

1
2
3 **Climate Change Science Program**
4 **Temperature Trends in the Lower Atmosphere**
5 *Understanding and Reconciling Differences*
6
7

8
9 **TABLE OF CONTENTS**

10
11 **Preface**

12
13 **Executive Summary**

14
15 **Chapters:**

- 16
17 **1** Why do temperatures vary vertically (from the surface to the stratosphere) and what do we understand
18 about why they might vary and change over time?
19
20 **2** What kinds of atmospheric temperature variations can the current observing systems detect and what are
21 their strengths and limitations, both spatially and temporally?
22
23 **3** What do observations indicate about the changes of temperature in the atmosphere and at the surface since
24 the advent of measuring temperatures vertically?
25
26 **4** What is our understanding of the contribution made by observational or methodological uncertainties to
27 the previously reported vertical differences in temperature trends?
28
29 **5** How well can the observed vertical temperature changes be reconciled with our understanding of the
30 causes of these temperature changes?
31
32 **6** What measures can be taken to improve the understanding of observed changes?
33

34 **Appendix A:** Statistical Issues Regarding Trends:

35 **Appendix B:** Members of the Assessment/Synthesis Product Team
36

37 **Glossary & Acronyms, Symbols, Abbreviations**
38
39
40
41

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

PREFACE

Report Motivation and Guidance for Using this Synthesis/Assessment Report

A primary objective of the U.S. Climate Change Science Program (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 has identified an initial set of 21 synthesis and assessment products that address its highest priority research, observation, and decision-support needs. This Synthesis/Assessment Report, the first of the 21 Reports, focuses on understanding the causes of the reported differences between independently produced data sets of atmospheric temperature trends from the surface through the lower stratosphere.

Background

Measurements of global surface air temperature show substantial increases over the past several decades. In the early 1990s, data from NOAA's polar orbiting satellites were analyzed for multi-decadal trends. These initial analyses indicated that temperatures in the troposphere showed little or no increase, in contrast with surface air measurements from ships, land-based weather stations, and ocean buoys. This result led some to question the reality and/or the cause of the surface temperature increase, on the basis that human influences, thought to be important contributors to observed change, were expected to increase temperatures both at the surface and in

24 the troposphere with larger increases expected in the tropical troposphere. This
25 surprising result led to an intensive effort by climate scientists to better understand
26 the causes of the apparent differences in the rates of temperature increase between the
27 surface and the troposphere.

28

29 Scientists analyzing the data knew that there were complex and unresolved issues
30 related to inadequacies of observing systems that could lead to misinterpretation of
31 the data. There were also uncertainties in our understanding of how the climate might
32 respond to various forcings as is often assessed through the use of climate models. In
33 an attempt to resolve these issues, in 2000 the National Research Council specifically
34 addressed the general issue of troposphere and surface derived temperature trends. In
35 its Report, the NRC concluded that “the warming trend in global-mean surface
36 temperature observations during the past 20 years is undoubtedly real and is
37 substantially greater than the average rate of warming during the twentieth century.
38 The disparity between surface and upper air trends in no way invalidates the
39 conclusion that surface temperature has been rising.” The NRC further found that
40 corrections in the Microwave Sounding Unit (MSU) processing algorithms brought
41 the satellite data record into slightly closer alignment with surface temperature trends.
42 They concluded that the substantial disparity that remains probably reflects a less
43 rapid warming of the troposphere than the surface in recent decades due to both
44 natural and human-induced causes.

45

46 In 2001, the Intergovernmental Panel on Climate Change (IPCC) Third Assessment
47 Report devoted additional attention to new analyses of the satellite, weather balloon,
48 and surface data to evaluate the difference in temperature trends between the surface
49 and the troposphere. Similar to the NRC, the IPCC concluded that it was very likely
50 that the surface temperature increases were larger and differed significantly from
51 temperature increases higher in the troposphere. They concluded, “during the past
52 two decades, the surface, most of the troposphere, and the stratosphere have
53 responded differently to climate forcings because different physical processes have
54 dominated in each of the regions during that time.” (IPCC; Climate Change 2001 The
55 Scientific Basis, Chapter 2, p. 122-123; Cambridge University Press).

56

57 **Focus of this Synthesis/Assessment Report**

58

59 The efforts of the NRC and IPCC to address uncertainties about the temperature
60 structure of the lower atmosphere (i.e., from the surface through the lower
61 stratosphere) have helped move us closer to a comprehensive understanding of
62 observed trends of temperature. Although these documents provided a great deal of
63 useful information, full resolution of the issue was hampered by the complexities
64 coupled with shortcomings of the available observing systems. To more fully address
65 remaining fundamental questions, a broader examination has been undertaken here to
66 answer the following questions:

- 67 1) Why do temperatures vary vertically (from the surface to the stratosphere)
68 and what do we understand about why they might vary and change over
69 time?

70

- 71 2) What kinds of atmospheric temperature variations can the current
72 observing systems measure and what are their strengths and limitations,
73 both spatially and temporally?
74
- 75 3) What do observations indicate about the changes of temperature in the
76 atmosphere and at the surface since the advent of measuring temperatures
77 vertically?
78
- 79 4) What is our understanding of the contribution made by observational or
80 methodological uncertainties to the previously reported vertical differences
81 in temperature trends?
82
- 83 5) How well can the observed vertical temperature changes be reconciled with
84 our understanding of the causes of these changes?
85
- 86 6) What measures can be taken to improve the understanding of observed
87 changes?
88

89 These questions provide the basis for the six main chapters in this
90 Synthesis/Assessment Report (the chapter numbers correspond to the question
91 numbers above). They highlight several of the fundamental uncertainties and
92 differences between and within the individual components of the existing
93 observational and modeling systems. The responses to the questions are written in a
94 style consistent with major international scientific assessments (e.g., IPCC
95 assessments, and the Global Ozone Research and Monitoring Project of the World
96 Meteorological Organization). The Executive Summary, which presents the key
97 findings from the main body of the Report, is intended to be useful for those involved
98 with the policy-related global climate change issues. The Chapters supporting the
99 Executive Summary are written at a more technical level suitable for non-climate
100 specialists within the scientific community and well-informed lay audiences.

101

102 To help answer the questions posed, climate model simulations of temperature change
103 based on time histories of the forcing factors thought to be important, have been
104 compared with observed temperature changes. If the models replicate the observed
105 temperature changes, this increases confidence in our understanding of the observed
106 temperature record and reduces uncertainties about projected changes. If not, then this
107 implies that the time histories of the important forcings are not adequately known, all
108 of the important forcings are not included, the processes being simulated in the
109 models have serious flaws, the observational record is incorrect, or some combination
110 of these factors.

111

112 This U.S. Climate Change Science Program Assessment/Synthesis Report assesses
113 the uncertainties associated with the data used to determine changes of temperature,
114 and whether such changes are consistent with our understanding of climate processes.
115 This requires a detailed comparison of observations and climate models used to
116 simulate observed changes, including an appreciation of why temperatures might
117 respond differently at the surface compared to higher levels in the atmosphere.

118

119 This CCSP Report addresses the accuracy and consistency of the temperature records
120 and outlines steps necessary to reconcile differences between individual data sets.
121 Understanding exactly how and why there are differences in temperature trends
122 reported by several analysis teams using different observation systems and analysis
123 methods is a necessary step in reducing the uncertainties that underlie current efforts
124 focused on the detection and quantification of surface and tropospheric temperature

125 trends.

126

127 **New observations and analysis since the IPCC and NRC Reports**

128

129 Since the IPCC and NRC assessments, there have been intensive efforts to create new
130 satellite and weather balloon data sets using a range of approaches. Having multiple
131 satellite data sets provides the opportunity for much greater understanding of
132 observed changes and their uncertainty than was possible in the previous assessments.
133 In addition, for the first time a suite of models simulating observed climate since
134 1979 (when satellite data began) has provided us a unique opportunity to inter-
135 compare observed trends from various data sets with model simulations using various
136 scenarios of historical climate forcings. Taken together, these advances lead to a
137 much greater understanding of the issues.

138

139 The science of upper air temperature issues is a rapidly evolving field. During the
140 preparation of this Report, new findings were published and have been included in the
141 current draft, causing numerous changes from draft to draft. The authors certainly
142 expect that new data and discoveries that follow the release of this Report, will
143 further improve our understanding. Some open questions originally discussed in the
144 first drafts of this Report were actually resolved during the deliberations. For
145 example, a recent article cleverly demonstrated a subtle problem in the method used
146 in one of the data sets to correct for satellite orbital drift. Since it was possible for the
147 error to be rectified fairly quickly, a new satellite-derived version of lower

148 tropospheric temperatures was available for this Report. At the same time, another
149 research team produced their first version of satellite-derived lower troposphere
150 temperature, and yet another team updated their tropospheric temperature time series
151 as the final drafts were written. All these results are included in this Report.

152

153 Factors that guided the authors in the selection of the climate records considered
154 extensively in this Report were (a) publication heritage, (b) public availability, (c) use
155 by the community at-large, (d) updated on a monthly basis, and (e) period of record
156 beginning in 1979 or earlier. The three surface analyses that were used have many
157 publications covering their construction methods. These data sets are readily
158 available, and are widely used. Two of the three satellite data sets used, while
159 relatively recent, are based on a heritage of published versions which have
160 incorporated new adjustments as discoveries have been made. Each of these data sets
161 allows ready access to the public and has been used in several research publications.
162 A third, more recently developed, data set has been updated during the preparation of
163 this Report. Two data sets used were based on weather balloon data. One of these
164 data sets publicly appeared in 2005, but the authors had made the preliminary
165 versions and methodology available to scientists as early as 2002 and have built upon
166 the extensive experience acquired from previous versions of these data sets. Another
167 data set has a heritage dating back several decades and was recently updated.

168

169 **How to use this Synthesis/Assessment Report**

170

171 This Report promises to be of significant value to decision-makers, and to the expert
172 scientific and stakeholder communities. Readers of this Report will find that new
173 observations, data sets, analyses, and climate model simulations enabled the Author
174 Team to resolve many of the perplexities noted by the NRC and the IPCC in their
175 earlier Reports. The Synthesis/Assessment Report already has had an important
176 impact on the content of the draft to the Fourth Assessment Report of the
177 Intergovernmental Panel on Climate Change (IPCC), due to be published in 2007. In
178 addition, we expect the information generated here will be used both nationally and
179 internationally e.g., by the Global Climate Observing System Atmospheric
180 Observation Panel to help identify effective ways to reduce observational uncertainty.
181 The findings regarding observations and model-observation comparisons of lower
182 stratospheric temperature trends will be useful for the 2006 WMO/UNEP Ozone
183 Assessment.

184

185 Some terms used in the Report may be unfamiliar to those without training in
186 meteorology; a glossary and list of acronyms is thus included at the end of the Report.
187 Two sets of terms are useful to define at the outset since they are particularly
188 fundamental to this Report. This includes a set of terms related to various levels of
189 agreement or disagreement on key issues and findings among the expert Lead
190 Authors as well as terminology describing their considered judgment about the
191 likelihood of critical key results.

192

193

194 To integrate a wide variety of information, this Report also uses a lexicon of terms to
195 express the team's considered judgment about the likelihood of results. Confidence in
196 results is highest at each end of the spectrum. Unless otherwise noted, all statements
197 are certain.

198

199 Preface figure 1

200 This illustration shows the layers of the atmosphere of primary interest to this
201 Synthesis/Assessment Report. The multi-colored line on this diagram indicates the
202 variations in temperature with altitude. The chart beneath the diagram defines the

203 terminology used in this Report for the layers of the atmosphere.

204

205 Preface figure 2

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

Terms for Layers of the Atmosphere Used in this Report

Preface table 1

226
227 **Note: Abbreviated terms** --- Subscript ‘S’, refers to the Surface. Subscripts ‘2’ and ‘4’
228 refer to MSU data from channels 2 and 4. Subscript ‘2LT’ refers to a modification of
229 channel 2 data to focus more directly on the Lower Troposphere and reduce the influence
230 of stratospheric temperatures on channel 2 data. Subscripts ‘850–300’ and ‘100–50’ are
231 specific atmospheric layers sampled by radiosondes. Subscript ‘*_G’ refers to a
232 combination of channel 2 and channel 4 data derived by Fu and co-workers, applicable to

¹ Only about 10% of this layer extends into the lower stratosphere

233 global averages, and ‘*_T’ refers to applicable tropical averages. For the model-
234 observation comparisons, the observation-based definitions as listed in the Table
235 were employed.

236

237

238

239 **The Authoring Team**

240 A full list of this Reports’ authoring team (in addition to a list of lead authors
241 provided at the beginning of each Chapter) is provided in an Appendix at the end of
242 this Report. The focus of this Report follows the Prospectus developed by the Climate
243 Change Science Program and posted on its website at <http://www.climatescience.gov>.

244

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

EXECUTIVE SUMMARY

Convening Lead Author: Tom M.L. Wigley

Lead Authors: V. Ramaswamy, J.R. Christy, J.R. Lanzante, C.A. Mears, B.D. Santer,
and C.K. Folland

22 *New Results and Findings*

23

24 This Report is concerned with temperature changes in the atmosphere, differences in these
25 changes at various levels in the atmosphere, and our understanding of the causes of these
26 changes and differences. Considerable progress has been made since the production of reports by
27 the National Research Council (NRC) and the Intergovernmental Panel on Climate Change
28 (IPCC) in 2000 and 2001. Data sets for the surface and from satellites and radiosondes
29 (temperature sensors on weather balloons) have been extended and improved, and new satellite
30 and radiosonde data sets have been developed¹. Many new model simulations of the climate of
31 the 20th century have been carried out using improved climate models¹ and better estimates of
32 past forcing changes, and numerous new and updated model/observed data comparisons have
33 been performed. The present Report reviews this progress. A summary of the main results is
34 presented first. Then, to address the issues in more detail, six questions that provide the basis for
35 the six main chapters in this Synthesis/Assessment Report are posed and answered.

36

37 The important new results presented in this Report include:

38

39

40 Global Average Temperatures

41

- 42 • Since the late 1950s, the start of the study period for this Report, the surface and troposphere
43 have warmed² substantially, while the stratosphere has cooled². These changes are in accord
44 with our understanding of the effects of radiative forcing agents and with model predictions.
- 45
- 46 • Since the late 1950s, the low and mid troposphere have warmed at a rate slightly faster than
47 the rate of warming at the surface.

48

49 • During the satellite era (1979 onwards), both the low and mid troposphere have warmed. The
50 majority of data sets show warming at the surface that is greater than in the troposphere.

51 Some data sets, however, show the opposite – tropospheric warming that is greater than that
52 at the surface.

53

54 • For global-mean temperature changes in the new climate model simulations, some show
55 more warming in the troposphere than at the surface, while a slightly smaller number of
56 simulations show the opposite behavior. Given the range of observed results and the range of
57 model results, there is no inconsistency between models and observations at the global scale.

58

59 • Studies to detect climate change and attribute its causes using patterns of observed
60 temperature change in space and time (rather than global averages) show clear evidence of
61 human influences on the climate system (due to changes in greenhouse gases, aerosols, and
62 stratospheric ozone).

63

64 • The observed patterns of change cannot be explained by natural processes alone, nor by the
65 effects of short-live species (such as aerosols and tropospheric ozone) alone.

66

67 Tropical Temperatures (20°S to 20°N)

68 • The majority of observed data sets show more warming at the surface than in the
69 troposphere, while some newer observed data sets show the opposite behavior. Almost all
70 model simulations show more warming in the troposphere than at the surface.

71

72 These results characterize important changes in our understanding of the details of temperature
73 changes at the surface and higher in the troposphere. In 2000 and 2001, the NRC and the IPCC
74 both concluded that global-mean surface temperature increases were larger and differed
75 significantly from temperature increases in the troposphere. The new and improved observed
76 data sets and new model simulations that have been developed require modifications of these
77 conclusions.

78

79 The crucial issue here is whether changes in the troposphere are greater or less than those at the
80 surface. Greater changes in the troposphere would mean that changes there are “amplified”
81 relative those at the surface. We use the short-hand notation “amplification” to refer to this
82 possibility. Studies of amplification in the tropics have considered changes on month-to-month,
83 year-to-year and decade-to-decade time scales.

84

85 At the global-mean level, observed changes from 1958 through 2004 exhibit amplification: i.e.,
86 they show greater warming trends in the troposphere compared with the surface. Since 1979,
87 however, the situation is different: most data sets show slightly greater warming at the surface.

88

89 Whether or not these results are in accord with expectations based on climate models is a
90 complex issue, one that we have been able to address more comprehensively now using new
91 model results. Over the period since 1979, the range of recent model simulations is almost
92 evenly divided among those that show greater global-mean warming at the surface and others

93 that show greater warming aloft. Given this range of results, there is no conflict between
94 observed changes and the results from climate models.

95
96 In the tropics, the agreement between models and observations depends on the time scale
97 considered. For month-to-month and year-to-year variations, models and observations both show
98 amplification (i.e., the month-to-month and year-to-year variations are larger aloft than at the
99 surface). The magnitude of this amplification is essentially the same in models and observations.
100 On decadal and longer time scales, however, while almost all model simulations show greater
101 warming aloft, most observations show greater warming at the surface.

102
103 These results have at least two possible explanations, which are not mutually exclusive. Either
104 amplification effects on short and long time scales are controlled by different physical
105 mechanisms, and models fail to capture such behavior; and/or remaining errors in some of the
106 observed tropospheric data sets adversely affect their long-term temperature trends. The second
107 explanation is judged more likely.

108

109 **1. How do we expect vertical temperature profiles to change?**

110 This Section considers the first question:

111 *Why do temperatures vary vertically (from the surface to the stratosphere) and what do we*
112 *understand about why they might vary and change over time?*

113

114 This is addressed in both Chapter 1 and Chapter 5 of this Report.

115

116 In response to this question, Chapter 1 notes the following ...

117

118 (1) Temperatures vary vertically.

119 The effects of solar heating of the surface of the planet combined with the physical properties of
120 the overlying air, lead to the highest temperatures, on average, occurring at the surface. Surface
121 heat is mixed vertically and horizontally, and this mixing, combined with the effects of various
122 physical processes, produces a decrease of temperature with height up to the tropopause
123 (marking the top of the troposphere, i.e., the lower 8 to 16 km of the atmosphere, depending on
124 latitude). Above this, the radiative properties of the air produce a warming with height through
125 the stratosphere (up to about 50 km).

126

127 (2) Temperature trends at the surface can be expected to be different from temperature trends
128 higher in the atmosphere because:

129 • The physical properties of the surface vary substantially according to location and
130 this produces strong horizontal variations in near-surface temperature. Above the
131 surface, these contrasts are quickly smoothed out so the patterns of change in the
132 troposphere must differ from those at the surface. Temperature trend variations with
133 height must, therefore, vary according to location.

134 • Changes in atmospheric circulation or modes of atmospheric variability (e.g., the El
135 Niño-Southern Oscillation [ENSO]) can produce different temperature trends at the
136 surface and aloft.

137 • Under some circumstances, temperatures may increase with height near the surface or
138 higher in the troposphere, producing a "temperature inversion." Such inversions are
139 more common at night, in winter over continents, and in the trade wind regions. Since

140 the air in inversion layers is resistant to vertical mixing, temperature trends can differ
141 between inversion layers and adjacent layers.

142 • Forcing factors, either natural or human-induced³, can result in differing temperature
143 trends at different levels in the atmosphere, and these vertical variations may change
144 over time.

145

146 As noted above, temperatures in the atmosphere vary naturally as a result of internal factors and
147 natural and human-induced perturbations (“forcings”³). These factors are expected to have
148 different effects on temperatures near the surface, in the troposphere, and in the stratosphere, as
149 summarized in Table 1. When all forcings are considered, we expect the troposphere to have
150 warmed and the stratosphere to have cooled since the late 1950s (and over the whole 20th
151 century). The relative changes in the troposphere and stratosphere provide information about the
152 causes of observed changes.

153

154 Table 1: Summary of the most important global-scale climate forcing factors and their likely individual effects on
155 global, annual-mean temperatures; based on Figure 1.3 (which gives temperature information) and Table 1.1 (which
156 gives information on radiative forcing) in Chapter 1, and literature cited in Chapter 1. The stated effects are those
157 that would be expected if the change specified in column 1 were to occur. The top two rows are the primary natural
158 forcing factors, while the other rows summarize the main human-induced forcing factors. The relative importance of
159 these different factors varies spatially and over time.
160
161

162
163

164 Within the troposphere, the relative changes in temperature at different levels are controlled by
165 different processes according to latitude. In the tropics, the primary control is the
166 thermodynamics of moist air (i.e., the effects of evaporation at the surface and the release of
167 latent heat through condensation that occurs in clouds as moist air rises due to convection), and
168 the way these effects are distributed and modified by the atmospheric circulation.
169 Thermodynamic principles require that temperature changes in the tropics will be larger in the
170 troposphere than near the surface (“amplification”), largely independent of the type of forcing. In
171 mid to high latitudes, the processes controlling how temperature changes in the vertical are more
172 complex, and it is possible for the surface to warm more than the troposphere. These issues are
173 addressed in Chapter 1 and Chapter 5.

174

175 **2. Strengths and limitations of the observational data**

176 The second question is:

177 *What kinds of atmospheric temperature variations can the current observing systems detect and*
178 *what are their strengths and limitations, both spatially and temporally?*

179

180 This is addressed in Chapter 2 of this Report. Chapter 2 draws the following main conclusions ...

181

182 (1) The observing systems available for this Report are able to detect small surface and upper air
183 temperature variations from year to year as well as trends⁴ in climate since the late 1950s (and
184 over the last century for surface observations), once the raw data are successfully adjusted for
185 changes over time in observing systems and practices, and micro-climate exposure.

186 Measurements from all systems require such adjustments. This Report relies solely on adjusted
187 data sets.

188

189 (2) Independently-performed adjustments to the land surface temperature record have been
190 sufficiently successful that trends given by different data sets are very similar on large (e.g.,
191 continental) scales. This conclusion holds to a lesser extent for the ocean surface record.

192

193 (3) Adjustments for changing instrumentation are most challenging for upper-air datasets. While
194 these show promise for trend analysis, it is not clear that current upper-air climate records have
195 achieved sufficient accuracy to resolve trend-related scientific questions.

- 196 • Upper-air datasets have been subjected to less scrutiny than surface datasets.
- 197 • Adjustments are complicated, large compared to the linear trend signal, involve
198 expert judgments, and cannot be stringently evaluated because of lack of traceable
199 standards.
- 200 • Unlike surface trends, reported upper-air trends vary considerably between research
201 teams beginning with the same raw data owing to their different decisions on how to
202 remove non-climatic factors.

203

204 Many different methods are used to measure temperature changes at the Earth's surface and at
205 various levels in the atmosphere. Near-surface temperatures have been measured for the longest
206 period, over a century, and are measured directly by thermometers. Over land, these data come
207 from fixed meteorological stations. Over the ocean, measurements are of both air temperature
208 and sea-surface (top 10 meters) temperature taken by ships or from buoys.

209

210 The next-longest records are upper-air data measured by radiosondes (temperature sensors
211 carried aloft by weather balloons). These have been collected routinely since 1958. There are still
212 substantial gaps in radiosonde coverage.

213

214 Satellite data have been collected for the upper air since 1979 with almost complete global
215 coverage. The most important satellite records come from Microwave Sounding Units (MSU) on
216 polar orbiting satellites. The microwave data from MSU instruments require calculations and
217 adjustments in order to be interpreted as temperatures. Furthermore, these satellite data do not
218 represent the temperature at a particular level, but, rather, the average temperature over thick
219 atmospheric layers (see Figure 2.2 in Chapter 2). Channel 2 data (mid to upper troposphere, T_2)
220 have a latitudinally-dependent contribution from the stratosphere, while Channel 4 data (lower
221 stratosphere, T_4) have a latitudinally-dependent contribution from the troposphere, factors that
222 complicate their interpretation.

223

224 All measurement systems have inherent uncertainties associated with: the instruments employed;
225 changes in instrumentation; and the way local measurements are combined to produce area
226 averages. All data sets require careful examination for instrument biases and reliability, and
227 adjustments to remove changes that might have arisen for non-climatic reasons. The term
228 “homogenization” is used to describe this adjustment procedure. Recent improvements in and
229 corrections to some of these adjustments have resulted in better agreement between data sets.

230

231

232 **3. What temperature changes have been observed?**

233 This Section combines information related to questions 3 and 4:

234 *What do observations indicate about the changes of temperature in the atmosphere and at*
235 *the surface since the advent of measuring temperatures vertically?*

236
237 *What is our understanding of the contribution made by observational or methodological*
238 *uncertainties to the previously reported vertical differences in temperature trends?*
239

240 These questions are addressed in Chapters 3 and 4 of this Report. The following conclusions are
241 drawn in these chapters. Supporting information is given in Figure 1 and Figure 2.

242

243 (1) Surface temperatures: For global-mean changes, as well as in the tropics (20°S to 20°N), all
244 data sets show warming at the surface since 1958, with a greater rate of increase since 1979.

245 Differences between the data sets are small.

246

247 • Global-mean temperature increased at about 0.12°C per decade since 1958, and about
248 0.16°C per decade since 1979. In the tropics, temperature increased at about 0.11°C per
249 decade since 1958, and about 0.13°C per decade since 1979.

250

251 • Local biases in surface temperatures may exist due to changes in station exposure and
252 instrumentation over land⁵, or changes in measurement techniques by ships and buoys in
253 the ocean. It is likely that these biases are largely random and therefore cancel out over
254 large regions such as the globe or tropics, the regions that are of primary interest to this
255 Report.

256

257 • Errors in observed surface/troposphere trend differences are more likely to come from
258 errors in tropospheric data than from errors in surface data.

259

260 (2) Tropospheric temperatures: All data sets show that the global-mean and tropical troposphere
261 has warmed from 1958 to the present, with the warming in the troposphere being slightly more
262 than at the surface. Since 1979, due to the considerable disagreements between tropospheric data
263 sets, it is not clear whether the troposphere has warmed more than or less than the surface.

264

- 265 • Global-mean tropospheric temperature increased at about 0.14°C per decade since 1958,
266 and between 0.10°C and 0.20°C per decade since 1979. In the tropics, temperature
267 increased at about 0.13°C per decade since 1958, and between 0.02°C and 0.19°C per
268 decade since 1979.

269

- 270 • It is very likely that trends in troposphere temperatures are affected by errors that remain
271 in the homogenized radiosonde data sets. Such errors arise because the methods used to
272 produce these data sets are only able to detect and remove the more obvious cases, and
273 involve many subjective decisions. The full consequences of these errors for large-area
274 averages, however, have not yet been fully resolved. Nevertheless, it is likely that a net
275 spurious cooling corrupts the area-averaged homogenized radiosonde data in the tropical
276 troposphere, causing these data to indicate less warming than has actually occurred there.

277

- 278 • For tropospheric satellite data, a primary cause of trend differences between different
279 versions is differences in how the data from different satellites are merged together.
280 Corrections required to account for drifting measurement times, and diurnal cycle
281 adjustments are also important.

282

- 283 • Comparisons between satellite and radiosonde temperatures for the mid to upper
284 tropospheric layer (MSU channel 2; T₂) are very likely to be corrupted by excessive
285 stratospheric cooling in the radiosonde data

286

287 (3) Lower stratospheric temperatures: All data sets show that the stratosphere has cooled
288 considerably from 1958 and from 1979 to the present, although there are large differences in the
289 linear trend values from different data sets.

290

- 291 • The largest differences between data sets are in the stratosphere, particularly between the
292 radiosonde and satellite-based data sets. It is very likely that the satellite/radiosonde
293 discrepancy arises primarily from uncorrected errors in the radiosonde data.

294 Figure 1 shows the various temperature time series examined in this Report.

295
296 **Figure 1:** Observed surface and upper air global-mean temperature records. From top to bottom: A, lower
297 stratosphere (denoted T_4) records from two satellite analyses (UAH and RSS) together with equivalently-weighted
298 radiosonde records based on HadAT2 and RATPAC data; B, mid- to upper-troposphere (T_2) records from three
299 satellite analyses (UAH, RSS and U.Md.) together with equivalently-weighted radiosonde records based on HadAT2
300 and RATPAC; C, lower troposphere (T_{2LT}) records from UAH and RSS (satellite), and from HadAT2 and RATPAC
301 (equivalently-weighted radiosonde); D, surface (T_S). All time series are based on monthly-mean data smoothed with
302 a 7-month running average, expressed as departures from the Jan. 1979 to Dec. 1997 average. Note that the T_2 data
303 (panel B) contain a small contribution (about 10%) from the lower stratosphere. Information here is from Figures
304 3.1, 3.2 and 3.3 in Chapter 3.
305

306 For the lower stratosphere, the cooling trend since the late 1950s (which is as expected due to the
307 effects of greenhouse-gas concentration increases and stratospheric ozone depletion) is
308 punctuated by short-term warming events associated with the explosive volcanic eruptions of Mt.
309 Agung (1963), El Chichón (1982) and Mt. Pinatubo (1991).

310

311 Both the troposphere and the surface show warming since the late 1950s. For the surface, most,
312 if not all of the temperature increase since 1958 occurs starting around the mid-1970s, a time
313 coincident with a previously identified climate regime shift. However, there does not appear to
314 be a strong jump up in temperature at this time; rather, the major part of the rise seems to occur
315 in a more gradual fashion. For the balloon-based tropospheric data, the major part of the
316 temperature increase since 1958 appears in the form of a rapid rise in the mid-1970s, apparently
317 in association with the climate regime shift that occurred at this time.

318

319 The dominant shorter time scale fluctuations are those associated with the El Niño-Southern
320 Oscillation phenomenon (ENSO). The major ENSO warming event in 1998 is obvious in all
321 records. Cooling following the eruptions of Mt. Agung and Mt. Pinatubo is also evident, but the
322 cooling effect of El Chichón is masked by an ENSO warming that occurred at the same time.

323 The changes following volcanic eruptions (i.e., surface and tropospheric cooling and
324 stratospheric warming) are consistent with our physical understanding and with model
325 simulations.

326

327 Global-mean temperature changes over the periods 1958 through 2004 and 1979 through 2004
328 are shown in Figure 2 in degrees Celsius and degrees Fahrenheit.

329

330

331

332

333

334

335

336

337

338

339

340

341

Figure 2: Total global-mean temperature changes for the surface and different atmospheric layers, from different data sets and over two periods, 1958 to 2004 and 1979 to 2004. The values shown are the total change over the stated period in both degrees Celsius (degC; lower scales) and degrees Fahrenheit (degF; upper scales). All changes are statistically significant at the 5% level except RSS T₄ and RATPAC, HadAT2 and UAH T₂. Total change in degC is the linear trend in degC per decade (see Tables 3.2 and 3.3 in Chapter 3) times the number of decades in the time period considered. Total change in degF is this number times 1.8 to convert to degF. For example, the Table 3.2 trend for NOAA surface temperatures over January 1958 through December 2004 is 0.11°C/decade. The total change is therefore 0.11 times 4.7 decades to give a total change of 0.53°C, Multiplying this by 1.8 gives a total change in degrees Fahrenheit of 0.93°F. Warming is shown in red, and cooling in blue.

342

343 **4. Are model simulations consistent with the observed temperature changes?**

344 Computer-based climate models encapsulate our understanding of the climate system and the
345 driving forces that lead to changes in climate. Such models are the only tools we have for
346 estimating the likely patterns of response of the climate system to different forcing mechanisms.

347 The crucial test of our understanding is to compare model simulations with observed changes.

348 The fifth question therefore is:

349 *How well can the observed vertical temperature changes be reconciled with our*
350 *understanding of the causes of these changes?*

351

352 This question is addressed in Chapter 5 of this Report. Chapter 5 draws the following

353 conclusions ...

354

355 **PATTERN STUDIES**

356

357 (1) Results from many different pattern-based “fingerprint”⁵ studies (see Box 5.5 in Chapter 5)
358 provide consistent evidence for human influences on the three-dimensional structure of
359 atmospheric temperature changes over the second half of the 20th century.

360

- 361 • Fingerprint studies have identified greenhouse gas and sulfate aerosol signals in observed
362 surface temperature records, a stratospheric ozone depletion signal in stratospheric
363 temperatures, and the combined effects of these forcing agents in the vertical structure of
364 atmospheric temperature changes.

365

366 (2) Natural factors have influenced surface and atmospheric temperatures, but cannot fully
367 explain their changes over the past 50 years.

368

369 LINEAR TREND COMPARISONS^{4,6}

370

371 (3) When models are run with natural and human-induced forcings, simulated global-mean
372 temperature trends for individual atmospheric layers are consistent with observations.

373

374 (4) Comparing trend differences between the surface and the troposphere exposes potential
375 model/observed data discrepancies in the tropics.

376

- 377 • In the tropics, the majority of observational data sets show more warming at the surface
378 than in the troposphere, while almost all model simulations have larger warming aloft
379 than at the surface.

380

381 AMPLIFICATION OF SURFACE WARMING IN THE TROPOSPHERE

382

383 (5) Amplification means that temperatures show larger changes aloft than at the surface.

384 In the tropics, on monthly and inter-annual time scales, both models and observations show

385 amplification of temperature variability in the troposphere relative to the surface. This

386 amplification is of similar magnitude in models and observations. For multi-decadal trends,

387 models show the same amplification that is seen on shorter time scales. A number of observed

388 data sets, however, do not show this amplification.

389

- 390 • These results have several possible explanations, which are not mutually exclusive. One
391 explanation is that “real world” amplification effects on short and long time scales are
392 controlled by different physical mechanisms, and models fail to capture such behavior. A
393 second explanation is that remaining errors in some of the observed tropospheric data sets
394 adversely affect their long-term temperature trends. The second explanation is more
395 likely in view of the model-to-model consistency of amplification results, the large
396 uncertainties in observed tropospheric temperature trends, and independent physical
397 evidence supporting substantial tropospheric warming.

398

399 OTHER FINDINGS

400

- 401 (6) Because of differences between different observed data sets, it is important to account for
402 observational uncertainty in comparisons between modeled and observed temperature changes.

403

- 404 • Large “construction” uncertainties in observed estimates of global-scale atmospheric
405 temperature change can critically influence the outcome of consistency tests between
406 models and observations.

407

- 408 (7) Inclusion of previously-ignored spatially-heterogeneous forcings in the most recent climate
409 models does not fundamentally alter conclusions about the amplification of warming in the
410 troposphere relative to the surface.

411

412 • Changes in sulfate aerosols and tropospheric ozone, which have spatially-heterogeneous
413 forcings, have been incorporated routinely in climate model experiments for a number of
414 years. It has been suggested that the spatially-heterogeneous forcing effects of black
415 carbon aerosols and land use/land cover may have had significant effects on regional
416 temperatures that might modify previous conclusions regarding vertical temperature
417 changes. These forcings have been included for the first time in about half of the global
418 model simulations considered here. Within statistical uncertainties, model simulations
419 that include these forcings show the same amplification of warming in the troposphere
420 relative to the surface at very large spatial scales (global and tropical averages) as
421 simulations in which these forcings are neglected.

422

423 Chapter 5 analyses state-of-the-art model simulations from 19 institutions globally, run using
424 combinations of the most important natural and human-induced forcings. The Chapter compares
425 the results of these simulations with a number of different observational data sets for the surface
426 and different atmospheric layers, resulting in a large number of possible model/observed data
427 comparisons.

428

429 Figures 3 and 4 summarize the new model results used in this Report, together with the
430 corresponding observations. Figure 3 gives global-mean results, while Figure 4 gives results for
431 the tropics (20°S to 20°N). Model and observed results are compared in these Figures using
432 linear trends over the period January 1979 through December 1999⁷ for the surface, for
433 individual layers, and (right-hand panels) for surface changes relative to the troposphere.

434 Rectangles are used to illustrate the ranges of both model trends (red rectangles) and observed
435 trends (blue rectangles). Individual observed-data trends are also shown.
436
437 Since statistical uncertainties (see Appendix) are not shown in these Figures, the rectangles do
438 not represent the full ranges of uncertainty. However, they allow a meaningful first-order
439 assessment of model/observed similarities and differences. Fully overlapping rectangles indicate
440 consistency, partially overlapping rectangles point to possible discrepancies, while rectangles
441 that either do not overlap or show minimal overlap indicate important model/observed data
442 inconsistencies. At the global-mean level, models and observations generally show fully
443 overlapping rectangles. The only potentially serious inconsistency is in the tropics (Figure 4)
444 where the troposphere warms more rapidly than the surface in all except two of the 49 individual
445 model simulations examined here, while, in the majority of observational data sets, the surface
446 has warmed more rapidly than the troposphere.

447
448
449
450
451
452
453
454
455
456
457
458
459
460
461

Figure 3: Comparison of observed and model-simulated global-mean temperature trends (left-hand panels) and trend differences (right-hand panels) over January 1979 through December 1999, based on Table 5.4A and Figure 5.3 in Chapter 5. The upper red rectangles in each box show the range of model trends from 49 model simulations. The lower blue rectangles show the range of observed trends, with the individual trends from different data sets indicated by the symbols. From bottom to top, the left-hand panels show trends for the surface (T_S), the lower troposphere (T_{2LT}), the troposphere (T^*), the mid troposphere to lower stratosphere (T_2), and the lower stratosphere (T_4). The right-hand panels show differences in trends between the surface and either the troposphere or the lower troposphere, with a positive value indicating a stronger warming at the surface. The red vertical lines show the average of all model results. The vertical black dashed lines show the zero value. For the observed trend differences, there are eight values corresponding to combinations of the four upper-air data sets (as indicated by the symbols) and either the HadCRUT2v surface data or the NASA/NOAA surface data (which have almost identical trends).

462
463
464
465

Figure 4: As Figure 3, but for the tropics (20°S to 20°N), based on Table 5.4B and Figure 5.4 in Chapter 5.

466
467

468

469

470 **5. Recommendations**

471 This Section addresses question 6:

472 *What measures can be taken to improve the understanding of observed changes?*

473

474 In answer to this question, drawing on the material presented in the first five chapters of this

475 Report, a set of primary recommendations has been developed and is described in detail in

476 Chapter 6. It should be noted that, rather than invent new proposals or recommendations, the

477 items described in Chapter 6 expand and build upon existing ideas, emphasizing those that are

478 considered to be of highest utility. The seven recommendations are:

479

480 (1) In order to encourage further independent scrutiny, data sets and their full metadata (i.e.,

481 information about instrumentation used, observing practices, the environmental context of

482 observations, and data-processing procedures) should be made openly available. Comprehensive

483 analyses should be carried out to ascertain the causes of remaining differences between data sets

484 and to refine uncertainty estimates.

485

486 (2) Efforts should be made to archive and make openly available surface, balloon-based, and

487 satellite data and metadata that have not previously been exploited. Emphasis should be placed

488 on the tropics.

489

490 (3) Efforts should be made to create climate quality data sets⁸ for a range of variables other than

491 temperature. These data sets should subsequently be compared with each other and with

492 temperature data to determine whether they are consistent with our physical understanding.

493

494 (4) Efforts should be made to create several homogeneous atmospheric reanalyses⁹. Particular
495 care needs to be taken to identify and homogenize critical input data. Identification of critical
496 data requires, in turn, observing system experiments where the impacts and relative importance
497 of different observation types from land, radiosonde, and space-based observations are assessed.

498

499 (5) Models that appear to include the same forcings often differ in both the way the forcings are
500 quantified and how these forcings are applied to the model. Hence, efforts are required to
501 separate more formally uncertainties arising from model structure from the effects of forcing
502 uncertainties. This requires running multiple models with standardized forcings, and running the
503 same models individually under a range of plausible scenarios for each forcing.

504

505 (6) The GCOS (Global Climate Observing System) climate monitoring principles should be fully
506 adopted. In particular, when any particular type of instrument is changed or re-sited, there should
507 be a period of overlap between old and new instruments or configurations that is sufficient to
508 allow analysts to adjust for the change with small uncertainties that do not prejudice the analysis
509 of climate trends. The minimum period is a full annual cycle of the climate.

510

511 (7) A small subset (about 5%) of the operational radiosonde network should be developed and
512 implemented as reference sites for all kinds of climate data from the surface to the stratosphere.

513

514

515 **Footnotes**

516

517 ¹ For details of new observed data see Table 3.1 in Chapter 3. For details of new models and
518 model simulations see Chapter 5 and http://www-pcmdi.llnl.gov/ipcc/model_documentation.

519

520 ² We use the words “warming” and “cooling” here to refer to temperature increases or decreases,
521 as is common usage. Technically, these words refer to changes in heat content, which may occur
522 through changes in either the moisture content and/or the temperature of the atmosphere. When
523 we say that the atmosphere has warmed (or cooled) over a given period, this means that there has
524 been an overall positive (or negative) temperature change based on a linear trend analysis.

525

526 ³ The main natural perturbations are changes in solar output and the effects of explosive volcanic
527 eruptions. The main human-induced (“anthropogenic”) factors are: the emissions of greenhouse
528 gases (e.g., carbon dioxide [CO₂], methane [CH₄], nitrous oxide [N₂O]); aerosols (tiny droplets
529 or particles such as smoke) and the gases that lead to aerosol formation (most importantly, sulfur
530 dioxide); and changes in land cover and land use. Since these perturbations act to drive or
531 “force” changes in climate, they are referred to as “forcings”. Tropospheric ozone [O₃], which is
532 not emitted directly, is also an important greenhouse gas. Tropospheric ozone changes occur
533 through the emissions of gases like carbon monoxide, nitrogen oxides and volatile organic
534 compounds, which are not important directly as greenhouse gases.

535

536 ⁴ Many of the results in this Report (and here in the Executive Summary) are quantified in terms
537 of linear trends, i.e., by the value of the slope of a straight line that is fitted to the data. A simple
538 straight line is not always the best way to describe temperature data, so a linear trend value may
539 be deceptive if the trend number is given in isolation, removed from the original data.
540 Nevertheless, used appropriately, linear trends provide the simplest and most convenient way to
541 describe the overall change over time in a data set, and are widely used. For a more detailed
542 discussion, see the Appendix.

543

544

545 ⁵ Some have expressed concern that land temperature data might be biased due to urbanization
546 effects. Recent studies specifically designed to identify systematic problems using a range of
547 approaches have found no detectable urban influence in large-area averages in the data sets that
548 have been “homogenized” (i.e., adjusted to remove non-climatic influences).

549

550 ⁶ Fingerprint studies use rigorous statistical methods to compare the patterns of observed
551 temperature changes with model expectations and determine whether or not similarities could
552 have occurred by chance. Linear trend comparisons are less powerful than fingerprint analyses
553 for studying cause-effect relationships, but can highlight important differences and similarities
554 between models and observations.

555

556 ⁷ This is the longest period common to all model simulations.

557

558 ⁸ Climate quality data sets are those where the best possible efforts have been made to identify
559 and remove non-climatic effects that might produce spurious changes over time.

560

561 ⁹ Reanalyses are mathematically blended products based upon as many observing systems as
562 practical. Observations are assimilated into a global weather forecasting model to produce
563 globally-comprehensive data sets that are most consistent with both the available data and the
564 assimilation model.

565

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

Chapter 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?

Convening Lead Author: V. Ramaswamy

Lead Authors: James W. Hurrell, Gerald A. Meehl

Contributing Authors: A. Phillips, B. Santer, M. D. Schwarzkopf, D. Seidel, S. Sherwood, P. Thorne

45

46

Summary

47

Temperatures vary vertically.

49 The solar heating of the surface of the planet, combined with the physical properties of the
50 overlying air, produce the highest temperatures, on average, at the surface; that heat is mixed
51 vertically and horizontally by various physical processes. Taking into account the distribution of
52 atmospheric moisture and the lower air pressure with progressively increasing altitude, there
53 results a decrease of temperature with height up to the tropopause. The tropopause marks the top
54 of the troposphere, i.e., the lower 8 to 16 km of the atmosphere depending on latitude. Above this
55 altitude, the radiative properties of the air produce a warming with height through the
56 stratosphere (extending from the tropopause to ~50 km).

57

Temperature trends at the surface can be expected to be different from temperature trends 59 higher in the atmosphere because:

- 60 • Physical properties of the surface vary depending on whether the location has land,
61 sea, snow, or ice. Near the surface, these differing conditions can produce strong
62 horizontal variations in temperature. Above the surface layer, these contrasts are
63 quickly smoothed out, contributing to varying temperature trends with height at
64 different locations.
- 65 • Changes in atmospheric circulation or modes of atmospheric variability (e.g., El
66 Niño-Southern Oscillation [ENSO]) can produce different temperature trends at the
67 surface and aloft.

- 68 • Under some circumstances, temperatures may increase with height near the surface or
69 higher in the troposphere, producing a "temperature inversion." Such inversions are
70 more common at night, in winter over continents, and in the trade wind regions. Since
71 the air in inversion layers is resistant to vertical mixing, temperatures trends can differ
72 between inversion layers and adjacent layers.
- 73 • Forcing factors, either natural (e.g., volcanoes and solar) or human-induced (e.g.,
74 greenhouse gas, aerosols, ozone, and land use) can result in differing temperature
75 trends at different altitudes, and these vertical variations may change over time. This
76 can arise due to spatial changes in the concentrations or properties of the forcing
77 agents.

78

79

80 This Chapter describes the temperature profile of the layers of the atmosphere from the surface
81 through the stratosphere and discusses the basic reasons for this profile. We also use results from
82 global climate model simulations to show how changes in natural and human-induced factors can
83 produce different temperature trends in the various layers of the atmosphere. This discussion
84 provides the background for the presentation of the observed changes (Chapters 2-4), and for the
85 understanding of their causes (Chapter 5). We also describe temperature changes in the
86 stratosphere in recent decades and the influences of these changes on the troposphere. Finally,
87 making use of surface and satellite observations, we examine the physical processes that can
88 result in different temperature trends at the surface and in the troposphere.

89

90 **1.1 The Thermal Structure of the Atmosphere**

91 Surface temperatures are at their warmest in the tropics, where the largest amount of solar
92 radiation is received during the course of the year, and decrease towards the Polar Regions where
93 the annual-mean solar radiation received is at a minimum (Oort and Peixoto, 1992). The
94 temperature contrast between summer and winter increases from the equator to the poles. Since
95 land areas heat up and cool more rapidly than oceans, and because of the preponderance of land
96 in the Northern Hemisphere, there is a larger contrast between summer and winter in the
97 Northern Hemisphere.

98
99 Figure 1.1 illustrates the climatological vertical temperature profiles for December, January,
100 February and June, July, August mean conditions, as obtained from the National Centers for
101 Environmental Prediction (NCEP) reanalyses (Kalnay et al., 1996; updated). It is convenient to
102 think first in terms of climatological conditions upon which spatial and temporal
103 variations/trends are superimposed. The solid line in the plot illustrates the tropopause, which
104 separates the troposphere below from the stratosphere above. The tropopause is at its highest
105 level in the tropics ($\sim 20^{\circ}\text{N}$ - 20°S). It descends sharply in altitude from ~ 16 km at the equator to
106 ~ 12 km at ~ 30 - 40° latitude, and to as low as about 8 km at the poles.

107

108 Figure 1.1. Global climatological vertical temperature profiles from surface to troposphere and extending into the
109 stratosphere for December-January-February and June-July-August mean conditions, as obtained from the National
110 Centers for Environmental Prediction reanalyses (Kalnay et al., 1996; updated). The solid line denotes the
111 tropopause which separates the stratosphere from the surface-troposphere system.

112

113 Temperatures generally decrease with height from the surface although there are important
114 exceptions. The rate at which the temperature changes with height is termed the “lapse rate.” The
115 lapse rate can vary with location and season, and its value depends strongly on the atmospheric

116 humidity, e.g., the lapse rate varies from $\sim 5^{\circ}\text{C}/\text{km}$ near the surface in the moist tropical regions
117 (near the equator) to much larger values ($\sim 8\text{-}9^{\circ}\text{C}/\text{km}$) in the drier subtropics ($\sim 20\text{-}30^{\circ}$). Important
118 departures from nominal lapse rate values can occur near the surface and in the upper
119 troposphere. In the equatorial tropics, the tropopause region (~ 16 km) is marked by a smaller
120 value of the lapse rate than in the lower troposphere.

121
122 The thermal structure of the lowest 2-3 km, known as the “planetary boundary layer,” can be
123 complicated, even involving inversions (in which temperature increases rather than decreases
124 with height) occurring at some latitudes due to land-sea contrasts, topographic influences,
125 radiative effects and meteorological conditions. Inversions are particularly common during
126 winter over some middle and high latitude land regions and are a climatological feature in the
127 tropical trade wind regions. The presence of inversions acts to decouple surface temperatures
128 from tropospheric temperatures on daily or even weekly timescales.

129
130 Above the tropopause is the stratosphere, which extends to ~ 50 km and where the temperature
131 increases with height. In the vicinity of the tropical tropopause, (i.e., the upper troposphere and
132 lower stratosphere regions, $\sim 15\text{-}18$ km), the temperature varies little with height. The
133 extratropical (poleward of 30°) lower stratosphere (at $\sim 8\text{-}12$ km) also exhibits a similar feature
134 (Holton, 1979). The lapse rate change with altitude in the upper troposphere/lower stratosphere
135 region is less sharp in the extratropical latitudes than in the tropics.

136
137 The global temperature profile of the atmosphere reflects a balance between the radiative and
138 dynamical heating/cooling of the surface-atmosphere system. From a global, annual-average

139 point of view, the thermal profile of the stratosphere is the consequence of a balance between
140 radiative heating and cooling rates due to greenhouse gases, principally carbon dioxide (CO₂),
141 ozone (O₃) and water vapor (H₂O) (Andrews et al., 1987). The vertical profile of the troposphere
142 is the result of a balance of radiative processes involving greenhouse gases, aerosols, and clouds
143 (Stephens and Webster, 1981; Goody and Yung, 1989), along with the strong role of moist
144 convection (Holton, 1979; Kiehl, 1992). An important difference between the troposphere and
145 stratosphere is that the stratosphere is characterized by weak vertical motions, while in the
146 troposphere, the vertical motions are stronger. Most significantly, the moist convective processes
147 that are a characteristic feature of the troposphere include the transfer of large amounts of heat
148 due to evaporation or condensation of water.

149
150 Convective processes are important in determining the temperature profile in the troposphere.
151 This is illustrated by the fact that radiative processes alone would cause the surface to be
152 significantly warmer than it is actually. This would occur because the atmosphere is relatively
153 transparent to the Sun's radiation. By itself, this would lead to a drastic heating of the surface
154 accompanied by a net radiative cooling of the atmosphere (Manabe and Wetherald, 1967).
155 However, the resulting convective motions remove this excess heating from the surface in the
156 form of sensible and latent heat, the latter involving the evaporation of moisture from the surface
157 (Ramanathan and Coakley, 1978). As air parcels rise and expand, they cool due to
158 decompression, leading to a decrease of temperature with height. The lapse rate for a dry
159 atmosphere, when there are no moist processes and the air is rising quickly enough to be
160 unaffected by other heating/cooling sources, is close to 10°C/km. However, because of moist
161 convection, there is condensation of moisture, formation of clouds and release of latent heat as

162 the air parcels rise, causing the lapse-rate to be much less, as low as 4°C/km in very humid
163 atmospheres (Houghton, 1977). In a more rigorous sense, the interactions between radiation,
164 convection, cloud physics, and dynamical motions (ranging from large- to meso- and small-
165 scales) govern the actual rate at which temperature decreases with height (the lapse-rate) at any
166 location. Large-scale dynamical mechanisms tend to homogenize temperatures above the
167 boundary layer over horizontal scales (Rossby radius) that vary from planetary scale near the
168 equator to a couple of thousand kilometers at midlatitudes and to a few hundred kilometers near
169 the poles.

170
171 Convective processes and vertical mixing of air can add complexity to the nominal thermal
172 profile in the tropics mentioned above. For example, a more detailed picture in subtropical
173 regions consists of a surface mixed layer (up to about 500 m) and a trade wind boundary layer
174 (up to about 2 km) above which is the free troposphere. Each of the boundary layers is topped by
175 an inversion which tends to isolate the region from the layer above (Sarachik, 1985). This
176 indicates the limitations in assuming nominal lapse rate values from the surface to the tropopause
177 everywhere.

178
179 The radiative-convective picture above is likely of dominant relevance only for the tropics. In the
180 extra-tropics (poleward of 30°), the lapse rate and tropopause height are mostly determined by
181 instabilities associated with the more familiar weather systems ("baroclinic instability"). The
182 rising motions of air parcels in the equatorial moist tropics associated with deep convection
183 descend in the subtropical regions leading to drier environments there (Hadley circulations). In
184 the polar regions (~60-90°), planetary-scale waves forced by the influences of mountains and

185 that of land-sea contrasts upon the flow of air play a significant role in the determination of the
186 wintertime temperatures at the poles.

187
188 Based on these simple ideas, the lapse rate can be expected to decrease with an increase in
189 humidity, and also to depend on the atmospheric circulation. As specific humidity is strongly
190 related to temperature, and is expected to rise with surface warming, the lapse rate (other things
191 being equal) can be expected to decrease with warming such that temperature changes aloft
192 exceed those at the surface.

193
194 The above simple picture of radiative-convective balance, together with the requirement of
195 radiative balance at the top-of-the-atmosphere (namely, equilibrium conditions wherein the net
196 solar energy absorbed by the Earth's climate system must be balanced by the infrared radiation
197 emitted by the Earth), can help illustrate the significance of long-lived infrared absorbing gases
198 in the atmosphere such as carbon dioxide. The presence of strongly infrared-absorbing
199 greenhouse gases (water vapor, carbon dioxide, methane, etc.) causes the characteristic infrared
200 emission level of the planet to be shifted to a higher altitude where temperatures are colder. The
201 re-establishment of thermal equilibrium leads to warming and communication of the added heat
202 input to the troposphere and surface (Goody and Yung, 1989; Lindzen and Emanuel, 2002).

203
204 In the tropical upper troposphere, moisture- and cloud-related features (e.g., upper tropospheric
205 relative humidity, cirrus cloud microphysics, and mesoscale circulations in anvil clouds) are
206 important factors in governing the thermal profile (Ramanathan et al., 1983; Ramaswamy and
207 Ramanathan, 1989; Donner et al., 2001; Sherwood and Dessler, 2003). In the upper troposphere

208 and especially the stratosphere, convective motions become weak enough that radiative (solar
209 and longwave) heating/cooling become important in establishing the lapse rate.

210

211 **1.2 Forcing of climate change**

212

213 Potentially significant variations and trends are superimposed on the basic climatological thermal
214 profile, as revealed by observational data in the subsequent chapters. While the knowledge of
215 the climatological mean structure discussed in the previous section involves considerations of
216 radiative, convective, and dynamical processes, understanding the features and causes of the
217 magnitude of changes involves a study of the perturbations in these processes which then frame
218 the response of the climate system to the forcing. While the understanding of climate variability
219 is primarily based on observations of substantial changes (e.g., sea-surface temperature changes
220 during El Niño), the vertical temperature changes being investigated in this report are changes on
221 the order of a few tenths of degrees on the global-mean scale (local changes could be much
222 greater), as discussed in the subsequent chapters.

223

224 “Unforced” variations, i.e., changes arising from internally generated variations in the
225 atmosphere-ocean-land-ice/snow climate system, can influence surface and atmospheric
226 temperatures substantially, e.g., due to changes in equatorial sea-surface temperatures associated
227 with ENSO. Climate models indicate that global-mean unforced variations on multidecadal
228 timescales are likely to be smaller than, say, the 20th Century global-mean increase in surface
229 temperature (Stouffer et al., 2000). However, for specific regions and/or seasons, this may not be

230 valid and the unforced variability could be substantial. Chapter 5 provides more detail on models
231 and their limitations (see particularly Box 5.1 and 5.2).

232
233 Because of the influence of radiative processes on the thermal structure, anything external to the
234 climate system that perturbs the planet's radiative heating distribution can cause climate changes,
235 and thus is potentially important in accounting for the observed temperature changes (Santer et
236 al., 1996). The radiative (solar plus longwave) heat balance of the planet can be forced by:

- 237 • natural factors such as changes in the Sun's irradiance, and episodic, explosive volcanic
238 events (leading to a build-up of particulates in the stratosphere);
- 239 • human-induced factors such as changes in the concentrations of radiatively active gases
240 (carbon dioxide, methane, etc.) and aerosols.

241 Potentially important external forcing agents of critical relevance for the surface and atmospheric
242 temperature changes over the 20th Century are summarized in Table 1.1 (for more details, see
243 Ramaswamy et al., 2001; NRC, 2005).

244 Table 1.1. Agents potentially causing an external radiative forcing of climate change in the 20th Century (based on
245 Ramaswamy et al., 2001). See notes below for explanations.
246

247
248 **Notes:**
249 Natural (N) and Anthropogenic (A) sources of the forcing agents. Direct aerosol forcing is to be contrasted with the
250 indirect effects; the latter comprise the so-called first, second, and semi-direct effects. Y denotes a significant
251 component, “small” indicates considerably less important but not negligible, while no entry denotes a negligible

252 component. Forcings other than well-mixed gases and solar are spatially localized, with the degree of localization
253 having considerable variations amongst the different agents, depending on their respective source locations. In
254 addition, for short-lived species such as ozone and aerosols, the long-range meteorological transport plays an
255 important role in their global distributions. Level of confidence is a subjective measure of the certainty in the
256 quantitative estimate of the global-mean forcing.

257 # Typically, the forcing becomes near-global a few months after an intense tropical eruption.

258 * * In the case of volcanic aerosols, the level of confidence in the forcing from the most recent intense eruption, that
259 of Mt. Pinatubo in 1991, is reasonably good because of reliable observing systems in place; for prior explosive
260 eruptions, observations were absent or sparse which affects the reliability of the quantitative estimates for the
261 previous volcanic events.

262 @ Although solar irradiance variations before 1980 have a very low level of confidence, direct observations of the
263 Sun's output from satellite platforms since 1980 are considered to be accurate (Lean et al., 2005). Thus, the forcing
264 due to solar irradiance variations from 1980 to present are known to a much greater degree of confidence than from
265 pre-industrial to present time.

266
267 The forcing agents differ in terms of whether their radiative effects are felt primarily in the
268 stratosphere or troposphere or both, and whether the perturbations occur in the solar or longwave
269 spectrum or both. The quantitative estimates of the forcing due to the well-mixed greenhouse
270 gases (comprised of carbon dioxide, methane, nitrous oxide and halocarbons) are known to a
271 higher degree of scientific confidence than the other forcings. Among aerosols, black carbon is
272 distinct because it strongly absorbs solar radiation (see also Box 5.3). In contrast to sulfate
273 aerosols, which cause a perturbation of solar radiation mainly at the surface (causing a cooling
274 effect), black carbon acts to warm the atmosphere while cooling the surface (Chung et al., 2002;
275 Menon et al., 2002). This could have implications for convective activity and precipitation
276 (Ramanathan et al., 2005), and the lapse rate (Chung et al., 2002; Erlick and Ramaswamy, 2003).
277 The response to radiative forcings need not be localized and can manifest in locations remote
278 from the perturbation. This is because atmospheric circulation tends to homogenize the effect of
279 heat perturbations and hence the temperature response. The vertical partitioning of the radiative
280 perturbation determines how the surface heat and moisture budgets respond, how the convective
281 interactions are affected, and hence how the surface temperature and the atmospheric thermal
282 profile are altered. "Indirect" aerosol effects arising from aerosol-cloud interactions can lead to
283 potentially significant changes in cloud characteristics such as cloud lifetimes, frequencies of
284

285 occurrence, microphysical properties, and albedo (reflectivity) (e.g., Lohmann et al., 2000;
286 Sherwood, 2002; Lohmann and Feichter, 2005). As clouds are important components in both
287 solar and longwave radiative processes and hence significantly influence the planetary radiation
288 budget (Ramanathan et al., 1987; Wielicki et al., 2002), any effect caused by aerosols in
289 perturbing cloud properties is bound to exert a significant effect on the surface-troposphere
290 radiation balance and thermal profile.

291
292 Estimates of forcing from anthropogenic land-use changes have consisted of quantification of the
293 effect of snow-covered surface albedos in the context of deforestation (Ramaswamy et al., 2001).
294 However, there remain considerable uncertainties in these quantitative estimates. There are other
295 possible ways in which land-use change can affect the heat, momentum and moisture budgets at
296 the surface (e.g., changes in transpiration from vegetation) (see also Box 5.4), and thus exert a
297 forcing of the climate (Pielke et al., 2004; NRC, 2005). In addition to the forcings shown in
298 Table 1.1, NRC (2005) has evoked a category of “nonradiative” forcings involving aerosols,
299 land-cover, and biogeochemical changes which may impact the climate system first through
300 nonradiative mechanisms, e.g., modifying the hydrologic cycle and vegetation dynamics.
301 Eventually, radiative impacts could occur, though no metrics for quantifying these nonradiative
302 forcings have been accepted as yet (NRC, 2005).

303
304 Not all the radiative forcings are globally uniform. In fact, even for the increases in well-mixed
305 greenhouse gases, despite their globally uniform mixing ratios, the resulting forcing of the
306 climate system is at a maximum in the tropics due to the temperature contrast between the
307 surface and troposphere there and therefore the increased infrared radiative energy trapping.

308 Owing to the dependence of infrared radiative transfer on clouds and water vapor, which have
309 substantial spatial structure in the low latitudes, the greenhouse gas forcing is non-uniform in the
310 tropics, being greater in the relatively drier tropical domains. For short-lived gases, the
311 concentrations themselves are not globally uniform so there tends to be a distinct spatial
312 character to the resulting forcing, e.g., stratospheric ozone, whose forcing is confined essentially
313 to the mid-to-high latitudes, and tropospheric ozone whose forcing is confined to tropical to
314 midlatitudes. For aerosols, which are even more short-lived than ozone, the forcing has an even
315 more localized structure (see also Box 5.3). However, although tropospheric ozone and aerosol
316 forcing are maximized near the source regions, the contribution to the global forcing from
317 remote regions is not negligible. The natural factors, namely solar irradiance changes and
318 stratospheric aerosols from volcanic eruptions, exert a forcing that is global in scope.

319
320 In terms of the transient changes in the climate system, it is also important to consider the
321 temporal evolution of the forcings. For well-mixed greenhouse gases, the evolution over the past
322 century, and in particular the past four decades, is very well quantified because of reliable and
323 robust observations. However, for the other forcing agents, such as aerosols, there are
324 uncertainties concerning their evolution that can affect the inferences about the resulting surface
325 and atmospheric temperature trends. Stratospheric ozone changes, which have primarily occurred
326 since ~1980, are slightly better known than tropospheric ozone and aerosols. For solar irradiance
327 and land-surface changes, the knowledge of the forcing evolution over the past century is poorly
328 known. Only in the past five years have climate models included time varying estimates of a
329 subset of the forcings that affect the climate system. In particular, current models typically
330 include GHGs, ozone, sulfate aerosol direct effects, solar influences, and volcanoes. Some very

331 recent model simulations also include time-varying effects of black carbon and land use change.
332 Other forcings either lack sufficient physical understanding or adequate global time- and space-
333 dependent datasets to be included in the models at this time. As we gain more knowledge of
334 these other forcings and are better able to quantify their space- and time-evolving characteristics,
335 they will be added to the models used by groups around the world. Experience with these models
336 so far has shown that the addition of more forcings generally tends to improve the realism and
337 details of the simulations of the time evolution of the observed climate system (e.g., Meehl et al.,
338 2004).

339
340 Whether the climate system is responding to internally generated variations in the atmosphere
341 itself, to atmosphere-ocean-land-surface coupling, or to externally applied forcings by natural
342 and/or anthropogenic factors, there are feedbacks that arise which can play a significant role in
343 determining the changes in the vertical and horizontal thermal structure. These include changes
344 in the hydrologic cycle involving water vapor, clouds, sea-ice, and snow, which by virtue of their
345 strong interactions with solar and longwave radiation, amplify the effects of the initial
346 perturbation (Stocker et al., 2001; NRC, 2003) in the heat balance, and thus influence the
347 response of the climate system. Convection and water vapor feedback, and cloud feedback in
348 particular, are areas of active observational studies; they are also being pursued actively in
349 climate modeling investigations to increase our understanding and reduce uncertainties
350 associated with those processes.

351

352 **1.3 Stratospheric forcing and related effects**

353

354 Observed changes in the stratosphere in recent decades have been large and several recent
355 studies have investigated the causes. WMO (2003) and Shine et al. (2003) conclude that the
356 observed vertical profile of cooling in the global, annual-mean stratosphere (from the tropopause
357 up into the upper stratosphere) can, to a substantial extent, be accounted for in terms of the
358 known changes that have taken place in well-mixed greenhouse gases, ozone, and water vapor
359 (Figure 1.2). Even at the zonal, annual-mean scale, the lower stratosphere temperature trend is
360 discernibly influenced by the changes in the stratospheric gases (Ramaswamy and Schwarzkopf,
361 2002; Langematz et al., 2003). In the tropics, there is considerable uncertainty about the
362 magnitude of the lower stratospheric cooling (Ramaswamy et al., 2001). In the high northern
363 latitudes, the lower stratosphere becomes highly variable both in the observations and model
364 simulations, especially during winter, such that causal attribution is difficult to establish. In
365 contrast, the summer lower stratospheric temperature changes in the Arctic and the springtime
366 cooling in the Antarctic can be attributed in large part to the changes in the greenhouse gases
367 (WMO, 2003; Schwarzkopf and Ramaswamy, 2002).

368

369 Figure 1.2. Global- and annual-mean temperature change over the 1979-1997 period in the stratosphere.
370 Observations: LKS (radiosonde), SSU and MSU (satellite) data.

371 Vertical bars on satellite data indicate the approximate span in altitude from where the signals originate, while the
372 horizontal bars are a measure of the uncertainty in the trend. Computed: effects due to increases in well-mixed
373 gases, water vapor, and ozone depletion, and the total effect (Shine et al., 2003).

374

375 Owing to the cooling of the lower stratosphere, there is a decreased infrared emission from the

376 stratosphere into the troposphere (Ramanathan and Dickinson, 1979; WMO, 1999), leading to a

377 radiative heat deficit in the upper troposphere, and a tendency for the upper troposphere to cool.

378 In addition, the depletion of ozone in the lower stratosphere can result in ozone decreases in the

379 upper troposphere due to reduced transport from the stratosphere (Mahlman et al., 1994). This

380 too affects the heat balance in the upper troposphere. Further, lapse rate near the tropopause can

381 be affected by changes in radiatively active trace constituents such as methane (WMO, 1986;

382 WMO/SROC Report, 2004).

383

384 The height of the tropopause (the boundary between the troposphere and stratosphere) is
385 determined by a number of physical processes that make up the integrated heat content of the
386 troposphere and the stratosphere. Changes in the heat balance within the troposphere and/or
387 stratosphere can consequently affect the tropopause height. For example, when a volcanic
388 eruption puts a large aerosol loading into the stratosphere where the particles absorb solar and
389 longwave radiation and produce stratospheric heating and tropospheric cooling, the tropopause
390 height shifts downward. Conversely, a warming of the troposphere moves the tropopause height
391 upward (e.g., Santer et al., 2003). Changes in tropopause height and their potential causes will be
392 discussed further in Chapter 5.

393
394 The episodic presence of volcanic aerosols affects the equator-to-pole heating gradient, both in
395 the stratosphere and troposphere. Temperature gradients in the stratosphere or troposphere can
396 affect the state of the polar vortex in the northern latitudes, the coupling between the stratosphere
397 and troposphere, and the propagation of temperature perturbations into the troposphere and to the
398 surface. This has been shown to be plausible in the case of perturbations due to volcanic aerosols
399 in observational and modeling studies, leading to a likely causal explanation of the observed
400 warming pattern seen in northern Europe and some other high latitude regions in the first winter
401 following a tropical explosive volcanic eruption (Jones et al., 2003; Robock and Oppenheimer,
402 2003; Shindell et al., 2001; Stenchikov et al., 2002). Ozone and well-mixed greenhouse gas
403 changes in recent decades can also affect stratosphere-troposphere coupling (Thompson and
404 Solomon, 2002; Gillett and Thompson, 2003), propagating radiatively-induced temperature
405 perturbations from the stratosphere to the surface in the high latitudes during winter.
406

1.4 Simulated responses in vertical temperature profile to different external forcings

407
408 Three-dimensional computer models of the coupled global atmosphere-ocean-land surface
409 climate system have been used to systematically analyze the expected effects of different
410 forcings on the vertical structure of the temperature response and compare these with the
411 observed changes (e.g., Santer et al., 1996; 2003; Hansen et al., 2002). A climate model can be
412 run with time-varying observations of just one forcing over the 20th Century to study the
413 temperature response in the vertical. Then, by running more single forcings, a picture emerges
414 concerning the effects of each one individually. The model can then be run with a combination
415 of forcings to determine the degree to which the simulation resembles the observations made in
416 the 20th Century. Note that a linear additivity of responses, while approximately valid for certain
417 combinations of forcings, need not hold in general (Ramaswamy and Chen, 1997; Hansen et al.,
418 1997; Santer et al., 2003; Shine et al., 2003). To first order, models are able to reproduce the time
419 evolution of globally averaged surface air temperature over the 20th Century, with the warming
420 in the first half of the century generally ascribed to natural forcings (mainly volcanoes and solar)
421 or unforced variations, and the warming in the late 20th Century mostly due to human-induced
422 increases of GHGs (Stott et al., 2000; Mitchell et al., 2001; Meehl et al., 2003; 2004; Broccoli et
423 al., 2003). Such modeling studies used various observed estimates of the forcings, but
424 uncertainties remain regarding details of such factors as solar variability (Frohlich and Lean,
425 2004), historical volcanic forcing (Bradley, 1988), and aerosols (Charlson et al., 1992; Anderson
426 et al., 2003).

427
428 Analyses of model responses to external forcings also require consideration of the internal
429 variability of the climate system for a proper causal interpretation of the observed surface

430 temperature record. For example, the mid-1970s saw the beginning of a significant increase of
431 global temperatures that was associated with an apparent regime shift in the Pacific (Trenberth
432 and Hurrell, 1994). While it has been argued that the warming could have been a delayed
433 response to a regime shift due to the heat capacity of the ocean (Lindzen and Giannitsis, 2002),
434 this increase in temperature starting in the 1970s is also simulated as a response due to changes
435 in external forcing in the models noted above. The relationship between external forcing and
436 internal decadal variability of the climate system (i.e., can the former influence the latter, or are
437 they totally independent) is an intriguing research problem that is being actively studied.

438
439 In addition to the analyses of surface temperatures outlined above, climate models can also show
440 the expected effects of different forcings on temperatures in the vertical. For example using
441 simplified ocean representations for equilibrium 2XCO₂, Hansen et al. (2002) show that changes
442 of various anthropogenic and natural forcings produce different patterns of temperature change
443 horizontally and vertically. Hansen et al. (2002) also show considerable sensitivity of the
444 simulated vertical temperature response to the choice of ocean representation, particularly for the
445 GHG-only and solar-only cases. For both of these cases, the “Ocean A” configuration (SSTs
446 prescribed according to observations) lacks a clear warming maximum in the upper tropical
447 troposphere, thus illustrating that there could be some uncertainty in our model-based estimates
448 of the upper tropospheric temperature response to GHG forcing (see Chapter 5).

449
450 An illustration of the effects of different forcings on the trends in atmospheric temperatures at
451 different levels from a climate model with time-varying forcings over the latter part of the 20th
452 Century is shown in Figure 1.3. Here, the temperature changes are calculated over the time

453 period of 1979-1999, and are averages of four-member ensembles. As in Hansen et al., this
454 model, the NCAR/DOE Parallel Climate Model (PCM, e.g., Meehl et al., 2004) shows warming
455 in the troposphere and cooling in the stratosphere for an increase of GHGs, warming through
456 most of the stratosphere and a slight cooling in the troposphere for volcanic aerosols, warming in
457 a substantial portion of the atmosphere for an increase in solar forcing, warming in the
458 troposphere from increased tropospheric ozone and cooling in the stratosphere due to the
459 decrease of stratospheric ozone, and cooling in the troposphere and slight warming in the
460 stratosphere from sulfate aerosols. The multiple-forcings run shows the net effects of the
461 combination of these forcings as a warming in the troposphere and a cooling in the stratosphere.
462 Note that these simulations may not provide a full accounting of all factors that could affect the
463 temperature structure, e.g., black carbon aerosols, land use change (Ramaswamy et al., 2001;
464 Hansen et al., 1997; 2002; Pielke, 2001; NRC, 2005; Ramanathan et al., 2005).

465

466

467 Figure 1.3. PCM simulations of the vertical profile of temperature change due to various forcings, and the effect due
468 to all forcings taken together (after Santer et al., 2000).

469

470 The magnitude of the temperature response for any given model is related to its climate

471 sensitivity. This is usually defined either as the equilibrium warming due to a doubling of CO₂

472 with an atmospheric model coupled to a simple slab ocean, or the transient climate response

473 (warming at time of CO₂ doubling in a 1% per year CO₂ increase experiment in a global coupled

474 model). The climate sensitivity varies among models due to a variety of factors (Cubasch et al.,
475 2001; NRC, 2004).

476
477 The important conclusion here is that representations of the major relevant forcings are important
478 to simulate 20th Century temperature trends since different forcings affect temperature differently
479 at various levels in the atmosphere.

480

481 **1.5 Physical factors, and temperature trends at the surface and in the troposphere**

482

483 Tropospheric and surface temperatures, although linked, are separate physical entities (Trenberth
484 et al., 1992; Hansen et al., 1995; Hurrell and Trenberth, 1996; Mears et al., 2003). Insight into
485 this point comes from an examination of the correlation between anomalies in the monthly-mean
486 surface and tropospheric temperatures over 1979-2003 (Figure 1.4). The correlation coefficients
487 between monthly surface and tropospheric temperature anomalies (as represented by
488 temperatures derived from MSU satellite data) reveal very distinctive patterns, with values
489 ranging from less than zero (implying poor vertical coherence of the surface and tropospheric
490 temperature anomalies) to over 0.9. The highest correlation coefficients (>0.75) are found across
491 the middle and high latitudes of Europe, Asia, and North America, indicating a strong
492 association between the surface and tropospheric monthly temperature variations. Correlations
493 are generally much less (~ 0.5) over the tropical continents and the North Atlantic and North
494 Pacific Oceans. Correlations less than 0.3 occur over the tropical and southern oceans and are
495 lowest (<0.15) in the tropical western Pacific. Relatively high correlation coefficients (>0.6) are

496 found over the tropical eastern Pacific where the ENSO signal is large and the sea-surface
497 temperature fluctuations influence the atmosphere significantly.

498
499 Figure 1.4. Gridpoint correlation coefficients between monthly surface and tropospheric temperature anomalies over
500 1979-2003. The tropospheric temperatures are derived from MSU satellite data (Christy et al., 2003).

501
502 Differences between the surface and tropospheric temperature records are found where there is
503 some degree of decoupling between the layers of the atmosphere. For instance, as discussed
504 earlier, over portions of the subtropics and tropics, variations in surface temperature are
505 disconnected from those aloft by a persistent trade-wind inversion. Shallow temperature
506 inversions are also commonly found over land in winter, especially in high latitudes on sub-
507 seasonal timescales, so that there are occasional large differences between monthly surface and
508 tropospheric temperature anomalies.

509
510 More important than correlations for trends, however, is the variability of the two temperature
511 records, assessed by computing the standard deviation of the measurement samples of each

512 record (Figure 1.5). The figure exhibits pronounced regional differences in variability between
513 the surface and tropospheric records. Standard deviations also help in accounting for the
514 differences in correlation coefficients, because they yield information on the size and persistence
515 of the climate signal relative to the noise in the data. For instance, large variations in eastern
516 tropical Pacific sea surface temperature associated with ENSO dominate over measurement
517 uncertainties, as do large month-to-month swings in surface temperatures over extratropical
518 continents.

519

520 Figure 1.5. Standard deviations of monthly mean temperature anomalies from the surface and tropospheric
521 temperature records over 1979-2003. The tropospheric temperatures are derived from MSU satellite data (Christy et
522 al., 2003).

523
524 The largest variability in both surface and tropospheric temperature is over the Northern
525 Hemisphere continents. The standard deviation over the oceans in the surface data set is much
526 smaller than over land except where the ENSO phenomenon is prominent. The standard
527 deviations of tropospheric temperature, in contrast, exhibit less zonal variability. Consequently,
528 the standard deviations of the monthly tropospheric temperatures are larger than those of the
529 surface data by more than a factor of two over the North Pacific and North Atlantic. Over land,
530 tropospheric temperatures exhibit slightly less variability than surface temperatures. These
531 differences in variability are indicative of differences in physical processes over the oceans
532 versus the continents. Of particular importance are the roles of the land surface and ocean as the
533 lower boundary for the atmosphere and their very different abilities to store heat, as well as the
534 role of the atmospheric winds that help to reduce regional differences in tropospheric
535 temperature through the movement of heat from one region to another.

536
537 Over land, heat penetration into the surface involves only the upper few meters, and the ability of
538 the land to store heat is low. Therefore, land surface temperatures vary considerably from
539 summer to winter and as cold air masses replace warm air masses and vice versa. The result is
540 that differences in magnitude between surface and lower-atmospheric temperature anomalies are
541 relatively small over the continents: very warm or cold air aloft is usually associated with very
542 warm or cold air at the surface. In contrast, the ability of the ocean to store heat is much greater
543 than that of land, and mixing in the ocean to typical depths of 50 meters or more considerably
544 moderates the sea surface temperature response to cold or warm air. Over the northern oceans,
545 for example, a very cold air mass (reflected by a large negative temperature anomaly in the

546 tropospheric record) will most likely be associated with a relatively small negative temperature
547 anomaly at the sea surface. This is one key to understanding the differences in trends between
548 the two records.

549
550 Long-term changes in the atmospheric circulation can be reflected by trends in indices of
551 patterns (or modes) of natural climate variability such as ENSO, the North Atlantic Oscillation
552 (NAO; also known as the Northern Hemisphere annular mode, or NAM), and the Southern
553 Hemisphere (SH) annular mode (SAM). The exact magnitudes of the index trends depend on the
554 period of time examined. Over the past several decades, for instance, changes in atmospheric
555 circulation (reflected by a strong upward trend in indices of the NAO) have contributed to a Cold
556 Ocean Warm Land (COWL) surface temperature pattern over the Northern Hemisphere (Hurrell,
557 1996; Thompson and Wallace, 2000). In the lower atmosphere, winds blowing from ocean to
558 land to ocean are much stronger than at the surface, and this moderating influence of the winds
559 contributes to less east-west variability in the tropospheric data (Figure 1.5). Thus, the recent
560 warm anomalies over the continents are roughly cancelled by the cold anomalies over the oceans
561 in the tropospheric dataset. This is not the case for the surface temperature record, which is
562 dominated by the warmth over the continents. The result is that the changes in the Northern
563 Hemisphere (NH) atmospheric circulation over the past few decades have produced a significant
564 difference in surface and tropospheric temperature trends (Hurrell and Trenberth, 1996).
565 Similarly, Thompson and Solomon (2002) showed that recent tropospheric temperature trends at
566 high southern latitudes were related to changes in the SAM.

567

568 Physical differences between the two measures of temperature are also evident in their dissimilar
569 responses to volcanic eruptions and ENSO (e.g., Santer et al. 2000). These phenomena have a
570 greater effect on tropospheric than surface temperature, especially over the oceans (Jones, 1994).
571 However, changes in ENSO over the past several decades do not explain long-term changes in
572 tropical tropospheric temperatures (Hegerl and Wallace, 2002). Changes in concentrations of
573 stratospheric ozone could also be important, as the troposphere is cooled more by observed
574 ozone depletion than is the surface ([Hansen et al., 1995](#); [Ramaswamy et al., 1996](#)). Another
575 contributing factor could be that at the surface, the daily minimum temperature has increased at a
576 faster rate than the daily maximum, resulting in a decrease in the diurnal temperature range over
577 many parts of the world (e.g., Easterling et al., 1997; Dai et al., 1999). Because of nighttime
578 temperature inversions, the increase in the daily minimum temperatures likely involves only a
579 shallow layer of the atmosphere that would not be evident in upper-air temperature records.

580 These physical processes provide indications of why trends in surface temperatures are expected
581 to be different from trends in the troposphere, especially in the presence of strong interannual
582 variability, even if both sets of measurements were perfect. Of course they are not, as described
583 in more detail in Chapter 2, which deals with the strengths and limitations of current observing
584 systems. An important issue implicit in Figure 1.5 is that of spatial sampling, and the
585 accompanying caveats about interpretation of the truly global coverage provided by satellites
586 versus the incomplete space and time coverage offered by radiosondes. These are discussed in
587 depth in Chapters 2 and 3.

588

589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630

References

- 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.
- Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: Middle atmosphere dynamics. Academic Press, Florida, 489pp.
- Bradley, R.S., 1988: The explosive volcano eruption signal in Northern Hemisphere continental temperature records. *Climatic Change*, **12**, 221—243.
- Broccoli, A., K. Dixon, T. Delworth, T. Knutson and R. Stouffer, 2003: Twentieth-century temperature and precipitation trends in ensemble climate simulations including natural and anthropogenic forcing.
- 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.
- Christy, J.R., R.W. Spencer, W.B. Norris, W.D. Braswell and D.E. Parker, 2003: Error estimates of Version 5.0 of MSU/AMSU bulk atmospheric temperatures. *J. Atmos. Oceanic Tech.*, **20**, 613-629.
- Chung, C., V. Ramanathan and J. Kiehl, 2002: Effects of the South Asian absorbing haze on the northeast monsoon and surface-air heat exchange. *J. Climate*, **15**, 2462-2476.
- Cubasch, U., 2001: Projections of future climate change. In Climate Change 2001: The Scientific Basis. *Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* [Houghton, J. T. et al (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 881 pp.
- Dai, A., K. E. Trenberth, and T. R. Karl, 1999: Effects of clouds, soil moisture, precipitation, and water vapor on diurnal temperature range. *J. Climate*, **12**, 2451– 2473.
- Donner, L., C. J. Seman, R. S. Hemler, and S. Fan, 2001: A cumulus parameterization including mas fluxes, convective vertical velocities, and mesoscale effects: thermodynamic

- 631 and hydrological aspects in a general circulation model. *J. Climate*, **14**, 3444-3463.
632
- 633 Easterling, D. R., and Coauthors, 1997: Maximum and minimum temperature trends for the
634 globe. *Science*, **277**, 364–367.
635
- 636 Erlick, C. E., and V. Ramaswamy, 2003: Sensitivity of the atmospheric lapse rate to solar
637 cloud absorption in a radiative-convective model. *J. Geophys. Res.*, **108(D16)**, 4522,
638 doi:10.1029/2002JD002966.
639
- 640 Frohlich, C. and J. Lean, 2004: Solar radiative output and its variability: Evidence and
641 mechanisms. *Astron. Astrophys. Rev.*, **12**, 273—320.
642
- 643 Gillett, N., and D. W. J. Thompson, 2003: Simulation of recent Southern Hemisphere climate
644 change. *Science*, **302**, 273-275.
645
- 646 Goody, R. M. and Y. L. Yung, 1989: Atmospheric radiation: A theoretical basis. Oxford
647 University Press, New York, 519pp.
648
- 649 Hansen, J. E. et al., 1995: satellite and surface temperature data at odds? *Clim. Change*, **30**,
650 103-117.
651
- 652 Hansen, J. E., M. Sato and R. Ruedy, 1997: Radiative forcing and climate response, *J.*
653 *Geophys. Res.*, **102**, 6831-6864.
654
- 655 Hansen, J. and co-authors, 2002: Climate forcings in Goddard Institute for
656 Space Studies SI2000 simulations, *J. Geophys. Res.*, **107(D18)**, 4347,
657 doi:10.1029/2001JD001143.
658
- 659 Hegerl, G. C., and J. M. Wallace, 2002: Influence of patterns of climate variability on the
660 difference between satellite and surface temperature trends. *J. Climate*, **15**, 2412- 2428.
661
- 662 Holton, J.R., 1979: An introduction to dynamic meteorology. Academic Press, New York,
663 p. 286.
664
- 665 Houghton, J. T., 1977: The physics of atmospheres. Cambridge University Press, UK, 203pp.
666
- 667 Hoskins, B. J., 2003: Atmospheric processes and observations. *Philos. Tr. Soc. S.-A* **363**,
668 1945-1960.
669
- 670 Hurrell, J. W., 1996: Influence of variations in extratropical wintertime teleconnections on
671 Northern Hemisphere temperatures. *Geophysical Research Letters*, **23**, 665-668.
672
- 673 Hurrell, J. W., and K. E. Trenberth, 1996: Satellite versus surface estimates of air temperature
674 since 1979. *Journal of Climate*, **9**, 2222-2232.
675
- 676 Jones, P. D., 1994: Hemispheric surface air temperature variations: a reanalysis and an update to

- 677 1993. *J. Climate*, **7**, 1794-1802.
678
- 679 Jones, P.D., A. Moberg, T.J. Osborn, and K.R. Briffa, 2003: Surface climate responses to
680 explosive volcanic eruptions seen in long European temperature records and mid-to-high
681 latitude tree-ring density around the Northern Hemisphere. In *Volcanism and the Earth's*
682 *Atmosphere*. A. Robock and C. Oppenheimer, eds. Washington D.C.: American
683 Geophysical Union, 239—254.
684
- 685 Kalnay, E. et al., 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor.*
686 *Soc.*, **77**, 437-471.
687
- 688 Kiehl, J. T., 1992: *Atmospheric general circulation modeling*. In *Climate System Monitoring*
689 (K. Trenberth, editor). Cambridge University Press, pp319-369.
690
- 691 Langematz, U., M. Kunze, K. Kruger, K. Labitzke, and G. Roff, 2003: Thermal and dynamical
692 changes of the stratosphere since 1979 and their link to ozone and CO2 changes. *J.*
693 *Geophys. Res.*, **108**, 4027, doi: 1029/2002JD002069.
694
- 695 Lean, J., G. Rottman, J. Harder and G. Kopp, 2005: *SORCE contributions to new understanding*
696 *of global change and solar variability. Solar Phys.*, in press.
697
- 698 Lindzen, R. S., and K. Emanuel, 2002: The greenhouse effect. Pp. 562-566 in *Encyclopedia of*
699 *Global Change, Environmental Change and Human Society, Vol. 1* (Andrew S. Goudie, ed.)
700 New York. Oxford University Press, 710 pp.
701
- 702 Lindzen, R. S., and C. Giannitsis, 2002: Reconciling observations of global temperature change.
703 *Geophys. Res. Lett.*, 29: 10.1029/2001GL014074.
704
- 705 Lohmann, U., and J. Feichter, 2005: Global indirect aerosol effects: a review. *Atmos. Chem.*
706 *Physics*, **5**, 715-737.
707
- 708 Lohmann, U., J. Feichter, J. Penner, and R. Leaitch, 2000: Indirect effect of sulfate and
709 carbonaceous aerosols: A mechanistic treatment. *J. Geophys. Res.*, **105**, 12193- 12206.
710
- 711 Mahlman, J. D., J. P. Pinto, and L. J. Umscheid, 1994: Transport, radiative, and dynamical
712 effects of the Antarctic ozone hole: A GFDL “SKYHI” experiment. *J. Atmos. Sci.*, **51**,
713 489-508.
714
- 715 Manabe, S. and R. Wetherald 1967: Thermal equilibrium of the atmosphere with a given
716 distribution of relative humidity. *J. Atmos. Sci.*, **24**, 241-259.
717
- 718 McClatchey, R. A., et al., 1972: *Optical properties of the atmosphere*, third ed.,
719 AFCRL-72-0497, Air Force Cambridge Research Labs., Hanscom, MA, 110 pp.
720
- 721 Mears, C. A., M. C. Schabel, and F. W. Wentz, 2003: A reanalysis of the MSU channel
722 tropospheric temperature record. *J. Climate*, **16**, 3650-3664.

- 723
724 Meehl, G.A., W.M. Washington, T.M.L. Wigley, J.M. Arblaster, and A. Dai,
725 2003: Solar and greenhouse gas forcing and climate response in the 20th century. *J.*
726 *Climate*, **16**, 426--444.
- 727
728 Meehl, G.A., W.M. Washington, C. Ammann, J.M. Arblaster, T.M.L. Wigley, and C.
729 Tebaldi, 2004: Combinations of natural and anthropogenic forcings and 20th century
730 climate. *J. Climate*, in press.
- 731
732 Menon, S., J. Hansen, L. Nazarenko, and Y. Luo, 2002: Climate effects of back carbon
733 aerosols in China and India. *Science*, **297**, 2250-
734
- 735 Mitchell, J. F. B. et al., 2001: Detection of climate change and attribution of causes. *In*
736 *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third*
737 *Assessment Report of the Intergovernmental Panel on Climate Change* [Houghton, J. T.
738 et al (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA,
739 881 pp.
- 740
741 National Research Council, 2003: *Understanding Climate Change Feedbacks*. The National
742 Academies Press, Washington, D. C., 152pp.
- 743
744 National Research Council, 2004: *Estimating Climate Sensitivity*, in preparation.
- 745
746 National Research Council, 2005: *Radiative Forcing of Climate Change: Expanding the*
747 *Concept and Addressing Uncertainties*, in preparation.
- 748
749 Oort, A. H. and J. Peixoto, 1992: *Physics of Climate*. American Institute of Physics, New York,
750 520pp.
- 751
752 Pielke, RA Sr., 2001: Influence of the spatial distribution of vegetation and soils on the
753 prediction of cumulus convective rainfall. *Review of Geophysics*, **39**, 151—177.
- 754
755 Pielke, RA Sr., C. Davey and J. Morgan, 2004: Assessing “global warming” with surface
756 heat content. *EOS*, **85**, no. 21, 210—211.
- 757
758 Ramanathan, V., and J. A. Coakley, 1978: Climate modeling through radiative-convective
759 models. *Rev. Geophys. Space Phys.*, **16**, 465-489.
- 760
761 Ramanathan, V. and R. Dickinson, 1979: The role of stratospheric ozone in the zonal and
762 seasonal radiative energy balance of the Earth-troposphere system. *J. Atmos. Sci.*, **36**,
763 1084-1104.
- 764
765 Ramanathan, V., P. Crutzen, J. Kiehl and D. Rosenfeld, 2001: Aerosols, climate and the
766 hydrological cycle. *Science*, **294**, 2119-2124.
- 767
768 Ramanathan, V. et al., 1989: Cloud radiative forcing and climate: Results from the earth budget

- 769 radiation experiment. *Science*, **243**, 57-63.
- 770
- 771 Ramanathan, V., E. Pitcher, R. Malone and M. Blackmon, 1983: The response of a general
772 circulation model to refinements in radiative processes. *J. Atmos. Sci.*, **40**, 605-630.
- 773
- 774 Ramanathan, V. et al., 2005: Atmospheric brown clouds: Impacts on South Asian climate and
775 hydrological cycle. *Proc. Natl. Acad. Sci.*, **102**, 5326-5333.
- 776
- 777 Ramaswamy, V. and C-T. Chen, 1997: Climate forcing-response relationships for greenhouse
778 and shortwave radiative perturbations. *Geophys. Res. Lett.*, **24**, 667-670.
- 779
- 780 Ramaswamy, V. and M. D. Schwarzkopf, 2002: Effects of ozone and well-mixed gases on
781 annual-mean stratospheric temperature trends. *Geophys. Res. Lett.*, **29**, 2064, doi:
782 10.1029/2002GL015141.
- 783
- 784 Ramaswamy, V. and V. Ramanathan, 1989: Solar absorption by cirrus clouds and the
785 maintenance of the tropical upper troposphere thermal structure. *J. Atmos. Sci.*, **46**,
786 2293-2310.
- 787
- 788 Ramaswamy, V. M. D. Schwarzkopf and W. Randel, 1996: Fingerprint of ozone depletion
789 in the spatial and temporal pattern of recent lower-stratospheric cooling. *Nature*, **382**,
790 616-618.
- 791
- 792 Ramaswamy, V. et al., 2001: *Radiative forcing of climate change. In Climate Change 2001: The*
793 *Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the*
794 *Intergovernmental Panel on Climate Change* [Houghton, J. T. et al (eds.)]. Cambridge
795 University Press, Cambridge, UK, and New York, NY, USA, 881 pp.
- 796
- 797 Ramaswamy, V. et al., 2001: Stratospheric temperature trends: Observations and model
798 simulations. *Reviews of Geophys.*, **39**, 71-122.
- 799
- 800 Robock, A. and C. Oppenheimer, 2003: *Volcanism and the Earth's Atmosphere*. AGU
801 Geophysical Monograph Series, 139. Washington D.C.: American Geophysical Union.
802 360pp.
- 803
- 804 Santer, B. D., K. E. Taylor, T. M. L. Wigley, T. C. Johns, P. D. Jones, D. J. Karoly, J. F. B.
805 Mitchell, A. H. Oort, J. E. Penner, V. Ramaswamy, M. D. Schwarzkopf, R. J. Stouffer, and
806 S. Tett, 1996: A search for human influences on the thermal structure of the atmosphere.
807 *Nature*, **382(6586)**, 39-46.
- 808
- 809 Santer, B. D. et al., 2000: Interpreting differential temperature trends at the surface and in the
810 lower troposphere. *Science*, **287**, 1227-1232.
- 811
- 812 Santer, B. D. et al., 2003: Contributions of anthropogenic and natural forcing to recent
813 tropopause height changes. *Science*, **301**, 479-483.
- 814

- 815 Sarachik, E., 1985: A simple theory for the vertical structure of the tropical atmosphere.
816 *Pure and Appl. Geophys.*, **123**, 261-271.
817
- 818 Schwarzkopf, M. D. and V. Ramaswamy, 2002: Effects of changes in well-mixed gases and
819 ozone on stratospheric seasonal temperatures. *Geophys. Res. Lett.*, **29**, 2184, doi:
820 10.1029/2002GL015759.
821
- 822 Schneider, T., 2004: The tropopause and the thermal stratification in the extratropics of a dry
823 atmosphere. *J. Atmos. Sci.*, **61**, 1317-1340.
824
- 825 Sherwood, S. C., 2002: A microphysical connection among biomass burning, cumulus clouds,
826 and stratospheric moisture. *Science*, **295**, 1271-1275.
827
- 828 Sherwood, S. C. and A. E. Dessler, 2003: Convective mixing near the tropical tropopause:
829 insights from seasonal variations. *J. Atmos. Sci.*, **60**, 2674-2685.
830
- 831 Shindell D. T., G. A. Schmidt, R. L. Miller and D. Rind, 2001: Northern Hemisphere winter
832 climate response to greenhouse gas, ozone, solar and volcanic forcing. *J. Geophys. Res.*,
833 **106**, 7193-7210.
834
- 835 Shine, K. P. et al., 2003: An alternative to radiative forcing for estimating the relative
836 importance of climate change mechanisms. *Geophys. Res. Lett.*, **30**, 2047, doi:
837 10.1029/2003GL018141.
838
- 839 Shine, K. P. et al., 2003: A comparison of model-simulated trends in stratospheric
840 temperatures. *Quart. J. Roy. Met. Soc.*, **129**, 1565-1588.
841
- 842 Stenchikov, G. et al., 2002: Arctic Oscillation response to the 1991 Mount Pinatubo eruption:
843 Effects of volcanic aerosols and ozone depletion. *J. Geophys. Res.*, **107**, 4803, doi:
844 10.1029/2002JD002090.
845
- 846 Stephens, G. L. and P. J. Webster, 1980: Clouds and climate: sensitivity of simple systems. *J.*
847 *Atmos. Sci.*, **38**, 235-247.
848
- 849 Stocker T. F. et al., 2001: Physical climate processes and feedbacks. In Climate Change 2001: The
850 Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the
851 Intergovernmental Panel on Climate Change [Houghton, J. T. et al. (eds.)]. Cambridge
852 University Press, Cambridge, UK, and New York, NY, USA, 881 pp.
853
- 854 Stott, P. A. et al., 2000: External control of 20th century temperature by natural and
855 anthropogenic forcings. *Science*, **290**, 2133-2137.
856
- 857 Stouffer, R. J., G. Hegerl and S. Tett, 2000: A comparison of surface air temperature variability
858 in three 1000-year coupled ocean-atmosphere model integrations. *J. Climate*, **13**, 513-
859 537.
860

- 861 Thompson, D. W. J. and J. M. Wallace, 2000: Annular modes in the extratropical circulation.
862 1. Month-to-month variability. *J. Climate*, 13, 1000-1016.
863
- 864 Thompson, D. W., and S. Solomon, 2002: Interpretation of recent Southern Hemisphere climate
865 change, *Science*, 296, 895—899.
866
- 867 Trenberth, K. E., and J. W. Hurrell, 1994: Decadal atmosphere-ocean variations in
868 the Pacific. *Climate Dynamics*, **9**, 303-319.
869
- 870 Trenberth, K. E., J. R. Christy and J. W. Hurrell, 1992: Monitoring global monthly mean surface
871 temperatures. *Journal of Climate*, **5**, 1405-1423.
872
- 873 Wielicki, B. A., et al., 2002: Evidence for large decadal variability in the tropical mean radiative
874 energy budget. *Science*, **295(5556)**, 841-844.
875
- 876 World Meteorological Organization, 1986: *Atmospheric ozone*. Global Ozone Research and
877 Monitoring Project, Report no. 16, 1095pp, Geneva.
878
- 879 World Meteorological Organization, 1999: *Scientific assessment of ozone depletion: 1998*.
880 Global Ozone Research and Monitoring Project, Report no. 44, chapter 5, Geneva.
881
- 882 World Meteorological Organization, 2003: *Scientific assessment of ozone depletion: 2002*.
883 Global Ozone Research and Monitoring Project, Report no. 47, 498pp, Geneva.
884
- 885 World Meteorological Organization, 2005: Special Report on Ozone and Climate. *In press*.
886
887

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

Chapter 2

What kinds of atmospheric temperature variations can the current observing systems detect and what are their strengths and limitations, both spatially and temporally?

Convening Lead Author: John Christy

Lead Authors: Dian Seidel, Steve Sherwood

Contributing Authors: Adrian Simmons, Ming Cai, Eugenia Kalnay, Chris Folland, Carl Mears, Peter Thorne, John Lanzante

42 Findings and Recommendations

43 • The observing systems available for this report are able to detect small surface and upper air
44 temperature variations from year to year, for example, those caused by El Niño or volcanic
45 eruptions.

46

47 • The data from these systems also have the potential to provide accurate trends in climate over
48 the last few decades (and over the last century for surface observations), once the raw data are
49 successfully adjusted for changes over time in observing systems, practices, and micro-climate
50 exposure to produce usable climate records. Measurements from all systems require such
51 adjustments and this report relies on adjusted datasets.

52

53 • Adjustments to the land surface temperature record have been sufficiently successful that
54 trends are reasonably similar on large (e.g., continental) scales, despite the fact that spatial
55 sampling is uneven and some errors undoubtedly remain. This conclusion holds to a lesser extent
56 for the ocean surface record, which suffers from more serious sampling problems and changes in
57 observing practice.

58

59 • Adjustments for changing instrumentation are most challenging for upper-air datasets. While
60 these show promise for trend analysis, it is likely that current upper-air climate records give
61 reliable indications of directions of change (e.g. warming of troposphere, cooling of stratosphere)
62 but some questions remain regarding the precision of the measurements.

63 • Upper-air datasets have been subjected to less scrutiny than surface datasets.

- 64 • Adjustments are complicated, sometimes as large as the trend itself, involve expert
65 judgments, and cannot be stringently evaluated because of lack of traceable standards.
- 66 • Unlike surface trends, reported upper-air trends vary considerably between research
67 teams beginning with the same raw data owing to their different decisions on how to
68 remove non-climatic factors.
- 69 • The diurnal cycle, which must be factored into some adjustments for satellite data, is
70 well observed only by surface observing systems.
- 71 • No available observing system has reference stations or multi-sensor instrumentation
72 that would provide stable calibration over time.
- 73 • Most observing systems have not retained complete metadata describing changes in
74 observing practices which could be used to identify and characterize non-climatic
75 influences.
- 76
- 77 • Relevant satellite datasets measure broad vertical layers and cannot reveal the detailed vertical
78 structure of temperature changes, nor can they completely isolate the troposphere from the
79 stratosphere. However, retrieval techniques can be used both to approximately isolate these
80 layers and to check for vertical consistency of trend patterns. Consistency between satellite and
81 radiosonde data can be tested by proportionately averaging radiosonde profiles.
- 82 • Reanalyses and other multi-system products have the potential for addressing issues of
83 surface and atmospheric temperature trends by making better use of available information and
84 allowing analysis of a more comprehensive, internally consistent, and spatially and temporally

85 complete set of climate variables. At present, however, they contain biases, especially in the
86 stratosphere, that affect trends and that cannot be readily removed because of the complexity of
87 the data products.

88

89 • There are as yet under-exploited data archives with potential to contribute to our
90 understanding of past changes, and new observing systems that may improve estimates of future
91 changes if designed for long-term measurement stability and operated for sufficient periods.

92

93

94 *Recommendation: Current and future observing systems should adhere to the principles for*
95 *climate observations adopted internationally under the Framework Convention on Climate*
96 *Change and documented in “NRC 2000b” and the “Strategic Plan for the U.S. Climate Change*
97 *Science Program (2003)” to significantly mitigate the limitations listed above.*

98

99 *Recommendation: The ability to fully and accurately observe the diurnal cycle should be an*
100 *important consideration in the design and implementation of new observing systems.*

101

102 *Recommendation: When undertaking efforts to retrieve data it is important to also to collect*
103 *detailed metadata which could be used to reduce ambiguity in the timing, sign and magnitude of*
104 *non-climatic influences in the data.*

105

106 *Recommendation: New climate-quality reanalysis efforts should be strongly encouraged and*
107 *specifically designed to minimize small, time-dependent biases arising from imperfections in*
108 *both data and forecast models.*

109
110 *Recommendation: Some largely overlooked satellite datasets should be reexamined to try to*
111 *extend, fortify or corroborate existing microwave-based temperature records for climate*
112 *research, e.g. microwave data from NEMS (1972) and SCAMS (1975), infrared from the HIRS*
113 *suite and radio occultation from GPS.*

114

115

116

117 **1. MAIN OBSERVING SYSTEMS AND SYNTHESIS DATA PRODUCTS**

118

119 Temperature is measured in three main ways; (1) *in situ*, where the sensor is immersed in the
120 substance of interest; (2) by *radiative emission*, where a remote sensor detects the intensity or
121 brightness of the radiation emanating from the substance; and (3) *radiative transmission*, where
122 radiation is modified as it passes through the substance in a manner determined by the
123 substance's temperature. All observations contain some level of random measurement error,
124 which is reduced by averaging; bias, which is not reduced by averaging; and sampling errors (see
125 Appendix).

126

127 a) Surface and near-surface air temperatures

128 Over land, “near-surface” air temperatures are those commonly measured about 1.5 to 2.0 meters
129 above the ground level at official weather stations, at sites run for a variety of scientific purposes,
130 and by volunteer (cooperative) observers (e.g., Jones and Moberg, 2003). These stations often
131 experience relocations, changes in instrumentation and/or exposure and changing observing
132 practices all of which can introduce biases into their long-term records. These changes are often
133 undocumented.

134

135 “Near-surface” air temperatures over the ocean (“Marine Air Temperatures” or MATs) are
136 measured by ships and buoys at various heights from 2 to more than 25m, with poorer coverage
137 than over land (e.g., Rayner et al., 2003). To avoid the contamination of daytime solar heating of
138 the ships’ surfaces that may affect the MAT, it is generally preferred to limit these to night MAT
139 (NMAT) readings only. Observations of the water temperature near the ocean surface or “Sea
140 Surface Temperatures” (SSTs) are widely used and are closely tied to MATs; ships and buoys
141 measure SSTs within a few meters below the surface.

142

143 Incomplete geographic sampling, changing measurement methods, and land-use changes all
144 introduce errors into surface temperature compilations. The spatial coverage, indicated in Figure
145 2.1, is far from uniform over either land or ocean areas. The southern oceans, polar regions and
146 interiors of Brazil and Africa are not well sampled by in-situ networks. However, creating
147 global surface temperature analyses involves not only merging land and ocean data but also

148 considering how best to represent areas where there are few or no observations. The most
149 conservative approach is to use only those grid boxes with data, thus avoiding any error
150 associated with interpolation. Unfortunately, the areas without data are not evenly or randomly
151 distributed around the world, leading to considerable uncertainties in the analysis, though it is
152 possible to make an estimate of these uncertainties. Using the conservative approach, the tropical
153 land surface areas would be under-represented, as would the southern ocean. Therefore,
154 techniques have been developed to interpolate data to some extent into surrounding data-void
155 regions. A single group may produce several different such datasets for different purposes. The
156 choice may depend on whether the interest is a particular local region, the entire globe, or use of
157 the dataset with climate models (Chapter 5). Estimates of global and hemispheric scale averages
158 of near-surface temperatures generally begin around 1860 over both land and ocean.

159

160 Figure 2.1 Top: Location of radiosonde stations used in the HadAT upper air dataset with those also in the LKS as
161 crosses. Bottom: Distribution of land stations (green) and SST observations (blue) reporting temperatures used in the
162 surface temperature datasets over the period 1979-2004. Darker colors represent locations for which data were
163 reported with greater frequency.

164 See chapter 3 for definitions of datasets.

165

166 Datasets of near-surface land and ocean temperatures have traditionally been derived from *in-situ*

167 thermometers. With the advent of satellites, some datasets now combine both *in-situ* and
168 remotely sensed data (Reynolds et al., 2002; Smith and Reynolds, 2005), or use exclusively
169 remotely sensed data (Kilpatrick et al., 2001) to produce more geographically complete
170 distributions of surface temperature. Because the satellite sensors measure infrared or microwave
171 emission from the earth's surface (a "skin" typically tens of microns thick that may have a
172 temperature different from either the air above or material at greater depths), calculations are
173 required to convert the skin temperature into the more traditional near-surface air or SST
174 observation (in this context SSTs are called "bulk sea surface temperatures", Chelton, 2005.)
175 Typically, in-situ observations are taken as "truth" and satellite estimates (which may be affected
176 by water vapor, clouds, volcanic aerosols, etc.) are adjusted to agree with them (Reynolds, 1993.)
177 With continued research, datasets with surface temperatures over land, ice, and ocean from
178 infrared and microwave sensors should provide expanded coverage of surface temperature
179 variations (e.g., Aires et al., 2004).

180

181 Sampling errors in ship and buoy SST data typically contribute more to large-scale averages than
182 random measurement errors as shown in Smith and Reynolds (2004), especially as the
183 temperature record extends backward in time. Biases depend on observing method. Most ship
184 observations since the 1950s were made from insulated buckets, hull contact sensors, and engine
185 intake temperatures at depths of one to several meters. Historic correction of ship data prior to
186 1942 is discussed in (Folland and Parker, 1995) and bias and random errors from ships are
187 summarized by (Kent and Taylor, 2004) and (Kent and Challenor, 2004). They report that engine
188 intake temperatures are typically biased 0.1-0.2°C warmer than insulated buckets. This is

189 primarily due to engine room heating of the water temperatures although there is also
190 evaporative cooling of the water in the insulated buckets. Hull contact sensors are the most
191 accurate though much less common. The bias correction of the ship SST data (Kent and Kaplan,
192 2004) requires information on the type of measurement (e.g. insulated bucket, etc.) which
193 becomes more difficult to determine prior to 1990s due to incomplete documentation. Kent and
194 Kaplan (2005) also found that insulated bucket temperatures may be too cold by 0.12 to 0.16°C.
195 When the bucket bias is used, engine intake temperatures in the mid-to-late 1970s and the 1980s
196 were found to be smaller than that suggested by previous studies, ranging from 0.09 to 0.18°C. In
197 addition, their study indicates that engine intake SSTs may have a cold bias of -0.13°C in the
198 early 1990s. The reliability of these biases are subject to revision due to small sample sizes that
199 sample sizes for these comparisons tend to be small with large random errors. Buoy observations
200 became more plentiful following the start of the Tropical Ocean Global Atmosphere (TOGA)
201 Program (McPhaden, 1995) in 1985. These observations are typically made by an immersed
202 temperature sensor or a hull contact sensor, and are more accurate because they do not have the
203 bias errors of ship injection or insulated bucket temperatures.

204

205 The global surface air temperature data sets used in this report are to a large extent based on data
206 readily exchanged internationally, e.g., through CLIMAT reports and the WMO publication
207 *Monthly Climatic Data for the World*. Commercial and other considerations prevent a fuller
208 exchange, though the United States may be better represented than many other areas. In this
209 report we present three global surface climate records, created from available data by NASA
210 Goddard Institute for Space Studies (GISS), NOAA National Climatic Data Center

211 (NOAA/NCDC) and the cooperative project of the U.K. Hadley Centre and the Climate
212 Research Unit of the University of East Anglia (HadCRUT2v). These will be identified as $T_{\text{Sfc-G}}$,
213 $T_{\text{Sfc-N}}$ and $T_{\text{Sfc-U}}$ respectively.

214

215 b) Atmospheric “upper air” temperatures

216 *1. Radiosondes*

217 Radiosonde or balloon-based observations of atmospheric temperature are *in-situ* measurements
218 as the thermometer (often a thermistor or a capacitance-based sensor), suspended from a balloon,
219 is physically carried through the atmospheric column. Readings are radio-transmitted back to a
220 data recorder. Balloons are released once or twice a day (00 and/or 12 Coordinated Universal
221 Time or UTC) at about 1,000 stations around the globe, many of which began operations in the
222 late 1950s or 1960s. These sites are unevenly distributed, with only the extratropical northern
223 hemisphere land areas and the Western Pacific Ocean/Indonesia/Australia region being well-
224 sampled in space and time. Useful temperature data can be collected from near the surface
225 through the lower and middle stratosphere (though not all balloons survive to these heights).
226 Radiosonde data in the first hundred meters or so above the surface are sometimes erroneous if
227 the sensors have not been allowed to reach equilibrium with the atmosphere before launch, and
228 may not be representative of regional conditions, due to microclimatic and terrain effects.

229

230 Although most radiosonde data are transmitted to meteorological centers around the world and
231 archived, in practice many soundings do not reach this system and are collected later. No
232 definitive archive of radiosonde data exists, but several archives in the U.S. and abroad contain

233 nearly complete collections, though several different schemes have been employed for quality
234 control. To monitor climate, it is desirable to have a long, continuous record of measurements
235 from many well-distributed fixed sites. There are about 700 radiosonde stations that have
236 operated in the same location for at least three decades; many of these are clustered in a few
237 areas, further reducing the effective coverage (Figure 2.1). Thus, a dilemma exists for estimating
238 long-term changes: whether to use a smaller number of stations having long segments of
239 continuous records, or a larger number of stations with shorter records that do not always overlap
240 well. Various analysis groups have approached this differently (see Chapters 3 and 4).

241
242 Typically, radiosonde-based datasets are developed for specific atmospheric pressure surfaces
243 known as “mandatory reporting levels” (Figure 2.2). Such data at discrete vertical levels provide
244 unique information for assessing changes in the structure of the atmosphere. Two such datasets
245 are featured in this report, The Hadley Centre Atmospheric Temperatures from the U.K.
246 (HadAT) and Radiosonde Atmospheric Temperatures Products for Assessing Climate
247 (RATPAC) from NOAA. A product such as $T_{850-300}$, for example, will be identified as $T_{850-300-U}$
248 and $T_{850-300-N}$ for HadAT and RATPAC respectively. ¹

¹ A third radiosonde dataset was generated by comparing radiosonde observations against the first-guess field of the ERA-40 simulation forecast model (Haimberger, 2004). Adjustments were applied when the relative difference between the radiosonde temperatures and the forecast temperatures changed by a significant amount. The data were not yet in final form for consideration in this report, although the tropospheric values appear to have general agreement with HadAT and RATPAC

249

250 Figure 2.2 Terminology and vertical profiles for the temperature products referred to in this report. Radiosonde-
251 based layer temperatures ($T_{850-300}$, T_{100-50}) are height-weighted averages of the temperature in those layers. Satellite-
252 based temperatures (T_{2LT} , T_2 , and T_4) are mass-weighted averages with varying influence in the vertical as depicted
253 by the curved profiles, i.e., the larger the value at a specific level, the more that level contributes to the overall
254 satellite temperature average. The subscript simply indicates the layer where 90% of the information for the satellite
255 average originates.

256
257 Notes: (1) because radiosondes measure the temperature at discrete (mandatory) levels, their information may be
258 used to create a temperature value that mimics a satellite temperature (Text Box 2.1), (2) layer temperatures vary
259 from equator to pole so the pressure and altitude relationship here is based on the atmospheric structure over the
260 conterminous U.S., (3) about 10% (5%) of the value of T_{2LT} (T_2) is determined by the surface character and
261 temperature, (4) T^*_T and T^*_G are simple retrievals, being linear combinations of 2 channels, T_2 and T_4 .

262

263 Throughout the radiosonde era there have been numerous changes in stations, types of
264 instrumentation, and data processing methods that can create data discontinuities. Because
265 radiosondes are expendable instruments, instruments are more easily changed than for the more
266 permanent surface sites. The largest discontinuities appear to be related to solar heating of the
267 temperature sensor and changes in design and/or data adjustments intended to deal with this
268 problem. These discontinuities have greatest impact at stratospheric levels (the stratosphere's
269 lower boundary is ~16 km in the tropics, dropping to < 10 km in the high latitudes, Figure 2.2),

270 where direct sunlight can cause radiosonde-measured temperatures to rise several °C above
271 ambient temperatures. For example, when Australia and U.S. stations changed instrumentation to
272 Vaisala RS-80, processed stratospheric temperatures shifted downward by 1 to 3°C (Parker et al.,
273 1997, Christy et al., 2003). Many other sources of system-dependent bias exist (which often
274 affect the day and night releases differently), including icing of the sensors in regions of super-
275 cooled water, software errors in some radiosonde systems, poor calibration formulae, and
276 operator errors. Documentation of these many changes is limited, especially in the earlier
277 decades.

278

279 *2. Passive Satellite Instrumentation*

280 Unlike radiosondes, passive satellite observations of microwave and infrared brightness
281 temperatures sample thick atmospheric layers (and may include surface emissions), depicted as
282 weighting functions in Figure 2.2. These measurements may be thought of as bulk atmospheric
283 temperatures, as a single value describes the entire layer. Although this bulk measurement is less
284 informative than the detailed information from a radiosonde, horizontal coverage is far superior,
285 and consistency can be checked by comparing the appropriate vertical average from a radiosonde
286 station against nearby satellite observations (see Box 2.2). Furthermore, because there are far
287 fewer instrument systems than in radiosonde datasets, it is potentially easier to isolate and adjust
288 problems in the data.

289

290 The space and time sampling of the satellites varies according to the orbit of the spacecraft,
291 though the longer satellite datasets are based on polar orbiters. These spacecraft circle the globe

292 from pole to pole while maintaining a nominally constant orientation relative to the sun (sun-
293 synchronous). In this configuration, the spacecraft completes about 14 roughly north-south orbits
294 per day as the earth spins eastward beneath it, crosses the equator at a constant local time, and
295 provides essentially global coverage. Microwave measurements utilized in this report begin in
296 late 1978 with the TIROS-N spacecraft using a 4-channel radiometer (Microwave Sounding Unit
297 or “MSU”) which was upgraded in 1998 to a 16-channel system (advanced MSU or “AMSU”)
298 with better calibration, more stable station-keeping (i.e., the timing and positioning of the
299 satellite in its orbit – see discussion of “Diurnal Sampling” below), and higher spatial and
300 temporal sampling resolution.

301
302 Laboratory estimates of precision (random error) for a single MSU measurement are 0.25 °C.
303 Thus with 30,000 observations per day, this error is inconsequential for global averages. Of far
304 more importance are the time varying biases which arise once the spacecraft is in orbit; diurnal
305 drifting, orbital decay, intersatellite biases and calibration changes due to heating of the
306 instrument in space (see section 3 below.)

307
308 While bulk-layer measurements offer the robustness of a large-volume sample, variations within
309 the observed layer are masked. This is especially true for the layer centered on the mid-
310 troposphere (T_2) for which the temperatures of both lower stratospheric and tropospheric levels,
311 which generally show opposite variations, are merged (Figure 2.2). Three MSU/AMSU-based
312 climate records are presented in this report, prepared by Remote Sensing Systems (RSS) of Santa
313 Rosa, California, The University of Alabama in Huntsville (UAH), and The University of

314 Maryland (UMd). Subscripts identify the team, for example, T_2 will be listed as T_{2-R} , T_{2-A} and
315 T_{2-M} for RSS, UAH and UMd respectively.

316

317 Some polar orbiters also carry the Stratospheric Sounding Unit (SSU), an infrared sensor for
318 monitoring deep layer temperatures above about 15 km. SSU data have been important in
319 documenting temperature variations at higher elevations than observed by MSU instruments on
320 the same spacecraft (Ramaswamy et al., 2001). Generally, the issues that complicate the creation
321 of long-term MSU time series also affect the SSU, with the added difficulty that infrared
322 channels are more sensitive to variations in atmospheric composition (e.g., volcanic aerosols,
323 water vapor, etc.).

324

325 Future observing systems using passive-satellite methods include those planned for the National
326 Polar-orbiting Operational Environmental Satellite System (NPOESS) series: the microwave
327 sensors Conical scanning Microwave Imager/Sounder (CMIS) (which will succeed the Special
328 Sensor Microwave/Imager [SSM/I]), Special Sensor Microwave Imager/Sounder (SSM/I-S) and
329 Advanced Technology Microwave Sounder (ATMS) (which will succeed the AMSU), and the
330 infrared sensor Cross-track Infrared Sounder (CrIS) (following the High-resolution Infrared
331 Radiation Sounder [HIRS]). Each of these will follow measuring strategies that are both similar
332 (polar orbit) and dissimilar (e.g., CMIS's conical scanner vs. AMSU's cross-track scanner) but
333 add new spectral and more detailed resolution.

334

335 *3. "Active" satellite instrumentation*

336 A relatively recent addition to temperature monitoring is the use of Global Positioning System
337 (GPS) radio signals, whose time of transmission through the atmosphere is altered by an amount
338 proportional to air density and thus temperature at levels where humidity can be ignored
339 (Kursinski et al., 1997). A key advantage of this technique for climate study is that it is self-
340 calibrating. Current systems are accurate in the upper troposphere and lower to middle
341 stratosphere where moisture is insignificant, but at lower levels, humidity becomes a
342 confounding influence on density. Future versions of this system may overcome this limitation
343 by using shorter wavelengths to measure humidity and temperature independently. Because of
344 the relatively short GPS record and limited spatial coverage to date, its value for long-term
345 climate monitoring cannot yet be definitively demonstrated.

346

347 c) Operational Reanalyses

348 Operational reanalyses (hereafter simply “reanalyses”) will be discussed here in chapter 2, but
349 their trends presented only sparingly in the following chapters because of evidence that they are
350 not always reliable, even during the recent period. All authors expressed concern regarding
351 reanalyses trends, a concern that ranged from unanimous agreement that stratospheric trends
352 were likely spurious to mixed levels of confidence regarding tropospheric trends (see chapter 3).
353 Surface temperature trends are a separate issue as reanalyses values are indirectly *estimated*
354 rather than *observed* (see below). However, reanalyses products hold significant potential for
355 addressing many aspects of climate variability and change.

356

357 Reanalyses are not separate observing systems, but are mathematically blended products based

358 upon as many observing systems as practical. Observations are assimilated into a global weather
359 forecasting model to produce analyses that are most consistent with both the available data
360 (given their imperfections) and the assimilation model. The model, which is constrained by
361 known but parameterized atmospheric physics, generates a result that could be more accurate and
362 physically self-consistent than can be obtained from any one observing system. Some data are
363 rejected or adjusted based on detected inconsistencies. Importantly, the operational procedure
364 optimizes only the accuracy of each near-instantaneous (“synoptic”) analysis. Time-varying
365 biases of a few hundredths or tenths of a degree, which contribute little to short time scale
366 weather error, present a major problem for climate trends, and these are not minimized (e.g.,
367 Sherwood, 2000). The two main reanalyses available at this time are the National Centers for
368 Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR)
369 reanalysis of data since 1948 (Kalnay et al., 1996) and the European Center for Medium-Range
370 Weather Forecasts Re-Analysis-40 (ECMWF ERA-40) beginning in 1957 (Simmons, 2004).

371
372 Because many observational systems are employed, a change in any one will affect the time
373 series of the final product. Reanalyses would be more accurate than lower-level data products for
374 climate variations only if the above shortcomings were outweighed by the benefits of using a
375 state-of-the-art model to treat unsampled variability. Factors that would make this scenario likely
376 include a relatively skillful forecast model and assimilation system, large sampling errors (which
377 are reduced by reanalysis), and small systematic discrepancies between different instruments.
378 However, current models tend to have significant intrinsic biases that can particularly affect
379 reanalyses when sampling is sparse.

380

381 Reanalysis problems that influence temperature trend calculations arise from changes over time
382 in (a) radiosonde and satellite data coverage, (b) radiosonde biases (or in the corrections applied
383 to compensate for these biases), (c) the effectiveness of the bias corrections applied to satellite
384 data and (d) the propagation of errors due to an imprecise formulation of physical processes in
385 the models. For example, since few data exist for the Southern Hemisphere before 1979,
386 temperatures were determined mainly by model forecasts; a cold model bias (in ERA-40, for
387 example) then produces a spurious warming trend when real data become available. Indirect
388 effects may also arise from changes in the biases of other fields, such as humidity and clouds,
389 which affect the model temperature (Andrae et al., 2004; Simmons et al., 2004.).

390

391 Different reanalyses do not employ the same data. NCEP/NCAR does not include surface
392 temperature observations over land but the analysis still produces estimated near-surface
393 temperatures based on the other data (Kalnay and Cai, 2003). On the other hand, ERA-40 does
394 incorporate these but only indirectly through their modeled impacts on soil temperature and
395 surface humidity (Simmons et al., 2004). Thus, the 2-meter air temperatures of both reanalyses
396 may not track closely with surface observations over time (Kalnay and Cai, 2003). SSTs in both
397 reanalyses are simply those of the climate records used as input.

398

399 For upper air reanalyses temperatures, simultaneous assimilation of radiosonde and satellite data
400 is particularly challenging because the considerably different instrument characteristics and
401 products make it difficult to achieve the consistency possible in theory. Despite data adjustments,

402 artifacts still remain in both radiosonde and satellite analyses; these produce the largest
403 differences in the lower stratosphere in current reanalysis datasets (e.g., Pawson and Fiorino,
404 1999; Santer et al., 1999; Randel, 2004). Some of these differences can now be explained, so that
405 future reanalyses will very likely improve on those currently available. However any calculation
406 of deep-layer temperatures from reanalyses which require stratospheric information are
407 considered in this report to be suspect (see Figure 2.2, T_T , T_2 , T_4 , and T_{100-50}).

408

409 d.) Simple retrieval techniques

410 A problem in interpreting MSU (i.e., broad-layer) temperature trends is that many channels
411 receive contributions from both the troposphere and stratosphere, yet temperatures tend to
412 change oppositely in these two layers with respect to both natural variability and predicted
413 climate change. In particular, MSU Channel 2 (T_2) receives 10-15% of its emissions from the
414 stratosphere (Spencer and Christy, 1992), which is a significant percentage because stratospheric
415 cooling in recent decades far exceeds tropospheric warming. It is impossible to eliminate all
416 physical stratospheric influences on MSU 2 by simply subtracting out MSU 4 (T_4) influences
417 because any linear combination of these two channels still retains stratospheric influence
418 (Spencer et al., 2005), which will lead to errors. However, it is possible to rely upon radiosonde-
419 measured correlations between tropospheric and stratospheric temperature fluctuations in order
420 to find what linear combination of these two channels leads to a near-cancellation of these errors,
421 i.e., where y is determined by regression:

422 *Tropospheric Retrieval* = $(1+y) \cdot (T_2) - (y) \cdot (T_4)$. The challenge here is that the resulting
423 relationship depends on the training dataset (radiosondes) being globally or tropically

424 representative (i.e., the troposphere/stratosphere boundary varies spatially and thus the
425 relationship between T_2 and T_4 does as well) and free from significant biases.

426

427 Fu et al. (2004) used a radiosonde dataset to estimate values for y (for the globe, tropical region,
428 and Northern and Southern Hemispheres) that most closely reproduced the monthly variability of
429 mean temperature from 850 to 300 hPa, spanning most of the troposphere. From physical
430 arguments, however, it is clear that the true physical contributions to the retrieval come from a
431 broader range of altitudes, which, in the tropics, approximately span the full troposphere (Fu and
432 Johanson, 2004; 2005). Although derived values of y are robust ($\pm 10\%$, Gillett et al., 2004,
433 Johanson and Fu, 2005), the veracity of the retrieval for climate change has been a subject of
434 debate (due to the accuracy and global representativeness issues mentioned above), and will be
435 further addressed in Chapter 4.

436

437 In the following chapters, two simple retrievals will be utilized in comparison studies with the
438 products of the observing systems. The tropospheric retrieval generated from global mean
439 values of T_2 and T_4 , is identified as T^*_G where $y = 0.143$ (Johanson and Fu, 2005), and when
440 applied to tropical mean values is identified as T^*_T where $y = 0.100$ (Fu and Johanson, 2005).

441

442 A summary of the sources of biases and uncertainties for the datasets and other products
443 described above is given at the end of this chapter. There are several datasets yet to be generated
444 (or not yet at a stage sufficient for climate analysis) from other sources that have the potential to
445 address the issue of vertical temperature distribution. A generic listing of these datasets with a

446 characterization of their readiness is given in Table 2.1.

447

448

449

450

451 Table 2.1 Dataset types and readiness for high quality climate monitoring related to the vertical temperature
452 structure of the atmosphere. "Usage of Data" indicates the level of application of the dataset to the vertical
453 temperature issue. "Understanding" indicates the level of confidence (or readiness) in the dataset to provide accurate
454 information on this issue.

455

I

456

457 : Adequate for long-term global climate variations =Green

458 Improvements or continued research needed for long-term global climate variations = Yellow

459 Problems exist or a lack of analysis to date inhibit long-term global climate variation studies = Red

460 P: Polar orbiter, twice per day per orbiter per ground location

461 G: Geostationary, many observations per day per ground location

462 2x Day: Twice daily at site

463 Hrly: Up to several times per day, many report hourly

464 Syn: Synoptic or generally up to 8 times per day. (Buoys continuous)

465

466 **2. ANALYSIS OF CLIMATE RECORDS**

467 Two factors can interfere with the accurate assessment of climate variations over multi-year
468 periods and relatively large regions. First, much larger variability (weather or “atmospheric
469 noise”) on shorter time or smaller space scales can, if inadequately sampled by the observing
470 network, bias estimates of relatively small climate changes. For example, an extended heat wave
471 in an un-instrumented region accompanied by a compensating cold period in a well-instrumented
472 region may be interpreted as a “global” cold period when it was not. Such biases can result from
473 either spatial or temporal data gaps (Agudelo and Curry, 2004). Second, instrumental errors,
474 particularly biases that change over time, can create erroneous trends. The seriousness of each

475 problem depends not only on the data available but also on how they are analyzed. Finally, even
476 if global climate is known accurately at all times and places, there remains the issue of what
477 measures to use for quantifying climate change; different choices can sometimes create different
478 impressions, e.g., linear trends versus low frequency filtered analyses that retain some
479 information beyond a straight line.

480

481 Upper air layers experience relatively rapid horizontal smoothing of temperature variations, so
482 that on annual mean time scales, the atmosphere is characterized by large, coherent anomaly
483 features, especially in the east-west direction (Wallis, 1998, Thorne et al., 2005b). As a result, a
484 given precision for the global mean value over, say, a year, can be attained with fewer, if
485 properly spaced, upper air measurement locations than at the surface (Hurrell et al., 2000). Thus,
486 knowledge of global, long-term changes in upper-air temperature is limited mainly by
487 instrumental errors. However, for some regional changes (e.g., over sparsely observed ocean
488 areas) sampling problems may compete with or exceed instrumental ones.

489

490 a) Climate Records

491 Various groups have developed long time series of climate records, often referred to as Climate
492 Data Records (CDRs) (NRC, 2000b; 2000c; 2004) from the raw measurements generated by
493 each observing system. Essentially, climate records are time series that include estimates of error
494 characteristics so as to enable the study of climate variation and change on decadal and longer
495 time scales with a known precision.

496

497 Long-term temperature changes occur within the context of shorter-term variations, which are
 498 listed in Table 2.2. These shorter changes include: periodic cycles such as day-night and seasonal
 499 changes; fairly regular changes due to synoptic weather systems, the Quasi-Biennial Oscillation
 500 (QBO), and the El Niño-Southern Oscillation (ENSO); and longer-term variations due to
 501 volcanic eruptions or internal climate dynamics. These changes have different vertical
 502 temperature signatures, and the magnitude of each signal may be different at the surface, in the
 503 troposphere, and in the stratosphere. Details are given in Table 2.2. Some of these signals can
 504 complicate the identification of temperature trends in climate records.

505 Table 2.2 Listing of atmospheric temperature variations by time scale and their properties. (Time scales and sources
 506 of global temperature variations)
 507

Variation	Description	Dominant Period	Approx. Magnitude	Detectability	Effect on Trend Estimates
Diurnal ¹	Warmer days than nights, due to earth's rotation on its axis affecting solar heating.	Daily (outside of polar regions)	Highly variable. Surface skin T changes up to 35K. Boundary layer changes <10K. Free tropospheric changes <1K. Stratospheric changes ~0.1-1 K.	Well detected in surface data. Poorly detected globally in the troposphere and stratosphere due to infrequent sampling (once or twice daily) and potential influence of measurement errors with their own diurnal signal. A few ground-based systems	Satellite data require adjustment of drift in the local equatorial crossing time of spacecraft orbits. Inadequate quantification of the true diurnal cycle hinders this adjustment. Different diurnal adjustments by different groups may

Variation	Description	Dominant Period	Approx. Magnitude	Detectability	Effect on Trend Estimates
				detect signal well.	partly account for differences in trend estimates.
Synoptic ²	Temperature changes associated with weather events, such as wave and frontal passages, due to internal atmospheric dynamics.	1-4 days	Up to ~15K or more at middle latitudes, ~3K in Tropics.	Well detected by observing systems designed to observe meteorological variability.	Not significant, but contributes to noise in climate data records.
Intraseasonal ³	Most notably, an eastward-and vertically-propagating pattern of disturbed weather in the tropical Indo-Pacific ocean region, of unknown cause. Also, atmospheric “blocking” and wet/dry land surface can cause intra-seasonal variations at mid-latitudes.	40-60 days (Tropics), < 180 days (mid-latitudes)	1-2 K at surface, less aloft (tropics), larger in mid-latitudes.	Temperature signals moderately well detected, with tropical atmosphere limited by sparse radiosonde network and IR-based surface temperature limited by cloud. Reanalysis data are useful.	Not significant due to short duration, but may be important if character of the oscillation changes over time.
Annual ⁴	Warmer summers than winters, and shift in position of major precipitation	Yearly	~2-30 K; greater over land than sea, greater at high than low latitudes,	Well observed.	Trends are often computed from “anomaly” data, after the

Variation	Description	Dominant Period	Approx. Magnitude	Detectability	Effect on Trend Estimates
	zones, due to tilt of the earth's axis of rotation affecting solar heating.		greater near the surface and tropopause than at other heights.		mean annual cycle has been subtracted. Changes in the nature of the annual cycle could affect annual-average trends.
Quasi-Biennial Oscillation (QBO) ⁵	Nearly periodic wind and temperature changes in the equatorial stratosphere, due to internal atmospheric dynamics.	Every 23-28 months (average of 27 months because occasionally periods of up to 36 months occur.)	Up to 10 K locally, ~0.5 K averaged over the tropical stratosphere.	Fairly well observed by equatorial radiosonde stations and satellites.	Like ENSO, can influence trends in short data records, but it is relatively easy to remove this signal.
Interannual ⁶	Multiannual variability due to interaction of the atmosphere with dynamic ocean and possibly land surfaces; most notably, ENSO. Can also be caused by volcanic eruptions.	ENSO events occur every 3-7 years and last 6-18 months; major volcanic eruptions, irregular but approximately every 5-20 years with effects lasting ~ 2 years.	Up to 3K in equatorial Pacific (ENSO), smaller elsewhere. Volcanic warming of stratosphere can exceed 5K in tropics cooling of surface <2K.	Fairly well observed, although the vertical structure of ENSO is not as well documented, due to sparseness of the tropical radiosonde network.	ENSO affects surface global mean temperatures by ±0.4K, and more in the tropical troposphere. Large ENSO events near the start or end of a data record can strongly affect computed trends, as was the case for the 1997-98 event. Changes in ENSO

Variation	Description	Dominant Period	Approx. Magnitude	Detectability	Effect on Trend Estimates
					frequency or strength affect (and may be coupled with) long-term trends.
Decadal to interdecadal oscillations and shifts ⁷	Like interannual, but longer time scales. Prominent example is the PDO/ Interdecadal Pacific Oscillation. Despite long time scale, changes can occur as abrupt shifts, for example, a warming shift around 1976. Others include regional changes in the North Atlantic, Pacific-North American, Arctic, and the Antarctic oscillations. Some changes also caused by 11-year solar cycle.	Poorly known; 50-year PDO cycle suggested by 20 th -century observations; others a decade or two; solar 11-year cycle detectable also.	Not well studied. The 1976-77 shift associated with a sharp warming of at least 0.2K globally, though difficult to distinguish from anthropogenic warming. 11-year cycle leads to stratospheric temperature changes of ~2K, and interacts with the Quasi-Biennial Oscillation (QBO).	Relatively large regional changes are well observed, but global expression is subject to data consistency issues over time and possible real changes.	Can account for a significant fraction of linear trends calculated over periods of a few decades or less regionally. Such trends may differ significantly from one such period to the next.
Sub-centennial 60-80 year fluctuation or "Atlantic	Fluctuates in instrumental and paleo data at least back to c.1600. Seems to	60-80 years	~ ±0.5C in parts of the Atlantic. Apparently detectable in	Detectable globally above the noise, clear in North Atlantic SST.	Effects small globally, but probably detectable in last few

Variation	Description	Dominant Period	Approx. Magnitude	Detectability	Effect on Trend Estimates
Multidecadal Oscillation ⁸	particularly affect Atlantic sector. Possible interhemispheric component.		global mean ~ ±0.1C		decades. Readily detectable over this period in North Atlantic Ocean where it clearly affects surface temperature trends and probably climate generally.
Centennial and longer variations ⁹	Warming during 20 th Century due to human influences, solar, and internal variability. Earlier changes included the “little ice age” and “medieval warm period.”	None confirmed, though 1500 year Bond cycle possible.	20 th century warming of ~0.6K globally appears to be as large or larger than other changes during the late Holocene.	Surface warming during 20 th century fairly well observed; proxies covering earlier times indicated 20 th century warmer than the past 5 centuries	Natural temperature variations occur on the longest time scales accessible in any instrumental record.

508

509

510 ¹ Christy et al., 2003; Mears et al., 2003; Vinnikov and Grody. 2003; Dai and Trenberth, 2004; Jin, 2004; Seidel et
 511 al., 2005.

512 ² Palmen and Newton, 1969

513
 514 ³ Duvel et al.,2004.

515
 516 ⁴ Wallace and Hobbs, 1977

517
 518 ⁵ Christy and Drouilhet, 1994; Randel et al., 1999; Baldwin et al., 2001

519
 520 ⁶ Parker and Brownscombe, 1983; Pan and Oort, 1983; Christy and McNider, 1994; Parker et al., 1996; Angell,

521 2000; Robock, 2000; Michaels and Knappenberger, 2000; Santer et al., 2001; Free and Angell, 2002a; Trenberth et
522 al., 2002; Seidel et al., 2004; Seidel and Lanzante, 2004

523
524 ⁷ Labitzke, K., 1987; Trenberth and Hurrell, 1994; Lean et al., 1995; Zhang et al., 1997; Thompson et al., 2000;
525 Douglass and Clader, 2002; Seidel and Lanzante, 2004; Hurrell et al., 2003; Folland et al., 1999; Power et al., 1999;
526 Folland et al. 2002.

527
528 ⁸ Schlesinger and Ramankutty, 1994; Mann et al., 1998; Folland et al., 1999; Andronova and Schlesinger, 2000;
529 Goldenberg et al., 2001; Enfield et al., 2001

530
531 ⁹ Folland et al., 2001a.

532 Our survey of known atmospheric temperature variations, how well they are measured, and their
533 impact on trend estimates suggests that most observing systems are generally able to quantify
534 well the magnitudes of change associated with shorter time scales. For longer time scale changes,
535 where the magnitudes of change are smaller and the stability requirements more rigorous, the
536 observing systems face significant challenges (Seidel et al., 2004).

537

538 b) Measuring Temperature Change

539

540 Over the last three to five decades, global surface temperature records show increases of almost
541 two tenths of a °C per decade. Explaining atmospheric and surface trends therefore demands
542 relative accuracies of a few hundredths of a degree per decade in global time series of both
543 surface and upper-air observations. As this and subsequent chapters will show, the effects of
544 instrumental biases on the global time series are significantly larger than a few hundredths of a
545 degree for the upper-air data, though the global surface temperature compilations do appear to
546 reach this level of precision in recent decades (Folland et al., 2001b). These biases, especially
547 those of the upper air, must therefore be understood and quantified rather precisely (see section 3
548 below). For this fundamental reason, reliable assessment of lapse rate changes remains a

549 considerable challenge.

550

551 Natural modes of climate variability on regional scales are manifested in decadal fluctuations in
552 (a) the tropical Pacific, e.g., ENSO, and (b) the northern latitudes, e.g., the North Atlantic,
553 Pacific-North American and the Arctic atmospheric oscillations (Table 2.2). Even fluctuations on
554 longer time scales have been proposed, e.g., the Atlantic Multidecadal Oscillation/60-80 year
555 variation (Schlesinger and Ramankutty, 1994; Enfield et al., 2001). Each of these phenomena is
556 associated with regions of both warming and cooling. Distinguishing slow, human-induced
557 changes from such phenomena requires identifying the patterns and separating the influences of
558 such modes from the warming signal (e.g., as attempted by SST by Folland et al., 1999.) In
559 addition, these oscillations could themselves be influenced by human-induced atmospheric
560 changes (Hasselmann, 1999).

561

562 **3. LIMITATIONS**

563

564 A key question addressed in this report is whether climate records built by investigators using
565 various components of the observing system can meet the needs for assessing climate variations
566 and trends with the accuracy and representativeness which allows any human attribution to be
567 reliably identified. Climate record builders have usually underestimated the overall uncertainty in
568 their products by relying on traditional sources of uncertainty that can be quantified using
569 standard statistical methods. For example, published linear trend values exist of the same
570 temperature product from the same observing system whose error estimates do not overlap,

571 indicating serious issues with error determination. Thus, in 2003, three realizations of T_2 (or
572 MSU channel 2) 1979-2002 global trends were published as $+0.03 \pm 0.05$, $+0.12 \pm 0.02$, and $+0.24$
573 ± 0.02 °C per decade (Christy et al., 2003; Mears et al., 2003; and Vinnikov and Grody, 2003,
574 respectively.) Over 40% of the difference between the first two trends is due to the treatment of a
575 single satellite in the 1984-1986 period, with a combination of lesser differences during later
576 satellite periods. The third dataset has more complex differences, though it is being superseded
577 by a version whose trend is now lower (Grody et al., 2004, Vinnikov et al. 2005).

578

579 This situation illustrates that it is very challenging to determine the true error characteristics of
580 datasets (see Chapter 4), although considerably less attention has been paid to this than to the
581 construction of the datasets themselves. In this report, we refer to systematic errors in the climate
582 data records as “construction errors.” Such errors can be thought of as having two fundamentally
583 different sources, *structural* and *parametric* (see Box 2.1). The human decisions that underlie the
584 production of climate records may be thought of as forming a *structure* for separating real and
585 artificial behavior in the raw data. Assumptions made by the experts may not be correct, or
586 important factors may have been ignored; these possibilities lead to *structural uncertainty*
587 (Thorne et al., 2005a) in any trend or other metric obtained from a given the climate record.
588 Experts generally tend to underestimate structural uncertainty (Morgan, 1990). The T_2 example
589 above shows that this type of error can considerably exceed those recognized by the climate
590 record builders. Sorting out which decisions are better than others, given the fact many
591 individual decisions are interdependent and often untestable, is challenging.

592

593 Structural uncertainty is difficult to quantify because this requires considering alternatives to the
594 fundamental assumptions, rather than just to the specific sampling or bias pattern in the available
595 data (the main source of parametric uncertainty). For example, is an apparent diurnal variation
596 due to (a) real atmospheric temperature change, (b) diurnal solar heating of an instrument
597 component, (c) a combination of both, or (d) something else entirely? If the answer is not known
598 *a priori*, different working assumptions may lead to a different result when corrections are
599 determined and applied.

600

601 There may be several ways to identify structural errors. First, it is well known in statistics that
602 one should examine the variability that is left over when known effects are removed in a data
603 analysis, to see whether the residuals appear as small and “random” as implied by the
604 assumptions. Even when the residuals are examined, it is often difficult to identify the cause of
605 any non-randomness. Second, one can compare the results with external or independent data
606 (such as comparing SST and NMAT observations). However, one then encounters the problem
607 of assessing the accuracy of the independent data; because, in the case of global atmospheric
608 temperature data there are no absolute standards for any needed adjustment. Christy et al. (2000)
609 demonstrate the use of internal and external methods for evaluating the error of their upper air
610 time series. They assumed that where agreement of independent measurements exists, there is
611 likely to be increased confidence in the trends. Third, one can try to assess the construction
612 uncertainty by examining the spread of results obtained by multiple experts working
613 independently (e.g., the T₂ example, Thorne et al., 2005a). Unfortunately, though valuable, this
614 does not establish the uncertainties of individual efforts, nor is it necessarily an accurate measure

615 of overall uncertainty. If all investigators make common mistakes, the estimate of construction
616 uncertainty may be too optimistic; but if some investigators are unaware of scientifically sound
617 progress made by others, the estimate can be too pessimistic.

618

619 A general concern regarding all of the datasets used in this analysis - land air temperature, sea
620 surface temperature, radiosonde temperature, and satellite-derived temperature – is the level of
621 information describing the operational characteristics and evolution of the associated observing
622 system. As indicated above, the common factor that creates the biggest differences between
623 analyses of the same source data is the homogeneity adjustments made to account for biases in
624 the raw data. All homogeneity adjustments would improve with better metadata - that is,
625 information (data) about the data (see chapter 6). For satellite-derived temperature, additional
626 metadata such as more data points used in the pre-launch calibration would have been helpful to
627 know, especially if done for differing solar angles to represent the changes experienced on orbit.
628 For the in situ data sets, additional metadata of various sorts likely exist in one form or another
629 somewhere in the world and could be acquired or created. These include the type of instrument,
630 the observing environment, the observing practices and the exact dates for changes in any of the
631 above.

632

633 Below we identify various known issues that led to errors in the datasets examined in this report,
634 and which have generally been addressed by the various dataset builders. Note that reanalyses
635 inherit the errors of their constituent observing systems, though they have the advantage of
636 seeking a degree of consensus among the various observing systems through the constraint of

637 model physics. The complex reanalysis procedure transforms these errors of output data into
638 errors of construction methodology that are hard to quantify.

639

640 **Errors primarily affecting *in situ* observing systems.**

641

642 **Spatial and temporal sampling:** The main source of this error is the poor sampling of oceanic
643 regions, particularly in the Southern Hemisphere, and some tropical and Southern
644 Hemisphere continental regions (see Text Box 2.1). Temporal variations in radiosonde
645 sampling can lead to biases, (e.g., switching from 00 to 12 UTC) but these are generally
646 documented and thus potentially treatable.

647 **Local environmental changes:** Land-use changes, new instrument exposures, etc., create new
648 localized meteorological conditions to which the sensor responds. These issues are most
649 important for land near-surface air temperatures but can also affect the lower elevation
650 radiosonde data. Some changes, e.g., irrigation, can act to increase nighttime minima
651 while decreasing daytime maxima, leaving an ambiguous signal for the daily mean
652 temperature. Such changes are sources of error only if the change in the immediate
653 surroundings of the station is unrepresentative of changes over a larger region.

654 **Changes in methods of observation:** A change in the way in which an instrument is used, as in
655 calibrating a radiosonde before launch, i.e., whether it is compared against a typical
656 outdoor sensor or against a traceable standard.

657 **Changes in data processing algorithms:** A change in the way raw data are converted to
658 atmospheric information can introduce similar problems. For radiosonde data, the raw

659 observations are often not archived and so the effects of these changes are not easily
660 removed.

661

662 **Errors primarily affecting satellite systems**

663

664 **Diurnal sampling:** It is common for polar orbiters to drift slowly away from their “sun-
665 synchronous” initial equatorial crossing times (e.g., 1:30 p.m. to 5 p.m.), introducing
666 spurious trends related to the natural diurnal cycle of daily temperature. The later polar
667 orbiters (since 1998) have more stable station keeping. Diurnal drift adjustments for T_{2LT}
668 and T_2 impact the trend by a few hundredths °C/decade. Changes in local observation
669 time also significantly afflict *in situ* temperature observations, with a lesser impact on the
670 global scale.

671 **Orbit decay:** Variations in solar activity cause expansion and contraction of the thin atmosphere
672 at the altitudes where satellites orbit, which create variable frictional drag on spacecraft.
673 This causes periods of altitude decay, changing the instrument’s viewing geometry
674 relative to the earth and therefore the radiation emissions observed. This issue relates
675 most strongly to T_{2LT} , which uses data from multiple view angles, and is of order 0.1
676 °C/decade.

677 **Calibration shifts/changes:** For satellite instruments, the effects of launch conditions or
678 changes in the within-orbit environment (e.g., varying solar shadowing effects on the
679 spacecraft components as it drifts through the diurnal cycle) may require adjustments to
680 the calibration equations. Adjustment magnitudes vary among the products analyzed in

681 this report but are on the order of 0.1 °C/decade for T_{2LT} and T_2 .

682 **Surface emissivity effects:** The intensity of surface emissions in observed satellite radiances
683 can vary over time due to changes in surface properties, e.g. wet vs. dry ground, rough vs.
684 calm seas, etc., and longer-term land cover changes, e.g., deforestation leading to higher
685 daytime skin temperatures and larger diurnal temperature cycles.

686 **Atmospheric effects:** Atmospheric composition can vary over time (e.g., aerosols), affecting
687 satellite radiances, especially the infrared.

688

689 **Errors affecting all observing systems**

690

691 **Instrument Changes:** Systematic variations of calibration between instruments will lead to
692 time-varying biases in absolute temperature. These involve (a) changes in instruments
693 and their related components (e.g., changes in housing can be a problem for *in situ*
694 surface temperatures), (b) changes in instrument design or data processing (e.g.,
695 radiosondes) and (c) copies of the same instrument that are intended to be identical but
696 are not (e.g., satellites).

697

698 **Errors or differences related to analysis or interpretation**

699

700 **Construction Methodology:** As indicated, this is often the source of the largest differences
701 among trends from datasets and is the least quantifiable. When constructing a
702 homogeneous, global climate record from an observing system, different investigators

703 often make a considerable range of assumptions as to how to treat unsampled or
704 undersampled variability and both random and systematic instrument errors. The trends
705 and their uncertainties that are subsequently estimated are sensitive to treatment
706 assumptions (Free et al., 2002b). For example, the trends of the latest versions of T_2 from
707 the three satellite analyses vary from +0.044 to +0.199 °C/decade (chapter 3), reflecting
708 the differences in the combination of individual adjustments determined and applied by
709 each team (structural uncertainty.) Similarly, the T_2 global trends of the radiosonde-
710 based and reanalyses datasets range from -0.036 to +0.067 °C/decade indicating
711 noticeable differences in decisions and methodologies by which each was constructed.
712 Thus the goal of achieving a consensus with an error range of a few hundredths
713 °C/decade is not evidenced in these results.

714 **Trend Methodology:** Differences between analyses can arise from the methods used to
715 determine trends. Trends shown in this report are calculated by least squares linear
716 regression.

717 **Representativeness:** Any given measure reported by climate analysts could under- or overstate
718 underlying climatic behavior. This is not so much a source of error as a problem of
719 interpretation. This is often called statistical error. For example, a trend computed for
720 one time period (say, 1979-2004) is not necessarily representative of either longer or
721 earlier periods (e.g., 1958-1979), so caution is necessary in generalizing such a result. By
722 the same token, large variations during portions of the record might obscure a small but
723 important underlying trend. (See Appendix for Statistical Uncertainties.)

724

725

726 4. IMPLICATIONS

727 The observing systems deployed since the late 1950s, and the subsequent climate records derived
728 from their data, have the capability to provide information suitable for the detection of many
729 temperature variations in the climate system. These include temperature changes that occur with
730 regular frequency, e.g., daily and annual cycles of temperature, as well as non-periodic events
731 such as volcanic eruptions or serious heat and cold waves. The data from these systems also have
732 the potential to provide accurate trends in climate over the last few decades (and over the last
733 century for surface observations), once the raw data are successfully adjusted for changes over
734 time in observing systems, practices, and micro-climate exposure to produce usable climate
735 records. Measurements from all systems require such adjustments and this report relies on
736 adjusted datasets. The details of making such adjustments when building climate records from
737 the uncorrected observations are examined in the following chapters.

738

739 Text Box 2.1: Comparing Radiosonde and Satellite Temperatures

740 Attempts to compare temperatures from satellite and radiosonde measurements are hindered by a
741 mismatch between the respective raw observations. While radiosondes measure temperatures at
742 specific vertical levels, satellites measure radiances which can be interpreted as the temperature
743 averaged over a deep layer. To simulate a satellite observation, the different levels of
744 temperature in the radiosonde sounding are proportionally weighted to match the profiles shown
745 in Figure 2.2. This can be done in one of two ways.

746

747 1. Employ a simple set of geographically and seasonally invariant coefficients or weights, called
748 a static weighting function. These coefficients are multiplied by the corresponding set of
749 temperatures at the radiosonde levels and the sum is the simulated satellite temperature. Over
750 land, the surface contributes more to the layer-average than it does over the ocean, and this
751 difference is taken into account by slightly different sets of coefficients applied to land vs. ocean
752 calculations. This same method may be applied to the temperature level data of global
753 reanalyses. We have applied the “static weighting function” approach in this report.

754

755 2. Take into account the variations in the air mass temperature, surface temperature and pressure,
756 and atmospheric moisture (Spencer et al., 1990). Here, the complete radiosonde temperature
757 and humidity profiles are ingested into a radiation model to generate the simulated satellite
758 temperature (e.g., Christy and Norris, 2004). This takes much more computing power to
759 calculate and requires humidity information, which for radiosondes is generally of poorer quality
760 than temperature information or is missing entirely. For climate applications, in which the time
761 series of large-scale anomalies is the essential information, the output from the two methods
762 differs only slightly.

763

764 There are practical difficulties in generating long time series of simulated satellite temperatures
765 under either approach. To produce a completely homogeneous data record, the pressure levels
766 used in the calculation must be consistent throughout time, i.e., always starting at the surface and
767 reaching the same designated altitude. If, for example, soundings achieved higher elevations as
768 time went on, there would likely be a spurious trend due to the effects of having measured

769 observations during the latter period of record, while by necessity, relying on estimates for the
770 missing values in the earlier period. We also note that HadAT utilizes 9 pressure levels for
771 simulating satellite profiles while RATPAC use 15, so differences can arise from these differing
772 inputs.

773

774 An additional complication is that many radiosonde datasets and reanalyses may provide data at
775 mandatory levels beginning with 1000 and/or 850 hPa, i.e., with no identifiable surface. Thus,
776 the location of the material surface, and its temperature, can only be estimated so that an
777 additional source of error to the anomaly time series may occur. There are a number of other
778 processing choices available when producing a time series of simulated satellite data for site-by-
779 site comparisons between actual satellite data and radiosondes (or reanalyses) and these also
780 have the potential to introduce non-negligible biases.

781

782 Averaging of spatially incomplete radiosonde observations for comparison of global and tropical
783 anomalies also introduces some error (Agudelo and Curry, 2004). In this report we have first
784 zonally averaged the data, then generated satellite-equivalent measures from these data and
785 finally calculated global and tropical averages. The spatial coverage differs markedly between
786 the two radiosonde datasets. However, as anomalies are highly correlated in longitude the
787 relative poor longitudinal sampling density of RATPAC (and HadAT outside of the NH mid-
788 latitudes) is not necessarily an impediment (Hurrell et al., 2000). Comparing global averages
789 estimated using only those zonally-averaged grids observed at RATPAC station sites by MSU
790 versus the globally complete fields from MSU, a sampling error of less than ± 0.05 °C/decade

791 was inferred for T_{2LT} . Satellite and reanalyses are essentially globally complete and thus do not
792 suffer from spatial subsampling.

793 References

794

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

797

798 Aires, F., C. Prigent, and W.B. Rossow, 2004: Temporal interpolation of global surface skin
799 temperature diurnal cycle over land under clear and cloudy conditions. *J. Geophys. Res.*,
800 109, doi:10.1029/2003JD003527.

801

802 Andrae, U., N. Sokka and K. Onogi, 2004: The radiosonde temperature bias corrections used in
803 ERA-40. ERA-40 Project Series #15. 37 pp. European Centre for Medium Range
804 Weather Forecasts. Available at <http://www.ecmwf.int/publications/>

805

806 Angell, J.K., 2000: Tropospheric temperature variations adjusted for El Niño, 1958-1998. *J.*
807 *Geophys. Res.*, 105, 11841-11849.

808

809 Andronova, N.G. and M.E. Schlesinger, 2000: Causes of global temperature changes during the
810 19th and 20th centuries. *Geophys. Res. Lett.*, 27, 2137-2140.

811

812 Baldwin, M.P., L.J. Gray, T.J. Dunkerton, K. Hamilton, P.H. Haynes, W.J. Randel, J.R. Holton,
813 M.J. Alexander, I. Hirota, T. Horinouchi, D.B.A. Jones, J.S. Kinnersley, C. Marquardt,
814 K. Sato, and M. Takahashi et al., 2001: The quasi-biennial oscillation. *Rev. Geophys.*, **39**,
815 179-229.

816

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

819

820 Christy, J.R., R.W. Spencer, and W.D. Braswell, 2000: MSU Tropospheric temperatures: Data
821 set construction and radiosonde comparisons. *J. Atmos. Oceanic Tech.* 17, 1153-1170.

822

823 Christy, J.R., R.W. Spencer, W.B. Norris, W.D. Braswell and D.E. Parker, 2003: Error estimates
824 of Version 5.0 of MSU/AMSU bulk atmospheric temperatures. *J. Atmos. Oceanic Tech.*
825 20, 613-629.

826

827 Christy, J.R. and W.B. Norris, 2004: What may we conclude about tropospheric temperature
828 trends? *Geophys. Res. Lett.* 31, L06211.

829

830 Christy, J.R. and R. T. McNider, 1994 Satellite greenhouse signal. *Nature*, 367, 325.

831

832 Christy, J.R. and S.J. Drouilhet, 1994 Variability in daily, zonal mean lower-stratospheric
833 temperatures. *J. Climate*, 7, 106-120.

834

- 835 Dai, A., and K. E. Trenberth (2004), The diurnal cycle and its depiction in the Community
836 Climate System model, *J. Climate*, 17, 930-951.
837
- 838 Douglass, D.H. and B.D. Clader, 2002: Determination of the climate sensitivity of the earth to
839 solar irradiance. *Geophys. Res. Lett.*, 29, 331-334.
840
- 841 Duval, J.P., R. Roca and J. Vialard, 2004: Ocean mixed layer temperature variations induced by
842 intraseasonal convective perturbations of the Indian Ocean. *J. Atmos. Sci.*, 9, 1004-
843 1023.
844
- 845 Enfield, D.B., Mestas-Nuñez, A.M. and P.J Trimble, 2001: The Atlantic Multidecadal
846 Oscillation and its relation to rainfall and river flows in the continental US. *Geophys.*
847 *Res. Lett.*, 28, 2077-2080.
848
- 849 Folland, C. K. and D. E. Parker (1995). "Correction of instrumental biases in historical sea
850 surface temperature data." *Q. J. Roy. Meteor. Soc.* **121**: 319-367.
851
852
- 853 Folland, C.K., Parker, D.E., Colman, A. and R. Washington, 1999: Large scale modes of ocean
854 surface temperature since the late nineteenth century. Refereed book: Chapter 4, pp73-
855 102 of *Beyond El Nino: Decadal and Interdecadal Climate Variability*. Ed: A. Navarra.
856 Springer-Verlag, Berlin, pp 374.
857
- 858 Folland, C.K., T.R. Karl, J.R. Christy, R.A. Clarke, G.V. Gruza, J. Jouzel, M.E. Mann, J.
859 Oerlemans, M.J. Salinger and S.-W. Wang, 2001a: Observed climate variability and
860 change. In: *Climate Change 2001: The Scientific Basis*. Contribution of Working
861 Group I to the Third Assessment Report of the Intergovernmental Panel on Climate
862 Change [Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X
863 Dai, K. Maskell, and C.A. Johnson (eds.)]. Cambridge University Press, Cambridge,
864 United Kingdom and New York, NY, USA, 881 pp.
865
- 866 Folland, C.K., N.A. Rayner, S.J. Brown, T.M. Smith, S.S. P. Shen, D.E. Parker, I. Macadam,
867 P.D. Jones, R.N. Jones, N. Nicholls and D.M.H. Sexton, 2001b: Global temperature
868 change and its uncertainties since 1861. *Geophys. Res. Lett.*, 28, 2621- 2624.
869
- 870 Folland, C.K., J.A. Renwick, M.J. Salinger and A.B. Mullan, 2002: Relative influences of the
871 Interdecadal Pacific Oscillation and ENSO on the South Pacific Convergence Zone.
872 *Geophys. Res. Lett.*, **29** (13): 10.1029/2001GL014201. Pages 21-1 - 21-4.
873
- 874 Free, M., and J. K. Angell, 2002: Effect of volcanoes on the vertical temperature profile in
875 radiosonde data. *J. Geophys. Res.*, 10.1029/2001JD001128.
876
- 877 Free, M., I. Durre, E.Aguilar, D. Seidel, T.C. Peterson, R.E. Eskridge, J.K. Luers, D. Parker, M.

- 878 Gordon, J. Lanzante, S. Klein, J. Christy, S. Schroeder, B. Soden, and L.M. McMillin,
879 2002: CARDS Workshop on Adjusting Radiosonde Temperature Data for Climate
880 Monitoring: Meeting Summary. *Bull. Amer. Meteor. Soc.*, 83, 891-899.
881
- 882 Fu, Q., C.M. Johanson, S.G. Warren, and D.J. Seidel, 2004: Contribution of Stratospheric
883 Cooling to Satellite-Inferred Tropospheric Temperature Trends. *Nature*, 429, 55-58.
884
- 885 Fu, Q., and C.M. Johanson, 2005: Satellite-derived vertical dependence of tropical tropospheric
886 temperature trends. *Geophys. Res. Lett.* (in press).
887
- 888 Gillett, N. P., B. D. Santer, A. J. Weaver, 2004, Stratospheric cooling and the troposphere,
889 *Nature*, doi:10.1038.
890
- 891 Goldenberg, S.B, Landsea, C.W., Mestas Nunez, A.M. and W.M. Gray. The recent increase in
892 Atlantic Hurricane activity: causes and implications. *Science*, 293, 474- 479.
893
- 894 Grody, Norman C., K. Y. Vinnikov, M. D. Goldberg, J. T. Sullivan, and J. D. Tarpley, 2004.
895 Calibration of multisatellite observations for climatic studies: Microwave Sounding Unit
896 (MSU), *J. Geophys. Res. – Atm.*, 109, D24104, doi:10.1029/2004JD005079, December
897 21, 2004.
898
- 899 Hasselmann, K., 1999: Linear and nonlinear signatures. *Nature*, 398, 755-756.
900
- 901 Haimberger, L., 2004: Homogenization of radiosonde temperature time series using ERA-40
902 analysis feedback information. ERA-40 Project Report Series No. 22. European Centre
903 for Medium Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England. 67
904 pp.
905
- 906 Hurrell, J., S.J. Brown, K.E. Trenberth and J.R. Christy, 2000: Comparison of tropospheric temperatures
907 from radiosondes and satellites: 1979-1998. *Bull. Amer. Met. Soc.*, **81**, 2165-2177.
908
- 909 Hurrell, J.W., Kushnir, Y., Ottensen, G. and M. Visbeck, Eds, 2003: The North Atlantic
910 Oscillation: Climatic Significance and Environmental Impacts. *American Geophysical
911 Union*, pp 279.
912
- 913 Jin, M. (2004), Analysis of land skin temperature using AVHRR observations, *Bull. Amer.
914 Meteorol. Soc.*, 85, 587–600, doi: 10.1175/BAMS-85-4-587.
915
- 916 Johanson, C.M. and Q. Fu, 2005: Robustness of Tropospheric Temperature Trends from MSU
917 channels 2 and 4. *J. Climate* (submitted).
918
- 919 Jones, P.D., T.J. Osborn, K.R. Briffa, C.K. Folland, E.B. Horton, L.V. Alexander, D.E. Parker
920 and N.A. Rayner, 2001: Adjusting for sampling density in grid box land and ocean

- 921 surface temperature time series. *J. Geophys. Res.*, 106, 3371-3380.
922
- 923 Jones, P.D and Moberg, A. 2003. Hemispheric and large scale surface air temperature
924 variations: an extensive revision and an update to 2001. *J. Clim.*, 16, 206-223.
925
- 926 Kalney, E. and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull Amer.*
927 *Metero. Soc.*, 77, 437-471.
928
- 929 Kalnay, E. and M. Cai, 2003: Impact of urbanization and land-use change on climate. *Nature*,
930 423, 528-531.
931
- 932 Kent, E. C. and P. K. Taylor, 2004: Towards Estimating Climatic Trends in SST Data, Part 1:
933 Methods of Measurement. *Journal of Atmospheric and Oceanic Technology*, submitted.
934
- 935 Kent, E. C. and P. G. Challenor, 2004: Towards Estimating Climatic Trends in SST Data, Part 2:
936 Random Errors. *Journal of Atmospheric and Oceanic Technology*, submitted.
937
- 938 Kent, E. C. and A. Kaplan, 2004: Towards Estimating Climatic Trends in SST Data, Part 3:
939 Systematic Biases. *Journal of Atmospheric and Oceanic Technology*, submitted.
940
- 941 Kilpatrick, K. A., G. P. Podesta, et al., 2001: Overview of the NOAA/NASA advanced very
942 high resolution radiometer pathfinder algorithm for sea surface temperature and
943 associated matchup database.” *J. Geophys. Res.* **106**(C5): 9179-9198.
944
- 945 Kursinski E.r., G.A. Hajj, J.T. Schofield, R.P. Linfield and K.R. Hardy, 1997: Observing the
946 Earth’s atmosphere with radio occultation measurements using the Global Positioning
947 System. *J. Geophys. Res.* 102, 23429-23465.
948
- 949 Labitzke, K., 1987: Sunspots, the QBO, and the stratospheric temperature in the north polar
950 region. *Geophys. Res. Lett.*, 14, 535-537.
951
- 952 Lean, J., J. Beer, and R. Bradley, 1995: Reconstruction of solar irradiance since 1610:
953 implications for climate change. *Geophys. Res. Lett.*, 22, 3195-3198.
954
- 955 Mann, M.E., R.S. Bradley and M.K. Hughes, 1998: Global-scale temperature patterns and climate
956 forcing over the past six centuries. *Nature*, 392, 779-787.
957
- 958 McPhaden, M. J., 1995: The Tropical Atmosphere Ocean array is completed. *Bulletin of the*
959 *American Meteorological Society.* 76: 739-741.
960
- 961 Mears, C.A., M.C. Schabel, and F.J. Wentz, 2003: A reanalysis of the MSU channel 2
962 tropospheric temperature record. *J. Climate*, 16, 3650-3664.
963

- 964 Michaels, P.J., and P.C. Knappenberger, 2000: Natural signals in the MSU lower tropospheric
965 temperature record. *Geophys. Res. Lett.* 27, 2905-2908.
966
- 967 Morgan, M. G., 1990: *Uncertainty : a guide to dealing with uncertainty in quantitative risk and*
968 *policy analysis.* Cambridge University Press, 332 pp.
969
- 970 NRC 2000a. *Reconciling Observations of Global Temperature Change.* National Academy
971 Press, 85 pp.
972
- 973 NRC 2000b. *Ensuring the Climate Record from the NPP and NPOESS Meteorological Satellites.*
974 National Academy Press. 51 pp.
975
- 976 NRC 2000c. *Issues in the Integration of Research and Operational Satellite Systems for Climate*
977 *Research II: Implementation.* National Academy Press. 82 pp.
978
- 979 NRC 2004. *Climate Data Records from Environmental Satellites.* National Academy Press. 136
980 pp.
981
- 982 Palmen, E. and C. Newton, 1969: *Atmospheric Circulation Systems: Their Structure and*
983 *Interpretation.* Academic Press.
984
- 985 Pan, Y.-H., and A.H. Oort, 1983: Global climate variations connected with sea surface temperature
986 anomalies in the eastern equatorial Pacific Ocean for the 1958-1973 period. *Mon. Weath.*
987 *Rev.*, 111, 1244-1258.
988
- 989 Parker, D.E. and Brownscombe, 1983: *Nature*, 301, 406-408.
990
- 991 Parker, D.E., M. Gordon, D.P.N. Cullum, D.M.H. Sexton, C.K. Folland and N. Rayner, 1997: A
992 new global gridded radiosonde temperature data base and recent temperature trends.
993 *Geophys. Res. Lett.*, 24, 1499-1502.
994
- 995 Parker, D.E., H. Wilson, P.D. Jones, J.R. Christy and C.K. Folland, 1996: The impact of Mount
996 Pinatubo on world-wide temperatures. *Int. J. Climatol.*, 16, 487-497.
997
- 998 Pawson, S. and M. Fiorino, A comparison of reanalyses in the tropical stratosphere. Part 3:
999 Inclusion of the pre-satellite data era, *Clim. Dyn.*, 1999, 15, 241-250.
1000
- 1001 Power, S., Casey, T., Folland, C.K., Colman, A and V. Mehta, 1999: Inter-decadal modulation of
1002 the impact of ENSO on Australia. *Climate Dynamics*, 15, 319-323.
1003
- 1004 Ramaswamy, V., M.-L. Chanin, J. Angell, J. Barnett, D. Gaffen, M. Gelman, P. Kekhut, Y.
1005 Koshelkov, K. Labitzke, J.-J. R. Lin, A. O'Neill, J. Nash, W. Randel, R. Rood, K. Shine,
1006 M. Shiotani, and R. Swinbank, 2001: Stratospheric temperature trends: Observations and

- 1007 model simulations. *Rev. Geophys.*, 39, 71-122.
- 1008
- 1009 Randel, W.J., F. Wu, R. Swinbank, J. Nash, and A. O'Neill, 1999: Global QBO circulation
1010 derived from UKMO stratospheric analyses. *J. Atmos. Sci.*, **56**, 457-474.
- 1011
- 1012 Rayner, N. A., D. E. Parker, et al., 2003: Global analyses of sea surface temperature, sea ice, and
1013 night marine air temperature since the late nineteenth century. *J. Geophys. Res.* 108(d14).
- 1014
- 1015 Reynolds, R. W., 1993: Impact of Mt Pinatubo aerosols on satellite-derived sea surface
1016 temperatures. *J. Climate*, 6, 768-774.
- 1017
- 1018 Reynolds, R.W., Rayner, N.A. Smith, T.H. Stokes, D.C. and Wang W., 2002: An improved in
1019 situ and satellite SST analysis for climate. *J. Clim.*, 15, 1609-1625.
- 1020
- 1021 Robock, Alan, 2000: Volcanic eruptions and climate. *Rev. Geophys.*, 38, 191-219.
- 1022
- 1023 Santer. B.D., J.J. Hnilo, T.M.L. Wigley, J.S. Boyle, C. Doutriaux, M. Fiorino, D.E. Parker, and
1024 K.E. Taylor, 1999: Uncertainties in observationally based estimates of temperature
1025 change in the free atmosphere. *J. Geophys. Res.*, **104**, 6305-6333.
- 1026
- 1027 Santer, B.D., T.M.L. Wigley, C. Doutriaux, J.S. Boyle and 6 others, 2001: Accounting for the
1028 effects of volcanoes and ENSO in comparisons of modeled and observed temperature
1029 trends. *J. Geophys. Res.*, 106, 28033-28059.
- 1030
- 1031 Schlesinger, M.E. and N. Ramankutty, An oscillation in the global climate system of period 65-
1032 70 years. *Nature*, 367, 723-726.
- 1033
- 1034 Seidel, D.J., J.K. Angell, J.R. Christy, M. Free, S.A. Klein, J.R. Lanzante, C. Mears, D. Parker,
1035 M. Schabel, R. Spencer, A. Sterin, P. Thorne and F. Wentz, 2004: Uncertainty in signals
1036 of large-scale climate variations in radiosonde and satellite upper-air temperature
1037 datasets. *J. Climate*, 17, 2225-2240.
- 1038
- 1039 Seidel, D.J., M. Free, and J. Wang, The diurnal cycle of temperature in the free atmosphere
1040 estimated from radiosondes, *J. Geophys. Res.* (submitted).
- 1041
- 1042 Seidel, D.J., and J.R. Lanzante, 2004: An assessment of three alternatives to linear trends for
1043 characterizing global atmospheric temperature changes, *J. Geophys. Res.* 109, XXX-
1044 XXX, doi:10.1029/2003JD004414, in press.
- 1045
- 1046 Sherwood, S. C., Climate signal mapping and an application to atmospheric tides, *Geophys. Res.*
1047 *Lett.*, 2000, 27, 3525-3528.
- 1048
- 1049 Simmons, A.J., Jones, P.D., da Costa Bechtold, V., Beljaars, A.C.M., Källberg, P., Saarinen, S.,

- 1050 Uppala, S.M., Viterbo, P. and N. Wedi, N. 2004: Comparison of trends and variability
1051 in CRU, ERA-40 and NCEP/NCAR analyses of monthly-mean surface air temperature.
1052 *J. Geophys. Res.*, **109**, No D24 D24115 <http://dx.doi.org/10.1029/2004JD005306>
1053 December 21, 2004.
- 1054
1055 Simmons, A.J., 2004: Development of the ERA-40 Data Assimilation System. 20 pp. European
1056 Centre for Medium Range Weather Forecasts. Available at
1057 <http://www.ecmwf.int/publications/>
1058
- 1059 Smith, T. M. and R. W. Reynolds, 2004: Improved Extended Reconstruction of SST (1854-
1060 1997). *J. Climate*, 17, 2466-2477.
- 1061
1062 Smith, T. M. and R. W. Reynolds, 2005: A global merged land and sea surface temperature
1063 reconstruction based on historical observations (1880-1997). *Journal of Climate*,
1064 Submitted.
- 1065
1066 Spencer, R. W. , J. R. Christy and N. C. Grody, 1990: Global atmospheric temperature monitoring with
1067 satellite microwave measurements: Method and results 1979-1985. *J. Climate*, **3**, 1111-1128.
1068
- 1069 Spencer, R. W. and J. R. Christy, 1992: Precision and radiosonde validation of satellite gridpoint
1070 temperature anomalies, Part II: A tropospheric retrieval and trends during 1979-90. *J. Climate*,
1071 **5**, 858-866.
- 1072
1073 Spencer, R.W., J.R. Christy and W.D. Braswell, 2005: On the estimation of tropospheric
1074 temperature trends from MSU channels 2 and 4. *J. Atmos. Oceanic Tech.*, submitted.
1075
- 1076 Strategic Plan for the U.S. Climate Change Science Program, 2003: A Report by the Climate
1077 Change Science Program and the Subcommittee on Global Change Research.
1078
- 1079 Thompson, D.W.J., J.M. Wallace and G.C. Hegerl, 2000: Annual modes in the extratropical
1080 circulation Part II: trends. *J. Climate*, 13, 1018-1036.
1081
- 1082 Thorne, P.W., D.E. Parker, J.R. Christy and C.A. Mears, 2005: Causes of differences in observed
1083 climate trends. *Bull. Amer. Meteor. Soc.*, in press.
1084
- 1085 Thorne, P.W., D.E. Parker, S.F.B. Tett, P.D. Jones, M. McCarthy, H. Coleman, P. Brohan, and
1086 J.R. Knight, 2005: Revisiting radiosonde upper-air temperatures from 1958 to 2002. *J.*
1087 *Geophys. Res.*, in press.
1088
- 1089 Trenberth, K.E. and J.W. Hurrell, 1994: Decadal atmosphere-ocean variations in the Pacific.
1090 *Clim. Dyn.*, 9, 303-319.
1091
- 1092 Trenberth, K.E., Carron, J.M., Stepaniak, D.P. and S. Worley, 2002: Evolution of the El Nino-

- 1093 Southern Oscillation and global atmospheric surface temperatures. *J. Geophys. Res.*,
1094 107, D8, 10.1029/2000JD000298.
1095
1096 Wallace, J.M. and P.V. Hobbs, 1977: *Atmospheric Science: An Introductory Survey*.
1097 Academic Press, New York, NY, 467 pp.
1098
1099 Wallis, T.W.R., 1998: A subset of core stations from the Comprehensive Aerological Reference
1100 Data Set (CARDS). *J. Climate*, 11, 272-282.
1101
1102 Vinnikov, K.Y., and N.C. Grody, 2003. Global warming trend of mean tropospheric temperature
1103 observed by satellites, *Science*, 302, 269-272.
1104
1105 Vinnikov, K.Y., N.C. Grody, A. Robock, R.J. Stouffer, P.D. Jones and M.D. Goldberg, 2005:
1106 Temperature trends at the surface and the troposphere. *J. Geophys. Res.* submitted.
1107
1108 Zhang, Y., J.M. Wallace, and D.S. Battisti, 1997: ENSO-like interdecadal variability: 1900-
1109 93, *J. Climate*, 10, 1004-1020.
1110

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

Chapter 3

What do observations indicate about the changes of temperature in the atmosphere and at the surface since the advent of measuring temperatures vertically?

Convening Lead Author: John R. Lanzante

Lead Authors: Thomas C. Peterson, Frank J. Wentz and Konstantin Y. Vinnikov

Contributing Authors: Dian J. Seidel, Carl A. Mears, John R. Christy, Chris Forest, Russell S. Vose, Peter W. Thorne and Norman C. Grody

Key Findings

Observed Changes - Surface

Globally, as well as in the tropics, the temperature of the air near the Earth's surface has increased since 1958, with a greater rate of increase since 1979. All three surface temperature datasets are consistent in these conclusions.

- Globally, temperature increased at a rate of about 0.12°C per decade since 1958, and about 0.16°C per decade since 1979.
- In the tropics, temperature increased at a rate of about 0.11°C per decade since 1958, and about 0.13°C per decade since 1979.
- Most, if not all of the surface temperature increase since 1958 occurs starting around the mid-1970s, a time coincident with a previously identified abrupt climate regime shift. However, there does not appear to be a strong jump up in temperature at this time, rather the major part of the rise seems to occur in a more gradual fashion.

Observed Changes - Troposphere

Globally, as well as in the tropics, both balloon-based datasets dating back to 1958 agree that the tropospheric temperature has increased slightly more than that of the surface. Since 1979, due to the considerable disagreement among tropospheric datasets, it is not clear whether the temperature of the troposphere has increased more or less than that of the surface, both globally and in the tropics.

- Globally, temperature increased at a rate of about 0.14°C per decade since 1958 according to the two balloon-based datasets. Since 1979, estimates of the increase from the two balloon and three satellite datasets range from about 0.10 to 0.20°C per decade.
- In the tropics, temperature increased at a rate of about 0.13°C per decade since 1958 according to the two balloon-based datasets. However, since 1979, estimates of the increase from the two balloon and three satellite datasets range from about 0.02 to 0.19°C per decade.
- For the balloon-based estimates since 1958, the major part of the temperature increase appears in the form of an abrupt rise in the mid-1970s, apparently in association with a climate regime shift that occurred at this time.

Observed Changes - Lower Stratosphere

Globally, the temperature of the lower stratosphere has decreased both since 1958 and since 1979. The two balloon-based datasets yield reasonably consistent estimates of the rates of cooling for both time periods. However, since 1979 the two balloon datasets estimate a considerably greater rate of cooling than the two satellite datasets, which themselves disagree.

- 80
- 81 • Globally, the rate of cooling since 1958 is about 0.37°C per decade based on
82 the two balloon datasets. Since 1979, estimates of this decrease are about
83 0.65°C per decade for the two balloon datasets, and from about 0.33 to 0.45°
C per decade for the two satellite datasets.
 - 84 • The bulk of the stratospheric temperature decrease occurred from about the
85 late 1970s to the middle 1990s. It is unclear whether the decrease was
86 gradual or occurred in abrupt steps in the first few years after each major
87 volcanic eruption.

88 Chapter 3: Recommendations

89

- 90 • *Although considerable progress has been made in explaining the causes of*
91 *discrepancies between upper-air datasets, both satellite and balloon-based,*
92 *continuing steps should be taken to thoroughly assess and improve methods used*
93 *to remove time-varying biases that are responsible for these discrepancies.*
- 94 • *New observations should be made available in order to provide more redundancy*
95 *in climate monitoring. Activities should include both the introduction of new*
96 *observational platforms as well as the necessary processing of data from*
97 *currently under-utilized platforms. For example IR and GPS satellite observations*
98 *have not been used to any great extent, the former owing to complications when*
99 *clouds are present and the latter owing to a short period of record. Additionally,*
100 *the introduction of a network of climate quality, reference stations, that include*
101 *reference radiosondes, would place future climate monitoring on a firmer basis.*

102

1. Background

103

104 In this chapter we describe changes in temperature at the surface and in the atmosphere
105 based on four basic types of products derived from observations: surface, radiosonde,
106 satellite and reanalysis. However, we limit our discussion of reanalysis products given
107 their more problematic nature for use in trend analysis (see Chapter 2); only a few trend
108 values are presented for illustrative purposes.

109

110 Each of these four generic types of measurements consists of multiple datasets prepared
111 by different teams of data specialists. The datasets are distinguished from one another by
112 differences in the details of their construction. Each type of measurement system as well
113 as each particular dataset has its own unique strengths and weaknesses. Because it is
114 difficult to declare a particular dataset as being “the best,” it is prudent to examine results
115 derived from more than one “credible” dataset of each type. Also, comparing results from
116 more than one dataset provides a better idea of the uncertainties or at least the range of
117 results. In the interest of clarity and conciseness, we have chosen to display and perform
118 calculations for a representative subset of all available datasets. We consider these to be
119 the “state of the art” datasets of their type, based on our collective expert judgment.

120

121 In selecting datasets for use in this report, we limit ourselves to those products that are
122 being actively updated and for which temporal homogeneity is an explicit goal in the
123 construction, as these are important considerations for their use in climate change
124 assessment. By way of a literature review, we discuss additional datasets not used in this

125 report. Since some datasets are derivatives of earlier ones, we mention this where
126 appropriate. One should not misconstrue the exclusion of a dataset from this report as an
127 invalidation of that product. Indeed, some of the excluded datasets have proved to be
128 quite valuable in the past and will continue to be so into the future.

129

130 Most of the analyses that we have performed involve data that were averaged over a large
131 region, such as the entire globe or the tropics. The spatial averaging process is
132 complicated by the fact that the locations (gridpoints or stations) at which data values are
133 available can vary fundamentally by data type (see Chapter 2 for details) and, even for a
134 given type, between data production teams. In an effort towards more consistency, the
135 spatial averages we use represent the weighted average of zonal averages¹ (i.e., averages
136 around an entire latitude line or zone), where the weights are the cosine of latitude². This
137 insures that the different latitude zones are given equal treatment across all datasets.

138

139 This chapter begins with a discussion of the four different data types, introducing some
140 temperature datasets for each type, and then discussing their time histories averaged over
141 the globe. Later we present more detail, concentrating on the analysis of temperature
142 trends for two eras: (1) the period since the widespread availability of radiosonde
143 observations in 1958, and (2) since the introduction of satellite data in 1979. We compare

¹ The zonal averages, which were supplied to us by each dataset production team, differ among datasets. We allowed each team to use their judgment as how to best produce these from the available gridpoint or station values in each latitude zone

² The cosine factor weights lower latitudes more than higher ones, to account for the fact that lines of longitude converge towards the poles. As a result, a zonal band in lower latitudes encompasses more area than a comparably sized band (in terms of latitude/longitude dimensions) in higher latitudes.

144 overall temperature trends from different measurement systems and then go into more
145 detail on trend variations in the horizontal and vertical.

146

147 **2. Surface Temperatures**

148

149 **2.1 Land-based temperature data**

150 Over land, temperature data come from fixed weather observing stations with
151 thermometers housed in special instrument shelters. Records of temperature from many
152 thousands of such stations exist. Chapter 2 outlines the difficulties in developing reliable
153 surface temperature datasets. One concern is the variety of changes that may affect
154 temperature measurements at an individual station. For example, the thermometer or
155 instrument shelter might change, the time of day when the thermometers are read might
156 change, or the station might move. These problems are addressed through a variety of
157 procedures (see Peterson et al., 1998 for a review) that are generally quite successful at
158 removing the effects of such changes at individual stations (e.g., Vose et al., 2003 and
159 Peterson, 2005) whether the changes are documented in the metadata or detected via
160 statistical analysis using data from neighboring stations as well (Aguilar et al., 2003).
161 Subtle or widespread impacts that might be expected from urbanization or the growth of
162 trees around observing sites might still contaminate a dataset. These problems are
163 addressed either actively in the data processing stage (e.g., Hansen et al., 2001) or
164 through dataset evaluation to ensure as much as possible that the data are not biased³

³ Changes in regional land use such as deforestation, afforestation, agricultural practices, and other regional changes in land use are not addressed in the development of these datasets. While modeling studies have

165 (e.g., Jones et al., 1990; Peterson, 2003; Parker, 2004; Peterson and Owen, 2005).

166

167 **2.2 Marine temperature data**

168 Data over the ocean come from moored buoys, drifting buoys, and volunteer observing
169 ships. Historically, ships have provided most of the data, but in recent years an increasing
170 number of buoys have been used, placed primarily in data-sparse areas away from
171 shipping lanes. In addition, satellite data are often used after 1981. Many of the ships and
172 buoys take both air temperature observations and sea surface temperature (SST)
173 observations. Night marine air temperature (NMAT) observations have been used to
174 avoid the problem that the Sun's heating of the ship's deck can make the thermometer
175 reading greater than the actual air temperature. Where there are dense observations of
176 NMAT and SST, over the long term they track each other very well. However, since
177 marine observations in an area may only be taken a few times per month, SST has the
178 advantage over air temperature in that water temperature changes much more slowly than
179 that of air. Also, there are twice as many SST observations as NMAT from the same
180 platforms as SSTs are taken during both the day and night and SST data are
181 supplemented in data sparse areas by drifting buoys which do not take air temperature
182 measurements. Accordingly, only having a few SST observations in a grid box for a
183 month can still provide an accurate measure of the average temperature of the month.

184

suggested over decades to centuries these effects can be important on regional space scales (Oleson et al., 2004), we consider these effects to be those of an external forcing to the climate system and are treated as such by many groups in the simulation of climate using the models described in Chapter 5. To the extent that these effects could be large enough to have a measurable influence on global temperature, these changes will be detected by the land-based surface network.

185 2.3 Global surface temperature data

186 Currently, there are three main groups creating global analyses of surface temperature
187 (see Table 3.1), differing in the choice of available data that are utilized as well as the
188 manner in which these data are synthesized. Since the network of surface stations
189 changes over time, it is necessary to assess how well the available observations monitor
190 global or regional temperature. There are three ways in which to make such assessments
191 (Jones, 1995). The first is using “frozen grids” where analysis using only those grid
192 boxes with data present in the sparsest years is used to compare to the full dataset results
193 from other years (e.g., Parker et al., 1994). The results generally indicate very small
194 errors on multi-annual timescales (Jones, 1995). The second technique is subsampling a
195 spatially complete field, such as model output, only where in situ observations are
196 available. Again the errors are small (e.g., the standard errors are less than 0.06°C for the
197 observing period 1880 to 1990; Peterson et al., 1998b). The third technique is comparing
198 optimum averaging, which fills in the spatial field using covariance matrices,
199 eigenfunctions or structure functions, with other analyses. Again, very small differences
200 are found (Smith et al., 2005). The fidelity of the surface temperature record is further
201 supported by work such as Peterson et al. (1999) which found that a rural subset of global
202 land stations had almost the same global trend as the full network and Parker (2004) that
203 found no signs of urban warming over the period covered by this report.

204

205

206

207

208

209 **Table 3.1: Temperature datasets utilized in this report.**

210

211 *Our Name* *Name given by Producers* *Producers*212 *Web Page*213 **-- Surface --**

214

215 NOAA ER-GHCN-ICODS NOAA's National Climatic Data Center
(NCDC)216 <http://www.ncdc.noaa.gov/oa/climate/monitoring/gcag/gcag.html>

217

218 GISS Land+Ocean Temperature NASA's Goddard Institute for Space Studies
(GISS)219 <http://www.giss.nasa.gov/data/update/gistemp/graphs/>

220

221 HadCRUT2v HadCRUT2v Climatic Research Unit of the University of East
222 Anglia and the Hadley Centre of the UK Met
223 Office.224 <http://www.cru.uea.ac.uk/cru/data/temperature>

225

226 **-- Radiosonde --**

227

228 RATPAC RATPAC NOAA's: Air Resources Laboratory (ARL),
229 Geophysical Fluid Dynamics Laboratory
230 (GFDL), and National Climatic Data Center
231 (NCDC)232 <http://www.ncdc.noaa.gov/>

233

234

235 HadAT2 HadAT2 Hadley Centre, UK

236 <http://www.hadobs.org/>

237

238

239 **-- Satellite --**

240

241 ***Temperature of the Lower Troposphere***242 T_{2LT-A} TLT University of Alabama in Huntsville (UAH)243 <http://vortex.nsstc.uah.edu/data/msu/t2lt>

244

245 T_{2LT-R} TLT Remote Sensing System, Inc. (RSS)246 http://www.remss.com/msu/msu_data_description.html

247

248

250 ***Temperature of the Middle Troposphere***251 T_{2-A} TMT University of Alabama in Huntsville (UAH)252 <http://vortex.nsstc.uah.edu/data/msu/t2>

253

254 T_{2-R} TMT Remote Sensing System, Inc. (RSS)255 http://www.remss.com/msu/msu_data_description.html

256

257 T_{2-M} Channel 2 University of Maryland and NOAA/NESDIS
258 (U.Md.)

259

260 ***Temperature of the Middle Troposphere minus Stratospheric Influences***261 T*_G (global) T₍₈₅₀₋₃₀₀₎ University of Washington, Seattle (UW) and262 T*_T (tropics) NOAA's Air Resources Laboratory (ARL)

263

264 ***Temperature of the Lower Stratosphere***265 T_{4-A} TLS University of Alabama in Huntsville (UAH)266 <http://vortex.nsstc.uah.edu/data/msu/t4>

267

268 T_{4-R} TLS Remote Sensing System, Inc. (RSS)269 http://www.remss.com/msu/msu_data_description.html

270

271 **-- Reanalysis --**

272

273 US NCEP50 National Center for Environmental Prediction,
274 NOAA and the National Center for Atmospheric
275 Research276 <http://wesley.ncep.noaa.gov/reanalysis.html>

277

278 European ERA40 European Center for Medium Range
279 Forecasting280 <http://www.ecmwf.int/research/era>

281

282

283 **2.3.1 NOAA**

284 The National Oceanic and Atmospheric Administration (NOAA) National Climatic Data

285 Center (NCDC) integrated land and ocean dataset (see Table 3.1) is derived from in situ

286 data. The SSTs come from the International Comprehensive Ocean-Atmosphere Data Set

287 (ICOADS) SST observations release 2 (Slutz et al., 1985; Woodruff et al., 1998; Diaz et

288 al., 2002). Those that pass quality control tests are averaged into monthly 2° grid boxes

289 (Smith and Reynolds, 2003). The land surface air temperature data come from the Global
290 Historical Climatology Network (GHCN) (Peterson and Vose, 1997) and are averaged
291 into 5° grid boxes. A reconstruction approach is used to create complete global coverage
292 by combining together the faster and slower time-varying components of temperature
293 (van den Dool et al., 2000; Smith and Reynolds, 2005).

294

295 **2.3.2 NASA (GISS)**

296 The NASA Goddard Institute for Space Studies (GISS) produces a global air temperature
297 analysis (see Table 3.1) known as GISTEMP using land surface temperature data
298 primarily from GHCN and the U.S. Historical Climatology Network (USHCN;
299 Easterling, et al., 1996). The NASA team modifies the GHCN/USHCN data by
300 combining at each location the time records of the various sources and adjusting the non-
301 rural stations in such a way that their long-term trends are consistent with those from
302 neighboring rural stations (Hansen et al., 2001). These meteorological station
303 measurements over land are combined with in situ sea surface temperatures and Infrared
304 Radiation (IR) satellite measurements for 1982 to the present (Reynolds and Smith, 1994;
305 Smith et al., 1996) to produce a global temperature index (Hansen et al., 1996).

306

307 **2.3.3 UK (HadCRUT2v)**

308 The UK global land and ocean dataset (HadCRUT2v, see Table 3.1) is produced as a
309 joint effort by the Climatic Research Unit of the University of East Anglia and the
310 Hadley Centre of the UK Meteorological (Met) Office. The land surface air temperature

311 data are from Jones and Moberg (2003) of the Climatic Research Unit. The global SST
312 fields are produced by the Hadley Centre using a blend of COADS and Met Office data
313 bank in situ observations (Rayner, et al., 2003). The integrated dataset is known as
314 HadCRUT2v (Jones and Moberg, 2003)⁴. The temperature anomalies were calculated on
315 a 5°x5° grid box basis. Within each grid box, the temporal variability of the observations
316 has been adjusted to account for the effect of changing the number of stations or SST
317 observations in individual grid-box temperature time series (Jones et al., 1997, 2001).
318 There is no reconstruction of data gaps because of the problems of introducing biased
319 interpolated values.

320

321 **2.3.4 Synopsis of surface datasets**

322 Since the three chosen datasets utilize many of the same raw observations, there is a
323 degree of interdependence. Nevertheless, there are some differences among them as to
324 which observing sites are utilized. An important advantage of surface data is the fact that
325 at any given time there are thousands of thermometers in use that contribute to a global or
326 other large-scale average. Besides the tendency to cancel random errors, the large number
327 of stations also greatly facilitates temporal homogenization since a given station may
328 have several “near-neighbors” for “buddy-checks.” While there are fundamental
329 differences in the methodology used to create the surface datasets, the differing
330 techniques with the same data produce almost the same results (Wuertz et al., 2005). The

⁴ Although global and hemispheric temperature time series created using a technique known as optimal averaging (Folland et al., 2001a; Parker et al., 2004), which provides estimates of uncertainty in the time series, including the effects of data gaps and uncertainties related to bias corrections or uncorrected biases, are available, we have used the data in their more basic form, for consistency with the other datasets.

331 small differences in deductions about climate change derived from the surface datasets
332 are likely to be due mostly to differences in construction methodology and global
333 averaging procedures.

334

335 **2.4 Global surface temperature variations and differences between the datasets**

336 Examination of the three global temperature anomaly time series (T_{sfc}) from 1958 to the
337 present shown in Figure 3.1 reveals that the three time series have a very high level of
338 agreement. They all show some temperature decrease from 1958 to around 1976,
339 followed by a strong increase. That most of the temperature change occurs after the mid
340 1970s has been previously documented (Karl et al., 2000; Folland et al., 2001b; Seidel
341 and Lanzante, 2004). The variability of the time series is quite similar as are their trends.
342 The signature of the El Niño-Southern Oscillation (ENSO), whose origin is in the tropics,
343 is responsible for many of the prominent short-term (several year) up and down swings of
344 temperature (Trenberth et al., 2002). The strong El Niño of 1997-98 stands out as an
345 especially large warm event within an overall upward trend.

346
347 Figure 3.1 - Time series of globally averaged surface temperature (T_s) for NOAA (violet), GISS (black),
348 and HadCRUT2v (green) datasets. All time series are 7-month running averages (used as a smoother) of
349 original monthly data, which were expressed as a departure ($^{\circ}\text{C}$) from the 1979-97 average.
350

351

352 **3 RADIOSONDE TEMPERATURES**

353

354 **3.1 Balloon-borne temperature data**

355 Since the beginning of the radiosonde era, several thousand sites have been used to
356 launch balloons. However, many of these were in operation for only short periods of
357 time. One approach has been to use a fixed station network consisting of a smaller
358 number of stations having long periods of record. A complimentary approach is to grid
359 the data, using many more stations, allowing stations to join or drop out of the network
360 over the course of time. Since each approach has advantages and disadvantages, we
361 utilize both. A further complication is that changes over time in instruments and
362 recording practices have imparted artificial changes onto the temperature records. Some
363 groups have developed methods that try to remove these artificial effects as much as

364 possible. We employ two radiosonde datasets (see Table 3.1), one station-based and one
365 gridded. Both datasets have been constructed using homogeneity adjustments in an
366 attempt to minimize the effects of artificial changes.

367

368 **3.2 Radiosonde temperature datasets**

369

370 **3.2.1 NOAA (RATPAC)**

371 For several decades the 63 station dataset of Angell (Angell and Korshover, 1975) was
372 the most widely used station-based radiosonde temperature dataset for climate
373 monitoring. Recently, due to concerns regarding the effects of inhomogeneities, that
374 network shrank to 54 stations (Angell, 2003). To better address these concerns, LKS
375 (Lanzante, Klein, Seidel) (Lanzante et al., 2003a,b) built on the work of Angell by
376 applying homogeneity adjustment to the time series from many of his stations, as well as
377 several dozen additional stations, to create better regional representation via a network of
378 87 stations. However, because of the labor-intensive nature of the homogenization
379 process on these 87 stations, extension of the LKS dataset beyond 1997 is impractical.
380 Instead, the adjusted LKS dataset is being used as the basis for a new product (see Table
381 3.1), Radiosonde Air Temperature Products for Assessing Climate (RATPAC), that will
382 be updated regularly (Free et al., 2003; Free et al., 2005). A NOAA group (a
383 collaboration between the Air Resources Laboratory, the Geophysical Fluid Dynamics
384 Laboratory, and NCDC) is responsible for the creation of RATPAC.

385

386 The RATPAC product consists of two parts: RATPAC-A and RATPAC-B⁵, both of
387 which use the adjusted LKS data, supplemented by an extension up to present using data
388 from the Integrated Global Radiosonde Archive (IGRA). The IGRA data used in
389 RATPAC are based on individual soundings that have been quality controlled and then
390 averaged into monthly station data (Durre, 2005). In this report we use RATPAC-B.
391 Generally speaking, based on data averaged over large regions such as the globe or
392 tropics, trends from RATPAC-A and RATPAC-B are closer to one another than they are
393 to the unadjusted (IGRA) data (Free et al., 2005).

394 **3.2.2 UK (HadAT2)**

395 For several decades the Oort (1983) product was the most widely used gridded
396 radiosonde dataset. With the retirement of Abraham Oort, and cessation of his product,
397 the dataset produced at the Hadley Centre, UK Met Office, HadRT (Parker et al., 1997)
398 became the most widely used gridded product. Because of concern about the effects of
399 artificial changes, this product incorporated homogeneity adjustments, although they
400 were somewhat limited⁶. As a successor to HadRT, the Hadley Centre has created a new
401 product (HadAT2, see Table 3.1) that uses all available digital radiosonde data for a
402 larger network of almost 700 stations having relatively long records⁷. Identification and

⁵ RATPAC-A uses the adjusted LKS data up through 1997 and provides an extension beyond that using a different technique to reduce the impact of inhomogeneities (Peterson et al., 1998). However, the RATPAC-A methodology can only be used to derive homogenized temperature averaged over many stations, and thus cannot be used to homogenize temperature time series at individual stations. RATPAC-B consists of the LKS adjusted station time series that have been extended beyond 1997 by appending (unadjusted) IGRA data.

⁶ Adjustments were made to upper levels only (300 hPa and above), and since they were based on satellite data, only since 1979.

⁷ High quality small station subsets, such as Lanzante et al. (2003a) and the Global Climate Observing System Upper Air Network, were used as a skeletal network from

403 adjustment of inhomogeneities was accomplished by way of comparison of neighboring
404 stations.

405

406 **3.2.3 Synopsis of radiosonde datasets**

407 The two chosen datasets differ fundamentally in their selection of stations in that the
408 NOAA dataset uses a relatively small number of highly scrutinized stations, while the
409 UK dataset uses a considerably larger number of stations. Compared to the surface, far
410 fewer thermometers are in use at any given time (hundreds or less) so there is less
411 opportunity for random errors to cancel, but more importantly, there are far fewer
412 suitable “neighbors” to aid in temporal homogenization. While both products incorporate
413 a common building-block dataset (Lanzante et al., 2003a), their methods of construction
414 differ considerably. Any differences in deductions about climate change derived from
415 them could be attributed to both the differing raw inputs as well as differing construction
416 methodologies. Concerns about poor temporal homogeneity are much greater than for
417 surface data.

418

419 **3.3 Global radiosonde temperature variations and differences between the datasets**

420 **3.3.1 Troposphere**

421 Figure 3.2a displays $T_{(850-300)}$ time series for the RATPAC and HadAT2 radiosonde

which to define a set of adequately similar station series used in homogenization. The dataset is designed to impart consistency in both space and time and, by using radiosonde neighbors rather than satellites or reanalyses, minimizes the chances of introducing spurious changes related to the introduction of satellite data and their subsequent platform changes (Thorne et al., 2005).

422 datasets. Several noteworthy features are common to both. First, just as for the surface,
423 ENSO signatures are clearly evident. Second, there is an apparent step-like rise of
424 temperature around 1976-77 associated with the well-documented climate regime shift
425 (Trenberth, 1990). Third, there is a long-term rise in temperatures, although a
426 considerable amount of it may be due to the step-like change (Seidel and Lanzante,
427 2004). To a first approximation, both datasets display these features similarly and there is
428 very little systematic difference between the two. Although a major component of the
429 RATPAC product is used in the construction of the HadAT2 dataset, it should be kept in
430 mind that the former utilizes a much smaller network of stations, although the length of
431 the station records tends to be relatively long. If the good agreement is not fortuitous, this
432 suggests that the reduced RATPAC station network provides representative spatial
433 sampling⁸.

⁸ This result is consistent with the relatively large spatial scales represented by a single radiosonde station at this level on an annual time scale demonstrated by Wallis (1998) and Thorne et al. (2005).

434
435 Figure 3.2a – Bottom: Time series of globally averaged tropospheric temperature ($T_{(850-300)}$) for RATPAC
436 (violet) and HadAT2 (green) radiosonde datasets. All time series are 7-month running averages (used as a
437 smoother) of original monthly data, which were expressed as a departure ($^{\circ}\text{C}$) from the 1979-97 average.
438

439 Figure 3.2b – Top: Time series of globally averaged stratospheric temperature ($T_{(100-50)}$) for RATPAC
440 (violet) and HadAT2 (green) radiosonde datasets. All time series are 7-month running averages (used as a
441 smoother) of original monthly data, which were expressed as a departure ($^{\circ}\text{C}$) from the 1979-97 average.
442

443

444 **3.3.2 Lower Stratosphere**

445 Figure 3.2b displays global temperature anomaly time series of $T_{(100-50)}$ from the
446 RATPAC and HadAT2 radiosonde datasets. Several noteworthy features are common to
447 both datasets. First is the prominent signature of three climatically important volcanic

448 eruptions: Agung (March 1963), El Chichon (April 1982), and Pinatubo (June 1991).
449 Temperatures rise rapidly as volcanic aerosols are injected into the stratosphere and
450 remain elevated for about 2-3 years before diminishing. There is some ambiguity as to
451 whether the temperatures return to their earlier values or whether they experience step-
452 like falls in the post-volcanic period for the latter two volcanoes, particularly Pinatubo
453 (Pawson et al., 1998; Lanzante et al., 2003a; Seidel and Lanzante, 2004). Second, there
454 are small amplitude variations associated with the tropical quasi-biennial oscillation
455 (QBO) with a period of ~ 2-3 years (Seidel et al., 2004). Third, there is a downward
456 trend, although there is some doubt as to whether the temperature decrease is best
457 described by a linear trend over the period of record. For one thing, the temperature series
458 prior to about 1980 exhibits little or no decrease in temperature. After that, the
459 aforementioned step-like drops represent a viable alternative to a linear decrease (Seidel
460 and Lanzante, 2004).

461

462 In spite of similarities among datasets, closer examination reveals some important
463 differences. There is a rather large difference between RATPAC and HadAT2 time series
464 for the peak volcanic warming associated with Agung in 1963. This may be a reflection
465 of differences in spatial sampling because the horizontal pattern of the response is not
466 uniform (Free and Angell, 2002). More noteworthy for estimates of climate change are
467 some subtle systematic differences between the two datasets that vary over time. A closer
468 examination reveals that the RATPAC product tends to have higher temperatures than the
469 HadAT2 product from approximately 1963-85, with the RATPAC product having lower

470 values before and after this time period⁹. As we will see later, this yields a slightly greater
471 decreasing trend for the RATPAC product. Poorer agreement between the RATPAC and
472 HadAT2 products in the stratosphere compared to the troposphere is not unexpected
473 because of the fact that artificial jumps in temperature induced by changes in radiosonde
474 instruments and measurement systems tend to increase in magnitude from the near-
475 surface upwards (Lanzante et al., 2003b). More details on this issue are given in Chapter
476 4, Section 2.1.

477

478 **4 SATELLITE-DERIVED TEMPERATURES**

479

480 **4.1 Microwave satellite data**

481 Three groups, employing different methodologies, have developed satellite Microwave
482 Sounding Unit (MSU) climate datasets (see Table 3.1). We do not present results from a
483 fourth group (Prabhakara et al., 2000), which developed yet another methodology, since
484 they are not continuing to work on MSU climate analyses and are not updating their time
485 series. One of the main issues that is addressed differently by the groups is the inter-
486 calibration between the series of satellites, and is discussed in Chapters 2 and 4.

487 **4.2 Microwave Satellite Datasets**

488

⁹ It is worth noting that prominent artificial step-like drops, many of which were associated with the adoption of a particular type of radiosonde (Vaisala), were found in stratospheric temperatures at Australian and western tropical Pacific stations in the mid to late 1980s by Parker et al. (1997), Stendel et al. (2000), and Lanzante et al. (2003a). Differences in consequent homogeneity adjustments around this time could potentially explain a major part of the difference between the NOAA and UK products, although this has not been demonstrated.

489 **4.2.1 University of Alabama In Huntsville (UAH)**

490 The first group to produce MSU climate products, by adjusting for the differences
491 between satellites and the effects of changing orbits (diurnal drift), was UAH (A). Their
492 approach (Christy et al., 2000; Christy et al., 2003) uses both an offset adjustment to
493 allow for the systematic average differences between satellites and a non-linear hot target
494 temperature¹⁰ calibration to create a homogeneous series. The UAH dataset has products
495 corresponding to three temperature measures: T_{2LT} , T_2 , and T_4 (see Chapter 2 for
496 definitions of these measures). In this report we use the most up to date versions available
497 to us at the time, which is version 5.1 of the UAH dataset for T_2 and T_4 , and version 5.2
498 for T_{2LT} ¹¹.

499

500 **4.2.2 Remote Sensing Systems (RSS)**

501 After carefully studying the methodology of the UAH team, another group, RSS (R)
502 created their own datasets for T_2 and T_4 using the same input data but with modifications
503 to the adjustment procedure (Mears et al., 2003), two of which are particularly
504 noteworthy: (1) the method of inter-calibration from one satellite to the next and (2) the
505 computation of the needed correction for the daily cycle of temperature. While the second
506 modification has little effect on the overall global trend differences between the two
507 teams, the first is quite important in this regard. Recently the RSS team has created their
508 own version of T_{2LT} (Mears and Wentz, 2005) and in doing so discovered a

¹⁰ In fact, two targets are used, both with temperatures that are presumed to be well known. These are *cold* space, pointing away from the Earth, Moon, or Sun, and an onboard *hot* target.

¹¹ The version number for T_{2LT} differs from that for T_2 and T_4 because an error, which was found to affect the former (and was subsequently corrected), does not affect the latter two measures.

509 methodological error in the corresponding temperature measure of UAH. The UAH T_{2LT}
510 product used in this report is based on their corrected method. In this report we use
511 version 2.1 of the RSS data.

512

513 **4.2.3 University of Maryland (U.Md.)**

514 A very different approach (Vinnikov et al., 2004) was developed by a team involving
515 collaborators from the University of Maryland and the NOAA National Environmental
516 Satellite, Data, and Information Service (NESDIS) and was used to estimate globally
517 averaged temperature trends (Vinnikov and Grody, 2003). After further study, they
518 developed yet another new method (Grody et al., 2004). As done by the other two groups,
519 the U.Md. (M) team's methodology also recalibrates the instruments based on
520 overlapping data between the satellites. However, the manner in which they perform this
521 recalibration differs. Also, they do not adjust for diurnal drift directly, but average the
522 data from ascending and descending orbits. In their second approach, they substantially
523 altered the manner in which target temperatures are used in their recalibration to a
524 scheme more consistent with that of the other two groups (UAH and RSS). The effect of
525 their revision was to reduce the global temperature trends derived from their data from
526 0.22-0.26 to 0.17 °C/decade. Very recently they have revised their method to produce a
527 third version of their dataset, which we use in this report, whose trends differ only
528 slightly with those from the second version. In this most recent version they apply the
529 nonlinear adjustment of Grody et al. (2004) and estimate the diurnal cycle as described in
530 Vinnikov et al. (2005). The U.Md. group produces only a measure of T_2 , hence there is
531 no stratospheric product (T_4) or one corresponding to the lower troposphere (T_{2LT}).

532

533 4.3 Synopsis of satellite datasets

534 The relationship among satellite datasets is fundamentally different from that for surface
535 or radiosonde products. For satellites, different datasets use virtually the same raw inputs
536 so that any differences in derived measures are due to construction methodology. The
537 excellent coverage provided by the orbiting sensors, more than half the Earth's surface
538 daily, is a major advantage over in situ observations. The disadvantage is that while in
539 situ observations rely on data from many hundreds or thousands of individual
540 thermometers every day, providing a beneficial redundancy, the satellite data typically
541 come from only one or two instruments at a given time. Therefore, any problem
542 impacting the data from a single satellite can adversely impact the entire climate record.
543 The lack of redundancy, compounded by occasional premature satellite failure that limits
544 the time of overlapping measurements from successive satellites, elevates the issue of
545 temporal homogeneity to the overwhelming explanation for any differences in deductions
546 about climate change derived from the three datasets.

547

548 4.4 Global satellite temperature variations and differences between the datasets

549

550 4.4.1 Temperature of the Troposphere

551 Two groups (UAH and RSS) produce lower tropospheric temperature datasets, T_{2LT} (see
552 Chapter 2 for definition of this and related temperature measures) directly from satellite
553 measurements. Their time series are shown in Figure 3.3a along with an equivalent

554 measure constructed from the HadAT2 radiosonde dataset (see Box 2.2 for an
555 explanation as to how these equivalent measures were generated). The three temperature
556 series have quite similar behavior, with ENSO-related variations accounting for much of
557 the up and down meanderings, for example the historically prominent 1997-1998 El
558 Niño. But over the full period of record, the amount of increase indicated by the datasets
559 varies considerably. A closer look reveals that as time goes on, the RSS product indicates
560 a noticeably greater increase of temperature than the other two. For comparison purposes,
561 in Figure 3.3b we show an alternate measure of lower tropospheric temperatures, T^*_G ,
562 derived from products produced by the same three groups. From comparison of Figures
563 3.3a and 3.3b we see that both measures of lower tropospheric temperature agree
564 remarkably well, even with regard to the more subtle differences relating to the longer-
565 term changes. We will return to the issue of agreement between T_{2LT} and T^*_G later when
566 we discuss trends (section 6).

567

568 Figure 3.3a– Bottom: Time series of globally averaged lower tropospheric temperature (T_{2LT}) as follows:
569 UAH (blue) and RSS (red) satellite datasets, and HadAT2 (green) radiosonde data. All time series are 7-
570 month running averages (used as a smoother) of original monthly data, which were expressed as a
571 departure ($^{\circ}\text{C}$) from the 1979-97 average.
572

573 Figure 3.3b– Third: Time series of globally averaged middle tropospheric temperature (T^*_G) as follows:
574 UAH (blue) and RSS (red) satellite datasets, and HadAT2 (green) radiosonde data. All time series are 7-
575 month running averages (used as a smoother) of original monthly data, which were expressed as a
576 departure ($^{\circ}\text{C}$) from the 1979-97 average.
577

578 Figure 3.3c – Second: Time series of globally averaged upper middle tropospheric temperature (T_2) as

579 follows: UAH) (blue), RSS (red), and U.Md. (black) satellite datasets, and HadAT2 (green) radiosonde
580 data. All time series are 7-month running averages (used as a smoother) of original monthly data, which
581 were expressed as a departure ($^{\circ}\text{C}$) from the 1979-97 average.

582
583 Figure 3.3d – Top: Time series of globally averaged lower stratospheric temperature (T_4) as follows: UAH
584 (blue) and RSS (red) satellite datasets, and HadAT2 (green) radiosonde data. All time series are 7-month
585 running averages (used as a smoother) of original monthly data, which were expressed as a departure ($^{\circ}\text{C}$)
586 from the 1979-97 average.

587

588 Time series corresponding to the temperature of the upper middle troposphere (T_2) are
589 shown in Figures 3.3c. The products represented in this figure are the same as for the
590 lower troposphere, except that an additional product, that from the U.Md. group is
591 available. Again, all of the time series have similar behavior with regard to the year to
592 year variations. However, closer examination shows that two of the products (U.Md. and
593 RSS satellite data) indicate considerable temperature increase over the period of record,
594 whereas the other two (UAH satellite and HadAT2 radiosonde) indicate slight warming
595 only. A more detailed discussion of the differences between the various products can be
596 found in Chapter 4.

597

598 We note that all of the curves for the various tropospheric temperature series (Figures
599 3.3a-c) exhibit remarkably similar shape over the period of record. For the common time
600 period, the satellite measures are similar to the tropospheric layer-averages computed
601 from radiosonde data. The important differences between the various series are with
602 regard to the more subtle long-term evolution over time, which manifests itself as
603 differences in linear trend, discussed later in more detail.

604

605

606 **4.4.2 Temperature of the Lower Stratosphere**

607 Figure 3.3d shows the temperature of the lower stratosphere (T_4); note that there is no
608 product from the U.Md. team for this layer. The dominant features for this layer are the
609 major volcanic eruptions: El Chichon in 1982 and Pinatubo in 1991. As discussed above,
610 the volcanic aerosols tend to warm the stratosphere for about 2-3 years before
611 diminishing. In contrast, ENSO events have little influence on the stratospheric
612 temperature. Both products show that the stratospheric temperature has decreased
613 considerably since 1979, as compared to the lesser amount of increase that is seen in the
614 troposphere. The T_{4-R} product shows somewhat less overall decrease than the T_{4-A}
615 product, in large part as a result of the fact that the former increases relative to the latter
616 from about 1992-94. As was the case for the troposphere, the radiosonde series show a
617 greater decrease than the satellite data. Again, the satellite and radiosonde series for the
618 lower-stratosphere exhibit the same general behavior over time.

619

620

621 **5 REANALYSIS TEMPERATURE “DATA”**

622

623 A number of agencies from around the world have produced reanalyses based on
624 different schemes for different time periods. We focus on two of the most widely
625 referenced, which cover a longer time period than the others (see Table 3.1). The U.S.
626 reanalysis represents a collaborative effort between NOAA’s National Center for
627 Environmental Prediction (NCEP) and the National Center for Atmospheric Research
628 (NCAR). For U.S. reanalysis, gridded air temperatures at the surface and aloft are

629 available from 1958 to present. Using a completely different system, the European Center
630 for Medium-Range Weather Forecasts (ECMWF) has produced similar gridded data from
631 September 1957 to August 2002 (Europe). Reanalyses are “hybrid products,” utilizing
632 raw input data of many types, as well as complex mathematical models to combine these
633 data. For more detailed information on the reanalyses, see Chapter 2. As the reanalysis
634 output does not represent a different observing platform, a separate assessment of
635 reanalysis data will not be made.

636

637 **6. COMPARISONS BETWEEN DIFFERENT LAYERS AND OBSERVING** 638 **PLATFORMS**

639

640 **6.1 During the radiosonde era, 1958 to the present**

641 **6.1.1 Global**

642 As shown in earlier sections, globally averaged temperature time series indicate
643 increasing temperature at the surface and in the troposphere with decreases in the
644 stratosphere over the course of the last several decades. It is desirable to derive some
645 estimates of the magnitude of the rate of these changes. The widely-used, least-squares,
646 linear trend technique is adopted for this purpose with the explicit caveat that long-term
647 changes in temperature are not necessarily linear, as there may be departures in the form
648 of periods of enhanced or diminished change, either linear or nonlinear, as well as abrupt,

649 step-like changes¹². While it has been shown that such constructs are plausible, it is
650 nevertheless difficult to prove that they provide a better fit to the data, over the time
651 periods addressed in this report, than the simple linear model (Seidel and Lanzante,
652 2004). Additional discussion on this topic is given in the Statistical Appendix.

653

654 Trends computed for the radiosonde era are given in Table 3.2 for the surface as well as
655 various tropospheric and stratospheric layer averages¹³. The surface products are quite
656 consistent with one another, as are the radiosonde products in the troposphere. In the
657 stratosphere, the radiosonde products differ somewhat, although there is an inconsistent
658 relationship involving the two stratospheric measures ($T_{(100-50)}$ and T_4) regarding which
659 product indicates a greater decrease in temperature¹⁴. The reanalysis products, which are
660 “hybrid-measures,” agree better with the “purer” surface and radiosonde measures at and
661 near the surface. Agreement degrades with increasing altitude such that the reanalyses
662 indicate more tropospheric temperature increase and considerably less stratospheric
663 decrease than do the radiosonde products. The disparity between the reanalyses and other
664 products is not surprising given the suspect temporal homogeneity of the reanalyses (see
665 Chapter 2, Section 1c).

¹² For example, the tropospheric linear trends in the periods 1958-1979 and 1979-2003 were shown to be much less than the trend for the full period (1958-2003), based on one particular radiosonde dataset (Thorne et al., 2005), due to the abrupt rise in temperature in the mid 1970s.

¹³ Note that it is instructive to examine the behavior of radiosonde and reanalysis temperatures averaged in such a way as to correspond to the satellite layers (T_{2LT} , T^*_G , T_2 , and T_4) even though there are no comparable satellite measures prior to 1979.

¹⁴ The reason for this inconsistency is that the HadAT2 product records data at fewer vertical levels than the RATPAC product, so the comparison is not one-to-one.

666 Table 3.2 - Global temperature trends in °C per decade from 1958 through 2004 (except for European
667 which terminates September 2001) calculated for the surface or atmospheric layers by data source. The
668 trend is shown for each, with the approximate 95% confidence interval (2 sigma) below in parentheses. The
669 levels/layers, from left to right, go from the lowest to the highest in the atmosphere. Bold values are
670 estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the
671 lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used
672 to assess significance (see Appendix for discussion of confidence intervals and significance testing).

673

674

675

676 Perhaps the most important result shown in Table 3.2 is that both the radiosonde and
677 reanalysis trends indicate that the tropospheric temperature has increased as fast as or
678 faster than the surface over the period 1958 to present. For a given dataset, the 3
679 measures (T_{2LT} , $T_{(850-300)}$, and T^*_G) always indicate more increase in the troposphere than
680 at the surface, although this is usually not true when the T_2 measure is considered. The
681 reason for the inconsistency involving T_2 is because of contributions to the layer that it
682 measures from stratospheric cooling, an effect first recognized by Spencer and Christy
683 (1992) (see discussion of this issue in Chapters 2 and 4). The development of T^*_G as a
684 global measure, and its counterpart, T^*_T for the tropics (Fu et al., 2004; Fu and Johanson,
685 2005; Johanson and Fu, 2005) was an attempt to remove the confounding effects of the
686 stratosphere using a statistical approach (see Chapter 2).

687 **6.1.2 Land vs. ocean**

688 Most of the land and ocean surface temperature increased during the radiosonde era, with
689 the exception of parts of the North Atlantic Ocean, the North Pacific Ocean, and a few
690 smaller areas. With a few exceptions, such as the west coast of North America, trends in
691 land air temperature in coastal regions are generally consistent with trends in SST over
692 neighboring ocean areas (Houghton et al., 2001). Because bias adjustments are performed
693 separately for land and ocean areas, before merging to create a global product, it is
694 unlikely that the land-ocean consistency is an artifact of the construction methods used in
695 the various surface analyses. However, land air temperatures did increase somewhat more
696 rapidly than SSTs in some regions during the past two decades. Possibly related to this is

697 the fact that since the mid-1970s, El Niño has frequently been in its “warm” phase, which
698 tends to bring higher than normal temperatures to much of North America, among other
699 regions, which have had strong temperature increases over the past few decades (Hurrell,
700 1996). Also, when global temperatures are rising or falling, the global mean land
701 temperature tends to both rise and fall faster than the ocean, which has a tremendous heat
702 storage capacity (Waple and Lawrimore, 2003).

703

704 **6.1.3 Marine air vs. sea surface temperature**

705 In ocean areas, it is natural to consider whether the temperature of the air and that of the
706 ocean surface (SST) increases or decreases at the same rate. Several studies have
707 examined this question. Overall, on seasonal and longer scales, the SST and marine air
708 temperature generally move at about the same rate globally and in many ocean basin
709 scale regions (Bottomley et al., 1990; Parker et al., 1995; Folland et al., 2001b; Rayner et
710 al., 2003). Differences between SST and marine air temperature in some regions were
711 first noted by Christy et al. (1998) and then examined in more detail by Christy et al.
712 (2001). The latter study found that in the tropics, SST increased more than NMAAT from
713 1979 –1999 derived from the Tropical Atmosphere Ocean (TAO) array of tropical buoys
714 and transient marine ship observations. But this difference may be related to changes in
715 surface fluxes associated with ENSO and the interdecadal Pacific oscillation (Folland et
716 al., 2003). Consistent results were found using two datasets, one with more widespread
717 observations from ships, and another, which sampled a more limited number of locations
718 using moored buoys. There were some indications that the accelerated increase of SST

719 compared to air temperature may have been concentrated in two periods: the early 1980s
720 and mid 1990s. So over the satellite era, there are some unexplained differences in these
721 trends that were also noted by Folland et al. (2003) in parts of the tropical south Pacific
722 using the Rayner et al. (2003) NMAT dataset which incorporates new corrections for the
723 effect on NMAT of increasing deck (and hence measurement) heights.

724

725 **6.1.4 Minimum vs. maximum temperatures over land**

726 Daily minimum temperature increased about twice as fast as daily maximum temperature
727 over global land areas during the radiosonde era (Karl et al., 1993; Easterling et al., 1997;
728 Folland et al, 2001b). However, a closer look at recent years has found that during the
729 satellite era, maximum and minimum temperature have been rising at the same rate (Vose
730 et al, 2005). Daily minimum temperature increased in virtually all areas except eastern
731 Canada, Eastern Europe, and other scattered regions often near coasts. Most regions also
732 witnessed an increase in the daily maximum, but over the longer time frame the rate of
733 increase was generally smaller, and decreasing trends were somewhat more common
734 (e.g., in eastern Canada, the southern United States, southern China, eastern Europe, and
735 portions of South America). The causes of this asymmetric warming are still debated, but
736 many of the areas with greater increases of minimum temperatures correspond to those
737 where cloudiness appears to have increased over the period as a whole (Dai et al., 1999;
738 Henderson-Sellers, 1992; Sun and Groisman, 2000; Groisman et al., 2004). This makes
739 physical sense since clouds tend to cool the surface during the day by reflecting incoming
740 solar radiation, and warm the surface at night by absorbing and reradiating infrared
741 radiation back to the surface.

742

743 6.2 During the satellite era, 1979 to the present

744

745 6.2.1 Global

746 A comparable set of global trends for the satellite era is given in Table 3.3. Comparison
747 between Tables 3.2 and 3.3 reveals that some of the relationships between levels and
748 layers, as well as among datasets, are different during the two eras. Comparing satellite
749 era trends with the radiosonde era trends for datasets that have both periods in common, it
750 is clear that the surface temperature increase (see Figure 3.1) has accelerated in recent
751 decades while the tropospheric increase (see Figure 3.2a) has decelerated. Since most of
752 the stratospheric decrease has occurred since 1979 (see Figure 3.2b) the rate of
753 temperature decrease there is close to twice as large as during the full radiosonde era.
754 Thus, care must be taken when interpreting results from only the most recent decades.
755 Agreement among different surface and radiosonde datasets is reasonable and about as
756 good as during the longer radiosonde era. The reanalysis datasets show poorer agreement
757 with surface data and especially with stratospheric radiosonde data for the European
758 product.

759 Table 3.3 - Global temperature trends in °C per decade from 1979 through 2004 (except for European
760 which terminates September 2001) calculated for the surface or atmospheric layers by data source. The
761 trend is shown for each, with the approximate 95% confidence interval (2 sigma) below in parentheses. The
762 levels/layers, from left to right, go from the lowest to the highest in the atmosphere. Bold values are
763 estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test, using the
764 lag-1 autocorrelation to account for the non-independence of residual values about the trend line, was used
765 to assess significance (see Appendix for discussion of confidence intervals and significance testing).

766

771 numerical trend values found in Table 3.3, which are also graphed in Figure 3.4a. While
772 the tropospheric satellite products from the UAH team have trends that are not too
773 dissimilar from the corresponding radiosonde trends, the two other satellite datasets show
774 a considerably greater increase in tropospheric temperature. In the stratosphere, there is a
775 large disagreement between satellite and radiosonde products, with the latter indicating
776 much greater decreases in temperature. Here too, the reanalyses are quite inconsistent,
777 with the European product closer to the satellites and the U.S. product closer to the
778 radiosondes.

779

780 Figure 3.4a (top) – Global temperature trends (°C/decade) for 1979-2004 from Table 3.3 plotted as
781 symbols. See figure legend for definition of symbols. Filled symbols denote trends estimated to be
782 statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1
783 autocorrelation to account for the non-independence of residual values about the trend line, was used to
784 assess significance (see Appendix for discussion of confidence intervals and significance testing).

785

786 Figure 3.4b (bottom) – Tropical (20°N-20°S) temperature trends (°C/decade) for 1979-2004 from Table 3.4
787 plotted as symbols. See figure legend for definition of symbols. Filled symbols denote trends estimated to
788 be statistically significantly different from zero (at the 5% level). A Student's t-test, using the lag-1
789 autocorrelation to account for the non-independence of residual values about the trend line, was used to
790 assess significance (see Appendix for discussion of confidence intervals and significance testing).

791

792 Perhaps the most important issue is the relationship between trends at the surface and in
793 the troposphere. As shown in Table 3.3 and Figure 3.4a, both radiosonde datasets as well
794 as the UAH satellite products indicate that, in contrast to the longer radiosonde era,
795 during the satellite era the temperature of the surface has increased more than that of the
796 troposphere. However, tropospheric trends from the RSS satellite dataset, based on both
797 measures of temperature having little or no stratospheric influence (T_{2LT} and T^*_G) yield
798 an opposing conclusion: the tropospheric temperature has increased as much or more than
799 the surface. For the third satellite dataset, comparisons with surface temperature are
800 complicated by the fact that the U.Md. team produces only T_2 , which is influenced by
801 stratospheric cooling (see Chapter 2). Nevertheless, we can infer that it too suggests more
802 of a tropospheric temperature increase than that at the surface¹⁵.

803

804 Since global change theory suggests more warming of the troposphere than the surface
805 only in the tropics (see Chapter 1), much of the interest in observed trends has been in
806 this region. Therefore, to compliment the global trends (Figure 3.4a and Table 3.3), we
807 present a similar plot of tropical trends in Figure 3.4b (with corresponding trend values in

¹⁵ The difference in trends, T^*_G minus T_2 , for the UAH and RSS datasets is about 0.06 to 0.08 °C/decade. Adding this amount to the U.Md. T_2 trend (0.20 °C/decade) yields an estimate of the U.Md. trend in T^*_G of about 0.26 to 0.28 °C/decade. In this calculation we are assuming that the effects of the stratospheric cooling trend on the U.Md. product are the same as from the UAH and RSS datasets.

808 Table 3.4). Compared to the global trends, the tropical trends show even more spread
809 among datasets, particularly in the lower stratosphere¹⁶. The result of the greater spread is
810 that the range of plausible values for the difference in trends between the surface and
811 troposphere is larger than that for the globe as a whole. Similar to the global case, in the
812 tropics the UAH satellite plus the two radiosonde datasets (RATPAC and HadAT2)
813 suggest more warming at the surface than in the troposphere, while the opposite
814 conclusion is reached based on the other two satellite products (RSS and U.Md.).
815 Resolution of this issue would seem to be of paramount importance in the interpretation
816 of observed climate change central to this synthesis assessment.

817

818

819 Table 3.4 – Tropical (20°N-20°S) temperature trends in °C per decade from 1979 through 2004 (except for
820 European which terminates September 2001) calculated for the surface or atmospheric layers by data
821 source. The trend is shown for each, with the approximate 95% confidence interval (2 sigma) below in
822 parentheses. The levels/layers, from left to right, go from the lowest to the highest in the atmosphere. Bold
823 values are estimated to be statistically significantly different from zero (at the 5% level). A Student's t-test,
824 using the lag-1 autocorrelation to account for the non-independence of residual values about the trend line,
825 was used to assess significance (see Appendix for discussion of confidence intervals and significance
826 testing).

827

¹⁶ The larger spread may be partially an artifact of the fact that when averaging over a smaller region, there is less cancellation of random variations. In addition, the fact that the networks of in situ observations are much sparser in the tropics than in the extratropics of the Northern Hemisphere may also contribute.

RATPAC	0.13 (0.068)	0.08 (0.119)	0.06 (0.136)	0.07 (0.153)	0.00 (0.140)	-0.75 (0.362)	-0.69 (0.289)
HadAT2	0.15 (0.115)	0.05 (0.152)	0.03 (0.164)	0.02 (0.176)	-0.04 (0.170)	-0.66 (0.304)	-0.64 (0.307)
Satellite:							
UAH		0.05 (0.176)		0.09 (0.191)	0.05 (0.167)		-0.37 (0.281)
RSS		0.15 (0.192)		0.18 (0.196)	0.14 (0.175)		-0.29 (0.303)
U.Md.					0.19 (0.159)		
Reanalyses:							
US	0.03 (0.163)	0.05 (0.172)	0.04 (0.173)	-0.03 (0.183)	-0.10 (0.166)	-0.89 (0.405)	-0.83 (0.340)
European	0.03 (0.211)	0.00 (0.234)	-0.03 (0.249)	0.06 (0.255)	0.05 (0.232)	-0.03 (0.453)	-0.05 (0.423)

828

829 **6.2.2 Latitude bands**

830 Globally averaged temperatures paint only part of the picture. Different layers of the
831 atmosphere behave differently depending on the latitude. Furthermore, even the
832 processing of the data can make for latitudinal difference in long-term trends. Figure 3.5
833 shows the trends in temperature for different datasets and levels averaged over latitude
834 bands. Each of these trends was created by making a latitudinally averaged time series of
835 monthly anomalies and then fitting that time series with a standard least-squares linear
836 regression slope.

837

838

839 Figure 3.5 -- Temperature trends for 1979-2004 ($^{\circ}\text{C}/\text{decade}$) by latitude.840 Left: stratospheric temperature (T_4) based on RSS (red) and UAH (blue) satellite datasets, and RATPAC
841 (violet) and HadAT2 (green) radiosonde datasets.842 Middle: mid-tropospheric temperature (T_2) based on U.Md. (orange), RSS (red) and UAH (blue) satellite
843 datasets, and RATPAC (violet) and HadAT2 (green) radiosonde datasets; and surface temperature (T_S)
844 from NOAA data (black).845 Right: surface temperature (T_S) from NOAA data (black) and lower tropospheric temperature (T_{2LT}) from
846 RSS (red) and UAH satellite data (blue), and from RATPAC (violet) and HadAT2 (green) radiosonde data.847 Filled circles denote trends estimated to be statistically significantly different from zero (at the 5% level). A
848 Student's t-test, using the lag-1 autocorrelation to account for the non-independence of residual values849 about the trend line, was used to assess significance (see Appendix for discussion of confidence intervals
850 and significance testing).

851

852

853

854

855 In the stratosphere (left panel of Figure 3.5), trend profiles for the two satellite datasets
856 are fairly similar, with a greater temperature decrease everywhere according to T_{4-A} than
857 T_{4-R} . Some of the largest temperature decrease occurs in the South Polar Region, where
858 ozone depletion is largest. A broad region of weaker decrease occurs in the deep tropics.
859 By contrast, the RATPAC and HadAT2 radiosonde datasets are quite different from the
860 satellite products, with much flatter profiles. It is worth noting that there is a fundamental
861 disagreement between the radiosonde and satellite products. Except for the mid-latitudes
862 of the Northern Hemisphere¹⁷, at most other latitudes the radiosonde products show more
863 of a temperature decrease than the satellite products, with the largest discrepancy in the
864 tropics¹⁸.

865

866 For the middle troposphere (middle panel of Figure 3.5) there is general agreement
867 among the radiosonde and satellite datasets in depicting the same basic structure. The
868 largest temperature increase occurs in the extratropics of the Northern Hemisphere, with
869 a smaller increase or slight decrease in the tropics, and even lesser increase or more
870 decrease in the extratropics of the Southern Hemisphere. At most latitudes, T_{2-M} indicates
871 the most increase (least decrease), followed next by T_{2-R} , then T_{2-A} , and finally the
872 radiosonde products with the least increase (most decrease).

¹⁷ The apparently better radiosonde-satellite agreement in the midlatitudes of the Northern Hemisphere may be the result of spurious stratospheric warming at stations located in countries of the former Soviet Union, offsetting the more typical spurious cooling bias of radiosonde temperatures (Lanzante et al., 2003a,b).

¹⁸ We note that in the tropics, where the radiosonde and satellite products differ the most, abrupt artificial drops in temperature appear to be particularly problematic for radiosonde data (Parker et al., 1997; Lanzante et al., 2003a,b). Other studies (Sherwood et al., 2005; Randel and Wu, 2005) also suggest spurious cooling for radiosonde temperatures, especially in the tropics.

873 For the lower troposphere and surface (right panel of Figure 3.5) the profiles are roughly
874 similar in shape to those for the middle troposphere with one major exception: the higher-
875 latitude temperature increase of the Northern Hemisphere is more pronounced compared
876 to the other regions. Comparing the surface temperature trend profile (black) with that
877 from the various tropospheric products in the middle and right panels of Figure 3.5
878 suggests that the sign and magnitude of this difference is highly dependent upon which
879 tropospheric measure is used.

880

881 **6.2.3 Maps**

882 Trend maps represent the finest spatial granularity with which different levels/layers and
883 observing platforms can be compared. However, since maps may not be the optimal way
884 in which to examine trends¹⁹, we present only a limited number of such maps for
885 illustrative purposes. Figure 3.6 presents maps of trends for the surface (bottom), lower
886 troposphere (second from bottom), upper middle troposphere (second from top), and
887 stratosphere (top). The surface map is based on the NOAA dataset²⁰ while those for the
888 troposphere and stratosphere are based on the RSS satellite dataset²¹. In examining these

¹⁹ Averaging over space (e.g., over latitudes, the tropics or the globe, as presented earlier) tends to reduce noise that results from the statistical uncertainties inherent to any observational measurement system. Furthermore, models that are used to study climate change have limited ability to resolve the smallest spatial scales and therefore there is little expectation of detection at the smallest scales (Stott and Tett, 1998). The formal methodology that is used to compare models with observations (“fingerprinting,” see Chapter 5) concentrates on the larger-scale signals in both models and observations in order to optimize the comparisons.

²⁰ Trend maps from other surface datasets (not shown) tend to be fairly similar to that of the NOAA map, differing mostly in their degree of spatial smoothness, which is a function of dataset construction methodology.

²¹ A comparison between UAH and RSS trend maps for tropospheric layers is given in Chapter 4.

889 maps it should be kept in mind that based on theory *we expect* the difference in trend
890 between the surface and troposphere to vary by location. For example, as shown in
891 Chapter 1, climate model projections typically indicate that human induced changes
892 should lead to more warming of the troposphere than the surface in the tropics, but the
893 opposite in the Arctic and Antarctic.

894

895 Figure 3.6 – Temperature trends for 1979-2004 ($^{\circ}$ C /decade).

896 Bottom (d): NOAA surface temperature (T_{S-N}).

897 Third (c): RSS lower tropospheric temperature (T_{2LT-R}).

898 Second (b): RSS upper middle tropospheric temperature (T_{2-R}).

899 Top (a): RSS lower stratospheric temperature (T_{4-R}).

900

901 The trend maps indicate both similarities and differences between the surface and
902 tropospheric trend patterns. There is a rough correspondence in patterns between the two.

903 The largest temperature increase occurs in the extratropics of the Northern Hemisphere,
904 particularly over landmasses. A decreases or smaller increase is found in the high
905 latitudes of the Southern Hemisphere as well as in the eastern tropical Pacific. Note the
906 general correspondence between the above noted features in Figures 3.6c,d and the zonal
907 trend profiles (middle and right panels of Figure 3.5). Note that the upper middle
908 tropospheric temperature is somewhat of a hybrid measure, being affected most strongly
909 by the troposphere, but with a non-negligible influence by the stratosphere.

910

911 In contrast to the surface and troposphere, a temperature decrease is found almost
912 everywhere in the stratosphere (Figure 3.6a). The largest decrease is found in the
913 midlatitudes of the Northern Hemisphere and the South Polar Region, with a smaller
914 decrease in the tropics. Again note the correspondence between the main features of the
915 trend map (Figure 3.6a) and the corresponding zonal trend profiles (left panel of Figure
916 3.5).

917

918 **7. CHANGES IN VERTICAL STRUCTURE**

919

920 **7.1 Vertical profiles of trends**

921 Up to this point, our vertical comparisons have contrasted trends of surface temperature
922 with trends based on different layer-averaged temperatures. Layers are useful because the
923 averaging process tends to reduce noise. The use of layer-averages is also driven by the
924 limitations of satellite measurement systems that are unable to provide much vertical
925 detail. However, as illustrated in Chapter 1, changes in various forcing agents can lead to
926 more complex changes in the vertical. Radiosonde data, because of their greater vertical
927 resolution, are much better suited for this than currently available satellite data.

928

929 Figure 3.7 shows vertical profiles of trends from the RATPAC and HadAT2 radiosonde
930 datasets for temperature averaged over the globe (top) or tropics (bottom) for the
931 radiosonde (left) and satellite (right) eras. The trend values of Figure 3.7 are also given in
932 Table 3.5. Each graph has profiles for the two radiosonde datasets. The tropics are of
933 special interest because many climate models suggest that under global warming
934 scenarios trends should increase from the lower troposphere upwards, maximizing in the
935 upper troposphere (see Chapters 1 and 5).

936 Table 3.5 – Temperature trends in °C per decade from the RATPAC and HadAT2 radiosonde datasets
937 corresponding to the plots in Figure 3.7 (see figure caption for further details). Global and tropical trends
938 are given for 1958 through 2004 and 1979 through 2004 (except for European which terminates September
939 2001). The HadAT2 dataset does not have temperatures for some of the levels, hence the empty table cells.
940 The trend is shown for each vertical level (hPa), with the approximate 95% confidence interval (2 sigma)
941 below in parentheses. Bold values are estimated to be statistically significantly different from zero (at the
942 5% level). A Student's t-test, using the lag-1 autocorrelation to account for the non-independence of
943 residual values about the trend line, was used to assess significance (see Appendix for discussion of
944 confidence intervals and significance testing).

945

947
948 Figure 3.7 -- Vertical profiles of temperature trend ($^{\circ}\text{C}/\text{decade}$) as a function of altitude (i.e., pressure in
949 hPa) computed from the RATPAC (violet) and HadAT2 (green) radiosonde datasets. Trends (which are
950 given in Table 3.5) have been computed for 1958-2004 (left) and 1979-2004 (right) based on temperature
951 that has been averaged over the globe (top) or the tropics, 20°N - 20°S (bottom). Surface data for the
952 HadAT2 product is taken from HadCRUT2v since the HadAT2 dataset does not include values at the
953 surface; the surface values have been averaged so as to match their observing locations with those for the
954 radiosonde data. By contrast, the surface temperatures from the RATPAC product are those from the
955 RATPAC dataset, which are surface station values reported with the radiosonde data. Note that these differ
956 from the NOAA surface dataset values (ER-GHCN-ICOADS) as indicated in Table 3.1. Filled symbols
957 denote trends estimated to be statistically significantly different from zero (at the 5% level). A Student's t-
958 test, using the lag-1 autocorrelation to account for the non-independence of residual values about the trend
959 line, was used to assess significance (see Appendix for discussion of confidence intervals and significance
960 testing).

961

962

963 For the globe, the figure indicates that during the longer period the tropospheric
964 temperature increased slightly more than that of the surface. By contrast, for the globe

965 during the satellite era, the surface temperature increased more than that of the
966 troposphere. Both datasets agree reasonably well in these conclusions. For the tropics, the
967 differences between the two eras are more pronounced. For the longer period there is
968 good agreement between the two datasets in that the temperature increase is smaller at the
969 surface and maximized in the upper troposphere. The largest disagreement between
970 datasets and least amount of tropospheric temperature increase is seen in the tropics
971 during the satellite era. For the RATPAC product, the greatest temperature increase
972 occurs at the surface with a slight increase (or decrease) in the lower and middle
973 troposphere followed by somewhat larger increase in the upper troposphere. The HadAT2
974 product also shows largest increase at the surface, with a small increase in the
975 troposphere, however, it lacks a distinct return to increase in the upper troposphere. In
976 summary, the two datasets have fairly similar profiles in the troposphere with the
977 exception of the tropics during the satellite era²². For the stratosphere, the decrease in
978 temperature is noticeably greater for both the globe and the tropics during the satellite
979 than radiosonde era as expected (see Figure 3.2b). Some of the largest discrepancies
980 between datasets are found in the stratosphere.

981

982

983 **7.2 Lapse rates**

984 Temperature usually decreases in the troposphere going upward from the surface. Lapse

²² However, the differences between datasets may not be meaningful since they are small compared to the statistical uncertainty estimates (see Table 3.5 and discussion in the Appendix).

985 rate is defined as the rate of decrease in temperature with increasing altitude and is a
986 measure of the stability of the atmosphere²³. Most of the observational work to date has
987 not examined lapse rates themselves, but instead has used an approximation in the form
988 of a vertical temperature difference²⁴. This difference has taken on the form of the surface
989 temperature minus some tropospheric temperature, either layer-averaged (in the case of
990 satellite data) or at some specific pressure level (in the case of radiosonde data)²⁵.

991

992 Much of the interest in lapse rate variations has focused on the tropics. Several studies
993 (Brown et al., 2000; Gaffen et al., 2000; Hegerl and Wallace, 2002; Lanzante et al.,
994 2003b) present time series related to tropical lapse rate based on either satellite or
995 radiosonde measures of tropospheric temperature. As examples, we present some such
996 time series in Figure 3.8, based on measures of lower tropospheric temperature from
997 three different datasets. Some essential low-frequency characteristics are common to all.
998 A considerable proportion of the variability of the tropical lapse rate is associated with
999 ENSO²⁶, a manifestation of which are the up and down swings of about 3-7 years in the

²³ A larger lapse rate implies more unstable conditions and a greater tendency towards vertical mixing of air.

²⁴ The reasons for this are two-fold: (1) satellite measurement systems are only able to resolve temperatures in deep layers rather than at specific levels, and (2) radiosonde measurements are consistently recorded at a fixed number of constant pressure rather than height levels.

²⁵ When constant pressure level data from radiosondes are used, the resulting lapse rate quantity may be influenced by changes in the thickness (i.e., average temperature) of the layer. However, some calculations by Gaffen et al. (2000) suggest that thickness changes do not have very much influence. Therefore, we consider vertical temperature differences to be a suitable approximation of lapse rate

²⁶ Lapse rate changes occur about five to six months after a particular change in ENSO (Hegerl and Wallace, 2002; Lanzante et al., 2003b). During a tropical warming event (El Niño) the tropical troposphere warms relative to the surface; the opposite is true during a tropical cooling event (La Niña).

1000 series shown in Figure 3.8. Another feature evident in the four studies cited above, and
1001 seen in Figure 3.8 as well, is an apparent strong association with the climate regime shift
1002 that occurred ~1976-77 (Trenberth and Hurrell, 1994). There is a rather sharp drop in
1003 tropical lapse rate at this time²⁷, coincident with an abrupt change in a measure of
1004 convective stability (Gettelman et al. 2002). Overall, the variation in tropical lapse rate
1005 can be characterized as highly complex, with rapid swings over a few years,
1006 superimposed upon persistent periods of a decade or more, as well as longer-term drifts
1007 or trends evident during some time periods.

1008
1009
1010

1011 Figure 3.8 - Time series of vertical temperature difference (surface minus lower troposphere) for the tropics
1012 (20°N-20°S). NOAA surface temperatures ($T_{S,N}$) are used in each case to compute differences with lower
1013 tropospheric temperature (T_{2LT}) from three different groups: HadAT2 radiosonde (green), RSS satellite
1014 (red), and UAH satellite (blue). All time series are 7-month running averages (used as a smoother) of
1015 original monthly data, which were expressed as a departure (°C) from the 1979-97 average.
1016

²⁷ Lanzante et al. (2003b) also noted an apparent decrease in the amplitude of ENSO-related tropical lapse rate variations after the ~1976-77 regime shift.

1017 The feature of the tropical lapse rate series that has drawn the most interest is the linear
1018 trend component during the satellite era. From a long historical perspective (see also
1019 Figure 3.8), this trend is a rather subtle feature, being overshadowed by both the ENSO-
1020 related variations as well as the regime shift of the late 1970s. Several studies (Brown et
1021 al., 2000; Gaffen et al., 2000; Hegerl and Wallace, 2002; Lanzante et al., 2003b) have
1022 estimated trends in lower tropospheric lapse rate while another (Christy et al., 2001) has
1023 estimated trends in the difference between SST and surface air temperature.

1024

1025 The different trend estimates vary considerably among the above-cited studies, being
1026 dependent upon the details of the calculations²⁸. From the cited studies, satellite-era
1027 trends in lapse rate based on temperatures averaged over the tropics range from nearly
1028 zero (no change) to about 0.20°C/decade (surface warms more than the troposphere). The
1029 time series of Figure 3.8 also exhibit a wide range of satellite-era trends²⁹. During the
1030 longer radiosonde era, the various studies found trends of opposite sign (i.e., air
1031 temperature at the surface increases more slowly than that of air aloft) and show less

²⁸These details include: time period, latitude zone, datasets utilized, station network vs. grid, time of day of observations, use of homogeneity adjustment, and whether or not measurements in the troposphere and surface were taken from the same locations. Particularly noteworthy is the fact that Lanzante et al. (2003b) found that during the satellite era, use of homogenized data could, depending on the other details of the analysis, either halve or eliminate the positive tropical lapse rate trend found using the unadjusted data.

²⁹ Trends from 1979 to 2004 (°C /decade) for the three time series in Figure 3.8 are: 0.11 (HadAT2 radiosonde), 0.08 (UAH satellite), and -0.02 (RSS satellite). While the first two of these trends are statistically significant at the 5% level, the third is not (see Appendix for discussion of significance testing).

1032 sensitivity, with a range of values of near-zero to about $-0.05^{\circ}\text{C}/\text{decade}$ ³⁰.

1033

1034 Spatial variations in lapse rate trends have also been examined. During the satellite era,
1035 some have found predominantly increasing trends in the tropics (Gaffen et al., 2000;
1036 Brown et al., 2000) while others have found a greater mixture, with more areas of
1037 negative trends (Hegerl and Wallace, 2002; Lanzante et al., 2003b). Outside of the
1038 tropics, both Hegerl and Wallace (2002) and Lanzante et al. (2003b) found complex
1039 spatial patterns of trend. Lanzante et al. (2003b) also found considerable local sensitivity
1040 to homogeneity adjustment in the tropics and even more so over the extratropics of the
1041 Southern Hemisphere, which is quite sparsely sampled.

³⁰ The trend from 1958 to 2004 for the HadAT2 radiosonde series shown in Figure 3.8 is $-0.02^{\circ}\text{C}/\text{decade}$. This trend is not statistically significant at the 5% level (see Appendix for discussion of significance testing).

1042 **REFERENCES**

1043

1044 **Aguilar**, E., I. Auer, M. Brunet, T. C. Peterson and J. Wieringa, 2003: *Guidelines on*
1045 *Climate Metadata and Homogenization*, WCDMP-No. 53, WMO-TD No. 1186.
1046 World Meteorological Organization, Geneva, 55 pp.

1047

1048 **Andersson**, E., and Coauthors, 1998: The ECMWF implementation of three-dimensional
1049 variational assimilation (3D-Var). Part III: Experimental results. *Quarterly*
1050 *Journal of the Royal Meteorological Society*, **124**, 1831-1860.

1051

1052 **Angell**, J. K., and J. Korshover, 1975: Estimate of the global change in tropospheric
1053 temperature between 1958 and 1973. *Monthly Weather Review*, **103**, 1007-1012.

1054

1055 **Angell**, J. K., 2003: Effect of exclusion of anomalous tropical stations on temperature
1056 trends from a 63-station radiosonde network, and comparison with other analyses.
1057 *Journal of Climate*, **16**, 2288-2295.

1058

1059 **Bottomley**, M., Folland, C.K., Hsiung, J., Newell, R.E. and D.E. Parker, 1990: *Global*
1060 *Ocean Surface Temperature Atlas "GOSTA"*. Joint project of Met Office and
1061 Massachusetts Institute of Technology supported by US Dept. of Energy, US
1062 National Science Foundation and US Office of Naval Research. HMSO, London,
1063 20pp+iv, 313 plates.

1064

1065 **Brown**, S., D. Parker, C. Folland, and I. Macadam, 2000: Decadal variability in the
1066 lower-tropospheric lapse rate. *Geophysical Research Letters*, **27**, 997-1000.

1067

1068 **Christy**, J., D. Parker, S. Brown, I. Macadam, M. Stendel, and W. Norris, 2001:
1069 Differential trends in tropical sea surface and atmospheric temperature since
1070 1979. *Geophysical Research Letters*, **28(1)**, 183-186.

1071

1072 **Christy**, J. R., R. W. Spencer, and E. S. Lobl, 1998: Analysis of the merging procedure
1073 for the MSU daily temperature time series. *Journal of Climate*, **11**, 2016-2041.

1074

1075 **Christy**, J. R., R. W. Spencer, and W. D. Braswell, 2000: MSU tropospheric
1076 temperatures: dataset construction and radiosonde comparisons. *Journal of*
1077 *Atmospheric and Oceanic Technology*, **17**, 1153-1170.

1078

1079 **Christy**, J. R., R. W. Spencer, W. B. Norris, W. D. Braswell, and D. E. Parker, 2003:
1080 Error estimates of version 5.0 of MSU-AMSU bulk atmospheric temperatures. .

- 1081 *Journal of Atmospheric and Oceanic Technology*, **20**, 613-629.
- 1082
- 1083 **Dai**, A., K.E. Trenberth, and T.R. Karl, 1999: Effects of clouds, soil moisture,
1084 precipitation, and water vapor on diurnal temperature range. *Journal of Climate*,
1085 **12**, 2451-2473.
- 1086 **Diaz**, H.F., Folland, C.K., Manabe, T., Parker, D.E., Reynolds, R.W. and Woodruff, S.D.,
1087 2002: Workshop on Advances in the Use of Historical Marine Climate Data
1088 (Boulder, Co., USA, 29th Jan - 1st Feb 2002). *WMO Bulletin*. **51**, 377-380
- 1089
- 1090 **Durre**, I., R. S. Vose, and D. B. Wuertz, 2005: Overview of the Integrated Global
1091 Radiosonde Archive. *Journal of Climate*, submitted.
- 1092
- 1093 **Easterling**, D.R., T.R. Karl, E.H. Mason, P.Y. Hughes, and D.P. Bowman, 1996: *United*
1094 *States Historical Climatology Network (U.S. HCN), monthly temperature and*
1095 *precipitation data*. (Environmental Sciences Division Publication no. 4500;
1096 ORNL/CDIAC 87; NDP-019/R3). Carbon Dioxide Information Analysis Center,
1097 Oak Ridge National Laboratory, Oak Ridge, 263 pp.
- 1098
- 1099 **Easterling**, D. R., B. Horton, P. D. Jones, T. C. Peterson, T. R. Karl, D. E. Parker, M. J.
1100 Salinger, V. Razuvayev, N. Plummer, P. Jamason, and C. K. Folland, 1997:
1101 Maximum and minimum temperature trends for the globe. *Science*, **277**, 364-367.
- 1102
- 1103 **Folland**, C.K., Rayner, N.A., Brown, S.J., Smith, T.M., Shen, S.S.P., Parker, D.E.,
1104 Macadam, I., Jones, P.D., Jones, R.N., Nicholls, N. and Sexton, D.M.H., 2001a:
1105 Global temperature change and its uncertainties since 1861. *Geophysical*
1106 *Research Letters*, **28**, 2621-2624.
- 1107
- 1108 **Folland**, C.K., T.R. Karl, J.R. Christy, R.A. Clarke, G.V. Gruza, J. Jouzel, M.E. Mann, J.
1109 Oerlemans, M.J. Salinger and S-W. Wang, and 142 other authors, 2001b:
1110 Observed Climate Variability and Change. In: *Climate Change 2001: The*
1111 *Scientific Basis. Contribution of Working Group I to the Third Assessment Report*
1112 *of the Intergovernmental Panel on Climate Change* [Houghton, J. T., Y Ding, D.
1113 J. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. I. Johnson
1114 (eds.)]. Cambridge University Press, 99-181.
- 1115
- 1116 **Folland**, C.K., M. J. Salinger, N. Jiang, and N.A. Rayner, 2003: Trends and variations in
1117 South Pacific Island and ocean surface temperatures. *Journal of Climate*, **16**,
1118 2859-2874.
- 1119

- 1120 **Free**, M., and J. K. Angell, 2002: Effect of volcanoes on the vertical temperature profile
1121 in radiosonde data. *Journal of Geophysical Research*, **107**, 4101, doi:
1122 10.1029/2001JD001128.
- 1123
- 1124 **Free**, M., J. K. Angell, I. Durre, J. R. Lanzante, T. C. Peterson and D. J. Seidel, 2003:
1125 Using first differences to reduce inhomogeneity in radiosonde temperature
1126 datasets. *Journal of Climate*, **21**, 4171-4179.
- 1127
- 1128 **Free**, M., I. Durre, E. Aguilar, D. J. Seidel, T. C. Peterson, R. E. Eskridge, J. K. Luers, D.
1129 Parker, M. Gordon, J. R. Lanzante, S. A. Klein, J. R. Christy, S. Schroeder, B. J.
1130 Soden, and L. M. McMillin, 2002: CARDS workshop on adjusting radiosonde
1131 temperature data for climate monitoring: Meeting summary, *Bulletin of the*
1132 *American Meteorological Society*, **83**, 891-899.
- 1133
- 1134 **Free**, M., D. J. Seidel, J. K. Angell, J. R. Lanzante, I. Durre, and T. C. Peterson, 2005:
1135 Radiosonde Atmospheric Temperature Products for Assessing Climate
1136 (RATPAC): A new dataset of large-area anomaly time series. *Journal of*
1137 *Geophysical Research*, submitted.
- 1138
- 1139 **Fu**, Q., and C. M. Johanson, 2005: Satellite-Derived Vertical Dependence of Tropical
1140 Tropospheric Temperature Trends. *Geophysical Research Letters*, **32**, L10703,
1141 doi:10.1029/2004GL022266.
- 1142
- 1143 **Fu**, Q., C. M. Johanson, S. Warren, and D. Seidel, 2004: Contribution of stratospheric
1144 cooling to satellite-inferred tropospheric temperature trends. *Nature*, **429**, 55-58.
- 1145
- 1146 **Gaffen**, D, B. Santer, J. Boyle, J. Christy, N. Graham, and R. Ross, 2000: Multi-decadal
1147 changes in the vertical temperature structure of the tropical troposphere. *Science*,
1148 **287**, 1239-1241.
- 1149
- 1150 **Gettelman**, A., D. J. Seidel, M. C. Wheeler and R. J. Ross, 2002: Multidecadal trends in
1151 tropical convective available potential energy. *Journal of Geophysical Research*,
1152 **107**, 4606, doi:10.1029/2001JD001082.
- 1153
- 1154 **Grody**, N. C., K. Y. Vinnikov, M. D. Goldberg, J. T. Sullivan, and J. D. Tarpley, 2004:
1155 Calibration of Multi-Satellite Observations for Climatic Studies: Microwave
1156 Sounding Unit (MSU). *Journal of Geophysical Research*. **109**, D24104,
1157 doi:10.1029/2004JD005079.
- 1158

- 1159 **Groisman, P.Y., R.W. Knight, T.R. Karl, D.R. Easterling, B. Sun, and J.H. Lawrimore,**
1160 2004: Contemporary changes of the hydrological cycle over the contiguous
1161 United States, trends derived from in situ observations. *Journal of*
1162 *Hydrometeorology*, **5** 64-85.
- 1163
- 1164 **Hansen, J., R. Ruedy, M. Sato, and R. Reynolds** 1996. Global surface air temperature in
1165 1995: Return to pre-Pinatubo level. *Geophysical Research Letters*, **23**, 1665-1668.
- 1166
- 1167 **Hansen, J., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and**
1168 **T. Karl,** 2001: A closer look at United States and global surface temperature
1169 change. *Journal of Geophysical Research*, **106**, 23,947-23,963.
- 1170
- 1171 **Hegerl, G., and J. Wallace,** 2002: Influence of patterns of climate variability on the
1172 difference between satellite and surface temperature trends. *Journal of Climate*,
1173 **15**, 2412-2428.
- 1174
- 1175 **Henderson-Sellers, A.,** 1992: Continental cloudiness changes this century. *GeoJournal*,
1176 **27**, 255–262.
- 1177
- 1178 **Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K.**
1179 **Maskell, and C. A. Johnson, Eds.,** 2001: Climate Change 2001: The Scientific
1180 Basis. Cambridge University Press, 881 pp.
- 1181
- 1182 **Hurrell, J.W.,** 1996: Influence of variations in extratropical wintertime teleconnections
1183 on Northern Hemisphere temperature. *Geophysical Research Letters*, **23**, 665-
1184 668.
- 1185
- 1186 **Johanson, C. M., and Q. Fu,** 2005: Robustness of tropospheric temperature trends from
1187 MSU Channels 2 and 4. *Journal of Climate*, submitted.
- 1188
- 1189 **Jones, P.D.,** 1995: Land surface temperatures - is the network good enough? *Climatic*
1190 *Change* **31**, 545-558.
- 1191
- 1192 **Jones, P. D., P. Ya. Groisman, M. Coughlan, N. Plummer, W-C. Wang, and T. R. Karl,**
1193 1990: Assessment of urbanization effects in time series of surface air temperature
1194 over land. *Nature*, **347**, 169-172.
- 1195
- 1196 **Jones, P.D., Osborn, T.J. and Briffa, K.R.,** 1997: Estimating sampling errors in large-
1197 scale temperature averages. *Journal of Climate*, **10**, 2548-2568.

1198

1199 **Jones**, P.D., New, M., Parker, D.E., Martin, S. and Rigor, I.G., 1999: Surface air
1200 temperature and its changes over the past 150 years. *Reviews of Geophysics*, **37**,
1201 173-199.

1202

1203 **Jones**, P.D., Osborn, T.J., Briffa, K.R., Folland, C.K., Horton, B., Alexander, L.V.,
1204 Parker, D.E. and Rayner, N.A., 2001: Adjusting for sampling density in grid-box
1205 land and ocean surface temperature time series. *Journal of Geophysical Research*
1206 **106**, 3371-3380.

1207

1208 **Jones**, P. D. and A. Moberg, 2003: Hemispheric and large-scale surface air temperature
1209 variations: An extensive revision and an update to 2001. *Journal of Climate*, **16**,
1210 206-223.

1211

1212 **Karl**, T. R, Jones, P. D, Knight, R. W, Kukla, G., Plummer, N., Razuvayev, V., Gallo,
1213 K. P., Lindsey, J., Charlson, R. J, and T. C. Peterson, 1993: A new perspective
1214 on recent global warming: Asymmetric trends of daily maximum and minimum
1215 temperature. *Bulletin of the American Meteorological Society*, **74**, 1007-1024.

1216

1217 **Karl**, T. R., R. W. Knight, and B. Baker, 2000: The record breaking global temperatures
1218 of 1997 and 1998: Evidence for an increase in the rate of global warming?
1219 *Geophysical Research Letters*, **27**, 719-722.

1220

1221 **Kistler**, R., and Coauthors, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means
1222 CD-ROM and documentation. *Bulletin of the American Meteorological Society*,
1223 **82**, 247-267.

1224

1225 **Lanzante**, J. R., S. A. Klein, and D. J. Seidel, 2003a: Temporal homogenization of
1226 monthly radiosonde temperature data. Part I: Methodology. *Journal of Climate*,
1227 **16**, 224-240.

1228 **Lanzante**, J. R., S. A. Klein, and D. J. Seidel, 2003b: Temporal homogenization of
1229 monthly radiosonde temperature data. Part II: Trends, sensitivities, and MSU
1230 comparison. *Journal of Climate*, **16**, 241-262.

1231

1232 **Mears**, C. A., M. C. Schabel, and F. J. Wentz, 2003: A Reanalysis of the MSU channel 2
1233 tropospheric temperature record. *Journal of Climate*, **16**, 3650-3664.

1234

1235 **Mears**, C. A., and F. J. Wentz, 2005: The effect of diurnal correction on satellite-derived
1236 lower tropospheric temperature. *Science*, submitted.

1237

1238 **Oleson**, K. W., G. B. Bonan, S. Levis and M. Vertenstein, 2004: Effects of land use
1239 change on North American climate: impact of surface datasets and model
1240 biogeophysics. *Climate Dynamics*, **23**,117 - 132.

1241

1242 **Oort**, A. H. (1983): Global Atmospheric Circulation Statistics, 1958-1973. NOAA Prof.
1243 Pap 4. US Government Printing Office, Washington, D.C., 180 pp.

1244

1245 **Parker** D.E., Folland, C.K. and M. Jackson, 1995: Marine surface temperature: observed
1246 variations and data requirements. *Climatic Change*, 31, 559-600, and in: *Long-*
1247 *term climate Monitoring by the Global Climate Observing System*, Ed: T. Karl,
1248 pp429-470, Kluwer, Dordrecht.

1249

1250 **Parker**, D. E., P. D. Jones, C. K. Folland and A. Bevan, 1994: Interdecadal changes of
1251 surface temperature since the late nineteenth century. *Journal of Geophysical*
1252 *Research*, **99**, 14,373-14,399.

1253

1254 **Parker**, D. E., M. Gordon, D. P. N. Cullum, D. M. H. Sexton, C. K. Folland, and N.
1255 Rayner, 1997: A new global gridded radiosonde temperature data base and recent
1256 temperature trends. *Geophysical Research Letters*, **24**, 1499-1502.

1257

1258 **Parker**, D.E., Alexander, L.V. and Kennedy, J., 2004: Global and regional climate in
1259 2003. *Weather*, **59**, 145-152.

1260

1261 **Parker**, D.E., 2004: Large-scale warming is not urban, *Nature*, **432**, 290-291,
1262 10.1038/432290b.

1263

1264 **Pawson**, S., K. Labitzke, and S. Leder, 1998: Stepwise changes in stratospheric
1265 temperature. *Geophysical Research Letters*, **25**, 2157-2160.

1266

1267 **Peterson**, Thomas C., 2003: Assessment of urban versus rural in situ surface
1268 temperatures in the contiguous U.S.: No difference found. *Journal of Climate*,
1269 **18**, 2941-2959.

1270

1271 **Peterson**, Thomas C., 2005: Examination of potential biases in air temperature caused by
1272 poor station locations. *Bulletin of the American Meteorological Society*,
1273 submitted.

1274

- 1275 **Peterson**, T. C., D. R. Easterling, T. R. Karl, P. Ya. Groisman, N. Nicholls, N. Plummer,
1276 S. Torok, I. Auer, R. Boehm, D. Gullett, L. Vincent, R. Heino, H. Tuomenvirta,
1277 O. Mestre, T. Szentimre, J. Salinger, E. Førland, I. Hanssen-Bauer, H.
1278 Alexandersson, P. Jones, D. Parker, 1998: Homogeneity adjustments of in situ
1279 atmospheric climate data: A review. *International Journal of Climatology*, **18**,
1280 1493-1517.
- 1281
- 1282 **Peterson**, T. C., T. R. Karl, P. F. Jamason, R. Knight, and D. R. Easterling, 1998b: The
1283 first difference method: Maximizing station density for the calculation of long-
1284 term global temperature change. *Journal of Geophysical Research*, **103**, 25,967-
1285 25,974.
- 1286
- 1287 **Peterson**, T. C., K. P. Gallo, J. Lawrimore, T. W. Owen, A. Huang, and D. A.
1288 McKittrick, 1999: Global rural temperature trends. *Geophysical Research Letters*,
1289 **26**, 329-332.
- 1290
- 1291 **Peterson**, T.C., T.W. Owen, 2005. Urban heat island assessment: Metadata are
1292 important. *Journal of Climate*, in press.
- 1293
- 1294 **Peterson**, Thomas C. and Russell S. Vose, 1997: An overview of the Global Historical
1295 Climatology Network temperature data base. *Bulletin of the American*
1296 *Meteorological Society*, **78**, 2837-2849.
- 1297
- 1298 **Prabhakara**, C., R. Iacovazzi Jr, J.-M. Yoo, and G. Dalu., 2000: Global warming:
1299 estimation from satellite observations. *Geophysical Research Letters*, **27**, 3517-
1300 3520.
- 1301
- 1302 **Randel**, W. J. and F. Wu, 2005: Biases in stratospheric and tropospheric temperature
1303 trends derived from historical radiosonde data. *Journal of Climate*, accepted.
- 1304
- 1305 **Rayner**, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell,
1306 E. C. Kent and A. Kaplan, 2003: Global analyses of sea surface temperature, sea
1307 ice, and night marine air temperature since the late nineteenth century. *Journal of*
1308 *Geophysical Research*, **108**, 4407, doi:10.1029/2002JD002670.
- 1309
- 1310 **Reynolds**, R. W. and T. M. Smith, 1994: Improved global sea surface temperature
1311 analyses using optimum interpolation. *Journal of Climate*, **7**, 929-948.
- 1312
- 1313 **Santer**, B. D., and Coauthors, 2004: Identification of anthropogenic climate change using
1314 a second-generation reanalysis. *Journal of Geophysical Research*, **109**, D21104,

- 1315 doi:10.1029/2004JD005075.
- 1316
- 1317 **Seidel**, D. J., J. K. Angell, M. Free, J. Christy, R. Spencer, S. A. Klein, J. R. Lanzante, C.
1318 Mears, M. Schabel, F. Wentz, D. Parker, P. Thorne, and A. Sterin, 2004:
1319 Uncertainty in signals of large-scale climate variations in radiosonde and satellite
1320 upper-air temperature datasets. *Journal of Climate*, **17**, 2225-2240.
- 1321
- 1322 **Seidel**, D.J., and J.R. Lanzante, 2004: An assessment of three alternatives to linear trends
1323 for characterizing global atmospheric temperature changes. *Journal of*
1324 *Geophysical Research*, **109**, doi: 10.1029/2003JD004414.
- 1325
- 1326 **Sherwood**, S., J. R. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late-
1327 20th Century warming. *Science*, **309**, 1556-1559.
- 1328
- 1329 **Slutz**, R. J., S. J. Lubker, J. D. Hiscox, S. D. Woodruff, R. L. Jenne, D. H. Joseph, P. M.
1330 Steurer, and J. D. Elms, 1985: *COADS: Comprehensive Ocean-Atmosphere*
1331 *Data Set. Release 1*, 262 pp. [Available from Climate Research Program,
1332 Environmental Research Laboratories, 325 Broadway, Boulder, CO 80303.]
- 1333
- 1334 **Smith**, T. M., R. W. Reynolds, R. E. Livezey and D. C. Stokes, 1996: Reconstruction of
1335 historical sea surface temperatures using empirical orthogonal functions. *Journal*
1336 *of Climate*, **9**, 1403-1420.
- 1337
- 1338 **Smith**, T.M., and R.W. Reynolds, 2003: Extended reconstruction of global sea surface
1339 temperatures based on COADS Data (1854-1997). *Journal of Climate*, **16**, 1495-
1340 1510.
- 1341 **Smith**, T.M. and R. W. Reynolds, 2005: A global merged land and sea surface
1342 temperature reconstruction based on historical observations (1880-1997). *Journal*
1343 *of Climate*, **18**, 2021-2036.
- 1344
- 1345 **Smith**, T. M., T. C. Peterson, J. Lawrimore, and R. W. Reynolds, 2005: New surface
1346 temperature analyses for climate monitoring. *Geophysical Research Letters*,
1347 submitted.
- 1348
- 1349 **Spencer**, R.W. and J. R. Christy, 1990: Precise monitoring of global temperature trends
1350 from satellites. *Science*, **247**, 1558-1562.
- 1351
- 1352 **Spencer**, R.W. and J. R. Christy, 1992: Precision and radiosonde validation of satellite
1353 gridpoint temperature anomalies. Part I: MSU channel 2. *Journal of Climate*, **5**,

- 1354 847-857.
- 1355
- 1356 **Stendel**, M., J. Christy, and L. Bengtsson, 2000: Assessing levels of uncertainty in recent
1357 temperature time series. *Climate Dynamics*, **16**, 587-601.
- 1358
- 1359 **Stott**, P. A., and S. F. B. Tett, 1998: Scale-dependent detection of climate change.
1360 *Journal of Climate*, **11**, 3282-3294.
- 1361
- 1362 **Sun**, B. and P. Ya. Groisman, 2000: Cloudiness variations over the former Soviet Union.
1363 *International Journal of Climatology*, **20**, 1097–1111.
- 1364
- 1365 **Thorne**, P.W., D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, and P.
1366 Brohan, 2005: Revisiting radiosonde upper-air temperatures from 1958 to 2002.
1367 *Journal of Geophysical Research*, **110**, D18105, doi:10.1029/2004JD00575.
- 1368
- 1369 **Trenberth**, K. E., 1990: Recent observed interdecadal climate changes in the Northern
1370 Hemisphere. *Bulletin of the American Meteorological Society*, **71**, 988-993.
- 1371
- 1372 **Trenberth**, K., and J. Hurrell, 1994: Decadal atmosphere-ocean variations in the Pacific.
1373 *Climate Dynamics*, **9**, 303-319.
- 1374
- 1375 **Trenberth**, K.E., Carron, J.M., Stepaniak, D.P. and S. Worley, 2002: Evolution of the El
1376 Nino-Southern Oscillation and global atmospheric surface temperatures. *Journal*
1377 *of Geophysical Research*, **107**, D8, 10.1029/2000JD000298.
- 1378
- 1379 **Van den Dool**, H. M., S. Saha, and A. Johansson, 2000: Empirical orthogonal
1380 teleconnections. *Journal of Climate*, **13**, 1421-1435.
- 1381
- 1382 **Vinnikov**, K. Y., and N. C. Grody, 2003: Global warming trend of mean tropospheric
1383 temperature observed by satellites. *Science*, **302**, 269-272.
- 1384
- 1385 **Vinnikov**, K. Y., A. Robock, N. C. Grody, and A. Basist, 2004: Analysis of diurnal and
1386 seasonal cycles and trends in climatic records with arbitrary observation times.
1387 *Geophysical Research Letters*, **31**, L06205, doi:10.1029/2003GL019196, 2004.
- 1388
- 1389 **Vinnikov**, K. Y., N. C. Grody, A. Robock, R. J. Stouffer, P. D. Jones, and M. D.
1390 Goldberg, 2005: Temperature Trends at the Surface and in the Troposphere.
1391 *Journal of Geophysical Research*, submitted.
- 1392

- 1393 **Vose**, R. S., C. N. Williams, T. C. Peterson, T. R. Karl, and D. R. Easterling, 2003: An
1394 evaluation of the time of observation bias adjustment in the U.S. historical climate
1395 network. *Geophysical Research Letters*, **30**, doi:10.1029/2003GL018111.
- 1396
- 1397 **Vose**, R.S., D.R. Easterling, and B. Gleason, 2005: Maximum and minimum temperature
1398 trends for the globe: An update through 2004. *Geophysical Research Letters*,
1399 accepted.
- 1400
- 1401 **Wallis**, T.W.R., 1998: A subset of core stations from the Comprehensive Aerological
1402 Reference Data Set (CARDS). *Journal of Climate*, **11**, 272-282.
- 1403
- 1404 **Waple**, A. M. and J. H. Lawrimore, eds, 2003: Climate of 2002. *Bulletin of the*
1405 *American Meteorological Society*, S1-S68.
- 1406
- 1407 **Woodruff**, S. D., H. F. Diaz, J. D. Elms, and S. J. Worley, 1998: COADS Release 2 data
1408 and metadata enhancements for improvements of marine surface flux fields.
1409 *Physics and Chemistry of the Earth*, **23**, 517-527.
- 1410
- 1411 **Wuertz**, D., R. S.Vose, T. C. Peterson and P. D. Jones, 2005: The GHCN-Jones
1412 comparison. *Geophysical Research Letters*, submitted.
- 1413

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

Chapter 4

What is our understanding of the contribution made by observational or methodological uncertainties to the previously reported vertical differences in temperature trends?

Convening Lead Author: Carl Mears

Lead Authors: Chris Forest, Roy Spencer, Russell Vose, and Dick Reynolds

Contributing Authors

Peter Thorne and John Christy

46 **Chapter 4: Key Findings**

47

48 *Surface*

49

50 **It is likely that errors in the homogenized surface air temperature data do not**
51 **contribute substantially to the large-scale differences between trends for**
52 **different levels because these errors are very likely to be smaller than those for**
53 **the upper air data.**

54 • Systematic local biases in surface trends may exist due to changes in station
55 exposure or instrumentation over land, and due to the small number of
56 measurements over a number of regions of the earth, including parts of the
57 oceans, sea ice areas, and some land areas. Such biases have been
58 documented at the local and regional scale, but no such effect has been
59 identified in the zonal and global averages presented in this report. On large
60 spatial scales, sampling studies suggest that these local biases in trends are
61 likely to mostly cancel through the use of many observations with differing
62 instrumentation.

63 • Since all known bias adjustments have not yet been applied to sea surface
64 temperature data, it is likely that errors remain in these data, though it is
65 generally agreed that these errors are likely to be small compared to errors in
66 radiosonde and satellite measurements of the upper air, especially for the
67 satellite era.

68

69 ***Troposphere***

70 **While all datasets indicate that the troposphere has warmed over both the**
71 **radiosonde era and the satellite era, uncertainties in the tropospheric data**
72 **make it difficult to determine whether the troposphere has warmed more than**
73 **or less than the surface. Some tropospheric datasets indicate that the**
74 **troposphere has warmed more than the surface, while others indicate the**
75 **opposite.**

- 76 • It is very likely that errors remain in the homogenized radiosonde datasets in
77 the troposphere since the methods used to produce them are only able to
78 detect and remove the more obvious errors, and involve many subjective
79 decisions. It is likely that a net spurious cooling corrupts the area-averaged
80 homogenized radiosonde data in the tropical troposphere in at least one and
81 probably both of the datasets, causing the data to indicate less warming than
82 has actually occurred.
- 83 • For tropospheric satellite data (T_2 and T_{2LT}), the primary cause of trend
84 discrepancies between different versions of the datasets is differences in
85 how the data from the different satellites are merged together.
- 86 • A secondary contribution to the differences between these datasets is the
87 difference between the diurnal adjustments that are used to account for
88 drifting measurement times. These differences in the diurnal adjustment are
89 more important for regional trends than for global trends, though regional
90 trend differences are also partly influenced by differences in merging
91 methods.

- 92 • Each tropospheric satellite dataset has strengths and weaknesses that are
93 coming into better focus. Improvements have occurred in several datasets
94 even during the drafting of this report, each moving it closer to the others,
95 suggesting that further convergence in the not-too-distant future is a strong
96 possibility.
- 97 • Comparisons between radiosonde data and satellite data for T_2 are very
98 likely to be corrupted by the excessive cooling in the radiosonde data from
99 the stratosphere which are used to help construct the radiosonde-derived T_2
100 data. Trend discrepancies between radiosonde and satellite datasets are
101 reduced by considering a multi-channel retrieval that estimates and removes
102 the stratospheric influence (T^*_G).

103

104 *Stratosphere*

105 **Despite their large discrepancies, all datasets indicate that the stratosphere has**
106 **cooled considerably over both the radiosonde era and the satellite era.**

- 107 • The largest discrepancies between datasets are in the stratosphere,
108 particularly between the radiosonde and satellite-based datasets. It is very
109 likely that the satellite-sonde discrepancy arises primarily from uncorrected
110 errors in the radiosonde data.
- 111 • There are also substantial discrepancies between the satellite datasets in the
112 stratosphere, indicating that there remain unresolved issues with these
113 datasets as well.

114

Chapter 4 recommendations

116

117 *All of the surface and atmospheric temperature datasets used in this report require*
118 *ongoing assessment to further quantify uncertainty and to identify and remove any*
119 *possible systematic biases that remain after the appropriate homogenization methods*
120 *have been applied.*

121

122 • *The diurnal cycles in both atmospheric and surface temperature need to be*
123 *accurately determined and validated to reduce uncertainties in the satellite data*
124 *due to the diurnal adjustment. Possible approaches include examining more*
125 *model or reanalysis data to check the diurnal adjustments currently in use,*
126 *concerted in situ measurement campaigns at a number of representative*
127 *locations, or operating a satellite-borne sounder in a non sun-synchronous orbit.*
128 *Information about the surface skin temperature diurnal cycle may be obtained by*
129 *studying data from existing satellites, or the upcoming Global Precipitation*
130 *Mission.*

131 • *The relative merits of different merging methods for satellite data for all relevant*
132 *layers need to be diagnosed in detail. Possible approaches include comparison*
133 *with other temperature data sources (radiosondes or IR satellites) over limited*
134 *time periods where the discrepancies between the satellite results are the greatest,*
135 *comparison with other ancillary data sources such as winds and integrated water*
136 *vapor, and comparison of trends on regional spatial scales, particularly in*
137 *regions where trends are large or well characterized by radiosonde data.*

- 138 • *The methods used to remove radiosonde inhomogeneities and their effects on*
139 *trends need to be rigorously studied. The detailed intercomparisons of the*
140 *methods used by different groups to construct satellite based climate records has*
141 *been beneficial to our understanding of these products, and similar parallel*
142 *efforts to create climate records from radiosonde data would be likely to provide*
143 *similar benefits.*
- 144 • *Possible errors in trends in spatially averaged surface temperature need to be*
145 *assessed further. On land these errors may arise from local errors due to*
146 *changes in instrumentation or local environment that do not completely cancel*
147 *when spatial averaging is performed. Over the ocean, these errors may arise*
148 *from the small number of samples available in many regions, and long-term*
149 *changes in measurement methods. For historical data, these assessments may*
150 *benefit from the recovery of additional metadata to better characterize possible*
151 *non-climatic signals.*
- 152 • *Tools and methods need to be developed to help reduce structural uncertainty by*
153 *providing methods to objectively differentiate between different datasets and*
154 *construction methods. To the extent possible, such tools should be based on*
155 *generally accepted physical principles, such as consistency of the temperature*
156 *changes at adjacent levels in the atmosphere, include physically-based*
157 *comparisons with external ancillary data, and take account of the consistency of*
158 *intermediate data generated while producing the datasets.*

159

160 **1. Background**

161

162 In the previous chapter, we have discussed a number of estimates of vertically resolved
163 global temperature trends. Different sources of data (e.g., surface measurements, vertical
164 profiles from radiosondes, and data from satellite borne sounding radiometers), as well as
165 different analysis methods applied to the same data, can yield long term (multi-decadal)
166 temperature trends that differ by as much as several tenths of a °C per decade. This is of
167 comparable magnitude to the actual climate change signal being searched for. In this
168 chapter we discuss these discrepancies in light of the observing system capabilities and
169 limitations described in Chapter 2. We note the degree to which estimates of uncertainty
170 can account for the differences in reported values for the temperature trends in given
171 layers, and differences in the trends of adjacent layers. Most of the time our focus will be
172 on the period from 1979-2004, during which atmospheric temperatures were observed
173 using multiple observing systems.

174

175 We begin our discussion in the stratosphere, and move to successively lower layers until
176 we reach the Earth's surface. We proceed in this order because the largest discrepancies
177 in trends between data sources occur in upper atmospheric layers, especially the
178 stratosphere. As mentioned in Box 2.2, when satellite-equivalent measures are made from
179 vertically resolved radiosonde data to facilitate comparisons between the two systems,
180 large stratospheric errors can significantly influence measures centered much lower in the
181 atmosphere.

182

183 **2. Uncertainty in stratospheric temperature trends**

184

185 Long-term observations of the stratosphere have been made by two observing systems:
186 radiosondes and satellite-borne sounders. On both the global and the zonally averaged
187 scale, there is considerably less variation between datasets derived from the same type of
188 observing system for this layer than between those from different observing systems.
189 This can be seen in the leftmost panel of Figure 3.5, which shows the zonally averaged
190 trends over the satellite era (1979-2004) for two radiosonde-based datasets, and two
191 satellite-based datasets. The radiosonde data ($T_{4\text{-HadAT2}}$ and $T_{4\text{-NOAA}}$) show more cooling
192 than datasets based on satellite data ($T_{4\text{-UAH}}$ and $T_{4\text{-RSS}}$), and also do not show the reduced
193 cooling in the tropics relative to the mid-latitudes that is seen in the satellite data.

194

195 *2.1 Radiosonde Uncertainty*

196 As discussed in Chapter 2, radiosonde data are plagued by numerous spurious
197 discontinuities in measured temperature that must be detected and removed in order to
198 construct a homogenized long-term record of atmospheric temperature, a task that is
199 particularly difficult in the absence of reliable metadata describing changes in
200 instrumentation or observing practice. A number of physical sources of such
201 discontinuities have larger effects in the stratosphere. The lower atmospheric pressure in
202 the stratosphere leads to reduced thermal contact between the air and the temperature
203 sensor in the radiosonde package. This in turn leads to increased errors due to daytime
204 solar heating and lags between the real atmospheric temperature and the sensor response
205 as the instrument rises through atmospheric layers with rapidly varying temperatures.

206 Such systematic errors are not important for trend studies provided that they do not
207 change over the time period being studied. In practice, as noted in Chapter 2, radiosonde
208 design, observing practices, and procedures used to attempt to correct for radiation and
209 lag errors have all changed over time.

210

211 Past attempts to make adjustments to radiosonde data using detailed physical models of
212 the instruments (Luers and Eskridge, 1998) improved data homogeneity in the
213 stratosphere, but not in the troposphere (Durre et al., 2002). Since it is important to use
214 the same methods for all radiosonde levels for consistency, scientists have tended to
215 instead use empirical methods to deduce the presence and magnitude of any suspected
216 discontinuity. Both of the homogenized radiosonde datasets used in this report make
217 these estimates using retrospective statistical analyses of the radiosonde data without
218 input from other measurements. The investigators who constructed these datasets have
219 attempted to identify and to adjust for the effects of suspected change points, either by
220 examination of station time series in isolation (NOAA), or by comparison with nearby
221 stations (UK). Both approaches can most successfully identify changes that are large and
222 step-like. While based in statistics, both these methods also include significant subjective
223 components. As a result, different investigators with nominally the same sets of
224 radiosonde data can calculate different trend estimates because of differences in
225 adjustment procedures (Free et al., 2002). The lack of sensitivity to small or gradual
226 changes may bias the resulting homogenized products if such changes are numerous and

227 predominantly of one sign or the other¹. The relative frequency of large step-like changes
228 and smaller changes that may be statistically indistinguishable from natural variability
229 remains an open question.

230

231 Since the adjustments needed to remove the resulting discontinuities tend to be larger for
232 the stratosphere than for lower levels (Parker et al., 1997; Christy et al., 2003; Lanzante
233 et al., 2003), the uncertainty associated with the homogenization procedures is very likely
234 to be larger in the stratosphere than at lower levels, as has been shown for the UK
235 radiosonde dataset (Thorne et al., 2005). The best estimate of the size of this source of
236 uncertainty is obtained by comparing the statistics (e.g., the trends) from the two adjusted
237 radiosonde datasets that are currently available. However, the UK group analysis is partly
238 based upon the NOAA dataset, so we may be under-estimating the uncertainty. Only
239 through increasing the number of independently produced datasets under different
240 working assumptions can we truly constrain the uncertainty (Thorne et al., 2005).

241

242 Differences in trends between daytime and nighttime observations in the uncorrected
243 radiosonde data used in constructing the NOAA and UK radiosonde datasets, suggest that
244 the biases caused by solar heating² have been reduced over time, leading to a spurious
245 cooling trend in the raw daytime data (Sherwood et al., 2005). Many of the changes in

¹ It is speculated that gradual changes could result from the same changes in instrumentation or practices that cause the step like changes, provided that these changes are implemented gradually (Lanzante et al., 2003).

² For some types of radiosondes, radiation adjustments based on information provided by the manufacturer are made as part of routine processing of radiosonde data by the observing station. The findings cited here refer to data that has already had these corrections performed. The reduction in daytime biases is likely to be due to a combination of improvements in instrument design, and improvements in the radiation adjustment procedure.

246 observing practice will affect both day and night time observations; e.g., a change in
247 practice may yield a spurious 0.5°C daytime cooling and 0.4°C night time cooling, so
248 day-night differences cannot be used in isolation to correct the observations. Whether the
249 NOAA and UK methods have successfully removed day-night and other effects, or if
250 sufficiently targeted are capable of doing so, is a matter for ongoing research. Randel and
251 Wu (2005) have shown for a subset of tropical stations in the NOAA dataset, there is
252 strong evidence for step-like residual cooling biases following homogenization, which
253 will cause a spurious cooling in the tropical area-averaged NOAA time series considered
254 here. They find that the effect is not limited to daytime launches, as would be expected
255 from discussions above, and that it is likely to affect at least the upper-troposphere as
256 well as the stratosphere. Finally, the balloons that carry the instruments aloft have
257 improved over time, so they are less likely to burst at high altitudes or in extreme cold.
258 This could also lead to a warm sampling bias within the stratosphere in early radiosondes
259 which has gradually ameliorated with time, introducing a spurious stratospheric cooling
260 signal (Parker et al., 1997). Taken together these results imply that any residual
261 systematic errors in the homogenized radiosonde products will likely lead to a spurious
262 cooling bias.

263

264 Since the radiosonde stations selected for inclusion in the homogenized datasets do not
265 cover the entire globe, there can be a bias introduced in to the global mean trend
266 depending on the locations of the chosen stations. On a global scale, this bias has been
267 estimated to be less than 0.02°C/decade for T_4 by sub-sampling globally complete

268 satellite or reanalysis datasets at the station locations³, and thus it is not an important
269 cause of the differences between the datasets on large spatial scales (Free and Seidel,
270 2004). Though they have not been explicitly calculated, sampling errors are likely to be
271 more important for the zonal radiosonde trends plotted in Figure 3.5, and may account for
272 some of the zone-to-zone variability seen in the radiosonde data in that figure that is not
273 duplicated in the smoother satellite data. The sampling effects also permeate in the
274 vertical – above 100hPa there is a significant reduction in the number of valid
275 measurements whereas below this level the number of measurements is relatively stable.
276 Because the trends vary with height, this can lead to errors, particularly when calculating
277 satellite-equivalent measures.

278

279 *2.2 Satellite Uncertainty*

280

281 The two satellite-based stratospheric datasets (T_{4-UAH} and T_{4-RSS}) have received
282 considerably less attention than their tropospheric counterparts (see section 4.3 below),
283 though they differ in estimated trend by roughly the same absolute amount
284 ($\sim 0.1^{\circ}\text{C}/\text{decade}$) as the corresponding tropospheric datasets produced by the same
285 institutions. However the importance of the differences is perceived to be much less
286 because the trend is much larger (a cooling over 1979-2004 of approximately 0.8°C). A
287 detailed comparison of the methods used to construct the two datasets has not yet been
288 performed. Despite the lack of such a study, it is very likely that in the stratosphere, like
289 the troposphere (discussed in section 4.3), structural uncertainty is the most important

³ This estimate is valid for the NOAA dataset and a previous version of the UK dataset. The estimated bias increases to about 0.05K for a tropical average. In the cited work the tropics were defined to be 30S to 30N – we would expect the sampled error to be a few hundredths of a degree per decade larger for the 20S to 20N definition of the tropics used in this report.

290 source of uncertainty. Two important types of structural uncertainty are likely to
291 dominate: those associated with the method of correcting for drifts in diurnal sampling
292 time, and those associated with the method of correcting calibration drifts associated with
293 the temperature of the hot calibration target. Section 3 discusses how these uncertainty
294 sources are treated in the troposphere.

295

296 Despite unresolved problems in the satellite datasets, the similarity of the satellite
297 measurement and homogenization methods suggest that the satellite measurements of the
298 stratosphere are no more uncertain than those of the mid-troposphere, where satellites and
299 radiosondes are in much closer agreement. This assessment, coupled with the evidence
300 presented above that residual artificial cooling is likely to exist in the stratospheric
301 radiosonde data, particularly in the tropics, implies that the discrepancy between
302 radiosondes and satellite estimates of stratospheric trends (see Table 3.3) during the
303 satellite era is very likely to be mostly due to uncorrected biases in the radiosonde
304 measurements.

305

306 **3. Uncertainty in tropospheric trends**

307 In contrast to the stratosphere, differences in reported tropospheric trends from the same
308 type of measurement are as large as or larger than differences in trends reported from
309 different data sources. This can be seen in Figure 3.5 and Tables 3.3 and 3.4. Also note
310 that the radiosonde data for the two tropospheric layers show the same general north-
311 south pattern (i.e. more temperature increase in the mid-latitudes than at the poles or in
312 the tropics) as the satellite data, in contrast to the stratospheric results.

313

314 *3.1 Radiosonde uncertainty*

315 The main sources of error in tropospheric radiosonde trends are similar to those
316 encountered in the stratosphere. The challenge is to assess to what extent these types of
317 errors, which in the stratosphere likely result in artificial cooling even in homogenized
318 datasets, extend down into the troposphere. Another important issue is that when
319 performing calculations to directly compare radiosonde data with satellite trends for the
320 T₂ layer, the contribution of errors in the stratospheric trends to the results for this layer
321 become important, since 10% to 15% of the weight for this layer comes from the
322 stratosphere.

323

324 *3.1.1 Removing non-climatic influences.*

325 There are several pieces of evidence that suggest that any residual bias in tropospheric
326 radiosonde data will be towards a cooling. First, the more obvious step-like
327 inhomogeneities that have been found tend to predominantly introduce spurious cooling
328 into the raw time series, especially in the tropics. This suggests that any undetected
329 change points may also favor spurious cooling (Lanzante et al., 2003). Second, solar-
330 heating-induced errors, while largest in the stratosphere have been found to bias daytime
331 measurements to higher temperatures at all levels, particularly in the tropics. Periodic
332 radiosonde intercomparisons (most recently at Mauritius in Feb. 2005) undertaken under
333 the auspices of WMO imply that the magnitude of these errors has been reduced over
334 time, and that radiosondes from independent manufacturers have become increasingly

335 similar (and presumably more accurate) over time⁴ (Silveira et al., 2003; Pathack et al.,
336 2005). If these effects have on average been uncorrected by the statistical procedures
337 used to construct the homogenized radiosonde datasets discussed in this report, they
338 would introduce an artificial cooling signal into the radiosonde records. Of course on an
339 individual station basis the picture is likely to be much more ambiguous and many
340 stations records, even following homogenization efforts, are likely to retain large residual
341 warm or cold biases. But on average, the evidence outlined above suggests that if there is
342 a preferred sign it is likely to be towards a residual cooling. It is important to stress that to
343 date the quantitative evidence to support such an argument, at least away from a small
344 number of tropical stations (Randel and Wu, 2005), is at best ambiguous.

345

346 *3.1.2 Sampling uncertainty*

347 The fact that most radiosonde data are primarily collected over Northern Hemispheric
348 land areas naturally leads to uncertainties about whether or not averages constructed from
349 radiosonde data can faithfully represent global trends. However, (Wallis, 1998) and
350 (Thorne et al., 2005) show that stations can be representative of much larger scale
351 averages above the boundary layer, particularly within the deep tropics. Spatial and
352 temporal sampling errors for the radiosonde datasets have been assessed by sub-sampling
353 trends in reanalyses or satellite data at the locations of radiosonde stations used in the
354 production of global datasets, and comparing the results to the full global average of the
355 reanalysis or satellite data (Free and Seidel, 2004). Typically, errors of a few hundredths
356 of a °C per decade have been estimated for global averages, too small to fully account for

⁴ These intercomparisons provide a source data about the differences between different type of sondes that have not yet been used to homogenize sonde data.

357 the differences between radiosonde and satellite trends, though it has been suggested that
358 the existing sampling could lead to a warm bias in the radiosonde record (Agudelo and
359 Curry, 2004). As is the case for the stratosphere, sampling errors may be part of the cause
360 for the zone to zone variability seen in the radiosonde data. Residual differences between
361 two radiosonde dataset global means are assessed to be approximately equally caused by
362 sampling error, choice of raw data, and choice of adjustments made⁵.

363

364 *3.1.3 The influence of uncertainty in stratospheric measurements*

365 To compare data that represent identical layers in the atmosphere, “satellite-equivalent”
366 radiosonde data products are constructed using a weighted average of radiosonde
367 temperatures at a range of levels (see Box 2.2). The T₂ radiosonde datasets are
368 constructed to match the weighting function for Microwave Sounding Unit (MSU)
369 channel 2. Since 10% to 15% of the weight for this channel comes from the stratosphere
370 (see Figure 2.1), it is important to keep in mind the suspected relatively large errors in the
371 stratospheric measurements made by radiosondes. It is possible that stratospheric errors
372 could cause the trends in the radiosonde-derived T₂ to be as much as 0.05°C/decade too
373 cool, particularly in the tropics, where the suspected stratospheric errors are the largest
374 (Randel and Wu, 2005) and therefore have a large impact on area-weighted averages.
375 This error source may be partly eliminated by considering the multi-channel tropospheric
376 retrievals discussed in section 5 below.

377

⁵ This comparison was made using a previous version of the UK dataset (HadRT), which uses a different set of stations than the current version. This difference is very unlikely to substantially alter these conclusions.

378 *3.2. Satellite uncertainty*

379 Satellite-derived temperature trends in the middle and upper troposphere have received
380 considerable attention. In particular, the causes of the differences between T_{2-UAH} and T_{2-}
381 RSS have been examined in detail; less work has been done concerning $T_{2-U.Md.}$ because
382 this dataset is newer. There are two potentially important contributions to the residual
383 uncertainty in satellite estimates of global trends for the satellite-based datasets: (1)
384 corrections for drifts in diurnal sampling, and (2) different methods of merging data from
385 the series of different satellites.

386

387 *3.2.1 Diurnal Sampling Corrections*

388 During the lifetime of each satellite, the orbital parameters tend to drift slowly with time.
389 This includes both a slow change of the local equator crossing time (LECT), and a decay
390 of orbital height over time due to drag by the upper atmosphere. The LECT is the time at
391 which the satellite passes over the equator in a northward direction. Changes in LECT
392 indicate corresponding changes in local observation time for the entire orbit. Because the
393 temperature changes with the time of day (e.g., the cycle of daytime heating and
394 nighttime cooling), slow changes in observation time can cause a spurious long-term
395 trend. These diurnal sampling effects must be estimated and removed in order to produce
396 a climate-quality data record.

397

398 The three research groups that are actively analyzing data from microwave satellite
399 sounders first average together the ascending and descending orbits, which has the effect
400 of removing most of the first harmonic of the diurnal cycle. For the purposes of this

401 report, “diurnal correction” means the removal of the second and higher harmonics. Each
402 group uses a different method to perform the diurnal correction.

403

404 The UAH group calculates mean differences by subtracting the temperature
405 measurements on one side of the satellite track from the other (Christy et al., 2000). This
406 produces an estimate of how much, on average, the temperature changes due to the
407 difference in local observation times from one side of the satellite swath to another,
408 typically about 40 minutes. This method has the advantage of not relying on data from
409 other sources to determine the diurnal cycle, but it has been shown to be sensitive to
410 satellite attitude errors (Mears and Wentz, 2005), and is too noisy to produce a diurnal
411 adjustment useable on small spatial scales.

412

413 The RSS group uses hourly output from a climate model in a microwave radiative
414 transfer model to estimate the diurnal cycle in brightness temperature at each grid point in
415 the satellite dataset (Mears et al., 2003). This method has the advantage that a diurnal
416 adjustment can be made at the data resolution. However, it is likely that the climate
417 model-based adjustment contains errors, both because models are often unable to
418 accurately represent the diurnal cycle (Dai and Trenberth, 2004), and because the
419 parameterization of the ocean surface temperature used as a lower boundary for the
420 atmospheric model used does not include diurnal variability. The model has been shown
421 to represent the first harmonic of the diurnal cycle for MSU channel 2 with less than 10%
422 error, but less is known about the accuracy of the second and higher harmonics that are
423 more important for adjusting for the diurnal sampling errors (Mears et al., 2003).

424

425 Both groups use their diurnal cycle techniques to adjust the satellite data before merging
426 the data from the different satellites. In contrast, the Maryland group averaged the
427 ascending and descending satellite data to remove only the first harmonic in the diurnal
428 cycle before merging, and used a fitting procedure to account for both the first and
429 second harmonic diurnal components when performing the trend analysis *after* merging
430 the data from different satellites (Vinnikov and Grody, 2003; Vinnikov et al., 2005).

431 Since they only accounted for the first harmonic diurnal component during the merging
432 of satellite data, errors in the diurnal cycle can cause errors in the data analysis following
433 the merging procedure. However, the removal of the diurnal cycle before merging may
434 also introduce some error into UAH and RSS merging procedures if the assumed diurnal
435 cycle is inaccurate, but physically, the removal of the diurnal harmonics before merging
436 seems to be a more logical approach as the diurnal harmonics will tend to add noise
437 unless removed.

438

439 On a global scale, the total impact of the diurnal correction applied by the RSS and UAH
440 groups to the microwave sounding data for the RSS data is to increase the decadal trend
441 by about 0.03°C/decade for T₂ (Christy et al., 2003; Mears et al., 2003). The impact of the
442 Maryland group's adjustment is almost negligible. For the RSS T₂ data, when a diurnal
443 correction is applied that is 50% or 150% as large as the best estimate, these adjustments
444 significantly worsen the magnitude of the intersatellite differences. Changes of this
445 magnitude in the diurnal cycle lead to temperature trends that differ by 0.015°C; so we
446 estimate that the uncertainty in trends due to uncertainty in the diurnal correction is about

447 0.015°C/decade for T_2 . The UAH group estimates that the diurnal correction for T_2 is
448 known to 0.01°C/decade (Christy et al., 2000). These estimates of residual uncertainty are
449 relatively small, and are considerably less than the structural uncertainties associated with
450 the satellite merging methodology described in the next section. Despite the global
451 agreement for the diurnal adjustment for the RSS and UAH results, significant
452 differences in the adjustments exist as a function of location (Mears and Wentz, 2005),
453 which may explain some of the difference on smaller spatial scales between these two
454 datasets⁶.

455

456 3.2.2 Satellite merging methodology

457

458 It is very likely that the most important source of uncertainty in microwave sounding
459 temperature trends is due to inter-satellite calibration offsets, and calibration drifts that
460 are correlated with the temperature of the calibration target (Christy et al., 2000; Mears et
461 al., 2003). When results from supposedly identical co-orbiting satellites are compared,
462 intersatellite offsets are immediately apparent. These offsets, typically a few tenths of a
463 °C, must be identified and removed or they will produce errors in long-term trends of
464 several tenths of a °C per decade. When constant offsets are used to remove the inter-
465 satellite differences, the UAH group found that significant differences still remain that
466 are strongly correlated with the temperature of the calibration target⁷ (Christy *et al.*,
467 2000). This effect has since been confirmed by the RSS group (Mears *et al.*, 2003). Both

⁶ See for example Figure 3.5 versus Figure 4.3.

⁷ The calibration target can change temperature by tens of °C over the course of the life of the satellite due to orbit- and season-dependent solar heating.

468 the UAH and RSS groups remove the calibration target temperature effect using a model
469 that includes a constant offset for each satellite, and an additional empirical “target
470 factor” multiplied by the calibration target temperature.

471

472 Despite the similarity in methods, the RSS and UAH groups obtain significantly different
473 values for the global temperature trends (see Table 3.3). In particular, the difference
474 between the trends for T_2 has received considerable attention. A close examination of the
475 procedures suggests that about 50% of the discrepancy in trends is accounted for by a
476 difference between the target factor for the NOAA-09 instrument deduced by the two
477 groups. This difference mainly arises from the subsets of data used by the two groups
478 when determining the satellite merging parameters (i.e., offsets and target factors). The
479 UAH group emphasizes pairs of satellites that have long periods of overlap, and thus uses
480 data from six pairs of satellites, while RSS uses all available (12) overlapping pairs of
481 satellites. Most of the remainder of the difference is due to a smaller difference in the
482 calibration target temperature proportionality constant for NOAA-11, and to small
483 differences in the diurnal correction. Both these differences primarily affect the
484 measurements made by NOAA-11 and NOAA-14, due to their large drifts in local
485 measurement time.

486

487 In Figure 4.1a, we plot the difference ($T_{2-RSS} - T_{2-UAH}$) between the RSS and UAH time
488 series. There is an obvious step that occurs in 1986, near the end of the NOAA-09
489 observation period, and a gradual slope that occurs during the observation periods of
490 NOAA-11 and NOAA-14. Note that the trend difference between these two datasets is

491 statistically significant at the 1% level, even though the error ranges quoted in Table 3.3
492 overlap, due to the presence of nearly identical short term fluctuations in the two datasets
493 (see Appendix A for more details).
494

495

496 Figure 4.1 (a) Time series of the difference between global averages of satellite-derived T_2 datasets. Both
497 the RSS and UMD datasets show a step-like feature relative to the UAH dataset during the lifetime of
498 NOAA-09. The difference between the RSS and the UAH datasets shows a slow drift during the NOAA-11
499 and NOAA-14 lifetimes. Both these satellites drifted more than 4 hours in observations time. (b) Time
500 series difference between global averages of satellite derived T_{2LT} datasets. A slow drift is apparent during
501 the lifetime of NOAA-11, but the analysis during the NOAA-14 lifetime is complicated because the T_{2LT} -
502 RSS dataset does not include data from the AMSU instruments on NOAA-15 and NOAA-16, while the T_{2LT} -
503 UAH dataset does. All time series have been smoothed using a Gaussian filter with width = 7 months.
504

505 The Maryland group data set ($T_{2-U.Md.}$), in its most recent version (Grody et al., 2004;
506 Vinnikov et al., 2005), implemented a more detailed, physically based error model to

507 describe the errors that correlated with a nonlinear combination of the observed
508 brightness temperature measurements and the warm target temperature used for
509 calibration⁸. They use a substantially different merging procedure to deduce values of the
510 parameters that describe the intersatellite differences. First, as noted above, only the first
511 harmonic diurnal component is accounted for during the satellite merging, possibly
512 causing errors in the retrieved parameters. Second, they only use the spatial variation
513 seen by the different MSU instruments to derive the calibration adjustments and perform
514 long-time-scale temporal averaging of the measured temperatures to reduce the noise in
515 the overlapping satellite measurements. This averaging procedure may attenuate the time
516 dependent signal that the UAH empirical error model was introduced to explain. The
517 large step in the $T_{2-U.Md.} - T_{2-UAH}$ difference time series that occurs in 1986 (see Figure
518 4.1a) suggests that uncertainty in the parameters for the NOAA-09 satellite are also
519 important for this dataset⁹. The cause of the large fluctuations in the difference during the
520 2000-2004 time period is not known, but may be related to the absence of AMSU data in
521 the $T_{2-U.Md.}$ dataset. Due to its relatively recent appearance, considerably less is known
522 about the reasons for the differences between the Maryland dataset and the RSS and
523 UAH datasets, thus the comments about these differences should be viewed as more
524 speculative than the statements about the RSS-UAH differences.

⁸ The Maryland group accounted for uncertainties in the radiometers non-linearity parameter as well as errors in the warm target radiation temperature (due to uncertainties in its emissivity and physical temperature) and errors in the cold space radiation temperature (due to uncertain antenna side lobe contributions for example). However, while all of these error sources are accounted for, they are assumed to be constant during the lifetime of a given instrument and thus do not take into account the possibility of contributions to the side lobe response from the earth or warm parts of the satellites whose temperature varies with time. These error sources lead, when globally averaged and linearized, to an expression where the target temperature is the most important factor. Thus while the exact physical cause of the observed effect is not known precisely, it is possible to accurately model and remove it on a global scale from the data using either method

⁹ The trend in this difference time series is statistically significant at the 1% level.

525

526 These differences are an excellent example of structural uncertainty, where identical
527 input data and three seemingly reasonable methodologies lead to trends that differ
528 significantly more than the amount expected given their reported internal uncertainties.
529 Since methodological differences yield data products showing differences in trends in T_2
530 of about 0.1 °C per decade, it is clear that the most important source of uncertainty for
531 satellite data are structural uncertainties and that these need to be included in any overall
532 uncertainties assessed for tropospheric temperature trends and lapse rates.

533

534 *3.2.3 Differences in spatial pattern.*

535 Only T_{2-UAH} and T_{2-RSS} have provided gridded results. Maps of gridded trends for these
536 products are shown in Fig 4.2, along with a map of the difference between the trends. The
537 overall pattern in the trends is very similar between the two datasets, aside from
538 difference in the globally averaged trends. Differences in the latitude dependence are due
539 to the use of zonally varying intersatellite offsets in the construction of T_{2-UAH} (in contrast
540 to the constant offsets in T_{2-RSS}) and to differences in the applied diurnal adjustment as a
541 function of latitude. Other differences may be caused by the spatial smoothing applied to
542 the T_{2-UAH} during the construction of the data set, and to differences in spatial averaging
543 performed on the diurnal adjustment before it was applied. This last difference will be
544 discussed in more detail in section 4.4 below because the effects are more obvious for the
545 T_{2LT} layer.

546

547

548 Figure 4.2 Global maps of trends from 1979-2004 for (a) T_{2-UAH} and (b) T_{2-RSS} . Except for an overall
549 difference between the two results, the spatial patterns are very similar. A map of the difference $T_{2-UAH} -$
550 T_{2-RSS} between trends for the two products shown in (c) reveals more subtle differences in the trend.
551
552

553 **4 Uncertainty in Lower Tropospheric Trends**

554

555 *4.1 Radiosonde Uncertainty*

556 Uncertainties in lower tropospheric trends measured by radiosondes are very similar to
557 those discussed above for the middle-upper troposphere. The most important difference is
558 that when comparing to the T_{2LT} satellite product, the contribution of the stratospheric
559 radiosonde trends, which is suspected to be erroneous to some extent, is substantially less
560 than for the T_2 data records. This decreases the likelihood that T_{2LT} data products
561 constructed from radiosonde data are biased toward excess cooling. However, it is
562 possible that undetected negative trend bias remains in all tropospheric levels (see section
563 3.1 above for more details), so radiosonde trends may still be biased cold.

564

565 *4.2 Satellite Uncertainty*

566 Currently, there are two lower tropospheric satellite data records, $T_{2LT-UAH}$ and $T_{2LT-RSS}$.
567 As discussed in the Preface, both datasets are relatively recent, thus little is known about
568 the specific reasons for their differences. Because of the noise amplification effects of the
569 differencing procedure¹⁰ used to construct the data record, the merging parameters tend to
570 be more sensitive to the methods used to deduce them. A number of different methods
571 were explored in the creation of $T_{2LT-RSS}$, leading to an estimate of the structural
572 uncertainty of 0.08°C/decade for global trends. When combined with internal uncertainty,
573 the estimated total global trend uncertainty for this dataset is 0.09°C/decade (Mears and
574 Wentz, 2005). Note that the difference between the global trends for $T_{2LT-RSS}$

¹⁰ The T_{2LT} datasets are constructed by subtracting 3 times the average temperature measured by the outermost 4 (near-limb) views from 4 times the average temperature measured by the 4 adjacent views, which are closer to nadir. This has the effect of removing most of the stratospheric signal, and moving the effective weighting function lower in the troposphere (Spencer and Christy 1992). Assuming that the errors in each measurement are uncorrelated, this has the effect of amplifying these errors by a factor of about 5 relative to T_2 (Mears and Wentz 2005). Even if some of the error is correlated between view, this argument still applies to the uncorrelated portion of the error.

575 (0.189°C/decade) and $T_{2LT-UAH}$ (0.115°C/decade) shown in
576 Table 3.3 is less than this estimated uncertainty. The estimated global trends in the
577 radiosonde datasets are also within the $T_{2LT-RSS}$ error range. In Figure 4.1b we plot the
578 difference ($T_{2LT-RSS} - T_{2LT-UAH}$) between the RSS and UAH time series. This time series
579 shows more variability than the corresponding T_2 difference time series, making it more
580 difficult to speculate about the underlying causes of the differences between them. The
581 step-like feature during the 1985-1987 period is less obvious, and while there appears to
582 be a slow drift during the NOAA-11 lifetime, a corresponding drift during the NOAA-14
583 lifetime is less obvious, perhaps because the RSS data do not yet include data from the
584 more recent AMSU satellites. We speculate that the drift during NOAA-11 is in part due
585 to differences in the diurnal correction applied. The UAH diurnal correction is based on
586 a parameterization of the diurnal cycle which is constrained by measurements made
587 during a time period with 3 co-orbiting satellites, , while RSS uses a model-based diurnal
588 correction analogous to that used for TMT.

589

590 In Figure 4.3, we show global maps of the gridded trends for $T_{2LT-UAH}$ and $T_{2LT-RSS}$, along
591 with a map of the trend differences. The spatial variability in the trend differences
592 between the two datasets is much larger than the variability for T_2 , though both datasets
593 show similar patterns in general, with the greatest temperature increase occurring in the
594 Northern Hemisphere, particularly over Eastern Asia, Europe, and Northern Canada. The
595 two datasets are in relatively good agreement north of 45°N latitude. In the tropics and
596 subtropics, the largest differences occur over land, particularly over arid regions.

597

598 Figure 4.3 Global maps of trends from 1979-2004 for (a) $T_{2LT-UAH}$ and (b) $T_{2LT-RSS}$. Except for an overall
599 difference between the two results, the spatial patterns are similar. A map of the difference $T_{2LT-UAH} - T_{2LT-}$
600 RSS between trends for the two products shown in (c) shows that the largest differences are over tropical and
601 subtropical land areas. Data from land areas with elevation higher than 2000m are excluded from the T_{2LT-}
602 RSS dataset and shown in white.
603

604

605 We speculate that this may be in part due to differences in how the diurnal adjustment is
606 done by the two groups. The UAH group applies an averaged diurnal adjustment for each

607 zonal band, based on different adjustments used for land and ocean. The RSS group uses
608 a grid-point resolution diurnal correction. The UAH method may lead to errors for
609 latitudes where the diurnal cycle varies strongly with longitude. More arid regions (e.g.,
610 subtropical Africa), which typically have much larger surface diurnal cycles, may be
611 under-adjusted when the zonally averaged correction is applied, leading to long-term
612 trends that are too low. Correspondingly more humid regions and oceans may be over-
613 adjusted, in some cases making up for the overall difference between the two datasets,
614 perhaps accounting for the good agreement in regions such as Southeast Asia, Southern
615 India, and Northern South America. Further analysis is required using a range of
616 alternative diurnal correction estimation techniques for definitive conclusions to be
617 reached. Other differences, such the north-south streaking seen in the RSS data, may be
618 caused by differences in spatial smoothing, and by the inclusion of AMSU data in T_{2LT-}
619 UAH , but not in $T_{2LT-RSS}$.

620

621 The decay of orbital height over each satellite's lifetime can cause substantial errors in
622 satellite-derived T_{TLT} because changes in height lead to changes in the earth incidence
623 angles for the near-limb observations used to construct the data record (Wentz and
624 Schabel, 1998). Both the RSS and UAH groups correct for this error by calculating the
625 expected change in observed temperature as a function of incidence angle, and then using
626 this estimate to remove the effect of orbital decay. The straight-forward method used to
627 make these corrections, combined with its insensitivity to assumptions about the vertical
628 structure of the atmosphere, leads to the conclusion that errors due to orbital decay have

629 been accurately removed from both datasets and are not an important cause of any
630 differences between them.

631

632 *4.3 Comparison between satellite and well characterized radiosonde stations*

633 Point-by-point comparisons between radiosonde and satellite data eliminate many
634 sources of sampling error normally present in radiosonde data. Also, since uniform global
635 coverage is less important when using radiosondes to validate satellite data locally,
636 stations can be chosen to minimize the contribution due to undocumented changes in
637 radiosonde instrumentation or observing practice. For instance, if one restricts
638 comparisons of the satellite and radiosonde data to 29 Northern Hemisphere radiosonde
639 stations that have consistently used a single type of instrumentation (the Viz sonde) since
640 1979, the average difference between these radiosonde trends and $T_{2LT-UAH}$ trends since
641 1979-2004 is only 0.03°C/decade (Christy et al., 2003; Christy et al., 2005). Similarly,
642 when this set of radiosondes is extended to include a set of Southern Hemisphere stations
643 where instrument changes were well documented, agreement between $T_{2LT-UAH}$ and
644 radiosonde trends is almost as good (Christy and Norris, 2004; Christy et al., 2005). This
645 suggests that, for the T_{2LT} layer, where the stratospheric problems with radiosonde data
646 are minimized, some level of corroboration can be attained from these two diverse
647 measurement systems.

648

649 **5 Multi-channel retrievals of tropospheric temperature.**

650 As mentioned above, the single channel satellite measurements commonly identified as
651 tropospheric temperature (T_2) are impossible to interpret as solely tropospheric

652 temperatures because 10% to 15% (seasonally and latitudinally varying) of the signal
653 measured by MSU channel 2 arises from the stratosphere. In principle, it is possible to
654 reduce the stratospheric contribution to Channel 2 by subtracting out a portion of the
655 stratospheric Channel 4, though the exact values of the weights used in this procedure are
656 controversial (see Chapter 2 for more details). Despite this controversy, there is little
657 doubt that the resulting trends are more representative of the troposphere than the T_2
658 datasets. The reduction in stratospheric signal also reduces the difference between trends
659 in the satellite data and the radiosonde data (see Table 3.3), because the error-prone
660 stratospheric levels in the stratosphere have reduced (but still non-zero) weight.

661

662 The existence of a stratosphere-corrected tropospheric retrieval allows tests for
663 consistency of temperature trends among the different datasets constructed by a research
664 group for different atmospheric layers. One test, when applied to an earlier version (v5.1)
665 of the UAH global average trends, did not prove inconsistency on the global scale,
666 because the difference between the $T_{2LT-UAH}$ trend and the retrieval-calculated T_{2LT} trend
667 was well within the published margin of error. However, a clearer inconsistency was
668 found for the tropics (Fu and Johanson, 2005; Johanson and Fu, 2005). In this case, the
669 difference between the retrieval-calculated trend and $T_{2LT-UAH}$ trend was larger than its
670 estimated error range, an indication of uncharacterized error in at least one of the UAH
671 products, or more generally that $T_{2LT-UAH}$, T_{2-UAH} and T_{4-UAH} were not strictly self-
672 consistent as a set. This inconsistency is now resolved (within error estimates) with the
673 introduction of a new version of the $T_{2LT-UAH}$ dataset. The RSS versions of the T_2 , T_4 and
674 T^* datasets were found to be consistent for both global and tropical averages (Fu and

675 Johanson, 2005). The trends in the RSS version of the TLT dataset (produced after Fu
676 and Johanson was submitted) is also consistent with the other RSS based datasets.

677

678 **6 Uncertainty in Surface Trends.**

679

680 *6.1 Sea surface temperature uncertainty*

681 Temperature analyses over the ocean are produced from sea surface temperatures (SST)
682 instead of marine air temperatures. This is because marine air temperatures are biased
683 from daytime ship deck heating (Folland and Parker, 1995; Rayner et al., 2003) and
684 because satellite observations are available for SST beginning in November 1981 to
685 augment *in situ* data (Reynolds and Smith, 1994). Spatially complete analyses of SSTs
686 can be produced by combining satellite and *in situ* data (from ships and buoys) (Reynolds
687 et al., 2002; Rayner et al., 2003), from *in situ* data alone (Smith and Reynolds, 2004), or
688 from satellite data alone (Kilpatrick et al., 2001).

689

690 *6.1.1 Satellite SST uncertainties*

691 Climate comparison analyses based on infra-red satellite data alone are not useful
692 because of possible large time-dependent biases. These biases have typically occurred
693 near the end of a satellite's life time when the instrument no longer works properly, or
694 during periods when assumptions made about the atmospheric profile in the satellite
695 algorithm are no longer valid, e.g., during periods immediately following volcanic
696 eruptions, when a large amount of dust from the eruption is present in the stratosphere
697 (Reynolds, 1993; Reynolds et al., 2004). These problems may be partially mitigated in

698 the future by use of the microwave SST sensors that became available starting with the
699 launch of the Tropical Rainfall Measuring Mission (TRMM) in 1987 (Wentz et al.,
700 2000), but these microwave SST data have not been available long enough to derive
701 meaningful trends, and are difficult to calibrate absolutely due to various instrument
702 related problems (Wentz et al., 2001; Gentemann et al., 2004). Thus, analyses now use
703 multiple satellite instruments blended with or anchored to *in situ* data that reduce the
704 overall analysis errors (e.g., Reynolds et al., 2002, Rayner et al., 2003).

705

706 *6.1.2 In Situ SST uncertainties*

707 As discussed in Chapter 2, the primary sources of uncertainty in *in situ* SST
708 measurements are non-climatic signals caused by changes in the mix of instrumentation
709 over time and sampling errors. Over time the measurements have typically evolved from
710 insulated bucket measurements to engine intake, through hull, and buoy mounted sensors
711 – these changes are not necessarily accurately recorded in the metadata Both non-climatic
712 signals and sampling error are thought to be largest in sparsely sampled regions, such as
713 the southern oceans, where a single erroneous or unrepresentative measurement could
714 bias the average for an entire measurement cell for the month in question. Both types of
715 errors have been calculated for the ERSST dataset and included in the quoted error range
716 (see Figure 4.4).

717

718 Figure 4.4. SST, Land Surface Air Temperature, and the Combined Temperature Data Record anomaly
719 averaged annually and between 60°S and 60°N (purple), with its estimated 95% confidence intervals
720 (dashed). Data are from the NOAA GHCN-ERSST dataset (Smith and Reynolds 2005).
721

722 *6.2 Land surface air temperature uncertainty*

723 The three surface temperature analyses exhibit similar warming rates since 1958. As the
724 surface data sets have many stations in common, they are not totally independent.

725 However, the MSU series take identical input, and radiosonde datasets have common
726 data also, so this problem is not unique to the surface records. The fact that the range in
727 trends is much smaller for the surface datasets than for these other datasets implies that
728 the structural uncertainty arising from dataset construction choices is much smaller at the
729 surface, in agreement with the arguments made in Thorne et al. (2005a). Also, a number
730 of studies e.g., (Peterson et al., 1999; Vose et al., 2004) suggest that long-term, large-
731 scale trends are not particularly sensitive to variations in choice of station networks. But
732 because most land networks were not designed for climate monitoring, the data contain
733 biases that dataset creators address with different detailed methods of analysis. The
734 primary sources of uncertainty from a land-surface perspective are (a) the construction
735 methods used in the analyses and (b) local environmental changes around individual
736 observing stations that may not have been addressed by the homogeneity assessments.

737

738 Because the stations are not fully representative of varying-within-area land surface,
739 coastal, and topographical effects, global data sets are produced by analyzing deviations
740 of temperature from station averages (anomalies) as these deviations vary more slowly
741 with a change in location than the temperatures themselves (Jones et al., 1997). Random
742 errors in inhomogeneity detection and adjustments may result in biased trend analyses on
743 a grid box level. However, on the relatively large space scales of greatest importance to
744 this Report, such problems are unlikely to be significant in current data sets in the period
745 since 1958 except where data gaps are still serious, e.g., in parts of central Africa, central
746 South America, and over parts of Antarctica. Note that for the contiguous United States,
747 the period 1958-2004 uses the greatest number of stations per grid box anywhere on the

748 Earth's land surface, generally upwards of 20 stations per grid box. For regions with
749 either poor coverage or data gaps, trends in surface air temperature should be regarded
750 with considerable caution, but do not have serious effects on the largest of scales as most
751 of the spatial variability is well sampled.

752

753 Local micro-climatological environmental changes around observing stations may be
754 problematic, particularly if a similar change occurred near many observing stations (e.g.,
755 Davey and Pielke, 2005). For instance, urbanization may have increased temperatures in
756 many locations. Numerous investigators have used a variety of approaches to study these
757 effects and most have shown that any bias is likely to be small in comparison to the
758 warming signal for large-scale means (e.g., Peterson, 2003; Parker, 2004). To insure that
759 potential urbanization effects do not impact analyses, NASA adjusts the data from all
760 urban stations so that their long-term trends are consistent with those from neighbouring
761 rural stations (Hansen et al., 2001). It is generally accepted that local biases in trends
762 mostly cancel through the use of many stations or ocean observations. Because such a
763 cancellation has not been rigorously proved, partly due to the lack of adequate metadata,
764 it is conceivable that systematic changes in many station exposures of a similar kind may
765 exist over the land during the last few decades, which may give biases in trends of one
766 sign over large land regions.

767

768 *6.3 Combined land-ocean analyses uncertainty.*

769 Global combined surface temperature products are computed by combining ocean and
770 land gridded datasets. The latest version of the UK surface dataset, HadCRUT2v, (Jones
771 and Moberg, 2003) has been optimally averaged with uncertainties for the globe and

772 hemispheres. The NOAA surface temperature dataset produced by (Smith and Reynolds,
773 2005), uses Global Historical Climatology Network (GHCN), merged with the *in situ*
774 Extended Reconstruction SST (ERSST) analysis of (Smith and Reynolds, 2004). The
775 analyses are done separately over the ocean and the land following the ERSST methods.
776 Error estimates include the bias, random and sampling errors.

777

778 As an example of uncertainties in a combined land-ocean analysis, near-global time
779 series (60°S to 60°N) are shown in Figure 4.4 for SST, land-surface air temperature, and
780 the combined SST and land-surface air temperature (Smith and Reynolds, 2005). (The
781 combined product is the GHCN-ERSST product used in Chapter 3). The SST has the
782 tightest (95%) uncertainty limits (upper panel). The land-surface air temperature (middle
783 panel) has a larger trend over the period since 1958, but its uncertainty limits are also
784 larger than for SST. Land surface air temperature uncertainty is larger than the
785 uncertainty for SST because of higher variability of surface air temperature over land (see
786 Chapter 1), persistently un-sampled regions, including central Africa and interior South
787 America, and because the calculations include an increasing urbanization bias-error
788 estimate. Merged temperature anomalies and their uncertainty (lower panel) closely
789 resemble the SST result, since oceans cover most of the surface area. Similar uncertainty
790 was found by (Folland et al., 2001) using different methods.

791

792

793 **7. Interlayer comparisons.**

794

795 *7.1 Troposphere/Stratosphere*

796

797 All data sources agree that on a global scale, the stratosphere has cooled substantially
798 while the troposphere has warmed over both the 1958-2004 and the 1979-2004 time
799 periods (note that this is not true for all 25-year time periods within the longer 1958-2004
800 time period). We suspect that the stratospheric cooling trends estimated from radiosondes
801 are larger in magnitude than the actual trend. Despite the uncertainty in the exact
802 magnitude of stratospheric cooling, we have very high confidence that the lower
803 stratosphere has cooled relative to the troposphere by several tenths of a °C per decade
804 over the past 5 decades.

805

806 *7.2 Lower Troposphere/Mid-Upper Troposphere*

807

808 The difference in trend between the lower troposphere and mid-upper troposphere is not
809 well characterized by the existing data. On a global scale, all data sets suggest that T_{2LT} is
810 warming relative to T_2 , but it is important to note that the T_2 data records have significant
811 stratospheric contributions that reduce their warming trends. Radiosonde measurements
812 suggest that the $T_{(850-300)}$ layer (which does not include the stratosphere) is warming at
813 about the same rate as T_{2LT} , while satellite data suggest that T^*_G is warming more rapidly
814 than T_{2LT} . The magnitude of these inter-dataset differences are typically less than their
815 individual estimates of uncertainty, substantially reducing confidence in our ability to
816 deduce even the sign of the lower troposphere-mid-upper troposphere trend difference.

817

818 *7.3 Surface/Lower Troposphere*

819

820 On a global scale, both radiosonde datasets and one of the satellite datasets ($T_{2LT-UAH}$)
821 indicate that the surface warmed more than the lower troposphere between 1979 and
822 2004, while one satellite dataset ($T_{2LT-RSS}$) suggests the opposite. The magnitude of these
823 differences is less than the uncertainty estimates for any one data record. The situation is
824 similar in the tropics. However, in some regions, such as North America and Europe
825 (regions where the most reliable radiosonde stations are located), the warming in the
826 surface and lower troposphere appears to be very similar in all datasets.

827

828 *7.4 Surface/Mid Troposphere*

829

830 It is also interesting to consider the trend differences between the surface and mid
831 troposphere since more satellite datasets are available for T_2 . Here, mostly due to the
832 large structural uncertainty in the trends in T_2 , the various datasets are unable to agree on
833 the sign of the trend difference over the 1979-2004 period. On a global scale, the two
834 radiosonde datasets and two of the satellite datasets suggest that T_2 has warmed less than
835 the surface, but the other satellite dataset suggests that the opposite is true. The situation
836 is similar in the tropics. For the longer 1958-2004 period, all available datasets agree that
837 T_2 warms more than the surface. When T^*_G is considered, the difference between the
838 surface and tropospheric trends is reduced, with two satellite datasets indicating more
839 warming than at the surface.

840

841 8. Resolution of Uncertainty

842

843 In almost all of the tropospheric and stratospheric data records considered, our
844 uncertainty is dominated by structural uncertainty arising through dataset construction
845 choices (Thorne et al., 2005). Differences arising as a result of different, seemingly
846 plausible correction models applied by different groups to create a climate-quality data
847 record are significantly larger than the uncertainties internal to each method, in the raw
848 data measurements, or in the sampling uncertainties. These structural uncertainties are
849 difficult to assess in an absolute sense. The best estimates we can currently make come
850 from examining the spread of results obtained by different groups analyzing the same
851 type of data. This “all datasets are equal” approach has been employed in our present
852 analysis. As outlined in Chapter 2, this estimate of uncertainty can either be too small or
853 too large, depending on the situation. Given this caveat, it is always better to have
854 multiple (preferably at least three) data records that purport to measure the same aspects
855 of climate with the same data, so we can get some idea of the structural uncertainty.

856

857 In reality, all datasets are not equally plausible realizations of the true climate system
858 evolution. The climate system has evolved in a single way, and some datasets will be
859 closer to this truth than others. Given that the importance of structural uncertainty,
860 particularly for trends aloft, has only recently been recognized, it is perhaps unsurprising
861 that we are unable to quantify this at present. We could make value-based judgments to
862 imply increased confidence in certain datasets, but these would not be unambiguous, may
863 eventually be proven wrong, and are not a tenable approach in the longer term from a

864 scientific perspective. Therefore tools need to be developed to objectively discriminate
865 between datasets. These may include (1) measures of the internal consistency of the
866 construction methods, (2) assessment of the physical plausibility of the merged products,
867 including consistency of vertically resolved trends, and (3) comparisons with vicarious
868 data – for example, changes in temperature need to be compared with changes in water
869 vapor, winds, clouds, and various measures of radiation to assess consistency with the
870 expected physical relationships between these variables. Taken together such a suite of
871 indicators can be used to provide an objectively based way of highlighting residual
872 problems in the datasets and gaining a closer estimate of the truth. Such an audit of
873 current datasets should be seen as very high priority and preferably undertaken
874 independently of the dataset builders in a similar manner to the model intercomparisons
875 performed at Lawrence Livermore National Laboratory. In addition to an agreed set of
876 objective analysis tools, such an effort requires full and open access to all of the datasets
877 including a full audit trail.

878

879 Some specific suggestions for resolving some of the issues brought forward in this
880 chapter are mentioned here, but these are not exhaustive and further investigation is
881 required.

882

883 *8.1 Radiosondes.*

884

885 A significant contribution to the long-term inhomogeneity of the radiosonde record
886 appears to be related to changes in radiative heating of the temperature sensor for various

887 radiosonde models, and changes in the adjustments made to attempt to correct for these
888 changes. Recent work suggests that such problems may account for much of the tropical
889 cooling shown in unadjusted data. Other recent work suggests that step-like changes in
890 bias may still remain, even in adjusted datasets. Suitable tests on radiosonde products
891 may therefore include: stability of day-night differences, spatial consistency, internal
892 consistency (perhaps including wind data that to date have not been incorporated), and
893 consistency with MSU-derived and other independent estimates.

894

895 *8.2 Satellites.*

896

897 The most important contributions to satellite uncertainty are merging methodology and
898 the diurnal adjustment. The satellite data are simple enough that considerable
899 understanding can result from examination of intermediate results in the merging process,
900 including intersatellite differences that remain after the merging adjustments are
901 complete. Consistent reporting of such results can help differentiate between methods. It
902 appears that the differences in merging methodology often result in sharp step-like
903 features in difference time series between datasets. Other datasets, such as spatially
904 averaged adjusted radiosonde data, might be expected to show more slowly changing
905 errors, since their errors are due to the overlap of many different, potentially step-like
906 errors that occur at different times. So comparisons of satellite data with radiosonde data
907 over short time periods may help differentiate between satellite datasets. The diurnal
908 adjustment can be improved by a more rigorous validation of model-derived diurnal
909 cycles, or by further characterization of the diurnal cycle using the TRMM satellite or

910 concerted radiosonde observing programs designed to characterize the diurnal cycle at a
911 number of representative locations.

912

913 *8.3 Surface.*

914

915 The uncertainty in the historical near-surface temperature data is dominated by sampling
916 uncertainty, systematic changes in the local environment of surface observing stations,
917 and by difficult-to-characterize biases due to changes in SST measurement methods. The
918 relative maturity of the surface datasets suggest that to a large degree, these problems
919 have been addressed to the extent possible for the historical data, due to the absence of
920 the required metadata (for the bias-induced uncertainties) or the existence of any
921 observations at all. However, it is likely that much of the relatively recent SST data can
922 be adjusted for measurement type as some of the needed metadata is available or can be
923 estimated.

924

925

References

926 **Agudelo**, P. A. and J. A. Curry, 2004: Analysis of spatial distribution in tropospheric
927 temperature trends. *Geophysical Research Letters*, **31**: L22207.

928

929 **Christy**, J. R. and W. B. Norris, 2004: What may we conclude about global tropospheric
930 temperature trends? *Geophysical Research Letters*, **31**: L06211.

931

932 **Christy**, J. R., R. W. Spencer, et al., 2000: MSU Tropospheric Temperatures: Dataset
933 Construction and Radiosonde Comparisons. *Journal of Atmospheric and Oceanic
934 Technology*, **17**(9): 1153-1170.

935

936 **Christy**, J. R., R. W. Spencer, W. Norris, W. Braswell, D. Parker, 2003: Error Estimates
937 of Version 5.0 of MSU/AMSU Bulk Atmospheric Temperatures. *Journal of
938 Atmospheric and Oceanic Technology*, **20**: 613-629.

939

- 940 **Dai**, A. and K. E. Trenberth, 2004: The Diurnal Cycle and Its Depiction in the
941 Community Climate System Model. *Journal of Climate* **17**: 930-951.
942
- 943 **Davey**, C. A. and R. A. Pielke, 2005: Microclimate exposures of surface-based weather
944 stations - implications for the assessment of long-term temperature trends. Bull.
945 Amer. Meteor. Soc. **86**(4): 497-504.
946
- 947 **Durre**, I., T. C. Peterson, R. Vose, 2002: Evaluation of the Effect of the Luers-Eskridge
948 Radiation Adjustments on Radiosonde Temperature Homogeneity. *Journal of*
949 *Climate* **15**(6): 1335-1347.
- 950 **Folland**, C. K. and D. E. Parker, 1995: Correction of instrumental biases in historical sea
951 surface temperature data. *Q. J. Roy. Meteor. Soc.* **121**: 319-367.
952
- 953 **Folland**, C. K., N. A. Rayner, S. J. Brown, T. M. Smith, S. S. P. Shen, D. E. Parker, I.
954 Macadam, P. D. Jones, R. N. Jones, N. Nicholls, D. M. H. Sexton, 2001: Global
955 temperature change and its uncertainties since 1861. *Geophysical Research*
956 *Letters* **28**(13): 2621-2624.
957
- 958 **Free**, M., I. Durre, E. Aguilar, D. Seidel, T. C. Peterson, R. E. Eskridge, J. K. Luers, D.
959 Parker, M. Gordon, J. Lanzante, S. Klein, J. Christy, S. Schroeder, B. Soden, L. M.
960 McMillin, 2002: Creating climate reference datasets: CARDS workshop on
961 adjusting radiosonde temperature data for climate monitoring. Bulletin of the
962 American Meteorological Society **83**: 891-899.
963
- 964 **Free**, M. and D. Seidel, 2004: Accounting for Differences between Radiosonde Upper-air
965 Temperature Datasets. Under review.
966
- 967 **Fu**, Q. and C. M. Johanson, 2004: Stratospheric influences on MSU-derived tropospheric
968 temperature trends: A direct error analysis. *Journal of Climate* **17**: 4636-4640.
969
- 970 **Fu**, Q. and C. M. Johanson, 2005: Satellite-Derived Vertical Dependence of
971 Tropospheric Temperature Trends. *Geophysical Research Letters* **32**: L10703.
972
- 973 **Fu**, Q., C. M. Johanson, F. Warren, D. Seidel, 2004: Contribution of stratospheric
974 cooling to satellite-inferred tropospheric temperature trends. *Nature* **429**: 55-58.
975
- 976 **Gentemann**, C. L., F. J. Wentz, C. Mears, D. Smith, 2004: In situ validation of Tropical
977 Rainfall Measuring Mission microwave sea surface temperatures. *Journal of*
978 *Geophysical Research* **109**: C04021.
979
- 980 **Gillett**, N. P., B. D. Santer, A. Weaver, 2004: Quantifying the influence of stratospheric
981 cooling on satellite-derived tropospheric temperature trends. *Nature* **432**:
982 Published online December 2, 2004 (doi: 10.1038/nature03209).
983

- 984 **Grody**, N. C., K. Y. Vinnikov, M. Goldberg, J.. Sullivan, J. Tarpley, 2004: Calibration of
985 Multi-Satellite Observations for Climatic Studies: Microwave Sounding Unit
986 (MSU). *Journal of Geophysical Research* **109**.
987
- 988 **Hansen**, J. E., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, T.
989 Karl, 2001: A Closer look at United States and global surface temperature change.
990 *Journal of Geophysical Research* **106**: 23947-23963.
991
- 992 **Jones**, P. D. and A. Moberg, 2003: Hemispheric and Large-Scale Surface Air
993 Temperature Variations: An Extensive Revision and an Update to 2001. *Journal*
994 *of Climate* **16**: 206–223.
- 995 **Jones**, P. D., T. J. Osborn, K. Briffa, 1997: Estimating Sampling Errors in Large-Scale
996 Temperature Averages. *Journal of Climate* **10**(10): 2548–2568.
997
- 998 **Kilpatrick**, K. A., G. P. Podesta, R. Evans, 2001: Overview of the NOAA/NASA
999 advanced very high resolution radiometer pathfinder algorithm for sea surface
1000 temperature and associated matchup database. *Journal of Geophysical Research*
1001 **106**(C5): 9179-9198.
1002
- 1003 **Lanzante**, J., S. Klein, D. Seidel, 2003: Temporal Homogenization of Monthly
1004 Radiosonde Temperature Data. Part II: Trends, Sensitivities, and MSU
1005 comparison. *Journal of Climate* **16**: 241-262.
1006
- 1007 **Luers**, J. K. and R. E. Eskridge, 1998: Use of Radiosonde Temperature Data in Climate
1008 Studies. *Journal of Climate* **11**: 1002-1019.
1009
- 1010 **Mears**, C. and F. Wentz, 2005: The Effect of Drifting Measurement Time on Satellite-
1011 Derived Lower Tropospheric Temperature. *Science* **309**: 1548-1551.
1012
- 1013 **Mears**, C. A., M. C. Schabel, F. W. Wentz, 2003: A reanalysis of the MSU channel 2
1014 tropospheric temperature record. *Journal of Climate* **16**: 3650-3664.
1015
- 1016 **Parker**, D., 2004: Large scale warming is not urban. *Nature* **432**: 290.
1017
- 1018 **Parker**, D. E., M. Gordon, D. P. N Cullum, D. M. H. Sexto, C. K. Folland, N. Rayner,
1019 (1997: A new global gridded radiosonde temperature data base and recent
1020 temperature trends. *Geophysical Research Letters* **24**: 1499-1502.
1021
- 1022 **Pathack**, B., J. Nash, R. Smout, and S. Kurnosenko, 2005: Preliminary Results of WMO
1023 Intercomparison of high quality radiosonde systems, Mauritius, February 2005.
1024 [www.wmo.ch/web/www/IMOP/meetings/Upper-Air/RSO-](http://www.wmo.ch/web/www/IMOP/meetings/Upper-Air/RSO-ComparisonMauritius/3(16)_UK_Nash_Vacoas.pdf)
1025 [ComparisonMauritius/3\(16\)_UK_Nash_Vacoas.pdf](http://www.wmo.ch/web/www/IMOP/meetings/Upper-Air/RSO-ComparisonMauritius/3(16)_UK_Nash_Vacoas.pdf)
1026
- 1027 **Peterson**, T. C., 2003: Assessment of Urban Versus Rural In Situ Surface Temperatures
1028 in the Contiguous United States: No Difference Found. *Journal of Climate*
1029 **16**(18): 2941-2959.

- 1030
1031 **Peterson**, T. C., K. P. Gallo, J. Lavrimore, T. W. Owen, A. Huang, D.A. McKittrick,
1032 (1999: Global rural temperature trends. *Geophysical Research Letters* **26**(3): 329-
1033 332.
1034
1035 **Randel**, W. J. and F. Wu, 2005: Biases in stratospheric temperature trends derived from
1036 historical radiosonde data. *Journal of Climate* (in press).
1037
1038 **Rayner**, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell,
1039 2003: Global analyses of sea surface temperature, sea ice, and night marine air
1040 temperature since the late nineteenth century. *Journal of Geophysical Research*
1041 **108**(d14).
1042
1043 **Reynolds**, R. W., 1993: Impact of Mt Pinatubo aerosols on satellite-derived sea surface
1044 temperatures. *Journal of Climate* **6**: 768-774.
1045
1046 **Reynolds**, R. W., C. L. Gentemann, F. W. Wentz, 2004: Impact of TRMM SSTs on a
1047 Climate-Scale SST Analysis. *Journal of Climate* **17**: 2938-2952.
1048
1049 **Reynolds**, R. W., N. A. Rayner, et al., 2002: An Improved In Situ and Satellite SST
1050 Analysis for Climate. *Journal of Climate* **15**: 1609-1625.
1051
1052 **Reynolds**, R. W. and T. M. Smith, 1994: Improved global sea surface temperature
1053 analyses using optimum interpolation. *Journal of Climate* **7**: 929-948.
1054
1055 **Sherwood**, S., J. Lanzante, Cathryn Meyer, 2005: Radiosonde daytime biases and late
1056 20th century warming. *Science* **309**: 1556-1559.
1057
1058 **da Silveira**, R. B., G. Fisch, L Machato, A.Dall'Antonia, L Sapuchi, D. Fernandes, J.
1059 Nash, 2003: *Executive Summary of the WMO Intercomparison of GPS*
1060 *Radiosondes*. Geneva, World Meteorological Organization: 17.
1061
1062 **Smith**, T. M. and R. W. Reynolds, 2004: Improved Extended Reconstruction of SST
1063 (1854-1997). *Journal of Climate* **17**: 2466-2477.
1064
1065 **Smith**, T. M. and R. W. Reynolds, 2005: A global merged land and sea surface
1066 temperature reconstruction based on historical observations (1880-1997). *Journal*
1067 *of Climate* **18**(12): 2021-2036.
1068
1069 **Spencer**, R. W. and J. R. Christy, 1992: Precision and Radiosonde Validation of Satellite
1070 Gridpoint Temperature Anomalies. Part II: A Tropospheric Retrieval and Trends
1071 during 1979-1990. *Journal of Climate* **5**: 858-866.
1072
1073 **Spencer**, R.J., W.D. Braswell and J.R. Christy, 2005: Document describing v5.2 lower
1074 tropospheric retrieval.
1075

- 1076 **Tett, S.** and P. Thorne, 2004: Atmospheric Science: Tropospheric temperature series
1077 from satellites. *Nature* **432** .
1078
- 1079 **Thorne, P. W.,** D. E. Parker, et al., 2005: Causes of differences in observed climate
1080 trends. *Bulletin of the American Meteorological Society*. (in press).
1081
- 1082 **Thorne, P. W.,** D. E. Parker, S. F. B. Tett, P. D. Jones, M. McCarthy, H. Coleman, P.
1083 Brohan, 2005: Revisiting radiosonde upper-air temperatures from 1958 to 2002.
1084 *Journal of Geophysical Research*. (in press).
1085
- 1086 **Vinnikov, K. Y.** and N. C. Grody, 2003: Global Warming Trend of Mean Tropospheric
1087 Temperature Observed by Satellites. *Science* **302**: 269-272.
1088
- 1089 **Vinnikov, K. Y.,** N. C. Grody, et al., 2005: Temperature Trends at the Surface and in the
1090 Troposphere. *Journal of Geophysical Research* (submitted).
1091
- 1092 **Vose, R. S.,** D. Wuertz, T. C. Peterson, P. D. Jones, 2004: An intercomparison of land
1093 surface air temperature analyses at the global, hemispheric, and grid-box scale.
1094 *Geophysical Research Letters* **32**: L18718.
1095
- 1096 **Wallis, T. W. R.,** 1998: A subset of core station from the Comprehensive Aerological
1097 Reference Dataset (CARDS). *Journal of Climate* **11**: 272-282.
1098
- 1099 **Wentz, F.,** M. Schabel, 1998: Effects of satellite orbital decay on MSU lower
1100 tropospheric temperature trends. *Nature* **394**: 661-664.
1101
- 1102 **Wentz, F.,** C. Gentemann, D. Smith, D. Chelton, 2000: Satellite measurements of sea
1103 surface temperature through clouds. *Science* **288**: 847-850.
1104
- 1105 **Wentz, F. J.,** P. D. Ashcroft, C. Gentemann, 2001: Post-launch calibration of the TMI
1106 microwave radiometer. *IEEE Transactions on Geoscience and Remote Sensing*
1107 **39**: 415-422.
1108
1109

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26

Chapter 5

How well can the observed vertical temperature changes be reconciled with our understanding of the causes of these changes?

Convening Lead Author:

Ben D. Santer

Lead Authors:

Joyce E. Penner, and Peter W. Thorne

Contributors:

W. Collins, K. Dixon, T.L. Delworth, C. Doutriaux, Chris K. Folland, Chris E. Forest, I. Held, John Lanzante, Gerald A. Meehl, V. Ramaswamy, Dian J. Seidel, M.F. Wehner, and Tom M.L. Wigley

27 **KEY FINDINGS**

28

29 PATTERN STUDIES

30

31 Fingerprint studies use rigorous statistical methods to compare spatial and temporal patterns
32 of climate change in computer models and observations.

33

34 ***1. Both human and natural factors have affected Earth’s climate. Computer models are the***
35 ***only tools we have for estimating the likely climate response patterns (“fingerprints”)***
36 ***associated with different forcing mechanisms.***

37

38 To date, most formal fingerprint studies have focused on a relatively small number of
39 climate forcings. Our best scientific understanding is that:

40

41 • Increases in well-mixed greenhouse gases (which are primarily due to fossil fuel
42 burning) result in large-scale warming of the Earth’s surface and troposphere and
43 cooling of the stratosphere.

44 • Human-induced changes in the atmospheric burdens of sulfate aerosol particles
45 cause regional-scale cooling of the surface and troposphere.

46 • Depletion of stratospheric ozone cools the lower stratosphere and upper troposphere.

47 • Large volcanic eruptions cool the surface and troposphere (over 3 to 5 years) and
48 warm the stratosphere (over 1 to 2 years).

- 49 • Increases in solar irradiance warm throughout the atmospheric column (from the
50 surface to the stratosphere).

51
52 2. ***Results from many different fingerprint studies provide consistent evidence for a human***
53 ***influence on the three-dimensional structure of atmospheric temperature over the second***
54 ***half of the 20th century.***

55
56 Robust results are:

- 57
58 • Detection of greenhouse-gas and sulfate aerosol signals in observed surface
59 temperature records.

- 60 • Detection of an ozone depletion signal in stratospheric temperatures.

- 61 • Detection of the combined effects of greenhouse gases, sulfate aerosols, and ozone
62 in the vertical structure of atmospheric temperature changes (from the surface to the
63 stratosphere).

64
65 3. ***Natural factors have influenced surface and atmospheric temperatures, but cannot fully***
66 ***explain their changes over the past 50 years.***

- 67
68 • The multi-decadal climatic effects of volcanic eruptions and solar irradiance
69 changes are identifiable in some fingerprint studies, but results are sensitive to
70 analysis details.

71

72

73 TREND COMPARISONS

74

75 Linear trend comparisons are less powerful than “fingerprinting” for studying cause-effect
76 relationships, but can highlight important differences (and similarities) between models and
77 observations.

78

79 **4. *When run with natural and human-caused forcings, model global-mean temperature***
80 ***trends for individual atmospheric layers are consistent with observations.***

81

82 **5. *Comparing trend differences between the surface and the troposphere exposes potential***
83 ***model-data discrepancies in the tropics.***

84

85 • Differencing surface and tropospheric temperature time series (a simple measure of
86 the temperature lapse rate) removes much of the common variability between these
87 layers. This makes it easier to identify discrepancies between modeled and observed
88 lapse-rate changes.

89 • For globally-averaged temperatures, model-predicted trends in tropospheric lapse
90 rates are consistent with observed results.

91 • In the tropics, most observational datasets show more warming at the surface than in
92 the troposphere, while most model runs have larger warming aloft than at the
93 surface.

94

95

96

AMPLIFICATION OF SURFACE WARMING IN THE TROPOSPHERE

97

98

6. *In the tropics, surface temperature changes are amplified in the free troposphere. Models and observations show similar amplification behavior for monthly- and interannual temperature variations, but not for decadal temperature changes.*

99

100

101

102

- Tropospheric amplification of surface temperature anomalies is due to the release of latent heat by moist, rising air in regions experiencing convection.

103

104

- Despite large inter-model differences in variability and forcings, the size of this amplification effect is remarkably similar in the models considered here, even across a range of timescales (from monthly to decadal).

105

106

107

- On monthly and annual timescales, amplification is also a ubiquitous feature of observations, and is very similar to values obtained from models and basic theory.

108

109

- For longer-timescale temperature changes over 1979 to 1999, only one of four observed upper-air datasets has larger tropical warming aloft than in the surface records. All model runs with surface warming over this period show amplified warming aloft.

110

111

112

113

- These results have several possible explanations, which are not mutually exclusive. One explanation is that “real world” amplification effects on short and long time scales are controlled by different physical mechanisms, and models fail to capture such behavior. A second explanation is that remaining errors in some of the

114

115

116

117 observed tropospheric data sets adversely affect their long-term temperature trends.
118 The second explanation is more likely in view of the model-to-model consistency of
119 amplification results, the large uncertainties in observed tropospheric temperature
120 trends, and independent physical evidence supporting substantial tropospheric
121 warming.

122

123 OTHER FINDINGS

124

125 7. *It is important to account for observational uncertainty in comparisons between modeled*
126 *and observed temperature changes.*

127

128 • There are large “construction uncertainties” in the process of generating climate data
129 records from raw observations. These uncertainties can critically influence the outcome
130 of consistency tests between models and observations.

131

132 8. *Inclusion of spatially-heterogeneous forcings in the most recent climate models does not*
133 *fundamentally alter simulated lapse-rate changes at the largest spatial scales.*

134

135 • Changes in black carbon aerosols and land use/land cover (LULC) may have had
136 significant influences on regional temperatures, but these influences have not been
137 quantified in formal fingerprint studies.

- 138 • These forcings were included for the first time in about half the global model
139 simulations considered here. Their incorporation did not significantly affect simulations
140 of lapse-rate changes at very large spatial scales (global and tropical averages).

141

142 **RECOMMENDATIONS**

143

- 144 *1. Separate the uncertainties in climate forcings from uncertainties in the climate response*
145 *to forcings.*

146

147 The simulations of Twentieth Century (20CEN) climate analyzed here show climate
148 responses that differ because of differences in:

149

- 150 • Model physics and resolution;
- 151 • The forcings incorporated in the 20CEN experiment;
- 152 • The chosen forcing history, and the manner in which a specific forcing was applied.

153

154 We consider it a priority to partition the uncertainties in climate forcings and model
155 responses, and thus improve our ability to interpret differences between models and
156 observations. This could be achieved by better coordination of experimental design,
157 particularly for the 20CEN simulations that are most relevant for direct comparison with
158 observations.

159

160 2. *Quantify the contributions of changes in black carbon aerosols and land use/land cover*
161 *to recent large-scale temperature changes.*

162
163 We currently lack experiments in which the effects of black carbon aerosols and LULC are
164 varied individually (while holding other forcings constant). Such “single forcing” runs will
165 help to quantify the contributions of these forcings to global-scale changes in lapse-rates.

166
167 3. *Explicitly consider model and observational uncertainty.*

168
169 Efforts to evaluate model performance or identify human-induced climate change should
170 always account for uncertainties in both observations and in model simulations of historical
171 and future climate. This is particularly important for comparisons involving long-term
172 changes in upper-air temperatures. It is here that current observational uncertainties are
173 largest and require better quantification.

174
175 4. *Perform the “next generation” of detection and attribution studies.*

176
177 Formal detection and attribution studies utilizing the new generation of model and
178 observational datasets detailed herein should be undertaken as a matter of priority.

179 1 ***Introduction***

180

181 A key scientific question addressed in this report is whether the Earth’s surface has warmed more
182 rapidly than the troposphere over the past 25 years (*NRC, 2000*). Chapter 1 noted that there are
183 good physical reasons why we do not expect surface and tropospheric temperatures to evolve in
184 unison at all places and on all timescales. Chapters 2, 3, and 4 summarized our current
185 understanding of observed changes in surface and atmospheric temperatures. These chapters
186 identified important differences between surface and tropospheric temperatures, some of which
187 may be due to remaining problems with the observational data, and some of which are likely to be
188 real.

189

190 In Chapter 5, we seek to explain and reconcile the apparently disparate estimates of observed
191 changes in surface and tropospheric temperatures. We make extensive use of computer models of
192 the climate system. In the real world, multiple “climate forcings” vary simultaneously, and it is
193 difficult to identify and separate the climate effects of individual factors. Furthermore, the
194 experiment that we are performing with the Earth’s atmosphere lacks a suitable control – we do not
195 have a convenient “parallel Earth” on which there are no human-induced changes in greenhouse
196 gases, aerosols, or other climate forcings. Climate models can be used to perform such controlled
197 experiments, or to simulate the response to changes in a single forcing or combination of forcings,
198 and thus have real advantages for studying cause-effect relationships. However, models also have
199 systematic errors that can diminish their usefulness as a tool for interpretation of observations
200 (*Gates et al., 1999; McAvaney et al., 2001*).

201

202 We evaluate published research that has made rigorous quantitative comparisons of modeled and
203 observed temperature changes, primarily over the satellite and radiosonde eras. Some new model
204 experiments (performed in support of the IPCC Fourth Assessment Report) involve simultaneous
205 changes in a wide range of natural and human-induced climate forcings. These experiments are
206 highly relevant for direct comparison with satellite-, radiosonde-, and surface-based temperature
207 observations. We review their key results here.

208

209 2 *Model Simulations of Recent Temperature Change*

210

211 Many different types of computer model are used for studying climate change issues (*Meehl*, 1984;
212 *Trenberth*, 1992; see Box 5.1). Models span a large range of complexity, from the one- or two-
213 dimensional energy-balance models (EBMs) through Earth System models of intermediate
214 complexity (EMICs) to full three-dimensional atmospheric General Circulation Models (AGCMs)
215 and coupled atmosphere-ocean GCMs (CGCMs). Each type has advantages and disadvantages for
216 specific applications. The more complex AGCMs and CGCMs are most appropriate for
217 understanding problems related to the atmosphere's vertical temperature structure, since they
218 explicitly resolve that structure, and incorporate many of the physical processes (*e.g.*, convection,
219 interactions between clouds and radiation) thought to be important in maintaining atmospheric
220 temperature profiles. They are also capable of representing the horizontal and vertical structure of
221 unevenly-distributed climate forcings that may contribute to differential warming of the surface and
222 troposphere. Examples include volcanic aerosols (*Robock*, 2000) or the sulfate and soot aerosols
223 arising from fossil fuel or biomass burning (*Penner et al.*, 2001; *Ramaswamy et al.*, 2001a,b).

224

225 BOX 5.1: Climate Models

226
227 Climate models provide us with estimates of how the real world’s climate system behaves and is
228 likely to respond to changing natural and human-caused forcings. Because of limitations in our
229 physical understanding and computational capabilities, models are simplified and idealized
230 representations of a very complex reality. The most sophisticated climate models are direct
231 descendants of the computer models used for weather forecasting. While weather forecast models
232 seek to predict the specific timing of weather events over a period of days to several weeks, climate
233 models attempt to simulate future changes in the *average distribution* of weather events.
234 Simulations of 21st Century climate are typically based on “scenarios” of future emissions of
235 GHGs, aerosols and aerosol precursors, which in turn derive from scenarios of population changes,
236 economic growth, energy usage, developments in energy production technology, *etc.*
237

238
239 Climate models are also used to “hindcast” the climate changes that we have observed over the 20th
240 Century. When run in “hindcast” mode, a climate model is not constrained by *actual* weather
241 observations from satellites or radiosondes. Instead, it is driven by our best estimates of changes in
242 some (but probably not all) of the major “forcings”, such as GHG concentrations, the Sun’s energy
243 output, and the amount of volcanic dust in the atmosphere. In “hindcast” experiments, a climate
244 model is free to simulate the full four-dimensional (latitude, longitude, height/depth and time)
245 distributions of temperature, moisture, *etc.* Comparing the results of such an experiment with long
246 observational records constitutes a valuable test of model performance.

247
248 AGCM experiments typically rely on an atmospheric model driven by observed time-varying
249 changes in sea-surface temperatures (SSTs) and sea-ice coverage. This is a standard reference
250 experiment that many AGCMs have performed as part of the Atmospheric Model Intercomparison
251 Project (“AMIP”; *Gates et al.*, 1999). The AMIP-style experiments discussed here also include
252 specified changes in a variety of natural and human-caused forcing factors (*Hansen et al.*, 1997,
253 2002; *Folland et al.*, 1998; *Tett and Thorne*, 2004).

254
255 In both observations and climate models, variations in the El Niño/Southern Oscillation (ENSO)
256 have pronounced effects on surface and tropospheric temperatures (*Yulaeva and Wallace, 1994;*
257 *Wigley, 2000; Santer et al., 2001; Hegerl and Wallace, 2002; Hurrell et al., 2003*). When run in an
258 AMIP configuration, an atmospheric model “sees” the same changes in ocean surface temperature
259 that the real world’s atmosphere experienced. The time evolution of ENSO effects on atmospheric
260 temperature is therefore very similar in the model and observations. This facilitates the direct
261 comparison of modeled and observed temperature changes.¹ Furthermore, AMIP experiments
262 reduce climate noise by focusing on the random variability arising from the atmosphere rather than
263 on the variability of the coupled atmosphere-ocean system (which is larger in amplitude). This
264 “noise reduction” aspect of AMIP runs has been exploited in efforts to identify human effects on
265 year-to-year changes in atmospheric temperatures (*Folland et al., 1998; Sexton et al., 2001*)
266

¹This does not mean, however, that the atmospheric model will necessarily capture the correct amplitude and horizontal and vertical structure of the tropospheric temperature response to the specified SST ice changes. Note also that even with the specification of ocean boundary conditions, the time evolution of modes of variability that are forced by both the ocean and the atmosphere (such as the North Atlantic Oscillation; see *Rodwell et al., 1999*) will not be the same in the model and in the real world (except by chance).

¹Volcanic forcing provides an example of the signal estimation problem. The aerosols injected into the stratosphere during a massive volcanic eruption are typically removed within 2-3 years (*Sato et al., 1993; Hansen et al., 2002; Ammann et al., 2003*). Because the large thermal inertia of the oceans cause a lag in response to this forcing, the cooling effect of the aerosols on the troposphere and surface persists for much longer than 2-3 years (*Santer et al., 2001; Free and Angell, 2002; Wigley et al., 2005a*). In the real world and in “AMIP-style” experiments, this slow, volcanically-induced cooling of the troposphere and surface is sometimes masked by the warming effects of El Niño events (*Christy and McNider, 1994; Wigley, 2000; Santer et al., 2001*), thus hampering volcanic signal estimation.

¹There are a variety of different spin-up strategies.

¹In most of the experiments reported on here, n is between 3 and 5.
and sea-ice changes. Note also that even with the specification of ocean boundary conditions, the time evolution of modes of variability that are forced by both the ocean and the atmosphere (such as the North Atlantic Oscillation; see *Rodwell et al., 1999*) will not be the same in the model and in the real world (except by chance).

267 One disadvantage of the AMIP experimental set-up is that significant errors in one or more of the
268 applied forcing factors (or omission of key forcings) are not “felt” by the prescribed SSTs. Such
269 errors are more obvious in a CGCM experiment, where the ocean surface is free to respond to
270 imposed forcings. The lack of an ocean response and the masking effects of natural variability
271 make it difficult to use an AMIP-style experiment to estimate the slow response of the climate
272 system to an imposed forcing change.² CGCM experiments are more useful for this specific
273 purpose (see Chapter 1, Figure 1.3).

274
275 The CGCM experiments of interest here involve a model that has been “spun-up” until it reaches
276 some quasi-steady climate state³. The CGCM is then run with estimates of how a variety of natural
277 and human-caused climate forcings have changed over the 20th century. We refer to these
278 subsequently as “20CEN” experiments. Since the true state of the climate system is never fully
279 known, the same forcing changes are applied n times,⁴ each time starting from a slightly different
280 initial climate state. This procedure yields n different realizations of climate change. All of these
281 realizations contain some underlying “signal” (the climate response to the imposed forcing
282 changes) upon which are superimposed n different manifestations of “noise” (natural internal
283 climate variability). Taking averages over these n realizations yields less noisy estimates of the
284 signal (Wigley *et al.*, 2005a).

²Volcanic forcing provides an example of the signal estimation problem. The aerosols injected into the stratosphere during a massive volcanic eruption are typically removed within 2-3 years (Sato *et al.*, 1993; Hansen *et al.*, 2002; Ammann *et al.*, 2003). Because the large thermal inertia of the oceans cause a lag in response to this forcing, the cooling effect of the aerosols on the troposphere and surface persists for much longer than 2-3 years (Santer *et al.*, 2001; Free and Angell, 2002; Wigley *et al.*, 2005a). In the real world and in “AMIP-style” experiments, this slow, volcanically-induced cooling of the troposphere and surface is sometimes masked by the warming effects of El Niño events (Christy and McNider, 1994; Wigley, 2000; Santer *et al.*, 2001), thus hampering volcanic signal estimation.

³There are a variety of different spin-up strategies.

⁴In most of the experiments reported on here, n is between 3 and 5.

285

286 In a CGCM, ocean temperatures are fully predicted rather than prescribed. This means that even a
287 (hypothetical) CGCM which perfectly captured all important aspects of ENSO physics would not
288 have the same timing of El Niño and La Niña events as the real world (except by chance). The fact
289 that ENSO variability – and its effects on surface and atmospheric temperatures – does not “line up
290 in time” in observations and CGCM experiments hampers direct comparisons between the two.⁵
291 This problem can be ameliorated by statistical removal of ENSO effects (*Santer et al.*, 2001;
292 *Hegerl and Wallace*, 2002; *Wigley et al.*, 2005a).⁶

293

294 The bottom line is that AMIP-style experiments and CGCM runs are both useful tools for exploring
295 the possible causes of differential warming.⁷ We note that even if these two experimental
296 configurations employ the same atmospheric model and the same climate forcings, they can yield
297 noticeably different simulations of changes in atmospheric temperature profiles. These differences
298 arise for a variety of reasons, such as AGCM-versus-CGCM differences in sea-ice coverage, SST
299 distributions, and cloud feedbacks, and hence in climate sensitivity (*Sun and Hansen*, 2003).⁸

300

301 Most models undergo some form of “tuning”. This involves changing poorly-known parameters
302 which directly affect key physical processes, such as convection and rainfall. Parameters are varied

⁵If n is large enough to adequately sample the (simulated) effects of natural variability on surface and tropospheric temperatures, it is not necessarily a disadvantage that the simulated and observed variability does not line up in time. In fact, this type of experimental set-up allows one to determine whether the single realization of the observations is contained within the “envelope” of possible climate solutions that the CGCM simulates.

⁶Residual effects of these modes of variability may remain in the data.

⁷Provided that comparisons with observations account for the specific advantages and disadvantages noted above.

⁸See, for example, the Ocean A and Ocean E results in Figure 3 of *Sun and Hansen* (2003).

303 within plausible ranges, which are generally derived from direct observations. The aim of tuning is
304 to reduce the size of systematic model errors and improve simulations of present-day climate.
305 Tuning does *not* involve varying uncertain model parameters over the course of a 20CEN
306 experiment, in order to improve a given model’s simulation of observed climate change over the
307 20th Century.⁹

308
309 Several groups are now beginning to explore model “parameter space”, and are investigating the
310 possible impact of parameter uncertainties on simulations of mean present-day climate and future
311 climate change (*Allen, 1999; Forest et al., 2002; Murphy et al., 2004; Stainforth et al., 2005*). Such
312 work will help to quantify one component of model uncertainty. Another component of model
313 uncertainty arises from differences in the basic structure of models.¹⁰ Section 5 considers results
314 from a range of state-of-the-art CGCMs, and thus samples some of the “structural uncertainty” in
315 model simulations of 20th Century climate change (Table 5.1). A further component of the “spread”
316 in simulations of 20th Century climate is introduced by uncertainties in the climate forcings with
317 which models are run (Table 5.2). These are discussed in the following Section.

318 Table 5.1: Acronyms of climate models referenced in this Chapter. All 19 models performed simulations of 20th
319 century climate change (“20CEN”) in support of the IPCC Fourth Assessment Report. The ensemble size “ES” is the
320 number of independent realizations of the 20CEN experiment that were analyzed here.

⁹Potentially, highly uncertain climate forcings (particularly those associated with the indirect effects of aerosol particles on clouds) could be adjusted to improve the correspondence between modeled and observed global-mean surface temperature changes over the 20th Century. Such tuning does not occur *per se* and would be an unacceptable procedure, quite different from the parameter adjustments that are made when improving AGCM and CGCM simulations of mean climate.

¹⁰The computer models constructed by different research groups can have quite different “structures” in terms of their horizontal and vertical resolution, atmospheric dynamics (so-called “dynamical cores”), numerical implementation (*e.g.*, spectral versus grid-point), and physical parameterizations. They do, however, share many common assumptions.

321 ¹CSIRO is the Commonwealth Scientific and Industrial Research Organization.

322 ²NIES is the National Institute for Environmental Studies.

323 ³JAMSTEC is the Frontier Research Center for Global Change in Japan.

324

325

326 Table 5.2: Forcings used in IPCC simulations of 20th century climate change. This Table was compiled using
327 information provided by the participating modeling centers (see [http://www-](http://www-pcmdi.llnl.gov/ipcc/model.documentation)
328 [pcmdi.llnl.gov/ipcc/model.documentation](http://www-pcmdi.llnl.gov/ipcc/model.documentation)). Eleven different forcings are listed: well-mixed greenhouse gases (G),
329 tropospheric and stratospheric ozone (O), sulfate aerosol direct (SD) and indirect effects (SI), black carbon (BC) and
330 organic carbon aerosols (OC), mineral dust (MD), sea salt (SS), land use/land cover (LU), solar irradiance (SO), and
331 volcanic aerosols (V). Shading denotes inclusion of a specific forcing. As used here, “inclusion” means specification of

332 a time-varying forcing, with changes on interannual and longer timescales. Forcings that were varied over the seasonal
333 cycle only are not shaded.

334

335

336

337 3 *Forcings in Simulations of Recent Climate Change*

338

339 In an ideal world, there would be reliable quantitative estimates of all climate forcings – both
340 natural and human-induced – that have made significant contributions to differential warming of
341 the surface and troposphere. We would have detailed knowledge of spatial and temporal changes in
342 these forcings. Finally, we would have used standard forcings to perform climate-change
343 experiments with a whole suite of numerical models, thus isolating uncertainties arising from
344 structural differences in the models themselves (see Box 5.2).

345

346

347

348 Box 5.2: Uncertainties in Simulated Temperature Changes

349

350 In discussing the major sources of uncertainty in observational estimates of temperature change,
351 Chapter 2 partitioned uncertainties into three distinct categories: “structural,” “parametric,” and
352 “statistical.” Uncertainties in simulated temperature changes fall into similar categories. In the
353 modeling context, “structural” uncertainties can be thought of as the uncertainties resulting from
354 the choice of a particular climate model, model configuration (Section 2), or forcing dataset
355 (Section 3).

356

357 Within a given model, there are small-scale physical processes (such as convection, cloud
358 formation, precipitation, *etc.*) which cannot be simulated explicitly. Instead, so-called
359 “parameterizations” represent the large-scale effects of these unresolved processes. Each of these
360 has uncertainties in the values of key parameters.¹¹ Varying these parameters within plausible
361 ranges introduces “parametric” uncertainty in climate change simulations (*Allen, 1999; Forest et*
362 *al., 2002; Murphy et al., 2004*). Finally (analogous to the observational case), there is statistical
363 uncertainty that arises from the unpredictable “noise” of internal climate variability, from the
364 choice of a particular statistical metric to describe climate change, or from the application of a
365 selected metric to noisy data.

366

¹¹Note that some of these parameters influence not only the climate response, but also the portrayal of the forcing itself. Examples include parameters related to the size of sulfate aerosols, and how aerosol particles scatter incoming sunlight.

367
368 Unfortunately, this ideal situation does not exist. As part of the IPCC Third Assessment Report,
369 *Ramaswamy et al.* (2001b) assigned subjective confidence levels to our current “level of scientific
370 understanding” (LOSU) of the changes in a dozen different climate forcings. Only in the case of
371 well-mixed greenhouse gases (“GHGs”; carbon dioxide [CO₂], methane, nitrous oxide, and
372 halocarbons) was the LOSU characterized as “high.” The LOSU of changes in stratospheric and
373 tropospheric ozone was judged to be “medium.” For all other forcings (various aerosols, mineral
374 dust, land use-induced albedo changes, solar, *etc.*), the LOSU was estimated to be “low” or “very
375 low” (see Chapter 1, Table 1).¹²

376
377
378 In selecting the forcings for simulating the climate of the 20th Century (20CEN), there are at least
379 three strategies that modeling groups can adopt. The first strategy is to incorporate only those
380 forcings whose changes and effects are thought to be better understood, and for which time- and
381 space-resolved datasets suitable for performing 20CEN experiments are readily available. The
382 second strategy is to include a large number of different forcings, even those for which the LOSU
383 is “very low.” A third strategy is to vary the size of poorly-known 20CEN forcings. This yields a
384 range of simulated climate responses, which are then used to estimate the levels of the forcings that
385 are consistent with observations (*e.g.*, *Forest et al.*, 2002).

386
387 The pragmatic focus of Chapter 5 is on climate forcings that have been incorporated in many
388 CGCM simulations of 20th century climate. The primary forcings that we consider are changes in

¹²We note that there is no direct relationship between the LOSU of a given forcing and the contribution of that forcing to 20th Century climate change. Forcings with “low” or “very low” LOSU may have had significant climatic impacts at regional and even global scales.

389 well-mixed GHGs, the direct effects of sulfate aerosol particles, tropospheric and stratospheric
390 ozone, volcanic aerosols, and solar irradiance. These are forcings whose effects on surface and
391 atmospheric temperatures have been quantified in rigorous fingerprint studies (see Section 4). This
392 does not diminish the importance of other climate forcings, whose global-scale contribution to
393 “differential warming” has not been reliably quantified to date.

394
395 Examples of these “other forcings” include carbon-containing aerosols produced during fossil fuel
396 or biomass combustion, human-induced changes in land surface properties, and the indirect effects
397 of tropospheric aerosols on cloud properties. There is emerging scientific evidence that such
398 spatially-variable forcings may have had important impacts on regional and even on global climate
399 (*NRC*, 2005). Some of this evidence is summarized in Box 5.3 and Box 5.4 for the specific cases of
400 carbonaceous aerosols and land use change. These and other previously-neglected forcings have
401 been included in many of the new CGCM simulations of 20th century climate described in Section
402 5 (see Tables 5.1 and 5.2).

403

404

405

406

407

408

409

410

411

412 Box 5.3: Example of a Spatially-Heterogeneous Forcing: Black Carbon Aerosols

413
414 Carbon-containing aerosols (also known as “carbonaceous” aerosols) exist in a variety of chemical
415 forms (*Penner et al.*, 2001). Two main classes of carbonaceous aerosol are generally distinguished:
416 “black carbon” (BC) and “organic carbon” (OC). Both types of aerosol are emitted during fossil
417 fuel and biomass burning. Most previous modeling work has focused on BC aerosols rather than
418 OC aerosols. Some of the new model experiments described in Section 5 have now incorporated
419 both types of aerosol in CGCM simulations of 20th century climate changes (see Tables 5.2 and
420 5.3).

421
422 Black carbon aerosols absorb sunlight and augment the GHG-induced warming of the troposphere
423 (*Hansen et al.*, 2000; *Satheesh and Ramanathan*, 2000; *Penner et al.*, 2001; *Hansen*, 2002; *Penner*
424 *et al.*, 2003).¹³ Their effects on atmospheric temperature profiles are complex, and depend on such
425 factors as the chemical composition, particle size, and height distribution of the aerosols (*e.g.*,
426 *Penner et al.*, 2003).

427
428 *Menon et al.* (2002) showed that the inclusion of fossil fuel and biomass aerosols over China and
429 India¹⁴ directly affected simulated vertical temperature profiles by heating the lower troposphere
430 and cooling the surface. In turn, this change in atmospheric heating influenced regional circulation
431 patterns and the hydrological cycle. *Krishnan and Ramanathan* (2002) found that an increase in
432 black carbon aerosols has reduced the surface solar insolation (exposure to sunlight) over the
433 Indian subcontinent. Model experiments performed by *Penner et al.* (2003) suggest that the net
434 effect of carbonaceous aerosols on global-scale surface temperature changes depends critically on
435 how aerosols affect the vertical distribution of clouds. On regional scales, the surface temperature
436 effects of these aerosols are complex, and vary in sign (*Penner et al.*, 2005).

437
438
439
440
441
442
443
444

¹³Note that soot particles are sometimes transported long distances by winds, and can also have a “far field” effect on climate by reducing the reflectivity of snow in areas remote from pollution sources (*Hansen and Nazarenko*, 2003; *Jacobson*, 2004).

¹⁴During winter and spring, black carbon aerosols contribute to a persistent haze over large areas of Southern Asian and the Northern Indian Ocean (*Ramanathan et al.*, 2001).

445 Box 5.4: Example of a Spatially-Heterogeneous Forcing: Land Use Change

446 Humans have transformed the surface of the planet through such activities as conversion of forest
447 to cropland, urbanization, irrigation, and large water diversion projects (see Chapter 4). These
448 changes can affect a variety of physical properties of the land surface, such as the albedo
449 (reflectivity), the release of water by plants (transpiration), the moisture-holding capacity of soil,
450 and the surface “roughness.” Alterations in these physical properties may in turn affect runoff, heat
451 and moisture exchanges between the land surface and atmospheric boundary layer, wind patterns,
452 and even rainfall (*e.g.*, *Pitman et al.*, 2004). Depending on the nature of the change, either warming
453 or cooling of the land surface may occur (*Myhre and Myhre*, 2003).

454
455 At the regional level, modeling studies of the Florida peninsula (*Marshall et al.*, 2004) and
456 southwest Western Australia (*Pitman et al.*, 2004) have linked regional-scale changes in
457 atmospheric circulation and rainfall to human transformation of the natural vegetation. Modeling
458 work focusing on North America suggests that the conversion of natural forest and grassland to
459 agricultural production has led to a cooling in summertime (*Oleson et al.*, 2004). The global-scale
460 signal of land use/land cover (LULC) changes from pre-industrial times to the present is estimated
461 to be a small net cooling of surface temperature (*Matthews et al.*, 2003, 2004; *Brovkin et al.*, 2004;
462 *Hansen et al.*, 2005a; *Feddema et al.*, 2005). Larger regional trends of either sign are likely to be
463 evident (*e.g.*, *Hansen et al.*, 2005a).¹⁵

464

465 Clearly, we will *never* have complete and reliable information on all forcings that are thought to
466 have influenced climate over the late 20th century. A key question is whether those forcings most
467 important for understanding the differential warming problem are reliably represented. This is
468 currently difficult to answer. What we *can* say, with some certainty, is that the expected
469 atmospheric temperature signal due to forcing by well-mixed GHGs alone is distinctly different
470 from the signal due to the combined effects of multiple natural and human forcing factors (Chapter
471 1; *Santer et al.*, 1996; *Tett et al.*, 1996; *Hansen et al.*, 1997, 2002; *Bengtsson et al.*, 1999; *Santer et*
472 *al.*, 2003a).

473

¹⁵Note that larger regional trends do not necessarily translate to enhanced detectability. Although the signals of LULC and other spatially-heterogeneous forcings are likely to be larger regionally than globally, the “noise” of natural climate variability is also larger at smaller spatial scales. It is not obvious *a priori*, therefore, how signal-to-noise relationships (and detectability of a given forcing’s climate effects) behave as one moves from global to continental to regional scales.

474 This is illustrated by the 20CEN and “single forcing” experiments performed with the Parallel
475 Climate Model (PCM; *Washington et al.*, 2000). In PCM, changes in the vertical profile of
476 atmospheric temperature over 1979 to 1999 are primarily forced by changes in well-mixed GHGs,
477 ozone, and volcanic aerosols (Figure 5.1). Changes in solar irradiance and the scattering effects of
478 sulfate aerosols are of secondary importance over this period. Even without performing formal
479 statistical tests, it is visually obvious from Figure 5.1 that radiosonde-based estimates of observed
480 stratospheric and tropospheric temperature changes are in better agreement with the PCM 20CEN
481 experiment than with the PCM “GHG only” run.

482

483

484

485 Figure 5.1: Vertical profiles of global-mean atmospheric temperature change over 1979 to 1999. Surface temperature
486 changes are also shown. Results are from two different radiosonde data sets (HadAT2 and RATPAC; see Chapter 3)
487 and from single forcing and combined forcing experiments performed with the Parallel Climate Model (PCM;
488 *Washington et al., 2000*). PCM results for each forcing experiment are averages over four different realizations of that
489 experiment. All trends were calculated with monthly mean anomaly data.
490

491 This illustrates the need for caution in comparisons of modeled and observed atmospheric
492 temperature change. The differences evident in such comparisons have multiple interpretations.
493 They may be due to real errors in the models,¹⁶ errors in the forcings used to drive the models, the
494 neglect of important forcings, and residual inhomogeneities in the observations themselves. They

¹⁶These may lie in the physics, parameterizations, inadequate horizontal or vertical resolution, *etc.*

495 may also be due to different manifestations of natural variability noise in the observations and a
496 given CGCM realization. All of these factors may be important in model evaluation work.

497

498 4. ***Published Comparisons of Modeled and Observed Temperature Changes***

499

500 A number of observational and modeling studies have attempted to shed light on the possible
501 causes of “differential warming”.¹⁷ We have attempted to organize the discussion of results so that

¹⁷We do not discuss studies which provide empirical estimates of “equilibrium climate sensitivity” – the steady-state warming of the Earth’s surface that would eventually be reached after the climate system equilibrated to a doubling of pre-industrial atmo tropospheric temperatures to massive volcanic eruptions (*Hansen et al.*, 1993; *Lindzen and Giannitsis*, 1998; *Dougllass and Knox*, 2005; *Wigley et al.*, 2005a,b; *Robock*, 2005); the “intermediate” (100- to 150-year) response of surface temperatures to natural and human-caused forcing changes over the 19th and 20th centuries (*Andronova and Schlesinger*, 2001; *Forest et al.*, 2002; *Gregory et al.*, 2002; *Harvey and Kaufmann*, 2002) or to solar and volcanic forcing changes over the past 1-2 millennia (*Crowley*, 2000), and the slow (100,000-year) response of Earth’s temperature to orbital changes between glacial and interglacial conditions (*Hoffert and Covey*, 1992; *Hansen et al.*, 1993). These investigations are not directly relevant to elucidation of the causes of changes in the vertical structure of atmospheric temperatures, which is the focus of our Chapter.

¹⁷It is useful to mention one technical issue relevant to model-data comparisons. As noted in Chapter 2, the satellite-based Microwave Sounding Unit (MSU) monitors the temperature of very broad atmospheric layers. To facilitate comparisons with observed MSU datasets, many of the studies reported on here calculate “synthetic” MSU temperatures from climate model experiments. Technical aspects of these calculations are discussed in Chapter 2, Box 2.

¹⁷The studies by *Jones* (1994) and *Christy and McNider* (1994) remove volcano and ENSO effects from T_{2LT} , and estimate residual trends of 0.093 and 0.090°C/decade over 1979 to 1993. A similar investigation by *Michaels and Knappenberger* (2000) obtained a residual trend of 0.041°C/decade over 1979 to 1999. The error bars on these residual trend estimates are either not given, or claimed to be very small (*e.g.*, $\pm 0.005^\circ\text{C}/\text{decade}$ in *Christy and McNider*). A fourth study removed combined ENSO, volcano, and solar effects from T_{2LT} , and estimated a residual trend of $0.065 \pm 0.012^\circ\text{C}/\text{decade}$ over 1979 to 2000 (*Dougllass and Clader*, 2000).

heric CO₂ levels. This is often referred to as $\Delta T_{2\times\text{CO}_2}$. Estimates of $\Delta T_{2\times\text{CO}_2}$ have been obtained by studying Earth’s temperature response to “fast”, “intermediate”, and “slow” forcing of the climate system. Examples include the “fast” (<10-year) response of surface and tropospheric temperatures to massive volcanic eruptions (*Hansen et al.*, 1993; *Lindzen and Giannitsis*, 1998; *Dougllass and Knox*, 2005; *Wigley et al.*, 2005a,b; *Robock*, 2005); the “intermediate” (100- to 150-year) response of surface temperatures to natural and human-caused forcing changes over the 19th and 20th centuries (*Andronova and Schlesinger*, 2001; *Forest et al.*, 2002; *Gregory et al.*, 2002; *Harvey and Kaufmann*, 2002) or to solar and volcanic forcing changes over the past 1-2 millennia (*Crowley*, 2000), and the slow (100,000-year) response of Earth’s temperature to orbital changes between glacial and interglacial conditions (*Hoffert and Covey*, 1992; *Hansen et al.*, 1993). These investigations are not directly relevant to elucidation of the causes of changes in the vertical structure of atmospheric temperatures, which is the focus of our Chapter.

502 investigations with similar analysis methods are grouped together.¹⁸ Our discussion proceeds from
503 simple to more complex and statistically rigorous analyses.

504

505 4.1 *Regression studies using observed global-mean temperature data*

506

507 One class of study that has attempted to address the causes of recent tropospheric temperature
508 change relies on global-mean observational data only (*Jones, 1994; Christy and McNider, 1994;*
509 *Michaels and Knappenberger, 2000; Douglass and Clader, 2002*). Such work uses a multiple
510 regression model to quantify the statistical relationships between various “predictor variables”
511 (typically time series of ENSO variability, volcanic aerosol loadings, and solar irradiance) and a
512 single “predictand” (typically T_{2LT}). The aim is to remove the effects of the selected predictors on
513 tropospheric temperature, and to estimate the residual trend that may arise from human-induced
514 forcings. The quoted values for this residual trend in T_{2LT} range from 0.04 to 0.09°C/decade.¹⁹

515

516 These studies make the unrealistic assumption that the uncertainties inherent in such statistical
517 signal separation exercises are very small. They do not explore the sensitivity of regression results
518 to uncertainties in the predictor variables, and generally use solar and volcanic *forcings* as
519 predictors rather than the climate *responses* to those forcings. Distinctions between forcing and

¹⁸It is useful to mention one technical issue relevant to model-data comparisons. As noted in Chapter 2, the satellite-based Microwave Sounding Unit (MSU) monitors the temperature of very broad atmospheric layers. To facilitate comparisons with observed MSU datasets, many of the studies reported on here calculate “synthetic” MSU temperatures from climate model experiments. Technical aspects of these calculations are discussed in Chapter 2, Box 2.

¹⁹The studies by *Jones (1994)* and *Christy and McNider (1994)* remove volcano and ENSO effects from T_{2LT} , and estimate residual trends of 0.093 and 0.090°C/decade over 1979 to 1993. A similar investigation by *Michaels and Knappenberger (2000)* obtained a residual trend of 0.041°C/decade over 1979 to 1999. The error bars on these residual trend estimates are either not given, or claimed to be very small (*e.g.*, $\pm 0.005^\circ\text{C}/\text{decade}$ in *Christy and McNider*). A fourth study removed combined ENSO, volcano, and solar effects from T_{2LT} , and estimated a residual trend of $0.065 \pm 0.012^\circ\text{C}/\text{decade}$ over 1979 to 2000 (*Douglass and Clader, 2000*).

520 response are important (*Wigley et al.*, 2005a). Accounting for uncertainties in predictor variables
521 (and use of responses rather than forcings as predictors) expands the range of uncertainties in
522 estimates of residual T_{2LT} trends (*Santer et al.*, 2001).²⁰

523
524 Regression methods have also been used to estimate the net effects of ENSO and volcanoes on
525 trends in global-mean surface and tropospheric temperatures. For T_{2LT} , both *Jones* (1994) and
526 *Christy and McNider* (1994) found that ENSO effects induced a small net warming of 0.03 to
527 0.05°C/decade over 1979 to 1993, while volcanoes caused a cooling of 0.18°C/decade over the
528 same period. *Michaels and Knappenberger* (2000) also reported a relatively small ENSO influence
529 on T_{2LT} trends.²¹ *Santer et al.* (2001) noted that over 1979 to 1997, volcanoes had likely cooled the
530 troposphere by more than the surface. Removing the combined volcano and ENSO effects from
531 surface and UAH T_{2LT} data helped to explain some of the observed differential warming: the “raw”
532 T_S -minus- T_{2LT} trend over 1979 to 1997 decreased from roughly 0.15°C/decade to 0.05°-
533 0.13°C/decade.²² Removal of volcano and ENSO influences also brought observed lapse rate trends
534 closer to model results, but could not fully reconcile modeled and observed lapse rate trends.²³

535

²⁰*Santer et al.* (2001) obtain residual T_{2LT} trends ranging from 0.06 to 0.16°C/decade over 1979 to 1999. Their regression model is iterative, and involves removal of ENSO and volcano effects only.

²¹Their T_{2LT} trends were 0.04°C/decade over 1979 to 1998 and 0.01°C/decade over 1979 to 1999. This difference in the net ENSO influence on T_{2LT} (with the addition of only a single year of record) arises from the El Niño event in 1997/98, and illustrates the sensitivity of this kind of analysis to so-called “end effects”.

²²The latter results were obtained with the HadCRUTv surface data (*Jones et al.*, 2001) and version d03 of the UAH T_{2LT} data. The range of residual lapse-rate trends arises from parametric uncertainty, *i.e.*, from the different choices of ENSO predictor variable and volcano parameters.

²³*Santer et al.* (2001) analyzed model experiments performed with the ECHAM4/OPYC model developed at the Max-Planck Institute for Meteorology in Hamburg (*Roeckner et al.*, 1999). The experiments included forcing by well-mixed greenhouse gases, direct and indirect sulfate aerosol effects, tropospheric and stratospheric ozone, and volcanic aerosols (Pinatubo only).

536 4.2 *Regression studies using spatially-resolved temperature data*

537

538 Other regression studies have attempted to remove natural variability influences using spatially-
539 resolved temperature data. Regression is performed “locally” at individual grid-points and/or
540 atmospheric levels. To obtain a clearer picture of volcanic effects on atmospheric temperatures,
541 *Free and Angell* (2002) removed the effects of variability in ENSO and the Quasi-Biennial
542 Oscillation (QBO) from Hadley Centre radiosonde data²⁴. Their work clearly shows that the
543 cooling effect of massive volcanic eruptions has been larger in the upper troposphere than in the
544 lower troposphere. The implication is that volcanic effects probably contribute to slow changes in
545 observed lapse rates.

546

547 *Hegerl and Wallace* (2002) used regression methods to identify and remove different components
548 of natural climate variability from gridded fields of surface temperature data, the UAH T_{2LT}, and
549 “synthetic” T_{2LT} calculated from radiosonde data. They focused on the variability associated with
550 ENSO and the so-called “cold ocean warm land” (COWL) pattern (*Wallace et al.*, 1995). While
551 ENSO and COWL variability made significant contributions to the month-to-month and year-to-
552 year variability of temperature differences between the surface and T_{2LT}, it had very little impact on
553 decadal fluctuations in lapse rate. The authors concluded that natural variability alone was unlikely
554 to explain these slow lapse-rate changes. However, the removal of ENSO and COWL effects more
555 clearly revealed a volcanic contribution, consistent with the findings of *Santer et al.* (2001) and

²⁴The HadRT2.1 dataset of Parker *et al.* (1997). Like *Santer et al.* (2001), *Free and Angell* (2002) also found some sensitivity of the estimated volcanic signals to “parametric” uncertainty.

556 *Free and Angell (2002)*. A climate model control run (with no changes in forcings) and a 20CEN
557 experiment were unable to replicate the observed decadal changes in lapse rate.²⁵

558

559 4.3 *Other studies of global and tropical lapse-rate trends*

560

561 Several studies have investigated lapse-rate trends without attempting to remove volcano effects or
562 natural climate noise. *Brown et al. (2000)* used surface, radiosonde, and satellite data to identify
563 slow, tropic-wide changes in the lower tropospheric lapse rate.²⁶ In their analysis, the surface
564 warmed relative to the troposphere between the early 1960s and mid-1970s and after the early
565 1990s. Between these two periods, the tropical troposphere warmed relative to the surface. The
566 spatial coherence of these variations (and independent evidence of concurrent variations in the
567 tropical general circulation) led *Brown et al. (2000)* to conclude that tropical lapse rate changes
568 were unlikely to be an artifact of residual errors in the observations.

569

570 Very similar decadal changes in lower tropospheric lapse rate were reported by *Gaffen et al.*
571 (2000).²⁷ Their study analyzed radiosonde-derived temperature and lapse rate changes over two
572 periods: 1960 to 1997 and 1979 to 1997. Tropical lapse rates decreased over the longer period²⁸

²⁵The model was the ECHAM4/OPYC CGCM used by *Bengtsson et al. (1999)*. The 20CEN experiment analyzed by *Hegerl and Wallace (2002)* involved combined changes in well-mixed greenhouse gases, the direct and indirect effects of sulfate aerosols, and tropospheric ozone. Forcing by volcanoes and stratospheric ozone depletion was not included.

²⁶The *Brown et al. (2000)* study employed UKMO surface data (HadCRUT), version d of the UAH T_{2LT}, and an early version of the Hadley Centre radiosonde dataset (HadRT2.0) that was uncorrected for instrumental biases.

²⁷*Gaffen et al. (2000)* used a different radiosonde dataset from that employed by *Brown et al. (2000)*. The two groups also analyzed different surface temperature datasets.

²⁸Corresponding to a tendency towards a more stable atmosphere.

573 and increased over the satellite era.²⁹ To evaluate whether natural climate variability could explain
574 these slow variations, *Gaffen et al.* (2000) computed lapse rates from the control runs performed
575 with three different CGCMs. Each control run was 300 years in length. These long runs provided
576 estimates of the “sampling variability” of modeled lapse rate changes on timescales relevant to the
577 two observational periods (38 and 19 years).³⁰ Model-based estimates of natural climate variability
578 could not explain the observed tropical lapse rate changes over 1979 to 1997. Similar conclusions
579 were reached by *Hansen et al.* (1995) and *Santer et al.* (2000). Including natural and anthropogenic
580 forcings in the latter study narrowed the gap between modeled and observed estimates of recent
581 lapse-rate changes, although a significant discrepancy between the two still remained.

582
583 It should be emphasized that *all* of the studies reported on to date in Section 4 relied on satellite
584 data from one group only (UAH), on early versions of the radiosonde data³¹, and on experiments
585 performed with earlier model “vintages.” It is likely, therefore, that this work may have
586 underestimated the structural uncertainties in observed and simulated estimates of lapse rate
587 changes. We will consider in Section 5 whether modeled and observed lapse rate changes can be
588 better reconciled by the availability of more recent 20CEN runs and more comprehensive estimates
589 of structural uncertainties in observations.

590

²⁹These lapse-rate changes were accompanied by increases and decreases in tropical freezing heights (which were inferred from the same radiosonde data).

³⁰This was done by generating, for each control run, distributions of 38-year and 19-year lapse rate trends. For example, a 300-year control run can be split up into 15 different “segments” that are each of length 19 years (assuming there is no overlap between segments). From these segments, one obtains 15 different estimates of how the lapse rate might vary in the absence of any forcing changes. The observed lapse rate change over 1979 to 1997 is then compared with the model trend distribution to determine whether the observed result could be explained by natural variability alone.

³¹These radiosonde datasets were either unadjusted for inhomogeneities, or had not been subjected to the rigorous adjustment procedures used in more recent work (*Lanzante et al.*, 2003; *Thorne et al.*, 2005).

591 4.4 *Pattern-based “fingerprint” detection studies*

592

593 Fingerprint detection studies rely on *patterns* of temperature change (Box 5.5). The patterns are
594 typically either latitude-longitude “maps” (*e.g.*, for T_4 , T_2 , T_S , *etc.*) or latitude-height cross-sections
595 through the atmosphere.³² The basic premise in fingerprinting is that different climate forcings have
596 different characteristic patterns of temperature response (“fingerprints”), particularly in the free
597 atmosphere (Chapter 1, Figure 1.3; *Hansen et al.*, 1997, 2002, 2005a; *Bengtsson et al.*, 1999;
598 *Santer et al.*, 1996; *Tett et al.*, 1996).

599

600

601

602

603

604

605

606

607

608

609

610

611

³²In constructing these cross-sections, the temperature changes are generally averaged along individual bands of latitude. Zonal averages are then displayed at individual pressure levels, starting at the lowest model or radiosonde level and ending at the top of the model atmosphere or highest reported radiosonde level (see, *e.g.*, Chapter 1, Figure 3).

612 Box 5.5: Fingerprint Studies

613

614 Detection and attribution (“D&A”) studies attempt to represent an observed climate dataset as a
615 linear combination of the climate signals (“fingerprints”) arising from different forcing factors and
616 the noise of natural internal climate variability (Section 4.4). A number of different fingerprint
617 methods have been applied to the problem of identifying human-induced climate change. Initial
618 studies used relatively simple pattern correlation methods (*Barnett and Schlesinger, 1987; Santer*
619 *et al., 1996; Tett et al., 1996*). Later work involved variants of the “optimal detection” approach
620 suggested by *Hasselmann (1979, 1993, 1997)*.³³ These are essentially regression-based techniques
621 that seek to estimate the strength of a given fingerprint pattern in observational data (*i.e.*, how
622 much a given fingerprint pattern has to be scaled up or down in order to best match observations).
623 For example, if the regression coefficient for a GHG-induced T_S fingerprint is significantly
624 different from zero, GHG effects are deemed to be “detected” in observed surface temperature
625 records. Attribution tests address the question of whether these regression coefficients are also
626 consistent with unity – in other words, whether the size of the model fingerprint is consistent with
627 its amplitude in observations (*e.g.*, *Allen and Tett, 1999; Mitchell et al., 2001*).

628

629 There are two broad classes of regression-based D&A methods (*Mitchell et al., 2001*). One class
630 assumes that although the fingerprint’s amplitude changes over time, its spatial pattern does not
631 (*Hegerl et al., 1996, 1997; Santer et al., 2003a,b, 2004*). The second class explicitly considers both
632 the spatial structure and time evolution of the fingerprint (*Allen and Tett, 1999; Allen et al., 2005;*
633 *Stott and Tett, 1998; Stott et al., 2000; Tett et al., 1999, 2002; Barnett et al., 2001, 2005*). This is
634 particularly useful if the time evolution of the fingerprint contains specific information (such as a
635 periodic 11-year solar cycle) that may help to distinguish it from natural internal climate variability
636 (*North et al., 1995; North and Stevens, 1998*).

637 A number of choices must be made in applying D&A methods to real-world problems. One of the
638 most important decisions relates to “reduction of dimensionality”. D&A methods require some
639 knowledge of the correlation structure of natural climate variability.³⁴ This structure is difficult to
640 estimate reliably, even from long model control runs, because the number of time samples available
641 to estimate correlation behavior is typically much smaller than the number of spatial points in the
642 field. In practice, the total amount of spatial information (the “dimensionality”) must be reduced.
643 This is often done by using a mathematical tool (Empirical Orthogonal Functions) to reduce a
644 complex space-time dataset to a very small number of spatial patterns (“EOFs”) that capture most
645 of the information content of the dataset.³⁵ Different analysts use different procedures to determine
646 the number of patterns to retain. Further decisions relate to the choice of data used for estimating

³³*Hasselmann (1979)* noted that the engineering field had extensive familiarity with the problem of identifying coherent signals embedded in noisy data, and that many of the techniques routinely used in signal processing were transferable to the problem of detecting a human-induced climate change signal.

³⁴The relationship between variability at different points in a spatial field.

³⁵The number of patterns retained is often referred to as the “truncation dimension”. How the truncation dimension should be determined is a key decision in optimal detection studies (*Hegerl et al., 1996; Allen and Tett, 1999*).

647 fingerprint and noise, the number of fingerprints considered, the selection of observational data, the
648 treatment of missing data, *etc.*³⁶

649

650 D&A methods have some limitations. They do not work well if fingerprints are highly uncertain, or
651 if the fingerprints arising from two different forcings are similar.³⁷ They make at least two
652 important assumptions: that model-based estimates of natural climate variability are a reliable
653 representation of “real-world” variability, and that the sum of climate responses to individual
654 forcing mechanisms is equivalent to the response obtained when these factors are varied in concert.
655 Testing the validity of both assumptions remains an important research activity (*Allen and Tett,*
656 *1999; Santer et al., 2003a; Gillett et al., 2004a*).

657

658 Most analysts rely on a climate model to provide physically-based estimates of each fingerprint’s
659 structure, size, and evolution. The model fingerprints are searched for in observational climate
660 records, using rigorous statistical methods to quantify the degree of correspondence with observed
661 patterns of climate change.³⁸ Fingerprints are also compared with patterns of climate change in
662 model control runs. This helps to determine whether the correspondence between the fingerprint
663 and observations is truly significant, or could arise through internal variability alone (Box 5.5).
664 Model errors in internal variability³⁹ can bias detection results, although most detection work tries
665 to guard against this possibility by performing “consistency checks” on modeled and observed
666 variability (*Allen and Tett, 1999*), and by using variability estimates from multiple models (*Hegerl*
667 *et al., 1997; Santer et al., 2003a,b*).

³⁶Another important choice determines whether global-mean changes are included or removed from the detection analysis. Removal of global means focuses attention on smaller-scale features of modeled and observed climate-change patterns, and provides a more stringent test of model performance.

³⁷This problem is known as “degeneracy”. Formal tests of fingerprint degeneracy are sometimes applied (*e.g., Tett et al., 2002*).

³⁸The fingerprint can be either the response to an individual forcing or a combination of forcings. One strategy, for example, is to search for the climate fingerprint in response to combined changes in a suite of different human-caused forcings.

³⁹For example, current CGCMs fail to simulate the stratospheric temperature variability associated with the QBO or with solar-induced changes in stratospheric ozone (*Haigh, 1994*). Such errors may help to explain why one particular CGCM underestimated observed temperature variability in the equatorial stratosphere (*Gillett, 2000*). In the same model, however, the variability of temperatures and lapse rates in the tropical troposphere was in reasonable agreement with observations.

668

669 The application of fingerprint methods involves a variety of decisions, which introduce uncertainty
670 in detection results (Box 5.5). Our confidence in fingerprint detection results is increased if they are
671 shown to be consistent across a range of plausible choices of statistical method, processing options,
672 and model and observational datasets.

673

674 *Surface temperature changes*

675

676 Most fingerprint detection studies have focused on surface temperature changes. The common
677 denominator in this work is that the model fingerprints resulting from forcing by well-mixed GHGs
678 and sulfate aerosols⁴⁰ are statistically identifiable in observed surface temperature records (*Hegerl*
679 *et al.*, 1996, 1997; *North and Stevens*, 1998; *Tett et al.*, 1999, 2002; *Stott et al.*, 2000). These results
680 are robust to a wide range of uncertainties (*Allen et al.*, 2005).⁴¹ In summarizing this body of work,
681 the IPCC concluded that “There is new and stronger evidence that most of the warming observed
682 over the last 50 years is attributable to human activities” (*Houghton et al.*, 2001, page 4). The
683 causes of surface temperature change over the first half of the 20th Century are more ambiguous
684 (*IDAG*, 2005).

685

686 Most of the early fingerprint detection work dealt with global-scale patterns of surface temperature
687 change. The positive detection results obtained for “GHG-only” fingerprints were driven by model-

⁴⁰Most of this work considers only the direct scattering effects of sulfate aerosols on incoming sunlight, and not indirect aerosol effects on clouds.

⁴¹For example, to uncertainties in the applied greenhouse-gas and sulfate aerosol forcings, the model responses to those forcings, and model-based estimates of natural internal climate variability.

688 data pattern similarities at very large spatial scales (*e.g.*, at the scale of individual hemispheres, or
689 land-versus-ocean behavior). Fingerprint detection of GHG effects becomes more challenging at
690 continental or sub-continental scales.⁴² It is at these smaller scales that spatially heterogeneous
691 forcings, such as those arising from changes in aerosol loadings and land use patterns, may have
692 large impacts on regional climate (see Box 5.3 and 5.4). This is illustrated by the work of *Stott and*
693 *Tett* (1998), who found that a combined GHG and sulfate aerosol signal was identifiable at smaller
694 spatial scales than a “GHG-only” signal.

695
696 Recently, *Stott* (2003) and *Zwiers and Zhang* (2003) have claimed positive identification of the
697 continental- or even sub-continental features of combined GHG and sulfate aerosol fingerprints in
698 observed surface temperature records.⁴³ Using a variant of “classical” fingerprint methods,⁴⁴ *Min et*
699 *al.* (2005) identified a GHG signal in observed records of surface temperature change over East
700 Asia. *Karoly and Wu* (2005) suggest that GHG and sulfate aerosol effects are identifiable at even
701 smaller spatial scales (“of order 500 km in many regions of the globe”). These preliminary
702 investigations raise the intriguing possibility of formal detection of anthropogenic effects at
703 regional scales that are of direct relevance to policymakers.

704

705 *Changes in latitude/longitude patterns of atmospheric temperature or lapse rate*

706

⁴²This is partly due to the fact that natural climate noise is larger (and models are less skillful) on smaller spatial scales.

⁴³Another relevant “sub-global” detection study is that by *Karoly et al.* (2003). This showed that observed trends in a variety of area-averaged “indices” of North American climate (*e.g.*, surface temperature, daily temperature range, and the amplitude of the seasonal cycle) were consistent with model-predicted trends in response to anthropogenic forcing, but were inconsistent with model estimates of natural climate variability.

⁴⁴Involving Bayesian statistics.

707 Fingerprint methods have also been applied to spatial “maps” of changes in layer-averaged
708 atmospheric temperatures (*Santer et al.*, 2003b; *Thorne et al.*, 2003) and lapse rate (*Thorne et al.*,
709 2003). The study by *Santer et al.* (2003b) compared modeled and observed changes in T_2 and T_4 .
710 Model fingerprints were estimated from 20CEN experiments performed with PCM (see Table 5.1),
711 while observations were taken from two different satellite datasets (UAH and RSS; see *Christy et*
712 *al.*, 2003, and *Mears et al.*, 2003). The aim of this work was to assess the sensitivity of detection
713 results to structural uncertainties in observed MSU data.

714
715 For the T_4 layer, the model fingerprint of combined human and natural effects was consistently
716 detectable in both satellite datasets. In contrast, PCM’s T_2 fingerprint was identifiable in RSS data
717 (which show net warming over the satellite era), but not in UAH data (which show little overall
718 change in T_2 ; see Chapter 3). Encouragingly, once the global-mean differences between RSS and
719 UAH data were removed, the PCM T_2 fingerprint was detectable in *both* observed datasets. This
720 suggests that the structural uncertainties in RSS and UAH T_2 data are most prominent at the global-
721 mean level, and that this global-mean difference masks underlying similarities in smaller-scale
722 pattern structure (Chapter 4; *Santer et al.*, 2004).

723
724 *Thorne et al.* (2003) applied a “space-time” fingerprint method to six individual climate variables.
725 These variables contained information on patterns⁴⁵ of temperature change at the surface, in broad
726 atmospheric layers (the upper and lower troposphere), and in the lapse rates between these layers.⁴⁶

⁴⁵The “patterns” are in the form of temperature averages calculated over large areas rather than temperatures on a regular latitude/longitude grid.

⁴⁶*Thorne et al.* calculated the lapse rate changes between the surface and lower troposphere, the surface and upper troposphere, and the lower and upper troposphere.

727 *Thorne et al.* explicitly considered uncertainties in the searched-for fingerprints, the observed
728 radiosonde data⁴⁷, and in various data processing/fingerprinting options. They also assessed the
729 detectability of fingerprints arising from multiple forcings.⁴⁸ The “bottom-line” conclusion of
730 *Thorne et al.* is that two human-caused fingerprints – one arising from changes in well-mixed
731 GHGs alone, and the other due to combined GHG and sulfate aerosol effects – were robustly
732 identifiable in the observed surface, lower tropospheric, and upper tropospheric temperatures.
733 Evidence for the existence of a detectable volcanic signal was more equivocal. Volcanic and
734 human-caused fingerprints were not consistently identifiable in observed patterns of lapse rate
735 change.⁴⁹

736

737 *Changes in latitude/height profiles of atmospheric temperature*

738

739 Initial detection work with zonal-mean profiles of atmospheric temperature change used pattern
740 correlations to compare model fingerprints with radiosonde data (*Karoly et al.*, 1994; *Santer et al.*,
741 1996; *Tett et al.*, 1996; *Folland et al.*, 1998; *Sexton et al.*, 2001). These early investigations found
742 that model fingerprints of the stratospheric cooling and tropospheric warming in response to
743 increases in atmospheric CO₂ were identifiable in observations (Chapter 1, Figure 1.3a). The
744 pattern similarity between modeled and observed changes generally increased over the period of
745 the radiosonde record.

⁴⁷The model fingerprint was estimated from 20CEN runs performed with two different versions of the Hadley Centre CGCM (HadCM2 and HadCM3). Observational data were taken from two early compilations of the Hadley Centre radiosonde data (HadRT2.1 and HadRT2.1s).

⁴⁸Well-mixed greenhouse gases, the direct effects of sulfate aerosols, combined greenhouse-gas and sulfate aerosol effects, volcanic aerosols, and solar irradiance changes.

⁴⁹The failure to detect volcanic signals is probably due to the coarse time resolution of the input data (five-year averages) and the masking effects of ENSO variability in the radiosonde observations. Note that the two models employed in this work yielded different estimates of the size of the natural and human-caused fingerprints.

746
747 The inclusion of other human-induced forcings in 20CEN experiments – particularly the effects of
748 stratospheric ozone depletion and sulfate aerosols – tended to improve agreement with observations
749 (*Santer et al.*, 1996a; *Tett et al.*, 1996; *Sexton et al.*, 2001). The addition of ozone depletion cooled
750 the lower stratosphere and upper troposphere. This brought the height of the “transition level”
751 between stratospheric cooling and tropospheric warming lower down in the atmosphere, and in
752 better accord with observations (Chapter 1, Figure 1.3F). It also improved the agreement between
753 simulated and observed patterns of T_4 (*Ramaswamy et al.*, 1996), and decreased the size of the
754 “warming maximum” in the upper tropical troposphere, a prominent feature of CO₂-only
755 experiments (compare Figures 1.3A and 1.3F in Chapter 1).

756
757 Early work on the direct scattering effects of sulfate aerosols suggested that this forcing was
758 generally stronger in the Northern Hemisphere (NH) than in the Southern Hemisphere (SH), due to
759 the larger emissions of sulfur dioxide in industrialized regions of the NH. This asymmetry in the
760 distribution of anthropogenic sulfur dioxide sources should yield greater aerosol-induced
761 tropospheric cooling in the NH (*Santer et al.*, 1996a,b). Other forcings can lead to different
762 hemispheric temperature responses. Increases in atmospheric CO₂, for example, tend to warm land
763 more rapidly than ocean (Chapter 1). Since there is more land in the NH than in the SH, the
764 expected signal due to CO₂ increases is greater *warming* in the NH than in the SH. Because the
765 relative importance of CO₂ and sulfate aerosol forcings evolves in a complex way over time (*Tett et*
766 *al.*, 2002; *Hansen et al.*, 2002),⁵⁰ the “imprints” of these two forcings on NH and SH temperatures
767 must also vary with time (*Santer et al.*, 1996b; *Stott et al.*, 2005).

768

⁵⁰See, for example, Figure 1a in *Tett et al.* (2002) and Figure 8b in *Hansen et al.* (2002).

769 Initial attempts to detect sulfate aerosol effects on atmospheric temperatures did not account for
770 such slow changes in the hemispheric-scale features of the aerosol fingerprint. They searched for a
771 *time-invariant* fingerprint pattern in observed radiosonde data (*Santer et al.*, 1996a). This yielded
772 periods of agreement and periods of disagreement between the (fixed) aerosol fingerprint and the
773 time-varying effect of aerosols on atmospheric temperatures. Some have interpreted the periods of
774 disagreement as ‘evidence of absence’ of a sulfate aerosol signal (*Michaels and Knappenberger*,
775 1996). However, subsequent studies (see below) illustrate that such behavior is expected if one uses
776 a fixed sulfate aerosol fingerprint, and that it is important for detection studies to account for large
777 temporal changes in the fingerprint.

778
779 “Space-time” optimal detection schemes explicitly account for time variations in the signal pattern
780 and in observational data (Box 5.5). Results from recent “space-time” detection studies support
781 previous claims of an identifiable sulfate aerosol effect on surface temperature (*Stott et al.*, 2005)
782 and on zonal-mean profiles of atmospheric temperature (*Allen and Tett*, 1999; *Forest et al.*, 2001,
783 2002; *Thorne et al.*, 2002; *Tett et al.*, 2002; *Jones et al.*, 2003). This work also illustrates that the
784 identification of human effects on atmospheric temperatures can be achieved using tropospheric
785 temperatures alone (*Thorne et al.*, 2002). Positive detection results are not solely driven by the
786 inclusion of strong stratospheric cooling in the vertical pattern of temperature change (as has been
787 claimed by *Weber*, 1996).

788
789 In summary, fingerprint detection studies provide strong and consistent evidence that human-
790 induced changes in greenhouse gases and sulfate aerosols are identifiable in radiosonde records of
791 free atmospheric temperature change. The fingerprint evidence is much more equivocal in the case

792 of solar and volcanic signals in the troposphere. These natural signals have been detected in some
793 studies (*Jones et al.*, 2003) but not in others (*Tett et al.* 2002), and their identification appears to be
794 more sensitive to specific processing choices that are made in applying fingerprint methods (*Leroy*,
795 1998; *Thorne et al.*, 2002, 2003).

796

797 5 *New Comparisons of Modeled and Observed Temperature Changes*

798

799 In this section, we evaluate selected results from recently-completed CGCM 20CEN experiments
800 that have been performed in support of the IPCC Fourth Assessment Report (AR4). The runs
801 analyzed here were performed with 19 different models, and involve modeling groups in 10
802 different countries (Table 5.1). They use new model versions, and incorporate historical changes in
803 many (but not all) of the natural and human forcings that are thought to have influenced
804 atmospheric temperatures over the past 50 years⁵¹ (Table 5.2). These new experiments provide our
805 current best estimates of the expected climate change due to combined human and natural effects.

806

807

808

809 The new 20CEN runs constitute an “ensemble of opportunity” (*Allen and Stainforth*, 2002). The
810 selection and application of natural and anthropogenic forcings was not coordinated across
811 modeling groups.⁵² For example, only seven of the 19 modeling groups applied time-varying

⁵¹This was not the case in previous model intercomparison exercises, such as AMIP (*Gates et al.*, 1999) and CMIP2 (*Meehl et al.*, 2000).

⁵²In practice, experimental coordination is very difficult across a range of models of varying complexity and sophistication. Aerosols are a case in point. Some modeling groups that contributed 20CEN simulations to the IPCC AR4 do not have the technical capability to explicitly include aerosols, and instead attempt to represent their net radiative effects by adjusting the surface albedo.

812 changes in LULC (Table 5.2). Groups that included LULC effects did not always use the same
813 observational dataset for specifying this forcing, or apply it in the same way (Table 5.3). Only six
814 models included some representation of the indirect effects of anthropogenic aerosols, which are
815 thought to have had a net cooling influence on surface temperatures through their effects on cloud
816 properties (*Ramaswamy et al.*, 2001b).

817

818

819

820

821

822

823

824

825

826

827

828

829 Table 5.3: Forcings used in 20CEN experiments performed with the PCM, CCSM3.0, GFDL CM2.1, and GISS-EH
830 models. Grey shading denotes a forcing that was included in the experimental design. Shading indicates a forcing that
831 was not incorporated or that did not vary over the course of the experiment.

832

833 ¹See *Dai et al.* (2001) for further details.834 ²See *Meehl et al.* (2005) for further details.835 ³The chemistry transport model (MOZART; see *Horowitz et al.*, 2003; *Tie et al.*, 2005) was driven by meteorology from the Middle
836 Atmosphere version of the Community Climate Model (“MACCM”; version 3). “1990” weather from MACCM3 was used for all
837 years between 1860 and 2000.838 ⁴See *Hansen et al.* (2005a) for further det

839

840

841

842 One important implication of Table 5.3 is that model-to-model differences in the applied forcings
843 are intertwined with model-to-model differences in the climate *responses* to those forcings. This
844 makes it more difficult to isolate systematic errors that are common to a number of models, or to
845 identify problems with a specific forcing dataset. Note, however, that the lack of a coordinated
846 experimental design is also an advantage, since the “ensemble of opportunity” spans a wide range
847 of uncertainty in current estimates of climate forcings.

848

849 In addition to model forcing and response uncertainty, the 20CEN ensemble also encompasses
850 uncertainties arising from inherently unpredictable climate variability. Roughly half of the
851 modeling groups that submitted 20CEN data performed multiple realizations of their historical
852 forcing experiment (see Section 2 and Table 5.1). For example, the five-member ensemble of
853 CCSM3.0 20CEN runs contains an underlying signal (which one might define as the ensemble-
854 average climate response to the forcings varied in CCSM3.0) plus five different sequences of

855 climate noise. Such multi-member ensembles provide valuable information on the relative sizes of
856 signal and noise. In all, a total of 49 20CEN realizations were examined here.

857
858 The following Section presents preliminary results from analyses of these 20CEN runs and the new
859 observational datasets described in Chapters 2-4. Our primary focus is on the tropics, since
860 previous work by *Gaffen et al.* (2000) and *Hegerl and Wallace* (2002) suggests that this is where
861 climate models have significant problems in simulating observed lapse rates changes. We also
862 discuss comparisons of global-mean changes in atmospheric temperatures and lapse rates. We do
863 not discount the importance of comparing models and data at much smaller scales (particularly in
864 view of the incorporation of regional-scale forcing changes in many of the runs analyzed here), but
865 comprehensive regional-scale comparisons were not feasible given the limited time available for
866 completion of this report.

867
868 In order to facilitate “like with like” comparisons between modeled and observed atmospheric
869 temperature changes, we calculate synthetic MSU T_4 , T_2 , and T_{2LT} from the model 20CEN results
870 (see Chapter, Box 2). Both observed and synthetic MSU T_2 data include a contribution from the
871 cooling stratosphere (*Fu et al.*, 2004a,b), and hence complicate the interpretation of slow changes
872 in T_2 . To provide a less ambiguous measure of “bulk” tropospheric temperature changes, we use
873 the statistical approach of *Fu et al.* (2004a, 2005) to remove stratospheric influences, thereby
874 obtaining T^*_G and T^*_T in addition to T_{2LT} .⁵³ As a simple measure of lapse-rate changes, we
875 consider temperature differences between the surface and three different atmospheric layers (T_{2LT} ,

⁵³There is still some debate over the reliability of T^*_G trends estimated with the *Fu et al.* (2004a) statistical approach (*Tett and Thorne*, 2004, *Gillett et al.*, 2004; *Kiehl et al.*, 2005; *Fu et al.*, 2004b; Chapter 4). T^*_T is derived mathematically (from the overlap between the T_4 and T_2 weighting functions) rather than statistically, and is now generally accepted as a reasonable measure of temperature change in the tropical troposphere.

876 T^*_G , and T^*_T). Each of these layers samples slightly different portions of the troposphere (Figure
877 2.2).

878
879 The trend comparisons shown in Sections 5.1 and 5.2 do not involve any formal statistical
880 significance tests (see Statistical Appendix). While such tests are entirely appropriate for
881 comparisons of individual model and observational trends,⁵⁴ they are less relevant here, where we
882 compare a 49-member ensemble of model trends with a relatively small number of observationally-
883 based estimates. The model ensemble encapsulates uncertainties in climate forcings and model
884 responses, as well as the effects of climate noise on trends. The observational range characterizes
885 current structural uncertainties in historical changes. We simply assess whether the simulated trend
886 distributions do or do not overlap with these observations. Our goal here is to determine where
887 model results are qualitatively consistent with observations, and where serious inconsistencies are
888 likely to exist. This does not obviate the need for the more rigorous statistical comparisons
889 described in Box 5.5, which should be a high priority (see Recommendations).

890

891 5.1 *Global-Mean Temperature and Lapse-Rate Trends*

892

893 In all but two of the 49 20CEN realizations, the global-mean temperature of the lower stratosphere
894 experiences a net cooling over 1979 to 1999 (Figures 5.2A, 5.3A).⁵⁵ The model average T_4 trend is
895 $-0.25^\circ\text{C}/\text{decade}$ (Table 5.4A). Most of this cooling is due to the combined effects of stratospheric

⁵⁴For example, such tests have been performed by *Santer et al.* (2003b) in comparisons between observed MSU trends (in RSS and UAH) and synthetic MSU trends in four PCM 20CEN realizations.

⁵⁵In the following, all inter-model and model-data comparisons are over January 1979 to December 1999. This is the longest period of overlap (at least during the satellite era) between the model experiments (which generally end in 1999) and the satellite data (which start in 1979).

896 ozone depletion and increases in well-mixed GHGs (*Ramaswamy et al.*, 2001a,b), with the former
897 the dominant influence on T_4 changes over the satellite era (*Ramaswamy et al.*, 1996; *Santer et al.*,
898 2003a). The model average cooling is larger ($-0.35^\circ\text{C}/\text{decade}$) and closer to the satellite-based
899 estimates if it is calculated from the subset of 20CEN realizations that include forcing by ozone
900 depletion. The range of model T_4 trends encompasses the trends derived from satellites, but not the
901 larger trends estimated from radiosondes. The most likely explanation for this discrepancy is a
902 residual cooling trend in the radiosonde data (Chapter 4).⁵⁶ The neglect of stratospheric water vapor
903 increases in most of the 20CEN runs considered here (*Shine et al.*, 2003) may be another
904 contributory factor.⁵⁷

905

906

907

908

909

910

911

912

913

⁵⁶Recent work suggests that this residual trend is largest in the lower stratosphere and upper troposphere, and is largely related to temporal changes in the solar heating of the temperature sensors carried by radiosondes (and failure to properly correct for this effect; see *Sherwood et al.*, 2005; *Randel and Wu*, 2005).

⁵⁷Recent stratospheric water vapor increases are thought to be partly due to the oxidation of methane, and are expected to have a net cooling effect on T_4 . To our knowledge, CH_4 -induced stratospheric water vapor increases were explicitly incorporated in only two of the 19 models considered here (GISS-EH and GISS-ER; *Hansen et al.*, 2005a).

914

915 Figure 5.2A: Modeled and observed changes in global-mean monthly-mean lower stratospheric temperature (T_4). A
916 simple weighting function approach (Box 2.2) was used to calculate a “synthetic” T_4 (equivalent to the MSU T_4
917 monitored by satellites) from model temperature data. Synthetic T_4 results are from “20CEN” experiments performed
918 with nine different models (see Table 5.1). These models were chosen because they satisfy certain minimum
919 requirements in terms of the forcings applied in the 20CEN run: all nine were driven by changes in well-mixed GHGs,
920 sulfate aerosol direct effects, tropospheric and stratospheric ozone, volcanic aerosols, and solar irradiance (in addition
921 to other forcings; see Table 5.2). Observed satellite-based estimates of T_4 changes were obtained from both RSS and
922 UAH (see Chapter 3). All T_4 changes are expressed as departures from a 1979 to 1999 reference period average, and
923 were smoothed with the same filter. To make it easier to compare the variability of T_4 in models with different
924 ensemble sizes (see Table 5.1), only the first 20CEN realization is plotted from each model. This also facilitates
925 comparisons of modeled and observed variability.
926

927 Figure 5.2B: As for Figure 5.2A, but for time series of global-mean, monthly-mean lower tropospheric temperature
928 anomalies (T_{2LT}).
929

930 Figure 5.2C: As for Figure 5.2A, but for time series of global-mean, monthly-mean surface temperature anomalies (T_S).
931

932 Figure 5.2D: As for Figure 5.2A, but for time series of global-mean, monthly-mean temperature differences between
933 the surface and T_{2LT} .
934

935

936 Figure 5.3: Modeled and observed trends in time series of global-mean T_4 (panel A), T_2 (panel B),
937 T^*_G (panel C), T_{2LT} (panel D), T_S (panel E), T_S minus T^*_G (panel F), and T_S minus T_{2LT} . All trends
938 were calculated using monthly-mean anomaly data. The analysis period is 1979 to 1999. Model
939 results are displayed in the form of histograms. Each histogram is based on results from 49
940 individual realizations of the 20CEN experiment, performed with 19 different models (Table 5.1).
941 The applied forcings are listed in Table 5.2. The vertical red line in each panel is the mean of the
942 model trends, calculated with a sample size of $n = 19$ (see Table 5.4A). Observed trends are
943 estimated from two radiosonde and three satellite datasets (T_2), two radiosonde and two satellite
944 datasets (T_4 , T^*_G and T_{2LT}), and three different surface datasets (T_S) (see Chapter 3). The bottom
945 “rows” of the observed difference trends in panels F and G were calculated with NOAA T_S data.
946 The top “rows” of observed results in F and G were computed with HadCRUT2v T_S data. The
947 vertical offsetting of observed results in these panels (and also in panels B-E) is purely for the

948 purpose of simplifying the visual display – observed trends bear no relation to the y-axis scale. To
 949 simplify the display, the Figure does not show the statistical uncertainties arising from the fitting of
 950 linear trends to noisy data. GISS T_S trends (not shown) are very close to those estimated with
 951 NOAA T_S data (see Chapter 3).

952
 953

954 Table 5.4A: Summary statistics for global-mean temperature trends calculated from 49 different realizations of 20CEN
 955 experiments performed with 19 different coupled models. Results are for four different atmospheric layers (T_4 , T_2 , T^*_G ,
 956 and T_{2LT}), the surface (T_S), and differences between the surface and the troposphere (T_S minus T^*_G and T_S minus T_{2LT}).
 957 All trends were calculated over the 252-month period from January 1979 to December 1999 using global-mean
 958 monthly-mean anomaly data. Results are in $^{\circ}\text{C}/\text{decade}$. The values in the “Mean” column correspond to the locations of
 959 the red lines in the seven panels of Figure 5.3A. For each layer, means, medians and standard deviations were
 960 calculated from a sample size of $n = 19$, *i.e.*, from ensemble means (if available) and individual realizations (if
 961 ensembles were not performed). This avoids placing too much weight on results from a single model with a large
 962 number of realizations. Maximum and minimum values were calculated from all available realizations (*i.e.*, from a
 963 sample size of $n = 49$).

964 Superimposed on the overall cooling of T_4 are the large stratospheric warming signals in response

Layer	Mean	Median	Std. Dev. (1σ)	Minimum	Maximum
T_4	-0.252	-0.281	0.194	-0.695	0.079
T_2	0.142	0.122	0.079	0.015	0.348
T^*_G	0.181	0.167	0.077	0.052	0.375
T_{2LT}	0.198	0.186	0.070	0.058	0.394
T_S	0.164	0.156	0.062	0.052	0.333
$T_S - T^*_G$	-0.017	-0.017	0.046	-0.110	0.083
$T_S - T_{2LT}$	-0.034	-0.031	0.030	-0.099	0.052

965

966 to the eruptions of El Chichón (in April 1982) and Pinatubo (in June 1991).⁵⁸ Nine of the 19 IPCC
 967 models explicitly included volcanic aerosols (Figure 5.2A and Table 5.2).⁵⁹ Seven of these nine
 968 models overestimate the observed stratospheric warming after Pinatubo. GFDL CM2.1 simulates
 969 the Pinatubo response reasonably well, but underestimates the response to El Chichón. Differences

⁵⁸These warming signals occur because volcanic aerosols absorb both incoming solar radiation and outgoing thermal radiation (*Ramaswamy et al.*, 2001a).

⁵⁹The documentation for the Russian INM-CM3.0 model claims that volcanic aerosols were incorporated in the 20CEN run, but does not show evidence of stratospheric warming signatures after massive volcanic eruptions. This suggests that volcanic cooling effects on surface temperature were implicitly incorporated by changing the surface albedo (a procedure that would not yield volcanically-induced stratospheric warming signals).

970 in the magnitude of the applied volcanic aerosol forcings must account for some of the inter-model
971 differences in the T_4 warming signals (Table 5.3).⁶⁰

972
973 Over 1979 to 1999, the global-mean troposphere warms in all 49 20CEN simulations considered
974 here (Figures 5.2B, 5.3B-D). The shorter-term cooling signals of the El Chichón and Pinatubo
975 eruptions are superimposed on this gradual warming.⁶¹ Because of the influence of stratospheric
976 cooling on T_2 , the model average trend is smaller for this layer than for either T_{2LT} or T^*_G , which
977 are more representative of temperature changes in the bulk of the troposphere (Table 5.4A).⁶² All of
978 the satellite- and radiosonde-based trends in T_{2LT} and T^*_G are contained within the spread of model
979 results. This illustrates that there is no fundamental discrepancy between modeled and observed
980 trends in global-mean tropospheric temperature.

981
982 In contrast, the T_2 trends in both radiosonde datasets are either slightly negative or close to zero,
983 and are smaller than all of the model results. This difference is most likely due to contamination
984 from residual stratospheric and upper-tropospheric cooling biases in the radiosonde data (Chapter
985 4; *Sherwood et al.*, 2005; *Randel and Wu*, 2005). The satellite-based T_2 trends are either close to

⁶⁰More subtle details of the forcing are also relevant to interpretation of inter-model T_4 differences, such as different assumptions regarding the aerosol size distribution, the vertical distribution of the volcanic aerosol relative to the model tropopause, *etc.* Note that observed T_4 changes over the satellite era are not well-described by a simple linear trend, and show evidence of a step-like decline in stratospheric temperatures after the El Chichón and Pinatubo eruptions (*Pawson et al.*, 1998; *Seidel and Lanzante*, 2004). Model-model differences in the applied ozone forcings and solar forcings may help to explain why the GFDL, GISS, and HadGEM1 models appear to reproduce some of this step-like behavior, particularly after El Chichón, while T_4 decreases in PCM are much more linear (*Dameris et al.*, 2005; *Ramaswamy et al.*, 2006).

⁶¹Because of differences in the timing of modeled and observed ENSO events (Section 5.2), the tropospheric and surface cooling caused by El Chichón is more noticeable in all models than in observations (where it was partially masked by the large 1982/83 El Niño; Figures 5.2B,C).

⁶²Because of ozone-induced cooling of the lower stratosphere, the model-average T_2 trend is slightly smaller (0.12°C/decade) and closer to the RSS result if it is estimated from the subset of 20CEN runs that include stratospheric ozone depletion. Subsetting in this way has little impact on the model-average T_{2LT} and T^*_G trends.

986 the model average (RSS and VG) or just within the model range (UAH; Fig. 5.3B). Even without
987 formal statistical tests, it is clear that observational uncertainty is an important factor in assessing
988 the consistency between modeled and observed changes in mid- to upper tropospheric temperature
989 (*Santer et al.*, 2003b).

990

991 Observed T_S trends closely bracket the model average (Figures 5.2C, 5.3E). There is no evidence of
992 a serious inconsistency between modeled and observed surface temperature changes. Structural
993 uncertainties in observed T_S trends are much smaller than for trends in T_4 or tropospheric layer-
994 average temperatures (see Chapter 4).

995

996 The model-simulated ranges of lapse-rate trends also encompass virtually all observational results
997 (Figures 5.3F,G).⁶³ Closer inspection reveals that the model-average trends in tropospheric lapse
998 rate are slightly negative,⁶⁴ indicating larger warming aloft than at the surface. Most combinations
999 of observed T_S , T^*_G , and T_{2LT} datasets yield the converse result, and show smaller warming aloft
1000 than at the surface. As in the case of global-mean T^*_G and T_{2LT} trends, RSS-based lapse-rate trends
1001 are invariably closest to the model average results. Both models and observations show a tendency
1002 towards positive values of T_S minus T_{2LT} for several years after the El Chichón and Pinatubo
1003 eruptions, indicative of larger cooling aloft than at the surface (Figure 5.2D; Section 5.4).

1004

1005 5.2 *Tropical Temperature and Lapse-Rate Trends*

1006

⁶³Note that the subtraction of temperature variability common to surface and troposphere decreases (by about a factor of two) the large range of model trends in T_S , T^*_G , and T_{2LT} (Table 5.4A).

⁶⁴Values are $-0.02^\circ\text{C}/\text{decade}$ in the case of T_S minus T^*_G and $-0.03^\circ\text{C}/\text{decade}$ for T_S minus T_{2LT} .

1007 The previous section examined whether simulated global-mean temperature trends were contained
1008 within current estimates of structural uncertainty in observations. Since ENSO is primarily a
1009 tropical phenomenon, its influence on surface and tropospheric temperature is more pronounced in
1010 the tropics than in global averages. Observations contain only one specific sequence of ENSO
1011 fluctuations from 1979 to present, and only one sequence of ENSO effects on tropical
1012 temperatures. The model 20CEN runs examined here provide many different sequences of ENSO
1013 variability. We therefore expect – and find – that these runs yield a wide range of trends in tropical
1014 surface and tropospheric temperature (Figure 5.4)⁶⁵. It is of interest whether this large model range
1015 encompasses the observed trends.

1016
1017 At the surface, results from the multi-model ensemble include all observational estimates of
1018 tropical temperature trends (Figure 5.4E; Table 5.4B). Observed results are close to the model
1019 average T_S trend of $+0.16^\circ\text{C}/\text{decade}$. There is no evidence that the models significantly over- or
1020 underestimate the observed surface warming. In the troposphere, all observational results are still
1021 within the range of possible model solutions, but the majority of model results show tropospheric
1022 warming that is larger than observed (Figures 5.4B-D). As in the case of the global-mean T_4 trends,
1023 the cooling of the tropical stratosphere in both radiosonde datasets is larger than in any of the

⁶⁵This would be true even for a hypothetical “perfect” climate model run with “perfect” forcings. This large model range of tropical temperature trends is not solely due to the effects of ENSO and other modes of internal variability. It also arises from uncertainties in the models and forcings (see Box 5.2 and Table 5.2). Note that the trends discussed here are calculated over a relatively short period of time (several decades).

1024 satellite datasets or model results (Figure 5.4A).⁶⁶ The UAH and RSS T₄ trends are close to the
1025 model average.⁶⁷

⁶⁶This supports recent findings of a residual cooling bias in tropical radiosonde data (*Sherwood et al., 2005; Randel and Wu, 2005*).

⁶⁷The model average is $-0.27^{\circ}\text{C}/\text{decade}$ when estimated from the subset of 20CEN runs that include stratospheric ozone depletion.

1026

1027 Figure 5.4: As for Figure 5.3, but for trends in the tropics (20°N-20°S).

1028

1029

1030 Table 5.4B: As for Table 5.4A, but for tropical temperature trends (calculated from spatial averages over 20°N-20°S).

1031

1032

1033 In the model results, trends in the two measures of tropical lapse-rate (T_S minus T_{2LT} and T_S minus
1034 T^*_T) are almost invariably negative, indicating larger warming aloft than at the surface (Figure
1035 5.4F,G). Similar behavior is evident in only one of the four upper-air datasets examined here
1036 (RSS).⁶⁸ The RSS trends are just within the range of model solutions.⁶⁹ Tropical lapse-rate trends in
1037 both radiosonde datasets and in the UAH satellite data are always positive (larger warming at the
1038 surface than aloft), and lie outside the range of model results.

1039

1040 This comparison suggests that discrepancies between our current best estimates of simulated and
1041 observed lapse-rate changes may be larger and more serious in the tropics than in globally-
1042 averaged data. Large structural uncertainties in the observations (even in the sign of the trend in

⁶⁸Note that the VG group do not provide either a stratospheric or lower-tropospheric temperature retrieval, and so could not be included in the comparison of modeled and observed trends in T_S minus T^*_T or T_S minus T_{2LT} .

⁶⁹Three of the four RSS-based results in Figures 5.4F and 5.4G are within two standard deviations of the model average values (see Table 5.4B).

1043 tropical lapse-rate changes) make it difficult to reach more definitive conclusions regarding the
1044 significance and importance of model-data discrepancies (see Section 5.4).

1045

1046 5.3 *Spatial Patterns of Lapse-Rate Trends*

1047

1048 Maps of the trends in lower tropospheric lapse rate help to identify geographical regions where the
1049 model-data discrepancies in Figures 5.4F and 5.4G are most pronounced. We focus first on four
1050 U.S. models: CCSM3.0, PCM, GFDL CM2.1, and GISS-EH (Table 5.3). These show qualitatively
1051 similar patterns of trends in T_S minus T_{2LT} (Figures 5.5A-D). Over most of the tropical ocean, the
1052 simulated warming is larger in the troposphere than at the surface. All models have some tropical
1053 land areas where the surface warms relative to the troposphere. The largest relative warming of the
1054 surface occurs at high latitudes in both hemispheres.

1055

1056 Figure 5.5: Modeled and observed maps of the differences between trends in T_S and T_{2LT} . All trends in T_S and T_{2LT}
1057 were calculated over the 252-month period from January 1979 to December 1999. Model results are ensemble means
1058 from 20CEN experiments performed with CCSM3.0 (panel A), PCM (panel B), GFDL CM2.1 (panel C), and GISS-EH
1059 (panel D). Observed results rely on NOAA T_S trends and on two different satellite estimates of trends in T_{2LT} , obtained
1060 from UAH (panel E) and RSS (panel F). White denotes high elevation areas where it is not meaningful to calculate
1061 synthetic T_{2LT} (panels A-D). Note that RSS mask T_{2LT} values in such regions, while UAH do not (*c.f.* panels F, E).
1062

1063 To illustrate structural uncertainties in the observed data, we show two different patterns of trends
1064 in T_S minus T_{2LT} . Both rely on the same NOAA surface data, but use either UAH (Figure 5.5E) or
1065 RSS (Figure 5.5F) as their source of T_{2LT} results. The “NOAA minus UAH” combination provides
1066 a picture that is very different from the model results, with coherent warming of the surface relative
1067 to the troposphere over much of the world’s tropical oceans. While “NOAA minus RSS” also has

1068 relative warming of the surface in the Western and tropical Pacific, it shows relative warming of
1069 the *troposphere* in the eastern tropical Pacific and Atlantic Oceans. This helps to clarify why
1070 simulated lapse-rate trends in Figures 5.4F and 5.4G are closer to NOAA minus RSS results than to
1071 NOAA minus UAH results.

1072
1073 As pointed out by *Santer et al.* (2003b) and *Christy and Spencer* (2003), we cannot use such
1074 model-data comparisons alone to determine whether the UAH or RSS T_{2LT} dataset is closer to (an
1075 unknown) “reality.” As the next section will show, however, models and basic theory can be used
1076 to identify aspects of observational behavior that require further investigation, and may help to
1077 constrain observational uncertainty.

1078

1079 5.4 *Tropospheric Amplification of Surface Temperature Changes*

1080

1081 When surface and lower tropospheric temperature changes are spatially averaged over the deep
1082 tropics, and when day-to-day tropical temperature changes are averaged over months, seasons, or
1083 years, it is evident that temperature changes aloft are larger than at the surface. This “amplification”
1084 behavior has been described in many observational and modeling studies, and is a consequence of
1085 the release of latent heat by moist convecting air (*e.g.*, *Manabe and Stouffer*, 1980; *Horel and*
1086 *Wallace*, 1981; *Pan and Oort*, 1983; *Yulaeva and Wallace*, 1994; *Hurrell and Trenberth*, 1998;
1087 *Soden*, 2000; *Wentz and Schabel*, 2000; *Hegerl and Wallace*, 2002; *Knutson and Tuleya*, 2004).⁷⁰

⁷⁰The essence of tropical atmospheric dynamics is that the tropics cannot support large temperature gradients, so waves (Kelvin, Rossby, gravity) even out the temperature field between convecting and non-convective regions. The temperature field throughout the tropical troposphere is more or less on the moist adiabat lapse rate set by convection over the warmest waters. This is why there is a trade wind inversion where this profile finds itself inconsistent with boundary layer temperatures in the colder regions.

1088
1089 A recent study by *Santer et al.* (2005) examined this amplification behavior in the same 20CEN
1090 runs and observational datasets considered in the present report. The sole difference (relative to the
1091 data used here) was that *Santer et al.* analyzed a version of the UAH T_{2LT} data that had not yet been
1092 adjusted for a recently-discovered error (*Mears and Wentz*, 2005).⁷¹ The amplification of tropical
1093 surface temperature changes was assessed on different timescales (monthly, annual, and multi-
1094 decadal) and in different atmospheric layers (T^*_T and T_{2LT}).

1095
1096 On short timescales (month-to-month and year-to-year variations in temperature), the estimated
1097 tropospheric amplification of surface temperature changes was in good agreement in all model and
1098 observational datasets considered, and was in accord with basic theory. This is illustrated in Figure
1099 5.6, which shows the standard deviations of monthly-mean T_S anomalies plotted against the
1100 standard deviations of monthly-mean anomalies of T_{2LT} (panel A) and T^*_T (panel B). All model
1101 and observational results lie above the black line indicating equal temperature variability aloft and
1102 at the surface. All have similar “amplification factors” between their surface and tropospheric
1103 variability.⁷² In the models, these similarities occur despite differences in physics, resolution, and
1104 forcings, and despite a large range (roughly a factor of 5) in the size of simulated temperature
1105 variability. In observations, the scaling ratios estimated from monthly temperature variability are
1106 relatively unaffected by the structural uncertainties discussed in Chapter 4.

⁷¹The error was related to the UAH group’s treatment of systematic drifts in the time of day at which satellites sample Earth’s diurnal temperature cycle (see Chapter 4).

⁷²Note that the slope of the red regression lines that has been fitted to the model results is slightly steeper for T^*_T than for T_{2LT} (*c.f.* panels 5.6A and 5.6B). This is because T^*_T samples more of the mid-troposphere than T_{2LT} (see Prospectus). Amplification is expected to be larger in the mid-troposphere than in the lower troposphere.

1107

1108 Figure 5.6: Scatter plots showing the relationships between tropical temperature changes at Earth's surface and in two
1109 different layers of the troposphere. All results rely on temperature data that have been spatially-averaged over the deep
1110 tropics (20°N-20°S). Model data are from 49 realizations of 20CEN runs performed with 19 different models (Table
1111 5.1). Observational results were taken from four different upper-air datasets (two from satellites, and two from
1112 radiosondes) and two different surface temperature datasets (see Chapter 3). The two upper panels provide information
1113 on the month-to-month variability in T_S and T_{2LT} (panel A) and in T_S and T^*_T (panel B). The two bottom panels
1114 consider temperature changes on multi-decadal timescales, and show the trends (over 1979 to 1999) in T_S and T_{2LT}
1115 (panel C) and in T_S and T^*_T (panel D). The red line in each panel is the regression line through the model points. Its
1116 slope provides information on the amplification of surface temperature variability and trends in the free troposphere.
1117 The black line in each panel is given for reference purposes, and has a slope of 1. Values above (below) the black lines
1118 indicate tropospheric amplification (damping) of surface temperature changes. There are two columns of observational
1119 results in C and D. These are based on the NOAA and HadCRUT2v TS (0.12 and 0.14°C/decade, respectively). Note
1120 that panel C show results from published and recently-revised versions of the UAH T_{2LT} data (versions 5.1 and 5.2).
1121 Since the standard deviations calculated from NOAA and HadCRUT2v monthly T_S anomalies are very similar,
1122 observed results in A and B use NOAA standard deviations only. The blue shading in the bottom two panels defines
1123 the region of simultaneous surface warming and tropospheric cooling.

1124

1125

1126 A different picture emerges if amplification behavior is estimated from decadal changes in tropical

1127 temperatures. Figures 5.6C and 5.6D show multi-decadal trends in T_S plotted against trends in T_{2LT}

1128 and T^*_T . The 20CEN runs exhibit amplification factors that are consistent with those estimated
1129 from month-to-month and year-to-year temperature variability.⁷³ Only one observational upper-air
1130 dataset (RSS) shows amplified warming aloft, and similar amplification relationships on short and
1131 on long timescales. The other observational datasets have scaling ratios less than 1, indicating
1132 tropospheric damping of surface warming (*Fu et al.*, 2005; *Santer et al.*, 2005).⁷⁴

1133
1134 These analyses shed further light on the differences between modeled and observed changes in
1135 tropical lapse rates described in Section 5.2. They illustrate the usefulness of comparing models
1136 and data on different timescales. On short timescales, it is evident that models successfully capture
1137 the basic physics that controls “real world” amplification behavior. On long timescales, model-data
1138 consistency is sensitive to structural uncertainties in the observations. One possible interpretation
1139 of these results is that in the real world, different physical mechanisms govern amplification
1140 processes on short and on long timescales, and models have some common deficiency in
1141 simulating such behavior. If so, these “different physical mechanisms” need to be identified and
1142 understood.

1143
1144 Another interpretation is that the same physical mechanisms control short- and long-term
1145 amplification behavior. Under this interpretation, residual errors in one or more of the observed

⁷³As in the case of amplification factors inferred from short-timescale variability, the factors estimated from multi-decadal temperature changes are relatively insensitive to inter-model differences in physics and the applied forcings (see Table 5.3). At first glance, this appears to be a somewhat surprising result in view of the large spatial and temporal heterogeneity of certain forcings (see Section 3). Black carbon aerosols, for example, are thought to cause localized heating of the troposphere relative to the surface (Box 5.3), a potential mechanism for altering amplification behavior. The fact that amplification factors are similar in experiments that include and exclude black carbon aerosols suggests that aerosol-induced tropospheric heating is not destroying the connection of large areas of the tropical ocean to a moist adiabatic lapse rate. Single-forcing experiments (see Recommendations) will be required to improve our understanding of the physical effects of black carbon aerosols and other spatially-heterogeneous forcings on tropical temperature-change profiles.

⁷⁴The previous version of the UAH T_{2LT} data yielded a negative amplification factor for multi-decadal changes in tropical temperatures.

1146 datasets must affect their representation of long-term trends, and must lead to different scaling
1147 ratios on short and long timescales. This explanation appears to be the more likely one in view of
1148 the large structural uncertainties in observed upper-air datasets (Chapter 4) and the complementary
1149 physical evidence supporting recent tropospheric warming (see Section 6).

1150

1151 “Model error” and “observational error” are not mutually exclusive explanations for the
1152 amplification results shown in Figures 5.6C and D. Although a definitive resolution of this issue
1153 has not yet been achieved, the path towards such resolution is now more obvious. We have learned
1154 that models show considerable consistency in terms of what they tell us about tropospheric
1155 amplification of surface warming. This consistency holds on a range of different timescales.
1156 Observations display consistent amplification behavior on short timescales, but radically different
1157 behavior on long timescales. Clearly, not all of the observed lapse-rate trends can be equally
1158 probable. Intelligent use of “complementary evidence” – from the behavior of other climate
1159 variables, from remote sensing systems other than MSU, and from more systematic exploration of
1160 the impacts of different data adjustment choices – should ultimately help us to constrain
1161 observational uncertainty, and reach more definitive conclusions regarding the true significance of
1162 modeled and observed lapse-rate differences.

1163

1164 5.5 *Vertical Profiles of Atmospheric Temperature Change*

1165

1166 Although formal fingerprint studies have not yet been completed with atmospheric temperature-
1167 change patterns estimated from the new 20CEN runs, it is instructive to make a brief qualitative
1168 comparison of these patterns. This helps to address the question of whether the inclusion of

1169 previously-neglected forcings (like carbonaceous aerosols and land use/land cover changes; see
1170 Section 2) has fundamentally modified the “fingerprint” of human-induced atmospheric
1171 temperature changes searched for in previous detection studies.

1172
1173 We examine the zonal-mean profiles of atmospheric temperature change in 20CEN runs performed
1174 with four U.S. models (CCSM3, PCM, GFDL CM2.1, and GISS-EH). All four show a common
1175 large-scale fingerprint of stratospheric cooling and tropospheric warming over 1979 to 1999
1176 (Figures 5.7A-D). The pattern of temperature change estimated from HadAT2 radiosonde data is
1177 broadly similar, although the transition height between stratospheric cooling and tropospheric
1178 warming is noticeably lower than in the model simulations (Figure 5.7E). Another noticeable
1179 difference is that the HadAT2 data show a relative lack of warming in the tropical troposphere,⁷⁵
1180 where all four models simulate maximum warming. This particular aspect of the observed
1181 temperature-change pattern is very sensitive to data adjustments (*Sherwood et al.*, 2005; *Randel*
1182 *and Wu*, 2005). Tropospheric warming in the observations is most obvious in the NH extratropics,
1183 where our confidence in the reliability of radiosonde records is greatest.

⁷⁵Despite the “end point” effect of the large El Niño event in 1997-1998 (see Chapter 3).

1184

1185

1186 Figure 5.7: Zonal-mean patterns of atmospheric temperature change in “20CEN” experiments performed with four
1187 different climate models and in observational radiosonde data. Model results are for CCSM3.0 (panel A), PCM (panel
1188 B), GFDL CM 2.1 (panel C), and GISS-EH (panel D). The model experiments are ensemble means. There are
1189 differences between the sets of climate forcings that the four models used in their 20CEN runs (Table 5.3). Observed
1190 changes (panel E) were estimated with HadAT2 radiosonde data (*Thorne et al.*, 2005, and Chapter 3). The HadAT2
1191 temperature data do not extend above 30 hPa, and have inadequate coverage at high latitudes in the Southern
1192 Hemisphere. All temperature changes were calculated from monthly-mean data and are expressed as linear trends (in
1193 °C/decade) over 1979 to 1999.
1194

1195
1196 Note that some of the details of the model fingerprint pattern are quite different. For example,
1197 GFDL’s cooling maximum immediately above the tropical tropopause is not evident in any of the
1198 other models. Its maximum warming in the upper tropical troposphere is noticeably larger than in
1199 CCSM3.0, PCM, or GISS-EH. While CCSM and GFDL CM2.1 have pronounced hemispheric
1200 asymmetry in their stratospheric cooling patterns, with largest cooling at high latitudes in the SH,⁷⁶
1201 this asymmetry is less apparent in PCM and GISS-EH.

1202
1203 Future work should consider whether the conclusions of detection studies are robust to such
1204 fingerprint differences. This preliminary analysis suggests that the large-scale “fingerprint” of
1205 stratospheric cooling and tropospheric warming over the satellite era – a robust feature of previous
1206 detection work – has not been fundamentally altered by the inclusion of hitherto-neglected forcings
1207 like carbonaceous aerosols and LULC changes (see Table 5.3). This does not diminish the need to
1208 quantify the individual contributions of these forcings in appropriate “single forcing” experiments.

1209
1210 6 *Changes in “Complementary” Climate Variables*

1211
1212 Body temperature is a simple metric of our physical well-being. A temperature of 40°C (104°F) is
1213 indicative of an illness, but does not by itself identify the cause of the illness. In medicine,
1214 investigation of causality typically requires the analysis of many different lines of evidence.
1215 Similarly, analyses of temperature alone provide incomplete information on the causes of climate
1216 change. For example, there is evidence that major volcanic eruptions affect not only the Earth’s

⁷⁶This may be related to an asymmetry in the pattern of stratospheric ozone depletion: the largest ozone decreases over the past 2-3 decades have occurred at high latitudes in the SH.

1217 radiation budget (*Wielicki et al.*, 2002; *Soden et al.*, 2002) and atmospheric temperatures (*Hansen*
1218 *et al.*, 1997, 2002; *Free and Angell*, 2002; *Wigley et al.*, 2005a), but also water vapor (*Soden et al.*,
1219 2002), precipitation (*Gillett et al.*, 2004c), atmospheric circulation patterns (see, *e.g.*, *Robock*, 2000,
1220 and *Ramaswamy et al.*, 2001a; *Robock and Oppenheimer*, 2003), ocean heat content and sea level
1221 (*Church et al.*, 2005), and even global-mean surface pressure (*Trenberth and Smith*, 2005). These
1222 responses are physically interpretable and internally consistent.⁷⁷ The combined evidence from
1223 changes in all of these variables makes a stronger case for an identifiable volcanic effect on climate
1224 than evidence from a single variable only.

1225
1226 A “multi-variable” perspective may also be beneficial in understanding the possible causes of
1227 differential warming. The value of “complementary” climate datasets for studying this specific
1228 problem has been recognized by *Pielke* (2004) and by *Wentz and Schabel* (2000). The latter found
1229 internally-consistent increases in SST, T_{2LT} , and marine total column water vapor over the 12-year
1230 period from 1987 to 1998.⁷⁸ Multi-decadal increases in surface and lower tropospheric water vapor
1231 were also reported in the IPCC Second Assessment Report (*Folland et al.*, 2001).⁷⁹ More recently,

⁷⁷The physical consistency between the temperature and water vapor changes after the Pinatubo eruption has been clearly demonstrated by *Soden et al.* (2002). The surface and tropospheric cooling induced by Pinatubo caused a global-scale reduction in total column water vapor. Since water vapor is a strong GHG, the reduction in water vapor led to less trapping of outgoing thermal radiation by Earth’s atmosphere, thus *amplifying* the volcanic cooling. This is referred to as a “positive feedback.” *Soden et al.* “disabled” this feedback in a climate model experiment, and found that the “no water vapor feedback” model was incapable of simulating the observed tropospheric cooling after Pinatubo. Inclusion of the water vapor feedback yielded close agreement between the simulated and observed T_{2LT} responses to Pinatubo. This suggests that the model used by *Soden et al.* captures important aspects of the physics linking the real world’s temperature and moisture changes.

⁷⁸The Wentz and Schabel study used NOAA optimally-interpolated SST data, a version of the UAH T_{2LT} data that had been corrected for orbital decay effects, and information on total column water vapor from the satellite-based Special Sensor Microwave Imager (SSM/I).

⁷⁹More specifically, *Folland et al.* (2001) concluded that “Changes in water vapour mixing ratio have been analysed for selected regions using *in situ* surface observations as well as lower-tropospheric measurements based on satellites and weather balloons. A pattern of overall surface and lower-tropospheric water vapour mixing ratio increases over the past few decades is emerging, although there are likely to be some time-dependent biases in these data and regional variations in trends. The more reliable data sets show that it is likely that total atmospheric water vapour has increased

1232 *Trenberth et al.* (2005) found significant increases in total column water vapor over the global
1233 ocean.⁸⁰ At constant relative humidity, water vapor increases nonlinearly with increasing
1234 temperature (*Hess*, 1959). Slow increases in tropospheric water vapor therefore provide
1235 circumstantial evidence in support of tropospheric warming. However, water vapor measurements
1236 are affected by many of the same data quality and temporal homogeneity problems that influence
1237 temperature measurements (*Elliott*, 1995; *Trenberth et al.*, 2005), so the strength of this
1238 circumstantial evidence is still questionable.⁸¹

1239
1240 Other climate variables also corroborate the warming of Earth's surface over the second half of the
1241 20th Century. Examples include increases in ocean heat content (*Levitus et al.*, 2000, 2005; *Willis et*
1242 *al.*, 2004), sea-level rise (*Cabanes et al.*, 2001), thinning of major ice sheets and ice shelves
1243 (*Krabill et al.*, 1999; *Rignot and Thomas*, 2002; *Domack et al.*, 2005), and widespread glacial
1244 retreat, with accelerated rates of glacial retreat over the last several decades (*Arendt et al.*, 2002;
1245 *Paul et al.*, 2004).⁸²

1246

several per cent per decade over many regions of the Northern Hemisphere since the early 1970s. Changes over the Southern Hemisphere cannot yet be assessed”.

⁸⁰*Trenberth et al.* (2005) reported an increase in total column water vapor over 1988 to 2001 of “ $1.3 \pm 0.3\%$ per decade for the ocean as a whole, where the error bars are 95% confidence intervals.” This estimate was obtained with an updated version of the SSM/I dataset analyzed by *Wentz and Schabel* (2000).

⁸¹Note, however, that SSM/I-derived water vapor measurements may have some advantages relative to temperature measurements obtained from MSU. *Wentz and Schabel* (2000) point out that (under a constant relative humidity assumption), the 22 GHz water vapor radiance observed by SSM/I is three times more sensitive to changes in air temperature than the MSU T₂ 54 GHz radiance. Furthermore, while drift in sampling the diurnal cycle influences MSU-derived tropospheric temperatures (Chapter 4), it has a much smaller impact on SSM/I water vapor measurements.

⁸²*Folland et al.* (2001) note that “Long-term monitoring of glacier extent provides abundant evidence that tropical glaciers are receding at an increasing rate in all tropical mountain areas”. Accelerated retreat of high-elevation tropical glaciers is occurring within the tropical lower tropospheric layer that is a primary focus of this report, and provides circumstantial support for warming of this layer over the satellite era”.

1247 Changes in some of these “complementary” variables have been used in detection and attribution
1248 studies. Much of this work has focused on ocean heat content. When driven by anthropogenic
1249 forcing, a number of different CGCMs capture the overall increase in observed ocean heat content
1250 estimated by *Levitus et al.* (2000; 2005), but not the large decadal variability in heat content
1251 (*Barnett et al.*, 2001; *Levitus et al.*, 2001; *Reichert et al.*, 2002; *Sun and Hansen*, 2003; *Pielke*,
1252 2003; *Gregory et al.*, 2004; *Hansen et al.*, 2005b).⁸³ It is still unclear whether this discrepancy
1253 between simulated and observed variability is primarily due to model deficiencies or is an artifact
1254 of how *Levitus et al.* (2000; 2005) “infilled” data-sparse ocean regions (*Gregory et al.*, 2004;
1255 *AchutaRao et al.*, 2005).

1256
1257 In summary, the behavior of complementary variables enhances our confidence in the reality of
1258 large-scale warming of the Earth’s surface, and tells us that the signature of this warming is
1259 manifest in many different aspects of the climate system. Pattern-based fingerprint detection work
1260 performed with ocean heat content (*Barnett et al.*, 2001; *Reichert et al.*, 2002; *Barnett et al.*, 2005;
1261 *Pierce et al.*, 2005), sea-level pressure (*Gillett et al.*, 2003), and tropopause height (*Santer et al.*,
1262 2003a, 2004) suggests that anthropogenic forcing is necessary in order to explain observed changes
1263 in these variables. This supports the findings of the surface- and atmospheric temperature studies
1264 described in Section 4.4. To date, however, investigations of complementary variables have not
1265 enabled us to narrow uncertainties in satellite- and radiosonde-based estimates of tropospheric
1266 temperature change over the past two-and-a-half decades.⁸⁴ Formal detection and attribution studies

⁸³Model control runs cannot generate such large multi-decadal increases in the heat content of the global ocean.

⁸⁴The tropopause is the transition zone between the turbulently-mixed troposphere, where most weather occurs, and the more stably-stratified stratosphere (see Preface and Chapter 1). Increases in tropopause height over the past 3–4 decades represent an integrated response to temperature changes above and below the tropopause (*Highwood et al.*, 2000; *Santer et al.*, 2004), and are evident in both radiosonde data (*Highwood et al.*, 2000; *Seidel et al.*, 2001) and reanalyses (*Randel et al.*, 2000). In model 20CEN simulations, recent increases in tropopause height are driven by the

1267 involving water vapor changes may be helpful in this regard, since observations suggest a recent
1268 moistening of the troposphere, consistent with tropospheric warming.

1269

1270 7 *Summary*

1271

1272 This chapter has evaluated a wide range of scientific literature dealing with the possible causes of
1273 recent temperature changes, both at the Earth's surface and in the free atmosphere. It shows that
1274 many factors – both natural and human-related – have probably contributed to these changes.
1275 Quantifying the relative importance of these different climate forcings is a difficult task. Analyses
1276 of observations alone cannot provide us with definitive answers. This is because there are
1277 important uncertainties in the observations and in the climate forcings that have affected them.
1278 Although computer models of the climate system are useful in studying cause-effect relationships,
1279 they, too, have limitations. Advancing our understanding of the causes of recent lapse-rate changes
1280 will best be achieved by comprehensive comparisons of observations, models, and theory – it is
1281 unlikely to arise from analysis of a single model or observational dataset.

1282

combined effects of GHG-induced tropospheric warming and ozone-induced stratospheric cooling (*Santer et al.*, 2003a). Available reanalysis products do not provide a consistent picture of the relative contributions of stratospheric and tropospheric temperature changes to recent tropopause height increases (*Pielke and Chase*, 2004; *Santer et al.*, 2004).

1283 **References**

- 1284
- 1285 AchutaRao, K.M., *et al.*, 2005: Variability of ocean heat uptake: Reconciling observations and
1286 models. *Journal of Geophysical Research (Oceans)* (in press).
- 1287
- 1288 Allen, M.R., *et al.*, 2005: Quantifying anthropogenic influence on recent near-surface temperature.
1289 *Surveys in Geophysics* (in press).
- 1290
- 1291 Allen, M.R., and D.A. Stainforth, 2002: Towards objective probabilistic climate forecasting.
1292 *Nature*, **419**, 228.
- 1293
- 1294 Allen, M., 1999: Do-it-yourself climate prediction. *Nature*, **401**, 642.
- 1295
- 1296 Allen, M.R., and S.F.B. Tett, 1999: Checking for model consistency in optimal fingerprinting.
1297 *Climate Dynamics*, **15**, 419-434.
- 1298
- 1299 Ammann, C.M., *et al.*, 2003: A monthly and latitudinally varying forcing dataset in simulations of
1300 20th century climate. *Geophysical Research Letters*, **30**, 1657, doi:10.1029/2003GL016875.
- 1301
- 1302 Andronova, N.G., and M.E. Schlesinger, 2001: Objective estimation of the probability density
1303 function for climate sensitivity. *Journal of Geophysical Research (Atmospheres)*, **106**,
1304 22605-22611.
- 1305
- 1306 Andronova, N.G., *et al.*, 1999: Radiative forcing by volcanic aerosols from 1850 to 1994. *Journal*
1307 *of Geophysical Research (Atmospheres)*, **104**, 16807-16826.
- 1308
- 1309 Arendt, A.A., *et al.*, 2002: Rapid wastage of Alaska glaciers and their contribution to rising sea
1310 level. *Science*, **297**, 382-386.
- 1311
- 1312 Barnett, T.P. *et al.*, 2005: Penetration of human-induced warming into the world's oceans. *Science*,
1313 **309**, 284-287.
- 1314
- 1315 Barnett, T.P., D.W. Pierce, and R. Schnur, 2001: Detection of anthropogenic climate change in the
1316 world's oceans. *Science*, **292**, 270-274.
- 1317
- 1318 Barnett, T.P., and M.E. Schlesinger, 1987: Detecting changes in global climate induced by
1319 greenhouse gases. *Journal of Geophysical Research (Atmospheres)*, **92**, 14772-14780.
- 1320
- 1321 Bengtsson, L., E. Roeckner, and M. Stendel, 1999: Why is the global warming proceeding much
1322 slower than expected? *Journal of Geophysical Research (Atmospheres)*, **104**, 3865-3876.
- 1323
- 1324 Brown, S.J., *et al.*, 2000: Decadal variability in the lower-tropospheric lapse rate. *Geophysical*
1325 *Research Letters*, **27**, 997-1000.
- 1326
- 1327 Brovkin, V., *et al.*, 2004: Role of land cover changes for atmospheric CO₂ increase and climate
1328 change during the last 150 years. *Global Change Biology*, **10**, 1253-1266.

- 1329
1330 Cabanes, C., A. Cazenave, and C. Le Provost, 2001: Sea level rise during past 40 years determined
1331 from satellite and in situ observations. *Science*, **294**, 840-842.
1332
- 1333 Chase, T.N., *et al.*, 2004: Likelihood of rapidly increasing surface temperatures unaccompanied by
1334 strong warming in the free troposphere. *Climate Research*, **25**, 185-190.
1335
- 1336 Christy, J.R., and R.T. McNider, 1994: Satellite greenhouse signal. *Nature*, **367**, 325.
1337
- 1338 Christy, J.R., and R. Spencer, 2003: Reliability of satellite data sets. *Science*, **301**, 1046-1047.
1339
- 1340 Christy, J.R., *et al.*, 2003: Error estimates of version 5.0 of MSU-AMSU bulk atmospheric
1341 temperatures. *Journal of Atmospheric and Oceanic Technology*, **20**, 613-629.
1342
- 1343 Church, J.A., N.J. White, and J.M. Arblaster, 2005: Significant decadal-scale impact of volcanic
1344 eruptions on sea level and ocean heat content. *Nature*, **438**, doi:10.1038/nature04237.
1345
- 1346 Collins, W.D., *et al.*, 2005: The Community Climate System Model: CCSM3. *Journal of Climate*
1347 (*accepted*).
1348
- 1349 Crowley, T.J., 2000: Causes of climate change over the past 1,000 years. *Science*, **289**, 270-277.
1350
- 1351 Dai., A., *et al.*, 2001: Climates of the Twentieth and Twenty-First centuries simulated by the
1352 NCAR Climate System Model. *Journal of Climate*, **14**, 485-519.
1353
- 1354 Dameris, M., *et al.*, 2005: Long-term changes and variability in a transient simulation with a
1355 chemistry-climate model employing realistic forcing. *Atmospheric Chemistry and Physics*,
1356 **5**, 2121-2145.
1357
- 1358 Delworth, T.L., *et al.*, 2005: GFDL's CM2 global coupled climate models – Part 1: Formulation
1359 and simulation characteristics. *Journal of Climate* (in press).
1360
- 1361 Domack, E., *et al.*, 2005: Stability of the Larsen B ice shelf on the Antarctic Peninsula during the
1362 Holocene epoch. *Nature*, **436**, 681-685.
1363
- 1364 Douglass, D.H., and R.S. Knox, 2005: Climate forcing by the volcanic eruption of Mount Pinatubo.
1365 *Geophysical Research Letters*, **32**, L05710, doi:10.1029/2004GL022119.
1366
- 1367 Douglass, D.H., B.D. Pearson, and S.F. Singer, 2004: Altitude dependence of atmospheric
1368 temperature trends: Climate models versus observation. *Geophysical Research Letters*, **31**,
1369 doi:10.1029/2004GL020103.
1370
- 1371 Douglass, D.H., and B.D. Clader, 2002: Climate sensitivity of the Earth to solar irradiance.
1372 *Geophysical Research Letters*, **29**, doi:10.1029/2002GL015345.
1373

- 1374 Elliott, W.P., 1995: On detecting long-term changes in atmospheric moisture. *Climatic Change*, **31**,
1375 349-367.
1376
- 1377 Feddema, J., *et al.*, 2005: An evaluation of GCM sensitivity to land cover change experiments, and
1378 their potential importance to IPCC scenario simulations. *Journal of Climate* (accepted).
1379
- 1380 Free, M., and J.K. Angell, 2002: Effect of volcanoes on the vertical temperature profile in
1381 radiosonde data. *Journal of Geophysical Research (Atmospheres)*, **107**,
1382 doi:10.1029/2001JD001128.
1383
- 1384 Folland, C.K., *et al.*, 2001: Observed climate variability and change. In: *Climate Change 2001: The*
1385 *Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the*
1386 *Intergovernmental Panel on Climate Change* [Houghton, J.T., *et al.*, (eds.)]. Cambridge
1387 University Press, Cambridge, United Kingdom and New York, NY, USA, 881 pp.
1388
- 1389 Folland, C.K., *et al.*, 1998: Influences of anthropogenic and oceanic forcing on recent climate
1390 change. *Geophysical Research Letters*, **25**, 353-356.
1391
- 1392 Forest, C.E., *et al.*, 2002: Quantifying uncertainties in climate system properties with the use of
1393 recent climate observations. *Science*, **295**, 113-117.
1394
- 1395 Forest, C.E., *et al.*, 2001: Constraining climate model properties using optimal fingerprint detection
1396 studies. *Climate Dynamics*, **18**, 277-295.
1397
- 1398 Fu, Q., and C.M. Johanson, 2005: Satellite-derived vertical dependence of tropical tropospheric
1399 temperature trends. *Geophysical Research Letters*, **32**, L10703,
1400 doi:10.1029/2004GL022266.
1401
- 1402 Fu, Q., *et al.*, 2004a: Contribution of stratospheric cooling to satellite-inferred tropospheric
1403 temperature trends. *Nature*, **429**, 55-58.
1404
- 1405 Fu, Q., *et al.*, 2004b: Reply to “Tropospheric temperature series from satellites” and “Stratospheric
1406 cooling and the troposphere”. *Nature*, doi:10.1038/nature03208.
1407
- 1408 Gaffen, D.J., *et al.*, 2000: Multidecadal changes in the vertical structure of the tropical troposphere.
1409 *Science*, **287**, 1242-1245.
1410
- 1411 Gates, W.L., *et al.*, 1999: An overview of the results of the Atmospheric Model Intercomparison
1412 Project (AMIP I). *Bulletin of the American Meteorological Society*, **80**, 29-55.
1413
- 1414 Gillett, N.P. *et al.*, 2004a: Testing the linearity of the response to combined greenhouse gas and
1415 sulfate aerosol forcing. *Geophysical Research Letters*, **31**, L14201,
1416 doi:10.1029/2004GL020111.
1417

- 1418 Gillett, N.P., B.D. Santer, and A.J. Weaver, 2004b: Quantifying the influence of stratospheric
1419 cooling on satellite-derived tropospheric temperature trends. *Nature*,
1420 doi:10.1038/nature03209.
- 1421
- 1422 Gillett, N.P., *et al.*, 2004c: Detection of volcanic influence on global precipitation. *Geophysical*
1423 *Research Letters*, **31**, doi:10.1029/2004GL020044.
- 1424
- 1425 Gillett, N.P., *et al.*, 2003: Detection of human influence on sea level pressure. *Nature*, **422**, 292-
1426 294.
- 1427
- 1428 Gillett, N.P., M.R. Allen, and S.F.B. Tett, 2000: Modelled and observed variability in atmospheric
1429 vertical temperature structure. *Climate Dynamics*, **16**, 49-61.
- 1430
- 1431 Gregory, J.M., *et al.*, 2004: Simulated and observed decadal variability in ocean heat content.
1432 *Geophysical Research Letters*, **31**, L15312, doi:10.1029/2004GL020258.
- 1433
- 1434 Gregory, J.M., *et al.*, 2002: An observationally based estimate of the climate sensitivity. *Journal of*
1435 *Climate*, **15**, 3117-3121.
- 1436
- 1437 Grody, N.C., *et al.*, 2004: Calibration of multisatellite observations for climatic studies: Microwave
1438 Sounding Unit (MSU). *Journal of Geophysical Research (Atmospheres)*, **109**, D24104,
1439 doi:10.1029/2004JD005079.
- 1440
- 1441 Haigh, J.D., 1994: The role of stratospheric ozone in modulating the solar radiative forcing of
1442 climate. *Nature*, **370**, 544-546.
- 1443
- 1444 Hansen, J., *et al.*, 2005a: Efficacy of climate forcings. *Journal of Geophysical Research*
1445 *(Atmospheres)*, **110**, D18104, doi:10.1029/2005JD005776.
- 1446
- 1447 Hansen, J., *et al.*, 2005b: Earth's energy imbalance: Confirmation and implications. *Science*, **308**,
1448 1431-1435.
- 1449
- 1450 Hansen, J., and L. Nazarenko, 2003: Soot climate forcing via snow and ice albedos. *Proceedings of*
1451 *the National Academy of Sciences*, **101**, 423-428, doi:10.1073/pnas.2237157100.
- 1452
- 1453 Hansen, J., 2002: A brighter future. *Climatic Change*, **52**, 435-440.
- 1454
- 1455 Hansen, J., *et al.*, 2002: Climate forcings in Goddard Institute for Space Studies SI2000
1456 simulations. *Journal of Geophysical Research (Atmospheres)*, **107**,
1457 doi:10.1029/20001JD001143.
- 1458
- 1459 Hansen, J., *et al.*, 2000: Global warming in the Twenty-First Century: An alternative scenario.
1460 *Proceedings of the National Academy of Sciences*, **97**, 9875-9880.
- 1461
- 1462 Hansen, J., *et al.*, 1997: Forcings and chaos in interannual to decadal climate change. *Journal of*
1463 *Geophysical Research (Atmospheres)*, **102**, 25679-25720.

- 1464
1465 Hansen, J., *et al.*, 1995: Satellite and surface temperature data at odds? *Climatic Change*, **30**, 103-
1466 117.
1467
1468 Hansen, J., *et al.*, 1993: How sensitive is the Earth's Climate? *Natl. Geog. Res. Explor.*, **9**, 142-158.
1469
1470 Harvey, L.D.D., and R.K. Kaufmann, 2002: Simultaneously constraining climate sensitivity and
1471 aerosol radiative forcing. *Journal of Climate*, **15**, 2837-2861.
1472
1473 Hasselmann, K., 1997: Multi-pattern fingerprint method for detection and attribution of climate
1474 change. *Climate Dynamics*, **13**, 601-612.
1475
1476 Hasselmann, K., 1993: Optimal fingerprints for the detection of time dependent climate change.
1477 *Journal of Climate*, **6**, 1957-1971.
1478
1479 Hasselmann, K., 1979: In: *Meteorology of Tropical Oceans* (Ed. D.B. Shaw). Royal
1480 Meteorological Society of London, London, U.K., pp. 251-259.
1481
1482 Hegerl, G.C., and J.M. Wallace, 2002: Influence of patterns of climate variability on the difference
1483 between satellite and surface temperature trends. *Journal of Climate*, **15**, 2412-2428.
1484
1485 Hegerl, G.C., *et al.*, 1997: Multi-fingerprint detection and attribution of greenhouse-gas- and
1486 aerosol-forced climate change. *Climate Dynamics*, **16**, 737-754.
1487
1488 Hegerl, G.C., *et al.*, 1996: Detecting greenhouse-gas-induced climate change with an optimal
1489 fingerprint method. *Journal of Climate*, **9**, 2281-2306.
1490
1491 Hess, S.L., 1959: *Introduction to Theoretical Meteorology*. Holt, Rinehart and Winston, New York,
1492 362 pp.
1493
1494 Highwood, E.J., B.J. Hoskins, and P. Berrisford, 2000: Properties of the Arctic tropopause.
1495 *Quarterly Journal of the Royal Meteorological Society*, **126**, 1515-1532.
1496
1497 Hoffert, M.I., and C. Covey, 1992: Deriving global climate sensitivity from paleoclimate
1498 reconstructions. *Nature*, **360**, 573-576.
1499
1500 Horel, J.D., and J.M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the
1501 Southern Oscillation. *Monthly Weather Review*, **109**, 813-829.
1502
1503 Horowitz, L.W. *et al.*, 2003: A global simulation of tropospheric ozone and related tracers:
1504 Description and evaluation of MOZART, version 2, *Journal of Geophysical Research*
1505 (*Atmospheres*), **108**, 4784, doi:10.1029/2002JD002853.
1506
1507 Houghton, J.T., *et al.*, 2001: *Climate Change 2001: The Scientific Basis*. Cambridge University
1508 Press, Cambridge, U.K., 881 pp.
1509

- 1510 Hoyt, D.V., and K.H. Schatten, 1993: A discussion of plausible solar irradiance variations, 1700-
1511 1992. *Journal of Geophysical Research (Atmospheres)*, **98**, 18895-18906.
- 1512
- 1513 Hurrell, J.W., *et al.*, 2003: *The North Atlantic Oscillation: Climatic Significance and*
1514 *Environmental Impact*. American Geophysical Union, Geophysical Monograph 134, 279
1515 pp.
- 1516
- 1517 Hurrell, J.W., and K.E. Trenberth, 1998: Difficulties in obtaining reliable temperature records:
1518 Reconciling the surface and satellite Microwave Sounding Unit records. *Journal of Climate*,
1519 **11**, 945-967.
- 1520
- 1521 Hurtt, G., *et al.*, 2006: Three centuries of gridded, global land-use transition rates and wood harvest
1522 statistics for Earth System Model applications. *Global Change Biology* (submitted).
- 1523
- 1524 International Detection and Attribution Group (IDAG), 2005: Detecting and attributing external
1525 influences on the climate system: A review of recent advances. *Journal of Climate*, **18**,
1526 1291-1314.
- 1527
- 1528 Jacobson, M.Z., 2004: Climate response of fossil fuel and biofuel soot, accounting for soot's
1529 feedback to snow and sea ice albedo and emissivity. *Journal of Geophysical Research*
1530 *(Atmospheres)*, **109**, D21201, doi:10.1029/2004JD004945.
- 1531
- 1532 Jones, P.D., *et al.*, 2003: Surface climate responses to explosive volcanic eruptions seen in long
1533 European temperature records and mid-to-high latitude tree-ring density around the
1534 Northern Hemisphere. In: *Volcanism and the Earth's Atmosphere*, A. Robock and C.
1535 Oppenheimer (Eds.), AGU Geophysical Monograph Series, **139**, Washington D.C.
1536 American Geophysical Union, 239-254.
- 1537
- 1538 Jones, P.D., *et al.*, 2001: Adjusting for sampling density in grid box land and ocean surface
1539 temperature time series. *Journal of Geophysical Research (Atmospheres)*, **106**, 3371-3380.
- 1540
- 1541 Jones, P.D., *et al.*, 1999: Surface air temperature and its changes over the past 150 years. *Reviews*
1542 *of Geophysics*, **37**, 173-199.
- 1543
- 1544 Jones, P.D., 1994: Recent warming in global temperature series. *Geophysical Research Letters*, **21**,
1545 1149-1152.
- 1546
- 1547 Jones, G.S., S.F.B. Tett, and P.A. Stott, 2003: Causes of atmospheric temperature change 1960-
1548 2000: A combined attribution analysis. *Geophysical Research Letters*, **30**,
1549 doi:10.1029/2002GL016377.
- 1550
- 1551 Karoly, D.J., and Q. Wu, 2005: Detection of regional surface temperature trends. *Journal of*
1552 *Climate* (submitted).
- 1553
- 1554 Karoly, D.J., *et al.*, 2003: Detection of a human influence on North American climate. *Science*,
1555 **302**, 1200-1203.

- 1556
1557 Karoly, D.J., *et al.*, 1994: An example of fingerprint detection of greenhouse climate change.
1558 *Climate Dynamics*, **10**, 97-105.
1559
- 1560 Kiehl, J.T., J.M. Caron, and J.J. Hack, 2005: On using global climate model simulations to assess
1561 the accuracy of MSU retrieval methods for tropospheric warming trends. *Journal of*
1562 *Climate*, **18**, 2533-2539.
1563
- 1564 Klein Goldewijk, K., 2001: Estimating global land use change over the past 300 years: The HYDE
1565 database. *Global Biogeochemical Cycles*, **15**, 417-433.
1566
- 1567 Knutson, T.R., and R.E. Tuleya, 2004: Impact of CO₂-induced warming on simulated hurricane
1568 intensity and precipitation: Sensitivity to the choice of climate model and convective
1569 parameterization. *Journal of Climate*, **17**, 3477-3495.
1570
- 1571 Koch, D., 2001: Transport and direct radiative forcing of carbonaceous and sulfate aerosols in the
1572 GISS GCM. *Journal of Geophysical Research (Atmospheres)*, **106**, 20311-20332.
1573
- 1574 Koch, D., *et al.*, 1999: Tropospheric sulfur simulation and sulfate direct radiative forcing in the
1575 Goddard Institute for Space Studies general circulation model. *Journal of Geophysical*
1576 *Research (Atmospheres)*, **104**, 23799-23822.
1577
- 1578 Krabill, W., *et al.*, 1999: Rapid thinning of parts of the Southern Greenland Ice Sheet. *Science*, **283**,
1579 1522-1524.
1580
- 1581 Krishnan, R., and V. Ramanathan, 2002: Evidence of surface cooling from absorbing aerosols.
1582 *Geophysical Research Letters*, **29**, doi:10.1029/2002GL014687.
1583
- 1584 Lean, J., 2000: Evolution of the Sun's spectral irradiance since the Maunder Minimum.
1585 *Geophysical Research Letters*, **27**, 2425-2428.
1586
- 1587 Lean, J., J. Beer, and R. Bradley, 1995: Reconstruction of solar irradiance since 1610: Implications
1588 for climate change. *Geophysical Research Letters*, **22**, 3195-3198.
1589
- 1590 Leroy, S.S., 1998: Detecting climate signals: Some Bayesian aspects. *Journal of Climate*, **11**, 640-
1591 651.
1592
- 1593 Levitus, S., J.I. Antonov, and T.P. Boyer, 2005: Warming of the world ocean, 1955-2003.
1594 *Geophysical Research Letters*, **32**, L02604, doi:10.1029/2004GL021592.
1595
- 1596 Levitus, S., *et al.*, 2001: Anthropogenic warming of Earth's climate system. *Science*, **292**, 267-270.
1597
- 1598 Levitus, S., *et al.*, 2000: Warming of the world ocean. *Science*, **287**, 2225-2229.
1599
- 1600 Lindzen, R.S., and C. Giannitsis, 1998: On the climatic implications of volcanic cooling. *Journal of*
1601 *Geophysical Research (Atmospheres)*, **103**, 5929-5941.

- 1602
1603 Manabe, S., and R.J. Stouffer, 1980: Sensitivity of a global climate model to an increase of CO₂
1604 concentration in the atmosphere. *Journal of Geophysical Research (Atmospheres)*, **85**,
1605 5529-5554.
- 1606
1607 Marshall, C.H., *et al.*, 2004: The impact of anthropogenic land-cover change on the Florida
1608 Peninsula sea breezes and warm season sensible weather. *Monthly Weather Review*, **132**,
1609 28-52.
- 1610
1611 Matthews, H.D., *et al.*, 2004: Natural and anthropogenic climate change: incorporating historical
1612 land cover change, vegetation dynamics and the global carbon cycle. *Climate Dynamics*, **22**,
1613 461-479.
- 1614
1615 Matthews, H.D., *et al.*, 2003: Radiative forcing of climate by historical land cover change.
1616 *Geophysical Research Letters*, **30**, 1055, doi:10.1029/2002GL016098.
- 1617
1618 McAvaney, B.J., *et al.*, 2001: Model evaluation. In: *Climate Change 2001: The Scientific Basis.*
1619 *Contribution of Working Group I to the Third Assessment Report of the Intergovernmental*
1620 *Panel on Climate Change* [Houghton, J.T., *et al.*, (eds.)]. Cambridge University Press,
1621 Cambridge, United Kingdom and New York, NY, USA, 881 pp.
- 1622
1623 Mears, C.A., and F.W. Wentz, 2005: The effect of diurnal correction on satellite-derived lower
1624 tropospheric temperature. *Science*, **309**, 1548-1551.
- 1625
1626 Mears, C.A., M.C. Schabel, and F.W. Wentz, 2003: A reanalysis of the MSU channel 2
1627 tropospheric temperature record. *Journal of Climate*, **16**, 3650-3664.
- 1628
1629 Meehl, G.A., *et al.*, 2005: Climate change in the 20th and 21st centuries and climate change
1630 commitment in the CCSM3. *Journal of Climate* (in press).
- 1631
1632 Meehl, G.A., *et al.*, 2000: The Coupled Model Intercomparison Project (CMIP). *Bulletin of the*
1633 *American Meteorological Society*, **81**, 313-318.
- 1634
1635 Meehl, G.A., 1984: Modeling the Earth's climate. *Climatic Change*, **6**, 259-286.
- 1636
1637 Menon, S., *et al.*, 2002: Climate effects of black carbon aerosols in China and India. *Science*, **297**,
1638 2250-2253.
- 1639
1640 Michaels, P.J., and P.C. Knappenberger, 2000: Natural signals in the MSU lower tropospheric
1641 temperature record. *Geophysical Research Letters*, **27**, 2905-2908.
- 1642
1643 Michaels, P.J., and P.C. and Knappenberger, 1996: Human effect on global climate? *Nature*, **384**,
1644 523-524.
- 1645

- 1646 Min, S.-K., A. Hense, and W.-T. Kwon, 2005: Regional-scale climate change detection using a
1647 Bayesian detection method. *Geophysical Research Letters*, **32**, L03706,
1648 doi:10.1029/2004GL021028.
1649
- 1650 Minschwaner, K., *et al.*, 1998: Infrared radiative forcing and atmospheric lifetimes of trace species
1651 based on observations from UARS. *Journal of Geophysical Research (Atmospheres)*, **103**,
1652 23243-23253.
1653
- 1654 Mitchell, J.F.B., *et al.*, 2001: Detection of climate change and attribution of causes. In: *Climate*
1655 *Change 2001: The Scientific Basis. Contribution of Working Group I to the Third*
1656 *Assessment Report of the Intergovernmental Panel on Climate Change* [Houghton, J.T., *et*
1657 *al.*, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY,
1658 USA, 881 pp.
1659
- 1660 Murphy, J.M., *et al.*, 2004: Quantification of modeling uncertainties in a large ensemble of climate
1661 simulations. *Nature*, **430**, 768-772.
1662
- 1663 Myhre, G., and A. Myhre, 2003: Uncertainties in radiative forcing due to surface albedo changes
1664 caused by land use changes. *Journal of Climate*, **16**, 1511-1524.
1665
- 1666 National Research Council, 2005: *Radiative Forcing of Climate Change: expanding the Concept*
1667 *and Addressing Uncertainties*. Board on Atmospheric Sciences and Climate, National
1668 Academy Press, Washington D.C., 168 pp.
1669
- 1670 National Research Council, 2000: *Reconciling Observations of Global Temperature Change*. Board
1671 on Atmospheric Sciences and Climate, National Academy Press, Washington D.C., 85 pp.
1672
- 1673 Nicholls, N., *et al.*, 1996: Observed climate variability and change. In: *Climate Change 1995: The*
1674 *Science of Climate Change. Contribution of Working Group I to the Second Assessment*
1675 *Report of the Intergovernmental Panel on Climate Change* [Houghton, J.T., *et al.*, (eds.)].
1676 Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 572
1677 pp.
1678
- 1679 North, G.R., and M.J. Stevens, 1998: Detecting climate signals in the surface temperature record.
1680 *Journal of Climate*, **11**, 563-577.
1681
- 1682 North, G.R., *et al.*, 1995: Detection of forced climate signals. Part I: Filter theory. *Journal of*
1683 *Climate*, **8**, 401-408.
1684
- 1685 Oleson, K.W., *et al.*, 2004: Effect of land use change on North American climate: impact of surface
1686 datasets and model biogeophysics. *Climate Dynamics*, **23**, 117-132.
1687
- 1688 Pan, Y.H., and A.H. Oort, 1983: Global climate variations connected with sea surface temperature
1689 anomalies in the eastern equatorial Pacific Ocean for the 1958-73 period. *Monthly Weather*
1690 *Review*, **111**, 1244-1258.
1691

- 1692 Paul, F., *et al.*, 2004: Rapid disintegration of Alpine glaciers observed with satellite data.
1693 *Geophysical Research Letters*, **31**, L21402, doi:10.1029/2004GL020816.
1694
- 1695 Pawson, S., K. Labitzke, and S. Leder, 1998: Stepwise changes in stratospheric temperature.
1696 *Geophysical Research Letters*, **25**, 2157-2160.
1697
- 1698 Penner, J.E., *et al.*, 2005: Effect of black carbon on mid-troposphere and surface temperature
1699 trends. In: *Integrated Assessment of Human Induced Climate Change*. [Schlesinger, M.E.
1700 (ed.)]. Cambridge University Press, Cambridge (in press).
1701
- 1702 Penner, J.E., S.Y. Zhang, and C.C. Chuang, 2003: Soot and smoke aerosol may not warm climate.
1703 *Journal of Geophysical Research (Atmospheres)*, **108**, 4657, doi:10.1029/2003JD003409.
1704
- 1705 Penner, J.E., *et al.*, 2001: Aerosols, their direct and indirect effects. In: *Climate Change 2001: The
1706 Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the
1707 Intergovernmental Panel on Climate Change* [Houghton, J.T., *et al.*, (eds.)]. Cambridge
1708 University Press, Cambridge, United Kingdom and New York, NY, USA, 881 pp.
1709
- 1710 Pielke Sr., R.A., 2004: Assessing “global warming” with surface heat content. *EOS*, **85**, 210-211.
1711
- 1712 Pielke Sr., R.A., and T.N. Chase, 2004: Comment on “Contributions of anthropogenic and natural
1713 forcing to recent tropopause height changes”. *Science*, **303**, 1771c.
1714
- 1715 Pielke Sr., R.A., 2003: Heat storage within the Earth system. *Bulletin of the American
1716 Meteorological Society*, **84**, 331-335.
1717
- 1718 Pierce, D.W., *et al.*, 2005: Anthropogenic warming of the oceans: Observations and model results.
1719 *Journal of Climate* (in press).
1720
- 1721 Pitman, A.J., *et al.*, 2004: Impact of land cover change on the climate of southwest Western
1722 Australia. *Journal of Geophysical Research (Atmospheres)*, **109**,
1723 doi:10.1029/2003JD004347.
1724
- 1725 Ramachandran, S., *et al.*, 2000: Radiative impact of the Mount Pinatubo volcanic eruption: Lower
1726 stratospheric response. *Journal of Geophysical Research (Atmospheres)*, **105**, 24409-24429.
1727
- 1728 Ramanathan, V., *et al.*, 2001: The Indian Ocean experiment: An integrated analysis of the climate
1729 forcing and effects of the great Indo-Asian haze. *Journal of Geophysical Research
1730 (Atmospheres)*, **106**, 28371-28398.
1731
- 1732 Ramankutty, N., and J.A. Foley, 1999: Estimating historical changes in global land cover:
1733 croplands from 1700 to 1992. *Global Biogeochemical Cycles*, **13**, 997-1027.
1734
- 1735 Ramaswamy, V., *et al.*, 2006: Anthropogenic and natural influences in the evolution of lower
1736 stratospheric cooling. *Science* (in preparation).
1737

- 1738 Ramaswamy, V., *et al.*, 2001a: Stratospheric temperature trends: observations and model
1739 simulations. *Reviews of Geophysics*, **39**, 71-122.
1740
- 1741 Ramaswamy, V., *et al.*, 2001b: Radiative forcing of climate change. In: *Climate Change 2001: The*
1742 *Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the*
1743 *Intergovernmental Panel on Climate Change* [Houghton, J.T., *et al.*, (eds.)]. Cambridge
1744 University Press, Cambridge, United Kingdom and New York, NY, USA, 881 pp.
1745
- 1746 Ramaswamy, V., M.D. Schwarzkopf and W.J. Randel, 1996: Fingerprint of ozone depletion in the
1747 spatial and temporal pattern of recent lower-stratospheric cooling. *Nature*, **382**, 616-618.
1748
- 1749 Randel, W.J., and F. Wu, 2005: Biases in stratospheric temperature trends derived from historical
1750 radiosonde data. *Journal of Climate* (in press).
1751
- 1752 Randel, W.J., F. Wu, and D.J. Gaffen, 2000: Interannual variability of the tropical tropopause
1753 derived from radiosonde data and NCEP reanalyses. *Journal of Geophysical Research*
1754 *(Atmospheres)*, **105**, 15509-15523.
1755
- 1756 Randel, W.J., and F. Wu, 1999: Cooling of the Arctic and Antarctic polar stratosphere due to ozone
1757 depletion. *Journal of Climate*, **12**, 1467-1479.
1758
- 1759 Reichert, B.K., R. Schnur, and L. Bengtsson, 2002: Global ocean warming tied to anthropogenic
1760 forcing. *Geophysical Research Letters*, **29**, doi:10.1029/2001GL013954.
1761
- 1762 Rignot, E., and R.H. Thomas, 2002: Mass balance of polar ice sheets. *Science*, **297**, 1502-1506.
1763
- 1764 Robock, A., 2005: Using the Mount Pinatubo volcanic eruption to determine climate sensitivity:
1765 Comments on “Climate forcing by the volcanic eruption of Mount Pinatubo” by David H.
1766 Douglass and Robert S. Knox. *Geophysical Research Letters* (in press).
1767
- 1768 Robock, A., and C. Oppenheimer, 2003: *Volcanism and the Earth’s Atmosphere*. AGU
1769 Geophysical Monograph Series, **139**, Washington D.C. American Geophysical Union, 360
1770 pp.
1771
- 1772 Robock, A., 2000: Volcanic eruptions and climate. *Reviews of Geophysics*, **38**, 191-219.
1773
- 1774 Rodwell, M.J., D.P. Rowell, and C.K. Folland, 1999: Oceanic forcing of the wintertime North
1775 Atlantic Oscillation and European climate. *Nature*, **398**, 320-323.
1776
- 1777 Santer, B.D., *et al.*, 2005: Amplification of surface temperature trends and variability in the tropical
1778 atmosphere. *Science*, **309**, 1551-1556.
1779
- 1780 Santer, B.D., *et al.*, 2004: Identification of anthropogenic climate change using a second-generation
1781 reanalysis. *Journal of Geophysical Research (Atmospheres)*, **109**,
1782 doi:10.1029/2004JD005075.
1783

- 1784 Santer, B.D., *et al.*, 2003a: Contributions of anthropogenic and natural forcing to recent tropopause
1785 height changes. *Science*, **301**, 479-483.
- 1786
- 1787 Santer, B.D., *et al.*, 2003b: Influence of satellite data uncertainties on the detection of externally-
1788 forced climate change. *Science*, **300**, 1280-1284.
- 1789
- 1790 Santer, B.D., *et al.*, 2001: Accounting for the effects of volcanoes and ENSO in comparisons of
1791 modeled and observed temperature trends. *Journal of Geophysical Research (Atmospheres)*,
1792 **106**, 28033-28059.
- 1793
- 1794 Santer, B.D., *et al.*, 2000: Interpreting differential temperature trends at the surface and in the lower
1795 troposphere. *Science*, **287**, 1227-1232.
- 1796
- 1797 Santer, B.D., *et al.*, 1999: Uncertainties in observationally-based estimates of temperature change
1798 in the free atmosphere. *Journal of Geophysical Research (Atmospheres)*, **104**, 6305-6333.
- 1799
- 1800 Santer, B.D., *et al.*, 1996a: A search for human influences on the thermal structure of the
1801 atmosphere. *Nature*, **382**, 39-46.
- 1802
- 1803 Santer, B.D. *et al.*, 1996b: Human effect on global climate? *Nature*, **384**, 522-524.
- 1804
- 1805 Satheesh, S.K., and V. Ramanathan, 2000: Large differences in tropical aerosol forcing at the top
1806 of the atmosphere and Earth's surface. *Nature*, **405**, 60-63.
- 1807
- 1808 Sato, M., *et al.*, 1993: Stratospheric aerosol optical depths, 1850-1990: *Journal of Geophysical*
1809 *Research (Atmospheres)*, **98**, 22987-22994.
- 1810
- 1811 Seidel, D.J., *et al.*, 2004: Uncertainty in signals of large-scale climate variations in radiosonde and
1812 satellite upper-air temperature datasets. *Journal of Climate*, **17**, 2225-2240.
- 1813
- 1814 Seidel, D.J., and J.R. Lanzante, 2004: An assessment of three alternatives to linear trends for
1815 characterizing global atmospheric temperature changes. *Journal of Geophysical Research*
1816 *(Atmospheres)*, **109**, doi: 10.1029/2003JD004414.
- 1817
- 1818 Seidel, D.J., R.J. Ross, J.K. Angell, and G.C. Reid, 2001: Climatological characteristics of the
1819 tropical tropopause as revealed by radiosondes. *Journal of Geophysical Research*
1820 *(Atmospheres)*, **106**, 7857-7878.
- 1821
- 1822 Sexton, D.M.H., *et al.*, 2001: Detection of anthropogenic climate change using an atmospheric
1823 GCM. *Climate Dynamics*, **17**, 669-685.
- 1824
- 1825 Sherwood, S.C., J. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late 20th century
1826 warming. *Science*, **309**, 1556-1559.
- 1827

- 1828 Shindell, D.T., G. Fulavegi, and N. Bell, 2003: Preindustrial-to-present-day radiative forcing by
1829 tropospheric ozone from improved simulations with the GISS chemistry-climate GCM.
1830 *Atm. Chem. Phys.*, **3**, 1675-1702.
1831
- 1832 Shine, K.P., *et al.*, 2003: A comparison of model-simulated trends in stratospheric temperatures.
1833 *Quarterly Journal of the Royal Meteorological Society*, **129**, 1565-1588.
1834
- 1835 Smith, S.J., H. Pitcher, and T.M.L. Wigley, 2005: Future sulfur dioxide emissions. *Climatic*
1836 *Change*, **73**, 267-318.
1837
- 1838 Smith, S.J., H. Pitcher, and T.M.L. Wigley, 2001: Global and regional anthropogenic sulfur dioxide
1839 emissions. *Global and Planetary Change*, **29**, 99-119.
1840
- 1841 Soden, B.J., *et al.*, 2002: Global cooling after the eruption of Mt. Pinatubo: A test of climate
1842 feedback by water vapor. *Science*, **296**, 727-730.
1843
- 1844 Soden, B.J., 2000: the sensitivity of the tropical hydrological cycle to ENSO. *Journal of Climate*,
1845 **13**, 538-549.
1846
- 1847 Stainforth, D.A., *et al.*, 2005: Uncertainties in predictions of the climate response to rising levels of
1848 greenhouse gases. *Nature*, **433**, 403-406.
1849
- 1850 Stott, P.A., *et al.*, 2005: Robustness of estimates of greenhouse attribution and observationally
1851 constrained predictions of global warming. *Journal of Climate* (accepted).
1852
- 1853 Stott, P.A., D.A. Stone, and M.R. Allen, 2004: Human contribution to the European heatwave of
1854 2003. *Nature*, **423**, 61-614.
1855
- 1856 Stott, P.A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural
1857 causes. *Geophysical Research Letters*, **30**, doi: 10.1029/2003GL017324.
1858
- 1859 Stott, P.A., *et al.*, 2000: External control of 20th century temperature by natural and anthropogenic
1860 forcings. *Science*, **290**, 2133-2137.
1861
- 1862 Stott, P.A., and S.F.B. Tett, 1998: Scale-dependent detection of climate change. *Journal of*
1863 *Climate*, **11**, 3282-3294.
1864
- 1865 Sun, S., and J.E. Hansen, 2003: Climate simulations for 1951-2050 with a coupled atmosphere-
1866 ocean model. *Journal of Geophysical Research (Atmospheres)*, **16**, 2807-2826.
1867
- 1868 Tett, S.F.B., and P.W. Thorne, 2004: Comment on tropospheric temperature series from satellites.
1869 *Nature*, doi: 10.1038/nature03208.
1870
- 1871 Tett, S.F.B., *et al.*, 2002: Estimation of natural and anthropogenic contributions to twentieth
1872 century temperature change. *Journal of Geophysical Research (Atmospheres)*, **107**,
1873 doi:10.1029/2000JD000028.

- 1874
1875 Tett, S.F.B., *et al.*, 1999: Causes of twentieth-century temperature change near the Earth's surface.
1876 *Nature*, **399**, 569-572.
1877
- 1878 Tett, S.F.B., *et al.*, 1996: Human influence on the atmospheric vertical temperature structure:
1879 Detection and observations. *Science*, **274**, 1170-1173.
1880
- 1881 Thorne, P.W., *et al.*, 2005: Revisiting radiosonde upper-air temperatures from 1958 to 2002.
1882 *Journal of Geophysical Research (Atmospheres)* (in press).
1883
- 1884 Thorne, P.W., *et al.*, 2003: Probable causes of late twentieth century tropospheric temperature
1885 trends. *Climate Dynamics*, **21**, 573-591.
1886
- 1887 Thorne, P.W., *et al.*, 2002: Assessing the robustness of zonal mean climate change detection.
1888 *Geophysical Research Letters*, **29**, doi: 10.1029/2002GL015717.
1889
- 1890 Tie, X., *et al.*, 2005: Assessment of the global impact of aerosols on tropospheric oxidants. *Journal*
1891 *of Geophysical Research (Atmospheres)* (in press).
1892
- 1893 Trenberth, K.E., J. Fasullo, and L. Smith, 2005: Trends and variability in column-integrated
1894 atmospheric water vapor. *Climate Dynamics*, doi:10.1007/s00382-005-0017-4.
1895
- 1896 Trenberth, K.E., and L. Smith, 2005: The mass of the atmosphere: A constraint on global analyses.
1897 *Journal of Climate*, **18**, 860-875.
1898
- 1899 Trenberth, K.E., and T.J. Hoar, 1996: The 1990-1995 El Niño-Southern Oscillation event: longest
1900 on record. *Geophysical Research Letters*, **23**, 57-60.
1901
- 1902 Trenberth, K.E., 1992: *Climate System Modeling*. Cambridge University Press, Cambridge, 788 pp.
1903
- 1904 Wallace, J.M., Y. Zhang, and J.A. Renwick, 1995: Dynamic contribution to hemispheric mean
1905 temperature trends. *Science*, **270**, 780-783.
1906
- 1907 Washington, W.M., *et al.*, 2000: Parallel Climate Model (PCM) control and transient simulations.
1908 *Climate Dynamics*, **16**, 755-774.
1909
- 1910 Weber, G.R., 1996: Human effect on global climate? *Nature*, **384**, 524-525.
1911
- 1912 Wehner, M.F., 2000: A method to aid in the determination of the sampling size of AGCM
1913 ensemble simulations. *Climate Dynamics*, **16**, 321-331.
1914
- 1915 Wentz, F.J., and M. Schabel, 2000: Precise climate monitoring using complementary satellite data
1916 sets. *Nature*, **403**, 414-416.
1917
- 1918 Wielicki, B.A., *et al.*, 2002: Evidence for large decadal variability in the tropical mean radiative
1919 energy budget. *Science*, **295**, 841-844.

- 1920
1921 Wigley, T.M.L., *et al.*, 2005a: The effect of climate sensitivity on the response to volcanic forcing.
1922 *Journal of Geophysical Research (Atmospheres)*, **110**, D09107,
1923 doi:10.1029/2004JD005557.
1924
1925 Wigley, T.M.L., *et al.*, 2005b: Using the Mount Pinatubo volcanic eruption to determine climate
1926 sensitivity: Comments on “Climate forcing by the volcanic eruption of Mount Pinatubo”, by
1927 David H. Douglass and Robert S. Knox. *Geophysical Research Letters* (in press).
1928
1929 Wigley, T.M.L., 2000: ENSO, volcanoes, and record-breaking temperatures. *Geophysical Research*
1930 *Letters*, **27**, 4101-4104.
1931
1932 Willis, J.K., D. Roemmich, and B. Cornuelle, 2004: Interannual variability in upper-ocean heat
1933 content, temperature, and thermosteric expansion on global scales. *Journal of Geophysical*
1934 *Research*, **109**, C12036, doi:10.1029/2003JC002260..
1935
1936 Yulaeva, E., and J.M. Wallace, 1994: The signature of ENSO in global temperature and
1937 precipitation fields derived from the microwave sounding unit. *Journal of Climate*, **7**, 1719-
1938 1736.
1939
1940 Zwiers, F.W., and X. Zhang, 2003: Towards regional-scale climate change detection. *Journal of*
1941 *Climate*, **16**, 793-797.
1942

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

Chapter 6

What measures can be taken to improve the understanding of observed changes?

Convening Lead Author:
Chris Folland

Lead Authors:
David Parker, Richard Reynolds, Steve Sherwood, and Peter Thorne

44

45

46

47 **Background**

48

49 There remain differences between independently estimated temperature trends for the surface,
50 troposphere and lower stratosphere, and differences between the observed changes and model
51 simulations, that are, as yet, not fully understood, although recent progress is reported in
52 previous chapters. This Chapter makes recommendations that address these specific problems
53 rather than more general climate research aims, building on the discussions, key findings, and
54 recommendations of the previous chapters. Because the previous chapters fully discuss the
55 many issues, we only provide a summary here. Furthermore, we only list key references to
56 the peer reviewed literature. To ensure traceability and to enable easy cross-referencing we
57 refer to the chapters by e.g., (C5) for Chapter 5. We do not specifically refer to sub-sections
58 of chapters.

59

60 Much previous work has been done to address, or plan to address, most of the problems
61 discussed in this Report. Rather than invent brand new proposals and recommendations, we
62 have tried to expand and build upon existing ideas emphasizing those we believe to be of
63 highest utility. Key documents in this regard are: the Global Climate Observing System
64 (GCOS) Implementation Plan for the Global Observing System (GCOS, 2004), the wider
65 Global Earth System of Systems (GEOSS) 10 year Implementation Plan Reference
66 Document (GEO, 2005) which explicitly includes the GCOS Implementation Plan as its
67 climate component; and the over-arching Climate Change Science Program plan (CCSP,
68 2004).

69

70 The remainder of this Chapter is split into six sections. Each section discusses requirements
71 under a particular theme, aiming to encapsulate the key findings and recommendations of the
72 earlier chapters and culminating in one main recommendation in each of Sections 1 to 5 and
73 two recommendations in Section 6. Sections 1 to 5 focus on key actions that should be
74 carried out in the near future, making use of existing historical data and current climate
75 models. Section 6 discusses future climate monitoring in relation to the vertical profile of
76 temperature trends in the atmosphere.

77

78 **1. Constraining observational uncertainty**

79

80 An important advance since recent in-depth reviews of the subject of this Report (NRC,
81 2000a, IPCC, 2001) has been a better appreciation of the uncertainties in our estimates of
82 recent temperature changes, particularly above the surface (**C2, C3, C4**). Many observations
83 that are used in climate studies are taken primarily for the purposes of operational weather
84 forecasting (**C2**). Not surprisingly, there have been numerous changes in instrumentation,
85 observing practices, and the processing of data over time. While these changes have
86 undoubtedly led to improved forecasts of weather, they add significant complexity to
87 attempts to reconstruct past climate trends, (**C2, C4**). The main problem is that such an
88 evolution tends to introduce artificial (non-climatic) changes into the data (**C2**).

89

90 Above the Earth's surface, the spread in independently-derived estimates of climate change,
91 representing what is referred to in this report as "construction" uncertainty (**C2, C4,**
92 **Appendix**) (Thorne et al., 2005), is of similar magnitude to the expected climate signal itself

93 **(C3, C4, C5)**. Changes in observing practices have been particularly pervasive aloft, where
94 the technical challenges in maintaining robust, consistent measurements of climate variables
95 are considerably greater than at the surface **(C2, C4, C5)**. This does not imply that there are
96 no problems in estimating temperature trends at the surface. Such problems include
97 remaining uncertainties in corrections that must be made to sea surface temperatures (SSTs)
98 in recent decades **(C2, C4)**, and uncertainties in accounting for changes in micro-climate
99 exposure for some individual land stations **(C2, C4)** or simply allowing for genuinely bad
100 stations (Davey and Pielke, 2005). Differences between surface data sets purporting to
101 measure the same variable become larger as the spatial resolution being considered decreases.
102 This implies that many problems tend to have random effects on climate analyses at the large
103 spatial scales, that are the focus of this Report, but can be systematic at much smaller scales
104 **(C2, C3, C4)**.

105

106 The climate system has evolved in a unique way, and, by definition the best analysis is that
107 which most closely approaches this actual evolution. However, because we do not know the
108 evolution of the climate system exactly, we have generally had to treat apparently well
109 constructed but divergent data sets, of atmospheric temperature changes in particular, as
110 equally valid **(C3, C4, C5)**. Clearly, this approach is untenable in the longer-term. Thus, it is
111 imperative that we reduce the uncertainty in our knowledge of how the three-dimensional
112 structure of atmospheric temperature has evolved **(C4)**.

113

114 To ascertain unambiguously the causes of differences in datasets generally requires extensive
115 metadata¹ for each data set (**C4**; NRC, 2000b). Appropriate metadata, whether obtained from
116 the peer-reviewed literature or from data made available on-line, should include, for data on
117 all relevant spatial and temporal scales:

- 118 • Documentation of the raw data and the data sources used in the data set construction
119 to enable quantification of the extent to which the raw data overlap with other similar
120 datasets;
- 121 • Details of instrumentation used, the observing practices and environments and their
122 changes over time to help assessments of, or corrections for, the changing accuracy of
123 the data;
- 124 • Supporting information such as any adjustments made to the data and the numbers
125 and locations of the data through time;
- 126 • An audit trail of decisions about the adjustments made, including supporting evidence
127 that identifies non-climatic influences on the data and justifies any consequent
128 adjustments to the data that have been made; and
- 129 • Uncertainty estimates and their derivation.

130 This information should be made openly available to the research community.

131

132 There is evidence, discussed in earlier chapters, for a number of unresolved issues in existing
133 data sets that should be addressed:

- 134 • Systematic, historically-varying biases in day-time relative to night-time radiosonde
135 temperature data are important, particularly in the tropics (**C4**). These are likely to

¹ Metadata are literally “data about data” and are typically records of instrumentation used, observing practices, the environmental context of observations, and data-processing procedures.

136 have been poorly accounted for by present approaches to quality controlling such data
137 (Sherwood et al., 2005) and may seriously affect trends.

138 • Radiosonde stratospheric records are strongly suspected of retaining a spurious long-
139 term cooling bias, especially in the tropics (C4).

140 • Diurnal adjustment techniques for satellite temperature data are uncertain (C2, C4).
141 This effect is particularly important for the 2LT retrieval (C4). Further efforts are
142 required to refine our quantification of the diurnal cycle, perhaps through use of
143 reanalyses, in-situ observations, or measurements from non-sun-synchronous orbiters
144 (C4).

145 • Different methods of making inter-satellite bias adjustments, particularly for satellites
146 with short periods of overlap, can lead to large discrepancies in trends (C4) (see also
147 Section 6).

148 • Variable biases in modern SST data remain that have not been adequately addressed
149 (C4). Some historical metadata are now available for the first time, but are yet to be
150 fully exploited (Rayner et al., 2005). Better metadata, better use of existing metadata,
151 and use of recently bias-adjusted day-time marine air temperature data are needed to
152 assess remaining artifacts (C4).

153 • Land stations may have had undocumented changes in the local environment that
154 could lead to their records being unrepresentative of regional- or larger-scale changes
155 (C2, C4).

156

157 In addition to making data sets and associated metadata openly available and addressing the
158 issues discussed above, it would be useful to develop a set of guidelines that can be used to

159 help assess the quality of data sets (C4). It is important that numerous tests be applied to
160 reduce ambiguity. There are three types of check that may be used:

161

162 1. *Internal consistency checks*

163 For example, we expect only relatively small real changes in the diurnal cycle of
164 temperature above the atmospheric boundary layer (C1) (Sherwood et al., 2005), so
165 an apparently homogenized data set that shows large changes in the diurnal cycle in
166 these regions should be closely scrutinized.

167 2. *Inter-dataset comparisons*

168 For example, comparisons are needed between radiosonde and MSU temperature
169 measures representing the same regions (Christy and Norris, 2004).

170 3. *Consistency with changes in other climate variables and parameters*

171 This is a potentially powerful but much under-utilized approach and is discussed
172 further in Section 3.

173

174

175 **RECOMMENDATION** *In order to encourage further independent scrutiny, data sets and*
176 *their full metadata (i.e., information about instrumentation used, observing practices, the*
177 *environmental context of observations, and data-processing procedures) should be made*
178 *openly available. Comprehensive analyses should be carried out to ascertain the causes of*
179 *remaining differences between data sets and to refine uncertainty estimates.*

180

181 **2. Making better use of existing observational data**

182

183 There is a considerable body of observational data that have either been under-utilized or not
184 used at all when constructing the data sets of historical temperature changes discussed in this
185 Report (**C2, Table 2.1**). Estimates of temperature changes can potentially be made from
186 several satellite instruments beside the (Advanced) Microwave Sounding Unit data
187 considered here (**C2, C3**). In particular, largely overlooked satellite datasets should be re-
188 examined to try to extend, fortify or corroborate existing microwave-based temperature
189 records for climate research, e.g., microwave data from other instruments such as the Nimbus
190 5 (Nimbus E) Microwave Spectrometer (NEMS) (1972) and the Nimbus 6 Scanning
191 Microwave Spectrometer (SCAMS) (1975), infra-red data from the High Resolution Infrared
192 Radiation Sounder (HIRS) suite, and radio occultation data from Global Positioning System
193 (GPS) satellites (**C2**). Some of these instruments may allow us to extend the records back to
194 the early 1970s. Many unused radiosonde measurements of a relatively short length exist in
195 regions of relatively sparse coverage and, with some effort, could be advantageously used to
196 fill gaps. Many additional surface temperature data exist, mainly over land over the period
197 considered in this Report, but are either not digitized or not openly available. This latter
198 problem is particularly common in many tropical regions where much of the interest in this
199 Report resides. Given the needed level of international cooperation, we could significantly
200 improve our current estimates of tropical temperature changes over land and derive better
201 estimates of the changing temperature structure of the lower atmosphere (**C2**).

202

203 In addition to the recovery and use of such existing data, we need to improve the access to
204 metadata for existing raw observations (**C2**). Additional information on when and how
205 changes occurred in observing practices, the local environment, etc., is potentially available
206 in national meteorological and hydrometeorological services. Such metadata would help

207 reduce current uncertainties in estimates of observed climate change. In the absence of
208 comprehensive metadata, investigators have to make decisions regarding the presence of
209 heterogeneities (non-climatic jumps or trends) using statistical methods alone. Statistical
210 methods of adjusting data for inhomogeneities have a very useful role, but are much more
211 valuable in the presence of good and frequent metadata that can be used to confirm the
212 presence, type, and timing of non-climatic influences. Metadata requirements will vary
213 according to observing system, but, if in doubt, all potentially important information should
214 be included. For example, surface temperature metadata may include:

- 215 • Current and historical photographs and site sketches to ascertain changes in micro-
216 climate exposure and their timing, collected during the routine site inspections made
217 by most meteorological services;
- 218 • The history of instrumentation changes;
- 219 • Changes in the way stations are maintained and in their immediate environment;
- 220 • Changes in observers; and
- 221 • Changes in observing and reporting practices.

222 For other instrument types, e.g., for humidity measurements, the detailed metadata
223 requirements will vary. A further discussion on the challenges of collecting climate data can
224 be found in Folland et al. (2000).

225

226 **RECOMMENDATION** *Efforts should be made to archive and make openly available*
227 *surface, balloon-based, and satellite data and metadata that have not previously been*
228 *exploited. Emphasis should be placed on the tropics and on recovery and inclusion of*
229 *satellite data before 1979 which may allow better characterization of the climate regime shift*
230 *in the mid-1970's*

231

232

233 **3. Multivariate analyses**

234

235 Temperature changes alone are a necessary, but insufficient, constraint on understanding the
236 evolution of the climate system. Even with a perfect knowledge of temperature changes,
237 knowledge about changes in the climate system would be incomplete. Consequently,
238 understanding temperature trends also requires knowledge about changes in other measures
239 of the climate system. For example, changes in atmospheric circulation and accompanying
240 dynamical effects, and also in latent heat transport, have significant implications for vertical
241 profiles of temperature trends **(C1)**.

242

243 Changes in variables other than temperature may be used to confirm the attribution of climate
244 change to given causes **(C5)** and to test the physical plausibility of reported temperature
245 changes **(C3, C4)**. It is likely that to fully understand changes in atmospheric temperature, it
246 will be necessary to consider changes in at least some of the following physical parameters
247 and properties of the climate system beside its temperature:

248

- Water vapor content **(C1, C5)**

249

- Ocean heat content **(C5)**

250

- The height of the tropopause **(C5)**

251

- Wind fields

252

- Cloud cover and the characteristics of clouds

253

- Radiative fluxes

254

- Aerosols and trace gases

- 255 • Changes in glacial mass, sea ice volume, permafrost and snow cover (**C5**)

256

257 Our current ability to undertake such multivariate analyses of climate changes is constrained
258 by the relative paucity of accurate climate datasets for variables other than temperature.
259 Furthermore, since our analysis of temperature datasets has highlighted the importance of
260 construction uncertainty in determining trends (**C2, C4, Appendix**), it is very likely that
261 similar considerations will pertain to these other data types. It is therefore necessary to
262 construct further independent estimates of the changes in these variables even where datasets
263 already exist. Similar considerations to those discussed in Section 1 are also important for
264 these additional data.

265

266 **RECOMMENDATION** *Efforts should be made to create climate quality² data sets for a*
267 *range of variables other than temperature. These should subsequently be compared with each*
268 *other and with temperature data to determine whether they are consistent with our physical*
269 *understanding. It is important to create several independent estimates for each parameter in*
270 *order to assess the magnitude of construction uncertainties.*

271

272 4. Climate quality reanalyses

273

274 Reanalyses are derived from Numerical Weather Prediction (NWP) (forecast) models run
275 retrospectively with historical observations to produce physically consistent, fully global
276 fields with high temporal and spatial resolution. As in NWP, reanalyses employ all available
277 observations to produce their analysis and minimize the instantaneous differences between

² “Climate quality” refers to a record for which the best possible efforts have been made to identify and remove non-climatic effects that produce spurious changes over time.

278 the available observations and a background forecast field initiated a number of hours earlier.
279 Reanalyses also use the same NWP model throughout the reanalysis period. However, as for
280 observed climate datasets, pervasive changes in the raw observations lead to discontinuities
281 and spurious drifts (C2). Because such discontinuities and drifts have been identified in the
282 temperature fields of the current generation of reanalyses, these have been deemed
283 inappropriate for the purpose of long-term temperature trend characterisation by this Report's
284 authors (C2, C3). **However**, it is recognised that some progress has been made (e.g.,
285 Simmons et al., 2004, C2). This does not preclude the usefulness of reanalyses for
286 characterizing seasonal to interannual timescale variability and processes, or trends in other,
287 related, variables such as tropopause height (C5). Indeed, they have proven to be a very
288 important tool for the climate research community.

289

290 A more homogeneous reanalysis that minimized time-dependent biases arising from changes
291 in the observational network would be of enormous benefit for multivariate analyses of
292 climate change (C2, C3). Advances in NWP systems, which will continue to happen
293 regardless of climate requirements, will inevitably lead to better future reanalyses of
294 interannual climate variability. Some advances, such as so-called 'feedback files'³ from the
295 data assimilation of reanalyses, could be uniquely helpful for climate reanalysis and should
296 be encouraged for this reason if no other. However, to determine trends accurately from
297 reanalyses will also require intensive efforts by the reanalysis community to understand
298 which observations are critical for trend characterization and to homogenize these data
299 insofar as possible to eliminate non-climatic changes before input to the reanalysis system.
300 This in turn requires observing system experiments (OSEs) where the impact on trends of

³ "Feedback files" are diagnostic summaries of adjustments applied to data during assimilation.

301 different observation types from land, radiosonde, and space-based observations are assessed.

302 A few examples (far from an exhaustive list) are:

- 303 • Run a short period (e.g., a year) of reanalysis with and without radiosonde
304 temperatures.
- 305 • Carry out experiments incorporating long radiosonde data records only, then repeat
306 with additional less certain but spatially more complete short radiosonde records.
- 307 • Successively include or remove specific satellite retrievals (e.g., MSU Channel 2).
- 308 • Carry out test reanalyses for one or more decades with different corrections to the
309 observed data for inhomogeneities within their construction uncertainty estimates.

310

311 Progress would depend on reanalyses and data construction experts from all the key groups
312 working closely together.

313

314

315 **RECOMMENDATION** *Consistent with Key Action 24 of GCOS (2004)⁴ and a 10 Year*
316 *Climate Target of GEOSS (2005), efforts should be made to create several homogeneous*
317 *atmospheric reanalyses. Particular care needs to be taken to identify and homogenize critical*
318 *input climate data, and to more effectively manage large-scale changes in the global*
319 *observing system to avoid non-climatic influences.⁵*

320

321 5. Better understanding of uncertainties in model estimates

⁴ Parties are urged to give high priority to establishing a sustained capacity for global climate reanalysis, to develop improved methods for such reanalysis, and to ensure coordination and collaboration among Centers conducting reanalyses.

⁵ A focal point for planning of future U.S. reanalysis efforts is the CCSP Synthesis and Assessment Product 1.3: "Re-analyses of historical climate data for key atmospheric features. Implications for attribution of causes of observed change". Ongoing progress in the planning of future U. S. reanalysis efforts can be found at: http://www.joss.ucar.edu/joss_psg/meetings/climatesystem/

322

323 New state-of-the-art global climate models have simulated the influences of natural and
324 anthropogenic climate forcings on tropospheric and surface temperature. The simulations
325 generally cover the period since the late nineteenth century, but results are only reported over
326 the period of primary interest to this Report, 1979-1999 (the satellite era), in Chapter 5.
327 Taken together, these models, for the first time, consider most of the recognized first-order
328 climate forcings and feedbacks as identified in IPCC (2001), NRC (2003), and NRC (2005).
329 This is an important step forward (C5).

330

331 However, most individual models considered in this Report still do not make use of all likely
332 important climatic forcings (C5, Table 5.2). In addition, many of the forcings are not yet well
333 quantified. Models that appear to include the same forcings often differ in both the way the
334 forcings are quantified and how these forcings are applied to the model. This makes it
335 difficult to separate intrinsic differences between models from the effects of different forcings
336 on predicted temperature trends. Thus, within the “ensemble of opportunity” considered in
337 this Report (C5), it is difficult to separate differences in:

338

- Model physics and resolution;

339

- The details of the way the forcings are applied in the experiments;

340

- The chosen history of the changes in the forcing.

341

342 To better quantify the impacts of the various forcings on vertical temperature trends, a further
343 suite of experiments is needed along the following lines:

344

- Runs with one forcing applied in a single experiment with a given model; these are
345 already required in some detection and attribution studies (C5). They have been

346 performed for a small number of models already. This approach is particularly
347 important for the recently developed and spatially heterogeneous land use / land cover
348 change and black carbon aerosol forcings (C5).

349 • Apply the same forcing in exactly the same manner to a suite of models so that the
350 differences that result are due unambiguously to model differences (C5).

351 • Apply the full range of important forcings, with their uncertainties explicitly sampled
352 to a small subset of the most advanced models to gain an overall estimate of the
353 effects on temperature trends of the uncertainties in these forcings.

354

355 It is recognized that there are many problems in achieving this, so a considerable effort will
356 be needed over a number of years. In addition, these model runs should be compared to the
357 full range of observational estimates to avoid ambiguity (C5). Finally, detection and
358 attribution studies should be undertaken using this new range of observations and model-
359 based estimates to reassess previous results; which have consistently identified a human-
360 induced influence (C5).

361

362

363 **RECOMMENDATION** *Models that appear to include the same forcings often differ in both*
364 *the way the forcings are quantified and how these forcings are applied to the model. Hence,*
365 *efforts are required to more formally separate uncertainties arising from model structure*
366 *from the effects of forcing uncertainties. This requires running multiple models with*
367 *standardized forcings, and running the same models individually under a range of plausible*
368 *scenarios for each forcing."*

369

370 **6. Future monitoring of climate**

371

372 Much of this Report hitherto has concerned historical climate measurements. However, over
373 the coming decades new, mainly space-based, observations will yield very large increases in
374 the volume and types of data available. These will come from many different instruments
375 making measurements with greater accuracy and detail, especially in the vertical direction,
376 and with greater precision (**C2, C3**). In fact, new types of more accurate data such as
377 temperature and moisture profiles from GPS radio-occultation measurements are already
378 available, although, as yet, few efforts have been made to analyse them (**C2, C3**). Current and
379 planned multi-spectral infra-red satellite sounders such as the Atmospheric InfraRed Sounder
380 (AIRS) and the Infrared Atmospheric Sounding Interferometer (IASI) have much finer
381 vertical resolution than earlier satellite sounders used in the Report. They have the potential
382 to resolve quite fine vertical and horizontal details of temperature and humidity through the
383 depth of much of the atmosphere. These higher spectral resolution data should also permit a
384 continuation of records equivalent to earlier coarser infrared satellite data (e.g., from the
385 HIRS satellite instruments). The new suite of satellite data will not only prove useful for
386 sensing changes aloft. For example, satellite data to remotely sense sea-surface temperatures
387 now include microwave products that can sense surface temperatures even in cloudy
388 conditions (C4). The Global Ocean Data Assimilation Experiment (GODAE) High-
389 resolution Sea Surface Temperature (SST) Pilot Project (GHRSSST- PP) has been established
390 to give international focus and coordination to the development of a new generation of global,
391 multi-sensor, high-resolution SST products (Donlon et al., 2005).

392

393 Many other agencies and bodies (e.g., NRC, 2000b; GCOS, 2004; GEOSS, 2005; CCSP,
394 2004) have already made recommendations for managing such new data developments.

395 These include such subjects as:

- 396 • Adherence to the GCOS Climate Monitoring Principles, needed to create and
397 maintain homogenous data sets of climate quality and for which there is a special set
398 for satellites (GCOS, 2004, Appendix 3)
- 399 • Continuation of records equivalent to current monitoring abilities: e.g., use new and
400 more detailed satellite data to create equivalent MSU measures of temperature to
401 allow the indefinite extension of the historical records used in this Report.
- 402 • Full implementation of national and international climate monitoring networks such
403 as the GCOS Upper-Air Network and the GCOS Surface Network.
- 404 • Overlap of measurement systems as they evolve in time.

405

406 This last point is of primary importance. It was given prominence by NRC (2000b) and is
407 emphasized in the GCOS Climate Monitoring Principles and leads to the following
408 recommendation. If this recommendation had been followed in the past, one of the major
409 problems in producing a homogeneous record of MSU temperatures would have been largely
410 removed (C4):

411

412 **RECOMMENDATION:** *The GCOS Climate Monitoring principles should be fully adopted.*
413 *In particular when any type of instrument for measuring climate is changed or re-sited, the*
414 *period of overlap between the old and new instruments or configurations should be sufficient*
415 *to allow analysts to adjust for the change with small uncertainties that do not prejudice the*
416 *analysis of climate trends. The minimum period is a full annual cycle of the climate. Thus,*

417 *replacement satellite launches should be planned to take place at least a year prior to the*
418 *expected time of failure of a satellite.*

419

420

421 Finally, we expand on a recommendation made in GCOS (2004) that is imperative for
422 successful future monitoring of temperatures at and above the Earth's surface. The main
423 lesson learned from this Report is that great difficulties in identifying and removing non-
424 climatic influences from upper-air observations have led to a very large spread in trend
425 estimates (**C2, C3, C4**). These differences can lead to fundamentally different interpretations
426 both of the extent of any discrepancies in trends between the surface and the troposphere
427 (**C3,C4**); and of the skill of climate models (**C5**). The problem has arisen because there has
428 been no high quality reference or "ground truth" data, however restricted in scope, against
429 which routine observations can be compared to facilitate rigorous removal of non-climatic
430 influences.

431

432 Our key recommendation in this regard is a set of widely distributed (perhaps about 5% of the
433 operational radiosonde network) reference sites that will provide high quality data for
434 anchoring more globally-extensive monitoring efforts (satellites, reanalyses, etc.). At such
435 reference sites (which could coincide with selected GCOS Upper Air Network (GUAN),
436 GCOS Surface Network (GSN) or Global Atmospheric Watch (GAW) sites) there would be
437 full, high-quality measurements of atmospheric column properties, both physical and
438 chemical. This requires a large suite of instrumentation and redundancy in measurements⁶.
439 These globally distributed reference sites should incorporate upward looking instruments

⁶ Measurement of the same parameter by two or more independent instruments

440 (radar, lidar, GPS-related data, microwave sensors, wind profilers, etc.) along with high-
441 quality temperature, relative humidity and wind measurements on balloons regularly
442 penetrating well into the stratosphere⁷ A key requirement is an end-to-end management
443 system including archiving of coincident observations made from over-flying satellites. The
444 data would be made openly available. The development of such a reference network is
445 recommended in outline by GCOS (2004). The ideas are currently being discussed in more
446 detail as part of an on-going process led by NOAA and WMO. Further details can be found at
447 <http://www.oco.noaa.gov/workshop/>.

448

449 **RECOMMENDATION** *Following Key Action 12⁸ of the GCOS Implementation Plan*
450 *(GCOS, 2004), develop and implement a subset of about 5% of the operational radiosonde*
451 *network as reference network sites for all kinds of climate data from the surface to the*
452 *stratosphere.*

453

454 **References**

455

456 Christy, J.R., and W.B. Norris, 2004: What may we conclude about tropospheric temperature
457 trends? *Geophysical Research Letters*, **31**, L06211

458

459 CCSP, 2004, Strategic Plan for the Climate Change Science Program Final Report, July 2003.

460 Available from

461 <http://www.climate-science.gov/Library/stratplan2003/final/default.htm>

462

463 Davey, C.A. and R.A. Pielke, 2005: Microclimatic exposures of surface-based weather
464 stations. *Bulletin of the American Met. Society*, **86**, 497-504.

465

466 Donlon, C J., 2005: Proceedings of the sixth Global Ocean Data Assimilation Experiment
467 (GODAE) High Resolution Sea Surface Temperature Pilot Project (GHRSSST-PP)
468 Science Team Meeting, Met Office, Exeter, May 16-20th 2005, Available from the
469 International GHRSSST-PP Project Office, Hadley Centre, Exeter, UK, pp 212.

470

⁷ Recent inter-comparisons under the auspices of WMO suggest that new operational sondes are as accurate as proposed reference sondes (C4; Pathack et al., 2005), which may reduce costs.

⁸ Parties need to: ... establish a high-quality reference network of about 30 precision radiosonde stations and other collocated observations. ...

- 471 Folland, C.K., N. Rayner, P. Frich, T. Basnett, D. Parker & E.B. Horton, 2000: Uncertainties
472 in climate data sets - a challenge for WMO. *WMO Bull.*, **49**, 59-68.
473
- 474 GCOS, 2004: Global Climate Observing System Implementation Plan for the Global
475 Observing System for Climate in support of the UNFCCC, GCOS-92/WMO TD
476 1219, WMO, Geneva, pp136, Available at
477 <http://www.wmo.int/web/gcos/gcoshome.html>. CD also available.
478
- 479 GEO, 2005: Global Earth Observation System of Systems, GEOSS. 10-Year Implementation
480 Plan Reference Document. GEO1000R/ESA SP-1284. ESA Publications, Noordwijk,
481 Netherlands.209pp. CD available.
482
- 483 IPCC (2001) *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to*
484 *the Third Assessment Report of the Intergovernmental Panel on Climate Change*
485 [Houghton, J. T., Y. Ding, D.J. Griggs, M. Noguer, P. van der Linden, X. Dai, K.
486 Maskell, and C.I. Johnson (eds.)]. Cambridge University Press, 881pp.
487
- 488 NRC, 2000a: Reconciling observations of Global Temperature Change. National Academy
489 Press, 85 pp.
490
- 491 NRC, 2000b: Improving atmospheric temperature monitoring capabilities. A letter report to
492 NOAA from the Panel on Reconciling Temperature Observations.
493
- 494 NRC, 2003: Understanding Climate Change Feedbacks. National Academy Press.
495
- 496 NRC, 2005: National Research Council, 2005: Radiative forcing of climate change:
497 Expanding the concept and addressing uncertainties. National Academy Press.
498
- 499 Pathack, B., J. Nash, R. Smout, and S. Kurnosenko, 2005: Preliminary Results of WMO
500 Intercomparison of high quality radiosonde systems, Mauritius, February 2005.
501 [www.wmo.ch/web/www/IMOP/meetings/Upper-Air/RSO-](http://www.wmo.ch/web/www/IMOP/meetings/Upper-Air/RSO-ComparisonMauritius/3(16)_UK_Nash_Vacoas.pdf)
502 [ComparisonMauritius/3\(16\)_UK_Nash_Vacoas.pdf](http://www.wmo.ch/web/www/IMOP/meetings/Upper-Air/RSO-ComparisonMauritius/3(16)_UK_Nash_Vacoas.pdf)
503
- 504 Rayner, N.A., P. Brohan, D.E. Parker, C K. Folland, J.J. Kennedy, M Vanicek, T Ansell, and
505 S.F.B. Tett, 2005: Improved analyses of changes and uncertainties in sea surface
506 temperature measured *in situ* since the mid-nineteenth century. Accepted by *J.*
507 *Climate*.
508
- 509 Seidel, D.J, J.K. Angell, J. Christy, M. Free, S.A. Klein, J.R. Lanzante, C. Mears, D. Parker,
510 M. Schabel, R. Spencer, A. Sterin, P. Thorne, and F. Wentz, 2004: Uncertainty in
511 signals of large-scale climate variations in radiosonde and satellite upper-air
512 temperature datasets. *J. Climate*, **17**, 2225-2240.
513
- 514 Sherwood, S.C., J. Lanzante, and C. Meyer, 2005: Radiosonde daytime biases and late 20th
515 century warming. *Science*, 309, 1556-1559.
516
- 517 Simmons, A.J., P.D. Jones, V. da Costa Bechtold, A.C.M. Beljaars, P.W. Kållberg,

- 518 S. Saarinen, S.M. Uppala, P. Viterbo and N. Wedi, 2004: Comparison of trends and
519 low-frequency variability in CRU, ERA-40 and NCEP/NCAR analyses of surface air
520 temperature. *J. Geophys. Res.*, **109**, D24115, doi:10.1029/2004JD006306
521
- 522 Thorne, P.W., D.E. Parker, J.R. Christy, and C.A. Mears, 2005: Causes of differences in
523 observed climate trends. *Bull. Amer. Meteor. Soc.*, in press

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

APPENDIX: STATISTICAL ISSUES REGARDING TRENDS

Tom M.L. Wigley

20 **Abstract:**

21

22 The purpose of this Appendix is to explain the statistical terms and methods used in this Report.

23 We begin by introducing a number of terms: mean, standard deviation, variance, linear trend,

24 sample, population, signal, and noise. Examples are given of linear trends in surface,

25 tropospheric, and stratospheric temperatures. The least squares method for calculating a best fit

26 linear trend is described. The method for quantifying the statistical uncertainty in a linear trend is

27 explained, introducing the concepts of standard error, confidence intervals, and significance

28 testing. A method to account for the effects of temporal autocorrelation on confidence intervals

29 and significance tests is described. The issue of comparing two data sets to decide whether

30 differences in their trends could have occurred by chance is discussed. The analysis of trends in

31 state-of-the-art climate model results is a special case because we frequently have an ensemble of

32 simulations for a particular forcing case. The effect of ensemble averaging on confidence

33 intervals is illustrated. Finally, the issue of practical versus statistical significance is discussed. In

34 practice, it is important to consider construction uncertainties as well as statistical uncertainties.

35 An example is given showing that these two sources of trend uncertainty can be of comparable

36 magnitude.

37

38 **(1) Why do we need statistics?**

39

40 Statistical methods are required to ensure that data are interpreted correctly and that apparent
41 relationships are meaningful (or “significant”) and not simply chance occurrences.

42

43 A “statistic” is a numerical value that describes some property of a data set. The most commonly
44 used statistics are the average (or “mean”) value, and the “standard deviation”, which is a
45 measure of the variability within a data set around the mean value. The “variance” is the square
46 of the standard deviation. The linear trend is another example of a data “statistic”.

47

48 Two important concepts in statistics are the “population” and the “sample”. The population is a
49 theoretical concept, an idealized representation of the set of all possible values of some measured
50 quantity. An example would be if we were able to measure temperatures continuously at a single
51 site for all time – the set of all values (which would be infinite in size in this case) would be the
52 population of temperatures for that site. A sample is what we actually see and can measure: i.e.,
53 what we have available for statistical analysis, and a necessarily limited subset of the population.
54 In the real world, all we ever have is limited samples, from which we try to estimate the
55 properties of the population.

56

57 As an analogy, the population might be an infinite jar of marbles, a certain proportion of which
58 (say 60%) is blue and the rest (40%) are red. We can only draw off a finite number of these
59 marbles (a sample) at a time; and, when we measure the numbers of blue and red marbles in the
60 sample, they need not be in the precise ratio 60:40. The ratio we measure is called a “sample

61 statistic”. It is an estimate of some hypothetical underlying population value (the corresponding
62 “population parameter”). The techniques of statistical science allow us to make optimum use of
63 the sample statistic and obtain a best estimate of the population parameter. Statistical science
64 also allows us to quantify the uncertainty in this estimate.

65

(2) Definition of a linear trend

67

68 If data show underlying smooth changes with time, we refer to these changes as a trend. The
69 simplest type of change is a linear (or straight line) trend, a continuous increase or decrease over
70 time. For example, the net effect of increasing greenhouse-gas concentrations and other human-
71 induced factors is expected to cause warming at the surface and in the troposphere and cooling in
72 the stratosphere (see Figure 1). Warming corresponds to a positive (or increasing) linear trend,
73 while cooling corresponds to a negative (or decreasing) trend. These changes are not expected to
74 be strictly linear, but the linear trend provides a simple way of characterizing the change and of
75 quantifying its magnitude.

76

77

78

79 Figure 1: Examples of temperature time series with best-fit (least squares) linear trends: A, global-mean surface
80 temperature from the UKMO Hadley Centre/Climatic Research Unit data set (HadCRUT2v); and B, MSU channel 4
81 data (T_4) for the lower stratosphere from the University of Alabama at Huntsville (UAH). Note the much larger
82 temperature scale on the lower panel. Temperature changes are expressed as anomalies relative to the 1979 to 1999
83 mean (252 months). Dates for the eruptions of El Chichón and Mt Pinatubo are shown by vertical lines. El Niños are
84 shown by the shaded areas.

85

86

87 Alternatively, there may be some physical process that causes a rapid switch or change from one
88 mode of behavior to another. In such a case the overall behavior might best be described as a
89 linear trend to the changepoint, a step change at this point, followed by a second linear trend
90 portion. Many temperature data sets show this type of behavior, arising from a change in the
91 pattern of variability in the Pacific that occurred around 1976 (a switch in a mode of climate
92 variability called the Pacific Decadal Oscillation).

93

94 Step changes can lead to apparently contradictory results. For example, a data set that shows an
95 initial cooling trend, followed by a large upward step, followed by a renewed cooling trend could
96 have an overall warming trend. To state simply that the data showed overall warming would
97 misrepresent the true underlying behavior.

98

99 A linear trend may therefore be deceptive if the trend number is given in isolation, removed from
100 the original data. Nevertheless, used appropriately, linear trends provide the simplest and most
101 convenient way to describe the overall change over time in a data set, and are widely used.

102

103 Linear temperature trends are usually quantified as the temperature change per year or per
104 decade (even when the data are available on a month by month basis). For example, the trend for
105 the surface temperature data shown below in Figure 1 is 0.169°C per decade. This is a more
106 convenient representation than the trend per month, which would be $0.169/120 = 0.00141^\circ\text{C}$ per
107 month, a very small number. An alternative method is to use the “total trend” over the full data
108 period – i.e., the total change for the fitted line from the start to the end of the record (see Figure
109 2 in the Executive Summary). In Figure 1, the data shown span January 1979 through December
110 2004 (312 months or 2.6 decades). The total change is therefore $0.169 \times 2.6 = 0.439^\circ\text{C}$.

111

112 **(3) Expected temperature changes: signal and noise**

113

114 Different physical processes generally cause different spatial and temporal patterns of change.

115 For example, anthropogenic emissions of halocarbons at the surface have led to a reduction in

116 stratospheric ozone and a contribution to stratospheric cooling over the past three or four

117 decades. Now that these chemicals are controlled under the Montreal Protocol, the

118 concentrations of the controlled species are decreasing and there is a trend towards a recovery of

119 the ozone layer. The eventual long-term effect on stratospheric temperatures is expected to be

120 non-linear: a cooling up until the late 1990s followed by a warming as the ozone layer recovers.

121

122 This is not the only process affecting stratospheric temperatures. Increasing concentrations of

123 greenhouse gases lead to stratospheric cooling; and explosive volcanic eruptions cause sharp, but

124 relatively short-lived stratospheric warmings (see Figure 1)¹. There are also natural variations,

125 most notably those associated with the Quasi-Biennial Oscillation (QBO)². Stratospheric

126 temperature changes (indeed, changes at all levels of the atmosphere) are therefore the combined

127 results of a number of different processes acting across all space and time scales.

128

129 In climate science, a primary goal is to identify changes associated with specific physical

130 processes (causal factors) or combinations of processes. Such changes are referred to as

131 “signals”. Identification of signals in the climate record is referred to as the “detection and

132 attribution” (D&A) problem. “Detection” is the identification of an unusual change, through the

133 use of statistical techniques like significance testing (see below); while “attribution” is the

134 association of a specific cause or causes with the detected changes in a statistically rigorous way.

135
136 The reason why D&A is a difficult and challenging statistical problem is because climate signals
137 do not occur in isolation. In addition to these signals, temperature fluctuations in all parts of the
138 atmosphere occur even in the absence of external driving forces. These internally-driven
139 fluctuations represent the “noise” against which we seek to identify specific externally-forced
140 signals. All climate records, therefore, are “noisy”, with the noise of this natural variability
141 tending to obscure the externally-driven changes. Figure 1 illustrates this. At the surface, a
142 primary noise component is the variability associated with ENSO (the El Niño/Southern
143 Oscillation phenomenon)¹, while, in the stratosphere, if our concern is to identify anthropogenic
144 influences, the warmings after the eruptions of El Chichón and Mt Pinatubo constitute noise.

145
146 If the underlying response to external forcing is small relative to the noise, then, by chance, we
147 may see a trend in the data due to random fluctuations purely as a result of the noise. The science
148 of statistics provides methods through which we can decide whether the trend we observe is
149 “real” (i.e., a signal associated with some causal factor) or simply a random fluctuation (i.e.,
150 noise).

151

(4) Deriving trend statistics

153

154 There are a number of different ways to quantify linear trends. Before doing anything, however,
155 we should always inspect the data visually to see whether a linear trend model is appropriate. For
156 example, in Fig. 1, the linear warming trend appears to be a reasonable description for the
157 surface data (top panel), but it is clear that a linear cooling model for the lower stratosphere
158 (lower panel) fails to capture some of the more complex changes that are evident in these data.
159 Nevertheless, the cooling trend line does give a good idea of the magnitude of the overall
160 change.

161

162 There are different ways to fit a straight line to the data. Most frequently, a “best fit” straight line
163 is defined by finding the particular line that minimizes the sum, over all data points, of the
164 squares of deviations about the line (these deviations are generally referred to as “residuals” or
165 “errors”). This is an example of a more general procedure called least squares regression.

166

167 In linear regression analysis, a predictand (Y) is expressed as a linear combination of one or
168 more predictors (X_i):

169

$$170 \quad Y_{\text{est}} = b_0 + b_1 X_1 + b_2 X_2 + \dots \quad \dots (1)$$

171

172 where the subscript ‘est’ is used to indicate that this is the estimate of Y that is given by the fitted
173 relationship. Differences between the actual and estimated values of Y, the residuals, are defined
174 by

175

$$176 \quad e = Y - Y_{\text{est}} \quad \dots (2)$$

177

178 For linear trend analysis of temperature data (T) there is a single predictor, time (t; t = 1,2,3, ...).

179 The time points are almost always evenly spaced, month by month, year by year, etc. – but this is

180 not a necessary restriction. In the linear trend case, the regression equation becomes:

181

$$182 \quad T_{\text{est}} = a + b t \quad \dots (3)$$

183

184 In equ. (3), ‘b’ is the slope of the fitted line – i.e., the linear trend value. This is a sample statistic,

185 i.e., it is an estimate of the corresponding underlying population parameter. To distinguish the

186 population parameter from the sample value, the population trend value is denoted β .

187

188 The formula for b is:

189

$$190 \quad b = [\sum((t - \langle t \rangle)T_t)] / [\sum(t - \langle t \rangle)^2] \quad \dots (4)$$

191

192 where $\langle \dots \rangle$ denotes the mean value, and the summation is over t = 1,2,3, ... n (i.e., the sample

193 size is n). T_t denotes the value of temperature, T, at time ‘t’. Equation (4) produces an unbiased

194 estimate³ of population trend, β .

195

196 For the usual case of evenly spaced time points, $\langle t \rangle = (n+1)/2$, and

197

$$\sum (t - \langle t \rangle)^2 = n(n^2 - 1)/12 \quad \dots (5)$$

199

200 When we are examining deviations from the fitted line the sign of the deviation is not important.

201 This is why we consider the squares of the residuals in least squares regression. An important

202 and desirable characteristic of the least squares method is that the average of the residuals is

203 zero.

204

205 Estimates of the linear trend are sensitive to points at the start or end of the data set. For

206 example, if the last point, by chance, happened to be unusually high, then the fitted trend might

207 place undue weight on this single value and lead to an estimate of the trend that was too high.

208 This is more of a problem with small sample sizes (i.e., for trends over short time periods). For

209 example, if we considered tropospheric data over 1979 through 1998, because of the unusual

210 warmth in 1998 (associated with the strong 1997/98 El Niño; see Figure 1), the calculated trend

211 may be an overestimate of the true underlying trend.

212

213 There are alternative ways to estimate the linear trend that are less sensitive to endpoints.

214 Although we recognize this problem, for the data used in this Report tests using different trend

215 estimators give results that are virtually the same as those based on the standard least-squares

216 trend estimator.

217

218

218 **(5) Trend uncertainties**

219

220 Some examples of fitted linear trend lines are shown in Figure 1. This Figure shows monthly
221 temperature data for the surface and for the lower stratosphere (MSU channel 4) over 1979
222 through 2004 (312 months). In both cases there is a clear trend, but the fit is better for the surface
223 data. The trend values (i.e., the slopes of the best fit straight lines that are shown superimposed
224 on monthly data) are $+0.169^{\circ}\text{C}/\text{decade}$ for the surface and $-0.452^{\circ}\text{C}/\text{decade}$ for the stratosphere.
225 For the stratosphere, although there is a pronounced overall cooling trend, as noted above
226 describing the change simply as a linear cooling considerably oversimplifies the behavior of the
227 data¹.

228

229 A measure of how well the straight line fits the data (i.e., the “goodness of fit”) is the average
230 value of the squares of the residuals. The smaller this is, the better is the fit. The simplest way to
231 define this average would be to divide the sum of the squares of the residuals by the sample size
232 (i.e., the number of data points, n). In fact, it is usually considered more correct to divide by $n - 2$
233 rather than n , because some information is lost as a result of the fitting process and this loss of
234 information must be accounted for. Dividing by $n - 2$ is required in order to produce an unbiased
235 estimator.

236

237 The population parameter we are trying to estimate here is the standard deviation of the trend
238 estimate, or its square, the variance of the distribution of b , which we denote $\text{Var}(b)$. The larger
239 the value of $\text{Var}(b)$, the more uncertain is b as an estimate of the population value, \square .

240

241 The formula for $\text{Var}(b)$ is ...

242

$$243 \quad \text{Var}(b) = [\sigma^2]/[\sum(t - \langle t \rangle)^2] \quad \dots (6)$$

244

245 where σ^2 is the population value for the variance of the residuals. Unfortunately, we do not in

246 general know what σ^2 is, so we must use an unbiased sample estimate of σ^2 . This estimate is

247 known as the Mean Square Error (MSE), defined by ...

248

$$249 \quad \text{MSE} = [\sum(e^2)]/(n - 2) \quad \dots (7)$$

250

251 Hence, equ. (6) becomes

252

$$253 \quad \text{Var}(b) = (\text{SE})^2 = \text{MSE}/[\sum(t - \langle t \rangle)^2] \quad \dots (8)$$

254

255 where SE, the square root of $\text{Var}(b)$, is called is called the “standard error” of the trend estimate.

256 The smaller the value of the standard error, the better the fit of the data to the linear change

257 description and the smaller the uncertainty in the sample trend as an estimate of the underlying

258 population trend value. The standard error is the primary measure of trend uncertainty. The

259 standard error will be large if the MSE is large, and the MSE will be large if the data points show

260 large scatter about the fitted line.

261

262 There are assumptions made in going from equ. (6) to (8): viz. that the residuals have mean zero

263 and common variance, that they are Normally (or “Gaussian”) distributed⁴, and that they are

264 uncorrelated or statistically independent. In climatological applications, the first two are
265 generally valid. The third assumption, however, is often not justified. We return to this below.
266

267 **(6) Confidence intervals and significance testing**

268

269 In statistics we try to decide whether a trend is an indication of some underlying cause, or merely
270 a chance fluctuation. Even purely random data may show periods of noticeable upward or
271 downward trends, so how do we identify these cases?

272

273 There are two common approaches to this problem, through significance testing and by defining
274 confidence intervals. The basis of both methods is the determination of the “sampling
275 distribution” of the trend, i.e., the distribution of trend estimates that would occur if we analyzed
276 data that were randomly scattered about a given straight line with slope β . This distribution is
277 approximately Gaussian with a mean value equal to β and a variance (standard deviation
278 squared) given by equ. (8). More correctly, the distribution to use is Student’s ‘t’ distribution,
279 named after the pseudonym ‘Student’ used by the statistician William Gosset. For large samples,
280 however (n more than about 30), the distribution is very nearly Gaussian.

281

282 ***Confidence intervals***

283

284 The larger the standard error of the trend, the more uncertain is the slope of the fitted line. We
285 express this uncertainty probabilistically by defining confidence intervals for the trend associated
286 with different probabilities. If the distribution of trend values were strictly Gaussian, then the
287 range $b - SE$ to $b + SE$ would represent the 68% confidence interval (C.I.) because the
288 probability of a value lying in that range for a Gaussian distribution is 0.68. The range $b -$
289 $1.645(SE)$ to $b + 1.645(SE)$ would give the 90% C.I.; the range $b - 1.96(SE)$ to $b + 1.96(SE)$

290 would give the 95% C.I.; and so on. Quite often, for simplicity, we use $b - 2(SE)$ to $b + 2(SE)$ to
291 represent (to a good approximation) the 95% confidence interval.

292

293 Because of the way C.I.s are usually represented graphically, as a bar centered on the best-fit
294 estimate, they are often referred to as “error bars”. Confidence intervals may be expressed in two
295 ways, either (as above) as a range, or as a signed error magnitude. The approximate 95%
296 confidence interval, therefore, may be expressed as $b \pm 2(SE)$, with appropriate numerical values
297 inserted for b and SE .

298

299 As will be explained further below, showing confidence interval for linear trends may be
300 deceptive, because the purely statistical uncertainties that they represent are not the only sources
301 of uncertainty. Such confidence intervals quantify only one aspect of trend uncertainty, that
302 arising from statistical noise in the data set. There are many other sources of uncertainty within
303 any given data set and these may be as or more important than statistical uncertainty. Showing
304 just the statistical uncertainty may therefore provide a false sense of accuracy in the calculated
305 trend.

306

307 *Significance testing*

308

309 An alternative method for assessing trends is hypothesis testing. In practice, it is much easier to
310 disprove rather than prove a hypothesis. Thus, the standard statistical procedure in significance
311 testing is to set up a hypothesis that we would like to disprove. This is called a “null hypothesis”.
312 In the linear trend case, we are often interested in trying to decide whether an observed data trend

313 that is noticeably different from zero is sufficiently different that it could not have occurred by
314 chance – or, at least, that the probability that it could have occurred by chance is very small. The
315 appropriate null hypothesis in this case would be that there was no underlying trend ($\square = 0$). If
316 we disprove (i.e., “reject”) the null hypothesis, then we say that the observed trend is
317 “statistically significant” at some level of confidence and we must accept some alternate
318 hypothesis. The usual alternate hypothesis in temperature analyses is that the data show a real,
319 externally-forced warming (or cooling) trend. (In cases like this, the statistical analysis is
320 predicated on the assumption that the observed data are reliable. If a trend were found to be
321 statistically significant, then an alternative possibility might be that the observed data were
322 flawed.)

323

324 An alternative null hypothesis that often arises is when we are comparing an observed trend with
325 some model expectation. Here, the null hypothesis is that the observed trend is equal to the
326 model value. If our results led us to reject this null hypothesis, then (assuming again that the
327 observed data are reliable) we would have to infer that the model result was flawed – either
328 because the external forcing applied to the model was incorrect and/or because of deficiencies in
329 the model itself.

330

331 An important factor in significance testing is whether we are concerned about deviations from
332 some hypothesized value in any direction or only in one direction. This leads to two types of
333 significance test, referred to as “one-tailed” (or “one-sided”) and “two-tailed” tests. A one-tailed
334 test arises when we expect a trend in a specific direction (such as warming in the troposphere due
335 to increasing greenhouse-gas concentrations). Two-tailed tests arise when we are concerned only

336 with whether the trend is different from zero, with no specification of whether the trend should
337 be positive or negative. In temperature trend analyses we generally know the sign of the expected
338 trend, so one-tailed tests are more common.

339

340 The approach we use in significance testing is to determine the probability that the observed
341 trend could have occurred by chance. As with the calculation of confidence intervals, this
342 involves calculating the uncertainty in the fitted trend arising from the scatter of points about the
343 trend line, determined by the standard error of the trend estimate (equ. (8)). It is the ratio of the
344 trend to the standard error (b/SE) that determines the probability that a null hypothesis is true or
345 false. A large ratio (greater than 2, for example) would mean that (except for very small samples)
346 the 95% C.I. did not include the zero trend value. In this case, the null hypothesis is unlikely to
347 be true, because the zero trend value, the value assumed under the null hypothesis, lies outside
348 the range of trend values that are likely to have occurred purely by chance.

349

350 If the probability that the null hypothesis is true is small, and less than a predetermined threshold
351 level such as 0.05 (5%) or 0.01 (1%), then the null hypothesis is unlikely to be correct. Such a
352 low probability would mean that the observed trend could only have occurred by chance one
353 time in 20 (or one time in 100), a highly unusual and therefore “significant” result. In technical
354 terms we would say that “the null hypothesis is rejected at the prescribed significance level”, and
355 declare the result “significant at the 5% (or 1%) level”. We would then accept the alternate
356 hypothesis that there was a real deterministic trend and, hence, some underlying causal factor.

357

358 Even with rigorous statistical testing, there is always a small probability that we might be wrong
359 in rejecting a null hypothesis. The reverse is also true – we might accept a null hypothesis of no
360 trend even when there is a real trend in the data. This is more likely to happen when the sample
361 size is small. If the real trend is small and the magnitude of variability about the trend is large, it
362 may require a very large sample in order to identify the trend above the background noise.

363

364 For the null hypothesis of zero trend, the distribution of trend values has mean zero and standard
365 deviation equal to the standard error. Knowing this, we can calculate the probability that the
366 actual trend value could have exceeded the observed value by chance if the null hypotheses were
367 true (or, if we were using a two-tailed test, the probability that the magnitude of the actual trend
368 value exceeded the magnitude of the observed value). This probability is called the ‘p-value’. For
369 example, a p-value of 0.03 would be judged significant at the 5% level (since $0.03 < 0.05$), but not
370 at the 1% level (since $0.03 > 0.01$).

371

372 Since both the calculation of confidence intervals and significance testing employ information
373 about the distribution of trend values, there is a clear link between confidence intervals and
374 significance testing.

375

376 *A complication; the effect of autocorrelation*

377

378 The significance of a trend, and its confidence intervals, depend on the standard error of the trend
379 estimate. The formula given above for this standard error (equ. (8)) is, however, only correct if
380 the individual data points are unrelated, or statistically independent. This is not the case for most

381 temperature data, where a value at a particular time usually depends on values at previous times;
382 i.e., if it is warm today, then, on average, it is more likely to be warm tomorrow than cold. This
383 dependence is referred to as “temporal autocorrelation” or “serial correlation”. When data are
384 autocorrelated (i.e., when successive values are not independent of each other), many statistics
385 behave as if the sample size was less than the number of data points, n .

386
387 One way to deal with this is to determine an “effective sample size”, which is less than n , and
388 use it instead of n in statistical formulae and calculations. The extent of this reduction from n to
389 an effective sample size depends on how strong the autocorrelation is. Strong autocorrelation
390 means that individual values in the sample are far from being independent, so the effective
391 number of independent values must be much smaller than the sample size. Strong autocorrelation
392 is common in temperature time series. This is accounted for by reducing the divisor ‘ $n - 2$ ’ in the
393 mean square error term (equ. (7)) that is crucial in determining the standard error of the trend
394 (equ. (8)).

395
396 There are a number of ways that this autocorrelation effect may be quantified. A common and
397 relatively simple method is described in Santer et al. (2000). This method makes the assumption
398 that the autocorrelation structure of the temperature data may be adequately described by a “first-
399 order autoregressive” process, an assumption that is a good approximation for most climate data.
400 The lag-1 autocorrelation coefficient (r_1) is calculated from the observed data⁵, and the effective
401 sample size is determined by

402
403
$$n_{\text{eff}} = n (1 - r_1) / (1 + r_1) \quad \dots (9)$$

404

405 There are more sophisticated methods than this, but testing on observed data shows that this
406 method gives results that are very similar to those obtained by more sophisticated methods.

407

408 If the effective sample size is noticeably smaller than n , then, from equs. (7) and (8) it can be
409 seen that the standard error of the trend estimate may be much larger than one would otherwise
410 expect. Since the width of any confidence interval depends directly on this standard error (larger
411 SE leading to wider confidence intervals), then the effect of autocorrelation is to produce wider
412 confidence intervals and greater uncertainty in the trend estimate. A corollary of this is that
413 results that may show a significant trend if autocorrelation is ignored are frequently found to be
414 non-significant when autocorrelation is accounted for.

415

416 **(7) Comparing trends in two data sets**

417

418 Assessing the magnitude and confidence interval for the linear trend in a given data set is
419 standard procedure in climate data analysis. Frequently, however, we want to compare two data
420 sets and decide whether differences in their trends could have occurred by chance. Some
421 examples are:

422

423 (a) comparing data sets that purport to represent the same variable (such as two versions of a
424 satellite data set) – an example is given in Figure 2;

425 (b) comparing the same variable at different levels in the atmosphere (such as surface and
426 tropospheric data); or

427 (c) comparing models and observations.

428

429

430

431

432 Figure 2: Three estimates of temperature changes for MSU channel 2 (T_2), expressed as anomalies relative to the
433 1979 to 1999 mean. Data are from: A, the University of Alabama at Huntsville (UAH); B, Remote Sensing Systems
434 (RSS); and C, the University of Maryland (U.Md.) The estimates employ the same ‘raw’ satellite data, but make
435 different choices for the adjustments required to merge the various satellite records and to correct for instrument
436 biases. The statistical uncertainty is virtually the same for all three series. Differences between the series give some
437 idea of the magnitude of structural uncertainties. Volcano eruption and El Niño information are as in Figure 1.
438

439

440 In the first case (Figure 2), we know that the data sets being compared are attempts to measure
441 precisely the same thing, so that differences can arise only as a result of differences in the
442 methods used to create the final data sets from the same ‘raw’ original data. Here, there is a
443 pitfall that some practitioners fall prey to by using what, at first thought, seems to be a
444 reasonable approach. In this naïve method, one would first construct C.I.s for the individual trend
445 estimates by applying the single sample methods described above. If the two C.I.s overlapped,

446 then we would conclude that there was no significant difference between the two trends. This
447 approach, however, is seriously flawed.

448

449 An analogous problem, comparing two means rather than two trends, discussed by Lanzante
450 (2005), gives some insights. In this case, it is necessary to determine the standard error for the
451 difference between two means. If this standard error is denoted ‘s’, and the individual standard
452 errors are s_1 and s_2 , then

453

$$454 \quad s^2 = (s_1)^2 + (s_2)^2 \quad \dots(10)$$

455

456 The new standard error is often called the pooled standard error, and the pooling method is
457 sometimes called “combining standard errors in quadrature”. In some cases, when the trends
458 come from data series that are unrelated (as in the model/observed data comparison case; (c)
459 above) a similar method may be applied to trends. If the data series are correlated with each
460 other, however (cases (a) and (b)), this procedure is not correct. Here, the correct method is to
461 produce a difference time series by subtracting the first data point in series 1 from the first data
462 point in series 2, the second data points, the third data points, etc. The result of doing this with
463 the microwave sounding unit channel 2 (MSU T₂) data shown in Figure 2 is shown in Figure 3.
464 To assess the significance of trend differences we then apply the same methods used for trend
465 assessment in a single data series to the difference series.

466

467

468

469 Figure 3: Difference series for the MSU T₂ series shown in Figure 2. Variability about the trend line is least for the
470 UAH minus RSS series indicating closer correspondence between these two series than between U.Md. and either
471 UAH or RSS.
472

473

474 Analyzing differences removes the variability that is common to both data sets and isolates those
475 differences that may be due to differences in data set production methods, temperature
476 measurement methods (as in comparing satellite and radiosonde data), differences in spatial
477 coverage, etc.

478

479 Figures 2 and 3 provide a striking example of this. Here, the three series in Figure 2 have very
480 similar volcanic and ENSO signatures. In the individual series, these aspects are noise that
481 obscures the underlying linear trend and inflates the standard error and the trend uncertainty.

482 Since this noise is common to each series, differencing has the effect of canceling out a large
483 fraction of the noise. This is clear from Figure 3, where the variability about the trend lines is
484 substantially reduced. Figure 4 shows the effects on the trend confidence intervals (taking due
485 account of autocorrelation effects). Even though the individual series look very similar in Figure
486 2, this is largely an artifact of similarities in the noise. It is clear from Figures 3 and 4 that there
487 are, in fact, very significant differences in the trends, reflecting differences in their methods of
488 construction.

489

490

491

492 Figure 4: 95% confidence intervals for the three MSU T_2 series shown in Figure 2 (see Table 3.3 in Chapter 3), and
493 for the three difference series shown in Figure 3.
494

495
496 Comparing model and observed data for a single variable, such as surface temperature,
497 tropospheric temperature, etc., is a different problem. Here, when using data from a state-of-the-
498 art climate model (a coupled Atmosphere/Ocean General Circulation Model⁶, or “AOGCM”),
499 there is no reason to expect the background variability to be common to both the model and
500 observations. AOGCMs generate their own internal variability entirely independently of what is
501 going on in the real world. In this case, standard errors for the individual trends can be combined
502 in quadrature (equ. (10)). (There are some model/observed data comparison cases where an
503 examination of the difference series may still be appropriate, such as in experiments where an
504 atmospheric GCM is forced by observed sea surface temperature variations so that ocean-related
505 variability should be common to both the observations and the model.)
506

507 For other comparisons, the appropriate test will depend on the degree of similarity between the
508 data sets expected for perfect data. For example, a comparison between MSU T_2 and MSU T_{2LT}
509 produced by a single group should use the difference test – although interpretation of the results
510 may be tricky because differences may arise either from construction methods or may represent
511 real physical differences arising from the different vertical weighting profiles, or both.
512

513 There is an important implication of this comparison issue. While it may be common practice to
514 use error bars to illustrate C.I.s for trends of individual time series, when the primary concern (as
515 it is in many parts of this Report) is the comparison of trends, individual C.I.s can be quite

516 misleading. In some cases in this Report, therefore, where it might seem that error bars should be
517 given, we consider the disadvantage of their possible misinterpretation to outweigh their
518 potential usefulness. Instead, we have chosen to express individual trend uncertainties through
519 the use of significance levels, which can be represented by a less obtrusive symbol. As noted in
520 Section (9) below, there are other reasons why error bars can be misleading.

521

(8) Multiple AOGCM simulations

523

524 Both models and the real world show weather variability and other sources of internal variability
525 that are manifest on all time scales, from daily up to multi-decadal. With AOGCM simulations
526 driven by historical forcing spanning the late-19th and 20th Centuries, therefore, a single run with
527 a particular model will show not only the externally-forced signal, but also, superimposed on
528 this, underlying internally-generated variability that is similar to the variability we see in the real
529 world. In contrast to the real world, however, in the model world we can perturb the model's
530 initial conditions and re-run the same forcing experiment. This will give an entirely different
531 realization of the model's internal variability. In each case, the output from the model is a
532 combination of signal (the response to the forcing) and noise (the internally-generated
533 component). Since the noise parts of each run are unrelated, averaging over a number of
534 realizations will tend to cancel out the noise and, hence, enhance the visibility of the signal. It is
535 common practice, therefore, for any particular forcing experiment with an AOGCM, to run
536 multiple realizations of the experiment (i.e., an ensemble of realizations). An example is given
537 in Figure 5, which shows four separate realizations and their ensemble average for a simulation
538 using realistic 20th Century forcing (both natural and anthropogenic).

539

540

541
542 Figure 5: Four separate realizations of model realizations of global-mean MSU channel 2 (T_2) temperature changes,
543 and their ensemble average, for a simulation using realistic 20th Century forcing (both natural and anthropogenic)
544 carried out with one of the National Centre for Atmospheric Research's AOGCMs, the Parallel Climate Model
545 (PCM). The cooling events around 1982/3 and 1991/2 are the result of imposed forcing from the eruptions of El
546 Chichón (1982) and Mt. Pinatubo (1991). Note that the El Chichón cooling is more obvious than in the observed
547 data shown in Fig. 1, because, in the model simulations, the ENSO sequences differed from the real world, and from
548 each other.
549

550
551 This provides us with two different ways to assess the uncertainties in model results, such as in
552 the model-simulated temperature trend over recent decades. One method is to express
553 uncertainties using the spread of trends across the ensemble members (see, e.g., Figures 3 and 4
554 in the Executive Summary). Alternatively, the temperature series from the individual ensemble
555 members may be averaged and the trend and its uncertainty calculated using these average data.

556
557 Ensemble averaging, however, need not reduce the width of the trend confidence interval
558 compared with an individual realization. This is because of compensating factors: the time series
559 variability will be reduced by the averaging process (as is clear in Figure 5), but, because

560 averaging can inflate the level of autocorrelation, there may be a compensating increase in
561 uncertainty due to a reduction in the effective sample size. This is illustrated in Figure 6.
562

563

564 Figure 6: 95% confidence intervals for individual model realizations of MSU T₂ temperature changes (as shown in
565 Fig. 5), compared with the 95% confidence interval for the ensemble (n=4) average.
566

567 Averaging across ensemble members, however, does produce a net gain. Although the width of
568 the C.I. about the mean trend may not be reduced relative to individual trend C.I.s, averaging
569 leaves just a single best-fit trend rather than a spread of best-fit trend values.

570 **(9) Practical versus statistical significance**

571

572 The Sections above have been concerned primarily with statistical uncertainty, uncertainty
573 arising from random noise in climatological time series – i.e., the uncertainty in how well a data
574 set fits a particular ‘model’ (a straight line in the linear trend case). Statistical noise, however, is
575 not the only source of uncertainty in assessing trends. Indeed, as amply illustrated in this Report,
576 other sources of uncertainty may be more important.

577

578 The other sources of uncertainty are the influences of non-climatic factors. These are referred to
579 in this Report as “construction uncertainties”. When we construct climate data records that are
580 going to be used for trend analyses, we attempt to minimize construction uncertainties by
581 removing, as far as possible, non-climatic biases that might vary over time and so impart a
582 spurious trend or trend component – a process referred to as “homogenization”.

583

584 The need for homogenization arises in part because most observations are made to serve the
585 short-term needs of weather forecasting (where the long-term stability of the observing system is
586 rarely an important consideration). Most records therefore contain the effects of changes in
587 instrumentation, instrument exposure, and observing practices made for a variety of reasons.
588 Such changes generally introduce spurious non-climatic changes into data records that, if not
589 accounted for, can mask (or possibly be mistaken for) an underlying climate signal.

590

591 An added problem arises because temperatures are not always measured directly, but through
592 some quantity related to temperature. Adjustments must therefore be made to obtain temperature

593 information. The satellite-based microwave sounding unit (MSU) data sets provide an important
594 example. For MSU temperature records, the quantity actually measured is the upwelling
595 emission of microwave radiation from oxygen atoms in the atmosphere. MSU data are also
596 affected by numerous changes in instrumentation and instrument exposure associated with the
597 progression of satellites used to make these measurements.

598

599 Thorne et al. (2005) divide construction uncertainty into two components: “structural
600 uncertainty” and “parametric uncertainty”. Structural uncertainty arises because there is no *a*
601 *priori* knowledge of the correct way to homogenize a given raw data set. Independent
602 investigators given the same raw data will make different seemingly sensible and defensible
603 adjustment choices based on their training, technological options at their disposal, and their
604 understanding of the raw data, amongst other factors. Differences in the choice of adjustment
605 pathway and its structure lead to structural uncertainties. Parametric uncertainty arises because,
606 once an adjustment approach or pathway has been chosen, additional choices may have to be
607 made with regard to specific correction factors or parameters.

608

609 Sensitivity studies using different parameter choices may allow us to quantify parametric
610 uncertainty, but this is not always done. Quantifying structural uncertainty is very difficult
611 because it involves consideration of a number of fundamentally different (but all plausible)
612 approaches to data set homogenization, rather than simple parameter “tweaking”. Differences
613 between results from different investigators give us some idea of the magnitude of structural
614 uncertainty, but this is a relatively weak constraint. There are a large number of conceivable
615 approaches to homogenization of any particular data set, from which we are able only to consider

616 a small sample – and this may lead to an under-estimation of structural uncertainty. Equally, if
617 some current homogenization techniques are flawed then the resulting uncertainty estimate will
618 be too large.

619

620 An example is given above in Figure 2, showing three different MSU T₂ records with trends of
621 0.044°C/decade, 0.129°C/decade, and 0.199°C/decade over 1979 through 2004. These
622 differences, ranging from 0.070°C/decade to 0.155°C/decade, represent a considerable degree of
623 construction uncertainty. For comparison, the statistical uncertainty, which is very similar for
624 each series and which can be quantified by the 95% confidence interval, is ±0.066 to
625 ±0.078°C/decade.

626

627 An important implication of this comparison is that statistical and construction uncertainties may
628 be of similar magnitude. For this reason, showing, through confidence intervals, information
629 about statistical uncertainty alone, without giving any information about construction
630 uncertainty, can be misleading.

631 Footnotes

632

633 ¹ Figure 1 shows a number of interesting features. In the stratosphere, the warmings following
634 the eruptions of El Chichón (April 1982) and Mt Pinatubo (June 1991) are pronounced. For El
635 Chichón, the warming appears to start before the eruption, but this is just a chance natural
636 fluctuation. The overall cooling trend is what is expected to occur due to anthropogenic
637 influences. At the surface, on short time scales, there is a complex combination of effects. There
638 is no clear cooling after El Chichón, primarily because this was offset by the very strong 1982/83
639 El Niño. Cooling after Pinatubo is more apparent, but this was also partly offset by the El Niño
640 around 1992/93 (which was much weaker than that of 1982/83). El Niño events, characterized by
641 warm temperatures in the tropical Pacific, have a noticeable effect on global-mean temperature,
642 but the effect lags behind the Pacific warming by 3-7 months. This is very clear in the surface
643 temperature changes at and immediately after the 1986/87 and 1997/98 El Niños, also very large
644 events. The most recent El Niños were weak and have no clear signature in the surface
645 temperatures.

646

647 ² The QBO is a quasi-periodic reversal in winds in the tropical stratosphere that leads to
648 alternating warm and cold tropical stratospheric temperatures with a periodicity of 18 to 30
649 months.

650

651 ³ An unbiased estimator is one where, if the same experiment were to be performed over and
652 over again under identical conditions, then the long-run average of the estimator will be equal to
653 the parameter that we are trying to estimate. In contrast, in a biased estimator, there will always

654 be some slight difference between the long-run average and the true parameter value that does
655 not tend to zero no matter how many times the experiment is repeated. Since our goal is to
656 estimate population parameters, it is clear that unbiased estimators are preferred.

657

658 ⁴ The “Gaussian” distribution (often called the “Normal” distribution) is the most well-known
659 probability distribution. This has a characteristic symmetrical “bell” shape, and has the property
660 that values near the center (or mean value) of the distribution are much more likely than values
661 far from the center.

662

663 ⁵ From the time series of residuals about the fitted line.

664

665 ⁶ An AOGCM interactively couples together a three-dimensional ocean General Circulation
666 Model (GCM) and an atmospheric GCM (AGCM). The components are free to interact with one
667 another and they are able to generate their own internal variability in much the same way that the
668 real-world climate system generates its internal variability (internal variability is variability that
669 is unrelated to external forcing). This differs from some other types of model (e.g, an AGCM)
670 where there can be no component of variability arising from the ocean. An AGCM, therefore,
671 cannot generate variability arising from ENSO, which depends on interactions between the
672 atmosphere and ocean.

673

674 **References:**

675

676 Santer, B.D., Wigley, T.M.L., Boyle, J.S., Gaffen, D.J., Hnilo J.J., Nychka, D., Parker, D.E. and
677 Taylor, K.E., 2000: Statistical significance of trends and trend differences in layer-average
678 temperature time series. *Journal of Geophysical Research* **105**, 7337–7356.

679

680 Thorne, P.W., Parker, D.W., Christy, J.R. and Mears, C.A., 2005: Uncertainties in climate
681 trends: lessons from upper-air temperature records. *Bulletin of the American*
682 *Meteorological Society* **86**, 1437–1442.

683

684 Lanzante, J.R., 2005: A cautionary note on the use of error bars. *Journal of Climate* **18**, 3699–
685 3703.

686

Appendix B

Climate Change Science Program Temperature Trends in the Lower Atmosphere Steps for Understanding and Reconciling Differences

Members of the Assessment/Synthesis Product Team

Director, Climate Change Science Program: James R. Mahoney
Director, Climate Change Science Program Office: Richard H. Moss

Chief Editor: Thomas R. Karl, NOAA
Associate Editor: Christopher D. Miller, NOAA
Associate Editor: William L. Murray, STG, Inc.
Technical Editor: Susan Joy Hassol, STG, Inc.
Graphic Designer: Sara Veasey, NOAA
Technical Support: Erin McKay, STG, Inc.

Preface

Thomas R. Karl, NOAA; Christopher D. Miller, NOAA; William L. Murray, STG Inc.

Executive Summary

Convening Lead Author: Tom M. L. Wigley, NSF NCAR
Lead Authors: V. Ramaswamy, NOAA; John R. Christy, Univ. of AL (Huntsville);
John Lanzante, NOAA; Carl A. Mears, Remote Sensing Systems; Ben D. Santer, DOE, LLNL
and Chris K. Folland, U.K. Met Office

Chapters:

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?

Convening Lead Author: V. Ramaswamy, NOAA
Lead Authors: James W. Hurrell, NSF NCAR and Gerald A. Meehl, NSF NCAR
Contributing Authors: A. Phillips, NCAR; Ben D. Santer, DOE, LLNL; M. D. Schwarzkopf, Dian J. Seidel, NOAA, Steve Sherwood, Yale Univ.; Peter W. Thorne, U.K. Met Office

2. What kinds of atmospheric temperature variations can the current observing systems measure and what are their strengths and limitations, both spatially and temporally?

Convening Lead Author: John Christy, Univ. of AL (Huntsville)
Lead Authors: Dian J. Seidel, NOAA; Steve Sherwood, Yale Univ.
Contributing Authors: Adrian Simmons, Ming Cai, FL State Univ.; Eugenia Kalnay, Univ. of MD;
Chris K. Folland, U.K. Met Office, Carl A. Mears, Remote Sensing Systems; Peter W. Thorne, U. K.
Met Office; John R. Lanzante, NOAA

3. What do observations indicate about the changes of temperature in the atmosphere and at the surface since the advent of measuring temperatures vertically?

Convening Lead Author: John R. Lanzante, NOAA

Lead Authors: Thomas C. Peterson, NOAA; Frank J. Wentz, Remote Sensing Systems; John Christy, Univ. of AL (Huntsville); Chris E. Forest, MIT and Konstantin Y. Vinnikov, Univ. of MD

Contributing Authors: Dian J. Seidel, NOAA; Carl A. Mears, Remote Sensing Systems; Russell S. Vose, NOAA; Peter W. Thorne, U. K. Met Office; Norman C. Grody, NOAA and John R. Christy, Univ. of AL (Huntsville)

4. What is our understanding of the contribution made by observational or methodological uncertainties to the previously reported vertical differences in temperature trends?

Convening Lead Author: Carl A. Mears, Remote Sensing Systems

Lead Authors: Chris E. Forest, MIT; Roy Spencer, Univ. of AL (Huntsville); Russell S. Vose, NOAA and Richard W. Reynolds, NOAA

Contributing Authors: Peter W. Thorne, U. K. Met Office and John R. Christy, Univ. of AL (Huntsville)

5. How well can the observed vertical temperature changes be reconciled with our understanding of the causes of these temperature changes?

Convening Lead Author: Ben D. Santer, DOE LLNL

Lead Author: Joyce E. Penner, Univ. of MI and Peter W. Thorne, U.K. Met Office

Contributing Authors: W. Collins, K. Dixon, T.L. Delworth, C. Doutriaux, Chris K. Folland, U.K. Met Office; Chris E. Forest, MIT; I. Held, GFDL; John Lanzante, NOAA; Gerald A. Meehl, NSF NCAR, V. Ramaswamy, NOAA; Dian J. Seidel, NOAA; M.F. Wehner, and Tom M.L. Wigley, NSF NCAR

6. What measures can be taken to improve the understanding of observed changes?

Convening Lead Author: Chris E. Folland, U.K. Met Office

Lead Authors: David Parker, U.K. Met Office; Richard W. Reynolds, NOAA; Steve Sherwood, Yale Univ.; Peter W. Thorne, U.K. Met Office

Appendix

Appendix A: Statistical Issues Regarding Trends: Tom Wigley, NCAR

Appendix B: Members of the Assessment/Synthesis Product Team: James R. Mahoney, CCSP; Richard H. Moss, CCSP; Thomas R. Karl, NOAA; Christopher D. Miller, NOAA; William L. Murray, STG, Inc.; Sara Veasey, NOAA; Erin McKay, STG, Inc.

Glossary & Acronyms, Symbols, Abbreviations

1	Glossary	
2		
3	Aerosols	tiny particles suspended in the air
4		
5	Adjusted	refers to time series data that has been homogenized for time
6		dependent biases; owing to uncertainties inherent in data bias
7		removal, the term “adjusted” is often used instead of “corrected”
8		
9	Albedo	the fraction of incident light that is reflected
10		
11	Anthropogenic	human-induced
12		
13	Black carbon	soot particles primarily from fossil fuel burning
14		
15	Climate sensitivity:	the equilibrium change in global-average surface air temperature
16		following a change in radiative forcing; in current usage, this term
17		generally refers to the warming that would result if atmospheric
18		carbon dioxide concentrations were to double from their pre-
19		industrial levels
20		
21	Contrails	condensation trails from aircraft
22		
23	Convection	motions in a fluid or the air that are predominantly vertical and
24		driven by buoyancy forces; a principal means of vertical energy
25		transfer
26		
27	Diurnal	occurring daily; varying within the course of a day
28		
29	Dewpoint	temperature at which water vapor condenses into liquid water
30	temperature	when cooled at constant pressure
31		
32	Error	the difference between an estimated or observed value and the true
33		value
34		
35	Forcing	a factor that influences climate
36		
37	Greenhouse gases	gases including water vapor, carbon dioxide, methane, nitrous
38		oxide, and halocarbons that trap infrared heat, warming the air near
39		the surface and lower levels of the atmosphere
40		
41	Homogenization	Removing changes in time series data that might have arisen for
42		non-climatic reasons
43		
44	Internal variability	natural cycles and variations in climate
45		
46		

47	Temperature inversion	
48		a condition in which the air temperature increases with height, in
49		contrast to the more common situation in which temperature
50		decreases with altitude
51		
52	Isothermal	constant temperature; often refers to a temperature profile meaning
53		constant temperature with height
54		
55	Lapse rate	the rate at which temperature decreases with increasing elevation
56		
57	Latent heat	the heat required to change the phase of a substance (solid to
58		liquid, liquid to vapor or solid to vapor); the temperature does not
59		change during this process; the heat is released for the reverse
60		process
61		
62	Latent heat of water	the energy released or gained by water during phase changes
63		
64	Metadata	supplemental records used to interpret measurements, such as how
65		and where measurements were collected and processed
66		
67	Moist enthalpy	sensible and latent heat content of the air; includes contributions
68		from temperature and from the amount of water vapor expressed as
69		specific humidity
70		
71	Parameterization	a mathematical representation of a process that cannot be explicitly
72		resolved in a climate model
73		
74	Radiosonde	a balloon carrying a thermometer or other sensing device that takes
75		measurements in the atmosphere and transmits them by radio to a
76		data recorder on the surface
77		
78	Reanalysis	a mathematically blended record that incorporates a variety of
79		observational data sets (with adjustments) in an assimilation model
80		
81	Reference networks	a small subset of sites consisting of multiple instruments that
82		independently measure the same variable which if well
83		coordinated could provide full characterization of instrument
84		errors and biases, significantly reducing uncertainty in observed
85		climate change
86		
87	Relative humidity	the percentage of water vapor in the air relative to what is required
88		for saturation to occur at a given temperature
89		
90	Sensible heat	heat that can be measured by a thermometer
91		

92 **Specific humidity** the amount of water vapor in the air in units of kilograms of water
 93 vapor per kilogram of air
 94

95 **Trend** a systematic change over time
 96
 97
 98

99 **Acronyms, Symbols, Abbreviations**

100	
101	AATSR Advanced Along-Track Scanning Radiometer
102	AGCM Atmospheric General Circulation Model
103	AIRS Atmospheric InfraRed Sounder
104	AMIP Atmospheric Model Intercomparison Project
105	AMSR Advanced Microwave Scanning Radiometer
106	AMSU Advanced Microwave Sounding Unit
107	AOGCM Atmosphere-Ocean General Circulation Model
108	ATMS Advanced Technology Microwave Sounder
109	ATSR Along-Track Scanning Radiometer
110	AVHRR Advanced Very High Resolution Radiometer
111	C20C Climate of the 20 th Century
112	CCSM Community Climate System Model
113	CDR Climate Data Record
114	CGCM Coupled Atmosphere-Ocean General Circulation Model
115	CH ₄ Methane
116	CLIVAR Climate Variability and Prediction
117	CMIP2 Coupled Model Intercomparison Project (version 2)
118	CMIS Conical scanning Microwave Imager/Sounder
119	CO ₂ Carbon Dioxide
120	COADS Comprehensive Ocean-Atmosphere Data Set
121	COWL Cold Ocean Warm Land
122	CrIS Cross-track Infrared Sounder
123	CRN Climate Reference Network
124	EBM Energy Balance Model
125	ECMWF European Center for Medium-Range Weather Forecasts
126	ENSO El Niño-Southern Oscillation
127	EOF Empirical Orthogonal Function
128	ERA ECMWF Re-Analysis
129	ERSST Extended Reconstruction Sea Surface Temperature
130	GCM General Circulation Model
131	GCOS Global Climate Observing System
132	GCSM Global Climate System Model (includes atmosphere, land, oceans, 133 glaciers)
134	GEOSS Global Earth Observation System of Systems
135	GFDL Geophysical Fluid Dynamics Laboratory
136	GHCN Global Historical Climatology Network
137	GHGs Greenhouse Gases

138	GHRSSST-PP	GODAE High-Resolution Sea Surface Temperature Pilot Project
139	GISS	Goddard Institute for Space Studies
140	GODAE	Global Ocean Data Assimilation Experiment
141	GPS	Global Positioning System
142	GSN	GCOS Surface Network
143	GUANGCOS	Upper Air Network
144	HadCM	Hadley Centre Climate Model
145	HadRTHadley	Centre Radiosonde Temperatures
146	hPa	hectoPascals, a measure of pressure
147	HIRS	High-resolution Infrared Radiation Sounder
148	IASI	Infrared Atmospheric Sounding Interferometer
149	ICOADS	International Comprehensive Ocean-Atmosphere Data Set
150	IGRA	Integrated Global Radiosonde Archive
151	IPCC	Intergovernmental Panel on Climate Change
152	IR	Infrared Radiation
153	ITCZ	Inter Tropical Convergence Zone
154	LECT	Local Equator Crossing Time
155	LKS	Lanzante, Klein, Seidel
156	LOSU	Level of Scientific Understanding
157	LTER	Long Term Ecological Research
158	LR	Linear Regression
159	MAT	Marine Air Temperatures
160	MPS	Median of Pair-wise Slopes
161	MSU	Microwave Sounding Unit
162	NAM	Northern Hemisphere Annual Mode
163	NAO	North Atlantic Oscillation
164	NASA	National Aeronautics and Space Administration
165	NCAR	National Center for Atmospheric Research
166	NCDC	National Climatic Data Center
167	NCEP	National Centers for Environmental Prediction
168	NEMS	Nimbus E Microwave Spectrometer
169	NESDIS	National Environmental Satellite, Data, and Information Service
170	NH	Northern Hemisphere
171	NMATNight	Marine Air Temperatures
172	N ₂ O	Nitrous Oxide
173	NOAA	National Oceanic and Atmospheric Administration
174	NPOESS	National Polar-orbiting Operational Environmental Satellite System
175	NRC	National Research Council
176	O ₃	Ozone
177	PCM	Parallel Climate Model
178	PDO	Pacific Decadal Oscillation
179	QBO	Quasi-Biennial Oscillation
180	RATPAC	Radiosonde Atmospheric Temperature Products for Assessing Climate
181	RSS	Remote Sensing Systems
182	SAM	Southern Hemisphere Annual Mode
183	SCAMS	SCAnning Microwave Spectrometer

184	SH	Southern Hemisphere
185	SMMR	Scanning Multi-spectral Microwave Radiometer
186	SO ₄	Sulfate
187	SRN	Surface Reference Network
188	SSM/I	Special Sensor Microwave Imager
189	SST	Sea Surface Temperature
190	SSU	Stratospheric Sounding Unit
191	TAO	Tropical Atmosphere Ocean
192	TIROS	Television InfraRed Observation Satellite
193	TLT	Temperature of the Lower Troposphere
194	TOGA	Tropical Ocean Global Atmosphere
195	TOVS	TIROS Operational Vertical Sounder
196	TRMM	Tropical Rainfall Measuring Mission
197	UAH	University of Alabama in Huntsville
198	UNFCCC	United Nations Framework Convention on Climate Change
199	USHCN	United States Historical Climatology Network
200	UTC	Coordinated Universal Time
201	VOS	Voluntary Observing Ships
202	VTPR	Vertical Temperature Profile Radiometer
203	WMO	World Meteorological Organization
204		