Climate Dynamics (2006) 27: 17–38
DOI 10.1007/s00382-006-0111-2

M. J. Webb · C. A. Senior · D. M. H. Sexton
W. J. Ingram · K. D. Williams · M. A. Ringer
B. J. McAvaney · R. Colman · B. J. Soden
R. Gudgel · T. Knutson · S. Emori · T. Ogura
Y. Tsushima · N. Andronova · B. Li · I. Musat
S. Bony · K. E. Taylor

On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles

Received: 1 April 2005 / Accepted: 7 December 2005 / Published online: 4 February 2006

Abstract Global and local feedback analysis techniques have been applied to two ensembles of mixed layer equilibrium CO_2 doubling climate change experiments, from the CFMIP (Cloud Feedback Model Intercomparison Project) and QUMP (Quantifying Uncertainty in Model Predictions) projects. Neither of these new ensembles shows evidence of a statistically significant change in the ensemble mean or variance in global mean climate sensitivity when compared with the results from the mixed layer models quoted in the Third Assessment Report of the IPCC. Global mean feedback analysis of these two ensembles confirms the large contribution made by inter-model differences in cloud feedbacks to those in climate sensitivity in earlier studies; net cloud feedbacks are responsible for 66% of the inter-model variance in the total feedback in the CFMIP ensemble and 85% in the QUMP ensemble. The ensemble mean global feedback components are all statistically indistinguishable between the two ensembles, except for the clear-sky shortwave feedback which is stronger in the CFMIP ensemble. While ensemble variances of the shortwave cloud feedback and both clear-sky feedback terms are larger in CFMIP, there is considerable overlap in the cloud feedback ranges; QUMP spans 80% or more of the CFMIP ranges in longwave and shortwave cloud feedback. We introduce a local cloud feedback classification system which distinguishes different types of cloud feedbacks on the basis of the relative strengths of their longwave and shortwave components, and interpret these in terms of responses of different cloud types diagnosed by the International Satellite Cloud Climatology Project simulator. In the CFMIP ensemble, areas where low-top cloud changes constitute the largest cloud response are responsible for 59% of the contribution from cloud feedback to the variance in the total feedback. A similar figure is found for the QUMP ensemble. Areas of positive low cloud feedback (associated with reductions in low level cloud amount)

M. J. Webb (✉) · C. A. Senior · D. M. H. Sexton
W. J. Ingram · K. D. Williams · M. A. Ringer
Hadley Centre for Climate Prediction and Research,
Met Office, FitzRoy Road, Exeter, EX1 3PB UK
E-mail: Mark.Webb@metoffice.gov.uk
Tel.: +44-1392-884515
Fax: +44-1392-885681

B. J. McAvaney · R. Colman
Bureau of Meteorology Research Centre (BMRC),
Melbourne, Australia

B. J. Soden
Rosenstiel School for Marine and Atmospheric Science,
University of Miami, Miami, FL, USA

R. Gudgel · T. Knutson
Geophysical Fluid Dynamics Laboratory (GFDL),
Princeton, NJ, USA

S. Emori · T. Ogura
National Institute for Environmental Studies (NIES),
Tsukuba, Japan

Y. Tsushima
Frontier Research Center for Global Change (FRCGC),
Japan Agency for Marine–Earth Science and Technology,
Kanagawa, Japan

N. Andronova
Department of Atmospheric,
Oceanic and Space Sciences,
University of Michigan,
Ann Arbor, MI, USA

B. Li
Department of Atmospheric Sciences,
University of Illinois at Urbana–Champaign (UIUC),
Urbana, IL, USA

I. Musat · S.Bony
Institut Pierre Simon Laplace (IPSL),
Paris, France

K. E. Taylor
Program for Climate Model Diagnosis and Intercomparison
(PCMDI), Livermore, CA, USA

contribute most to this figure in the CFMIP ensemble, while areas of negative cloud feedback (associated with increases in low level cloud amount and optical thickness) contribute most in QUMP. Classes associated with high-top cloud feedbacks are responsible for 33 and 20% of the cloud feedback contribution in CFMIP and QUMP, respectively, while classes where no particular cloud type stands out are responsible for 8 and 21%.

1 Introduction

Estimates of the equilibrium near-surface warming resulting from a doubling of CO_2 relative to pre-industrial levels (often referred to as the climate sensitivity) vary among comprehensive general circulation models by several degrees. The Third Assessment Report (TAR) of the Intergovernmental Panel on Climate Change (IPCC; Cubasch et al. 2001) reported the range in climate sensitivity from mixed layer ocean experiments with contemporary climate models to be 2.0–5.1 K (see their Table 9.4, p 560).

Cess et al. (1990, 1996), Le Treut and Li (1991), Senior and Mitchell (1993), Le Treut and McAvaney (2000), Colman (2003) and others have concluded that differences in cloud feedback make a large contribution to differences in climate sensitivity in GCMs. However, few of these intercomparisons have gone beyond global or zonal mean feedback analysis, and the lack of consistent methods or detailed cloud diagnostics in models has made it difficult to compare modelled cloud responses in a quantitative manner, meaning that the reasons for the differences in cloud response remain unclear. A number of recent developments however warrant a re-visit of the issue.

There have been significant advances in methods used to analyse and understand model feedbacks in relation to each other and to satellite observations. Yu et al. (1996) developed an approach for detailed and quantitative comparisons of clouds in climate models with observations from the International Satellite Cloud Climatology Project (ISCCP; Rossow and Schiffer 1999) which mimics the satellite view from space (along with certain ISCCP retrieval assumptions), yielding diagnostics that can be quantitatively compared with the observational retrievals of cloud amount, cloud top pressure and cloud optical thickness. Klein and Jakob (1999) and Webb et al. (2001) developed the 'ISCCP simulator', an implementation of this approach which can be run within models and which has been used in a number of studies (e.g. Tselioudis et al. 2000; Zhang et al. 2005; Williams et al. (2005b). Other studies (some of which use the ISCCP simulator) have developed compositing techniques to relate present day variations in clouds and/or radiative fluxes to variations in thermodynamic and dynamical variables in the models (Bony et al. 2004; Williams et al. 2003, 2005a, b; Bony and Dufresne 2005; M.C. Wyant et al., submitted). Studies of local feedbacks in individual climate models have also emerged; Colman (2002) applied methods derived from Wetherald and Manabe (1988; the so-called partial radiative perturbation (PRP) approach) to the Bureau of Meteorology Research Centre (BMRC) model, while a local analysis of feedbacks in the Canadian Climate Model (Boer and Yu 2003) separated the effects of clouds and clear-sky effects using a measure of cloud feedback based on cloud radiative forcing method of Cess and Potter (1988). These developments in cloud diagnosis and feedback analysis techniques are now being applied across a range of models as part of the Cloud Feedback Model Intercomparison Project (CFMIP; McAvaney and Le Treut 2003; http://www.cfmip.net). CFMIP is a World Climate Research Programme (WCRP)/Working Group on Coupled Modelling (WGCM) endorsed project which aims to perform a systematic intercomparison of cloud feedbacks in climate models and to evaluate aspects of their performance that are relevant to uncertainty in those feedbacks. The CFMIP experimental protocol requires participants to include the ISCCP simulator.

The so-called perturbed physics ensembles have explored the dependence of climate sensitivity of a single mixed-layer climate model on perturbations in model parameters. Murphy et al. (2004) report a 5–95% range of 1.9–5.3 K, estimated statistically by linearly combining the responses of single parameter perturbations in 53 versions of HadSM3 in the first ensemble from the Quantifying Uncertainties in Model Predictions project (QUMP). Stainforth et al. (2005) analysed a grand ensemble in which 6 of the 29 QUMP parameters were perturbed in combination. A second QUMP ensemble is now available which perturbs 29 parameters in combination in 128 ensemble members, with a comprehensive set of diagnostics including the ISCCP simulator (see Sect. 2 for details). The ranges in climate sensitivity in both QUMP ensembles come close to encompassing the equivalent range from the mixed layer models reported in the TAR, but previous studies have not established what feedback processes contribute most to this range, and whether or not these are in any way similar to those found in multi-model ensembles.

Drawing on these developments, we identify the local feedbacks which contribute most to the inter-model differences in climate sensitivity in the CFMIP and QUMP ensembles, and relate these to the responses in different cloud types (e.g. low and high top). We also quantify the extent to which the ranges of various feedbacks overlap in the two ensembles. We consider the direct GCM results only, and do not attempt a formal quantification of the uncertainty in climate sensitivity or individual climate feedbacks in the probabilistic sense used by Murphy et al. (2004).

2 Model descriptions and experimental design

Each of the CFMIP and QUMP experiments was carried out with a mixed layer 'slab' ocean model coupled to an atmospheric GCM. In each case the steady-state climate was simulated for a 'control' concentration of CO_2 and for a doubled concentration.

A brief description of the CFMIP ensemble follows. The UIUC model is described in Yang et al. (2000) and Andronova et al. (1999). Its atmospheric resolution is N36/L24 and it has prognostic equations for cloud liquid water and ice. The BMRC model is described in Colman et al. (2001). Its atmospheric resolution is T47/L17 and it has prognostic equations for cloud liquid water and ice (Rotstayn 1997). The GFDL AM2 'slab model' used in this study incorporates an atmospheric model very similar, though not identical, to the AM2 atmospheric model described in GFDL GAMDT (2004) and uses atmospheric and sea ice components very similar those in the CM2.0 coupled climate model described in Delworth et al. (2006). Atmospheric resolution is N72/L24, and it has prognostic equations for cloud liquid water and ice. HadSM3 is a version of the Hadley Centre climate model and is described by Pope et al. (2000) and Williams et al. (2001a, b). Atmospheric resolution is N48/L19, and the model has prognostic equations for total water (vapour plus liquid/ice) and liquid/ice water temperature. Cloud water/ice and cloud fraction are diagnosed using a symmetric probability density function for subgrid total water variations (Smith 1990). HadSM4 is a development version of HadSM3, described in Webb et al. (2001). Atmospheric resolution is N48/L38, and this model has a prognostic equation for cloud ice (Wilson and Ballard 1999), while warm clouds are as in Smith (1990), with modifications to allow cloud to partly fill a model layer in the vertical. HadGSM1 is the mixed layer version of HadGEM1, the first of a new generation of Hadley Centre climate models (Martin et al. 2005; Johns et al. 2006). Atmospheric resolution is N96/L38, and while there are changes to many areas of the atmospheric physics and dynamics, the treatment of clouds is similar to HadSM4. The Institut Pierre Simon Laplace (IPSL) model is described in Hourdin et al. (2006). The atmospheric resolution is N48/L19, and the cloud parametrisation is based on Le Treut and Li (1991). Cloud cover is predicted through a statistical cloud scheme that uses a (skewed) generalised log-normal distribution to describe the subgrid-scale variability of total water. In convective areas the cloud scheme is coupled to the cumulus convection scheme after Bony and Emanuel (2001). The slab model version submitted to CFMIP uses climatological sea ice. Two model versions of the T42/L20 MIROC3.2 model have been submitted to CFMIP from the CCSR/NIES/FRCGC group in Japan (K-1 model developers 2004). The cloud parametrisation is based on Le Treut and Li (1991). The 'lower-sensitivity' version (MIROC low) is exactly the same as the 'medium-resolution' version of this model submitted to the IPCC Fourth Assessment Report archive and described in the above reference. The 'higher-sensitivity' version (MIROC high) is the same as the above except for the following two points in the treatment of clouds: (1) the temperature range for the existence of mixed phase (liquid + ice) cloud is −25 to −5°C in the higher-sensitivity version, while it is −15 to 0°C in the lower-sensitivity version; (2) when cloud ice melts, it is converted to cloud water in the higher-sensitivity version, while it is converted to rain in the lower-sensitivity version. For a fuller description of the two versions, see T. Ogura et al. (submitted). Note that the sea ice scheme employed in the mixed layer model is a thermodynamic scheme, while the version run in the coupled version of this model is more sophisticated.

The second QUMP ensemble is made up of 128 versions of HadSM3, where a selection of 29 uncertain parameters were perturbed in combination; i.e. each ensemble member was run with multiple parameter perturbations (MPPs). The parameters used were the same as those in the first QUMP ensemble (Murphy et al. 2004), in which only one parameter was perturbed in each ensemble member (denoted as SPP for single parameter perturbation). The parameters were selected from the major components of atmospheric and surface physics in the GCM, namely large scale cloud, convection, radiation, boundary layer, dynamics, land surface processes and sea ice, in order to sample uncertainties in all the major surface and atmospheric processes in the model. The MPPs were chosen to cover a range of climate sensitivities while maximising, as far as possible, coverage of parameter space and model skill for a given climate sensitivity. Model skill is defined in this case as the ability to reproduce time-averaged present day climate according to the climate prediction index (CPI) introduced in Murphy et al. (2004). The MPPs for the second ensemble were generated using the following procedure. The size of the ensemble was chosen to be 128 members, on the basis of the computing resources available. Based entirely on information contained in the SPP ensemble, an earlier version of the linear prediction method of Murphy et al. (2004) was used to predict the climate sensitivities and CPI values associated with 3.6 million possible MPPs generated by random sampling from (assumed) uniform prior distributions for each parameter within its expert-specified limits. These possible perturbations were then split into 64 equi-probable bins in climate sensitivity, and the 20 runs with the best predicted CPI from each bin were then selected. From this 1,280 member set of MPPs, the combination with the highest predicted skill was selected first, and the second was chosen to be the combination that was the furthest away from the first in parameter space. Subsequent MPPs were selected such that they were in turn the farthest from all previously selected MPPs, with the added constraint that no more than 2 could be selected from any one bin, ensuring the selection of 128 MPPs in total. Full slab ocean experiments including calibration,

control and equilibrium CO_2 doubling components were then run for each of the 128 MPPs, with the addition of an interactive sulphur cycle as in Williams et al. (2001a). Sexton et al. (2004) provide additional detail on the design of the MPP ensemble.

In the CFMIP models, various methods were used to estimate the radiative forcing due to a doubling of CO_2. The values reported here are estimates at the tropopause with some allowance for stratospheric adjustment, except where stated otherwise below. The estimates for the Hadley Centre and CCSR/NIES/FRCGC model versions use the method of Tett et al. (2002). The GFDL model estimate uses the method of Hansen et al. (2002) diagnosing the net forcing at the surface. The BMRC model uses a method based on that of Colman et al. (2001) at the top of the atmosphere, while the IPSL model uses the method of Joshi et al. (2003). The UIUC group calculate an instantaneous forcing only, diagnosed at the tropopause. This value is used unadjusted, although reducing the forcing by 15% (a typical stratospheric adjustment) does not substantially affect the results presented below. For all of the QUMP model versions, the same value of CO_2 forcing is assumed (3.74 Wm^{-2}) based on the unperturbed version of the model, as in Murphy et al. (2004).

3 Analysis of global means

Figure 1a and b shows the climate sensitivity due to doubling CO_2 for the CFMIP models and a subset of the QUMP model versions. Estimates for the radiative forcing due to CO_2 doubling are also shown for the CFMIP and QUMP models. Figure 1c shows the climate sensitivities from the 16 mixed layer model versions in Table 9.1 (p 538) of the TAR. Various statistics of climate sensitivity are shown in Table 1 for the three ensembles. We consider whether there is evidence of a statistically significant change in the statistics of climate sensitivity between the TAR models and the newer

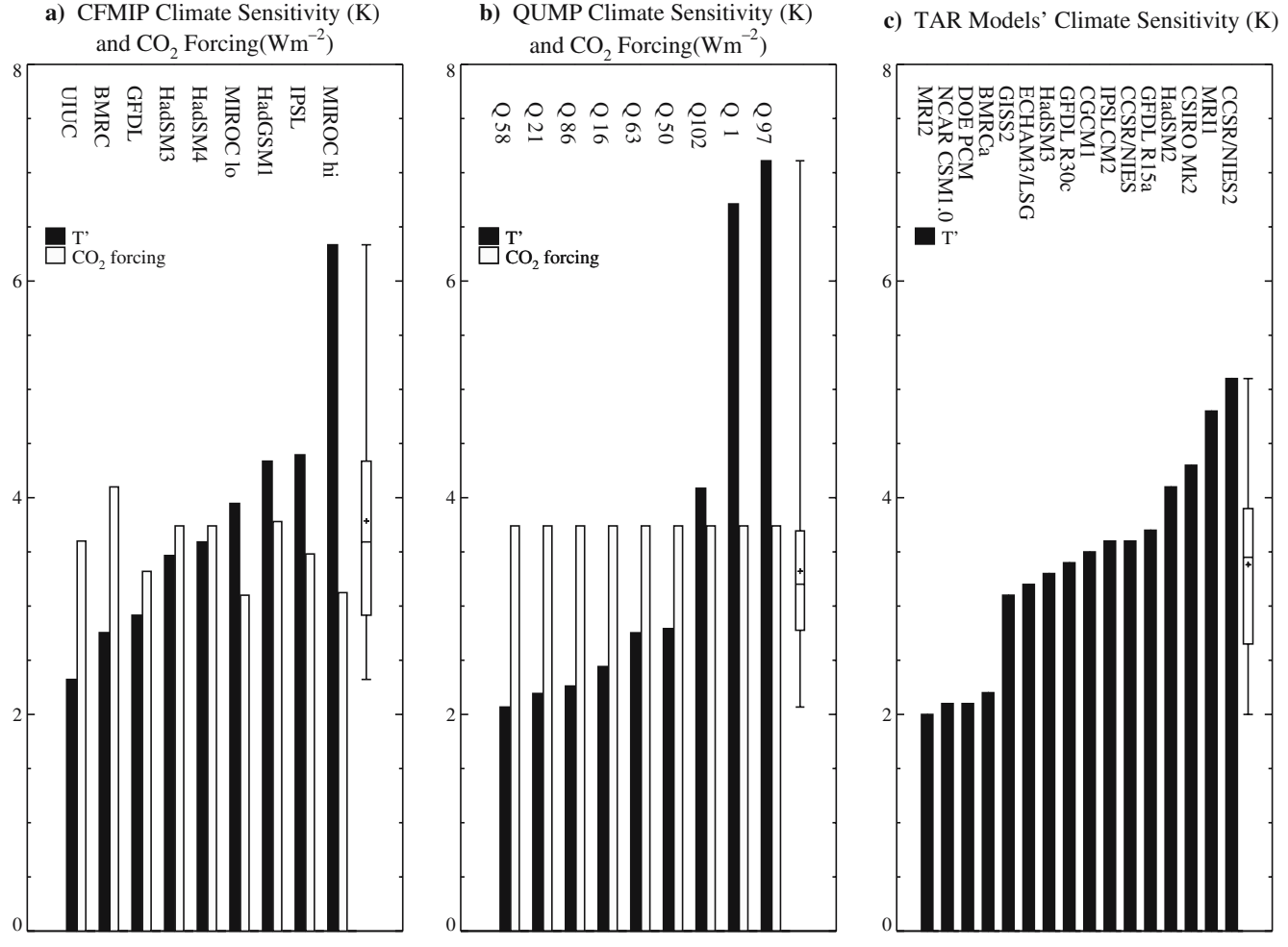

Fig. 1 Climate sensitivities of **a** CFMIP, **b** selected QUMP and **c** TAR mixed layer models (K). Also shown are the forcing estimates for CO_2 doubling in the CFMIP and QUMP models. The models are shown in the order of increasing climate sensitivity. The *box* and *whisker plots* summarise the climate sensitivity distributions and show the boundaries between the quartiles of the distributions, along with the mean (*marked with a* +) and the median (*horizontal line*) values. In (**b**) these summarise all 128 QUMP model versions. See caption for Fig. 2 for description of QUMP selection method

Table 1 Statistics of climate sensitivity for CFMIP, QUMP and TAR mixed layer models

	n	Mean	Median	SD	Range	Spread
TAR	16	3.4	3.5	0.9	2.0–5.1	3.1
CFMIP	9	3.8	3.6	1.2	2.3–6.3	4.0
QUMP	128	3.3	3.2	0.8	2.1–7.1	5.0

The spread is the difference between the maximum and minimum values
n Number of models in each ensemble

model versions in the CFMIP ensemble. The CFMIP ensemble mean climate sensitivity is 0.4 K higher than the TAR equivalent. The median is also slightly higher. In a two-tailed Student's *t* test we cannot reject the null hypothesis that the TAR and CFMIP model climate sensitivities are drawn from underlying distributions with the same means at the 10% confidence level. Although we assume that the climate sensitivities from the two ensembles represent independent random samples, reducing the number of degrees of freedom in the test would only strengthen this result. We deduce that the smaller change in the median is also unlikely to be statistically significant. Table 1 also shows the standard deviation and spread (max–min) values of climate sensitivity which give some measure of the inter-model differences in climate sensitivity in each ensemble. Both are approximately 30% larger in the CFMIP ensemble compared to the TAR ensemble, but an *F* test shows that the differences in standard deviations are not statistically significant at the 10% confidence level. The 128 member QUMP ensemble mean climate sensitivity is close to that of the TAR ensemble and the differences are again not significant. However, the CFMIP standard deviation is significantly larger than the QUMP standard deviation. It is interesting to note that even though the CFMIP ensemble has the larger variance, the QUMP ensemble has the wider range.

Using the notation of Boer and Yu (2003), any change R' in the global mean net downward radiation at the top of the atmosphere R may be expressed as the sum of the global radiative forcing and response. The response term may be expressed as a product of the near surface air temperature response T' and the signed feedback parameter Λ (which is negative for a stable system) to give:

$$R' = f + \Lambda T' \tag{1}$$

where f is the global mean radiative forcing due to doubling CO_2. Once the slab models in question have been run to equilibrium, R' will be close to zero, and we can use the relationship:

$$\Lambda = -1/s = -f/T' \tag{2}$$

(which illustrates the relationship between sensitivity s and feedback Λ) to evaluate Λ for each of the CFMIP models. Fixing Λ at the ensemble mean value ($\bar{\Lambda} = -1.03\,\mathrm{Wm^{-2}K^{-1}}$) and solving for T' we find that only a small part of the range in climate sensitivity (3.0–4.0 K out of 2.3–6.3 K) is caused by forcing differences. Using an average forcing value ($\bar{f} = 3.55\,\mathrm{Wm^{-2}}$) the range actually increases to 2.3–7.2 K. This shows not only that the differences in Λ can themselves explain much of the range in climate sensitivity, but that forcing differences in the models act to reduce the range slightly. Repeating this exercise on the standard deviations gives equivalent results.

Re-ordering Eq. 1, the response term can be separated into a clear-sky response and a difference between the all-sky and clear-sky responses:

$$\Lambda T' = (R' - f) = (\tilde{R}' - \tilde{f}) + [(R' - f) - (\tilde{R}' - \tilde{f})] \tag{3}$$

where $\tilde{R}'$ is the change in the clear-sky net downward flux at the top of the atmosphere, and $\tilde{f}$ is the forcing calculated for clear-sky conditions. The second term may be expressed as the sum of the change in the net cloud radiative forcing $R'_C = R' - \tilde{R}'$ and the difference in the all-sky and clear-sky forcings to give:

$$\Lambda T' = (\tilde{R}' - \tilde{f}) + [R'_C + (\tilde{f} - f)] \tag{4}$$

As clear-sky forcings are not available for the CFMIP models, we use f in the place of $\tilde{f}$ in the analysis of both the CFMIP and QUMP ensembles, which gives:

$$\Lambda T' \approx (\tilde{R}' - f) + R'_C \tag{5}$$

With this approximation the total response is decomposed into clear-sky and cloud feedbacks, consistent with the approach of Cess and Potter (1988). We then define the clear-sky and cloud feedback terms to be:

$$\Lambda = \Lambda_A + \Lambda_C \approx (\tilde{R}' - f)/T' + R'_C/T' \tag{6}$$

The subscripts C and A indicate 'cloud' and 'atmosphere', although the clear-sky terms do of course also include contributions from surface feedback processes. These components can in turn be separated into shortwave and longwave components, for instance:

$$\Lambda_A = \Lambda_{SA} + \Lambda_{LA} \approx (\tilde{R}'_S - f_S)/T' + (\tilde{R}'_L - f_L)/T' \tag{7}$$

As separate longwave and shortwave forcings f_S and f_L are not available from the CFMIP models, we use f in the place of f_L and take f_S to be zero in the analysis of both ensembles:

$$\Lambda_{SA} + \Lambda_{LA} \approx \tilde{R}'_S/T' + (\tilde{R}'_L - f)/T' \tag{8}$$

Similarly, the cloud feedback term can be decomposed into shortwave and longwave components

$$\Lambda_C = \Lambda_{SC} + \Lambda_{LC} \approx R'_{SC}/T' + R'_{LC}/T' \tag{9}$$

as can Λ itself:

$$\Lambda = \Lambda_S + \Lambda_L = [\Lambda_{SA} + \Lambda_{SC}] + [\Lambda_{LA} + \Lambda_{LC}] \tag{10}$$

An alternative approach would be to employ a decomposition based on the PRP methods of Wetherald

and Manabe (1988) and Colman (2002) which give a somewhat cleaner separation of forcing and feedback components, avoiding the so-called cloud masking effect which (when the cloud forcing method is used) leads to an overestimate of the strength of clear-sky feedback terms and an underestimate of the strength of cloud terms (Zhang et al. 1994; Colman 2003; Soden et al. 2004). This approach however requires suitable diagnostics which are not currently available for a wide range of models. In this study therefore, the term 'cloud feedback' will (unless stated otherwise) refer to that component of the feedback measured by the change in the cloud radiative forcing.

Table 2 presents statistics based on those of Boer and Yu (2003), but applied to inter-model differences in global mean feedback components Λ_i taken across the CFMIP ensemble, rather than across the spatial structure of the feedback components in a single model. For each component Λ_i we show the ensemble mean $\bar{\Lambda}_i$, the ensemble sample standard deviation σ_{Λ_i}, and V_i, the fractional contribution of the global mean feedback component Λ_i to σ_Λ^2 (the ensemble variance of Λ), defined as

$$V_i = \sum_{k=1,n} \frac{\Lambda_k^+ \Lambda_{ik}^+}{\sigma_\Lambda^2 (n-1)} \tag{11}$$

where (Λ_{ik}^+, $k = 1,n$) are the deviations of the individual models from the ensemble mean. The contributions sum to unity, but may take negative as well as positive values. The table also shows correlations of Λ_i with Λ.

Figure 2a shows Λ for the CFMIP ensemble members, while Fig. 2b shows the cloud feedback components Λ_{LC}, Λ_{SC} and Λ_C. The models are presented in the order of increasing Λ so as to highlight any relationship between Λ and its components. The visual impression is that the inter-model differences in the shortwave cloud feedback Λ_{SC} make the largest contribution to those in Λ via Λ_C, and this is confirmed by Table 2 which shows that Λ_{SC} contributes 59% of σ_Λ^2 (out of a total of 66% from Λ_C) and that both have a statistically significant correlation with Λ. Λ_{SA} contributes a smaller amount (35%) and is also well correlated with Λ (Table 2; Fig. 2c).

Figure 2d and e shows Λ and the cloud feedback components from a subset of members of the QUMP ensemble, selected to span the ranges of the various globally meaned feedback terms. The visual impression is that (as with CFMIP) inter-model differences in Λ are mainly due to Λ_C, and statistics from the full 128 member ensemble in Table 3 confirm this (Λ_C contributes 85% to σ_Λ^2 and is well correlated with Λ). However, unlike the CFMIP ensemble, differences in Λ_{LC} contribute the larger part of this (57%). The clear-sky components (Fig. 2f, Table 3) make a smaller contribution to σ_Λ^2 than is the case in CFMIP, mostly from Λ_{LA}.

Comparing Table. 2 and 3 shows that the differences in the ensemble mean values for the globally meaned feedback components Λ_{SC}, Λ_{LA} and Λ_{LC} are all less than 0.1 Wm^{-2} K^{-1}. These differences are not significant at the 10% confidence level. However the ensemble mean value of Λ_{SA} is approximately 0.2 Wm^{-2} K^{-1} larger in CFMIP than QUMP. This difference is statistically significant, and is large enough to explain the difference in ensemble mean values of Λ in itself. The CFMIP standard deviations are larger than their QUMP equivalents (in a statistically significant sense) for all of the feedback components except Λ_{LC}. The Λ_{LC} standard deviation in QUMP is larger than in CFMIP, but not significantly. The statistical tests applied here are based on an assumption that the members of each ensemble are independent. This assumption will tend to exaggerate the statistical significance of any differences by an amount that is difficult to quantify, so the more marginal differences should be treated with a degree of caution.

To quantify the extent to which the QUMP ensemble spans the range of the CFMIP ensemble, we express the size of the overlap in the ranges of the two ensembles as a percentage of the spread in the CFMIP ensemble. Table 4 shows that QUMP spans 80% or more of the CFMIP range in the global mean values of Λ_{SC}, Λ_{LA} and Λ_{LC}, but only 25% of the range in Λ_{SA}. The net effect of these overlaps is that the QUMP ensemble spans more than 90% of the CFMIP range in globally meaned Λ. Conversely, CFMIP spans approximately 90% or more of the QUMP range in Λ_{SC} and Λ_{LA}, but less of the QUMP range in Λ_{SA} and Λ_{LC}.

4 Classification of local cloud feedbacks

4.1 Local feedback terms and feedback classes

In this section, a feedback classification scheme is introduced which identifies areas of the globe where qualitatively similar cloud feedback mechanisms are acting. Following Boer and Yu (2003), a local feedback parameter can be defined simply as the local contribution to the global feedback parameter by re-writing the first part of Eq. 3 in terms of global means of spatially varying local values Λ_l, f_l, etc.

$$\Lambda = \langle \Lambda_l \rangle = (\langle R'_l \rangle - \langle f_l \rangle) / \langle T'_l \rangle \tag{12}$$

and so

$$\Lambda_l = (R'_l - f_l) / \langle T'_l \rangle \tag{13}$$

where angular brackets denote a global mean. Note that the denominator contains the globally averaged near air temperature change rather than the local value. In the following text, Λ, f, T', etc. will now be taken to mean local values, and global means will be denoted by angle brackets. In the local context, the equation

$$\Lambda = \Lambda_A + \Lambda_C = \Lambda_{SA} + \Lambda_{LA} + \Lambda_{SC} + \Lambda_{LC} \tag{14}$$

still applies.

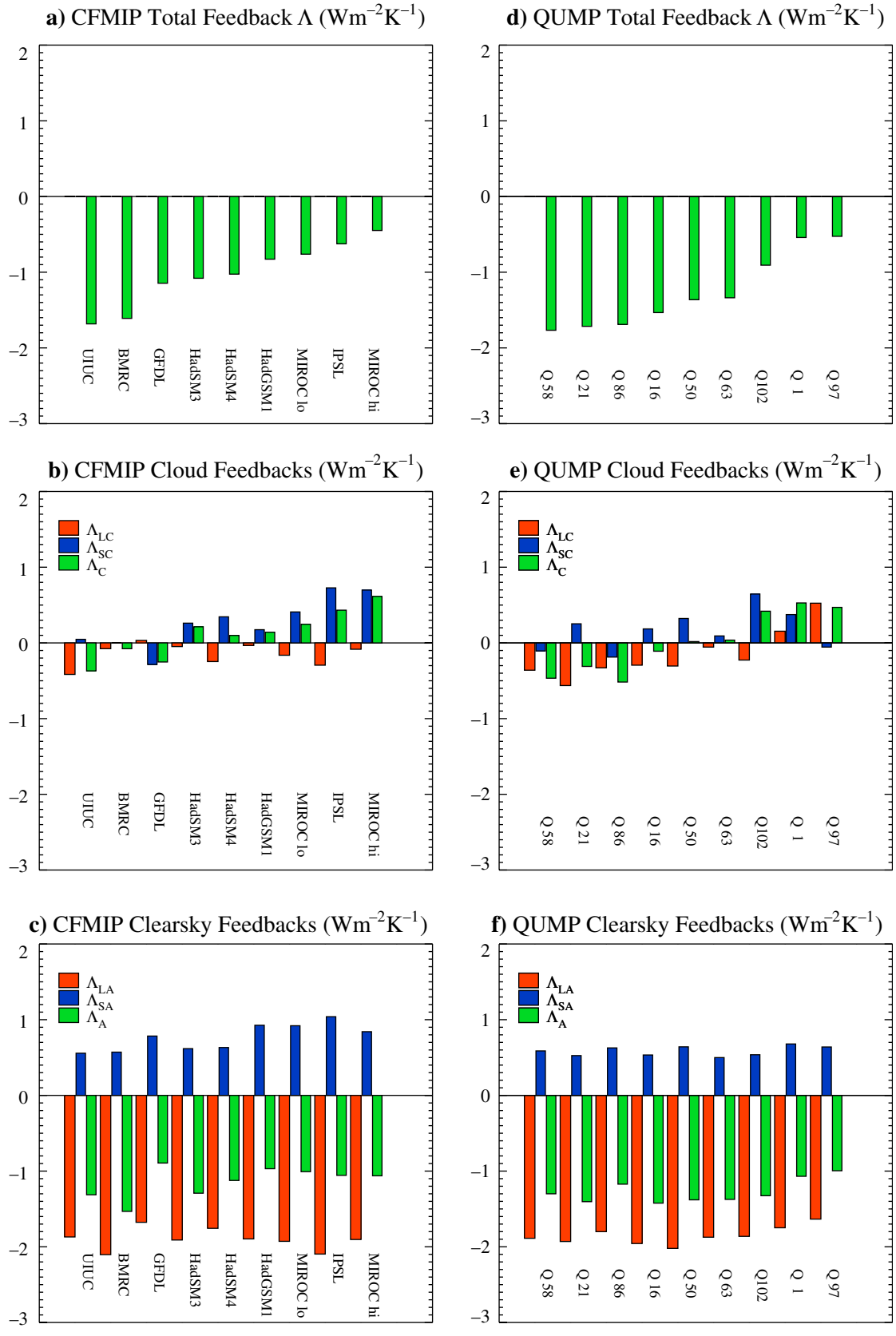

Fig. 2 Global mean feedbacks: **a** CFMIP values of Λ (Wm^{-2} K^{-1}); **b** CFMIP cloud feedbacks Λ_{LC}, Λ_{SC} and Λ_{C} (Wm^{-2} K^{-1}); and **c** CFMIP clear-sky feedbacks Λ_{LA}, Λ_{SA} and Λ_{A} (Wm^{-2} K^{-2}). The parts (**d**–**f**) show the same quantities from nine selected QUMP ensemble members, selected to include the maximum and minimum values of T', Λ_{LA}, Λ_{SA}, Λ_{A}, Λ_{LC}, Λ_{SC} and Λ_{C} from the full 128 member ensemble. The selected models are used in the figures only

Table 2 Ensemble statistics of the global mean feedback components in the CFMIP models

	Λ	Λ_C	Λ_{LC}	Λ_{SC}	Λ_A	Λ_{LA}	Λ_{SA}
$\overline{\Lambda_i}$	−1.02	0.12	−0.15	0.26	−1.14	−1.90	0.77
σ_{Λ_i}	0.42	0.31	0.15	0.33	0.20	0.14	0.18
V_i	**1.00**	**0.66**	0.08	**0.59**	**0.34**	−0.01	**0.35**
r_Λ	**1.00**	**0.88**	0.22	**0.75**	0.69	−0.03	**0.82**

The ensemble mean and standard deviation is shown for each component. V_i is the contribution of the component to σ^2_Λ, the variance of Λ. Components which contribute more than 20% to σ^2_Λ are in bold. r_Λ is the correlation of the feedback component with Λ. For nine ensemble members the significance threshold for the correlations are 0.67 for the 10% confidence level (with a two-tailed test). Correlations greater than or equal to 0.7 are highlighted in bold

Table 4 Overlap between the ranges of values in global mean feedback components taken by the CFMIP and QUMP ensembles

	Λ	Λ_A	Λ_C	Λ_{SA}	Λ_{SC}	Λ_{LA}	Λ_{LC}
QUMP (% CFMIP)	93	66	91	25	82	80	100
CFMIP (% QUMP)	93	100	85	67	100	89	41

The size of the overlap between the two ensembles' ranges expressed as a percentage of the spread in the CFMIP ensemble, which is a measure of the extent to which the QUMP ensemble spans the range of the CFMIP ensemble in each feedback component. The overlap is also expressed as a percentage of the spread in the QUMP ensemble to give an indication of the extent to which the CFMIP ensemble spans the QUMP ranges

The classification we adopt is illustrated in the schematic in Fig. 3 and is based on the idea of grouping local cloud feedbacks together on the basis of the relationship between their longwave and shortwave components. For example, consider an idealised situation where, in a CO_2 doubling experiment, the amount of low level cloud in a particular region decreases while other variables (such as the clear-sky albedo) remain unchanged. This would lead to a decrease in the magnitude of the local shortwave cloud radiative forcing which, (the shortwave cloud forcing being negative) would constitute a positive local cloud feedback with a corresponding positive value of Λ_{SC}. This decrease in low level cloudiness would have little effect on the longwave CRF, and so this case would fall in feedback class $A(S^+L^N$: orange) on the schematic in Fig. 3, where the S^+L^N indicates a positive shortwave cloud feedback coinciding with a (relatively) neutral longwave cloud feedback. Conversely, if low level cloudiness was to increase, we would expect this cloud feedback to fall into class $E(S^-L^N$: *dark blue*). If low level cloud optical thickness was to change, this cloud feedback would also be expected to fall into cloud feedback class $A(S^+L^N$: orange) or $E(S^-L^N$: dark blue), depending on the sign of the change.

Considering another case, observations show that the longwave and shortwave components of the cloud radiative forcing (positive and negative in sign, respectively) tend to cancel in parts of the tropics where deep convection produces optically thick clouds with tops near to the tropopause. A reduction of the amount of this type of cloud might lead to reductions in the magnitudes of the local longwave and shortwave cloud forcing of comparable size, and this case would then fall into class $B(S^+L^-$: yellow), while an equal and opposite increase would fall into class $F(S^-L^+$: purple). Similarly, decreases in the amount of thin or subvisual cirrus would be expected to fall into class $C(S^NL^-$: green), while increases would be expected to fall into class $G(S^NL^+$: dark red). For completeness, classes $H(S^+L^+$: red) and $D(S^-L^-$: light blue) are also included. Cloud feedbacks associated with changes in clouds at different levels in combination could fall into these classes: for instance increases in high thin cirrus coinciding with reductions in low level clouds could fall into class $H(S^+L^+$: red), while the reverse situation could fall into class $D(S^-L^-$: light blue).

Areas where the local values of Λ_{SA} are comparable in magnitude to those of Λ_{SC} are particularly difficult to interpret as it is not clear whether values of Λ_{SC} in these areas are indicative of a true cloud feedback, or a consequence of using the cloud radiative concept for the separation of cloud and atmospheric feedbacks (see below). For this reason, in much of the analysis that follows we remove areas where $\Lambda_{SA} > 1$ Wm^{-2} K^{-1} from the eight cloud feedback classes and place them into a cloud feedback class of their own, namely class $I(SA^+)$.

Although the examples described above are simplistic, we argue that they can be used as a basis for interpreting the various cloud feedback classes in terms of the primary cloud response. By this we mean the cloud type which is showing the largest changes in amount or cloud optical depth within the area covered by the class. We argue that classes $A(S^+L^N$: orange) and $E(S^-L^N$: dark blue) can be interpreted as areas where the primary cloud response is due to changes in the amount or optical thickness of low-top clouds. Similarly we argue that classes $B(S^+L^-$: yellow) and $F(S^-L^+$: purple) indicate a primary cloud response due to changes in amount, optical thickness or height of optically thick high-top clouds, and similarly associate classes $C(S^NL^-$: green) and $G(S^NL^+$: dark red) with a primary response in optically thin high-top clouds. Classes $D(S^-L^-$: light blue) and $H(S^+L^+$: red) may be interpreted primarily in terms of coincident changes in low and optically thin high-top clouds. Given that the ISCCP simulator diagnostics are available in both ensembles we are able to test these interpretations using the actual cloud responses in the models.

Table 3 As Table 2 but for the 128 QUMP model versions

	Λ	Λ_C	Λ_{LC}	Λ_{SC}	Λ_A	Λ_{LA}	Λ_{SA}
$\overline{\Lambda_i}$	−1.18	0.09	−0.10	0.19	−1.27	−1.85	0.59
σ_{Λ_i}	0.25	0.22	0.19	0.17	0.07	0.05	0.04
V_i	**1.00**	**0.85**	**0.57**	**0.29**	0.15	0.12	0.03
r_Λ	**1.00**	**0.97**	**0.77**	0.43	0.56	0.62	0.19

The significance threshold for correlations at the 10% confidence level is 0.17 for an ensemble of this size. Correlations greater than or equal to 0.7 are highlighted in bold

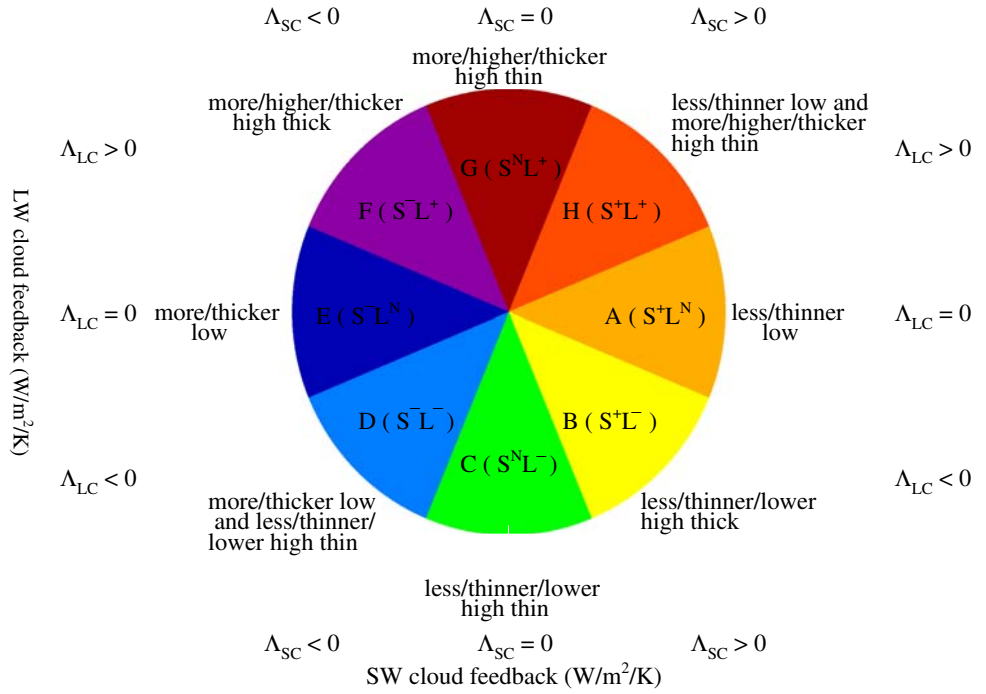

Fig. 3 Cloud feedback classification. Local values of the shortwave and longwave cloud feedback terms Λ_{SC} and Λ_{LC} are classified in a two-dimensional space with Λ_{SC} in the X dimension and Λ_{LC} in the Y dimension. The space is divided into eight sectors which sweep out equal angles of 45° from the origin. Classes $A(S^+L^N$: orange) and $E(S^-L^N$: dark blue) contain values which are in the two sectors lying along the line $\Lambda_{LC}=0$. In these classes, $|\Lambda_{LC}|$ is always $\leq \tan(22.5°)|\Lambda_{SC}|$. As $\tan(22.5°)=0.42$ we describe these classes as containing cloud feedbacks where Λ_{LC} is 'relatively neutral' compared to Λ_{SC}. Classes $C(S^NL^-$: green) and $G(S^NL^+$: dark red) contain values which are in the two sectors which include the line $\Lambda_{SC}=0$, and contain values of Λ_{SC} which are relatively neutral compared with Λ_{LC}. Classes $D(S^-L^-$: light blue) and $H(S^+L^+$: red) contain values which are in the two sectors which include the line $\Lambda_{SC}=\Lambda_{LC}$. Here the values of Λ_{SC} and Λ_{LC} are of the same sign, and as $|\Lambda_{LC}| \geq 0.42|\Lambda_{SC}|$ and $|\Lambda_{SC}| \geq 0.42|\Lambda_{LC}|$ we describe the values of Λ_{SC} and Λ_{LC} in these classes as having 'comparable magnitude'. Classes $B(S^+L^-$: yellow) and $F(S^-L^+$: purple) comprise the two sectors on the line $\Lambda_{SC}=-\Lambda_{LC}$ and contain values where Λ_{LC} and Λ_{SC} are of comparable magnitude but opposite sign. The cloud types which show the largest responses in each class are also indicated (see Sect. 4.2)

4.2 Interpretation of cloud feedback classes

We analyse changes in the frequency of occurrence (or amount) of nine ISCCP cloud types (see Fig. 2 of Rossow and Schiffer 1999) diagnosed by the ISCCP simulator in the models. The total cloud amount is broken down into three bins of cloud top pressure P_c: high-top ($P_c<440$ mb) mid-level-top (440 mb $\leq P_c<680$ mb) and low-top clouds ($P_c \geq 680$ mb). These are combined with three bins of cloud optical thickness τ: optically thin ($0.3 \leq \tau<3.6$), medium ($3.6 \leq \tau<23$) and optically thick ($\tau \geq 23$) to give the nine cloud types. We define the cloud response of cloud type i within a given class j to be $[C_i]_j=[c'_i]_j/\langle T' \rangle$ where c'_i is the local amount of that cloud type and where the square brackets denote a spatial mean taken within the class. Ensemble mean values of this quantity are shown in Table 5 to give an indication of the relative sizes of the responses in the different cloud types in each cloud feedback class. Figure 4a–i shows the spatial distributions of the nine cloud feedback classes for the CFMIP models, ordered in terms of increasing globally averaged Λ.

Class $A(S^+L^N$: orange) covers areas where a positive shortwave cloud feedback coincides with a relatively neutral longwave cloud feedback, which we argue is indicative of a cloud feedback primarily due to an decrease in the amount or optical thickness of low clouds. Table 7 shows the ensemble average class areas as global means and also broken down into low, mid and high latitude regions. This shows that class $A(S^+L^N$: orange) covers about a quarter of the globe in the ensemble mean, occurring mostly in the low and mid-latitude regions. Figure 4 shows that it is commonly seen off the western coasts of subtropical continental areas and in mid-latitude storm track areas, and Table 5 shows that reductions in the amount of low-top cloud (primarily of medium optical thickness) are over twice the size of the other cloud responses in this class, which is consistent with our interpretation given above. A secondary contribution from decreases in the amount of mid-level-top clouds is also noted.

Conversely, class $E(S^-L^N$: dark blue) indicates a negative shortwave cloud feedback with a relatively neutral longwave cloud feedback, which we argue is indicative of cloud feedback primarily due to an increase in the amount or optical thickness of low clouds. It covers less than a tenth of the globe in the ensemble mean, and this is spread roughly evenly across the three

Table 5 CFMIP ensemble mean cloud responses $\overline{[C_i]}$ within each feedback class (%/K)

	GL	$A(S^+L^N)$	$BS^+L^-)$	$C(S^NL^-)$	$D(S^-L^-)$	$E(S^-L^N)$	$F(S^-L^+)$	$G(S^NL^+)$	$H(S^+L^+)$	$I(\mathrm{SA}^+)$
Total	**-0.25**	**-0.95**	**-0.86**	−0.22	0.09	**0.27**	**0.56**	**0.50**	**-0.30**	**0.82**
Low	−0.21	**-0.81**	−0.20	0.19	0.23	0.09	−0.16	**-0.26**	**-0.90**	**0.60**
Mid	−0.22	**-0.38**	−0.24	−0.05	−0.07	−0.16	−0.15	−0.08	**-0.28**	−0.16
High	0.18	0.25	**-0.42**	**-0.36**	−0.07	**0.34**	**0.86**	**0.83**	**0.87**	**0.38**
Low/thin	−0.15	−0.11	−0.02	−0.01	−0.03	−0.24	−0.14	−0.05	−0.19	**-0.45**
Low/intermediate	−0.14	**-0.47**	−0.12	0.05	0.05	−0.15	−0.20	−0.16	**-0.57**	**0.54**
Low/thick	0.07	−0.23	−0.06	0.14	0.21	**0.48**	0.19	−0.05	−0.13	**0.52**
Mid/thin	−0.09	−0.05	−0.03	−0.01	−0.02	−0.09	−0.08	−0.07	−0.06	**-0.31**
Mid/intermediate	−0.12	−0.14	−0.09	−0.04	−0.05	−0.16	−0.12	0.00	−0.13	−0.15
Mid/thick	−0.02	−0.19	−0.11	−0.00	0.01	0.10	0.06	−0.01	−0.09	**0.30**
High/thin	0.04	0.18	−0.09	−0.17	−0.11	−0.06	0.19	**0.29**	**0.40**	−0.11
High/intermediate	0.03	0.06	−0.16	−0.15	−0.07	0.04	**0.26**	**0.31**	0.23	0.07
High/thick	0.12	0.01	−0.17	−0.04	0.10	**0.36**	**0.41**	0.23	0.24	**0.42**

Cloud responses of 0.25%/K or greater are highlighted in bold

latitude regions (Table 7). Figure 4 shows that it is commonly seen in the trade cumulus regions (e.g. to the west of south America), but also occasionally in the tropical west Pacific. It is also seen in oceanic regions at high latitudes, adjacent to the class $I(\mathrm{SA}^+)$ areas marked by the black contours, and stretches from there into the mid-latitude region in some of the models. Table 5 shows that the largest component of the cloud response in this class is an increase in the amount of optically thick low-top cloud, which along with the reductions in the amount of optically thin and intermediate optical thickness low-top clouds indicates an increase in the average low-cloud optical thickness in this class, consistent with our interpretation above. A secondary contribution from increases in the amount of optically thick high-top cloud is also noted.

Class $B(S^+L^-$: yellow) covers areas where a positive shortwave cloud feedback coincides with a negative longwave cloud feedback of comparable magnitude, which we argue is indicative of a cloud feedback primarily due to decreases in the amount, optical thickness or height of reflective high-top clouds. According to Table 7, class $B(S^+L^-$: yellow) covers about a quarter of the globe in the ensemble mean, and occurs mostly at low latitudes. Figure 4 shows that it is commonly seen in the South Pacific Convergence Zone (SPCZ), on the western sides of the subtropical Pacific, Indian and Atlantic oceans, and also occasionally in the mid-latitude storm track regions. Table 5 shows that the largest cloud responses in this class are due to reductions in the amounts of optically thick and intermediate optical thickness high-top clouds, which supports our interpretation above. Smaller contributions from reductions in the amounts of optically thin high-top, mid-level and low-top clouds are also seen.

Conversely, class $F(S^-L^+$: purple) covers areas where a negative shortwave cloud feedback coincides with a positive longwave cloud feedback of comparable magnitude [the opposite of class $B(S^+L^-$: yellow)]. We argue that this class is indicative of a cloud feedback primarily due to increases in the amount, optical thickness or height of reflective high-top clouds. This class covers less than a tenth of the globe in the ensemble mean, and occurs mostly at low and high latitudes. Figure 4 shows that it is commonly seen across the tropical Pacific, along or a few degrees north of the equator, and also occasionally over the Indian Ocean and parts of Africa. Table 5 shows that the increases in the amounts of optically thick and intermediate optical thickness high-top cloud are the largest in this class, and that combined they are three times larger than any of the other responses. Again this supports our interpretation above. A secondary contribution from increases in the amount of optically thin high-top cloud in this class is also seen.

Class $C(S^NL^-$: green) covers areas where a negative longwave cloud feedback coincides with a relatively neutral shortwave cloud feedback, which we argue is indicative of a cloud feedback primarily due to a decrease in the amount, optical thickness or height of the optically thinner high-top clouds. Table 7 shows that this class covers less than 5% of the globe in the ensemble mean, mostly in the tropics, while Fig. 4 shows that this class commonly appears in small amounts between areas of class $B(S^+L^-$: yellow) and $D(S^-L^-$: light blue). The UIUC model (which has the lowest climate sensitivity) is distinguished by having larger areas of class $C(S^NL^-$: green) than the other models, across much of the tropics. According to Table 5, the largest cloud responses in this class are reductions in the amount of optically thin and intermediate optical thickness clouds with high tops, which supports our interpretation. A secondary contribution is made by increases in the amount and optical thickness of the low-top clouds.

Conversely, class $G(S^NL^+$: dark red) covers areas where a positive longwave cloud feedback coincides with a relatively neutral shortwave cloud feedback, which we argue is indicative of a cloud feedback primarily due to an increase in the amount, optical thickness or height of optically thinner high-top clouds. Table 7 shows that this class covers just over 5% of the globe in the ensemble mean, split evenly between the tropics and the high latitudes, while Fig. 4 shows that this class com-

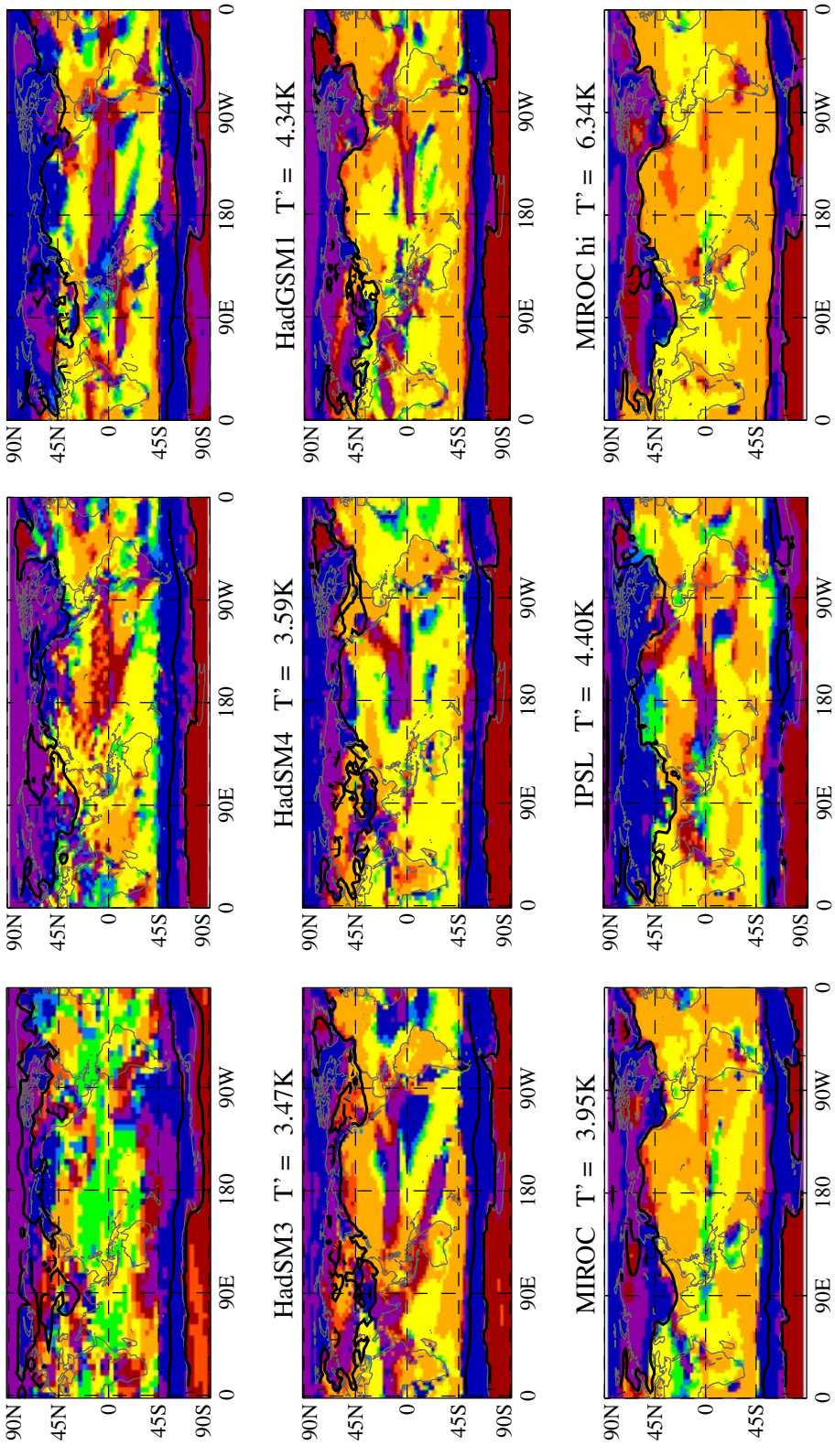

Fig. 4 Spatial distribution of CFMIP cloud feedback classes. The areas of the globe covered by each of the eight cloud feedback classes are shown by plotting the assigned colour of each class at locations where the local feedback components fall into that class. For example, the *orange areas* show where class $A(S^+L^N$: orange) behaviour is present—i.e. where Λ_{SC} is positive but Λ_{LC} is relatively neutral. Class $I(SA^+)$ areas are enclosed by *black contours*, but colour coded consistently with whichever of the eight original cloud feedback classes that they were initially placed in

Table 6 As Table 5 but for the QUMP ensemble

	GL	$A(S^+L^N)$	$B(S^+L^-)$	$C(S^NL^-)$	$D(S^-L^-)$	$E(S^-L^N)$	$F(S^-L^+)$	$G(S^NL^+)$	$H(S^+L^+)$	$I(\mathrm{SA}^+)$
Total	−0.19	**−1.02**	**−0.95**	**−0.27**	0.14	**0.61**	**0.87**	**0.53**	−0.18	**0.58**
Low	−0.14	**−0.84**	−0.09	**0.40**	**0.55**	**0.49**	−0.18	**−0.34**	**−0.85**	**0.30**
Mid	**−0.27**	**−0.35**	−0.23	−0.15	−0.18	**−0.26**	−0.16	−0.13	−0.23	**−0.39**
High	0.23	0.17	**−0.63**	**−0.51**	−0.23	**0.38**	**1.21**	**0.99**	**0.90**	**0.66**
Low/thin	−0.20	−0.16	−0.07	−0.03	−0.04	−0.21	−0.13	−0.13	−0.20	**−0.58**
Low/intermediate	−0.05	**−0.50**	−0.09	0.16	**0.29**	**0.29**	−0.08	−0.16	**−0.52**	**0.46**
Low/thick	0.11	−0.18	0.07	**0.26**	**0.30**	**0.41**	0.03	−0.05	−0.14	**0.42**
Mid/thin	−0.11	−0.09	−0.05	−0.04	−0.06	−0.15	−0.10	−0.12	−0.07	−0.22
Mid/intermediate	−0.12	−0.10	−0.09	−0.06	−0.06	−0.10	−0.08	0.01	−0.09	**−0.32**
Mid/thick	−0.05	−0.16	−0.10	−0.05	−0.06	−0.01	0.02	−0.02	−0.07	0.15
High/thin	0.07	0.18	−0.21	**−0.27**	−0.19	−0.05	**0.46**	**0.45**	**0.44**	0.00
High/intermediate	0.07	0.06	−0.07	−0.04	0.01	0.09	0.25	**0.29**	0.17	0.11
High/thick	0.09	−0.07	**−0.35**	−0.20	−0.05	**0.34**	**0.51**	**0.26**	**0.29**	**0.55**

Table 7 CFMIP class area statistics by region

Class area	$A(S^+L^N)$	$B(S^+L^-)$	$C(S^NL^-)$	$D(S^-L^-)$	$E(S^-L^N)$	$F(S^-L^+)$	$G(S^NL^+)$	$H(S^+L^+)$	$I(\mathrm{SA}^+)$
Global	0.25	0.25	0.05	0.03	0.07	0.08	0.06	0.04	0.16
30°N–30°S	0.13	0.18	0.04	0.02	0.03	0.04	0.03	0.03	0.00
30°–50° N/S	0.10	0.07	0.01	0.01	0.02	0.01	0.01	0.01	0.03
50°–90° N/S	0.01	0.00	0.00	0.00	0.03	0.02	0.03	0.01	0.13

The ensemble means of the fractional area covered by each class are shown for the whole globe, and for low latitude (30°N–30°S), mid-latitude (30°N–50°N, 30°S–50°S) and high latitude (50°N–90°N, 50°S–90°S) regions. The fractions of the globe covered by each region of latitude are 50, 26 and 24%, respectively

monly appears near areas of class $F(S^-L^+$: purple) in the tropics, and over Antarctica. Table 5 shows increases in high-top cloud amount are more than twice the size of the mid- and low-top cloud responses, and that this is mainly due to increases in the amount of intermediate/thin optical thickness high-top cloud, which is consistent with our interpretation. Reductions in the amount of low-top cloud play a secondary role.

Class $H(S^+L^+$: red) covers areas where a positive longwave cloud feedback coincides with a positive shortwave cloud feedback of comparable magnitude. It is not clear that changes in a single cloud type could be responsible for a cloud feedback of this type. However, as class $H(S^+L^+$: red) sits between classes $A(S^+L^N$: orange) and $G(S^NL^+$: dark red) in the cloud feedback phase space, we argue that this class is indicative of a combination of the cloud responses seen in these two classes—i.e. reductions in low-top cloud amount and/or optical thickness combined with increases in the amount, optical thickness or height of optically thin high-top clouds. Table 7 shows that this class covers just under 5% of the globe in the ensemble mean, half in the tropics and a quarter in each of the mid-latitude and high latitude regions, while Fig. 4 shows it to commonly appear between class $A(S^+L^N$: orange) and $H(S^+L^+$: red) areas. According to Table 5, the largest cloud responses are due to an increase in the amount of (primarily optically thin) high-top cloud and a reduction in low-top cloud amount (primarily of intermediate optical thickness), which supports our interpretation.

Conversely, class $D(S^-L^-$: light blue) covers areas where a negative longwave cloud feedback coincides with a negative shortwave cloud feedback of comparable magnitude. As with class $H(S^+L^+$: red) above, we argue that this is indicative a cloud feedback due to a combination of the cloud feedback behaviour from the adjacent classes [primarily reductions in amount or optical thickness of the optically thinner high-top clouds as in class $C(S^NL^-$: green) and increases in low-top cloud amount and/or optical thickness as in class $E(S^-L^N$: dark blue)]. Table 7 shows that this class covers less than 5% of the globe in the ensemble mean, two-thirds in the tropics and one-third in the mid-latitude region, and Fig. 4 shows that it appears mainly between areas of class $C(S^NL^-$: green) and $E(S^-L^N$: dark blue). Table 5 shows the largest cloud responses in this class indicate an increase in the amount and optical thickness of low-top clouds, and a reduction in the optically thin and intermediate optical thickness high-top cloud amounts, which is consistent with the interpretation above. There is also a smaller increase in the optically thick high-top cloud, as in class $E(S^-L^N$: dark blue).

Finally we turn to class $I(\mathrm{SA}^+)$, which we have defined to contain all cloud feedbacks which coincide with a positive clear-sky shortwave feedback where $\Lambda_{\mathrm{SA}} > 1\ \mathrm{Wm}^{-2}\ \mathrm{K}^{-1}$. This threshold was chosen experimentally to select areas (mostly at high latitudes) where changes in sea-ice and snow mean that they reflect less sunlight in the warmer climate. This class covers about 15% of the globe in the ensemble mean, the vast majority of which is in the high latitude region, and Fig. 4 shows that in these areas the CFMIP models show mostly class $E(S^-L^N$: dark blue) and $F(S^-L^+$: purple)-like behaviour, indicating negative values of

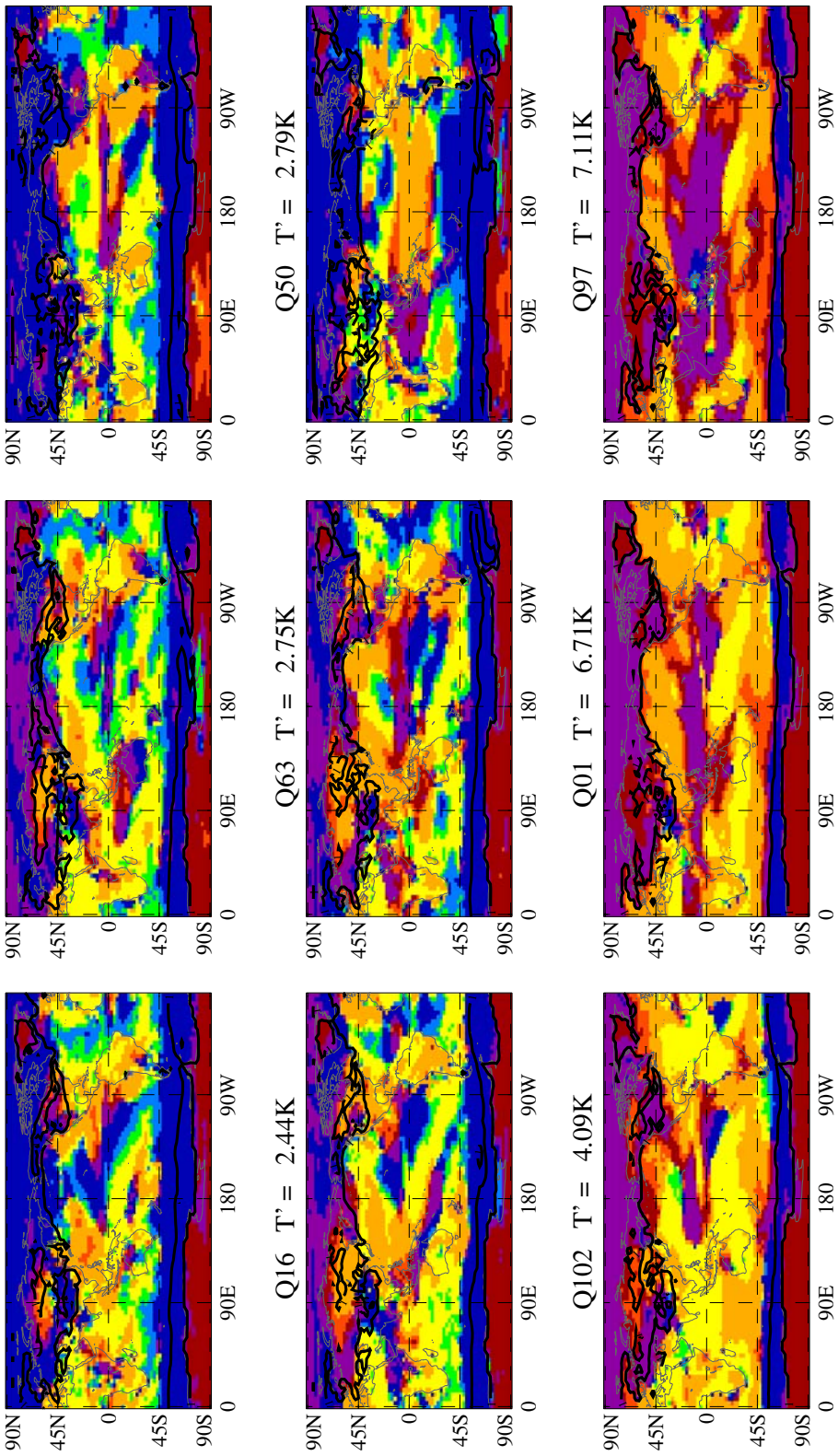

Fig. 5 As Fig. 4 but for the selected QUMP models

Λ_{SC}. However, this alone is not evidence that the clouds in these areas are necessarily showing increases in their amount or cloud optical thickness. This is because, even with no change in cloud properties, a reduction in clear-sky albedo can give a negative value of Λ_{SC}. This effect can be illustrated using a simple idealised model, where we consider the effects of surface reflection and scattering of shortwave radiation by clouds only. For instance, if:

$$R_S = -S[\alpha_c + (1-\alpha_c)\alpha_s] \quad (15)$$

and

$$\tilde{R}_S = -S\alpha_s \quad (16)$$

where S is the incoming solar radiation, α_s is the surface albedo and α_c represents the cloud albedo, then the shortwave cloud forcing

$$\begin{aligned} R_{SC} &= R_S - \tilde{R}_S = -S[\alpha_c + (1-\alpha_c)\alpha_s - \alpha_s] \\ &= -S\alpha_c(1-\alpha_s) \end{aligned}$$

In the case where α_s decreases while α_c remains unchanged, the magnitude of the shortwave cloud radiative forcing will increase (which will give a positive value of Λ_{SA} and a negative value of Λ_{SC}). This is an example of the cloud masking effect discussed by Colman (2003) and Soden et al. (2004). It is because of this difficulty in interpreting the meaning of Λ_{SC} in areas of clear-sky shortwave feedback that we place areas where $\Lambda_{SA} > 1$ Wm^{-2} K^{-1} into a separate class. We can however interpret the cloud feedbacks in this class using our understanding of the cloud feedbacks in other classes. Table 5 shows that the largest ensemble mean cloud responses in this class are for the low-top clouds, with increases in the amounts of optically thick and intermediate optical thickness low-top cloud and reductions in the amount of optically thin low-top clouds. These changes indicate that the primary cause of the cloud feedback in this class is an increase in the amount and optical thickness of low-top clouds which is consistent with the cloud response in class $E(S^-L^N$: dark blue) where a truly negative shortwave cloud feedback is operating. Increases in the amount of high-top optically thick cloud play a secondary role, which is also consistent with class $E(S^-L^N$: dark blue). From this we conclude that the ensemble mean cloud responses in class $I(SA^+)$ are fully consistent with the presence of a truly negative shortwave cloud feedback.

Turning now to the QUMP ensemble, Fig. 5 shows the spatial distributions of the nine cloud feedback classes for the nine selected model QUMP versions, while Table. 8 and 6 show the ensemble mean class areas and cloud responses for full 128 member QUMP ensemble. Comparing Table. 8 and 7 we see that class $F(S^-L^+$: purple) covers slightly larger areas of the globe in QUMP than in CFMIP, the differences coming mainly from the tropics, and this difference is particularly evident in the higher sensitivity models in Fig. 5. Statistical tests show that these differences are significant at the 10% confidence level. The areas of the other classes are statistically indistinguishable in the two ensembles, both globally and in the various latitude bands. A comparison of Table. 6 and 5 shows that while some differences are seen in the cloud responses in various classes, our interpretations of the feedback classes in terms of their primary cloud responses are also supported by the QUMP ensemble in each case. Some differences between the two ensembles are seen in the secondary cloud responses.

5 Cloud feedback class contributions

In this section we decompose the global cloud feedback components from the CFMIP and QUMP ensembles into contributions from each of the feedback classes so as to establish which feedback classes make the largest contribution to the variance in Λ in each ensemble. We also consider the extent to which the range of values taken in each feedback class overlap in the two ensembles. Table 9 shows the ensemble mean contributions $\overline{\langle\Lambda_i\rangle}_j$ of each feedback class j to each of the ensemble global mean cloud feedback components $\overline{\langle\Lambda_i\rangle}$ in the CFMIP ensemble, as well as the ensemble mean values within each class ($\overline{[\Lambda_i]}_j$). Also shown are the associated standard deviations, contributions to σ_Λ^2 and correlations with $\langle\Lambda\rangle$. If we sum the variance contributions to σ_Λ^2 from Λ_C in the cloud feedback classes in which low clouds constitute the primary cloud responses [classes $A(S^+L^N$: orange), $E(S^-L^N$: dark blue) and $I(SA^+)$] we find that these contribute 39% to σ_Λ^2, which is 59% of the 66% contributed by the global values of Λ_C. In each of these classes the contribution to the variance from Λ_{SC} is several times the size of the contribution from Λ_{LC}, which is consistent with the inter-model differences in the cloud feedbacks in these classes being largely due to those in low-top clouds. The cloud feedback classes which have high cloud changes as their primary causes [classes $B(S^+L^-$: yellow), $C(S^NL^-$: green), $F(S^-L^+$: purple) and $G(S^NL^+$: dark red)] contribute 22% to σ_Λ^2, which is 33% of the contribution from Λ_C. Classes $D(S^-L^-$: light blue) and $H(S^+L^+$: red) contribute 5% to σ_Λ^2.

Class $A(S^+L^N$: orange) is responsible for the largest contribution to σ_Λ^2, and more than half of the contribution from the cloud feedback in total, mostly because of its contribution from Λ_{SC}. Values of Λ_C and Λ_{SC} in this class are well correlated with $\langle\Lambda\rangle$, with correlation coefficients of 0.8 or more. The relationship between class $A(S^+L^N$: orange) and the total feedback is also apparent in Fig. 6 which shows the contributions from each of the feedback classes to global mean values of each of the cloud feedback components for the individual CFMIP models, presented in the order of increasing $\langle\Lambda\rangle$. The larger areas covered by class $A(S^+L^N$: orange) in the higher sensitivity CFMIP

Table 8 As Table 7 but for the QUMP ensemble

Class area	$A(S^+L^N)$	$B(S^+L^-)$	$C(S^NL^-)$	$D(S^-L^-)$	$E(S^-L^N)$	$F(S^-L^+)$	$G(S^NL^+)$	$H(S^+L^+)$	$I(SA^+)$
Global	0.23	0.23	0.04	0.03	0.10	0.10	0.06	0.04	0.16
30°N–30°S	0.11	0.15	0.03	0.02	0.05	0.07	0.03	0.02	0.00
30°–50° N/S	0.11	0.08	0.01	0.01	0.01	0.01	0.00	0.01	0.03
50°–90° N/S	0.02	0.00	0.00	0.00	0.04	0.02	0.03	0.01	0.13

Table 9 CFMIP cloud feedback component class statistics

	GL	$A(S^+L^N)$	$B(S^+L^-)$	$C(S^NL^-)$	$D(S^-L^-)$	$E(S^-L^N)$	$F(S^-L^+)$	$G(S^NL^+)$	$H(S^+L^+)$	$I(SA^+)$
Λ_{LC}										
$\overline{\langle\Lambda_i\rangle}$	−0.15	−0.06	−0.27	−0.07	−0.01	0.01	0.08	0.05	0.03	0.10
$\overline{[\Lambda_i]}$	−0.15	−0.24	−1.10	−0.94	−0.47	0.14	0.86	0.83	0.69	0.64
σ_{Λ_i}	0.15	0.03	0.11	0.13	0.01	0.01	0.06	0.03	0.02	0.03
V_i	0.08	−0.05	0.00	**0.20**	0.02	−0.01	−0.04	−0.06	0.00	0.02
r_Λ	0.22	**-0.83**	0.00	0.66	**0.80**	**-0.72**	−0.32	**-0.76**	0.03	0.26
Λ_{SC}										
$\overline{\langle\Lambda_i\rangle}$	0.26	0.42	0.34	0.01	−0.01	−0.11	−0.08	−0.00	0.04	−0.34
$\overline{[\Lambda_i]}$	0.26	1.70	1.35	0.10	−0.48	−1.22	−0.95	−0.05	0.79	−2.11
σ_{Λ_i}	0.33	0.21	0.13	0.01	0.01	0.08	0.06	0.00	0.02	0.10
V_i	**0.59**	**0.41**	0.10	−0.02	0.02	0.13	0.05	0.00	0.01	−0.10
r_Λ	**0.75**	**0.83**	0.32	−0.65	**0.74**	**0.72**	0.30	0.12	0.13	−0.43
Λ_C										
$\overline{\langle\Lambda_i\rangle}$	0.12	0.36	0.07	−0.07	−0.03	−0.09	−0.01	0.05	0.07	−0.23
$\overline{[\Lambda_i]}$	0.12	1.45	0.25	−0.84	−0.96	−1.07	−0.09	0.78	1.48	−1.47
σ_{Λ_i}	0.31	0.18	0.05	0.11	0.02	0.07	0.01	0.03	0.05	0.10
V_i	**0.66**	**0.35**	0.10	0.18	0.04	0.12	0.00	−0.06	0.01	−0.08
r_Λ	**0.88**	**0.80**	**0.86**	0.66	**0.77**	**0.70**	0.06	**-0.76**	0.08	−0.35

$\overline{\langle\Lambda_i\rangle_j}$ is the ensemble mean of the contribution $\langle\Lambda_i\rangle_j$ from the class j to the global feedback component $\langle\Lambda_i\rangle$. $\overline{[\Lambda_i]_j}$ is the ensemble mean of the feedback strength $[\Lambda_i]_j$ within class j. σ_{Λ_i} is the ensemble standard deviation of $\langle\Lambda_i\rangle_j$. V_i is the contribution of inter-model differences in $\langle\Lambda_i\rangle_j$ to σ^2_Λ. r_Λ is the correlation of $\langle\Lambda_i\rangle_j$ with $\langle\Lambda\rangle$. Values of V_i over 20% are in bold. Values of r_Λ greater than or equal to 0.7 are also in bold

models constitute a striking feature of Fig. 4, and the area covered by this class is strongly anti-correlated with $\langle\Lambda\rangle$ (r_Λ = -0.85). To quantify the contribution of inter-model differences in class area to σ^2_Λ, we recalculate the values of V_i in Table 9, but substitute $\overline{[\Lambda_i]}_j a_j$ for $\langle\Lambda_i\rangle_j$ so as to suppress the contribution of inter-model differences feedback strengths $[\Lambda_i]$ within each class to σ^2_Λ. We find that differences in class $A(S^+L^N$: orange) area contribute 38% out of the 41% of σ^2_Λ that is due to differences in Λ_{SC}. Figure 7 shows the global mean cloud responses in the low-, mid- and high-top cloud amount, and the contributions to these values from each of the feedback classes, ordered according to the model's values of $\langle\Lambda\rangle$. The higher sensitivity models tend to show larger reductions in the global mean low-top cloud amount, and much of this is due to the contribution from class $A(S^+L^N$: orange). The low-top/intermediate cloud response in this class is strongly anti-correlated with $\langle\Lambda\rangle$ ($r_\Lambda = -0.91$), as is the total low-top cloud response ($r_\Lambda = -0.84$). There is also a quite strong ($r_\Lambda = -0.79$) anti-correlation with the (smaller) intermediate optical thickness mid-level cloud reductions, which is consistent with these making a secondary contribution within this class.

The next largest contribution to the variance in σ^2_Λ from the low cloud feedback classes comes from class $E(S^-L^N$: dark blue), in which Λ_C contributes 12% to σ^2_Λ, again mostly from differences in Λ_{SC}. Figure 6 shows a tendency for stronger negative values of Λ_{SC} and Λ_C for the lower sensitivity models in this class, and the table shows that these values correlate with $\langle\Lambda\rangle$ with coefficients of 0.7 or more. We find that approximately three quarters of this contribution is due to differences in the class area in this case. Class $I(SA^+)$ makes a negative contribution to σ^2_Λ of 8%, again mostly from Λ_{SC}. This negative contribution is due to more strongly negative values of Λ_{SC} in the higher sensitivity versions (Fig. 6), and Fig. 7 indicates that this is at least partly due to larger increases in low cloud amount.

Table 10 is the equivalent of Table 9, but for the full 128 member QUMP ensemble. If we sum the variance contributions to σ^2_Λ from Λ_C in the cloud feedback classes which have low cloud feedbacks as their primary cause [classes $A(S^+L^N$: orange), $E(S^-L^N$: dark blue) and $I(SA^+)$] we find that these contribute 50% to σ^2_Λ, which is 59% of the 85% contributed by the global values of Λ_C. The cloud feedback classes which have high cloud changes as their primary cloud response [classes $B(S^+L^-$: yellow), $C(S^NL^-$: green), $F(S^-L^+$: purple) and $G(S^NL^+$: dark red)] contribute 17% to σ^2_Λ (or 20% of the contribution from Λ_C). This excludes classes $D(S^-L^-$: light blue) and $H(S^+L^+$: red) which contribute

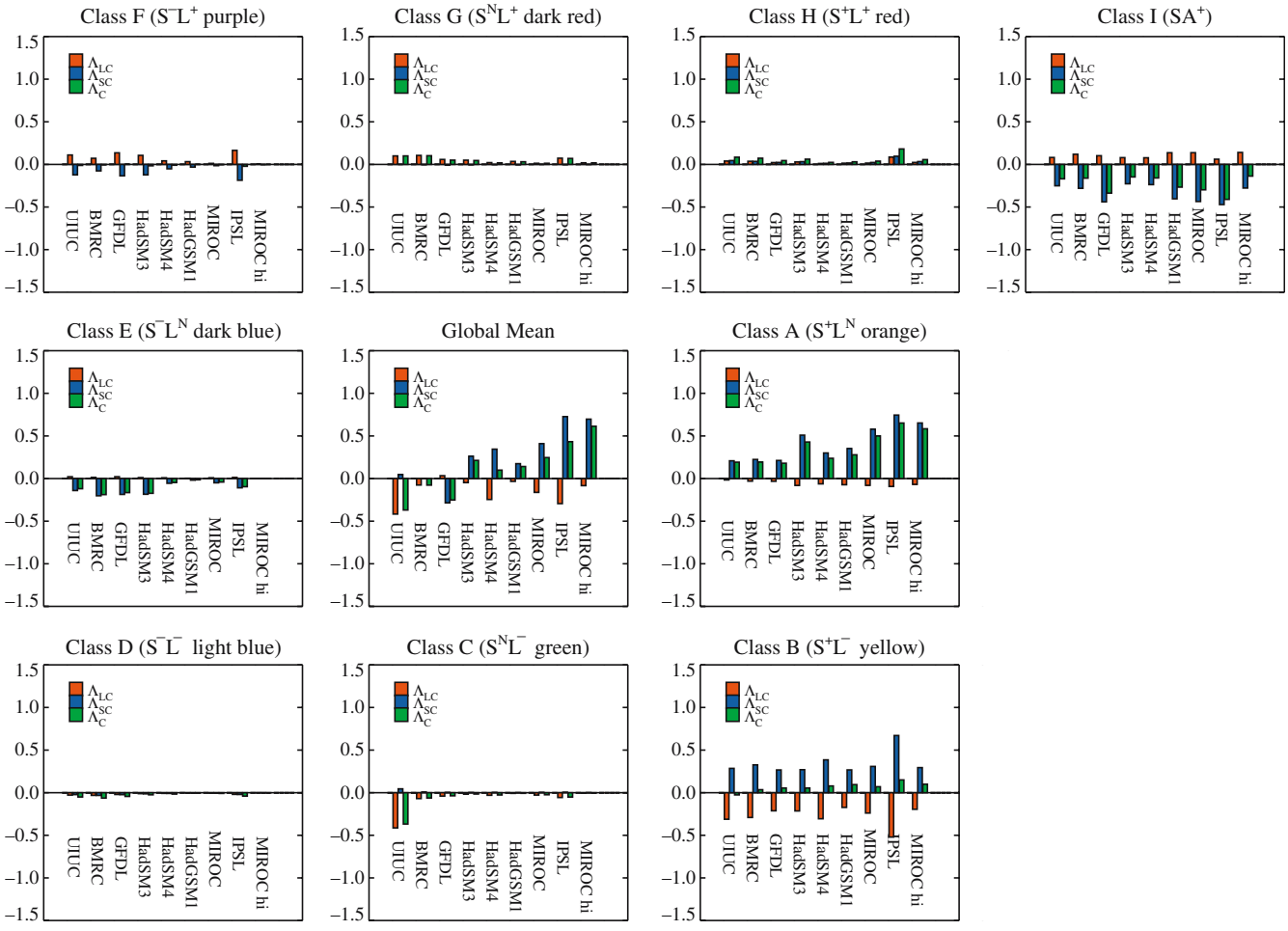

Fig. 6 Contribution to global feedback terms $\langle \Lambda_{SC} \rangle$, $\langle \Lambda_{LC} \rangle$ and $\langle \Lambda_{C} \rangle$ from cloud feedback classes for the CFMIP models, presented in the order of increasing $\langle \Lambda \rangle$

19% to σ^2_Λ (22% of the cloud feedback contribution). The fact that the low cloud feedback classes contribute more than the high cloud classes in the QUMP ensemble is initially surprising, given that our global analysis showed the larger contribution to σ^2_Λ to come from Λ_{LC}. However, a closer examination of Table 10 shows that classes $B(S^+L^-$: yellow) and $F(S^-L^+$: purple) (associated with changes in high-top optically thick clouds) make significant contributions to σ^2_Λ via Λ_{LC} which are almost completely cancelled in the net by equal and opposite contributions from Λ_{SC}. The sum of the contributions from the other classes are larger in Λ_{SC} than Λ_{LC}.

The class with the largest contribution to σ^2_Λ in the QUMP ensemble is class $E(S^-L^N$: dark blue), in which Λ_{SC} contributes 33% to σ^2_Λ (39% the cloud feedback contribution). We calculate that approximately 60% of this contribution is due to differences in class area, with 40% due to differences in feedback strength within the class. Figure 8 shows the equivalent of Fig. 6 but for the nine selected QUMP model versions, and shows stronger negative shortwave cloud feedbacks in this class for the lower sensitivity model versions, which is supported by strong correlations of Λ_{SC} and Λ_C with $\langle \Lambda \rangle$ in Table 10. Figure 9 shows that the increases in low-top clouds (and the smaller increases in high-top clouds) tend to be larger in the lower-sensitivity models, and we find that both the optically thick low-top cloud increases and optically thick high-top cloud increases are anti-correlated with $\langle \Lambda \rangle$ with correlation coefficients of −0.76 and −0.88, respectively. This is consistent with inter-model differences in low cloud response being the primary cause of the differences in the cloud feedbacks in this class, with high-top cloud responses playing a secondary role.

Class $I(SA^+)$ is responsible for 21% of the cloud feedback contribution to σ^2_Λ in the QUMP ensemble (again mostly from Λ_{SC}) and Fig. 8 shows evidence of stronger negative values of Λ_{SC} at lower sensitivities, coinciding with larger increases in low cloud amount (Fig. 9). In this case however we find that the contribution is almost completely due to differences in the strengths of the feedbacks within the class, and class area contributes little. Class $A(S^+L^N$: orange) contributes very little to σ^2_Λ in the QUMP ensemble, unlike the CFMIP ensemble.

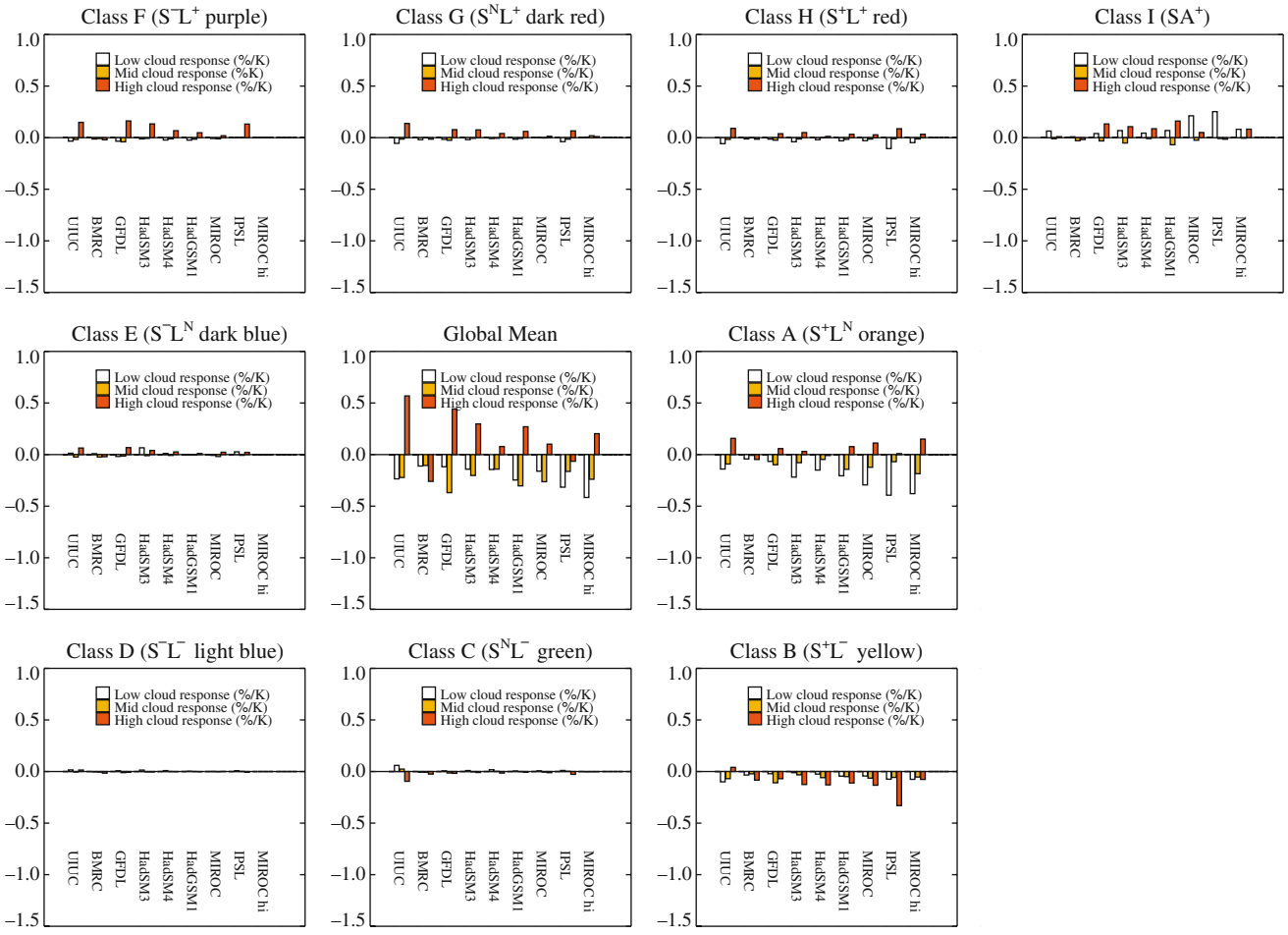

Fig. 7 Area weighted responses of low-, mid- and high-top clouds $\langle C_i \rangle_j$ in each feedback class j in the CFMIP models. Each plot shows the cloud response low-top (> 680 hPa), mid-level-top (680–440 hPa) and high-top (< 440 hPa) clouds (%/K)

Table 10 As Table 9 but for the QUMP ensemble

	GL	$A(S^+L^N)$	$B(S^+L^-)$	$C(S^NL^-)$	$D(S^-L^-)$	$E(S^-L^N)$	$F(S^-L^+)$	$G(S^NL^+)$	$H(S^+L^+)$	$I(SA^+)$
Λ_{LC}										
$\overline{\langle\Lambda_i\rangle}$	−0.10	−0.06	−0.25	−0.04	−0.02	0.01	0.11	0.04	0.02	0.08
$[\Lambda_i]$	−0.10	−0.26	−1.05	−0.83	−0.52	0.13	1.04	0.67	0.54	0.53
V_i	**0.57**	0.00	**0.20**	0.10	0.07	−0.00	0.09	0.05	0.02	0.04
r_Λ	**0.77**	0.07	0.56	**0.80**	**0.79**	−0.03	0.44	0.52	0.39	**0.71**
Λ_{SC}										
$\overline{\langle\Lambda_i\rangle}$	0.19	0.39	0.32	0.00	−0.02	−0.15	−0.12	−0.00	0.03	−0.25
$[\Lambda_i]$	0.19	1.74	1.35	0.07	−0.57	−1.47	−1.15	−0.06	0.63	−1.61
σ_{Λ_i}	0.17	0.09	0.10	0.00	0.03	0.10	0.05	0.00	0.01	0.06
V_i	**0.29**	−0.02	**-0.20**	−0.00	0.08	**0.33**	−0.07	−0.00	0.02	0.14
r_Λ	0.43	−0.05	−0.48	−0.56	**0.78**	**0.84**	−0.33	−0.52	0.37	0.62
Λ_C										
$\overline{\langle\Lambda_i\rangle}$	0.09	0.33	0.07	−0.04	−0.04	−0.14	−0.01	0.04	0.05	−0.17
$[\Lambda_i]$	0.09	1.48	0.30	−0.76	−1.09	−1.35	−0.11	0.61	1.18	−1.07
σ_{Λ_i}	0.22	0.09	0.03	0.03	0.05	0.10	0.01	0.02	0.03	0.06
V_i	**0.85**	−0.01	0.00	0.10	0.15	**0.33**	0.03	0.04	0.04	0.18
r_Λ	**0.97**	−0.04	0.03	**0.80**	**0.78**	**0.82**	0.46	0.51	0.38	**0.73**

Comparing the ensemble mean feedback components of the CFMIP and QUMP ensembles in Table. 9 and 10 shows that the largest differences in Λ_C are in classes $E(S^-L^N$: dark blue) and $I(SA^+)$, with QUMP showing more strongly negative values (largely from Λ_{SC}) in class $E(S^-L^N$: dark blue), while reverse is the case in class $I(SA^+)$. QUMP also shows stronger feedbacks in Λ_{SC} and Λ_{LC} in class $F(S^-L^+$: purple), although the effects

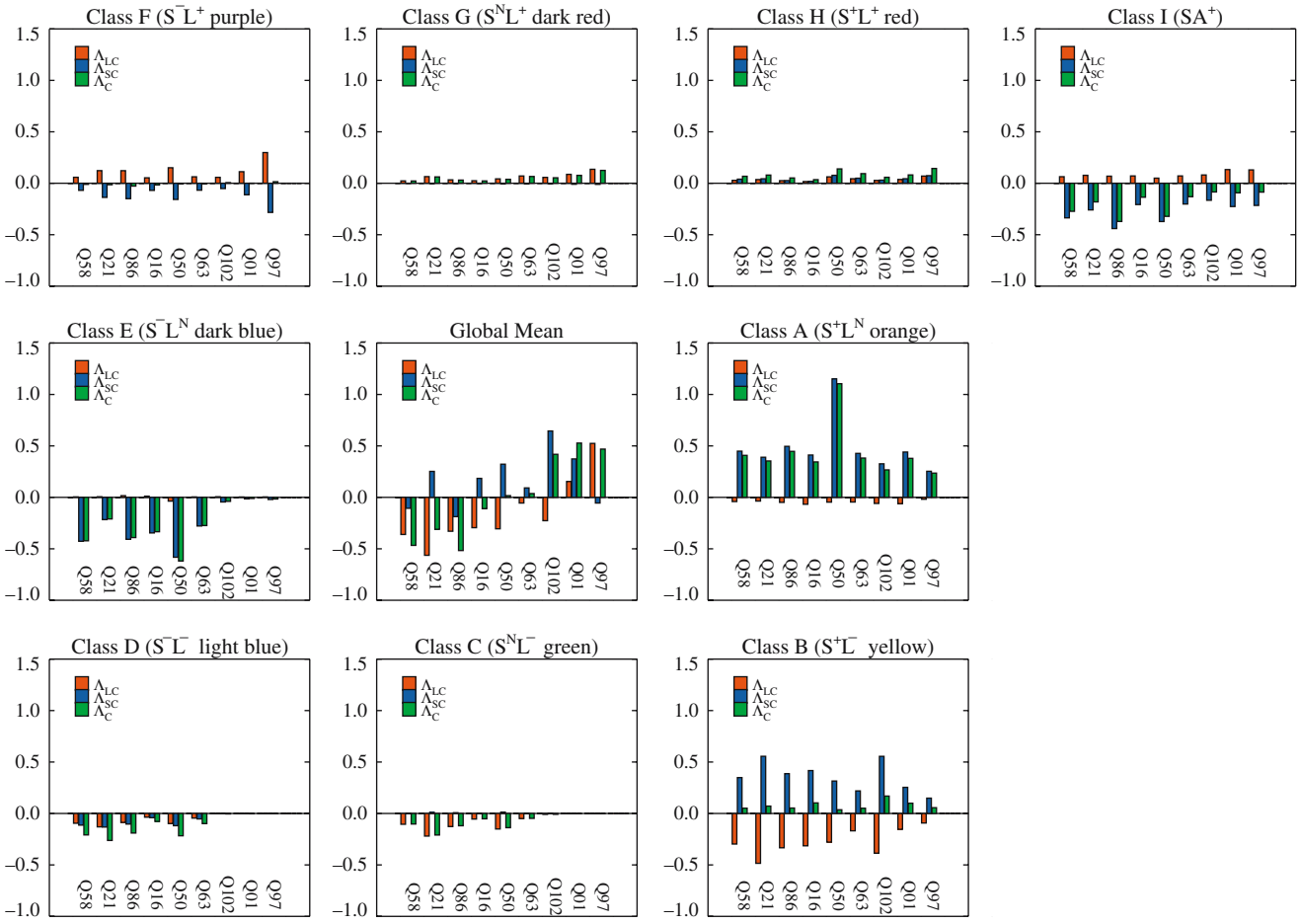

Fig. 8 As Fig. 6 but for the selected QUMP models

largely cancel in Λ_C. The CFMIP standard deviations in Table 9 are significantly larger than those for QUMP in Table 10 in classes $A(S^+L^N$: orange), $B(S^+L^-$: yellow), $C(S^NL^-$: green), $G(S^NL^+$: dark red), $H(S^+L^+$: red) and $I(\mathrm{SA}^+)$. Conversely the standard deviations in classes $D(S^-L^-$: light blue) and $F(S^-L^+$: purple) are significantly larger in the QUMP ensemble than those in CFMIP. (The same caveats noted when comparing the global mean statistics above also apply here.) However, while the ensemble standard deviations of the feedbacks taken in each feedback class are mostly not statistically equivalent, Table 11 shows that QUMP spans 75% or more of the CFMIP range within all classes but one, the exception being class $C(S^NL^-$: green). The CFMIP models typically span a smaller fraction of the QUMP ranges.

6 Summary and discussion

Global and local feedback analysis techniques have been applied to two ensembles of mixed layer equilibrium CO_2 doubling climate change experiments, from the CFMIP and QUMP projects. Neither of these new ensembles shows evidence of a statistically significant change in the ensemble mean or variance in global mean climate sensitivity when compared with the results from the mixed layer models quoted in the Third Assessment Report of the IPCC. A global mean feedback analysis shows that the largest contribution to inter-model differences in climate sensitivity in both ensembles is due to cloud feedbacks, which confirms the findings of studies with earlier models (e.g. Cess et al. 1996; Colman 2003). In the CFMIP ensemble, two-thirds of the inter-model variance in the total feedback Λ (and hence in the climate sensitivity) is due to differences in the strengths of the net cloud feedbacks. The clear-sky shortwave feedbacks contribute most of the remaining one-third. Cloud feedbacks contribute 85% to the variance in the total feedback in the QUMP ensemble. The ensemble mean global feedback components are statistically indistinguishable between the two ensembles (at the 10% confidence level), the exception being the clear-sky shortwave feedback Λ_{SA} which is stronger in the CFMIP ensemble. The ensemble variances of Λ_{SA}, Λ_{LA} (the clear-sky longwave feedback) and Λ_{SC} (the shortwave cloud feedback) are significantly larger in CFMIP than QUMP. However, there is considerable overlap in the ranges of

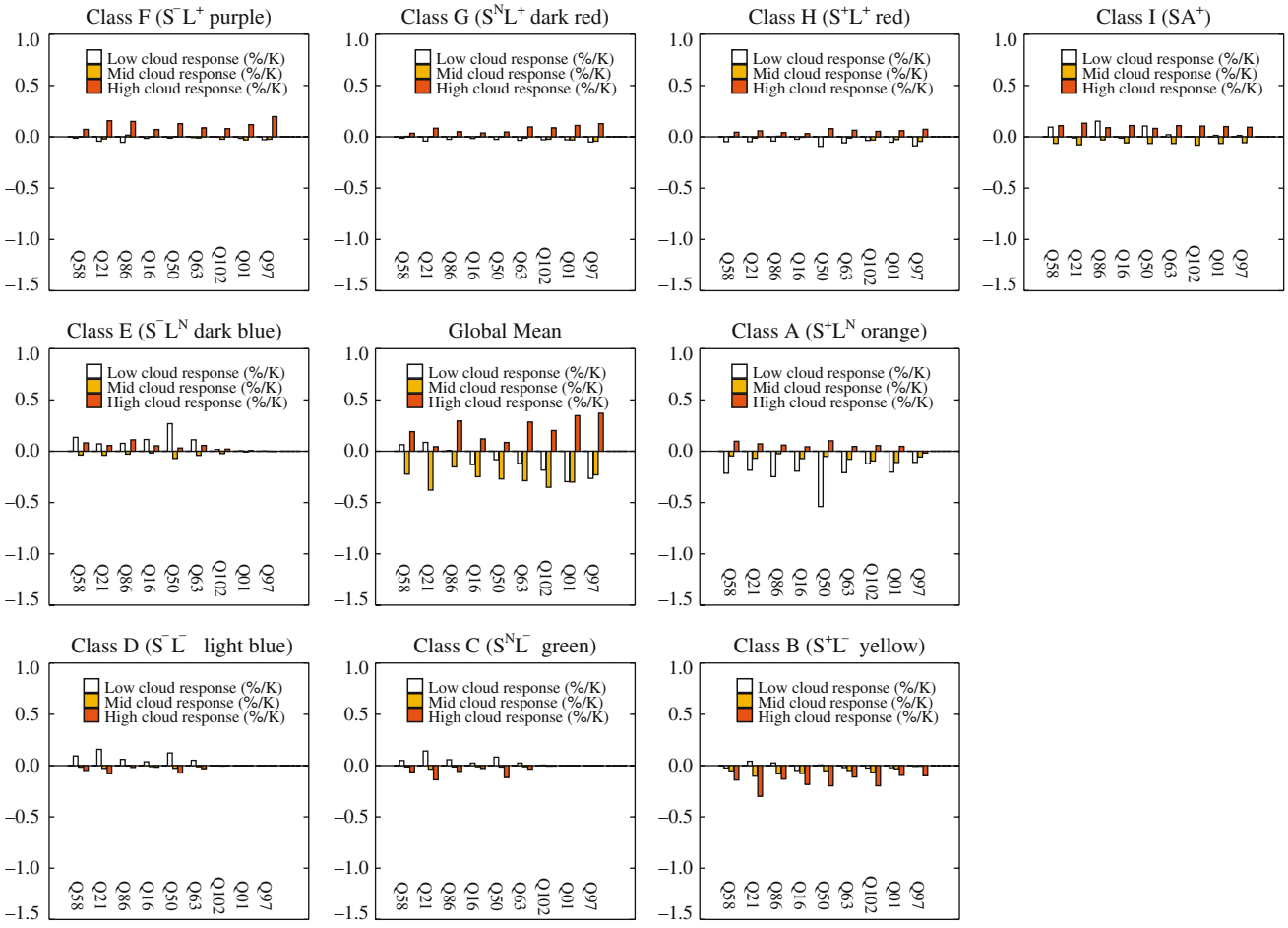

Fig. 9 As Fig. 7 but for the selected QUMP models

Table 11 Overlap in ranges of feedback components taken in QUMP and CFMIP ensemble in the cloud feedback classes

	GL	$A(S^+L^N)$	$B(S^+L^-)$	$C(S^NL^-)$	$D(S^-L^-)$	$E(S^-L^N)$	$F(S^-L^+)$	$G(S^NL^+)$	$H(S^+L^+)$	$I(SA^+)$
QUMP (% CFMIP)										
Λ_{LC}	100	100	100	52	99	100	86	100	78	90
Λ_{SC}	82	98	89	26	98	97	85	94	77	87
Λ_{C}	91	100	70	56	99	97	100	100	75	85
CFMIP (% QUMP)										
Λ_{LC}	41	92	79	98	23	34	51	74	89	86
Λ_{SC}	100	56	75	86	23	33	61	61	89	76
Λ_{C}	79	50	63	55	23	29	31	77	69	68

The overlap is expressed as a percentage of the spread in the CFMIP ensemble, which is a measure of the extent to which the QUMP ensemble spans the range of the CFMIP ensemble in that feedback class. The overlap is also expressed as a percentage of the spread in the QUMP ensemble to give an indication of the extent to which the CFMIP ensemble spans the QUMP ranges

the feedback values taken in the two ensembles. The QUMP ensemble spans 80% or more of CFMIP range in Λ_{SC}, Λ_{LA} and Λ_{LC}, (but only a quarter of the range in Λ_{SA}), while CFMIP spans close to 90% or more of the QUMP range in Λ_{SC} and Λ_{LA}, but 41% and 67% of the QUMP ranges in Λ_{LC} and Λ_{SA}, respectively.

We have introduced a local cloud feedback classification system which distinguishes different types of cloud feedbacks on the basis of the relative strengths of their longwave and shortwave components. Interpretations for each class are provided in terms of various cloud feedback mechanisms (e.g. decreases in low-top cloud amount/optical thickness, increases in high-top cloud amount). These interpretations are tested by examining the actual cloud responses in each cloud feedback class, using diagnostics from the ISCCP simulator. We find that the largest cloud responses in each class are consistent with the physical interpretations provided in each case, both in the CFMIP and QUMP ensembles.

In the CFMIP models, decomposition of the global feedback parameters into contributions from the different cloud feedback classes shows that classes $A(S^+L^N$: orange), $E(S^-L^N$: dark blue) and $I(\mathrm{SA}^+)$ (those in which low cloud changes constitute the primary cloud response) are responsible for 59% of the contribution of inter-model differences in the net cloud feedback to the variance in the total feedback. Areas where the shortwave cloud feedback is positive and the longwave cloud feedback is (relatively) small [class $A(S^+L^N$: orange)] are responsible for most of this contribution (53%). These areas coincide with regions of the globe commonly associated with the presence of low-top clouds, e.g. along the eastern sides of the tropical ocean basins and in the mid-latitude storm track regions, and the largest cloud responses within them are found to be primarily due to reductions in low-top cloud amount (although reductions in mid-level cloud make a secondary contribution). The inter-model differences in the contribution from this feedback class are almost entirely due to low cloud feedback operating over increasingly large areas in the higher sensitivity models, rather than differences in local feedback strengths within the class. Areas where the shortwave cloud feedback is negative and the longwave equivalent is relatively small [class E (S^-L^N: dark blue)] contribute 18%, and are associated with increases in the optical thickness of low-top clouds. In the tropics this class is commonly seen in the trade cumulus regions, (e.g. to the west of South America), and the opposite sign of the cloud feedback here (compared to that in class $A(S^+L^N$: orange) closer to the coasts) may reflect differing cloud feedback processes taking place in convectively and turbulently mixed boundary layers. Class $I(\mathrm{SA}^+)$ represents areas of non-negligible clear-sky albedo feedback, in which the 'cloud radiative forcing' measure of local shortwave cloud feedback requires careful interpretation. Increases in low cloud amount and optical thickness are seen in these areas which are consistent with those in classes which show a negative low cloud feedback in other areas. In the CFMIP ensemble this class acts to reduce the contribution of inter-model differences in cloud feedback to the variance in the total feedback by 12%. The equivalent contribution from classes in which the high-top cloud responses are strongest [classes $B(S^+L^-$: yellow), $C(S^NL^-$: green), $F(S^-L^+$: purple) and $G(S^NL^+$: dark red)] is considerably smaller than the low-top total (33% compared with 59%). Classes $D(S^-L^-$: light blue) and $H(S^+L^+$: red) which are due to a combination of low- and high-top cloud changes contribute 8%.

In the QUMP ensemble, the three low-top cloud feedback classes [$A(S^+L^N$: orange), $E(S^-L^N$: dark blue) and $I(\mathrm{SA}^+)$] are responsible for 59% of the contribution from the net cloud feedback to the variance in the total feedback (in close agreement with the CFMIP ensemble). However, it is class $E(S^-L^N$: dark blue) that is responsible for the largest part of this (39%), while class $A(S^+L^N$: orange) (which was largest in CFMIP) makes a small contribution (−1%). Class $I(\mathrm{SA}^+)$ contributes 21% compared with a negative contribution in CFMIP. As in the CFMIP ensemble, these classes are responsible for a larger proportion than the high-top cloud feedback classes which contribute 20% in total, or that from the remaining classes (21%).

Bony and Dufresne (2005) examine 1% compound CO_2 coupled climate change experiments from the IPCC AR4 models in the tropics. They find that the radiative response of clouds to changes in surface temperatures under climate change in the models differs most in regimes of large scale subsidence, and suggest that changes in marine boundary layer clouds are responsible. Our results confirm this by providing direct evidence of considerable low-top cloud responses in areas which contribute most to inter-model differences in global cloud feedback and climate sensitivity, in both the CFMIP and QUMP ensembles.

We also find evidence of increases in low-top cloud amount and optical depth in high latitude areas where it is difficult to separate the effects of cloud and clear-sky albedo feedbacks. The application of PRP techniques to diagnose local feedbacks would enable a a cleaner separation of these feedbacks, but it is unlikely that these will be applied consistently across all models in the foreseeable future. The evaluation and application of approximate PRP methods (e.g. Taylor et al. 2000; Yokohata et al. 2005; Soden and Held 2005) to local feedback analysis in the CFMIP and QUMP ensembles will be a subject of future study.

The ensemble mean values of the net cloud feedbacks in seven of the cloud feedback classes are statistically indistinguishable between the two ensembles, while classes $E(S^-L^N$: dark blue) and $I(\mathrm{SA}^+)$ show significant differences (at the 10% confidence level). Standard deviations are significantly larger in CFMIP compared to QUMP in six out of nine classes, the exceptions being classes $D(S^-L^-$: light blue) and $F(S^-L^+$: purple) where QUMP is larger, and class $E(S^-L^N$: dark blue) where there is no significant difference. However, QUMP spans 75% or more of the CFMIP range within all classes except class $C(S^NL^-$: green), while the CFMIP models typically span a smaller fraction of the QUMP ranges in each class. Given the differences in the statistics of the local cloud feedbacks in the two ensembles, it is to be expected that there will also be differences in the ranges of regional surface temperature responses seen in QUMP compared to those from multi-model ensembles. This does not necessarily mean, however, that the CFMIP local feedbacks and temperature responses are any more or less realistic than those in the QUMP ensemble. Further work is required not only to establish the reasons for these differences, but also to find ways to integrate the complementary information on uncertainty provided by multi-model- and parameter-perturbed ensembles. In future probabilistic predictions from the QUMP project we plan to use information from the CFMIP models to include an estimate of the uncertainty in a prediction which is not represented because of structural constraints that may arise from the use of a

single model. See Craig et al. (2001) for an example of this approach applied in a different field.

The extent to which we can quantify our confidence in the climate predictions from models is a subject of ongoing study. Confidence measures such as those applied in Murphy et al. (2004) could in principle be applied to multi-model ensembles as well as perturbed physics ensembles, but it would be beneficial also to include measures that have been shown to be directly relevant to the simulation of climate feedbacks (e.g. as in Williams et al. 2005a, b; Bony and Dufresne 2005). It is interesting to note that the six model versions in the mid-upper climate sensitivity range of the CFMIP ensemble all have statistical cloud schemes based on either Smith (1990) or Le Treut and Li (1991) (schemes which themselves have a common heritage). The next phase of the CFMIP project will involve examination of 5 years of daily cloud diagnostics from high and low sensitivity CFMIP models in an attempt to understand the differing feedback processes in terms of the physical assumptions on the models. The Atmospheric Radiation Measurement (ARM) and Gewex Cloud System Study (GCSS; Randall et al. 2003) programmes apply a wealth of techniques to evaluate cloud parametrisations using regional models, single column models, cloud/eddy resolving models, in situ measurements, operational analyses and satellite data (e.g. Xie et al. 2005). Daily cloud diagnostics which capture the key cloud feedbacks in climate models will allow us to apply such techniques more directly to evaluating cloud climate feedback processes.

Acknowledgements This work was funded in part by the UK Department of the Environment, Food and Rural Affairs, under contract PECD 7/12/37. Thanks are due also to IPSL for hosting of CFMIP data, to Jean-Louis Dufresne for providing the IPSL radiative forcing calculations, and to Mat Collins, James Murphy and Jonathan Gregory for helpful comments on the manuscript. We are also grateful to the anonymous reviewers whose comments led to improvements in this piece of work.

References

Andronova NG, Rozanov EV, Yang F, Schlesinger ME, Stenchikov GL (1999) Radiative forcing by volcanic aerosols from 1850 through 1994. J Geophys Res 104:807–816

Boer GJ, Yu B (2003) Climate sensitivity and response. Clim Dyn 20:415–429

Bony S, Dufresne J-L (2005) Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. Geophys Res Lett 32:L20806. DOI 10.1029/2005GL023851

Bony S, Emanuel KA (2001) A parameterization of the cloudiness associated with cumulus convection; evaluation using TOGA COARE data. J Atmos Sci 58:3158–3183

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

Cess RD, Potter GL (1988) A methodology for understanding and intercomparing atmospheric climate feedback processes in general circulation models. J Geophys Res 93:8305–8314

Cess RD, Potter GL, Blanchet JP, Boer GJ, Del Genio AD, Deque M, Dymnikov V, Galin V, Gates WL, Ghan SJ, Kiehl JT, Lacis AA, Le Treut H, Li ZX, Liang XZ, Macavaney BJ, Meleshko VP, Mitchell JFB, Morcrette J-J, Randall DA, Rikus L, Roeckner E, Royer JF, Schelse U, Scheinin DA, Slingo A, Sokolov AP, Taylor KE, Washington WM, Wetherald RT, Yagai I, Zhang MH (1990) Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. J Geophys Res 95:16601–16615

Cess RD, Zhang MH, Ingram WJ, Potter GL, Alekseev V, Barker HW, Cohen-Solal E, Colman RA, Dazlich DA, Del Genio AD, Dix MR, Dymnikov V, Esch M, Fowler LD, Fraser JR, Galin V, Gates WL, Hack JJ, Kiehl JT, Treut HL, Lo KKW, McAvaney BJ, Meleshko VP, Morcrette JJ, Randall DA, Roeckner E, Royer JF, Schlesinger ME, Sporyshev PV, Timbal B, Volodin EM, Taylor KE, Wang W, Wetherald RT (1996) Cloud feedback in atmospheric general circulation models: An update. J Geophys Res 101:12791–12794

Colman R (2002) Geographical contributions to global climate sensitivity in a general circulation model. Global Planet Change 32:211–243

Colman R (2003) A comparison of climate feedbacks in general circulation models. Clim Dyn 20:865–873

Colman R, Fraser J, Rotstayn L (2001) Climate feedbacks in a general circulation model incorporating prognostic clouds. Clim Dyn 18:103–122

Craig PM, Goldstein M, Rougier J, Sehueult A (2001) Bayesian forecasting for complex systems using computer simulators. J Am Stat Assoc 96:717–729

Cubasch U, Meehl GA, Boer GJ, Stouffer RJ, Dix M, Noda A, Senior CA, Raper SCB, Yap KS (2001) Projections of future climate change. In: Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden P, Dai X, Maskell K, Johnson CI (eds) Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, London, pp 525–582

Delworth TL, Broccoli AJ, Rosati A, Stouffer RJ, Balaji V, Beesley JT, Cooke WF, Dixon KW, Dunne J, Dunne KA, Durachta JW, Findell KL, Ginoux P, Gnanadesikan A, Gordon CT, Griffies SM, Gudgel R, Harrison MJ, Held IM, Hemler RS, Horowitz LW, Klein SA, Knutson TR, Kushner PJ, Langenhorst AR, Lee H-C, Lin S-J, Lu J, Malyshev SL, Milly PC D, Ramaswamy V, Russell J, Schwarzkopf MD, Shevliakova E, Sirutis JJ, Spelman MJ, Stern WF, Winton M, Wittenberg AT, Wyman B, Zeng F and Zhang R (2006). GFDL's CM2 Global Coupled Climate Models—Part 1: Formulation and simulation characteristics. J Climate (in press)

GFDL Global Atmospheric Model Development Team (GAMDT) (2004) The new GFDL global atmosphere and land model AM2–LM2: evaluation with prescribed SST simulations. J Climate 17(24):4641–4673

Hansen J, Sato M, Nazarenko L, Ruedy R, Lacis A, Koch D, Tegen I, Hall T, Shindell D, Santer B, Stone P, Novakov T, Thomason L, Wang R, Wang Y, Jacob D, Hollandsworth-Frith S, Bishop L, Logan J, Thompson A, Stolarski R, Lean J, Willson R, Levitus S, Antonov J, Rayner N, Parker D, Christy J (2002) Climate forcings in Goddard Institute for Space Studies SI2000 simulations. J Geophys Res 107. DOI 10.1029/2001JD001143

Hourdin F, Musat I, Bony S, Braconnot P, Codron, F, Dufresne J-L, Fairhead L, Filiberti M-A, Friedlingstein P, Grandpeix J, Krinner G, LeVan P, Li Z-X, Lott F (2006) The LMDZ general circulation model: climate performance and sensitivity to parameterized physics with emphasis on tropical convection. Clim Dyn (in revision)

Johns TC, Durman CF, Banks HT, Roberts MJ, McLaren AJ, Ridley JK, Senior CA, Williams KD, Jones A, Coauthors (2006) The new Hadley Centre climate model HadGEM1: evaluation of coupled simulations in comparison to previous models. J Climate (in press)

Joshi M, Shine K, Ponater M, Stuber N, Sausen R, Li L (2003) A comparison of climate response to different radiative forcings in three general circulation models: towards an improved metric of climate change. Clim Dyn 20:843–854. DOI 10.1007/s00382-003-0305-9

K-1 model developers (2004) K-1 coupled model (MIROC) description. In: Hasumi H, Emori S (eds) K-1 technical report, 1. Center for Climate System Research, University of Tokyo, Tokyo, 34 pp (available from http://www.ccsr.u-tokyo.ac.jp/kyosei/hasumi/MIROC/tech-repo.pdf)

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

Le Treut H, Li ZX (1991) Sensitivity of an atmospheric general circulation model to prescribed SST changes: feedback effects associated with the simulation of cloud optical properties. Clim Dyn 5:175–187

Le Treut H, McAvaney B (2000) Equilibrium climate change in response to a CO_2 doubling: an intercomparison of AGCM simulations coupled to slab oceans. Technical Report, Institut Pierre Simon Laplace 18:20

Martin GM, Ringer MA, Pope VD, Jones A, Dearden C, Hinton T (2005) The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model, HadGEM1. Part 1: Model description and global climatology. J Climate (in press)

McAvaney BJ, Le Treut H (2003) The cloud feedback intercomparison project: (CFMIP). In: CLIVAR Exchanges—supplementary contributions. 26:March

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

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parametrizations in the Hadley Centre climate model—HadAM3. Clim Dyn 16:123–146

Randall D, Krueger S, Bretherton C, Curry J, Duynkerke P, Moncrieff M, Ryan B, Starr D, Miller M, Rossow W, Tselioudis G, Wielicki B (2003) Confronting models with data—the GEWEX cloud systems study. Bull Am Meteorol Soc 84:455–469

Rossow WB, Schiffer RA (1999) Advances in understanding clouds from ISCCP. Bull Am Meteorol Soc 80:2261–2287

Rotstayn LD (1997) A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. Part 1: Description and evaluation of the microphysical processes. Q J R Meteorol Soc 123:1227–1282

Senior CA, Mitchell JFB (1993) Carbon dioxide and climate: the impact of cloud parameterization. J Clim 6:393–418

Sexton DMH, Collins M, Harris G, Murphy JM, Webb MJ (2004) Improvements in sampling parameter space for probabilistic climate predictions. DEFRA Report (available on request from the Hadley Centre)

Smith RNB (1990) A scheme for predicting layer clouds and their water content in a general circulation model. Q J R Meteorol Soc 116:435–460

Soden BJ, Held IM (2005) An assessment of climate feedbacks in coupled ocean-atmosphere models. J Climate (in revision)

Soden BJ, Broccoli AJ, Hemler RS (2004) On the use of cloud forcing to estimate cloud feedback. J Climate 17(19):3661–3665. DOI 10.1175/1520-0442(2004)017

Stainforth DA, Aina T, Christensen C, Collins M, Frame DJ, Kettleborough JA, Knight S, Martin A, Murphy J, Piani C, Sexton D, Smith LA, Spicer RA, Thorpe AJ, Allen MR (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. Nature 433:403–406

Taylor KE, Hewitt CD, Braconnot P, Broccoli AJ, Doutriaux C, Mitchell JFB (2000) Analysis of forcing, response and feedbacks in a paleoclimate modeling experiment PMIP, 2000: Paleoclimate Modelling Intercomparison Project (PMIP). In: Braconnot P (eds) Proceedings of the third PMIP workshop, Canada, 4–8 October 1999, WCRP Report WCRP-111, WMO/TD-No. 1007, pp 43–49

Tett SFB, Jones GS, Stott PA, Hill DC, Mitchell JFB, Allen MR, Ingram WJ, Johns TC, Johnson CE, Jones A, Roberts DL, Sexton DMH, Woodage MJ (2002) Estimation of natural and anthropogenic contributions to 20th century temperature change. J Geophys Res 107. DOI 10.1029/2000JD000028

Tselioudis G, Zhang Y, Rossow WB (2000) Cloud and radiation variations associated with northern midlatitude low and high sea level pressure regimes. J Climate 13(2):312–327

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

Wetherald RT, Manabe S (1988) Cloud feedback processes in a general circulation model. J Atmos Sci 45:1397–1415

Williams KD, Jones A, Roberts DL, Senior CA, Woodage MJ (2001a) The response of the climate system to the indirect effects of anthropogenic sulfate aerosol. Clim Dyn 17:845–856

Williams KD, Senior CA, Mitchell JFB (2001b) Transient climate change in the Hadley Centre models: the role of physical processes. J Climate 14(12):2659–2674

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, Ringer MA, Senior CA, Webb MJ, McAvaney BJ, Andronova N, Bony S, Dufresne J-L, Emori S, Gudgel R, Knutson T, Li B, Lo K, Musat I, Wegner J, Slingo A, Mitchell JFB (2005a) Evaluation of a component of the cloud response to climate change in an intercomparison of climate models. Clim Dyn DOI:10.1007/s00382-005-0067-7

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

Wilson DR, Ballard SP (1999) A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. Q J R Meteorol Soc 125:1607–1636

Xie S, Zhang M, Branson M, Cederwall RT, Del Genio AD, Eitzen ZA, Ghan SJ, Iacobellis SF, Johnson KL, Khairoutdinov M, Klein SA, Krueger SK, Lin W, Lohmann U, Miller MA, Randall DA, Somerville RCJ, Sud YC, Walker GK, Wolf A, Wu X, Xu K-M, Yio JJ, Zhang G, Zhang JH (2005) Simulations of midlatitude frontal clouds by single-column and cloud-resolving models during the Atmospheric Radiation Measurement March 2000 cloud intensive operational period. J Geophys Res 110(D15S03). DOI 10.1029/2004JD005119

Yang F, Schlesinger ME, Rozanov EV (2000) Description and performance of the UIUC 24-layer stratosphere-troposphere general-circulation model. J Geophys Res 105(D14):17925–17954

Yokohata T, Emori S, Nozawa T, Tsushima Y, Ogura T, Kimoto M (2005) A simple scheme for climate feedback analysis. Geophys Res Lett 32:L19703

Yu W, Doutriaux M, Seze G, Le Treut H, Desbois M (1996) A methodology study of the validation of clouds in GCMs using ISCCP satellite observations. Clim Dyn 12:389–401

Zhang MH, Cess RD, Hack JJ, Kiehl JT (1994) Diagnostic study of climate feedback processes in atmospheric GCMs. J Geophys Res 99:5525–5537

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)