

Observed and Simulated Temperature–Humidity Relationships: Sensitivity to Sampling and Analysis

MIKE BAUER

Columbia University, New York, New York

ANTHONY D. DEL GENIO

NASA Goddard Institute for Space Studies, New York, New York.

JOHN R. LANZANTE

NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 6 April 2001, in final form 30 July 2001)

ABSTRACT

Recent studies have demonstrated that the correlation between interannual variations of large-scale average temperature and water vapor is stronger and less height dependent in one GCM than in an objective analysis of radiosonde observations. To address this discrepancy, a GCM with a different approach to cumulus parameterization is used to explore the model dependence of this result, the effect of sampling biases, and the analysis scheme applied to the data.

It is found that the globally complete data from the two GCMs produce similar patterns of correlation despite their fundamentally different moist convection schemes. While this result concurs with earlier studies, it is also shown that this apparent model–observation discrepancy is significantly reduced (although not eliminated) by sampling the GCM in a manner more consistent with the observations, and especially if the objective analysis is not then applied to the sampled data. Furthermore, it is found that spatial averages of the local temperature–humidity correlations are much weaker, and show more height dependence, than correlations of the spatially averaged quantities for both model and observed data. The results of the previous studies are thus inconclusive and cannot therefore be interpreted to mean that GCMs greatly overestimate the water vapor feedback.

1. Introduction

Climate change simulations using general circulation models (GCMs) depict a large positive feedback on the greenhouse effect because of changes in water vapor concentration and its vertical distribution. For a given relative humidity, changes in specific humidity are determined by temperature via the Clausius–Clapeyron equation. Because of this dependence, large changes in specific humidity will accompany a climate change in temperature if relative humidity remains constant. In principle, however, dynamic transports and microphysical processes can change relative humidity significantly. Climate warming is likely to increase the specific humidity in the lower troposphere, as it is close to the surface source of water vapor (Del Genio et al. 1991; Lindzen 1990). The ultimate contribution of cloud, convection, and other dynamic processes to the water vapor

budgets of the mid- and upper troposphere is less clear, though, and thus uncertainty exists in the overall magnitude of the water vapor feedback.

Given the wide variety of approaches to the parameterization of moist convection, cloud formation, and microphysical processes in GCMs, one might expect water vapor feedbacks to differ from model to model. But this is not the case; all GCM climate change experiments agree that the water vapor feedback is strongly positive (cf. Cess et al. 1990; Houghton et al. 1990). If this is correct, then it might be possible to find examples of strong positive correlations between specific humidity and temperature in some (though not all) types of observed climate variability as well.

Indeed, many examples of locally positive temperature–humidity (T – q) correlations exist (e.g., Inamdar and Ramanathan 1994; Raval and Ramanathan 1989; Rind et al. 1991; Stephens and Greenwald 1991). In general, though, these findings can be explained as a geographic correlation of humidity with large-scale ascending motions, which by themselves may not reflect the sense of a global climate change in specific humidity

Corresponding author address: M. Bauer, NASA GISS 2880 Broadway, New York, NY 10025.
E-mail: mbauer@giss.nasa.gov

(Sun and Oort 1995). Several studies attempt to avoid this problem by averaging over the Tropics (or the whole globe) and thereby include settings characterized by ascending and descending motion. Such studies thus capture the large-scale specific humidity response to large-scale temperature fluctuations in time while incorporating the net effect of the large-scale dynamics on the humidity distribution. The result again is a strong positive relationship between temperature and humidity (Soden and Fu 1995; Yang and Tung 1998).

The companion studies of Sun and Oort (1995, hereafter SO) and Sun and Held (1996, hereafter SH) lead to a different conclusion. In their study, SO examined the interannual fluctuations of zonal average and tropical average temperature and specific humidity derived from the global radiosonde climatology. In their treatment, SH did the same but with output from the Geophysical Fluid Dynamics Laboratory (GFDL) GCM. While SO found a positive correlation between temperature and humidity at all altitudes, they noted that its strength decreases dramatically above the trade inversion to a midtropospheric minimum that only recovers again to moderate values in the upper troposphere. In contrast, SH noted strong positive T - q correlations throughout the troposphere of the GFDL GCM. This discrepancy places the magnitude of the water vapor feedback in this and all other GCMs in doubt and has been blamed on the inadequate parameterization of moist convection in these GCMs (Hu et al. 2000; Sun and Held 1996). In line with this trend in ideas is a recent finding that many GCMs maintain stronger correlations between surface humidity variations and those throughout the atmospheric column than are observed (Sun et al. 2001). However, attributing this to a faulty moist convection scheme is problematic given the diversity of approaches used to model this process (Rind 1999; Sun et al. 2001).

Perhaps some of these discrepancies can be reconciled by examining the limitations of the observations themselves. The radiosonde network, for example, provides only a sparse and biased sample of the Tropics; one that is particularly lacking over oceanic regions characterized by descending motion (Raval et al. 1994; Waliser et al. 1999). Indeed, Oort (1978) showed that the climatological statistics of a GCM are altered by the sampling characteristics of the global radiosonde network. He cautioned that this network might be inadequate to resolve longitudinal variability. Because interannual variations in the Tropics are dominated by El Niño variability, which involves spatial redistributions over both well- and poorly sampled locations, it is doubtful that the radiosonde climatology captures this kind of variability adequately. Indeed, Soden and Lanzante (1996) demonstrated such a bias by resampling satellite humidity measurements with a radiosonde-like network. Moreover, the radiosonde climatology used by SO was processed by an objective analysis scheme that interpolates observed values into the unobserved regions;

the impact of this operation on the resulting T - q correlation is unknown.

Inevitably the value of such comparison (SO and SH) rests on the reliability of the observations that the GCM is being asked to match. However, studies of radiosonde data quality indicate numerous potential distortions and inaccuracies, including procedural and instrument changes, instrument biases and limitations, reporting and archiving errors and more (e.g., Elliott and Gaffen 1991; Gutzler 1993; Luers and Eskridge 1998). However, SO argue that the properties of linear correlation and of large-scale averaging downplay these problems and uncertainties. In part this argument depends on random and systematic errors in the temperature record being comparatively small (Sun and Oort 1995). This argument is probably not correct for specific humidity, but SO reason that it is also unlikely that humidity errors covary in time with temperature. It follows then, that radiosonde data can be used to estimate the actual T - q correlation accurately (Sun and Oort 1995). Implicit in this is the assumption that radiosonde data are a representative sample of the atmosphere. Thus, SO rely on large-scale averaging to mitigate this concern. This is not an unqualified solution, however, as it makes no provision for the uncertainties introduced by missing and nonexistent data; that is, cases in which temperature measurements exist when or where humidity measurements do not (or less commonly the reverse). This also includes circumstances in which the atmosphere goes unobserved altogether. Such omissions seem especially problematic for correlations based on quantities that have been averaged separately (as in SO and SH). In practice then, we should not interpret T - q correlations as anything more than estimates of the behavior of the available data. In any case, tropical or zonal averages are probably not the best approach for understanding the water vapor feedback given the extremely nonlinear dependence of outgoing longwave radiation on temperature and humidity (Yang and Tung 1998). Other approaches have been used (e.g., Hu et al. 2000; Yang and Tung 1998), but a detailed account of the strengths and limitations of each approach, and what they tell us about the water vapor feedback, has not been conducted.

In this paper, we use the Goddard Institute for Space Studies (GISS) GCM to explore some of these issues and make more appropriate comparisons between the model and the radiosonde dataset. We compare T - q correlations for the full GCM, a version that is sampled in the same way as the data, and a version that is both sampled and objectively analyzed in the same way as the data used by SO. We also compare results from the original in situ sampled radiosonde dataset and the objectively analyzed version discussed by SO. We describe our data, model, and analysis methods in section 2. Our results are presented in section 3, and we discuss their implications for the question of water vapor feedback in section 4.

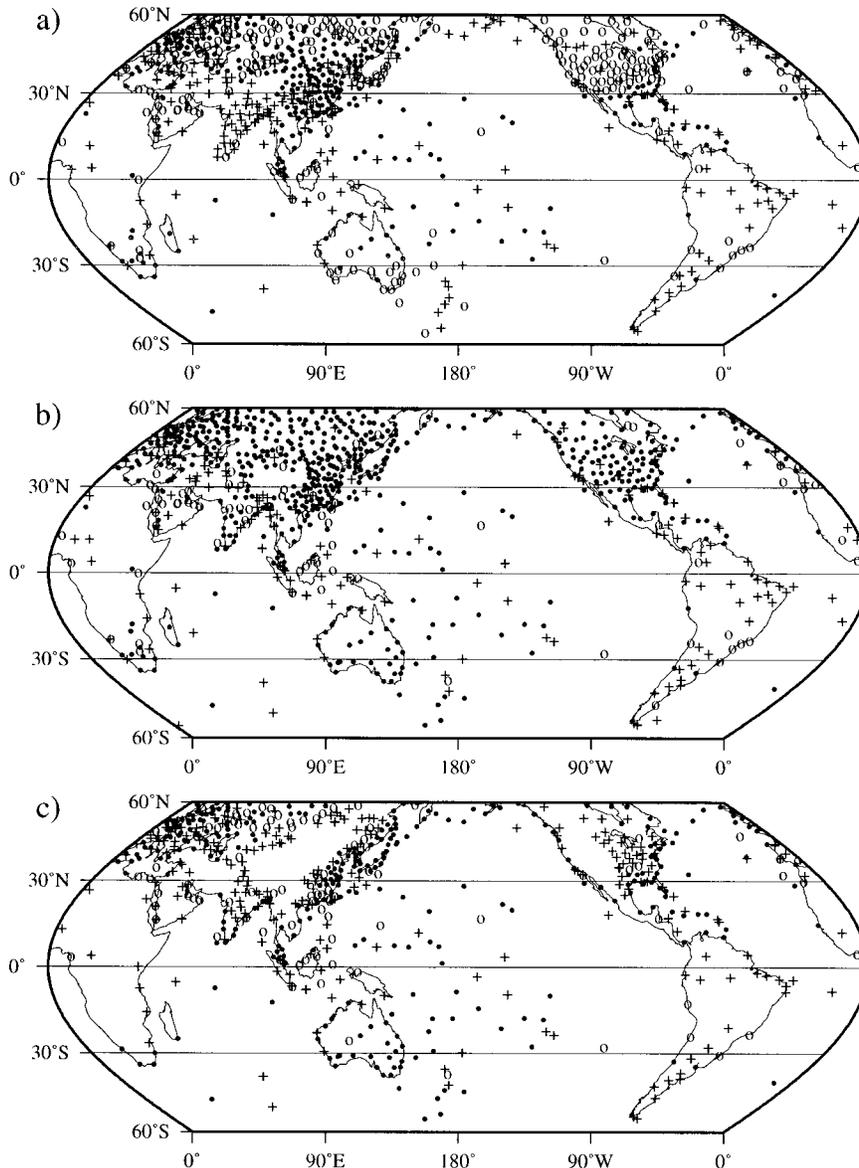


FIG. 1. Distribution of valid radiosonde stations at (a) 300, (b) 700, and (c) 950 hPa. Here a valid station contains no less than 10 days of T and q data (0000 UTC) during the calendar month. Stations reporting more than 70% of the available months (Jan 1979–Dec 1989) are marked as a closed circle. Stations reporting more often than 30% but less than 70% of the available months are marked as an open circle. Finally, stations reporting less than 30% of the available months are marked with a plus.

2. Data and methods

a. Data and model

Here we compare 11 yr (January 1979–December 1989) of observations against a GCM simulation of this period. The observations come from an updated version of the radiosonde archive documented by Oort (1983). Using this archive we compiled a set of station time series detailing the variations of monthly average temperature (T) and specific humidity (q). To qualify as a valid monthly average, at least 10 days of *both* T and

q must have been reported at each location. Figure 1 shows the resulting temporal coverage and spatial distribution of these radiosonde stations. Regular tropical radiosonde measurements are only available as a sparse and irregular network, this being concentrated in south Asia, the west Pacific, and the Caribbean.

Note that these observations are only those from 0000 UTC soundings, whereas SO reports the average of the 0000 and 1200 UTC soundings if both are available (Oort 1983). In addition, we analyze only a subset of the time frame covered by SO and SH (1963–89). Sim-

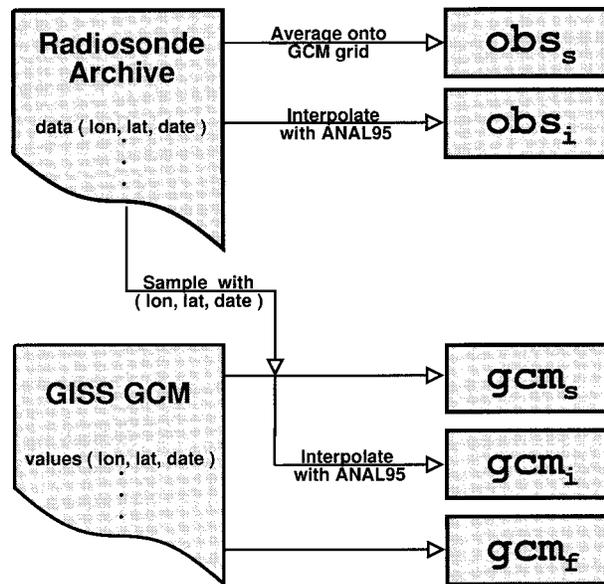


FIG. 2. Schematic view of the research design.

ilarities with SO and SH suggest that these differences do not greatly influence our conclusions.

The simulation analyzed here was done with the GISS GCM as part of the second phase of the Atmospheric Model Intercomparison Project (AMIP; Gates et al. 1999). The AMIP protocol prescribes a common set of radiative forcings and surface boundary conditions, including fixed greenhouse gases and insolation, climatological aerosol concentrations, and observed monthly variations of ozone, sea surface temperature, and sea ice. As suggested by AMIP, the SST and sea-ice fields were modified such that the interpolated daily values preserved the monthly average of the observed fields. Standard AMIP model output is based on continuously accumulated monthly average values. Consequently, the GCM values are not consistent with the intramonth and diurnal sampling of the observations.

This version of the GISS GCM operates on a $4^\circ \times 5^\circ$ latitude–longitude grid that spans 12 vertical sigma-coordinate layers. Moist convection in the GISS GCM is represented by a mass-flux parameterization with a cloud-base neutral buoyancy closure, convective downdrafts, entrainment, detrainment of condensate into anvils, and the reevaporation of precipitation (Del Genio and Yao 1993). Stratiform clouds are treated using a prognostic cloud water scheme that includes parameterizations of all important microphysical sources and sinks (Del Genio et al. 1996). Advection of heat and moisture uses the quadratic upstream scheme (cf. Prather 1986), which reduces the impact of numerical diffusion of water vapor. We use this full-field output (gcm_f) as a benchmark to assess the effects of sampling and interpolation.

The effect of the GISS cumulus parameterization on the humidity field has been documented by Del Genio

TABLE 1. Correspondence between radiosonde mandatory levels and GCM average layer pressures (reported in hPa).

Level	Obs	GCM
Upper	250	248
	300	338
	500	470
Middle	700	635
	850	894
Lower	950	959

et al. (1994, their Fig. 2b). Broadly, these findings are as follows: 1) parameterized subsidence dries and warms the atmosphere at most altitudes and latitudes, and 2) deep and shallow convection, respectively, moisten the upper troposphere and the boundaries of the subtropical trade inversion while having little effect on temperature. The GISS scheme should thus reduce or be neutral toward T – q correlations. Convective adjustment (cf. Manabe et al. 1965), on the other hand, adjusts an unstable temperature profile while maintaining saturated or fixed relative humidity conditions. That is, it dries the atmosphere while releasing latent heat whenever the reference humidity is exceeded. As in the GISS scheme, this action reduces T – q correlations. However, unlike the GISS scheme, convective adjustment also directly links changes in q to changes in T whenever q reaches its reference value. This likely contributes to positive T – q correlations in the GFDL GCM (Sun and Held 1996).

b. Sampling

A GCM sample (gcm_s) was created in the simplest manner; for each valid radiosonde station (month, location, and pressure), we substituted the corresponding GCM output (Fig. 2). We also constructed an analogous radiosonde sample (obs_s) by averaging the in situ data from all valid radiosonde stations onto the GISS GCM grid domain.

To avoid certain numerical ambiguities (Trenberth 1995) the GISS GCM accumulates each monthly average from values interpolated into pressure coordinates at the end of each model time step. From these we have selected six levels that best match the six mandatory reporting levels in the data (Table 1). The GCM creates “missing” values wherever interpolated pressure surfaces intersect the ground. Topography also occasionally precludes radiosonde observations at pressures above 850 hPa (Oort and Liu 1993); however, the highly smoothed topography in the GCM unrealistically eliminates low-level information from several key tropical locations. For the most part, these lost stations are located along coastal South America, where high topography only abuts the actual stations (Fig. 1). In such cases we replaced these missing GCM values with the average of neighboring grid values. This substitution was only applied to the fields input into the interpolation

described next, and does not affect the pure sampled fields gcm_s and obs_s .

c. Objective analysis

Next we interpolated the monthly average temperature and specific humidity from each valid radiosonde station to fill the entire GISS GCM grid domain; we refer to these fields as obs_i and gcm_i (Fig. 2). This interpolation was done with the objective analysis scheme ANAL95. ANAL95 is an updated, but algorithmically similar, version of the scheme used by Oort (1983, i.e., ANAL68).

Creating a guess field for the unobserved spaces between the radiosonde stations is the first, and in some ways the most influential, step in applying ANAL95 (Oort 1978). For this we used the zonal average of the valid radiosonde stations (Oort 1978; Sun and Held 1996). Next, this guess field is subtracted from the station values, creating an anomaly field, which is then passed into the ANAL95 software. The initial ANAL95 fit is zonally symmetric; created by binning (in 12° lat intervals) and then latitudinally smoothing (using a fourth-order polynomial) the anomaly field (Raval et al. 1996). The fitted field is then adjusted to the station values. For this, each station can influence the fitting field over a radius of 1750 km (the default for ANAL95). The adjusted fit is then smoothed. This process is then repeated using a smaller radius of influence (700 km). Afterward, the adjusted fit undergoes a final cross-equatorial smoothing as each hemisphere is analyzed separately on a polar stereographic grid. The original guess field is then added back to give the final product. This procedure is designed to mimic that used to produce the data products analyzed by SO and SH. It is worth noting, though, that the interpolation of a highly variable parameter such as specific humidity is bound to introduce significant errors relative to a strategy that interpolates relative humidity instead.

d. Analysis procedure

We focus on interannual anomalies of T and q by removing the seasonal cycle and trend using the same procedures as SO and SH. We calculate correlations from these anomalies in the same manner as SO, using the Pearson or linear correlation coefficient. We examine two methods for finding the large-scale correlation of the data. The first method follows SO and SH, as the correlations are based on spatially averaged temperature and humidity. The second method follows Hu et al. (2000) by examining spatially averaged local correlations. Strictly speaking correlation coefficients are not additive as assumed by Hu et al. (2000). A formal statistical procedure exists for this however, the Fisher-Z transformation, which we applied when averaging correlations (Dunlap et al. 1983).

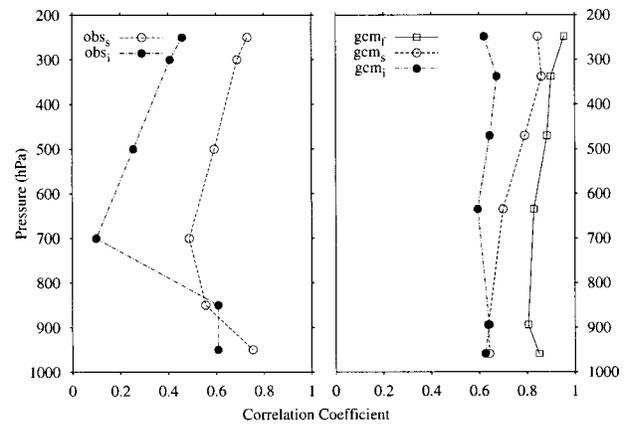


FIG. 3. Correlation coefficient as a function of pressure. Each panel represents the correlated interannual variation of the tropical average ($\pm 32^\circ$) temperature and specific humidity, i.e., $r(\langle T \rangle, \langle q \rangle)$, for either the observations (left) or the GISS GCM (right). Each curve represents one of three treatments: a solid line (open square) for the full GCM, a dashed line (open circle) for products of sampling, and a dash-dot line (filled circle) for products of objective analysis.

3. Results

a. Tropical averages

Figure 3 compares vertical profiles of the correlation between tropical average (32°N – 32°S) T and q for the various data and model versions. The correlations are strongest at all levels in the complete GCM (gcm_f). There is little difference between this profile and that from the GFDL GCM (cf. Sun and Held 1996) despite fundamentally different approaches to convective parameterization. Attributing the strength of the GFDL correlation profile to that model's convective adjustment scheme is therefore dubious. Thus the more likely scenario is that it is the vertical transport of water vapor by the resolved dynamics, which should be qualitatively similar in the two models, that determines this result (cf. Del Genio et al. 1991, 1994).

The weakest correlations in Fig. 3 appear in the interpolated data (obs_i). The contrast of this kind of comparison, that is, obs_s versus gcm_f , led SH and Sun et al. (2001) to improperly conclude that GCMs in general overestimate the coupling of temperature and humidity variations. Figure 3 demonstrates that this conclusion is overstated as much of the contrast between obs_s and gcm_f is an artifact of the objective analysis procedure; without interpolation, correlations in the data are larger by 0.2–0.3 at most levels. This dramatic difference is a consequence of the objective analysis procedure compensating for the spatial sampling pattern of the radiosonde stations and gaps in their time series. There is little difference, for example, between the correlation profile of obs_s and obs_i when the interpolated data are sampled at the places and times where data are present (i.e., the objective analysis does little to the actual data). On the other hand, when we sample obs_i at the same radiosonde locations but include the complete time se-

ries (i.e., allow dates filled in by the analysis), the resulting correlation profile (not shown) diverges from obs_s by about half the difference between the obs_s and obs_i profiles seen in Fig. 3. The remaining difference is thus due to contributions from locations without any radiosonde stations, where the time series is purely interpolated from other locations. The GCM is only sensitive to this latter effect, implying that interannual anomalies in the GCM are more spatially coherent than are those in the observations at locations where stations exist. This probably accounts for the relative similarity of the sampled and interpolated version of the GCM in Fig. 3.

The most direct and fair model-data comparison, that is, the purely sampled fields without interpolation (gcm_s vs obs_s), suggests that simulated T - q correlations at places and times where data exist are only about 0.2 greater than observed. The distinctive midtropospheric minimum in the sampled data is not present in the GCM, however. This result depends on paired observations, that is, that each station contributing to the tropical average reports both T and q . If we instead adopt the criteria of SO and SH, that is, use all available T and q for their respective tropical averages, the impression of a midtropospheric minimum is largely removed. In other words, relative to Fig. 3, upper-tropospheric correlations decrease by 0.2–0.3, while those at higher pressures remain unchanged in the sampled data and GCM. Primarily, this reflects the relative lack of humidity measurements in the upper troposphere. Other procedural changes have less effect on our findings. For instance, doubling the number of daily reports needed to form a valid monthly average does little except decrease midtropospheric correlations by about 0.1 (cf. Fig. 3) in both the sampled data and GCM (not shown). Limiting the analysis to locations reporting 70% or more of the possible months is a bit more influential as it decreases all correlations by about 0.1–0.2 in the sampled data and by about half this in the sampled GCM (not shown). In either case this primarily results from small shifts in the station distribution (in time and space) that go into the analysis.

The preceding discussion suggests that the objectively analyzed radiosonde climatology does not realistically capture tropical average temperature and humidity variability (and in fact, it was never intended for this purpose). The radiosonde climatology without objective analysis is probably lacking as well because of the spatial, temporal, and vertical sampling pattern of the radiosonde stations. Thus, while tropical averages of thermodynamic quantities are in theory better proxies for climate change than are local or regional variations, in practice this is not the case with the existing radiosonde network. Furthermore, outgoing longwave radiation is a highly nonlinear function of specific humidity and tropical averages of specific humidity must summarize fields containing substantial local variations. Thus examining the temporal variability of tropical averages

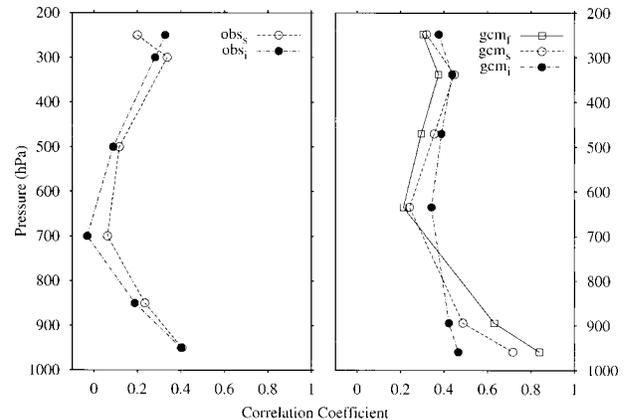


FIG. 4. Same as in Fig. 3, but for the tropical averaged local correlations between temperature and specific humidity, i.e., $\langle r(T, q) \rangle$. Note that each local correlation coefficient (having at least 10 valid monthly values) was first subjected to the Fisher-Z transformation, averaged, and then back-transformed into units of correlation.

may not be the best strategy for reaching conclusions relevant to climate change.

Consider Fig. 4, which shows instead the tropical average of the local T - q correlations for the various data and model versions. Overall, these correlations are weaker than are their counterparts in Fig. 3; a decrease in keeping with the statistical notion that individual elements of a population are less strongly correlated than are their average values (Freedman et al. 1978). The similarity of the interpolated data in both figures contradicts this pattern. However, the weak correlations reported in Fig. 3 reflect a bias in the analysis more than a property of the data themselves, as we argued earlier. This does not appear to be the case in Fig. 4. That is, the interpolated and sampled data are similar in this figure. Thus, while the objective analysis alters the local values such that the correlations of their tropical averages are distinct from those of just the data, it does not affect the local relationships between these values, and hence the tropical average of their local correlations resembles that for just the data. Indeed, the method of averaged correlations reduces the effects of sampling and objective analysis in both the data and the GCM, a more robust result. This method also draws out a midtroposphere minimum in the GCM correlations, much like that in the data, but at most altitudes the GCM correlations remain about 0.2 higher than observed, similar to the conclusion we reached by comparing the sampled model and data versions in Fig. 3.

b. Zonal averages

Figures 5 and 6 show cross sections of the correlation between zonally averaged T and q for both the data and the GCM. Generally the GCM is more positively correlated than are the data. Comparison with SH suggests that both GCMs share similar T - q correlations at all

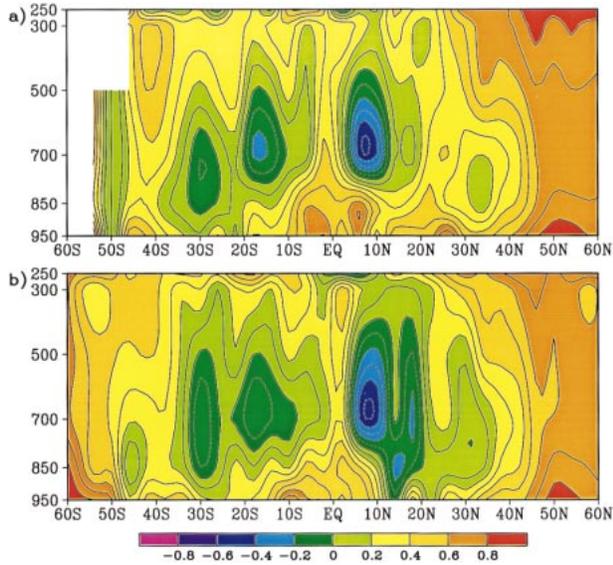


FIG. 5. Cross sections of the correlation coefficient of interannual variations of the zonal average temperature and specific humidity, i.e., $r[T, q]$. Shown here are (a) the sampled observations and (b) the interpolated observations.

latitudes and pressures, and thus, diverge from the observations in the same way. Again this suggests that the parameterization of moist convection in these GCMs is not a dominant factor in T - q correlation.

The sampled data and GCM agree near the equator (Figs. 5a and 6b). However, in the GCM this agreement only reflects the longitudinal distribution of radiosonde stations, which primarily sample the warmest regions of the Tropics, places with little interannual temperature variability (southern Asia, Malaysia, the western Pacific). Thus, these sampled equatorial correlations are lower than those from the complete GCM (Fig. 6a) which also includes cooler east Pacific locations and their large variations in both T and q due to ENSO. We cannot tell if this bias also affects the observations, although Soden and Lanzante (1996) found a similar shortcoming in a radiosonde-like sample of satellite data with ENSO variations.

Spatial sampling cannot directly explain the large model-data differences around 6°N , where midtropospheric correlations are opposite in sign (Figs. 5a and 6b). This difference can be traced to the 1982/83 El Niño. As SO noted, this was an unusual El Niño because the humidity decreased more in the west Pacific than it rose in the east Pacific, resulting in a negative zonal average humidity anomaly. The GISS and GFDL GCMs in contrast, give the more usual El Niño response (cf. Pan and Oort 1983) of a positive zonal average humidity anomaly (Sun and Held 1996). Indeed much of the model-data difference at 6°N depends on the uniqueness of this event; removing January–April 1983 decreases the most negative correlations in Fig. 5a (700 hPa, 6°N) to nearly zero, leaving a discrepancy of about 0.2 with the

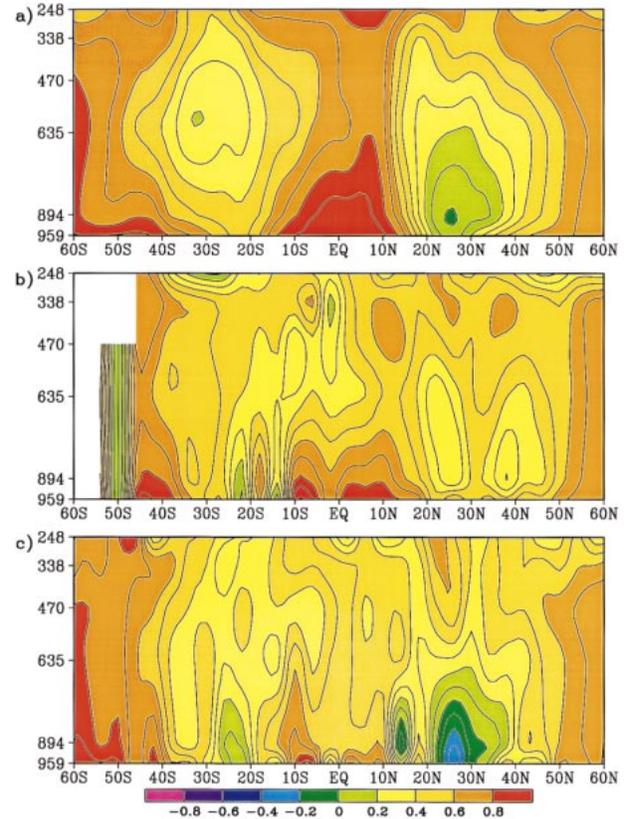


FIG. 6. As in Fig. 5, but for the GISS GCM. Shown here are (a) the complete GCM, (b) the sampled GCM, and (c) the interpolated GCM.

GCM. It is likely that this discrepancy is not an artifact of the radiosonde dataset because satellite measurements suggest a similar negative humidity anomaly at this time (Bates et al. 1996). Moreover, another example of a negative temperature and humidity relationship exists between 1965 and 1968 (see Sun and Oort 1995, their Fig. 3). Given the well-known sensitivity of linear correlation to bivariate outliers, these episodes likely weakened the correlations reported by SO, and thus, increased the model-data discrepancies reported by SH (see Lanzante 1996, his Figs. 10–11).

Most tropical grids contain only the products of objective analysis, which in many cases is simply a smoothed version of the initial-guess field for the analysis—the zonal average of the station values. As a result, the zonal averages of the sampled and interpolated GCM agree more with each other than either one does with the full GCM, as do the correlations derived from them. The same arguments probably apply to the data. Differences exist, however, particularly in the humidity field. To understand these differences, we artificially manipulated the radiosonde network locations and the parameters in the ANAL95 software in a series of sensitivity experiments. Overall, we found that the objective analysis scheme has two shortcomings for our pur-

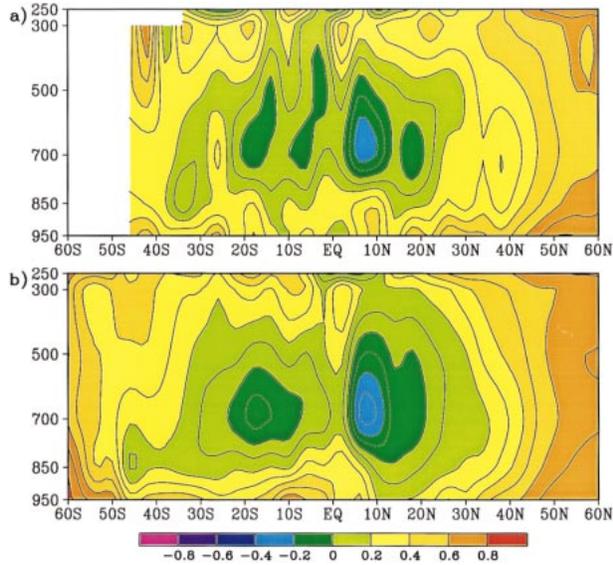


FIG. 7. As in Fig. 5, but for the zonal averaged local correlations (having at least 10 valid monthly values) between temperature and specific humidity, i.e., $[r(T, q)]$. Shown here are (a) the sampled observations and (b) interpolated observations. Note that a Fisher-Z transformation has been applied to these values.

poses: 1) it mixes information between latitudes, and 2) it extends the influence of isolated stations across vast regions of space. The decreased correlations produced by the analysis at 30°N (e.g., Fig. 5a vs 5b) can be traced to both the incorporation of extratropical “noise” into the fragmentary time series of some south Asian stations, and to the enhanced influence of isolated stations at Midway Island and in north Africa. The emergence of negative correlations around 18°N (Fig. 5) is likewise dependent on the enhanced influence of the Hawaiian, Wake, and Marshall Islands. Combined, these differences account for much of the midtropospheric minimum correlation seen in Fig. 3. However, this sort of bias depends on the meteorological setting being sampled by these isolated stations. For instance, stations from French Polynesia bias Fig. 5b toward more positive correlation at 18°S. The impact of these isolated stations can be reduced by setting both influence radii in the ANAL95 software to their smallest values (350 km—about half the longitudinal GCM-grid spacing), although this merely fills the data-void regions with the zonal average of the available data. Generally, data from the Southern Hemisphere, as well as output from the GCM in both hemispheres, are less affected by objective analysis. This pattern mirrors that of the standard deviation of the zonal anomaly (not shown); the Northern Hemisphere is more variable than the Southern Hemisphere and the data are more variable than the GCM. That is, sampling and objective analysis are most damaging where significant spatial inhomogeneity exists and observations are few. Moreover, objective analysis often

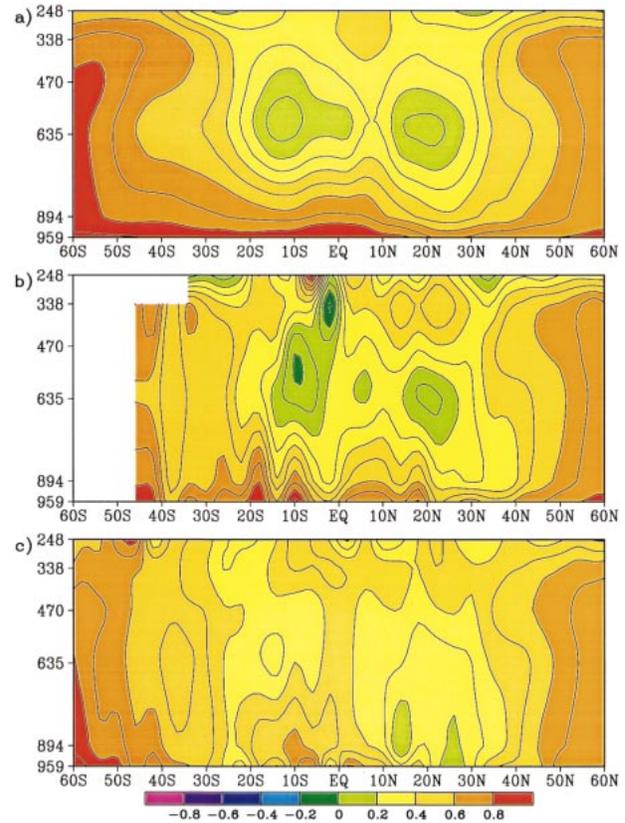


FIG. 8. As in Fig. 7, but for the GISS GCM. Shown here are (a) the complete GCM, (b) the sampled GCM, and (c) the interpolated GCM.

reduces longitudinal variability compared to the sample (or the complete field in the GCM).

Figures 7 and 8 show the zonally averaged T - q correlations for the GCM and the observations. As before, these averaged correlations are weaker than their counterparts based on correlated zonal averages. This decrease is most dramatic in the complete GCM. Key to this is the east Pacific, where large interannual variations whose local correlations are strongly positive sway the zonal average values, but not the zonally averaged correlation. Downplaying the influence of the east Pacific allows for the midtropospheric minimum and the insensitivity to sampling seen in the GCM with Fig. 4. The influence of the 1982/83 El Niño is reduced in a similar way. Nevertheless, the biases associated with objective analysis discussed earlier are only diminished, but not eliminated, by the method of averaged correlations.

c. Correlation maps

Figures 9 and 10 present correlation maps representative of the lower, mid-, and upper troposphere (950, 700, and 300 hPa). Figure 11 presents correlation maps as in Figs. 9 and 10 but for interpolated values. Gen-

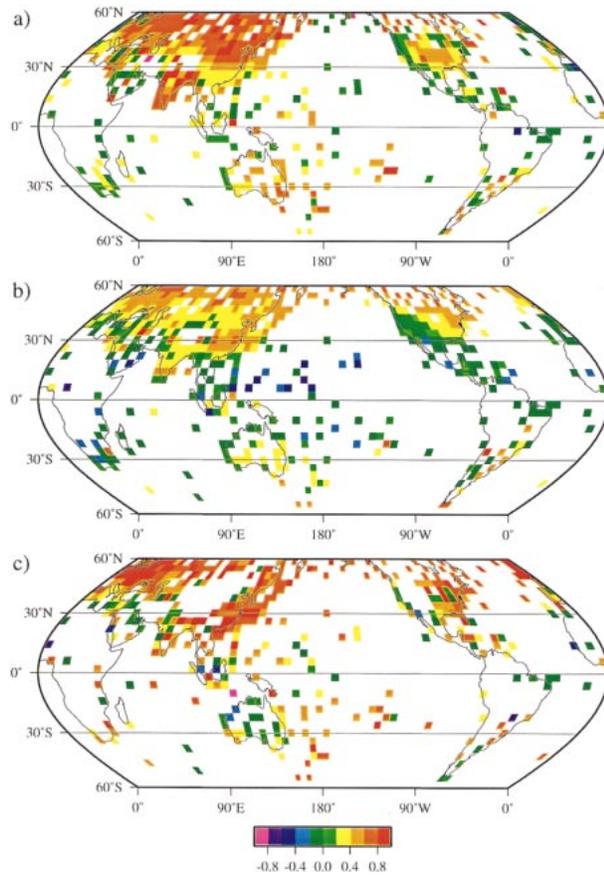


FIG. 9. Correlation coefficient maps for the interannual variations of temperature and specific humidity (having at least 10 valid monthly values) on constant pressure surfaces. Shown here are the sampled observations: (a) 300, (b) 700, and (c) 950 hPa.

erally, local T - q correlation exhibits large spatial and vertical variability. This, coupled with the station coverage shown in Fig. 1, argues that sampling must be an important consideration for studies using radiosonde data, even for those appealing to tropical averages.

The GCM diverges greatly from the observations in some locations. In the lower troposphere, the data are less correlated and more variable at places where correlation in the GCM is consistently strong and positive (Figs. 9c and 10c). This difference cannot be explained by our sampling method, that is, because of the incompleteness of the resulting time series in gcm_s (not shown). Indeed, generally gcm_f equals gcm_s at the same locations as indicated for obs_s . But then, this focus on the very local may strain the assumption that the radiosonde data and GCM-grid values represent similar things. For instance, GCM grids are not point samples. These GCM grids also capture the complete diurnal cycle rather than only 0000 UTC and do not suffer from intramonth sampling effects. Besides these sampling issues, we note that GCM grids likely lack meaningful local features such as islands, steep topography, and

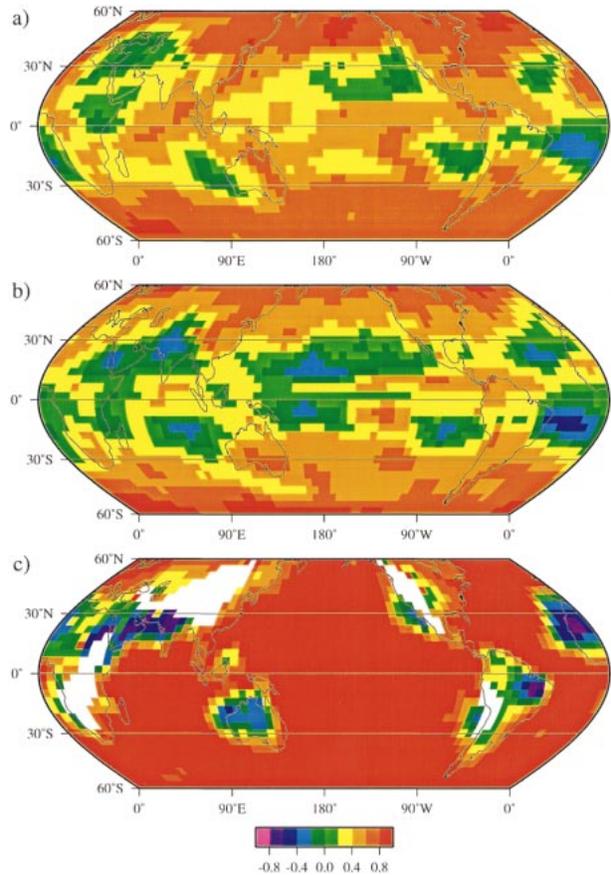


FIG. 10. As in Fig. 9, but for the complete GISS GCM: (a) 338, (b) 635, and (c) 959 hPa.

mesoscale weather events. All of these things may bring the issue of data quality and representativeness to prominence. In any case, objective analysis cannot reconstruct the actual local correlation field over the poorly sampled GCM oceans (Figs. 10c vs 11b). This comes about because the large spatial inhomogeneities in both T and q at this level cannot be fully recovered from the sparse sample. Instead, external information from well-sampled continental regions with differing T - q correlation are imposed on unobserved nearby ocean locations. This is likely to happen with the interpolated data as well.

The midtroposphere also contains many model-data differences, most particularly over India and in the west Pacific (Figs. 9b and 10b). The high degree of inhomogeneity in these figures partly explains why sampling and interpolation have their greatest impact at this level. However, inconsistencies in the radiosonde data values themselves may contribute to this inhomogeneity, and perhaps for some of the model-data differences as well. For instance, Gaffen (1992) notes numerous instrumental changes in the 1980s that might affect the continuity of humidity measurements taken over Australia (mainland plus islands). Moreover, work in progress by one

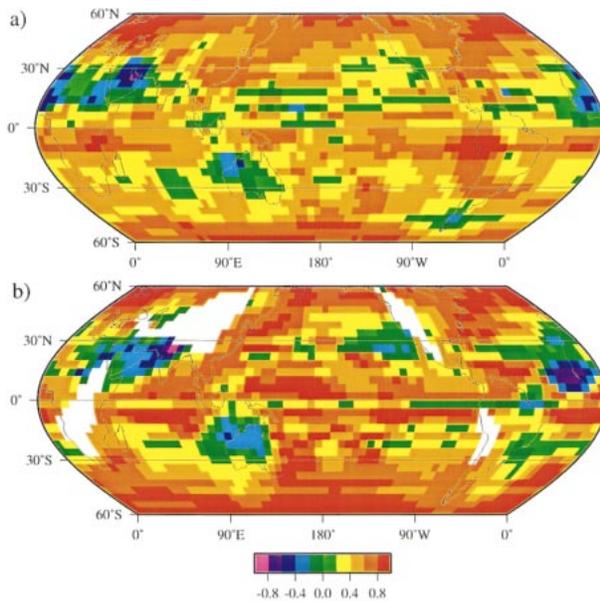


FIG. 11. As in Figs. 9c and 10c, but for interpolated values: (a) the observations (950 hPa) and (b) the GISS GCM (959 hPa).

of us (JRL) finds large spurious trends in tropospheric temperature at several Indian stations during the 1980s. Gaffen (1994) also reports problems with temperatures from Australian and Indian stations. Indeed, Gandin et al. (1993, their Table 3) found Indian radiosonde reports the most error-prone worldwide. Data from the western tropical Pacific are much less problematic in comparison (Gandin et al. 1993). Objective analysis also contributes to the model–data differences in the midtroposphere. For instance, objective analysis underestimates the prevalence of weak-to-negative correlation in the GCM midtroposphere. However, the sample does the same. That is, regions containing negative correlation in the GCM are poorly sampled. This suggests that the relative spatial offset of the GCM climatology with respect to the observing network introduces another uncertainty into the comparison. For example, the impression of widespread weak-to-negative correlation in the interpolated GCM can be greatly enhanced (not shown) by simply exchanging the time series of a Hawaiian radiosonde location with GCM output from a grid box farther to the southwest (one that is not included in the radiosonde sample). Nonetheless, the complete GCM appears more positively correlated in the Tropics than is observed, and even where negative correlation exists in the GCM it is often weak and isolated.

In contrast, the upper troposphere is remarkably alike for all versions of the GCM and the data. Apparently, the radiosonde network is adequate to capture the relatively spatially invariant fields of T and q at this level. Over well-sampled continental regions (United States, Argentina, Australia, South Africa), the correlations are slightly lower in the data than in the GCM (Figs. 9a vs 10a). However, any disagreement (agreement) with the

data is as likely to be explained by inadequacies in radiosonde humidity sensors, or reporting practices, at cold temperatures as by any defect in the GCM (Elliott and Gaffen 1991; Soden and Lanzante 1996).

4. Discussion

Radiosonde observations of the tropical troposphere are fragmentary and widely scattered. Reasonable efforts to account for this must be made if comparisons with global climate models are to be equitable (e.g., Jones et al. 1997; Kidson and Trenberth 1988; Oort, 1978). By doing this, we find that the correlation of tropical average temperature and specific humidity is closer to what is observed than if the complete GCM fields are used instead. Primarily, this is because the complete GCM includes information from the tropical eastern Pacific that a radiosonde-like sample does not.

Previous authors used a different tactic to make similar comparisons with GCMs; they used an objective analysis scheme (ANAL95) to interpolate the sparse radiosonde data onto a GCM-like grid (Sun and Held 1996; Sun and Oort 1995; Sun et al. 2001). We have demonstrated that this solution is unsatisfactory. Generally, we find that objective analysis lowers T – q correlations relative to those in the actual data and greatly exaggerates model–data discrepancies. There are several reasons for this systematic error. First, objective analysis assigns values from isolated stations to vast areas of the Tropics (see Raval et al. 1994, their Fig. 1). Second, objective analysis introduces spatially interpolated values into the incomplete time series of many tropical stations (grid cells). And third, objective analysis diffuses information across climatologically distinct boundaries, for example by importing extratropical influences into the Tropics. From this, we conclude that the objective analysis scheme ANAL95 is inadequate for this type of study and that any inferences about global climate variability drawn from such interpolated data should be regarded with caution.

It is perhaps inescapable that efforts to account for unobserved “data” will meet with such shortcomings, whether they are based on sophisticated numerical schemes (e.g., Trenberth and Caron 2001; Trenberth and Solomon 1994) or make use of physical modeling (e.g., Trenberth et al. 2001). Thus, we may be left with no choice but to compare irregular and fragmentary samples of the atmosphere and GCM. We have evaluated two methods by which this comparison has been made: correlated averages (e.g., Sun and Oort 1995) and averaged correlations (Hu et al. 2000). Generally, these methods lead to complementary conclusions. However, the method of correlated averages is sensitive to sampling and the bias of objective analysis. Moreover, correlated averages are also sensitive to whether all available information for each variable, or paired observations are used in the analysis. The method of averaged correlations inherently downplays these issues by en-

capsulating local variations into correlations before they are spatially averaged.

Sun and Held (1996) report a discrepancy between observed and simulated $T-q$ correlations of ~ 0.5 . Our attempts to account for the impact of spatial and inter-month sampling and the artifacts of analysis reduce this to ~ 0.2 . The impact of incomplete sampling within a given month of the data, which we do not account for, may explain part of the remaining difference. In the upper troposphere, where radiosonde humidity measurements are questionable, such differences are probably not significant. In the lower and midtroposphere, however, it is worth considering possible causes of the remaining GCM overestimate. Near the surface, model-data disagreements are largest over the tropical oceans (Figs. 9c vs 10c). The GCM captures the observed negatively or weakly correlated variability over most land locations, while producing uniformly strong correlations throughout the tropical ocean boundary layer (Fig. 10c). While this may point to flaws in the GCM boundary layer parameterization, it may also simply contrast inherent differences between actual radiosonde stations and the formulation of the GCM used in this analysis. For instance, so-called ocean stations in the radiosonde network are actually small islands, with at least some land surface influence on the overlying atmosphere; whereas the equivalent GCM grid boxes are truly oceanic in character, since they do not resolve islands much smaller than the grid box size. In practice, this may weaken the $T-q$ correlations over these islands relative to the surrounding ocean or all-ocean GCM grid, as surface temperature over land responds nearly instantaneously to passing weather systems whereas soil moisture deficits may not. Unfortunately, widespread, reliable and historical radiosonde-like data (i.e., with high vertical resolution) do not exist for nonisland ocean locations. The Oort (1983) archive does in fact include some weather ship data, but these ships did not report during the 1979–89 time period covered by our analysis. Nonstationary ship data are unlikely to contribute to our analysis either as the requirement that valid monthly means contain at least 10 days of data constrains their usefulness. Finally we note that the GCM used in this study, and those analyzed by SH and Sun et al. (2001), were forced by prescribed sea surface temperatures that vary smoothly on a monthly timescale. Thus, atmospheric processes that in the real world could produce temporary ~ 1 K sea surface temperature fluctuations (e.g., in the wake of passing convective systems) are not allowed to feed back on the ocean in these GCMs. This suggests that coupled ocean–atmosphere GCMs might be a more appropriate platform for testing water vapor feedback hypotheses.

On the other hand, the coarse vertical resolution of all current climate GCMs, and grid-scale noise in the vertical velocity field downwind of topography, are likely to lead to a certain amount of numerical diffusion that may explain the tendency of the models to over-

estimate correlations in the midtroposphere. The inability of the GFDL and GISS GCMs to produce the change in correlation sign during the especially strong 1982/83 ENSO event, ostensibly the result of a strengthened resolved circulation (Hadley and Walker cells), may be one example of deficiencies in their dynamical water vapor transports. Then again, this discrepancy may only suggest that the AMIP experimental design has shortcomings for model-data comparisons (K. Trenberth 2001, personal communication). For example, AMIP simulations lack the significant changes in radiative forcing that followed the 1982 eruption of El Chichón. In principle the observed SST field incorporates these forcings, but it seems unlikely that an AMIP GCM could manifest a complete atmospheric response to a perturbation like El Chichón on this basis alone (Mao and Robock 1998). Some model–data uncertainty may also be attributable to inaccuracies and other problems with the observed SST field itself (Hurrell and Trenberth 1999; Mao and Robock 1998). Finally we note that radiosonde data quality remains an unresolved issue even in the lower and midtroposphere, where the long-term continuity of radiosonde temperatures (Gaffen 1994; Gaffen et al. 2000) and humidity (Elliott and Gaffen 1991; Garand et al. 1992) cannot be ruled out as a source of some model–data differences.

The unstated premise underlying studies like SO and SH is that ENSO variations (the dominant contributor to tropical variability in the time series analyzed) are a useful proxy for long-term anthropogenic climate change. Equating conclusions reached from ENSO variations with implications for global climate change at the very least necessitates isolating those aspects of ENSO that vary over the largest spatial scales (net tropical changes) from those that reflect a simple spatial redistribution within the tropical domain (Lau et al. 1996). At first glance, the method of correlated averages seems better suited to this goal (Sun and Oort 1995). In practice, this is not the case for several reasons. First, the radiosonde network provides a sparse and west Pacific biased view of ENSO (Soden and Lanzante 1996). Correlation maps suggest large-scale horizontal and vertical structure in the ENSO signal. These correlation features appear well represented as large-scale averages, thus suggesting that the method of averaged correlations is a robust alternative to correlated averages. Second, water vapor feedback depends on the highly nonlinear relationship between humidity and longwave radiation, but tropical average humidity fluctuations take place in a climate regime characterized by both high and low extreme relative humidity values (that vary both spatially and temporally), thus these variations cannot be translated into tropical average perturbations in outgoing longwave radiation in any straightforward fashion.

It is unlikely, though, that ENSO is a good proxy for long-term anthropogenic climate change in any case. During ENSO, equatorial sea surface temperature becomes more zonal and the meridional gradient increases,

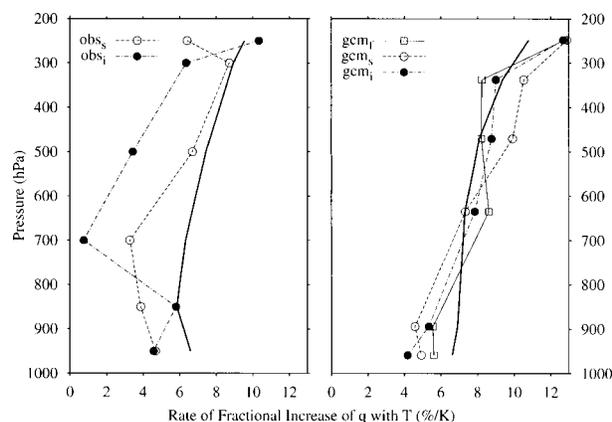


FIG. 12. Rate of fractional increase in tropical average specific humidity $\langle q_a \rangle$ with temperature $\langle T_a \rangle$ as a function of pressure. Following Sun and Oort (1995), this rate is estimated by regressing the ratio of the anomaly and time-average humidity $\langle q_a \rangle / \langle \bar{q} \rangle$ against the temperature anomaly $\langle T_a \rangle$. The observations are shown on the left and the GISS GCM on the right. Each curve represents one of three treatments: a solid line (open square) for the full GCM, a dashed line (open circle) for products of sampling, and a dash-dot line (filled circle) for products of objective analysis. The solid-thick lines (without symbols) give a reference rate based on the assumption of fixed relative humidity.

which leads to increased diabatic heating of the equatorial atmosphere, a stronger Hadley circulation (Pan and Oort 1983) and a decrease in subtropical humidity. In contrast, for simulations of long-term climate change the meridional component of the surface temperature gradient remains fairly constant in the Tropics, and the Hadley cell changes little or even weakens, because enhanced upper-level equatorial warming by moist convection allows the Hadley cell to export more moist static energy poleward without a change in the strength of the circulation itself (Yao and Del Genio 1999). Thus, subtropical humidity increases as the existing tropical circulation transports along a stronger water vapor gradient (Del Genio et al. 1991). In any case, the changes in radiative forcing associated with anthropogenic climate change, and the resulting shifts in the global hydrologic cycle, have no obvious counterparts in current climate variability.

Finally, such model–data differences as do exist may not be large enough to suggest that the water vapor feedback in GCMs is much too strong even if the differences are taken at face value. Consider for instance that SH estimated only a 15% reduction in total warming if the GFDL GCM had the same rate of fractional increase of specific humidity with temperature as is observed. Figure 12 shows that objective analysis also greatly underestimates the observed “water vapor feedback” defined in this fashion. That is, both the sampled data and the GISS GCM appear to act as nearly fixed relative humidity atmospheres under ENSO variations (see also Fig. 12 of Lanzante 1996). Thus, even if ENSO has some limited value as a proxy for global climate change, and this type of analysis is appropriate for un-

derstanding water vapor feedback, then at best we can say that GCMs are in broad agreement with observations, and at worst, that the data provide an inconclusive test for the GCMs.

Acknowledgments. The authors thank the reviewers for their many constructive comments. We especially thank Kevin Trenberth for sharing some of the results of his current research concerning the impact of radiosonde data quality (particularly in terms of its consistency with height) on correlation-based model–data comparisons and for making available to us an in-press manuscript on objective analysis limitations. We also thank William Ingram for his comments on an early draft of this paper. This research was supported by the DOE Atmospheric Radiation Measurement Program and the NASA Global Modeling and Analysis Program.

REFERENCES

- Bates, J., X. Wu, and D. Jackson, 1996: Interannual variability of upper-tropospheric water vapor band brightness temperature. *J. Climate*, **9**, 427–438.
- Cess, R., and Coauthors, 1990: Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *J. Geophys. Res.*, **95** (D10), 16 601–16 615.
- Del Genio, A., and M.-S. Yao, 1993: Efficient cumulus parameterization for long-term climate studies: The GISS scheme. *The Representation of Cumulus Convection in Numerical Models of the Atmosphere*, Meteor. Monogr., No. 46, Amer. Meteor. Soc., 181–184.
- , A. Lacis, and R. Ruedy, 1991: Simulations of the effect of a warmer climate on atmospheric humidity. *Nature*, **351**, 382–385.
- , W. J. Kovari, and M.-S. Yao, 1994: Climatic implications of the seasonal variation of upper troposphere water vapor. *Geophys. Res. Lett.*, **21**, 2701–2704.
- , M.-S. Yao, W. Kovari, and K.-W. Lo, 1996: A prognostic cloud water parameterization for global climate models. *J. Climate*, **9**, 270–304.
- Dunlap, W., M. Jones, and A. Bittner, 1983: Average correlations vs. correlated averages. *Bull. Psychonomic Soc.*, **21**, 213–216.
- Elliott, W. P., and D. J. Gaffen, 1991: On the utility of radiosonde humidity archives for climate studies. *Bull. Amer. Meteor. Soc.*, **72**, 1507–1520.
- Freedman, D., R. Pisani, and R. Purves, 1978: *Statistics*. Norton, 506 pp.
- Gaffen, D., 1992: Observed annual and interannual variations in tropospheric water vapor. NOAA Tech. Memo. ERL-ARL 198, Silver Spring, MD, 162 pp. [Available from NOAA Air Resources Laboratory, 1315 East-West Highway, Silver Spring, MD 20910.]
- , 1994: Temporal inhomogeneities in radiosonde temperature records. *J. Geophys. Res.*, **99** (D2), 3667–3676.
- , M. Sargent, R. Habermann, and J. R. Lanzante, 2000: Sensitivity of tropospheric and stratospheric temperature trends to radiosonde data quality. *J. Climate*, **13**, 1776–1796.
- Gandin, L., L. Morone, and W. Collins, 1993: Two years of operational comprehensive hydrostatic quality control at the National Meteorological Center. *Wea. Forecasting*, **8**, 57–72.
- Garand, L., C. Grassotti, J. Hallé, and G. L. Klein, 1992: On differences in radiosonde humidity reporting practices and their implications for numerical weather prediction and remote sensing. *Bull. Amer. Meteor. Soc.*, **73**, 1417–1423.
- Gates, W., and Coauthors, 1999: An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bull. Amer. Meteor. Soc.*, **80**, 29–56.

- Gutzler, D. S., 1993: Uncertainties in climatological tropical humidity profiles: Some implications for estimating the greenhouse effect. *J. Climate*, **6**, 978–982.
- Houghton, J. T., C. Jenkins, and J. Ephraums, Eds., 1990: *Climate Change: The IPCC Scientific Assessment*. Cambridge University Press, 365 pp.
- Hu, H., R. Oglesby, and B. Saltzman, 2000: The relationship between atmospheric water vapor and temperature in simulations of climate change. *Geophys. Res. Lett.*, **27**, 3513–3516.
- Hurrell, J. W., and K. E. Trenberth, 1999: Global sea surface temperature analysis: Multiple problems and their implications for climate analysis, modeling, and reanalysis. *Bull. Amer. Meteor. Soc.*, **80**, 2661–2678.
- Inamdar, A. K., and V. Ramanathan, 1994: Physics of greenhouse effect and convection in warm oceans. *J. Climate*, **7**, 715–731.
- Jones, P. D., T. J. Osborn, and K. R. Briffa, 1997: Estimating sampling errors in large-scale temperature averages. *J. Climate*, **10**, 2548–2568.
- Kidson, J., and K. Trenberth, 1988: Effects of missing data on estimates of monthly mean general circulation statistics. *J. Climate*, **1**, 1261–1275.
- Lanzante, J. R., 1996: Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *Int. J. Climatol.*, **16**, 1197–1226.
- Lau, K.-M., C.-H. Ho, and M.-D. Chou, 1996: Water vapor and cloud feedback over the tropical oceans: Can we use ENSO as a surrogate for climate change? *Geophys. Res. Lett.*, **23**, 2971–2974.
- Lindzen, R., 1990: Some coolness concerning global warming. *Bull. Amer. Meteor. Soc.*, **71**, 288–299.
- Luers, J., and R. E. Eskridge, 1998: Use of radiosonde temperature data in climate studies. *J. Climate*, **11**, 1002–1019.
- Manabe, S., J. Smagorinsky, and R. F. Strickler, 1965: Simulated climatology of a general circulation model with a hydrologic cycle. *Mon. Wea. Rev.*, **93**, 769–798.
- Mao, J., and A. Robock, 1998: Surface air temperature simulations by AMIP general circulation models: Volcanic and ENSO signals and systematic errors. *J. Climate*, **11**, 1538–1552.
- Oort, A., 1978: Adequacy of the rawinsonde network for global circulation studies tested through numerical model output. *Mon. Wea. Rev.*, **106**, 174–195.
- , 1983: *Global Atmospheric Circulation Statistics, 1958–1973*. NOAA Professional Paper 14, 180 pp.
- , and H. Liu, 1993: Upper-air temperature trends over the globe, 1958–1989. *J. Climate*, **6**, 292–307.
- Pan, Y., and A. H. Oort, 1983: Global climate variations connected with sea surface temperature anomalies in the eastern equatorial Pacific Ocean for the 1958–73 period. *Mon. Wea. Rev.*, **111**, 1244–1258.
- Prather, M., 1986: Numerical advection by conservation of second order moments. *J. Geophys. Res.*, **91**, 6671–6681.
- Raval, A., and V. Ramanathan, 1989: Observational determination of the greenhouse effect. *Nature*, **342**, 758–761.
- , A. Oort, and V. Ramaswamy, 1994: Observed dependence of outgoing longwave radiation on sea surface temperature and moisture. *J. Climate*, **7**, 807–821.
- , ———, and J. Lanzante, 1996: A new interpolation routine (“ANAL95”) for obtaining global fields from irregularly spaced meteorological data. Geophysical Fluid Dynamics Laboratory report, 25 pp. [Available from mbauer@giss.nasa.gov.]
- Rind, D., 1999: Water vapor feedback. *Frontiers in the Science of Climate Modeling: A Symposium in Honor of Professor Robert D. Cess*, Cambridge University Press, in press.
- , E. W. Chiou, W. Chu, J. Larsen, S. Oltmans, J. Lerner, M. P. McCormick, and L. McMaster, 1991: Positive water vapour feedback in climate models confirmed by satellite data. *Nature*, **349**, 500–503.
- Soden, B. J., and R. Fu, 1995: A satellite analysis of deep convection, upper-tropospheric humidity, and the greenhouse effect. *J. Climate*, **8**, 2333–2351.
- , and J. R. Lanzante, 1996: An assessment of satellite and radiosonde climatologies of upper-tropospheric water vapor. *J. Climate*, **9**, 1235–1250.
- Stephens, G. L., and T. J. Greenwald, 1991: The Earth’s radiation budget and its relation to atmospheric hydrology. Part 1. Observations of the clear sky greenhouse effect. *J. Geophys. Res.*, **96** (D8), 15 311–15 324.
- Sun, D.-Z., and A. Oort, 1995: Humidity–temperature relationships in the tropical troposphere. *J. Climate*, **8**, 1974–1987.
- , and I. M. Held, 1996: A comparison of modeled and observed relationships between interannual variations of water vapor and temperature. *J. Climate*, **9**, 665–675.
- , C. Covey, and R. Lindzen, 2001: Vertical correlations of water vapor in GCMs. *Geophys. Res. Lett.*, **28**, 259–262.
- Trenberth, K. E., 1995: Truncation and use of model-coordinate data. *Tellus*, **47A**, 287–303.
- , and A. Solomon, 1994: The global heat balance: Heat transports in the atmosphere and ocean. *Climate Dyn.*, **10**, 107–134.
- , and J. M. Caron, 2001: Estimates of meridional atmosphere and ocean heat transports. *J. Climate*, **14**, 3433–3443.
- , D. P. Stepaniak, and J. W. Hurrell, 2001: Quality of reanalyses in the Tropics. *J. Climate*, **14**, 1499–1510.
- Waliser, D., Z. Shi, J. Lanzante, and A. Oort, 1999: The Hadley circulation: Assessing NCEP/NCAR reanalysis and sparse in-situ estimates. *Climate Dyn.*, **15**, 719–735.
- Yang, H., and K. K. Tung, 1998: Water vapor, surface temperature, and the greenhouse effect—A statistical analysis of tropical-mean data. *J. Climate*, **11**, 2686–2697.
- Yao, M.-S., and A. Del Genio, 1999: Effects of cloud parameterization on the simulation of climate changes in the GISS GCM. *J. Climate*, **12**, 761–779.