Radiosonde Daytime Biases and Late-20th Century Warming

Steven C. Sherwood, *et al.*

Science **309**, 1556 (2005);
DOI: 10.1126/science.1115640

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

realizations reduces noise and facilitates signal estimation.
25. Materials and methods are available as supporting material on *Science* Online.
26. F. J. Wentz, M. Schabel, *Nature* **403**, 414 (2000).
27. T. M. Smith, R. W. Reynolds, *J. Clim.* **18**, 2021 (2005).
28. P. D. Jones, A. Moberg, *J. Clim.* **16**, 206 (2003).
29. HadCRUT2v is the designation for version 2 of the (variance-corrected) Hadley Centre/Climatic Research Unit surface temperature data set.
30. Q. Fu, C. M. Johanson, S. G. Warren, D. J. Seidel, *Nature* **429**, 55 (2004).
31. P. H. Stone, J. H. Carlson, *J. Atmos. Sci.* **36**, 415 (1979).
32. Here, we define $\overline{X}$ as the arithmetic mean of the ensemble means, i.e., $\overline{X} = \frac{1}{N}\sum_{j=1}^{N} \overline{X}_j$, where N is the total number of models in the IPCC archive and $\overline{X}_j$ is the ensemble mean signal of the jth model. This weighting avoids undue emphasis on results from a single model with a large number of realizations.
33. One measure of ENSO variability is $s(T_{\mathrm{NIÑO-3.4}})$, the standard deviation of sea-surface temperatures in the Niño 3.4 region of the equatorial Pacific. Values of $s(T_S)$ in the 49 IPCC realizations are closely correlated with $s(T_{\mathrm{NIÑO-3.4}})$ (correlation coefficient $r = 0.92$).
34. The theoretical expectation plotted in Fig. 3 was computed by taking the difference of two pseudo-adiabats calculated from surface air parcels with temperatures of 28.0° and 28.2°C and 80% relative humidity. These are conditions typical of deep convective regions over the tropical oceans. The pseudo-adiabats correspond to equivalent potential temperatures of 353.2 and 354.1 K. The assumed temperature difference of 0.2°C corresponds approximately to the total change in tropical ocean temperature over the years 1979 to 1999. Theoretical scaling ratios are relatively insensitive to reasonable variations in the baseline values of surface air temperature and relative humidity, as well as to the magnitude of the surface air temperature increase.
35. V. Ramaswamy *et al.*, in *Climate Change 2001: The Scientific Basis*, J. T. Houghton *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 2001), pp. 349–416.
36. B. D. Santer *et al.*, *Science* **301**, 479 (2003).
37. Work at Lawrence Livermore National Laboratory (LLNL) was performed under the auspices of the U.S. Department of Energy (DOE), Environmental Sciences Division, contract W-7405-ENG-48. A portion of this study was supported by the U.S. DOE, Office of Biological and Environmental Research, as part of its Climate Change Prediction Program. T.M.L.W. was supported by NOAA Office of Climate Programs (Climate Change Data and Detection) grant NA87GP0105. P.W.T. and G.J. were funded by the UK Department of the Environment, Food, and Rural Affairs. We acknowledge the international modeling groups for providing their data for analysis, the Joint Scientific Committee/Climate Variability and Predictability Working Group on Coupled Modeling and their Coupled Model Intercomparison Project and Climate Simulation Panel for organizing the model data analysis activity, and the IPCC WG1 TSU for technical support. The IPCC Data Archive at LLNL is supported by the Office of Science, U.S. DOE. The static MSU weighting functions and UAH MSU data were provided by J. Christy (UAH). We thank I. Held, T. Delworth (both Geophysical Fluid Dynamics Laboratory), D. Easterling (National Climatic Data Center), B. Hicks (NOAA Air Resources Laboratory), and two anonymous reviewers for useful comments. O. Boucher (Hadley Centre), G. Flato (Canadian Climate Centre), and E. Roeckner (Max-Planck Institute for Meteorology) supplied information on the historical forcings used by CNRM-CM3, CCCma-CGCM3.1(T47), and ECHAM5/MPI-OM.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1114867/DC1
Materials and Methods
Fig. S1
Table S1
References and Notes

16 May 2005; accepted 27 July 2005
Published online 11 August 2005;
10.1126/science.1114867
Include this information when citing this paper.

Radiosonde Daytime Biases and Late–20th Century Warming

Steven C. Sherwood,[1]* John R. Lanzante,[2] Cathryn L. Meyer[1]

The temperature difference between adjacent 0000 and 1200 UTC weather balloon (radiosonde) reports shows a pervasive tendency toward cooler daytime compared to nighttime observations since the 1970s, especially at tropical stations. Several characteristics of this trend indicate that it is an artifact of systematic reductions over time in the uncorrected error due to daytime solar heating of the instrument and should be absent from accurate climate records. Although other problems may exist, this effect alone is of sufficient magnitude to reconcile radiosonde tropospheric temperature trends and surface trends during the late 20th century.

Atmospheric models and simple thermodynamic arguments indicate that tropospheric and surface temperature changes should be closely linked (*1*). Radiosonde data during the late 20th century, however (*2*–*5*), have not shown warming commensurate with that reported for the surface (*1*, *6*, *7*). The main discrepancy is in the Tropics during the last two decades of the 20th century.

A number of design changes to radiosonde systems over the years may have affected trends (*8*). Indeed, the spread of trends among stations well exceeds that implied by satellite data (*9*), suggesting that trends in the observation bias typically exceed those of the actual temperature at individual stations.

Among the most serious known problems is bias due to solar heating of the temperature sensor (*10*). For many radiosonde designs this can elevate the temperature several °C above ambient during daylight, an effect that must be removed via an estimated correction. For other designs no correction is standard even though the effect may not be completely absent. Adjustment of climate records for instrument changes using their documented histories is problematic (*8*, *11*).

One can try to remove undocumented artifacts by careful examination of the data itself. Several such efforts have detected hundreds or thousands of apparent artifacts (*3*–*5*, *12*). Their net effect on trends was found to be large only in the stratosphere. Revised trends were still lower than those indicated by the Microwave Sounding Unit (MSU) in both the troposphere and stratosphere (*13*). Because empirical separation of artificial discontinuities from genuine variability is extremely challenging in correlated time series (*14*, *15*), especially as changes can probably occur in many small steps (*16*), it is not clear how successful the above efforts may have been in detecting discontinuities—or avoiding false adjustments—of amplitudes well below 1°C.

Here we adopt a strategy for quantifying trend errors that does not require identifying specific change events. The strategy applies only to the solar heating error and does not detect other errors. It relies on the fact that the diurnal temperature range in the free troposphere, hence its expected trend, is small and has known characteristics that differ from those expected from a radiation error.

The diurnal temperature variation in Earth's atmosphere is a tide arising from its direct solar heating and from diurnal variations of convective heating driven by the diurnal variation of surface temperature. Atmospheric heating, which occurs primarily in the stratosphere via ozone absorption, drives migrating resonant oscillations that cause temperature fluctuations of several °C in the upper stratosphere. In the troposphere, weaker solar heating occurs due mainly to near-infrared absorption by water with a contribution from dark aerosols. These influences produce diurnal temperature fluctuations of 1°C or less in the free troposphere (*17*). Near the land surface, variations of 5° to 15°C occur due to surface diurnal heating (*18*); over oceans, variations are ~1°C.

Because atmospheric tides are a linear phenomenon (*19*), the diurnal variation of temperature is proportional to that of the heating, though the two need not be in phase. Trends of ~ −0.2°C per decade are evident in the land surface diurnal temperature range (DTR) (*20*), which amount to about 2% of the mean DTR per decade. Tropospheric water vapor and stratospheric ozone changes do not exceed a few percent per decade in recent decades (*21*, *22*), and absorption increases

[1]Department of Geology and Geophysics, Yale University, New Haven, CT 06520, USA. [2]National Oceanic and Atmospheric Administration/Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, NJ 08542, USA.

*To whom correspondence should be addressed. E-mail: ssherwood@alum.mit.edu

weakly with concentration due to line saturation (*23*). It follows that tides could not have changed by more than a few percent, or ~ 0.01° to 0.02°C per decade. Because of this, trends in the observed day minus night difference in radiosonde temperatures should provide a sensitive detector of changes in the daytime observation bias.

We examine the diurnal range using the CARDS data set with no adjustments (*24*). We calculated a quantity ΔT equal to the temperature difference between adjacent 0000 UTC and 1200 UTC sonde flights, wherever such pairs were available. Pairs were used regardless of which time of day came first, but ΔT was always defined as temperature at 0000 minus temperature at 1200. At all CARDS stations with sufficient data, we fitted linear trends to ΔT for the same periods (1959 to 1997 and 1979 to 1997) documented by Lanzante *et al*. (LKS) (*3*). LKS considered temperature trends at 87 stations denoted here as the "LKS subset."

Figure 1 shows the 1979 to 1997 trend in stratospheric ΔT at tropical stations plotted by longitude, together with a sinusoid representing the local time of day at 0000. These data show that the trend is in phase with solar heating, with daytime readings growing cooler compared to nighttime readings, and is pervasive.

Although clearest in the stratosphere, these characteristics appear also at tropospheric levels. Indeed, tropospheric and stratospheric ΔT trends are highly correlated in general: for example, $r = 0.85$ between 50 and 300 hPa over 1959 to 1997. This is not true for natural temperature variability, which tends to be anticorrelated below and above the tropopause in both low and high latitudes (*25*), nor is it true of the tide itself. According to wind data, tidal fluctuations in the troposphere should lag those at 50 hPa by about 6 hours (*26*); this is also simulated by the National Center for Atmospheric Research (NCAR) Community Atmosphere Model (CAM3) (not shown) and appears (albeit with slightly less shift) in carefully selected radiosonde temperature data (*17*). Consequently, we expect peak ΔT magnitudes near 90°E and 90°W. However, before the 1980s ΔT peaked broadly around 0° and 180°, where solar heating was greatest. Only by the late 1990s did the pattern in the troposphere begin to appear as expected.

To quantify the anomalous signal, we defined an additional quantity $\Delta T'$, equal to $\pm\Delta T$ with sign determined by longitude to make it daytime (6 am to 6 p.m.) minus nighttime. To minimize sunrise-time ambiguities, we did not compute $\Delta T'$ at stations within 10° of the 90°E/W meridians.

A map of the trend in upper tropospheric $\Delta T'$ (Fig. 2) reveals regional variations. The largest trends occurred in the Tropics, particularly among Indian, African, and island stations where transitional problems have been reported previously (*3*, *4*, *27*). Trends were small in North America and most of Asia. We see no evidence in Fig. 2 that the ΔT trends at stations in the LKS subset differed systematically from those at neighboring, non-LKS stations. However, the most affected stations tend to be in sparsely sampled areas where they would be strongly weighted in any spatially representative climatology. We omitted all Indian stations from subsequent analysis, because these show anomalously large ΔT and have other problems (*3*, *4*).

Following LKS, we averaged $\Delta T'$ over three belts: the Tropics, the Northern Hemisphere extratropics (NH), and the Southern Hemisphere extratropics (SH). Because tropospheric temperature is expected to lag insolation by about 6 hours, the zonal means $\langle \Delta T' \rangle$

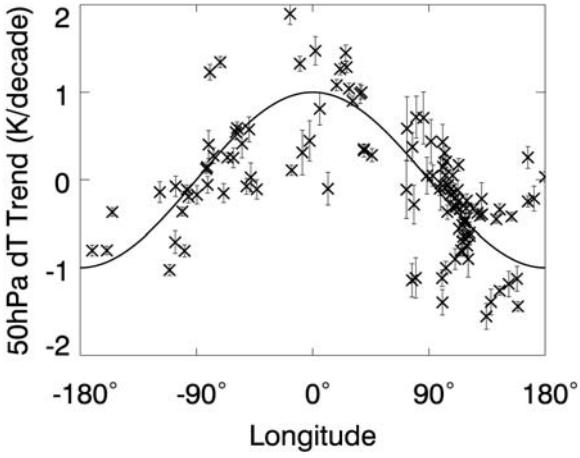

Fig. 1. Trend in 50-hPa ΔT (0000 UTC T minus 1200 UTC T) during 1979 to 1997 versus longitude at all Tropical stations. Sine wave (not a curve fit) represents the negative of solar forcing of ΔT, peaking where 0000 UTC falls at midnight and troughing where it falls at noon. Error bars are 1σ sampling uncertainties.

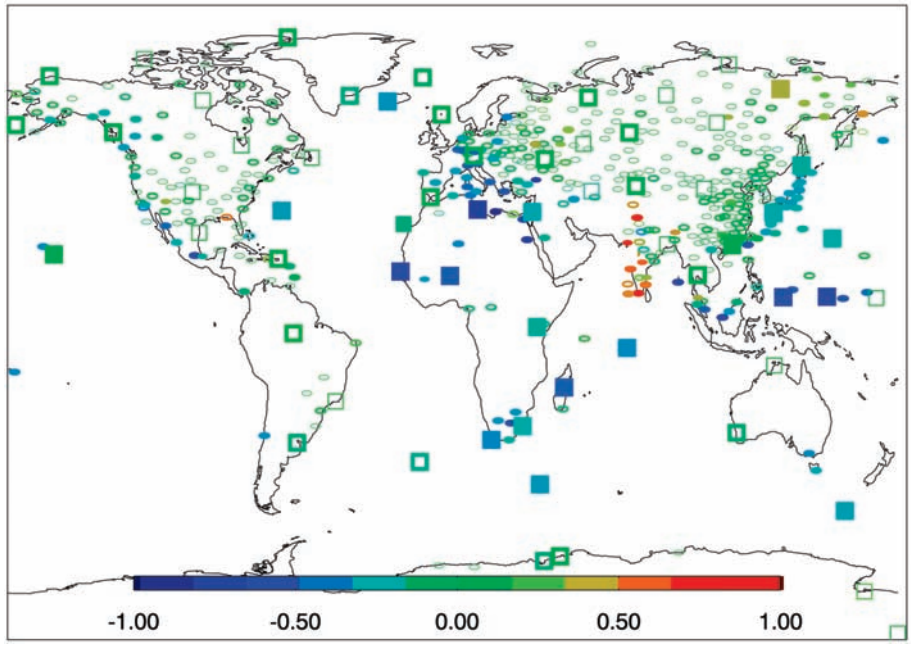

Fig. 2. Trends in 300-hPa day-night difference $\Delta T'$ during 1979 to 1997, in K per decade. LKS station subset is indicated by large squares. One station (Mumbai) is off scale (not shown). Solid symbols are significant at 95% confidence level; thick open symbols do not pass the test at 300 hPa but are significant in the stratosphere (50 hPa).

Table 1. Mean difference in ΔT trend from 1979 to 1997, vertically weighted according to the MSU channel 2 profile, sonde minus MSU (first two columns) among LKS stations; the differences in this quantity between the two station types (third column); and prediction of the latter based on assumptions in text (last column). All quantities are in °C per decade. Figures in parentheses are the number of stations used (*29*).

Daytime only	Twice daily	Difference	Predicted difference
	Tropics		
−0.228 (17)	−0.102 (8)	0.130	0.120 (18)
	Extratropics		
−0.052 (4)	−0.029 (43)	0.023	0.050 (38)

should be small due to near-cancellation of different longitudes.

The time series of tropical upper tropospheric $\langle \Delta T' \rangle$ (Fig. 3), however, shows significant long-term variations. Daytime temperatures warmed before about 1971, reaching values near 0.5°C above nighttime temperatures, then began a slow cooling trend. By the mid- to late 1990s, $\langle \Delta T' \rangle$ finally dropped to a level commensurate with predictions. The trend was particularly strong during the satellite era beginning in 1979. Since 1997 the trend has leveled off.

The linear trend in $\langle \Delta T' \rangle$ is shown by altitude in Fig. 4 for the two LKS time periods, for all three belts. It increases rapidly in the stratosphere, is weak in NH but strong in the other two belts, and is much stronger during the 1979 to 1997 period than the longer period starting in 1959.

This trend appears unrealistic in several respects. First, it is almost two orders of magnitude larger than can be justified physically based on the known forcings. (A run of the CAM3 general circulation model with half-normal ozone, an unrealistically large change, caused tropospheric ΔT to change by only 13%.) Indeed, if a 0.5°C change in diurnal temperature range were caused by a change in daytime heating from any source, then the radiative relaxation time scale of ~1 month for deep perturbations (*28*) would imply a change in equilibrium temperature of 10° to 20°C. Clearly, nothing like this has happened. Moreover, the spatial patterns of this trend are inconsistent with absorbing aerosol (which decreases with height and is scanty in SH) or convective heating (absent in the stratosphere) as a cause. Finally, the strong correlation of the ΔT trend between the troposphere and stratosphere is unnatural.

We are left to propose that the trends are caused by decreases over time in the uncorrected heating of the sensor. This is plausible a priori given the history of radiosonde development and improvement efforts and is fully consistent with all characteristics of the trend here documented: strong in the stratosphere (due mainly to the low thermal diffusivity of thin air) and in phase with solar heating. The smaller effect in NH is consistent with the expected superior stability of those stations.

The trend reported from a particular set of stations can be adjusted to a nighttime-only value by adding an adjustment δ_{sol} equal to the trend in $\langle \Delta T' \rangle$ multiplied by a factor f representing the fraction of the reported trend coming from daytime data (*29*). This assumes that stations that do not collect nighttime data are just as susceptible to spurious daytime trends, on average, as those that do.

MSU Channel 2 data can be used to test this assumption. We require only trend differences between sites, which are much more robust to analysis method than the overall MSU trend itself. We use diurnal-mean MSU trends from the University of Alabama at Huntsville at LKS station locations (*3*). Our assumption implies that daytime-only stations will cool more compared to colocated MSU retrievals than will twice-daily stations. The calculated differences, given in Table 1 (we combine SH and NH here because there are no daytime-

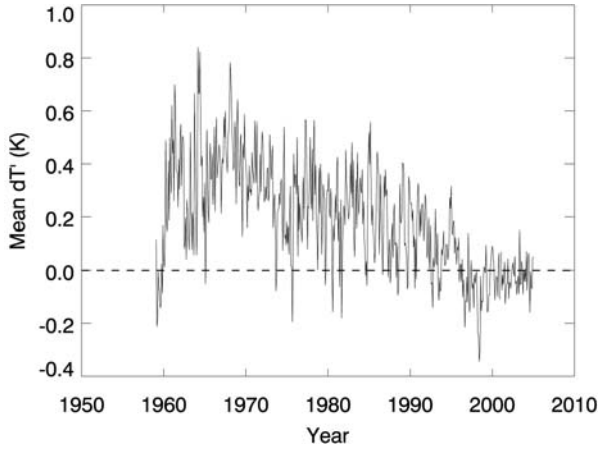

Fig. 3. Monthly mean 300-hPa $\langle \Delta T' \rangle$, the average day-night temperature difference, at the 10 LKS tropical stations spanning the 1959 to 1997 period.

Table 2. Layer-average tropospheric and stratospheric temperature trends (in K per decade) reported by LKS for unhomogenized data ("orig"), and with solar heating bias removed ("new"). Factor f is specific to LKS station subset. Uncertainties are 1σ sampling uncertainty in the solar heating bias correction only.

	Tropics	NH extratropics	SH extratropics
	1979–1997		
f	0.84	0.50	0.67
50–100 hPa (orig)	−1.30	−0.85	−1.04
50–100 hPa (new)	−0.81 ± 0.08	−0.78 ± 0.03	−0.67 ± 0.08
850–300 hPa (orig)	−0.02	+0.10	−0.07
850–300 hPa (new)	+0.14 ± 0.04	+0.14 ± 0.02	+0.01 ± 0.04
	1959–1997		
50–100 hPa (orig)	−0.71	−0.43	−0.50
50–100 hPa (new)	−0.52 ± 0.06	−0.38 ± 0.02	−0.30 ± 0.07
850–300 hPa (orig)	+0.17	+0.06	+0.25
850–300 hPa (new)	+0.23 ± 0.03	+0.07 ± 0.01	+0.30 ± 0.03

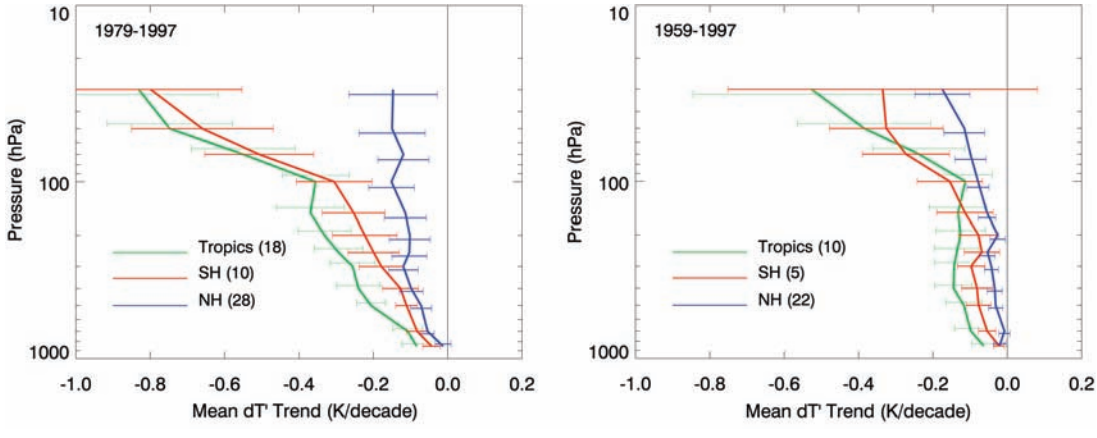

Fig. 4. Trend in $\langle \Delta T' \rangle$ during 1979 to 1997 (top) and 1959 to 1997 (bottom) at LKS stations. Green, Tropics (30°N to 30°S); red, Southern Hemisphere (90°S to 30°S); blue, Northern Hemisphere (30°N to 90°N). Error bars are 1σ sampling uncertainty. Figures in parentheses are the number of stations used.

only LKS stations in NH), are fully consistent with this assumption, particularly for the tropical stations. In the extratropics there are only four daytime-only stations so the MSU test is less meaningful, but the two independent estimates do agree within 0.03°C per decade.

To illustrate the importance of the heating bias, we have computed its impact δ_{sol} on the trends at LKS stations. The LKS f factors, unhomogenized trends, and trends adjusted only for solar heating are given for the middle troposphere and lower stratosphere in Table 2. In the stratosphere, our δ_{sol} is similar to the total adjustments by LKS and others, with trends moving closer to those from MSU (*13*). At the tropical tropopause (of relevance to stratospheric water vapor), δ_{sol} is somewhat smaller than LKS's. In the troposphere, however, δ_{sol} is much larger than previous adjustments. Indeed, the tropical trend with this adjustment (0.14°C per decade over 1979 to 1997) would be consistent with model simulations driven by observed surface warming, which was not true previously (*1*). One independent indication that the solar-adjusted trends should be more accurate is their consistency across latitude belts: for the period 1979 to 1997, the spread of values fell by 70% in the lower stratosphere and 25% in the troposphere.

Though this is encouraging, our confidence in these nighttime trends is still limited given that other radiosonde errors have not been addressed. SH trends from 1958 to 1997 seem unrealistically high in the troposphere, especially with the δ_{sol} adjustment, although this belt has by far the worst sampling. Previous homogenization efforts typically produced small changes to mean tropospheric trends, which could mean that other error trends cancel out δ_{sol} in the troposphere. In our judgment, however, such fortuitous cancellation of independent errors is unlikely compared to the possibility that most solar artifacts were previously either missed or their removal negated by other, inaccurate adjustments. To be detected easily, a shift must be large and abrupt, but δ_{sol} was spread out over so many stations (79% of stations during 1979 to 1997 and 90% during 1959 to 1997 experienced ΔT trends significant at 95% level), at such modest levels, and of sufficient frequency at many stations that many may have been undetectable. Most important, jumps in the difference between daytime and nighttime monthly means would be detectable at only a few tropical stations because most lack sufficient nighttime data. In any case, we conclude that carefully extracted diurnal temperature variations can be a valuable troubleshooting diagnostic for climate records, and that the uncertainty in late–20th century radiosonde trends is large enough to accommodate the reported surface warming.

References and Notes

1. B. D. Santer *et al.*, *Science* **309**, 1551 (2005); published online 11 August 2005 (10.1126/science.1114867).
2. J. K. Angell, *J. Clim.* **16**, 2288 (2003).
3. J. R. Lanzante, S. A. Klein, D. J. Seidel, *J. Clim.* **16**, 241 (2003).
4. D. E. Parker *et al.*, *Geophys. Res. Lett.* **24**, 1499 (1997).
5. P. W. Thorne *et al.*, *J. Geophys. Res.*, in press.
6. D. H. Douglass, B. D. Pearson, S. F. Singer, P. C. Knappenberger, P. J. Michaels, *Geophys. Res. Lett.* **31**, L13207 (2004).
7. D. J. Gaffen *et al.*, *Science* **287**, 1242 (2000).
8. D. E. Parker, D. I. Cox, *Int. J. Climatol.* **15**, 473 (1995).
9. M. Free, D. J. Seidel, *J. Geophys. Res.* **110**, D07101 (2005).
10. J. K. Luers, R. E. Eskridge, *J. Appl. Meteorol.* **34**, 1241 (1995).
11. I. Durre, T. C. Peterson, R. S. Vose, *J. Clim.* **15**, 1335 (2002).
12. L. Haimberger, "Homogenization of radiosonde temperature time series using ERA-40 analysis feedback information," Tech. Rep. European Center for Medium Range Weather Forecasting (2005), ERA-40 Project Report Series 23.
13. D. J. Seidel *et al.*, *J. Clim.* **17**, 2225 (2004).
14. P. R. Krishnaiah, B. Q. Miao, *Handbook of Statistics*, P. R. Krishnaiah, C. R. Rao, Eds. (Elsevier, New York, 1988), vol. 7.
15. M. Free *et al.*, *Bull. Am. Meteorol. Soc.* **83**, 891 (2002).
16. W. J. Randel, F. Wu, in preparation.
17. D. J. Seidel, M. Free, J. Wang, *J. Geophys. Res.* **110**, D09102 (2005).
18. A. Dai, K. E. Trenberth, T. R. Karl, *J. Clim.* **12**, 2451 (1999).
19. S. Chapman, R. S. Lindzen, *Atmospheric Tides* (D. Reidel, Norwell, MA, 1970).
20. D. R. Easterling *et al.*, *Science* **277**, 364 (1997).
21. D. J. Gaffen, R. J. Ross, *J. Clim.* **12**, 811 (1999).
22. W. J. Randel *et al.*, *Science* **285**, 1689 (1999).
23. K. N. Liou, T. Sasamori, *J. Atmos. Sci.* **32**, 2166 (1975).
24. R. E. Eskridge *et al.*, *Bull. Am. Meteorol. Soc.* **76**, 1759 (1995).
25. H. Riehl, *Tropical Meteorology* (McGraw Hill, New York, 1954).
26. S. C. Sherwood, *Geophys. Res. Lett.* **27**, 3525 (2000).
27. J. R. Christy, R. W. Spencer, W. B. Norris, W. D. Braswell, D. E. Parker, *J. Atmos. Oceanic Technol.* **20**, 613 (2003).
28. T. Sasamori, J. London, *J. Atmos. Sci.* **23**, 543 (1966).
29. Data files and further information on methods, uncertainty, and interpretation of our results are available as supporting material on *Science* Online.

30. S.C.S. thanks J. Risbey and K. Braganza for useful discussions. This work was supported by the National Oceanic and Atmospheric Administration Climate and Global Change Program award NA03OAR4310153, and by NSF ATM-0134893.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1115640/DC1
Methods
SOM Text
Data files
References and Notes

2 June 2005; accepted 27 July 2005
Published online 11 August 2005;
10.1126/science.1115640
Include this information when citing this paper.

The Transcriptional Landscape of the Mammalian Genome

The FANTOM Consortium* and RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group)*

This study describes comprehensive polling of transcription start and termination sites and analysis of previously unidentified full-length complementary DNAs derived from the mouse genome. We identify the 5′ and 3′ boundaries of 181,047 transcripts with extensive variation in transcripts arising from alternative promoter usage, splicing, and polyadenylation. There are 16,247 new mouse protein-coding transcripts, including 5154 encoding previously unidentified proteins. Genomic mapping of the transcriptome reveals transcriptional forests, with overlapping transcription on both strands, separated by deserts in which few transcripts are observed. The data provide a comprehensive platform for the comparative analysis of mammalian transcriptional regulation in differentiation and development.

The production of RNA from genomic DNA is directed by sequences that determine the start and end of transcripts and splicing into mature RNAs. We refer to the pattern of transcription control signals, and the transcripts they generate, as the transcriptional landscape. To describe the transcriptional landscape of the mammalian genome, we combined full-length cDNA isolation (*1*) and 5′- and 3′-end sequencing of cloned cDNAs, with new cap-analysis gene expression (CAGE) and gene identification signature (GIS) and gene signature cloning (GSC) ditag technologies for the identification of RNA and mRNA sequences corresponding to transcription initiation and termination sites (*2*, *3*). A detailed description of the data sets generated, mapping strategies, and depth of coverage of the mouse transcriptome is provided in supporting online material (SOM) text 1 (Tables 1 and 2). We have identified paired initiation and termination sites, the boundaries of independent transcripts, for 181,047 independent transcripts in the transcriptome (Table 3). In total, we found 1.32 5′ start sites for each 3′ end and 1.83 3′ ends for each 5′ end (table S1). Based on these data, the number of transcripts is at least one order of magnitude larger than the estimated 22,000 "genes" in the mouse genome (*4*) (SOM text 1), and the