IPCC Working Group I Fourth Assessment Report

Expert Review Comments on First-Order Draft

Chapter 8

The following compilation of review comments and author responses is supplied by the Working Group I Technical Support Unit as a record of the process used to prepare the Working Group I report. These comments and responses are not to be edited and/ or re-distributed in part or in full to others.

Please note that under IPCC procedures authors are required to take account of all substantive review comments in both review rounds. Thus responses to individual comments may be influenced by comments from other reviewers.

Batch AB

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-1	A	0:0	0:0	most instances of "climate forcing" could be probably be replaced by "radiative forcing" for more consistency with rest of report [Piers Forster]	Taken into account
8-2	A	0:0	0:0	This is really a tour de force. Far better than past model evaluation chapters. Really informative and well written. An example to us all of how to assess rather than list results. It should win a prize but I'm not going to share any of my chocolates with you [Piers Forster]	Noted
8-3	A	0:0		TSU NOTE: Please see supplementary review material [Richard Allan]	Taken into account: see response to 8-778
8-4	A	0:0		the chapter is very interesting and very useful even for scientist very involved in this matter, therefore I like sincerely congratulate with all the authors; however for the aims of AR4 this chapter is too long and don't take in the right consideration , in many analises, the role of the ocean; I suggest to control the references, that perhaps are too many and many were missed (Sun, 2005; Holt et al...; Santos....). [vincenzo artale]	Noted. Taken into account in SOD
8-5	A	0:0		The Executive Summary focuses on findings that address the capabilities of models, rather than projections of climate change. However, some of the assessments of climate change that have been made in the course of evaluating models are of great interest and should be included in the Executive Summary. One such finding is presented on Pg 52, lines 4-6: "There is no agreement among the models whether global warming will make tropical cyclones more or less intense. There seems to be some agreement among models that the frequency of tropical cyclones will be reduced." It should be repeated in the Executive Summary [Lenny Bernstein]	Accepted. Action – Modify ES.
8-6	A	0:0		My impression of the first draft of this chapter is very positive. It covers most of the aspects of model evauation in a direct and concise way. However, its reading is sometimes difficult because it focuses on changes/progress since TAR. [Alejandro Bodas-Salcedo]	Noted but space limit precludes fuller discussion of TAR results
8-7	A	0:0		I consider the IPCC reports as one of the best sources of information in climate research, and therefore care should be taken to provide a list of references as accurate as possible. Although I have not done a thorough check, I have identified some missing references (below). In addition, the style should be homogeneous throught the list of references. [Alejandro Bodas-Salcedo]	Accepted
8-8	A	0:0		Congratulations for a very well-written Chapter. My minor comments focus on areas where I have been working in over the past years. [Wenju Cai]	Noted – thanks

No.	Batch	Page:line From	To	Comment	Notes
8-9	A	0:0		I am surprised that there is no mention on the Indian Ocean variability, for example, the Indian Ocean Dipole, given its great importance in the neighbouring ountries. I am happy to provide a summary, if this is agreed to be an important aspect. [Wenju Cai]	Rejected. The major aspects of large-scale variability have been dealt with
8-10	A	0:0		I could find essentially no information on the evaluation of climate models ability to simulate sea level rise. [John Church]	Now discussed more explicitly in 8.3.
8-11	A	0:0		Many of the figures comparing zonal mean characteristics of the models vs. observations omit uncertainty estimates on the observations -- e.g., figure 8.3.9 for the global precipitation. The observations from Xie and Arkin should be accompanied by some type of graphical indication of the uncertainty range relative to other estimates from, say TRMM. The absence of an uncertainty range leads gives the impression of a spurious level of uncertainty regarding true zonal-mean precipitation. This is especially true in figure 8.3.13 for surface flux into the ocean. [William Collins]	Noted. The text discusses the uncertainty of the observations in many places. The quantification of the uncertainty in the observations has not been attempted in general.
8-12	A	0:0		INTRODUCTION Software used by regulatory bodies/organizations for use in decisions that affect the health and safety of the public and public policy are always production-level software. Production-level Software is characterized by high degrees of documentation, independent Verification and Validation, and Quality Assurance. Additionally, use of production-level software will reduce the number of re-runs and the distribution of incorrect calculated results and greatly increase the confidence in the reported results. In contrast, research-level software tends to be under development and in states of flux in several major areas of modeling and methods. At the present time, the major AOGCM software seems to be research-level software and not yet at production level status. The objective of these comments is to determine that the major AOGCM codes meet generally accepted industry standard requirements for application to analyses that might affect public policy decisions. Specific software requirements that are generally accepted practice for other software used to support public policy decisions are briefly discussed in these comments. Positive responses to the questions given at the end of these comments are necessary in order for the software to be considered to be suitable for supplying information for public-policy decisions. In the absence of positive responses, this Chapter should include a discussion of the status of the codes relative to production vs. research status and the suitability of the calculated results relative to decision making that will impact	Reject. IPCC Working Group 1 assesses recent scientific research on climate change. It is therefore necessary and appropriate that the report be based results obtained with research codes.

No.	Batch	Page:line From	To	Comment	Notes
				public policy. REFERENCES P. J. Roache, "Verification and Validation in Computational Science and Engineering," Hermosa Publishers, Albuquerque, 1998. N. Oreskes, K. Shrader-Fechette, and K. Belitz, "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences," Science, Vol. 263, pp. 641-646, February 4, 1994. This is a longish comment in which several issues are addressed. The issues are as follows: I. SOFTWARE AND PUBLIC POLICY DECISIONS The Third Annual Report, TAR, in 2001 included a Summary for Policymakers. I assume that a similar report will be a part of the 2006 Annual Report. These comments address several important issues relative to the use of computer software as a basis for information that might be used to make, or change, public policy. The potential impact of policy changes to address climate-change issues are enormous; probably unmatched in recent history. Policy makers will use the information to determine public policies that have the potential for enormous impact on millions of citizens. Software that is used to support decisions that might impact the health and safety of the public, and public policy decisions, must be supported by significant independent Verification and Validation, maintained under an approved Quality Assurance Plan, the users of the software must be shown to be qualified to apply the software to the analyses of interest, and the software must be shown to be Qualified for application to the analyses of interest. The software and associated calculational results presented in this chapter are part of the foundation for the information for the Summary for Policymakers. As such, the software itself requires examination relative to its Verification, Validation, Quality Assurance, and Qualification of the users of the software, Relative to policy information, these IPCC reports are unique because it is a true and universal fact that policy information is generally obtained under the direction of formally organized regulatory bodies and organizations. These offices are almost universally attached to the governments of individual countries so that government-specific requirements can be associated with the regulatory bodies. Some of those requirements are associated with the software from which calculated results that provide information to decision makers are obtained. Several aspects of computer software from the viewpoint of decision making are discussed in the following paragraphs. II. PRODUCTION-LEVEL SOFTWARE VS. RESEARCH-GRADE SOFTWARE There are distinct characteristics for production-level and research-level computer	

No.	Batch	Page:line From	To	Comment	Notes
				software. Research-level software, for example, are generally used as learning tools in efforts to gain additional, deeper understanding of the basis processes involved in the applications of interest. Research software is generally under constant change and development and many studies are used to determine the effects and properties of the changes. Many times changes to very basic aspects of the software are the focus of the development efforts; changes to numerical solution methods are examples. Changes at such a basic level usually have wide-ranging impacts on the calculated results and thus parameterizations and tuning must usually be changed in order to accommodate the effects of the changes in the software. It is only after the changes have been fully understood and qualified that the changes are moved into the production-level versions of the codes. In contrast, production-level software undergoes changes at a much reduced rate. The proven elements from the research-level development efforts are factored into the production software over long time scales. Production-level software is usually described as "frozen" or "fixed" and no local updates are incorporated into the code during the course of application calculations. The general AOGCM codes used for the calculations reported in this chapter are basically research codes. The evidence for this statement is contained throughout the chapter in that changes to many of the basic aspects of the codes are discussed. Generally, many of these changes are mentioned in the context of how the codes have continued to evolve since the TAR, and the new capabilities and calculations with the changed codes. Research codes are not tools to be used to provide information for public policy decision making. Instead, production-level software is generally used for these kinds of applications. Production-level software is basically characterized by a high degree of documentation, independent reviews of the documentation and the code, independent Verification of the coding, Validation of the code for its applications, maintenance under an approved Software Quality Assurance Plan, determination of the Qualification of the Users of the code, and demonstration of the Qualifications of the Code for analyses in its intended areas of applications. Software that is used to obtain information that will be a part of decisions that affect the health and safety of the public, and decisions that will affect public policy, must, by law in some cases, be independently determined to meet production-level software requirements. Some required characteristics of production-level software include: 1. Documentation of the software is extremely comprehensive and detailed. The documentation is usually to the extent that independent replication of both the software and its calculated results can be carried out. Software manuals are the usual method of	

No.	Batch	Page:line From	To	Comment	Notes
				presenting the documentation. Generally these manuals will include the following: A Theory Manual in which the details of all models and methods used in the software are described. A User Manual that describes how to use the software. A Computer Programmer Manual in which the details of the code structure and the coding are described. A Verification and Validation Manual in which the Verification and Validation activities associated with the software are described. Other manuals and reports in which the application of the software in its intended application areas are described. 2. Verification of Software Verification means that the equations coded in the software are correctly and accurately solved. "The equations are solved correctly." A brief description of Verification is: Verification is the process of ensuring that the equations are solved correctly and accurately. The focus of verification is the actual coding of the software with an objective to determine: (1) that the coding corresponds to the equations given in the specification document, (2) the order of accuracy of the numerical methods, and (3) the order of convergence of the numerical methods. In general, the latter two objectives are purely mathematical and go to the heart of the coding of the solution methods. The terminology Verification is used in the sense of Roache in the Reference listed above. It is not in the sense of Oreskes et al listed above. Roache has given a discussion of the differences between the Oreskes et al paper and the practical necessity of Verification and Validation of software. Complex software, based on comprehensive mathematical descriptions of inherently complex natural phenomena and processes, and designed for applications to complex phenomena and processes can in fact be and are in fact Verified. Verification is standard operating procedure for all kinds of software. Verification must always precede Validation and Validation must always precede applications of the code. Computer software that is used in part as a basis for public policy decision making must in fact be independently determined to have been independently Verified, Validated, Qualified, and maintained under an approved Quality Assurance Plan. The users of the software must also be qualified for application of the software to an analysis of interest. In some cases these Verification, Validation, Quality Assurance and User Qualification requirements for production-level software, especially software that is used as a basis for decisions that affect the health and safety of the public, are codified in the laws of the country. 3. Validation of Software Validation means that the predictions of the code compare favorably with data. "The	

No.	Batch	Page:line From	To	Comment	Notes
				correct equations are solved." A brief description of Validation: Validation is the process of ensuring that the correct equations have been solved. The objective of Validation is to determine that the mathematical models of physical phenomena and processes are correct. Comparisons of code predictions with experimental data is the basic approach to Validation. Code-calculated results are compared with experimental data. While Validation of AOGCM software is an extremely difficult problem, it is not an impossible problem. There is an abundant literature on Validation methodologies for complex codes designed for analyses of complex natural phenomena and processes. The many difficult issues associated with Validation of the major AOGCM codes are beyond the scope of these comments. The important concept is that independent Verification must always precede Validation. 4. Quality Assurance Plans for Climate-Change Software Software Quality Assurance Plans are methods and procedures to ensure that the code and its applications maintain Verification and Validation status throughout its life cycle. A brief description of Quality Assurance: Quality assurance procedures are implemented into a Quality Assurance Plan in order to maintain an approved production-level code under QA procedures. Maintenance activities include: (1) investigating and disposing of user-reported problems with the code, (2) updating existing models and methods to provide extended capabilities, and (3) incorporation of new models and methods to provide for new capabilities. Distribution of controlled versions of the code to users and maintaining the official QA records for the code, and the storage requirements for the records, are also performed under the QA plan. All the above discussions of software V&V and QA, while given in the framework of the AOGCM codes, in fact apply to all software that enter the decision-making processes. From small one-off codes used for data acquisition and analysis, to codes that operate on the output from other codes, etc.; all software used in the analysis processes. III. SPECIFIC COMMENTS TO BE ADDRESSED BY COMMENT REVIEWERS The comments on this Chapter are the following and require specific answers to the following questions. For the AOGCM codes the results of which are reported in this chapter: (1) Identify the codes that have readily available documentation sufficient to allow independent replication of the software. If there are none, so state. (2) Identify the codes that have undergone independent Verification of the coding. If there are none, so state. (3) Identify the codes for which it has been demonstrated that the equation sets for the	

No.	Batch	Page:line From	Page:line To	Comment	Notes
				models are correctly solved. If there are none, so state. (4) Identify the codes for which the order of accuracy of the numerical methods has been demonstrated by calculations. If there are none, so state. (5) Identify all calculations reported in this chapter that have been shown to be grid-independent. If there are none, so state. (6) Identify all calculations reported in this chapter for which the effects of changing initial and boundary conditions have been investigated. If there are none, so state. (7) Identify all calculations reported in this chapter for which the effects of all iterative stopping criteria have been investigated. If there are none, so state. (8) Identify all calculations reported in this chapter for which the sensitivity of changes in the parameterizations in the models have been investigated. If there are none, so state. (9) Identify the codes that are maintained under an approved Software Quality Assurance Plan. If there are none, so state. Note that a source-control system is not a Software Quality Assurance Plan. (10) Identify the analyses for which it has been independently verified that the input (all data and all parameters) have been correctly specified. If there are none, so state. (11) Identify the analyses for which it has been independently verified that the output files (text, plots, graphics) have been correctly constructed. If there are none, so state. (12) Present the discussion that will be in the final report about the status of the codes relative to production vs. research status and the suitability of the calculated results relative to decision making that will impact public policy. If none, so state. [Hughes Dan]	
8-13	A	0:0		Summary: I found the organization of this first order draft of chapter 8 to be good but the execution somewhat uneven. Some unevenness is understandable and probably inevitable, since the document has multiple audiences. But the non-uniform writing styles of the different sections can throw off a reader and could be improved. I did not detect any significant errors or glaring omissions in the assessment of the state of the science, though I assume numerous additions will be made to the reference list, as more papers are submitted to the TSU and reviewers suggest other possible additions. So, the majority of my comments touch on issues of clarity of the presentation. In my reading of the chapter, I tried to keep in mind 3 audience segments; a) the policymakers and other stakeholders who will read only the executive summary, b) the WG1 and WG2 scientists who will use this as a reference book for ~5 years after it is published, and c) the ideologues of various political hues who will read this looking for bits that they can	Noted. The text has been modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				extract to promote their particular cause. Clear, unambiguous text and figures will help minimize the chance of misinterpretation by group a or distortion by group c. I realize that the scientific content typically is the primary concern of early drafts and style concerns are tended to later the process. So I am not particularly concerned by the unevenness of the writing in this draft, but will make the following two observations in case they may be of any use. One type of writing style inhomogeneity is evident in the approach different authors took when selecting papers to cite. Some seemed to focus on just a few key papers per topic, offering these key papers as representative of the work done on that topic by the broader community. Others seemed to go out of their way to cite virtually any and all recent papers that touched on a particular subject, providing a more comprehensive literature list, albeit with little context. A case can be made for each of these styles, but having both appear in different sections of the same document is not ideal. Similarly, some sections seem to be written more from the perspective of the modeling accomplishments vs. remaining challenges glass being half full, while for others it seemed half empty at best. For me, the sections of the first order draft of chapter 8 that were most effective were those that opened with an introduction that provided some background information (including a description of the importance of the topic at hand and how it fits into the larger climate picture), broadly described the state of the modeling science (with an emphasis on the advances made since the TAR) and then transitioned into the more detail oriented literature review part, citing key papers that document and illustrate both the strengths and weaknesses of the models and providing some context (as opposed to merely reading like a laundry list), and then finished with a brief, forward-looking summary. (Sections that I found to not “read well” and that might benefit from improved introductory and/or summary text include 8.2 and especially 8.2.1, 8.2.2, and 8.2.6.). [Keith Dixon]	
8-14	A	0:0		Chapter 8 provides a generally comprehensive overview of climate-model developments and applications since TAR. It is largely organized as a review of the relevant literature. As such, it will be very useful, although the literature review is often uncritical. An assessment could aspire to much more, though, given the resources and organization of IPCC, the practical possibilities for doing so are obviously limited. One aspect of the chapter stands out especially. Most of the discussion describes the abilities of climate models to simulate contemporary climate and its variability. The inclusion of this material is justified on the grounds that recent improvements in these areas increase confidence in the abilities of climate models to project future climate. A burning assessment question is by how much do these improvements increase that confidence? The issue is discussed on	Noted. Chapter presents state of current knowledge in this area

No.	Batch	Page:line From	Page:line To	Comment	Notes
				p. 8-9, where it is actually suggested (18-22) that the fidelity of contemporary climate simulations may not provide strong constraints on future climate. One could imagine a chapter much more sharply focused on that issue with considerably more succint treatment, for example, of variability in Section 8.4. The latter consists of a nearly exhaustive review of capabilities in simulation of variability in current climate, with some predictions of how this variability might change in the future. Rarely, in this specific case as elsewhere in the chapter, does this meet the stated chapter goal on p. 8-3 (3-4) "to assess the capacity of global climate models...for projecting future climate change." [Leo Donner]	
8-15	A	0:0		Although Stainforth et al.'s (2005, Nature) study appears in several appropriate locations in the chapter, one aspect which is very important to the issue of assessing the capabilities of climate models to project future climate change is not emphasized. Their study showed explcitly how wide the range of climate outcomes can be as critical, poorly known parameters are varied under the constraint of producing realistically at least some aspects of contemporary climate. This provides an important way of assessing climate-change uncertainty for models and deserves further discussion in the chapter, including the Executive Summary. [Leo Donner]	The uncertainty of climate sensitivity estimates is discussed at length in Ch 10. In section 8.6, Stainforth et al. (2005) is cited as one study (among many others) showing that "in many climate models, details in the representation of clouds can substantially affect the model estimates of cloud feedback and climate sensitivity".
8-16	A	0:0		In at least one aspect, the chapter is unbalanced in the space devoted to particular topics. Section 8.4 devotes relatively great detail to design of studies on climate variability, to the extent where on p. 8-48 (23) it lists the starting months in one model's seasonal-interannual experiments. This occurs despite the fact the chapter is never able to provide much qualititative or quantitative characterization of how these results help us understand the ability of these models to project furture climate. On the other hand, the chapter clearly identifies cloud-feedback uncertainty in Section 8.6.3 as a crucial impediment to such future projections. Yet, the chapter fails to discuss in any detail progress in the AR4 model parameterizations important to simulating these and other important feedbacks. The parameterization section 8.2.1.3 briefly describes why parameterizations are necessary and is strangely defensive in tone ("Cloud parameterizations are not simply curve fits or collections..."). It spends nearly a quarter of its space on superparameterization and ultra-high resolution calculations which are not used in any of the models in the AR4 assessment. Except for briefly mentioning the Lock (2001) boundary-layer parameterization and semi-Lagrangian advection, there is no discussion (much less assessment) of advances in numerics or parameterizations in the models used for AR4, despite those advances being listed as the first highlight on p. 8-3 (14-16). This is a very serioius omission which should be corrected as the draft is revised. [Leo Donner]	Text has been modified to include more discussion of models used in AR4, including recent changes to both numerical methods and physical parameterizations. Text not directly relevant to AR4 models has been shortened.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-17	A	0:0		The chapter is very uneven in its assessment of the consequences for projecting future climate of the deficiencies in simulation of contemporary climate. Positive examples are on p. 8-34, where likely errors in transient climate response are linked to problems in the simulation of Southern Ocean mixing(4-8) and errors in sea-ice extent are linked to sensitivity (48-49). The atmospheric component section has some especially glaring deficiencies. Examples: (1) On p. 8-26 (12-20), individual model errors around 20 W/m**2 and a multi-model mean error around 13 W/m**2 are noted for outgoing SW. There is no indication as to the consequences for future model projections, even though, if errors in feedbacks related to this field are an order of magnitude smaller than the errors themselves, those feedbacks will be comparable to radiative forcing due to increased GHG forcing. (2) On p. 8-26 (56-57), errors in zonal-mean implied energy tranpsort are described as "encouraging." Errors in implied transport at the latitudes of maximum transport are up to 20% in magnitude. Are these errors sufficiently small, relative to the objective of projecting future climate change, to be spoken of as "encouraging"? (3) On p. 8-27 (32-41), the discussion of the precipitation errors also fails to indicate consequences of these errors. The model negative precipitation bias over the Amazon is not even noted, despite its signficance for coupling these models to carbon-cycle models, an important development thrust for earth-system simulation. (4) On p. 8-37 (7-18), very large differences in surface SW among the models are noted. Again, there is no mention that, if feedback errors involving this quantity, which are very likely given the roles of clouds and aerosols, are even an order of magnitude smaller than the model range, they will be substantial relative to changes in radiative forcing associated with changes in GHG, with large implications for projecting future climate. [Leo Donner]	Concerning the Reviewer's examples 1 & 3, it is possible that the errors in the mean fluxes bear little relationship to errors in the feedbacks (although this cannot be rigorously proved). The absolute error will certainly affect the basic state, and in the absence of compensating errors will lead to biases in temperature. The feedback, on the other hand, may be largely independent of basic states that are not too unrealistic. Many feedbacks in models seem to be proportional to the forcing (and also the surface temperature change), so a 20% error in SW may not imply any error in climate sensitivity. The text has been rewritten.
8-18	A	0:0		The overall structure and comment of the chapter is good. As someone approaching this from a different speciality, the use of the "average" model continues to mystify me. In some cases, it does serve its purpose - to identify biases that are common to all models. In others, though, it seems to bury issues. For example, Figure 8.3.17 the SH peak is at a different latitude for each model - some higher than observed, some lower. The mean is close to the median. The argument that this fact has significance is simply unclear to the outsider - the physical processes in the models are represented differently, and to a greater or lesser extent. I recognize that within this field the use of the "mean model" is accepted and thought to be important, but to the outsider it just fails to convey a sense that the model results are strongly connected to any model physics. [Anne Douglass]	Accepted. The text has been modified.
8-19	A	0:0		The Executive Summary should include the finding from page 52, lines 4-6: "There is no agreement among the models whether global warming will make tropical cyclones more or less intense. There seems to be some agreement among models that the frequency of	Rejected. This is the purview of Chapter 10.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				tropical cyclones will be reduced." [Howard Feldman]	
8-20	A	0:0		Model comparisons in this chapter are very valuable to the climate community. This is a fundamental difference between AR4 and TAR. This chapter adequately addresses intercomparison and should be regarded as a fundamental role of the IPCC report. [Melanie Fitzpatrick]	Noted
8-21	A	0:0		The chapter is well organized and cover all aspects important to consider for climate modeling, especially from the most complex to simpler models. However, a better understanding would not be achieved if modelers dod not give suggestions or reccommendations on what are relevant observations enabling validations and investigations of model limitations. [Savitri GARIVAIT]	The chapter assesses current scientific knowledge but cannot make pleas for research.
8-22	A	0:0		Generally a thoughtful and balanced chapter. Congratulations. Some debatable philosophical issues, and some inattention to oceanographic uncertainties. Comments follow. [Chris Garrett]	Noted
8-23	A	0:0		It is some years now since I commented on one of your drafts that you were not entitled to use the word “validation” because none of your nodels had ever been subjected to a proper validation procedure, which must involve strong evidence that models are capable of actual future prediction. At the time you changed the word “validation” to “evaluation” in that draft no less than fifty times and you have now settled down to a situation where you dare not test your models against what is happening now in the climate and all you do is “evaluate” them . This whole Chapter is therefore redundant. However much “confidence” you display the only way you can justify it is by showing that your models actually succeed in predicting the future. [Vincent Gray]	Rejected. See discussion in 8.1.1, 8.1.2, 8.4.11, Question 8.1
8-24	A	0:0		If you were serious you would grade models in terms of their relative plausibility. You are scared to do this as it would antagonise those at the bottom of the pile and they might not co-operate with you. As a result you are in the embarassing situation of having equal “confidence” in every model, however absurd. [Vincent Gray]	Rejected. The current state of understanding of this issue is discussed in 8.1.
8-25	A	0:0		It is a pretty hopeless task to persuade us that the models make sense when so many of them include absurd assumptions, such as the belief that carbon dioxide in the atmosphere is increasing by 1% a year. [Vincent Gray]	Rejected. No model makes such an assumption.
8-26	A	0:0		I think the chapter would be helped if several tables were added to the text listing, for example: (1) the names and basic characteristics (resolution, model components, model	(1) is already covered by Table 8.2.1 (2) Space precludes giving too much

No.	Batch	Page:line From	Page:line To	Comment	Notes
				developers, reference to a descriptive paper, etc.) vof all of the coupled GCMs reviewed in the chapter; (2) how each of the models "qualitatively" performs on many of the model output characteristics (such as model performance on simulating Southern Hemisphere ocean circulation, tropical storms, temperature extremes, precipitation extremes, extra-tropical storms, meridional overturning circulation, etc. Perhaps each model's performance could be "color-coded", in order to give the reader a quick look at which models are included when statements such as "many of the models have been found to reliably simulate the XX process" are made in the chapter. [Chuck Hakkarinen]	detail on this, though much of the information could be deduced by examining the figures and supplementary material.
8-27	A	0:0		Would have been easier to interpret some figures (e.g., 8.3.1, 8.3.4, 8.3.19, 8.3.22) if they were in Celcius rather than Kelvin [Anthony Hirst]	Accepted.
8-28	A	0:0		Considering the length limit and highly condensed style of the review, I'd say the balance of this chapter is generally good overall and it is well written and presented (though I am not capable of providing an expert review of all sections). I support the scientific conclusions in the executive summary. [Timothy Johns]	Noted
8-29	A	0:0		The correct name for the HadGEM1 model is "HadGEM1", not "HADGEM1". Please correct any uses of "HADGEM1" throughout this chapter (and elsewhere in the report if they occur). [Timothy Johns]	Accepted.
8-30	A	0:0		I have not read the chapter from the beginning to the end. However, there seems to be no evaluation how the models perform with respect to transient forcing (e.g. over the 20th century). Please include such an evaluation. [Fortunat Joos]	Discussed in Chapter 9
8-31	A	0:0		There should also be an evaluation how climate models perform with respect to ocean heat uptake. [Fortunat Joos]	Accepted – new 8.3.2. Simulation of observed transient heat content change is discussed in Chapter 9
8-32	A	0:0		Ocean ventilation time scales can be assessed by comparing modeled and observed distributions of ventilation tracers such as CFCs or radiocarbon. Realistic ventialation time scales are very important if it comes to model the upatek of heat, the transient climate sensitivity and carbon uptake. This chapter should assess the available studies. (See for example:. J.-C. Dutay, Bullister, J. L., Doney, S. C., Orr, J. C., Najjar, R., Caldeira, K., Campin, J.-M., Drange, H., Follows, M., Gao, Y., Gruber, N., Hecht, M. W., Ishida, A., Joos, F. , Lindsay, K., Madec, G., Maier-Reimer, E., Marshall, J. C., Matear, R. J., Monfray, P., Plattner, G.-K., Sarmiento, J., Schlitzer, R., Slater, R., Totterdell, I. J., Weirig, M.-F., Yamanaka, Y. and Yool, A. Evaluation of ocean model ventilation with	Noted, but chapter focus is on the specific models used in AR4. Some material on ventilation added in 8.3.2

No.	Batch	Page:line From	Page:line To	Comment	Notes
				CFC-11: comparison of 13 global ocean models. Ocean Modelling , 4, 89-120, 2002; K. Matsumoto, J. L. Sarmiento, R. M. Key, O. Aumont, J. L. Bullister, K. Caldeira, J.-M. Campin, S. C. Doney, H. Drange, J.-C. Dutay, M. Follows, Y. Gao, A. Gnanadesikan, N. Gruber, A. Ishida, F. Joos, K. Lindsay, E. Maier-Reimer, J. C. Marshall, R. J. Matear, P. Monfray, R. Najjar, G.-K. Plattner, R. Schlitzer, R. Slater,, P. S. Swathi, I. J. Totterdell, M.-F. Weirig, Y. Yamanaka, A. Yool, and J. C. Orr. Evaluation of ocean carbon cycle models with data-based metrics. Geophysical Research Letters , 31, doi:10.1029/2003GL018970, 2004; S. A. Müller, F. Joos, N. R. Edwards, and T. F. Stocker. Water mass distribution and ventilation time scales in a cost-efficient, 3-dimensional ocean model. J. Climate , submitted, 2005. [Fortunat Joos]	
8-33	A	0:0		There seems to be no evaluation of carbon cycle models, despite the effort by the Ocean Carbon Cycle Model Intercomparison Project. Please clarify with chapter 7 which chapter will evaluate biogeochemical ocean and terrestrial models. [Fortunat Joos]	Reject: A sentence has been added, but since the carbon cycle is not included in the Chapter 10 AR4 models, it is not relevent to this chapter's goals.
8-34	A	0:0		This is a thorough and detailed review of the ability of current climate models to simulate the mean climate and some aspects of its variability. I have not had time to prepare detailed comments, so I have focussed on some general issues in my comments below. [David Karoly]	Noted
8-35	A	0:0		The chapter is already very long and I found it not easy to digest the vast amount of material it contains. I have no easy solution for this. However, I noted a number of topics in which there are systematic errors in model simulations of climate variability or mean climate that are not mentioned. In particular, I think that there should be greater coverage of the following topics, in a rough priority order below: [David Karoly]	Taken into account in revision.
8-36	A	0:0		1. Near-surface air temperature variability either on global, continental or regional scales. A critical question for climate models is how well they simulate the mean surface temperatures and their variability. The simulation of mean surface temperature is addressed in section 8.3.1 and variability of surface temperatures is hardly addressed at all, except very briefly in section 8.4.4. A common question is how well to models simulate the internal variability of global mean temperature (addressed somewhat in chapter 9) and continental and regional scale temperature variability on interannual and decadal timescales. A common model problem in the pas has been a model overestimate of temperature variability over land and an underestimate over the oceans. Braganza et al (2003) show that for large-area average temperature variability, the models do a good job of simulating the interannual and decadal variability of temepratures over the land and ocean now. However, if the variability of temperatures is considered at individual grid-	On SAT variance. Reject. 8.4 deals with modes of variability. On ENSO: Accept. A new figure in 8.4 directly compares the ENSO frequency in the TAR and FAR simulations.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				boxes, models generally have substantially too much temperature variability at middle and high latitudes (Karoly and Wu, J Clim, Nov 2005). The discussion of the model simulation of ENSO in section 8.4.7 is mainly descriptive and provides little quantitative information on whether modes typically underestimate or overestimate the SST variability in the tropical eastern Pacific associated with ENSO, or how good (or poorly) are the simulated variability of the ENSO frequency. [David Karoly]	
8-37	A	0:0		2. Atmospheric general circulation. This is covered in section 8.3.1 but again I could find little discussion of common systematic errors, like the underestimate of the ampltude of the standing waves in the NH winter, leading to a too strong zonal flow. Also, while it is noted that the zonal winds in the SH are too strong in many of the IPCC simulations, there is no discussion of transient eddy statistics. I think that the eddy heat and momentum fluxes and eddy variability are generally too strong in the higher resolution models in the SH, although they are about corrrect in the NH. These aspects of the general circulation are described in some of the papers arising from analysis of hte new AR4 model simulations. [David Karoly]	Accepted.
8-38	A	0:0		The assessment of climate sensitivity and feedbacks is thorough and more complete than in previous IPCC assessments. However, the discussion of climate sensitivity in section 8.6 needs to be consistent with that in other chapters. For example, the definition of TCR on page 52, line 47 is valid only for model simulations with a 1% per year increase in CO2. [David Karoly]	Accepted. We now write in the text: "..the TCR (Cubasch et al. 2001) is defined as the globally averaged surface air temperature difference for the 20-year period around the time of CO2 doubling minus the control run in a 1%/yr atmospheric CO2 increase scenario".
8-39	A	0:0		As a very general comment, chapter 8 evaluates the multi model ensemble, and chapter 10/11 uses it for projections, but some discussion of what it actually is and to what extent it is useful, is missing. First, it should be stated somewhere what the criteria are to submit model results. Second, it should be clear that several models or versions can be submitted by one group. Thus the models will not be independent, and any diagnostic or projection is biased towards which models are submitted (two models with the same physics but different resolution will count twice, thus their bias will be counted twice). Third, the performance of the models varies greatly, from good to extremely poor (see e.g. BCC-CM1 in Fig. 8.3.9, or the MOC in some models). Treating them equally is probably the only feasible way here, but some comments on the limitations would be helpful. Fourth, there is no reason whatsoever to a priori believe that the models are spanning the full range of uncertainty, or that they are on average distributed around the true climate, such	A brief discussion has been added, but this topic is most appropriately discussed in Chapter 10.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				that all errors cancel out, in particular given the point above that a few models are way off from the rest. Models have many common problems. Although errors will probably cancel in some cases, this is not a priori given. Therefore what I think is missing in the chapter is a more thorough discussion (as far as possible) to which extent and in what diagnostics the multi-model mean is more useful, closer to observations or better in any other respect for projecting climate change, because chapter 10 builds on that. Some figures include a RMS error of the model mean, which is helpful, and this might be extended to other diagnostics. Finally, even if the multi model mean is much better than any single one, it is not a priori clear that the multi model mean of the future will be useful. We are not integrating the mean, but each single model. Errors that cancel today do not necesessarily cancel in the future. Many of these issues are not easily quantified, but should be mentioned in chapter 8 or 10 or both, such that it is clear that we are aware of the limitations of that approach. I do support the use of the multi model approach, and I do believe it is better than any single model, but to be safe, it should be noted that these arguments are largely based on expert judgement and the experience of modellers, and not so much on quantitative studies. There are a few from weather forecasts, but at least I'm not aware of any for climate models. [Reto Knutti]	
8-40	A	0:0		Blind' evaluation vs. diagnostics relevant for future climate: It would be helpful if more emphasis could be given to discuss why the diagnostics shown are relevant to the climate change problem. For some figures, it is not really clear why the quantity shown should matter, and a short explanation of what feedback it tell us about would be helpful. Beyond that, only the mean climate and variability is evaluated. However, the ability of accurately simulating the observed warming over the last century would seem crucial to me for the models to be credible to project into the future. Not a single word is spent on the discussion of whether models capture the trend and time evolution of global mean temperature, the pattern of it, observed vs. modelled vertical temperature trends in the atmosphere, the observed ocean warming trend and pattern, the observed trends in sea ice, trends in sea level, etc. An ocean evaluation would also provide more information when using tracers (CFCs, C14) instead of just temperature and salinity. The AR4 GCMs do not provide that, but other intercomparisons (OCMIP) have studied that in detail. A few figures of the trends are given in chapter 9, but only for surface temperature. In one of the chapters, this needs more emphasis. [Reto Knutti]	Chapter represents current state of knowledge of this issue. Simulation of 20th Century climate variations is discussed in Ch 9 (with a pointer to this in 8.1.2) Ocean tracers: Some material added in 8.3.2 but focus of chapter is on the actual models used in this report.
8-41	A	0:0		Since this chapter is about model evaluation, the Executive Summary highlights findings on model capabilities. However, some of the projections of climate change that have been developed as part of this model evaluation will be of high interest to policymakers and should be included in the Executive Summary. Given the intense interest in tropical	Rejected. This is the purview of Chapter 10.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				cyclones in many parts of the world the conclusion presented on Pg 52, lines 4-6: "There is no agreement among the models whether global warming will make tropical cyclones more or less intense. There seems to be some agreement among models that the frequency of tropical cyclones will be reduced." is an example of the type of informtion that should be included in the Executive Summary [Jeffrey Kueter]	
8-42	A	0:0		Opening Comment: In the Chapters that I am reviewing, I choose to not provide an anonomous review. This choice allows the various Chapter authors to contact me directly on matters of errors, concepts, or questions of disagreement. I have already performed thorough reviews of chapters 1-5. Due to the looming November 4th deadline for reviews, I am choosing to review Chapters 6-11 in a drastically shortened way . Rather than going through all of them as I did before, I am choosing to review only the Executive Summaries of chapters 6-11. There are some clear advantages for this strategy, independent of the obvious one of speeding up the very tedious reading and reviewing process. In the previous chapters I have reviewed, I have seen some significant disconnects between two obviously differering reporting strategies. First, it seems obvious to me that the fundamental purpose of these IPCC FAR reviews is to establish the case, or lack therof, for many of the diverse aspects of the human-caused global warming problem. Second, it is noteworthy that this draft WG1 report is roughly twice as long as the WG1 IPCC TAR report. Third, it seems very obvious that the key IPCC assessment-relevant punchlines are hardly double those of IPCC TAR. It seems clear to me that the global-warming research-advancement doubling time scale is a lot closer to twenty years than it is to five years. The obvious conclusion for me is that we don't really need or desire to double the length of the WG1 chapter assessment every five years! For these nearly obvious reasons, and to help me and the other reviewers refocus on the fundamentally important conclusions that are centrally relevant to the IPCC's human-caused climate assessment's goals, I am thus choosing to reduce drastically my own submitted WG1 reviews. And, most importantly, this gives me a good shot at reviewing meaningfully all of remaining chapters 6-11 by the daunting November 4th reviewers' deadline. [Jerry Mahlman]	Noted
8-43	A	0:0		It was gratifying to see the growth and emergence of a world-wide ethic of co-operation and intercomparison of climate models and climate modellers. This evolving era of co-operation in climate systems modelling is a great testament to the climate modelling community worldwide. Their surprisingly willful sacrifice of personal ego to make possible a world-wide ethical statement concerning what is expected to happen to earth's climate is historic, simply because this new ethic of co-operation is now intrinsically global at a totally unprecedented level. Perhaps the U.S. and the U.N. could learn some	Noted

No.	Batch	Page:line From	Page:line To	Comment	Notes
				invaluable lessons from these ground-breaking climate scientists. These climate scientists, perhaps without thinking about it very much, have "reinvented themselves" to create a new era of scientific co-operation in their global-scale service to the planet. If only our global politicians could be even a fraction as effective. I was impressed by the level of humility that these scientists have brought to attack this intrinsically global problem. Historically, this magnitude of self-sacrificing co-operation has not been thought possible on the global scale, or even within the wealthier countries. If only the politicians who pay a militantly hostile attention to this ground breaking phenomenon and world-wide environmental challenge could comprehend the power of this level of international co-operation in pursuit of service to the planet, they might even learn to empower themselves to address co-operatively their equally daunting challenges in the global political world. [Jerry Mahlman]	
8-44	A	0:0		This is a well-written chapter which, despite the inevitable difficulty of combining sections written by different authors, forms a coherent picture of the current state of climate modelling. [Gill Martin]	Noted
8-45	A	0:0		It is unclear, however, whether the overall aim is to provide an overview of the current state of climate modelling or to describe how modelling has advanced since the TAR. Indeed, the Executive Summary concentrates on the latter while many (but not all) of the sub-sections take the form of the former. Perhaps both aspects should be covered in all sections, including the Executive Summary. [Gill Martin]	Taken into account. However ES needs to focus on new developments in order to be manageable.
8-46	A	0:0		The Chapter is well written and organized. However, it is suggested that issues in the Executive Summary such as the level of natural variability directly related to the climate system on long time scales & the ability of models to accurately simulate natural variability in those long time scales should be address more precisely. It will help policy makers to assimilate the science. [Luis Jose Mata]	Noted.
8-47	A	0:0		The Chapter 8 Figures are excellent compared to those of Chapters 1 and 2; however, some figures (e.g., 8.6.4, 8.6.5) could use improvement. [Lourdes Maurice]	Accepted. Figures 8.6.4 and 8.6.5 (now called 8.6.3 and 8.6.4) have been slightly improved (the resolution of the figures has been increased, more explanations are given on the figures).
8-48	A	0:0		much more detailed evaluation than previously, and the scrutiny of outside workers was a great idea. Sections 8.1 and 8.2 constitute an exceedingly long introduction and should be	Taken into account.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				substantially trimmed. The chapter is much more verbose than need be. I have made a few specific suggestions for trimming it but there are many more opportunities for conciseness. [Philip Mote]	
8-49	A	0:0		Another, more serious issue concerns the time spans used for comparison in the Figures. We have CMIP baseline of 1980-99 (why not 1971-2000?), observed periods 1961-90 (8.3.1-3), an unspecified 40-year mean (8.3.4, 8.3.12, 8.3.14), 1985-89 (8.3.5-8), 1979-93 (8.3.9-11), 1945-89 (8.3.13), no observations (8.3.14-16), 1960-2000 (8.3.17), and FINALLY a matched comparison in Figure 8.3.18. Then it's back to 1961-90 (8.3.19), then CMIP model years 1950-99 compared with unspecified observed years (8.3.20-23)... I could go on. Some effort to standardize periods of comparison, or at least reduce the number of different periods of comparison and ensure similar numbers (e.g. 30 years for both model and obs, even if it's a different 30 years) would be an improvement. [Philip Mote]	The WG1 lead authors suggested that when possible the climatology should be based on the years 1981-2000 (inclusive), but for several models output was only available through 1999, so 1980-1999 was the default period chosen for the models. When possible, observational climatologies for this period will be used in revised versions of the figures. When this is impossible, and the climatology is for a period greater than 20 years, we shall compute the matching climatology from the models. If the observed climatology is based on a period less than 20 years, we shall compare it to the 20 year model climatologies based on years 1980-1999.
8-50	A	0:0		model comparisons should be evaluated objectively, and subjective terms like "well" and "poor" should be avoided in favor of quantitative comparison (e.g., "within 2K") [Philip Mote]	Rejected. Evaluation includes an element of subjective judgement (e.g. feature recognition). This reflects the current state of development of the science.
8-51	A	0:0		a note on semantics: "capture" is often misused in a modeling context, including 24 times in this chapter; possibly it is a corruption of one of the more obscure definitions of capture, "to record or preserve, as in a photograph". Other words that may serve better: replicate, reproduce, simulate. [Philip Mote]	We think this is a fairly widely understood usage.
8-52	A	0:0		The chapter needs better coordination with observational chapters and with chapter 9, with much more cross-referencing, elimination of inconsistencies (examples below), and perhaps a reordering of section 8.4 to match the order in 3.6-3.7 and 9.5. [Philip Mote]	Noted. Too late for the FOD but we will attempt to coodinate in future drafts.
8-53	A	0:0		The paper shows a clear bias towards the view that, fundamentally, climate models are	Taken into account in revision of ES.

No.	Batch	Page:line From	To	Comment	Notes
				OK. When problems are mentioned, the language becomes rather obscure, or at least, indirect. Hence p8-3 line 28 "..precipitation remains elusive" and p8-5 lin22-24 "obtaining a completely accurate…..continues to present a challenge." Why such roundabout language? [Timothy Palmer]	
8-54	A	0:0		The authors have made a great effort to compile a lot of information. Nevertheless, I think the chapter has some serious problems. First, I think it is too long. The authors seem to not have stuck to the tight discipline the CLAs of our chapter kept, throwing out anything not policy relevant and to the point. A lot of this chapter reads like a technical review written for colleagues, and will impenetrable for the non-specialist. It was a hard read even for me. I cite the invitation for review, which says: "the authors are required to work within strict length constraints and must present a concise assessment of current knowledge, not a scientific review of how that knowledge was derived". Much of this chapter is in fact the latter, and in my opinion requires a major rewrite. Section 8.2 is particularly far too technical and too long; it should be possible to summarise the main advances in modeling in much simpler and shorter terms. Section 8.3 is a richly illustrated picture book of a model intercomparison - extremely useful for other modelers like myself, but how many of those pictures are actually needed and useful for a wider audience? Perhaps some can be condensed in multi-panel figures that don't take so much space? And the text could be crisper. Section 8.4 on the other hand suffers from lack of figures, so one cannot get a real feeling for how the variability modes are modeled. [Stefan Rahmstorf]	Length and policy relevant focus: taken into account in revision. 8.2 length/technicality: taken into account in revisor. 8.3 figures: The number of figures has been reduced. Re 8.4 figures: Accept. Section 8.4 has new figures.
8-55	A	0:0		Figures with south on the right are counterintuitive and non-standard (usually a coordinate system runs positive towards the north, with negative latitudes in the SH), and are inconsistent with other chapters. Particularly bad for the MOC graph 8.3.24 where positive stream function maxima now imply counter-clockwise flow against all conventions. [Stefan Rahmstorf]	Taken into account but may not be practical to standardise completely.
8-56	A	0:0		Observed time mean for a number of Figures is calculated for 1961-1990, while the model climatology is calculated for 1980-1999. Why this inconsistency has been introduced and what is an implications of this for the comparison. I guess, that due to rather fast temperature changes during 1990-1999 and different ENSO regime the model's surface air temperature can be biased high. Please, comment on this. [Eugene Rozanov]	Time-means: Noted. Possible SAT bias: We have addressed the inconsisencies between the two climatologies.
8-57	A	0:0		It would interesting to estimate statistical significance of the model's climatology deviation from the observations. It can point to the area where the deviation of the simulated quantities from observed is significant. I think it is much better then to compare	Noted. Egorova et al. computed climatology based on many years and several different observationally based

No.	Batch	Page:line From	Page:line To	Comment	Notes
				the model deviation with the range of the variability of the considered parameter. This approach was applied by Egorova et al., (2005, ACP) and helped to identify some missing processes in the model. [Eugene Rozanov]	products. The standard deviation (S.D.) over years and models provided some measure of observational uncertainty folded in with interannual variability. Including the S.D. over different obs. datasets provides a measure of obs. uncertainty that would be useful. In our figures, we typically have only 1 or 2 obs. datasets (which usually aren't independent), so it does not seem worth doing this for our figures. Also, the uncertainty in the true climatology, arising because it is based on a finite number of years, is much smaller than the model errors almost everywhere, so nearly all differences in climatology turn out to be statistically significant.
8-58	A	0:0		Throughout. I suggest avoid first person plural. Generally ambiguous. The review authors or the general climate scientist public? Examples: Page 8, our present sccientific understanding; better "present scientific understanding" but page 3, line 11, "we focus on areas of progress" is fine. Page 62, line 6-7 is especially awkward: To better weight our confidence in the different model estimates of climate sensitivity, one may apply two kinds of observational tests to climate models. [Stephen E Schwartz]	Taken into account, but the chapter ultimately does represent the judgement of the Lead team.
8-59	A	0:0		Chapter 8 is now generally well balanced and the authors have done a good job in bringing together all of the parts to make the chapter flow without seeming disjointed (and I know how hard this is!). I have not fully reveiwed the whole chapter due to time constraints (sorry), so my detailed comments focus on sections 8.1,8.2,8.31,8.35,8.6 and 8.7. However, I still feeel the balance between variability (9.5 pages) and extremes (3 pages) is wrong. [Catherine Senior]	Reject. The imbalance reflects that the fact that there is much more literature on variability than on extremes.
8-60	A	0:0		Overall the chapter reads well and the authors should be congratulated [Keith Williams]	Noted
8-61	A	0:0		Although more of the chapter has been devoted to evaluating aspects of climate variability, extremes and climate sensitivity compared with previous IPCC reports, there is still too much of the chapter devoted to evaluating mean climate. There is growing	Accepted. The number of figures in Section 8.3 has been reduced, and the balance is improved as a result.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				evidence (mentioned in several places the chapter - particularly section 8.6) that evaluation of processes occurring as part of changes in the current climate (e.g. through climate variability) may be of more relevance to climate change prediction than being able to simulate the mean climate correctly. This should be reflected in the structure and allocation of space within the chapter. [Keith Williams]	
8-62	A	0:0		Considerably more of the chapter is spent evaluating modes of variability compared with extremes. This may reflect the volume of science which has been undertaken on the two subjects. If that is the case, it would be useful to have it stated in order to point out that more scientific effort is required on evaluating extremes in the future. [Keith Williams]	The imbalance between variability and extremes reflects that there is much more literature on the former.
8-63	A	0:0		This chapter does not seem to cover the evaluation of simulated trends and variability in ocean heat content and freshwater storage as well as the hydrological cycle. I would thingk these should be included to assess coupled model's capability of simulating future climate change and feedbacks. [Peili Wu]	These are discussed in Chapter 9
8-64	A	1:0	1:	suggestion: please add two Contributing Authors who are Dr. Y Xu a key expert of BCC-CM1, Dr.Y Yu a key expert of FGOALS-g1.0 [Zong-Ci Zhao]	Rejected. Drs. Xu and Yu have not contributed in the sense of CAs..
8-65	A	1:1	1:1	In the discussion about regional trends, there was no reference to Douglass et al. (2004) who showed that current climate models do not capture the zonal trends in atmospheric temperatures measured with MSU and with radiosondes. This tpye of reference and discussion are need for balance and a more complete assessment. Douglass et al., Geophys. Res. Lett., 31, L13208, doi:10.1029/2004GL020103 [Patrick Minnis]	Comment passed to Ch 9
8-66	A	1:9	1:19	The list of authors should be given with the full name of the people, as in Chapter 1 and without their affiliation. [Philippe Tulkens]	Will be dealt with as an editorial issue
8-67	A	1:14	1:14	Please update my name as contributing author; (Albert A.M. Holtslag rather than Bert Holtslag) [Albert A. M. Holtslag]	Accepted
8-68	A	1:31		Section "8.11" should be 8.1 [Philip Mote]	Accepted
8-69	A	3:0	6:	executive summary could be shortened by reducing or eliminating redundancy. For example, climate drift is mentioned both page 3, lines 18-19 and page 4, lines 50-52. Effectively the list on page 3 is like an executive summary within an executive summary. Better to eliminate the list on page 3 and shorten the rest of the summary. The word	Accepted. ES shortened.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				"however" is overused, especially to begin sentences - 10 instances in the executive summary alone. [Philip Mote]	
8-70	A	3:1	7:5	It is proposed to avoid repetitions in the executive summary in order to keep the text as short and informative as possible. [Klaus Radunsky]	Accepted. ES shortened
8-71	A	3:1		Overall, I found the organizational structure of the executive summary to work well. I was surprised that no bulleted entry was devoted to the fact that the number of models, modeling centers, and nations with modeling centers represented in AR4 is significantly greater than it was in the TAR. That consensus is being built from a larger and more diverse modeling base seems something that might qualify as a highlight to those who read the executive summary. [Keith Dixon]	Accept. Add quantitative bullet
8-72	A	3:1		Exectutive Summary. This is well balanced and easy to read. I think the section on feedbacks should be labelled as such and not lumped under 'developments in analysis techniques' [Catherine Senior]	Accept . 'Developments in analysis methods to include just 1st and last paras. Separeate section on feedbacks
8-73	A	3:1		Section executive summary. The executive summary extends over more than 5 pages, this might be too long. There are some repetitions and in general, too much explanation is given on the findings. The next version could concentrate more on the findings and leave their justification to the body of the text. [Philippe Tulkens]	Shortened
8-74	A	3:3	3:3	The goal of the chapter as stated here is too narrow compared with the approved chapter outline and the additional notes made by the governments on the outline. Notable points made by the governments include: -Model evaluation will draw heavily on comparisons with observations but also what can be learned from model intercomparisons. -Consider range of information beyond mean climate parameters. The chapter seem to miss the opportunity to evaluate biogeochemical models, to evaluate models with respect to ventilation time scales as evidenced by transient tracers, ocean heat uptake data, or data on the transient evolution of the climate-biogeochemical system. [Fortunat Joos]	The two points requested by governments are indeed considered. Re transient tracers: Noted. Oceanic tracers were not included in the specifications for the models in the PCMDI database. This chapter mainly evaluates those model integrations. A sentence was added noting the advantages of including oceanic tracers. Transient heat uptake is discussed in Ch 9 Reject the evaluation of carbon cycle in terms of terrestrial processes (see 8-33)

No.	Batch	Page:line From	To	Comment	Notes
8-75	A	3:3		might want to highlight what you mean by "climate models" - for example, do you evaluate those of intermediate complexity, those w/ ice sheet dynamics, interactive veg or biogeochem? Where you don't evaluate a component of earth models, best say where in the report readers should go - e.g., for ice sheet dynamic modeling [Jonathan Overpeck]	Reject. See chapter intro and roadmap
8-76	A	3:4	3:5	There is no evidence for this statement., No model has ever been tested successfully against its future presdiction. There is, therefore, no basis for “confidence that it can be done. [Vincent Gray]	Disagree. Decades ago, climate models predicted warming in the late twentieth century, strongest near the poles, and this has been observed. As a second example, a climate model predicted the cooling due to the Pinatubo before it occurred. The text has been modified to make these points.
8-77	A	3:4	3:6	I agree with this assertion that climate modelling skill has indeed advanced notably since the TAR. Much of this advance has come from the global-scale enhancement and diagnosis of the growing suite of highly capable climate models worldwide. [Jerry Mahlman]	Noted
8-78	A	3:8	3:9	Again, there is no evidence for this statement. No model has ever successfully predicted a future climate change. You only talk about the “long term” to get you out of the task of actually testing it. Your “confidence” is misplaced . [Vincent Gray]	Reject. The chapter lays out the evidence.
8-79	A	3:8	3:10	The growing skill of climate modelling "at larger scales" is a very important advance. It even includes some humility in its stated recognition that regional and sub-regional scales remain challenging, particularly in regions with high, and rough, topography. [Jerry Mahlman]	Noted
8-80	A	3:8	3:11	Given the very large range of possible climate sensitivities now quoted in Chapter 10, even larger than in past reports, what does it mean to say that "models are reliable enough to provide useful projections?" What is a useful projection? [Peter Stone]	Climate models are based on accepted physical principles. They can reproduce many observed features of current climate and past climate changes. They can be and have been used to make physically-based projections of future climate change, particularly at larger scales. In this summary we focus on areas of progress since the TAR. Taken into account in redrafting

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-81	A	3:9	3:11	You do not include in your list an actual test for the capability o future prediction. See my paper on the subject (Gray. V R 1998 “The IPCC future projections : are they plausible Climate Research 10 155-162) which shows that they do not work [Vincent Gray]	See response to 8-76
8-82	A	3:11		Insert after “changes”, “They have, so far, not been shown capable of confirmed future projections” [Vincent Gray]	See response to 8-76
8-83	A	3:13	3:50	It looks like a collections of sentences. Perhaps, it will look better if organized by topics such as development in model formulation, in model climate simulation, etc [Luis Jose Mata]	Taken into account
8-84	A	3:13	3:50	Of the three chapters I have looked at, chapter 8 seems to me to be the clearest, and easiest to read. I was pleased that a section (#8.8) was devoted to the use of EMICs and simpler models consdering their importance in results presented in other chapters in studies of past and future climates. However, despite this and their extended use since TAR, EMICs and simpler models have been omitted from "highlights since the TAR" in the Executive Summary. I think there should be one bullet point which states that these models are being used more and more to produce some important policy-relevant results. [David Sexton]	Taken into account
8-85	A	3:13		Two of the key highlights since the TAR (mentioned later in the chapter, but missing in the bullets) are (i) that many models have in the meantime been tested on past, different climates and climate evolution, e.g., Eemian interglacial, LGM, last millennium, etc. - see chapter 6, and (ii) the coming of age of a variety of EMICS as useful tools for long runs and large ensembles. [Stefan Rahmstorf]	(i) is a Ch 6 issue (ii) Taken into account
8-86	A	3:14	3:15	These model physics improvements since the TAR have indeed been impressive. It seems that each new generation of model building has taken the step to improve the incorporation of new and better physical processes. [Jerry Mahlman]	Noted
8-87	A	3:16	3:19	This statement needs a gentle dose of "modelling humility". Indeed, it is true that the use of empirical "flux adjustments" has declined significantly in response to the realization that such adjustments do constrain the models to what we currently is our "best guess" of what is current reality. The alternative to flux adjustments, however, remains conceptually problematic, and is not discussed meaningfully here. No thanks to the enormous heat capacity of the world ocean, we still struggle to find the most acceptible ways to initialize the climate system for global warming studies in a way that is a good facsimile of the real world. We can choose ways to initialize climate models that constrain the upper ocean, but constrain the deeper layers to not participate meaningfully,	Taken into account but space precludes a detailed discussion in ES – see 8.2.7 and 8.3.2 for this.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				or we can do our best to describe the current ocean-atmosphere balance and then hang on and hope that the climate will not drift away from its initial state too quickly. Another strategy is to run the climate model for a few thousand years for the current(or possibly, pre-industrial) climate, at considerable cost and hope that it will equilibrate to a climate that is very similar to today's(or pre-industrial) climate. This is, in my experience, much easier said than done. If, indeed, new initialization techniques have been invented that solves this very challenging problem, such a "breakthrough" should be highlighted prominently in the Executive Summary. I do suspect, however, that the IPCC modelling community has virtually been forced to simply initialize the global upper ocean at its current state and then hang on through the model integrations for a hundred years or two, and hope that the poorly initialized deeper layers, say, below 1000 meters, do not come back to "bite" the modelers before very long-term drift sets in. I thus recommend that this discussion be made clearer for the Executive Summary readers. [Jerry Mahlman]	
8-88	A	3:18	3:19	Here it is mentioned that climate drift remains an issue; one would thus expect to find a further ellaboration on this topic in the subsequent sections. However, it is not mentioned again in any of the sections of this chapter. See comment #15. [Marisa Montoya]	Taken into account but space precludes a detailed discussion in ES – see 8.2.7 and 8.3.2 for this.
8-89	A	3:18		can you say more about the relevance (to AR4 results and society) of this drift? How much does it matter. [Jonathan Overpeck]	Taken into account but space precludes a detailed discussion in ES – see 8.2.7 and 8.3.2 for this.
8-90	A	3:18		It should be noted that climate drift is a problem only for models that are not run into an equilibrium, hence not for EMICS. These are generally equilibrated and do not drift any more. [Stefan Rahmstorf]	Noted
8-91	A	3:23	3:24	"At least some of …" : this sentence is redundant. It conveys the same information as the previous sentence. [Ileana Bladé]	Accepted
8-92	A	3:25	28:	this one is too vague? Seems it (and many other of your bullets) could be used to say "in Chap 8, the IPCC says their models deficient in many ways, so why should we have any faith in Cap 10 projections?" Need to give the significance of what models do well, and not so well. So what if the models can't simulate major modes? Sounds bad to me - even thought I know better, many readers (policy makers) might not. Can you say something explicitly about drought and how well our models simulate it? Note the recent papers that CAN simulate 20th century drought (e.g., 1930's and 50's of US) with perscribed SST's, but NOT with coupled models. [Jonathan Overpeck]	Accept first part, and will reword in ES. Reject suggestion re droughts, because this is discussed in Chapter 9.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-93	A	3:29	3:34	It is good to get better bounds on the equilibrium climate sensitivity, although most assessment scientists are more focussed on measures of the transient climate sensitivity, e.g., at the time of CO2 doubling. Clearly, we are doing a rather good job of that, but we still seem to be lacking on the bounds of the higher-side sensitivity. Maybe this is because we tend to "over-believe" the ill-constrained PDFs on the high-sensitivity side. At any rate, this discussion on climate sensitivity can use some clarification that speaks more clearly to the goals of FAR. [Jerry Mahlman]	Noted. This is done in Chapter 10.
8-94	A	3:29	33:	even your average smart college kid knows models have cloud issues - please say more - do current models treat clouds better than before, or? Just saying it is an issue is too vague. What is the significance of this bullet? Models better or not? [Jonathan Overpeck]	Discussed later in ES.
8-95	A	3:31	3:31	New "observational" evidence ? [Ileana Bladé]	Accepted: clarify by changing to 'new observational and modelling evidence'
8-96	A	3:35	3:39	8-3, lines 35 – 39, on the Atlantic overturning, it might be worth mentioning that while there is a large spread on its response, no model produces an increasing overturning in response to global warming. [Wenju Cai]	Discussed in Ch 10
8-97	A	3:35	3:40	This is a nice, useful, and balanced discussion. [Jerry Mahlman]	Noted
8-98	A	3:37		how big a range of possible MOC responses are possible? What about what we have learned from simulating past (paleo) observed abrupt changes in MOC? Best coordinate w/ Bette Otto-Bliesner, David Rind, Stafan Rahmstorf and Dick Peltier of Chap 6 ("Chap 6 model team" refered to below). [Jonathan Overpeck]	Discussed in Ch 6 and 10
8-99	A	3:38		Start new bullet for SO bias, it seems a point not related to the first part. [Stefan Rahmstorf]	Text changed
8-100	A	3:41	3:43	I agree that the use of multi-model / multi-ensemble integrations is a very valuable tool that should be even more useful by the time of the Fifth Assessment Report. [Jerry Mahlman]	Noted
8-101	A	3:42	3:43	It would be interesting to learn more about the requirements for a proven model metric and what the remaining gaps are. [Klaus Radunsky]	See discussion later in ES and chapter body.
8-102	A	3:42		"the goal of a proven model metric" - the meaning of this is unclear even to me, I suspect the general reader will not know what you mean here. [Stefan Rahmstorf]	Text modified

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-103	A	3:44	3:47	This needs clarification. Obviously, the planetary-scale modes have predictability beyond the ENSO scale, and ocean-scale anomolies can have decades of predictibility. Whether or not that predictibility is useful in some sense is less obvious. [Jerry Mahlman]	Noted. Text modified.
8-104	A	3:48	3:49	Replace "in the main climate pojections" by " in most climate projections" [Michel Petit]	Accept
8-105	A	3:48		what about beyond the "next few decades"? What about methane clathrates or release of major carbon from high northern hemisphere terrestrial soils? These could be large, no? [Jonathan Overpeck]	Accept – text modified.
8-106	A	3:49	3:50	Chapter 7 stresses the importance of feedback mechanism. For the next few decades, the impacts of these feedbacks may not be very significant. However, it is likely to increase. Therefore, the executive summary of Chapter 8 could mention (on line 50) that although the impact of the feedbacks may not be large in the next few decades, it needs to be accounted for in longer term predictions (of about a century?) as emphasized in Chapter 7 and elsewhere in the executive summary of Chapter 8 (p. 8-4 L. 23 to 28). [Philippe Tulkens]	Accept – text modified to avoid touching on projection
8-107	A	3:52	4:9	The "Developments in model formulation" section does a good job of hitting the main points. Since this part of the executive summary, it might be useful to specifically mention that these developments are the outgrowth of continuing advancement in a) our understanding of the climate system, b) our ability to translate that understanding into computer model codes, and c) access to bigger and faster computer resources. Somewhere in this section the connection should be made between increased computational resources and the ability to run models at higher resolution, larger ensembles, and with more complex representations of more climate system components. The phrase "improved computation strategies" (P6, line 7) could be incorrectly construed as referring only to clever experimental designs that were independent of computational resources (i.e., doing more with a fixed resource). [Keith Dixon]	Rejected. Goes beyond focus of chapter.
8-108	A	3:52	7:5	It should look more appealing to the general reading in particular to policy makers [Luis Jose Mata]	Taken into account
8-109	A	3:52		Is this a subheading of the Exec summary? [Stefan Rahmstorf]	Yes
8-110	A	4:1	4:9	I suggest to cut the sentence in the bullet because don't improve the clarity of the sub-section [vincenzo artale]	Text modified
8-111	A	4:4	4:9	This list is supposed to be a list of changes in climate model formulations and methods of evaluation. Items 4 (better simulations) and 9 (improved understanding) don't seem to	Text modified

No.	Batch	Page:line From	Page:line To	Comment	Notes
				belong here. [Ileana Bladé]	
8-112	A	4:10		You left out "International Co-operation in Climate Model Development", a major achievement, given the previous barriers. Now that the U.S. Earth Systems Modelling Framework has gone international, it clearly deserves a mention. [Jerry Mahlman]	Text modified
8-113	A	4:11	4:11	Suggest to replace semi-Lagrangian by „improved transport numerics“ the specific type of transport scheme is probably not the point here. [Rolf Müller]	Text modified
8-114	A	4:11		I am not at all convinced that Semi-Lagrangian Advection is an improvement when incorporated into climate models. On longer time scales, it is quite diffusive, thus failing to reproduce the conservation properties that are intrinsic to the 3-D advection process. Indeed, the longer the time scales of importance for advected quantities, the worse it gets. It is thus fine for shorter-term weather forecasts, but that is not what is at stake here. [Jerry Mahlman]	Text modified
8-115	A	4:13	4:14	aerosol effects are "now more widely simulated" ?? Does that mean more people simulate them or their affects are apparent over a wider model domain? [Anne Douglass]	Text modified
8-116	A	4:18	4:19	This sentence is confusing and should be rewritten, specifically, what type of significant development occurred in modeling if the best model in TAR already represented these processes adequately. [Robert Molinari]	Major revisions in the Executive Summary has modifed this text
8-117	A	4:19	4:19	As there is no metric to identify which model is the "best", I suggest replacing 3best" with "most comprehensive" . [Philippe Tulkens]	Major revisions in the Executive Summary has modifed this text
8-118	A	4:21	4:22	The double negative construction is awkward. Simpler would be better " ..good enough that we think they capture the main processes …." [Ileana Bladé]	Accept – text rewritten
8-119	A	4:21	4:22	eliminate double negative and royal pronoun in sentence. Suggested wording is "Representation of terrestrial processes in climate models has improved in recent years, and the models should be able to simulate large-scale climate changes over the next few decades." [Chuck Hakkarinen]	Accept – text rewritten
8-120	A	4:21	4:22	is generally... cannot capture" - this is a flaccid way of saying "is adequate to simulate [Philip Mote]	Accept – text rewritten
8-121	A	4:21	4:22	What is the evidence for saying that the representation of terrestrial processes in climate	Accept – text rewritten

No.	Batch	Page:line From	Page:line To	Comment	Notes
				models is generally good enough that we have no reason to think that they cannot capture the main processes…? [Catherine Senior]	
8-122	A	4:27	4:27	"varies from small to significantly positive" seems to contradict earlier statement in Page 3, line 50 (positively and relatively small). [Ileana Bladé]	Accept – text rewritten
8-123	A	4:32		Positive definite advection schemes will be invaluable when they can also conserve variance, and accuracy. I know of no such schemes at this time. Also, are "adiabatic isopycnal" mixing schemes a near oxymoron? [Jerry Mahlman]	Noted.
8-124	A	4:34	4:34	"rigid lid - virtual salt flux" is a somewhat obscure description and I would have thought it more properly described as a simplified boundary condition rather than a parameterization? [Timothy Johns]	Agree. Text changed.
8-125	A	4:34		the em dash seems to be shorthand for a conjunction - perhaps “and”? [Philip Mote]	Agree. Text changed.
8-126	A	4:42		the line about terrestrial snow processes might belong more under the land surface processes paragraph (lines 18ff), as this is where it actually appears in the chapter (pages 18-20) [Philip Mote]	Accept – text rewritten
8-127	A	4:43	4:43	"snow ripening" may not be a commonly-understood phenomenological term - explain? [Timothy Johns]	Reject. No space to expand.
8-128	A	4:45	4:45	capturing: reproducing ? [Ileana Bladé]	Accept – text rewritten
8-129	A	4:47		article missing “only a few” [Philip Mote]	Accept – text rewritten
8-130	A	4:50	4:52	This is a very important statement. Clearly, the simulated climates still drift without controls. The question, however, is whether they do not drift meaningfully on the time scales of a particular climate models' integration? [Jerry Mahlman]	Text modified
8-131	A	4:50	4:51	These sentences are identical to lines 16 to 19 on page 3, this repetition could be avoided [Philippe Tulkens]	Text modified
8-132	A	4:52		This is repeated here from above. [Stefan Rahmstorf]	Text modified
8-133	A	5:1	5:2	It is not clear how the lack qualitative changes since TAR in projections of ocean changes increases confidence. The models still can have significant problems and robustness can	Text deleted

No.	Batch	Page:line From	Page:line To	Comment	Notes
				only be determined through favorable comparisons with observations. [Robert Molinari]	
8-134	A	5:1	5:1	"Development of the new AR4 models". Have some models been developed for the AR4 ? I thought that model developments take place in different contexts and that the results of some model simulations are used for AR4. Model intercomparison projects also take place but that is different from developing new models for AR4. [Philippe Tulkens]	Rejected. The term 'AR4 models' is being used to indicate the models that are primarily used in AR4 for detection/attribution and projections (especially in Chapters 9 and 10).
8-135	A	5:3	5:5	This is an encouraging discussion. [Jerry Mahlman]	Noted
8-136	A	5:3	5:5	The importance of the Southern Ocean for heat uptake is a model result. In the observations there is more heat uptake in the Northern Hemisphere around 20 to 40 N (Levitus et al., 2005) . No model gives this pattern. In addition the models are overestimating the rate at which heat is being mixed into the deep ocean (Forest et al., 2005). The models are in fact doing a poor job of simulating the heat uptake, and the uncertainty in the transient response is much greater than is implied here. [Peter Stone]	Noted. There is quite a bit of uncertainty associated with the Levitus data sets. The sentence is reworded. Simulation of historical changes in heat content is discussed in Chapter 9.
8-137	A	5:8	5:8	"over thick" or "overly thick"? [Ileana Bladé]	Agree. Text reworded.
8-138	A	5:15	5:15	"simulations" or "components" ? [Ileana Bladé]	Rejected. Could be "components", but "simulations" is better here. While it is difficult to separate improvements in components from improvements in simultations, it is simulation of driving fields (winds, currents, etc.) which is central in this context.
8-139	A	5:18		Terminology: so far we talk about AOGCM, now suddenly "coupled GCMs" - is this meant to be synonymous, and how would the general reader know? [Stefan Rahmstorf]	Change to AOGCM
8-140	A	5:23	5:25	8-5, lines 23 – 25, I would reverse the sentence: Change to: In the tropics, development in model formulation since….. led to improvements in the simulation of ENSO and MJO in the amplitude, structure….. persist. Obtaining a completely representation of ….. Challenge. [Wenju Cai]	Reject. We prefer the present wording.
8-141	A	5:28	5:30	This is an important discussion. It seems clear that this should be regarded as an expected result. It results from the fact that convecton that produces precipitation occurs typically on spatial scales well below most atmospheric models. Indeed, the effect even shows up in mesoscale weather forecast model skill(or lack therof). At the rain-guage level, this is	Added reference to paper entitled "How Often Does It Rain?" by Sun et al., J. Climate in press.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				an important effect. At the AOGCM scale, it is what we have to live with. Downscaling techniques probably would be useful for addressing this problem. Or, it could be clarified through use of mesoscale subregions in higher-resolutions climate models. [Jerry Mahlman]	
8-142	A	5:28	5:30	The statement about extremes needs more information. Are temperature extremes better simulated since the TAR? Why are precip extremes not well simulated, is it resoultion? [Catherine Senior]	Accepted.
8-143	A	5:38	5:39	This seems like a no-brainer to me. [Jerry Mahlman]	Noted.
8-144	A	5:38	5:39	this warning about EMICs is puzzling - this is not a temptation most readers would face. [Philip Mote]	Noted.
8-145	A	5:39		It would not be sensible…" This seems to be a (superfluous) piece of advice to EMIC users rather than a policy-relevant summary of knowledge statement.What about: "Because of their reduced resolution …. EMICs only allow inferences about very large scales. [Stefan Rahmstorf]	Accepted
8-146	A	5:41	6:41	most of this section is results, not methods. [Philip Mote]	Text modified
8-147	A	5:43	5:49	This is an inspiring achievement, well worthy of note by real global-warming policymakers. [Jerry Mahlman]	Noted
8-148	A	5:46	8:50	Enumerating the numerous benefits of the PCMDI/IPCC WG-1 data archive is a good thing to do in the executive summary. However, to list the benefits without alluding to the substantial personnel and computation costs (at PCMDI and the individual modeling centers) could leave readers wih the incorrect impression that this was done with little cost or impact on the community. Acknowledging in the executive summary that costs were incurred (as is done nicely in section 8.1.2.2) could help make the point that this effort would benefit from increased resource support in the future. [Keith Dixon]	Rejected. IPCC cannot lobby for research
8-149	A	5:48		delete the words: "gross modeling" [Stefan Rahmstorf]	Accepted
8-150	A	5:52	6:3	This is a very important achievement. It produces major clarification of one of the most fundamental processes in the climate system. [Jerry Mahlman]	Noted
8-151	A	5:53	5:54	"anti-correlation between them" is unclear. Does "them" refer to "models"? To "water vapor and lapse rates"?	Accepted: Reworded to remove this ambiguity: changed "between them" to

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Chuck Hakkarinen]	"between these feedbacks"
8-152	A	5:56	5:57	quantify this statement that the "model is consistent with the range of uncertainty of the observations" by stating what IS the range of uncertainty in the observations. [Chuck Hakkarinen]	Noted: It is impossible to summarise these in the ES due to lack of space. These are discussed in the section for each particular observational test. Sentence removed from ES for space constraints.
8-153	A	5:57	6:3	state this result more forcefully: something like “Furthermore, most evidence confirms the model result that in a changing climate the relative humidity will not change.” [Philip Mote]	Noted: this sentence was removed due to space constraints in the ES, and degree of duplication.
8-154	A	6:5	6:14	This is an eloquent statement concerning our continuing vexing difficulties in quantifying the role of cloud-radiation feedbacks in determining the "true" climate sensitivity. I suspect that we all may be forced to learn the answer empirically. [Jerry Mahlman]	Noted
8-155	A	6:7	6:8	I'm not sure I agree that we can categorically say that tropical low clouds are the main contributor to intermodel differences in global cloud feedbacks. The paper by Bony and Dufresne, 2005 shows that tropical low clouds are the main contributor to tropical cloud feedbacks. Webb et al, 05 suggest that low cloud increases in a feedback class with clouds from both the tropics and mid-latitudes are highly correlated with global cloud feedback. The important point is really whether it is a particular cloud *process* that is important rather than its geographical location. [Catherine Senior]	Accepted
8-156	A	6:7	6:7	The wording implies that processes governing changes in tropical low cloud are more uncertain than those governing mid-latitude low cloud. This has not been demonstrated. I suggest removing the word 'tropical' (Also see comment 6 below). [Keith Williams]	Accepted
8-157	A	6:8	6:9	"eastern tropical oceans" is unclear -- does this mean eastern sides of tropical atlantic, tropical pacific and tropical indian oceans? What is the nature of the "reason for concern"? What are the specific implications of poor cloud simulation by models in these geographic regions? [Chuck Hakkarinen]	Text modified
8-158	A	6:36	6:41	This extended deterministic skill, of course, is limited to the very largestspace and time scale, but is interesting, nevertheless. [Jerry Mahlman]	Noted
8-159	A	6:43	7:5	The reliability of climate projections needs to address the ability to model climate forcings, feedbacks, and time constants. There are two aspects of climate forcings, i.e., accurate modeling of the radiative effects for prescribed changes in atmospheric	Noted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				constituents, and accurate estimation of future changes in radiative forcing agents. Feedbacks encompass all of the physical processes and interactions that occur in the climate system in response to the applied radiative forcings. Many are well understood and are being modeled with good accuracy. Others are poorly understood and may be difficult or too costly to include in GCM parameterizations. Accurate rendering of climate response time depends on accurate modeling of heat exchange between the atmosphere and deep ocean. The reliability of any climate projection depends on the demonstrated reliability of each contributor. So, any 'model metric' should be addressing the individual reliability of the contributing components as well as their performance as a whole. One should be cautious about "canceling errors" that yield a "good" result. [Andrew Lacis]	
8-160	A	6:43	7:5	This section could go at the start of the Executive summary, before the Highlights (e.g. page 3, between lines 6 and 8). Otherwise the reader is left unclear as to whether the model evaluation has improved the reliability of climate predictions, whereas we want to reassure him/her that the model improvements have indeed increased our confidence. [Gill Martin]	Text has been deleted for space reasons
8-161	A	6:43		Replace Heading with “Reliability of climate projections” [Vincent Gray]	Text has been deleted for space reasons
8-162	A	6:45	6:56	This is a very nice discussion. It is clear and it is honest. [Jerry Mahlman]	Text has been deleted for space reasons
8-163	A	6:45		don't understand the "robust model metric" - is this really required here, not just something for technical talk amongst modelers? [Stefan Rahmstorf]	Text has been deleted for space reasons
8-164	A	6:46	6:47	A few studies show that available observational tests potentially have value . . .given the large international investment in obtaining such observations, and the large pressure to continue/expand the observational base, this seems like a weak statement. [Anne Douglass]	The statement reflects our assessment of the state of the science. Text has been modified
8-165	A	6:54	6:55	add also ocean water masses characteristics in the analysis suggested [vincenzo artale]	Text has been shortened- comment now irrelevant
8-166	A	7:1	7:5	Delete this paragraph. It says that currently the evaluation of climate models is based on expert judgement. This is clear from the rpeceding paragraph, where the lack of a metric to measure model reliability is described. The strengths and weakness of climate models are well described in earlier parts of the executive summary, so this final statement adds no new information. [Lenny Bernstein]	Text deleted
8-167	A	7:1	7:1	Quantitative light??? [Rolf Müller]	Text has been deleted

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-168	A	7:6		What about models ability to simulate abrupt change dynamics? OR ice sheet dynamics, and hence the potentially major contributions of ice sheets to future sea level rise. I think you might be able to refer elsewhere - e.g., chapters 6 and 4, but chap 8 should perhaps at least provide a short summary, so that it stands as the place in the AR4 to start with when worrying about the realism of AR4 models. [Jonathan Overpeck]	Material from Ch 6,8,9 will be drawn together in the TS
8-169	A	8:0		section 8.1: Aspects of this section sound too much like a tutorial. I suggest sharply trimming it [Philip Mote]	Accepted. The text has been shortened.
8-170	A	8:1	1:1	Please, delete the word Philosophy [Luis Jose Mata]	Accepted
8-171	A	8:8	8:19	I suggest to cut for the same motivation above [vincenzo artale]	To be shortened
8-172	A	8:8	8:8	a" ." Should be deleted. [Philippe Tulkens]	Accepted
8-173	A	8:8	12:	Is there really (p8-12 line 8) “considerable confidence that models are reliable enough”? It bothers me that there is little discussion in the text of the fundamental problem with the current formulation of current climate models - that the bulk-formula parametrisation approach to representation of sub-grid processes are making the models very dissipative, much too dissipative in my view, if we look at the energy spectra in models and compare with observations, eg Nastrom and Gage. It is not inconceivable that if sub-grid parametrisations were much less dissipative then: 1) Some of the fundamental long-standing systematic biases in models, typically an over-emphasis of the dominant westerly regimes in the atmosphere may be reduced. 2) Some of the poorly-simulated phenomena such as the MJO, ENSO, the structure of the ITCZs etc would be improved 3) Estimates of internal variability of the atmosphere would be seen to have been underestimated in current versions of climate models. Points 1) and 2) could lead to rather different responses to doubling CO2. Point 3), as discussed below, could impact substantially on detection/attribution studies. There is a whole class of parametrisations now under development which break this bulk-formula approach to the respresentation of sub-grid processes. Some of these issues are discussed in Palmer, T.N., 2001: A nonlinear dynamical perspective on model error: a proposal for nonlocal stochastic-dynamic parametrisation in weather and climate prediction models. Q.J.R.Meteorol.Soc., 127, 685-708.	Noted. There has been some modification to the text, however the overall level of confidence is unchanged. Space limitations in the common question preclude discussion of dissipative issues in particular. However, text has been modified to be more explicit on strengths and weaknesses of models in modes of variability etc (also covered in responses to comments 8-224ff) modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Timothy Palmer]	
8-174	A	8:23	8:31	this is not how I would define evaluation. Results are rarely so clear-cut as to be "right or wrong". Evaluation compares models with observed properties of climate, either of state variables or of derived quantities like OLR, in a quantitative and objective fashion. Saying that a model is "right" or does a "good job" is a subjective evaluation. Also, the comparison with weather forecasting is a needless distraction, or at least carelessly executed; a better point would be that climate is about statistics of weather, and the goal of climate modeling is to reproduce means and higher moments of important observed variables. The goal of evaluation is to determine the extent to which models, singly and collectively, reproduce those statistics. [Philip Mote]	Text has been modified.
8-175	A	8:33		Section 8.1.2: This section seems to short and a bit "handwavy" given that a good review and balanced review of the methods is a key pre-requisite to enable understanding and strengths and weaknesses of the evaluations in subsequent sections of the chapter. In particular, there is one line of approach to model testing and evaluation which is not treated: it is published in papers such as Goody et al., BAMS, vol 79, no 11, Nov 1998, 2541-2549; Haskins et al., J. Climate, vol 12, May 1999, 1409-1422, Goody et al, BAMS, vol 83, 873-878, 2002. None of these papers is quoted/cited and the relevance of systematic model testing by climate benchmark measurements is also not adequately treated (see, e.g., Goody et al. 2002). [Gottfried Kirchengast]	The text has been rewritten.
8-176	A	8:33		Section 8.1.2 should be trimmed to about half a page. Seems to me that the essential points are (1) the importance of both component-level and system-level evaluation; (2) the complication of using 20th century climate, with its transient aspect, to compare against "control" runs, and the related issue of constraining the climate sensitivity (a subject whose details are best left to chapters 9 and 10). These subject could be discussed much more succinctly, and other aspects like using climate models for weather forecasting could be omitted. [Philip Mote]	The text has been modified.
8-177	A	8:39	8:41	Please use an analogy from the physical sciences -- the airplane analogy is colloquial. [William Collins]	The text has been modified.
8-178	A	8:39	8:41	Aircraft design practices are fairly mature. Comparing climate models o aircraft models is misguiding and a reader might infer a comparable level of fidelity, which is of course inaccurate. [Lourdes Maurice]	The text has been modified.
8-179	A	8:44	8:45	Probably this paper should be cited: Trigo R.M., Garcia-Herrera R., Diaz J., Trigo I.F. and	Rejected. This comment appears to be

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Valente M.A. (2005). 'How Exceptional Was the Early August 2003 Heatwave in France?'. Geophysical Research Letters, 32(10), Art. No. L10701. In this paper a very complete analysis was presented estimating the effects of the 2003 summer heat wave over mortality in Europe, the mechanisms causing this situation and comparing the measured temperature values with values from the previous 500 years. [Pedro Ribera]	misplaced.
8-180	A	8:55	8:55	delete one 'in their' [Reto Knutti]	Accepted
8-181	A	8:55	8:55	Two "in their" [Eugene Rozanov]	Accepted
8-182	A	9:10	9:10	I realise it is contentious, but I would replace the word 'process' with 'mechanism' here. Given that you are differentiating between component and system level evaluation, I think 'mechanism' is more appropriate for what is essentially a physically based system-level assessment. [Catherine Senior]	Accepted.
8-183	A	9:23	9:24	"sensitive to the inclusion of a .. process … which is absent" ? Unclear. "Sensitive to whether a particular process or feedback is absent in …. " ? [Ileana Bladé]	Accepted
8-184	A	9:46	9:53	The use of different aerosol forcings with models of different sensitivities to reproduce historical temperature records is an important qualfication on those studies. This paragraph treats the issue well. Similarly, the summary on the benefits and limits of using NWP to learn about climate models are clearly stated on p. 8-10 (10-15). [Leo Donner]	Noted.
8-185	A	9:47		Since the climate forcing, particularly the aerosol forcing, is not perfectly known over that period (Chapter 2), such tests [model-observtion comparisons] cannot be regarded as unambiguous. This statement is more than a bit disingenuous and inconsistent with statements elsewhere in this review. Aerosol forcing is in fact quite uncertain in the context of the total forcing; it therefore makes the total forcing all the more uncertain, perhaps by a factor of several fold. For example, Chapter 9 Page 60, line 51 a high sensitivity cannot be ruled out because it is possible that a high aerosol forcing could nearly cancel greenhouse gas forcing. Moreover the statement implies that other forcing such as GH gas forcing is "perfectly" known, whereas in fact there is substantial uncertainty with that as well Chapter 2 noted (page 2-12) a substantial range of forcing associated with doubled CO2 in different models:	Accepted.Text clarified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				A recent comparison of line-by-line and GCM radiation schemes found that clear sky instantaneous RF and surface forcing agreed very well (better than 10%) among the 5 line-by-line models investigated, using the same single atmospheric background profile. The GCM radiation schemes were less accurate, with ~20% errors in the CO2 RF ... (Collins et al., 2005 and Chapter 10). Nevertheless, the current set of Atmosphere and Ocean GCMs (AOGCMs) used in Chapter 10 of this report found values for RF, for a doubling of CO2 that ranged between 3.5 and 4.2 W m –2 , in good agreement with the TAR RF value of 3.7 W m –2 (see Chapter 10 and Forster, 2005). Webb et al (2005) compare forcing for doubled CO2 in 9 models, with that forcing ranging from 3 to 4 W m-2. Webb, M. J., C. A. Senior, D. M. H. Sexton, 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, and K. Taylor, 2005: On uncertainty in feedback mechanisms controlling climate sensitivity in two GCM ensembles. Clim. Dyn., in revision. A similar estimate of uncertainty is reached in Table 10.2.1, for which the average and standard deviation forcing for doubled CO2 for 9 models is 3.71 ± 0.48 W m-2, or ±13% (range 2.99 to 4.23). [Stephen E Schwartz]	
8-186	A	9:49	9:49	"model versions" ? How about just "models" ? [Ileana Bladé]	Accepted
8-187	A	9:55		the proper use is “different from”, not “different to” or “different than [Philip Mote]	Accepted
8-188	A	9:57	9:57	I strongly suggest avoiding use of the word "tuned" in this paragraph, since it is vague, and open to misinterpretation by non-modelers. One could ask, if models in fact can be "tuned to reproduce recent climate" then why don't all models reproduce recent climate with negligible errors, including changes in Arctic sea ice, Amazon precipitation, annual surface air temperature variations in Tokyo, Southern Ocean heat uptake, the Atlantic MOC circulation, sea level in Tuvalu, snowfall in Lake Placid, and Atlantic hurricane frequency? Answer: because that's not what the authors meant by "tuned to reproduce recent climate", but the reader doesn't know that from this paragraph. (Such an interpretation is not discouraged until page 12, lines 25-27.) The somewhat more detailed description of 'tuning' in section 8.1.3.1, though incomplete, is less objectionable. [Keith Dixon]	Accepted
8-189	A	10:2		it is also true of the observed record that both the forcing and state variables are “imperfectly known” - perhaps a more nuanced statement is needed (e.g., whether they are known well enough to reconcile the one with the other)	Accepted

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Philip Mote]	
8-190	A	10:4	10:4	The implication of this statement is that climate sensitivity of not constant, but is a function of climate state. If this is true, then it represents change in the way climate sensitivity is defined. For example, the TAR (WG I, Pg.789) states: More generally, equilibrium climate sensitivity refers to the equilibrium change in surface air temperature following a unit change in radiative forcing. The text needs to explain whether this definition is still complete and valid. If not, a new definition of climate sensitivity should be presented. The topic is addressed again in section 8.6.2.1, but only to add more information as to the limitations of the concept. [Lenny Bernstein]	Now discussed in Glossary definitions
8-191	A	10:4		the proper use is “different from”, not “different to” or “different than [Philip Mote]	Accepted
8-192	A	10:8	10:19	I feel that the testing climate models through weather forecasting is the weakest link of this section, and one that calls for more discussion of its limitations. To me, the caveat found in lines 13 to 17 is insufficient. More detailed disclosure of the limitations of what can and can not be tested in this manner would be prudent. Even later in section 8.4.11 the emphasis is on the "fruitful new avenue" with no real mention of the limitations. The world's best weather forecast model can be the best even if it's long term climate drift and climate sensitivity to doubling CO2 render it unacceptable for dec-cen climate modeling applications. Similarly, the value of a climate model's 21st century projections chouldn't be discounted because it doesn't produce highly credible synoptic scale features (e.g., Atlantic hurricanes) that matter to weather forecasts. These limitations, obvious to most in the modeling community but not to all readers, are not readily apparent from this text. [Keith Dixon]	(Mostly) Rejected. In our view the discussion of testing climate models in weather forecast mode, as one method of evaluating them, does not imply that other tests no longer need to be done. Both in this section and in 8.4.11 caveats concerning limitations are included. We did reword the phrase "fruitful new avenue" to read "potentially fruitful new avenue".
8-193	A	10:8		The NWP test is also very useful to test how well weather systems/regimes, such as blocking, that are very important in climate are handled. [Brian Hoskins]	Noted.
8-194	A	10:12	11:24	This section should, probably, include a reference to the devasting effects of Katrina Hurricane in USA, a fully developed country, and supposed to be better prepared to avoid the effects of natural disasters. [Pedro Ribera]	Rejected. Outside remit of chapter
8-195	A	10:19		use --> used; also it may be worth noting in this paragraph that "weather" is a perfectly valid mode of variability thereby providing a test of model processes. [Richard Allan]	Noted.
8-196	A	10:21	8:29	Mentioning the subjectivity of indices. It seems to me that indices are developed, and models tested until they can meet the observation with some level of skill. Then that particular ceases to have value, except as something that should be "maintained" as other	Rejected. Space limitations prevent detailed discussion of model development process.

No.	Batch	Page:line From	To	Comment	Notes
				indices are developed. the subjective nature is true, but this eliminates the important part of the process - and also that indices are selected based on a combination of data availability and their ability to discriminate among simulations. [Anne Douglass]	
8-197	A	10:21	10:53	these sub-sections could be collocate in the other part of the chapter or eliminated [vincenzo artale]	Rejected. We have modified the text, but it has not been moved or eliminated.
8-198	A	10:21	10:35	Section 8.1.2.1 The second paragraph, reads as if mechanism based system-level evaluation is not common place. I think it has become much more so since the TAR and this is one of the reasons we are progressing on understanding the physical reasons for differences in climate sensitivity between models and hence able to give greater confidence in the model projections. [Catherine Senior]	Rejected. Second paragraph points to 8.6 and 8.7, and the importance of these advances is brought out in the ES
8-199	A	10:38		In model intercomparison activities, were the program codes exchanged and available to the community of researchers involved in the activity? In some cases (based on my own experience), publications on models differ somewhat from what is indeed programmed in the code. Exercices involving the exchange of source codes and simulations performed by other teams than those who programmed the code would be of interest and would demonstrate the full transparency of the process. In the early years, of climate modeling, access to the codes was not an issue. Today, the access to source codes is unfortunately very restricted and it may impact negatively the image of community of modelers. Should exchanges of source codes have occurred in the intercomparison activities I would be worth a mention it in the text. [Philippe Tulkens]	Noted. Codes are sometimes but not generally exchanged. No modification to the text.
8-200	A	11:2	11:7	Another downside of model intercomparison is that it can create consistent biases in models. If the majority of models obtain the same wrong output, then there will be pressure on the models that obtain a different output to "correct" their models to match the majority, even though the minority position may be correct. While gross errors, i.e., predicting global average cooling with increased GHG concentration, will be detectable, more subtle errors may not be. An overemphasis on model intercomparison could lead to less, not more, accurate models. Models should be tested against observations, not against each other. [Lenny Bernstein]	Noted. The text has been rewritten.
8-201	A	11:2	11:7	The discussion of the downsides of the modeling efforts in support of IPCC is perhaps tangential and perhaps represents an editorial comment that should be reworded. [William Collins]	Noted. The text has been rewritten.
8-202	A	11:2	11:7	Too much emphasis on model intercomparison can lead to poorer rather than better	Noted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				models. There is a natural tendency in all human activities to try to be part of the majority. If the majority of models have an incorrect characterization of some aspect of the climate system, the few modelers who were using a different, but correct, approach will tend to adjust their models to the majority's incorrect approach. These biases are likely to be subtle, since a modeling error that led to an obviously wrong output would soon be corrected. However, the combination of a number of subtle errors in modeling can lead to unrealistic models. Models need to be validated against observations. Testing them against each other is useful, but does not demonstrate that any of the models are valid. [Jeffrey Kueter]	
8-203	A	11:6	11:6	some disagreement EXISTS [Ileana Bladé]	Accepted
8-204	A	11:12	11:13	The conservation laws on which models are based (for mass, energy and momentum) should be explicitly mentioned instead of just Newton's law [Marisa Montoya]	Rejected due to space limitations.
8-205	A	11:20		An important point to be made about tuning is that it must be done not just to present values of a parameter, but to formulae which allow correctly for feedbacks and changes in the parameter. For example, the vertical mixing rates in the ocean interior may depend on things like stratification and wind speed, so that tuning to values representative of the present climate may exclude important feedbacks. [Chris Garrett]	Noted.
8-206	A	11:25		Tuning "justifiable"? This is far too defensive - tuning is an essential part of good modeling practice, using a model for projections which has not been properly tuned would be unjustifiable and irresponsible. [Stefan Rahmstorf]	The text has been rewritten.
8-207	A	11:28	11:40	I feel this discussion with the 2 conditions falls behind what actually is in the literature about this. I suggest to replace these two points with the "rules for good tuning practice" defined by Petoukhov et al. 2000 (already in the ref list of this chapter), to quote them: 1. Parameters which are known empirically or from theory must not be used for tuning. 2. Whereever possible parametrizations should be tuned separately against observed data, not in the context of the whole model. 3. Parameters must relate to physical processes, not to specific geographic regions (hidden flux adjustments). 4. The number of tuning parameters must be much smaller than the degrees of freedom predicted by the model. [Stefan Rahmstorf]	The text has been rewritten.
8-208	A	11:28		missing full stop after "exceeded" [Richard Allan]	Accept.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-209	A	11:34	11:40	The distinction between "observable" and "certain key observations" might not be clear to the reader. The author appears to be saying that if you tune to one variable, then just because the model agrees well with observations for this variable doesn't mean that the model is a good predictor; but that if you tune to several "key observations", this means that the model can be a good predictor. I'm afraid I don't follow the argument. Also, if chapter 10 is referred to for clarification, the specific section of chapter 10 should be mentioned. [Kevin Walsh]	Text clarified in revison.
8-210	A	11:36	11:40	There is to my knowledge no published basis for the statement that tuning to key observations increases the predictive capability of a model, especially when one is talking about climate change. A model that can replicate key observed *processes*, however, should be expected to have greater predictive capability. [Anthony Del Genio]	Rejected – see refs p9 ll 6-28.
8-211	A	11:42	11:44	The words "various optimization procedures" might be supplemented with the words " using data assimilation methods". Now, automated parameter tuning efforts have been made for not only low-resolution GCMs also common CGCMs. They are promising and could be specified in the article. [Toshiyuki Awaji]	Accepted – see response to 8-212
8-212	A	11:42	11:46	I am not certain which paper Annan et al 2003 is since it is not in the references (and also no paper of that name exists). Also, the same techniques we applied to EMICs have now also been applied to MIROC3.2 AGCM at T21. Suggested changes: "Given sufficient computer time the 'tuning' procedure can in principle be automated using various optimisation procedures; however this has only been feasible to date for EMICs (Hargreaves et al, 2004) and low-resolution GCMs (Jones at al., 2001, Annan et al 2005b). Ensemble techniques (Annan et al., 2005a; Murphy et al., 2004; Stainforth et al., 2005) allow in principle a range of parameter settings to be generated, each giving equally 'good' climate simulations according to some chosen measure." [Julia Hargreaves]	Accepted – reference corrected.
8-213	A	11:42	:46	References to the above: Hargreaves et al 2004 - you already have this paper in the references. This is the paper in which the model is tuned to present day climatology for the first time. Annan et al 2005a (I think this is the one you mean - it is the first paper on the EnKF outlining the method and predating Hargreaves et al 2004, but it is labelled as 2005 because Ocean Modelling have a numbering problem!) J. D. Annan, J. C. Hargreaves, N. R. Edwards and R. Marsh. Parameter estimation in an intermediate complexity Earth System Model using an ensemble Kalman filter. Ocean	Accepted – reference corrected.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Modelling, 2005, Volume 8, Issues 1-2, Pages 135-154. Annan et al 2005b J. D. Annan, J. C. Hargreaves, R. Ohgaito, A. Abe-Ouchi, S. Emori. Efficiently constraining climate sensitivity with paleoclimate simulations. SOLA, 2005, Vol.1, 181-184. [Julia Hargreaves]	
8-214	A	11:43	11:44	Annan et al. (2003) is cited twice here but not included in the reference list [Marisa Montoya]	Accepted – reference corrected.
8-215	A	11:44	11:44	I am not sure whether Jones et al (2001) exists and it is not in the references. However, a better reference, and maybe the one intended here, is Jones, C., J. Gregory, R. Thorpe, P. Cox, J. Murphy, D. Sexton, and P.Valdes, 2005: Systematic optimisation and climate simulation of FAMOUS, a fast version of HadCM3. Climate Dynamics, 25, 189-204. [David Sexton]	Accepted – reference corrected.
8-216	A	11:44	11:46	"Ensembles…chosen measure" is not an accurate statement, due to the fact that ensemble members are not equally "good". Murphy et al sampled parameter space where runs were considered equally plausible prior to any comparison with observations. Stainforth et al sampled parameter space but did not specify any prior probabilities to their runs. Annan et al's method samples the posterior probability distribution i.e. the distribution which has accounted for how well the models compare to the observations, and some runs here will definitely be "better" than the others. Webb et al (2005), cited elsewhere in chapter, selects parameter combinations which are expected to be of differing quality in how they represent present-day climate, although runs which are predicted to have the same sensitivity are approximately equal in quality. Given this is in a paragraph about tuning models, maybe the following sentence is more appropriate - "Ensemble technqiues (Annan et al, 2003, Murphy et al 2004, Stainforth et al, 2005, Webb et al, 2005), which explore model output over a variety of parameter settings, show that a number of models in disparate regions of parameter space, can produce equally "good" climate simulations according to some chosen measure". [David Sexton]	Taken into account in redrafting text
8-217	A	11:44		Palmer and Raisanen discussed probabilistic ensemble-forecast techniques using multi-model ensembles. (Palmer, T.N. and J. Räisänen, 2005: Quantifying the risk of extreme seasonal precipitation events in a changing climate. Nature, 415, 512-514.). A comparison of the multi-model and perturbed parameter approaches to representing model uncertainty (as well as the stochastic physics approach, see below) will be tested in seasonal forecast mode as part of the EU ENSEMBLES project (www.ensembles-eu.org). [One method of validating the use of (fast-physics) perturbed parameters, as in Stainforth et al and Murphy et al, is through the budget residual technique (Klinker, E. and P.D.	Rejected. Interesting developments but will miss publication deadline.

No.	Batch	Page:line From	To	Comment	Notes
				Sardeshmukh, 1992: The diagnosis of mechanical dissipation in the atmosphere from large-scale balance requirements. J.Atmos.Sci., 49, 608-627), ie assimilating observations into the perturbed model, running the perturbed models for just a few timesteps from different initial conditions, and examining the net tendency, which is a measure of the imbalance of the model against observations. Initial tests suggests that this will be a very discriminating test, much more than testing models by running in 20th century mode. There is a paper: Rodwell, M., Palmer, T.N. and Stainforth, D., 2005: Using initial weather prediction imbalances to assess climate models. In preparation. Unfortunately this won't be accepted by the time of deadline of AR4, hence the reason for putting this paragraph in parentheses. It will, however, be an ECMWF Technical Memorandum within a month or so.] [Timothy Palmer]	
8-218	A	11:45		Why "equally good"? Why can't some ensemble members give better simulations than others? [Stefan Rahmstorf]	Accepted. Text clarified.
8-219	A	11:46	11:46	Recently Severijns and Hazeleger (2005, J Clim, 18,3527-3535, Optimizing parameters in an atmospheric general circulation model) published a method that is efficient and can be used to optimize parameters in general circulation models. [Wilco Hazeleger]	Accepted.
8-220	A	11:52	11:52	IN that ? [Ileana Bladé]	Accepted
8-221	A	11:52	11:52	Please define 'emergent properties". [William Collins]	Accepted
8-222	A	11:55		Replace "They also have" with "Some also have", as this does not apply to all [Stefan Rahmstorf]	Accepted
8-223	A	11:57		Replace this line with: "… so that particular care has to be taken when interpreting the results" Otherwise, the statement might equally apply to GCMs - for those it is also not clear whether all results apply to the real world. In fact I'm sure they don't, if I look at the results from some of those GCMs included in the figures, which show some fundamental climate characteristics for which we would send our EMIC back to the drawing board. I would say that even more care has to be taken with EMICs than with GCMs, but this is a difference of degree, not a fundamental one. [Stefan Rahmstorf]	Accepted. Text modified
8-224	A	12:0	13:	Question 8.1: It may be worth reiterating that: "A realistic simulation of climatology or variability, while increasing confidence in model skill, does not necessarily mean that the processes, and in particular feedback mechanisms, important for climate prediction are	Noted: although space limitations preclude this explicitly, the point is included by implication in the

No.	Batch	Page:line From	To	Comment	Notes
				adequately represented." [Richard Allan]	comment: “models display a substantial range of global temperature change in response to greenhouse gas forcing, and assessment of model skill in the representation of current climate has not, to date, been able to significantly reduce this range”.
8-225	A	12:0		Question 8.1: The use of the phrase "climate features" appears rather vague. [Richard Allan]	Accepted: this is now made explicit by listing a number of particular features, such as temperature and precipitation variability, ocean currents, ...
8-226	A	12:0		are the Questions to be collected in the SPM or technical summary? they seem to be on a more basic level than the chapters, and too buried for such crucial overview-type issues. The answer to this question is well written and hits some important points. [Philip Mote]	The answer is yes, common questions are to be collected together. Also yes, they are designed to be at a more "basic" level i.e. less technical than the rest of the chapter.
8-227	A	12:8	8:44	this discussion is repetitious with that given on pages 10 and 11. [Anne Douglass]	Rejected: CQ needs to "stand alone", and repetition is indeed necessary to some degree so as to make it consistent with other parts of the chapter.
8-228	A	12:8	12:11	Statements like this, asserting there is "considerable confidence" about models providing "useful projections" are of very limited value. There is no quantification of what "considerable condidence" means, and "useful" is highly context-dependent. [Leo Donner]	The statement has been slightly rewritten both here and in the executive summary, but overall level of confidence has been maintained. Note that the use of "useful", in particular however, denotes appropriateness for policy applications of model output, and was used in the TAR summary. The degree of confidence assessed is considered appropriate, for the target audience of the "common question".
8-229	A	12:8	12:11	This paragraph is an exact repetition of lines 3-6 of page 8_3. It is unnecessary and should be deleted [Vincent Gray]	Rejected: the "Common Question" sits separately from the Executive Summary, and all CQs will also be collected together and published separately. A degree of repetition is

No.	Batch	Page:line From	Page:line To	Comment	Notes
					acceptable (in fact unavoidable) because the question is in many ways a synopsis of much of the chapter, and must closely parallel the Executive Summary.
8-230	A	12:8	12:9	Is this statement really true,if we consider the wide range of climate sensitivity among available models? Especialy considering the potentials to predict the relationship between the temperature rise and the amount of CO2 emission. If we consider the CO2 mitigation target for the situation where future temperature rise is limited under 2 degreeC, the actual level of mitigation target changes with climate sensitivity. It seems to me that there is some misunderstanding to see the implicaiton of range in climate sensitivity. Some people misunderstood that this range represents the uncertainty associated with a specific model. The implication of range in climate sensitivity should be clearly described somewhere in this chapter. [Shigeki Kobayashi]	Noted: the wording has been changed somewhat, although the overall level of confidence expressed remains fairly consistant. This statement, the "headline answer" is the overall assessment of the chapter on confidence in models. Issues to do with uncertainty (for example of climate sensitivity), are acknowledged, and are dealt with in subsequent paragraphs of the answer to the Common Question. Also it seems unlikely that there is really widespread misunderstanding on this issue as the AR4 takes pains to say that the range of sensitivity comes from a range of models (see e.g. Ch 10). The implication of the range in sensitivity is also described fully in Ch 10.
8-231	A	12:8	12:11	Question 8.1: This first paragraph (the "headline answer") should be in italics. [David & David Wratt & Fahey]	Change done.
8-232	A	12:8	:9	There is considerable confidence that models are reliable enough to provide useful projections of future climate change, particularly at larger scales. This statement needs to be made more quantitative. The term "useful projections" is unacceptibly vague. What accuracy is required? What is the present accuracy? Similar comment to lines 24-25: Models show significant, and increasing, skill in representing many climate features, particularly at larger spatial scales, and this increases our confidence in their use for simulating future climates. "significant skill" is a "weasel" phrase. "Significant" in the sense of "better than random"? [Stephen E Schwartz]	First part, noted, and text slightly modified: see response to 8-228. Second part, rejected: the common question is not written in statistical language, e.g. using " significant" to denote statistical significance, but in its more colloquial meaning. The language has been chosen as required to explain to the required level of technical detail for the Common

No.	Batch	Page:line From	Page:line To	Comment	Notes
					Questions..
8-233	A	12:14	:16	Model fundamentals are based on physical laws, such as conservation of mass, energy and momentum, along with a wealth of scientific observations. This is true but somewhat disingenuous. Models also contain many parameterizations of processes, and the skill of the model depends on the skill of these parameterizations. Often slight changes in parameterizations can change model sensitivity and other model observables. Modeled climate sensitivity is highly dependent on parameterizations; for example, a change in cloud parameterization in the UK Meteorological Office model changed modeled climate sensitivity to doubled CO2 by a factor of 2.8, from 1.9 to 5.4 K. Senior, C. A. and Mitchell, J. F. B. Carbon dioxide and climate: the impact of cloud parameterization. J. Climate 6, 393-418 (1993). I would note the large variance of models and observations eg figures 8.3.5, 8.3.6, (10 W m-2) and even in global mean 8.3.28 of several W m-2. The question that needs to be addressed is whether differences such as these in climate properties lead to substantial differences in capability of models to predict climate change in response to specific forcings. [Stephen E Schwartz]	First part: rejected. Model parameterisations, as well as dynamics are also constrained by these fundamental physical laws. Furthermore, parameterisations often or usually draw from observations, such as field experiments. So the comment is correct as it stands, in our opinion. In addition, the need for, and limitations introduced by parameterisations is explicitly addressed (8-12, line 48 onwards). Second part: rejected. The dependence of sensitivity on parameterisations is fully acknowledged and discussed. The importance of clouds is also emphasised (paragraph starting 8-12, line 48) Third part: Noted. However there is already extensive, and we feel adequate, discussion of why there is a range of sensitivity, and what it means in the 6th and 8th paragraphs.
8-234	A	12:16		Comparison with aircraft design is misleading. In that case there is little, if any, need for parameterization of unresolved processes. In the climate system, there are many, many, orders of magnitude separating resolved and unresolved processes. [Chris Garrett]	Accepted: Dropped aircraft analogue.
8-235	A	12:24	12:25	"Models shows … skill .. particularly ….and this increases .. " is redundant. The same statements are made above in lines 8, 10 and 18. [Ileana Bladé]	Taken into account: Lines 8 & 10 form part of the "headline answer", and therefore form a summary with necessary repetition. Line 18 is also the "headline sentence" for the paragraph starting on line18 However wording has been changed slightly in first sentences of paragraphs 2, 3 and 5 to enumerate the “headline” confidence

No.	Batch	Page:line From	Page:line To	Comment	Notes
					criteria discussed in each paragraph.
8-236	A	12:24	12:27	The assertion that the ability of climate models to reproduce variability indicates "real" skill and does not arise from adjusting or tuning models is of dubious accuracy. Most climate model development groups are very cognizant of variability in their models during the development process and definitely make development choices based on thier implications for variability. [Leo Donner]	Taken into account: comment on tuning now dropped due to space limitations.
8-237	A	12:24	12:24	Question 8.1 You say the models show significant and increasing skill…so are the models more reliable than the TAR? [Catherine Senior]	Noted: For the Common Question increasing skill is assessed as an ongoing feature of models, and not just restricted to development since the TAR. Other parts of the chapter assess specific developments and confidence changes in models since the TAR.
8-238	A	12:31	12:32	Insert "However, many problems still remain in simulating ENSO, the most important mode of interannual variability. These are described in Section 8.4.7. The statement that model are becoming more skillful in predicting features of interannual variability, e.g. ENSO, is technically correct, since there have been improvements in ENSO modeling. However, the implication is highly misleading. Section 8.4.7, particularly Pg. 44, lines 30-42, details the problems remaining in simulation of ENSO. The section concludes with the observation (Pg, 45, lines 9-10): "Finally it remains unclear how changes in mean climate will ultimately impact ENSO predictability." [Lenny Bernstein]	Accepted: text modified (although not using the suggested wording) to be made consistent with section 8.4.7. : limitations in representing ENSO are now mentioned explicitly.
8-239	A	12:31	12:32	This is an overly optimistic statement. An objective assessment of the state of models is that they cannot very well simulate the frequency, strength, vertical structure and clouds associated with synoptic midlatitude storms (Lambert, S., J. Sheng and J. Boule, 2002: Winter cyclone frequencies in 13 models participating in the Atmospheric Model Intercomparison Project (AMIP1). Clim. Dyn., 19, 1-16; Zhang, M., et al., 2004 (already cited - but the year should be 2005); Bauer, M., and A.D. Del Genio, 2005: Composite analysis of winter cyclones in a GCM: Influence on climatological humidity. J. Clim., in press.), nor can they simulate the MJO (Lin, J.-L.. et. al., 2005 (already cited, but now in press), nor can they simulate well the cloud/water vapor response to ENSO (Soden, 2000, already cited). [Anthony Del Genio]	Accepted, although the level of detail is necessarily limited due to the scope and space considerations of the Common Question. Text modified to add caveats on ENSO and tropical convective organisation as examples of large scale errors.
8-240	A	12:31	12:32	While there have been improvements in simulating ENSO, this text gives an overly optimistic assessment of current capabilities. Section 8.4.7 details the many problems still	Accepted: made consistent with section 8.4.7. : limitations in representing

No.	Batch	Page:line From	To	Comment	Notes
				remaining in ENSO simulation. In addition, as Section 8.4.7 (Pg. 45, lines 9-10) concludes: "Finally it remains unclear how changes in mean climate will ultimately impact ENSO predictability." Policymakers need a balanced assessment of climate science's capabilities, not general statement of improvement. This is particularly true for climate phenomena, such as ENSO, that have significant social and economic impact. [Jeffrey Kueter]	ENSO are mentioned explicitly.
8-241	A	12:32	12:34	The inability to predict weather beyond one or two weeks DOES limit our abilty to predict climate changes in two ways. First, it introduces interannual fluctuations of several tenths of a degree C in global mean surface temperature which can on occasion yield trends of as much as 0,5 C over periods of several decades, which are unpredictable. To be sure this is not a serious limitation on global warming projections for 2100, but the second limitation is potentially serious. The second is that, when th climate system approaches a bifurcation, the chaotic component of weather can make unpredictable exactly when the system crosses the bifurcation. An interseting example of this is given in Wang et al. (1999, J. Clim., 12, 71) and Marotzke (2000, Proc. Nat'l. Acad. Sci., 97, 1347). In this case the exact time of collapse of the MOC in the North Atlantic cannot be predicted within a period of several 100 years, even though the forcing is known exactly. [Peter Stone]	Both points: noted, but all discussion is now removed on this point for space considerations.
8-242	A	12:32		Too optimistic here on ENSO treament compared with statements elsewhere. [Brian Hoskins]	Accepted: text modified to make consistent (see response to 8-240).
8-243	A	12:34	12:34	We suspect the reference here to "Question 8.2" should actually be to Question 1.2 (there is no Question 8.2). [David & David Wratt & Fahey]	Accepted, text modified.
8-244	A	12:34		Question 8.2 - where is it? [Stefan Rahmstorf]	Accepted, text modified to “1.2”.
8-245	A	12:36	12:44	this discussion is out of the aims of this chapter is more related to the chapter 6 [vincenzo artale]	Rejected: to properly answer the Common Question, some aspects of other chapters must be discussed (A number of the Common Questions touch upon several chapters).
8-246	A	12:36	12:43	While models can reproduce many features of paleo-climate and changes in historical temperature records, they have generally been unsuccessful at simulating massive changes from one paleo-climate state to another. This should be noted in this section of the text, as it is an important test which the models cannot yet pass. [Leo Donner]	Taken into account: Text changed to models “can reproduce many features”. Further paleo comments not made due to space limitations within CQ.
8-247	A	12:36	12:38	It is not correct to say that climate models have been able to reproduce the LGM. In the simulations the continental ice sheets and their properties are specified, whereas a true	Rejected: the support for projecting anthropogenic climate change over the

No.	Batch	Page:line From	Page:line To	Comment	Notes
				climate model would be able to simulate these features of the LGM from first principles, and would not have to take them as given. [Peter Stone]	next century does not require models capable of growing continental ice sheets from scratch. However some support from LGM for modelling comes from GCM ability to simulate large-scale temperature changes in the presence of forcing.
8-248	A	12:36	39:	this discussion of paleo is encouraging, but I wonder if you could be more thorough - perhaps giving some indication of what paleo model eval suggests models do well and not so well (e.g., with the warmer (than what??) Holocene and last glacial maximum. Of course, you might be able to just cite some papers and chap 6 (please discuss/coordinate w/ the chap 6 model team mentioned above). [Jonathan Overpeck]	Noted: however more detail is impossible due to space constraints in the Common Question. The Common question can spare no more than a sentence or two on this, and given its target audience it is not appropriate to cite papers. However, a reference to chapter 6 has been inserted.
8-249	A	12:37		replace Holocene by mid-Holocene (we are also in the Holocene right now) [Stefan Rahmstorf]	Accepted: text modified.
8-250	A	12:39	:42	Models also simulate many observed aspects of climate change over the instrumental record, such as the global temperature trend over the past century (Figure 1), although uncertainties in the magnitude of the cooling associated with sulphate particles provide significant limitations to this test. This statement is more than a bit disingenuous. The uncertainties in present day forcing by aerosols (not just sulfate aerosols) at the time of the calculation referred to in Question 8.1, Figure 1 were quite large and within that uncertainty quite a range of temperature trends might have been obtained. The resultant impression left by the figure is of a much greater model skill than if the figure showed a set of runs with forcings that run the gamut of uncertainty in forcing. This point has been made, inter alia, by Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott and N. A. Rayner, 2002: An Observationally based estimate of the climate sensitivity. Journal of Climate, 15(22), 3117-3121 Boucher O. and Haywood J. (2001) On summing the components of radiative forcing of climate change. Climate Dynamics 18, 297-302. Schwartz S. E., Uncertainty requirements in radiative forcing of climate change. J. Air Waste Management Assoc. 54, 1351-1359 (2004). [Stephen E Schwartz]	Rejected: Comment is already made (starting 8-12, line 41) that aerosol uncertainties provide "significant limitations to this test", so that it is made clear that there are substantial uncertainties here. Note also that the figure is taken from the TAR Summary for Policymakers, with similar levels of caveats.
8-251	A	12:41	12:41	Figure 1 is mentioned here but not included in the text	Figure is included at the end of the

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Marisa Montoya]	chapter figures.
8-252	A	12:48		Many aspects of the larger-scale organisation of tropical convection are poorly handled. [Brian Hoskins]	Accepted: comments added on this point, and also on ENSO limitations.
8-253	A	12:55	12:55	temperature change IN RESPONSE to greenhouse gas forcing [Ileana Bladé]	Accept grammatical correction
8-254	A	12:55	12:56	Whilst it is true that we have yet to actually reduce the uncertainty, I think you should qualify this comment by saying that we are now gaining knowledge of some of the causes of uncertainty (e.g. low cloud) and so there is hope! [Catherine Senior]	Noted: Good point, but space limitations preclude more detail on this point.
8-255	A	13:0		Choice of references questionable: the results of Tziperman 97 have been shown (in Rahmstorf & Ganopolski, J. Clim. 1999, http://www.pik-potsdam.de/~stefan/Publications/Journals/rahmstorf_ganopolski_jclim99.pdf) to be an artefact of an unphysical experimental design; this paper should either not be cited, or together with a caveat pointing to the rebuttal paper. Rind et al 2001 on the other hand does not demonstrate thresholds. [Stefan Rahmstorf]	Accepted. Text modified.
8-256	A	13:5		Suggest omitting or explaining 'downscaling methods' [David & David Wratt & Fahey]	Accepted: text now refers to CQ 11.1 in relation to 'downscaling methods'.
8-257	A	13:15	13:24	This summary paragraph speaks in extremely qualitative terms. It should identify more concretely the sensitivity range of the models. [Leo Donner]	Noted: however the purpose of CQ 8.1 is not to provide a quantitative range of climate sensitivity, but to assess our confidence in the use of models for projection, and is expressed qualitatively with non-specialist target audience in mind.
8-258	A	13:15	:16	"In summary, confidence in models comes from their physical basis, and their skill in representing observed climate and past climate changes." The fact that present AOGCM's exhibit a substantial range of sensitivity yet reproduce the observed change in global mean temperature over the past 100 years quite well suggests that the above statement would have to be qualified. Schwartz S. E., Uncertainty requirements in radiative forcing of climate change. J. Air Waste Management Assoc. 54, 1351-1359 (2004). The following statements from Hansen (2005) seems quite a propos: A caveat accompanying our analysis concerns the uncertainty in climate forcings. A good fit of observed and modeled temperatures (Fig. 1) also could be attained with smaller forcing and larger climate sensitivity, or with the converse. Hansen, J., L. Nazarenko, R. Ruedy, Mki. Sato, J. Willis, A. Del Genio, D. Koch, A.	Rejected: this point is already adequately covered in the discussion of observed temperature changes in our opinion – in particular caveats on the reproduction of observed warming due to forcing uncertainties are explicitly mentioned. Skill in observed climate change is not the only aspect being referred to here. Models also have ability to represent e.g. decrease in diurnal temperature range, Pinatubo response.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Lacis, K. Lo, S. Menon, T. Tovakov, Ju. Perlwitz, G. Russell, G.A. Schmidt, and N. Tausnev 2005. Earth's energy imbalance: Confirmation and implications. Science 308, 1431-1435, doi:10.1126/science.1110252. [Stephen E Schwartz]	Note, however, that text has been modified slightly: "the degree of their skill" to replace "their skill", as is making it clear that they have a degree of skill, but not claiming that they are without errors and problems
8-259	A	13:19	13:19	Change "a degree of uncertainty" to "substantial uncertainty." The TAR finding that the range in projections of future global average temperature was as large when a single scenario was used in the range of models as when the ranage of emissions scenarios were used in a single model is still valid. Models outputs for a variety of critical parameters, e.g. climate sensitivity, still vary over a factor of 3 or more. [Lenny Bernstein]	Accepted in part: text changed from "a degree of uncertainty" to "uncertainty", which reflect a more significant size of range.
8-260	A	13:19	13:19	The phrase "a degree of uncertainty" is too vague, and should be replaced by some measure of the uncertainty in projecting the magnitude and timing of climate change. The TAR projected a rise of 1.4 - 5.8 C in global average surface temperature between 1990 and 2100. Policymakers will want to know whether this projection has changed, and if so, what is the basis for changing the projection? While making new projections is the responsibility of Chapter 10, this section should use Chapter 10's projections and refer the reader to the appropriate section of Chapter 10 for details. [Jeffrey Kueter]	Rejected: the purpose of CQ 8.1 is not to provide a quantitative range of climate sensitivity, but to assess our confidence in the use of models for projection. Discussion of changes in time of the range of climate sensitivity are also outside the scope of the question.
8-261	A	13:20	13:20	magnitude, timing and regional patterns of predicted climate change [Catherine Senior]	Accepted: add "regional patterns"
8-262	A	13:22	13:24	Just as I thought: To the limited extent to which we can believe the big models, we didn't need them! This needs a bit more thought to show why the big models are really more reliable than a back of the envelope estimate, given the huge uncertainties in the effects of unresolved processes. [Chris Garrett]	Rejected: the text did not say that models were not needed or useful, but simply that models are not the only line of evidence. These are very different statements.
8-263	A	13:22	24:	this sentence/assertion needs some citations [Jonathan Overpeck]	Rejected: Citations for particular points not appropriate in the Common Question due to style of answer and target audience.
8-264	A	13:34	13:35	…,and new physical processes have been added to the models.' How is this different from the second category? [Catherine Senior]	Accepted. The text has been modified.
8-265	A	13:35		Relaxation of the rigid-lid approximation in the equations for the barotropic mode in ocean models is an aspect of the dynamical core, not a parameteriation of a physical	Accepted. Text will be revised

No.	Batch	Page:line From	Page:line To	Comment	Notes
				process. Replacement of virtual salt fluxes by the natural fresh water flux boundary condition, while physically better posed, is a relatively minor factor in the improvements of the veracity of solutions of ocean components of climate models. In the period since the TAR, the single most important factor in the improvement of OGCMs has probably been the nearly universal (with the exception of PCM?) adoption of adiabatic closures for eddy mixing. The shift to natural b.c. should be de-emhasied relative to this aspect, or perhaps issues like vertical mixing schemes, BBL, etc. Also, note that use of a free-surface in the dynamics does not necessarily imply that the natural b.c. on fresh water has been adopted. Similarly, in table 8.2.1, the form of the lateral mixing operator would be a more valuable piece of information than the form of the surface B.C. [Frank Bryan]	
8-266	A	13:39	13:40	The statement "...it is imposible to state that any existing model is fully adequate to make projections of future climate." is a highly important statement that must be reflected in the Executive Summary with whatever explanation the authors feel is needed. [Lenny Bernstein]	Sentence is open to misinterpretation. Redrafted
8-267	A	13:39	13:40	The finding "...it is imposible to state that any existing model is fully adequate to make projections of future climate." needs to be retained in future drafts. It also needs to be repeated in the Executive Summary and in the higher level summaries (SPM, Synthesis Report) that will be prepared later in the writing process. [Jeffrey Kueter]	Same response as for 8-266
8-268	A	13:40		This line will be loved and trumpeted by the sceptics. I also disagree with its sweeping conclusion. What does "fully adequate" actually mean? This is too wide open for all kinds of interpretations. [Stefan Rahmstorf]	Same response as for 8-266
8-269	A	13:44	13:45	Multi model more robust than single model': Please specify what robust means here. Does it mean in better agreement with observations? If yes, what observations? This seems to be the case for some diagnostics, but in other cases many models have common biases. Please provide references. This statement has strong implications for the use of a multi model mean in chapter 10, and should be very well supported. I doesn't seem to me that this has been established well enough in the literature and in chapter 8. [Reto Knutti]	Accepted. Sentence deleted.
8-270	A	13:49	15:36	There is no section on radiation in this Chapter. Perhaps it may be implicitly assumed that radiation modeling is being done so accurately that there are no outstanding issues, hence mention is necessary. That would be wishful thinking. Radiation is the key process that determines the energy exchange between the different radiative constituents of the climate system and defines the equilibrium temperature of climate. Chapter 2 describes and compares some of the details of radiative forcings by different GHGs and	Accepted. A discussion of radiation has been added.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				aerosols. However, there should be an overview discussion of radiative modeling issues pertinent to climate and climate change modeling. Some questions to raise and answer include: How can we tell if the radiation modeling is adequate? If the TOA radiative fluxes do not agree with ERBE, is that a radiation problem, or maybe the model clouds and hydrology are at fault? If the tropopause is too warm or too cold, is that a radiation problem, or inadequate model dynamics or ozone distribution? Is daily-mean radiation OK, or is an explicit diurnal cycle dependence required? [Andrew Lacis]	
8-271	A	13:49	15:37	Discussion of parameterizations of microphysical processes of clouds and precipitation is lacking, which seems inadequate in view of their importance in models of various scales from CRMs to GCMs. Some progress has been made in this aspect. For example, in a serious of papers, a new analytical autoconversion parameterization has been theoretically derived, which not only provides a firm physical basis but also eliminates tunable parameters existing in traditional parameterization schemes (Liu and Daum, J. Atmos. Sci., 61, 1539-1548, 2004; Liu et al., Geophys. Res. Lett. 31, doi:10.1029/2003GL019117, 2004; Liu et al. Geophys. Res. Lett. 32, doi:10.1029/2005GL022636, 2005; Liu et al., Parameterization of autoconversion process. Part II: Generalization of Sundqvist-type parameterizations, J. Atmos. Sci., in press, 2005). Application of the new scheme leads to a 61% decrease of the second indirect aerosol effect (Rotstayn and Liu, 32, doi:10.1029/2004GL021922, 2005). [Yangang Liu]	Noted. Discussion is limited to parameterizations in use in AR4 models.
8-272	A	13:50		The Executive Summary highlights should probably be slightly tempered to include a few negative aspects associated with some of the points. Specifically: (1) (29-34) In noting substantial progress in identifying cloud feedbacks, it should also be noted that little progress has been made on developing a strategy for representing those feedbacks correctly. The draft highlights already note the lack of progress on cryospheric feedbacks, so it seems only consistent to provide a fuller status on the cloud feedbacks, even though this point will be expanded upon later in the summary. (2) (45-47) It's claimed that forecasting successes with climate models increase confidence in processes important for longer-term climate prediction. Can even a qualtitative indication be provided as to the nature or extent of this increased confidence? [Leo Donner]	(1) Rejected. The text already states that cloud feedbacks are a large source of uncertainty. Focus of ES is on current state of knowledge. (2) Rejected. Insufficient space here. Fuller discussion is provded in the body of the chapter.
8-273	A	13:52	13:56	This discussion reads as if spectral advection schemes should be replaced by semi-Lagrangian schemes. While spectral (or centered space grid based) schemes are certainly not favoured any more, semi-Lagrangian schemes are known to be probelmatic as well beacuse of their lack of mass conservation. Recent GCMs (Steil et al., JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 108, NO. D9, 4290, doi:10.1029/2002JD002971,	Noted. Text is revised.

No.	Batch	Page:line From	To	Comment	Notes
				2003) have moved away from semi-Lagrangian schemes and use schemes like SPITFIRE. [Rolf Müller]	
8-274	A	13:52	14:2	I do not see the value of this paragraph's level of detail (e.g., "negative water" or the GFDL and MRI model histories). [Keith Dixon]	Noted
8-275	A	13:54	13:54	Reference Williamson and Rasch, 199x does not appear in the list of references. [Alejandro Bodas-Salcedo]	Noted
8-276	A	13:54	13:54	Williamson and Rasch (199x) is incomplete and not included in the reference list [Marisa Montoya]	Noted
8-277	A	13:54	13:54	No year in the reference [Eugene Rozanov]	Noted
8-278	A	13:54	13:54	"ref. Williamson and Rash, 199x" should be completed. [Philippe Tulkens]	Noted
8-279	A	13:54	13:54	missing the reference year for Williamson and Rasch, 199x. [Jin-Yi Yu]	Noted
8-280	A	13:55	13:55	See for example Muller, Mon. Wea. Rev, , p1407,1992, Sokol, QJ, 1999, or Thuburn, QJ, 1993, for comparisons of various positive definite, mass conserving transport schemes and application in GCMs. The discussion should not solely concentrate on the semi-Lagrangian method. and the problematic point of this scheme (no mass conservation) should not be neglected. [Rolf Müller]	Noted.
8-281	A	14:7		ECMWF run at much higher resolutions than climate models using the spectral technique and have no particular problems with transform times or Gibbs problems. Runs at T2047 (10km grid) show only 20% of the CP time being spent on transforms. [Brian Hoskins]	Taken into account in the text.
8-282	A	14:8	14:8	Give reference for the Gibbs phenomenon [Marisa Montoya]	No need.
8-283	A	14:9	14:10	Gibbs phenomena are a fairly serious problem even at coarser resolutions. The current wording, "become non-negligible," at increasing resolution, connotes this is not so. [Leo Donner]	Noted
8-284	A	14:19	14:36	Section 8.2.1.2. In discussing the increases in resolution, mention should be made of the need for sufficient horizontal resolution when semi-Lagrangian advection schemes are adopted. Studies (e.g. Chen et al.,1997 [J Clim, 10, 2374-2389]; Williamson et al., 1998 [Mon. Weath. Rev. 126, 1001-1012]) have shown that semi-Lagrangian dynamical cores tend to have less transient eddy activity than equivalent Eulerian dynamics at typical clijmate resolutions (~2.5 deg). For example, in HadGEM1, at a resolution of 3.75 by 2.5	Notd

No.	Batch	Page:line From	To	Comment	Notes
				deg the transient eddy kinetic energy was underestimated, storm tracks were weaker and the frequency of blocking reduced compared to the previous model version (HadCM3) in which an Eulerian dynamical core was used (Martin et al., 2005; Ringer et al., 2005 [see comment 24]). [Gill Martin]	
8-285	A	14:19	14:36	Section 8.2.1.2. The impacts of vertical resolution changes are much more difficult to evaluate, because model physics is particularly sensitive to it and some schemes, e.g. boundary layer, convection, cloud) behave very differently at different resolutions.This should be mentioned. There appear to be more problems with increasing lower tropospheric resolution, while upper tropospheric and stratospheric resolution increases are generally beneficial. [Gill Martin]	Noted
8-286	A	14:20	14:26	I do not see the value of this paragraph's level of detail. That the models have higher resolution is interesting, but I doubt that many readers will find the jargon filled resolution details of just a few of the AR4 models particularly useful. [Keith Dixon]	Rejected. Necessary for model developers.
8-287	A	14:20	14:25	will the notation of triangular truncation be explained somewhere? I don’t have the glossary... perhaps a footnote here would be good. Also, what is TL959 (page 14 line 25) [Philip Mote]	Noted.
8-288	A	14:24	14:31	Another example that could be cited: Diaz J., Alberdi J. C., Pajares M. S., Lopez C., Lopez R., Lage M. B. and Otero A. (2001) 'A Model for Forecasting Emergency Hospital Admissions: Effect of Environmental Variables'. Journal of Environmental Health, 64(3), 9-15. This paper describes a model where environmental variables, like pollutants concentratios -ozono included- are used to predict the admissions in Madrid (Spain) hospitals. Ozone influenced the admissions for circulatory deseases. [Pedro Ribera]	Rejected
8-289	A	14:28	14:36	I suggest to live only the references, the discussion is too local [vincenzo artale]	Rejected. It is a good example of the high resolution climate model.
8-290	A	14:28	14:36	while this is an interesting example, a more fundamental evaluation of the effect of resolution is warranted here. Section 8.2.1.2 is out of balance with 8.2.2.2; the latter is far too long, the former misses the main point. [Philip Mote]	Noted. 8.2.2.2 will be shortened.
8-291	A	14:28		This paragraph could be rewritten in a more logical order. Keep all the regional effects together (move last sentence next to the one on Hawaian Island effect) and then move on to global effects. Are there other examples of improvements at the global scale which would make the progress more compelling ? [Ileana Bladé]	Noted. See the text.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-292	A	14:29	14:30	Explain far-reaching effect of the Hawaiian Islands [Marisa Montoya]	Noted. See the text.
8-293	A	14:29	14:29	"… far reaching effect of the Hawaiian islands…" Need to state specifically what this is. [Kevin Walsh]	Noted. See the text.
8-294	A	14:38	14:47	"parametrizaion" is I believe the most widely used spelling: lines 47, 43x2 and 47 should be corrected. [Richard Allan]	Noted Accepted.
8-295	A	14:38		Section 8.2.1.3 This section currently lacks any mention of the recent and relevant work of Shutts and Palmer (e.g. "Representing model uncertainty in weather and climate prediction" Author(s): Palmer TN, Shutts GJ, Hagedorn R, Doblas-Reyes E, Jung T, Leutbecher M. Source: ANNUAL REVIEW OF EARTH AND PLANETARY SCIENCES 33: 163-193 2005) on new stochastic-dynamic schemes to represent unresolved physical processes, although it is alluded to in Chapter 10 (section 10.5.1) in the context of representing/sampling uncertainty in climate model predictions. [Timothy Johns]	Noted, but a space is limited and all references cannot be included.
8-296	A	14:39	14:39	Please check whether "radiational" is a legitimate adjective. [William Collins]	Text modified.
8-297	A	14:48		Suggest modification: "The climate system includes a variety of physical mechanisms, such as cloud, radiation and boundary layer processes, which interact…" [Richard Allan]	Accepted.
8-298	A	14:50	14:50	Add reference to Martin et al., 2000 [Mon Weath Rev, 128, 3200-3217], where the tests of the new boundary layer scheme in HadAM3 are shown. [Gill Martin]	Accepted.
8-299	A	14:53	14:55	These lines enumerate some of the cloud processes that affect climate. One of these refers to the regulation of the flow of radiation at the top of the atmosphere. This might be understood as the only level where clouds have radiative impact. I consider relevant to point out that clouds affect the distribution of radiative fluxes at any level within the atmosphere, not only at the top of the atmosphere. [Alejandro Bodas-Salcedo]	Accepted. Text modified.
8-300	A	15:1	15:1	No year in the reference [Eugene Rozanov]	Accepted.
8-301	A	15:1	15:1	1990?. To che checked. [Philippe Tulkens]	Accepted.
8-302	A	15:4	15:4	Is the curiously defensive statement on that cloud parameterizations are not simply "collections of adjustable parameters" needed? (who said they were?). [Keith Dixon]	Noted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-303	A	15:6	15:8	It is strange to see TOGA-COARE listed as a project conducted to include cloud parameterizations, when it included virtually no cloud data except for precipitation and lightning, while other field experiments and long-term programs designed specifically with cloud parameterization in mind, such as FIRE and ARM, are not mentioned. [Anthony Del Genio]	Accepted. Text modified.
8-304	A	15:28	15:29	The statement that several more decades are needed for global cloud resolving models to be applied to full climate simulations is very discouraging in view of lines 1 and 2 that state realistic parameterizations of cloud processes are considered to be essential to produce good climate simulations and reliable projections. If global cloud resolving models are not necessary to satisfy the stated requirement in lines 1 and 2 this should be stated explicitly. [Robert Molinari]	Text modified.
8-305	A	15:31	15:36	If the role of aerosols will be discussed further in 8.2.5, this paragraph can be deleted. [Robert Molinari]	Rejected. This is a brief introductory survey.
8-306	A	15:32	15:32	specified BUT fully interactive aerosols …. [Ileana Bladé]	Accepted. Text modified.
8-307	A	15:32	15:33	GFDL_CM2 models, as identified in Table 8.2.1, do not contain fully interactive aerosol models, as stated here. Aerosols are specified from a chemical transport model, and only their direct effects are treated. As of this writing, a treatment of warm-cloud aerosols has been incorporated in the GFDL atmospheric component model only. [Leo Donner]	Accepted. Text modified.
8-308	A	15:45	15:47	Give reference for the distortion of water masses [Marisa Montoya]	Accepted
8-309	A	15:46	15:47	this is the only place in which the authors mention the marginal seas, may be useful to continue to undertake their importance in other part of the chapter [vincenzo artale]	Noted. General water mass characteristics are discussed in 8.2. More is precluded by lack of literature on these aspects of AR4 models.
8-310	A	15:52		typo: "that" should be "than" [Philip Mote]	Accepted
8-311	A	16:1	16:2	Capturing the salinity adequately at river mouths is a problem not only for models using a virtual salt flux [Marisa Montoya]	Rejected. Agree with comment but the sentence describes a specific advance.
8-312	A	16:11	17:4	Some attention to the role of eddies in the ocean could be given here. Eddies are of large importance to the ocean circulation. Their impact on the meridional mean circulation and ventilation is most pronounced in the Southern Ocean and in the tropics. In the Southern Ocean eddies compensate for a large part the Deacon Cell (Doos and Webb, 1994, The	Rejected. Discussion of specific studies with eddy permitting coupled models is given. Broader, tutorial material is beyond the scope of the

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Deacon Cell and other meridional cells in the Southern Ocean, J. Phys. Oceanogr. 24, 429-442) and in the tropics the near equatorial Tropical Cells are compensated by eddy transports associated with tropical instability waves (Hazeleger et al, 2001, Do tropical cells ventilate the Indo-Pacific equatorial thermocline? Geoph. Res. Letters., 28, 1763-1766). Since eddies are not resolved in most CGCMs they need to be parameterized. [Wilco Hazeleger]	chapter.
8-313	A	16:11		Section 8.2.2.2 seems somewhat disjointed, and not a very coherent discussion of the benefits realized by recent improvements in ocean model resolution. Vertical resolution improvements are largely ignored, though colleagues have told me that the kind of increased vertical resolution seen in the several of the AR4 coupled models was a prerequisite for adding ocean biogeochemical model components. [Keith Dixon]	Noted. Discussion of vertical resolution will be added to the extent that literature is available.
8-314	A	16:15	16:27	the first paragraph says eddy-resolving models have hardly been used, the second paragraph rephrases that statement. The paragraphs seem to be redundant. This helps contribute to a section that is much too long. [Philip Mote]	Taken into account in redrafting.
8-315	A	16:29	16:39	8-16, lines 29 – 39, the timescale of MOC response is determined by vertical mixing strength. [Wenju Cai]	Rejected. MOC response is not what is being discussed.
8-316	A	16:33	16:33	Mention standard resolution in the HADCM3 model [Marisa Montoya]	Rejected. Think this is clear.
8-317	A	16:34	16:41	The terms MOC and THC are used just a few lines apart. Elsewhere this draft says that they are the same thing. Chose one term or the other, preferabley MOC, or explain why the terms are used the way they are. [Lenny Bernstein]	Noted. Terms are used correctly, both are now defined in Glossary.
8-318	A	16:41	16:46	Why the potential benefits resolving the Bering Strait and Canadian Archipelago are mentioned but not other important straits (e.g., Indonesian Throughflow) seems odd. Would it be better to first make a more general statement about the role of relatively narrow passages in the circulation of the world ocean and then, if deemed necessary, give examples? [Keith Dixon]	Accepted
8-319	A	16:53	16:53	Text refers to Fig. 8.2.1 to show a change due to increased radiative forcing. Fig. 8.2.1 does not show responses to changes in radiative forcing. [Leo Donner]	Accepted
8-320	A	17:0	19:	section 8.2.3: too long. Should be no longer than atmosphere section, or shorter [Philip Mote]	Accept - text significantly modified to meet space limitations

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-321	A	17:7	18:18	Vertical mixing is the most uncertain of oceanic parameters. This should be mentioned explicitly. [Marisa Montoya]	Rejected. Not sure this is true. (and no evidence provided)
8-322	A	17:7	18:18	Recently physically based parametrizations of vertical mixing based on the dissipation of tidal energy have been implemented (Simmons et al. 2004). This should be mentioned. Reference: Simmons, H. L., Jayne, S. R., St Laurent, L C., and Weaver, A. J. 2004: Tidally driven mixing in a numerical model of the ocean circulation. Ocean Modelling 6: 254-263. [Marisa Montoya]	Accepted.
8-323	A	17:12	17:13	The sentence "Representation …." might refer to Noh et al. 2002. (Noh Y, Jang CJ, Yamagata T, et al.: Simulation of more realistic upper-ocean processes from an OGCM with a new ocean mixed layer model, JOURNAL OF PHYSICAL OCEANOGRAPHY 32 (5): 1284-1307 MAY 2002) [Toshiyuki Awaji]	Will check and include ref if it supports the sentence, otherwise will remove the sentence.
8-324	A	17:15	17:15	YET, while there have been … [Ileana Bladé]	Rejected. No onconsistency with previous sentence is implied.
8-325	A	17:15	17:18	To what extent are the model predictions invalidated by the failure of models to account for ocean mixing correctly? This difficult issue is downplayed and somewhat neglected in this chapter. [Chris Garrett]	Rejected. Four studies are cited here which discuss this issue. The question as directly posed by the comment has not been answered but caveat is implicit in the text.
8-326	A	17:29		The following section details the incorporation of carbon cycle processes in land surface models. A parallel development has taken place in ocean components. Excluding even a mention of this (perhaps with a cross-reference to chapter 7) would creates a misperception of the state of model development. [Frank Bryan]	Accepted.
8-327	A	17:30	17:30	Would it be appropropriate to mention here the parameterization of barotropic tidal mixing effects? If so, a reference could be Lee, H-C., A. Rosati, and M.J. Spelman, "Barotropic tidal mixing effects in a coupled climate model: Oceanic conditions in the Northern Atlantic", Ocean Modelling, (2006), Vol. 11, pp. 464-477. (PDF available from http://nomads.gfdl.noaa.gov/CM2.X/references/) [Keith Dixon]	Accepted.
8-328	A	17:40	17:40	Betts et al. (2004) not listed in references. [Alejandro Bodas-Salcedo]	Reference now removed
8-329	A	17:40		Betts et al. is not in the list of references. I didn’t check many references so I don’t know if others are missing. [Philip Mote]	Reference now removed

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-330	A	17:43	17:43	CAUSING further warming" or "ENHANCING WARMING FURTHER [Ileana Bladé]	Accept – text modified
8-331	A	17:52		"requires" implies that models do not simulate any regional warming without high resolution. Is the point that regional detail requires smaller-scale processes? [Philip Mote]	Accept – text modified
8-332	A	18:6	18:6	uncertainty CONCERNING the role [Ileana Bladé]	Accept – text modified
8-333	A	18:18	18:25	The land surface processes are strongly controlled by the physiological activities of vegetation. The physiology-based ecosystem modelling has been advanced significantly in last decade. Unfortunately, the current climate models have not taken full advantage of this advancement. The development of the physiology controlled water balance equation (Eq. 7 in Wang et al., 2002), the dynamic coupling of this water balance equation with surface energy balance equation (Eq. 8 in Wang et al., 2002), and the simulation of root dynamic growth (Wang et al., 2001) and its impact on canopy transpiration through root water uptake (Fig 4 and Fig 5, and Eq. 3-6 in Wang et al., 2002), are significant progress in addressing the physiological dimension of the land surface process and its modelling and should be included in this paragraph. [Shusen Wang]	Reject – this is process based comment that is more appropriate for Chapter 7
8-334	A	18:20	18:20	Based on the above comment, I suggest adding, after "2004), "coupling dynamic water balance equation with surface energy balance equation (Wang et al., 2002)". (Wang, S., Grant, R.F., Verseghy, D.L., and Black, T.A. 2002, Modelling carbon-coupled energy and water dynamics of a boreal aspen forest in a General Circulation Model land surface scheme. International Journal of Climatology 22: 1249-1265.) [Shusen Wang]	Accept – despite comment above – there is value in adding this reference
8-335	A	18:20	18:20	The model development of soil carbon and nitrogen processes and plant carbon and nitrogen processes is a key component in ecosystem models for carbon simulation and it has attracted more and more attention in climate models. The model development of soil and plant carbon and nitrogen by Wang et al. (2001, 2002) in the Canadian Land Surface Scheme (CLASS) is still one of the robust algorithms found in literatures. I suggest to add, after "high latitude organic soils", "and the soil carbon simulations (Wang et al., 2002).". (Wang, S., Grant, R.F., Verseghy, D.L., and Black, T.A. 2002, Modelling carbon dynamics of boreal forest ecosystems using the Canadian Land Surface Scheme. Climatic Change, 55: 451-477.) [Shusen Wang]	Accept – text modified
8-336	A	18:22	18:	I would like to suggest that a sub-grid scale soil heterogeneity on Horton and Dunne runoff parameterization (Liang et al., 2001; Xie et al., 2003, Liang et al., 2003a) is included in line 21 of the page 8-18. Also, I would like to suggest that the papers be cited	Reject – I tried to select one or two papers that were formulative – there is no space to list all those that have

No.	Batch	Page:line From	Page:line To	Comment	Notes
				such as Yang and Xie (2003), Liang et al. (2003) in the line 22. Cited papers here: Liang X., Z. Xie, 2003a: Important factors in land-atmosphere interactions: surface runoff generactions and interactions between surface and groundwater. Global Planetary Change, 38,101-114. Liang X., Z. Xie, 2001: A New Surface Runoff Parameterization with Subgrid -Scale Soil Heterogeneity for Land Surface Models. Advances in water resources, 24(9-10), 1173-1193. Xie Z., F. Su, X. Liang, Q. Zeng, Z. Hao, Y. Guo, 2003: Applications of a surface runoff model with Horton and Dunne runoff for VIC. Advances in Atmospheric Sciences. 20(2), 165-172, 2003. Yang H., Z. Xie, 2003: A new method to dynamically simulate groundwater table in land surface model VIC, Progress in Natural Progress,13(11), 819-825. [Zhenghui Xie]	contributed in this area
8-337	A	18:27		which are the two basins? Only the Rhone is mentioned, that I could find. [Philip Mote]	Accept – text modified
8-338	A	18:36	18:37	This sentence could be interpreted as suggesting that the two models mentioned are the only AR4 coupled models to include a river routing scheme, though that is not the case, unless I'm misinterpreting what is meant by "river routing scheme". Perhaps they are the only two to use the scheme reported on in Oki and Sud (1998) or the only ones to have published papers that specifically compare results with and without such a scheme. [Keith Dixon]	Accept – text modified by deleting sentence since the major advances were for the TAR not AR4
8-339	A	18:36	18:38	Text lists just two models with river routing incorporated, while Table 8.2.1 indicates that all but a couple of the models used in this assessment have routing. [Leo Donner]	Accept – text modified
8-340	A	18:38	18:38	Kanae (2005) is not included in the reference list. [Marisa Montoya]	Rejected (sentence now deleted)
8-341	A	18:38		this sentence says nothing. Suggest deleting or expanding [Philip Mote]	Accept – text deleted
8-342	A	18:40		A very powerful tool with which climate models can be evaluated is the extent to which they are able to simulate the spatial and temporal variability of isotopes, especially stable water isotopes, in the various parts of the climate system. The distribution of these in the biosphere, atmosphere, ocean and cryosphere depends upon many processes including atmospheric transport, cloud, condensation, evaporation, moisture representation, temperature, fractionation etc. All the atmospheric components associated with these must be simulated with reasonable accuracy before accurate water isotope distribution can be obtained (and interpreted). I was surprised that under this Chapter on evaluation there was virtually no mention of this 'modern' component of evaluation (except in this short paragraph). The issue is of relevance to each of the Atmospheric Processes, Ocean(ic)(?) Processes, Terrestrial Processes, Cryospheric Processes, and Aerosol Modelling and	Accepted – some text is added elsewhere that highlights isotopes, but since the AR4 models are not isotope enabled we cannot use isotopes in evaluation at this time.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Atmospheric Chemistry subsections in the Chapter. It could be suggested that a smallish subsection be devoted to this issue with the appropriate coverage of the recent literature. [Ian Simmonds]	
8-343	A	18:42	18:42	The phrase "more recent biophysically based models and very recent biophysically based models" is nearly meaningless. [Leo Donner]	Reject – the reference explains the terminology
8-344	A	19:9		“contribution” in what sense? the latent and sensible heat fluxes? Be explicit [Philip Mote]	Reject – there is no space to expand tis assessment – and this statement reflects our assessment
8-345	A	19:12	19:13	The first sentence of this section seems too generally worded for a section specifically concerning soil moisture feedbacks. Omit or move? [Timothy Johns]	Accept – sentence deleted
8-346	A	19:16	19:18	too many references. [Philip Mote]	Accept – several now deleted.
8-347	A	19:16		To present a SH case it may be worth adding the study of Simmonds, I., and P. Hope, 1998: Seasonal and regional responses to changes in Australian soil moisture conditions. International Journal of Climatology, 18, 1105-1139. [Ian Simmonds]	Reject – this is pre-TAR and the other pre-TAR references in this paragraph have been deleted.
8-348	A	19:19	19:19	Schubert et al., 2004a, Reference is missing under References [Christoph Frei]	Accept – reference added
8-349	A	19:36		say more about the figure or omit [Philip Mote]	Reject – no space and the details can easily be sourced from the references
8-350	A	19:43	19:43	Need to be specific about which Hadley model [Catherine Senior]	Accept – text modified
8-351	A	19:54	20:2	Neither the Robock [should be et al.] (2000) nor the Reichle [should be et al.] (2004) references are in the reference list. The correct references are Robock, Alan, Konstantin Y. Vinnikov, Govindarajalu Srinivasan, Jared K. Entin, Steven E. Hollinger, Nina A. Speranskaya, Suxia Liu, and A. Namkhai, 2000: The Global Soil Moisture Data Bank. Bull. Amer. Met. Soc., 81, 1281-1299. and Reichle, R. F., R. D. Koster, J. Dopng, and A. A. Berg (2004), Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation, J. Hydrometeorol., 5, 430–442. [Alan Robock]	Accept – reference added – hte Reichle reference is deleted following 8-353
8-352	A	19:54	20:2	It is not correct that nobody has evaluated climate models with observations. Srinivasan et al. (2000) showed that GCMs did not do a good job of simulating soil moisture and Li et al. (2005) showed that not even reanalysis does a good job of simulation soil moisture. Our latest work shows essentially the same for AR4 models, but the paper has not been	Reject – the text is explicit “post TAR” and these references are pre-TAR.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				completed yet. refs: Srinivasan, G., Alan Robock, Jared K. Entin, Lifeng Luo, Konstantin Y. Vinnikov, Pedro Viterbo, and Participating AMIP Modeling Groups, 2000: Soil moisture simulations in revised AMIP models. J. Geophys. Res., 105, 26,635-26,644. and Li, Haibin, Alan Robock, Suxia Liu, Xingguo Mo, and Pedro Viterbo, 2005: Evaluation of reanalysis soil moisture simulations using updated Chinese soil moisture observations. J. Hydrometeorol., 6, 180-193. [Alan Robock]	
8-353	A	19:56	20:1	Mentioning the results of Reichle et al. (2004) here is misleading. There results are really a statement about the ability of remote sensing of soil moisture and alnd surface data assimilation schemes, and this is a separate issue from whether models can do a good job of simulating soil moisture. [Alan Robock]	Accept – text modified
8-354	A	20:0		section 8.2.4.2: shorten, and refer to section 4.4.2 (sea ice) [Philip Mote]	Accepted. Text modified.
8-355	A	20:6		Section 8.2.4.1. For completeness, perhaps you could mention glaciers in here. They are not currently coupled interactively in any AOGCM (as far as I know) because they are very sub-gridscale and probably do not have a large climate feedback on large scales. Chapter 10 describes offline models. [Jonathan Gregory]	Accepted. Text added.
8-356	A	20:7	20:8	Ice-sheet models have been included in EMICs. This should be mentioned here. One example is: Calov, R., Ganopolski, A., Petoukhov, V., Claussen, M., Greve, R., 2002: Large-scale instabilities of the Laurentide ice sheet simulated in a fully coupled climate-system model. Geophys. Rev. Lett., 29 (24), 2216 doi:10.1029/2002GL016078. [Marisa Montoya]	Accepted. Text modified.
8-357	A	20:10	20:11	Missing reference: "Ridley et al., ()" [Richard Allan]	Noted
8-358	A	20:19	20:21	A similar lament can be expressed for sea ice albedo which typically is rather arbitrarily prescribed with only crude dependence on ice thickness, snow cover, and puddling effects. [Andrew Lacis]	Accepted. Text modified.
8-359	A	20:20	20:20	sea ice salinity is prognostic in GISS-ER and GISS-EH (schmidt et al 2004) [Gavin Schmidt]	Accepted. Text modified.
8-360	A	20:33	20:36	A candidate to add to the list of sea ice models that have variable heat capacity, though not as sophisticated as some that are cited, is Winton's SIS model used in the GFDL CM2.x AOGCMs. Winton, Michael, 2000: A reformulated three-layer sea ice model. Journal of Atmospheric & Oceanic Technology, 17(4), 525-531. (PDF available at http://www.gfdl.noaa.gov/reference/bibliography/authors/winton.html)	Rejected due to space limits

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Keith Dixon]	
8-361	A	20:38		"Snow models…" Why only discussed in context of sea-ice? Just as important for land. Snow in some land models already includes compaction. [Robert E. Dickinson]	Accepted. Text added.
8-362	A	20:44	20:44	similarly, snow ice formation (with distinct salinities) is incorporated into GISS-ER and GISS-EH. [Gavin Schmidt]	Accepted. Text modified.
8-363	A	21:18	21:19	Not clear to me what is meant here. What are these "functions normally included in the sea ice component" [Anthony Hirst]	Accepted. Text modified
8-364	A	21:44	21:47	GFDL_CM2 models listed in Table 8.2.1 do not contain an interactive aerosol sub-component model, as stated here. Aerosols in those models are specified from a chemical transport model. [Leo Donner]	Accepted. Text modified.
8-365	A	21:44	21:46	The GISS ModelE-H and ModelE-R also have the ability to simulate aerosol changes interactively, although aerosol amounts were held fixed in the IPCC simulations [Ron Miller]	Accepted. Text modified.
8-366	A	21:49	21:56	I think it is important to show the progress with Chemistry-Climate models, this kind of models are developing rather fast. [Eugene Rozanov]	Noted
8-367	A	21:50	21:50	There is no such such thing as the CTM, rather is is a type of model like e.g. MOZART or STOCHEM [Rolf Müller]	Rejected. Such models do exist. Reference added.
8-368	A	21:53	21:53	Which interaction, which process? Sulphur chemistry (i.e. chemistry --> aerosol) or heteorogenous chemistry on aerosol particles(i.e. aerosol --> chemistry) [Rolf Müller]	The text has beeen reworded.
8-369	A	22:6	22:28	this section is very relevant and deserve more references and discussion [vincenzo artale]	Noted.
8-370	A	22:6		Section 8.2.6 I think it would be worth recording the progress made towards improving/standardising the software infrastructures to support coupling in Earth System models, specifically in the EU PRISM project (Reference: S. Valcke, E. Guilyardi, C. Larsson, 2005. PRISM and ENES: A European approach to Earth system modelling. Concurrency Computat.: Pract. Exper., 17:1-16), which builds on the OASIS coupler. [Timothy Johns]	Agree. Reference added.
8-371	A	22:8	22:8	Not being a tautology afficionado, I would think one could improve on the sentence that reads 'A "coupler" couples the various components of a climate model.'	Agree. Text modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Keith Dixon]	
8-372	A	22:21		In nature the components are coupled instantaneously, and internal gravity waves are excited by the wind, providing a key source of energy for mixing. The sentence here suggests that this is a flaw in the model. Should perhaps clarify that high-frequency forcing may excite IGWs that are poorly resolved. [Frank Bryan]	Agree. Text added.
8-373	A	22:24	22:25	This is a surprising statement. My understanding is that performing ensemble realizations allows for a better sampling of the internal variability. Initial conditions can't be important on these climatic time scales. [Ileana Bladé]	Noted. Text deleted.
8-374	A	22:24	22:28	this paragraph seems to fit better with the next section [Philip Mote]	Noted. Text deleted.
8-375	A	22:30		Section # 8.2.7 provides an unbalanced argument about the use of flux adjustments. It should be recognised that not using flux adjustments can cause large biases in SST, say. Therefore, some modellers still opt to use flux adjustments where they feel that the the reduction in SST biases outweighs the need to avoid flux adjustment e.g. studies of regional climate change. Collins et al (2005) provides a strong justifcation for the use of flux adjustemnts in their work. Reference is Collins, M., Booth, B.B.B., Harris, G., Murphy, J.M., Sexton, D.M.H., Webb, M. Towards Quantifying Uncertainty in Transient Climate Change. Clim Dyn, submitted. I would add a couple of sentences at the end of the first paragraph something like: "However, by not using flux adjustments, the models develop worse regional SST biases. Collins et al 2005 still use flux adjustments because they believe that the reduction is regional SST biases is the most important consideration in their study of regional climate change." Of course, there may be other studies still using flux corrections but I am not aware of these. [David Sexton]	Noted. Some text was modified, but the elimination of flux adjustments in some model and the small climate drifts found in those models, is one of the major new findings of AR4.
8-376	A	22:32	22:57	Connecting with comment #1, the fact that climate drift remains an issue should be mentioned here with some explanation. [Marisa Montoya]	Noted. Some text has been added addressing this issue.
8-377	A	22:40		The references to the World Ocean Atlas should be corrected. The "Levitus" data set is an analysis, not the data itself. [Frank Bryan]	Agree. Text modified.
8-378	A	22:48		what does this mean? [Philip Mote]	Agree. Text clarified.
8-379	A	22:52		While the initialization procedure may be similar, I do not think the outcome (the I.C. of the transient experiments) could be classified as "relatively consistent'. Even for a single	Agree. Text deleted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				model, drift in the control could produce quite incosnsistent I.C.s if they are sampled at long (multi-century) intervals. The TOA imbalance across the models can be wildly different depending on the particular choices made in the models and ocean I.c.s, again not 'relatively consistent'. [Frank Bryan]	
8-380	A	23:0	28:	Section 8.3.1 lacks references to relevant publications. There have been a large number of analyses of model performances since TAR, especially with IPCC AR4 models, as evidenced by the March 2005 Workshop in HI (I believe the PCMDI website or the IPCC WG1 Office has a list of these recent model evaluation papers submitted before May 31, 2005). [Aiguo Dai]	Rejected: We are not writing a review and it would therefore be inappropriate to cite all papers that include a model evaluation aspect. We will check that we haven't missed any critical studies.
8-381	A	23:0	28:	Section 8.3.1 also seems to have too much on what models should do (i.e., general knowledge) and too little on how models perform (i.e., specific results from model evaluation studies). [Aiguo Dai]	Rejected: We attempted to discuss those aspects of performance that might most directly affect projections (i.e., fundamentals).
8-382	A	23:0	39:	I regret that in this part the sensitivity of climate models to turbulent mixing in stratified conditions (over land and ice) is completely overlooked, in particular the achievements within the GEWEX Atmospheric Boundary Layer Study, GABLS (see Holtslag, Bound. Lay. Meteorol. 2006, in press and cited references for more information)! [Albert A. M. Holtslag]	Accepted. Reference added.
8-383	A	23:1		Section 8.3, with its numerous subtopics and figures, certainly poses a challenge to the authors who clearly are trying to provide a look at a range of model simulation features in an efficient manner. The presentations of the numerous models' results, often compared to observational estimates, is generally well done, though it is a bit surprising how few of the figures have direct counterparts in the TAR. That the comparisons document the many AR4 model results and available observations, but do not compare the TAR generation of models with the AR4 generation is understandable, but will nonetheless disappoint some. As an aside, if I'm not the only one who would find such inter-generational comparisons informative, perhaps this is an argument for maintaining PCMDI's AR4 WG1 archive for several years so that such inter-generational comparisons could be done for (dare I mention it?) the AR5 report. [Keith Dixon]	Noted.
8-384	A	23:1		Evaluation of mean climate, in particular of the atmosphere: I have trouble to see why the models are evaluated almost exclusively based on annual means. Stainforth et al (Nature 2005) have shown that annual mean climatology provides hardly any constraint on a model or its future projection, i.e. a lot of the errors tend to cancel in an annual mean. See e.g. Knutti and Meehl, submitted to J.Climate for an evaluation of the same dataset based	This is covered in Sections 8.4 and 8.5. Taken into account. Note that other sections evaluate other aspects of climate. Agree that compensating errors can be a problem in some

No.	Batch	Page:line From	Page:line To	Comment	Notes
				on seasonal means, showing that on seasonal means model performance is much more obvious. I'm pretty sure looking at winter and summer would reveal many more details on model performance also on the AR4 models. For example, what is the meaning of looking at annual mean precipitation in mid latitudes (e.g. Europe) when winter precipitation is determined by storms and summer precipitation is mostly convective? Models might have both processes wrong but the annual mean right. I suggest showing less diagnostics but at least for a few show winter and summer. In the figures where errors are given, an alternative would be to give the mean (or sum) of the winter and summer error. I imagine that would make the point more clear that a few models are really performing much worse than the rest. [Reto Knutti]	regions, but even in the annual mean, model biases are evident. We might prepare winter and summer figures for some selected fields to be included as supplementary material. We chose not to focus on individual model performance (partly because the relevance of skill to future projections has not yet been established), so designing figures to highlight such differences is not a high priority.
8-385	A	23:1		Section 8.3. The aims of the section are stated as "to highlight where models generally perform well and to identify their deficiencies" and to "quantify the evolution in model skill that has been seen over the last several years". The latter is not discussed in several sub-sections, although there is a separate section 8.3.5 which looks at this in an overall statistical sense. I feel that some mention of how the model evaluations of each the different climate characteristics have changed in recent years would be very helpful to the reader. [Gill Martin]	Noted. We agree that it would be helpful, but the model output database for earlier model versions will not easily support such an analysis
8-386	A	23:16	23:17	What is the difference between a “new” model and one that has been improved since the TAR? [Philip Mote]	A new model is not necessarily an improved model.
8-387	A	23:23	23:23	Reference to section 8.2 should perhaps be 8.5? [Gill Martin]	Reject. The reference is correct.
8-388	A	23:23	23:23	See comment #8. I think this would read better as '...various aspects of unforced variability and more mechanistic criteria (see 8.2 and 8.6)' [Catherine Senior]	Taken into account. This sentence was eliminated in meeting the length guidelines.
8-389	A	23:25	8:29	If analysis showed that a set of models was more reilable to use for predictions, would it not be an international service to point that out. I don't think that the scientific community is serving itself well by not using their skill to identify models that, through better representation of physical processes, actually make more reliable predictions. This is part of the argument that we as a community are gaining skill in our predictions - that through analysis and use of observations we can develop better models, identify outliers and reduce uncertainty in prediction. On the other hand, obviously if no model represents a particular process well, or produces an (observationally) anticipated response, then this is something that needs to be pointed out as well. [Anne Douglass]	Rejected. Our assessment is that in the area of climate, metrics determining the reliability of predictions have not yet been established. Although we agree that there are some obvious shortcomings of individual models, how much those matter is rarely known.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-390	A	23:25	23:29	This paragraph might point out that currently we cannot quantify these improvements, due to the infancy of developing well agreed model metrics. However, this is considered a fruitful avenue for the future. [Catherine Senior]	Taken into account. This issue was touched on in the first paragraph of this section, but space limitations preclude expanding the discussion
8-391	A	23:37	23:37	The phrase 'chose to archive' suggests that some centres made a conscious decision not to archive particular diagnostics, possibly due to known errors in their model. I don't believe this is the case for any modelling centre. In most cases, a diagnostic will not have been submitted from a particular centre either because they did not have code in their model to produce the diagnostic, or because the diagnostic was saved but problems were encountered when trying to convert it to the required format, or because the diagnostic was saved but has not been submitted yet due to a lack of time/resources. I suggest replacing 'chose to archive' with 'submitted' and add 'in sufficient time' to the end of the sentence. [Keith Williams]	Accepted (but without the phrase 'in sufficient time').
8-392	A	23:40	23:42	?accid writing. Keep it short and clear [Philip Mote]	Accepted. Moved, shortened and revised.
8-393	A	23:46	29:32	Section 8.3.1. There is little indication in this section of the range of error in the observations. The certainty of the observations is critical when they are being used to evaluate the models. I would have expected observational error to be potentially especially significant in evaluation based on Figures like 8.3.5 through 8.3.9, where the observations are based on only a five-year ERBE period (1985-1989). Such a short period allows the possibility of substantial sampling error. Some idea of the sampling error could be estimated from the interannul variability between the years within the ERBE period. Errors from sampling and from other sources (e.g., instrumential and algorithmic limitations) need to be considered and included, e.g., as error bars in the appropriate Figures. [Anthony Hirst]	Rejected. Except in rare cases, published error estimates do not seem to be available. When available, they often omit estimates of systematic errors, which often dominate, we suspect. When possible, we shall try to compare to two different observational datasets.
8-394	A	23:52		Missing "a" at the end of the line. Also, the reponse of the surface to solar heating for example evaporation, must also influence the surface temperature [Richard Allan]	Accepted. Sentence will be modified to include the other surface heat fluxes.
8-395	A	24:6	24:7	“Away...” is a convoluted sentence, although the quantitative final phrase is great. What is the source of the “observed” field and what is the connection between the error field and the sparseness of observations? [Philip Mote]	Accepted. Reworded. Observed field source is noted in figure caption. Will note the sparseness of data at high latitudes.
8-396	A	24:8		“typically” misspelled [Philip Mote]	Accepted.
8-397	A	24:21		missing comma after "Still"	Taken into account. This sentence was

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Richard Allan]	reworded.
8-398	A	24:28	24:28	The largest periodic cycle is the diurnal cycle (min to max ampl. up to 40deg.C) over Northern Africa, Tibet and other low-latitude high terrain, it is not the annual cycle over these regions. [Aiguo Dai]	Taken into account. Sentence rewritten.
8-399	A	24:28	24:38	This paragraph generally proclaims the high quality of surface temperature simulated by the multi-model mean. However, there are numerous regions where the multi-model mean error is a very large fraction of the observed standard deviation (e.g., of the east coasts of North America and Asia and over much of South America). These deficiencies should be noted in the text. [Leo Donner]	Accepted. Deficiencies of this kind will be noted at the end of the paragraph.
8-400	A	24:36	24:36	More accurate would be to say "over eastern Siberia", as there appears an overestimate of the range in western Siberia in Fig. 8.3.2 [Anthony Hirst]	Accepted. text reworded.
8-401	A	24:42	24:53	The diurnal cycle in convective cloud is well known to be poorly simulated in models and is likely to influence and be influenced by errors in diurnal temperature range [Richard Allan]	Accepted. Note to that effect added.
8-402	A	24:42	24:	I think it should be more directly acknowledged that the amplitude of the diurnal cycle is seriously underestimated at low and high latitudes. [Ileana Bladé]	Rejected. Already essentially stated.
8-403	A	24:42	28:53	Please note that Dai and Trenberth (2004, J. Climate) showed observed and CCSM2-simulated diurnal amplitude and phase of surface air temperature over the globe. They found that the temperature diurnal cycle is generally well represented by the CCSM2 over land, but over the oceans models tend to have no SST diurnal cycles which results in weak diurnal cycles in marine air temperature and other fields. I think the lack of diurnal cycles in SST and the associated weak diurnal cycle in marine boundary layer is a common error in all models that need to be addressed here. [Aiguo Dai]	Rejected. The cited paper is an analysis of a single model and there is not enough space to discuss this aspect of a single model. We will, however, try to reconcile the apparent discrepancy between the results shown in fig. 8.3.3 and the results in Dai and Trenberth which shows that the predecessor to NCAR's IPCC model does not appear to radically underestimate diurnal temperature range over land.
8-404	A	24:49		how do dry surface conditions suppress daytime cooling? doesn’t the desert typically warm until just before sunset? this sentence is unclear. [Philip Mote]	Taken into account. This sentence was removed due to space limitations. (Note that "cooling by evaporation and transpiration" refers to the heat lost by the surface which tends to keep the

No.	Batch	Page:line From	Page:line To	Comment	Notes
					temperature from rising as much during the day.
8-405	A	24:52	24:53	number agreement between “parameterizations” and “is” [Philip Mote]	Accepted.
8-406	A	25:0		. As Richard Goody puts it in his 2000 QJRMS paper: “Our knowledge of atmospheric sources and sinks of entropy is evidently poor and, even if they were known, it would be difficult to construct a climate model with sources and sinks as precise as Johnson considers to be necessary.” [Richard Anthes]	Noted. [part of this comment seems to be missing.]
8-407	A	25:0		Goody, R., 2000: Sources and sinks of climate entropy. Q.J.R. Meteorol. Soc., 126, 1953-1970. Johnson, D. R., 1997: ‘General coldness of climate models’ and the Second Law: Implications for modeling the earth system. J. Climate, 10, 2826-2846. Johnson, D. R., A. J. Lenzen, T. H. Zapotocny, and T. K. Schaack, 2000: Numerical Uncertainties in the Simulation of Reversible Isentropic Processes and Entropy Conservation. J Climate, 13, 3860-3884. Johnson, D. R., A. J. Lenzen, T. H. Zapotocny, and T. K. Schaack, 2002: Numerical Uncertainties in Simulation of Reversible Isentropic Processes and Entropy Conservation: Part II. J. Climate, 15, 1777-1804. [Richard Anthes]	Noted.
8-408	A	25:1	25:22	The discussion on temperature gives an impression that these simulations are reasonable and provide little cause for concern regarding future climate projection, touting zonal mean errors of less than 2K against a field spanning more than 100K. This totally omits the issue of trends in temeprature lapse rate in the models, which may be problematic, have proved very difficult to compare against uncertain satellite and radiosonde observations, and have been the topic of extensive discussion. (See, for example, Thorne et al. (2005, Bulletin of the American Meteorological Society, 86, 1471-1476.)) [Leo Donner]	Taken into account. This topic should be covered in chapter 9.
8-409	A	25:8	25:9	Errors … are of special concern because they indicate model shortcomings": this is a commonplace statement. I suggest combining this and the next sentence. "Deficiencies in the simulation of the vertical profile of atmospheric temperature are of special concern as they impact both the surface temperature and ..... [Ileana Bladé]	Accepted.
8-410	A	25:11	25:22	In the same vein as the comments above discussing the importance of referencing	Accepted. There are several possible

No.	Batch	Page:line From	Page:line To	Comment	Notes
				previous work, as well as alternative points of view, a significant omission in this chapter is a reference to Donald R. Johnson's critique of climate models and their "cold bias" and the advantages of isentropic coordinates (Johnson, 1997; Johnson et al., 2000; Johnson et al., 2002. Johnson argues that the non-conservation of moist entropy in virtually all models causes a cold bias in the models. The main emphasis in these works is on developing strategies to assess the accuracies of models with respect to dealing with entropy, which is the primary property that needs to be addressed in sorting out the magnitude of aphysical sources relative to other diabatic processes that ultimately constitute the forcing of atmospheric circulation. The "cold bias" is still there according to lines 12-22 page 8.25 and illustrated in Fig. 8.3.4. The cold bias from all these models throughout most of the atmosphere is remarkable, and Johnson's work and his hypothesis should be discussed here [Richard Anthes]	causes of the cold bias, which cannot be discussed in depth. The non-conservation of moist entropy is one factor, and is now mentioned along with others.
8-411	A	25:16	25:16	"and this is reflected in the mean model error": redundant. [Ileana Bladé]	Accepted.
8-412	A	25:17	25:17	More accurate would be to say "persisted for many years" [Anthony Hirst]	Accepted.
8-413	A	25:26	25:26	tmosphere -> atmosphere [Reto Knutti]	Accepted.
8-414	A	25:26	25:26	A typo! The a from atmosphere is missing. [Hendrik M. van Aken]	Accepted.
8-415	A	25:26		Missing "a" in "atmosphere" [Richard Allan]	Accepted.
8-416	A	25:26		"Atmosphere" missing an A. [Philip Mote]	Accepted.
8-417	A	25:27	25:31	Since the top of atmospheric radiative fluxes are often tuned to agree with ERBE data, the use of zonal mean fluxes to evaluate models is of limited usefulness and prone to generating compensating errors (for example not enough cloud that is too bright but gives a reasonable radiation balance at the top of the atmosphere). The radiation budget data is however useful in evaluating variability and regional behaviour of clouds, water vapour and surface properties simulated by models. [Richard Allan]	Noted. This is true, but space limitations preclude a discussion of this.
8-418	A	25:27	25:27	change "horizontal" to "latitudinal". "seasonally AND meridionally varying pattern" [Ileana Bladé]	Accepted. But only the first part. pattern implies meridionally varying in this case.
8-419	A	25:30	25:30	the impact ON TEMPERATURE	Accepted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Ileana Bladé]	
8-420	A	25:33	25:38	As has been shown recently by Raschke (unpublished, I think, but shown in a recent issue of GEWEX News), models have errors even in basic TOA insolation due to such things as having only 30-day months, a 365-day rather than 365.25-day year, and incorrect equinox timing, that are not much less than 10 W/m2. [Anthony Del Genio]	Noted. In our judgment, these monthly mean differences due largely to incorrect definition of months in models, are not likely to be important for future projections, whereas annual mean biases indicate a systematic error in the representation of the effects of radiation by models, which more likely will affect future projections.
8-421	A	25:33	25:46	The study by Wang et al. (2005) specifically investigated the model performance of 17 GCMs for AR4 in simulating surface albedo of the high latitudes, and compared these model outputs with satellite observations using ISCCP FD. In their publication they discussed the impact of model simulated ice/snow on surface albedo as well as the models albedo treatment regarding to the seasonal changes in vegetation, solar zenith angle, etc. [Shusen Wang]	Noted. Severe space limitations precluded discussion of the surface albedo and eliminated the discussion of shortwave clear sky fluxes, which compared to other factors was judged not to be of primary concern.
8-422	A	25:33		What does this mean, “if not for...”? [Philip Mote]	Taken into account. This paragraph was largely removed to meet length constraints.
8-423	A	25:36	25:36	Footnote 2 -- The footnote should be reworded -- there are strong interactions in blue and UV wavelengths with oxygen ions and compounds from the stratosphere through the thermosphere, so it's not clear what is meant by interactions becoming "trivial". Perhaps one could say where the (Rayleigh) scattering of solar radiation by atmospheric molecules becomes negligible. [William Collins]	Accepted. Reworded.
8-424	A	25:36	25:36	How exactly the top of the atmosphere was defined? [Eugene Rozanov]	Rejected. See footnote 2 in Section 8.3.
8-425	A	25:39	25:39	Based on the above comment, I suggest, in line 39 after "primarily affects higher latitudes", to add "Comparisons of model simulated surface albedo in the northern high latitudes by 17 GCMs with satellite observations showed that most models simulated seasonal albedo variations reasonably well, but large differences existed among the climate models particularly in winter (Wang et al., 2005). Wang et al. (2005) also found that all of the 17 GCMs failed to simulate the large interannual variations in albedo and the systematic decreasing trend over the study period of 1984-1999 as being observed from satellite data". (Wang, S., Trishchenko, A., Khlopenkov, K., Davidson, A., 2005, Comparison of IPCC AR4 climate model simulations of surface albedo with satellite products over Northern Latitudes. Journal of Geophysical Research - Atmospheres	Rejected. See response to 8-421.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				(revised). [Shusen Wang]	
8-426	A	25:46		It may be worth noting that the clear-sky sampling is only a problem if the clear-sky atmosphere is different to cloudy atmosphere. This is generally the case for clear-sky OLR: for example clear regions are generally drier and warmer during daylight hours and colder by night compared to cloudy regions. [recent refs for clear-sky sampling issues for SW and LW respectively: (1) Erlick, C., and V. Ramaswamy (2003), Note on the definition of clear sky in calculations of shortwave cloud forcing, J. Geophys. Res., 108(D5), 4156, doi:10.1029/2002JD002990; (2) Allan, R. P. and M. A. Ringer, Inconsistencies between satellite estimates of longwave cloud forcing and dynamical fields from reanalyses, Geophys. Res. Lett., 30(9), 1491, doi:10.1029/2003GL017019, 2003.] [Richard Allan]	Noted. If space permits, a clarifying statement will be added, with a possible reference.
8-427	A	25:55	26:55	Isn't this agreement merely consistent with the agreement on TOA LW and SW radiative fluxes? i.e., it does not provide an independent test of the models's performance. [Ileana Bladé]	Accepted. This will be noted.
8-428	A	26:3	26:3	Ok, but this seems to disprove the earlier statement (Page 25, line 33) that "if not for clouds … cliamte models should be able to simulate with reasonable accuracy SW radiation". That having clouds degrades the accuracy of the SW simulation is actually not shown any where in this section. [Ileana Bladé]	Accepted. Text will be reworded for consistency.
8-429	A	26:10	26:12	Text states that multi-model RMS error in outgoing SW is less than that of individual models at all latitudes. Fig. 8.3.6 does not show this. HadGEM1 has smaller errors than the multi-model mean around 30N and at most latitudes from EQ to 60S. [Leo Donner]	Accepted. Will be reworded as a generalization, not as invariably the case.
8-430	A	26:12	26:12	delete "at each latitude" and change to "weighted average of the mean-square error AT EACH LATITUDE" [Ileana Bladé]	Accepted.
8-431	A	26:28		avoid “well” and be quantitative - combine with the next sentence [Philip Mote]	Accepted.
8-432	A	26:43	26:47	Perhaps a small thing, but the climate is obviously not in equilibrium in the late 20th century (Fig. 8.3.8), which is contrary to the equilibrium assumption mentioned in the first sentence of this paragraph. [Keith Dixon]	Noted.
8-433	A	26:47	26:56	“well” and “encouraging” - be objective, quantitative [Philip Mote]	Accepted. Reworded.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-434	A	27:1		Omit Figure 8.3.8, or combine with 8.3.14 to tell a more complete story. [Philip Mote]	Accepted.
8-435	A	27:3		Section 8.3.1.2. The section on moisture and precipitation is one where an assessment of how modelling has improved since the TAR would be very helpful. Many of the improvements in climate modelling made in recent years have targetted the hydrological cycle (e.g. changes to boundary layer and convection schemes, cloud parametrisations and even resolution). Precipitation is notoriously hard to model and systematic errors remain in most models. However, Section 8.3.5 suggests that skill in modelling precipitaiton has increased. More detail on how and where these improvements have been made would be most enlightening. [Gill Martin]	Rejected. This kind of in depth analysis is beyond the scope of this assessment.
8-436	A	27:7	27:8	this section needs to be coordinated with section 9.5.3, especially with regard to the statements about surface temperature vs insolation (globally, precipitation shows a temporal response to shortwave but not longwave forcing - Chapter 9, page 48, line 1) [Philip Mote]	Rejected. Concerning what governs the large-scale precipitation pattern, we say that "this is more directly related to temperature than insolation," which does not contradict there being a global temporal response to shortwave forcing (which directly impacts evaporation).
8-437	A	27:8	27:21	adequate to note only once the decrease in precip with latitude [Philip Mote]	Rejected. One introduces the idea, the second expands on it.
8-438	A	27:9	27:9	Please remove the statement about air's "capacity to hold water vapor," which is an incorrect statement of the physics of saturation (see Bohren and Albrecht's Atmospheric Thermodynamics textbook; this is one of their pet peeves), and replace it with a correct statement about the higher vapor pressure in equilibrium with the condensed phase. [Anthony Del Genio]	Accepted. Reworded.
8-439	A	27:11	27:11	"of various sorts" is colloquial [Anthony Hirst]	Accepted. Reworded.
8-440	A	27:20	27:53	Please note that Dai (2005, J. Climate, in revision, pdf available from http://www.cgd.ucar.edu/cas/adai/papers/Dai-cmep-paper.pdf) evaluated various aspects of precipitation characteristics in 18 IPCC AR4 models. He found that most non-flux corrected models still have large deficiencies in simulating tropical rainfall patterns (e.g., the ITCZ and SPCZ) that are related errors in SST fields. Most models produce too much convective and too little stratiform precipitation. They reproduce the percentage contribution (to total precipitaiton) and frequency for moderate precipitation (10-20 mm/day), but underestiamte the contribution and frequency for heavy precipitation (>20 mm/day) and overestimate them for light precipitation (<10 mm/day). The new models still rains too frequently at reduced intensity, and warm-season precipitation occurs too	Accepted. Main points now included.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				early and too frequently during the day, consistent with an previous analysis of NCAR CCSM2 (Dai and Trenberth 2004, J. Clim). The above frequency analysis results are consistent with Sun et al. (2005: Sun, Y., S. Solomon, A. Dai, and R. Portmann, 2005: How often does it rain? J. Climate, in press) who also analyzed daily precipitation frequency over land for the IPCC AR4 models. [Aiguo Dai]	
8-441	A	27:20		Section 8.3.1.2a. Mention must be made here of the uncertainty in "observed" precipitation estimates over the ocean, where they are derived from satellite measurements and sometimes with the addition of model estimates. [Gill Martin]	Accepted.
8-442	A	27:21	8:41	In spite of your egalitarian intentions it does not seem right or valid to fail to mention the obvious outlier in Figure 8.3.9. [Anne Douglass]	Rejected. It is possible that this outlier is due simply to a mistake in processing. Even if after checking this possibility (which will be done), we think the reader can easily see this without it being pointed out. If we understood why the model is an outlier, it would be worth mentioning.
8-443	A	27:24	27:24	Actually the picture reflects the tendency for the ITCZ to stay north of the equator over the oceans, leading to a much stronger annual mean zonal mean NH ITCZ, as well as the models's failure to reproduce that. Change to "reflecting a tendency for the ITCZ to reside off the Equator". [Ileana Bladé]	Accepted. Text will be reworded.
8-444	A	27:32	27:41	Would be worthwhile to also mention here the systematic insufficient rainfall over the Amazon [Anthony Hirst]	Accepted. Will add text.
8-445	A	27:37		there is no “mean simulation” - omit “simulation” and you’ve got a reasonably good sentence [Philip Mote]	Accepted.
8-446	A	27:39	27:40	Are the features of the SPCZ being too zonal and the bredth of the tropical Atlantic precipitation maximum features only of the multi-model mean or of the individual models as well? [Catherine Senior]	Accepted. Text modified.
8-447	A	27:41	27:41	There is also a notable underestimate of the observed rainfall over Amazonia by the multi-model ensemble. [William Collins]	Accepted. Will add text.
8-448	A	27:45	27:45	Considerable effort by who? The community - references? Or by this report?	Accepted. Text modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Catherine Senior]	
8-449	A	27:50	27:53	The impact of the systematic error in precipitation in the eastern Pacific and, indeed, in the western and central Pacific, on ENSO variability and teleconnections must be known in current climate, and it is certainly discussed in Section 8.4.7. [Gill Martin]	Accepted. Text modified.
8-450	A	27:53	27:53	and THEIR GLOBAL IMPACT ("influence" has just been used) THROUGH TELECONNECTIONS TO MIDLATITUDES [Ileana Bladé]	Accepted. Text modified.
8-451	A	27:55		discussion of Figure 8.3.11 - the peak precipitation amounts in the model mean are much lower than observed, a point worth noting. To what extent is that due to the effect of averaging several models together? if they had the right strength of ITCZ but it were a few grid points off in some of the models, that could reduce the average substantially. [Philip Mote]	Rejected. It is clear that this is true in the max. north of the equator, but clearly most of the models have peaks that are too low, so it's not due to averaging. This should be obvious from the figure, but we don't understand why the peak is too low, so we don't mention it.
8-452	A	28:1	28:15	Specific humidity errors are presented relative to ERA 40 with no discussion of the problematic nature of NWP-based analyses for upper-tropospheric humidity (especially). Curiously, the problems of using analyses for tropical storms are discussed, later on the page (50-56). The limitations of analyses should probably receive general treatment earlier in the assessment, before they are routinely compared with models. [Leo Donner]	Taken into account. This discussion has been substantially modified. A general treatment of problems with reanalyses has be rejected because errors depend on the field considered.
8-453	A	28:8	27:9	ERA-40 is NOT an obvservation of the humidity field. As shown by Bengtsson, L., K. Hodges and S. Hagemann, 2004: Sensitivity of large-scale atmospheric analyses to humidity observations and its impact on the global water cycle and tropical and extratropical weather systems in ERA-40. Tellus, 56A, 202-217), ERA-40 water vapor, especially in the upper troposphere, is controlled mostly by the model cumulus parameterization, not by observations. It's fine to do the comparison but unwise to conlcude that the models are in error based on Fig. 8.3.12. Compare to TOVS water vapor if you want to make such statements. [Anthony Del Genio]	Taken into account (see comment immediately above.)
8-454	A	28:8	28:8	Considered ERA-40 data should not be called observed. [Eugene Rozanov]	Taken into account (See comment immediately above.)
8-455	A	28:8		Fig.8.3.12 does not show an observed field - it is from reanalysis (e.g. "an observationally based estimate"). [Richard Allan]	Taken into account (See comment immediately above.)
8-456	A	28:14	28:14	The shading in the figure indicates the multi-model mean error in the tropical tropopause	Taken into account (See comment

No.	Batch	Page:line From	Page:line To	Comment	Notes
				can be as high as 40% (not 50%). However, this begs the question of how large is the error in individual models. [Ileana Bladé]	immediately above.)
8-457	A	28:24	28:27	These lines emphasize that the strength of water vapour feedback is determined by changes in water vapour. This is generally true if the system is linear. In a more general case, when the system is not linear, the strength of the feedback will also depend on the initial state of the system. This emphasises the relevance of a good representation of present-day climate in climate models. The clarification on the relevance of the initial state in non-linear systems is done in the first paragrhaph of Section 8.3. However, I would consider important a careful treatment of statements that apply to linear/non-linear systems throughout the text. [Alejandro Bodas-Salcedo]	Noted. In this context, no studies have suggested that initial conditions are particularly important in determining the strength of water vapor feedback.
8-458	A	28:29		Section 8.3.1.3 This section may fit better in the extremes section 8.5, as has been done in chapter 10 section 10.3.6. [Ruth McDonald]	Rejected. Extra-tropical cyclones are a fundamental element of the climate system and are neither rare nor extreme.
8-459	A	28:35	28:48	Other relevant references include Lambert et al. (2002), Bauer and Del Genio (2005), full citations given in a previous comment. [Anthony Del Genio]	Rejected. The discussion of extra-tropical cyclones has been considerably reduced and, so these references were omitted due to space limitations.
8-460	A	28:36	28:40	Another method of storm track analysis is a storm frequency index based on daily maximum 10m wind speed. This type of analysis has been applied to the ECHAM4/HOPE-G model by Fischer-Bruns et al. (2005) (Fischer-Bruns I, von Storch H, Gonzalez-Rouco JF and Zorita E Modelling the variability of midlatitude storm activity on decadal to centruy time scales. Climate Dynamics (2005) 25:461-476). [Ruth McDonald]	Rejected. The discussion of storm track methodology has been severely reduced and this method was not applied to the AR4 models, so it was omitted due to space limitations.
8-461	A	28:38		too many references - this is not a review [Philip Mote]	Taken into account. Section rewritten.
8-462	A	28:38		To give this a more modern feel I would suggest adding citations to Simmonds, I., K. Keay and E.-P. Lim, 2003: Synoptic activity in the seas around Antarctica. Monthly Weather Review, 131, 272-288. Wernli, W., and C. Schwierz, 2005: Surface cyclones in the ERA40 data set (1958-2001). Part I: Novel identification method and global climatology. Journal of the Atmospheric Sciences, (accepted) [Ian Simmonds]	Rejected. This is not intended to be a review, and this section had to be reduced in length, so these references, along with many others, were omitted
8-463	A	28:50		Replace “validation” by “evaluation”, As I explain in “General” validation must include successful future prediction, and none of the models have passed this test'	Taken into account. Section rewritten.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Vincent Gray]	
8-464	A	28:56	28:56	About the impact of the observing system changes: Dell'Aquila et al., 2005 (Dell'Aquila, A., V. Lucarini, P.M. Ruti, S. Calmanti, 2005: Hayashi spectra of the northern hemisphere mid-latitudes atmospheric variability in the NCEP-NCAR and ECMWF reanalyses. Climate Dynamics, DOI 10.1007/s00382-005-0048-x) have found significant differences between the two reanalyses in representing the baroclinic available energy conversion processes in the pre-satellite period. [SUSANNA CORTI]	Noted. The discussion of observational uncertainty was reduced substantially, and space limitations precluded inclusion of much discussion and many relevant publications.
8-465	A	28:56	29:2	The NCEP southern hemisphere cyclone tracks are also discussed by Simmonds (2003). Simmonds I (2003) Modes of atmospheric variability over the Southern Ocean J Geophys Res 108 doi:10.1029/2000JC000542 [Ruth McDonald]	Noted. Because this paper did not address the AR4 ensemble of models, it had to be passed over.
8-466	A	29:1	29:1	in the representation? How about "in the statistics", or "in the climatologies" ? [Ileana Bladé]	Taken into account. Section rewritten.
8-467	A	29:5	25:8	"All models...' this sentence does not make sense. Why mention reanalyses at the end? Isn't that what the other comparisons used as well? And what's the difference between storm tracks and the distribution of cyclones? [Philip Mote]	Accepted/Taken into account. Reference to reanalysis dropped; Also, sentence reworded.
8-468	A	29:10	29:26	These two paragraphs could use some tidying up. It is not clear what increasing vertical resolution (line 25) is important for (diagnostics or simulation ?). The following paragraph starts by referring to the models's ability to simulate the cyclone response to ENSO-induced changes in SST, but in very vague terms (where? what changes?), and then the text goes back to citing conclusions from the studies listed in the previous paragraph. Also, "using a different analysis method" is mentioned twice in the same sentence (lines 13-14). [Ileana Bladé]	Taken into account. Section rewritten.
8-469	A	29:13	29:13	The Martin et al. (2005) referred to here is ambiguous, but is the first of the two given in the References I think. However, see also the next related comment. [Timothy Johns]	Taken into account. Section rewritten.
8-470	A	29:18	29:18	Text states that repsonse of tropical cyclones to ENSO is "also" reproduced well; in fact, a balanced reading of the preceding sections presents a mixed picture on the ability of models to simulate tropical cyclones. The term "also" is thus not approriate here. [Leo Donner]	Taken into account. Section rewritten.
8-471	A	29:18	29:18	"The correct response" of what? Extra-tropical storm tracks? [Anthony Hirst]	Taken into account. Section rewritten.
8-472	A	29:30	29:32	It is likely that remaining problems in simulating tropical cyclones are not just attributable	Taken into account. Section rewritten.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				to specification of boundary conditions, as stated. Ongoing problems with inadequate resolution and physical parameterizations are also very important. [Leo Donner]	
8-473	A	29:30	29:32	Text here implies that still higher model resolution will not help in the simulation of storm tracks and storm properties (apart from improved representation of the "boundary conditions"). Is there a consensus on this? [Anthony Hirst]	Taken into account. Section rewritten.
8-474	A	29:30	29:32	This sentence reads as if there is no further role for increased resolution or better physics! [Catherine Senior]	Taken into account. Section rewritten.
8-475	A	29:34		Section 8.3.2. I am not an expert in ocean modelling so I have not read this section in great detail. However, I notice that there are 11 figures relating to the discussion on temperature and salinity, which seems rather excessive (see also my comment 12). [Gill Martin]	Noted. The number of figures in 8.3.2 has been greatly reduced due to space limitations.
8-476	A	29:34		This section does not evaluate ocean components of models but typically only lists various ocean phenomena and model successes and deficiencies. A true evaluation would indicate the importance of the deficiencies in simulating climate change and provide priorities for improvement. [Robert Molinari]	Disagree. The framework for the discussion is given at the start of 8.3.2. Quantitative measures of the effect of these errors on the response are not known.
8-477	A	29:34		Vertical mixing is the most uncertain of oceanic parameters. This should be mentioned explicitly at some point in this section. [Marisa Montoya]	Agree. Models simulate mixing on a wide range of time and space scales. It is true that the subgrid scale mixing coefficients and parameterizations can have a large impact on the simulation. The text is modified to make this point clearer.
8-478	A	29:36		Weak way to start an important sentence. [Philip Mote]	Agree. The paragraph is restructured.
8-479	A	29:37	29:38	Reconsider the wording here. Stating that the surface temperature change is determined by a model's sensitivity seems tautological. [Leo Donner]	Noted. Text is added to make the point clearer.
8-480	A	29:38	29:38	Rapier -> Raper [Reto Knutti]	Agree. Corrected.
8-481	A	29:41	29:44	Wind stress (curl) fields could be also essentially instrumental in determining the magnitude of the oceanic heat uptake and so on. The sentence "The sea surface…." might be modified. [Toshiyuki Awaji]	Agree. The text is modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-482	A	29:42	29:42	merdional -> meridional [Reto Knutti]	Agree. Corrected.
8-483	A	29:52	29:54	Some justification for "our assessment" should be given. [Robert Molinari]	Noted. Text modified
8-484	A	29:56	30:2	These sentences are specific to a draft report, they we'll be revised for the next draft after receiving the reviewers comments on what to keep in the next draft. [Philippe Tulkens]	Noted. These statements of space limitations will always apply to any analysis presented in an IPCC report.
8-485	A	30:0	31:	Figures 8.3.14 through 8.3.16 seem to lack observations - in which case they are not properly model evaluation. In the interest of brevity they should be omitted. Figure 8.3.17 is just about wind, which really belongs in the atmosphere section. It would be far more useful to have a map of surface winds, and vector differences from observations (or alternatively, the model sea level pressure field compared with observed). This would inform subsequent discussions. [Philip Mote]	Noted. The surface fluxes are important in determining the model response. We believe that they should be presented here. Observations will be added to the plots. Much of the flux figures has been moved to the supplemental material.
8-486	A	30:12	30:20	There is no uncertainty estimate on the figure, so the argument/discussion does not seem obvious.The mean model is not obvious and is distracting from the discussion. Also it does not appear in the next figure. If you are going to use it at all it seems like it should be used consistently. [Anne Douglass]	Noted. The mean model will be added to the line plots.
8-487	A	30:12	30:33	Figure 8.3.14, the "observed" heat transport estimate using NCEP or ERA reanalyses north of 60 N is lower then the heat transport of all except one models. This probably means that north of 60 N, the NCEP and ERA net surface heat flux are higher (are closer to 0) then the model net surface heat flux. Therefore it may be interesting to add the NCEP and ERA net surface heat flux on figure 8.3.13. Most of the model results would be included between NCEP or ERA data and COADS "observations". Comment on figure 8.3.13 (lines 12-20) may be slightly modified in consequence. [Jean-Louis Dufresne]	Noted. Text has been added to give a second observational estimate north of 45N.
8-488	A	30:12		COADS seems a poor choice here, especially since the following figure of the closely related meridional heat flux uses the Trenberth-Caron analyses. The apparent large discrepancy in the high latitudes of the SH is almost certainly a problem in COADS. [Frank Bryan]	Noted. Plot will be modified to show Trenberth-Caron if possible.
8-489	A	30:18		From the figure, it looks more like 30-45 N, which is not the subtropics. [Philip Mote]	Agree. Text modified.
8-490	A	30:29		How does the ocean transport heat to higher latitudes in the ITCZ? [Philip Mote]	Noted. Text seems correct.
8-491	A	30:32	30:32	Text jumps suddenly from one sub-topic discussion into the middle of the next. An	Agree. Sentence added.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				introductory sentence here would make for smoother reading. [Anthony Hirst]	
8-492	A	30:32	30:33	It should be stated whether the overestimation of the northward oceanic heat transport is due to the Atlantic or the Indo-Pacific Oceans. It would be worth showing the heat transports by the individual oceans (Atlantic, Indo-Pacific) together with that by the the global ocean. [Marisa Montoya]	The basin transports are being computed and will be added to the supplementary material.
8-493	A	30:32	30:34	Saying that the models transport too much heat north of 45N is not justified. Not only have the errors in the observational estimates shown in fig. 8.3.14 not been shown, but the writer has ignored the estimates of heat transport based on hydrographic data (Ganachaud and Wunsch, 2003, J. Clim., 16, 696). These yield an estimate of 0.6 PW at 47 N which is in much better agreement with the models. [Peter Stone]	Agree. Text added.
8-494	A	30:32		Direct oceanographic estiimates of heat transport (e.g. Kolterman et al, Deep Sea Res. II, 46, 109-138, 1999) also exceed the Trenberth and Caron estimates at high latitudes. It would be useful , in fact, to add direct ocean estimates to Figure 8.3.14 in order to further quantify the uncertainties in the "observational" analyses. [Frank Bryan]	Agree. Text added and Ganachaud and Wunsch, 2003 reference.
8-495	A	30:39	30:39	Figure 8.3.14 (and 8.3,8, and 8.3.16) omit a number of models in the AR4 archive. It is possible that the ocean data was not available at the time of the creation of these figures, but would it be possible to update them for completeness? [William Collins]	Accepted. Figures will be updated.
8-496	A	30:41	30:55	Are any observations available to plot on Figs. 8.3.15 and 8.3.16? The text speaks of errors, e.g. on line 55; it would be good if the figures could show the basis for that assertion. [Leo Donner]	Noted. Observed estimate of freshwater transport will be added if possible.
8-497	A	31:0	32:	Figure 8.3.19 has similar information but is much more useful than 8.3.18; the latter should be omitted. The same goes for 8.3.20 and 8.3.21 - omit 8.3.20. [Philip Mote]	Disagree. As noted in the introduction of this subsection both temperature and salinity are important in determining the density errors and therefore the circulation errors.
8-498	A	31:4	31:10	The performance of the models' zonal wind-stress is evaluated compared to the ERA40 reanalyses. The errors are mentioned to be largest in the Southern Hemisphere, but in this region the spread of the observations is also quite large (e.g. NCEP, Trenberth, Hellerman and Rosenstein). This should be taked into account. [Marisa Montoya]	Agree. Text modified.
8-499	A	31:23		The cold bias east of the Grand Banks in the subpolar region of the Atlantic in many	Agree. Text modeified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				models is attributable to a poor simulation of the path of the North Atlantic Current. It is noteworthy that this is an issue with the representation of ocean dynamics in the models, not a bias in the atmospheric surface forcing. [Frank Bryan]	
8-500	A	31:26	31:27	The model mean on Fig. 8.3.18 does not show evidence of a warm bias just south of the equator, as stated in the text. [Leo Donner]	Agree. Text deleted.
8-501	A	31:45	31:45	One of the most important biases for the Atlantic is not mentioned: The zonal temperature gradient and the thermocline slope in the tropical Atlantic is reversed from the observed gradients (Davey et al. 2002 STOIC: a study of coupled model climatology and variability in tropcial ocean regions. Clim Dyn. 18, 403-320). This has leads to failure in simulating the Atlantic Nino mode and associated errors in precipitation. The lack of stratiform cloud in the southeastern Atlantic, the wind-errors induced from the Pacific or errors in local upwelling feedbacks can cause this. Recently Hazeleger and Haarsma (2005, Sensitivity of tropical Atlantic climate to mixing in a coupled ocean-atmosphere model, Clim Dyn, 25, 487-399) showed that upper ocean mixing parameterizations can be tuned to produce the correct tropical Atlantic climate in a coupled general circulation model. [Wilco Hazeleger]	Noted. It is not clear that these errors are more important than other errors also not mentioned in the text due to space limitation
8-502	A	31:55	31:57	Some discussion of the deficiencies in WOA-2004 (I.e., data coverage, smoothing, particularly near coasts and averaging on depth surfaces instead of isopycnal surfaces) is required. The effect of this deficiencies on the comparisons (I.e., comparisons in data areas of the southern oceans) should be described. [Robert Molinari]	Agree. Text added.
8-503	A	32:17	32:23	Same comment as for previous line. [Robert Molinari]	Agree. Text added.
8-504	A	32:38	32:41	It is mentioned that the errors in the zonally averaged potential temperature are partly related to errors in formation and mixing of NADW and referenced to 8.2 but in 8.2 there is no mention to this issue. [Marisa Montoya]	Agreed. Reference deleted.
8-505	A	32:42		no such thing as the “mean model”; also, a zonal mean can’t show the differences at the Mediterranean sea. Omit Figure 8.3.23. [Philip Mote]	Agree. Text changed. Disagree with the suggestion to delete the figure. It is important to discuss both the T and S errors as they both impact the the density errors and therefore the ocean circulation.
8-506	A	32:43	32:43	the actual figure is 8.3.23 [vincenzo artale]	Noted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-507	A	32:44		The high salinity bias in the vicinity of 30N is equally apparent in the Pacific maps in the supplementary material. There is not a consistent picture of a bias in Med outflow water. This statement is misleading. [Frank Bryan]	Agree. Text changed.
8-508	A	33:1	33:10	Since Subtropical Cells (STCs) transfer most of the oceanic heat poleward (more than the THC) and since they determine the ventilation and therefore the subsurface structure of the tropical thermocline (and therefore El Nino properties) some attention to these wind-driven cells could be given here. The results of Hazeleger 2005 (Can global warming affect tropical ocean heat transport? Geophys. Res. Lett. in press) show that the South Atlantic STC does not change but the heat transport responds to the weakening MOC. In the Pacific STCs do change, but the heat transport remains constant due to compensating gyre and overturning transports (see also Hazeleger et al. 2004. How can Pacific heat transport vary? J. Phys. Oceanogr. 34, 320-333). [Wilco Hazeleger]	Noted. There is some discussion of the STC in section 8.5.
8-509	A	33:9	33:10	Suggest to add Dai et al. (2005) to the citation: Dai, A., A. Hu, G. A. Meehl, W. M. Washington, and W. G. Strand, 2005: Atlantic thermohaline circulation in a coupled model: Unforced variations vs. forced changes. J. Climate, 18, 2990–3013, which shows both decreased northward heat transport and increased vertical stability. [Aiguo Dai]	Agree. Reference added.
8-510	A	33:12		It should be clarified whether the figures show only the Eulerian overturning or the combined Eulerian/eddy-induced transport. If the latter, are models with and without GM averaged together to produce the multi-model means? This would give a very distorted piture. [Frank Bryan]	Noted. Figure caption changed.
8-511	A	33:28	34:10	I think this reference should be included: Trigo R.M., Garcia-Herrera R., Diaz J., Trigo I.F. and Valente M.A. (2005). 'How Exceptional Was the Early August 2003 Heatwave in France?'. Geophysical Research Letters, 32(10), Art. No. L10701. In this paper a very complete analysis was presented estimating the effects of the 2003 summer heat wave over mortality in Europe, the mechanisms causing this situation and comparing the measured temperature values with values from the previous 500 years [Pedro Ribera]	Noted. Page numbers (or chapter) for comment seem incorrect.
8-512	A	33:29	33:29	What are "Fig. 10.x" and "Fig. 10.y". [Anthony Hirst]	Noted. Figure number from chapter 10 have been added.
8-513	A	33:33	33:34	the concept is very interesting and deserve more deeper analysis and may be integrated in the executive summary [vincenzo artale]	Noted.
8-514	A	33:33	33:34	Actually in the models considered by Gregory et al (2005) there is a positive correlation	Agree. Text deleted. Statements like

No.	Batch	Page:line From	Page:line To	Comment	Notes
				between the control MOC strength and the magnitude of the weakening. [Jonathan Gregory]	this belong in chapter 10.
8-515	A	33:36	33:37	In addition to model resolution and mixing schemes, couldn't better air-sea fluxes in water formation regions also contribute to the MOC simulation improvement? [Keith Dixon]	Agree. Text added.
8-516	A	33:44	33:44	"of the new AR4 models". Ase these models werenot developped for AR4 specifically, I would suggest replacing these words by "of the models selected for AR4". [Philippe Tulkens]	Agree. Text modified.
8-517	A	33:48	33:49	I believe the statement here is correct, but can another or additional more recent reference be added? The Southern Ocean simulation in the cited reference had significant deficiencies which have been largely mitigated in more recent models. [Anthony Hirst]	Agree. Reference added.
8-518	A	34:0		Section 8.3.3 on sea ice is a good addition to the AR4 (there was little discussion of sea ice in the TAR). This section should remain in the report. [Melanie Fitzpatrick]	Noted.
8-519	A	34:1	4:1	The reference to Kamenkovich and Sloyan (2005) is not included in the reference list. [Marisa Montoya]	Agree. Reference added.
8-520	A	34:1	34:1	The Kamenkovich and Sloyen reference is missing in the bibliography. [Peter Stone]	Agree. Reference added.
8-521	A	34:4	8:8	I realy like the statement about the lifely effect of these errors - I presume "these" refers only to the immediate subsection. [Anne Douglass]	Agree. Text modified to make discussion clearer.
8-522	A	34:4	34:7	The importance of the Southern Ocean for heat uptake is a model result. In the observations there is more heat uptake in the Northern Hemisphere around 20 to 40 N (Levitus et al., 2005) . No model gives this pattern. In addition the models are overestimating the rate at which heat is being mixed into the deep ocean (Forest et al., 2005). The models are in fact doing a poor job of simulating the heat uptake, and the uncertainty in the transient response is much greater than is implied here. [Peter Stone]	Noted. Key is quality of the heat content observations.
8-523	A	34:10	34:19	Noting that errors in T and salinity in the thermocline have decreased - any chance that this reduces uncertainty or contributes to reduction in spread of results? [Anne Douglass]	Noted. No, because of the importance of the Southern Ocean in heat uptake.
8-524	A	34:13		The references mentioned here don't support the statement of fact, but rather (I presume) a description of flux adjustments. Suggest rewording to clarify. [Philip Mote]	Agree. References not needed and are deleted.
8-525	A	34:14	34:5	In addition to model resolution and parameterization schemes, couldn't better air-sea	Agree. Text added.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				fluxes in water formation regions also contribute to the overall ocean simulation improvement? After all, these are "coupled" models and the ocean circulations are being driven by the surface forcing. [Keith Dixon]	
8-526	A	34:14		There are still serious problems with simulating the tropical Pacific thermocline that should be mentioned: Fevrier, S., Frankignoul, C., Sirven, J., Davey, M.K., Delecluse, P., Ineson, S., Macias, J., Sennechael, N., and D.B. Stephenson, 2000: A multivariate intercomparison between three oceanic GCMs using observed current and thermocline depth anomalies in the tropical Pacific during 1985-1992, J. Marine Systems, 24, pp. 249-275. [David Stephenson]	Noted. It appears that many of the best AR4 model simulation are better than those found in Fevrier et al. The error still exists and that fact is noted in the text.
8-527	A	34:42	34:43	Some explanation is required to explain why a mean of models with errors that agrees with observations is significant. Since few if any of the models agree with observations they can not be used in simulations or projections. However, if it has been shown that these means of model results can not only reproduce climatologies but also generate valid predictions these findings should be described. [Robert Molinari]	Accepted. A brief discussion is given in Section 8.1.
8-528	A	34:44	34:44	It is not immediately clear what two things the word "both" is referring to. [Keith Dixon]	Accepted. Text modified.
8-529	A	34:44		mean" model should be model mean; what's the meaning of "both by the beginning and at the end [Philip Mote]	Accepted. Text modified.
8-530	A	35:3		this is another reason it would be great to have a plot of global wind fields (with an inset for the Arctic in polar stereo) [Philip Mote]	Figure showing sea level pressure now included as supplementary material.
8-531	A	35:5		As written, the sentence implies that advanced (better ?) sea ice dynamics are a potential cause of the problems? [Frank Bryan]	Yes, the worse are the driving fields the worse may be errors in geographical distributions of sea ice mass – even worse than in the models without sea ice dynamics.
8-532	A	35:16		Why is IAP FGOALS missing in the table? Is it missing in all the sea ice figures? If yes, please state so and explain why. There are rumours that it was excluded because it showed very poor performance. For a credible report, data from all models available should be included and no model should be excluded because it does not agree with the rest, unless it is stated clearly why it is excluded. [Reto Knutti]	Accepted. The table has been removed. Text modified.
8-533	A	35:20	35:21	I assume the authors mean the ABILITY to evaluate the land surface component is	Accept – text modified

No.	Batch	Page:line From	Page:line To	Comment	Notes
				HAMPERED by the LACK of available observations. [Chuck Hakkarinen]	
8-534	A	35:27	35:29	the extension of acronyms could be eliminated [vincenzo artale]	Accept – text modified
8-535	A	35:32	35:32	word missing in this line [Reto Knutti]	Accept – text modified
8-536	A	35:32	35:33	Same comment as for previous line. [Robert Molinari]	Unclear what this comment refers to
8-537	A	35:41	35:41	Figure 8.3.27 is referenced at the end of this line, but the authors meant to reference 8.3.26. [William Collins]	Accepted.
8-538	A	36:1	36:10	I found this figure hard to understand - it seems like the mean should fall between the individual points. Since it doesn't, a line or two more explanation would help. [Anne Douglass]	Reject – its hard to add this text for space limitations.
8-539	A	36:7	36:18	It would be useful to summarize what's known about the solar budget, say using the ISCCP FD and/or CERES SARB products. The TOA absorption should be compared to ERBE or CERES. [William Collins]	Reject – this is not clearly in the literature and we cannot therefore assess it (no reference provided by referee)
8-540	A	36:9		the statement about "unrealistically dampened" is off the mark - it's a simple fact that averaging reduces noise, and unforced variability constitutes noise. [Philip Mote]	Reject – this is as stated by the authors of the paper
8-541	A	36:32	36:43	This paragraph has several unclear statements and needs to be rewritten, particularly the second sentence (line 35) and fourth sentence (line 37ff). Refer back to the systems approach and say the hydrological component has been evaluated both separately (refs.) and in a systems framework (refs). The main point of the paragraph is muddied, but appears to be that errors in the precipitation field can dominate the evaluation of runoff when the evaluation is performed at the systems level. [Philip Mote]	Accept – text modified and simplified
8-542	A	36:40	36:43	LPJ acronym is not defined. [Leo Donner]	Accept – text modified
8-543	A	36:45		what is a "Chapter 10 model"??? [Philip Mote]	Accept – text modified
8-544	A	37:1	37:4	This is a duplicate of p. 18, line 54 to p. 20, line 2 [see above two comments] and should be removed. [Alan Robock]	Accept – text modified
8-545	A	37:7	37:18	Two references are missing in the list (Wild et al., 2001 and Wild(2005)). There is no	Accept – text modified

No.	Batch	Page:line From	Page:line To	Comment	Notes
				observed data in the Figure 8.3.28 and the model abbreviations are not clear. [Eugene Rozanov]	
8-546	A	37:7	37:18	The cited references of Wild 2005 and Wild et al. 2001 are missing in the publication list on page 8-102: The references are Wild, M., 2005: Solar radiation budgets in atmospheric model intercomparisons from a surface perspective, Geophys. Res. Lett, 32, doi:10.1029/2005GL022421. Wild, M., Ohmura, A., Gilgen, H., Morcrette, J.J., and Slingo, A., 2001: Downward longwave radiation in General Circulation Models. J. Climate, 14, 3227-3239 [Martin Wild]	Accept – text modified
8-547	A	37:18		simulation **of** sensible… [Philip Mote]	Accept – text modified
8-548	A	37:19	37:44	The whole 8.8 section (key uncertainties) is a little difficult to follow. In general the chapter is clearly written and it is easy to find the important facts of every section. It is not so in this case. I would suggest a clearer rewritting of the 4 paragraphs of this section (or at least paragraphs 1 and 2). Another possibility could be simply present the key uncertainties the authors want to highlight in a schematic way. [Pedro Ribera]	Misplaced comment – refers to Section 8.8. Comment noted.
8-549	A	37:34	37:34	"my observed climatologies"? [Leo Donner]	Accept – text modified
8-550	A	37:34		forced **by** observed… [Philip Mote]	Accept – text modified
8-551	A	37:35	35:48	The detailed discussion of a single model in this context is inappropriate. Similar studies have been carried out with other models. Either the passage should be abbreviated, or made more generic. [Frank Bryan]	Accept – text modified by using more recent analyses based on C4MIP.
8-552	A	37:52	38:19	some of this is redundant with section 8.1.2.2 and should be shortened and moved there, with a cross-reference. [Philip Mote]	Accepted. Text modified?
8-553	A	38:11	38:21	The 'standardization' is in fact, much looser than impled here. For example, the solar constant, GHG concentrations (present day vs. preindustrial), and many other critical model choices are left unspeficied in the 'standard'. This leaves interpretation of model differences far from 'easily assessed'. [Frank Bryan]	Rejected. The statement 'more easily assessed' (not 'easily assessed') is accurate.?

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-554	A	38:48		Figure 8.3.29 is missing results for absorbed solar radiation, but has OLR. Absorbed solar should be added as a key metric. [Bruce Wielicki]	Accepted.
8-555	A	39:0		General Comment: In the initial draft outline of the AR4 the chapter addressing "Evaluation of Large-Scale Climate Variability as Simulated by Coupled Global Models" was meant to discuss the ability of today's climate models to represent both: the general statistics of the system and the specific skill to predict the large scale variability. In the current draft of the AR4 the majority of the chapter is concerned with the ability to represent the statistics of the system. Only one and a half pages focus on the skill linked with the initial value climate predictions. Since in a number of countries seasonal climate predictions has become part of the operational Met Office task I assume that for many readers from the governmental side it would be very helpful to better bridge the gape between the model used for seasonal climate predictions and the models used to make projections of future climate change. Excellent reference would be: Palmer, T.N., and J. Shukla, 2000: Editorial (for special issue on DSP/PROVOST). Quart. J. Roy. Meteor. Soc., 126, p1989-1990. 40 Palmer, T. N., and Coauthors, 2004: Development of a European multimodel ensemble system for seasonal to interannual prediction (DEMETER). Bull. Amer. Met. Soc., 85, 853–872 On the topic ENSO many references exist below just an example: Oldenborgh, G.J. van, M.A. Balmaseda, L. Ferranti, T.N. Stockdale and D.L.T. Anderson, Did the ECMWF seasonal forecast model outperform statistical ENSO forecast models over the last 15 years? J. Climate, 2005, 18, 16, 2960-2969 On the topic NAO prediction see e.g. The skill of multi-model seasonal forecasts of the wintertime North Atlantic Oscillation Doblas-Reyes FJ, Pavan V, Stephenson DB CLIMATE DYNAMICS 21 (5-6): 501-514 NOV 2003 Muller WA, Appenzeller C, Schar C Probabilistic seasonal prediction of the winter North Atlantic Oscillation and its impact on near surface temperature CLIMATE DYNAMICS 24 (2-3): 213-226 FEB 2005 [Christof Appenzeller]	Reject. We feel that we have discussed weather, seasonal, and decadal prediction appropriately.
8-556	A	39:1	39:35	interesting but too long [vincenzo artale]	Taken into account. Section rewritten and shortened.
8-557	A	39:1	39:35	This is a big discussion about the "mean model". At the end of this section, it seems that comment that various processes are related and therefore it is possible to characterize	Taken into account. Section rewritten and shortened.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				performance using a limited number of fields - this very statement argues against the physical information that can be gained from a "mean model" - in which physical processes are related in the sub-models but probably not in the ensemble average. [Anne Douglass]	
8-558	A	39:12		maybe I'm missing something, but two statements seem contradictory: models have improved (by what measure? I thought RMS error was the most comprehensive) but decreases in RMS error are small. Is the message that "models have improved slightly, as measured by RMS error"? [Philip Mote]	Accepted. Yes models have improved slightly, as measured by RMS error, and text has been revised to clarify.
8-559	A	39:16	39:19	Same comment as for previous line. [Robert Molinari]	Taken into account. Section rewritten and discussion of mean model result clarified.
8-560	A	39:21	39:21	Figure 8.3.30; It looks like all improvement is obtained due to massive improvement of one model? Is it correct? T200 is not improved, could you, please, comment on this? [Eugene Rozanov]	Taken into account. Section rewritten. The median is relatively insensitive to individual models, and the figure clearly shows that most models seem to improve, at least a little. The T200 error is dominated by the cold bias which remains (and is even a little worse) on average.
8-561	A	39:25		Isn't the evaluation of interannual variability the focus of section 8.4? [Philip Mote]	Taken into account. Section rewritten.
8-562	A	39:34	39:35	It is therefore perhaps not unreasonable to characterize" is VERY awkward and should be re-written. How about "Overall changes in model fields can be characterized by … [Chuck Hakkarinen]	Taken into account. Section rewritten.
8-563	A	39:37		Section 8.4. There is a noticeable lack of figures in sections 8.4.1 to 8.4.10. Although I appreceiate the need to keep the number of figures to a minimum, this seems a rather large gap. Where possible, it would be useful to have an illustration of some of these modes of large-scale climate variability. [Gill Martin]	Accepted. New figures have been added.
8-564	A	39:39	39:41	How can one conclude that improved simulations of shorter term variability lead to increased confidence in longer term climate projections? Longer term projections are dependent in all their aspects on getting the climate sensitivity right, but Chapter 10 is now concluding that there is a 8 % chance that the sensitivity is less than 1 C and a 28 % chance that it is larger than 6 C. This is a larger range than the 1.5 to 4.5 C range that has long been cited. This change implies greater uncertainty and therefore less confidence in the climate models' long term projections.	Rejected. Improved simulations of variability directly point to improvements in model performance and enhances our confidence in climate change projections.

No.	Batch	Page:line From	To	Comment	Notes
				[Peter Stone]	
8-565	A	39:43		A recent study on the AO reveals that the AO can be excited by various kinds of external forces (H. L. Tanaka & M. Matsueda, J. Meteorol. Soc. Jpn., 83, 611-619 (2005)). In this regard, a relation between solar wind and AO index is worth to mention here: [1] Palamara & Bryant ("Geomagnetic activity forcing of the Northern Annular Mode via the stratosphere," Annales Geophysicae (2004) 22: 725–731) as well as [2] Boberg & H. Lundstedt ("Solar Wind Variations Related to Fluctuations of the North Atlantic Oscillation," Geophys. Res. Lett., VOL. 29, NO. 15, 1718, 10.1029/2002GL014903, 2002). These reports deal with this relation although the latter is on NAO. This kind of relation may be useful because it may solve the discrepancy between the surprisingly small change in the solar luminosity and the large fluctuation in the global historical temperature records (Esper et al, 2002; Moberg et al., 2005). [Kiminori Itoh]	Reject. Sufficient references.
8-566	A	39:43		Section 8.4.1. This is an interesting section. Has there been an improvement in the simulation of the NAM and SAM with the increased horizontal and vertical resolution, and improved dynamical cores, in new models compared with those in the TAR? [Gill Martin]	Noted. To our knowledge this has not been systematically addressed in the literature.
8-567	A	39:45	:51	This section fails to acknowledge the ongoing debate about the physical meaning (if any?!) of NAM. There is certainly not a consensus view of scientists who believe that NAM is a leading mode of variability – some of us think it is an artefact of doing EOFs over a hemispheric domain! This ongoing debate is important and should be fairly treated by citing recent papers on the subject such as: Ambaum, M.H.P., B.J. Hoskins, and D.B. Stephenson, 2001: Arctic Oscillation or North Atlantic Oscillation?, J. Climate, 14, 3495-3507. Ambaum, M.H.P., B.J. Hoskins, and D.B. Stephenson, 2002: Corrigendum: Arctic Oscillation or North Atlantic Oscillation?, J. Climate, 15, 553. and recent references therein. [David Stephenson]	Taken into account. The debate is discussed in detail in Chapter 3.
8-568	A	39:49	39:51	Since we were talking about the NAM, why are we suddenly shifting to "models's leading modes of variability" ? It seems more appropiate to simply point out the fact that many models's projections of climate change project onto the NAM and refer the reader to Chapter 10 (Section 10.3.5.3.1). [Ileana Bladé]	Reject. The important point is that because of the strong projection successful simulation of the NAM affects our confidence in the Chapter 10 climate model projections.
8-569	A	39:49	39:49	8-39, line 49, The reference should add Cai, Whetton, & Karoly paper: Cai, W. J., Whetton, P. H., and Karoly, D. J. (2003). The response of the Antarctic Oscillation to increasing and stabilized atmospheric CO2. Journal of Climate, 16 (10):	Reject. Sufficient references.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				1525-1538. < http://ams.allenpress.com/amsonline/?request=get-abstract&issn=1520-0442&volume=016&issue=10&page=1525 > [Wenju Cai]	
8-570	A	39:49	39:51	Shindell et al (1999) made this point at the same time as Fyfe et al (1999) and should be referenced: Shindell, D. T., Miller†R. L., Schmidt, G. A. and Pandolfo, L., 1999, Simulation of recent northern winter climate trends by greenhouse-gas forcing, Nat, 399, 452-455. [Ron Miller]	Accepted. Reference added.
8-571	A	39:49	39:49	The reference to Fyfe et al. (1999) is not included in the reference list. [Marisa Montoya]	Noted
8-572	A	39:52		too many references. [Philip Mote]	Accepted. Number of references reduced.
8-573	A	40:2	40:3	Split the sentence after (Osborn, 2004)."In some models this is related to a bias towards ….." [Ileana Bladé]	Accepted. Text modified.
8-574	A	40:2		In addition to Osborn (2004), Stephenson and Pavan (2003) also noted that the models tended to overestimate the correlation between NAO and ENSO. [David Stephenson]	Noted.
8-575	A	40:6	40:7	As the source of this sentence (Miller et al 2005, revised), I need to correct it. Please delete: `underestimate the observed temporal variability of atmospheric pressure'. The remainder of the sentence is correct. I apologize! I will send an updated copy of the article to the chapter coordinating lead authors. [Ron Miller]	Accepted. Text modified.
8-576	A	40:8	40:8	Recent papers using results from several GCMs conclude that observed multidecadal variability is inconsistent with that found in coupled GCMs and is not accounted for in simulations with external forcings. New work shows that NAM trends can be simulated if stratospheric conditions are specified according to observations. Suggest changing lines 11 and 12 to:..lower in coupled GCM control simulations than is observed, and can also not be reproduced in current model simulations with external forcings (Osborn 2004). However, Scaife et al (2005) show that the observed multidecadal trend in the surface NAO and NAM can be reproduced in a model if observed trends in the lower stratospheric circulation are prescribed in the model.Troposphere-stratosphere coupling processes may therefore need to be included in models to fully simulate NAM variability. T.J.Osborn, 2004 Simulating the winter North Atlantic Oscillation: the roles of internal variability and greenhouse gas forcing Clim. Dyn., 22, 605-623. A.A.Scaife, J.R.Knight,	Accepted. Text modified.

No.	Batch	Page:line From	To	Comment	Notes
				C.K.Folland and G.K.Vallis 2005, A stratospheric Influence on the winter NAO and North Atlantic surface climate. Geophys.Res. Lett., 32, L18715. [Chris Folland]	
8-577	A	40:8		"well simulated" -- subjective statement. Be quantitative. [Philip Mote]	Accepted. Changed "well simulated" to "correctly simulated".
8-578	A	40:12	40:16	Miller et al (2005, revised draft) show that the multi-model NAM response of 9 IPCC AR4 models to volcanic forcing is statistically distinct from zero and of the correct sign. However, it is significantly smaller than the observed response. Variability of individual models is too large to distinguish the volcanic response from zero. As for the effect of SST anomalies, I believe that Selten et al 2004 have shown a statistically significant NAM response to this forcing, at least in the NCAR PCM. [Ron Miller]	Noted.
8-579	A	40:14		In addition to Alexander et al. (2004), a strong NAO model response to sea-ice was also noted by Kvamsto et al. (2004): Kvamsto, N.G., P. Skeie, D.B. Stephenson 2004: Impact of Labrador sea-ice on the North Atlantic Oscillation, International J. of Climatology, 24, 603-612. [David Stephenson]	Reject. Sufficient references.
8-580	A	40:16		A recent study of the NAO simulated by 20 CMIP2 models (Stephenson et al. 2005) has found that the majority of models show a slight increasing trend of NAO with increasing amounts of CO_2. However, there is a large amount of model uncertainty in the sensitivity of the response and some of the models presented in earlier studies (such as HadCM3, ECHAM3, ECHAM4) appear to have overly strong responses compared to the other models. Stephenson et al. (2005) also found that the European precipitation and temperature responses to CO_2 increase showed less variation across models despite the large differences in NAO response. This recent work should be cited: Stephenson, D.B., Pavan, V, Collins, M., Junge, M., Quadrelli, R., 2005: North Atlantic Oscillation Response to Greenhouse Gas Forcing and its Impact on Climate Change in Europe: An Assessment of 18 CMIP2 Coupled Climate Model Simulations. Climate Dynamics, submitted. There has also been progress on using multi-model ensembles to make seasonal forecasts of the NAO: Doblas-Reyes, F.J., V. Pavan, and D.B. Stephenson 2003: Multi-model seasonal hindcasts of the North Atlantic Oscillation, Climate Dynamics, 21, 501-514. [David Stephenson]	Reject. This discussion is best left for Chapter 9 and/or 10.
8-581	A	40:22	40:23	8-40, lines 22-23, a wrong paper is referred to. Change the Cai and Watterson 2002 to Cai, Whetton, and Karoly:	Accepted. Reference changed.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Cai, W. J., P. H. Whetton, and D. J. Karoly (2003). The response of the Antarctic Oscillation to increasing and stabilized atmospheric CO2. Journal of Climate, 16 (10): 1525-1538. < http://ams.allenpress.com/amsonline/?request=get-abstract&issn=1520-0442&volume=016&issue=10&page=1525 > [Wenju Cai]	
8-582	A	40:22	40:23	too many references. [Philip Mote]	Accepted. Number of references reduced.
8-583	A	40:24	40:25	If you would like to see the SAM pattern in 14 IPCC AR4 models compared to NCEP, see Figure 4 of Miller et al (2005, revised draft). [Ron Miller]	Accepted. This figure is now included.
8-584	A	40:25	40:30	The wording needs to be changed a bit so that line 27 "are captured well" doesn't appear to be in open contradiction with the following paragraph as far as the GFDL simulation of the SAM is concerned. [Ileana Bladé]	Accepted. Text modified.
8-585	A	40:33	40:35	`do not always compare well' is a vague criticism. Miller et al (2005, revised draft) shows that the pattern correlation between 14 IPCC AR4 models compared to NCEP is generally above 0.95, although the amplitude of variability is too large by up to 50%. The standard deviation of the correlation among individual members of each ensemble is 0.02, which argues that the SAM simulated by the individual members is highly correlated with NCEP. [Ron Miller]	Accepted. Text modified.
8-586	A	40:41	40:44	I suggest rewriting the first part of this sentence: `The SAM extends through both the troposphere and stratosphere, although it can be captured, for example...' The current version makes it sound like the stratosphere is not involved, although its influence is discussed later in the paragraph. [Ron Miller]	Accepted. Text modified with the focus on the tropospheric SAM.
8-587	A	40:41	40:54	some of this applies to the NAM too. [Philip Mote]	Noted.
8-588	A	40:41	40:44	Please add a reference after "in atmospheric GCMs with a poorly resolved stratsophere and driven by prescribed SSTs (e.g., …; Zhou and Yu, 2004). For detail, see : Zhou Tianjun, Rucong Yu, 2004, Sea-surface temperature induced variability of the Southern Annular Mode in an atmospheric general circulation model?Geophysical Research Letters, 31,L24206,doi:10.1029/2004GL021473 [Tianjun ZHOU]	Rejected. Sufficient references already
8-589	A	40:44		In fact even minimal stochastic models with no real eddy-mean flow interaction dynamics	Rejected. We need to keep focused on

No.	Batch	Page:line From	Page:line To	Comment	Notes
				at all (just conversation of momentum) can yield a NAM/SAM-like leading EOFs, which do not reflect a dynamical oscillation (i.e., Gerber and Vallis, 2005, J. Climate 18, 2102-2118, and Witman, Charlton and Polvani, 2005, J. Climate 18, 2119-2112). [Ileana Bladé]	coupled simulations.
8-590	A	40:47	40:48	The ocean influences SAM variability in important ways, which has been proved by AGCM expriment. Reference: Zhou Tianjun, Rucong Yu, 2004, Sea-surface temperature induced variability of the Southern Annular Mode in an atmospheric general circulation model?Geophysical Research Letters, 31,L24206,doi:10.1029/2004GL021473 [Rucong Yu]	Rejected . We need to keep focused on coupled simulations.
8-591	A	40:47	40:48	The ocean influences SAM variability in important ways, and there exists significant correlation between tropical ocean SST and SAM in terms of interannual variability, which has been proved by AGCM expriment. Reference: Zhou Tianjun, Rucong Yu, 2004, Sea-surface temperature induced variability of the Southern Annular Mode in an atmospheric general circulation model?Geophysical Research Letters, 31,L24206,doi:10.1029/2004GL021473 [Tianjun ZHOU]	Rejected . We need to keep focused on coupled simulations.
8-592	A	40:53	40:53	8-40, line 53, before "Thus", Recent studies (Cai et al. 2005, and Cai 2005) show that there is a vertically integrated impact of SAM on climate from the troposphere, through Earth surface, to ocean circulation. Cai, W. J. (2005), Antarctic Ozone depletion causes an intensification of the Southern Ocean super-gyre circulation. Geophys. Res. Lett. (in press) Cai, W. J., G. Shi, T. Cowan, D. Bi, and J. Ribbe (2005), The response of the southern annular mode, the East Australian Current, and the southern mid-latitude ocean circulation to global warming, Geophys. Res. Lett. (in press). [Wenju Cai]	Rejected. We are space constrained.
8-593	A	40:55	40:55	8-40, line 55, SAM. ? SAM and its impacts. [Wenju Cai]	Rejected. We prefer the current text.
8-594	A	41:1		Section 8.4.2. I am unclear as to the conclusion of this section; is it that case that the presence (or lack thereof) of an IPO in the IPCC AR4 coupled models is not known? [Gill Martin]	Rejected. The section states that "coupled models do not seem to have difficulty in simulating IPO-like variability". The last sentence is included to highlight that a closer

No.	Batch	Page:line From	To	Comment	Notes
					examination of the details of IPO-like variability in simulations would be useful.
8-595	A	41:1		section 8.4.2: fewer references, refer to chapter 3. [Philip Mote]	Accepted. References reduced. Links to other chapters may appear in future drafts,
8-596	A	41:12	41:12	"the PDO-like mode they examined": where? , what model? [Ileana Bladé]	Accepted. Text modified.
8-597	A	41:12	41:39	It has been suggested that oceanic extratropical-tropical connections generate the memory for the PDO (Gu and Philander, 1997 Science 275, 805-807), but model results suggest that temperature anomalies that subduct in the extratropics do not reach the tropics (e.g. Hazeleger et al., 2001, J. Geoph. Res., 106, 8971-8988). Variations in strength of the subtropical cells that connect the extratropics with the tropics may have an impact (Kleeman et al., 1999, Geoph. Res. Lett., 1743-1746). [Wilco Hazeleger]	Noted. Space limitations prevent us from discussing this point.
8-598	A	41:14	41:16	The PDO may be the North Pacific expression of the IPO. However, I do not think it is authorized. The relationship between PDO and IPO strongly depends on period, analytical method, and dataset. The sentence might be modified not to be misunderstood. [Toshiyuki Awaji]	Accepted. Text modified.
8-599	A	41:21	41:21	8-41, line 21, add Cai and Whetton, 2000: Cai, W. J., and Whetton, P. H. (2000). Evidence for a time-varying pattern of greenhouse warming in the Pacific Ocean. Geophysical Research Letters, 27 (16): 2577-2580. < http://www.agu.org/journals/gl/gl0016/1999GL011253/0.html > [Wenju Cai]	Accepted. Text modified.
8-600	A	41:26	41:26	aren't equatorially-trapped waves "definitively" important for ENSO dynamics ? [Ileana Bladé]	Accepted. Text modified.
8-601	A	41:47	41:52	"The occurrence…flow pattern" does not seem policy relevant and could be removed. [Richard Allan]	Rejected. Discussion of the processes contributing to the formation of the PNA pattern is essential to this subsection.
8-602	A	41:47	41:47	This sentence is a bit misleading: I think it should be rephrased to indicate that the PNA can be internally generated in GCMs, although the PNA appears to be a preferred pattern of response to external forcing. Since this seems to be a common misperception, perhaps this would work: "Although the PNA pattern is commonly associated with the response to anomalous boundary forcing, GCMs do not require the presence of external forcings to produce a PNA-like pattern or mode of variability".	Accepted. Text clarified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Ileana Bladé]	
8-603	A	41:57	42:8	Particularly noteworthy…results demonstrate that…": Listing projects is not so useful here. I recommend retaining last sentence: "Climate models do not respond…ensemble members of a given model (Palmer and Shukla, 2000). [Richard Allan]	Rejected. Listing of these projects would allow interested readers to gain access to more delailed information on model responses to anomalous SST forcing.
8-604	A	42:8	42:8	mayeb add "indicating that extratropical variability is only weakly constrained by tropical SST forcing". Also the sentence "atmospheric climate models do respond to the prescribed SST forcing" is a little flat and vague. Can we make the statement more forceful? [Ileana Bladé]	Accepted. Text modified and made more forceful.
8-605	A	42:12		"This system uses a 2-tiered approach..." this could be removed since the results are the important component to emphasise. [Richard Allan]	Rejected. Brief mention of this 2-tier method serves to distinguish the NCEP approach from the fully-coupled experiments in DEMETER, which are described in the latter half of the paragraph.
8-606	A	42:15		"good agreement": be quantitative [Philip Mote]	Accepted. Statement made quantitative.
8-607	A	42:23	42:23	delete "to be" [Ileana Bladé]	Accepted. Text modified.
8-608	A	42:29	42:37	This paragraph seems a little weak. For starters, there is not one single reference. Also, the text states what the goal of these simulations is not but not what the actual goal is. The finding that ENSO events are associated with a PNA-like response is not specific to these kinds of multi-century simulations. Finally, I object to the last statement. Many GCMs forced with prescribed SSTs produce a PNA-like response which is shifted west relative to the observed pattern. Also, I think it is more the distribution of tropical rainfall that is relevant, rather than the distribution of SST. [Ileana Bladé]	Accepted. Text modified and reference to Wittenberg et al. (2006) added.
8-609	A	42:29	42:37	A figure could be included here to illustrate the PNA pattern and the impact of a poor ENSO simulation on its spatial configuration. [Gill Martin]	Rejected. Space constraints do not allow another figure.
8-610	A	42:34		westward relative to" could be replaced by "west of [Philip Mote]	Accepted. Text modified.
8-611	A	42:49	42:49	MEAN temperature and sea level pressure [Ileana Bladé]	Rejected. The suggested change would make the statement too narrow. Wu

No.	Batch	Page:line From	Page:line To	Comment	Notes
					and Strauss (2004b) also link the COWL pattern to the spatial pattern of observed temperature trends.
8-612	A	43:2	43:6	This is a fascinating punch line! What is the interpretation for model evaluation? [Philip Mote]	Noted. As with any other observational result, it can be a target for model evaluation, but we are not aware of any attempt to replicate the Quadrelli and Wallace (2004) results using model output.
8-613	A	43:6	43:6	add "in observations (reference ?) and models (Wu and Strauss 2004)". [Ileana Bladé]	Rejected. The statement in the draft applies to the findings of Quadrelli and Wallace (2004), which are based on observed data. Wu and Straus (2004b) also looked at reanalysis data (not models), but the patterns they considered are not the same as those of Quadrelli and Wallace. We are unaware of any modeling studies on this topic, and thus leave the statement be left unchanged.
8-614	A	43:8		Section 8.4.5. Reference could be made here to Ringer et al. (2005; previously Martin et al., 2005b, see comment 24) in which an evaluation of synoptic-scale weather regimes over Europe and of northern hemisphere blocking is carried out with HadGEM1. It was found that HadGEM1 is capable of reproducing synoptic variabilty over Europe which is comparable with reality, at least in terms of weather regime distributions, and shows a significant improvement over HadAM3. HadGEM1 reproduces the observed preferred areas of northern hemisphere blocking and the blocking frequency in the north Atlantic sector is realistic. However, the blocking frequency over the Pacific in the coupled model is much lower than observed, which is though to be related to a cold SST bias in the equatorial Pacific. [Gill Martin]	Rejected. Reference not available.
8-615	A	43:10	43:28	Need a more explicit definition of regimes, and of "sectorial" - this paragraph ends up being rather vague (and with too many references) [Philip Mote]	Accepted. Text modified.
8-616	A	43:24	43:24	After (Corti et al. 1999) I would add the following (or a similar) sentence: However the previous paradigm apply well when weak forcing variations are considered; strong forcing variations, on the other hand, can alter the number (and the “shape”) of flow	Accepted. Text modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				regimes (Molteni and Corti, 1998 ("Long term fluctuations in the statistical properties of low-frequency variability: dynamical origin and predictability" Q. J. R. Meteorol. Soc. 124, 495-526)), Straus and Molteni 2004 (: Circulation regimes and SST forcing: Results from large GCM ensembles. J. Climate, 17, 1641-1656) [SUSANNA CORTI]	
8-617	A	43:24	43:24	SUGGESTION FOR A SUPPLEMENTARY STATEMENT: However the previous paradigm apply well when modest forcing variations are applied; strong forcing variations, on the other hand, can alter the structure of the phase space (Molteni and Corti, 1998 "Long term fluctuations in the statistical properties of low-frequency variability: dynamical origin and predictability" Q. J. R. Meteorol. Soc. 124, 495-526) [Paolo Michele Ruti]	Taken into account. See 8-616.
8-618	A	43:24	43:24	SUGGESTION FOR A SUPPLEMENTARY STATEMENT: On the other hand, the sub-tropical jet could play a relevant role in defining the uni-modal or multi-modal state of the mid-latitude planetary waves (Ruti, P.M., V. Lucarini, A. Dell'Aquila, S. Calmanti, A. Speranza, 2005: "Does the subtropical jet catalyze the mid-latitude atmospheric regimes? ", GRL, revised submission.). [Paolo Michele Ruti]	Rejected. Reference not available.
8-619	A	43:24		The "ideas" tested observationally in Corti et al were based on Palmer, T.N. 1999: A Nonlinear Dynamical Perspective on Climate Prediction. J.Clim., 12, 575-591. [Timothy Palmer]	Accepted. Text modified.
8-620	A	43:25	43:28	SUGGESTION FOR A SUPPLEMENTARY STATEMENT: Recently, Christiansen, Bo, (2005 - On the bimodality of planetary-scale atmospheric wave amplitude index. J. Atmos. Sci.,. In press) has confirmed the bimodal signature of the mid-latitude planetary waves. [Paolo Michele Ruti]	Rejected. Present references deemed adequate.
8-621	A	43:26	43:28	"About the sentence: "the statistical significance of the regimes has been questioned". I would change this sentence with something like "the statistical significance of regimes has been discussed and (still) represents an open issue." Furthermore the following reference should be added: Molteni et al. 2005: Molteni, Kuchraski and Corti, On the predictability of flow-regime properties on interannual to interdecadal timescales. In Predictability of Weather and Climate, Cambridge Press, Palmer and Hagedorn Eds. Cambridge 2005 [SUSANNA CORTI]	Accepted. Text modified. Reference not included as it is not available.
8-622	A	43:54	43:54	"a period of 50 to 100 years" ? Maybe "timescales" is better [Ileana Bladé]	Accepted. Text modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-623	A	43:56		Update this with the contents of Chapter 6. [Philip Mote]	Noted. Will consider an update in a future version.
8-624	A	44:1	44:2	Dipole is typically used to describe oppositely signed features with similar amplitudes. This is not the case of multidecadal variability in the Atlantic the North Atlantic SST signal is considerable larger than the South Atlantic signal (Enfield et al., 2001) and dipole is not an appropriate term. Recent studies have shown that the cross-equatorial SST gradient is the important variable for regional climate. [Robert Molinari]	Rejected. There is indeed a dipole, with anomalies of opposite signs in the North and South Atlantic. See Fig.1 of Latif et al. 2006 (J. Climate, in press). However, this dipole should not be confused with the tropical Atlantic “dipole” which does not really exist. Only for the latter, the term gradient mode would be justified. The Atlantic multidecadal variability, however, involves a real dipole (see also the early papers by Folland and colleagues).
8-625	A	44:2		Recent studies of daily date from the HadCM3 coupled model have demonstrated that there is a statistically significant impact of this SST pattern on the potential predictability of NAO: the SST pattern has been found to be Granger causal for NAO. See: Mosedale, T.J., D.B. Stephenson, M. Collins, and T.C. Mills, 2005: Granger Causality of Coupled Climate Processes: Ocean Feedback on the North Atlantic Oscillation, J. Climate, (in press). Mosedale, T.J., D.B. Stephenson and M. Collins 2005: Atlantic Atmosphere-Ocean Interaction: A Stochastic Climate Model-Based Diagnosis, J. Climate, 18, 1086-1095. [David Stephenson]	Rejected. This is very controversial. Some studies show some impact on the NAO, others, however, show that the NAO drives the Atlantic multidecadal variability (Latif et al. 2006, J. Climate, in press). We prefer to stick to the well established relationships, e.g. Sahelian rainfall and Atlantic hurricane activity.
8-626	A	44:5	44:5	Suggest to add Dai et al. (2005) to the citation: Dai, A., A. Hu, G. A. Meehl, W. M. Washington, and W. G. Strand, 2005: Atlantic thermohaline circulation in a coupled model: Unforced variations vs. forced changes. J. Climate, 18, 2990–3013, which shows a sharp 24yr THC and NAO cycle in the NCAR PCM caused by density anomalies associated THC-induced SST and SSS changes. [Aiguo Dai]	Rejected. The simulated period of 24 years is too far off the observed 50-100 years timescale.
8-627	A	44:19	45:10	As evidence of advances of ENSO modeling in IPCC class models, it might be worth mentioning here that, in at least one case, the same model used for IPCC deccen studies is also being used in operational ENSO-related seasonal-to-interannual forecasts applications. In a somewhat different context, this is already mentioned later on page 48, lines 16-26, and is noted in Wittenberg et al., "GFDL's CM2 global coupled climate models - Part 3: Tropical Pacific climate and ENSO", accepted by J. Climate. (PDF at http://nomads.gfdl.noaa.gov/CM2.X/references/)	Accepted. Text modified and reference to Wittenberg et al. (2006) included.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Keith Dixon]	
8-628	A	44:19	45:11	The discussion of ENSO in this model evaluation seems cursory. As ENSO is the dominant mode of natural variability outside of the annual cycle, and has enormous implications for extreme events and their impacts, particularly in highly populated tropical regions, this phenomenon should receive thorough discussion in both the model evaluation and climate projection chapters (8, 10). The fact that the models do not simulate well this important aspect of present climate variability is important for the evaluation of projected climate. A figure or two illustrating the models' performance is needed. [Anji Seth]	Accepted. Figure from AchutaRao and Sperber (2006) showing simulated power spectra has been included.
8-629	A	44:19	47:35	I feel a few phrases about scale interactions in the tropics are mising. New work has pointed out the role of diurnal cycle on intraseasonal variability (Bernie et al, J. Clim 2005) as well as the role of intrasesonal variability on El Nino (See review by Lengaigne et al. 2004 - Ocean-Atmosphere Interaction and Climate Variability. AGU Monograph) or the role of the seasonal cycle on El Nino (Guilyardi, Clim Dyn 2005 in press). Certainly more will come out in time for AR5 but this an important new area of research. [Eric Guilyardi]	Rejected. This issue is mentioned in the last paragraph. Additional comments are beyond the scope of this section.
8-630	A	44:19		Section 8.4.7. Given the importance of ENSO, it would be appropriate to include a figure showing the IPCC AR4 models' performance, e.g. reproduce one from AchutaRao and Sperber (2005) [Gill Martin]	Taken into account. See 8-628.
8-631	A	44:19		This section would be greatly strengthened if a figure could be constructed to indicate model skill at ENSO variability, perhaps a power spectrum of Nino3.4.(as in the TAR but with all the models). [Philip Mote]	Taken into account. See 8-628.
8-632	A	44:22	44:24	The intercomparisions of CGCM simulations of ENSO by Latif et al. (2001, Climate Dynamics) should also be refered for the "steady progress in simulating and predicting ENSO". [Jin-Yi Yu]	Accepted. Latif et al. (2001) reference now included.
8-633	A	44:22	:23	The following recent modeling studies are missing from this paragraph and should be included: Martin, G.M., K. Arpe, F. Chauvin, L. Ferranti, K. Maynard, J. Polcher, D.B. Stephenson, P. Tschuck (2000) Simulation of the Asian Summer Monsoon in Five European General Circulation Models. Atmospheric Science Letters, Vol. 1, Issue 1, pp. 37-55. DOI:10.1006/asle.2000.0004 Douville, H., J-F. Royer, J. Polcher, P. Cox, N. Gedney, D.B. Stephenson, and P.J.Valdes, 2000: "Impact of CO2 doubling on the Asian Summer Monsoon: Robust versus model-	Rejected. Not relevant to this sub-section.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				dependent responses", J. Met. Soc. of Japan, Vol. 78, No. 4, pp. 421-439. Douville, H., S. Planton, J.-F. Royer, D.B. Stephenson, S. Tyteca, L. Kergoat, S. Lafont, and R.A. Betts, 2000: "Importance of vegetation feedbacks in doubled-CO2 climate experiments", J. Geophys. Res., Vol. 105, No. D11, pp. 14,841-14,861. [David Stephenson]	
8-634	A	44:27	44:28	The statement beginning "and the application of observations ….." does not seem relevant to the climate change problem and I suggest be omitted or moved to section 8.4.11. [Anthony Hirst]	Rejected. The model's ability to use observations to improve predictions is an indication of improved fidelity. Better models are clearly relevant to the climate change problem.
8-635	A	44:34	44:34	8-44, line 34, add after zonal SST gradient, the equatorial Pacific cold tongue structure, which is too cold, too equatorially confined and extend too far west (Cai et al. 2003), Cai, W. J., Collier, M. A., Gordon, H. B., and Waterman, L. J. (2003). Strong ENSO variability and a Super-ENSO pair in the CSIRO mark 3 coupled climate model. Monthly Weather Review, 131 (7): 1189-1210. < http://ams.allenpress.com/amsonline/?request=get-abstract&issn=1520-0493&volume=131&issue=07&page=1189 > [Wenju Cai]	Accepted. Text modified and reference included.
8-636	A	44:35	44:41	Another important model defeciency should be mentioned is that most CGCMs can not produce the phase locking of ENSO to the seaonal cycle or phase lock into a wrong season. This defeciency still exists in may CGCMs. [Jin-Yi Yu]	Accepted. Text modified.
8-637	A	44:36	44:36	"extent" should be "extend". [Aiguo Dai]	Reject. Correct as is.
8-638	A	44:40	44:40	8-44, line 40, add and too high biennial signals. [Wenju Cai]	Rejected. The original senetence has been removed.
8-639	A	44:40	44:40	8-44, line 40, add and Cai et al. (2003). On the other hand, the statistical relationship between thermocline and Nino temperature appears to be reasonably simulated (Cai et al. 2004). Cai, W. J., McPhaden, M. J., and Collier, M. A. (2004). Multidecadal fluctuations in the relationship between equatorial Pacific heat content anomalies and ENSO amplitude. Geophysical Research Letters, 31: L01201, doi:10.1029/2003GL018714. < http://www.agu.org/pubs/crossref/2004/2003GL018714.shtml >	Rejected. Reference is not required.

No.	Batch	Page:line From	To	Comment	Notes
				[Wenju Cai]	
8-640	A	44:41	44:42	Many models also fail to capture the spatial and temporal structure of the El Nino-La Nino asymmetry (Monahan, A.H. and A. Dai, 2004: The spatial and temporal structure of ENSO nonlinearity. J. Climate, 17, 3026-3036). [Aiguo Dai]	Accepted. Text modified and reference included.
8-641	A	44:41	44:42	In additional to Davey et al. (2002), Latif et al. (2001; Climate Dynamics) also give a very complete summary of model defeciencies in ENSO simulations and should be mentioned here. [Jin-Yi Yu]	Taken into account. See 8-632.
8-642	A	44:42		Coupled climate models have also been found to generally underestimate the skewness seen in observed SST ENSO indices as discussed in these two recent papers: Hannachi, A., D.B. Stephenson, and K.R. Sperber, 2003: Probability-based methods for quantifying nonlinearity in ENSO, Climate Dynamics, 20 (2-3), 241-256. Hannachi, A., D.B. Stephenson, K.R. Sperber, 2004: Corrigendum: Probability-based methods for quantifying nonlinearity in ENSO, Climate Dynamics, 22 1: 69-70. [David Stephenson]	Taken into account. See 8-640. Additional references not required.
8-643	A	44:47	44:47	. Ocean [Ileana Bladé]	Rejected.
8-644	A	44:51	45:10	The relevance of nearly all this paragraph to the climate change problem is not made clear. This discussion of ENSO prediction developments seems rather tangental and distracts the reader from the problem at hand. I think this paragraph should be omitted (except for the final sentence which is clearly relevant) or moved in modified form to section 8.4.11. [Anthony Hirst]	Rejected. The success (or improvements) in ENSO prediction directly points to improvements in model performance and enhances our confidence in climate change projections.
8-645	A	44:53		It is hardly a new breakthrough to claim that weather and climate forecasts should be issued with uncertainty estimates (or in other words as probability forecasts). This statement reveals a staggering level of ignorance in our subject. What is meant by “probabilistic measures of skill” – do you mean skill measures for probability forecasts? The review in this section of recent multi-model methods is far from complete – there has been much progress recently in developing ways to calibrate and combine multi-model predictions (e.g. Coelho et al., JClim, 2003 and references therein). [David Stephenson]	Accepted. Text modified. The assessment is not intended to be complete, but merely to suggest that this is the direction the field has moved since the last assessment. While the reviewer is up to date regarding the issue of issuing uncertainty estimates with climate forecasts, deterministic forecast without uncertainty estimates continue to be made and issued.
8-646	A	44:56	45:3	Initialization using a variational data assimilation method with a coupled system could lead to much improvement in skill. This point might be stressed in this article.	Rejected. Not really a key point for this sub-section.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Toshiyuki Awaji]	
8-647	A	45:1	45:10	Models can reproduce the changes in the tropical radiation balance over the ENSO cycle (Wielicki et al. 2002a). [Richard Allan]	Not sure this reference is needed.
8-648	A	45:1	45:10	is this really model evaluation in a climate context? Seems like a dynamical discussion of ENSO that's not tied to the task at hand. [Philip Mote]	Rejected. The issue of how ENSO predictability might change in a changing climate is clearly relevent. Given this, the potential sources limiting predictability is worth of some note.
8-649	A	45:15		mention here the timescale (30-60 days) [Philip Mote]	Accepted. Text modified.
8-650	A	45:21	45:23	"while not necessarily predictable ….. in northern summer". So what? Suggest deletion. [Anthony Hirst]	Accepted. Deleted as suggested.
8-651	A	45:28		Remove superfluous text: "was (at the time of the TAR) and still" [Richard Allan]	Accepted. Removed as suggested.
8-652	A	45:40	45:46	The first three sentences essentially repeat the same notion twice, or at least it feels that way … If not, then the distinction between both statements should be made more clear. [Ileana Bladé]	Accepted. Text modified and made more concise.
8-653	A	45:45	45:45	Add reference to Ringer et al.(2005) where analysis of the MJO in HadGEM1 shows a more realistic simulation than in HadCM3. [Gill Martin]	Rejected. Reference not available.
8-654	A	45:47		replace "under-simulation" with "under estimate" [Richard Allan]	Accepted. Replaced as suggested.
8-655	A	45:55		note that the coupling is between atmosphere and ocean [Philip Mote]	Accepted. Text modified.
8-656	A	46:13		Figure 8.3.10 doesn't show the double ITCZ in the Indian ocean [Philip Mote]	Rejected. Not relevant.
8-657	A	46:14		this is another reason it would be great to have a plot of global wind fields [Philip Mote]	Noted.
8-658	A	46:19	46:21	The statement that a consensus is emerging that MJO is most realistic when convective parameterizations are based on local vertical stability with a trigger directly contradicts Lin et al. (2005, J. Climate, in press), who found that convective parameterizations linked to moisture convergence seemed to generate the best MJOs. [Leo Donner]	Accepted. We presume that the reviewer is referring to Liu et al. We have added a caveat, but the Liu et al study is far from conclusive. Little detail of the implementation of the schemes is provided in their paper.

No.	Batch	Page:line From	To	Comment	Notes
					And, specific experiments whereby the threshold for deep convection was increased, were not performed. It is also not clear how similar the thresholds are in the two schemes used in their paper
8-659	A	46:25	46:25	Is there QBO in AR4 models? It would be interesting to know. [Eugene Rozanov]	Accepted. Statement added to the end of the subsection.
8-660	A	46:25		Section 8.4.9. It is not clear from this section whether any of the IPCC AR4 models reproduce the QBO, despite improvements in vertical resolution and parametrisations. [Gill Martin]	Accepted. Statement added to the end of the subsection.
8-661	A	46:27		Section 8.4.9. This section does currently not mention that coupled AOGCMs, as employed for IPCC AR4, in general do not resolved the stratosphere and therefore do not simulate the QBO. [Marco A. Giorgetta]	Accepted. Statement added to the end of the subsection.
8-662	A	46:28	46:28	The QBO extends higher than 10 hPa, which is the upper limit of classical rawinsonde observations of the QBO. Rocket measurements and model simulations indicate that the QBO jets form near 3 hPa from where they propagate downwards. The amplitude maximum of the QBO is between 10 and 20 hPa. [Marco A. Giorgetta]	Accepted. Text modified.
8-663	A	46:29	46:29	It should be mentioned that the QBO is important for understanding the interannual variability of trace gases like ozone in the middle atmosphere. [Marco A. Giorgetta]	Rejected. This is the implication of the statement at 46:28
8-664	A	46:30	46:33	This statement should be formulated differently. It should be stated first that the current knowledge of the QBO forcing assumes that a broad spectrum of waves is required. A realistic forcing of the QBO therefore can be achieved either by combining resolved wave forcing with parameterized gravity wave drag (e.g. Scaife et al., 2000; Giorgetta et al., 2002, 2005) or by very high resolution simulations that do not need to parameterize gravity waves relevant for the QBO (Watanabe et al., SOLA,, 2005). Equatorial oscillations (“QBO-like” oscillations) can be driven also by a narrower spectrum, if the provided momentum flux is sufficient, as demonstrated in simulations at standard resolution without gravity wave drag parameterization (Takahashi 1996, 1999, Horinouchi and Yoden, 1998; Hamilton et al., 2001). [McLandress 2002 discusses rather the sensitivity of the diurnal tide to lower equatorial oscillations than the “QBO-like” oscillation used in this study.] [Marco A. Giorgetta]	Rejected. The Watanabe et al., 2005 study (SOLA, vol 1, 189-192) is of moderate horizontal resolution (T106), has fairly coarse vertical resolution, and extends to only 40km elevation. Further, the simulations discussed in this paper do not include a modeled QBO. 46:30-46:33 simply represents a statement about current advances in modeling the QBO at climate model resolutions
8-665	A	46:46	46:46	An equatorial enhancement of gravity wave sources is applied in some, but not in all	Accepted. Text modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				models that produce realistic QBOs. Hence this is not a general requirement for modelling the QBO. [Marco A. Giorgetta]	
8-666	A	46:51	46:51	"and in turn the amount …": redundant. Delete [Ileana Bladé]	Accepted. Text deleted.
8-667	A	46:55		Section 8.4.10. This section describes a number of studies which show different model problems in simulating various aspects of the monsoon. However, apart from the first sentence, there is no overall conclusion from the section. How has the simulation of monsoon precipitation improved since the TAR? Is there any indication that the simulation of interannual variability has improved? Is there any consistency in which monsoon regions are simulated better? How do such errors in these major climate phenomena relate to systematic errors in the precipitation climatologies of the models, which are shown to have improved in section 8.3.5? A comment should be made about the lack of confidence in any predictions of future changes in monsoon variability given these results. [Gill Martin]	Noted. Section rewritten for the second order draft.
8-668	A	47:9	47:9	Before the sentence starting with: “This indicates..”. I would refer another study based on an ensemble of forced atmospheric seasonal integrations (Molteni et al. 2003 (F. Molteni, S. Corti, L. Ferranti and J. M. Slingo, 2003: “Predictability experiments for the Asian summer monsoon: impact of SST anomalies on interannual and intraseasonal variability.” Journal of Climate, 16, 4001-4021). A poor interannual predictability was found for the dominant mode of interannual variability of the Asian summer monsoon, while the second mode (clearly associated with ENSO) was successfully simulated. [SUSANNA CORTI]	Rejected. Existing references are sufficient.
8-669	A	47:9	47:9	SUGGESTION FOR A SUPPLEMENTARY STATEMENT: “In a study based on an ensemble of AMIP-kind seasonal integrations Molteni et al. (F. Molteni, S. Corti, L. Ferranti and J. M. Slingo, 2003: “Predictability experiments for the Asian summer monsoon: impact of SST anomalies on interannual and intraseasonal variability.” Journal of Climate, 16, 4001-4021) poor interannual predictability was found as far as the dominant mode of interannual variability of the Asian summer monsoon is concerned, while the second mode (clearly associated with ENSO) was successfully simulated. [Paolo Michele Ruti]	Rejected. Existing references are sufficient.
8-670	A	47:21	47:21	West Africa or North Africa? [Ileana Bladé]	Accepted. Text modified.
8-671	A	47:34	47:35	Can something be said about the reasons for those differences? [Ileana Bladé]	Noted. The authors do not comment on the reasons for the differences.
8-672	A	47:37	48:50	Section 8.4.11 Why is this section in 8.4?	Taken into account. Subsection too

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Catherine Senior]	short to warrant separate section. Pointer to this material in 8.1 will be made more specific.
8-673	A	48:14	48:14	The link of ENSO with the NAO is a model bias (it occurs in nearly all models), there is no evidence from observations that this link exists. [Wilco Hazeleger]	Rejected. No link between ENSO and NAO is claimed.
8-674	A	48:26	48:26	I'm confused. The Anderson et al. (2004) paper refers to simulations with a GFDL atmospheric/land model with prescribed SST but this paper is quoted in reference to *coupled* GFDL simulations ? [Ileana Bladé]	Accepted. Text clarified.
8-675	A	48:26	48:26	Anderson et al. (2004) should be GFDL Global Atmosphere Development Team (2004). This reference appears twice in the reference list, once as Anderson et al. (2004) (p. 70, l. 48-50) and again (correctly) as GFDL Global Atmosphere Development Team (2004) (p. 79, l. 10-11). Only the latter listing should remain. [Leo Donner]	Accepted. Text modified. Reference is "*The* GFDL Global Atmosphere Development Team"
8-676	A	48:29	28:29	Year missing in ref. [Reto Knutti]	Accepted. Text modified.
8-677	A	48:29	48:29	Supply date for Smith et al reference [Andrew Lacis]	Accepted. Text modified.
8-678	A	48:29	48:29	Reference to Smith et al. The year of publication is to be filled in. [Philippe Tulkens]	Accepted. Text modified.
8-679	A	48:29		Smith et al () missing ref. [Richard Allan]	Accepted. Text modified.
8-680	A	48:29		Year required for Smith et al. citation [Ian Simmonds]	Accepted. Text modified.
8-681	A	48:40	48:43	I don't understand the increase in spread. I thought the spread between ensemble members tapered off much earlier than at 1 year lag [Ileana Bladé]	Noted. Section rewritten for second order draft.
8-682	A	49:0		Section 8.5. There are some statements on 20C trends of extremes, which should be moved to or coordinated with Ch.9, and on future projection of extremes, which should be moved to or coordinated with Ch.10. Specific parts will be suggested below separately. [Seita Emori]	Accepted.
8-683	A	49:8		Section 8.5: Brief comment required about the observational datasets available for model vaildation e.g. there have been improvements in dataset quality/coverage since the TAR. But also that deficiencies exist in the use of reanalysis products for model validation and trend analysis, and also in the short term nature of satellite derived products. Perhaps add	Accepted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				a reference to the Appnedix 3.A.5 in Chapter 3. [John Caesar]	
8-684	A	49:8		Any material on wind extremes? [Brian Hoskins]	No
8-685	A	49:8		Section 8.5 - introductory paragraphs could be trimmed. A bit of redundancy with what follows (e.g., discussion of frost days) [Philip Mote]	Noted
8-686	A	49:23	49:31	The first 3 sentences are difficult to understand (How do you expect the extremes are insensitive to global warming?). The following sentences are monstly on 20C trend. [Seita Emori]	Noted, sentences on trend will be moved to chapter 9
8-687	A	49:23		This is an uncommon description of extreme events - aren't they more commonly described as being extremes of some statistical distribution, the rare outliers, rather than the product of instabilities? [Philip Mote]	Noted
8-688	A	49:28	49:31	The style is a little too personal here. How about a more impersonal "There is no evidence that …" (at least that's how the rest of the chapter reads). [Ileana Bladé]	Accepted
8-689	A	49:28	49:31	The statement that there is no evidence that trends in extreme events can be simulated without anthropogenic forcing should be qualified. The section discusses model capabilities in simulating a wide range of extreme events, some of which, like precipitation, cannot be simulated well. Only the discussion of Kiktev et al. (2003) states clearly that extremes were simulated with and without anthropgenic forcing. The disucssion of Meehl et al. (2004) does not indicate what happens to simulations of frost days and heat waves (defined as "three consecutive warmest nights") if anthrogenic forcing is removed, and without the comparison no attribution can be implied. Further, frost days are hardly an extreme event in many regions for some of the seasons, and there will always be "three consecutive warmest nights;" it's changes in the temperatures of those warmest nights due to anthropogenic forcing that is relevant. It's acknowledged (p. 8-51, l. 33-39) that there is substantial disagreement about the effects of global warming on tropical cyclones, and, if that is the case, there must also be uncertainty about the effects of anthropogenic forcing on tropical cyclones. [Leo Donner]	Noted, attribution remarks moved to chapter 9
8-690	A	49:28		"well simulated" -- subjective statement. Be quantitative. [Philip Mote]	Noted
8-691	A	49:29	49:31	this sentence is good for the executive summary [vincenzo artale]	Noted

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-692	A	49:35	49:35	Kharin ref missing from reference list [Simon Brown]	Accepted
8-693	A	49:35	49:35	Kharin ref is used here in a general sense but it seems odd it is not then used in any specific extremes section. If it is significant enough to use in a general sense surely it must have material useful for the sections later? [Simon Brown]	Accepted
8-694	A	49:41	49:49	Brabson 05, Clark 05 and Kharin 05 find extremes are very sensitive to soil moisture, particularly the drying out of soils. This seems to contradict the conclusion made here that the simulation of extremes is insensitive to the modelling of surface processes. if models incorrectly simulate soil moisture extreme heat events will not be properly simulated. [Simon Brown]	Accepted, sentence is deleted
8-695	A	49:46	49:49	FAR -> AR4? The following sentence ("There is therefore…") is difficult to understand. MOST climate models (NOT ALL of them) explicitly model the processes, so you can compare with the rest. [Seita Emori]	Accepted, sentence is deleted
8-696	A	49:46	49:47	I just noticed an inconsistency between chapters in the use of the acronym "FAR" Chapter 8 uses it to mean "Fourth Assessment Report", but Chapter 1 uses it to mean "First Assessment Report" [Chuck Hakkarinen]	Accepted, no change required in this chapter
8-697	A	49:47	49:50	"There is therefore no evidence that the capacity of climate models to simulate temperature and rainfall extremes I slimiated by uncertainty in how the terrestrial surface is modelled." This statement is based on the fact that most of the AR4 models employ canopy conductance and interception in their land parameterizations. However, in Chapter 7, there is ample discussion of biome shifts, diebacks, etc (see pages 7-7 to 7-12) which would affect albedo, roughness, and partiioning of energy at the surface. Certainly, the inclusion of vegetation response to climate changes would have feedbacks that affect extremes? [Anji Seth]	Accepted, sentence is deleted
8-698	A	50:1		Section 8.5.1. Is there any evidence of a dependence of extreme temperature simulation on model resolution or land surface parametrisation? [Gill Martin]	Yes, accepted
8-699	A	50:3	50:44	Mostly 20C trend. [Seita Emori]	Accepted, trend moved to chapter 9
8-700	A	50:13		Citation incorrect - should be 2005. Again on line 19. [Philip Mote]	Accepted

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-701	A	50:21	50:26	Some caution should be raised in the implication that NCEP reanalysis can be considered as obsrvations for assessing other models. For example, NCEP sites recent work over Canada by this reviewer (http://wwwt.emc.ncep.noaa.gov/mmb/rreanl/) showing substantial biases in NCEP surface fields for both regional and global reanalysis products. [Richard Fernandes]	Accepted
8-703	A	50:23	50:23	Year missing in ref. [Reto Knutti]	Accepted
8-704	A	50:23		Holt et al. () missing ref [Richard Allan]	Accepted
8-705	A	50:25		"well simulated" -- subjective statement. Be quantitative. [Philip Mote]	Noted
8-706	A	50:31		reword as "reproduce the location and magnitude of cold air outbreaks in the current climate." [Philip Mote]	Accepted
8-707	A	50:37	50:38	"heat waves … were associated with the 500 hPa circulation pattern": this is too vague. [Ileana Bladé]	Rejected
8-708	A	50:40	50:44	This paragraph, being more closely related to drought than temperature, would perhaps be better placed in the following section on precipitation, where Burke and Brown (2005) is also cited anyway. [John Caesar]	Accepted, this para moved to section 8.5.2
8-709	A	50:40	50:44	Indicate whether the drought trend described here has been shown to require the presence of anthropogenic forcing. [Leo Donner]	Noted
8-710	A	50:42	50:42	Replace 'area' with 'areas'. [John Caesar]	Accepted
8-711	A	50:42	50:42	Replace 'Hadley Center' with 'Hadley Centre'. [John Caesar]	Accepted
8-712	A	50:44		What observed drying trend? See Figure 3.3.1 - significant drying trend globally only if you choose a certain starting point. PDSI is a lousy measure of drought except for the cornfields of Iowa for which it was developed. [Philip Mote]	Rejected
8-713	A	50:48	50:54	Kimoto et al. (2005) also compared the daily precipitation (over Japan) in an AOGCM at two different resolution and found better statistics at the higher resolution. Kimoto, M., N. Yasutomi, C. Yokoyama and S. Emori, 2005: Projected changes in precipitation characteristics around Japan under the global warming, SOLA , 1, 85-88, doi: 10.2151/sola. 2005-023.	Accepted

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Seita Emori]	
8-714	A	50:50	50:54	Text uses terms here of "coarse resolution" with a quantitative definition, and T239 for "highest resolution" with no indication of what this equates to in the unit typically used in earlier sections of the chapter (such as degrees of latitude or kilometers) [Chuck Hakkarinen]	Noted
8-715	A	50:54	50:57	20C trend. [Seita Emori]	Rejected
8-716	A	51:1	51:1	Did May only look at differences over India ? [Ileana Bladé]	Yes.
8-717	A	51:13	51:15	The way of citing Emori et al. (2005) is not to the point. This work is on parameterization dependence of daily precipitation statistics. So, it would be better cited just after Iorio et al. (2004), i.e., Line 54. Moreover, I would supplement its implication as "Emori et al. (2005) have shown ... below 80%, suggesting that modeled extreme precipitation can be strongly prameterization-dependent." [Seita Emori]	Accepted
8-718	A	51:17	51:20	20C trend. [Seita Emori]	Rejected
8-719	A	51:17		dry or wet areas where? Globally? [Philip Mote]	Globally. The text will be modified.
8-720	A	51:22	51:45	The section on tropical cyclones is not sufficient, given the huge interest this will generate after the 2005 season. It just juxtaposes a couple of contradictory model results without any attempt at an assessment. This also needs to be coordinated with other chapters - there should be one place for a comprehensive discussion of cyclones - observations, model validation and future projections discussed together. [Stefan Rahmstorf]	Accepted, section modified
8-721	A	51:22	51:54	Section 8.5.3 There is still a lot of information in here about changes due to global warming. These belong in Chapter 10. [Catherine Senior]	Accepted, section modified
8-722	A	51:22	51:54	I tried to make some comment on tropical cyclone part (subsection 8.5.3) in the first-order draft of IPCC WG1 report. Three papers are referred in the text. However, I cannot find these in the Reference section or in press. Therefore, I have no comments to submit. I think that the text in subsection 8.5.3 are written well so that I can understand what the author wants to say. However, I cannot evaluate scientific results without more information in the form of written papers and reports which include some detail of experimetal results. [Masanori Yamasaki]	Accepted, reference included

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-723	A	51:22	52:10	Perhaps cite Kerry Emmanuel's recent Nature paper in this discussion of possible GHG-induced changes in hurricanes. [Keith Dixon]	Rejected, GHG induced changes in chapter 9 & 10
8-724	A	51:22		Section 8.5.3 - first paragraph says resolution "not high enough to resolve tropical cyclones, especially their sensitivity" and the third paragraph seems to contradict that. [Philip Mote]	Noted, language modified
8-725	A	51:22		Section 8.5.3. I feel that this section is too certain in its conclusions. At the start, the authors state that GCMs cannot simulate tropical cyclone intensities well, but then go on to make conclusions about them using GCMs, referring also to the results of regional rather than global models (Knutson and Tuleya 2004). This study was run at a high resolution (9 km). In contrast, the GCM study of Bengtsson et al. (2005), which to my knowledge is still under review, was run at a horizontal resolution of T63, which despite the quality of other results achieved is quite inadequate for the generation of tropical cyclones with substantial intensities. This makes one wonder about how reliable are predictions with this model of tropical cyclone intensities. Bengtsson et al. (2005) shows good agreement with intensities as represented in ERA-40, but of course this reanalysis data set has a limited horizontal resolution itself. It is not clear which ERA-40 resolution was used in their paper, but even the full resolution data set is too coarse to resolve tropical cyclone intensities fully. At least the regional models are run at much finer resolutions, and even they do not yet capture the full extent of tropical cyclone intensity variations. Thus I suggest the following changes: Paragraph from lines 33 to 39. Change this to the following: "While the results of regional models and theoretical techniques suggest an increase in tropical cyclone intensity in a warmer world (section 9.5.2.6), the resolution of current CGCMs needs to be increased for them to be used to make robust inferences on this issue. For example, the CGCM study of Bengtsson et al. (2005) concluded that there were no indications of intensity increases in their study. Several idealized experiments have suggested that global tropical cyclone numbers may decrease due to global warming (Bengtsson et al. 2005; Yoshimura and Sugi 2005)." Certainly, this section should refer to the discussion in section 9.5.2.6. Because of their much better simulation of actual cyclone intensities, the regional models are arguably more reliable than GCMs for this purpose. An exception, of course, is recent work performed with the Earth System Simulator, which was run at a global resolution comparable to the regional model studies, but I'm not sure if that work made it into this report due to the deadline for submission of published articles. Next paragraph, lines 41-46. "Bengtsson et al. (2005) have shown that the global metrics of tropical cyclones (tropical or hemispheric averages) are broadly reproduced by the	Accepted

No.	Batch	Page:line From	Page:line To	Comment	Notes
				ECHAM model, even as a function of intensity". Not true for tropical cyclones, as the comparison in this paper was to data that was already well degraded in resolution. May be true for extratropical cyclones, which are much larger. "Surprisingly, the current generation models have a remarkable ability to simulate the statistics and the geographical distributions of tropical cyclones". It is not surprising at all if the detection thresholds used to pick out the tropical cyclones from the model output are tuned to get a good climatological distribution of cyclones! As such studies have usually done. However, the authors do have a point: tropical cyclones can be generated by climate models in roughly the right places. Therefore, replace this paragraph with the following, which can then be added to start of the following paragraph. "Although there have been few studies, the current generation of climate models can generate substantial numbers of tropical cyclones in regions where they are observed to form (Bengtsson et al. 2005; Yoshimura and Sugi 2005)." Also, Yoshimura and Sugi (2005) is not in the reference list. [Kevin Walsh]	
8-726	A	51:24	51:31	This paragraph would be greatly enhanced if the authors used some specific quantification of resolution. For example, add phrases like "models used in the IPCC have horizontal resolutions of x to y km, but simulation of tropical cyclones will require resolutions finer than z km. Boundary conditions from GCM runs at T106 (100 km resolution) are now being applied to regional models run at Txxx (J km) resolution by (authors) to simulate tropical cyclone formation and evolution. [Chuck Hakkarinen]	Rejected, not possible to make definitive statement
8-727	A	51:33	51:39	Future projection. [Seita Emori]	Accepted, some text to move to chapter 10
8-728	A	51:33	51:39	<In page 51 of chapter 8, limited number of past studies should be referred for 8.5.3 Tropical Cyclones because KATRINA and RITA were very important events. Therefore, I recommend the follwoing table should be included.> Modeling studies using GCMs. -- paper model description freq change comment -- Broccoli and Manabe (1990) 5~3-deg atm, slab ocean +17%, -12% dependent on cloud physics Haarsmaa et al. (1993) 2.5x3.75-deg atm, slab ocean +50% max intensity +15%	Rejected, not an acceptabl reference by IPCC

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Bengtsson et al. 1.1-deg atm, -37% no intensity change (1996) slab ocean Sugi et al. 1.1-deg atm, -34%, no intensity change (2002) specified SST +60% in Atlantic Tsutsui (2002) 2.8-deg atm, no significant slight increase in specified SST change intensities Ouchi et al. 20-km atm, -30%, increased intense (2005) specified SST +30% in Atlantic (more than 45 m/s) storms -- This table is partly based on an AMS seminar "New Orleans, Hurricanes and Climate Change: A Question of Resiliency". The presentation material is available from http://www.ametsoc.org/atmospolicy/documents/6-20-05NewOrleansandHurricanes-ThomasKnutson.pdf <Please add the following papers in the reference> Bengtsson., L, M. Botzet, and M. Esch, 1996:Will Greenhouse Gas-Induced Warming over the Next 50 Years Lead to Higher Frequency and Greater Intensity of Hurricanes? , Tellus, 48A, 57-73. Broccoli, A. J., and S. Manabe, 1990:Can existing climate models be used to study anthropogenic changes in tropical cyclone climate? ,Geophys. Res. Lett., 17, 1917-1920. Haarsma, R. J., J. F. B. Mitchell, and C. A. Senior, 1993: Tropical disturbances in a GCM.Climate Dyn., 8, 247-257. Oouchi, K., J. Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda, 2005: Tropical cyclone climatology in a global-warming climate as simulated in a 20km-mesh global atmospheric model. J. Meteor. Soc. Japan, in revision. Sugi, M., A. Noda, and N. Sato, 2002:Influence of the global warming on tropical cyclone climatology: An experiment with the JMA global model.,J. Meteor. Soc. Japan, 80, 249-272. Tsutsui., J., 2002:Implications of anthropogenic climate change for tropical cyclone activity: A case study with the NCAR CCM2.,J. Meteor. Soc. Japan, 80, 45-65. [Koki Maruyama]	
8-729	A	51:33	51:34	The disagreements between models of future intensity changes does not appear to be consisitent with the conclusions of chapter 10 (Q10.1). [Ruth McDonald]	Noted, to coordinate with chapter 10
8-730	A	51:33	51:38	It would be interesting here to compare how well the models did in simulating the increased intensity of hurricanes in the past 50 years, as compared to the observational studies of Emmanuel (Nature, 2005) and Webster et al. (Science, 2005). [Alan Robock]	Noted

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-731	A	51:33	51:39	The results of projection of climate change should be described in Chapter 10. Recent model results seem to have reached some consensus: reduction of total number of tropical cyclones and intensification of tropical cyclones (increase in the number of intense tropical cyclones) in the future warmer climate. [Masato Sugi]	Accepted
8-732	A	51:34	51:34	Knutson and Tuleya (2004), using a 9-km grid regional model, find that.. [Thomas Knutson]	Accepted
8-733	A	51:35	51:36	…conditions." Oouchi et al. (2005), using a 20-km grid global model, report that the number of intense tropical cylones, and the intensities of the strongest tropical cylones, increase in their greenhouse warming "time-slice" experiments. In contrast, Bengtsson et al. (2005), using a <GIVE BRIEF DETAIL ON THEIR MODEL>, conclude that... [Thomas Knutson]	Accepted
8-734	A	51:35	51:37	I think it should be noted here that that no warm core criteria is applied when the cyclones are located and tracked in the Bengtsson et al (2005) study. Therefore the cyclones are not necessarily tropical cyclones. [Ruth McDonald]	Noted
8-735	A	51:36	51:39	The references to Bengtsson et al., (2005) seem rather contradictory here. [Gill Martin]	Rejected
8-736	A	51:37	51:39	Although most models show global decreases in tropical storm frequency, there are increases in tropical storm frequency in some basins. The models disagree on the sign of the frequency changes in the individual basins. [Ruth McDonald]	Noted
8-737	A	51:39	51:39	I cannot find this citation: Yoshimura and Sugi, 2005. I suggest that all citations be cross-checked for accuracy. [Hughes Dan]	Noted
8-738	A	51:41	51:46	Hasegawa and Emori (2005) successfully validated the simulated mean daily precipitation intensity associated with tropical cyclones over the western North Pacific basin. This work would fit in this paragraph. Hasegawa, A. and S. Emori, 2005: Tropical cyclones and associated precipitation over the western North Pacific: T106 atmospheric GCM simulation for present and doubled CO2 climates, SOLA, 1, 145-148, doi:10.2151/sola.2005-038. [Seita Emori]	Accepted
8-739	A	51:41	51:46	<In the context of the simulation of tropical cyclones in the current climate, I would like to include the following sentence around line 46 of page 51>. "Tsutsui et al. (2004) presented results from the ensemble climate simulations with different horizontal resolutions in the range fromT42 to T341, implying that SST-forcing is partly responsible for interannual variations of observed TC frequencies".	Accepted

No.	Batch	Page:line From	Page:line To	Comment	Notes
				<Please add the following paper in the reference.> Tsutsui, J., H. Hatsushika, and H. Kitabata, 2004:Interannual variability of tropical cyclone frequencies implied from an ensemble climate simulation with the NCAR Community AtmospherebModel., 26th Conference on Hurricanes and Tropical Meteorology,Amer. Meteor. Soc., May 3-7, 2004, Miami, Florida, 433-434, accepted. [Koki Maruyama]	
8-740	A	51:41	51:46	McDonald et al. (2005) found that the geographic distribution of tropical cyclones in an N144 version of HadAM3 (the atmospheric component of HadCM3) compared well with observations. McDonald RE, Bleaken DG, Cresswell DR, Pope VD and Senior CA (2005) Tropical storms: representation and diagnosis in climate models and the impacts of climate change.Climate Dynamics 25: 19-36 [Ruth McDonald]	Noted
8-741	A	51:41	51:46	Validation of the specific models (Sugi et al. 2002, McDonald et al. 2005, Hasegawa and Emori, 2005, Yoshimura and Sugi, 2005, Yoshimura et al. 2005) used for climate change experiments on tropical cyclone activity described in Chapter 10 should be given here. Bengtsson (1995, 1996) already have shown a remarkable ability of 100km resolution AGCM to reproduce the climatological feature of tropical cyclones such as the geographical distribution. Recently several models with the same resolution have been used for the climate change experiment on tropical cyclone activity and their overall ability to simulate the tropical cyclones is found to be reasonable. On the other hand, their ability to simulate the very strong wind near the cyclone center and the low central pressure of tropical cyclones is not sufficient, even in the 20km resolution model. However, the simulation results using these models give reasonable information on the impact of climate change on tropical cyclone activity if a proper interpretation of the results is made. [Masato Sugi]	Noted
8-742	A	51:48	51:54	References needed after "… able to simulate those differences" (Line 50) and after "… during the past 50 years" (Line 53). Also note that the latter sentence is on 20C trend. [Seita Emori]	Accepted
8-743	A	51:52	51:54	20C trend. [Seita Emori]	Accepted
8-744	A	52:4	52:6	Given the interest in the effects of climate change on tropical cyclones, the finding "There is no agreement among the models whether global warming will make tropical cyclones more or less intense. There seems to be some agreement among models that the frequency of tropical cyclones will be reduced." should be repeated in the Executive Summary	Noted

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Lenny Bernstein]	
8-745	A	52:4	52:6	Future projection. [Seita Emori]	Noted
8-746	A	52:4	52:6	the finding "There is no agreement among the models whether global warming will make tropical cyclones more or less intense. There seems to be some agreement among models that the frequency of tropical cyclones will be reduced." should be repeated in the Executive Summary [Howard Feldman]	Deleted. (moved to Chapter 10)
8-747	A	52:4	52:6	Tropical cyclones are among the most deadly of climate phenomena and a subject of intense interest in many parts of the world. Therefore, the conclusion: "There is no agreement among the models whether global warming will make tropical cyclones more or less intense. There seems to be some agreement among models that the frequency of tropical cyclones will be reduced." should appear in the Executive Summary and in the higher level summaries (SPM, Synthesis Report) that will be prepared later in the writing process. [Jeffrey Kueter]	Deleted. (moved to Chapter 10)
8-748	A	52:4	52:5	Inconsistent with the statement in chapter 10, page 5, line 46-48 :"New results from global models of around 1 degree resolution, and a new global model with 20 km resolution, show a future global decrease of tropical cyclones of around 30% (but with regional increases in the North Atlantic), and an increase of precipitation in those storms. [Michel Petit]	Noted
8-749	A	52:6	52:10	I agree, in principle, but feel it's too discouraging. Some high-resolution (~1deg) models can reproduce fairly realistic statistics of daily precipitation averaged over ~1deg (Emori et al, 2005; Kimoto et al, 2005; Hasegawa and Emori, 2005 and maybe more others), which should be more stressed here. [Seita Emori]	Noted
8-750	A	52:12	52:35	The first quantitative discussion of radiative forcings and feedback effects is given by Hansen et al. (1984) (see also Hansen et al., 1997). Hansen et al. (1984) express their radiative forcing for doubled CO2 (and 2% solar irradiance increase) in terms of Delta-T-zero, which is the equivalent of adjusted forcing, but expressed in terms of a global surface temperature change with no feedbacks allowed to operate. For estimating global climate change, this is actually a more robust quantity than adjusted forcing. Lacis and Mishchenko (1995) show that for a globally uniform forcing, such as doubled CO2, Delta-T-zero is essentially independent of latitude while the adjusted flux has a significant latitudinal dependence because it depends directly on the magnitude of the local Planck radiation, whereas Delta-T-zero has already taken that into account. [Andrew Lacis]	Discussion of this point (if included at all) belongs in Chapter 2, where radiative forcing is discussed or defined.

No.	Batch	Page:line From	To	Comment	Notes
8-751	A	52:12		It is disappointing in this section not to see any mention of the current range of model climate sensitivities. Yes, this is given in a later chapter as 2.1-4.4 deg. C, but it will be very disconcerting to the reader looking to see whether the canonical 1.5-4.5 deg. C range has changed to not see any mention of it in a section titled "Climate Sensitiivty and Feedbacks," nor to have any discussion of what the current thinking is. Granted, the global number is a simplistic distillation of a lot of different things going on, but some real changes in understanding have taken place relative to previous WG1 ARs, and it is important that AR1 be clear about this. Specifically, if one looks at previous ARs, the high end of the sensitivity range, which some models always simulated, was discounted when prognostic cloud schemes came into fashion, because the early ones always predicted reduced sensitivity - incorrectly, it turns out, because they either had low clouds or anvil clouds brightening with warming, or they had unrealistic ice fallspeeds. I don't think there was ever an IPCC model that actually got a 1.5 deg. C sensitivity, but now, there are none that get a sensitivity below 2.0 deg. C. And there are good reasons for that - we now know that in most of the world, low clouds get optically thinner with warming (Tselioudis and Rossow 1994 and Del Genio and Wolf 2000, already cited). We know that cumulus anvils do not exhibit either thermostat or adaptive iris behavior, and simple thermodynamics and microphysics explains why (Del Genio, A.D., W. Kovari, M.-S. Yao and J. Jonas, 2005: Cumulus microphysics and climate sensitiivty. J. Clim., 18, 2376-2387). We know that on ENSO time scales at least, the SW cloud forcing of low clouds in subsidence regimes becomes less negative and that high sensitiivty models come closer to that than low sensitivity models (Bony and Dufresne 2005, already cited). And we know that a single GCM subjected to an ensemble of perturbations of physical parameters produces a pdf of climate sensitivities that makes the low sensitivity end less probable and the high sensitivity end more probable than the canonical 1.5-4.5 deg. C range (Stainforth et al.2005, already cited). It is time now in AR4 to integrate all these developments and to make a clear statement that the community's best estimate is that the low end of the previously quoted range seems unlikely and a high sensitivity is more probable. And this is the chapter and section in which to do it, rather than Chapter 10, because this is where the advances in understanding of physical processes and comparisons to cloud observations is discussed. [Anthony Del Genio]	Noted. The different issues addressed in this comment are either addressed elsewhere in the report (in particular in the Box on climate sensitivity in chapter 10), or not supported enough by the current litterature. More specifically: (1) Climate sensitivity estimates of individual models are now given in Table 8.1, as well as the range. (2) The range of climate sensitivity estimates is discussed in several chapter: chapter 8 discusses mostly our understanding of this range, chapter 9 the observational constraints on climate sensitivity, and chapter 10 the relative probability of the different sensitivity estimates. A synthesis of these different discussions and an expert judgment on the range of climate sensitivity estimates are presented in Box 10.2. (3) It remains very difficult to constrain climate sensitivity estimates from our understanding (and the evaluation) of feedback processes. Even if observations suggest a neutral cloud feedback from cumulus anvil clouds and a positive feedback from low-level clouds, nobody knows currently how it translates into a climate sensitivity number. We would need to apply these different process studies and observational tests to GCMs to know whether it constrains climate change cloud feedbacks and climate sensitivity. Unfortunately,

No.	Batch	Page:line From	Page:line To	Comment	Notes
					studies of this type are currently too rare in the litterature to allow us to constrain robustly climate sensitivity estimates from feedback or process studies.
8-752	A	52:12		On the sensitivity measurements. There are several interesting reports on the sensitivity measurements. 1) O. Kärner, J. Geophys. Res., 2002 VOL. 107, NO. D20, 4415, doi:10.1029/2001JD002024, 2002, "On nonstationarity and antipersistency in global temperature series." 2) D. H. Douglass, E. G. Blackman, and R. S. Knox, Phys. Lett A 323, 3/10/04, 315-322 (2004) and its Erratum, "Temperature response of Earth to the annual solar irradiance cycle." 3) D. H. Douglass and B. D. Clader, Geophys. Res. Lett., VOL. 29, NO. 16, 0.1029/2002GL015345, 2002 "Climate sensitivity of the Earth to solar irradiance." These reports may be probably not welcome in the climate community and much criticism may exit, but comments should be made because they appeared in good journals. More importantly, a very interesting report of Forster & Gregory (2005) ("The climate sensitivity and its components diagnosed from Earth radiation budget data," J. Climate, submitted), which is cited in this section, points out that the climate sensitivity of models tend to be too large. They also say that among models in TAR only one out of ten gives sensitivity patterns similar to their estimation. [Kiminori Itoh]	Rejected: the AR4 process is an assessment, not a literature review, so the fact that papers have appeared in "good journals" alone does not warrant their inclusion. Douglass and Clader (2002) which suggests a positive feedback to solar forcing belongs (if at all) in Chapter 9. Kärner (2002) is assessed as not presenting a clear and coherent case challenging the amplification of global warming by positive feedbacks compared with the 'no feedbacks' situation. Discussion of the solar influence on observed climate change in this paper belongs (if at all) in Chapter 9. The Forster and Gregory (2005) discussion on strength of water vapour feedback is relevant to the current chapter, but was not included as error bars were so large that significant conclusions could not be drawn (along with a caveat that cloud feedback would need to have been close to neutral, which did not have strong evidence).
8-753	A	52:12		Sect.8.6. It's not clear to me how the scope of this section was chosen and limited. This broad-level discussion of our understanding and role of certain feedbacks is very useful. I would like to see it extended to include carbon cycle feedbacks (likely very important for global temperature/ climate sensitivity), and soil moisture feedbacks (very important for regional temperature/ precip in some areas).	Noted. We now explain in the introduction to 8.6 that climate feedbacks associated with chemical or biochemical processes are addressed in other

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Dave Rowell]	chapters (7 and 10), and that we focus on radiative feedbacks (i.e. associated with changes in climate variables that directly affect the global Earth's radiation budget and surface temperature). Moreover, although we recognize that carbon cycle feedbacks and soil moisture feedbacks can substantially affect the magnitude, the timing or the patterns of climate warming, (1) the carbon cyle feedback affects the rate of increase of CO2 but not the "climate sensitivitiy" (which is defined for a doubling of the CO2 concentration) and (2) there is no evidence in the literature that soil moisture feedbacks substantially affect the "climate sensitivity" (which is defined from global mean temperature and radiation changes).
8-754	A	52:16	52:18	Suggest modification: "Climate sensitivity is a metric used to characterize the response of the global climate system to a given forcing and is broadly defined as the equilibrium global mean surface temperature change following a doubling of atmospheric CO2 concentration." [Richard Allan]	Accepted: text modified.
8-755	A	52:16		The concept of climate sensitivity, which is broadly defined as the equilibrium global mean surface temperature change following a doubling of atmospheric CO2 concentration, is being used to characterize the response of the global climate system to a given forcing. Not necessarily or exclusively equilibrium. Transient sensitivity is important, perhaps more important. We live in a transient world, not an equilibrium world. See also line 47. [Stephen E Schwartz]	Rejected: the concept of equilibrium climate sensitivity remains useful, and has remained consistent in definition across the 4 IPCC reports. For the transient response, other measures of response are available, such as the transient climate response (defined in 8.6.2.1)
8-756	A	52:16		The concept of climate sensitivity, which is broadly defined as the equilibrium global mean surface temperature change following a doubling of atmospheric CO2 concentration, is being used to characterize the response of the global climate system to a given forcing.	Noted: The current definition of ‘climate sensitivity’ is deemed useful because it has a long history, permitting direct

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Consideration should be given to abandoning the definition of sensitivity based on a doubling of CO_2 concentration. The reason for this is that sensitivity will not shift every time further research refines the forcing per doubling of CO_2; this has happened several times during the lifetime of IPCC and will surely occur again. In fact doubling implies an initial concentration, which also changes. Chapter 2 noted (page 2-12) a substantial range of forcing associated with doubled CO_2 in different models: A recent comparison of line-by-line and GCM radiation schemes found that clear sky instantaneous RF and surface forcing agreed very well (better than 10%) among the 5 line-by-line models investigated, using the same single atmospheric background profile. The GCM radiation schemes were less accurate, with ~20% errors in the CO_2 RF ... (Collins et al., 2005 and Chapter 10). Nevertheless, the current set of Atmosphere and Ocean GCMs (AOGCMs) used in Chapter 10 of this report found values for RF, for a doubling of CO_2 that ranged between 3.5 and 4.2 W m –2 , in good agreement with the TAR RF value of 3.7 W m –2 (see Chapter 10 and Forster, 2005). Webb et al (2005) compare forcing for doubled CO_2 in 9 models, with that forcing ranging from 3 to 4 W m-2. Webb, M. J., C. A. Senior, D. M. H. Sexton, 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, and K. Taylor, 2005: On uncertainty in feedback mechanisms controlling climate sensitivity in two GCM ensembles. Clim. Dyn., in revision. A similar conclusion is reached in Table 10.2.1, for which the average and standard deviation forcing for doubled CO_2 for 9 models is 3.71 ± 0.48 W m-2, or ±13% (range 2.99 to 4.23). If the basis for the expression of sensitivity (that is forcing associated with doubled CO_2) is itself uncertain to 20%, it will be impossible to ascertain whether reports of different sensitivities in different AOGCM's are due to different model sensitivities or different forcings associated with doubled CO_2. The Webb paper shows a strong correlation among different models of the increase in equilibrium warming for doubled CO_2 with the increase in forcing for doubled CO_2. How much of this is due to a difference in forcing versus a difference in sensitivity? All the more reason to define sensitivity as K per W m-2. The use of systematic units is to be encouraged throughout science. Use of sensitivity referenced to CO_2 doubling is akin to measuring the density of substances relative to water, and then having to change the density of various substances if further research refines the density of water. Except of course that the forcing due to doubling of CO_2 is a lot less well known than the density of water.	comparisons with previous reports, and for the practical reason that radiative forcing is often or usually not known. It has proved a useful and consistent measure of model response, and there is no evidence that there have been systematic shifts in model climate sensitivity due to changes in radiative forcing from CO_2 increases. Nevertheless, it is important to note (as pointed out in this comment) that part of the spread in climate sensitivity will be due to differences in the calculated "forcing", and that this is a limitation on the concept of "climate sensitivity". A sentence is now added: “Some differences in climate sensitivity will also result simply from differences in the particular radiative forcing calculated by different radiation codes (refer sections 2.3.1 and 8.6.2.3).” The relative magnitude of the contribution of forcing and feedbacks is already discussed in 8.6.2.3.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				I therefore urge that sensitivity be defined as change in global mean surface temperature in response to a radiative forcing of 1 W m-2. Sooner or later as the science is refined so that differences of 20% are important that decision will be made. I urge that it be made sooner in order to advance the science. [Stephen E Schwartz]	
8-757	A	52:17		Concentration: In scientific writing the term "concentration" means amount or mass per volume, typical units mol m-3 or kg m-3, respectively. The measure of abundance of CO2 and other GHGs is mole fraction or mixing ratio (with respect to dry air) with typical unit ppm (µmol mol-1). The use of the term "concentration" in lieu of mole fraction or mixing ratio is of long standing and consequently replacing it throughout might lead to confusion and be net detrimental, but perhaps a footnote stating all of this might be appropriate. [Stephen E Schwartz]	Noted: however footnote belongs in Chapter 2 (if anywhere).
8-758	A	52:36	52:36	Hansen et al. (1984) show that the radiative equivalent of different climate feedback contributions can be readily identified and quantified. Upon running the doubled CO2 model to equilibrium, precise changes in water vapor distribution, clouds, lapse rate, and surface albedo can be tabulated from the GCM diagnostic output, which can then be easily evaluated using a 1D radiative-convective model. Lacis and Mishchenko (1995) perform this evaluation with a 2D model and include also the latitudinal contribution of advective feedbacks which, by definition, average to zero globally. This analysis is made possible because the total equilibirum surface temperature response in the doubled CO2 exeriment has to be completely sustained by the radiative effects due to the changes in water vapor, lapse rate, clouds, and surface albedo that were induced by the total temperature change. (This shows, in effect, that the change in global temperature can serve as a "medium of exchange" between different feedback effects, even though cloud formation is not directly a function of temperature.) [Andrew Lacis]	Noted: this point is, however, not discussed explicitly due to lack of space.
8-759	A	52:37	53:11	Hansen et al. show that while the feedback efficiencies of the different feedback processes can be compared in linear fashion, the feedback effects on the global surface temperature are multiplicative in nature and do not combine linearly. Thus, while the radiative effects of atmospheric constituents can be evaluated with good accuracy, the model physics involved in producing the different feedback processes are necessarily more complex, and thus differ more widely between different GCMs. [Andrew Lacis]	Noted: this point is, however, not discussed explicitly due to lack of space.
8-760	A	52:39	53:10	This draft does not change the definition of climate sensitivity used in the TAR. However that definition implies that climate sensitivity is constant, whereas the text in this section states that it is dependent on the type of forcing applied, and on the mean climate state. If this is so, it is difficult to see the utility of the concept. What this text seems to imply is	The remaining utility of climate sensitivity is explicitly addressed in lines 5-10, p8-53. In our view, the present level of discussion is adequate

No.	Batch	Page:line From	Page:line To	Comment	Notes
				that to understand an estimate of climate sensitivity one must know all the details of case for which it was calculated. Some explanation of the remaining utility of the climate sensitivity concept should be included in this section. [Lenny Bernstein]	considering space restrictions.
8-761	A	52:39	53:10	This draft does not change the definition of climate sensitivity used in the TAR. However that definition implies that climate sensitivity is constant, whereas the text in this section states that it is dependent on the type of forcing applied, and on the mean climate state. If this is so, it is difficult to see the utility of the concept. What this text seems to imply is that to understand an estimate of climate sensitivity one must know all the details of case for which it was calculated. Some explanation of the remaining utility of the climate sensitivity concept should be included in this section. [Jeffrey Kueter]	Identical response to 8-760.
8-762	A	52:42	52:43	climate sensitivity: is estimated approximately from atmos models with slab ocean, or computed by simulating a coupled equilibrium state with doubled CO2 (standard practice in EMICs) [Stefan Rahmstorf]	Text unchanged: The section refers only to climate sensitivity in GCMs, so the calculation of climate sensitivity in EMICs is not directly relevent here.
8-763	A	52:47	52:52	Change Cubash to Cubasch twice [Reto Knutti]	Done.
8-764	A	52:48	52:48	add 'in a 1%/yr atmospheric CO2 increase scenario' in the definition of TCR [Reto Knutti]	Done.
8-765	A	52:49	52:51	Effective climate sensitivity can be defined at any point in time, and hence suggesting that it links the equilibrium climate sensitivity to the TCR could be misleading. [Catherine Senior]	Accepted: wording changed to remove the link between the two. New definition: “An estimate of the equilibrium climate sensitivity in transient climate change integrations is obtained from the *effective climate sensitivity* (Murphy 1995)….”
8-766	A	52:55	52:55	Gregory et al (2004) might be relevant here too, as it shows that HadCM3 has half the sensitivity to solar forcing as to CO2 forcing. That's not the same Gregory et al (2004) as in the reference list of chapter 8, but author = {Gregory, J. M. and Ingram, W. J. and	Noted: however the point is already adequately referenced -- extra references not included due to space

No.	Batch	Page:line From	To	Comment	Notes
				Palmer, M. A. and Jones, G. S. and Stott, P. A. and Thorpe, R. B. and Lowe, J. A. and Johns, T. C. and Williams, K. D.}, title = {A new method for diagnosing radiative forcing and climate sensitivity}, journal = {Geophys. Res. Lett.}, year = {2004}, volume = {31}, pages = {L03205, doi:10.1029/2003gl018747}. [Jonathan Gregory]	limitations.
8-767	A	52:57	52:57	Could also reference Senior and Mitchell 2000 [Catherine Senior]	Noted: however the point is already adequately referenced -- extra references not included due to space limitations.
8-768	A	53:0		Replace Heading with "Changes in model estimates since the TAR" [Vincent Gray]	Rejected: heading to 8.6.2.2 chosen deliberately to address issues of *cause* of sensitivity changes.
8-769	A	53:15	53:15	Improved parameterization of clouds, boundary layer, and convection in the models used for AR4 is cited as the first factor explaining changes in model estimates since TAR, and Section 8.2 is referenced. However, Section 8.2 contains very little information about these changes in parameterizations, with only the Lock (2001) boundary-layer parameterization mentioned there. [Leo Donner]	Accepted. Section 8.2 now includes more information about changes in these parameterizations.
8-770	A	53:20	53:20	the overall quality of the model simulation [Ileana Bladé]	Change made.
8-771	A	53:24	53:37	The references to Williams et al 2005b will have to be changed to Johns et al 2005 [Catherine Senior]	Change made.
8-772	A	53:26	53:26	Sensitivity for HadCM3 unchanged, but large for HadGEM1. Please specify clearly. [Reto Knutti]	Accepted. The climate sensitivity of HadSM1 is larger than that of HadSM3, but this difference (primarily due to different ice-albedo feedbacks) concerns only "slab" models. The coupled models (HadGEM1 and HadCM3) have very similar sensitivities. To clarify, we mention the Hadley Center twice: once for the slab model, and once for the full coupled model.
8-773	A	53:30	53:37	The Johns et al. paper (submitted to J Climate, as mentioned in comment 3), if accepted by the journal, should be cited here intsead of Williams et al. (2005b). [Timothy Johns]	Change made.
8-774	A	53:30	53:37	Williams et al. (2005b) has been rejected for publication in ASL. However all of the	Change made.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				results/discussion in the paper relevant for this paragraph have been transferred to Johns et al (2005). Hence the references to Williams et al. (2005b) can simply be replaced with Johns et al. (2005). [Keith Williams]	
8-775	A	53:32	53:33	At end of sentence "Also, the parameterization…A+B" cite Stainforth et al (2005) as they explicitly demonstrate this nonlinear interaction between parameters. [David Sexton]	Citation added.
8-776	A	53:37	53:37	finish the sentence (what's the impact in the GFDL model?) [Ileana Bladé]	Rejected. We do not write explicitely that the introduction of the Lock parameterization has decreased the sensitivity of the GFDL model because of space restriction, and because the important point of this sentence is that the impact on climate sensitivity of a parameterization change is model dependent (the sign does not matter much here). The reference Soden et al. (2004) is given, in which more information about the impact of the Lock scheme may be found.
8-777	A	53:54	53:54	Scheider, Cash and Bengtsson have a paper, currently in press in J. Climate (sorry, I'm unable to provide the title) in which they compare the greenhouse sensitivities of CCM3 and ECHAM4.5 by effectively "transplanting" pieces of one model's parametrization into the other. They find that replacing the radiation scheme associated with liquid cloud water (I believe) of CCM3 with that of ECHAM4.5 greatly reduces most of the difference in their global mean temperature projection (2xC02) as well as in the spatial structure of those differences ! [Ileana Bladé]	Noted. But this paper will not be accepted on time to be cited in the AR4.
8-778	A	54:0	57:	Section 8.6.3.1: There is repetition between sections 8.6.3.1 and 8.6.3.1.1 and I believe that combining the sections will help to elucidate and distil the assessment since there is no need to separate the model and observational findings. Specific suggestions are noted below and I also include my suggestion for a combined section 8.6.3.1-8.8.3.1.1 further below: (1) State in first paragraph of 8.6.3.1 that surface and atmospheric temperature are observed to be coupled (Section 3.4.1) and that the combined water vapour-lapse rate feedback is relatively insensitive to changes in the temperature lapse rate if RH is constant due to compensation between temperature and water vapour thereby explaining why it is	Overall: taken into account: There are a number of helpful comments here, and separate responses are listed below. Structure in 8.6.3.1 left unchanged (although some wording changes adopted from supplementary material). The structure in 8.6.3.1 was chosen in

No.	Batch	Page:line From	To	Comment	Notes
				logical to treat the two feedbacks together (2) combine the paragraphs from both sections dealing with (i) humidity distribution (ii) humidity variability and combine the paragraphs dealing with vertical humidity correlations with lapse rate (3) restructure the paragraph dealing with Pinatubo, stating more clearly the real advance which is that models can only reproduce the observed cooling and drying following Pinatubo if water vapour feedback is switched on and then noting caveavts and additional studies. (4) Reduce the paragraph on stratospheric water vapour since this is dealt with in section 3.4.2.4 and it is not clear that this is a feedback process. [Richard Allan]	part for consistency with the structure in 8.6.3.2 on cloud feedback, and also to seperately highlight the progress that has been made in both observational evaluation of water vapour feedback, and on the evaluation of feedbacks/feedback processes in models. It was also chosen so as to provide some structure through separating conceptual and simple model studies from GCM evaluation as far as possible. (1) First part: accepted: coupling now mentioned. Second part: accepted: short comment inserted linking water vapour/temperature impact on OLR in introduction. (2) First part: see comment above. Second part: accepted – paragraphs on vertical profiles and lapse rate combined. (3) Paragraph structure left unchanged: The paragraph was structured so as to put together the 'trends' evaluation (compared with the earlier variability evaluation). The level of emphasis on model agreement with Pinatubo cooling only with WV feedback present is not greater because although it is an important finding, it is a single modelling result only, and the degree of agreement is subject to significant uncertainties from a number of sources (e.g. climate noise, possible cloud feedbacks, uncertainties in forcing). (4) Accepted in part: Paragraph reduced, and more cautious on observed feedback. (Note however

No.	Batch	Page:line From	Page:line To	Comment	Notes
					section 3.4.2.4 does not discuss the possible feedback mechanism at any length, or differing responses to different forcing, such as CO2 or O3).
8-779	A	54:0		Section 8.6.3. This detailed and thoughtful section would benefit considerably from more consideration of observational constraints on climate sensitivity. In addition, the section would benefit from tighter organization, as the text seems to backtrack occasionally, making it hard for the reader to follow. Some of the phrasing is a little convoluted: for instance, section 8.6.3.1.2. "…no substantial evidence suggests that the broadscale RH response of models to climate change constitutes an artefact of GCMs." This is an obscure way of saying that the water vapour feedback in the GCMs is consistent with observationally-based estimates. Rephrase, here and elsewhere in this section. [Kevin Walsh]	First part: rejected – observational constraints on climate sensitivity covered in chapter 9. Second part: no specific organisational changes suggested. Third part: Noted: and changes made to the sentence to make it less convoluted. Note, however that sentence was not changed to that suggested, since the general wording was chosen deliberately, and does not mean the same thing as "water vapour feedback in models is consistent with observationally based estimates". The first statement includes evidence from various other types of models, such as cloud resolving and mesoscale models, and includes GCMs with very high vertical resolution.
8-780	A	54:15		An estimate of the albedo effect is shown in Chapter 4, box 4.1, Figure 1. [Philip Mote]	Noted. Reference to Box 4.1 has been added in the last paragraph before 8.6.3 and at the beginning of the 2nd paragraph of section 8.6.3.4.
8-781	A	54:21		Section 8.6.3: this is an important section. Keep it intact! [Philip Mote]	Noted.
8-782	A	54:30	54:32	This sentence is slightly confusing. Earlier in this chapter climate sensitivity is defined as the temperature change with doubling of CO2. Therefore it is not the "defined with respect to a specified CO2 forcing (e.g., a CO2 doubling)", which implies the doubling of CO2 is just one specification used in climate sensitivity. This should read more like : "defined with respect to a forcing from a doubling of CO2" [Gareth S. Jones]	Accepted: change made.
8-783	A	54:32	54:32	CO2 forcing' is unclear, suggest 'specified atmospheric CO2 concentration'	Noted: change also made to sentence in

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Reto Knutti]	response to 8-782
8-784	A	54:34	54:34	this section might want to mention the trade-off between lapse rate a water vapour feedbacks. For me this was the striking element of the Colman 2003 paper and earlier work e.g. Shina 1996(?) [Piers Forster]	Taken into account: see comment 8-778.
8-785	A	54:34	57:22	Suggested new section 8.6.3.1 which combines 8.6.3.1 and 8.6.3.1.1 (TSU saved suggested section to ALLENRichard_8_6_3_1.doc Chapter 8 Supplemental Material) [Richard Allan]	Noted: see response to 8-778 above.
8-786	A	54:35	54:41	Are LW and RH defined? Is RH defined with respect to water or water/ice depending on temperature? [Richard Allan]	First part: accepted: expand LW to 'long wave'. 'RH' defined in 8.6.2.2 (8-54 line 7). Second part, no change made: RH is not restricted to being with respect to water in the discussion, and distinction not important at this level of approximation.
8-787	A	54:40		How slightly is the feedback strength reduced? About 5%? [Richard Allan]	It was decided not to include the figure of around 5% -- quoted by Soden and Held (2005) – in the text, as there are uncertainties associated with their method of calculating radiative impacts of water vapour changes, and it was unclear what the uncertainty was associated with this number.
8-788	A	54:43	44:51	Please note that A. Dai (2005: Recent climatology, variability and trends in global surface humidity. J. Climate, accepted with minor revision) shows that RH over the globe has relative small temporal and spatial variability and that the NCAR PCM produces the historical surface q increases associated with recent warming. [Aiguo Dai]	Paper now cited.
8-789	A	55:5	55:23	It seems to me that if the water vapor response is similar across models, and does not depend on such implementation considerations as the vertical resolution, this is because the physical process itself is represented in a way that it will behave in the same way regardless of vertical resolution. the fact that all lthe models get it still probably means that they are all at their core representing this in the same way. Now if they all do it the same way, is this because we really know the process and therefore all include it the same way, or it has few degrees of freedom the way we think of it? It would be much more valuable to state/discuss WHY the models all get the same response rather than inferring	Rejected: the consistency across GCMs regardless of parametrisation or resolution specification is only one piece of evidence supporting confidence in water vapour feedback and only one small part of the assessment as is made clear in the text. The approach taken is to discuss this,

No.	Batch	Page:line From	Page:line To	Comment	Notes
				some truth from the fact that they all DO get the same resonse. Essentially same comment for section 8.6.3.1.2 p. 8-57. [Anne Douglass]	along with other lines of evidence in evaluating confidence in water vapour feedback. Other evidence is assessed at length and includes: (1) degree of confidence in the representation of important physical processes in models, (2) GCM ability to represent different time and spatial scales of observed water vapour variability (3) model response to different types of forcing and (4) GCM consistency with other types of model, such as cloud resolving or mesoscale models.
8-790	A	55:33		Repetition: Chapter 3, page 35, lines 7-9 also describe the Minschwaner and Dessler findings but note that these results may be caused by spurious changes in NCEP temperature used in the calculations. [Richard Allan]	Noted: Reference is now made to Chapter 3 in 8.6.3.1.1. Note that cross chapter issues are now addressed by the addition of Box 8.1 covering UTRH and water vapour feedback across all chapters.
8-791	A	55:47		Also reference Huang et al. 2005 [Huang, X., B. J. Soden, and D. L. Jackson (2005), Interannual co-variability of tropical temperature and humidity: A comparison of model, reanalysis data and satellite observation, Geophys. Res. Lett., 32, L17808, doi:10.1029/2005GL023375.] [Richard Allan]	Reference cited.
8-792	A	55:49	55:49	I would add that the humidity is simulated with +/- 40% accuracy (it follows from Figure 8.3.12) [Eugene Rozanov]	This figure now dropped, due to lack of confidence in reanalyses for this purpose.
8-793	A	55:51		UTH --> UTRH [Richard Allan]	Correction made.
8-794	A	56:39	56:40	It's stated that the range of observed estimates of water vapor feedbacks "covers that of models." But on Fig. 8.6.2, the observed estimates seem to span a smaller range than even the estimates of just HadCM3. [Leo Donner]	Accepted: the reference of "covering the range in models" refers to the range noted in Fig 8.6.1, diagnosed from annual mean water vapour feedbacks in GCMs The text has been changed to make this reference explicit. The range of feedback estimated from HadCM3 is calculated for individual

No.	Batch	Page:line From	Page:line To	Comment	Notes
					months, which shows higher variability.
8-795	A	57:8		reference Cess (1975) here. [Cess, R.D. (1975) Global climate change: an investigation of atmospheric feedback mechanisms, Tellus, 27, 193-198.] [Richard Allan]	Reference cited.
8-796	A	57:28	57:28	Fix spelling of "apparent" [Alan Robock]	Correction made
8-797	A	58:3		change "to understand" to "understanding of" [Richard Allan]	Accepted. Correction made.
8-798	A	58:13		on --> of? [Richard Allan]	Accepted. Correction made.
8-799	A	58:17	58:17	lockstep? [Catherine Senior]	Accepted. "Lockstep" removed.
8-800	A	58:23	58:27	See also Del Genio, Kovari, Yao and Jonas 2005, full citation in previous comment. [Anthony Del Genio]	Accepted. Citation added.
8-801	A	58:46	59:6	There is no mention in the discussion of midlatitude storm clouds that models as a whole overpredict the optical thickness of storm clouds in upwelling regions, as summarized by Zhang et al., 2004 (already cited, but the correct year is 2005) and also by Del Genio, A.D., A.B. Wolf and M.-S. Yao, 2005: Evaluation of regional cloud feedbacks using Single Column Models. J. Geophys. Res., 110, D15S13, doi:10.1029/2004JD005011. An explanation for this in terms of the coarse resolution of climate GCMs and their resulting inability to correctly simulate the strength, and therefore the tilt, of the ageostrophic frontal circulation has been given by Bauer and Del Genio 2005 (full citation given in a previous comment). [Anthony Del Genio]	First part rejected: The overprediction of the optical thickness of clouds (whatever the latitude) is discussed in 8.6.3.2.3 (2nd paragraph). Second part accepted: We added a sentence in this paragraph about the interpretation of this bias for middle latitudes: "for mid-latitudes, these biases have been interpreted as the consequence of the coarse resolution of climate GCMs and their resulting inability to correctly simulate the strength of ageostrophic circulations (Bauer and Del Genio 2006)."
8-802	A	58:46		in --> into [Richard Allan]	Accepted. Change made.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-803	A	59:0	60:	Section 8.6.3.2.3: There could be a reference made to the problems in comparing model cloud with global/tropical observations of cloudiness which, as discussed in Section 3.4.3 and 3.4.4, are at lthe limit of the observational capability. [Richard Allan]	Noted. We add in the 2nd paragraph of 8.6.3.2.3: "(note however that uncertainties remain in the observational determination of the relative amounts of the different cloud types)". On the other hand, we do not refer explicitely to chapter 3 because the observational difficulties discussed in chapter 3 primarily concern the detection of cloud trends.
8-804	A	59:28	59:34	It has not been demonstrated that cloud feedback processes involving tropical clouds are different/more uncertain than extra-tropical clouds. Webb et al. (2005) indicate that a large part of the uncertainty in climate sensitivity is due to differences in cloud feedback from low cloud. It happens to be the case that most of this cloud type occurs in the sub-tropics, hence using the definition of the tropics as 30N-30S, then most of the variance appears to be in this region in figure 8.6.3. However it has not been demonstrated that the processes involved in tropical low cloud feedback are fundamentally different to those occurring elsewhere. If the tropics were defined at 15N-15S then most of the uncertainty would be considered to come from the extra-tropics. [Keith Williams]	Accepted. - For AR4 models, former figure 8.6.3 showed clearly that the spread of model cloud feedbacks is larger in the tropics than in the extratropics (each region representing half of the globe by definition). However, this result may depend on the ensemble of models considered. For instance in the ensemble of models considered in CFMIP and by Webb et al. (2005), the contrast between the tropical and extratropical spreads is weaker (probably because the NCAR, FGOALS and INM AR4 slab models did not participate in CFMIP). - Here is how we changed the text: now (1) we say more explicitely that we discuss the spread of cloud feedbacks *among AR4 models*, (2) we say that "the spread is substantial both in the tropics and in the extratropics, and tends to be larger in the tropics", (3) *we don't show anymore the former figure 8.6.3* (to emphasize less this

No.	Batch	Page:line From	To	Comment	Notes
					point) but refer to Webb et al. 2005 and Bony et al. 2006, where the figure is shown), (4) we highlight the dominant role of low-level clouds in the diversity of cloud feedbacks without emphasizing the tropical or extratropical origin of these clouds.
8-805	A	59:48		Reference Section 8.3.1 [Richard Allan]	Accepted. Reference added.
8-806	A	60:12		At the end of the paragraph, add "Unfortunately, large uncertainties exist in the relative amounts of clouds in different layers as well as their optical properties due to inherent difficulties determining the cloud layers using any passive satellite observations especially for overlapped clouds (Chang and Li 2005a). The latest global cloud statistics obtained from MODIS (Chang and Li 200b) showed much less mid-level clouds and more low-level clouds than those obtained from the ISCCP (Rossow and Schiffer 1999) due to different treatments of overlapped clouds. In comparison with the new MODIS product, problems suffered by GCMs seem to be much less serious in generating mid-level clouds than low-level clouds. More accurate information on cloud vertical structure will be forthcoming from the CloudSat (Stephens et al. 2002)." [Zhanqing Li]	Accepted. We add in the 2nd paragraph of 8.6.3.2.3: "(note however that uncertainties remain in the observational determination of the relative amounts of the different cloud types)". We are not more explicit about the MODIS-ISCCP comparison because of space limitations and because the main point of this paragraph is to point out problems in the simulation of cloud types in general (which would be confirmed -especially for low-level clouds- if models were compared to MODIS instead of ISCCP).
8-807	A	60:31	60:35	Actually, there are now two global observational assessments of the model assumptions used to predict cloud water phase: One from POLDER polarization data (Doutriaux-Boucher, M., and J. Quaas, 2004: Evaluation of cloud thermodynamic phase parameterizations in the LMDZ GCM by using POLDER satellite data. Geophys. Res. Letters, 31, L06126, doi: 10.1029/2003GL019095), and one from MODIS (Naud, C.M., and A.D. Del Genio, 2005: Observational constraints on cloud thermodynamic phase in midlatitude storms. J. Climate, submitted). Both of these agree that the 0-15 deg. C range for transition from liquid to ice assumed in some GCMs is biased warm, and that such models will therefore be biased low in their climate sensitivity. [Anthony Del Genio]	Accepted. We changed "few evaluations of the assumptions used in current models are available" in: "the evaluation of these assumptions is just beginning (Doutriaux-Boucher and Quaas 2004)." We cannot cite Naud and Del Genio (2005) because the paper will not be accepted for publication on

No.	Batch	Page:line From	Page:line To	Comment	Notes
					time. We do not write that "models that use the 0-15 deg range for transition from liquid to ice are biased low in their climate sensitivity" because the sign of the impact on sensitivity is difficult to predict a priori (it depends on how liquid and ice cloud properties are parameterized in models).
8-808	A	61:4	61:8	The assertion that snow-albedo feedback on climate time scales can be evaluated by comparing a model's seasonal snow-albedo feedback with observed seasonal snow-albedo feedback rests on the high correlation between the models' seasonal and climate-scale snow-albedo feedbacks. If this correlation in the models is not realistic-and it's obviously not easy to test-this argument breaks down. [Leo Donner]	Accepted. Text modified.
8-809	A	61:13	61:15	The text cites studies claiming that surface processes are more important than cloud fields in explaining differences in simulations of surface albedo feedback. But, cloud fields can play a large role in determining surface radiation, which is a key surface process. How can the two be separated? [Leo Donner]	Noted. The main point of Qu and Hall (2005) is that the perturbations to the cloud fields in snow-regions associated with climate change do not significantly change the attenuation effect of the atmosphere on surface albedo anomalies, and so therefore are not a major source of divergence in snow albedo feedback. To the extent they have a large effect on radiation fields, they are properly though of as contributing to global cloud feedback. Winton came to the same conclusion independently, by scattering surface albedo feedback strength in AR4 models against the surface component of surface albedo feedback (neglecting cloud effects). He found

No.	Batch	Page:line From	Page:line To	Comment	Notes
					the two are highly-correlated, indicating that intermodel variations in cloud fields or perturbations to cloud fields aren't the principle source of divergence in surface albedo feedback.
8-810	A	61:43		"…suggest that feedbacks.." Too vague - is this just referring to the mechanisms of high latitude positive lapse rate feedback that has long been known (about 30 years at least) ? [Robert E. Dickinson]	Accepted. Text modified.
8-811	A	61:53		Typos [Brian Hoskins]	Accepted. Correction made.
8-812	A	62:12	62:13	To my knowledge, very few studies have demonstrated that simulation of any aspects of present-day climate are 'necessary'. Ones that demonstrate a link such as Bony et al. (2005) and Williams et al (2005) point to diagnostics which appear to be relevant to some aspects of the climate change response. Against this, the use of the phrase 'necessary' appears to be a little strong. I suggest replacing with '...feedbacks (8.6.2), some processes appear relevant (although probably not sufficient) and should be considered...' [Keith Williams]	Accepted. We now write: "some processes appear relevant and should be considered..".
8-813	A	62:25	62:26	Although it is too early to use metrics in this report, I think it would be nice if we could encourage their development in time fo AR5. [Catherine Senior]	Rejected. It is not the role of IPCC assessments (however the discussion will suggest it indirectly).
8-814	A	62:28	67:5	Section 8.7 I still think it is hard to justify this ection, given the discussion in Chapter 10 [Catherine Senior]	Noted. By agreement with the TSU, chapters 6, 8 and 10, this section remains.
8-815	A	62:28		Section 8.7. There is considerable overlap here with chapter 6 and chapter 10 - section 8.7 need not cover everything about abrupt climate change, e.g. the definition. Chapter 6 already deals with how well models simulate past abrupt events, and chapter 10 with the possibility of future abrupt events. Hence, it is not clear what the role of the section here is - the title suggests it might be a closer look at processes and their representation in models. Then, everything could be cut up until page 63 line 42, i.e. the model intercomparison of idealised freshwater experiments. [Stefan Rahmstorf]	Noted. The small cluster breakout on abrupt climate change suggested the current arrangement.
8-816	A	62:34	62:34	(reference?). To be cpmpleted. [Philippe Tulkens]	Agree. Reference added.
8-817	A	62:34		I presume the reference required here is to: R. B Alley and colleagues: 2002: Abrupt Climate Change: Inevitable Surprises. US National Research Council Report, National Academy Press, 230 pp.	Agree. Reference added.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Ian Simmonds]	
8-818	A	63:1	64:13	The title of this section refers to thermohaline circulation, but the text to meridional overturning circulation. One term should be chosen and used consistently. [Lenny Bernstein]	Agree. Text changed.
8-819	A	63:1		On thermohaline circulation changes. I think that a statement of Carl Wunsch on the thermohaline circulation is important ("What Is the Thermohaline Circulation?" Science, Vol 298, Issue 5596, 1179-1181, 8 November 2002). He points out the following: "The conclusion from this and other lines of evidence is that the ocean's mass flux is sustained primarily by the wind, and secondarily by tidal forcing." Since broad readers of Science magazine know his statement, there should be a comment in this section. [Kiminori Itoh]	Agree. Text added.
8-820	A	63:3	63:3	The ice sheets are also part of the physical climate system, and have longer response times [Reto Knutti]	Agree. Text modified.
8-821	A	63:3		Should specify that of the components of the physical climate system that are explicitly modelled in current generation coupled systems, the ocean carries the long time scales. Ice caps, mantle rebound and other aspects of interaction with the solid Earth have yet longer timescales. [Frank Bryan]	Agree. Text modified.
8-822	A	63:8	14:	would be good to update refrences in this short para - these are mostly circa TAR. Coordinate w/ Chap 6 model team. This comment applies to most paleo in chap 8 [Jonathan Overpeck]	Agree. References updated.
8-823	A	63:25	63:31	This discussion treats as quite credible the possibility of the MOC becoming unable to sustain itself with resultant abrupt cooling. It is quite at variance with the behavior of 12 models reported recently at the Aspen Workshop on Abrupt Climate Change (Kerr, 2005, Science, 310, pp. 432-433). In those models, which undoubtedly overlap strong the AR4 models, the MOC weakened by 10% to 50% but did not collapse, and none of the models showed any cooling. [Leo Donner]	Disagree. Next PP discusses the likeihood of extreme climate change due to a shut down.
8-824	A	63:25	63:31	This paragraph is based on the assumption that the MOC is self-sustained, through the positive salinity advection feedback. This should be stated. [Marisa Montoya]	Disagree. The weakening may be forced through heat and/or freshwater fluxes.
8-825	A	63:33	63:40	This discussion does not belong in this chapter, and also seems to be a knee-jerk reaction to a Hollywood film - why discuss a Hollywood desaster movie scenario? As far as I know, no scientist has ever suggested that greenhouse warming could cause an ice age - as witnessed by the fact that the reference given (Joyce and Keigwin) is only to a web page, and this page does not even say that an ice age could be caused. (It does speak about	Disagree. This extreme scenario is widely reported in the press. Some brief discussion is needed here.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				a "little ice age", refer to Chapter 6 if there is any confusion here between LIA and a real ice age.) [Stefan Rahmstorf]	
8-826	A	63:42	63:44	The sentence stating that changes in the MOC are able to produce abrupt climate change requieres a reference. [Marisa Montoya]	Agree. Reference added.
8-827	A	63:44	45:	need a citation for this assertion [Jonathan Overpeck]	Agree. Reference added.
8-828	A	63:44	63:46	Some of the idealized studies should be cited as an example. [Marisa Montoya]	Agree. Reference added.
8-829	A	63:44	63:46	It is important to clarify that the CMIP idealized freshwater simulations reported by Stouffer et al., 2005 use present-day forcings and boundary conditions so a clear comparison to the paleo-record is not possible. [Bette Otto-Bliesner]	Agree. Text added.
8-830	A	63:53	63:55	Does this say: we can currently produce reliable forecasts? That would be wrong. Or does it say: some time in this century we will learn how to do reliable forecasts? [Stefan Rahmstorf]	Agree. Text deleted.
8-831	A	64:8	64:8	I'm not sure that Gregory et al (2003) is relevant here - it is about hysteresis in response to freswater forcing, not about the surface freshwater budget. [Jonathan Gregory]	Agree. Reference deleted.
8-832	A	64:12		When listing idealised model intercomparisons, include Rahmstorf et al. (now in press at GRL - in the ref list as submitted) [Stefan Rahmstorf]	Agree. Reference added.
8-833	A	64:13	64:13	If space permits, it would be good to give some more results from Stouffer et al (2005) on the response to 0.1 Sv forcing. [Jonathan Gregory]	Noted. Space limitation are severe.
8-834	A	64:24		chapters 4 and 6 in addition to chap 10 [Jonathan Overpeck]	Agree. References added.
8-835	A	64:31	64:49	I think this is largely chapter 10 material, and is covered there - it probably doesn't belong in chapter 8, as it's about projections. [Jonathan Gregory]	Agree. Text deleted.
8-836	A	64:32		this study might not be that robust - see paper submitted by Otto-Bliesner et al. Even if this paper isn't published in time, it might be a stretch to say that this process will slow retreat of the GIS significantly. [Jonathan Overpeck]	Agree. Text deleted.
8-837	A	64:38	64:40	The reference of the study referred to should be given.	Noted. Text deleted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Philippe Tulkens]	
8-838	A	64:45	49:	IF you are right (you sound confident), then this should be communicated for inclusion in chap 10! But, it might not be good to be so confident - check with chap 6 model team (esp. Otto-Bliesner and Rahmstorf). Even then, I'd coordinate w/ chap 10 - this is relevent enough to be mentioned in their chap too. Also, as I read this chap 8 section, it seems less like evaluation, and more like projection. [Jonathan Overpeck]	Noted. Text deleted.
8-839	A	64:47		"threshold for major weakening"? I think major weakening, e.g. by 30 or 50%, is not a threshold process, but there is (at least in many models) a threshold for shutdown. [Stefan Rahmstorf]	Noted. Text deleted.
8-840	A	64:55	64:55	Reference to Lunt et al. The year of publication is to be filled in. [Philippe Tulkens]	Agree. Reference added.
8-841	A	65:1	65:7	Weaver et al (2002) suggested that meltwater pulse IA at 14,600 yr BP could have come from Antarctica; by freshening the S Ocean it could have stimulated the Atlantic overturning and caused N Atlantic warming. The relevant mechanism was investigated by Saenko et al 2003. Meltwater Pulse 1A from Antarctica as a Trigger of the Bølling-Allerød Warm Interval Andrew J. Weaver, Oleg A. Saenko, Peter U. Clark, Jerry X. Mitrovica. SCIENCE VOL 299 14 MARCH 2003 1709-1713 author = {Oleg A. Saenko and Andrew J. Weaver and Jonathan M. Gregory}, title = {On the link between the two modes of the ocean thermohaline circulation and the formation of global-scale water masses}, journal = {J. Climate}, year = {2003}, volume = {16}, pages = {2797-2801} [Jonathan Gregory]	Agree. Text added.
8-842	A	65:1		as you discuss the WAIS, I'd check/coordinate with chaps 4, 6 and 10 - each has only part of the story, as does chap 8. Again, it seems like chap 8 is into the projection game? I would go through each section in the entire chapter, and ask the question - "is this model evaluation, or something more akin to climate projection" [Jonathan Overpeck]	Agree. Text modified and references added to point the reader to chapters 4, 6, and 10.
8-843	A	65:8	65:8	I suggest inserting "basal" before "melting", as there will not be much surface meltwater from Antarctica unless there is very large warming. [Jonathan Gregory]	Noted. Text deleted.
8-844	A	65:11	65:11	I'd suggest Shepherd et al (2002) instead of or as well as Thomas et al (2004). Shepherd, A., D.J. Wingham, and J.A.D. Mansley, 2002: Inland thinning of the Amundsen Sea sector, West Antarctica. Geophys. Res. Lett., 29, 1364, 10.1029/2001GL014183. 12 [Jonathan Gregory]	Noted. Text deleted.
8-845	A	65:22	65:27	Large volcanoes with radiative forcings in excess of -3 W/m2 do occur - Pinatubo, Krakatoa, Tambora. Their radiative forcing is clearly greater in magnitude than the	Noted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				accumulated GHG forcing, but such large volcanic eruptions are rare. With an e-folding time of about a year, their radiative effects are effectively gone in about 3 years. The heat capacity of the ocean dictates the magnitude of the actual global temperature response and its recovery to normal. [Andrew Lacis]	
8-846	A	65:22	65:36	For example Section 8.7.2.3 is vague and has very little information in it, notably the 1st and third paragraphs [Catherine Senior]	Noted.
8-847	A	65:29	65:32	It is odd to be thinking in terms of volcanic eruptions counteracting GHG warming over the next century. Citing a "modeling study" that this would be extremetly unlikely hardly adds credence to the conjecture. Volcanic eruptions larger than Pinatubo have occured in the past and could occur at any time in the future - they are a fact of life. The Tambora eruption (1815) had significant impact on global climate for several years. The Laki (1783) eruption in Iceland had severe regional impact. [Andrew Lacis]	Noted. Text deleted.
8-848	A	65:30	65:32	As the "modelling study of a super volcano" is mentioned here, could the reference to the study be included please:- GS Jones, JM Gregory, PA Stott, SFB Tett, RB Thorpe, "An AOGCM simulation of the climate response to a volcanic super-eruption", Climate Dynamics, 2005, doi:10.1007/s00382-005-0066-8 [Gareth S. Jones]	Noted. Text deleted.
8-849	A	65:30	65:32	Need to include reference here to Gareth Jones et al. (2005),and also include the caveat that this experiment assumed a 100 times Pinatubo aerosol loading, but did not consider the more realistic assumptions that the aerosol loading would last longer than that for Pinatubo nor land surface/vegetation feedbacks, both of which would make the climate response larger and longer lasting. [Although I am not sure what this has to do with the main subject of the AR4]. [Alan Robock]	Noted. Text deleted.
8-850	A	65:32	65:32	Jones et al. (2005) show this. author={G. S. Jones and J. M. Gregory and P. A. Stott and S. F. B. Tett and R. B. Thorpe}, title = {An {AOGCM} simulation of the climate response to a volcanic super-eruption}, journal = {Clim. Dyn.}, year = {2005}, pages = {doi:10.1007/s00382-005-0066-8} [Jonathan Gregory]	Noted. Text deleted.
8-851	A	65:32		"stadial period"? These are phases during glacial times. [Stefan Rahmstorf]	Noted. Text deleted.
8-852	A	65:34	65:36	They may be similar, but they are also different. GHGs through their greenhouse effect impact thermal radiation throughout the atmosphere. Volcanic aerosols are localized in the stratosphere. They primarily scatter (reflect) solar radiation, but also have a	Noted. Text modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				substantial thermal greenhouse component that acts to heat the stratosphere. The size of the volcanic aerosol is important. If the volcanic aerosols were larger than 2.2 microns effective radius (typical size is 0.5 microns), the greenhouse effect of the volcanic aerosol will exceed the albedo effect, and the volcanic aerosol will warm rather than cool the ground surface (Lacis and Mishchenko, 1995). [Andrew Lacis]	
8-853	A	65:34	65:36	However, the volcanic aerosol produces winter warming over the land and it is heavily depends on the model ability to reasonably simulate PNJ (e.g., Rozanov et al., 2002). I guess, there is no such limitation for GHG increase [Eugene Rozanov]	Noted. Text modified to be more specific to abrupt climate changes.
8-854	A	65:34		This statemnet is patently false. It is much more difficult to simulate the response to a near point-source of aerosols than to changes in well mixied greehouse gases. A different set of processes, that are in general poorly represented or excluded all together, are involved. [Frank Bryan]	Agree. Text modified to be more specific to abrupt climate changes.
8-855	A	65:38	53:	Case in point - is this section about model evaluation, or just highlighting an issue. Seems it would be appropriate to reveal how well we simulate clathrate and permafrost dynamics - should we have confidence in simulations of these processes? [Jonathan Overpeck]	Noted. Highlighting a process. Text modified.
8-856	A	65:38	65:38	Number 8.7.2.3 is wrong, should be 8.7.2.4 [Eugene Rozanov]	Agree. Corrected.
8-857	A	66:1	31:	where is the evaluation in this section? I'm sure there is much to be had, but it isn't in this section. Should it be here or in another chapter (6 and 7?). If so, you should summarize here, and perhaps enhance in some way [Jonathan Overpeck]	Agree. Text modified to reflect a process evaluation.
8-858	A	66:15	66:20	Multiple-equilibria literature reviewed here all predates AR4. More current papers reviewed in Chapter 7. How are these 2 sections connected? [Robert E. Dickinson]	Taken into account in revision
8-859	A	66:28		Same as comment number 15. This statement is false. [Frank Bryan]	Agree. The paragraph is deleted.
8-860	A	66:33		Section 8.7.3: I propose to cut it - a very nice and intriguing model finding, worthy of further study, but currently lacking policy relevance. What should policy makers conclude from the fact that in one model, in one 15,000 year long experiment, a couple of weird cooling events happened? We refrained from discussing many interesting paleo-modelling and data results in chapter 6 if they were not clearly policy-relevant. [Stefan Rahmstorf]	Disagree. The section points out the role of natural variability in climate change detection. If a large event such as the one found in these model integrations were to occur in the future, it would make attribution very difficult. That said, the subsection has

No.	Batch	Page:line From	Page:line To	Comment	Notes
					been greatly shortened.
8-861	A	66:35	40:	Does this discussion jive with the rest of the chapters? Need batter coordination on abrupt change for sure. [Jonathan Overpeck]	Noted. Several links to other chapters have been added.
8-862	A	66:42	49:	where is the evaluation in this para? We see things in models, but are they realistic? Please help. The entire section is not that helpful. What the reader wants to know is whether models can simulate abrupt climate change in a realistic manner. You point out that long simulations/obs series are needed, so I guess that means that we can only use paleo? Better cite chap 6 in any case, but perhaps there is more chap 8 can add? [Jonathan Overpeck]	Noted. The paragraph is deleted and links to other chapters have been added.
8-863	A	67:7	69:56	this section is very important but it's too short, deserve more attention and needs more link with chapter 6 and also with other sections of this chapter; otherwise the authors may also to decide to eliminate totally this section making an agreement with the author of chapter 6 placing EMIC in the right evidence. [vincenzo artale]	Noted. Space limitation is very severe. Links with other Chapters are now more explicit.
8-864	A	67:11		Section 8.8.1 could be shortened. The main points could be made in far fewer words to bring it more in line with the rest of the introductory text in this chapters vrious sections. [Keith Dixon]	Taken into account. The first para of Section 8.8.1 has been shortened. Authors believe that the text in the other paras is justified and very useful for other Chapters.
8-865	A	67:17	67:17	I wonder if Fichefet et at 2003 is the best reference in this case for providing an overview of GCMs. I did not read that paper, however I thought that concerned research team worked with one GCM. Isn't there a reference available that gives a broad overview on GCMs and their components ? [Philippe Tulkens]	Accepted. We now refer to Section 8.2.
8-866	A	67:28		"Simple climate models contain modules…" This is just one simple climate model, others are different from what is described here. [Stefan Rahmstorf]	Accepetd. Text modified.
8-867	A	67:43	67:48	Perhaps the Hansen et al. (1997) sector GCM (the "Wonderland" climate model) might be included in the category of "simpler" EMIC models. The Wonderland model is 120 degree sector GCM with idealized topography that uses the same model physics as the full GISS GCM, but runs 3 time faster. With this model Hansen et al. performed and analyzed a series of radiative forcing and climate response experiments that included doubled CO2, solar consant changes, ozone, cloud, aerosol, CFC, and surface albedo variations. [Andrew Lacis]	Rejected. Table 8.8.2 focuses on EMICs that are used in Chapter 10 (see caption).
8-868	A	67:44		replace "proposed" with "developed" - these models exist.	Accepted. Text modified.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Stefan Rahmstorf]	
8-869	A	67:49	68:11	Any flux adjustments in EMICs needs to be stated. [Bette Otto-Bliesner]	Accepted. Information on flux adjustments in EMICs is now included in Table 8.8.2.
8-964	B	68:49	69:6	This is a highly biassed description on EMICs. One could be under the impression that Claussen invented EMICs. As in section 8.8.2, some reference to the historical development would be in order. [Thomas Stocker]	Accepted. We now mention that the references (some have been modified) provide the reader with reviews on simple climate models and EMICs.
8-870	A	69:25		EMICS have a smaller climate sensitivity than GCMs?? This is just chance, the sample shown is far too small to make this a general conclusion or a systematic difference. [Stefan Rahmstorf]	Accepted. Statement removed.
8-871	A	74:0		References to be added to Chapter 8: [Zhanqing Li]	Noted
8-872	A	74:8		Chang, F.-L., ad Z, Li, 2005a: A new method for detection of cirrus overlapping water clouds and determination of their optical properties, J. Atmos. Sci., 62, 3993–4009. Chang, F.-L., and Z. Li, 2005b, A near-global climatology of single-layer and overlapped clouds and their optical properties retrieved from Terra/MODIS data using a new algorithm, J. Climate, 18, 4752-4771. Rossow, W. B., and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP, Bull. Amer. Meteor. Soc., 80, 2261-2287. Stephens, G. L., and coauthors, 2002: The CLOUDSAT mission and the A-Train, Bull. Amer. Meteor. Soc., 83, 1771-1790. [Zhanqing Li]	Taken into account in making reference list consistent with citations.
8-873	A	75:51	75:52	Many of the starting letters should be low-case in the article title. [Aiguo Dai]	Accepted
8-874	A	84:24	84:29	The reference to the Johns et al. (2004) technote report, as documentation of the HadGEM1 coupled model, will be superseded by the following papers, subject to them being accepted by the journals: T. C. Johns, C. F. Durman, H. T. Banks, M. J. Roberts, A. J. McLaren, J. K. Ridley, C. A. Senior, K. D. Williams, A. Jones, G. J. Rickard, S. Cusack, W. J. Ingram, M. Crucifix, D. M. H. Sexton, M. M. Joshi, B-W. Dong, H. Spencer, R. S. R. Hill, J. M. Gregory, A. B. Keen, A. K. Pardaens, J. A. Lowe, A. Bodas-Salcedo, S. Stark, and Y. Searl: The new Hadley Centre climate model HadGEM1: Evaluation of coupled simulations. J. Climate (submitted); and A. J. McLaren, H. T. Banks, C. F. Durman, J. M. Gregory, T. C. Johns, A. B. Keen, J. K. Ridley, M. J. Roberts, W. H. Lipscomb, W. M. Connolley, and S. W. Laxon: Evaluation of the sea ice simulation in a new coupled atmosphere-ocean climate model (HadGEM1), J. Geophys.	Accepted. Only the Johns et al paper will be cited (McLaren et al still under review.

No.	Batch	Page:line From	To	Comment	Notes
				Res. (Oceans) (submitted). [Timothy Johns]	
8-875	A	84:24	84:29	The reference to the technical report should be replaced be the paper (Johns et al., 2005) [Keith Williams]	Accepted.
8-876	A	87:38	87:39	Lin, J.-L. et al. (2005, J. Climate) has been accepted. [Leo Donner]	Noted
8-877	A	89:30	89:39	Both of the two "Martin et al." papers referred to are now accepted by J Climate I understand. Note, however, that the second of the two is superseded by the following reference, which has a different first author: M. A. Ringer, G. M. Martin, C. Z. Greeves, T. J. Hinton, P. M. James, V. D. Pope, A. A. Scaife, R. A. Stratton, P. M. Inness, J. M. Slingo, and G.-Y. Yang: The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model, HadGEM1. Part 2: Aspects of variability and regional climate, J Climate (accepted). Any references in the text to the second "Martin et al. (2005)" paper should therefore be replaced by "Ringer et al. (2005)". [Timothy Johns]	Accepted.
8-878	A	89:30	89:33	The reference to the technical report should be replaced be the paper (Martin et al., 2005) which also appears in the bibliography as the next entry. [Keith Williams]	Accepted
8-879	A	89:37	89:39	This reference should be replaced by: Ringer, M. A., G. M. Martin, C. Z. Greeves, T. J. Hinton, P. M. James, V. D. Pope, A. A. Scaife, R. A. Stratton, P. M. Inness, J. M. Slingo, and G.-Y. Yang, 2005. The physical properties of the atmosphere in the new Hadley Centre Global Environmental Model, HadGEM1. Part II: Aspects of variability and regional climate. (Submitted to J. Climate) [Gill Martin]	Accepted.
8-880	A	92:7	92:8	Updated Reference: Norris, J. R., and S. F. Iacobellis, 2005: North Pacific cloud feedbacks inferred from synoptic-scale dynamic and thermodynamic relationships. J. Climate, 18, 4862–4878. [Joel Norris]	Accepted
8-881	A	98:20	98:23	Update this reference to: Sokolov, A.P., C.A. Schlosser, S. Dutkiewicz, S. Paltsev, D.W. Kicklighter, H.D. Jacoby, R.G. Prinn, C.E. Forest, J. Reilly, C. Wang, B. Felzer, M.C. Sarofim, J. Scott, P.H. Stone, J.M. Melillo & J. Cohen, 2005: The MIT Integrated Global System Model (IGSM) Version 2: Model Description and Baseline Evaluation, MIT JP Report 124. http://web.mit.edu/globalchange/www/MITJPSPGC_Rpt124.pdf [Ronald Prinn]	Accepted. Reference updated.
8-882	A	103:8	103:9	Typing errors in Williamson (1968) [Alejandro Bodas-Salcedo]	Accepted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
8-883	A	106:0		Table 8.2.1. Botton line on page. Primary reference to the ocean component of the CSIRO-Mk3.0 model is the Gordon et al., 2002 reference, not "Packanowski 1996". [Anthony Hirst]	Accepted.
8-884	A	106:0		Table 8.2.1. Botton line on page. Primary reference to the sea ice component of the CSIRO-Mk3.0 model is O'Farrell 1998. Suggest leave that as sole reference. If Semtner 1976 is to be cited for the ice thermodynamics, then it is essential that Flato-Hibler 1992 also be cited to specify the ice dynamics [Anthony Hirst]	Accepted.
8-885	A	106:10	106:10	The reference for the ocean component of the BCC-CM1 coupled model should be: Jin Xiangze?Xuehong Zhang, and Tianjun Zhou?1999?Fundamental framework and experiments of the Third Generation of IAP/LASG World Ocean General Circulation Model?Advances in Atmospheric Sciences?16?197-215. [Tianjun ZHOU]	Accepted.
8-886	A	107:0	107:	Table 8.2.1: entries for GISS-EH and GISS-ER models are in error: The atmospheric model top in both cases are 0.1 hPa, and the number of levels in the atmospheric code is 20. The ocean for GISS-ER is at a 4x5 resolution, with 13 levels. GISS-EH uses HYCOM for the ocean (16 levels and 2x2 resolution). The sea ice in both cases is the same (reference BOTH Liu et al 2003 and Schmidt et al 2004). [Gavin Schmidt]	Accepted.
8-887	A	107:0	107:	For both GFDL models (CM2.0 and CM2.1) the "?" after "leads" in the "Sea Ice" column should be removed. Both models have leads. [Michael Winton]	Accepted.
8-888	A	107:0	107:	The reference for the ocean component of the FGOALS-g1.0 coupled model should include: Jin X.Z.?X.H. Zhang? T. J. Zhou? 1999: Fundamental Framework and Experiments of the Third Generation of IAP/LASG World Ocean General Circulation Model? Adv. Atmos. Sci.? 16?197-215. Liu? H.L.? X.H. Zhang? W. Li? Y.Q. Yu? and R.C. Yu? 2004: An Eddy-Permitting Oceanic General Circulation Model and Its Preliminary Evaluation? Adv. Atmos. Sci.? 21? 675-690. [Tianjun ZHOU]	Accepted.
8-889	A	107:0	107:	The land and sea ice component model of FGOALS-g1.0 are exactly the same as those of NCAR CCSM2 except for the horizontal resolution, hence the corresponding references should be: Bonan? G. B.? K. W. Oleson? M. Vertenstein? S. Levis? X. Zeng? Y. Dai? R. E. Dickinson? and Z.-L. Yang? 2002: The land surface climatology of the Community Land Model coupled to the NCAR Community Climate Model. J. Climate? 15? 3123-3149. Briegleb? B. P.? C. M. Bitz?E. C. Hunke? W. H. Lipscomb? M. M. Holland? J. L.	Accepted.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				Schramm? and R. E. Moritz? 2004: Scientific description of the sea ice component in the Community Climate System Model? Version Three. NCAR Tech. Note NCARTN-463+STR? 70 pp. [Tianjun ZHOU]	
8-890	A	108:0		Table 8.2.1: some features of the IPSL-CM4 model should be modified. Atmospheric top: 4 hPa The reference for the atmospheric model should be Hourdin et al. 2005. Hourdin, F., I. Musat, S. Bony, P. Braconnot, F. Codron, J.-L. Dufresne, L. Fairhead, M.-A. Filiberti, P. Friedlingstein, J.-Y. Grandpeix, G. Krinner, P. LeVan, Z.-X. Li and F. Lott. The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, submitted to climate dynamics, 2005. Ocean resolution: L31 [Jean-Louis Dufresne]	Accepted.
8-891	A	110:0	111:	Table 8.8.2. Flux adjustments used in EMICs should be noted where appropriate. [Bette Otto-Bliesner]	Accepted. Information on flux adjustments in EMICs is now included in Table 8.8.2.
8-892	A	110:0		Table 8.8.2: For the MIT-IGSM2, the references for the land biosphere (BT, BV) have been omitted whereas they are reported for the other models. The reference for both BT and BV is B. Felzer, J. Reilly, J. Melillo, D. Kicklighter, Q. Zhuang, C. Wang, R. Prinn and M. Sarofim, Global and future implications of ozone on net primary production and carbon sequestration using a biogeochemical model, Climatic Change, 73, 345-373, 2005. [Ronald Prinn]	Accepted. Reference added.
8-893	A	110:0		Table 8.8.2: For the MIT model details line please change MIT-IGSM2 to MIT-IGSM2.3. Also for the MIT ocean biosphere (BO) replace McKinley et al, 2004 with Parekh et al, 2005. The new reference details are: P. Parekh, M.J. Follows and E. Boyle, Decoupling of iron and phosphate in the global ocean, Global Biogeochem. Cycles, 19, doi: 10.1029/2004GB002280, 2005. [Ronald Prinn]	Accepted. Acronym and reference updated.
8-894	A	110:1	110:6	Table 8.8.2. It is good to see that the IPCC is now recognizing the value of EMICs. One of the values of EMICs is that they can include complexities that are not possible with the AOGCMs when applied to multiple century to millenium-scale integrations. However, this Table downgrades the importance of atmospheric chemistry (gases and aerosols) by mentioning them only as "CHEM" and giving no references for them. Chemistry demands at least the same attention as the other headings in this table. For the MIT model, for example, there are even separate interactive submodels for urban and large-scale	Accepted. Reference added.

No.	Batch	Page:line From	Page:line To	Comment	Notes
				atmospheric chemistry. I understand that chemistry was surprisingly not included in the exercise in Chapter 10, but that is no excuse to not properly report here the important elements of each EMIC. The chemistry in the MIT model is described in M. Mayer, C. Wang, M. Webster and R. Prinn, Linking local air pollution to global chemistry and climate, J. Geophys. Res., 105, 22869-22896, 2000. [Ronald Prinn]	
8-895	A	110:4	110:4	Table 8.8.2: The reference for the CLIMBER-3? model is missing. This is: Montoya, M., A. Griesel, A. Levermann, J. Mignot, M. Hofmann, A. Ganopolski, and S. Rahmstorf. The Earth System Model of Intermediate Complexity CLIMBER-3?. Part I: description and performance for present day conditions. Climate Dynamics, 25, (2005), 237-263. [Marisa Montoya]	Accepted. Reference added.
8-896	A	114:0		FIG. 8.2.1. To be fair, these figure only shows that there is more structure at high resolution and not that the simulation is improved at high resolution. Could a panel showing corresponding observations be included? [Ileana Bladé]	Accepted.
8-897	A	114:0		Figure 8.2.1 This is an incredibly complex and detailed figure to choose as the first figure in a chapter on Climate Models. Why not choose something that represents the comparison between observed and modelled climate? It would make more logical sense to start this chapter with Figures 8.3.1, 8.3.2 and 8.3.3. [Melanie Fitzpatrick]	Rejected. 8.2 and 8.3 represent different approaches to evaluation (see 8.1), and we do not feel there is a clear case for 8.3 to precede 8.2
8-898	A	114:0		Give references [Vincent Gray]	Accepted.
8-899	A	114:0		Fig. 8.2.1 Label each graph, e.g., (a), (b) and (c); refer to each separately in caption. Add labels and units to X and Y axes of each graph. Add label to color bar. [Melinda Marquis]	See response to 8-902
8-900	A	115:0		Give references [Vincent Gray]	Accepted.
8-901	A	115:0		Fig. 8.2.2 Add labels and units to X and Y axes, including inset graphs. [Melinda Marquis]	Rejected (see 8-902).
8-902	A	116:0		Fig. 8.3.1 Add labels and units to X and Y axes and to color bars. [Melinda Marquis]	Rejected. On plots of this kind both the labels ("Latitude" and "Longitude"), as well as the units ("degrees north" and "degrees east") are often left off the plot so that more area can be devoted to displaying the information of real interest. The outline of the continents

No.	Batch	Page:line From	Page:line To	Comment	Notes
					make it clear that this is Earth and the coordinate values make it obvious what the longitude and latitude units are. As for the color bar, the information (units, quantity plotted) is given in the figure caption, and is not included as part of the figure proper, which is again common practice.
8-903	A	117:0		Fig. 8.3.2 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	See response to 8-902.
8-904	A	118:0		Figure 8.3.3 At least some of the models listed in the key (e.g. HadCM3, HadGEM1) don't appear to be shown in the figure yet - I hope they can be added. [Timothy Johns]	Accepted. Figures will be updated.
8-905	A	118:0		Fig. 8.3.3 Add label and units to X axis. [Melinda Marquis]	See response to 8-902.
8-906	A	119:0		Fig. 8.3.4 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Accepted. Y access to be labeled; see response to comment 8-902 and 8-908.
8-907	A	120:0		Figure 8.3.5 There is no discussion of this Figure in the text. One drawback of this plot is that it compares models with only one satellite data set. The satellite data retrievals (particularly for polar regions) are known to be fairly poor inthe SW - suggest that this be mentioned or the plot be deleted. [Melanie Fitzpatrick]	Taken into account. Plot moved to supplementary material.
8-908	A	120:0		Fig. 8.3.5 Add label and units to X axis. [Melinda Marquis]	Rejected. The figure caption and the x-axis coordinate values (90N, EQ, 90S) already make it obvious that this is latitude. The label "Latitude" is omitted to conserve space.
8-909	A	121:0		Fig. 8.3.6 Add label and units to X axis. [Melinda Marquis]	Rejected. see response to 8-908.
8-910	A	122:0		FIG. 8.3.7 is missing the legend for the ERBE observed curve [Ileana Bladé]	Accepted.
8-911	A	122:0		Fig. 8.3.7 Add label and units to X axis. [Melinda Marquis]	Rejected. see response to 8-908.
8-912	A	123:0		Figure 8.3.8, 8.3.14 and 8.3.16: the number of models shown in Figure 8.3.8 can be extended. Indeed, the total energy transport can be estimate for all the models for which the TOA fluxes are available, i.e. for all the models shown in figure 8.3.7. For the same	Accepted. All figures to be updated with the latest data in the archive.

No.	Batch	Page:line From	To	Comment	Notes
				reasons, all the models on figures 8.3.13 and 8.3.15 should be present on figures 8.3.14 and 8.3.16. [Jean-Louis Dufresne]	
8-913	A	123:0		Fig. 8.3.8 Add units to X axis. Clarify Y axis label and units. Define PW. [Melinda Marquis]	Accepted.
8-914	A	124:0		Fig. 8.3.9 Add label and units to X axis. [Melinda Marquis]	Rejected. See response to 8-908.
8-915	A	125:0		Figure 8.3.10. Panel b overlaps with the colorbar of panel a. [Alejandro Bodas-Salcedo]	Accepted.
8-916	A	125:0		Fig.8.3.10 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Rejected. see response to comment 8-902
8-917	A	125:10	125:12	Fig. 8.3.10: Need to use common period such as 1979-1999. Need to state what models are shown here (e.g. is it a model-ensemble mean?). Errors in data periods. [Aiguo Dai]	Accepted. Caption rewritten. Also see response to 8-49.
8-918	A	125:12	125:12	19801999 should read 1980-1999. [Philippe Tulkens]	Accepted.
8-919	A	126:0		Fig. 8.3.11 Add label and units to X axis. [Melinda Marquis]	Rejected. See response to 8-908.
8-920	A	126:1	126:8	Since discussion of Fig. 8.3.11 is limited to the tropics (page 27, lines 45-53), why not modify the x-axis to show only the low latitudes? To do so would help one see through the blizzard of lines an make it easier to identify individual model results. [Keith Dixon]	Rejected. There is no great value here in actually identifying individual models, since we are discussing a systematic error in the models. The higher latitude information also shows that inter-model differences are smaller there, even without specific discussion.
8-921	A	127:0		Figure 8.3.12. What are the units on the key? (percent error?) In the caption, "mean fractional error" is here more accurately written "mean percentage error". [Anthony Hirst]	Accepted.
8-922	A	127:0		Fig. 8.3.12 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Accepted. Y access to be labeled; see response to comment 8-902 and 8-908 for further response.
8-923	A	128:0		Fig. 8.3.13 Add label and units to X axis. [Melinda Marquis]	Rejected. see response to 8-908.
8-924	A	129:0		Give references [Vincent Gray]	Accepted.
8-925	A	129:0		Fig. 8.3.14 Add units to X axis. Clarify Y axis label and units. Define PW.	Accepted. (except for X axis units; see

No.	Batch	Page:line From	Page:line To	Comment	Notes
				[Melinda Marquis]	response to 8-908).
8-926	A	130:0		Fig. 8.3.15 Add label and units to X axis. [Melinda Marquis]	Rejected. see response to 8-908.
8-927	A	131:0		Give references [Vincent Gray]	Rejected. There are no observations shown in this figure.
8-928	A	131:0		Fig. 8.3.16 Add units to X and Y axes. [Melinda Marquis]	Rejected. units already appear on Y-axis. On x-axis, they're obvious (see response to 8-908).
8-929	A	132:0		Fig. 8.3.17 Add label and units to X axis. [Melinda Marquis]	Rejected. see response to 8-908.
8-930	A	133:0		Add "Error" to ordinate label. [Leo Donner]	Accepted.
8-931	A	133:0		Fig. 8.3.18 Add label and units to X axis. [Melinda Marquis]	Rejected. see response to 8-908.
8-932	A	134:0		Fig.8.3.19 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Rejected. see response to comment 8-902
8-933	A	134:0		Fig. 8.3.19. This is quite similar to Fig. 8.3.1. Maybe delete. [Kevin Walsh]	Accepted.
8-934	A	135:0		Add "Error" to ordinate label. [Leo Donner]	Accepted.
8-935	A	135:0		Fig. 8.3.20 Add label and units to X axis. [Melinda Marquis]	Rejected. see response to 8-908.
8-936	A	135:0		Figs 8.3.20 and 8.3.21. Is it my imagination, or are these figures inconsistent with each other? Fig. 8.3.20 shows that the ocean models have a fresh bias in the Arctic, whereas Fig. 8.3.21 shows that they have a salty bias, unless the fresh bias near Murmansk utterly dominates. Text also says that the bias is salty. [Kevin Walsh]	Accepted. This is being carefully checked.
8-937	A	136:0		Fig. 8.3.21 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Rejected. see response to comment 8-902
8-938	A	137:0		Fig.8.3.22 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Accepted. Y access to be labeled; see response to comment 8-902 and 8-908 for further response.
8-939	A	138:0		Fig.8.3.23 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Accepted. Y access to be labeled; see response to comment 8-902 and 8-908 for further response.

No.	Batch	Page:line From	To	Comment	Notes
8-940	A	139:0		Give references [Vincent Gray]	Rejected. No observations are shown.
8-941	A	139:0		Fig.8.3.24 Add labels and units to X and Y axes and to color bar. [Melinda Marquis]	Accepted. Y access to be labeled; see response to comment 8-902 and 8-908 for further response.
8-942	A	139:5		Supplementary figure of streamfunction: the FGOALS model has the opposite sign in the wind driven cells near the equator compared to all other models and observations. Is something wrong in this panel or is the model so badly wrong? [Reto Knutti]	Accepted. This is being carefully checked.
8-943	A	140:0		Fig. 8.3.25 Add label and units to color bar. [Melinda Marquis]	Accepted. Figure modified.
8-944	A	141:0		Fig.8.3.26 Add label and units to color bar. [Melinda Marquis]	Accepted. Fifure modified.
8-945	A	142:0		Fig.8.3.27 Add units to Y axis. [Melinda Marquis]	Noted.
8-946	A	143:0		Give references [Vincent Gray]	Noted.
8-947	A	143:0		Fig.8.3.28 Add labels to X and Y axes. [Melinda Marquis]	Noted.
8-948	A	144:0		FIG 8.3.29: TAS not defined [Ileana Bladé]	Accepted.
8-949	A	144:0		Fig.8.3.29Clarify labels of X and Y axes, e.g, s.d. of what … [Melinda Marquis]	Taken into account. This is clarified in the revised version of the text.
8-950	A	144:1	145:9	Use of similar format for both Taylor diagrams recommended, instead of normalized standard deviation for Fig. 8.3.29 and standard deviation for Fig. 8.3.30. To accommodate a wider readership, provide an interpreation of the Taylor diagrams, especially of the green curves if the format of Fig. 8.3.30 is used. [Leo Donner]	Accepted. Green contours indicating "skill" have been removed. One Taylor diagram deleted.
8-951	A	144:9	144:9	860 hPa should probably be 850 hPa [Leo Donner]	Accepted.
8-952	A	145:0		FIG 8.3.30: the green lines represent a measure of skill, i'm guessing, but the caption does not indicate so [Ileana Bladé]	Accepted. Green contours indicating "skill" have been removed.
8-953	A	145:0		Fig. 8.3.20 Clarify labels of X and Y axes, e.g, s.d. of what … [Melinda Marquis]	Taken into account. This is clarified in the revised version of the text.
8-954	A	146:0		FIG. 8.4.1 Lines in top panel are too faint. Also iI'm confused by the legend. Simulations	Accepted. Green contours indicating

No.	Batch	Page:line From	Page:line To	Comment	Notes
				starting in different seasons were combined ? [Ileana Bladé]	"skill" have been removed. Text revised to clarify.
8-955	A	146:0		Give references [Vincent Gray]	Accepted.
8-956	A	149:0		Fig.8.6.3 Add Y axis labels and units to each graph. [Melinda Marquis]	This figure has been removed.
8-957	A	149:0		Connected with comment 6, figure 8.6.3 does not add much value to the chapter as the result is an artifact of the choice of boundary between the tropics and extra-tropics. I believe a more useful figure would be one demonstrating that low cloud feedback is a major source of uncertainty. (Figure 8.6.4 provides circumstantial evidence that this is the case, however another figure demonstrating the link directly would be of value. Possibly a figure from Webb et al., 2005, using the ISCCP simulator diagnostics in the feedback classes would be suitable). [Keith Williams]	Noted. (former) Figure 8.6.3 has been removed, and it will not be replaced because of space limitation (the draft has to be shortened). The fact that low-cloud feedback is a major source of uncertainty is already illustrated by Figure 8.6.4 (now Figure 8.6.3) from the paper Bony and Dufresne, GRL, 2005: "Marine boundary-layer clouds at the heart of tropical cloud feedback uncertainties in climate models".)
8-958	A	151:0		Give references [Vincent Gray]	Rejected. Reference already given
8-959	A	151:0		Fig.8.6.5 Add units to X and Y axes. Try to shorten this lengthy caption by putting info in caption into text of chapter. [Melinda Marquis]	Accepted. Units have been added and the captions shortened.
8-960	A	152:0		Fig.8.8.1 Each graph needs to have X and Y axis labels and units. [Melinda Marquis]	Accepted. Figure redrawn.
8-961	A	153:0	153:	the legend of figure 8.8.2 is wrong [vincenzo artale]	Figure removed because of space limitation.
8-962	A	153:0		Figure 8.8.2. Update GCM results (gray crosses) with new AR4 model results. [Bette Otto-Bliesner]	Figure removed because of space limitation.
8-963	A	154:0		FIG. 8.1 . "for the instrumental record climate in response" … ? Not very smooth wording. "Globally averaged surface temperature, and as simulated in climate models in response to ….. Also "the red line" (not "the multiple model lines" and "for the observed trends" (not "the simulated trends"). [Ileana Bladé]	Accepted: wording changed.