The peer reviewers for 3.1 are as below:

Dr. Kerry Cook
Professor, Atmospheric Science
Director of Graduate Studies, Atmospheric Sciences
Department of Earth and Atmospheric Sciences
3114 Snee Hall
Cornell University
Ithaca NY 14853

Tel: (607) 255 9716
Fax: (607) 254 4780

Dr. Carlos R. Mechoso
University of California, Los Angeles
Department of Atmospheric and Oceanic Sciences
7127 Math Sciences Building
405 Hilgard Avenue
Los Angeles, California 90095-1565

Tel: (310) 825 3057
Fax: (310) 206 5219

Dr. Gerald Meehl
Research Scientist
Climate and Global Dynamics Division
National Center for Atmospheric Research
PO Box 3000
Boulder Colorado 80307
USA
Tel: (303) 497 1331
Fax: (303) 497 1333

Dr. Philip Mote
State Climatologist
CSES Climate Impacts Group, Box 354235
University of Washington
Seattle WA 98195

Tel: (206) 616 5346

Fax: (206) 616 5775

Dr Brad Udall
Director
CU-NOAA Western Water Assessment
NOAA Earth System Research Laboratory
325 Broadway, R/PSD
Boulder, CO 80305-3328

Tel: (303) 497 4573
Fax: (303) 497 6449

Dr. John Walsh
President's Professor of Climate Change & Chief Scientist
IARC 405, International Arctic Research Center
930 Koyukuk Drive
P.O. Box 757340
Fairbanks, Alaska 99775-7340

Tel: (907) 474 2677
Fax: (907) 474 5662

REVIEWER 1:

Pg. 5, line 9, that *models*

Pg. 5., lines 7-11.
Climate Model Construction, first paragraph. The second and third sentences in this paragraph seem to me to gloss over a basic issue about the development of climate models. These sentences suggest that climate model development has been driven by the need for applications related to people, and this really isn't the case – historically. For example, there has been very little attention paid to "storminess" over the decades of GCM development since the models can't resolve storms. GCMs were interesting to construct because they taught us how the climate system works – their development has much more of a "basic science" motivation than this paragraph indicates. More recently, of course, development is being driven by the global warming problem.

Pg. 5, lines 14-18. To say that a "good" climate model "must" accomplish these feats is not correct. Plenty of climate models – probably all of them, depending on the strictness of your measures of success – fail at most of these tasks.

Pg. 5., line 20. A more complete definition of "climate", carefully distinguishing it from weather, would be good here. People often need to understand how we can claim to have some skill in predicting climate change in 100 years or more when we can't produce very skillful seasonal or interannual forecasts.

Pg. 6, lines 1-2 and 7. The tone of the piece is uneven. These lines give an example, with lines 1-2 seeming to be directed to other scientists (maybe a physicist, for example) and line 7 sounding like a middle-school text book (particularly "by scientists").

Pg. 7, line 7. Suggested rewording: … and the physical laws that govern the exchanges of mass and energy

Pg. 7, line 13, *the* primitive equations … with *the* hydrostatic

Pg. 9, lines 4 – 5: Cumulus *convection* … *is* …

pg. 11, line 13, referred to *as* a …

pg. 12, line, 9: delete *separate*

pg. 17, line 26: define and use "OGCM", similar to AGCM

pg. 28, lines 1-21. As I think is implied by an author's parenthetic note, these paragraphs are out of place and/or redundant with material at the beginning. There are some fresh thoughts here, though, they should not be lost.

Pg. 28, lines 24-31. I like this paragraph and think it conveys something useful about how GCMers work and think.

Pg. 47, lines, 5 and 9, and numerous other spots: Regional climate model applications should not be uniformly referred to as "downscaling" simulations, since this gives the impression that their only use is to provide more detail in conjunction with GCM simulations (as fancy interpolators, for example). This is one use of regional models, but they are also used to simulate climate and climate change independently of GCMs as well – in present day, future, and paleoclimate applications.

Pg. 50, line 21-23. RCMs have been successfully run without convection parameterizations with grid spacing on the order of 5 km. This is noted on pg 51, lines 12-13, contradicting the pg 50, line 21-23 statement.

Pg. 79, lines 1-2. messed up text

General comments:

1. While the detailed comparison of the U.S. models is useful to provide depth, singling out 3 U.S. models (modeling groups, more accurately) does not give an accurate overview of the ability of GCMs in general. The U.S. models are all terrific and absolutely state-of-the-art, but GCM modeling is a global endeavor at this point – requiring international collaboration - and that could be more strongly related in the text.

2. It would be useful to strengthen the discussion throughout of how observations support modeling activity, with specific examples. For example, a statement that satellite observations are essential for validating models and understanding climate processes in remote land areas and over the oceans. Also, as global models evolve to finer grid spacings, and currently for regional model simulations, information about structure at the surface (soil moisture, vegetation, soil temperatures) is going to become increasing important – either for constraining the models or for validating. This will require both land-based and space-based observing systems. Another possible example concerns observations of the large-scale ocean circulation, e.g., the THC. How accurately is it known, including the features of its natural variability? In general, how does a scarcity of observations map onto model development and improvement?

REVIEWER 2:

1. Will this report be useful to its readers?

I think that the report will be very useful to the readers. It covers most of the major topics and provides a comprehensive list of past successes, talks about issues of current concern, and reviews future directions.

As in most cases with this type of reports, there is some uncertainty on the targeted audience. The recently published Encyclopedia of the Atmospheric Sciences (Holton and Curry, Editors) includes several papers on climate modeling for the general readership. If one builds on this foundation, then one could go deeper on the technical aspects of climate modeling both scientifically and technically. The text falls in the middle ground, and mentions many intercomparisons of results and little critical contrast of modeling approaches by different groups.

The report focuses primarily on the physical climate models that were used for the most recent international Coupled Model Intercomparison Project's (CMIP) coordinated experiments (Meehl, et al., 2006), sponsored by the World Climate Research Programme (WCRP). Nevertheless, several other models are mentioned along the text. In my opinion, this limitation is too restrictive as discussed in my answer to question 6.

2. Is the charge clearly described in the report? Are the aspects of the charge as outlined in the prospectus for 3.1 fully addressed? Do the authors go beyond their charge or their expertise?

The authors in the team are highly qualified to speak authoritatively about climate models, their uses, limitations, sensitivity, feedbacks, and uncertainties.

3. Are the report's exposition and organization effective? Have the authors effectively communicated the material?

I do not have elements to give a final opinion on this matter. What I have in front of me is a long text in a very bland format, similar to the one used for a scientifically oriented readership. A good technical editor can make the text much more readable by highlighting key sentences and inserting attractive figures. The process will lead to an evaluation of the report's, as outstanding aspects become isolated from the large amount of information available.

In the same way, the text consistency can be greatly improved. The text refers to a different number of models in different parts. Also, the models selected for discussion are referred to in different ways: American models, leading models, US AOGCMs…One can recognize pieces of other reports in the overall text. This roughness can be easily smoothed out.

4. Are the conclusions adequately supported by evidence, analysis, and argument? How well is uncertainty recognized and discussed? Does the report effectively recognize and communicate the strengths and weaknesses of current climate models?

I find many statements without an adequate reference. It is understood that if all references are included in a text dealing with so many diverse topics, the list may become longer than the paper itself. Nevertheless, in some cases the need of a reference is obvious. (For example, the statement of the anti-correlation between rainfall over the Sahel and the Amazon requires a reference.) There is a very nice discussion on climate sensitivity. The list of strengths and weaknesses of current models is long and there is always room for one more. My answer to question 6 mentions other possible candidates.

5. Does the executive summary concisely and accurately describe the key findings and recommendations? Is it consistent with other sections of the report?

My version does not have an executive summary.

6. What other significant improvements, if any, might be made in the report?

As I indicated in the answer to question 1, the report focuses primarily on climate models that were used for the most recent international Coupled Model Intercomparison Project's (CMIP). Thus, the emphasis is mostly on the coupled atmosphere-ocean system. The decision to restrict the scope of the report is a good one, and the way it is done is justified at the present time. I am sure that the authors are expecting some challenges to their decision, so here they go in the form of questions.

a. What are the fundamental differences in the modeling approaches of the different institutions? Are some efforts more innovative than others? Can we get the feeling of an integrated national approach to climate modeling and simulation?

b. Why aren't any university lead efforts mentioned? It is acknowledged that university groups played a leading role in the development of climate models. What is happening nowadays? Are there universities producing new modeling paradigms, and if not, why not? (I think they are!) Where is instruction on climate modeling happening?

c. Why isn't there a section on the stratosphere? The Antarctic Ozone Hole is a success story since science motivated an international agreement. The role of climate models in this problem has not been, to my knowledge, properly discussed. Obviously the model could not predict the feature due to the lack of the proper chemistry. The problem is not completely gone; can climate models help to understand why?

d. The access by users of a computational infrastructure to run large codes can be briefly reviewed. This is, of course, in the understanding that work with GCMs is not confined to the large national laboratories. Even if this were the case, are national laboratories satisfied with their computer facilities?

e. The efforts lead by NASA and NCAR to create a software infrastructure to facilitate the use of climate models can be mentioned. The Earth System Modeling Framework (ESMF) promises to enhance the use of climate models.

Chapter I. Introduction

I do not have any comments of note on this chapter. The text gives a feeling for a more or less monotonic improvement of models from the point of view of science, in a way that increased complexity results almost exclusively from increased computer power. I believe that mentioning just one of the milestones in climate modeling (e.g., the Phillips two-layer experiment) will enhance the reader's appreciation of the science issues.

Chapter II. Description of Global Climate System Models

Atmospheric general circulation models

The descriptions in this chapter are authoritative. However, they fall in a middle ground that is of little use to either the general reader and the specialist. The clearest example is in the paragraph on cumulus convection. The expert reader will learn little from the description of the schemes. For the non-expert, a mention of the quasi-equilibrium assumption without a minimum context will be meaningless.

Why is it that the majority of AGCMs use variants of the Arakawa-Schubert parameterization? This magnificent accomplishment is almost 30 year old; what has happened in the meantime? Is this an issue that ought to be brought up in this review?

Ocean general circulation models

The text refers to the relatively coarse horizontal resolution of the models. It is indicated that "eddy scales" are parameterized. I think it is important to clarify that these scales do not correspond to the turbulent eddies that are parameterized in AGCMs. Contemporary OGCMs do not resolve mesoscale eddies, which can be originated by baroclinic-barotropic instabilities of ocean currents. These can play an important role in closing the ocean mixed layer budget by providing shoreward heat and material transport that balance the upwelling supply of cold water and the air-sea heat exchange. There are also standing eddies associated with alongshore coastline and bathymetric irregularities. The difficulties in closing the budgets may be key in many places, such as the eastern part of the tropical oceans.

Evaluation of AGCMs and OGCMs

This section is one paragraph long and is not balanced with the others in the report. Perhaps it could be merged with the longer discussion in Chapter V.

Land Surface Models

This is a straifgtforward description of the different aspects of land surface modeling.

I find intriguing that "PILPS has lead to a better agreement among land models". Is it implied that the models were basically the same except for "tunable" parameters? The statement that "The latest generation of land surface models exhibit relatively smaller differences compared to previous generations" reinforces this impression. Are there major differences between land surface models?

Sea Ice Models, including parameterizations and evaluation

The two dominant paradigms in sea ice modeling are discussed here practically side by side. This is a useful strategy.

Component coupling and coupled model simulations

This section includes the development plans at the 3 US groups that contributed to the 4th Assessment of the IPCC. It is good to find the plans in one place, but is unclear whether this compilation adds to the information already on the institutions web sites.

Reductive vs holistic evaluation of models

This section is very different from the others in terms of scope and style. The speculative style seems to be at odds with the matter-of-fact style in the remainder of the text. I gather that the concepts to be transmitted are three. First, ensemble simulations must be performed in order to consider the spread and characteristics of variability of the individual realizations. Second, our "confidence in its explanatory and predictive power of climate models grows based on their ability to simulate many aspects of the climate system simultaneously with the same set of physically based rules." Third, one cannot "tune" the model for one region of the world since all regions are simulated. Perhaps this can be done very efficiently in a few sentences.

Chapter III – The added values of regional climate simulations

Types of downscaling simulations

I liked this section; it brings up many of the concerns on the topic and that are not easy to find in a single source.

There are a couple of spots that I found to be rough. In reference to the different performance of parameterizations in global and regional models, it is stated, "This factor is part of a larger issue, that parameterizations may have regime dependence, performing better for some conditions than others" (page 51). I can understand dependence on grid size, but I am not quite clear on different physical regimes for the same grid size. Or, doesn't the difference sensitivity of parameterizations to physical process have an impact in all grid sizes and the impact becomes exaggerated for some grid sizes?

More on this subject in the comments to Chapter VI.

Chapter IV– Model Climate Sensitivity

This is a very important chapter, and I believe that the job was well done.

Chapter V – Model simulation of major climate features

There is a lot of information in this chapter and I will be selective on a few matters that caught my attention.

Mean climate

This section has a long paragraph on the “double ITCZ” problem without a single reference. This is an important problem of high relevance to climate simulation and prediction. The links to the model difficulties with ENSO are evident.

The last paragraph of the section is that “AOGCMs generally simulate large-scale mean climate with considerable accuracy, but the models are not reliable for aspects of mean climate in some regions, especially precipitation.” The last paragraph at the end of a long section will attracts a lot of attention from the readers, and requires more elaboration and an attempt to synthesis.

Monsoons

A more current view describes monsoons as involving both atmosphere and oceans. The presentation here is more traditional and looks at the atmosphere as reacting to changes in different time scales.

A reference is needed on the processes that limit the extent of the monsoons. Are the authors referring to the ventilation paradigm?

It seems to me that one basic problem in monsoon simulation is not addressed, and that this problem poses serious questions on whether the climate model monsoons are proxies of reality. Monsoons comprise processes at the planetary, continental and meso scales. Among the latter are the “low-level jets”. These differ in the monsoons: 1) The Somali Jet, which flows in summer at all times, 2) the South American Low Level Jet, which flows along the lee of the Andes during the entire year, and 3) the Great Plains Low Level Jet, which flows in North America at night during the warm season. These mesoscale features are captured poorly by global GCMs. The associated problem is that water advection is underestimated. If simulated precipitation is realistic, then local processes such as evaporation have to be exaggerated. Consistently, the role of soil processes may be over-emphasized. Can monsoon projections be trusted in view of these uncertainties in the water budget?

Another problem that is attracting a lot of attention is the GCM difficulties in simulating the diurnal cycle and its variability in monsoon regions. It has become clear that the peak precipitation amplitude is too early in the day. This feature is likely associated with the PBL parameterization, which receives little attention in the text.

Monsoon researchers have recognized that tropical cyclones contribute significantly to precipitation, primarily in the North American monsoon. AGCMs mentioned in the report cannot resolve such features, but others are claiming that they do it to some extent. Any opinions on this?

Polar Climates

Add “in the polar regions” before the reference to Uotila et al. (2007).

Please clarify in which way “stable boundary layers remain an important area for model improvement.” There is little discussion of PBL in the report and this may be a place where this limitation can be at least partially addressed.

The well-know problem of the “cold lower stratosphere of GCMs” receives little or no attention. This affects the zonal wind and planetary wave behavior, and hence low-frequency variability. The paper by Pawson et al (BAMS 2000) addresses this problem.

Please clarify what is meant by “Because both the northern and southern polar regions are within circumpolar atmospheric circulations, their synoptic coupling with other regions is more limited than is the case with midlatitude regions embedded in the westerlies”.

Modes of variability

I believe that the section on El Niño-Southern Oscillation (ENSO) must be adjusted a little since it appears to be originally intended to discuss many more (15) models than the ones selected for this report. We read a very important statement: “We find that even among the models with the most realistic simulation of ENSO *and* seasonal variability there is no consensus on the anticipated change in climate within the tropical Pacific.” (Presumably the realistic ENSOs are those obtained for current climate conditions.) This is difficult to justify by inspection of just the three models that selected at the beginning.

I am unclear on the argument about the upwelling “dilution” associated with the coarse grid spacing of CGCMs. According to the argument, the dilution limits the amplitude of resulting ocean temperature fluctuations. A current hypothesis is that the most important aspect of coarse horizontal resolution is the inability to resolve mesoscale ocean eddies that result from the baroclinic-barotropic instability along the upwelling front. The mesoscale eddies transport cold and fresh water off shore, thus extending the effect of upwelling in a scale far larger than the grid size.

In regard to climate prediction, what is the relative skill of physical models based on coupled GCMs in relation to simpler dynamical models and to statistical models? This is an important issue, although the report aims to time scales longer than the interannual.

Extreme events

"Extreme events" here refers to largest simulated values of precipitation or surface temperature. The limitations of current GCMs in this area are so clear (i.e. inability to simulate tropical cyclones, at least in climate simulations) that the text can be trimmed to emphasize issues of consensus on the information provided. Several researchers have already evaluated model performance in the context of "extreme events". I would like to see a discussion on the usefulness of these studies in reference to 1) climate science, 2) model performance.

This is actually attempted in parts of the text. For example, the issue that thunderstorms are responsible for many intense events is raised and could be discussed further. It is mentioned that this is related to the parameterization of convection, which is only one aspect of the problem.

Chapter VI – Future Model Development

The description of CRMs and their future reads well. Maybe it gives the impression that these models are more ready for climate studies than they actually are. Higher resolution changes the parameterization problem, but it doesn't make it necessarily easier. The full effect of increased model resolution provided by CRMs will be experienced when comparable or even higher resolution is also available in the boundary conditions, which must be provided by other models.

There is another paradigm for multiscale problems that will be likely attempted in the next decade. This is the nesting of coupled regional models of the atmosphere and the ocean within global coupled GCMs. The difficulties in nesting regional and global models are discussed in Chapter III of the report. However, some of those difficulties may be reduced in regions that are key to the climate system, and yet interactions with other regions at the synoptic scale are not intense. I am referring to the eastern part of the tropical oceans, where coupled GCMs fail with the stratocumulus and their radiative effects. It seems to me that, in the near future, there will be a strong interest in coupling regional models of the atmosphere-ocean system. Some of the work has started, but the full potential of the approach remains to be evaluated.

REVIEWER 3:

The version I printed out to review did not have page numbers or a list of authors. To facilitate the review, I've numbered the pages, such that "Chapter 1 Introduction" is page 1.

General comment

The report is a quite thorough overview of the state of current climate modeling. In fact there is probably a bit too much text book type material that could be trimmed. There is some duplication of material (e.g. ENSO is described in two different places). With the exception of the section that had implications for ethical practices by modeling groups that included speculation and hearsay that modeling groups essentially cheat by tuning equilibrium climate sensitivity, I found the draft to be an otherwise high quality and comprehensive review of climate modeling.

Specific comments:

P. 7, top: There is a statement here that "typical AGCMs have spatial resolution of 200 kilometers in the horizontal and 20 levels…" This certainly isn't "typical" of current AGCMs used in coupled climate models. The current crop is closer to 150 km with about 30 levels (more details are given on the PCMDI CMIP web site).

P. 8, top: Perhaps it could be mentioned that very recent work on CRMs will be covered later in the report.

P. 10, bottom: It is stated here that these ocean models have resolution of "about 1/3 of a degree at the equator". However, this is not a complete portrayal, and it should be mentioned that usually these models have increasing resolution in the equatorial tropics usually between about 5N and 5S

P. 18, middle: When talking about the "bucket", it should be mentioned that the so-called bucket is actually meant to represent a physical quantity, namely field capacity of the soil

P. 26, bottom: This should read "IPCC Fourth Assessment Report (AR4)"

P. 27, bottom: This discussion is at best inaccurate in implying that modeling groups "engineer" a particular value of climate sensitivity. The particularly regrettable sentence is: "Especially if one is willing to compromise on some measures of fitness, one can control the models' sensitivity to some extent (ref to Hadley center)". No reputable modeling group I am aware of does this. In fact, this discussion of modeling groups that "hold various views on the most likely value of climate sensitivity, but rarely with much conviction [sic]" is outdated given the analysis of equilibrium climate sensitivity in Ch. 10 of the IPCC AR4 where a best estimate of actual climate sensitivity is 3.0C, with a likely range of 2.0 to 4.5C. Modeling groups end up with climate sensitivity of their model at the end of their model development process. To imply that somehow groups tune their climate sensitivity at the outset is inaccurate, and, to the best of my knowledge, is simply not true. This falls into the category of speculation and is not appropriate in a CCSP report. In fact, what is implied on p. 27 (that modeling groups "cheat") is directly refuted by description of an actual model development process on P. 29 where indeed climate sensitivity was an outcome of model development, not an a priori goal. I suggest the authors avoid speculation on model developers' ethics, and stick to a discussion of the facts regarding current assessment of climate sensitivity as given, for example, in the IPCC AR4.

P. 30, near bottom: The authors use the term "transient climate sensitivity". This is incorrect. The actual term in common usage (see the TAR and the AR4) is "transient climate response", or TCR.

P. 32-37: This section, titled "reductive vs. holistic evaluation of models" sits uneasily in this report. I suggest it be revised to reflect current usage of terminology and common practice. What is actually described here, more or less, are model sensitivity experiments, and this term is commonly used in the field. In fact, this section is overly long and could be reduced by at least a factor of two. A simple discussion of the methodology of sensitivity experiments where various factors are altered in systematic ways to assess model response and the role of physical processes could be summarized in a page or two. A lot of this arcane discussion complicates a fairly simple procedure commonly used to study processes and responses in climate models. Also, value judgements such as "hidden behind the surface of this seemingly unremarkable time series is the profound imprint of these variations on economies and societies, in this case especially the stark human suffering associated with the drought period in the 70s and 80s" strikes me as inappropriate in a CCSP report on climate science. Additionally, the discussion at the bottom of p. 36 confuses model evaluation with model analysis that occurs after model development.

P. 41, section titled "idealized climate simulations": I suggest this section be revised such that it reflects a more appropriate title for this section, namely "climate model response metrics". Then a simple discussion of the two main metrics, namely equilibrium climate sensitivity and transient climate response, would then follow.

P. 42, near bottom: The authors have used a word here that is not common usage to my knowledge: "paleocalibrate". Such inventions should be discouraged in a CCSP report that reflects current practice and terminology, unless use of this word has slipped past me. If it is in common usage, perhaps the authors could provide a few substantiating references. In fact, this section should stress that paleoclimate simulations are an important part of model evaluation, since it is a severe test of a climate model to be able to simulate a past climate accurately.

P. 43, near top: In a discussion of "numerical downscaling", the method of statistical downscaling should be mentioned

P. 43, about half way down: It is stated here that RCMs “require lateral boundary conditions from observations”. However, the application being discussed here is when RCMs are embedded in AOGCMs, in which case the RCMs require lateral boundary conditions from the global model in which they are embedded

P. 44, top: An important recent modeling study with a global 20 km model should be described here: Oouchi, K., J. Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda, 2006: Tropical cyclone climatology in a global-warming climate as simulated in a 20km-mesh global atmospheric model: Frequency and wind intensity analyses. J. Met. Soc. Japan, 84, 259-276.

P. 52: Somewhere here the authors should discuss the prospects and obstacles involved with two-way nesting with an RCM embedded in an AOGCM, with the AOGCM forcing the RCM, and the RCM giving information back to the AOGCM, and so on.

P. 56, near top: The authors err here in not using current terminology. They use the term “equilibrium warming”, but the TAR and AR4, in assessing the literature on this topic for the past 10 years or so, use the term “equilibrium climate sensitivity”. This CCSP report should be consistent with that usage and change “equilibrium warming” to “equilibrium climate sensitivity” everywhere in this report.

p. 56, middle: The authors need to include the recent assessment of equilibrium climate sensitivity from the IPCC AR4, Ch. 10.

P. 59-60-61-62: See comment immediately above; this entire discussion needs to be updated and replaced given the AR4 assessment of equilibrium climate sensitivity following multiple lines of evidence from a host of models and observational studies (see AR4, Ch. 10)

P. 65 and elsewhere: This draft perpetuates terminology that the WGCM is trying to correct, and I suggest the authors follow their request to call the multi-model dataset at PCMDI the “CMIP3 multi-model dataset assessed in the IPCC AR4”, in place of the “AR4 coupled models” (here and throughout the report)

P. 66, top: Somewhere here the authors should discuss results from a major project to assess cloud forcing called CFMIP (e.g. Webb M.J., C.A. Senior, D.M.H. Sexton, W.J. Ingram, K.D. Williams, M.A. Ringer, B.J. McAvaney, R. Colman, B.J. Soden, R. Gudgel, T. Knutson, S. Emori, T. Ogura, Y. Tsushima, N. Andronova, B. Li, I. Musat, S. Bony and K.E. Taylor, 2006: On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Clim. Dyn.* **27** (1): 17-38 doi:10.1007/s00382-006-0111-2.)

P. 66, middle: The authors should recognize that black carbon aerosols can also absorb solar radiation

P. 72, near bottom: Here is yet another variation, “PCMDI/AR4 simulations”; they should be called the “CMIP3 multi-model dataset assessed in the IPCC AR4”

P. 80: Since this is a U.S. CCSP report, it is odd that the authors use a somewhat dated result from the U.K here. Perhaps a more appropriate figure would be a more recent one from a U.S. model: Meehl, G. A., W. M. Washington, C. Amman, J. M. Arblaster, T. M. L. Wigley, and C. Tebaldi, 2004: Combinations of natural and anthropogenic forcings and 20th century climate. *Journal of Climate,* **17**, 3721–3727; see Fig. 2d.

P. 81, top: The authors should cite the attribution results by region shown in the IPCC AR4 ch. 9 which is the most recent and complete assessment of this topic

P. 86: Here is the first ENSO discussion…

P. 97: The section on monsoons includes way too much text book material, in my opinion. The authors should concentrate the text on the topic at hand, climate model simulations. This would shorten this section by at least a factor of two or more.

P. 99: Another significant result is a projected increase in monsoon interannual variability (Hu, Z.-Z., M. Latif, E. Roeckner, and L. Bengtsson, 2000: Intensified Asian summer monsoon and its variability in a coupled model forced by increasing greenhouse gas concentrations. Geophys. Res. Lett., 27, 2681-2684.; Räisänen, J., 2002: CO2-induced changes in interannual temperature and precipitation variability in 19 CMIP2 experiments. J. Clim., 15, 2395-2411.; Meehl, G.A., and J.M. Arblaster, 2003: Mechanisms for projected future changes in south Asian monsoon precipitation. Clim. Dyn., 21, 659-675.)

P. 102: A recent study has attributed the intensification of the SAM mostly to decreases in stratospheric ozone, with some contributions from increasing GHGs (Arblaster J. M., and G. A. Meehl, 2006: Contribution of various external forcings to trends in the Southern Annular Mode, *Journal of Climate,* **19**, 2896–2905.)

P. 104, top: In the discussion of melting permafrost, the relevant key reference should be Lawrence and Slater (2005). Oddly, this reference is given later on this page in regards to vegetation changes, though its key and most important result pertains to permafrost melting.

P. 112: …and here is the second ENSO discussion

The “El Nino-like” response to increasing CO2 was first identified by Meehl and Washington in 1996 (Meehl, G. A., and W. M. Washington, 1996: El Niño-like climate change in a model with increased atmospheric CO2 concentrations. Nature, 382, 56–60). This type of response has subsequently been addressed by, for example, by Cubasch, U., G.A. Meehl, G.J. Boer, R.J. Stouffer, M. Dix, A. Noda, C.A. Senior, S. Raper, and K.S. Yap, 2001: Projections of future climate change. In: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the

Intergovernmental Panel on Climate Change [J.T. Houghton, et al. (eds.)]. Cambridge University Press, Cambridge, pp. 525-582;
Collins, M., and The CMIP Modelling Groups, 2005: El Niño- or La Niña-like climate change? Clim. Dyn., 24, 89-104; and
Yamaguchi, K., and A. Noda, 2006: Global warming patterns over the North Pacific: ENSO versus AO. J. Met. Soc. Japan, 84, 221-241.

In the IPCC AR4, Ch. 10, there is a figure (10.6) relating El Nino-like vs. La Nina-like response across a number of the current models. It is instructive that the majority show an El Nino-like response to increasing CO2.

P. 114, table: it should be noted that the spacing of grid points is at the equator, since the models typically increase the resolution starting near 5N and 5S to maximum resolution at the equator

Near P. 117-118: This is very much in the nature of a text book discussion and could be trimmed quite a bit

P. 126: Some papers on extremes have model evaluation of either climatology of extremes (e.g. heat wave intensity over North America and Europe in Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent and longer lasting heat waves in the 21st century. *Science,* **305**, 994–997), or observed frost day trends over the U.S. (e.g. Meehl, G. A., C. Tebaldi, and D. Nychka, 2004: Changes in frost days in simulations of 21st century climate. *Climate Dynamics,* **23**, 495–511. doi: 10.1007/s00382-004-0442-9).

P. 132: It should be noted that in all models studied so far, inclusion of carbon cycle introduces a positive feedback to the system from increase CO2 (i.e. Friedlingstein, P., et al., 2006: Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison. J. Clim., 19, 3337-3353.)

P. 140: use "end-to-end" example from ACPI Pilot Project described in the special issue of *Climatic Change*, vol. **62**

REVIEWER 4:

Overall comments.
The report needs considerable additional polishing to make it readable. Figures are improperly labeled and captioned, placeholder comments pepper the document, and the document needs careful proofreading and copy-editing. In addition, the readability -- particularly the use of disciplinary jargon -- of the document is uneven. The section on ocean modeling is rife with words that would be unfamiliar to a generally well-educated audience, like enthalpy. Another point concerns the organization of the report. Portions of Chapter 2 would fit better in Chapter 1; Chapter 7 is just a snippet, too short to be a chapter. Navigation would be much easier with a table of contents and numbered sections and subsections.

Specific comments.
Pg 2 line 2: “The way” - surely not the only way? One way? Analyzing observations is another way.
Pg 2, line 26: list the three here
Pg 4 line 9: what about Manabe and Weatherald 1967?
Pg 4 line 21: wasn’t USGCRP established 1990?
Pg 5 line 3: discuss this figure or don’t include it
Pg 6 lines 2-3: use of infinitives getting repetitive
Pg 6 line 7: “This” antecedent unclear. Compromise?
Pg 6 line 10: chapters, not sections. This paragraph would be unnecessary if the report had a table of contents
(which would be more useful)
Pg 7 line 21: explain “primitive” to the lay reader
Pg 9, line 4: “convections” not usually plural
Pg 9, line 13: “initiation condition” - explain
Pg 9 line 18: “habitat” - explain
Pg 9 line 24-25 “layer *within* tens of meters of the surface”
Pg 11 line 11 “between” should be “among”
Pg 12 line 11 GFDL
Pg 12 line 26-27 could delete rest of sentence after “category”
Page 13 lots of jargon on this page - potential enthalpy, isopycnal, diapycnal. Either explain these or don’t use them.
Page 14 define the abbreviations in this table
Page 15 Figure may be too complicated to explain properly for a lay audience, but if included it needs lots more explanation - what is potential vorticity? σ? B?
Page 16 lines 1-16 - too much detail for this setting. Trim substantially and say why it matters.
Page 17 lines 1-5: this is a good link between numerics and results. Need this kind of statement every time the text gets detailed about numerics (e.g. previous page)
Page 17 lines 21-22 already mentioned the straits - move previous passage to here or vice versa
Page 19 line 16 how can water drain away from the model? should it read “from the grid point”

Page 19 line 23 “between” should be “among”
Page 19 line 27, is the objective “agreement among land models” or an improvement in realism?
Page 20 lines 11-13 the second clause repeats the first without adding anything.
Page 21 line 6 define sensible and latent heat fluxes
Page 21 lines 8-16 mention how important snow is for the albedo feedback; lines 18-30 mention applications of streamflow simulations
Page 23 line 9 - semantic point, the programmer constrains the behavior of the model. Observations cannot - they only provide a reality check.
Page 23 line 11 delete “relatively” (smaller already is a comparative); differences relative to what, among models or from obs? line 12 replace “compared to” with “than”
Page 23 line 28 what does “focus processes” mean?
Page 24 line 16 define rheology for the lay reader.
Page 25 line 5 explain “solves for the ice stress tensor” and “implicit iterative approach”
Page 25 lines 26-28 use consistent terms for ice categories or explain differences between ice categories and ice-thickness categories
Page 26 line 10 “open water” - should that be “fraction of area represented by open water” or something along those lines?
Page 27 lines 9ff - early in this section, perhaps in the third paragraph, say something about the technical aspects of coupling - time steps, fluxes, coupling, platforms, flux adjustments.
Page 28 lines 26-31 include refs to, and discussion of, Stainforth et al., Forest et al., Knutti et al. on climate sensitivity. Much of the material in this paragraph is repeated a bit later and the two parts should be combined.
Also, discuss tuning here - goal is to improve simulation of present climate, not to produce a certain value of sensitivity.
Page 28 line 31, word choice “confrontation with observations”
Page 29 line 12-13 explain flux adjustments
Page 29 line 26 don’t need quotes around hills.
Page 30 lines 4-6 need a verb
Page 30 line 13 sp maintaining, define “sufficient strength”
Page 30 line 21 - “modifications” - direclty or as a consequence of pbl scheme?
Page 30 lines 29-30 sentence fragment
Page 31 lines 1-11 are these separate paragraphs? should be formatted so, if so.
Page 33 line 1 why?
Page 33 line 31 to page 34 line 1 - important point - is that true for other models too?
Page 34 line 15 - a model with a top at 10mb can’t really be said to have a stratospheric circulation
Page 34 line 23 phenomena is plural, should be phenomenon
Page 35 line 2 define TOA
Page 35 line 16 - define “this period”
Page 36 line 27 - this section isn’t really about evaluation but on models dynamical vs empirical
Page 37 line 5 - Number the figures please.
Page 40 line 1 - ref for “It has been suggested”
Page 42 fig needs a caption

Pages 45-46 this could go in chapter 1
Page 52 line 6 - multiple sources of future climate - be explicit, are they scenarios from GCMs
Page 56 line 16 - RCMs do not "correct biases" but merely change
Page 56 line 18 - need some articles in this sentence
Page 59 line 13 - Graeme Stephens had a great review paper in Journal of Climate (2005?) on this
Page 60 line 27 - "to" should be "of"
Page 64 - very similar to pg 35. Mention also the constraints on climate sensitivity calculated using observations.
PP 65-66 this paragraph is incongruously colloquial and should be rewritten.
Page 66 second paragraph, mention also other constraints on sensitivity calculated from observations (Hegerl et al. 2006, Knutti et al. 2002, 2006).
Page 67 line 15 explain a little more how PRP is calculated, just a sentence as is done for CRF.
Page 67 line 19 not "radiative transfer" (passage through the atmosphere) but more accurately radiation
Page 69 line 2 "in a climate change"?
Page 69 second paragraph, I wonder if an equation would clarify the difference between adding up feedbacks and the total sensitivity.
Page 69 line 19 could delete the needless jargon "with super-parameterization"
Page 69 line 26 resolution is singular; line 27, delete "amount of" without loss of meaning
Page 70 line 5 "accurately"
Page 71 line 9, insert a paragraph break here - the LBL calculation deserves its own paragraph. line 11, rather than stating in line 23 that clouds are excluded, frame more positively here "since the goal is to compute the radiative effects of greenhouse gases alone, the calculation is performed for clear-sky conditions"
Page 71 line 21 replace "compared to" with "and"
Page 73 line 10 replace "compared to" with "than"
Page 76 line 2 stray 't'
Page 77 line 7 - "freedom to choose" could be replaced by more scientific language, like "despite the fact that surface temperatures are calculated from physical principles" or something like that
Page 78 line 10 variations on all timescales (dominated by seasonal) or after seasonal removed?
Page 79 line 1 needs attention
Page 80 need more information in caption
Pages 81-82 how about precipitation and temperature from the same model. Page 82, left column, is that supposed to be "CM2.0 minus observed"?
Page 83 line 11 needs attention
Page 84 update with a figure from AR4
Page 84 line 17 it's not merely (or even mostly) the uncertainty about initial conditions as the chaotic nature of the climate system
Page 85 figure needs a caption, and the numbers in the upper right of each panel need explanation.

Page 86 Annular modes might belong better under section C on page 101. Explain "annular" for a general reader.
Page 87 lines 6-8 this sentence is unclear - why would comparison by itself lead to improvement?
Page 87 line 25 missing "of" before heat? and "the" at the end of the line
Page 87 lines 28-29 this is debatable - see Seager et al. 2001 (IJOC I think)
Page 88 line 4 ref. for this suggestion. IPCC TAR and AR4 simulations downplay the impacts of thermohaline shutdown outside of the waters of the North Atlantic
Page 90-91 this ENSO discussion should be combined with that on pp 116ff
Pages 94-100 be more selective about figures in this section. Some don't seem to be mentioned in the text (though it's a little hard to match them up since they're not numbered.) Explain the difference between the figure on page 94 and that on page 81.
Page 101 lines 19ff this is a welcome and interesting diversion.
Page 106 line 22 quantify "skill"
Page 107 line 17ff turbulent fluxes may need explanation for the general reader
Page 108 line 12 awkward
Pages 109-110 the figure needs some explanation or should be omitted.
Page 111 lines 4ff - this section could be moved to chapter 2 where model components are discussed
Page 113 line 2 could add the same panel for NH. What are the years used to form the figure, and forcing used in the model simulations?
Page 114 I suggest discussing ENSO first.
Page 114 line 29 Explain "US IPCC"
Page 115 line 27 - unclear what it means to transfer a simulation to another model.
Page 117 could include a figure for one of these points.
Page 118 line 21 clarify that these are internal waves not surface waves (right?)
Page 143 read the caption carefully...
Pages 144-145 this "chapter" is quite short, and should either be beefed up or eliminated. Since the title of the report is "strengths and limitations for user applications", I would recommend greatly expanding this chapter and truly addressing the issue of user applications, beyond "additional uncertainty is introduced" (p 144 line 29)

REVIEWER 5:

Reviewer's Personal Background

First, to put my comments in perspective some background is in order. I am the Director of the Western Water Assessment (WWA), one of the NOAA Regional Integrated Sciences and Assessments programs. While I am on the University of Colorado research faculty through an appointment with the NOAA Cooperative Instiitute for Research in the Environmental Sciences, my office is in the NOAA Earth System Research Laboratory. I am an engineer by training and my background in climate science has been acquired while being embedded with climate scientists of all persuasions over the last four years. Although I have read literally hundreds of journal articles and numerous books on various aspects of climate during this time, I am not a climate scientist. WWA works closely with major water providers in the Rocky Mountain West such as Denver Water, water provider to approximately one quarter of the state of Colorado, the Bureau of Reclamation, the operator of the major facilities on the Colorado River and various water conservancy districts. For many of our stakeholders over the past year climate change has been transformed from a theoretical construct to an issue requiring planning efforts *now*. Many of our stakeholders want information about climate models. One of our fundamental roles is to act as a boundary organization between scientists and decision makers and hence this CCSP SAP is very important.

Given this background, I am in most cases not proficient to comment on factual aspects of climate modeling as presented in the document. This review will thus be limited to very high level question of does the Draft meet the requirements of the Prospectus.

Two portions of the Prospectus deserve highlighting and will be discussed below. The Prospectus specifically states: *"The topics addressed by this Climate Change Science Program (CCSP) product are the strengths and limitations of climate models at different spatial and temporal scales. Its purpose is to provide this information on the strengths and limitations of the results from climate models,* ***in ways that will allow the potential user of the information to evaluate how best it may be applied or not applied*** *(CCSP Strategic Plan, page 19)."* (Emphasis mine.)

It also states: ***The intended audiences of this CCSP product are decisionmakers and researchers who use climate model output as input to studies or analyses in their respective, non-climatic disciplines (e.g., ecosystem science, hydrology and water resources, economics, human health, and agriculture and forestry)****. In order to facilitate application and decisionmaking using climate model information, an evaluation and assessment of the state of science of climate models is essential. This product is directed towards this goal.* (Emphasis mine.)

General Review Comments

This title of document is "Climate Models: An Assessment of Strengths and Limitations **For User Applications**." Unfortunately, much of the draft does not address user applications and appears to be directed at climate scientists, not decision makers. In addition, this first version feels very much like a simple collation of facts, rather than a true 'Synthesis and Assessment'. This is understandable given the magnitude of the task, but this product needs substantial work if it is truly to be of use to decision makers. Many of the critical questions posed in the 3.1 Prospectus have not been discussed, let alone answers provided. I am sympathetic to the authors of this document – these are extremely difficult questions to answer -- I would suggest that some have never been answered anywhere – and answering them may require the assistance of experts currently not present on the panel. Nevertheless, these questions are the key part of the document and if they can't be answered such acknowledgement needs to be forthrightly made and some form of guidance for decision making provided.

In order to be useful to decision makers, the document needs several additions. It needs tables that synthesize answers to key questions. It needs an executive summary designed to be read by the heads of water management agencies with a body that can be understood by the technical staff of those agencies. The diagrams and tables in this document should ultimately tell much of the storyline, yet, with the exception of chapter 5, there are very few diagrams and/or tables. It also needs a section on where to get other basic information on climate models and a glossary. All too often, e.g. Chapter 2, the material is presented as a highly technical data dump with no introductory material. Detailed technical information is fine, but the document needs a hierarchy of detail that is navigable by readers with different backgrounds. Finally the document needs to provide less informed readers with a road map of other basic documents on climate modeling including for example, the relevant IPCC chapters. I understand that many of these suggested enhancements are rarely found in first drafts, but I want to make sure that these considerations are incorporated in future versions.

The Prospectus states: *"The intended use of this CCSP product is to provide information to those who use climate model outputs about the strengths and limitations associated with using models to project the potential effects of human activities on climate and sea-level rise.* ***A discussion of appropriate and inappropriate uses of model output will be included.*** *The product will address scientific issues on a comprehensive, objective, open, and transparent basis. While based on the peer-reviewed scientific literature,* ***it will be written to be accessible and useful to the well-informed general reader and decisionmaker.****"* I failed to find a discussion about appropriate and inappropriate uses of model output. As currently written, the material is not accessible to the well-informed general reader and decision-maker.

Finally, the document has a substantial emphasis on US models. Given the expertise of the authors this is understandable, but decision makers are not just interested in US models. While these models are high quality and deserve discussion, other models of similar quality exist, and it seems increasingly clear that multi-model results do a better job of constraining the future than do smaller subsets. Some discussion or acknowledgement of other high quality models would substantially enhance this document.

Specific Comments on the Six Prospectus Questions

The Prospectus includes six key questions which I provide below in italics along with my analysis of whether these questions have been adequately answered. I will note that as a reviewer, sometimes these questions do not directly map to a single part of the Draft document. While I do not desire to tell the authors how best to structure this document, the existing organization makes it difficult to assess if these questions have been addressed. (And as a minor sidelight, the lack of a Table of Contents in the draft was also an unfortunate barrier to analysis.)

1) What are the major components and processes of the climate system that are included in present state-of-the-science climate models, and how do climate models represent these aspects of the climate system? This section will include descriptions of crucial processes such as tropical convection and major feedbacks in the climate system (e.g., clouds, atmospheric water vapor, surface albedo, and soil moisture). This section will evaluate the ability of the current generation of models to simulate key processes, and identify gaps in understanding. It will also include brief discussion of crucial processes that are likely to play an important role in climate that are not yet incorporated in the models.

This question maps to Chapter II, Description of Global Climate System Models. This chapter deserves a less technical introduction, and the section in the Introduction entitled 'A Brief History of Climate Model Development' belongs in this chapter. (The Introduction also needs significant work.) Simple concepts like "parameterization" need to be introduced. A simple table or two that answer the questions posed above would be extremely useful.

2) How are changes in the Earth's energy balance incorporated into climate models? How sensitive is the Earth's (modeled) climate to changes in the factors that affect the energy balance? This section will explain current approaches for incorporating changes in radiative forcing from both natural and human factors since the pre-industrial era. These include changes resulting from greenhouse gas and trace constituent emissions into the atmosphere, volcanic eruptions, and variations in the sun's intensity. This section will present a brief overview of the response of the global climate system, as derived from climate model results, for the various forcings (e.g., solar, volcanic, aerosols, anthropogenically derived greenhouse gases). The relative contributions of natural variability and human-caused factors for the period under consideration will be examined.

These questions are addressed in Chapter IV, Model Climate Sensitivity. As currently structured, Chapter IV deals with uncertainties due to variability in solar radiative forcing, cloud and water vapor feedbacks, and aerosols as well as the general concepts associated with differing model sensitivities. Like most of the chapters, an introduction with a roadmap and a simple explanation of the magnitude of each of these factors would enhance the understandability for high level decision makers. Alternatively, a chapter summary could provide the same material.

3) How uncertain are climate model results? In what ways has uncertainty in model-based simulation and prediction changed with increased knowledge about the climate system? This section will provide a discussion of the major sources of uncertainty in climate model results, as estimated through structured intercomparisons to observations, including the identification of the major sources of uncertainty in model assumptions and the characterization of radiative forcing. A description (or acknowledgement) of how increased knowledge can lead to greater uncertainty by increasing the number and complexity of processes included in climate models will be included.

Answers to these questions are found in Change V, Model Simulation of Major Climate Features. This is an enormous question and the Prospectus doesn't give much guidance on exactly what results this question should apply to. From the perspective of Rocky Mountain water users, temperature, precipitation and runoff are three key variables of interest but other decision makers are likely interested in the complete suite of potential model results including sea level rise. A summary table is critical.

In chapter V, uncertainties associated with mean and other statistical moments are discussed. The discussion needs some more regional focus, including graphics for the US. Modeled precipitation biases need much greater discussion.

I do not believe the second question and the last question above are addressed in the Draft. These are critical questions.

4) How well do climate models simulate natural variability and how does variability change over time? The ability of climate models to simulate the climatology and interannual variability is crucial for their use by the impacts and applications community. This section will describe efforts to evaluate these aspects of model performance. This section will also discuss the ability of climate models to simulate known patterns of natural variability, such as the Madden-Julian Oscillation, the El Niño Southern Oscillation, the North Atlantic Oscillation, and the Pacific Decadal Oscillation. A section on how these modes of variability have changed over time will be included.

This material is covered in Chapter V, Model Simulation of Major Climate Features, under the Modes of Variability subsection. The section includes some areas not spelled out by the Prospectus including extremes but lacks information on the NAO and PDO. A summary of this information is critical.

5) How well do climate models simulate regional climate variability and change? This section will discuss how changes in certain regions (e.g., the North Atlantic or

Tropical Pacific) can influence global climate change. It will also discuss limitations of "downscaling" methodologies—including regional climate modeling— used to obtain regional information from global simulations.

For the most part, answers to these questions appear to be entirely missing from the document. Chapter III discusses The Added Value of Regional Climate Model Simulations, but that discussion is mostly technical in nature and does not discuss the output of such efforts. A discussion about the merits of dynamical vs. statistical downscaling is implied by the last part of this question. Such a discussion would be extremely useful to decision makers.

6) What are the tradeoffs to be made in further climate model development (e.g., between increasing spatial/ temporal resolution and representing additional physical/ biological processes)? This section will consider the opportunities and constraints on future model development (e.g., additional computational cycles and lack of process knowledge). It will outline prospects for improvements potentially important to policymaking and decisionmaking.

Presumably this question maps entirely to Chapter VI, Future Model Development. The current version of the chapter discusses a variety of technical modeling issues and provides a general overview associated with Cloud Resolving Models, the Carbon Cycle and other biogeochemical cycles, Land Cover and Land Management Practice Changes and Ocean Biogeochemistry. Unfortunately, it doesn't address any of the questions posed in the Prospectus. These are very important questions and must be addressed in future drafts.

Summary

This document has the potential to be very useful to decision makers coping with the various uncertainties surrounding future climate change including how well current climate models work and their applicability for planning purposes. The first draft of this document is a good start but it needs to be written in a way that is truly accessible to decision makers. This means that adequate introductory material needs to be added and simple summaries provided. In addition, the authors need to provide answers to all questions in the Prospectus.

REVIEWER 6:

At first glance, I wondered about the need for this 140-page report, given that (1) the new Fourth Assessment Report of the IPCC's WG1 contains a chapter assessing current climate models, and (2) Parkinson and Washington's climate modeling textbook -- 2nd edition, 2005 -- and Trenberth's *Earth System Modeling* volume cover the subject of climate modeling rather comprehensively. However, after a careful reading, I concur that the report serves a useful purpose. The focus on the climate models of the U.S. gives the report its niche. The most effective portions of the present draft adhere to this focus; the least effective do not.

It is also commendable that the report largely restricts itself to simulations of recent (20th century) and present climate. The inclusion of projected changes and/or paleoclimate simulations would have opened cans of worms and made the report unwieldy.

An inevitable limitation of the report is that its shelf life will be rather short. As the next generation of models comes on the scene over the next few years, progress in climate modeling will likely make much of this report outdated. Periodic updates, *a la* IPCC, would be a remedy.

As noted in the instructions to reviewers, this report provided to me for review in February was clearly a first draft. This draft was missing some sections (e.g., Executive Summary; p. 30); not all figures were included; and the need for further input was mentioned parenthetically on many occasions. Although a reviewer can contribute more effectively when a first draft is polished, I am nevertheless providing detailed comments on this first draft, while ignoring many editorial details and obvious typos that will likely be fixed in the next draft.

First, a general comment. On many occasions the text states that models "have improved" in their simulations of (fill in the blank). Those statements should always indicate what the improvements are relative to -- and ideally should provide specific references to substantiate the statements.

[For reference purposes, I have numbered the pages of my draft, beginning with "Chapter 1 -- Introduction" as p. 1]; my version of the draft did not have page numbers].

1. Page 1, middle: The text states that the report will focus primarily on the CMIP models. Given the immense coordination and archival of models for the IPCC's AR4 (2007) and the many journal papers now reporting diagnostic assessments of those models, an emphasis on the AR4 models would certainly be more timely. Some sections (e.g., the assessment of model-simulated ENSO variability on p. 112-121) do utilize the AR4 models.

2. Page 57, bottom: The sentence "The models are also adjusted in different ways…so as to optimize so as to optimize the fit to observations deemed to be of particular importance" seems extremely significant and deserving of elaboration -- perhaps with

a few examples. This issue is fundamental to an understanding of the climate modeling enterprise.

3. The text should address the issue of flux adjustment (or other restoring methodologies) by defining these procedures and saying where they are -- or are not -- used in the models. The only such comment I could find is on p. 28 and pertains specifically to the GFDL model. If flux adjustment/restoring is not used anywhere in any U.S. model, then say so – that's a major advance over the coupled models of ten years ago.

4. Pages 101-106: The "Polar climate" section does not focus on U.S. models, creating some inconsistency with other sections. The subsequent section on sea ice (p. 107-108) does emphasize the U.S. models.

5. Pages 101-109: Missing from the presentation are assessments of the models' simulations of permafrost and snow cover (including snow cover over the larger non-permafrost areas). Even generic background information is missing -- for example, do any of the models include permafrost thaw or, more generally, the effects of soil freeze/thaw on hydrology?

6. Related to snow cover: Somewhere in the text (Under "Land surface models"?), there should be some indication of how snow is masked by vegetation in the models. Snow/vegetation interaction is important for the albedo-temperature feedback seasonally, interannually and over longer timescales.

7. How is glacial runoff (seasonal melt) included or not included? Page 82, middle, implies that CCSM3 does not include this process. Presumably none of the models include ice sheet dynamics. Given this issue's high visibility during the recent release of the IPCC's SPM (WG1), the text should include more explicit information on the treatment of ice sheets and glaciers in climate models.

8. The section on sea ice models (p. 23-26) fails to do justice to the role of high-frequency ice deformation, which controls the fractions of open water and thin ice where the vast majority of surface/atmosphere exchanges occur over polar oceans.

9. Pages 28-30: The sections on "Recent development paths" at the U.S. modeling centers will have an especially short shelf life. When the missing sections are provided, they should be made as brief as possible -- and the GFDL section could be shortened.

10. Pages 6-16: The description of atmosphere and ocean GCMs seems <u>way</u> more technical than necessary. This is material for textbooks, not for this report. I suggest trimming by 50% or more.

11. Pages 32-40 are confusing and difficult to review (e.g., with figures missing). More specifically, the presentation of results from one ensemble member "that agrees best

with the observed…time series" seems like a very dubious practice -- planners and policymakers can easily misinterpret and misuse such selective results.

12. Page 80: Is the mid-century cooling, as simulated by the models, due to the prescribed aerosols? to prescribed solar variability? or both? Given the recent paper by M. Wang et al. (2007, *J. Climate*) arguing that the mid-century warming/cooling is consistent with natural variability in some IPCC AR4 models, elaboration seems to be needed here.

13. Pages 112-121: This ENSO section is <u>way</u> too long. Simulation of ENSO was already covered on p. 86-87, where the models' underestimation of equatorial Pacific SST variability is reported.

<u>Minor points</u>:

14. Page 79, first sentence of "B. 20th century trends": Why not identify the model? Models are identified everywhere else in the report.

15. Page 82: A reference is needed for the stratosphere-troposphere coupling of volcanic effects.

16. Page 84: The statement that the Gulf Stream warms Europe needs to be placed in the context provided by R. Seager et al. (2006, "The source of Europe's warm climate", *American Scientist, 94*, 334-341).

17. Page 84: What is the uncertainty in estimates of the integrated transport by the AMOC? The text shows uncertainties only in transports at specific latitudes (p. 93-94, figure from Schmidt et al., 2006).

18. Page 99, lines 10-11: Where is the circulation strength in both winter and summer…expected to weaken?"

19. Page123: Reference to Burke et al. (2006) is missing.

20. Page 124: How is the "4th largest precipitation event" equivalent to the 99th percentile?

21. Page 131: I suggest omitting the second paragraph on this page ("Research with CRM falls into two categories…"). The present report does not need this paragraph.

22. Page 138: The source for Figure VI-1 needs to be provided.