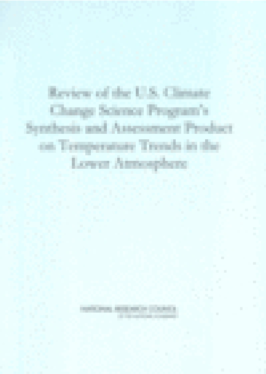

Review of the U.S. Climate Change Science Program's Synthesis and Assessment Product on Temperature Trends in the Lower Atmosphere

Committee to Review the U.S. Climate Change Science Program's Synthesis and Assessment Product on Temperature Trends in the Lower Atmosphere, Climate Research Committee, National Research Council

ISBN: 0-309-55109-9, 74 pages, 8 1/2 x 11, (2005)

This free PDF was downloaded from:
http://www.nap.edu/catalog/11285.html

Visit the National Academies Press online, the authoritative source for all books from the National Academy of Sciences, the National Academy of Engineering, the Institute of Medicine, and the National Research Council:

- Download hundreds of free books in PDF
- Read thousands of books online, free
- Sign up to be notified when new books are published
- Purchase printed books
- Purchase PDFs
- Explore with our innovative research tools

Thank you for downloading this free PDF. If you have comments, questions or just want more information about the books published by the National Academies Press, you may contact our customer service department toll-free at 888-624-8373, visit us online, or send an email to comments@nap.edu.

This free book plus thousands more books are available at http://www.nap.edu.

Copyright © National Academy of Sciences. Permission is granted for this material to be shared for noncommercial, educational purposes, provided that this notice appears on the reproduced materials, the Web address of the online, full authoritative version is retained, and copies are not altered. To disseminate otherwise or to republish requires written permission from the National Academies Press.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

Review of the U.S. Climate Change Science Program’s Synthesis and Assessment Product on Temperature Trends in the Lower Atmosphere

Committee to Review the U.S. Climate Change Science Program’s Synthesis and Assessment Product on Temperature Trends in the Lower Atmosphere

Climate Research Committee

Board on Atmospheric Sciences and Climate

Division on Earth and Life Studies

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

THE NATIONAL ACADEMIES PRESS • 500 Fifth Street, N.W. • Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was supported by a contract between the National Academy of Sciences and the National Oceanic and Atmospheric Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

International Standard Book Number 0-309-09674-X

Copies of this report are available from the program office:
Board on Atmospheric Sciences and Climate
500 Fifth Street, N.W.
Washington, DC 20001
(202) 334-3512

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Wm. A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. Wm. A. Wulf are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

COMMITTEE TO REVIEW THE U.S. CLIMATE CHANGE SCIENCE PROGRAM'S SYNTHESIS AND ASSESSMENT PRODUCT ON TEMPERATURE TRENDS IN THE LOWER ATMOSPHERE

WILLIAM J. RANDEL (*Chair*), National Center for Atmospheric Research, Boulder, Colorado
JUDITH A. CURRY, Georgia Institute of Technology, Atlanta
DENNIS L. HARTMANN, University of Washington, Seattle
PHIL JONES, University of East Anglia, Norwich, United Kingdom
KENNETH KUNKEL, Illinois State Water Survey, Champaign
RICHARD S. LINDZEN, Massachusetts Institute of Technology, Cambridge
RICHARD L. SMITH, University of North Carolina, Chapel Hill
JOHN MICHAEL WALLACE, University of Washington, Seattle
JUNHONG WANG, National Center for Atmospheric Research, Boulder, Colorado

NRC Staff

CHRIS ELFRING, Director, Board on Atmospheric Sciences and Climate
PARIKHIT SINHA, Study Director
RACHAEL SHIFLETT, Senior Program Assistant

CLIMATE RESEARCH COMMITTEE

ANTONIO J. BUSALACCHI, JR. (*Chair*), University of Maryland, College Park
LEE E. BRANSCOME, Climatological Consulting Cooperation, Palm Beach Gardens, Florida
JAMES A. COAKLEY, JR., Oregon State University, Corvallis
JULIA E. COLE, University of Arizona, Tucson
JUDITH A. CURRY, Georgia Institute of Technology, Atlanta
CLARA DESER, National Center for Atmospheric Research, Boulder, Colorado
DAVID KAROLY, University of Oklahoma, Norman
ROBERT J. LEMPERT, RAND Corporation, Santa Monica, California
LINDA O. MEARNS, National Center for Atmospheric Research, Boulder, Colorado
GERALD A. MEEHL, National Center for Atmospheric Research, Boulder, Colorado
PETER B. RHINES, University of Washington, Seattle
W. JAMES SHUTTLEWORTH, University of Arizona, Tucson
LYNNE D. TALLEY, Scripps Institution of Oceanography, La Jolla, California
HASSAN VIRJI, International START Secretariat, Washington, D.C.
YUK YUNG, California Institute of Technology, Pasadena

NRC Staff

AMANDA STAUDT, Senior Program Officer
ROB GREENWAY, Senior Program Assistant

BOARD ON ATMOSPHERIC SCIENCES AND CLIMATE

ROBERT J. SERAFIN (*Chair*), National Center for Atmospheric Research, Boulder, Colorado
M. JOAN ALEXANDER, Colorado Research Associates, Boulder
FREDERICK R. ANDERSON, McKenna Long & Aldridge LLP, Washington, D.C.
MICHAEL L. BENDER, Princeton University, New Jersey
ROSINA M. BIERBAUM, University of Michigan, Ann Arbor
MARY ANNE CARROLL, University of Michigan, Ann Arbor
CAROL ANNE CLAYSON, Florida State University, Tallahassee
WALTER F. DABBERDT, Vaisala Inc., Boulder, Colorado
KERRY A. EMANUEL, Massachusetts Institute of Technology, Cambridge
DENNIS L. HARTMANN, University of Washington, Seattle
PETER R. LEAVITT, Weather Information, Inc., Newton, Massachusetts
JENNIFER A. LOGAN, Harvard University, Cambridge, Massachusetts
VERNON R. MORRIS, Howard University, Washington, D.C.
WILLIAM J. RANDEL, National Center for Atmospheric Research, Boulder, Colorado
THOMAS H. VONDER HAAR, Colorado State University/CIRA, Fort Collins
ROGER M. WAKIMOTO, University of California, Los Angeles

Ex Officio Members

ANTONIO J. BUSALACCHI, JR., University of Maryland, College Park
ERIC F. WOOD, Princeton University, New Jersey

NRC Staff

CHRIS ELFRING, Director
PAUL CUTLER, Senior Program Officer
AMANDA STAUDT, Senior Program Officer
MARIA UHLE, Program Officer
ELIZABETH A. GALINIS, Senior Program Assistant
ROB GREENWAY, Senior Program Assistant
DIANE GUSTAFSON, Administrative Coordinator
ANDREAS SOHRE, Financial Associate

Acknowledgments

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

James Angell, National Oceanic and Atmospheric Administration, Silver Spring, Maryland
Richard Levine, San Diego State University, California
Jerry Mahlman, National Center for Atmospheric Research, Boulder, Colorado
Steven Running, University of Montana, Missoula
Siegfried Schubert, NASA Goddard Space Flight Center, Greenbelt, Maryland
Kevin Trenberth, National Center for Atmospheric Research, Boulder, Colorado

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the conclusions or recommendations nor did they see the final draft of the report before its release. The review of this report was overseen by Eric Barron, Pennsylvania State University. Appointed by the National Research Council, he was responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Contents

Introduction

The U.S. Climate Change Science Program (CCSP) was established in 2002 to coordinate climate and global change research conducted in the United States. Building on and incorporating the U.S. Global Change Research Program of the previous decade, the program integrates federal research on climate and global change, as sponsored by 13 federal agencies and overseen by the Office of Science and Technology Policy, the Council on Environmental Quality, the National Economic Council and the Office of Management and Budget. A primary objective of the CCSP is to provide the best possible scientific information to support public discussion and government and private sector decision-making on key climate-related issues. To help meet this objective, the CCSP is producing a series of synthesis and assessment products that address its highest priority research, observation, and decision-support needs. At this time, the CCSP plans to conduct 21 such activities over the next 3 years, covering topics such as the North American carbon budget and implications for the global carbon cycle, coastal elevation and sensitivity to sea-level rise, and trends in emissions of ozone-depleting substances and ozone recovery and implications for ultraviolet radiation exposure. Each of these documents will be written by a team of authors selected on the basis of their past record of interest and accomplishment in the given topic.

The National Oceanic and Atmospheric Administration (NOAA) is the lead agency for the first CCSP synthesis and assessment product, which focuses on both understanding reported differences between independently produced data sets of temperature trends for the surface through the lower stratosphere and comparing these data sets to model simulations (see Appendix A). In trying to understand these differences, the assessment attempts to answer six fundamental questions (see Box 1). This synthesis and assessment product builds on and extends the results of a 2000 National Research Council (NRC) report, *Reconciling Observations of Global Temperature Change*, relevant parts of the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC, 2001), and other advances in our understanding of this issue. This assessment is expected to contribute to the IPCC Fourth Assessment Report (due to be published in 2007) and to the Global Climate Observing System (GCOS) Atmospheric Observation Panel by identifying effective ways to reduce observational uncertainty. The assessment is to be written in a style consistent with major international scientific assessments.

In a recent review of the U.S. CCSP Strategic Plan, the NRC recommended that synthesis and assessment products should be produced with independent oversight and review from the wider scientific and stakeholder communities (NRC, 2004). To ensure credibility and quality, NOAA has requested an independent review of its synthesis and assessment product on temperature trends by the NRC. The NRC committee’s statement of task is included in Appendix B. The committee conducted its work by reading the CCSP report *Temperature Trends in the Lower Atmosphere: Understanding and Reconciling Differences* (draft dated 2/9/2005) carefully, meeting with the authoring team to ask questions, and then compiling this summary of reactions. In addition, a public comment period is scheduled to occur after this review has been completed and revisions have been made by the authoring team.

BOX 1
Questions to be Addressed in the First CCSP Synthesis and Assessment Product

The first CCSP synthesis and assessment product focuses on both understanding reported differences between independently produced data sets of temperature trends for the surface through the lower stratosphere and comparing these data sets with model simulations. The fundamental questions posed in the assessment are:

1. Why do temperatures vary vertically (from the surface to the stratosphere) and what do we understand about why they might vary and change over time?
2. What kinds of atmospheric temperature variations can the current observing systems measure and what are their strengths and limitations, both spatially and temporally?
3. What do observations indicate about the changes of temperature in the atmosphere and at the surface since the advent of measuring temperatures vertically?
4. What is our understanding of the contribution made by observational or methodological uncertainties to the previously reported vertical differences in temperature trends?
5. How well can the observed vertical temperature changes be reconciled with our understanding of the causes of these changes?
6. What measures can be taken to improve the understanding of observed changes?

Major Review Comments

Understanding temperature trends in the lower atmosphere over recent decades is important to our overall understanding of the Earth's climate and its evolution. Such understanding requires both analysis and comparison of changes derived from different observing platforms and comparison of observed changes to those simulated by models. The vertical profile of temperature changes is particularly important because it provides a fingerprint of the mechanisms responsible for past changes. This topic is of substantial interest to the climate research community, especially the key issues of data quality, data-model agreement, and monitoring of future climate change. Thus the Climate Change Science Program (CCSP) synthesis and assessment product on this topic stands to be a timely and useful document, and the committee commends the CCSP for initiative and leadership in this area. *Temperature Trends in the Lower Atmosphere: Understanding and Recognizing Differences* is a good first draft that covers an appropriate range of issues. It reflects extensive effort and coordination by a team of experienced, knowledgeable authors. By nature, this review of the report is critical, but our comments are intended as constructive input to strengthen the document.

This review begins with a summary of the most important, major comments for the document as a whole. Subsequently comments and specific suggestions for each of the chapters of the CCSP synthesis and assessment product, called the Temperature Trends report for brevity, are presented.

1. The Temperature Trends report should include an improved discussion of the motivation for this report, which will increase the report's effectiveness for a variety of audiences. The committee also suggests more explicit clarification of the context and intended audience for this report. This background should occur in the preface or in an introduction and address the specific scientific issues that motivated the work (which surprisingly go unmentioned in the draft), what has been done previously, and what are the key outstanding issues. This section should be accessible to general educated readers and also scientifically sound.
2. To help the report communicate effectively to an array of audiences, the committee suggests changes to the presentation style within each chapter. The key findings for each chapter, should be brought to the front of that chapter possibly in the form of bulleted highlights, with a one-sentence summary and brief discussion for each key point (similar to the format within Intergovernmental Panel on Climate Change and

World Meteorological Orginization Assessments). The key points should be based on the detailed discussions within each chapter. The chapters should in some cases include more scientific rigor (as detailed in the chapter reviews that follow) and be aimed at the broad climate science community. One possible strategy for including scientific rigor in the chapters is to use footnotes to describe technical details, as was done in Chapter 5.

3. More explicit discussion of the statistical characterization of uncertainty in trends is needed, with emphasis on several specific topics.

    a. The conclusions reached in the report are often based on estimates of trends, neglecting uncertainty levels, and many statements on comparisons are inaccurate because of this. Uncertainties are relatively large because of small trend signals, large natural variability, and data records that are relatively short. Conclusions within the report should more accurately reflect the uncertainties inherent in the statistical trend calculations.

    b. A more thorough discussion of the detailed statistical trend calculations for the various data sets is needed. This discussion might be appropriately placed within an appendix.

    c. When comparing trend differences between two estimates of the same quantity (e.g., tropospheric temperatures from radiosondes and satellites), it is more appropriate to examine the trend of the difference time series, rather than trends for each time series individually (because the data contain similar overall variability). Specific instances where this is relevant are identified in the chapter reviews that follow.

4. The report would benefit from a more critical evaluation of the trend differences between the University of Alabama, Huntsville (UAH) and Remote Sensing Systems (RSS) satellite microwave data sets. The report presents results from UAH, RSS, and University of Maryland (UMD) satellite data and effectively treats the UAH and RSS results as equal independent estimates with a lesser discussion of UMD results because the latter do not include a diurnal correction and because their newest analysis appeared only recently. An expanded discussion and critical assessment of the UAH and RSS differences is needed in particular, as well as a discussion of the implication of the differences. The committee would like to see a resolution of this difference or, if that is not possible within the current report, some discussion of what sorts of analyses would be required to resolve the differences between the two groups' trend estimates. Similar scrutiny of the UMD product and its associated uncertainties would be valuable.

5. Satellite trends derived from combining separate microwave channels (the so-called Fu et al. technique, as described in Fu and Johanson [2004 and 2005], and Fu et al. [2004a and b]) should be discussed in more detail. The Fu et al. results are potentially of key importance to the issue of tropospheric temperature trends and should be discussed more thoroughly. Currently, there is minimal discussion, and the data are labeled "controversial". A more thoughtful and balanced appraisal is necessary, and more recent references should be included.

6. A more thorough discussion of biases in trends derived from historical radiosonde data is needed. The report concludes that stratospheric trends derived from radiosondes are biased so badly as to be unbelievable (a result for which there appears to be a

community consensus). However, details of the causes of these biases should be clearly described and a stronger explicit case made for discounting the stratospheric results. More critically, it is important to assess if the stratospheric bias problem extends into the middle and upper troposphere, especially in the tropics. The vertical range for which the radiosonde data have least the uncertainties for trend analyses should be more carefully discussed. Finally, the prospects for removing the biases or more effectively utilizing these data should be discussed.

7. The recommendations in Chapter 6 are too non-specific, unprioritized, and largely disconnected from the findings in Chapters 1-5. We suggest that Chapter 6 be reorganized into two parts:

   a. The first part should take findings from Chapters 1-5 to recommend specific opportunities to improve understanding of vertical temperature trends. These recommendations should focus on understanding remaining uncertainties in existing satellite and radiosonde data sets.
   b. The second part should focus on future measurement opportunities in the context of the specific goals of the report for reconciling observation and understanding of temperature trends.

8. Changes in the presentation and content of the Executive Summary are needed to make key results more accessible to a wide audience and ensure traceability to the results in Chapters 1-6. A possible strategy is to bring forward key bullet points from each chapter as answers to the six main chapter questions, followed by brief explanatory text, key figures, and implications for understanding within each chapter. The Executive Summary should reflect an appropriate balance of new results and outstanding uncertainties. The first page of the Executive Summary should concisely summarize the key results of the report in a short abstract.

One further summary comment is relevant for the CCSP synthesis and assessment reports in general. The committee feels that the current report suffered to some degree from the author group assessing their own work and excluding other independent work. This is evidenced by a lack of critical evaluations on some key data issues and numerous citations to their own work. To the extent possible, the authors should not be put in a position of assessing work where they have a vested interest in the outcome. While it is not reasonable to revise author teams for this report, those preparing future CCSP reports should carefully consider this issue.

Review of the Executive Summary

The Executive Summary is intended to present the key findings from the main body of the report, written in a style intelligible to a technically literate lay audience. The most obvious take-home messages and quotable text should be presented in a clear and concise manner. An effective summary is especially important because many readers of the report will carefully read only the Executive Summary, merely browsing the individual chapters.

In the committee's opinion, the draft Executive Summary does not communicate the most important points in an effective manner. It is quite long and inefficient in conveying information, with many restatements of various methodological issues that result in no conclusions. In fact, it tends to read as if it were a chapter itself, not a crisp summary of the key major conclusions, accomplishments, and future challenges in this important topic. Changes in the length, presentation, and content of the Executive Summary are needed to make key results more accessible to a wide audience and to ensure traceability to the results developed in Chapters 1-6.

MAJOR COMMENTS

1. A possible strategy to structure the Executive Summary is to bring forward key bullet points from each chapter as answers to the six main chapter questions, followed by brief explanatory text, key figures, and implications for understanding within each chapter. The Executive Summary should reflect an appropriate balance of conclusions and outstanding uncertainties. The first page of the Executive Summary should concisely summarize the key results of the report in a short abstract.
2. The first two paragraphs of the Executive Summary say essentially the same thing—that models and observations are now more "consistent". It is not clear what these statements really mean. A clear statement about what the observations show, independent of model projections, including specifically what is new in this report that has not been presented in previous reports cited would provide clarification. A second conclusion might address how new observations and new model runs affect our perception of whether the models are consistent with observed trends or basic theory. As currently written, the observational work and the modeling work seem muddled together.

3. It appears that some subjectivity is necessary in making the optimistic-sounding statements in the first two paragraphs. Given the very small absolute changes that have been observed over the 20 years of the focus period of 1979-1999 (~ 0.2K) and the larger natural variability that has occurred over that time (~ 0.5K), it seems that the data can neither reject nor confirm the hypothesis that the models are in some sense reliable. The Executive Summary focuses almost exclusively on the 1979-1999 period, a rather short period for which the trends discussed would not be significant at the 5% level, if the significance testing were done in the usual way. The choice to present the results without any statistical significance testing or confidence intervals is highly questionable and ordinarily not allowed in the scientific literature.

4. Some additional support for the conclusions in the first two paragraphs might be gained if the issue of separating the stratospheric and tropospheric signals in the Microwave Sounding Unit (MSU) channels were addressed more directly in the body of the report and carried forward to the Executive Summary. In this regard, the Fu et al. results appear important but are basically ignored in the explicit arguments used to formulate the conclusions. This should be remedied. The Fu et al. work may be central to the issue of measuring and interpreting the vertical profile of temperature change and is no more difficult or controversial than addressing the spurious radiosonde trends or the differences between the Remote Sensing Systems (RSS), University of Alabama, Huntsville (UAH), and University of Alabama (UMD) interpretations of the MSU data.

5. The traceability of the conclusions in the Executive Summary to the detailed arguments presented in the chapters is not entirely clear. In some cases the Executive Summary does not appear consistent with the corresponding chapters where the conclusions should have been developed. A way to clarify the source of the conclusions might be to develop major conclusions in the chapters and then move these statements forward to the Executive Summary.

SPECIFIC COMMENTS

1. The Executive Summary should start with a statement about why the reader should care about the subject of the report.

2. The first two paragraphs appear to be the abstract of the document and contain the main takeaway messages. Both paragraphs conclude with statements about the comparison of models with observations. In the second paragraph, it is not clear what consistent means. The observations seem to give qualitatively different conclusions on several key points. Does consistency merely mean that the observations and the model results have overlapping uncertainties?

3. In lines 71-73, the conclusion that the report increases confidence in our understanding of recent climate change seems optimistic and inconsistent with the supporting evidence. The evidence concerning this conclusion needs to be spelled out more clearly in the document and succinctly summarized in the Executive Summary. It must be possible and convenient to trace these conclusions back to arguments that can be evaluated scientifically.

4. Figure 1 is not effective. For example, the color scheme should be changed because the "multi-colored line" is mostly just one shade of blue.

5. Figure 2 is the basis for the major new conclusions summarized in the first two paragraphs of the report; however, it is not developed in detail and appears only in the Executive Summary and not in the supporting chapters. It seems that additional analysis or logic is being applied in the Executive Summary that is not present in the supporting chapters, suggesting a lack of traceability between the major conclusions of the Executive Summary and the supporting chapters. A number of critical decisions were made about which data sets to include and exclude and whether or not to show sampling uncertainty. These decisions are not adequately discussed in the report. Figure 2 deals only with the 20-year period from 1979-1999. This is a rather short period so the sampling errors are large relative to the absolute changes. However, Figure 2 has merit in that it does include error bars that are generally missing from the report’s figures.

6. Figures 2-5, and 7: All the arguments are formulated in terms of the satellite weighting functions. A rather strange nomenclature is developed based on the MSU weighting functions shown in Figure 3, Low-Trop, Trop-UW, Mid-Trop and Low-Strat. The labeling is misleading because Mid-Trop has a significant contribution from the stratosphere, and Low-Strat has a considerable contribution from the troposphere. It is potentially misleading to call these estimates tropospheric and stratospheric as they are a blend of tropospheric and stratospheric trends. It might be more accurate to use the satellite nomenclature T2LT, TFU, T2 and T4. Projecting the radiosonde data onto these weighting functions forces one to combine potentially good lower tropospheric radiosonde data with potentially bad radiosonde data above 200 mb.

7. An important statement in the report is “the climate models simulate greater warming in the troposphere than at the surface which is not apparent in the observations” (lines 67-68). Whether this statement is true depends on whether the results by Fu et al. are correct. As discussed elsewhere, the Fu et al. results should be discussed in the report chapters and then distilled in the Executive Summary.

   a. The Fu et al. retrieval is mentioned briefly in Chapters 3, 4, and 5, but it is not included in the broader discussion. Given that the Fu et al. results appear in the peer reviewed literature, it is inadequate to dismiss them as “controversial”.

   b. It is only fair to include the published Fu et al. 850-300 mb temperature trend estimate in the figures presented in Chapter 5 and in the Executive Summary (and to modify the text accordingly). The authors may state their reservations about the Fu et al. method, provided that they distinguish clearly between what is published and what is (as yet) unpublished, and that they incorporate Fu et al.’s published reply to the one published paper (thus far) that is critical of their method (Tett and Thorne, 2004).

   c. Failure to account for the stratospheric contribution in the comparisons between data and models may compromise the report’s conclusions. This should be addressed in Chapter 5 and also summarized in the Executive Summary.

8. In Figure 4, why were the negative weights for TLow-Trop truncated off the plot at the left edge?

9. Figure 5 should include error bars.

10. lines 123-153 The section on suggests “Motivation for this Report” that the main purpose of the report is to address the single issue of surface versus tropospheric temperature trends over the past 20 years, yet the six questions that were to be addressed

and the main body of the report seem to be a somewhat more general approach to the question of temperature trends. Is the Temperature trends report intended to be a summary and extension of the 2000 NRC report or a more general statement of knowledge about temperature trends?

11. In lines 163-169, it would be helpful to specify exactly which are the new data sets that have lead to new interpretations. This information is not clearly stated anywhere in the Executive Summary. What specifically happened since NRC (2000) and IPCC (2001) and how has this changed perceptions? The two paragraphs on this page should be replaced with simple, direct statements of fact, if possible.

12. In lines 353-359, it seems misleading to use the phrase: “at any one level”. The comparison has not been made for levels, but for very deep layers as specified by the MSU channels.

13. In Section 3.2, “Radiosonde data” no specific information is given about uncertainties in the two radiosonde datasets.

14. Does the statement on lines 426-427 include an assessment of the sampling uncertainty and a statistical confidence level?

15. In lines 557-579, rather, discuss succinctly and in common language the meaning of Figures 8 and 9. The discussion of fingerprinting is not really necessary. Technique description, jargonizing, and philosophizing should be eliminated in favor of straightforward, accurate, descriptions of the significant results or conclusions.

16. In line 586 the statement that not including the indirect effect of aerosols is the most important deficiency of the global model simulations should be better justified. In particular, other uncertainties, such as those associated with cloud or water feedbacks, could be of similar magnitude.

17. Because Figure 6 is repeated in the top half of Figure 8, perhaps Figure 6 can be removed.

18. In lines 620-626 the statistical significance is very important, so it would be helpful to highlight the statistical significant areas in Figures 8 and 9, if they are known. If they are not known, then that is important too. Throughout the Executive Summary, the reader is invited to take the values presented literally, even though they may have very large statistical uncertainty in addition to residual structural uncertainty.

19. In Figure 9, it is somewhat of a misnomer to call this “Mid-Tropospheric Temperature”. It contains a significant contribution from stratospheric trends that increases in magnitude with latitude. Much of the apparent agreement in this figure is simply the result of negative trends in the stratosphere and the increasing fraction of the stratosphere that the weighting function samples as one moves toward the poles.

20. In lines 620-626, are the only trends that are statistically significant the ones above 100 mb, and are they believed to be spurious trends associated with the radiosonde instrument? If so, how are the conclusions about different trends in the surface and troposphere supported? It seems reasonable to suppose that the spurious negative trend in the radiosonde data extends below 100 mb, albeit with decreased magnitude.

21. In lines 639-640, given the uncertainties in both the direct aerosol forcing and the indirect aerosol effects, the report should provide better justification for the conclusion that the aerosol effects have almost certainly been underestimated.

22. Are the statements in lines 646-652 true for the radiosonde era, or just for the satellite era?

23. Section 6 "Improving our understanding" starts with rather general philosophical statements, whose connection to the recommendations that follow is unclear. The recommendations are broad and unspecific, despite the fact that the report raises some very specific problems. Some of the recommendations are not argued clearly elsewhere in the report, in particular Chapter 6. For example, it is not clear that Chapter 6 calls for "efforts to better understand and reconcile differences between climate data records that purport to measure the same variable," the second of the specific recommendations listed on page 32 of the Temperature Trends report. More specific recommendations that could immediately be acted upon could be to initiate an action to determine whether the RSS-UAH MSU discrepancy can be resolved or to better characterize the instrumental contribution to the radiosonde trends. In addition, more recognition that planning activities have gone on internationally to design an effective global climate monitoring system should be given.

24. It would be useful to have a summary table showing all datasets and models used in this report to avoid long figure captions.

Review of Chapter 1

Chapter 1 asks the following: Why do temperatures vary vertically (from the surface to the stratosphere) and what do we understand about why they might vary and change over time? The chapter begins with a limited description of the vertical temperature structure of the atmosphere and devotes most of its space to a consideration of forcing mechanisms that affect the local temperature structure. This chapter provides an excellent opportunity to explain the implications of differences in temperature trends between the troposphere and the surface. Such an explanation will be extremely important because, as the first of the Climate Change Science Program (CCSP) synthesis and assessment products, this report will be closely scrutinized with respect to scientific methodology, accuracy, and awareness of current understanding and theory.

The present text is inadequate on several counts, as listed in detail below. Of particular importance, the primary implications of differences or lack thereof in atmospheric and surface trends are never identified. In addition, the chapter pays insufficient attention to the effect of dynamics on temperature structure. The role of dynamics in smoothing temperature both horizontally (over the Rossby radius) and vertically (over the convective depth) is largely ignored. Finally, the chapter would be strengthened if it had further discussion of theory, which would complement the chapter's current emphasis on bringing models and observations into agreement. It should be recognized that we are discussing temperature changes of tenths of a degree—a challenge not only for instruments, but also for theory and modeling. Dealing with this challenge calls for a more sophisticated conceptual framework.

MAJOR COMMENTS

1. The explanation of the greenhouse effect should more clearly describe its effect on the atmospheric temperature structure. In particular, the chapter should explain how the addition of infrared absorbing gases causes the characteristic emission level to be at a higher altitude, where temperatures are colder and where the reestablishment of radiative balance with space calls for warming at this level and communication of this warming to the surface (Goody and Yung, 1989; Lindzen and Emanuel, 2002). Thus, for example, the absence of any warming within the troposphere might suggest that the greenhouse effect is not responsible for the surface warming. A related topic concerns the question of

whether temperature changes originating at the surface necessarily lead to temperature changes within the troposphere.

2. Similarly, a discussion of the relation between cumulus convection and the moist adiabat provides an opportunity to use such differential trends to understand the coupling between the surface and the lifting condensation level. Indeed, in the tropics, the temperature structure consists of a surface mixed layer (up to about 500 m) and a trade wind boundary layer (up to about 2 km) above which is the free troposphere. Each of the boundary layers is topped by an inversion which tends to isolate the layer from the region above (Sarachik, 1985). Outside the tropics, the surface communicates with upper levels primarily by quasi-horizontal motions along isentropic surfaces (e.g., Hoskins, 2003). Consequently, the report and the scientific community should move beyond the naïve notion that the lapse rate is a rigid constraint operating from the surface to the tropopause. Instead the observations this report is concerned with should be exploited in order to answer important questions about climate. This objective provides meaningful motivation for ascertaining the accuracy of the temperature measurements and the resulting time series. That said, it should be emphasized that the temperature changes being considered are changes on the order of tenths of a degree (although local changes may be much greater), and current theories may prove inadequate for such small changes.

3. In general, spatial and temporal sampling is not adequately dealt with in the Temperature Trends CCSP report. Given the fact that horizontal temperature variability at the surface tends to get smoothed as one rises to the free troposphere, there may be serious issues of sampling. Horizontal smoothing over large scales occurs above the boundary layer, but that at the surface and within the boundary layer, there can be much more horizontal variation of temperature. Thus, much more data may be required at the surface to get characteristic temperatures.

4. For Chapter 1, explanations of the processes involved in determining vertical profiles of temperature should represent the current state of understanding or lack thereof. The chapter should focus less on details of the vertical profile of temperature that are not resolved by the observations that are the focus of the report. For example, the satellite data are only reported in coarse vertical layers.

5. For discussions that are felt to be too detailed for the body of the text, footnotes are a reasonable device.

SPECIFIC COMMENTS

1. The chapter should include more discussion of theories that provide physical constraints on the apparent differences between surface and tropospheric records.

2. The discussion on lines 69-80 should be replaced with a more accurate figure as well as a description of the differences between the tropics, the extratropics, and the polar regions. In the tropics, the temperature is hardly linear with height, given that the lapse rate associated with the moist adiabat goes from about 5 K/km near the surface to almost 9.8 K/km at the tropopause near 16 km. It should also be noted that the tropopause descends sharply to 12 km near 30 degrees latitude and to around 8 km near the poles. The existence of the near surface inversion layer at high latitudes should also be noted as well as its dependence on meteorological conditions.

3. Relatedly, lines 100-131 should be replaced by a more complete discussion wherein it is noted that a radiative-convective balance is only likely to be of dominant relevance in the tropics, while in the extratropics, the lapse rate and the tropopause height are mostly determined by the same baroclinic instability that gives rise to weather systems (Schneider, 2004). Planetary-scale forced waves in winter and other circulation features, such as the Hadley and Walker cells, should be mentioned.

4. The authors should provide further discussion of the role played by dynamics. The discussion of dynamics in lines 128-131 should introduce the concept of Rossby radius. This vital concept shows that dynamics tend to homogenize temperatures (above the boundary layer) over horizontal scales that vary from the planetary scale near the equator to a couple of thousand kilometers at midlatitudes and to a few hundred kilometers near the poles.

5. The discussion in lines 133-137 should be strengthened, in particular so that it distinguishes between specific and relative humidity.

6. Remove "especially critical" from line 141.

7. In lines 150-156, the question of internal variability needs to be improved and clarified. For one thing, there can be internal variability without external forcing, and even without air-sea interaction. Further, there are limitations associated with using numerical models to examine the importance of internal variability because such models poorly characterize such things as El Niño/Southern Oscillation (ENSO), the 1976 regime shift, and the quasi-biennial oscillation (QBO) at the levels of tenths of a degree. It should be emphasized that most rules of thumb used for atmospheric structure may not be appropriate at the level of the small temperature changes being considered in this report.

8. It would be worth stressing that the temperature changes that are being discussed are only a few tenths of a degree. Much of our thinking is based on more substantial changes. There is an extensive literature arguing for and against the relevance of the moist adiabat in the tropics (e.g., Xu and Emanuel, 1989). However, even those arguing for its relevance would not argue that it should hold to better than a few tenths of a degree. Similarly, it might be argued that the role of motions should cancel when averaged over the earth. But the above is not strictly true. The existence of radiation leads to irreversibility, and when the strong changes in water vapor with latitude are taken into account, changes in circulation can lead to changes in global mean temperature that might be on the order of a few tenths of a degree.

9. In lines 194-202 and in lines 283-291, the report should be more cautious in arguing that local changes in radiative constituents can lead to local changes in temperature profile in light of such processes as the mean circulation in the tropics, which homogenizes temperature, and quasi-geostrophic dynamics in the extratropics.

10. In line 206, while the radiative impact of clouds is undoubtedly very important, further explanation is needed if one is to attribute to them a role as a "regulator".

11. In lines 222-224, it should mentioned that greenhouse gas forcing in the tropics is not uniform owing to the current distribution of clouds and water vapor. Thus, greenhouse gas forcing from anthropogenic sources is greatest in dry regions.

12. The claim of local radiative influence in lines 233-234 should either be explained or omitted.

13. In lines 251-252, caution should be suggested in adding unknown forcings because these can easily become nothing more than adjustable parameters. Of course, care should also be taken to include all forcings that are quantitatively known.

14. In lines 254-255, it should be noted that while the air-sea interaction can play a role in internal variability, such variability can also occur in the atmosphere alone.

15. In lines 258-261, while water vapor and clouds are indeed critical to the high climate sensitivity of many models, the references cited (Stocker et al., 2001; NRC, 2003) carefully note that water vapor and especially clouds are areas of major uncertainty in models, and even in nature.

16. The discussion of volcanic influence on lines 309-312 should be reworked to include additional work that has been done on this subject. For example, there is more on the effects of volcanoes on European temperatures in Jones et al. (2003) and in Robock and Oppenheimer (2003). The most affected region is Northern Europe—not North America and certainly not Siberia. The two studies cited in lines 309-312 are also basically model studies, and evidence from observations is less convincing.

17. The claim on lines 331-336 should note the substantial uncertainty of such factors as solar variability (Frohlich and Lean, 2004), historical volcanic forcing (Bradley, 1988), and aerosols (Charlson et al., 1992; Anderson et al., 2003).

18. The report appropriately notes that the radiosondes show an abrupt increase in temperature in the troposphere around 1976 and the fact that this is missed in the satellite data which starts in 1979. It has been argued that the surface warming is simply the response to this jump with a delay due the heat capacity of the ocean (Lindzen and Giannitsis, 2002). This is distinctly relevant to the present report.

Review of Chapter 2

Chapter 2 asks the following two questions: (1) what kinds of atmospheric temperature variations can the current observing systems measure, and (2) what are their strengths and limitations, both spatially and temporally? The chapter concludes that most observing systems are generally able to quantify well the magnitudes of temperature change associated with shorter time scales, such as diurnal and seasonal cycles, quasi-biennial oscillation (QBO), El Niño Southern Oscillation (ENSO), and volcanic eruptions. However, for longer time scale changes, where the magnitudes of change are smaller and the stability requirements more rigorous, the observing systems face significant challenges to document climate variations and trends with the accuracy and representativeness that allows attribution of change to human causes to be reliably identified. Therefore, many sources of errors in climate temperature data records must be identified and eliminated or significantly reduced.

This chapter did a reasonably good job summarizing the main observing systems for measuring surface and upper air temperatures and showing what kinds of atmospheric temperature variations these observing systems can measure using Table 2.2. The strengths and limitations of these observing systems (the second question), however, are not presented as well, mainly because of a lack of quantitative information and some redundancy between Chapters 2 and 4. The chapter contains a fairly lengthy discussion of statistical issues associated with measuring trends in time series, but it omits some key issues such as autocorrelation. This discussion also seems out of place in Chapter 2 since the main treatment of "uncertainty" is in Chapter 4. The statistical discussion should be strengthened and possibly moved elsewhere in the document. One possibility is a self-contained appendix devoted to trends in time series.

MAJOR COMMENTS

1. In Chapter 4, the following is asked: what is our understanding of the contribution made by observational or methodological uncertainties to the previously reported vertical differences in temperature trends? The nature of the discussion in Chapters 2 and 4 needs to be focused to reduce redundancy and avoid omissions in both Chapters 2 and 4. Chapter 2 appears to be focused on answering Chapter 4's question, rather than Chapter 2's questions. In general, some of the topical divisions between Chapters 2 and 4 are

artificial, so some redundancy in material presented is inevitable. However, the committee suggests that Chapter 2 focus on the various observing systems and Chapter 4 focus on trends in the observations, as differentiated in the following:

Chapter 2 should focus on:

a. explaining the measuring systems and instrumentation, their accuracy and precision, and spatial temporal variability for global measurements of temperature;
b. addressing measurement issues both for surface temperature measurements and atmospheric temperature measurements;
c. addressing spatial and temporal sampling errors; and
d. discussing any particular geographic regions where measurement and retrieval errors are particularly large.

Chapter 4 should focus on:

a. errors associated with trends; and
b. assessing which of the bias errors in Chapter 2 could influence the trends, and why they do or do not do so.

The discussion related to trend estimation and uncertainties in Chapter 2 should be moved to Chapter 4. Text on reanalysis trends from lines 266-277 should be moved to Section 7 ("Reanalysis") in Chapter 4. Also, Chapter 4 should add a section on "Methodological uncertainties" by including from Chapter 2 most of the text about linear trends in Section 2b (lines 385-460), discussions on structural uncertainty from lines 480-521, and the summary on "Errors or differences related to analysis or interpretation" from lines 584-601.

Alternatively, all material on trend estimation and uncertainties may be brought together in an appendix to the Temperature Trends report. In addition to the above material, discussion of statistical uncertainty in Chapter 3 (pages 39-40) could be included in the appendix.

2. Quantitative information is needed about the strengths and limitations of the observing systems. Specifically, quantitative discussion of the following sources of uncertainties should be included: accuracy and precision of the sensor, uncertainties in converting the fundamental measurement into temperature, and spatial and temporal sampling errors. There should be a summary of studies (with references) in which the different measurement types (e.g., radiosondes, active sensors, different satellite retrievals) have been intercompared and evaluated on a pixel level.

3. Increased discussion is needed on surface temperature measurements and trends, to parallel the detailed discussion provided on atmospheric temperatures. From reading this document, the impression is given that global surface temperature measurement is a solved problem, but this is not the case. Description of skin and bulk sea surface temperature (SST) in Chapter 4.5.1 should be moved to Chapter 2 and should reference the recent work of Chelton (2005). Errors associated with sea surface temperature measurement are not adequately covered in either Chapters 2 or 4. Discussion of microwave SST and blended infrared/microwave products should be included. Note, the

bulk SST is probably the suitable variable for trend estimation, but the skin SST values are needed to understand the variations in both bulk SST and atmospheric temperatures. Issues related to land-surface temperature measurement (skin versus screen) are not adequately addressed (see Jin and Dickinson, 2002).

4. Because we need to understand the processes contributing to the trends as well as measure the trends themselves, geographical regions having particularly large uncertainty should be addressed. For example, regional problems in surface temperature measurement should be discussed, including the Arctic Ocean and Southern Ocean, warm current regions, and the Indian Ocean.

5. Four and a half pages (pages 9-13) are devoted to "Reanalysis". Uncertainties in reanalysis trends are nicely summarized, and it is shown how the data are used by National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis and European Center for Medium-Range Weather Forecasts (ECMWF) Reanalysis (ERA-40). The conclusion is that there are considerable uncertainties in reanalysis trends, so reanalysis results are downplayed and not used in drawing conclusions in this report. The committee agrees with the authors' decision to deemphasize reanalysis data in the trend analyses in their report. However, this long discussion of reanalysis should be moved from Chapter 2 to page 19 in Chapter 3, where reanalysis temperature "data" are presented. Chapter 2 focuses on observing systems instead of particular datasets, and reanalysis products are not in fact "datasets". In addition, the four and a half pages of reanalysis discussions seem overly long in comparison to the approximately two pages for surface air temperatures and approximately four pages for upper air temperatures.

6. Scientific justifications for future observing systems are listed in Chapter 6, such as why we need reference radiosondes, but this is not mentioned in Chapter 2. For example, after "no absolute standards" in line 512, one sentence can be inserted to state that reference instruments are needed for future networks, such as the global reference radiosonde network proposed by the Global Climate Observing System (GCOS).

7. The report states that two main methods are widely used for calculating trends: linear regression and a "nonparametric" method attributed to Gilbert (1987). In the statistics literature, a technique is said to be "robust" if it is insensitive to violations of the underlying assumptions (the presence of outliers is one example of how underlying assumptions could be violated). In this sense, linear (least squares) regression is not robust, though it is not clear that this is an issue in any of the climatic time series under discussion. Gilbert's method does not seem to be widely used, but there are other methods (e.g., methods based on minimizing the sum of absolute deviations instead of squared deviations as in least squares) that have a large literature and should be referenced. These include the use of R functions to perform robust regression (Venables and Ripley, 2002), semi-parametric regression methods (Ruppert et al., 2003) and additive models (e.g., Hastie and Tibshirani, 1990). The distinction between least-squares and robust methods is not likely the main source of uncertainty in analyzing climatic time series. Non-linear trends are discussed on pages 18-19 of Chapter 2, though without making a clear-cut recommendation. It is self-evident that the trend is non-linear over any respectably long time interval, but nevertheless, fitting a linear trend could be the best thing to do if one is simply interested in coming up with one number to represent a trend over a stated time period. In the view of the committee, it is not unreasonable to use

linear trends in this kind of analysis, with two caveats: (i) it is important to remember that linear trends for different time periods will be different, and (ii) such linear trends should not be used for predicting future values. A further important issue is that when comparing observations and coupled model results, ENSO can appear in different sequences and magnitudes, making sampling a major issue. While linear removal of ENSO can ameliorate this problem, it is in fact impossible to remove all ENSO aspects even with multiple indices. As for ENSO (and similar) effects, the report discusses these on pages 18-19 of Chapter 2 but does not mention the most direct solution, that is, including ENSO (or other "natural variability" components) as additional covariates in the regression. There are arguments both for and against doing this, but comparing both analyses could be a useful reality check on the results.

8. The report barely mentions the issues of autocorrelation, i.e., the fact that correlations in time series could severely affect the estimation of a trend, especially in the calculation of standard error. Chapter 2 discusses error bars extensively without mentioning this issue. Chapter 3 mentions it tangentially, with discussion of error bars on lines 876-886 and a passing reference to the first-order autocorrelation in the captions of Tables 6.1 and 6.2, but with no details about the method. Given the importance of correct treatment of autocorrelation in the assessment of linear trends, this seems to be a major omission. The report should acknowledge that autocorrelation is a problem, as it is generally done incorrectly, and recommend how to properly account for its influence. The standard errors of estimated trends, allowing correctly for autocorrelation and other effects, are likely comparable to the "uncertainties" due to instrument shifts and effects of that nature quoted at numerous places in the report. This could lead to a quite different perspective on the relative importance of "structural" as opposed to simple statistical errors.

In fact, the method of Santer et al. (2000) seems to rely on the assumption that after subtracting trends, the time series is of AR1 form, which can indeed be characterized by the first-order autocorrelation. However, the AR1 assumption may not be correct and is certainly unnecessary as it is possible to fit a general ARMA (autoregressive, moving average) model with scarcely any more work. The "arima" function in the freely available R statistical package allows for fitting a linear regression component with ARMA errors, where the autoregressive and moving average components are of arbitrary order. The method is exact maximum likelihood, and standard errors are calculated for both the regression coefficients and the ARMA parameters. It should be noted that earlier versions of this method have been in use in the climatology literature for some time (Karl et al. 1996, 1998). Earlier discussions of time series approaches (e.g., including those based on fractional ARIMA models) have been given by Bloomfield (1992) and Bloomfield and Nychka (1992).

Another issue is whether to include an ENSO signal directly as a covariate in the analysis. In an analysis of annual hemispheric temperature averages, Smith et al. (2003) argued that inclusion of the Southern Oscillation Index as a covariate, though not having a great effect on the estimated trend, allows for specification of a lower-order AR model (AR1 rather than AR4) and in this sense simplifies the analysis. It would be of interest to see whether the same applies with the time series under discussion in this report.

Another method mentioned in Chapter 6 of the report is the adaption of methods from longitudinal data analysis (e.g. medical data in which individual subjects are followed for some period of time) a book by Diggle et al. (1996) is mentioned in this respect. While it is conceivable that these methods could be adapted to the estimation of trends in climatological time series, it also seems unnecessary, given that the AR/ARMA/ARIMA approach is quite well established. Therefore, discussion of this method should be omitted.

9. Direct discussions with the authors of the report made it clear that they had given more consideration to statistical assessment of trends in time series than is apparent in the written report, but neveretheless, it was the strong view of the committee that the issues should be dealt with explicitly in the report. Based on the overall structure of the document, such discussion would logically belong with the "uncertainty" discussion in Chapter 4 rather than Chapter 2 but the authors might alternatively consider writing a separate appendix on the statistical issues associated with estimating trends in climatic time series.

10. There are insufficient bibliographic references to the technical aspects of temperature measurements and error determination and far too many references associated with climate variability and trends (these are more suitable to other chapters). Recent references (since NRC, 2000) should especially be included.

11. Cross evaluation and intercomparison of different technologies (including surface-based remote sensing) to measure temperature should be described.

SPECIFIC COMMENTS

1. Observations not used in this report should be mentioned**,** such as why Television Infrared Observation Satellite Program (TIROS) Operational Vertical Sounder/Infrared (TOVS/IR) was never used for trend analysis. This probably should be mentioned after line 193. A discussion of TOVS temperature profiles is needed. Note, the tuned regression type analysis used by the National Oceanic and Atmospheric Administration (NOAA) is not the only temperature available from TOVS. The Pathfinder effort and French 3I effort represent research-quality retrievals. John Bates at NOAA is in the process of doing a careful calibration of TOVS so that trends can be determined.

2. Table 2.2 is used to answer the first question of this chapter but provides insufficient emphasis on the long-term temperature changes due to anthropogenic effects (i.e., temperature trends), which is the sole focus of this report. It would be useful to add one column to list the "Outstanding issues" regarding specific variation, which includes inconsistencies among different datasets (or observing systems) and what future data are needed for better characterizing and understanding this variation. The column "Effect on trend estimates" needs more quantitative information if available, such as how much the temperature trends change before and after removing ENSO signals in the time series.

3. It appears that Table 2.1 and the text on pages 22-24 were used to try to answer the second question for this chapter. The information given here is too general and too qualitative. More quantitative information and some references should be given on pages 22-24. For example, the authors can summarize how insufficient spatial sampling of the radiosonde network affects the temperature trend from Agudelo and Curry (2004) and others. The authors can give more information about radiosonde errors in the upper troposphere and lower stratosphere—such as radiation errors, their magnitudes and

characteristics—errors in existing radiation corrections and how they affect the trends. Table 2.1 should include specific instruments and pixel size for satellite measurements. Humidity and wind measurements should be excluded from the table, although the authors may want to discuss how these measurements can be useful proxy diagnostics if measured carefully with climate-quality monitoring.

4. In lines 88-89, near-surface air temperatures over land are measured about 1.5-2 m above the ground level at official weather stations, rather than 1.5 m.
5. A reference is needed for this statement.
6. The caption for Figure 2.2 should state that the pressure levels at the y-axis are radiosonde "mandatory reporting levels".
7. In lines 217-226, the reference for Global Positioning System-Radio Occultation (GPS-RO) is Kursinski et al. (1997). The comparison between GPS-RO and radiosonde data has shown that the GPS-RO soundings are of sufficiently high accuracy to differentiate performance among the various radiosonde types (Kuo et al., 2005). Also, the report should discuss the findings of Schroder et al (2003) on MSU versus GPS. In particular, Schroder et al. (2003) found that UAH T4 retrievals in the Arctic lower stratosphere in winter were biased relative to temperatures derived from GPS Radio Occultation measurements.
8. The statement "the method of calibrating a radiosonde before launch may introduce time-varying biases" in lines 543-544 needs clarification.
9. The references at the end of this review include several additional papers that should be considered for inclusion in Chapter 2 of the Temperature Trends report.

Review of Chapter 3

Chapter 3 asks the following: what do observations indicate about the changes of temperature in the atmosphere and at the surface since the advent of measuring temperatures vertically? This chapter describes observed temperature trends for three classes of data sets: surface observations (mainly in situ but including some satellite data), radiosonde measurements, and microwave sounding unit (MSU) observations for various levels in the atmosphere. Linear trends were computed for two periods: 1958-2004 for the surface and radiosonde data and 1979-2002 for all three data sets. The analysis includes three compilations of surface data, two compilations of radiosonde data, and three analyses of MSU data. A fourth data set, reanalyses products, are discussed but downplayed.

In general, the discussion is comprehensive and in most cases reflects an accurate interpretation of the results. The discussion of the three global surface temperature trend analyses is readable and yet discusses some of the important details. The general conclusion of a consistent warming signal at the surface seems well justified. Consistency between the two analyses of radiosonde data, for 1958-2004, is encouraging. The question to be addressed by the chapter is answered. However, the chapter needs to clarify a few points.

MAJOR COMMENTS

1. A major issue is the drop in temperature associated with the introduction of the Vaisala sonde. It is stated that this affects the stratosphere, but it is unclear how deeply this systematic bias might extend into the troposphere. This is an important research problem that should be addressed.
2. Mentioning the similarity of the basic data in the surface dataset while highlighting many of the potential problems (almost all of which have been adequately handled) sows doubts in the minds of readers. This gets picked up and emphasized in Chapter 6. The report should provide more explanation of how various problems in the data have been addressed and how this leads to some level of confidence in examining trends. For example, it can be easily shown by sub-sampling the surface data that the resulting hemispheric and global trends from the sub-samples would be almost exactly

the same. References should be made to the frozen grid analyses work done in the Climate Research Unit (CRU) in the mid-1980s to the mid-1990s.

3. In general, the chapter would benefit from a more careful dissection of the global mean and a recognition that radiosondes are not near global. It does not address global mapping or the need for evaluations at co-located sites of sondes (see Hurrell et al., 2000; Agudelo and Curry, 2004; and Free and Seidel, 2005). There are no latitude-time series presented. This chapter does identify differences over high latitude land as being the main reason for surface being larger than troposphere trends in the extratropics (page 30), but this is not carried forward to the Executive Summary or Chapter 5. Weakening or removal of inversions over cold land or ice is a very good reason why the surface should warm more and a good example of why the global mean should be dissected.

4. The chapter also places too much emphasis on linear trends. Only linear trends as a function of latitude are presented, however this presentation can hide many things. The claimed agreement between radiosondes is not shown except for the linear trends (e.g., Table 3.6.1) (see Free and Seidel, 2005). There is nothing on root mean square differences, which are very revealing (Hurrell et al., 2000), or on monthly differences (smoothing the time series can be misleading).

5. In a number of places, the assertion is made that the troposphere has warmed more than the surface. However, the differences in trends are often quite small, particularly for the 1958-2004 period. It is not clear that these differences are statistically significant. Although statistical significance is assessed for the trends themselves, no analysis of the significance of trend differences is presented. When comparing trend differences between two estimates of the same quantity (e.g., tropospheric temperatures from radiosondes and satellites), it is more appropriate to examine the trend of the difference time series, rather than trends for each time series individually (because the data contain similar overall variability). This is an omission that should be corrected, and the text should reflect the results of such an analysis. In particular, any statement about differences in warming should be weakened considerably if the differences are not statistically significant.

6. The trends calculated from reanalyses are downplayed because the input data sets are not homogenized. Although there is potential independent value of reanalysis products, it is not clear what trends from a reanalysis model mean in the context of temporally varying inputs. Therefore, the committee agrees with the decision of the authors to downplay this source of information.

7. The issue of regional land-use and land-cover changes is brought up in a number of places, but the implications are not clearly addressed. For example, in lines 94-96 it is suggested that regional land-use change must be considered in the development of land-based data sets. However, if regional changes are large enough to have a measurable influence on global temperature, then these changes will be sampled and detected by the existing land-based networks. As such, why is this an issue when analyzing the differences among the data sets? There is an issue related to land-use and land-cover changes that could be addressed here or in other chapters. In the modeling discussions in Chapters 1, 5, and 6, land-use and land-cover is considered to be a forcing (with uncertain magnitude in the past) that is incorporated in some models and not in others. The committee believes this is correct and that land-use and land-cover should be considered as a forcing. Any land-use and land-cover effects in observational datasets should

therefore be left in and not commented upon as a problem in Chapter 6. In other words, Chapter 6 cannot have it both ways the data are affected by land-use and land-cover change, so they are somehow wrong, yet this forcing is omitted from many models.

8. The Fu et al. results have the potential to be centrally important to the issue of tropospheric temperature trends and should be discussed more thoroughly in lines 863-868. Attempts to separate tropospheric and stratospheric contributions to trends are reasonable. They should not be rejected with the value statement that they are "controversial". The only published criticism of the Fu et al. approach is by Tett and Thorne (2004), with other criticisms in the grey literature. The Fu et al. method has since been followed up by several studies which show that it is robust, including further research by Fu and colleagues and Gillett et al. (2004). The potential clarification that the Fu et al. method can contribute to the central issues is very significant.

9. The difference between the Remote Sensing Systems (RSS) and University of Alabama, Huntsville (UAH) trends is left as an open issue, with no relative value given. It is important to resolve this discrepancy, if possible. The trend difference in the mid-troposphere is the same size as the signal: zero for UAH and +0.1 K/decade for RSS. If no distinction can be made, then no conclusion can be drawn. Statements in lines 355-359 and elsewhere about discrepancies between RSS and UAH as being mostly due to the NOAA 9 satellite are misleading as can be seen by looking solely at the post 1987 period. In fact, examining differences between the two datasets, which are not shown in the report, reveals major issues remaining on adjustments for other satellites and diurnal cycle issues (especially as a function of latitude and in the tropics).

10. In lines 791-816, if the tropical tropospheric temperature profile behaves as a moist adiabat, which to an approximation it does, then the lapse rate is expected to decrease as temperature increases (i.e., as the surface warms, the troposphere is expected to warm more). This is the "global change theory" the authors refer to in Section 6.2.1. Therefore, it is no surprise that when the surface warms due to ENSO, the troposphere warms relative to the surface (line 798), or that when the atmosphere warmed in 1976-1977, the lapse rate dropped (line 802). These results are currently presented with no link to physical theory. The authors say that "the variation in tropical lapse rate can be characterized as highly complex, with rapid swings over a few years, superimposed on persistent periods of a decade or more", but our guess is that much of this variation can be explained by changes in the mean temperature. Further, the authors say that the enhanced warming of the troposphere associated with surface warming gives "enhanced static stability" (lines 799 and 803). A reference should be provided for this statement. It should be noted that the troposphere did warm relative to the surface in the tropics during the 1997-98 El Niño event, which is a large signal. Also, the report should reference a study by Gettelman et al. (2002) on changes in stability. This study highlights the observed increases in Convective Available Potential Energy (CAPE) that are not replicated by models (which remove all CAPE), and so it is also relevant to Chapter 5 of the report.

SPECIFIC COMMENTS

1. The numerical system for numbering the figures is overly complicated and inconsistent. It would be simpler to number the figures 3.1, 3.2, 3.3 etc., rather than 2.4, 3.3, 4.4, 6.2, 6.2.2, 6.2.3, 7.1. In all the figures, the notations used to label the curves in the diagrams are different from the descriptions in the captions. For example, Figure 2.4 has the labels N, G, and U, and these are not defined in the caption. The same is true in different ways for 3.3, 4.4, etc. Also, without a very good color print, the different colored lines can be difficult to distinguish.
2. In lines 53-55, comparing results from more than one dataset also provides a better idea of the uncertainties or at least the range of results.
3. In lines 86-88, the statement that homogenization procedures are "quite successful" at addressing these issues should be more nuanced. While we are in agreement with the statement with regard to biases introduced by changes in time of observation, we are less confident that other issues (e.g., exposure changes) can so readily be addressed because there is often a lack of metadata.
4. In lines 107-111, the benefits of sea surface temperature (SST) over night marine air temperature (NMAT) are discussed without saying anything about what the relationship between SST and NMAT is likely to be (e.g., is SST a good proxy for NMAT?).
5. There should be a reference in line 163 to Jones et al. (1997, 2001). These papers give details of the procedure for allowing for changing numbers of observations through time.
6. This text in lines 180-183 is a bit wordy and does not follow on well from the previous sentence. The paper by Vose et al. (2005) should show that the differing techniques with the same data produce almost the same results.
7. In line 205, the text "since neither choice is optimal" suggests that there is a single optimal approach. This should be rephrased to "since each approach has advantages and disadvantages."
8. In lines 229-231, the Radiosonde Atmospheric Temperature Product for Assessing Climate (RATPAC) data set incorporates different homogeneity adjustments before and after 1997. Has anyone evaluated the extent to which this might introduce an inhomogeneity into this data set?
9. Lines 294-296 state, "There is some ambiguity about whether the temperatures return to their earlier values or whether they experience step-like falls". Surely this is just a matter of how best to describe the curves. A more important question is whether the observations agree with particular models (global circulation models or theoretical models). Has anyone suggested a plausible mechanism that would give a step-like cooling after a volcano (e.g., Douglass and Knox, 2005)?
10. In lines 297-299, is the interannual variability really mainly due to the Quasi-Biennial Oscillation (QBO)? If so, a reference should be provided.
11. The text in lines 299-301 makes it sound like the stratospheric cooling trend has been completely explained as a combination of the responses to stratospheric ozone depletion and cooling due to carbon dioxide. This is in disagreement with the Executive Summary, which indicates that the cooling cannot be fully explained by these forcings.

12. In lines 301-305, there are various descriptions of the curve including "the aforementioned step-like drops represent a viable alternative to a linear decrease". What do the authors mean by "a viable alternative"? Presumably, they do not mean one based on a physically-plausible mechanism. Again, this seems to just be a discussion of how best to describe the curve, whereas the real issue is whether the observations agree with theoretical predictions.

13. In lines 320-323, the change to the Vaisala radiosonde in certain tropical areas is given as a possible reason for the differences in the two radiosonde data sets. What analysis has been done to suggest this possibility? Or is this statement made simply because the timing is coincidental?

14. The nomenclature of TMid-Trop-R and TMid-Trop-A are introduced without definition in line 358. At least a reference to Chapter 2, Figure 2.2 and related discussion should be included for those who may start reading here. Are these just the Microwave Sounding Unit (MSU) channels and their radiosonde integral equivalents, or something else? Also, the nomenclature in the figures and captions is inconsistent and not sufficiently defined.

15. TTrop-UW-A is introduced without definition in line 396.

16. In line 447, the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalyses go back to 1948. It is probably best to ignore the period between 1948-57 as this study only goes back to 1958.

17. A reference to Simmons et al. (2004) might be needed in line 475, or a reference back to Chapter 2.

18. The Pielke and Chase (2004) reference in line 488 is missing from the reference list.

19. It is not completely clear what is meant in lines 502-505. Presumably this relates to the abrupt change in the late 1970s.

20. What do the authors mean by "it has been shown that such constructs are plausible" in lines 505-508? What criteria are used to judge their plausibility? Presumably it is just how well they fit the data. In this case, you could make a perfect fit to the data by regressing it on itself. Again, the real issue is whether the observations fit with theoretical predictions.

21. It is unclear what is meant in lines 515-518.

22. In lines 521-527, The reanalysis models tend to agree better with the climate model predictions than do the raw observations. Is this an alternate explanation of the differences between the reanalyses and the raw observations (i.e., if the reanalysis model has similar physics to the climate models, then its troposphere will warm more than its surface)?

23. It is not entirely clear what is meant in lines 542-544. Do the authors mean something like "trends in land air temperature in coastal regions are generally consistent with trends in SST over neighboring ocean areas"?

24. In lines 548-552, the authors do not mention the most obvious explanation for enhanced warming over land, namely the smaller effective heat capacity over land than ocean. Enhanced warming over land is seen in every climate change simulation and does not relate primarily to the phase of ENSO, though this could be a contributor. Better justification should be provided for a link between warmer temperature over Siberia and ENSO. Siberia encompasses a large area, so be more specific and provide a reference.

25. In lines 561-563, SSTs and NMAT have different trends for short periods owing to ENSO and changes in surface fluxes, as shown in other works.

26. In lines 568-570, these differences might be related to an increase in mean ship height above the sea surface.

27. While the explanation in lines 589-591 sounds plausible, has it ever actually been shown? Are the free tropospheric temperatures more highly correlated with maximum surface temperatures than with minimum temperatures? If there is not a published reference on this, then should be removed.

28. In lines 600-601, it is not clear whether this relative change in trend in the troposphere and surface is statistically significant in the recent era. Visually, it does not seem that impressive or obvious.

29. The comparison the authors make in lines 627-629 is equivalent to assuming that the Maryland stratospheric trend is the same as that in the other two datasets (since the Fu et al. approach is just to fit to a regression model).

30. In lines 655-656, "TLow-Strat-A" and "TLow-Strat-B" need definitions.

31. In lines 656-657, why is the cooling at the South Pole not more dramatic, especially given known problems over sea ice (Swanson, 2003) and the high ice sheet of Antarctica that greatly impacts channel 2? In fact, it looks like the cooling is larger in the northern hemisphere midlatitudes.

32. Replace "Soviet stations" in line 672 with "stations located in Russia and other countries of the former Soviet Union".

33. The word "granularity" should be replaced in line 693.

34. In line 696, replace "noisy patterns that result" with "noise that results".

35. The figure labeling (a, b, c and d) in line 704 is incorrect.

36. No mention is made of the Antarctic in line 713.

37. In lines 722-724, the sharp contrasts only seem to be around the western coasts of the Americas.

38. The unit for a lapse rate trend looks wrong in line 823. Surely it should be K km^{-1} $decade^{-1}$ or something with the same dimensions.

39. Are there missing crosses for the surface in Figure 6.2b, or do they all overlap?

Review of Chapter 4

Chapter 4 raises the following question: what is our understanding of the contribution made by observational or methodological uncertainties to the previously reported vertical differences in temperature trends? The chapter gives separate discussion to three main layers of the atmosphere (stratosphere, mid/upper troposphere, and lower troposphere) and to surface data on both sea and land. For the atmospheric data, there is separate discussion of radiosonde and satellite data, although less attention is given to radiosonde data because of the well-known historical bias in radiosonde stratosphere measurements (which also affects averages computed for the troposphere). Much of the discussion is devoted to the conflict between the University of Alabama, Huntsville (UAH) and Remote Sensing Systems (RSS) reconstructions of the middle and upper troposphere, with particular emphasis on the consequences of different calibration corrections the two groups make for one satellite (NOAA-9). The possibility of combining Microwave Sounding Units (MSU) channels (work by Fu et al.) is mentioned, though dismissed as controversial. The chapter also mentions, but similarly dismisses, the use of reanalyses. The last part of the chapter is about the different sources of bias in surface data.

As discussed previously, there is some overlap between Chapters 2 and 4. The report would benefit from some reorganization to more clearly focus the two chapters; a suggestion for what material belongs in which chapter is provided in the committee’s comments on Chapter 2. The emphasis throughout is on “structural” sources of uncertainty, as opposed to statistical uncertainty which has been dealt with in Chapter 2 (for reasons that are not quite clear because it would seem more logical to deal with all the sources of uncertainty together). Within its thus-defined scope, the chapter does a generally good job of describing the different sources of bias in temperature reconstructions. The following comments are aimed at improving the chapter discussion.

MAJOR COMMENTS

1. There should be better discussion of the Fu et al. approach. Simply stating that this is controversial is a value judgment and not an adequate reason for dismissing the approach. The review panel sensed that some of the authors had more specific objections to the approach, but these are not adequately documented. For example, why should it be

a problem that the approach uses negative weights for part of the signal? The ultimate goal here is to eliminate or reduce stratospheric contributions to middle troposphere temperature trends. As described in Line 184, about 10 percent of the weight of Channel 2 comes from the stratosphere, but the integrated weight for Fu et al. weighting function is near zero. As stated on Line 187, the stratospheric contamination on $T_{\text{Mid-Trop}}$ trends is about 0.05 K/decade, while the trend uncertainty due to the uncertainty in derived coefficients in the Fu et al. method is only about 0.01 K/decade. The potential for incorrect stratospheric temperatures to corrupt the mid-tropospheric values should receive greater emphasis in Chapter 4 and Chapter 5. In conclusion, the Fu et al. method appears to reduce stratospheric contributions and may represent a valuable resource for this report. The report could, if appropriate, include references to more recent work of Fu et al. and possibly other authors. The new papers might give more insights on controversial issues of negative weights in the Fu et al. method and its impacts on trends. The Fu et al. (2004) reference on line 487 is missing in the chapter's references.

2. The report gives a very even-handed discussion of the reasons for different trend estimates by UAH and RSS. Is there any way to go further, for example by stating which approach is better or proposing ways to reconcile the two approaches? Would the authors recommend further statistical analyses? If so, what form should these take? It appears to be the case that, although issues with diurnal corrections and the calibration target are important, these are not the major reason why the two groups obtain different trend estimates. These differences appear to hinge on the different treatments of the NOAA-9 satellite. It may be possible to do better using Bayesian statistical methods. For example, one could treat the unknown shift of the time series (resulting from the change of satellites) as a parameter with a prior distribution, construct a posterior distribution by analyzing data from both satellites, and then integrate out that posterior distribution using Monte Carlo methods to derive a reconstructed time series that allows for uncertainty in the shift. This method could potentially work better than current methods when there is only a very limited amount of overlapping data. In lines 294-300, the authors use the lack of a diurnal correction in the University of Maryland (UMD) dataset as an excuse for not discussing it. Because of the differences between UAH and RSS and the small residual uncertainty from diurnal sampling, it could be informative to use the UMD dataset as an independent check to understand and possibly reconcile the differences between UAH and RSS. The suggestion that the correction for target temperature is a function of latitude (or orbit relative to the Sun), as done by the UMD group (Grody et al., 2004), but not by UAH and RSS, is an interesting one and builds in some diurnal cycle corrections. These issues ought to be discussed openly.

3. Overall, the satellite uncertainty is summarized in detail and in depth, while the radiosonde uncertainty is described in less detail and less quantitatively (see below for more detailed comments). There is no discussion of the strengths and weaknesses of homogenized methods used by different dataset groups. There is a lack of attention to developing physical-based correction schemes. For example, radiosonde radiation error is the main source of errors for upper troposphere and lower stratosphere temperatures. it appears that none of the groups has implemented radiation corrections to non-corrected historical data or adjusted applied corrections. It is true that the trend analysis relies more on long-term homogeneity than on the absolute accuracy. But accurate data throughout the period would minimize the temporal inhomogeneity and can be used for other studies.

Also, the report has no discussion of missing data within a month for radiosonde data. In Hadley Center Radiosonde Temperature (HadAT) only 12 soundings are required to make a monthly mean and two monthly means to make a season; there is no allowance for this in the error bars. Missing months are especially an issue in the tropics, where records are woefully incomplete, as shown by Hurrell et al. (2000). Free and Seidel (2005), however, find missing monthly data to have a fairly minor effect on trends.

4. There is no discussion of statistical uncertainties in methodologies for calculating trends, calculating monthly mean values and creating global time series (i.e., spatial averaging techniques for radiosonde data). Some of this discussion appears instead in Chapter 2. Somewhere (Chapter 2, Chapter 4, or a separate appendix) there should be a separate section on statistical methods for estimating trends in time series, including standard errors or other measures of statistical uncertainty.

5. The largest discrepancy between radiosondes and satellite estimates of trends is in the stratosphere. More detailed discussions on the stratosphere discrepancy are needed in Section 2. Section 2.1 briefly describes two uncertainties associated with undetected changes in instrumentation and early bursting of balloons in early radiosondes. There can be significant biases in the radiosonde temperature data in the stratosphere due to radiation errors. Both radiosonde datasets do not include physical models for radiation adjustments. Durre et al. (2002) show that Luers and Eskridge (1998) adjustments make radiosonde temperatures more homogeneous in the stratosphere, although it frequently amplifies the discontinuities in the troposphere. Regarding the statement "The discrepancy … is likely to be mostly due to pervasive uncorrected biases in the radiosonde measurements" on lines 96-98, can the authors be more specific about what those uncorrected biases are? What about time lag errors of radiosonde data that could cause a cold bias in the stratosphere? There are minimal discussions on the largest disparities in the tropics between two radiosonde datasets and between radiosonde and satellite data in Figure 6.2.2. in Chapter 3. How does the difference in station distributions between these two radiosondes contribute to the largest discrepancy in the tropics? Is the enhanced cooling in the tropics relative to the midlatitudes in the stratosphere in radiosonde datasets due to a lack of sampling over open oceans, or is it due to larger adjustments associated with the switch to Vaisala radiosondes for most of tropical stations? It seems that the former has minor impacts because Figure 6.2.3 in Chapter 3 shows that the stratosphere trends in the tropics are zonally uniform.

6. It seems that the difference in homogeneity adjustment methods is the main contributor to disagreements in trends among different radiosonde datasets presented in Sections 2.1 and 3.1. Do the adjustments reduce or increase the discrepancies in trends (by comparing the trends before and after the adjustments)?

SPECIFIC COMMENTS

1. In lines 176-193, does the bias in the radiosonde-derived T_{Mid_Trop} from stratospheric errors have the same magnitudes of about 0.05 K/decade for NOAA and UK-Met datasets? As shown in the middle panel of Figure 6.2.2 in Chapter 3, the difference between $T_{Mid\text{-}Trop\text{-}U}$ and $T_{Mid\text{-}Trop\text{-}N}$ at around 5°N is about 0.1 K/decade. Adding

~0.05 K/decade to both datasets still cannot explain the large disparity between two datasets at this latitude.

2. In lines 335-347, how can the uncertainty of the lower troposphere temperature record be consistent with the mid-troposphere uncertainty, especially given that the mid tropospheric record is biased low by contaminating lower stratosphere influences?
3. Section 4.3 fails to examine root mean square (RMS) differences (e.g., Hurrell et al., 2000) and only deals with average trends.
4. The surface record also has problems that are not discussed in Section 5.1 of Chapter 4. In particular, no error bars are assigned to the systematic corrections.

Review of Chapter 5

Chapter 5 asks the following: how well can the observed vertical temperature changes be reconciled with our understanding of the causes of these changes? This chapter aims to explain the different observational surface and tropospheric temperature trends through using state-of-the-art modeling results, principally from integrations that include multiple climate forcing factors. Most of the model simulations analyzed are relatively new, using model integrations performed for the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4) assessment.

Overall the committee liked this chapter. It is the clearest and most lucid of all of the chapters. However, there are several important issues that should be addressed, especially related to the correct use of statistical uncertainties and comparisons with satellite data. This chapter has copious footnotes unlike all the others. This approach seems better for the presumed audience and is therefore recommended for all other chapters. Thus, sufficient detail can be presented without damaging the flow for the more general reader.

MAJOR COMMENTS

1. The conclusions reached are often based on estimates of trends, neglecting uncertainty levels, and many statements on comparison are inaccurate because of this. The report should be more explicit about the choices made regarding the treatment of trend confidence intervals in model-data comparisons. If the authors believe that including error bars could hide model-data discrepancies, inadequate understanding of model uncertainties, or both, then this view should be discussed, possibly within the first conclusion or earlier. The second and third conclusions regarding the influence of volcanoes and El Niño Southern Oscillation (ENSO) could be removed and only briefly mentioned earlier. The volcano aspect could be included as an example of the single forcing whose signal can be detected in observational temperatures in the first concluding point. There should also be discussion of volcanoes in the context of Douglass and Knox (2005) and Lindzen and Giannistis (1998, 2002).
2. Error bars are essential on the plots, notably Figures 5.3 and 5.4, and all dots should be horizontal bars to allow for sampling uncertainty. This is important because ENSO is not in the same sequence in the coupled models. This is partly discussed in lines

641- 659, but model simulations cannot be definitive given the exceptional nature of the 1997-98 event. Even in Figure 5.7 only a single set of error bars is given.

3. The chapter notes the importance of the stratospheric contribution to the channel 2 temperatures and refers to Fu et al. in lines 580-583, but then never allows for this in subsequent comparisons. As a result, Figures 5.2B, 5.3, 5.4, and 5.5 and discussion are all misleading because the models clearly have different cooling in the stratosphere; discussions in Chapter 4 suggest that this accounts for 0.05 K/decade trend in channel 2 discrepancy. Several parts of the text ought to be substantially revised as a result of this (including lines 614-624, 631-633, and others).
4. Regarding the focus on the global means and some zonal means, regional trends differ a lot from global values (Agudelo and Curry, 2004). For instance, the large increase in surface temperature over northern land and the smaller decrease in the troposphere, which is related to changes in surface inversions (Chapters 1 and 3), are not examined in the models and not picked up. The chapter comes closest with Figure 5.5, but that fails to account for the stratospheric contamination. The fact that sondes are not global is also not dealt with. Subsampling of the modeling data at sonde locations is not done.
5. There should be more explicit discussions of the specific responses to individual forcings and how these combine together. This could be done in the first of the conclusions that have "some confidence"—see details towards the end of the specific comments. There should also be a discussion of the use of multiple regional forcings in models.
6. In the presentation by B. Santer during the February 23, 2005 NRC meeting (Chicago, Illinois), the committee liked the two model plots (of standard deviations and trends at the surface and at low and middle troposphere) and would hope that these can be included in a revised chapter. Also included should be as many results from additional models as time allows.
7. The committee had some discussion of how the basic methodology of "Detection and Attribution" should be presented in the report. What is needed is not a full mathematical description of the method (for that one can refer to the original source papers) but a discussion of the main principles behind the methodology that would be appropriate for a climate scientist who does not work directly in this area of research. There needs to be better understanding of the strengths and limitations of detection and attribution analyses. What follows is a tentative suggestion of how to do this. In addition, the authors may find the work of Levine and Berliner (1999) useful in revising this discussion.

> Detection and attribution methods try to represent an observed climatic data set in terms of signals due to forcing factors such as greenhouse gases, aerosols and solar fluctuations, plus correlated random noise. The methods are also called "fingerprint analysis" because it is possible to think of the method as identifying specific fingerprints (spatial patterns of climate change due to specific forcing factors) in the observational climate record. The climate data typically consist of temperature or rainfall averages over grid boxes and are very high-dimensional.

There are two versions of the method, one developed by Santer et al. (1995, 1996) based on estimating pattern correlations between observational data and fingerprints and the other developed primarily by Hegerl and Allen and their co-authors, (Allen and Tett, 1999; Hegerl and Allen, 2002; Allen and Stott, 2003; Allen et al., 2004), which uses regression analysis to decompose climate data as a linear combination of forcing factors, plus correlated noise. If the regression coefficient due to a forcing factor is statistically significantly different from zero, then we can claim to have "detected" that factor in the climate record. "Attribution" refers to the process of attributing the observed climate change to the different forcing factors. The two approaches are mathematically equivalent, though they differ in specific details because of different implementation decisions made by the two groups.

A critical feature to both versions of the method is how to estimate the spatial correlation structure of the unforced internal variability. Standard techniques—such as estimating the correlation between each pair of grid boxes from the observational data and assembling the resulting pair-wise correlations into a correlation matrix—fail completely because the number of data vectors available for estimating the correlation matrix is so much smaller than the dimension of the data. Therefore, an indirect approach is used. Samples, typically of around 1,000 years in length, are generated from control runs of the climate model where all forcing factors are kept constant. The covariance matrix is estimated from these control runs together with an orthogonal decomposition to reduce dimension (a technique variously known as Empirical Orthogonal Function (EOF) analysis, principal components analysis, or Karhunen-Loève expansion). Typically around 10-15 orthogonal components are used. Based on this decomposition, it is possible to greatly reduce the dimension of the original data and hence to estimate the pattern correlations or regression coefficients, with realistic approximations to the sampling distributions of those estimates.

Some potential difficulties with the methodology should be noted. There are various technical issues such as how many orthogonal components to choose (or the broader question of how well covariances calculated from control model runs represent those in real data). The method does assume that the signals or fingerprints are known—Myles Allen has proposed an extension that allows for random error in the signals but this makes the analysis much more complicated and it is not clear that it works well in this situation. As a result, the methodology is probably not appropriate for incorporating climate change effects where there is large uncertainty about the signal (one might argue that land-use and lan-cover effects fall in this category). Also, the methodology should probably not be used with too many different signal components; most successful applications have used just the major signals mentioned above (greenhouse gases, aerosols and solar fluctuations, possibly including also volcanic forcings, but always bearing in mind that some of these forcings are not well known and have large error estimates). It is a feature with any regression analysis that including too many collinear regressors lowers the precision of estimated regression coefficients, and this aspect is only made worse by the difficulties associated with estimating covariances.

The method assumes that the response to a combination of different forcing factors is a linear combination of the responses to the individual forcing factors (a property that statisticians call additivity). In principle one could get around this assumption by running, for example, climate models under different combinations of forcing factors and using these combined signals in the detection and attribution analysis. This is not done because of the computational expense of obtaining such multiple model runs and the statistical difficulties just mentioned of using detection and attribution techniques to select among a large number of possible model-based signals. The assumption of linearity in response should also be evaluated, as many of the forcings are not orthogonal to each other.

In summary, detection and attribution methods are an extremely powerful technique. They are essentially the only method available for formally analyzing the agreement between climate models and observational data. However, they cannot be expected to do everything. In particular, they cannot be expected to detect a poorly defined signal or to discriminate among a very large number of possible signals that might represent different explanations of climate change.

SPECIFIC COMMENTS

1. The word "lockstep" in line 52 should be replaced with "evolve together" or "in unison".
2. References should be provided for the differences of opinion discussed in lines 93-94.
3. In lines 108-109, while the model will simulate similar ENSO to the real world when run in Atmospheric Model Intercomparison Project (AMIP) mode, this is much smaller for North Atlantic Oscillation (NAO), which is not ocean forced.
4. The sentence in lines 112-113 could be expanded to be more informative.
5. In lines 122-123, the reason for using ensemble forecasts is not only because the full state of the climate system is not known. Producing a deterministic forecast of the climate would also require a perfect model.
6. Add regional aspects in footnote 11?
7. Evidence of the 0.3ºC cooling since the 1970s over India should be provided in line 216. This cooling is not evident in the maps the IPCC AR4 will use for 1979-2004. Over this timeframe most of India shows warming.
8. The urban heat island should be mentioned in lines 233-237 because it is a major effect in urban areas and of opposite sign to rural land-use and land-cover effects.
9. In line 237, Matthews et al. (2003) is not in the reference list (and it should actually be Matthews et al., 2004.). The cooling is small in global terms, but it might be large regionally, although it is hard to pin down because the noise levels are much higher.

In general, regional changes on a scale smaller than the Rossby radius tend to be confined to the boundary layer. Also, Table 1 of Chapter 1 says land-use change effects are small.

10. Another reference relevant to lines 281-282 is Jones (1994).
11. Are the very small estimates of error reported in footnote 18 still believed?
12. Lines 320-321 state, "volcanic effects probably contribute to slow changes in lapse rate variability". Do the authors mean changes in lapse rate variability, or changes in lapse rate here?
13. In footnote 21, can the HC/CRU surface data be referred to as HadCRUT2v and have the Jones et al. (2001) reference? This should be done elsewhere in the other chapters and can be in one of the footnotes.
14. Section 4.3 would benefit from more synthesis and assessment, rather than just reporting of results.
15. Which climate forcings does line 357 refer to?
16. In line 397, "various datasets" should be "various models".
17. The IDAG reference in line 411 is missing.
18. In lines 424-425, are there also different variables than temperature? Are there some more recent references regarding pressure detection? There could be a better link to the preceding paragraph.
19. In lines 428-432, positive detection results obtained in the absence of some forcing should not be taken as evidence of absence of that forcing. The same argument could be made about volcanic forcing, solar forcing, or even sulfate forcing, yet we know they have had an effect on the climate. This sentence should certainly be deleted. (In fact four pages later, the authors themselves argue that our inability to detect sulfate in some studies should not be taken as evidence of absence of a sulfate signal). This also reflects the potential non-orthogonality of the various forcings.
20. "This apparent contradiction" in line 433 is not really a contradiction, for the reasons given above.
21. What is the relevance of the final sentence in lines 498-500?
22. In lines 502-509, the difficulty of detecting the sulfate response could also be explained by degeneracy between the sulfate and greenhouse gas response patterns. The inability of some studies to detect the sulfate response should not be taken as evidence of absence of a sulfate signal, but at the same time, by itself it does not show that "it is important for detection work to account for large temporal changes in the fingerprint pattern".
23. None of the model runs for IPCC AR4 mentioned in lines 541-542 have been written up yet. As long as the references detailing the new integrations are submitted this is not a problem.
24. HadCM3 has also run all the experiments discussed in lines 550-551. Is it possible to include it, or are you hoping to use HadGEM?
25. In lines 553-554, the use of different forcings in the different IPCC models is sometimes presented as an advantage, since it folds in some approximation of forcing uncertainty into the analysis.
26. In lines 591-592, insert "partly" before "due". As far as we are aware, no one has shown that water vapor changes can explain the full difference between simulated and observed trends.
27. Insert "Body" before "temperature" at the start of line 782.

28. Line 790 should also refer to Robock and Oppenheimer (2003), which looks more at circulation patterns. There is a paper in that book by Jones et al. (2003).

29. Lines 800-802 states, "At constant relative humidity, water vapor is expected to increase nonlinearly with temperature (Soden et al., 2002)." Water vapor does increase nonlinearly with temperature at constant relative humidity. This is just the Clausius-Clapeyron equation. If a reference is cited it should be to Clausius and Clapeyron.

30. In line 808, ocean temperature data has very ambiguous implications as noted in Lindzen and Giannitsis (2002).

31. More discussion should be provided in line 811 about the widespread and accelerating glacial retreat. It is those mountain glaciers in low and midlatitudes that are melting systematically. This is a terrific proxy measurement because the mountain glacier melting is unprecedented in modern history and is now happening within the lower atmosphere that is a primary focus of this report. The report needs to point out that just glaciers that respond to summer temperatures are retreating. Many glaciers that respond to winter precipitation are advancing.

32. Trends are influenced by ozone, greenhouse gases, aerosols, etc. In addition to addressing combined forcings in lines 849-863, also discuss the response of the model to each of the forcings individually and the uncertainties in the various forcings.

33. Volcanoes and ENSO do not make much difference to the trend. As these series are now slightly longer than earlier studies so that they no longer end with a major (1997-1998) ENSO event, the points about volcanoes and ENSO could be removed. There could be a brief discussion in lines 865-881 or preferably earlier regarding their effects.

34. The section in lines 903-910 should be more explicit in saying which of the various components contribute separately to the agreement. It should also be more explicit about exactly which forcings are included in the "all" integration.

Review of Chapter 6

Chapter 6 asks: what measures can be taken to improve the understanding of observed changes? This chapter purports to respond to issues and shortcomings raised in Chapters 1-5 and develops a list of seven comprehensive recommendations. These recommendations address: (1) the need for improved observing standards that are rigorously implemented, (2) better use of existing data, (3) expanded use of regional and global climate models for assessing the impacts of forcings and feedbacks on temperature trends, (4) continued assessment of tropospheric trends using a full range of statistical techniques and modeling tools, (5) enhanced development of reanalyses, (6) improved metadata, and (7) development of scientific talent.

The committee finds that the recommendations in Chapter 6 are insufficiently specific and not clearly prioritized. Furthermore, the seven recommendations seem largely disconnected from the findings in Chapters 1-5, and even from the text in Chapter 6. This chapter needs a substantial rewrite, including re-organization of the text and reformulation of the recommendations.

MAJOR COMMENTS

1. Chapter 6 should be reorganized into two parts:

    a. The first part should take findings from Chapters 1-5 to recommend specific opportunities to improve understanding of vertical temperature trends. These should focus on addressing remaining uncertainties in existing satellite and radiosonde data sets.
    b. The second part should focus on future measurement opportunities in the context of the specific goals of the report for reconciling observations and understanding of temperature trends.

2. Also in reorganizing Chapter 6, the committee recommends starting with the Global Climate Observing System (GCOS) implementation plan and reinforcing, adding to, or modifying that plan, rather than starting from scratch. It is important that the community speak with a unified voice as much as possible. The authors should also discuss the current efforts to improve the relevant temperature measurements in addition to GCOS, including Global Ocean Data Assimilation Experiment (GODAE),

SEAFLUX, and Global Energy and Water Cycle Experiment (GEWEX) Radiation Panel efforts to develop technologies for reference radiosondes, and discuss international efforts, not just U.S. efforts.

3. The organization of the chapter is centered around data types, such as "surface", "tropospheric", and "reanalyses". A variety of new issues that do not directly map to the seven recommendations are brought up in these sections (much of which is not relevant or belongs in previous chapters). An alternate organization would be to have seven sections with section headings that are the first sentence of each recommendation. Then the text of each section would tie directly back to a need documented in the earlier chapters, would include discussion of the adequacy of current national and international plans to address this need, and make further specific recommendations for implementation of this recommendation.

4. A substantial amount of new information is introduced for the first time in Chapter 6, including material that should have been introduced in earlier chapters if it is deemed relevant and material that does not directly map to the seven recommendations. The following is specific information that is redundant or should be moved to previous chapters:

- The material in lines 54-71 should be mentioned in the context of Chapters 1 and 5.
- Text on snow and sea ice and sampling inadequacies in lines 179-187 should be moved to Chapter 2.
- For lines 138-177, lines 240-254, and lines 300-317, text on combining surface temperature and dew point temperature is far too wordy, and the main point is lost. This concept should be included in a general recommendation on the need to evaluate and interpret the temperature data in the context of other data sets (e.g., humidity, winds, ocean heat content, etc.) and to understand issues such as the impact of changing land use on temperature trends, as stated in lines 347-355.
- The text in lines 447-494 about recommending specific improved climate model parameterizations is not directly relevant to the present study, although it is appropriate to state in earlier chapters that inadequate parameterizations in numerical weather prediction models contribute to potential problems in using the reanalyses to determine temperature trends.
- The text in lines 98-105 should be moved to Chapters 2 and 4 as these points were not adequately made in those chapters.

5. As far as the current recommendations in Chapter 6 still appear after the chapter is revised, here are comments on each of the current recommendations. The seven recommendations in Chapter 6 have been said numerous times before in other reports. Also, given the relative lack of traceability of these recommendations to the previous five chapters, it may be that a significant recommendation was omitted.

a. The first recommendation concerns reference measurements. The recommendation should be formulated to account for the adequacy or inadequacy of current national and international plans to address this need. If inadequate recommendations are made in previous documents (e.g., GCOS), then very specific recommendations should be made to address the sensor design, sampling, or other needs.

b. The second recommendation concerns making better use of existing data. See comments for the first recommendation. This section should focus specifically on reprocessing of radiosonde data, resolving the differences between the different MSU analyses, and use of the TIROS Operational Vertical Sounder (TOVS) data, including some very specific recommendations to address the key issues. It should also discuss detailed intercomparison (at the pixel level) of the different data sets and cross checking with other variables. Better scientific uncertainty analysis of the data sets should be part of this recommendation. Specific recommendations here would add considerable value to this document.

c. The third recommendation concerns the use of climate models to interpret the cause of temperature trends. This recommendation needs to be reformulated or perhaps eliminated because it is too broad and inappropriate for the present study. What is recommended here should follow directly from Chapter 5 and any uncertainties or inconsistencies in the analyses that were identified. An alternative recommendation would be to "Improve the scientific understanding of the variations of the vertical temperature structure of the atmosphere". It should also be clearly emphasized that data is being used to test models and not vice-versa.

d. The fourth recommendation concerns statistical trend analysis. A clear case has not been made in the previous chapters (or in Chapter 6) that there is a need for new research in the statistical analysis of trends. Rather, the committee would prefer that the report give explicit discussion to existing methods for dealing with such issues as autoregressive behavior and nonlinearities in trends, as already discussed in review comments on Chapter 2.

e. The fifth recommendation concerns climate quality reanalyses. Just as for the third recommendation, this one needs to be reformulated or perhaps eliminated. It is not useful to state such a broad recommendation that has already been made in other contexts. If there are any specific recommendations that would help address the temperature trend problem, then they should be formulated. Possibilities would include careful documentation about what assimilation data is actually assimilated into the model as a function of space and time, data assimilation experiments, etc.

f. The sixth recommendation concerns metadata. It seems that this issue is (or easily could be) covered in the first recommendation. It is not clear that accessibility of the data is a major issue.

g. The seventh recommendation concerns education. This recommendation is very diffuse and is not motivated by the previous chapters. It is hard to disagree with the statement that education in our field should include a stronger emphasis on the proper use of statistics and error analysis. However, this point could easily be incorporated into the second recommendation.

h. An outstanding omission in terms of recommendations is the need for better methods to sense temperature or related variables from satellites, such as using instruments that are self calibrating, sounders with more channels for better vertical resolution, and the use of proxy measures such as refractive index and spectral TOA radiance.

SPECIFIC COMMENTS

1. In lines 79-80, it is the committee's understanding that the U.S. Climate Reference Network been shelved or at least stalled.
2. Most countries do not know about the GCOS Monitoring Principles, mentioned in lines 87-88.
3. Many of the recommendations in lines 150-160 may be difficult to achieve based on cost considerations. The GCOS aim is to get the data first, then work on metadata. Getting pictures of sites will only be useful if they are taken at regular intervals.
4. In lines 198-199, there is a GCOS working group of the Ocean Observing Panel for Climate (OOPC) and the Atmospheric Observing Panel for Climate (AOPC) looking at Sea Surface Temperature biases.
5. In line 287, locating the reference sonde stations for comparison with satellite overpasses requires observations at different times at each station. Thus, the CCSP authors may want to reconsider this recommendation.
6. Better use of statistics is needed in lines 366-369.
7. In lines 402-407, there are at least two comments on Kalnay and Cai (2003) and there should also be a reference to Simmons et al. (2004). Several criticisms of the Kalnay and Cai approach have been identified.
8. The recommendations for "tightly constraining" the dataset for reanalyses in line 425 is not possible or wise owing to continual changes in all observations, including sondes.

References

Agudelo, P. A. and J. A. Curry. 2004. Analysis of spatial distribution in tropospheric temperature trends. *Geophys. Res. Lett*. 31, L222207.

Allen, M. R. and S. F. B. Tett. (1999) Checking for model consistency in optimal fingerprinting. *Clim. Dynam.* 15: 419-434.

Allen, M. R. and P. A. Stott. 2003. Estimating signal amplitudes in optimal fingerprinting, Part I: Theory. *Clim. Dynam.* 21: 477-491.

Allen, M. R., N. P. Gillett, J. A. Kettleborough, R. Schnur, G. S. Jones, T. Delworth, F. Zwiers, G. Hegerl, T. P. Barnett. 2005. Quantifying anthropogenic influence on recent near-surface temperature. *Surv. Geophys.* (in press).

Anderson, T. L., R. J. Charlson, S. E. Schwartz, R. Knutti, O. Boucher, H. Rodhe, and J. Heitzenberg. 2003. Climate forcing by aerosols – a hazy picture. *Science* 300: 1103-1104.

Armstrong, E., and J. Vazquez-Cuervo. 2001. A New Global Satellite-Based Sea Surface Temperature Climatology. *Geophys. Res. Lett.* 28 (22): 4199–4202.

Bloomfield, P. 1992. Trends in global temperature. *Climatic Change* 21: 1-16.

Bloomfield, P. and D. Nychka. 1992. Climate spectra and detecting climate change. *Climatic Change* 21: 275-288.

Bradley, R. S. 1988. The explosive volcano eruption signal in Northern Hemisphere continental temperature records. *Climatic Change* 12: 221-243.

Castro, S. L., G.A. Wick, and W.J. Emery. 2003. Further refinements to models for the bulk-skin sea surface temperature difference. *J. Geophys. Res.* 108 (C12) 3377.

Charlson, R. J., S. E. Schwartz, J. M. Hales, R. D. Cess, J. A. Coakley, J. E. Hansen, and D. J. Hoffman. 1992. Climate forcing by anthropogenic aerosols. *Science* 255: 423-430.

Chedin, A., E. Pequignot, S. Serrar, and N. A. Scott. 2004. Simultaneous determination of continental surface emissivity and temperature from NOAA 10/HIRS observations: Analysis of their seasonal variations. *J. Geophys. Res.-Atmos.* 109: D20110.

Chelton, D. B. 2005. The impact of SST specification on ECMWF surface wind stress fields in the eastern tropical Pacific. *J. Climate* 18: 530-550.

Chou, C.-B. and H.-P. Huang. 2004. A new procedure for estimating observation errors in AMSU data and its application to retrieval. *Q. J. Roy. Meteoro. Soc.* 130 (596): 79-101 Part A.

Curry, J. A., A. Bentamy, M. A. Bourassa, D. Bourras, E.F. Bradley, M. Brunke, S. Castro, S.H. Chou, C.A. Clayons, W.J. Emery, C.W. Fairall, M. Kubota, B. Lin, W. Perrie, R.A. Reeder, I.A. Renfrew, W.B. Rossow, J. Schulz, S.R. Smith, P.J. Webster, G.A. Wick, and X. Zeng. 2004. SEAFLUX. *B. Am. Meterorol. Soc* 85: 409-424.

Diggle, P. J., K. Liang, and S. L. Zeger. 1996. Analysis of Longitudinal Data. New York: Oxford University Press.

Douglass, D. and R. Knox, R. 2005. Climate forcing by the volcanic eruption of Mount Pinatubo. *Geophys. Res. Lett.* 32, 10.1029/2004GL022119.

Durre, I., T. C. Peterson, and R. S. Vose, 2002. Evaluation of the effect of the Luers-Eskridge radiation adjustments on radiosonde temperature homogeneity. *J. Climate* 15: 1335-1347.

Goody, R. M. and Y. L. Yung. 1989. Atmospheric Radiation, Theoretical Basis. New York: Oxford University Press.

Free, M, and D. J. Seidel. 2005. Causes of differing temperature trends in radiosonde upper air ata sets. *J. Geophys. Res.* 110, DOI:10.1029/2004JD005481 (in press).

Frohlich, C. and J. Lean. 2004. Solar radiative output and its variability: evidence and mechanisms *Astron. Astrophys. Rev.* 12: 273-320.

Fu, Q., and C. M. Johanson. 2004. Stratospheric influence on MSU-derived tropospheric temperature trends: A direct error analysis. *J. Climate* 17: 4636-4640.

Fu, Q. and C. M. Johanson. 2005. Satellite-derived vertical dependence of tropical tropospheric temperature trends. *Geophys. Res. Lett.* submitted.

Fu, Q., C. M. Johanson, S. G. Warren, and D. J. Seidel. 2004a. Contribution of stratospheric cooling to satellite-inferred tropospheric temperature trends. *Nature* 429: 55-58.

Fu, Q., D. J. Seidel, C. M. Johanson, and S. G. Warren. 2004b. Stratospheric cooling and the troposphere (reply). *Nature* 432, DOI:10.1038/nature03210.

Gentemann, C., C. Donlon, A. Stuart-Menteth, and F. Wentz. 2003. Diurnal signals in satellite sea surface temperature measurements. *Geophys. Res. Lett.* 30 (3)1140.

Gettelman, A., D. J. Seidel, M. C. Wheeler and R. J. Ross. 2002. Multi-decadal trends in tropical convective available potential energy. *J. Geophys. Res.* 107, 4606, DOI: 10.1029/2001JD001082.

Gilbert, R. O. 1987. Statistical methods for environmental monitoring. New York: Van Nostrand Reinhold. 384 p.

Gillett, N. P., B. D. Santer, and A. J. Weaver. 2004. Stratospheric cooling and the troposphere. *Nature* 432, 10.1038/nature03209.

Goldberg, M., Y. Qu, L. McMillin, W. Wolf, L. Zhou, and M. Divakarla. 2003. AIRS Near-Real-Time Products and Algorithms in Support of Operational Numerical Weather Prediction. IEEE Transactions on Geoscience and Remote Sensing 41 (2): 379-389.

Grist, J., and S. Josey. 2003. Inverse Analysis Adjustment of the Soc Air-Sea Flux climatology using ocean heat transport Constraints. *J. Climate* 16: 3724-3295.

Grody, N. C., K. Y. Vinnikov, M. D. Goldberg, J. T. Sullivan, and J. D. Tarpley. 2004. Calibration of multi-satellite observations for climate studies: Microwave Sounding Unit (MSU). *J. Geophys. Res.* 109, DOI: 10.1029/2004JD005079.

Guan, L., and H. Kawamura. 2004. Merging Satellite Infrared and Microwave SSTs: Methodology and Evaluation of the New SST. *J. Oceanography* 60 (5): 905-912.

Hastie, T. J. and R. J. Tibshirani. 1990. Generalized Additive Models. Monographs on Statistics and Applied Probability, volume 43. Boca Roton, U.S.: Chapman and Hall.

Hegerl G. C., K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz. 1997. Multi-fingerprint detection and attribution analysis of greenhouse gas-, greenhouse gas-plus-aerosol- and solar-forced climate change. *Clim. Dynam.* 13: 613-634.

Hegerl, G. C., and M. R. Allen. 2002. Origins of model-data discrepancies in optimal fingerprinting. *J. Climate* 15: 1348-1356.

Hosoda, K., and H. Kawamura. 2004. Examination of the Merged Sea Surface Temperature Using Wavelet Analysis. *J. Oceanography* 60 (5): 843-852.

Hoskins, B. J. 2003. Atmospheric processes and observations. *Philos. Tr. R. Soc. S.-A* 363: 1945-1960.

Hurrell, J. W., S. J. Brown, K. E. Trenberth, and J. R. Christy. 2000. Comparison of tropospheric temperatures from radiosondes and satellites: 1979–98. *Bull. Amer. Met. Soc*. 81: 2165-2177.

IPCC (Intergovernmental Panel on Climate Change). 2001. Third Assessment Report of the Intergovernmental Panel on Climate Change 2001: The Scientific Basis. Cambridge, U. K.: Cambridge University Press.

Jin, M. and R. E. Dickinson. 2002. New Observational Evidence for Global Warming from Satellite Data Set. *Geophys. Res. Lett.* 29 (10): 10.1029/2001GLO013833.

Jones, P. D., T. J. Osborn, K. R. Briffa, C. K. Folland, B. Horton, L. V. Alexander, D.E. Parker, and N. A. Rayner. 2001. Adjusting for sampling density in grid-box land and ocean surface temperature time series. *J. Geophys. Res.* 106: 3371-3380.

Jones, P. D. 1994. Recent warming in global temperature series. *Geophys. Res. Lett.* 21: 1149-1152.

Jones, P. D., A. Moberg, T. J. Osborn, and K. R. Briffa. 2003. Surface climate responses to explosive volcanic eruptions seen in long European temperature records and mid-to-high latitude tree-ring density around the Northern Hemisphere. Pp. 239-254 in Volcanism and the Earth's Atmosphere, A. Robock and C. Oppenheimer, eds. Washington D.C.: American Geophysical Union.

Jones, P. D., T. J. Osborn, and K. R. Briffa. 1997. Estimating sampling errors in large-scale temperature averages. *J. Climate* 10: 2548-2568.

Kalnay, E. and M. Cai. 2003. Impact of urbanization and land-use change on climate. *Nature* 423: 528 - 531 (29 May 2003); DOI:10.1038/nature01675

Karl, T. R. and R. W. Knight. 1998. Secular trends of precipitation amount, frequency, and intensity in the USA. *Bull. Amer. Meteor. Soc*. 79: 231-241.

Karl, T. R., R. W. Knight, D. R. Easterling and R. G. Quayle. 1996. Indices of climate change for the United States. *Bull. Amer. Meteor. Soc*. 77: 279-292.

Kuo, Y.-H., W. S. Schreiner, J. Wang, D. L. Rossiter, and Y. Zhang. 2005. Comparison of GPS radio occultation soundings with radiosondes. *Geophys. Res. Lett.* 32, L05817, 10.1029/2004GL021443.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy. 1997. Observing the Earth's Atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.* 102: 23429-23465.

Lakshmi, V., and D. Zehrfuhs. 2002. Normalization and Comparison of Surface Temperatures Across a Range of Scales. IEEE Transactions on Geoscience and Remote Sensing 40 (12): 2636-2646.

Levine, R. A. and L. M. Berliner. 1999. Statistical Principles for Climate Change Studies, J. Climate 12: 565-574.

Lindzen, R. S. and K. Emanuel. 2002. The greenhouse effect. Pp. 562-566 in Encyclopedia of Global Change, Environmental Change and Human Society, Volume 1, Andrew S. Goudie, ed. New York: Oxford University Press. 710 pp.

Lindzen, R. S. and C. Giannitsis. 1998. On the climatic implications of volcanic cooling. *J. Geophys. Res.* 5929-5941

Lindzen, R.S. and C. Giannitsis. 2002. Reconciling observations of global temperature change. *Geophys. Res. Lett.* 29: 10.1029/2001GL014074.

Luers, J. K., and R. E. Eskridge. 1998. Use of radiosonde temperature data in climate studies. *J. Climate* 11: 1002-1019.

Matthews, H. D., A. J. Weaver, K.J. Meissner, N.P. Gillett, and M. Eby. 2004. Natural and anthropogenic climate change: Incorporating historical land cover change, vegetation dynamics and the global carbon cycle. *Clim. Dynam.* 22: 461-479.

Menglin, J. 2004. Analysis of land skin temperature using AVHRR observations. *B. Am. Meteorol. Soc.* 85 (4): 587-600.

Menglin, J., and R. Dickinson. 2002. New observational evidence for global warming from satellite. *Geophys. Res. Lett.* 29 (10) 1400.

NRC (National Research Council). 2000. Reconciling Observations of Global Temperature Change. Washington, D.C.: The National Academies Press.

NRC. 2003. Understanding Climate Change Feedbacks. Washington, D.C.: The National Academies Press.

NRC. 2004. Implementing Climate and Global Change Research: A Review of the Final U.S. Climate Change Science Program Strategic Plan. Washington, D.C.: The National Academies Press.

Pielke, R. A., Sr., and T. N. Chase. 2004. Technical comment: Contributions of anthropogenic and natural forcing to recent tropopause height changes. Science 303:1771b.

Reynolds, R., N. Rayner, T. Smith, D. Stokes, D., and W. Wang. 2002. An improved in Situ and Satellite SST Analysis for climate. *J. Climate* 15 (13): 1609-1625.

Reynolds, R., C. Gentemann, and F. Wentz. 2004. Impact of TRMM SSTs on a Climate-Scale SST Analysis. *J. Climate* 17 (15): 2938-2952.

Robock, A. and C. Oppenheimer. 2003. Volcanism and the Earth's Atmosphere. AGU Geophysical Monograph Series, 139. Washington, DC: American Geophysical Union. 360 pp.

Ruppert, D., M. P. Wand, and R. J. Carroll. 2003. Semiparametric Regression. Cambridge, U. K.: Cambridge University Press.

Santer, B. D., K. E. Taylor, T. M. L. Wigley, J. E. Penner, P. D. Jones, and U. Cubasch. 1995. Towards the detection and attribution of an anthropogenic effect on climate. *Clim. Dynam.* 12: 77-100.

Santer, B. D., K. E. Taylor, T. M. L. Wigley, T. C. Johns, P. D. Jones, D. J. Karoly, J. F. B. Mitchell, A. H. Oort, J. E. Penner, V. Ramaswamy, M. D. Schwarzkopf, R. J. Stouffer, and S. Tett. 1996. A search for human influences on the thermal structure of the atmosphere. *Nature* 382: 39-46.

Santer, B. D., T. M. L Wigley, D. J. Gaffen, L. Bengtsson, C. Doutriaux, J. S. Boyle, M. Esch, J.J. Hnilo, P.D. Jones, G. A. Meehl, E. Roeckner, K. E. Taylor, and M. F. Wehner, 2000. Interpreting differential temperature trends at the surface and in the lower troposphere. *Science* 287: 1227-1232.

Sarachik, E. 1985. A simple theory for the vertical structure of the tropical atmosphere. *Pure and Appl. Geophys.* 123: 261-271.

Schneider, T. 2004. The Tropopause and the Thermal Stratification in the Extratropics of a Dry Atmosphere. *J. Atmos. Sci.* 61: 1317-1340.

Schroder, T., S. Leroy, M. Stendel, and E. Kaas, 2003. Validating the microwave sounding unit stratospheric record using GPS occultation. *Geophys. Res. Lett*. 30, 1734, DOI 10.1029/2003GL017588.

Schweiger, A., R. Lindsay, J. Francis, J. Key, J. Intrieri, and M. Shupe. 2002. Validation of TOVS Path-P data during SHEBA. *J. Geophys. Res.-Oceans* 107 (C10) 8041.

Simmons, A. J., P. D. Jones, V. da Costa Bechtold, A. C. M. Beljaars, P. Kållberg, S. Saarinen, S. M. Uppala, P. Viterbo, and N. Wedi. 2004. Comparison of trends and variability in CRU, ERA-40 and NCEP/NCAR analyses of monthly-mean surface air temperature. *J. Geophys. Res.* 109, No D24 D24115.

Smith, R. L., T. M. L. Wigley, and B. D. Santer 2003. A bivariate time series approach to anthropogenic trend detection in hemispheric time series. *J. Climate* 16: 1228-1240.

Smith, S., D. Legler, and K. Verzone. 2001. Quantifying Uncertainties in NCEP Reanalyses Using High-Quality Research Vessel Observations. *J. Climate* 14 (20): 4062-4072.

Soden, B. J., R. T. Wetherald, G. L. Stenchikov, and A. Robock. 2002. Global cooling after the eruption of Mount Pintubo: A test of climate feedback by water vapor. *Science* 296: 727-730.

Stocker, T. F., G. K. C. Clarke, H. Le Treut, R. S. Lindzen, V. P. Meleshko, R. K. Mugara, T. N. Palmer, R. T. Pierrehumbert, P. J. Sellers, K. E. Trenberth, J. Willebrand, R. B. Alley, O. E. Anisimov, C. Appenzeller, R. G. Barry, J. J. Bates, R. Bindschadler, G. B. Bonan, C. W. Böning, S. Bony, H. Bryden, M. A. Cane, J. A. Curry, T. Delworth, A. S. Denning, R. E. Dickinson, K. Echelmeyer, K. Emanuel, G. Flato, I. Fung, M. Geller, P. R. Gent, S. M. Griffies, I. Held, A. Henderson-Sellers, A. A. M. Holtslag, F. Hourdin, J. W. Hurrell, V. M. Kattsov, P. D. Killworth, Y. Kushnir, W. G. Large, M. Latif, P. Lemke, M. E. Mann, G. Meehl, U. Mikolajewicz, W. O'Hirok, C. L. Parkinson, A. Payne, A. Pitman, J. Polcher, I. Polyakov, V. Ramaswamy, P. J. Rasch, E. P. Salathe, C. Schär, R. W. Schmitt, T. G. Shepherd, B. J. Soden, R. W. Spencer, P. Taylor, A. Timmermann, K. Y. Vinnikov, M. Visbeck, S. E. Wijffels, and M. Wild. 2001. Physical climate processes and feedbacks. 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, Y. Ding, D. J.

Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, eds., Cambridge, U.K.: Cambridge University Press. 881 pp.

Swanson, R. E. 2003. Evidence of possible sea-ice influence on Microwave Sounding Unit tropospheric temperature trends in polar regions. *Geophys. Res, Lett.* 30, 2040, DOI:10.1029/2003GL017938.

Tett, S. F. B. and P. W. Thorne. 2004. Tropospheric temperature series from satellites. *Nature* DOI:10.1038/nature03208.

Venables, W. N. and B. D. Ripley. 2002. Modern Applied Statistics with S (fourth edition). New York: Springer Verlag.

Xu, K.-M. and K. A. Emanuel. 1989. Is the tropical atmosphere conditionally unstable? *Mon. Weather Rev.* 117: 1471-1479.

Zhang, H.-M., R. Reynolds, and T. Smith. 2004. Bias characteristics in the AVHRR sea surface temperature. *Geophys. Res, Lett.* 31 L01307.

Appendixes

A

Prospectus for the Synthesis and Assessment Product

Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences

Chief Editor:	Thomas R. Karl
Associate Editors:	Christopher D. Miller, William L. Murray

1. Overview: Description of Topic, Audience, Intended Use, and Questions to Be Addressed

Independently produced data sets that describe the four-dimensional temperature structure from the surface through the lower stratosphere provide different temperature trends. These differences are seen in varying degrees in comparisons of separate *in situ* (surface and weather balloon) data sets, in comparisons of separate space-based data sets, and in comparisons of individual data sets drawn from the different observational platforms and different trend analysis teams.

This CCSP synthesis and assessment product will address the accuracy and consistency of these temperature records and outline steps necessary to reconcile differences between individual data sets. Understanding exactly how and why there are differences in temperature trends reported by several analysis teams using differing observation systems and analysis methods represents a necessary step in reducing the uncertainties that underlie current efforts focused on the detection and quantification of surface and tropospheric temperature trends. Consequently, this synthesis and assessment product promises to be of significant value to decisionmakers, and to the expert scientific and stakeholder communities. For example, we expect this assessment to be a major contributor to the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (due to be published in 2007). In addition, we expect the information generated will be used by the Global Climate Observing System Atmospheric Observation Panel to help identify effective ways to reduce observational uncertainty.

Recent efforts to address the uncertainties regarding the temperature structure of the lower atmosphere (i.e., from the surface through the lower stratosphere) have included release of a report under the auspices of the National Research Council (NRC) entitled

"Reconciling Observations of Global Temperature Change" (NRC, 2000) and the IPCC Third Assessment Report (IPCC, 2001, pp 101-123). Although these documents provided a great deal of useful information, the complexities of the issue coupled with shortcomings of the available observing systems prevented resolution of a number of fundamental questions, including:

1. Why do temperatures vary vertically (from the surface to the stratosphere) and what do we understand about why they might vary and change over time?
2. What kinds of atmospheric temperature variations can the current observing systems measure and what are their strengths and limitations, both spatially and temporally?
3. What do observations indicate about the changes of temperature in the atmosphere and at the surface since the advent of measuring temperatures vertically?
4. What is our understanding of the contribution made by observational or methodological uncertainties to the previously reported vertical differences in temperature trends?
5. How well can the observed vertical temperature changes be reconciled with our understanding of the causes of these changes?
6. What measures can be taken to improve the understanding of observed changes?

These questions provide the basis for the six main chapters in the synthesis and assessment product. They highlight several of the fundamental uncertainties and differences between and within the individual components of the existing observational and modeling systems. The responses to the questions will be written in a style consistent with major international scientific assessments [e.g., IPCC assessments, and the Global Ozone Research and Monitoring Project (WMO, 1999)].

2. Contact Information: Email and Telephone for Responsible Individuals at the Lead and Supporting Agencies

NOAA is the lead agency for this synthesis product. Relevant agency personnel are presented in the following table:

CCSP Member Agency	Agency Leads
DOC (NOAA)	Tom Karl/Chris Miller/Bill Murray
DOE	Rick Petty
NASA	Eric Fetzer
NSF	Jay Fein

3. Lead Authors: Required Expertise and Biographical Information

A list of lead author nominees was identified based on past records of interest and accomplishment in framing the core issues related to changes, trends, and uncertainties in the lower atmospheric temperature records, advancing relevant scientific arguments, and contributing to increased understanding of the behavior of respective components of the end-to-end system that provides the required data sets. Past contributions to relevant scientific assessments, success in peer-reviewed proposal funding competitions, and publication records in refereed journals are among the measures used in the selection process. The lead authors selected on the basis of these criteria are listed below. Chapter assignments and biographical information are presented in Appendix A.

Lead Authors

John Christy (University of Alabama/Huntsville)
Chris Folland, (Hadley Centre, U.K. Met Office)
Chris Forest (Massachusetts Institute of Technology)
Jim Hurrell (National Center for Atmospheric Research)
John Lanzante (NOAA/Geophysical Fluid Dynamics Laboratory)
Carl Mears (Remote Sensing Systems)
Jerry Meehl (National Center for Atmospheric Research)
David Parker (U.K. Met Office)
Joyce Penner (U. Michigan)
Thomas C. Peterson (NOAA/National Climatic Data Center)
Roger Pielke Sr. (Colorado State University)
V. Ramaswamy (NOAA/ Geophysical Fluid Dynamics Laboratory)
Dick Reynolds (NOAA/ National Climatic Data Center)
Ben Santer (Lawrence Livermore National Laboratory)
Dian Seidel (NOAA Air Resources Laboratory)
Steve Sherwood (Yale University)
Roy Spencer (U. Alabama-Huntsville)
Peter Thorne (U.K. Met Office/Hadley Centre)
Kostya Vinnikov (University of Maryland)
Russell S. Vose (NOAA/ National Climatic Data Center)
Frank Wentz (Remote Sensing Systems
Tom M.L. Wigley (National Center for Atmospheric Research)

4. Stakeholder Interactions

The questions addressed by the report were framed by the lead agency with the benefit of consultation from members of the Climate Change Science Program (CCSP) Office, the NOAA Science Advisory Board Climate Monitoring Working Group[1], and participants at a workshop on Reconciling Vertical Temperature Trends that was held at NOAA's National Climatic Data Center (NCDC) on 27-29 October 2003, and attended by 55 scientific experts from academia, the U.S. government, the private sector, and several

---

[1] The NOAA Science Advisory Board Climate Monitoring Working Group, which has since been merged with the Climate and Global Change Working Group, was charged to provide, in the context of national and international activities, scientific advice and broad program direction to NOAA on the condition and capabilities of NOAA's observing systems/data management systems for the purpose of climate monitoring.

scientific experts from other countries. The workshop was designed to address a broad range of issues related to vertical temperatures trends, and it provided a scientific foundation for the development of this CCSP synthesis product. The workshop presentations and results of breakout groups are posted on <http://www.ncdc.noaa.gov/oa/rvtt.html>. The workshop assessed the current state of knowledge on this topic, identified near-term and long-term steps to address existing uncertainties, and provided a framework for a synthesis and assessment product structured around the six questions listed above.

In addition, Principals on the CCSP Interagency Committee provided input from a governmental perspective during the CCSP review, and other stakeholders provided input during the public comment period (see <http://www.climatescience.gov/Library/sap/sap1-1/sap1-1prospectus-comments.htm> for a collation of the comments submitted during public comment period).

5. Drafting, Including Materials to be Used in Preparing the Product

The lead NOAA focal point is the Product 1.1 Chief Editor. The assistant NOAA focal points serve as Associate Editors. The core of a scientific author team presented in Appendix A has been drawn from the participants in the workshop described above. This core group has been supplemented with a number of individuals who have made major contributions to our present understanding of the issues related to vertical temperature change.

Under the leadership of a convening lead author for each of the six chapters, this group of lead authors and contributors is charged with the preparation of the scientific/technical analysis section of the synthesis report. They will draw upon published, peer-reviewed scientific literature in the drafting process.

The synthesis and assessment product will include an Executive Summary which will present key findings from Chapters 1-6. It will be written by a team consisting of a convening lead author assisted by the convening lead authors from each of the six chapters.

The synthesis product will identify disparate views that have significant scientific or technical support, and will provide confidence levels for key findings, as appropriate.

This synthesis and assessment product will pay special attention to addressing uncertainties and confidence levels in our statements regarding the temperature trends. We note that increased understanding of the complexities of the vertical temperature variability can lead to increased uncertainties regarding long-term behavior patterns. Just as independent data sets must be used for comparisons of results, the basic evaluation process must maintain appropriate degrees of separation; for example, data set developers should not be the only evaluators of data reliability in their products.

The communication of uncertainties will be quantitative in many instances but, from discussion during the Asheville workshop, it is clear that expert judgment will also be used because standard statistical methods alone do not reflect the full range of uncertainty. Our intent is to follow the protocol developed in the IPCC (2001) assessment and subsequent updates provided by IPCC.

6. Review

NOAA, the lead agency for this product, plans to present the document to an NRC expert committee for scientific review. The NRC Proposal (NAS Proposal No. 04-DELS-385-01) to conduct the review states that the review will address the following issues:

1. Are the goals, objectives, and intended audience of the product are clearly described in the document? Does the product address all the questions outlined in the prospectus?
2. Are findings and recommendations are adequately supported by evidence and analysis? If any recommendations are based on value judgments or the collective opinions of the authors, is this acknowledged and are adequate reasons given for reaching those judgments?
3. Are the data and analyses handled competently? Are the statistical methods applied appropriately? Are the uncertainties and confidence levels evaluated and communicated appropriately?
4. Are the document's presentation and is organization effective? Are the questions outlined in the prospectus addressed and communicated in a manner that is appropriate for the intended audience?
5. Is the document scientifically objective and policy neutral? Is it consistent with the scientific literature, including recent NRC reports and other scientific assessments on the same topic?
6. Does the summary concisely and accurately describe the content, key findings, and recommendations? Is it consistent with other sections of the document?
7. What other significant improvements, if any, might be made in the document?

The period of performance for the review is expected to be approximately January to April 2005.

Following expert review, the lead authors will revise the draft product by incorporating comments and suggestions from the reviewers, as the lead authors deem appropriate.

Following this revision, the draft product will be released for public comment. The public comment period will be 45 days and will take place from 1 June to 15 July 2005.

The lead authors will prepare a third draft of the product, taking into consideration the comments submitted during the public comment period. The scientific judgment of the lead authors will determine responses to the comments.

Once the revisions are complete, the lead agency will submit the synthesis and assessment product to the CCSP Interagency Committee for approval. If the CCSP Interagency Committee determines that further revision is necessary, their comments will be sent to the lead agency for consideration and resolution by lead authors. If needed, the NRC will be asked to provide additional scientific analysis to bound scientific uncertainty associated with specific issues.

If the CCSP Interagency Committee review determines that no further revisions are needed and that the product has been prepared in conformance with the *Guidelines for*

Producing CCSP Synthesis and Assessment Products (see <http://www.climatescience.gov/Library/sap/sap-guidelines.htm>) and the Data Quality Act (including ensuring objectivity, utility, and integrity as defined in 67 FR 8452), they will submit the product to the National Science and Technology Council (NSTC) for clearance. Clearance will require the concurrence of all members of the Committee on Environment and Natural Resources. Comments generated during the NSTC review will be addressed by the CCSP Interagency Committee in consultation with the lead and supporting agencies and the lead authors.

7. Related Activities: Coordination with Other National or International Assessment Processes

This CCSP synthesis and assessment product has been coordinated internationally with a U.K. Met Office workshop on understanding vertical profiles of temperature trends conducted in September 2004 in Exeter, England. The coordination included presentations in Exeter by the synthesis and assessment product lead authors to provide an interim look at progress on addressing each of the key questions. There is also ongoing coordination with a newly constituted Global Climate Observing System (GCOS)/Atmospheric Observations Panel for Climate (AOPC) Working Group on Reconciling Vertical Temperature Trends. The synthesis and assessment product is expected to provide input to the IPCC Fourth Assessment Report.

8. Communications

NOAA, the lead agency, will produce and release the completed product using a standard format for all CCSP synthesis and assessment products. The final product and the comments received during the expert review and the public comment period will be posted, without attribution (unless specific reviewers agree to attribution), on the CCSP web site.

The lead authors will also be encouraged to publish their findings in the scientific literature.

9. Proposed Timeline

Preparation of this synthesis and assessment product has been underway during completion of this prospectus because of the time required to finalize the overall *Guidelines for Producing CCSP Synthesis and Assessment Products*. This approach was taken in order to coordinate work on the product with other international efforts, in particular, so the product could be completed in time to provide an input to the IPCC Fourth Assessment Report. Comments received on the draft prospectus were taken into account in the process, and all procedures used in preparing the report have been adjusted to be consistent with those mandated by the Guidelines.

The timeline is divided into two phases. The planned completion date for Phase 1, which will result in the submission of the first draft of the synthesis product for scientific review by the National Research Council, is January 2005. The planned completion date for Phase 2, which will culminate with approval of the synthesis product by the President's National Science and Technology Council, is October 2005. Specific milestones follow.

PHASE 1
- Lead authors nominated – July 04
- Synthesis product prospectus released for public comment – July 04
- First lead author meeting – August 04
- Second lead author meeting – October 04
- Third lead author meeting – December 04
- Synthesis product first draft submitted for NRC scientific review – January 05

PHASE 2
- NRC review completed – April 05
- Synthesis product second draft released for public comment – 1 June 2005
- Public comment period completed– 15 July 2005
- Synthesis product third draft and compilation of comments submitted to CCSP Principals – August 05
- Synthesis product accepted by CCSP and submitted to NSTC for final review and approval – September 05
- Synthesis product approved by NSTC – October 05

References

IPCC. 2001. Observed climate variability and 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, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, eds. Cambridge University Press, U.K.

NRC. 2000. *Reconciling Observations of Global Temperature Change*. Washington D.C.: National Academy Press. 85 pp.

WMO. 1999. *Scientific Assessment of Ozone Depletion: 1998*. Global Ozone Research and Monitoring Project, Report #44. World Meteorological Organization, Geneva, Switzerland.

B

Committee to Review the U.S. Climate Change Science Program’s Synthesis and Assessment Product on Temperature Trends in the Lower Atmosphere Statement of Task

The Committee will review the Climate Change Science Program’s draft synthesis and assessment product on “Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences.” The committee will consider the following questions, among others, as it conducts its review:

1. Are the goals, objectives, and intended audience of the product clearly described in the document? Does the product address all the questions as outlined in the prospectus?
2. Are the findings and recommendations adequately supported by evidence and analysis? If any recommendations are based on value judgments or the collective opinions of the authors, is this acknowledged and are adequate reasons given for reaching those judgments?
3. Are the data and analyses handled competently? Are statistical methods applied appropriately? Are uncertainties and confidence levels evaluated and communicated appropriately?
4. Are the document’s presentation and organization effective? Are the questions outlined in the prospectus addressed and communicated in a manner that is appropriate for the intended audiences?
5. Is the document scientifically objective and policy-neutral? Is it consistent with the scientific literature, including recent NRC reports and other scientific assessments on the same topic?
6. Does the summary concisely and accurately describe the content, key findings, and recommendations? Is it consistent with other sections of the document?
7. What significant improvements, if any, might be made in the document?

C

Committee and Staff Biographies

William J. Randel (*Chair*), is a senior scientist at the National Center for Atmospheric Research. His research interests include dynamic variability and climatology of the stratosphere and the observed variability of trace constituents in the middle atmosphere using satellite observations. He has contributed to the WMO/UNEP (World Meteorological Organization/United Nations Environment Programme) Assessments of ozone and temperature trends in the stratosphere and is actively involved with a number of SPARC (Stratospheric Processes and their Role in Climate) activities. He is a lead author on the recently completed IPCC Special Report on Safeguarding the Ozone Layer and the Global Climate System, and a member of the scientific steering group for the Network for the Detection of Stratospheric Change (NDSC). He has also served as chair of the American Geophysical Union's Committee on Atmospheric Dynamics and the American Meteorological Society's Committee on the Middle Atmosphere and is a member of the NRC's Board on Atmospheric Sciences and Climate. Dr. Randel received his Ph.D. in physics from Iowa State University.

Judith A. Curry is Chair of the School of Earth and Atmospheric Sciences at the Georgia Institute of Technology. Her research interests include remote sensing, climate of the polar regions, atmospheric modeling, and air/sea interactions. She participates in the World Meteorological Organization's World Climate Research Program, was a member of the Science Steering Group of the Arctic Climate System (ACSYS) Program, and chairs the Global Energy and Water Cycle Experiment (GEWEX) Cloud System Studies Working Group on Polar Clouds. She co-chaired the Surface Heat Budget of the Arctic Ocean (SHEBA) program's Science Working Group. Dr. Curry previously served on two NRC Committees: the Polar Research Board's Committee to Review NASA's Polar Geophysical Data Sets and the Board on Atmospheric Sciences and Climate's Panel on Coastal Meteorology. She holds a Ph.D. in Geophysical Sciences from the University of Chicago. She is currently a member of the Board on Atmospheric Sciences and Climate's Climate Research Committee.

Dennis L. Hartmann is a professor and chair of the Department of Atmospheric Sciences at the University of Washington. His research includes low-frequency variability in the atmosphere and climate system and research climate change. His climate research focuses on the interaction of dynamics, radiation, and cloud processes, and their roles in determining the sensitivity of the global climate to forcings such as increasing carbon dioxide or aerosol burden. Recent work has resulted in a “Fixed Anvil Temperature” or FAT Hypothesis, that indicates that on the basis of fundamental physical processes the temperature at the tops of tropical anvil clouds should remain about the same during climate change. In the area of dynamics, he is currently interested in the atmospheric dynamical processes that give rise to intrinsic low-frequency variability, especially the interaction of transient and stationary waves with zonal jets in middle latitudes. Dr. Hartmann was a member of the NRC’s Climate Research Committee and Chair of the Panel on Climate Change Feedbacks. He received his Ph.D. in geophysical fluid dynamics from Princeton University.

Phil Jones is the Director of the Climatic Research Unit (CRU) and a Professor in the School of Environmental Sciences at the University of East Anglia in Norwich, England. Dr. Jones completed a B.A. in Environmental Sciences at the University of Lancaster and an M.Sc. and Ph.D. at the Department of Civil Engineering at the University of Newcastle-upon-Tyne. His research has focused in instrumental climate change, paleoclimatology, detection of climate change and the extension of riverflow records in the United Kingdom using long rainfall records. Dr. Jones is recognized for the time series of hemispheric and global surface temperatures, which he updates on a monthly basis. He has coedited four books: *Climate Since A.D. 1500* (with Ray Bradley); *Climatic Variations and Forcing Mechanisms of the Last 2000 Years* (with Ray Bradley and Jean Jouzel); *History and Climate: Memories of the Future* (with Astrid Ogilvie, Trevor Davies and Keith Briffa) and *Improved Understanding of Past Climatic Variability from Early European Instrumental Sources* (with Dario Camuffo). Dr. Jones has been a fellow of the Royal Meteorological Society since 1992 and was on the Editorial Committee of the International Journal of Climatology until 1995. Currently, he is on the editorial board of Climatic Change, an elected member of Academia Europaea since 1998 and a member of the American Meteorological Society since 2001. He was jointly awarded the Hugh Robert Mill Medal in 1995 by the Royal Meteorological Society for work on U.K. rainfall variability, and in 1997 the Outstanding Scientific Paper Award by the Environmental Research Laboratories/NOAA for being a coauthor on the paper “A search for Human Influences on the Thermal Structure of the Atmosphere,” by Ben Santer et al. in Nature, 382, 39-46 (1996). Most recently Dr. Jones was awarded the first Hans Oesschger Medal from the European Geophysical Society (now the European Geosciences Union) in 2002 and the International Journal of Climatology prize of the Royal Meteorological Society for papers published in the last five years, also in 2002.

Kenneth Kunkel is Head of the Atmospheric Environment Section at the Illinois State Water Survey. He has also served as Director of the Midwestern Regional Climate Center and Director of the Office of Applied Climatology. Dr. Kunkel’s research interests include climate variability and change, climate extremes, and boundary layer meteorology. He has considerable experience working with U.S. surface climate datasets and has published a number of papers on analysis of such datasets. In addition, Dr. Kunkel has knowledge of the limitations and proper use of surface temperature data.

Richard S. Lindzen is Sloan Professor of Meteorology in the Department of Earth, Atmospheric and Planetary Sciences at Massachusetts Institute of Technology. Professor Lindzen is a dynamical meteorologist with interests in the broad topics of climate, planetary waves, monsoon meteorology, planetary atmospheres, and hydrodynamic instability. His research involves studies of the role of the tropics in mid-latitude weather and global heat transport, the moisture budget and its role in global change, the origins of ice ages, seasonal effects in atmospheric transport, stratospheric waves, and the observational determination of climate sensitivity. He has made major contributions to the development of the current theory for the Hadley Circulation, which dominates the atmospheric transport of heat and momentum from the tropics to higher latitudes, and has advanced the understanding of the role of small scale gravity waves in producing the reversal of global temperature gradients at the mesopause. He pioneered the study of how ozone photochemistry, radiative transfer and dynamics interact with each other. He is currently studying the ways in which unstable eddies determine the pole to equator temperature difference, and the nonlinear equilibration of baroclinic instability and the contribution of such instabilities to global heat transport. He has also been developing a new approach to air-sea interaction in the tropics, and is actively involved in parameterizing the role of cumulus convection in heating and drying the atmosphere. He has developed models for the Earth's climate with specific concern for the stability of the ice caps, the sensitivity to increases in CO_2, the origin of the 100,000 year cycle in glaciation, and the maintenance of regional variations in climate. In cooperation with colleagues and students, he is developing a sophisticated, but computationally simple, climate model to test whether the proper treatment of cumulus convection will significantly reduce climate sensitivity to the increase of greenhouse gases. Professor Lindzen is a recipient of the AMS's Meisinger and Charney Awards, and the AGU's Macelwane Medal. He is a consultant to the Global Modeling and Simulation Group at NASA's Goddard Space Flight Center and a Distinguished Visiting Scientist at California Institute of Technology's Jet Propulsion Laboratory. He received his Ph.D. from Harvard University. Dr. Lindzen is a member of the National Academy of Science.

Richard L. Smith is a Professor of Statistics at the University of North Carolina, Chapel Hill. Dr. Smith received his Ph.D. from Cornell University in 1979 and has previously held academic positions at Imperial College (London), the University of Surrey (Guildford, England) and Cambridge University. His principal areas of research are spatial statistics, time series analysis, extreme value theory, and Bayesian statistics. Specific areas of expertise include spatial and time series modeling of environmental pollutants, the health effects of atmospheric pollution, the statistics of global climate change, and extreme values in insurance and finance. He is a Fellow of the American Statistical Association and the Institute of Mathematical Statistics, an Elected Member of the International Statistical Institute, and has won the Guy Medal in Silver of the Royal Statistical Society, and the Distinguished Achievement Medal of the Section on Statistics and the Environment, American Statistical Association. In 2004 he was the J. Stuart Hunter Lecturer of The International Environmetrics Society (TIES). He is also a Chartered Statistician of the Royal Statistical Society. Dr. Smith is a statistician with a particular interest in climatology and environmental change with published papers in both statistics and climatology journals on testing the significance of climate trends and their relation to climate models.

John Michael Wallace is a Professor in the Department of Atmospheric Sciences at the University of Washington. Dr. Wallace's research has been directed at improving understanding of global climate and its year-to-year and decade-to-decade variations, making use of observational data. He has contributed to the identification and understanding of a number of atmospheric phenomena, including the vertically propagating planetary waves that drive the quasi-biennial oscillation in zonal winds in the equatorial stratosphere, the 4-5-day period easterly waves that modulate daily rainfall over the tropical oceans, and the dominant spatial patterns in month-to-month and year-to-year climate variability, including the one through which the El Niño phenomenon in the tropical Pacific influences climate over North America. He has also contributed to the methodology for isolating systematic space-time patterns in noisy geophysical data. Presently, Dr. Wallace is attempting to assess the extent to which human activities are contributing to recent climatic trends such as the pronounced wintertime warming over Russia and Alaska. In 2000, he chaired the NRC Panel on Reconciling Temperature Observations. Dr. Wallace is a member of the National Academy of Sciences.

Junhong Wang is a scientist at the National Center for Atmospheric Research. Dr. Wang earned her Ph.D. in Atmospheric Science from Columbia University. Her expertise on climate observations and measurement, especially related to radiosonde observations, and on weather and climate variability is particularly relevant to this study. Dr. Wang has used global radiosonde data to study cloud vertical structure and has been working on understanding and improving radiosonde humidity measurements. Currently, she is continuing her work towards developing future reference radiosondes for global climate observations.

STAFF

Chris Elfring is director of the Polar Research Board (PRB) and the Board on Atmospheric Sciences and Climate (BASC). She is responsible for all aspects of strategic planning, project development and oversight, financial management, and personnel for both units. Since joining the PRB in 1996, Ms. Elfring has overseen or directed studies that produced the following reports: *Frontiers in Polar Biology in the Genomics Era* (2003), *Cumulative Environmental Impacts of Oil and Gas Activities on Alaska's North Slope* (2003), *A Century of Ecosystem Science: Planning Long-term Research in the Gulf of Alaska* (2002), and *Enhancing NASA's Contributions to Polar Science* (2001). In addition, she is responsible for the Board's activities as the U.S. National Committee to the Scientific Committee on Antarctic Research.

Parikhit Sinha was a Program Officer for the Board on Atmospheric Sciences and Climate (BASC) until April 2005. He received a B.A. in environmental engineering sciences from Harvard University and a Ph.D. in atmospheric sciences from the University of Washington, Seattle. His doctorate research involved airborne measurements and chemical transport modeling of trace gas and particle emissions from savanna fires in southern Africa. Since joining the National Academies in 2004, he has worked on studies addressing radiative forcing of climate change, rapid climate variability and change in Asia, and climate change indicators in the United States.

Rachael Shiflett is a senior program assistant with the Polar Research Board. She received her M.Sc. in environmental law from Vermont Law School in 2001 and will complete her J.D. at Catholic University in May 2007. Ms. Shiflett has coordinated National Research Council studies that produced the reports: *A Vision for the International Polar Year 2007-2008* (2004), and *International Polar Year 2007-2008 Report of the Implementation Workshop* (2004).