Interpreting Differential Temperature Trends at the Surface and in the Lower Troposphere

B. D. Santer,[1]* T. M. L. Wigley,[2] D. J. Gaffen,[3] L. Bengtsson,[4] C. Doutriaux,[1] J. S. Boyle,[1] M. Esch,[4] J. J. Hnilo,[1] P. D. Jones,[5] G. A. Meehl,[2] E. Roeckner,[4] K. E. Taylor,[1] M. F. Wehner[1]

Estimated global-scale temperature trends at Earth's surface (as recorded by thermometers) and in the lower troposphere (as monitored by satellites) diverge by up to 0.14°C per decade over the period 1979 to 1998. Accounting for differences in the spatial coverage of satellite and surface measurements reduces this differential, but still leaves a statistically significant residual of roughly 0.1°C per decade. Natural internal climate variability alone, as simulated in three state-of-the-art coupled atmosphere-ocean models, cannot completely explain this residual trend difference. A model forced by a combination of anthropogenic factors and volcanic aerosols yields surface-troposphere temperature trend differences closest to those observed.

Over the past century, atmospheric temperatures within a few meters of Earth's surface have been measured at thousands of locations worldwide (*1*). The record of temperature changes in the free atmosphere is considerably shorter. Since the late 1940s, radiosondes have made direct measurements of tropospheric and lower stratospheric temperatures at several hundred stations (*2*). The radiosonde and surface thermometer networks have incomplete coverage in space and time, which introduces errors in estimates of global-scale temperature changes (*3, 4*).

Since 1979, polar-orbiting satellites have monitored atmospheric temperatures on a global scale. Satellite temperature measurements are mass-weighted averages of the microwave emissions from deep atmospheric layers (*5*). They are not the same physical quantity as the near-surface temperatures monitored by thermometers (*6*).

There is an apparent difference between the thermometer-estimated surface warming of roughly 0.20°C per decade since 1979 and the much smaller temperature trend in the lower troposphere estimated from satellites and radiosondes (*7*) (Fig. 1). This difference has at least three possible interpretations. First, it may be an artifact, primarily related to residual data quality problems in either the surface (*8, 9*) and/or radiosonde (*2*) and satellite data (*7, 10*). Second, the difference may be real and due to the effects of natural internal variability (*11*) and/or external forcing (*11–13*). A third possibility is that some portion of the observed discrepancy is related to coverage differences between the satellite and surface temperature data. These three interpretations are not mutually exclusive.

Our focus here is on the second and third interpretations. We show that the observed difference between surface and tropospheric temperature changes cannot be fully explained by coverage differences between satellite- and surface-based measurement systems and/or the effects of natural internal climate variability. However, we find that both effects may make substantial contributions to the observed trend difference. We also discuss a recent model result that suggests that the observed warming of the surface relative to the lower troposphere may be a response to combined forcing by well-mixed greenhouse gases, sulfate aerosols, stratospheric ozone, and the effects of the Pinatubo eruption in June 1991 (*12*).

Observational data. Our analysis uses tropospheric temperature data from the satellite-based Microwave Sounding Unit (MSU), which measures the upwelling microwave radiation from oxygen molecules (*5, 8*). We rely on the most recent version (MSUd03, henceforth MSUd) of the MSU lower tropospheric temperature retrieval (2LT), which includes corrections for orbital decay effects and interannual variations in instrument-body temperature, as well as additional adjustments for east-west satellite drift and improved calibration coefficients for the MSU instrument on the NOAA-12 satellite (*14*).

The surface data used here are a combination of marine sea surface temperatures (SSTs) and land surface air temperatures (SATs). Numerous publications have documented the procedures involved in the quality control, correction, and gridding of the raw station data, in the merging of SST and SAT data sets, and in the estimation of random and systematic errors (*1, 3*). This data set has been used extensively by the Intergovernmental Panel on Climate Change (*15*), and we refer to it henceforth as IPCC.

Effect of coverage differences. The MSUd 2LT data have global coverage. In contrast, coverage of the IPCC surface data varies in time and space, with large gaps in the interior of Africa and poleward of 50°S and 70°N (Fig. 2). We performed three tests to investigate the effect of MSU-IPCC coverage differences on the estimated surface minus 2LT (surface – 2LT) trend differential

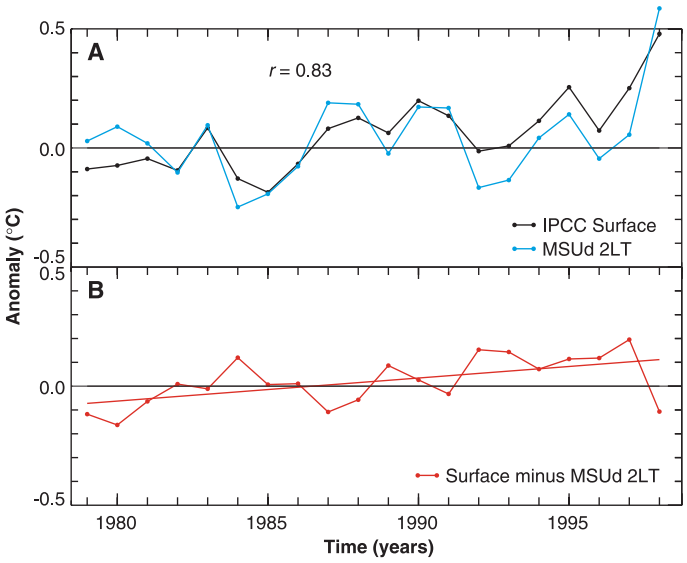

Fig. 1. (**A**) Anomaly time series for observed surface and lower tropospheric annual temperature data from the IPCC (*15*) and MSUd 2LT (*14*) data sets, respectively. (**B**) The observed surface – 2LT difference time series, together with its least-squares linear trend. Observed data were sampled with the FIXMASK method (see text) and are expressed as spatially averaged annual mean anomalies relative to 1979–1993 climatological annual means.

[1]Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. [2]National Center for Atmospheric Research, Boulder, CO 80307, USA. [3]NOAA Air Resources Laboratory, Silver Spring, MD 20910, USA. [4]Max Planck Institute for Meteorology, D-20146 Hamburg, Germany. [5]Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK.

*To whom correspondence should be addressed. E-mail: santer1@llnl.gov

between the two data sets. The first test, NOMASK, preserves the actual coverage differences between the MSUd 2LT and IPCC surface data. The second, VARMASK, imposes the space- and time-varying mask of the IPCC annual mean coverage on the MSU data. The third test, FIXMASK, identifies the subset of IPCC grid points that always have data over the period 1979–1998, and then uses this fixed mask to subsample the MSU data (*16*).

The aim of the subsampling exercise is not to obtain the best estimates of global-scale surface and lower tropospheric temperatures and temperature differences. If this were our objective, we would attempt to fill the observed surface coverage gaps (Fig. 2) by statistical techniques or by "blending" the observed in situ surface data with either satellite-estimated surface temperatures or results from numerical weather prediction models (*17, 18*). There are uncertainties inherent in such infilling approaches (*19*). Instead, we compare the surface and lower tropospheric data over areas of common data availability, and ask whether subsampling improves or degrades the statistical correspondence between the IPCC and MSUd 2LT data (*20*).

We find that this correspondence is improved. Accounting for MSU-IPCC coverage differences increases the correlation and reduces the root mean square (RMS) difference and the trend difference between the surface and lower troposphere data (Table 1). In NOMASK, the surface and 2LT least-squares linear trends over 1979–1998 are 0.196° and 0.057°C per decade, respectively, yielding a differential of 0.139°C per decade. The corresponding differentials in VARMASK and FIXMASK are reduced by 25 to 30%, to 0.106° and 0.097°C per decade, respectively. All three surface – 2LT trend differences are statistically significant (*21*), despite the large, overlapping 95% confidence intervals estimated for the individual IPCC and MSUd 2LT trends (Table 1) (*22*). The individual 20-year trends in surface temperature are significantly different from zero, whereas MSUd 2LT trends are not.

Effects of natural internal variability. After accounting for surface − 2LT coverage differences, we explored whether the residual trend difference of roughly 0.1°C per decade could be explained by natural variability of the climate system on decadal time scales. It is difficult to address this question with observational data, given the limited duration of satellite and radiosonde records and the difficulties involved in separating any putative anthropogenic signal from natural variability. In a model-based study, Hansen *et al.* (*11*) used a 3000-year control integration (with no changes in external forcings) of a global climate model to generate a frequency distribution of surface – 2LT trend differences. Twenty-year "unforced" trend differences never exceeded 0.1°C per decade, and differences larger than 0.05°C per decade were infrequent. We revisit this study below, using results from more recent coupled model simulations.

We examined data from 300-year control integrations performed with three models: the ECHAM4/OPYC model (henceforth ECHAM) (*23*) of the Max Planck Institute for Meteorology in Hamburg (*12, 24*), the Parallel Climate Model (PCM) (*25*) of the National Center for Atmospheric Research (NCAR) and Los Alamos National Laboratory (LANL), and the NCAR Climate System Model (CSM) (*26*). PCM and CSM use similar versions of the same atmospheric model (*27*) but are coupled to different ocean general circulation models. All three atmospheric models have comparable resolution (roughly 2.8° latitude × 2.8° longitude and 18 or 19 atmospheric levels). ECHAM uses annual mean adjustments of heat and freshwater fluxes, whereas PCM and CSM have no flux adjustments.

The model global mean time series of unforced fluctuations in surface, 2LT, and surface – 2LT temperatures (Fig. 3) reveal

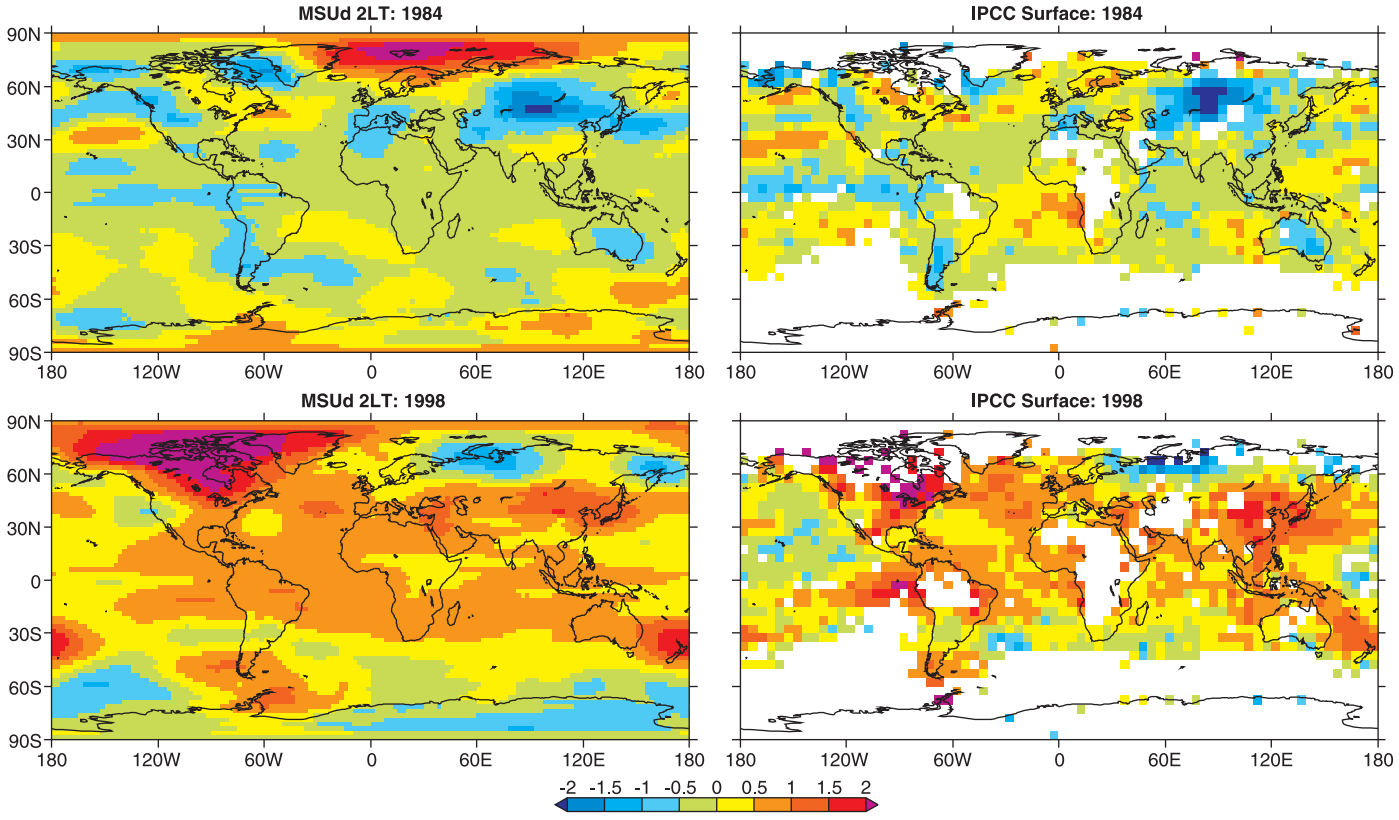

Fig. 2. Spatial coverage of lower tropospheric (*5, 14*) and surface (*1, 15*) temperatures in MSUd 2LT and IPCC data sets (white areas indicate no data). The MSU data have global, time-invariant coverage, whereas the IPCC data are spatially incomplete. These differences are accounted for in the VARMASK and FIXMASK sampling methods. The annual mean temperature anomalies shown are for the warmest (1998) and coolest (1984) years in the MSUd 2LT record, each with different IPCC coverage, and are expressed (°C) relative to climatological annual means over 1979–1993.

marked differences (*28*). Variability on time scales exceeding 50 years is noticeably smaller in PCM than in either CSM or ECHAM. The low-frequency variability differences between PCM and CSM (despite highly similar atmospheric models) may be related to systematic differences in thermocline structure and thermal inertia of the mixed layer, with a generally shallower thermocline in PCM. On time scales of one to two decades, ECHAM and PCM exhibit greater variability than CSM. Note also that ECHAM shows larger anomalies aloft than at the surface; such amplification is less apparent in PCM and CSM.

In computing sampling distributions of surface – 2LT 20-year trend differences, the model control run data were treated in an analogous way to the observations and were sampled using the NOMASK, VARMASK, and FIXMASK methods (*29*). We show results for the FIXMASK method only (Fig. 4), but our conclusions are generally insensitive to this choice (*30*). The three models yield sampling distributions of different shapes, largely reflecting the variability differences discussed above. In all three cases, however, even after accounting for coverage differences, the effects of model-simulated natural internal variability cannot explain the observed surface – 2LT trend difference of roughly +0.1°C per decade. The largest positive 20-year trend differential ever obtained is +0.08°C per decade in the ECHAM control run. The empirical probability of obtaining the observed surface trend and surface – 2LT trend difference is zero in the three model control runs, independent of the masking method (*29*).

Effects of external forcing. We next examined results from a set of three perturbation experiments recently performed with ECHAM (*12, 24*). The baseline experiment that we consider is GSDIO (*31*), which uses estimated observed changes in greenhouse gases, sulfate aerosols (direct scattering effects and indirect effects on clouds), and tropospheric ozone over the period 1860–1990. Changes in these atmospheric constituents over the period 1991–2050 follow IPCC scenario IS92a (*32*). The other two experiments, GSO and GSOP, cover the period 1979–1997 only. GSO uses all GSDIO forcings but also includes the effects of stratospheric ozone depletion (*33*). Two GSO realizations were available (GSO1 and GSO2), each starting from slightly different atmospheric initial conditions in 1979. GSOP is identical with GSO1 until June 1991, after which it additionally includes the effects of volcanic aerosols from the Pinatubo eruption (*34*).

In each experiment, we used our FIXMASK sampling method and then calculated surface, 2LT, and surface – 2LT trends and their associated 95% confidence intervals over model years 1979–1998 (GSDIO) and 1979–1997 (GSO1, GSO2, and GSOP). For surface trends, the observed confidence intervals overlap with those in all four perturbation experiments (Fig. 5A). Modeled and observed confidence intervals also overlap for all 2LT trends, although the overlap is minimal for GSDIO (Fig. 5B). In contrast, there is no overlap between modeled and observed results for surface – 2LT trend differences (Fig. 5C).

To facilitate interpretation of these results, we performed a simple test of the significance of modeled and observed trend differences (*35*). This test confirmed that modeled and observed surface trends are not significantly different (*36*). For 2LT trends, only GSDIO is significantly different from observations. The 2LT trend in GSOP (the exper-

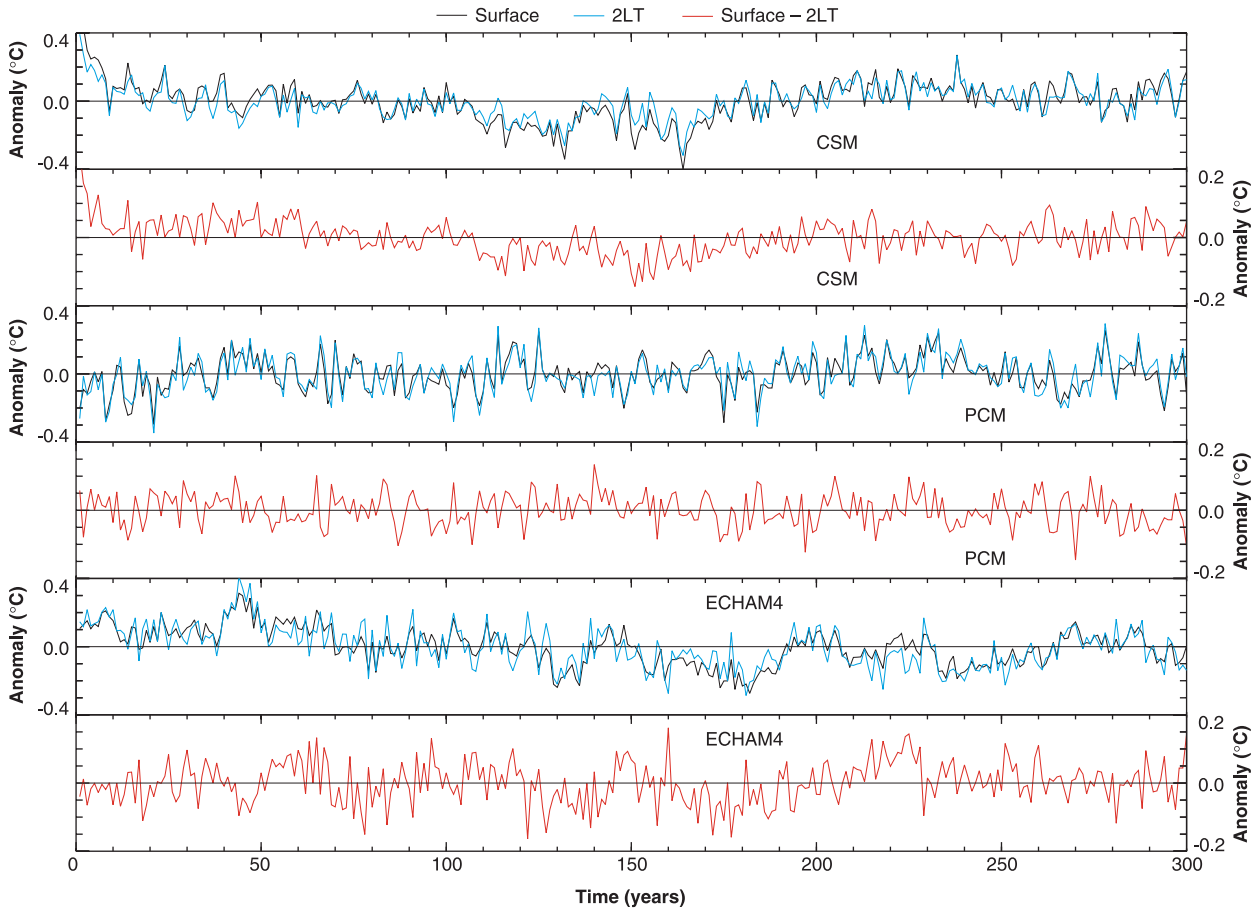

Fig. 3. Time series of global mean–annual mean anomalies in surface and lower tropospheric temperatures in three coupled model control runs (*24–26*). Surface and equivalent 2LT anomalies (*28*) are expressed relative to climatological annual averages over the entire control integration. Also shown are the time series of surface – 2LT differences for each model. No masking was applied to the model data.

iment with the most realistic combination of anthropogenic and natural forcings) is very close to the observed value, thus refuting recent claims that modeled and observed tropospheric trends are fundamentally inconsistent (*37*).

For surface – 2LT results, however, the trend difference in each model experiment is significantly different from the observed result (*36*). All model surface – 2LT trend differences are negative, unlike the observations (Fig. 5C). In the model results, greenhouse gas influences alone tend to yield larger temperature increases aloft than at the surface (*13*). The resulting large negative surface – 2LT trend difference is noticeably (but not completely) offset in the most realistic model simulation, GSOP, which includes both greenhouse gases and the effects of Pinatubo and stratospheric ozone depletion. This suggests that in GSOP, both Pinatubo and stratospheric ozone depletion have larger cooling signatures in the troposphere than at the surface (*11–13, 38*). The net result is a very small difference in GSOP [–0.035°C per decade (*39*)] between the surface and 2LT trends over the period 1979–1997 (*40*).

We can use these results to quantify, for this specific period, the relative contributions of Pinatubo and stratospheric ozone depletion to the simulated surface – 2LT temperature differences. We first estimate the effects of stratospheric ozone depletion by subtracting GSDIO trends from GSO1 and GSO2. In each of the GSO realizations, ozone depletion cools both the surface and the lower troposphere. The ozone-induced cooling is larger in the lower troposphere than at the surface—by 0.052° ± 0.031°C per decade in GSO1 and 0.026° ± 0.027°C per decade in GSO2 (*39*).

The effect of Pinatubo can be quantified by subtracting GSO1 trends from GSOP. As in the case of stratospheric ozone, Pinatubo cools the lower troposphere by more than it cools the surface. For the specific period considered here (*21*), the relative cooling (by 0.015° ± 0.035°C per decade) is smaller than for ozone depletion. The combined effect of Pinatubo and stratospheric ozone depletion over the period 1979–1997, estimated by subtracting the baseline GSDIO simulation from GSOP, is to cool the model lower troposphere by 0.067° ± 0.030°C per decade relative to the model surface (*39*).

Conclusions. We have shown that the combination of natural and anthropogenic forcings in the ECHAM GSOP experiment yields 2LT and surface – 2LT trends that are closest to observations. Even in GSOP, however, there is still a significant difference between models and data in terms of their relative temperature changes at the surface and in the lower troposphere. This discrepancy is probably related to a combination of four factors: forcing uncertainties, model errors, residual uncertainties in the surface and MSU 2LT data, and signal estimation problems. We consider these in turn below.

First, there are large uncertainties in the strength, patterns, and evolution of the forcings associated with direct and indirect sulfate aerosol effects and tropospheric and stratospheric ozone (*41*). This study has not explored the effects of such uncertainties or of assumptions made in the applied forcing (*33*). Furthermore, it neglects other forcings that may be important to an understanding of remaining discrepancies between modeled and observed atmospheric

Table 1. Effect of various masking options on spatially averaged annual mean surface (IPCC) and lower tropospheric (MSUd 2LT) temperature anomalies over 1979–1998, and on the IPCC minus MSU difference time series. The tabulated values show the effects of NOMASK, VARMASK, and FIXMASK on the RMS difference and on the correlation between surface and 2LT time series, and also on trends and estimated 95% trend confidence intervals. Trends and trend differences that are significantly different from zero at the 1% level or better are denoted by an asterisk. (There are no results significant at the 5% level that are not also significant at the 1% level.) Both confidence intervals and estimated significance levels account for temporal autocorrelation effects (*21*).

Statistic	Data set	NOMASK	VARMASK	FIXMASK
RMS difference (°C)	IPCC versus MSUd 2LT	0.117	0.108	0.102
Correlation	IPCC versus MSUd 2LT	0.742	0.810	0.831
Trend and 95% confidence interval (°C per decade)	IPCC	+0.196 ± 0.104*	+0.196 ± 0.104*	+0.201 ± 0.106*
	MSUd 2LT	+0.057 ± 0.157	+0.089 ± 0.166	+0.104 ± 0.166
	IPCC minus MSUd 2LT	+0.139 ± 0.061*	+0.106 ± 0.068*	+0.097 ± 0.070*

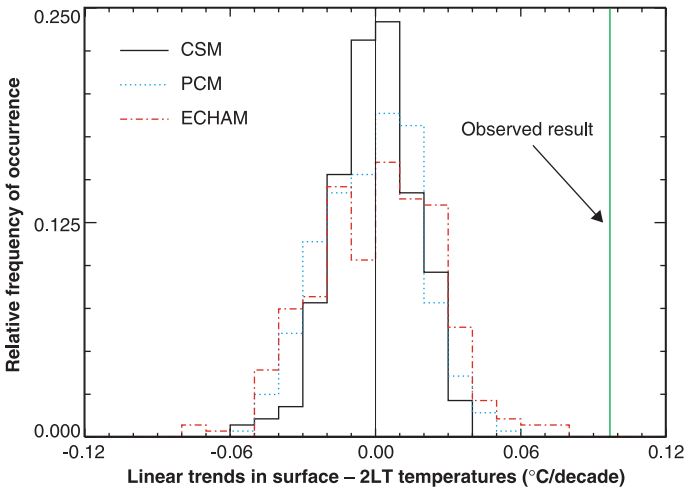

Fig. 4. Sampling distributions of 20-year trend differences between the surface and lower troposphere in three coupled model control runs. Distributions were computed as described in (*29*), with FIXMASK masking. The observed trend difference of +0.096°C per decade (Table 1) is also indicated.

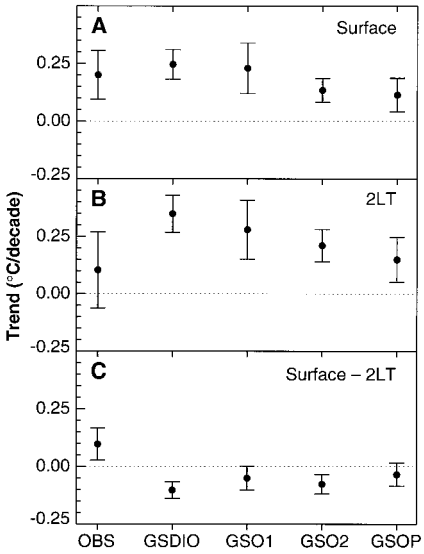

Fig. 5. Least-squares linear trends and associated 95% confidence intervals in modeled and observed surface (**A**), 2LT (**B**), and surface – 2LT (**C**) temperature time series. Observed trends (OBS) and confidence intervals over 1979–1998 were computed with annual mean, spatially averaged IPCC surface and MSUd 2LT data, using FIXMASK masking (*16*). Model-based results are from experiments with anthropogenic and natural forcings performed with ECHAM (*12, 24*). Model data were processed as described (*28, 29*), but with the difference that results are now given for only one 20-year period in GSDIO ("1979–1998" in model years) and for one 19-year period in GSO1, GSO2 and GSOP ("1979–1997" in model years). Model results are also based on FIXMASK sampling. Confidence intervals are adjusted to account for temporal autocorrelation in the data (*21*).

temperature profiles (*42*).

Second, there are errors in the model-predicted responses. In GSOP, for example, the model-predicted warming of the lower stratosphere induced by the Pinatubo aerosol is roughly twice as large as observed (*12*).

Third, decisions made in correcting for inhomogeneities in the observational data introduce further uncertainties (*1, 8, 14*). These may be of order ± 0.05°C per decade for the MSU 2LT data (*8*). If the MSU 2LT trend were 0.05°C per decade warmer than in the current version of the MSU data (MSUd), the observed surface – 2LT trend differential would not be significantly different from that in GSOP (see Fig. 5C). A nonsignificant trend differential would also occur if the surface warming had been overestimated by 0.05°C per decade in the IPCC data (*8*). The relative likelihood of such errors in the MSU and IPCC data is difficult to assess (*43*).

Fourth, we find substantial differences between two realizations of the GSO experiment (*36*), consistent with the broad distributions of "unforced" surface – 2LT temperature differences in coupled model control runs (Fig. 4). These results highlight the difficulty of reliably estimating the climate responses to different forcing mechanisms without multiple realizations for each forcing experiment.

All of these factors make it difficult to determine the precise cause or causes of recent observed surface-troposphere temperature trend differences. To better understand these causes, we urgently require additional simulations of the climate of the past two decades. Such simulations should be performed with a variety of models and should explore current uncertainties in key natural and anthropogenic forcings, using multiple realizations of each experiment.

We have examined the surface – 2LT trend differential over a relatively short (20-year) period of the observational record. Analyses of tropical radiosonde data available since the late 1950s suggest that the troposphere may have warmed relative to the surface until the last 10 to 20 years (*44*). One interpretation of this apparent change is that it reflects the complex space-time evolution of natural and anthropogenic forcings. If this is the case, studies of model-data consistency should consider full space-time information rather than focusing on linear trends over a comparatively short period of record (*21, 45*).

References and Notes

1. P. D. Jones, M. New, D. E. Parker, S. Martin, I. G. Rigor, *Rev. Geophys.* **37**, 173 (1999); P. D. Jones, *J. Clim.* **7**, 1794 (1994); D. E. Parker, P. D. Jones, A. Bevan, C. K. Folland, *J. Geophys. Res.* **99**, 14373 (1994).
2. D. J. Gaffen, *J. Geophys. Res.* **99**, 3667 (1994); D. E. Parker *et al*., *Geophys. Res. Lett.* **24**, 1499 (1997); D. E. Parker and D. I. Cox, *Int. J. Clim.* **15**, 473 (1995).
3. P. D. Jones, T. J. Osborn, K. R. Briffa, *J. Clim.* **10**, 2548 (1997).
4. B. D. Santer *et al*., *J. Geophys. Res.* **104**, 6305 (1999).
5. R. W. Spencer and J. R. Christy, *Science* **247**, 1558 (1990); *J. Clim.* **5**, 847 (1992); *J. Clim.* **5**, 858 (1992).
6. J. W. Hurrell and K. E. Trenberth, *J. Clim.* **9**, 2222 (1996).
7. _____, *Nature* **386**, 164 (1997); P. D. Jones, T. J. Osborn, T. M. L. Wigley, P. M. Kelly, B. D. Santer, *J. Geophys. Res.* **102**, 30135 (1997); J. K. Angell, *Geophys. Res. Lett.* **26**, 2761 (1999). The satellite-estimated lower tropospheric temperature trend shows slight warming when data for 1998 are included.
8. Nonrandom temporal changes in the density of buoy SST measurements may have biased observed SST data [J. R. Christy, R. W. Spencer, E. S. Lobl, *J. Clim.* **11**, 2016 (1998)].
9. P. D. Jones, P. M. Kelly, C. M. Goodess, T. R. Karl, *J. Clim.* **2**, 285 (1989); P. D. Jones *et al*., *Nature* **347**, 169 (1990).
10. F. J. Wentz and M. Schabel, *Nature* **394**, 661 (1998).
11. J. E. Hansen *et al*., *Clim. Change* **30**, 103 (1995).
12. L. Bengtsson, E. Roeckner, M. Stendel, *J. Geophys. Res.* **104**, 3865 (1999).
13. J. E. Hansen *et al*., *J. Geophys. Res.* **102**, 25679 (1997).
14. J. R. Christy, R. W. Spencer, W. D. Braswell, *J. Atmos. Oceanic Tech.*, in press. This study finds that artificial cooling induced by the orbital decay effect identified in (*10*) is partially offset in MSUd 2LT by adjustments relating to instrument body temperature and diurnal drift effects. The MSUd 2LT data used here are monthly means on a 2.5° × 2.5° latitude-longitude grid, spanning January 1979 through December 1998. The maximum of the 2LT weighting function is at roughly 740 hPa.
15. N. Nicholls *et al*., in *Climate Change 1995: The Science of Climate Change,* J. T. Houghton *et al*., Eds. (Cambridge Univ. Press, Cambridge, 1996), pp. 133–192; see also (*1*). The IPCC surface data were available in the form of monthly mean anomalies (relative to climatological means over 1961–1990) on a 5° × 5° latitude-longitude grid, spanning the period January 1856 through December 1998. For comparison with MSU, we use data for 1979–1998 only.
16. In VARMASK, the MSU and IPCC areal coverage is identical and varies from a maximum of 77% in 1987 to a minimum of 70% in 1992. FIXMASK has smaller coverage (~60%) than the minimum coverage in VARMASK. Although NOMASK involves no masking of observed data, it does entail masking of model surface data with observed surface coverage.
17. J. K. Gibson *et al*., *ECMWF Re-Analysis Project Report Series. 1. ERA Description* (European Centre for Medium Range Weather Forecasts, Reading, UK 1997).
18. E. Kalnay *et al*., *Bull. Am. Meteorol. Soc.* **77**, 437 (1996).
19. J. W. Hurrell and K. E. Trenberth, *Bull. Am. Meteorol. Soc.* **80**, 2661 (1999).
20. Because the strength of the surface-troposphere temperature coupling shows considerable spatial variability (*6*), subsampling tropospheric temperature data with surface coverage would be less useful if one considered limited geographical areas. Our analysis, however, deals with temperatures over at least 60 to 70% of Earth's surface.
21. The method for assessing statistical significance of trends and trend differences is described by B. D. Santer *et al*. (*J. Geophys. Res.*, in press). It involves the standard parametric test of the null hypothesis of zero trend, modified to account for lag-1 autocorrelation of the regression residuals [see J. M. Mitchell Jr. *et al*., *Climatic Change, World Meteorological Organization Tech. Note 79* (World Meteorological Organization, Geneva, 1966)]. The adjustments for autocorrelation effects are made both in computation of the standard error and in indexing of the critical t value.
22. This apparent contradiction has a simple explanation. The significance of trend differences is determined from $d(t)$, the time series of paired differences between the surface and 2LT data. Use of $d(t)$ markedly reduces noise levels by subtracting variability components common to the surface and 2LT data, thus facilitating the identification of relatively small trend differences embedded in noisy but covarying time series.
23. The acronym denotes version 4 of the atmospheric model jointly developed by the European Centre for Medium-Range Weather Forecasts and the Max Planck Institute for Meteorology in Hamburg, coupled to the Hamburg isopycnal ocean model.
24. E. Roeckner, L. Bengtsson, J. Feichter, J. Lelieveld, H. Rodhe, *J. Clim.* **12**, 3004 (1999).
25. W. M. Washington *et al*., in preparation.
26. B. A. Boville and P. R. Gent, *J. Clim.* **11**, 1115 (1998).
27. Version 3 of the NCAR Community Climate Model (CCM3) [B. A. Boville and J. W. Hurrell, *J. Clim.* **11**, 1327 (1998)].
28. We calculated equivalent 2LT temperatures from the model monthly mean data using both a radiative transfer code and a static global mean weighting function (*4, 5*). The two methods yield highly similar global mean anomaly time series. Here we show results for the static weighting function approach only. Model SSTs were merged with air temperatures measured at 2 m over land to form surface temperature data sets directly comparable with the IPCC observations.
29. Sampling distributions were computed in the following way. For model surface and 2LT data, annual means were calculated and expressed as anomalies relative to the 300-year climatological annual mean. These anomalies were transformed to the IPCC grid. We then processed maximally overlapping 20-year data segments (i.e., model years 1 through 20, 2 through 21, etc.). For each 20-year segment, annual-mean anomalies were recomputed relative to the average over years 1 through 15 of each 20-year segment. This mimics the definition of anomalies in the observations (Fig. 1). For each control run, we used the NOMASK, VARMASK, and FIXMASK methods to sample the 2LT and surface data in each of the 281 20-year segments of model data. Finally, we computed area-weighted surface and 2LT spatial means and fitted least-squares linear trends to these and to the surface – 2LT difference time series. This yields nine sampling distributions of "unforced" 20-year trends and trend differences (3 control runs × 3 masking methods) for each of three "variables" (surface, 2LT, and surface – 2LT). From these distributions, we estimated the empirical probability of obtaining a model result that exceeds the observed trend or trend difference given in Table 1.
30. Although VARMASK has greater common coverage (between the surface and 2LT) than FIXMASK, VARMASK results are more difficult to interpret because part of the observed surface – 2LT trend difference in VARMASK is due to nonrandom changes in coverage over the period 1979–1998. This is not the case with FIXMASK.
31. Our terminology follows that used in (*12*).
32. J. A. Leggett, W. J. Pepper, R. J. Swart, in *Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment,* J. T. Houghton, B. A. Callander, S. K. Varney, Eds. (Cambridge Univ. Press, Cambridge, 1992), pp. 69–95.
33. This was done by linearizing observed monthly mean latitude-height profiles of stratospheric ozone change from November 1978 to April 1993, and then extrapolating the computed trends to December 1997 (*12*).
34. Bengtsson *et al*. (*12*) described further GSOP experiments that represent linear combinations of responses to various forcings. We considered the one GSOP experiment that was not a linear combination of individual responses. Although the GSOP linear combinations cover the period of the satellite record, their construction inflates the high-frequency variability, which hampers estimation of relatively small trends.
35. There are 12 pairs of observed-minus-simulated trend differences (e.g., observed surface trend minus GSDIO surface trend, etc.). Let b_x and b_y represent any single pair of modeled and observed trends, with standard errors s_{bx} and s_{by}. The latter are both adjusted for temporal autocorrelation effects, as in (*21*). We assume that the normalized trend difference $d = (b_x - b_y)/[(s_{bx})^2 + (s_{by})^2]^{1/2}$ has a Gaussian distribution, and that values of $d > 1.96$ (2.58) indicate trend differences significant at the 5% (1%) level.
36. See Web table 1 at *Science* Online (www.sciencemag.org/feature/data/1046025.shl).

37. S. F. Singer, *EOS* **80**, 183 (1999); P. J. Michaels, *World Clim. Rep.* **4**(17), 1 (1999); testimony before the Committee on Government Reform, Subcommittee on National Economic Growth, Natural Resources and Regulatory Affairs, and Committee on Science, Subcommittee on Energy and Environment, U.S. House of Representatives, joint hearing on "Is CO_2 a Pollutant and Does EPA Have the Power to Regulate It?", Washington, DC, 6 October 1999.
38. J. R. Christy and R. T. McNider, *Nature* **367**, 325 (1994); P. D. Jones, *Geophys. Res. Lett.* **21**, 1149 (1994); S. F. B. Tett, J. F. B. Mitchell, D. E. Parker, M. R. Allen, *Science* **274**, 1170 (1996).
39. See Web table 2 at *Science* Online (www.sciencemag.org/feature/data/1046025.shl).
40. Bengtsson *et al.* (*12*), in their analysis of various realizations (*34*) of the GSOP experiment, show slight warming (by roughly +0.01° to +0.03°C per decade) of the model surface temperature relative to temperatures at 850 and 500 hPa. These results are not inconsistent with our finding of a slight cooling (by −0.035°C per decade) of the surface relative to lower troposphere in the single GSOP experiment we consider (Fig. 5C). These small differences arise from our use of a vertically weighted lower tropospheric temperature (the 2LT retrieval) rather than discrete temperatures at 850 and 500 hPa, and from our masking procedure, which excludes areas of large surface warming at high latitudes in both hemispheres [see (*12*), plate 3].
41. K. P. Shine, Y. Fouquart, V. Ramaswamy, S. Solomon, J. Srinivasan, in *Climate Change 1995: The Science of Climate Change*, J. T. Houghton *et al.*, Eds. (Cambridge Univ. Press, Cambridge, 1996), pp. 108–131; J. E. Hansen, M. Sato, R. Ruedy, *J. Geophys. Res.* **102**, 6831 (1997); T. K. Berntsen *et al.*, *J. Geophys. Res.* **102**, 28101 (1997).
42. Examples include the effects of El Chichón and other volcanic eruptions [N. G. Andronova, E. V. Rozanov, F. Yang, M. E. Schlesinger, G. L. Stenchikov, *J. Geophys. Res.* **104**, 16807 (1999)], as well as aerosols resulting from biomass burning [J. E. Penner, R. Dickinson, C. O'Neill, *Science* **256**, 1432 (1992)] and mineral dust [I. Tegen and I. Fung, *J. Geophys. Res.* **100**, 18707 (1995)].
43. National Research Council, Board on Atmospheric Sciences and Climate, Commission on Geosciences, Environment and Resources, *Reconciling Observations of Global Temperature Change* (National Academy Press, Washington, DC, 2000).
44. D. J. Gaffen *et al.*, *Science* **287**, 1242 (2000).
45. B. D. Santer *et al.*, *Nature* **384**, 523 (1996); T. M. L. Wigley, P. J. Jaumann, B. D. Santer, K. E. Taylor, *Clim. Dyn.* **14**, 781 (1998); S. F. B. Tett *et al.*, *Nature* **399**, 569 (1999); K. Hasselmann, *J. Clim.* **6**, 1957 (1993).
46. Supported by NOAA Office of Global Programs ("Climate Change Data and Detection") grant NA87GP0105 (T.M.L.W.) and U.S. Department of Energy (DOE) grant DE-FG02-98ER62601 (P.D.J. and T.M.L.W.). Work at Lawrence Livermore National Laboratory was performed under the auspices of the Environmental Sciences Division of DOE (contract W-7405-ENG-48). We thank J. R. Christy for the MSU data and static MSU weighting functions, and C. F. Keller, G. C. Hegerl, J. M. Wallace, M. C. MacCracken, A. Robock, and two anonymous reviewers for useful comments and suggestions that substantially improved the manuscript.

5 October 1999; accepted 29 December 1999

Crystal Structure of the Ribonucleoprotein Core of the Signal Recognition Particle

Robert T. Batey,[1] Robert P. Rambo,[1] Louise Lucast,[2] Brian Rha,[2] Jennifer A. Doudna[2*]

The signal recognition particle (SRP), a protein-RNA complex conserved in all three kingdoms of life, recognizes and transports specific proteins to cellular membranes for insertion or secretion. We describe here the 1.8 angstrom crystal structure of the universal core of the SRP, revealing protein recognition of a distorted RNA minor groove. Nucleotide analog interference mapping demonstrates the biological importance of observed interactions, and genetic results show that this core is functional in vivo. The structure explains why the conserved residues in the protein and RNA are required for SRP assembly and defines a signal sequence recognition surface composed of both protein and RNA.

Cells communicate with their surroundings by means of proteins that either reside within cellular membranes or are secreted to the outside. One of the principal pathways for targeting these proteins uses the SRP to catalyze cotranslational transport of nascent secretory and membrane proteins to the endoplasmic reticulum (ER) in eukaryotes and to the plasma membrane in prokaryotes. The well-characterized eukaryotic SRP, an evolutionarily conserved ribonucleoprotein complex, recognizes the NH_2-terminal signal sequence of targeted proteins as they emerge from the ribosome. Binding to the SRP arrests polypeptide elongation and mediates docking of the translating ribosome with receptors on the ER in a guanosine 5′-triphosphate (GTP)–dependent process (Fig. 1A) (*1*, *2*).

In bacteria, the SRP is essential for cell viability and efficient protein export (*3*) and consists of the 4.5*S* RNA and the Ffh protein. These components as well as the SRP receptor, FtsY, share sequence and functional homology with their eukaryotic counterparts 7SL RNA, SRP54, and SRα, respectively (*4*–*8*). The evolutionary conservation of this fundamental cellular component is demonstrated by the ability of human SRP54 to bind with high affinity to the *Escherichia coli* 4.5*S* RNA and to rapidly hydrolyze GTP in the presence of the SRP receptor (*5*). Similarly, Ffh is able to replace SRP54 in a chimeric mammalian SRP that is capable of elongation arrest and signal sequence recognition (*6*). Thus, the Ffh-4.5*S* RNA complex appears to be a minimized structural and functional homolog of the eukaryotic SRP, which makes it attractive for detailed structural and mechanistic studies.

The Ffh/SRP54 proteins contain three domains: N, G, and M. The NH_2-terminal N domain, a four-helix bundle, is closely associated with the adjacent G domain, a Ras-like guanosine triphosphatase (GTPase) (*9*) responsible for mediating the interaction of the SRP with its receptor and regulating SRP function through hydrolysis of GTP (*7*, *10*). Structurally related N and G domains are also present in the SRP receptor (*11*). The methionine-rich M domain near the COOH-terminus of Ffh/SRP54 contains recognition sites for both the signal peptide and the SRP RNA (*12*–*14*).

The Ffh binding site on the 4.5*S* RNA is localized to domain IV, a ~50-nucleotide (nt) region whose highly conserved secondary structure consists of two internal loops that include noncanonical base pairings and unpaired nucleotides (Fig. 1, B and C) (*15*–*18*). In bacterial and human SRP, several lines of evidence suggest that domain IV (Fig. 1, B and C) stabilizes the particle and its interaction with the signal peptide (*19*, *20*). Outside of domain IV, the size, sequence, and secondary structure of SRP RNAs vary widely, even among bacterial species (*21*), and it is unclear if these regions of the RNA are essential for SRP function.

Here we present the 1.8 Å resolution crystal structure of the universally conserved ribonucleoprotein core of the SRP, a complex between domain IV of 4.5*S* RNA and the M domain of Ffh. Nucleotide analog interference mapping and genetic results show that the structure represents the functional complex in vivo. In the structure, unique RNA-protein interactions characterize the molecular interface, and a network of highly ordered waters and metal ions also mediate key contacts. The structure explains why phylogenetically conserved residues in the protein and RNA are required for SRP assembly, and it suggests a possible role for the RNA in molecular recognition of signal peptides.

Structure determination and overview. To obtain well-ordered crystals of the M domain in complex with SRP RNA, we

[1]Department of Molecular Biophysics and Biochemistry, [2]Howard Hughes Medical Institute, Yale University, New Haven, CT 06511, USA.

*To whom correspondence should be addressed. E-mail: doudna@csb.yale.edu