric Research, Aspendale, Vic., Australia, 134 pp. Available online at http://www.cmar.csiro.au/e-print/open/gordon_2002a.pdf.

Gordon, H.B., O'Farrell, S.P., Collier, M.A., Dix, M.R., Rotstayn, L.D., Kowalczyk, E.A., Hirst, A.C. and Watterson, I.G. 2010. The CSIRO-Mk3.5 Climate Model, Technical Report No. 21, The Centre for Australian Weather and Climate Research, Aspendale, Vic., Australia, 62 pp. Available online at http://www.cawcr.gov.au/publications/technical-reports.php.

Hansen, J. and Nazarenko, L. 2004. Soot climate forcing via snow and ice albedos. *Proc. Natl. Acad. Sci., 101,* 423–428, doi:10.1073/pnas.2237157100.

Lamarque, J.-F., Bond, T.C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M.G., Shindell, D., Smith, S.J., Stehfest, E., Van Aardenne, J., Cooper, O.R., Kainuma, M., Mahowald, N., McConnell, J.R., Naik, V., Riahi, K. and van Vuuren, D.P. 2010. Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmos. Chem. Phys., 10,* 7017–7039, doi: 10.5194/acp-10-7017-2010.

Lean, J. 2000. Evolution of the Sun's spectral irradiance since the Maunder Minimum. *Geophys. Res. Lett., 27,* 2425–2428, doi:10.1029/2000GL000043.

Morice, C.P., Kennedy, J.J., Rayner, N.A. and Jones, P.D. 2012. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset, *J. Geophys. Res., 117,* 2156–202. D08101, doi:10.1029/2011JD017187.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, J.F.B., Nakicenovic, N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson, A.M., Weyant, J.P. and Wilbanks, T.J. 2010. The next generation of scenarios for climate change research and assessment. *Nature, 463,* 747–756, doi: 10.1038/nature08823.

O'Farrell, S.P. 1998. Investigation of the dynamic sea-ice component of a coupled atmosphere-sea-ice general circulation model. *J. Geophys. Res., 103,* 15751–15782.

Pacanowski, R.C. 1996. MOM 2 Version 2, Documentation, User's Guide and Reference Manual, GFDL Ocean Technical Report 3.2, Geophysical Fluid Dynamics Laboratory/NOAA, Princeton, NJ, USA, http://www.gfdl.noaa.gov/cms-filesystem-action/model_development/ocean/manual2.2.pdf.

Rotstayn, L.D., Jeffrey, S.J., Collier, M.A., Dravitzki, S.M., Hirst, A.C., Syktus, J.I. and Wong, K.K. 2012. Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a study using single-forcing climate simulations. *Atmos. Chem. Phys., 12,* 6377–6404, 2012, doi:10.5194/acp-12-6377.

Rotstayn, L.D., Collier, M.A., Feng, Y., Gordon, H.B., O'Farrell, S.P., Smith, I.N. and Syktus, J.I. 2010. Improved simulation of Australian climate and ENSO-related rainfall variability in a GCM with an interactive aerosol treatment. *Int. J. Climatol., 30,* 1067–1088, doi:10.1002/joc.1952.

Sato, M., Hansen, J.E., McCormick, M.P. and Pollack, J.B. 1993. Stratospheric aerosol optical depth, 1850–1990. *J. Geophys. Res., 98,* 22987–22994.

Sen Gupta, A., Muir, L.C., Brown, J.N., Phipps, S.J., Durack, P.J., Monselesan, D. and Wijffels, S.E. 2012. Climate drift in the CMIP3 models. *J. Clim., 25,* 4621–4640, doi:10.1175/JCLI-D-11-00312.1.

Taylor, K.E., Stouffer, R.J. and Meehl G.A. 2012. An Overview of CMIP5 and the experiment design. *Bull. Amer. Meteorol. Soc., 93,* 485–98 doi:10.1175/BAMS-D-11-00094.1.

Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S.J. and Rose, S.k. 2011. The representative concentration pathways: an overview. *Clim. Change, 109, 5–3,* DOI 10.1007/s10584-011-0148-z.

Wang, Y.-M., Lean, J.L. and Sheeley, Jr. N.R. 2005. Modeling the Sun's Magnetic Field and Irradiance since 1713. *Astrophys. J., 625,* 522–538, doi:10.1086/429689.

Williams, D.N., Ananthakrishnan, R., Bernholdt, D.E., Bharathi, S., Brown, D., Chen, M., Chervenak, A.L., Cinquini, L., Drach, R., Foster, I.T., Fox, P., Fraser, D., Garcia, J., Hankin, S., Jones, P., Middleton, D.E., Schwidder, J., Schweitzer, R., Schuler, R., Shoshani, A., Siebenlist, F., Sim, A., Strand, W.G., Su, M. and Wilhelmi, N. 2009. The earth system grid: Enabling access to multimodel climate simulation data. *Bull. Amer. Meteorol. Soc, 90,* 195–205, doi: 10.1175/2008BAMS2459.1.

Appendix 1. Experiment initialisation

Experiments were initialised as shown in Table 2. The branch date refers to when the experiment branched from the parent experiment using the year coordinates of the parent. The branch point includes an ensemble number if the experiment branched from a parent experiment having multiple ensemble members.

Table 2. Initialisation used in CSIRO-QCCCE CMIP5 experiments.
Notes: † data was not submitted to CMIP5 for spin-up periods; †† years in piControl which correspond to the years in the target experiment; ††† data provided for piControl is limited to years 1–500; ^ the historical experiment ends in 2005, all other experiments end in 2012; ^^ ensemble members 1–3 of experiments rcp45 and rcp85 extend to 2300. *E1*, *E2* etc. denotes the ensemble member in a multi-member ensemble.

Experiment	*Parent experiment*	*Ens.*	*Branch point*	*Spin-up*[†]	*Alignment*[††]
piControl	non-CMIP5 control experiment	1	Dec 429	160 years	Not applicable
historical	piControl	1	Dec 80		Jan 81–Dec 236^
historicalNat		2	Dec 91		Jan 92–Dec 247^
historicalGHG		3	Dec 103		Jan 104–Dec 259^
historicalMisc^{p1}		4	Dec 116		Jan 117–Dec 272^
historicalMisc^{p3}		5	Dec 126		Jan 127–Dec 282^
historicalMisc^{p4}		6	Dec 137		Jan 138–Dec 293^
historicalMisc^{p5}		7	Dec 152		Jan 153–Dec 308^
historicalMisc^{p6}		8	Dec 168		Jan 169–Dec 324^
		9	Dec 185		Jan 186–Dec 341^
1850–2005(2012)^		10	Dec 199		Jan 200–Dec 355^
historicalMisc^{p2}	historical	1	Dec 1949 *E1*		Jan 181–Dec 243
		2	Dec 1949 *E2*		Jan 192–Dec 254
		3	Dec 1949 *E3*		Jan 204–Dec 266
		4	Dec 1949 *E4*		Jan 217–Dec 279
		5	Dec 1949 *E5*		Jan 227–Dec 289
		6	Dec 1949 *E6*		Jan 238–Dec 300
		7	Dec 1949 *E7*		Jan 253–Dec 315
		8	Dec 1949 *E8*		Jan 269–Dec 331
		9	Dec 1949 *E9*		Jan 286–Dec 348
1950–2012		10	Dec 1949 *E10*		Jan 300–Dec 362
AMIP	piControl	1	Dec 91	30 years	Jan 122–Dec 152
		2	Dec 103		Jan 134–Dec 164
		3	Dec 116		Jan 147–Dec 177
		4	Dec 126		Jan 157–Dec 187
		5	Dec 137		Jan 168–Dec 198
		6	Dec 152		Jan 183–Dec 213
		7	Dec 168		Jan 199–Dec 229
		8	Dec 185		Jan 216–Dec 246
		9	Dec 199		Jan 230–Dec 260
1979–2009		10	Dec 214		Jan 245–Dec 275
midHolocene 1–100	piControl	1	Dec 80	300 years	Jan 381–Dec 480
rcp45	historical	1	Dec 2005 of *E1*		Jan 237–Dec 331 (531[†††])
rcp85		2	Dec 2005 of *E2*		Jan 248–Dec 342 (542[†††])
rcp26		3	Dec 2005 of *E3*		Jan 260–Dec 354 (554[†††])
rcp60		4	Dec 2005 of *E4*		Jan 273–Dec 367
		5	Dec 2005 of *E5*		Jan 283–Dec 377
		6	Dec 2005 of *E6*		Jan 294–Dec 388
		7	Dec 2005 of *E7*		Jan 309–Dec 403
		8	Dec 2005 of *E8*		Jan 325–Dec 419
		9	Dec 2005 of *E9*		Jan 342–Dec 436
2006-2100(2300) ^^		10	Dec 2005 of *E10*		Jan 356–Dec 450
1pctCO2 1–140	piControl	1	Dec 103		Jan 104–Dec 243
sstClim sstClim4xCO2 sstClimAerosol sstClimSulfate 1–30					Jan 104–Dec 133

Experiment	*Parent experiment*	*Ens.*	*Branch point*	*Spin-up*†	*Alignment*††
abrupt4xCO2	piControl	1	Dec 103		Jan 104–Dec 253
		2	Jan 104		Feb 104–Jan 109
		3	Feb 104		Mar 104–Feb 109
		4	Mar 104		Apr 104–Mar 109
		5	Apr 104		May 104–Apr 109
		6	May 104		Jun 104–May 109
		7	Jun 104		Jul 104–Jun 109
		8	Jul 104		Aug 104–Jul 109
		9	Aug 104		Sep 104–Aug 109
		10	Sep 104		Oct 104–Sep 109
E1: 1–150		11	Oct 104		Nov 104–Oct 109
E2–12: 1–5		12	Nov 104		Dec 104–Nov 109

Appendix 2. Datasets available

CMIP5 datasets have been prepared and published on the Earth System Grid for the parameters listed in Table 3. Three-dimensional atmospheric datasets use the vertical coordinate specified by CMIP5, which in most cases is either the model's native vertical coordinate or a prescribed set of pressure levels.

Table 3. CSIRO-Mk3.6 datasets.
Notes: † model levels; †† pressure levels; ††† data provided at 2 m (CMIP5 requested data at 10 m); ^ provided for all experiments, including AGCM experiments; ^^ data not available for years 1–36; ^^^ corresponds to years 1979–2009 in the historical experiments; ~ data for E1–E3 of RCPs 4.5 and 8.5 ends at 2100; ~~ pixels containing sea-ice are masked out; * time slices correspond to years 1986–2012 in the historical experiments; ** one-year time slice taken every decade; *** time slices correspond to years 1850, 1870, 1890, 1910, 1930, 1950, 1960, 1970, 1980, 1990, 2000, 2010, 2020, 2040, 2060, 2080 and 2100 in the historical experiments. E1, E2 etc. denotes the ensemble member in a multi-member ensemble; AGCM denotes atmosphere-only experiment; and AOGCM denotes coupled atmosphere-ocean experiment.

See over page.

Frequency	*Type*	*Variable(s)*	*Data available* *Experiments*	*Ens.*	*Years*
Six-hourly	2D	*Atmosphere variables* Sea level pressure.	All historical AMIP RCPs 4.5/8.5 piControl midHolocene	E1-10 E1-10 E1-10 E1 E1	1950–2005/2012 1979–2009 2006–2100 210–240^^^ 71–100
		Surface air pressure.	All historical AMIP RCPs 4.5/8.5	E1 E1 E1	1950–2005/2012 1979–2009 2006–2100
	3D m-levels†	*Atmosphere variables* Air temperature, eastward wind, northward wind, specific humidity	All historical AMIP RCPs 4.5/8.5	E1 E1 E1	1950–2005/2012 1979–2009 2006–2100
	3D p-levels††	*Atmosphere variables* Air temperature, eastward wind, northward wind	All historical AMIP RCPs 4.5/8.5 piControl midHolocene	E1-10 E1-10 E1-10 E1 E1	1950–2005/2012 1979–2009 2006–2100 210–240^^^ 71–100
Daily	2D	*Atmosphere variables* Daily maximum near-surface air temperature, daily-mean near-surface wind speed, daily minimum near-surface air temperature, near-surface specific humidity, near-surface air temperature, precipitation, sea level pressure.	All	All	All
		Ocean variables Sea surface temperature^{--}.	All	All	All
		Atmosphere variables Convective precipitation, daily maximum near-surface wind speed, eastward near-surface wind, near-surface relative humidity, northward near-surface wind, solid precipitation, surface daily maximum relative humidity, surface daily minimum relative humidity, surface downwelling longwave radiation, surface downwelling shortwave radiation, surface upward latent heat flux, surface upward sensible heat flux, surface upwelling longwave radiation, surface upwelling shortwave radiation, top-of-atmosphere outgoing longwave radiation, total cloud fraction.	All historical AMIP RCPs 4.5/8.5 RCPs 2.6/6.0 piControl	E1 E1 E1 E1 E1	1950–2005/2012 1979–2009 2006–2100, 2181–2200, 2281–2300 2006–2100 217–243*
		Land variables Moisture in upper portion of soil column, snow area fraction, surface snow amount, total runoff.	…	…	…
		Ocean variables Sea-ice area fraction, sea-ice thickness.	…	…	…
	3D p-levels	*Atmosphere variables* Air temperature, eastward wind, northward wind, omega, relative humidity, specific humidity.	All historical AMIP RCPs 4.5/8.5 RCPs 2.6/6.0 piControl	E1 E1 E1 E1 E1	1950–2005/2012 1979–2009 2006-2100, 2181–2200, 2281–2300 2006–2100 217–243*

Frequency	Type	Variable(s)	Data available: Experiments	Ens.	Years
Monthly	2D	*Aerosol variables*			
		Ambient aerosol absorption optical thickness at 550 nm, ambient aerosol optical thickness at 550 nm, ambient aerosol optical thickness at 870 nm, ambient fine aerosol optical thickness at 550 nm, cloud droplet number concentration of cloud tops, cloud-top effective droplet radius, column integrated cloud droplet number, dry deposition rate of dust, dry deposition rate of SO_2, dry deposition rate of SO_4, emission rate of black carbon aerosol mass, load of black carbon aerosol, load of dry aerosol organic matter, load of dust, load of seasalt, load of SO_4, rate of emission and production of dry aerosol total organic matter, surface concentration of black carbon aerosol, surface concentration of dry aerosol organic matter, surface concentration of dust, surface concentration of SO_4, surface diffuse dowwelling shortwave radiation, total direct emission rate of SO_4, total emission of primary aerosol from biomass burning, total emission rate of dms, total emission rate of dust, total emission rate of SO_2, wet deposition rate of dust, wet deposition rate of SO_2, wet deposition rate of SO_4.	All	All	All
		Atmosphere variables (other than aerosol variables)			
		Condensed water path, convective precipitation, daily maximum near-surface air temperature, daily minimum near-surface air temperature, eastward near-surface wind, evaporation, ice water path, near-surface air temperature, near-surface relative humidity, near-surface specific humidity, near-surface wind speed[†††], net downward flux at top of model, northward near-surface wind, precipitation, sea level pressure, snowfall flux, surface air pressure, surface downward eastward wind stress, surface downward northward wind stress, surface downwelling clear-sky longwave radiation, surface downwelling clear-sky shortwave radiation, surface downwelling longwave radiation, surface downwelling shortwave radiation, surface temperature, surface upward latent heat flux, surface upward sensible heat flux, surface upwelling clear-sky shortwave radiation, surface upwelling longwave radiation, surface upwelling shortwave radiation, top-of-atmosphere incident shortwave radiation, top-of-atmosphere outgoing clear-sky longwave radiation, top-of-atmosphere outgoing clear-sky shortwave radiation, top-of-atmosphere outgoing longwave radiation, top-of-atmosphere outgoing shortwave radiation, total cloud fraction, water vapor path.	All	All	All
		Land variables			
		moisture in upper portion of soil column, snow area fraction, snow depth, surface runoff, surface snow amount, total runoff, total soil moisture content.	All	All	All
		Ocean variables			
		downward heat flux at sea water surface, ocean barotropic mass streamfunction, ocean mixed layer thickness defined by sigma T, sea-ice area fraction[^], sea-ice thickness[^], sea surface height above geoid, sea surface salinity, sea surface temperature[~~], snow depth, surface downward x-stress, surface downward y-stress, surface snow area fraction, water flux into sea water, water flux into sea water from rivers[^^], x-component of sea-ice mass transport, y-component of sea-ice mass transport.	All AOGCM	All	All

Frequency	*Type*	*Variable(s)*	*Data available*		
			Experiments	*Ens.*	*Years*
	3D m-levels	*Aerosol variables* Concentration of black carbon aerosol, concentration of dry aerosol organic matter, concentration of dust, concentration of SO_4, mole fraction of DMS, mole fraction of SO_2.	All historical AMIP All RCPs piControl	E1-10 E1-10 E1-10 E1	1850–2005/2012** 1980–2009** 2010, 2020, 2040, 2060, 2080, 2100 81, 101, 121, 141, 161,181, 191, 201, 211, 221, 231, 241, 251,271, 291, 311, 331***
		Atmosphere variables (other than aerosol variables) Cloud area fraction, convective mass flux, mass fraction of cloud ice, mass fraction of cloud liquid water.	All	All	All
		Ocean variables Ocean mass x-transport, ocean mass y-transport, sea water potential density, sea water potential temperature, sea water salinity, sea water x-velocity, sea water y-velocity, upward ocean mass transport.	All	All	All
	3D p-levels	*Atmosphere variables* Air temperature, eastward wind, geopotential height, mole fraction of ozone, northward wind, omega, relative humidity, specific humidity.	All	All	All
	Area average	*Ocean variables* Global average sea water potential temperature, global average thermosteric sea level change~, global mean sea water salinity.	All AOGCM	All	All
Fixed	2D	Atmosphere grid-cell area^, fraction of grid cell covered with glacier^, land area fraction^, ocean grid-cell area, region selection index, sea area fraction, sea floor depth, surface altitude^.	All AOGCM	N/A	N/A
	3D	Ocean grid-cell volume, ocean model cell thickness.	All AOGCM	N/A	N/A

Appendix 3. Data versions

CMIP5 datasets include a version number to enable users to track modifications to the data. The version number is included in the HTTP path to each file and (depending on the modelling group), may also be specified in the metadata in each file. The path is visible in wget download scripts and is also accessible via the ESG web interface (see the 'file details' section once a dataset has been selected). The version number that is displayed when browsing the CMIP5 catalogue is derived from the date when the dataset was added to the catalogue; it may have no relation to the data version.

Table 4. Data versions in the CSIRO-QCCCE CMIP5 archive.
Note: † The vertical interpolation technique was modified so that extrapolation was used to generate data at the 1000 hPa level.

Version	*Description*
v20110518	Default
v20110829	Monthly near-surface air temperature (Amon:tas) datasets for the historical and rcp45 (2006–2100) experiments were replaced.
v20111029	Daily relative humidity (day:hur) datasets replaced.
v20111123	Monthly condensed water path (Amon:clwvi) datasets replaced.
v20111221	Daily sea-ice area fraction (day:sic) datasets replaced.
v20111222	Daily sea surface temperature (day:tos) datasets replaced.
v20120213	Monthly sea surface temperature (Omon:tos) datasets replaced.
v20120323	All six-hourly, daily and monthly atmospheric data on pressure levels were replaced†. Monthly and daily snow area fraction (OImon:snc, LImon:snc) datasets replaced.
v20130205	Monthly mass transport (Omon:umo, Omon:vmo and Omon:wmo) datasets replaced.