The Cloud Feedback Model Inter-comparison Project - Plans for CFMIP-2

Mark Webb (mark.webb@metoffice.gov.uk)
Sandrine Bony (Sandrine.Bony@lmd.jussieu.fr)
George Tselioudis (gtselioudis@giss.nasa.gov)
Chris Bretherton (breth@atmos.washington.edu)

August 22, 2007

Summary of key issues for WGCM meeting (3-5 September, 2007).

The IPCC AR4 reaffirms the spread in equilibrium climate sensitivity and in transient climate response estimates among current models. Recent studies show that inter-model differences in cloud feedbacks remain the primary source of this spread, with low clouds making the largest contribution.

The main objective of CFMIP-2 is to make, by the time of the AR5, an improved assessment of climate change cloud feedbacks by making progress in the (1) evaluation of clouds simulated by climate models and the (2) understanding of cloud-climate feedback processes. An international CFMIP workshop was organized in Paris in Spring 2007 to lay the foundations for a CFMIP-2 proposal, and to strengthen the links between the CFMIP, GEWEX/GCSS and the US CLIVAR Process Team communities. The CFMIP coordination committee is now composed of: Mark Webb, Sandrine Bony, George Tselioudis and Chris Bretherton.

The main activities of CFMIP-2, which are detailed in the body of this document (most recent version available from http://www.cfmip.net) are:

1. Development of the ISCCP simulator and of a CFMIP ISCCP/CloudSat/CALIPSO simulator (CICCS) to be distributed to the modelling groups to evaluate model clouds (and thus contribute to the model development process) using satellite observations from the new generation of spaceborne radar and lidar instruments and existing passive instruments. They are required if effective cloud-climate model metrics are to be applied to CMIP4 GCMs.

2. Design and analysis of short atmosphere-only CFMIP-2 experiments, requiring CICCS and other diagnostics, to better understand the physical mechanisms underlying the different cloud-climate feedbacks in climate models.

3. Collaboration with GEWEX-GCSS to assess the credibility of cloud-climate feedbacks: CFMIP-GCSS CRM/LES/SCM case studies focused on the sensitivity of low-level clouds to changes in climate, process studies based on the analysis of grid-point high-frequency outputs, and development of cloud-climate metrics.

We ask whether WGCM continues to endorse CFMIP-2 plans, and seek the following recommendations from WGCM (see body of report for details):

- requiring the use of the ISCCP simulator (and strongly encouraging the use of CICCS) in AMIP, 20C3M and 1%/yr CO_2 (plus control) experiments of CMIP4,
- increasing the number of cloud diagnostics in the CMIP4 output, storing 3D model outputs on model levels, and 3D global fields daily for selected periods,
- hosting the CFMIP-2 experiments together with the CMIP4 archive,
- saving some high frequency instantaneous model output from selected point locations in CMIP4 and adding them to the standard output.

We also request some clarification on the plans for 1%/yr CO_2 and slab experiments in CMIP4, as this has implications for the design of the CFMIP-2 experiments.

We believe that granting these requests will help to reduce systematic errors in the simulation of clouds in the present-day climate and to assess the credibility of the different cloud feedbacks produced by GCMs, making it easier to assess the reliability of climate projections.

Revision history.

- Aug 22nd Final version. Added section summary at end of introduction. Updated items 2 and 5 in section 2.2. Spell check, minor corrections.
- Aug 21st Minor edits from Chris Bretherton. Added comment on other cloud types to section 1.2 item 3 in response to comment from Brian Soden. Added text suggested by Steve Sherwood to section 1.3 item 1. Added figures from Bretherton et al 2006 and Bony and Dufresne 2005.
- Aug 20th Calipso figure added, plus minor changes suggested by Sandrine. Added Package 3 to 1% runs.
- Aug 17th New version of WGCM summary and updated schematic provided by Sandrine.
- Aug 15th Minor changes.
- Aug 14th Some changes to summary, intro and evaluation sections
- Aug 13th Various modifications to sections 1.1 and 1.2 to make more consistent with the body of the document. Added metrics section, with figure from Williams et al 2006, and table of RMS statistics.
- Aug 10th version GPCI diurnal cycle figure added. CCTD figure added. 1st draft summary for WGCM.
- Changes in Aug 8th version. Updated ISCCP simulator section and added Williams and Tselioudis figure. Updated Minghua Zhang figure and caption. Added reference to NICAM and CAM-SP in idealized experiments section. Added reference to pilot study and retuning to sensitivity experiments section.
- Changes in July 25th version. Bibliography section updated, and Minghua Zhang's figure added. Other minor corrections.
- Changes in July 23rd version. First full draft of section 4. Section 3.1 updated, and point diagnostic location figure added. Item 5 in Section 3.3 added.
- Version July 13th, 2007 - First draft of most sections, but section 4 in note form only. List of figures but not all figures present. Bibliography incomplete.

1 Introduction.

For almost two decades, climate models have exhibited a large range of climate sensitivities (estimates differ by a factor of 2 to 3). Early studies suggested that inter-model differences in the treatment of climate feedbacks, and of cloud feedbacks in particular, were the primary reason for this range (Cess et al. 1989). It is in this context that in 2003, the Working Group on Coupled Modelling (WGCM) of the World Climate Research Program (WCRP) launched the Cloud Feedback Model Intercomparison Project (CFMIP) to encourage coordinated research in the area of cloud feedbacks in climate models. To promote both systematic comparisons of cloud feedbacks among climate models and comparisons between model clouds and observations, McAvaney and Le Treut (2003) proposed a set of coordinated atmospheric and slab GCM experiments using the ISCCP simulator (a tool that diagnoses model clouds in a way that mimics the satellite view from space, (Klein and Jakob (1999),Webb et al (2001)), that led to several analyses of cloud feedbacks. (See http://www.cfmip.net for more information.) In parallel, intercomparisons of climate feedbacks produced by coupled ocean-atmosphere models participating in the Fourth Assessment Report (AR4) of the IPCC took place within the framework of the phase 3 of the Coupled Model Intercomparison Project (CMIP3, Meehl et al (2005)), as well as within the framework of the low-Latitude cloud feedbacks Climate Process Team (CPT, Bretherton et al (2004)). The IPCC AR4 reaffirms the large spread in climate sensitivity estimates among current models (Randall et al, 2007). Thanks to the different projects mentioned above, however, substantial progress has been made recently in our assessment of the different sources of spread, and in the evaluation of modelled clouds with satellite observations (see Bony et al (2006) for a review). Yet, these different activities also give us a better appreciation of where problems and challenges remain in the area of cloud feedbacks. Issues of progress and key remaining problems are reviewed below. Then, on the basis of these lessons, we propose an outline of plans for the second phase of CFMIP.

1.1 Areas of progress.

Several studies (Forster and Taylor (2006), Ringer et al (2006), Soden and Held (2006), Webb et al (2006), Winton (2006)) have diagnosed global climate sensitivities and feedbacks in a large ensemble of GCMs participating in the AR4. Although the radiative forcing estimates calculated by models substantially differ (Collins et al, 2006), inter-model differences in cloud feedbacks still constitute the main source of spread of equilibrium climate sensitivity estimates, as well as transient climate response estimates (Dufresne and Bony, in preparation). New methodologies of feedback analysis suggest that the responses of all types of clouds contribute to this spread, with a dominant role of inter-model differences in the response of low-level clouds (Bony and Dufresne (2005), Webb et al (2006), Wyant et al (2006), Medeiros et al (submitted), Williams and Tselioudis (2007)) (see for example Fig. 2.) Model-to-satellite approaches using the ISCCP simulator (Klein and Jakob (1999),Webb et al (2001), Zhang et al (2005)), and new compositing methodologies of model-data comparison stratifying cloud properties in dynamical regimes (Williams et al (2003), Bony et al (2004), Williams et al (2006), Wyant et al (2006)) or cloud clusters (Jakob and Tselioudis (2003), Rossow et al (2005), Williams et al (2005), Williams and Tselioudis (2007)) have pointed out some systematic biases in the simulation by GCMs of mean cloud properties and the response of clouds to changing environmental conditions. For instance, most GCMs underestimate the occurrence of low-level clouds and overestimate that of optically thick clouds (Zhang et al, 2005), indicating that agreement between modelled and observed cloud radiative forcing at the top of the atmosphere is due to compensating errors in the simulated vertical structure of cloud fraction and cloud water. Also, it has been established that virtually all coupled models underestimate the sensitivity of the shortwave cloud radiative forcing to interannual SST changes in suppressed dynamical regimes predominantly covered by low-level clouds (Bony and Dufresne, 2005). Some of these approaches can be used as the basis for climate metrics which, if applied widely, have the potential to constrain cloud-climate feedbacks.

1.2 Remaining problems and challenges.

Recent intercomparison studies have thus identified dominant sources of uncertainty in model estimates of climate sensitivity and cloud feedbacks. These suggest some priorities for how observations might be used to evaluate some components of cloud feedbacks, and point out some physical processes that should be better understood and evaluated in climate models. However, the following problems and challenges remain:

1. The advent of the CMIP3/AR4 database means that models are now faced with increasing levels of scrutiny from a wide range of analysts and analytical techniques. In the case of clouds, modellers need tools which allow them to trace systematic errors back to the physical schemes responsible for them. Standard output from CMIP3 only supports quantitative evaluation of model clouds using two-dimensional satellite observations of radiative fluxes (eg ERBE) and total cloud fraction (e.g. ISCCP.) The ISCCP simulator can be used to quantify errors in the prevalence of difference cloud types, the heights of their tops and their optical properties, which remain as compensating errors in models' control simulations after tuning, undermining their credibility. Evaluation of multi-level information on clouds and precipitation (with more information on low level clouds previously obscured by those above) will be also possible in the near future with the arrival of new observations from space-borne radar and lidar (CloudSat and CALIPSO) in synergy with other instruments from the A-Train constellation of satellites (Stephens et al, 2002). Until simulators become widely used to quantify errors in models' clouds (both in the assessment and development process), compensating errors will continue to undermine the credibility of models.

2. Projects such as GEWEX/GCSS aim at evaluating and improving the representation of cloud processes in large-scale models. However, it is still unclear how modelling assumptions in the parametrization of cloud, boundary layer and convective processes affect their responses to climate change. Conversely, if we better understand which processes are critical for climate change cloud feedbacks (e.g. the transition between different cloud types, entrainment and detrainment processes, geometrical thickness, cloud phase changes, etc), parametrization efforts may be focused more directly on better representation of those particular processes. This suggests that bottom-up and top-down approaches will be more effective if applied together, through a closer collaboration between GEWEX/GCSS and CFMIP activities.

3. Low-level clouds are the primary contributors to inter-model differences in cloud feedbacks (with other cloud types making a secondary but still significant contribution.) Why do models predict such different low-cloud responses? Further analysis is required to address this question and determine which are the key processes that determine the sign and magnitude of cloud-climate feedbacks. For such process studies, it is necessary to revisit the list (and the frequency) of cloud diagnostics that are required in coordinated

model experiments to support a better understanding of the physical cloud-climate feedback mechanisms operating in them. It is also necessary to design cloud-climate feedback experiments (for global climate models and single column models) which are relatively inexpensive to run, to form the basis for exploratory sensitivity experiments and to test hypotheses relating to different feedback mechanisms in models.

4. GCMs exhibit a broad range of cloud feedbacks in climate change and they cannot all be right. Despite numerous model-data comparisons and the development of new methodologies of analysis and evaluation of cloud feedbacks, it is still unknown which of the model cloud feedbacks are the most credible. It is possible that cloud feedbacks in some models are unintended consequences of the modelling assumptions that go into them, for instance where parametrizations exploit empirical relationships between observed variables which break down in the warmer climate. It is also possible that coarse vertical resolution leads to an unsatisfactory representation of boundary layer cloud processes and their sensitivity to climate change. The modelling community that uses CRMs (Cloud Resolving Models) and LES (Large Eddy Simulation) models is familiar with the study of cloud processes in the context of present-day climate, but few CRM/LES studies focused on the analysis of cloud feedbacks in a climate change context have been carried out so far. For this reason, and also because CRM/LES models simulate clouds more explicitly than GCMs, coordinated (and idealized) climate sensitivity experiments performed with these models would allow us to see which climate models exhibit feedback mechanisms that can be reproduced in models with fewer assumptions.

1.3 Plans for CFMIP-2.

Drawing lessons from past intercomparison studies of GCM cloud feedbacks, recognizing the remaining problems and challenges, and looking to the Fifth Assessment Report (AR5) of the IPCC, discussions for a second phase of CFMIP have been underway since Spring 2006. New coordinators took over the project and started to think about the future needs of CFMIP. First, we expressed our wish to gather the key participants of the past CFMIP-1, CMIP3 and CPT projects to encourage the scrutiny of model simulations by a wider community. Second, we took the initiative to develop a combined CFMIP ISCCP/CloudSat/CALIPSO simulator (CICCS) for use in climate models. Third, we expressed our wish to collaborate closely with GEWEX/GCSS. These different initiatives were endorsed by WGCM following their Victoria meeting in September 2006 [1], and by the GEWEX SSG panel (supporting our plan for GCSS-CFMIP collaboration, February 2007 GEWEX Newsletter[2].) We organized an international CFMIP workshop at LMD/IPSL in Paris on April 2007 (held jointly with the European ENSEMBLES [3] project) to discuss these plans further and lay the foundations for a CFMIP-2 proposal. To strengthen the links between CFMIP, GCSS and the US CPT community, we have invited new members to join the CFMIP coordination committee. This is now made up of: Mark Webb, Met Office Hadley Centre (lead coordinator), Sandrine Bony, LMD/IPSL, George Tselioudis, NASA GISS and Chris Bretherton, U. Washington (joint coordinators) and Bryant McAvaney (project adviser.)

The main objectives and components of the CFMIP-2 strategy are illustrated in Fig. 1. The main objective of CFMIP-2 is to make, by the time of the AR5, an improved assessment of climate change cloud feedbacks by making progress in our understanding of cloud-climate feed-

[1] see meeting summary at http://www.clivar.org/organization/wgcm/wgcm.php

[2] see http://www.gewex.org/Feb2007.pdf

[3] see http://ensembles-eu.metoffice.com/

back processes and in our evaluation of cloud properties and behaviour using observations. To achieve these objectives, we plan :

- to develop, maintain and distribute to the modelling groups tools for evaluating modelled clouds using satellite products. This includes in particular a CFMIP ISCCP/CloudSat/ CALIPSO simulator (CICCS), clustering software, and user-friendly subsets of satellite data. An important condition for the successful development of cloud-climate metrics for AR5 is for CMIP4 to require the ISCCP simulator output as standard, and for simulations be run using CICCS where possible. (Modelling groups may put CICCS into their models, or run it "offline" locally and provide the output to the CMIP4 database).

- to run short atmosphere-only CFMIP-2 experiments in parallel to CMIP4, suitable for diagnosing CO_2 forcing and cloud feedback (and the interactions between them) in conventional climate models and in global cloud resolving/super-parametrized models. These will include CICCS as standard, and additional diagnostics to support a better understanding of cloud-climate feedback mechanisms (point diagnostics and cloud tendency terms.) They will form the basis for sensitivity experiments to establish the impact of various physical processes on cloud feedback, test hypotheses on various cloud feedback mechanisms operating in climate models, or assess the impact of vertical resolution on low cloud feedback and climate sensitivity.

- to develop close collaboration with GEWEX/GCSS to enrich and deepen the analysis of cloud feedback processes in GCMs, to compare the cloud feedback processes in GCMs and in CRMs/LES models within the context of idealized climate change experiments, and to explore the relationship between cloud modelling assumptions and cloud feedbacks in GCMs.

- to develop the analysis of GCM simulations by enhancing the community that scrutinizes the simulations and by increasing the number and the frequency of cloud diagnostics. For this purpose, CFMIP will analyse simulations from coupled models that will be performed within CMIP4, and we will make recommendations on cloud diagnostics to the CMIP panel.

These different actions are detailed below. Section 2 describes our our plans for the development of cloud simulators and their use in model development and evaluation (including the development of metrics.) Section 3 describes our plans for gaining a better understanding of cloud feedback mechanisms, through the use of additional diagnostics and lightweight climate model experiments. Section 4 describes our plans for assessing the credibility of cloud-climate feedback mechanisms in climate models through CFMIP-GCSS collaboration. Section 5 details our recommendations for cloud related standard outputs from CMIP4 and CFMIP-2 experiments.

2 Evaluation of clouds simulated by climate models

CFMIP-2 will continue to develop techniques and tools for evaluating modelled cloud-climate feedbacks using satellite products. Current efforts in this area are mainly focused on development of the CFMIP ISCCP/CloudSat /CALIPSO Simulator (CICCS), some minor enhancements to the ISCCP simulator, and the development and application of cloud-climate feedback evaluation techniques and metrics based on comparisons with ISCCP and CloudSat/CALIPSO data using these simulators.

2.1 ISCCP simulator development

The ISCCP simulator diagnostics were central to CFMIP-1 and featured in a number of studies (see introduction). CFMIP-1 data are now available via the PCMDI data portal, and more daily data is still coming in from various modelling groups (see http://www.cfmip.net for details). Five years of daily ISCCP simulator output were requested for the control and equilibrium CFMIP-1 slab experiments. These were used to evaluate the present-day simulation of cloudiness and to estimate the impact of systematic errors in the control simulations on inter-model spread in climate sensitivity by applying a cloud clustering technique to identify cloud regimes (Williams and Tselioudis, 2007). For example, characteristics of the tropical stratocumulus regime are found to vary considerably between GCMs in the simulation of the present-day climate (see Fig. 3). Several models simulate this regime with cloud which is too optically thick leading to incorrect radiative properties, and the frequency of the regime differs dramatically between the models analysed. These sorts of errors undermine the credibility of climate models, but the use of simulators as part of the model development process can help to target efforts on the physical schemes responsible. Williams and Tselioudis (2007) argue that eliminating these errors would potentially reduce the spread in climate sensitivity by one third in the six model versions analysed.

The simulator is currently being adapted to save three 'lightweight' diagnostics (grid-box mean ISCCP cloud occurrence, cloud top pressure and optical depth) to serve as an alternative to the 49 cloud types currently in use. This will allow clustering techniques to be applied for longer periods and to a wider range of experiments in the future. A new release of the simulator is planned for use in CMIP4 and CFMIP-2, which will include the above diagnostics, as well as a more accurate algorithm for diagnosing cloud top pressure, and an improved pseudo-random number generator. The work involved for modelling centres to upgrade to the new release of the ISCCP simulator will be much smaller than the work that was required to install it initially (this has already been done in the Hadley Centre model). (Keith Williams, Mark Webb (Met Office Hadley Centre), Steve Klein (LLNL))

2.2 CFMIP ISCCP/CloudSat/CALIPSO simulator (CICCS) development

The CloudSat and CALIPSO spaceborne cloud profiling radar and lidar instruments in the A-Train constellation of satellites are providing new information on clouds and their vertical structure (e.g. Winker et al (GEWEX news Nov 2006 [4] , Zhang et al (2007).) The main goals of the CloudSat and CALIPSO missions (which should lead to improved weather prediction and understanding of climatic processes) are to measure the vertical structure of clouds and precipitation in the atmosphere, to quantify cloud ice and water contents, to quantify the relationship between cloud profiles and radiative heating and to improve the understanding of aerosol indirect effects on clouds and precipitation.

Although active sensors see more of the 3D structure of clouds than can be seen by passive sensors, the effects of instrument sensitivity and attenuation by clouds and precipitation mean that not all clouds are detected. For this reason, CloudSat and CALIPSO simulators are being developed to support quantitative 'like with like' evaluation of model output with CloudSat and CALIPSO data (e.g. (Haynes et al, in press), Chepfer et al (submitted).) CICCS, which is being developed to support consistent comparisons with ISCCP, CloudSat and CALIPSO data across climate models, will be made up of a number of modules:

1. The first version of the CICCS infrastructure layer was released in November 2006. This

[4] see http://www.gewex.org/Nov2006.pdf p4

includes the ISCCP simulator subgrid cloud overlap module (SCOPS) and a prototype radar reflectivity code (Alejandro Bodas-Salcedo, Met Office Hadley Centre.)

2. The QuickBeam radar reflectivity code developed at CSU, PNNL/UW (John Haynes, Roj Marchand) has been adapted to plug in to CICCS (August 2007, Yuying Zhang, Steve Klein LLNL)

3. A lidar reflectivity code (ACTSIM) has been developed by IPSL/LMD (Chiriaco, Chepfer, Bony, Dufresne) and is currently being adapted to plug in to the CICCS framework (Alejandro Bodas-Salcedo).

4. A subgrid precipitation overlap module is being developed by Yuying Zhang, Steve Klein and Alejandro Bodas-Salcedo.

5. Post-processing modules to simulate the CloudSat cloud/hydrometeor mask and to produce statistical summaries such as Cloud Frequency Altitude Diagrams (CFADS) are in development at LLNL and LMD/IPSL (e.g. see Fig. 4). Joint CloudSat/CALIPSO simulator summaries are also envisaged and will be based on forthcoming joint CloudSat/CALIPSO products when they become available.

6. The ISCCP simulator will also be provided as a module within CICCS to avoid the overhead of plugging two simulators into the same model separately. (Mark Webb, Keith Williams)

Some preliminary comparisons of simulator output with CloudSat and CALIPSO are shown in Figs. 5 and 7.

Fig. 5 shows a transect through a mid-latitude depression in the North Atlantic on July 7th, 2006. The upper panel shows the results of the CloudSat simulator applied to the output of the UK Met Office global forecast model, which has a horizontal resolution of approximately 40 km at mid latitudes. It shows the radar reflectivities computed at sub-grid scale, with 20 subcolumns per gridbox, which gives an effective resolution similar to the observations. The contour lines in this panel are isotherms, the solid line denoting the freezing level. The impact of convective precipitation can be clearly observed as it is concentrated in a small number of subcolumns, whereas the boxes dominated by large-scale cloud are more homogeneous. The lower panel shows the radar reflectivity from CloudSat (dBZ). The observations show a more homogeneous structure, with less sub-grid scale variability than simulated by the GCM.

An example comparison of the three dimensional distribution of the cloud fraction derived from CALIPSO observations and from the LMD/IPSL GCM is shown in Fig. 7. Modelled atmospheric profiles have been converted to an ensemble of subgrid-scale attenuated backscatter lidar signals from which a cloud fraction has been derived using a procedure similar to that used to deduce a cloud fraction from CALIPSO lidar signals. The comparison points out the underestimate (overestimate) by the GCM of the predicted low-level cloudiness in tropical (extra-tropical, respectively) regions.

Pilot studies are planned to demonstrate the use of CICCS with a small number of GCMs. One part of this will be to build CICCS into the Met Office Hadley Centre, LMD/IPSL, GFDL and NCAR models to run 'in line', as is now the case for the ISCCP simulator in most GCMs. The code has been vectorized with this in mind. For CMIP4 and CFMIP-2, modelling groups will be encouraged to run the simulator in line, or to run off line locally. This is preferable to running the simulator centrally, as it allows its use as part of the model development process as well as for model inter-comparison. The feasibility of running the simulator locally offline using daily meaned or time sampled instantaneous model outputs will also be assessed.

LMD/IPSL will coordinate a pilot model inter-comparison study using the lidar component of CICCS. The Met Office Hadley Centre will coordinate a pilot CloudSat simulator model intercomparison pilot study. A possible CICCS pilot intercomparison using climate models in forecast mode is also under discussion, which could pave the way for the adoption of the approach is projects such as transpose AMIP/CAPT (Phillips et al, 2004). It is hoped that a complete prototype beta version of CICCS will be released by the end of 2007, with the first production version available mid-2008 for use in CMIP4/CFMIP-2. The ISCCP module is expected to be included in the production release but not the beta release.

2.3 Development of metrics relevant to cloud-climate feedback.

In recent years, metrics have been proposed for assessing the relative quality of different climate models compared to observational datasets. Many of these (e.g. Murphy et al (2004),Reichler and Kim (submitted), Sexton et al (in preparation)) are based on climatological variables, and do not take advantage of the compositing techniques described above. As part of CFMIP-2 we plan to develop cloud-climate metrics which use monthly and daily cloud diagnostics to more effectively target the cloud feedback processes understood to contribute most to current uncertainties climate model sensitivity. An example of this approach is illustrated in Fig. 6 and Table 1. Fig. 6 is taken from Williams et al (2006) and shows that present day spatio-temporal composites can capture some aspects of the cloud-climate feedback/response pattern in climate models. Table 1 shows that a metric based on this compositing method tends to favour higher sensitivity models (consistent with the findings of Bony and Dufresne (2005).) However, metrics based on 2D radiative fluxes will not penalise models that get plausible values of cloud radiative forcing with compensating errors in cloud fraction, cloud optical depth, etc. Newer approaches based on simulator output show the potential to provide stronger constraints on cloud-climate feedbacks in models. Williams and Tselioudis (2007) argue that metrics based on the clustering approach (see Fig. 3) have the potential to reduce the inter-model spread in climate sensitivity by a third. However, a necessary pre-requisite for this is the inclusion of the ISCCP simulator as standard in a wide range of models. This not only supports the calculation of metrics, but gives a wealth of information that can be used in the model development process to remove compensating errors and improve the credibility of the models' cloud simulations in the longer term.

The development of cloud-climate feedback metrics is one part of a wider activity to develop metrics to assess all aspects of climate models performance. This topic was discussed at the Paris meeting in a presentation by Robert Pincus [5] and a plenary discussion chaired by Karl Taylor and Robert Pincus [6]. A session on metrics is planned for the PAN-GCSS meeting in Toulouse, June 2008.

3 Understanding cloud feedback mechanisms

CFMIP-1 was mainly concerned with the evaluation of model clouds using satellite data, and the identification of those cloud regimes contributing most to the inter-model spread in cloud feedback and climate sensitivity. CFMIP-2 aims to build upon this work, but also aims for a better understanding of the different cloud feedback mechanisms acting in models. This is to

[5] see http://cfmip.metoffice.com/PincusParis.pdf
[6] see http://cfmip.metoffice.com/MetricsParis.pdf

be achieved through a combination of diagnostic techniques and 'lightweight' climate change experiments.

3.1 Instantaneous diagnostics at selected locations.

Due to the high frequency variability of clouds, time averaged model output gives a fairly limited picture of the physical mechanisms underlying cloud simulations. High frequency, instantaneous diagnostics are likely to give more insight into the physical processes operating and the interactions between them (e.g. convective intermittency and convective/boundary layer interactions). They also support the diagnosis of any unphysical behaviour related to numerical noise, vertical discretisation effects, etc. An example of this approach is the high frequency data saved at selected points by modelling groups involved in the US Climate Process Team (CPT) on low latitude cloud feedbacks. Two of these models showed very different cloud simulations in stratocumulus regions in spite of similar values of net cloud forcing (see Fig. 8 .)

The WGNE-GCSS Pacific Cross Section Intercomparison (GPCI) led by J. Teixeira (GEWEX News, Nov 2006 [7]) has saved high frequency cloud diagnostics from twenty NWP and climate models along a section sampling the stratocumulus regime off the coast of California, the shallow cumulus to the south west and the deep convection in the ITCZ (as well as the transitions between them). Differences in the mean state, variability and diurnal cycle of cloudiness in these models in the present climate are being examined (see the project website [8] or Joao Teixeira's presentation from the Paris meeting [9].) For example, Fig. 9 shows a comparison of the diurnal cycle in cloud cover along the GPCI for three models, which can be evaluated using ISCCP data (not shown.) To better understand cloud-climate feedback mechanisms in climate models, we propose that high frequency instantaneous cloud data should be saved at representative locations around the globe in present day and climate change experiments for CMIP4 and CFMIP-2. These diagnostics will give a clearer picture of the relationships between cloud variables (e.g. cloud water, cloud fraction) and environmental variables (humidity, stability, vertical velocity, etc), of convective/boundary later interactions, and of the diurnal cycle of cloudiness in models and its impact on cloud feedback.

Work is currently underway to identify a representative set of locations (see Fig. 10). Likely candidates so far include the GPCI locations, the selection of points saved by the CPT (including ARM sites, ocean weather ships and buoys), selected locations associated with relevant observational field campaigns (e.g. VOCALS, Wood et al (2007)) and a set of locations chosen to sample the range of inter-model spread in low-cloud feedbacks in climate models. To make GCM data accessible to the 'process' community, the GEWEX SSG recommended that this data should be mirrored at the GCSS-DIME site along side observational data in a consistent form. (See also discussion in section 4.) (Joao Teixeira, George Tselioudis, Mark Webb, Sandrine Bony, Rob Wood)

3.2 Analysis of cloud tendency terms

We plan to use cloud condensate tendency diagnostics (CCTD) to gain insight into the physical mechanisms responsible for cloud feedbacks in the CFMIP-2 experiments. This technique has been used to understand the mid-latitude mixed-phase feedbacks in different versions of the MIROC 3.2 model (Ogura at 2007). A pilot model inter-comparison study has also been

[7] see http://www.gewex.org/Nov2006.pdf p17
[8] see http://www.igidl.ul.pt/cgul/projects/gpci.htm for details
[9] see http://cfmip.metoffice.com/TeixeiraParis.pdf

performed with the MIROC 3.2 model and the HadGEM1 version of the Met Office Hadley Centre Model. Fig. 11 shows temporal correlations between the cloud condensate response to increasing CO_2 and the CCTD terms for HadGEM1 and MIROC3.2. Positive values indicate a positive correlation with increasing cloud condensate, while negative values indicate a positive correlation with decreasing condensate. The condensation-evaporation and deposition-sublimation processes dominate the cloud response in HadGEM1 (Figs. 11(a) and 11(c)). In MIROC3.2, most of the cloud response pattern is related to the condensation-evaporation term (Fig. 11(d)). However, the ice sedimentation term also shows a considerable impact on the mixed phase cloud increase (Fig. 11(e)). (Ogura et al, submitted [10])

A pilot study with ECHAM5 and HadGEM2+PC2 (a new cloud scheme with prognostic condensate and cloud fraction) has been performed to use this technique to examine sub-tropical low cloud feedback mechanisms in SST forced experiments. Analysis of HadGEM2/PC2 along the GCSS Pacific Cross Section shows evidence for positive feedback in large-scale sub-tropical low clouds, with reduced condensation from cloud top cooling dominating nearer the coast, but reduced detrainment from shallow convection dominating further west [11]. It is planned to extend this study to include results from the MIROC model. (Mark Webb, Johannes Quaas, Tomoo Ogura)

3.3 A lightweight experimental framework for studying cloud-climate feedbacks

At the Paris meeting, a working group led by Mark Ringer (Met Office Hadley Centre) and Brian Soden (U. Miami) discussed and made recommendations for the experimental design for CFMIP-2 experiments. The full recommendations are available on the CFMIP website. A brief summary is given here.

To complement the CMIP4 experiments, a set of parallel CFMIP-2 experiments are proposed. These are shorter, SST forced atmosphere only experiments with more diagnostics relevant to cloud feedback processes. This 'lightweight' experimental design will also support the inter-comparison of conventional climate models with global cloud resolving models such as the NICAM [12] and CAM-SP [13], which are presently too expensive to run as AOGCMs.

The following set of CFMIP-2 'reference' experiments is proposed:

1. An AMIP-like experiment running from the start of the ERBE data period (1985 - date) including the CloudSat/CALIPSO era. This experiment will include all the CFMIP-2 diagnostics (see section 5.2) and will form the basis for evaluation against satellite data using the simulators, and studies relating cloud climate feedbacks to interannual variability.

2. A control experiment forced with a climatological observed SSTs and sea ice (as already provided by PCMDI) to form the basis for SST forced climate change experiments and CO_2 forcing diagnosis experiments.

3. A patterned SST forced climate change experiment. This will be based on 2) above, with an SST perturbation pattern based on a composite of coupled model SST responses taken from 1% coupled model experiments at time of CO_2 doubling, as developed by the CPT

[10] see also http://cfmip.metoffice.com/OguraParis.pdf
[11] see http://cfmip.metoffice.com/WebbSensitivityParis.pdf
[12] see http://cfmip.metoffice.com/TsushimaParis.pdf
[13] see http://cfmip.metoffice.com/KhairoutdinovParis.pdf

(Wyant et al, 2006). Although these experiments are not expected to reproduce exactly the global mean cloud feedbacks as in a coupled experiment or slab experiments, they are expected to explore the same range of cloud feedback processes (Wyant et al (2006), Ringer et al (2006)).

4. An experiment where CO_2 is doubled while SSTs remain fixed (Hansen method based on 2)) will be used to estimate the CO_2 forcing including the effects of any rapid cloud responses to changes in atmospheric structure in response to CO_2 doubling. Gregory and Webb (2007) have shown that a significant fraction of the inter-model spread in cloud 'feedback' in slab models occurs shortly after CO_2 doubling, and is not in fact related to the global mean surface temperature response. The combination of SST and CO_2 forced experiments will allow these two aspects of cloud 'feedback' to be separately quantified.

5. Some other experiments were also proposed at the Paris workshop. The first was an experiment where CO_2 levels in 3) are 'tuned' to bring the system into radiative balance. This would allow any nonlinear interactions between 3) and 4) to be examined. The second was the aquaplanet climate change experiment described in Medieros et al 2007 (submitted). Our recommendation is that these should remain outside the scope of CFMIP, but that a common experimental design should be agreed so that interested modelling groups can compare results if they wish.

6. Slab model experiments were not recommended for CFMIP-2, because of the technical difficulties that modelling centres have had getting a consistent representation of sea ice in slab and coupled model versions. The majority view at the Paris meeting was that if CMIP4 runs 1% experiments with suitable cloud output, then these should be sufficient to track the evolution of the effective climate sensitivity between the AR4 and AR5 model versions and to assess locally coupled cloud/SST interactions. (Note that the effective climate sensitivity (Murphy, 1995) estimates the eventual equilibrium response of a coupled model and is largely independent of ocean heat uptake rate.) However, these recommendations for the CFMIP-2 experimental design would need to be revisited if 1% and slab experiments were not to be run in CMIP4.

3.4 Sensitivity experiments

The lightweight nature of these experiments makes the prospect of running sensitivity experiments more attractive. Various types of sensitivity experiments are possible ; the following are currently under consideration.

First, modelling groups may wish to assess the impact that perturbing key model processes have on the cloud feedbacks in their models. An example of this approach was demonstrated in Pier Siebesma's presentation at the Paris meeting [14], and is the basis for probabilistic climate predictions such as Murphy et al (2004)

Second, sensitivity experiments may be used to test hypotheses that are made about the physical mechanisms responsible for cloud feedbacks in models. For example, reductions in shallow convective detrainment and longwave cloud top cooling were proposed as potential causes of subtropical cloud reductions at the Paris meeting (see section 3.2). These hypotheses could be tested in sensitivity experiments where the proposed feedback loops are cut.

Third, modelling groups could assess the impact of increasing boundary layer resolution on the cloud feedbacks in their models. SCM/CRM comparisons from DYCOMS II show SCMs

[14] see http://cfmip.metoffice.com/SiebesmaParis.pdf

to compare more favourably with CRM results when vertical resolution is increased (see Wyant et al (in press), and Adrian Lock's Paris presentation.[15]) If the fundamental character of the cloud feedback mechanism operating in a model changes when vertical resolution is increased, then this raises questions about its validity at the lower resolution.

For many of these sensitivity experiments, modelling groups may feel it necessary to retune their models to maintain a neutral energy balance in the present day simulation. This may be done using the usual procedure, or alternatively by applying a global scaling factor to the cloud fraction seen by the radiation scheme. The latter approach may well have a less unpredictable effect on the cloud feedback and would also be less work than the former.

Given the potentially large number of sensitivity experiments that are possible, it would not be feasible to arrange a coordinated inter-comparison/data exchange exercise for all of them. The vertical resolution sensitivity experiments would probably be the strongest candidate for a coordinated inter-comparison, as the first two options would be approached differently from model to model. These options are currently being explored as part of a pilot study involving the Met Office Hadley Centre, MPI and MIROC models. (Mark Webb, Johannes Quaas and Tomoo Ogura.)

4 Assessing the credibility of model feedbacks - CFMIP-GCSS collaboration

In recent years the aims of CFMIP and the GEWEX Cloud System Study (GCSS) have shown increasing overlap. As CFMIP-2 focuses more on understanding and assessing the credibility of cloud feedback mechanisms in climate models, there are benefits to comparing with models that resolve cloud feedback processes more explicitly. CRMs and LES models can give more consistently plausible simulations than Single Column Model (SCM) versions of NWP and climate models given the same forcings (e.g. Brown et al (2002) cf Lenderink et al (2004).) These have been routinely used by GCSS Working Groups to simulate cloud systems in present day case study mode, for comparison with SCMs and field observations. However, despite the explicit mandate of GCSS to improve cloud representation in climate models, GCSS case studies do not routinely examine the large-scale context and the climate model relevance of the simulated cloud systems. This context can be provided by CFMIP, which not only documents climate model cloud-type deficiencies but also assesses their impact on model climate sensitivity and its inter-model spread. This synergy led to the decision to organize a working group at the Paris meeting to explore practical routes for a CFMIP/GCSS collaboration (led by Pier Siebesma and George Tselioudis.)

The outcome of the session was a set of recommendations to CFMIP and GCSS on future collaboration that both extends the scope of existing CFMIP and GCSS initiatives to better address common issues, and proposes new initiatives to integrate research components of the two projects (see the working group summaries at www.cfmip.net.) Drawing on these recommendations (and subsequent discussions), we propose the following plan for future CFMIP-GCSS collaboration.

[15] see http://cfmip.metoffice.com/LockParis.pdf

4.1 Derivation and storing of high frequency diagnostics from selected locations in the CFMIP-2 and CMIP4 experiments.

The plan is to select a set of points around the globe which sample the range of cloud-climate model feedbacks important for climate sensitivity (e.g. the GPCI, see section 3.1). These will give a clearer picture of the cloud feedback mechanisms operating in climate models, and will make it possible to draw on the expertise of the GCSS community when assessing their credibility. They will also inform the future design of cloud-climate feedback CRM/LES/SCM studies by the GCSS community (see below). One additional consideration is the selection of grid points that fall within the domain of previous or upcoming field experiments and that have formed the basis of previous or ongoing GCSS case studies. This will make it possible to make use of both detailed field study observations and already existing GCSS model experiments. Discussions for the selection of the optimal set of grid points to be extracted from climate models are ongoing. The resulting model dataset will be stored at the GCSS-DIME web site along with observational data for the same set of grid points. Those datasets will form the basis for future individual and group studies of cloud feedback processes. (Joao Teixeira, George Tseliousdis, Mark Webb, Chris Bretherton, Rob Wood, Sandrine Bony.)

4.2 CFMIP-GCSS case study.

A common CFMIP-GCSS low cloud feedback case study, suitable for running CRM/LES models and SCMs is to be organized by the GCSS Boundary Layer working group led by Adrian Lock. The aim is to use CRM/LES runs to explore the physical processes involved in the range of climate model cloud responses, and to compare these with SCM versions of climate models. A number of approaches are proposed, and these are summarized below:

- A Lagrangian model experiment that includes CRM, LES, and SCM simulations along a section that encompasses the main extremes of model cloud feedback values would serve to investigate the wide range of cloud feedback processes. The GPCI is a natural candidate for such an experiment, since the cross section was defined to include the major tropical cloud types and their geographic transitions. However, alternative cross-sections are also being considered, in order to better capture the range of low cloud feedback responses in the climate models. (Pier Siebesma, Joao Teixeira, Adrian Lock, Chris Bretherton.)

- A methodology for developing idealized cloud feedback case studies for the sub-tropics has been developed as part of the US CPT (Bretherton et al 2006). For example, an SCM/LES trade cumulus transition case with an idealized SST/subsidence climate change forcing (Zhang and Bretherton, 2007) was presented by Minghua Zhang at the Paris meeting [16] (see Fig. 12). This case is currently being run by groups at Stony Brook (Minghua Zhang), GFDL (Ming Zhao) and the Met Office (Adrian Lock). The approach has also been applied to a stratocumulus case by Caldwell and Bretherton (2007), and a trade cumulus case is under development at the University of Washington (Matt Wyant, Peter Blossey and Chris Bretherton.)

- Cloud Resolving Model (CRM) studies in an Eulerian setting, for a number of locations on the GPCI-cross section, fed by ECMWF model fields and advection tendencies can also form the basis of a climate feedback case study (Kuan-Man Xu.)

[16] see http://cfmip.metoffice.com/ZhangParis.pdf

- Cloud feedback model studies using a simplified aquaplanet framework (see Medeiros et al, 2007 and his Paris presentation [17]) can also be used to derive forcings for idealized cloud feedback case studies.

Several issues related to the set-up of the proposed feedback study are still under consideration. These include whether climatological, idealized, or instantaneous forcing data are preferable, and how to generate climate change forcings for the experiments which are representative of those in the climate models. An important concern that needs to be addressed is the extent to which the results of a somewhat localized feedback study can be generalized to address global cloud feedback issues. Observational cloud clustering techniques such as those used in Williams and Tselioudis (2007) can be used to isolate the major modes of variability of the tropical and subtropical cloud fields and to test the representativeness of the regions selected for the cloud feedback studies. (George Tselioudis)

These various approaches will be developed in the coming months and presented at the next Pan-GCSS meeting (Toulouse, June 2008), after which a concrete design for a CFMIP-GCSS case study should be selected.

5 Recommendations for standard output from CMIP4 and CFMIP-2.

5.1 CMIP4 and CFMIP-2 Experiments

CFMIP-1 requested atmosphere–mixed-layer ocean (slab) experiments with a comprehensive diagnostic list aimed at understanding and evaluating cloud feedback. In parallel, the same experiment was requested by CMIP3 (IPCC AR4 data collection) with a slightly different diagnostic list. For for CFMIP-2, we propose that the CFMIP diagnostic requirements are included in the standard CMIP4 (IPCC AR5) request so that modelling centres are only required to run the additional CFMIP-2 experiments which don't form part of the CMIP4 experimental design. In planning CFMIP-2, it has been assumed that AMIP, coupled climate of the 20th century, pre-industrial control and 1%/year increasing CO_2 experiments using base physical models (i.e. without carbon cycle feedbacks) will again form a principal part of the CMIP4 request. We will request modelling groups to submit data from the proposed CFMIP-2 experiments (see sections 3.3 and 3.4) at the same time (or in advance of) the deadline for submission of CMIP4 output. Given that CFMIP-2 experiments will be much shorter than CMIP4 experiments (25 or so years in length) this will mean that modelling groups can start them later, giving more time to set up any extra diagnostics required.

5.2 Output recommendations for CMIP4 and CFMIP-2

The CFMIP-2 diagnostic recommendations for CMIP4 and CFMIP-2 experiments are summarized in Table 2. The requested outputs are organized into four diagnostics packages combined with a number of temporal and spatial sampling options to support a range of analytical techniques in the published literature. Each package builds upon the previous one so that studies using Package 3 tend to also require Packages 1 and 2. The packages are:

[17] see http://cfmip.metoffice.com/MedeirosParis.pdf

- Package 1: CFMIP-1 2D fields. This is the same as the CMIP3 table A1a except with the addition of 3 x 2D ISCCP simulator fields (see section 2.1), 500hPa vertical velocity and temperature and humidity at 700hPa (to calculate lower tropospheric stability).

- Package 2: CFMIP-1 3D fields. This is the same as the CMIP3 table A1c except with the addition of the full 4D ISCCP simulator diagnostics as standard, 3D cloud amount, liquid water and ice (total or stratiform and convective) and convective mass flux (separated into shallow and deep if possible). Most of the non-cloud fields in CMIP3 table A1c (in which the vertical profile is fairly smooth in the monthly mean) were requested on pressure levels. However, to see the relationships between cloud and environmental variables (temperature, humidity, etc) it is necessary to save all 3D fields at daily and higher temporal resolution on model levels. Saving these on standard pressure levels only would lose information on humidity/temperature inversions which can have a significant impact on the simulation of low level cloudiness (e.g. Lock et al (2000).) We therefore request that 3D fields being collected by CMIP4 at daily and higher temporal resolution are saved on model levels. A post processing tool to convert the daily fields to pressure levels for some users may also be necessary.

- Package 3: CloudSat/CALIPSO simulator diagnostics. The exact nature of these is still being discussed, however they are likely to include statistical summaries similar to the 4D ISCCP simulator diagnostics (see for example Fig. 4).

- Package 4: Cloud tendency diagnostics. These give insight into the processes responsible for the different cloud responses in GCMs. However, as they will vary from model to model the proposal is for these diagnostics not to form part of the CMIP4 request and will instead be confined to the additional CFMIP-2 experiments

The number of additional variables being requested over the existing CMIP monthly tables is quite small. However, CFMIP-2 also requests these diagnostics as daily means and 3 hourly instantaneous data for a selected number of grid points from the CMIP4 experiments (see section 3.1). It is recognized that this has potential to increase requested data volumes, so we limit these to particular time periods and forcing scenarios. Table 2 shows that CFMIP-2 is not requesting any daily or 3 hourly data from the CMIP4 scenario or Earth System Model experiments (only from present day and 1% CO_2 experiments) and that these data are requested for limited periods of no more than 25 years in length. Saving 3 hourly data at sixty or so point locations for a single model variable produces about half as many data points as required for an equivalent monthly mean gridded global field at 1.5° resolution. However, these variables would be saved for at most 20 years (at equilibrium or time of CO_2 doubling) so they would cause at most a 10% or so increase in the data volume produced from an O(100) CMIP4 year integration. For the CFMIP-2 experiments, these data may be saved at timestep frequency (eg half hourly.) Many of the modelling groups taking part in CMIP4 have taken part in the WGNE-GCSS GPCI project, and so already have experience of submitting point diagnostics to model inter-comparison projects.

The standard output requested by CMIP3 was split into 'higher priority' and 'lower priority' categories. The ISCCP simulator output was in the 'lower priority' category. In practice, many modelling centres chose not to submit 'lower priority' output. Only one modelling group submitted ISCCP simulator output to CMIP3, even though many of the same groups submitted simulator output to CFMIP (for which it was mandatory) during the same period. As modelling groups are in practice free to submit on not submit whatever output they wish, we request that any such separation is avoided in the CMIP4 request if at all possible.

Table 2 was developed at the joint CFMIP/ENSEMBLES meeting (Paris, 2007) during a break-out group lead by Keith Williams (Met Office) and Karl Taylor (PCMDI). Representatives from many modelling centres were present, together with many of those who are likely to use the data. The group was satisfied that the additional diagnostics and sampling proposed would be valuable if adopted as part of the CMIP4 request. We will continue to work with modelling centres to illustrate the benefits of the new diagnostics and encourage their inclusion and we are requesting WGCM's support for adopting the diagnostic requirements outlined above into the standard CMIP4 request.

5.3 Data management issues

Cloud diagnostics in CMIP4 are expected to be hosted as part of the data serving strategy in place for AR5, presumably coordinated by PCMDI. CFMIP-2 will deliver the greatest benefit to the scientific community if its experiments can be made available via the same route (as has now been arranged for CFMIP-1 data.) Hence we request that output data from the CFMIP-2 experiments should be hosted as part of same solution developed for CMIP4/AR5.

6 Acknowledgements

We would like to acknowledge the following people for their help in the CFMIP-2 planning process: Alejandro Bodas-Salcedo, Rob Colman, Jean-Louis Dufresne, Christian Jakob, Steve Klein, Adrian Lock, Bryant McAvaney, Tomoo Ogura, Robert Pincus, Johannes Quaas, Mark Ringer, Cath Senior, Steve Sherwood, Pier Siebesma, Brian Soden, Bjorn Stevens, Karl Taylor, Joao Teixeira, Keith Williams, Rob Wood, Minghua Zhang and Yuying Zhang.

References

Bony S, Dufresne JL (2005) Marine boundary layer clouds at the heart of cloud feedback uncertainties in climate models. Geophys Res Lett 32(20):Vol. 32, No. 20, L20806.

Bony S, Dufresne JL, Le Treut H, Morcrette JJ, Senior CA (2004) On dynamic and thermodynamic components of cloud changes. Clim Dyn 22:71–86 doi:10.1007/s00382-003-0369-6.

Bony S, Colman R, Kattsov VM, Allan RP, Bretherton CS, Dufresne JL, Hall A, Hallegate S, Holland MM, Ingram WJ, Randall DA, Soden BJ, Tselioudis G, Webb MJ (2006) How well do we understand and evaluate climate change feedback processes? J Climate 19(15):3445–3482.

Bretherton CS, Ferrari R, Legg S (2004) Climate process teams: A new approach to improving climate models. US Clivar Variations 2:1–6.

Bretherton CS (2006) The climate process team on low-latitude cloud feedbacks on climate sensitivity. US Clivar Variations 4:7–12.

Brown AR, Cederwall RT, Chlond A, Duynkerke PG, Golaz JC, Khairoutdinov M, Lewellen DC, Lock AP, MacVean MK, Moeng CH, Neggers RAJ, Siebesma AP, Stevens B (2002) Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. Q J R Meteorol Soc 128:1075–1093.

Chepfer H, Bony S, Chiriaco M, Dufresne JL, Seze G, Winker D (submitted) Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. Geophys Res Lett .

Collins WD, Ramaswamy V, Schwarzkopf MD, Sun Y, Portmann RW, Fu Q, Casanova SEB, Dufresne JL, Fillmore DW, Forster PMD, Galin VY, Gohar LK, Ingram WJ, Kratz DP, Lefebvre MP, Li J, Marquet P, Oinas V, Tsushima Y, Uchiyama T, Zhong WY (2006) Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). J Geophys Res 111:D14317 doi:10.1029/2005JD006713.

Dufresne JL, Bony S (in preparation) Relative contribution of the different radiative feedbacks to global warming in equilibrium and transient climate change simulations. J Climate.

Forster PMDF, Taylor KE (2006) Climate forcings and climate sensitivities diagnosed from coupled climate model integrations. J Climate 19(23):6181–6194.

Haynes JM, Marchand RT, Luo Z, Bodas-Salcedo A, Stephens GL (in press) A multi-purpose radar simulation package: Quickbeam. Bull Am Meteorol Soc (preprint available at http://reef.atmos.colostate.edu/haynes/radarsim/bams_simulator.pdf).

Jakob C, Tselioudis G (2003) Objective identification of cloud regimes in the Tropical Western Pacific. Geophys Res Lett 30(21) doi:10.1029/2003GL018367.

Klein SA, Jakob C (1999) Validation and sensitivities of frontal clouds simulated by the ECMWF model. Mon Weather Rev 127(10):2514–2531.

Lenderink G, Siebesma AP, Cheinet S, Irons S, Jones CG, Marquet P, Muller F, Olmeda D, Calvo J, Sanchez E, Soares PMM (2004) The diurnal cycle of shallow cumulus clouds over land: A single-column model intercomparison study. Q J R Meteorol Soc 130(604):3339–3364.

Lock AP, Brown AR, Bush MR, Martin GM, Smith RNB (2000) A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. Mon Weather Rev 128(9):3187–3199.

McAvaney BJ, Le Treut H (2003) The cloud feedback intercomparison project: (CFMIP). In CLIVAR exchanges - supplementary contributions.

Medeiros B, Stevens B, Held IM, Zhao M, Williamson DL, Olson JG, Bretherton CS (submitted) Aquaplanets, climate sensitivity, and low clouds. J Climate.

Meehl GA, Covey C, McAvaney B, Latif M, Stouffer RJ (2005) Overview of the coupled model intercomparison project. Bull Am Meteorol Soc 86(1).

Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430:768–772.

Murphy JM (1995) Transient response of the Hadley Centre coupled ocean-atmosphere model to increasing carbon dioxide: Part III. Analysis of global-mean response using simple models. J Climate 8:496–514.

Phillips TJ, Potter GL, Williamson DL, Cederwall RT, Boyle JS, Fiorino M, Hnilo JJ, Olson JG, Xie S, Yio JJ (2004) Evaluating parameterizations in General Circulation Models: Climate simulation meets weather prediction. Bull Am Meteorol Soc 85:1903–1915.

Randall DA, Wood RA, Bony S, Colman R, Fichefet T, Fyfe J, Kattsov V, Pitman A, Shukla J, Srinivasan J, Stouffer RJ, Sumi A, Taylor KE (2007) Climate models and their evaluation. In Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. C. Marquis, K. B. Averyt, M. Tignor and H. L. Miller (eds.)].

Reichler T, Kim J (submitted) How well do coupled models simulate today's climate? Bull Am Meteorol Soc .

Ringer MA, McAvaney B, Andronova N, Buja L, Esch M, Ingram W, Li B, Quaas J, Roeckner E, Senior C, Soden B, Volodin E, Webb M, Williams K (2006) Global mean cloud feedbacks in idealized climate change experiments. Geophys Res Lett 33:L07718 doi:10.1029/2005GL025370.

Rossow WB, Tselioudis G, Polak A, Jakob C (2005) Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. Geophys Res Lett 32(21) doi:10.1029/2005GL024584.

Sexton DMH, Rougier J, Murphy J, Webb M (in preparation) Part II: Using a large number of observations to constrain a climate prediction with an imperfect model. J Climate.

Soden BJ, Held IM (2006) An assessment of climate feedbacks in coupled ocean-atmosphere models. J Climate 19:33543360.

Stephens GL, Vane DG, Boain RJ, Mace GG, Sassen K, Wang Z, Illingworth AJ, O'Connor EJ, Rossow WB, Durden SL, Miller SD, Austin RT, Benedetti A, Mitrescu C, The CloudSat Science Team (2002) The CloudSat mission and the A-Train. Bull Am Meteorol Soc 83:1771–1790.

Taylor KE, Crucifix M, Braconnot P, Hewitt CD, Doutriaux C, Broccoli AJ, Mitchell JFB, Webb MJ (2007) Estimating shortwave radiative forcing and response in climate models. J Climate 20:2530–2543.

Webb M, Senior C, Bony S, Morcrette JJ (2001) Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. Clim Dyn 17:905–922.

Webb MJ, Senior CA, Sexton DMH, Ingram WJ, Williams KD, Ringer MA, McAvaney BJ, Colman R, Soden BJ, Gudgel R, Knutson T, Emori S, Ogura T, Tsushima Y, Andronova NG, Li B, Musat I, Bony S, Taylor KE (2006) On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. Clim Dyn 27(1):17–38 doi:10.1007/s00382-006-0111-2.

Williams KD, Tselioudis G (2007) GCM intercomparison of global cloud regimes: Present-day evaluation and climate change response. Clim Dyn doi:10.1007/s00382-007-0232-2.

Williams KD, Ringer MA, Senior CA (2003) Evaluating the cloud response to climate change and current climate variability. Clim Dyn 20:705–721 doi:10.1007/s00382-002-0303-3.

Williams KD, Senior CA, Slingo A, Mitchell JFB (2005) Towards evaluating cloud response to climate change using clustering technique identification of cloud regimes. Clim Dyn 24:701–719.

Williams KD, Ringer MA, Senior CA, Webb MJ, McAvaney BJ, Andronova N, Bony S, Dufresne JL, Emori S, Gudgel R, Knutson T, Li B, Lo K, Musat I, Wegner J, Slingo A, Mitchell JFB (2006) Evaluation of a component of the cloud response to climate change in an intercomparison of climate models. Clim Dyn 26:145–165 doi:10.1007/s00382-005-0067-7.

Winton M (2006) Surface albedo feedback estimates for the AR4 climate models. J Climate 19:359–365.

Wood R, Mechoso CR, Bretherton C, Huebert B, Weller R (2007) The VAMOS ocean-cloud-atmosphere-land study (VOCALS). US Clivar Variations 5:1–5.

Wyant MC, Bretherton CS, Bacmeister JT, Kiehl JT, Held IM, Zhao M, Klein SA, Soden BJ (2006) A comparison of low-latitude cloud properties and their response to climate change in three agcms sorted into regimes using mid-tropospheric vertical velocity. Clim Dyn 27(2-3):261 – 279 doi:10.1007/s00382-006-0138-4.

Wyant MC, Bretherton C, Chlond A, Griffin BM, Kitagawa H, Lappen CL, Larson VE, Lock AP, Park S, de Roode SR, Uchida J, Zhao M, Ackerman AS (in press) A single-column-model intercomparison of a heavily drizzling stratocumulus topped boundary layer. J Geophys Res .

Zhang M, Bretherton CS (2007) Mechanisms of climate feedback from low clouds: analysis of idealized simulations using the cam3 single-column model. In preparation .

Zhang MH, Lin WY, Klein SA, Bacmeister JT, Bony S, Cederwall RT, Del Genio AD, Hack JJ, Loeb NG, Lohmann U, Minnis P, Musat I, Pincus R, Stier P, Suarez MJ, Webb MJ, Wu JB (2005) Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. J Geophys Res 110(D15).

Zhang Y, Klein S, Mace GG, Boyle J (2007) Cluster analysis of tropical clouds using cloudsat data. Geophys Res Lett 34:L12813 doi:10.1029/2007GL029336.

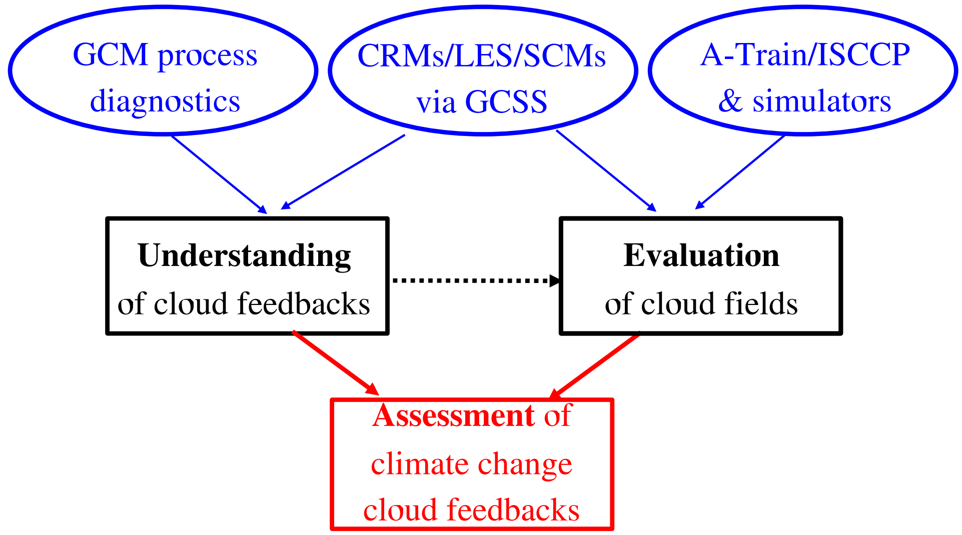


Figure 1: Schematic of CFMIP-2 project and initiatives.

Figure 2: Sensitivity of shortwave cloud radiative forcing changes in response to long term SST changes predicted in 1% CO_2 scenarios from 15 CMIP3/AR4 AOGCMs, separated into dynamical regimes. Dotted lines show the maximum and minimum values. The red squares and lines show the mean and standard deviation of the 8 higher sensitivity versions. The 7 lower sensitivity versions are shown in blue. From Bony and Dufresne (2005).

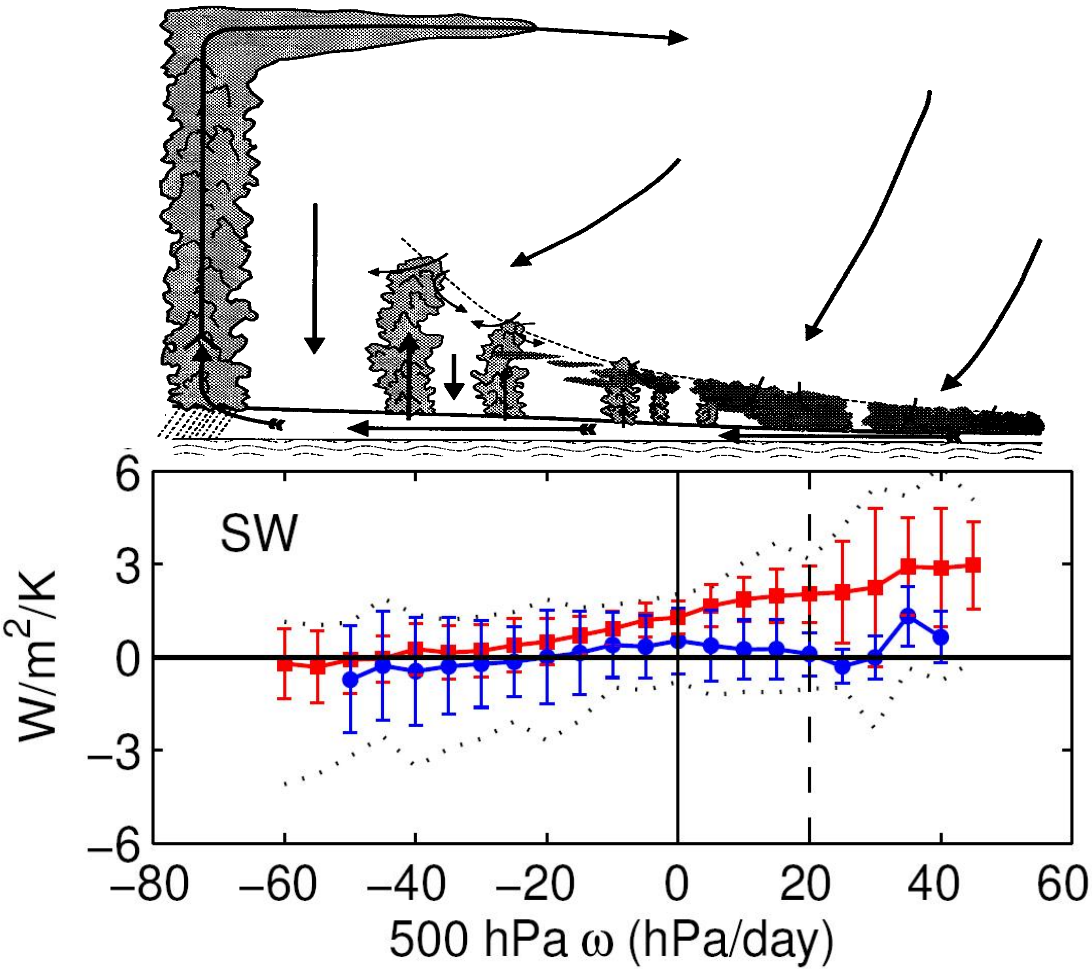


Figure 3: Top: Mean cloud-top pressure – optical depth histograms of cloud amount for the tropical stratocumulus cluster regime from ISCCP observations and from five GCMs. Bottom: The relative frequency of occurrence (RFO) of the tropical stratocumulus regime (as a fraction of the total number of days.) From Williams and Tselioudis (2007).

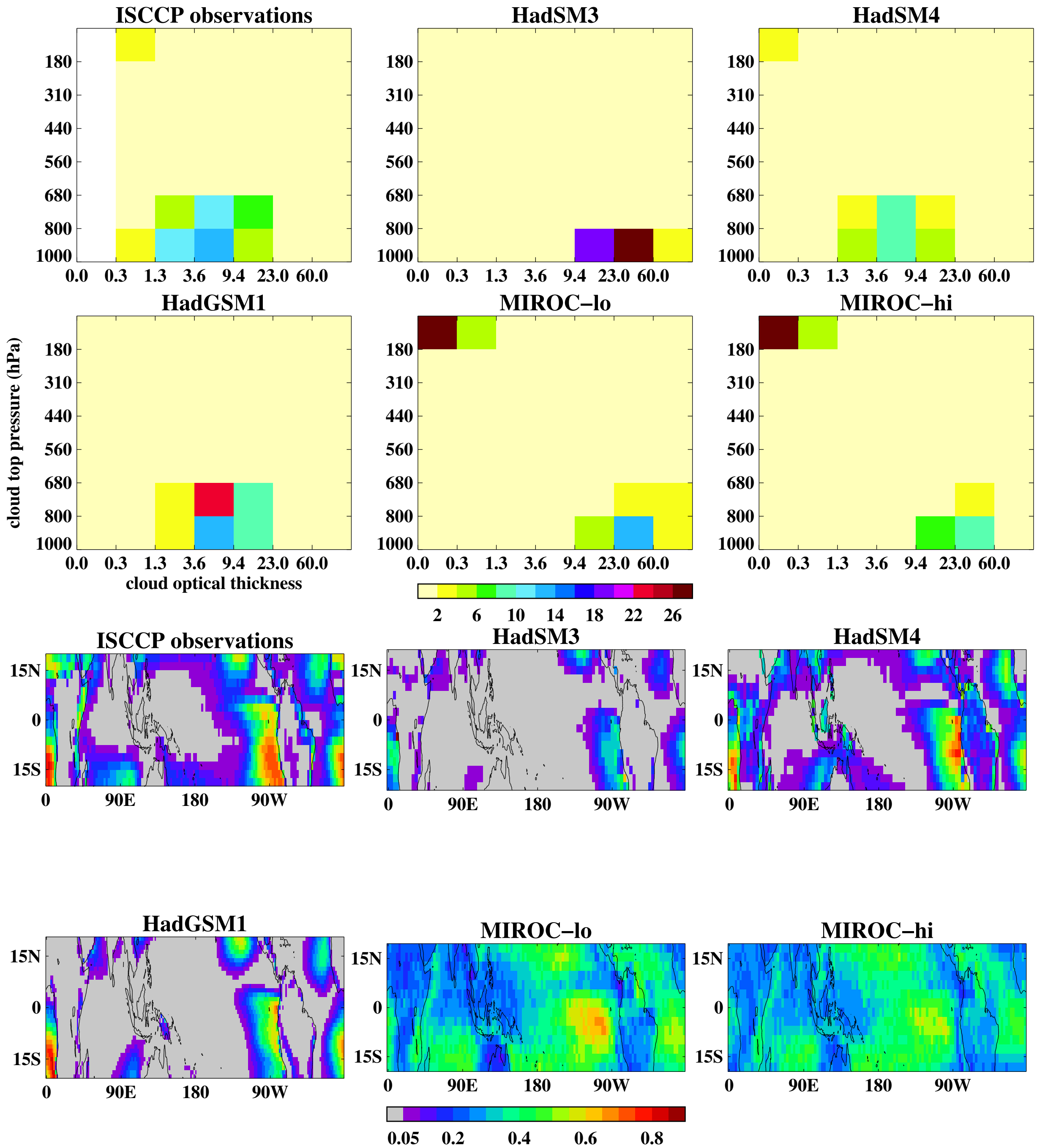


Figure 4: Example of CloudSat Simulator statistical summary output. The figure shows a height-reflectivity histogram calculated from one GCM gridbox of the MMF using CICCS with Quickbeam. Two hydrometeor species are used, cloud ice and water. These are amenable to cluster analysis as in Zhang et al, 2007 and can be used to validate models in free running climate mode. (Yuying Zhang, LLNL).

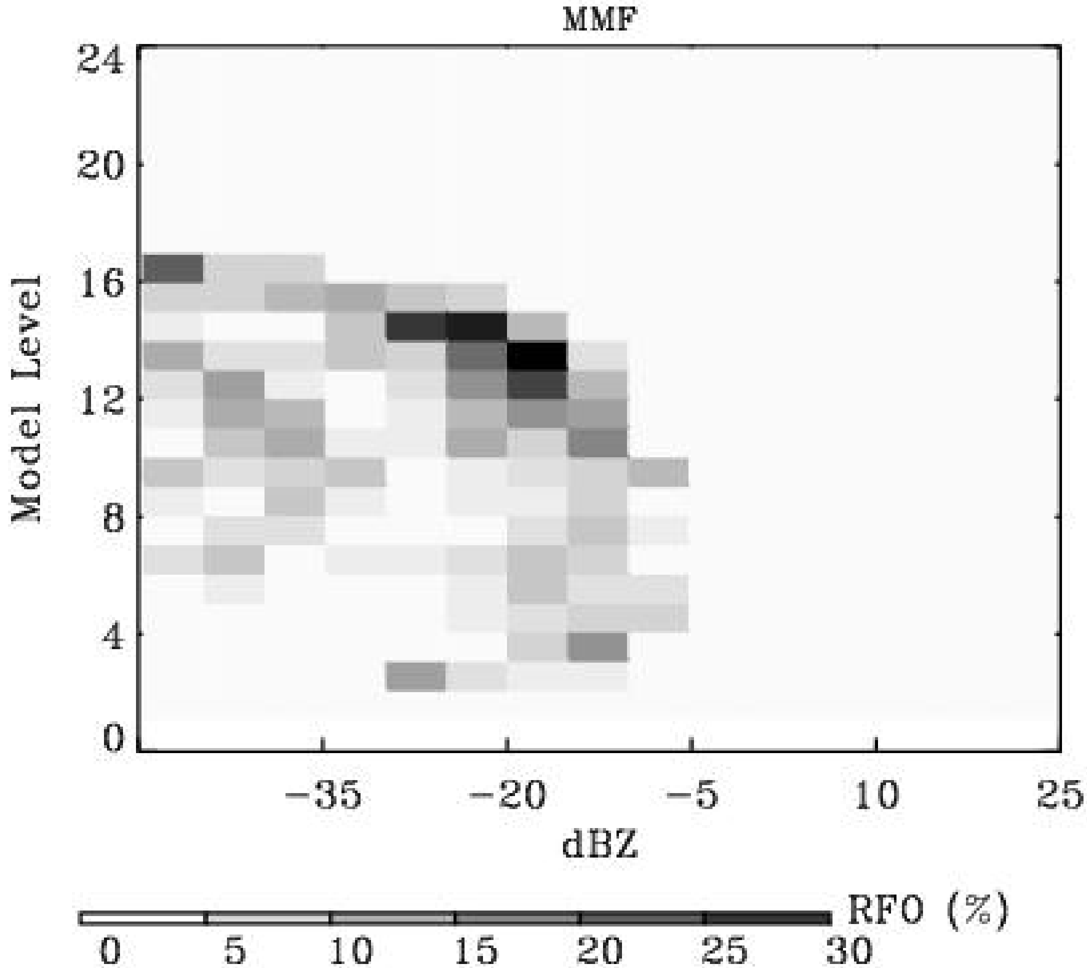


Figure 5: Example of a simulated mid-latitude system in the North Atlantic by the UK Met Office global forecast model on July 7th, 2006. The upper panel is the simulated reflectivity from the model outputs at sub-grid scale. The number of subcolumns per gridbox is 20, which gives a an effective resolution of 2 km at midlatitudes, comparable to the CloudSat footprint size. The lower panel shows the radar reflectivity (in dBZ) observed by CloudSat. (Alejandro Bodas-Salcedo, Met Office Hadley Centre).

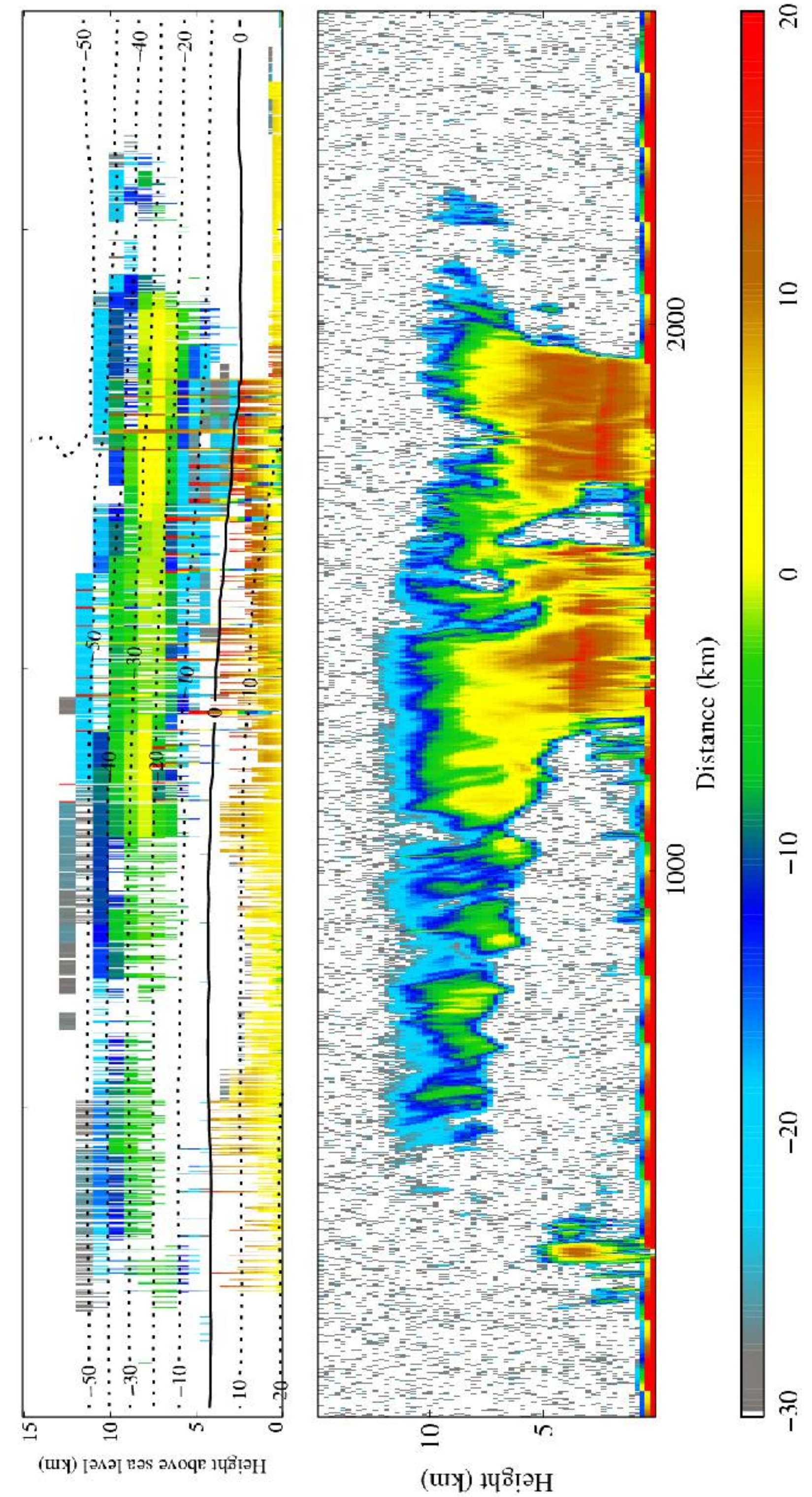


Figure 6: Ensemble mean composites from nine slab models using changes in 500mb vertical velocity (ω_{500}) and saturated lower tropospheric stability (Θ'_{es}) (Williams et al 2006). a) Relative frequency of occurrence (RFO) of changes in ω_{500} and Θ'_{es} at CO_2 doubling. b) SW cloud radiative forcing response composited by ω_{500} and Θ'_{es}. c) and d) as b) but for LW and Net components. e) - f) as a) - d) but for present day spatio-temporal variations.

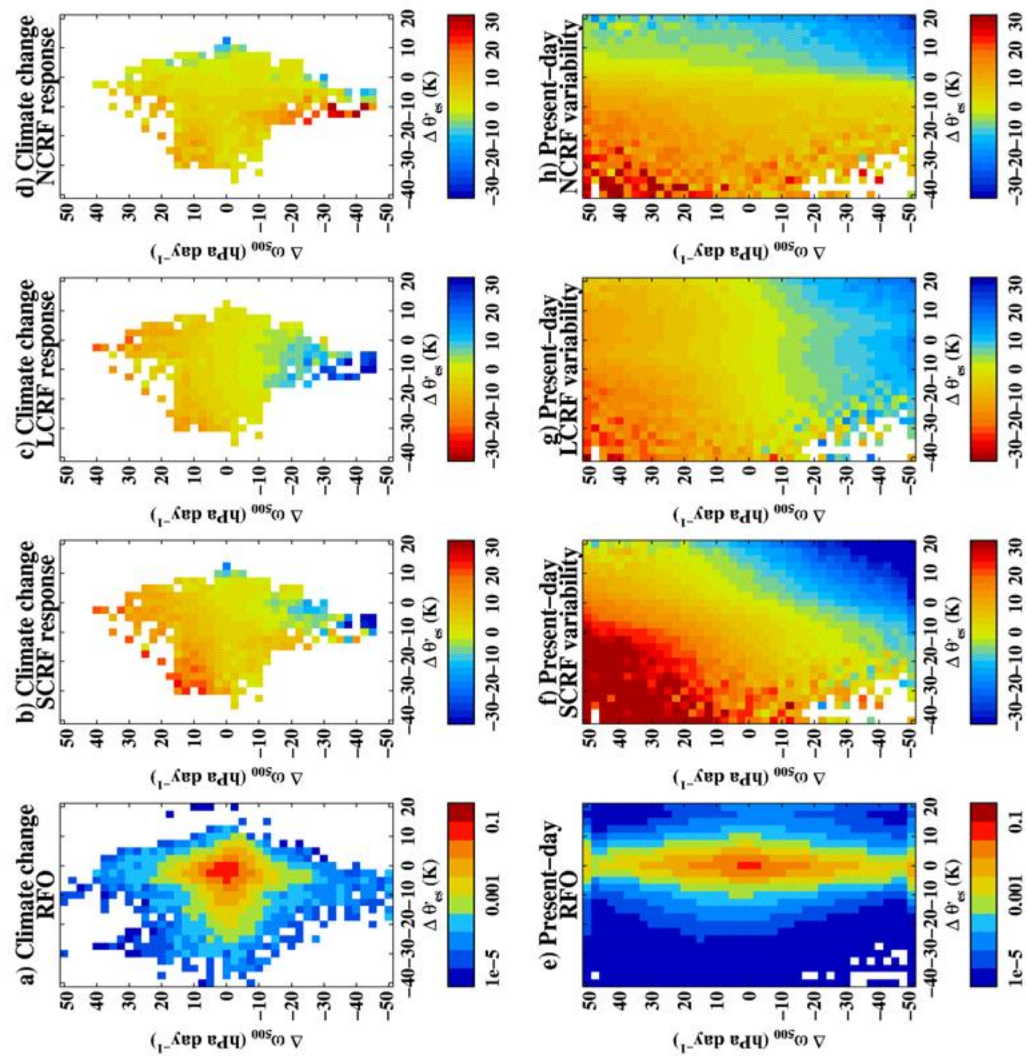


Table 1: RMS differences of ten models' present-day ω_{500} and Θ'_{es} cloud radiative forcing composites with ERBE/NCEP, adapted from Williams et al 2006. In each column, the five models with the smallest RMS differences are in blue, with the remaining five in red. The five models with the smallest average RMS difference have (on average) higher climate sensitivities.

ERBE /ERA SW	ERBE /ERA LW	ERBE/ NCEP SW	ERBE/ NCEP LW	Avg RMS	Avg Climate Sens.& Range
1.2	1.0	1.6	0.9	1.2	3.8K
1.3	1.2	1.3	1.1	1.2	
1.5	1.0	1.9	1.0	1.4	2.9
1.8	1.2	2.1	1.1	1.5	-4.4K
1.7	1.1	2.2	1.2	1.5	
2.1	1.1	2.3	1.0	1.6	3.0K
2.4	1.0	3.0	1.0	1.9	
1.7	1.6	2.2	1.9	1.9	2.3
2.2	1.4	2.8	1.5	2.0	-3.5K
3.4	1.3	3.7	1.1	2.4	

Figure 7: Zonal Mean for January-February-March zonal mean (top left) attenuated backscatter signal measured by Caliop/CALIPSO (in m-1sr-1), (top right) cloud fraction deduced from the observed lidar signal, (bottom left) cloud fraction predicted by the LMD GCM and (bottom right) cloud fraction diagnosed from the attenuated backscatter signal simulated from the GCM outputs. From Chepfer et al. (2007).

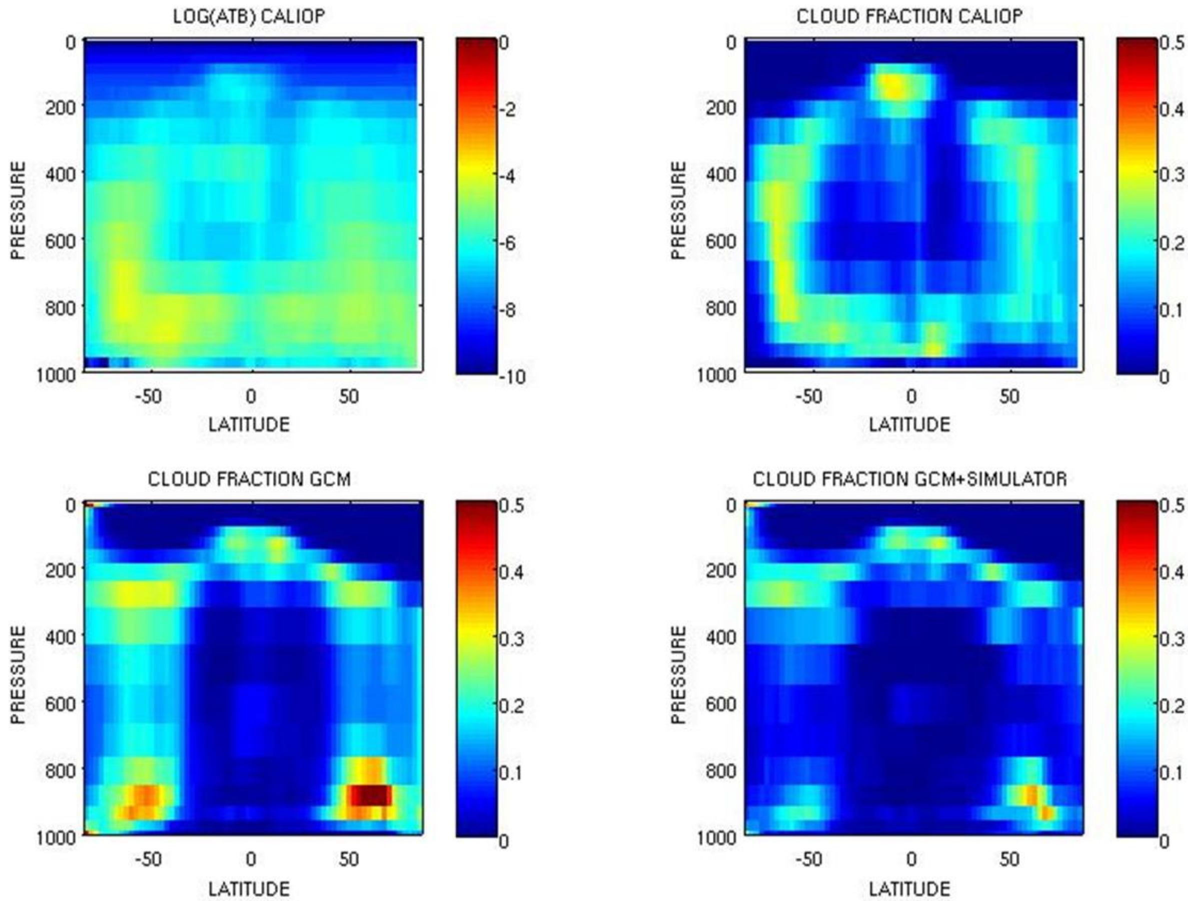


Figure 8: October time-height sections of relative humidity (shading, darker grey = more humid) and cloud fraction (blue contours every 10%) at 85W, 20S in the SE Pacific stratocumulus regime from climatological CAM3 (top) and AM2 (bottom). From Bretherton et al. (2006).

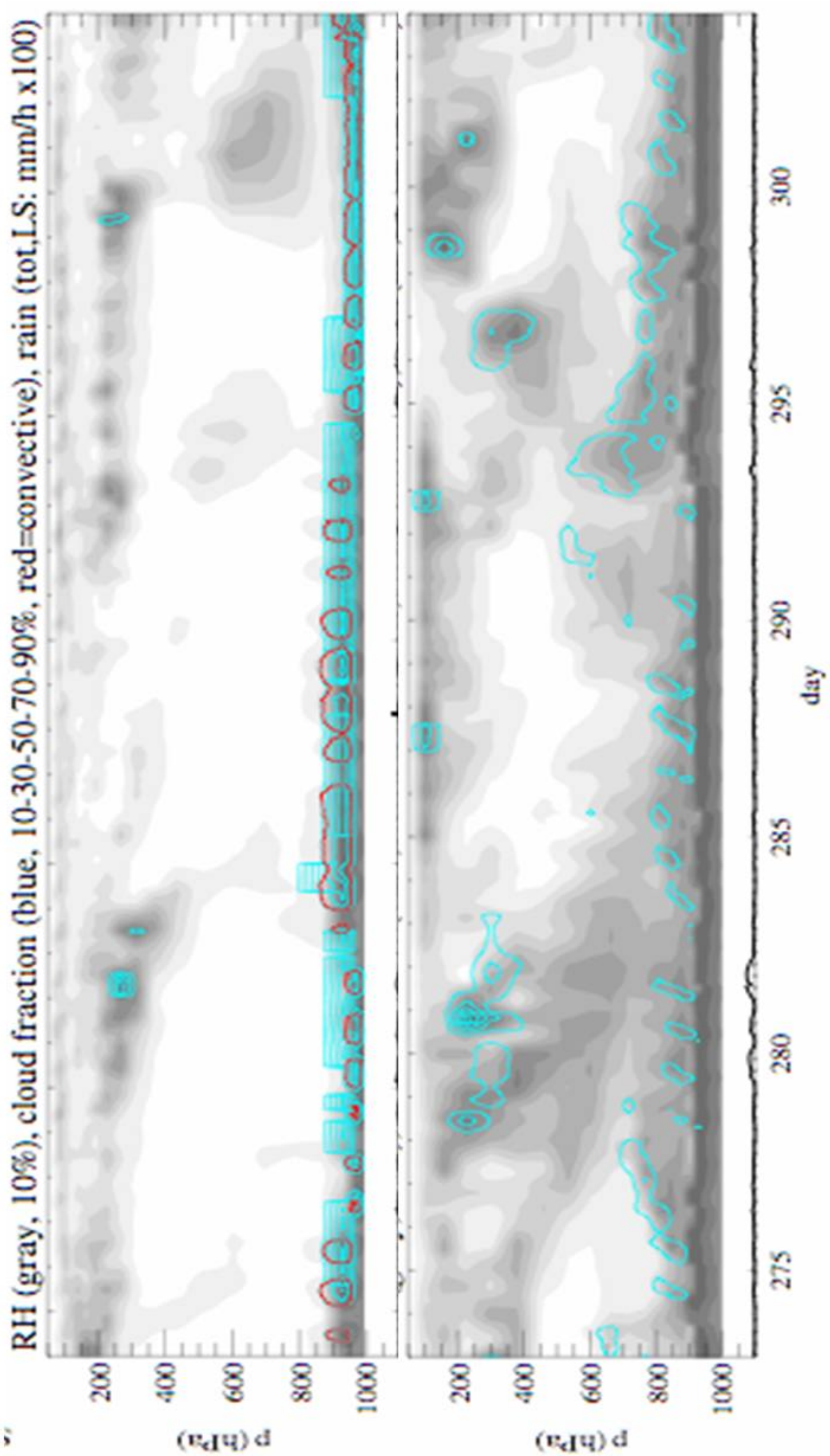


Figure 9: Diurnal cycle in low cloud cover in three models along the GCSS Pacific Cross Section (125W,35N) to (187W,1S). From the WGNE-GCSS Pacific Cross Section Inter-comparison (GPCI) Project. (Joao Teixeira)

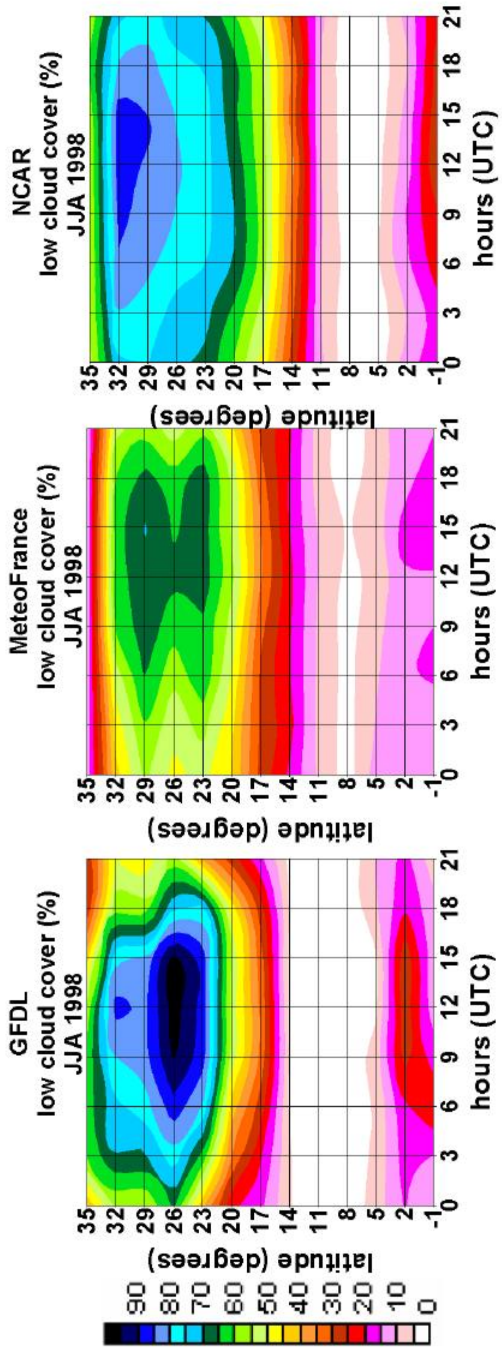
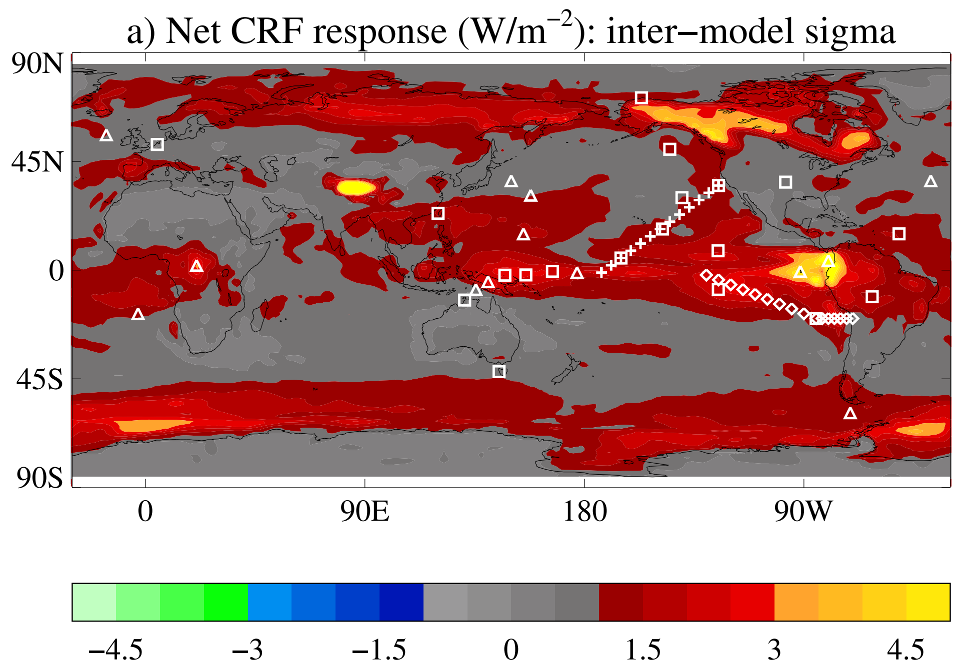


Figure 10: Potential locations for high frequency model output. a) Shows the inter-model standard deviation in the net CRF (Cloud Radiative Forcing) response to CO_2 doubling in twelve slab models. b) Shows the correlation of the local net CRF responses for these models with the global mean. The GPCI cross section locations are marked with +'s. The CPT locations are marked with squares. The diamonds show a South Pacific cross section produced by extending the VOCALS transect (75-85W,20S) towards the north west to intersect with a region of large inter-model spread which samples a range of positive and negative low cloud feedbacks. The triangles sample a number of locations with CRF responses that correlate well with the global means, many of which are in the climatological trade cumulus regions.



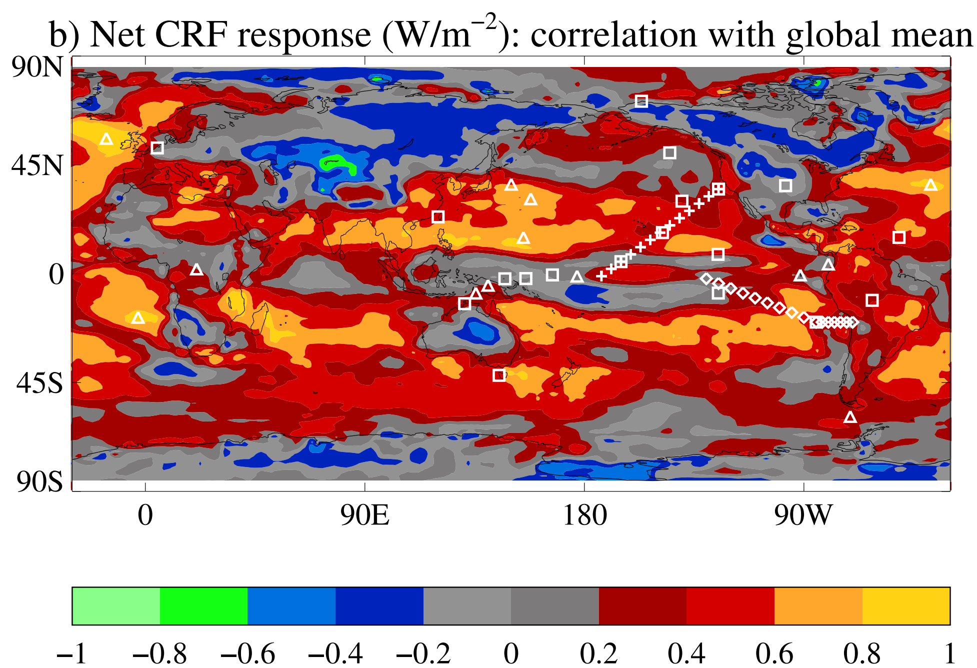


Figure 11: Temporal correlations between cloud condensate and sources of the cloud tendency equation in response to instantaneous CO2 doubling (zonal and annual averages). Cloud sources: (a,d) condensation minus evaporation, (b,e) ice sedimentation, (c) deposition minus sublimation. (a-c) HadGEM1, (d-e) MIROC3.2. Zero values are assigned to negative correlation coefficients, and the sign is changed to negative when correlation coefficient is positive and cloud response is negative. Contours denote temperature (0 C: 1xCO2,-15C and -40C: 2xCO2). (Tomoo Ogura)

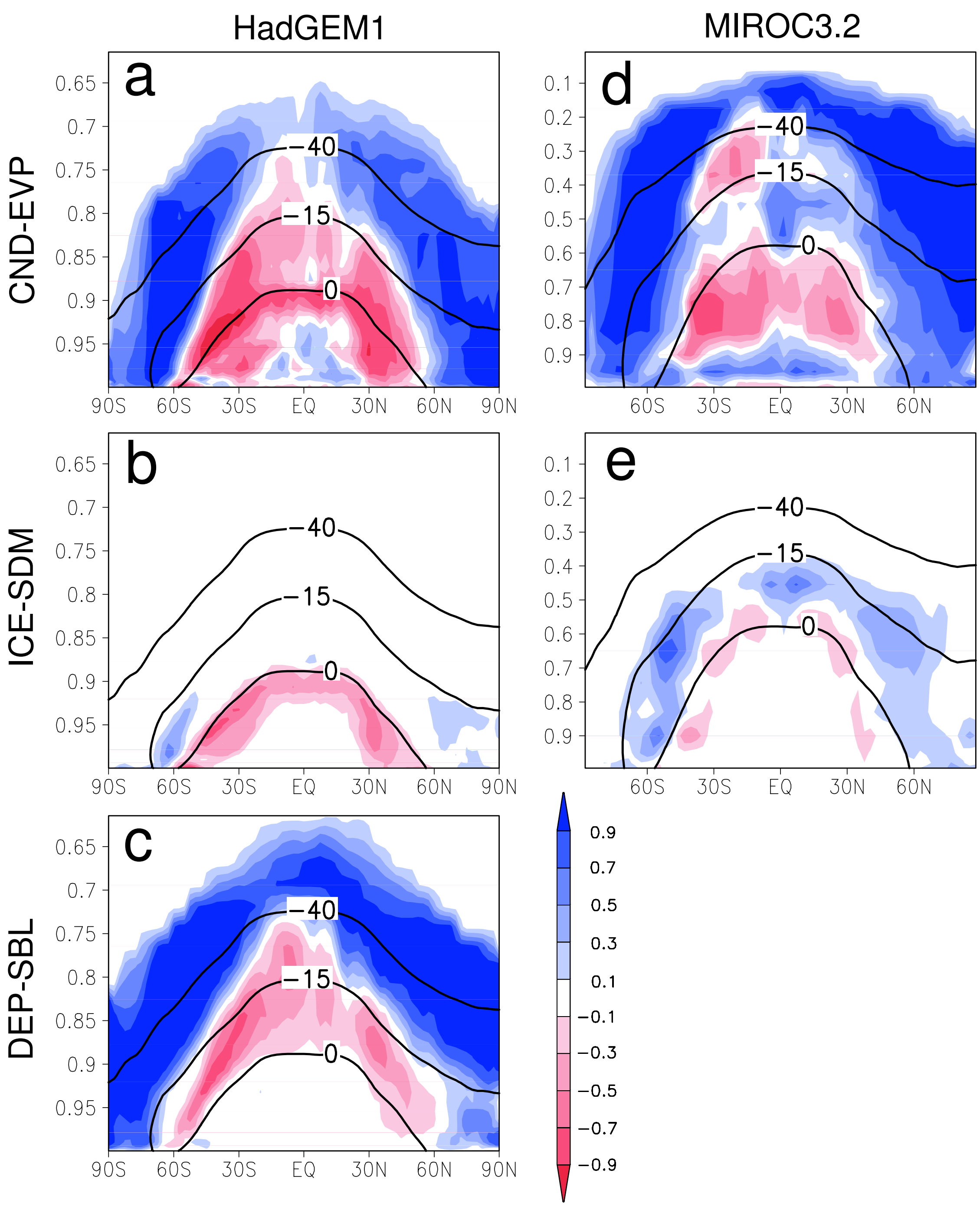


Figure 12: (a) Schematic setup of large-scale advective forcing of temperature and water vapour over low clouds. (b) Profiles of subsidence rate over the cold pool for the control climate (blue) and the warmer climate. In the warmer climate, sea surface temperature is raised by 2 degrees. (Minghua Zhang)

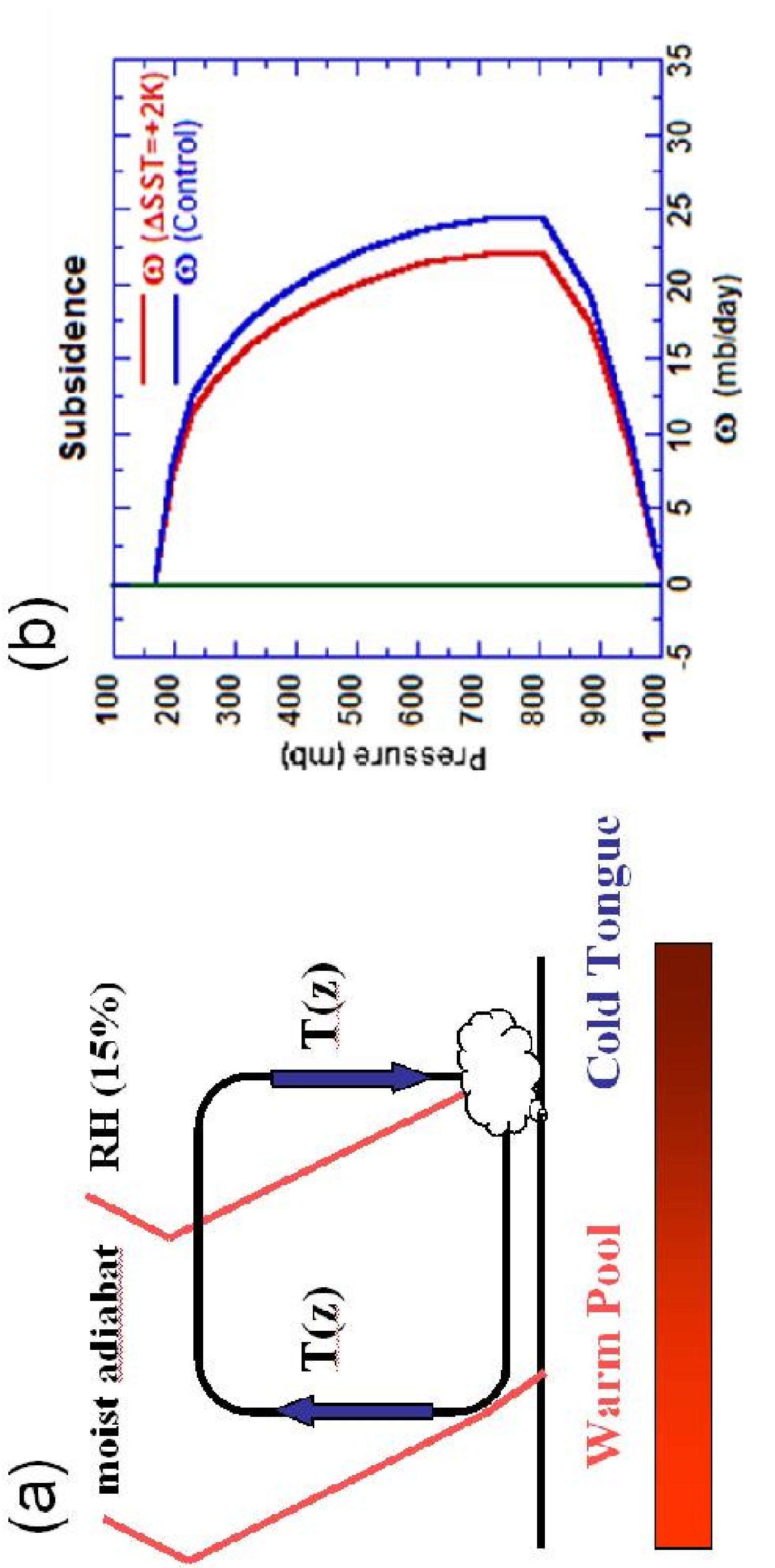


Table 2: Diagnostic recommendations for CFMIP-2 and CMIP4 experiments.

		Package 1	Package 2	Package 3	
	Time Sampling Frequency	CFMIP I 2D fields incl. 2D ISCCP simulator, fluxes, water paths, clt, stability, w500	CFMIP I 3D fields incl. ISCCP, cl,clw, omega, u,v,t,q	CloudSat /CALIPSO Simulator	Cloud/T/q tendencies
1) CMIP4/AR5 Present day experiments (AMIP, 20C)					
	Monthly	All (A)	All (B)	All (G)	
16 yr daily = 85-90 & 98-07	Daily	All (D)	16 (C)	16 (H)	
23 year sub-period 85-07	3hr/daily Point	23 (F)	23 (F)	23	
2) CMIP4/AR5 Coupled model 1%/yr simulations (and control)					
	Monthly	All (A)	All (B)	All (G)	
(Sub-periods centred on time of	Daily	All (D)	16 (C)	16 (H)	
CO2 doubling and equiv in ctrl)	3hr/daily Point	20 (F)	20 (F)	20	
3) CMIP4/AR5 scenarios	Monthly	All (A)	All (B)		
4) CFMIP-2 reference experiments (Short AMIP, AMIP Clim, AMIP Clim+Soden SST, AMIP Clim+2CO2)					
	Monthly	All (A)	All (B)	All (G)	All (E)
	Daily	All (D)	10\16 (C)	10\16 (H)	
	3hr/daily Point	20\23 (F)	20\23 (F)	20\23	20\23
5) CFMIP-2 sensitivity tests (Short AMIP, AMIP Clim, AMIP Clim+Soden SST, AMIP Clim+2CO2)					
	Monthly	All (A)	All (B)	All (G)	All (E)
	Daily	All (D)			
	3hr/daily Point	20\23 (F)	20\23 (F)	20\23	20\23

A: required for Bony and Dufresne 2005, Taylor et al 2007, Gregory and Webb 2007
B: required for Webb et al 2006, Ringer et al 2006, Tsushima et al 2006, Wyant et al 2006
C: required for Williams and Tselioudis 2007
D: for lightweight vn of Williams and Tselioudis 2007
E: for Ogura et al 2007
F:for GPCI studies (Teixeira), and Bretherton et al 2006 timeseries
G: For equivalent of Mace et al, Chepfer et al 2007
H: For equivalent of Zhang et al 2007