Synthesis WCRP Coupled Model Intercomparison Project Phase 5 (CMIP5) Survey

Veronika Eyring (DLR, Germany) and Ron Stouffer (GFDL)

Incoming and Current CMIP Panel Chairs

WGCM 17th Session

World Climate Research Programme

October 1-3, 2013

Victoria, Canada

DLR

Overview

I. Background and Goal CMIP5 Survey

II. Synthesis of Responses we have received so far

1. CMIP5 experiment design
2. Emissions / forcing for historical and RCP experiments
3. Standard output and specially prepared output
4. Model and experiment documentation
5. Data search and support
6. Timeline
7. Key science gaps that should be addressed by CMIP6
8. What else is missing?
9. Additional Comments

III. Summary and Way Ahead

I. Background and Goal CMIP5 Survey

WCRP Coupled Model Intercomparison Project Phase 5 (CMIP5) Survey

Prepared by the CMIP Panel and Co-Chairs of the Working Group of Coupled Modelling (WGCM)

Veronika Eyring, Ron Stouffer, Sandrine Bony, Natalie Mahowald, Jerry Meehl, Bjorn Stevens, Karl Taylor

Background:

- As most of the simulations for CMIP5 have been completed, and their analysis is in full swing it seems timely to ask, while experiences are still fresh, as to what went well, what didn't, and what gaps in the science are emerging.
- In particular, are gaps emerging that could be filled or bridged by a coordinated set of model experiments, and thus should be considered as a component of CMIP6.
- With an eye on CMIP6, a workshop "Next generation climate change experiments needed to advance knowledge and for assessment of CMIP6" was held early August 2013 in Aspen, USA.
- This workshop is the first, of what we hope will be a series of workshops and meetings, to assess the accomplishments and outstanding issues with the CMIP5 process and will help inform the design of CMIP6.

World Climate Research Programme

WCRP Coupled Model Intercomparison Project Phase 5 (CMIP5) Survey

Goal of this Survey: To learn from those most active in CMIP5 what went well and what didn't and to provide input for the Aspen meeting and future CMIP6 planning workshops.

Addressees: This survey was sent to representatives of the climate community on *28 June 13*

1. Point of Contacts **CMIP5 model groups**

USERS:

2. Co-Chairs **WCRP working groups** (http://www.wcrp-climate.org/index.php/about-implementation/28-mplementation-working-groups)
3. Co-Chairs **WGCM-Endorsed Community Coordinated Project** (http://www.wcrp-climate.org/wgcm/cmip_coord.shtml)
4. Co-Chairs **Model Intercomparison Project** (http://www.wcrp-climate.org/wgcm/projects.shtml)
5. Point of Contacts **Integrated Assessment Model groups**
6. Co-Chairs **Related IGBP group or activity**
7. Contacts **ESG Federation**
8. Members of **Climate Services Partnership Coordinating Group** (http://www.climate-services.org/coordinating-group)
9. Some selected **Additional People**

Outcome: Synthesis and dissemination of the survey and consideration of the responses in the planning of CMIP6.

WCRP World Climate Research Programme

Timeline

- First presented at the workshop “Next generation climate change experiments needed to advance knowledge and for assessment of CMIP6” in Aspen, August 2013
- Now updated for the WGCM meeting with several more responses
  - If you want to contribute to this survey and haven’t yet done so, or if you want to update some of your responses, please email *Veronika Eyring (Veronika.Eyring@dlr.de) and Ron Stouffer (Ronald.Stouffer@noaa.gov)*
  - If we missed one of your important points in the synthesis of the survey, please also contact us.

- An up-to-date version of this synthesis will be placed on the CMIP Panel website at **http://www.wcrp-climate.org/wgcm/cmip.shtml**
- Eventually turn this into a peer-reviewed paper on “CMIP5 lessons and looking forward to CMIP6”

World Climate Research Programme

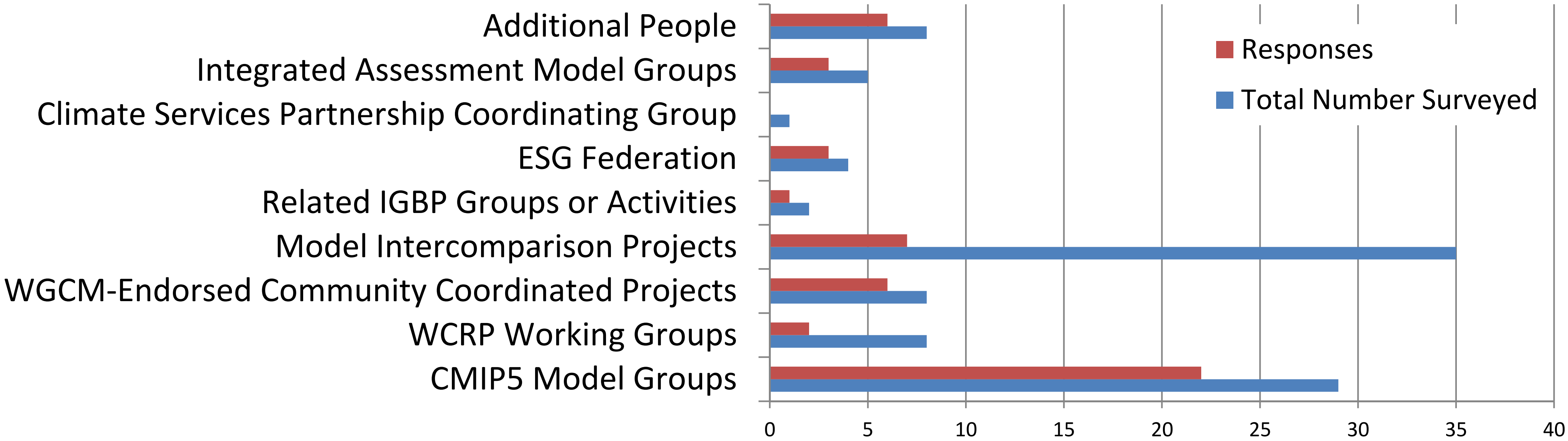
II. Synthesis of Responses we have received so far

5. CMIP5 experiment design
6. Emissions / forcing for historical and RCP experiments
7. Standard output
8. specially prepared output
9. Model and experiment documentation
10. Data search and support
11. Timeline
12. Key science gaps that should be addressed by CMIP6
13. What else is missing?
14. Additional Comments

Responses received so far (Date: 1 October 2013)

Virtually all responses very positive on the value of CMIP

Lots of enthusiasm to improve/change CMIP experiments



Total	101	**Responses**	50	50%

Many thanks to everyone who replied!

Your comments (some very detailed) will be very useful in the planning of CMIP6

World Climate Research Programme

5. CMIP5 experiment design (see Taylor et al., BAMS, 2012)

(i) General scientific focus like balance between simulations primarily focused on projections versus those designed more for advancing understanding

INDICATE WHAT WENT WELL?

CMIP5 Model Groups	Users
Overall: • **CMIP5 has created an enormous resource and enabled a large amount of science.** This science has fed into many papers, and hence has been influential in IPCC AR5 (across working groups). • **The balance was ok, others more critical** • **Even though development of the RCP process very complex – across all 3 IPCC WGs – it produced 4 fairly reasonable** future scenarios for GHG **Simulations** • **The idealized experiments** were extremely useful for model evaluation, conceptual work, and as a benchmark for comparing with other models or comparing among other experiments.	**Overall:** • **This was a success, resulting in a breadth of experiments run with consistent model versions. Large number of coherent simulations allow to study a lot of things.** **Simulations** • **Large number of idealized simulations very useful.** • **Emissions vs concentration simulations very useful for understanding carbon cycle uncertainty in RCP simulations.** **Models** • **High-top models and models with interactive carbon cycle, aerosols, chemistry part of CMIP5** **Number of Models and Simulations** • **Large number of models helped in the assessment of the robustness of future projections.**

WCRP World Climate Research Programme

5. CMIP5 experiment design (see Taylor et al., BAMS, 2012)

(i) General scientific focus like balance between simulations primarily focused on projections versus those designed more for advancing understanding

INDICATE WHAT WENT NOT SO WELL?

CMIP5 Model Groups	Users
Simulations • **Given their limited cost, idealized experiments could have been encouraged more strongly** and/or multiplied (e.g., aquaplanet). • **Scenario runs gave at most an illusion of greater understanding**. Settling on one standard scenario and complementary, but more idealized, experiments (8xCO2, ramp-up/ramp-down, specified land cover changes), would be much more useful than having many scenarios. • **Probably should also have had the low-end case (in this case RCP2.6) as a core simulation**, as it has proven very widely used in the literature.	**Simulations** • **Too many RCPs (e.g., one medium scenario would have been sufficient)** • **AMIP simulations not systematic** • **Abrupt4xCO2 should have been run to equilibrium** **Number of Models and Simulations** • **Large number of experiments is overwhelming for many groups, and results in a sparsely populated matrix**, which makes analysis of many interesting questions difficult.

World Climate Research Programme

5. CMIP5 experiment design (see Taylor et al., BAMS, 2012)

(i) General scientific focus INDICATE SUGGESTIONS FOR IMPROVEMENTS / FOCUS FOR CMIP6

CMIP5 Model Groups

Simulations

- **Even stronger focus on understanding (i.e. idealized experiments) and model evaluation**
- **The whole CMIP effort could be split into several –MIPs**, so as to allow individual groups to take part only to those experiments which are of interest.
- **Some of the new experiments were still at a pioneering stage. In CMIP6, we should repeat these experiments, revising the experimental design** based on the lessons we learned

Number of Models and Simulations

- Consider a reduction in the number of simulations – both the mandatory core and other tiers (for example, are 4 scenarios instead of 2 that useful?)

Users

Simulations

- **More focus on idealized experiments to tackle CMIP5 science gaps**, in close **connection with WCRP Grand Challenges**
- **Make the core idealized experiments easier for modeling groups to run** – e.g. simplify idealized experimental design to reduce the need different SST forcings
- **Ensure, as much as possible, continuity with CMIP5.** CMIP6 should be an opportunity for modelling groups to provide the CMIP5 expts/outputs they did not provide at the time of CMIP5.
- **Special focus on the "Now" period (1970-2030**)
- **Greater focus on diagnosing aerosol forcing and in performing standard simulations driven by aerosols alone.**

Number of Models and Simulations

- **Keep the core small / reduce number of simulations and increase the level of mandatory cases to have enough models in those**.
- **Submitting several near identical versions of a model not very informative.**
- **Highlighting which experiments would benefit from many model centers performing them and which could be satisfied by a smaller number**. More generally the question of how many models we need is an open question.

5. CMIP5 experiment design (see Taylor et al., BAMS, 2012)

(ii) Multi-tier approach

- **The core/tiers approach was a very good choice and should be maintained**

- The large number of experiments had very large resource implications. While the prioritization in "tiers" was helpful, there was still a sense that doing more was better.
- Large concern on the data storage needs for CMIP6.
- Single set of core experiments, with a single second tier, rather than near-term and long term experiment categories, both with their own tiers and cores.
- Balance between "Tier-1" and "Tier-2" experiment groups were not well designed. The number of "Tier-1" experiments was pretty heavy compared to "Tier-2".
- Modeling centers with limited resources should be encouraged to contribute usefully (e.g. with "large enough" ensembles and/or "high enough" resolution) to some experiments (e.g. D&A and/or decadal) rather than to do a bit of everything.

- **Strengthening and slim-lining the multi-tiered approach**
- identifying a few **scientific goals** for the intercomparison would be helpful
- i.e., a stronger but more conceptual and **hence small group of core experiments**; with a **second (but not third) tier of experiments reflecting a select group of important emerging science questions**.
- Here taking advantage of the **WCRP grand challenge** initiative seems particularly opportune.

World Climate Research Programme

5. CMIP5 experiment design (see Taylor et al., BAMS, 2012)

(iii) Decadal/long-term/atmosphere-only options

- Good but, given their cost, decadal hindcasts could appear as tier experiments.

- **Coordinated experiments work most poorly a coordinating body comes up with a new experimental design, largely untested and not well constrained, that seeks to break new scientific ground.**
- The decadal prediction experiments did not seem as well thought through as one would have liked.
- The documentation on the experimental design was not sufficiently clear (e.g. starting points).
- The 30-year initialized runs are not that useful. Better many more start dates and shorter
- Groups were asked to increase the number of start-dates (from every 5-year to – possibly - every year) at a very advanced stage of the "production" of CMIP simulations
- Large amount of effort went in to create the high frequency CMIP5 decadal prediction output. Do we really need so many sub daily fields?

Recommendations WGSIP (to be further discussed):

- **Start initialized decadal predictions on 1 November** to allow analysis of the first winter and indeed the first year without weather predictability giving a false impression of long range predictability.
- **Start initialized predictions every year rather than every 5 years** for statistical sampling reasons
- **Decadal experiment with idealized future volcanic eruption**
- Given that there is scant evidence for any skill from initialization beyond 5 year lead times could we **not simply make 5 year hindcasts next time for near term prediction**?

Others:

- More model variations (High/low res of same model / More models using different parameterizations)
- Standardized metrics and methods for verification

5. CMIP5 experiment design (see Taylor et al., BAMS, 2012)

(viii) Adequacy of experiment descriptions

Overall good; What went not so well:

- **There were inevitable departures from the experiment design/protocol** by some groups. The experiment plan was a reasonable overview for scientists, but **hides some complexity** when used as a specification for configuring and managing model runs. => **hidden work at some modelling centres, inconsistencies in the delivered data**
- Changes of experiment protocol (addition of decadalYYY[2-4,6-9], extension of historical) when the simulations were in operation was somehow disturbing
- **Processes that finalize the CMIP5 protocol were not very open to all the modeling centers**
- Point of confusion was the use of b**oth an experiment numbering scheme and experiment naming scheme.**
- The specification of anthropogenic aerosols in AMIP experiments could have been precised.
- Better specification of possible MISC D&A experiments.

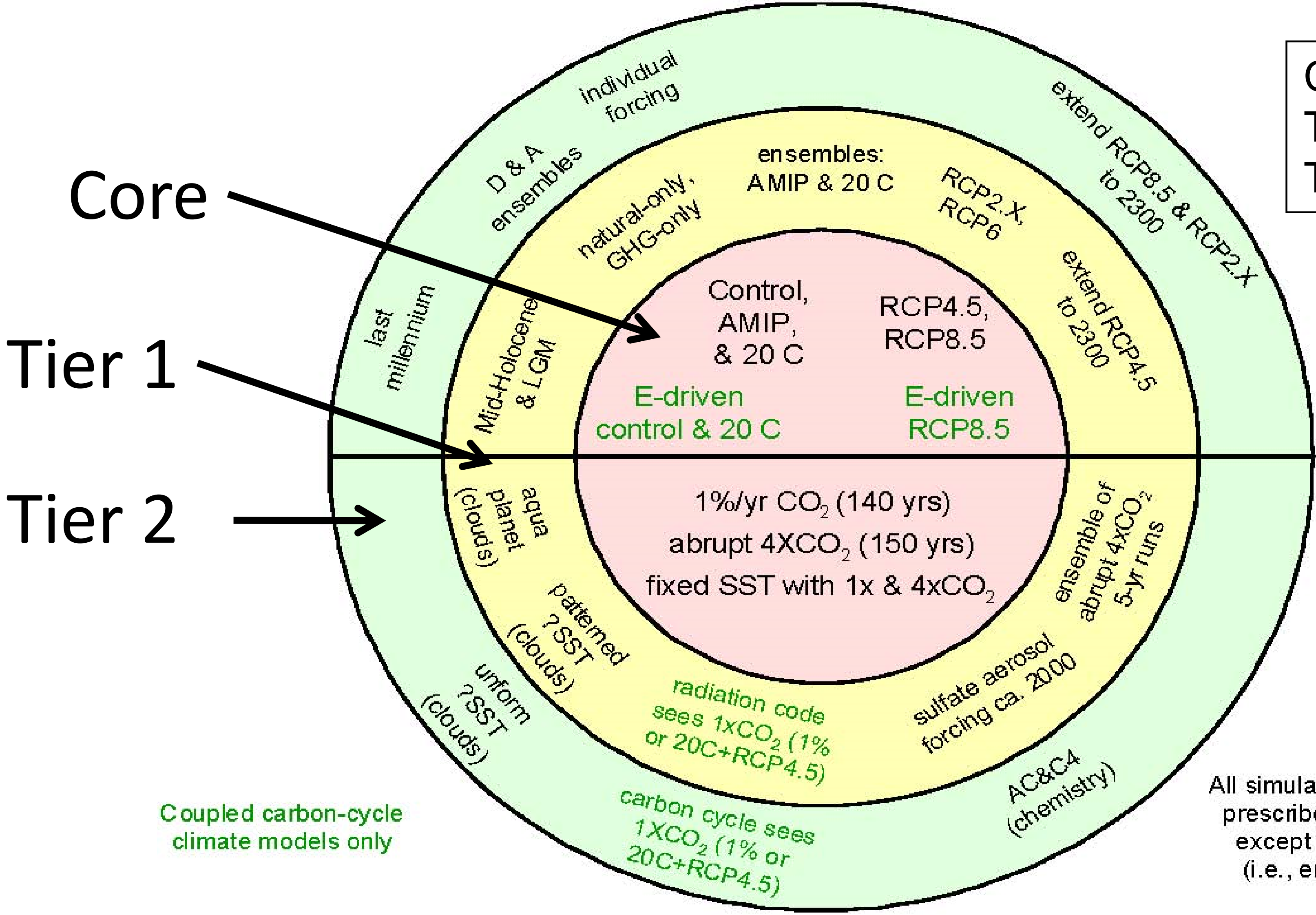
WCRP World Climate Research Programme

5. CMIP5 experiment design (see Taylor et al., BAMS, 2012)

(iii) Missing Experiments

- **More long term experiments** (SSPs included for CMIP6, aerosol range, LU range, CH4 range, 3.7w/m**2 forcing exp, overshoot exp, target temperature scenarios, …)
- **D&A – individual forcing runs:** much more of an emphasis on attribution of changes over the past century, with a lot more attention paid to AMIP simulations, to larger ensembles of historical coupled runs, and to the length of the control simulations provided.
- **For AMIP: single forcing attribution runs**, runs with fixed forcing, runs with the inclusion of the carbon cycle and atmospheric chemistry, and runs focusing on stratospheric ozone. All attribution runs require large ensembles (> 10), even in AMIP mode, to be optimally useful.
- **Aerosols – experiments to yield more details of response**
- **Additional simulations to estimate forcings** and GWP should be added
- **Component-only experiments like AMIP – Ocean (CORE) and land surface (GSWP)**
- **Idealized Runs also for ESMs (e.g.** ESMs with idealized carbon emissions) Idealized runs with the full models rather than aqua-planet configurations more appropriate for CMIP
- PMIP runs with full carbon cycle.
- **Decadal Prediction** – future volcanoes rethought, data assimilation exp, SI-like exp
- Future volcanoes in long term exp.
- Parametric studies – table of IC vs model parameters – evaluate impact on simulation

CMIP5 Long Term Experiments



Core:	≥1718 yrs
Tier 1:	≥1727 yrs
Tier 2:	≥2038 yrs

All simulations are forced by prescribed concentrations except those “E-driven” (i.e., emission-driven).

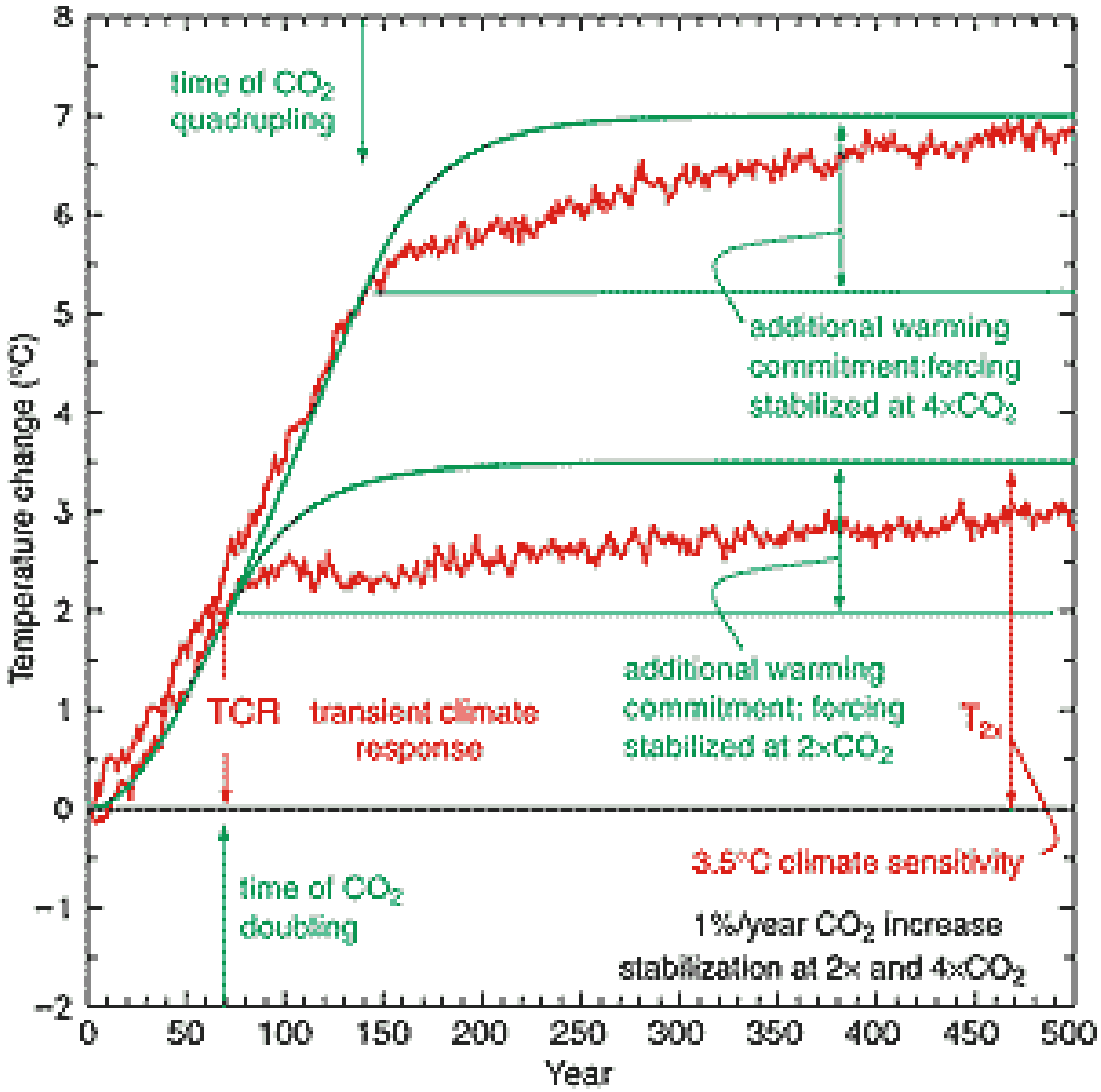
World Climate Research Programme

Core Idealized Experiments

- 1%
- 4XCO2 switch-on
- Atm-only using today's SSTs and 4xCO2 SSTs
- Suggestions
  - D&A – Run individual forcings alone
  - Looking for non-linear responses (How?)
    - 2x vs 8x?
  - Better radiative forcing estimates
    - Run atm-only using today and future SSTs
- Idealized aerosol exps using 1 RCP as base
- Investigating regional forcing/response relationship
- Design experiments which help identify cause for differences in model feedbacks
  - Radiative forcing:
    - Cloud = CFMIP
    - Snow/ice albedo = Who "owns" it?
    - Lapse rate = ?
    - Oceanic heat uptake = ?

WCRP World Climate Research Programme

Idealized 1% CO2 Increase

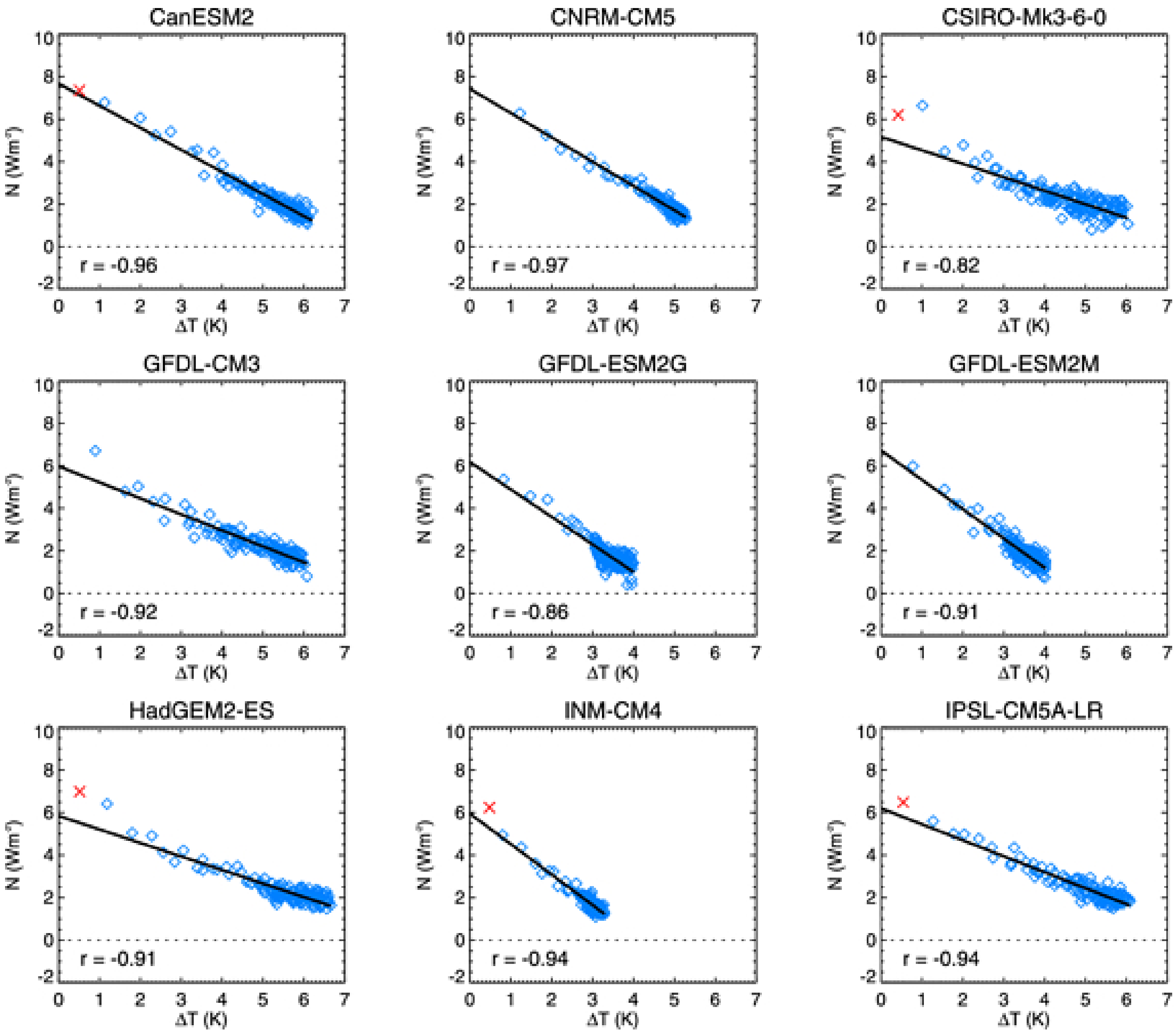


- Forcing nearly the same across models
- Easy to compare difference in model response
- Defines TCR
- Many interesting science questions

IPCC WG1 AR4

World Climate Research Programme

4xCO2 Switch-on



- Used to estimate equilibrium climate sensitivity
- Note: There are problems with method.

Andrews et al. GRL 2012

Suggestions Misc

- Resolution increase on atm side (<20km)
- Glacial-Interglacial – PMIP issue
- Observed extreme event studies
  - Initialize models 1-2 yrs prior past event
  - Investigate predictability
  - Covered if start dates become once or twice per year
- Ocean – diagnose eddy fluxes (WGOMD)
  - Param eddies/eddy permitting/eddy resolving

WCRP
World Climate Research Programme

Implementation of AMIP runs in CMIP6

- What to use for SST and SEA ICE fields?
  - Historical observations for SST
  - What to use for historical Sea Ice?
  - Where to get SSTs and Sea Ice?
    - Typically use center's own results (remember other models not available to end of process)
  - Monthly averaged coupled model output
    - Inter-model differences in lower boundary conditions (SST and Sea Ice) will affect results

World Climate Research Programme

6. Emissions / Forcing for historical and RCP experiments

INDICATE WHAT WENT WELL AND WHAT WENT NOT SO WELL?

- Switching from SRES to RCP was a good point.
- Common concentration and emissions data available for aerosols and short-lived species
- Future Land Use scenario exist
- Harmonization between future and past emissions
- CMIP5 models all consider past ozone depletion and future ozone recovery
- 4 RCP scenarios were simple and comprehensive in terms of GHG (but not aerosols)

- **Present-day (2005) too far in the past**
- **Forcing data were produced very late in the CMIP5 process.**
- **Radiative forcing not properly diagnosed**
- **Aerosol and land use future scenarios did not span uncertainty range**
- **Multiple realizations help address natural variability, but no counterpart in historical forcings**
- **Documentation of forcings was mixed, reasonable in some cases, not in others**
- **Land use implementations vary a lot from model to model => big impact on C changes**
- **Despite so many gases available, some major ones slipped through the net – e.g. gridded CO2 emissions**
- **No scientific check on consistency across datasets (e.g. does the land-use forcing match the biomass burning forcing regionally?)**
- **Some odd choices in forcing extensions** (such as zero wood harvest for RCP extensions that we are using for ESM integrations)
- **Not all models used specified forcing concentrations.**
- **Not all of forcing data were specified on the CMIP5 website (**e.g. stratospheric aerosols missing)
- **Standardized no policy scenario not available**
- **Possible changes in geographical distribution in the future not properly considered**
- Future Fossil Fuel emissions not gridded
- Land use beyond 2100 are not available

6. Emissions / Forcing for historical and RCP experiments

INDICATE SUGGESTIONS FOR IMPROVEMENTS / FOCUS FOR CMIP6

- **Provide standard ways/suggestions on diagnosing forcing in the models**
- **Establish a working group** early on in the process that defines how to properly diagnose forcings in CMIP6.
- **For CMIP6, all observed forcings should be updated until 2015**
- **Incorporate historical forcing uncertainty into experiment design**
- **Suggest the background level of stratospheric volcanic aerosol in the piControl be set at the mean over the historical period.**
- **Include RCPs featuring an appropriate plausible spread in future aerosol emissions trends.**
- **Revise the low-end scenario (currently RCP2.6) taking into account recent actual and likely near-future CO2 emissions increases.**
- **The forcings that the models actually see were not available in CMIP3 or CMIP5. They should be included in the standard diagnostics.**
- **New land use and aerosol scenarios**
- Natural emissions, biomass burning, require more attention in scenarios
- Volcanic forcing is assumed zero in the future which is not correct.
- More well-defined protocols for natural forcings (volcanic, solar).
- **Data available in timely manner and gridded**
- All models to include all forcings.
- Merging past/future forcing for “past1000” should be provided and not left to the groups.
- Provide spatial and temporal CO2 emissions for the future scenarios

7. Standard Output

Things that went well

- Clear description of most vbls
- Basic aerosol data archived
- Huge volumes of data available in stand format

Things that went poorly

- Software for moving data and/or analzing data without moving (BADC)
- Newer vbls (biogeochemical and aerosols) need more work
- Groups did not provide grid interpolators (hori and vert)
- Hard to interpolate unto common grid
- Ocean grid information incomplete (box edges, partial boxes – need volumes)
- Large amount of data to be processed
- No prioritization for atmospheric variables

WCRP World Climate Research Programme

7. Standard Output

Suggestions

- **Dedicated effort to anticipate and thoroughly address needed metadata**
- **Prioritization and better description**
- **Need a prioritisation of atmospheric fields**, to allow modelling groups to prioritise their processing. Suggest survey the usage of fields in CMIP5, as a guide to determine priority of fields for CMIP6.
- Better aerosol diagnostics
- Provide info on missing atm vertical coordinate
- Is high frequency short term data needed?
- Have an early release of files documenting model grids
- Need archived vbls to have observational framework/footing, particularly in land (and land C)
- Provide grid interpolators
- Save monthly mean 3-D fields of moisture transport uq and vq. The moisture transport is a key property, but can currently only be calculated (laboriously) for those periods where 6-hourly data are saved.
- Separate physical and biogeochemical variable tables
- Atm CO2 emission vbls need work – eg definition of natural and anthro land CO2 emissions
- For concentration runs, archive atm CO2 value – in restoring runs this value is different from the input CO2 value
- Some ocean biogeochem vbls missing – phosphorus example
- Standardize time axis (days or months or years since XXX)
- Standard chunk length (time)
- Packing option for vbls to save space

World Climate Research Programme

8. Specifically Prepared Output

Thing that went well

- COSP output useful
- 3-D radiative fluxes in 1X and 4X useful

Things that went poorly

- COSP buggy at first

Suggestions

- The inclusion of COSP has been very helpful and should be strengthened & continued.
- The potential value of additional "simulators" should be assessed
- Freeze simulator codes before modeling groups freeze model codes (Japan)

9. Model and experiment documentation

Things that went well

- Massive increase in quality of technical documentation
- Okay first attempt at model documentation

Things that went poorly

- METAFOR Late
- Hard to use
- No peer review of information
- In end, little help to IPCC authors
- Forms did not “fit” our model framework in many areas

Suggestions

- METAFOR / ES-DOC
  - Improve process, Improve match to model physics and biology
  - Finish development of software sooner
  - More group fill in forms (and sooner in process)
- Is it useful? Who “owns” it now?
- More information about the spin-up methodologies used by the modelling groups would help analysts study residual climate drifts in the piControl simulations
- There is a need for a community-wide consultation on how the models used in MIPs and in model-based climate studies should be documented. In my opinion, this requires a much broader view of "documentation" and some formalized online web- or wiki-based system to catalogue this information.
- Document the properties of a model version (as was attempted by the METAFORE effort for CMIP5). In addition, it is arguably more important to document the response/behaviour of the models

10. Data search and support

- Overall: Improve logistics of CMIP6, as a software, analysis and management problem. These things - how to get the data, how to analyze it, how to manage it - were fundamentally limiting in CMIP5 for many groups.

Things that went well

- **Distributed data sources**
- ESGF browser (later in process)
- Documentation

Things that went poorly

- **ESGF data download was just not working properly when it was needed.**
- Funding for nodes is expensive
- Finding rarely used aerosol data difficult
- Browser very difficult for non-climate scientist to use
- Late data (modeling centers late, unclear is missing data due to center or ESGF problem, buggy early software)
- WGET not working

Suggestions

- Governance needs settled soon – before CMIP6
- Have sfc temperature and precipitation easy to find for non-climate folks
- **Further improvements to download/scripting capabilities**
- Server-side data aggregation to reduce data transmission load
- **WCRP to help convince funders how important it is for the infrastructure to have stable funding.**
- Some comments that the concept of "downloading the data" is fundamentally wrong. Data should be located where its generated, with meta data, a doi, date stamp etc etc. Scripts should be able to access those data and pull down just the sub-data required.

11. Timeline

Things that went well

- General timeline went well (some responses, others see below…)
- There was plenty of time between CMIP3 and CMIP5. Problems with last minute development changes is a center management problem, not a CMIP problem
- Core physical data available in time for IPCC

Things that went poorly

- **From a perspective of IPCC AR5, the timeline of CMIP5 did not work out well:**
  - Small amount of time between the availability of CMIP data and IPCC deadlines
  - a CMIP5 archive was built up in parallel to the ESG at ETH, with a huge effort and very quickly which served many of the analysis for AR5, including the Atlas.
  - And had we all not written many of the CMIP5 papers ourselves, there would have been nothing to assess on CMIP5 in AR5.
- Experimental design details late
- List of output variables/units/etc still changing late in the process
- Caused problems for existing/archived CMORized data
- Some forcing vbls late
- There was no real timeline for model and experiment documentation

Suggestions

- **The timeline should be moved up 2 years, assuming that it will slip by a year.**
- **Set deadline for everything (also the forcings)**
- Vbls list compiled earlier, Exp design frozen at earlier date, Model output request must be made sufficiently far in advance to allow groups to implement and test diagnostics

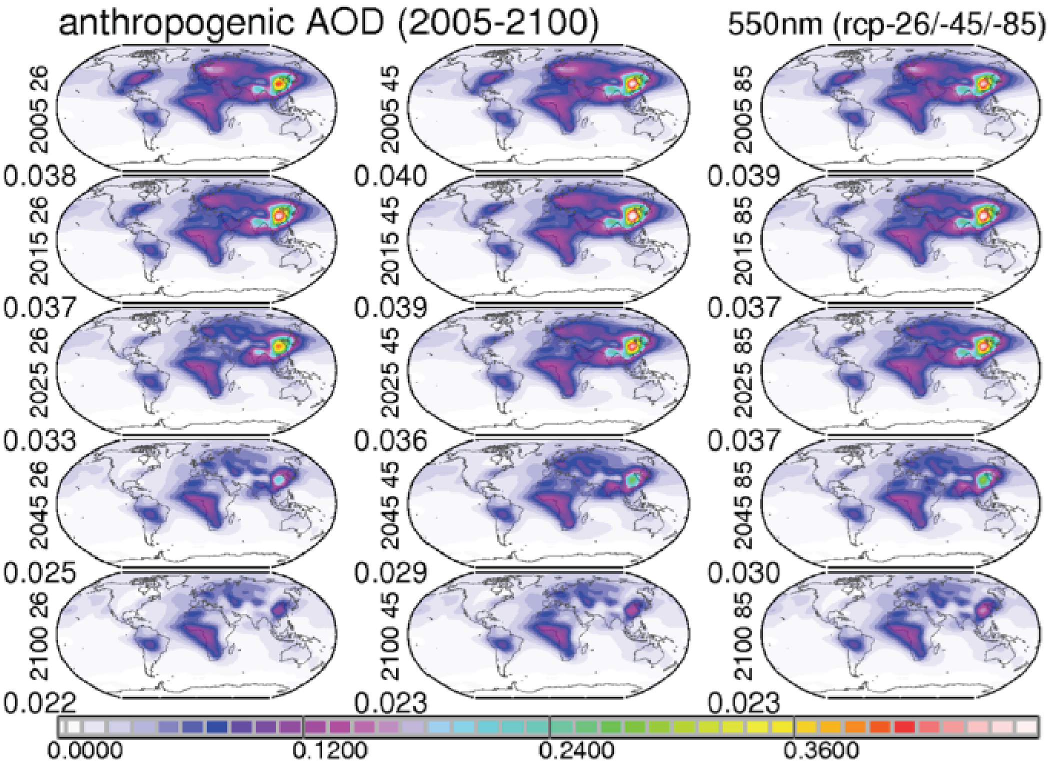
12. Science Gaps

- Some examples that have been mentioned in the responses –

Scientific Gaps 1: Aerosols, short-lived species, and methane

- **Aerosols do not span a good range in the RCPs**
- **Diagnosing indirect aerosol effects and absorbing aerosol effects**
- **More experiments** (like simplified aerosols) should be devoted to investigate the role of atmospheric aerosols

Aerosol Scenarios difficult to distinguish



Shown is mid-visible AOD for RCP 26 (left), RCP45 (middle) and RCP85 (right) for different timeslices in the future, ranging from 2005-2030 (top row) to 2100-2125 (bottom row). Global average is indicated by number on bottom left of each plot

Max-Planck-Institut für Meteorologie

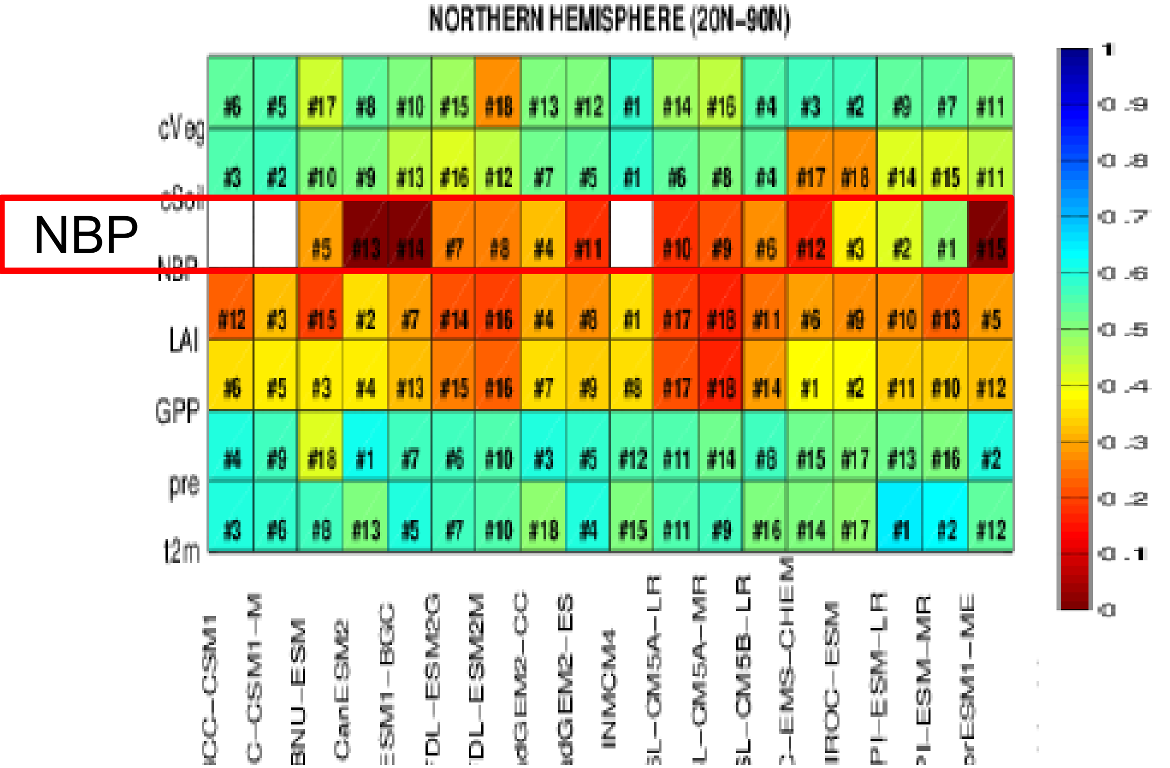
Kinne et al., *JAI*

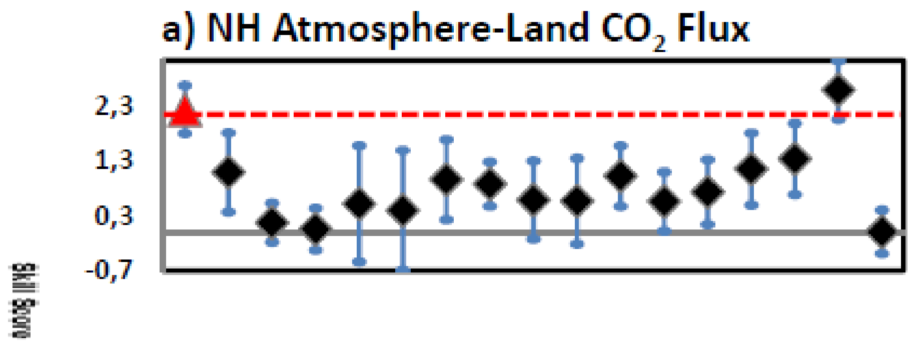
- Quantifying the role of chemistry and aerosols in near-time and long-term temperature projections
- Impact of methane emissions on climate

WCRP World Climate Research Programme

Scientific Gaps Example 2: Model evaluation, performance metrics

- **Detailed and systematic model evaluation** during the development process could be facilitated by CMIP and the metrics panel.
- **Obs4MIPs** seen as very positive way to improve regular model-obs evaluation, **should be grown.**
- The continued push for **standard performance metrics**, **readily published and viewable on a central website** (also for providing guidance for impact analyses).
- **Process-oriented evaluation** to understand model biases and error compensation.
- **Centrally coordinated model assessment**
- **Code repository** (e.g. at WGNE/WGCM Climate Metrics Panel Website)
- **CMIP Standardized diagnostic and performance metrics package that runs on the ESGF**





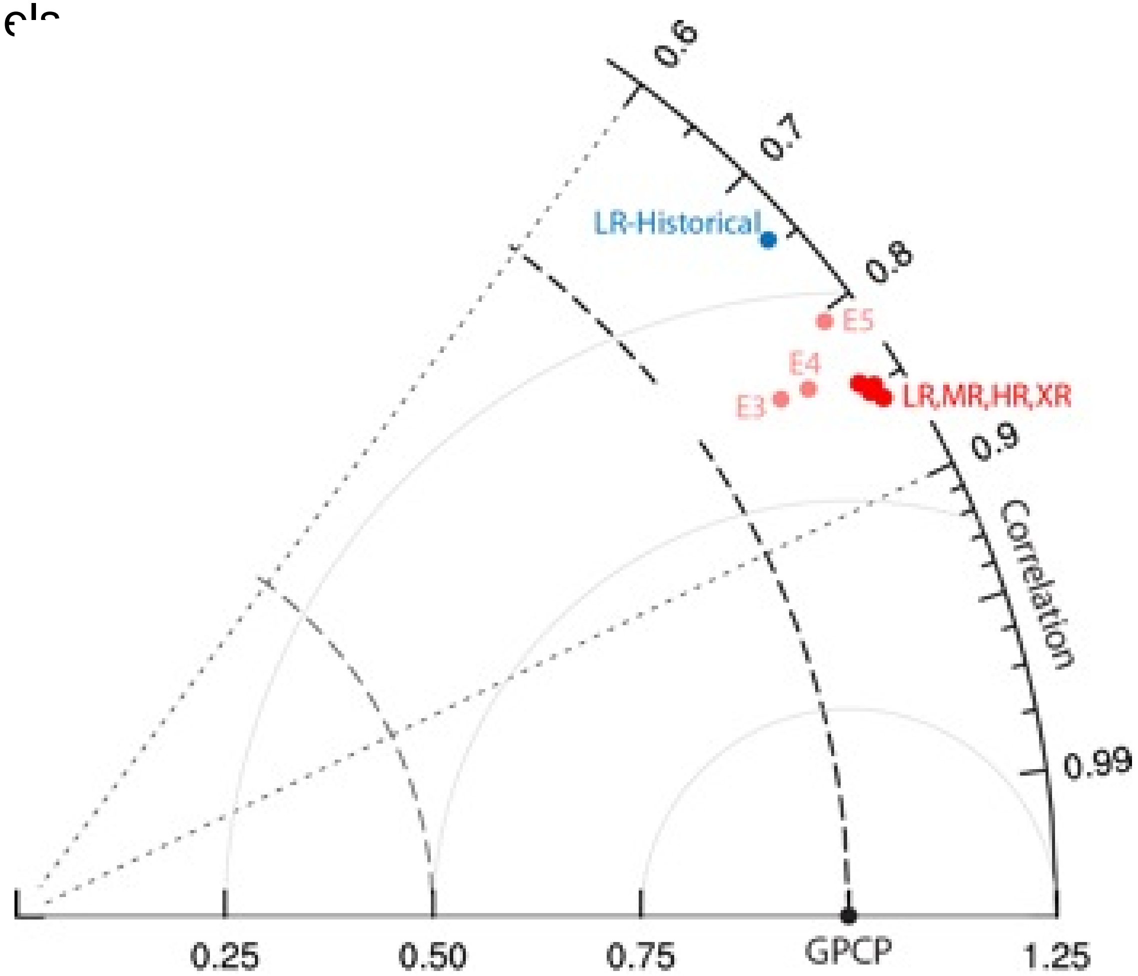
- In the NH over land almost all models systematically underestimate the sink, reflecting absence of sinks due to nitrogen deposition or poor representation of forest regrowth.

Anav et al., 2013

Scientific Gaps Example 3: Systematic Biases

- **Focus on persistent systematic errors could have been better (e.g. tropical biases)**

Example: biases in precipitation from four-five generations of ECHAM, dating back to ECHAM3 in the Early 1990s. It shows that improvements in the representation of precipitation have been very small. The red dots are AMIP runs, the blue is the coupled (worse of course). Typical behaviour also in other models



Stevens et al., 2013

World Climate Research Programme

Scientific Gaps Example 4: Better Diagnosing of Forcings

CMIP5/IPCC AR5 state of play

1. Offline **RF** estimates from MAGICC
2. CO2 **ERF** from Gregory method and 4xCO2 runs
3. Forster and Taylor method for time-varying **ERF** (historical and RCPs)
4. ACCMIP and AEROCOM time-slices for model subset and some fixed SST runs **(mix of RF and ERF estimates)**

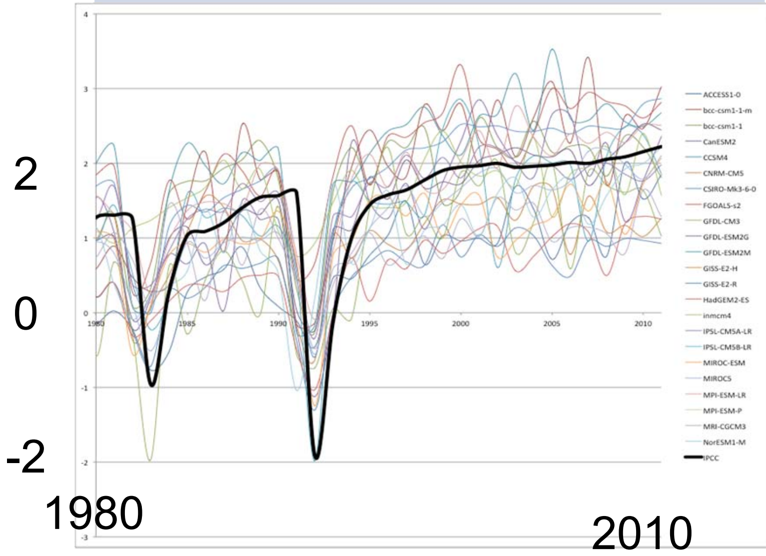
Slide Courtesy of Piers Forster

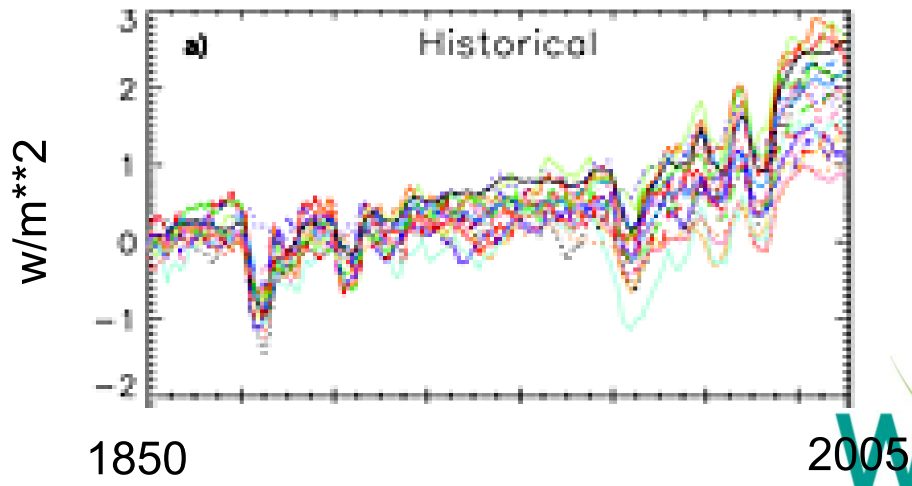
World Climate Research Programme

Scientific Gaps Example 4: Better Diagnosing Forcings

2.Consistency of approaches: Often not making a like-for-like comparison when comparing studies

Method	Pros	Cons
RF – double radiation call ideally with stratospheric adjustment	Good for offline tests of radiative transfer and line by line model comparisons	No one really does this now in a AOGCM environment - too much human time
ERF- from regression	Easy to do with standard diagnostics	Doesn't work for small forcings, noisy method, no regional forcing pattern
ERF from fixed SST	Less noisy than regression, gives regional patterns	Time-slice forcing only; very few runs done in CMIP5





Figures Courtesy of Piers Forster

World Climate Research Programme

Scientific Gaps Example 5: CO2 (ESM Issues)

- More and more ESMs with carbon cycle
- More and more analysis and publications
- **ESMs historical and future land carbon pools and fluxes are still embarrassingly all over the place.**
- Compared to AR4 (C^4MIP) more processes are included (land use change, nitrogen cycle) ☺ but this *artificially* enhances the models spread ☹
- Model evaluation is quite embarrassing…
- Obviously, not enough tuning/validation has been done in the model development phase (lack of time ?)
- What if … the way to get the carbon cycle right was to have it interactive so it could potentially screw up your simulation ;-), i.e. Emission driven runs should be the rule not the exception as it was in CMIP5.
- Better diagnostic of land use change (LUMIP)
- Specific simulations for GWP ?
- Diagnosis of land-use CO_2 fluxes was difficult in the CMIP5 experimental design – *how can we make it easier?*
- New components of carbon cycle models are likely for CMIP6 – more common use of N-limitations on land; better treatment of land-use and forest regrowth; O_3 effects on vegetation; more complex ocean ecosystem models with Fe-limitations.

Slide Courtesy of Pierre Friedlingstein and Peter Cox

World Climate Research Programme

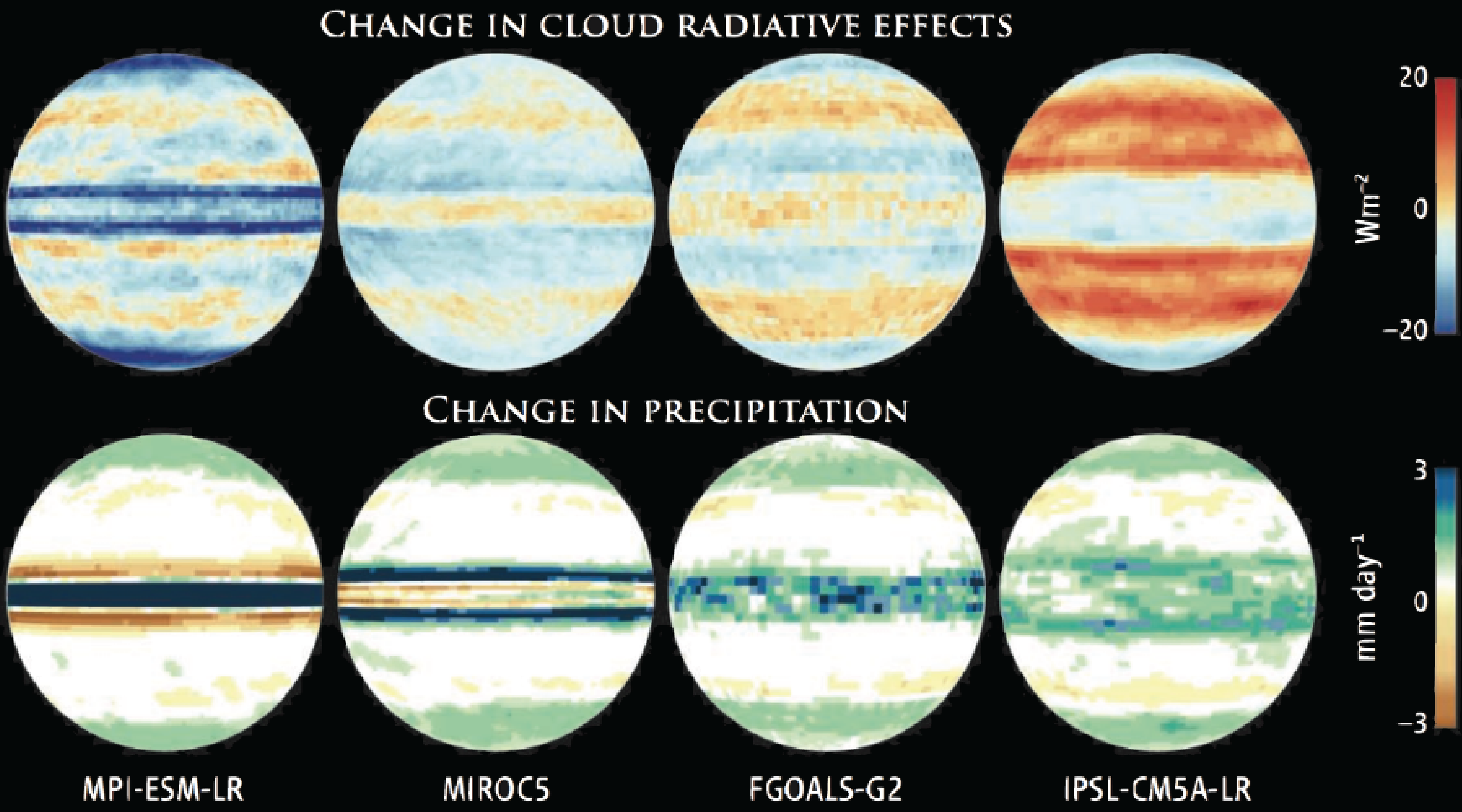
Scientific Gaps 6: Example Clouds (CFMIP)

- Incomplete implementation of CFMIP experiments/outputs
- Lack of process diagnostics in coupled idealized experiments.
- Inability to accurately diagnose time varying forcings and feedbacks from models during the 20th Century period
- Understanding the role of cloud and precipitation CO2 adjustments in coupled model response patterns
- Impact of clouds and cloud changes on regional temperatures, circulation and precipitation
- Identifying which parts of model physics contribute most to inter-model differences in cloud feedbacks.
- A CMIP approach to developing and testing physical hypotheses and observational constraints using sensitivity experiments.
- Identifying causal links between cloud controlling factors and cloud responses to climate forcings and warming. New idealized experiments?
- The gap in the experimental hierarchy between global GCM experiments and cloud resolving model/super-parametrization experiments is still large.

World Climate Research Programme

Coupled vs Earth System Modeling

Idealized Aquaplanet



Stevens and Bony, *Science* (2013)

Scientific Gaps Example 7: Decadal Predictions

- **Initialization Provides an Addition Source of Skill over some regions**
  - Skill Assessment Challenging
- **Provides Additional Information over the Forced Component Alone**
  - Particularly useful for the “warming hiatus”
- **Prediction Protocol Problematic**
  - 5-year vs. Every Year Initialization
  - Not Accurate Assessment of Skill – Volcanoes
- **Systematic Errors**
  - Large
  - Removing Systematic Error Not Simple
  - Surprising regions where initialization reduces skill
- **Main component is the production of decadal prediction hindcast data**
  - Extension and improvement of CMIP5 protocol
  - Actual forecast approach, i.e. no information from the future
  - Data protocol aligned with rest of CMIP with proposed addition of basic, quick access data set
- **Special purpose experiments under consideration but not (yet) formulated**
  - Sources and Mechanisms of Skill
  - Idealized Predictability …

Slide Courtesy of Ben Kirtman

WCRP
World Climate Research Programme

Scientific Gaps Example 8: Spread in model response

- Spread in model response still large even when natural variability taken into account even on the largest scales. Why?
    - Differing sub-grid scale physics implementation from model to model
        - Aerosols – some models predict own concentrations, some do not
        - Aerosol-cloud – some include parts of the indirect aerosol-cloud effect
            - Parts included vary model to model
            - None have “complete” aerosol-cloud physics (rjs opinion only!)
        - Land use change and more generally land carbon components vary a lot from model to model – especially important for C fluxes
        - Model resolution in all components
            - Particularly in the ocean (ocean eddies)
        -
        -
    - All of above will continue to be CMIP6 issues

- Importance of 1% CO2 increase run and potentially a standard 20C run

World Climate Research Programme

Scientific Gaps Example 9: Scenarios

- Covered earlier in talk
- Just list a few here
  - Aerosol and land use changes range in future
  - New future scenarios to 2100
  - Scenarios beyond 2100 (stylized/idealized?)
  - Historical/Future boundary (2005, 2010, 2013, 2015, ??)
  - Land ice scenario?
    - Coupling meltwater to ocean will have potentially large impact
    - GEIOD changes incorporated?

13. What else is missing?

Model Evaluation

- **Basic Evaluation of models along the way, see also previous slide**
- **International approach to evaluation** which is readily accessible would have been helpful. Establishment of such an approach in the near future would be helpful in the model development process for CMIP6.
- **Clearly models results were submitted before the modeling groups had a chance to look at their own results.**

Model Documentation

- **A coordinated description of the models in the literature**
- Inclusion of modeling groups in CMIP description.
- **More transparency in terms of model tuning/calibration**

Data handling

- **Better feedback to modeling groups on the use of their data.**
- **Some better ability to inform data users of when errata for published data sets are available**
- **A summary of available model output in the CMIP archive**: details of variables archived in each case, would be helpful.
- **DOI-like label for data**
- **Better recognition for modeling groups when their data is used in analyses (DOI, co-authorship).**
- **Having a common portal for accessing CMIP3, CMIP5 and CMIP6 data**, de-emphasizing different CMIP generations, and even perhaps having some simple capacity for filtering models based on meta data or climate indices so as to develop ensembles of data from the CMIP family would be very helpful.

14. Additional Comments

Model Groups	Users
• **CMIP5 has been a huge activity from concept up to present day and is ongoing – our many thanks to all who have made it so successful.** • The CMIP5 experimental design was very large, binding lots of resources. This was a chance to **get access to more resources** than the institute has available normally and to collaborate more closely across the institute. • **Possibly have major revisions to CMIP be once a generation (on a 15 year timescale)** and in between to augment the protocol around specific MIPs that address science gaps. • **“Recommendation report” for scientists, providing guidelines for future model developments and planning of the next CMIP**. • Maintain the **paleo-historical-future** link as key source data for weighting projections • **Data doi's** should be applied now with versioning. • encourage more frequent core additions to archive (every 2 years?) • Find ways to increase feedback from user analysis to model groups so it can affect development paths	• **From our perspective, CMIP5 has been a huge improvement on CMIP3, and we should celebrate its success.** • **Identify which data have been least accessed** to help prioritize diagnostics and reduce data volumes for transfer in CMIP6. • **Focus on attribution** • **There needs to be a clearer recognition of the importance of delivering societally relevant information efficiently.** • **IAMs employ socio-economic information in their scenarios, and increasingly GCMs-ESMs also use such information**. Ensure that, when running different IAM scenarios, the appropriate socio-economic variable s are consistent across the entire model. • Any MIP as part of CMIP6 should pre-determine what would constitute a model failure. what must a model be able to do to be included ? Any experiment needs some falsification criteria. • All models should be open source and freely available on-line for independent verification. Our models are now important - they affect major decisions nationally and internationally. • if data cannot be made available for commercial use free of charge, the responsible modelling groups should be strongly encouraged (perhaps required) to provide a mechanism for commercial groups to buy access.

Summary and Way Ahead

- Responses to the survey cover input from a variety of different communities
- The up-to-date version of this synthesis will be placed on the CMIP Panel website at **http://www.wcrp-climate.org/wgcm/cmip.shtml**
- Helps guiding the planning of CMIP6
- Thanks to everyone who has replied