

Sampling Errors in Rawinsonde-Array Budgets

BRIAN E. MAPES

NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado

PAUL E. CIESIELSKI AND RICHARD H. JOHNSON

Department of Atmospheric Sciences, Colorado State University, Fort Collins, Colorado

(Manuscript received 18 May 2002, in final form 19 May 2003)

ABSTRACT

Rawinsonde data used for sounding-array budget computations have random errors, both instrumental errors and errors of representativeness (here called sampling errors). The latter are associated with the fact that radiosondes do not measure large-scale mean winds and state variables, but are contaminated by small-scale variations as well. Data from the western Pacific and the summer monsoon of southeast Asia are used to estimate these random errors, and to propagate them through budget computations to assign error bars to derived quantities.

The statistics of sampling errors in directly measured variables are estimated from station pair analysis, in which variance is partitioned into contributions by resolved and unresolved scales. Resolved scales contribute the portion that is contained in averages of adjacent sounding stations and/or adjacent launch times (6-h intervals), while the rest of the total variance is defined as unresolved. Magnitudes of unresolved variability for typical rawinsonde-array spacings are ~ 0.5 K for temperature; $\sim 5\%$ for relative humidity at low levels, rising to nearly 15% in the middle-upper troposphere; and ~ 2 m s⁻¹ for winds, rising to 3 m s⁻¹ in the upper troposphere. These are much larger than random instrumental errors, as estimated from pairs of simultaneous rawinsondes launched very close together. Vertical correlation scales of unresolved variability are 100–200 hPa. Up to 50% of the variance of humidity is unresolved, while for zonal wind the unresolved portion is only a few percent. Spatial and temporal sampling errors become about equal for 6-hourly rawinsondes ~ 200 km apart.

The effects of sampling errors on budget computations are estimated by a perturbed-observation ensemble approach. All computations are repeated 20 times, with random realizations of unresolved variability added to the rawinsonde data entering the analysis. The ensemble standard deviation serves as an estimate of sampling error, which naturally decreases as the results are averaged over larger areas and longer time periods. For example, rainfall estimates on ~ 500 km scales have sampling errors of ~ 5 mm day⁻¹ in daily means, and ~ 1 mm day⁻¹ in monthly means. The ensemble spread of 120-day time integrations of the vertically averaged moist enthalpy equation with rawinsonde-array-derived advective sources exceeds 20 K, implying that sampling error could be responsible for substantial biases in column models forced with such source terms.

1. Introduction

A rawinsonde makes measurements of temperature, humidity, wind, and pressure averaged over just meters in space and seconds in time. Meanwhile, an array or network of rawinsonde stations only resolves horizontal scales (half-wavelengths) larger than the interstation spacing (typically hundreds of kilometers) and temporal scales larger than the launch interval (typically 6 or 12 h). The space scales and timescales in between contain vigorous phenomena including vertical convection, mesoscale circulations, gravity waves, and fine structure produced by deformation. Phenomena at these scales

contribute both important signals and confounding noise to rawinsonde-array budget computations.

The study of tropical convection with rawinsonde arrays has been mainly based on evaluating source and sink terms (residuals) in large-scale budget equations for heat and moisture [for a review, see Yanai and Johnson (1993)]. Momentum budget residuals are also sometimes computed from these data, albeit with greater uncertainty, owing to the larger number of terms involved in the budget and their propensity for cancellation. In constructing these large-scale budget estimates, an analysis is made in which point measurements from rawinsondes are taken as representative of a space-time region comparable to the distance between soundings. In this analysis process, a random representativeness error is incurred, along with a random instrumental error. These errors cannot be avoided and have, to our knowledge, never been systematically estimated in the context of budget computations.

Corresponding author address: Brian Mapes, NOAA-CIRES Climate Diagnostics Center, 325 S. Broadway, R/CDC1, Boulder, CO 80305.
E-mail: mapes@colorado.edu

Input data errors propagate through calculations of multiple-term budget equations in complex ways. For example, consider the gridded approach to the computation of rainfall rate in a moisture budget. Wind data are analyzed onto a grid and differentiated in the horizontal to yield wind divergence, which is integrated in the vertical to yield vertical velocity, subject to a non-local mass-continuity adjustment. This vertical velocity is multiplied by the vertical derivative of specific humidity. The resulting vertical advection estimate is added to horizontal advection and time tendency estimates to yield the apparent moisture sink. The apparent moisture sink is then integrated in the vertical and assumed equal to precipitation minus evaporation. Finally, an independent evaporation estimate is needed, for example from bulk aerodynamic formulae, to determine precipitation. Afterward, the resulting estimates are averaged in space and time. Errors tend to increase with differentiation and multiplication, and decrease with integration and averaging, but it is not easy to attach error bars to the resulting budget estimates simply by intuition or inspection.

Rawinsonde budget calculations also have other sources of error, including instrument biases (Wang et al. 2002), systematic errors associated with network geometry (Haertel 2002), and uncertainties in analysis assumptions (Molinari and Skubis 1988). To the extent that systematic errors can be identified, they can typically be removed or reduced. Random errors tend to reflect more fundamental uncertainties. An estimate of random errors is useful, both for warning users and interpreters about product limitations, and for informing data analysts about how much elaboration is worthwhile in input data quality control, analysis system refinements, etc. These errors may also be informative for designing future measurement arrays, or for constraining the choices made in variational adjustments to rawinsonde array data to optimize budget balances (Zhang and Lin 1997).

In order to estimate the role of random errors in rawinsonde-array budgets, we take a brute-force computational approach, generating an explicit ensemble of analyses. First, we estimate the random errors in rawinsonde array data by dividing the total variance of a set of soundings into two parts: a “resolved” part, reflecting fluctuations with space scales and timescales larger than the array spacing and launch interval; and the remainder, the “unresolved” part. The unresolved part includes both accurately measured but undersampled variance on scales too small to be resolved by the array, and other errors such as nonconstant instrumental errors. The variance decomposition analysis is described in section 3. Next, we perturb the radiosonde data going into budget computations, by adding realizations of random noise that statistically resemble the unresolved variability, as described in section 4. We then run identical data analysis and budget computation procedures using the original and several perturbed datasets, generating

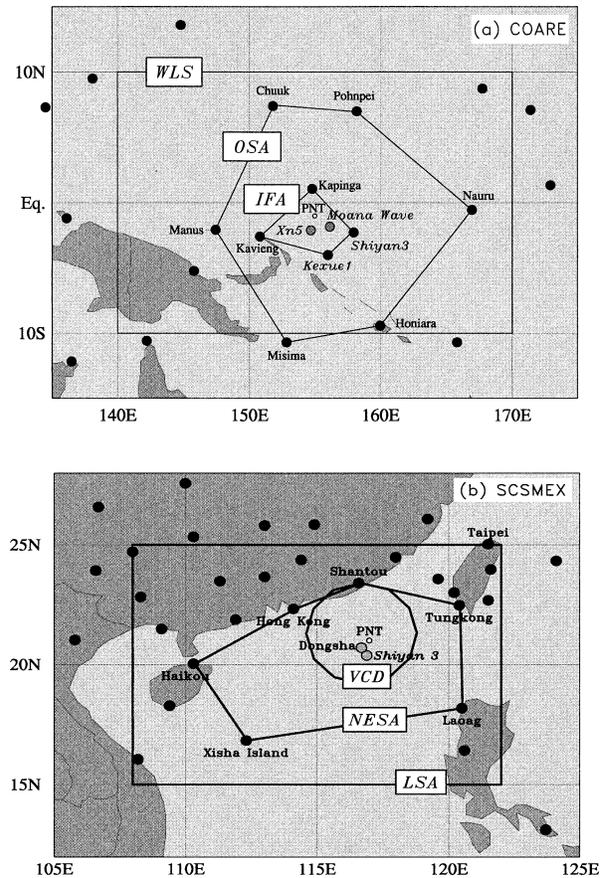


FIG. 1. Base map for sounding arrays: (a) Tropical Ocean Global Atmosphere (TOGA) COARE; (b) SCSMEX.

an ensemble of estimates for all budget quantities (section 5). Section 6 contains a concluding discussion.

2. Data

Figure 1a shows the Coupled Ocean–Atmosphere Response Experiment (COARE; Webster and Lukas 1992) rawinsonde array in the tropical western Pacific during November 1992–February 1993. The distance between sounding sites varies from over 1000 km in the Outer Sounding Array (OSA) to <200 km in the densest area of the Intensive Flux Array (IFA). We examine data from this densest region, the ships of the eastern IFA (gray dots in Fig. 1a). The resolved–unresolved separation of variability in these data will give a relatively conservative estimate of the amount of unresolved variability in the much sparser OSA. The use of ship-based soundings minimizes the distortions associated with local effects near land.

There were three main time periods in COARE (“cruises”) during which more than one ship was on station. The ships *Shiyun 3* and *Kexue 1* (also called xp3 and sc1, respectively, in this paper) offer the longest period of joint sampling, covering most of the three

cruises, so our COARE analysis focuses on them. These ships were about 200 km apart. Since the goal of the present study is to obtain trustworthy statistics, and not to extract all possible detailed information from the dataset, the sounding data were subject to stringent quality control. The data were obtained on 5-hPa pressure levels, updated with the latest humidity corrections (Wang et al. 2002). Some quality control was done by Zuidema (1998), from whom the data other than corrected humidity were obtained. Quality conditions used here included rejection of all data flagged as questionable and rejection of outliers beyond three standard deviations from the mean of each variable at each pressure level. For the synthetic noise in section 4, wind soundings were supplemented with information from wind profilers (Ciesielski et al. 1997) to overcome the lack of wind measurements at low levels by the now-defunct Omega navigation system used on the COARE rawinsondes.

In order to distinguish instrumental error from small-scale space–time variability, it is useful to examine data from very closely spaced rawinsonde launch sites. Toward this end, and to assess the applicability of COARE results to other tropical regions and weather regimes, data from two other research deployments of rawinsondes have also been analyzed. The South China Sea Monsoon Experiment (SCSMEX) included a special sounding network (Fig. 1b) during May–June of 1998 (Johnson and Ciesielski 2002). This network included soundings from two sites less than 50 km apart (the *Shiyan 3* ship and Dongsha Island, near 20°N, 117°E). These data, at 6-h intervals, were interpolated to 25-hPa vertical levels and were discarded if interpolation or extrapolation spanned pressure ranges exceeding 200 hPa. Outliers beyond three standard deviations were also rejected. The Nauru99 experiment (Westwater et al. 2003) included rawinsonde launches from Nauru (a small, flat island near the equator at 167°E, seen on Fig. 1a) and the R/V *Ronald H. Brown*, located immediately offshore (within 10 km) from 4 July to 16 July 1999. These soundings were interpolated to 5-hPa pressure levels and binned to 6-hourly time categories, and again outliers were rejected.

The Colorado State University (CSU) objective analysis system (Johnson and Ciesielski 2000, 2002) was used to interpolate the sounding data from the COARE and SCSMEX networks to a regular 1° horizontal grid at equally spaced 25-hPa vertical levels, 4 times a day (at 0000, 0600, 1200, and 1800 UTC nominal launch times), using a multiquadric analysis scheme (Nuss and Titley 1994). Each field (zonal and meridional wind, temperature, and specific humidity) at each pressure level was analyzed independently at each hour, with no first-guess or other background fields. Budget computations were then performed using centered differencing. Budget results were examined from the grid points labeled PNT in Figs. 1a and 1b, as well as from averages over the areas labeled IFA, OSA, and WLS for COARE and VCD, NESAs, and LSA in SCSMEX.

3. Resolved and unresolved variances

The variance of a set of data values is the mean square of their difference from the mean. The variance of smoothed or filtered (internally averaged) versions of the same dataset is necessarily smaller. In the present application, we consider rawinsonde data as a set of point samples taken at a sequence of measurement times (nominally 6 h apart here), from pairs of stations some distance apart. We define the *fully resolved* variance, that is, the variance of fluctuations that are large-scale with respect to the station spacing and measurement interval, as the variance of averages over the two stations, further averaged at two adjacent measurement times. *Unresolved* variance is the difference between the total variance and this fully resolved variance. Intermediate definitions are also useful: *spatially unresolved* and *temporally unresolved* variance.

To express these quantities mathematically, we follow the notation of Neter et al. (1990), and indicate the set of values of variable Y (temperature, humidity, or wind) at some particular pressure level as Y_{xst} . Here, the three subscripts indicate a three-dimensional indexing of the dataset, with two spatial locations (x), two adjacent time levels (t), and the hundred or so samples (s). An average over one or more dimension is indicated by replacement of the corresponding subscript(s) with a dash. The sample variance of a set of data values is indicated as a function $s^2(\cdot)$. Then we have

total variance:	$s^2(Y_{xst})$,
fully resolved variance:	$s^2(Y_{-st})$,
unresolved variance:	$s^2(Y_{xst}) - s^2(Y_{-st})$,
spatially unresolved variance (SUV):	$s^2(Y_{x-s}) - s^2(Y_{-s})$,
temporally unresolved variance (TUV):	$s^2(Y_{-ts}) - s^2(Y_{-s})$.

To clarify the meaning of these names, let us consider the example of spatially unresolved variance. As the station separation gets very small, the two measurements should converge toward the same value. For a dataset consisting of pairs of identical values, averaging each pair does not destroy any variance, so SUV will decline toward zero with vanishing station separation.¹ Conversely, the greater the station spacing, the more variance the x averaging destroys, and hence the larger the value of SUV. Similar logic holds for TUV as regards the time interval between measurements. To ensure consistency of the sample entering these different partial variances, data entering this variance analysis are restricted to situations in which all the necessary data are available to construct the fully resolved variance.

The weather during COARE included extended periods of convectively active (rainy) and inactive (dry)

¹ If there is random instrument error, SUV will not tend toward zero for vanishing measurement separation, but rather toward half the instrumental error variance.

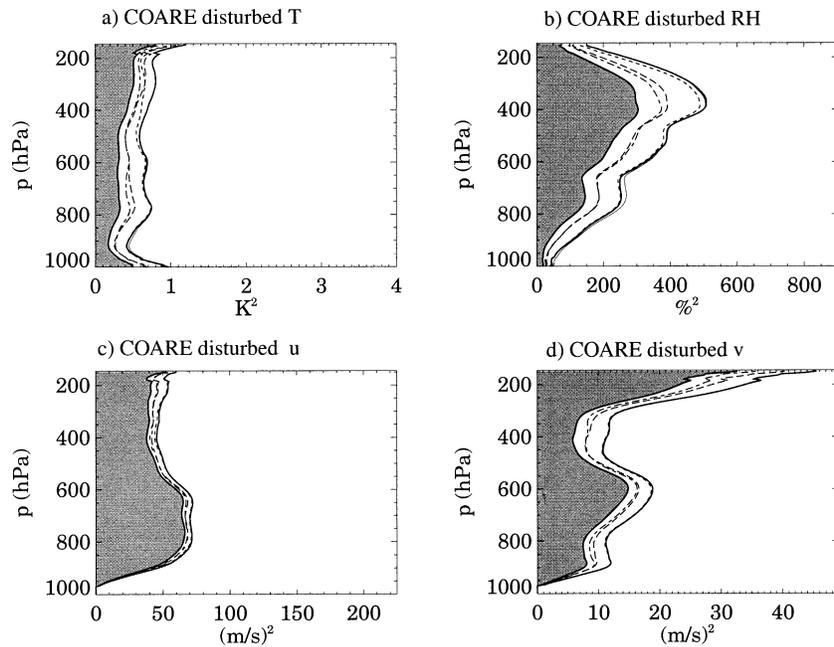


FIG. 2. Well-resolved (leftmost curve) and total (rightmost curve) variances for the COARE ship pair *Kexue 1* and *Shiyan 3*, in disturbed conditions. Broken curves show variances of various averaged and altered versions of the data, corresponding to the equivalent broken curves in Fig. 3. See text for details. (a) Temperature; (b) relative humidity; (c) zonal wind; (d) meridional wind.

weather, associated with the Madden–Julian or intra-seasonal oscillation (Lin and Johnson 1996a,b). We therefore subdivide the data into “disturbed” and “undisturbed” periods, based on visual inspection of the soundings’ upper-tropospheric relative humidity data. Upper-tropospheric humidity is well correlated with rainfall (Fig. 5 of Brown and Zhang 1997). The undisturbed periods, expressed in days after 0000 UTC 1 November 1992, are 0–20, 30–45, and 67.5–77.5. The remainder of the COARE 120-day period for which the ships were on station is considered disturbed. Lin and Johnson (1996a,b) offer additional context about COARE weather.

Figure 2 shows the variances of temperature (T), relative humidity (RH), and zonal (u) and meridional (v) wind for the *Kexue 1* and *Shiyan 3* pair of stations in disturbed conditions. The shaded region to the left of the leftmost curve indicates the well-resolved variance, while the greatest values at each level (rightmost curve) outline the total variance. The gap between these is the unresolved variance, and in that gap lie the various partially unresolved variances. The total temperature variance is less than 1 K^2 at all levels, with approximately one-half to two-thirds of the total variance being well resolved. A similar or slightly larger fraction of relative humidity variance is well resolved. The humidity variance profile has its maximum near 350 hPa, with a secondary peak near 750 and slightly enhanced unresolved variability near the freezing level (~ 550 hPa). Above the 400-hPa level, measurement problems as-

sociated with a slow sensor response time likely distort the true profile of humidity variance. The zonal wind variance is mostly well resolved, with a broad peak in the lower troposphere and artificially small (zero) values near the surface, where winds could not be reliably estimated by the now defunct Omega navigation system used by the rawinsondes. The meridional wind variance is about a quarter as great as zonal wind variance, but has a fractionally much larger unresolved component. Meridional wind variance has a strong peak near 600 hPa, also detectable in zonal wind variance.

The careful reader may notice a light line extending very slightly to the right of the heavy line in the RH and T variance plots of Fig. 2. This light line is the total variance of the set of measurements from the two stations, while the heavy line is the average of the variances calculated at each station separately. The difference between these is a measure of time-mean bias between the two stations. It is comforting that this bias is very small. Its largest magnitude is in low-level relative humidity, where *sc1* is known to have had a moist bias, of unknown origin and hence not removed in subsequent data corrections.

The broken curves on Fig. 2, reflecting partially resolved variances, can be more clearly seen on the expanded scale of Fig. 3, in which the well-resolved variance is subtracted from each corresponding curve in Fig. 2. The solid curves in Fig. 3 show the unresolved variance. In standard deviation units (the square root of variance), the magnitudes of unresolved variations are

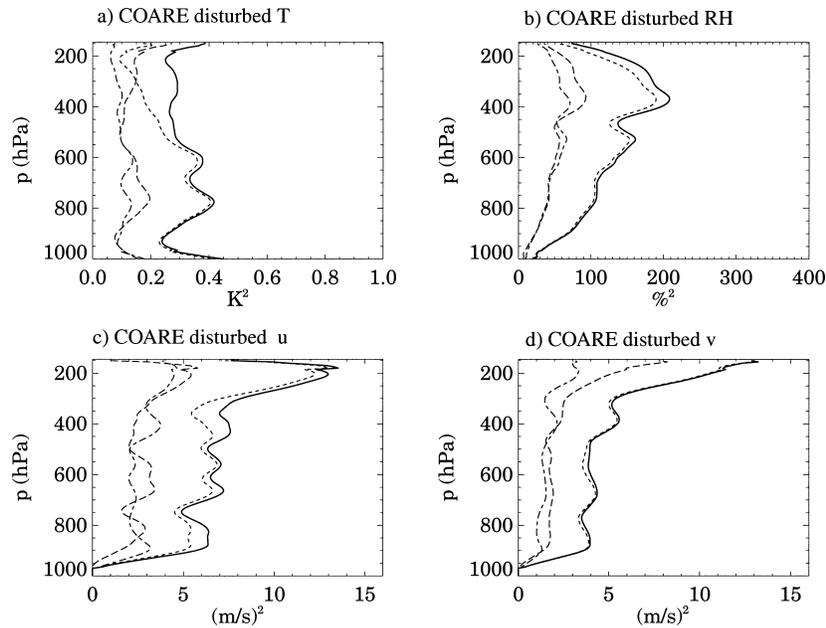


FIG. 3. Unresolved and partially resolved variances for ships *Kexue 1* and *Shiyan 3* during COARE disturbed conditions, created by subtracting off the well-resolved variance (which in this case is the shaded region in Fig. 2). See text for details. Solid line: unresolved variance. Dotted line: unresolved variance after mean diurnal cycles are removed from each station. Dashed line: temporally unresolved variance. Dashed-dotted: spatially unresolved variance. (a) Temperature, (b) relative humidity, (c) zonal wind, and (d) meridional wind.

T about 0.5 K; RH about 5% near the surface, increasing with altitude to nearly 15% near 400 hPa; and wind about 2 m s^{-1} , increasing to 3 m s^{-1} high in the upper troposphere. These values are comparable (discussed below) with those found by Sobel et al. (2004) for another tropical location. The dotted line indicates the unresolved variance calculated after subtracting the composite diurnal cycle of each variable. The closeness of this dotted line to the solid line indicates that the mean diurnal cycle is a small contributor to the total variance of any variable, with the exception of temperature in the upper troposphere (see Fig. 5 of Mapes and Zuidema 1996), where it reaches about 20% of the total variance and 50% of the unresolved variance. Radiation-induced instrument errors are always a worry and may increase with height due to sensor ventilation effects, but a real signal is also likely to exist, given that general circulation models show diurnal temperature cycles of comparable amplitude to the observations (Fig. 9 of Randall et al. 1991).

The dashed line on Fig. 3 indicates the temporally unresolved variance (TUV; variance that survives spatial averaging and hence is “large scale,” but is destroyed by time averaging), while the dashed-dotted curve is spatially unresolved variance (SUV, which persists in time averages and hence is “low frequency,” but disappears with spatial averaging). These two are roughly equal for 6-h launch intervals at these ships $\sim 200 \text{ km}$ apart, suggesting that spatial and temporal

resolution shortcomings are fairly well balanced for a space–time ratio of $200 \text{ km}/6 \text{ h} \sim 9 \text{ m s}^{-1}$. As expected, TUV is a larger fraction of the unresolved variance for 6-hourly launches much closer together in space (discussed below for SCSMEX and Nauru99 soundings), while SUV is a more dominant fraction when three-point (12 h) time averages are considered instead of two-point (6 h) averages (not shown).

Figures 4 and 5 repeat Figs. 2 and 3, but for undisturbed conditions (dry in the upper troposphere). In this case, well-resolved variance is generally smaller in all four variables, except for zonal wind in the extreme upper troposphere. Unresolved variance (Fig. 5) is slightly smaller in winds and temperature at most levels, consistent with the notion of less fast, small-scale dynamical activity such as convection. However, in relative humidity, unresolved variance is actually greater at most levels, and in fact makes up the majority of total humidity variance (Fig. 4b). Again, the diurnal cycle is small except for the temperature cycle in the upper troposphere. The spatially unresolved variances of temperature and wind components (dashed-dotted lines) are in this case smaller both in absolute terms and in relation to the temporally unresolved variances, presumably because the weaker and less-organized convection in undisturbed conditions creates less vigorous circulations at spatial scales $< 200 \text{ km}$. However, the opposite is true for humidity. In other words, undisturbed conditions are more spatially uniform than disturbed conditions in the

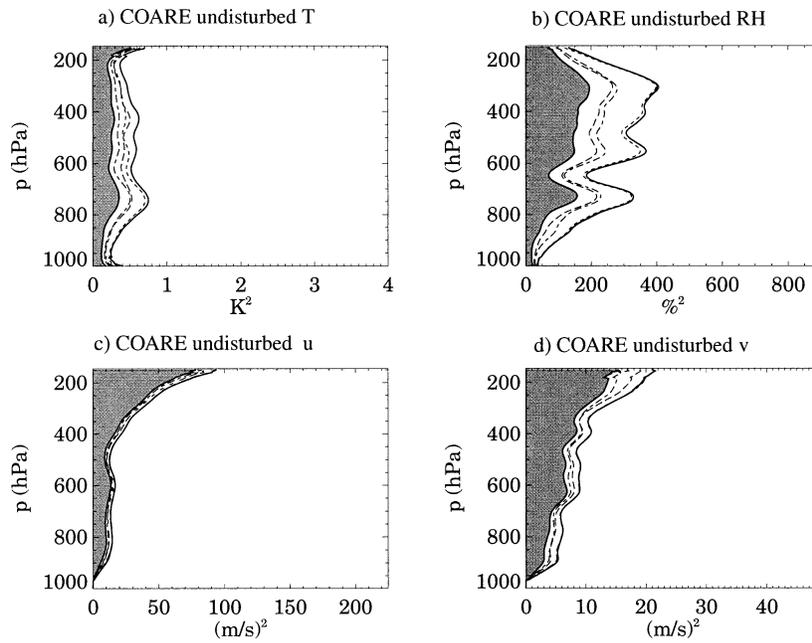


FIG. 4. As in Fig. 2, but for undisturbed conditions.

dynamic fields (wind and density) but are more spatially inhomogeneous in humidity. Perhaps this latter fact is because the weaker winds, on both large and small scales, are less efficient at homogenizing humidity through shear and mixing.

Figures 6 and 7 are also like Figs. 2 and 3, but for the two closest sounding stations in SCSMEX, Dongsha Island and *Shiyan 3*, less than 50 km apart. The data have not been separated into disturbed and undisturbed

conditions in this case, as the sample size is rather small (~ 100 or less, depending on variable and altitude). The fully resolved temperature and humidity variances (Figs. 6a,b) are both considerably greater at this higher latitude ($20^\circ N$) than in the near-equatorial COARE data, even when the COARE disturbed and undisturbed data are pooled (not shown). The SCSMEX upper-tropospheric winds (near 200 hPa) also have a larger variance, most of it fully resolved (Figs. 6c,d). Perhaps the most notable

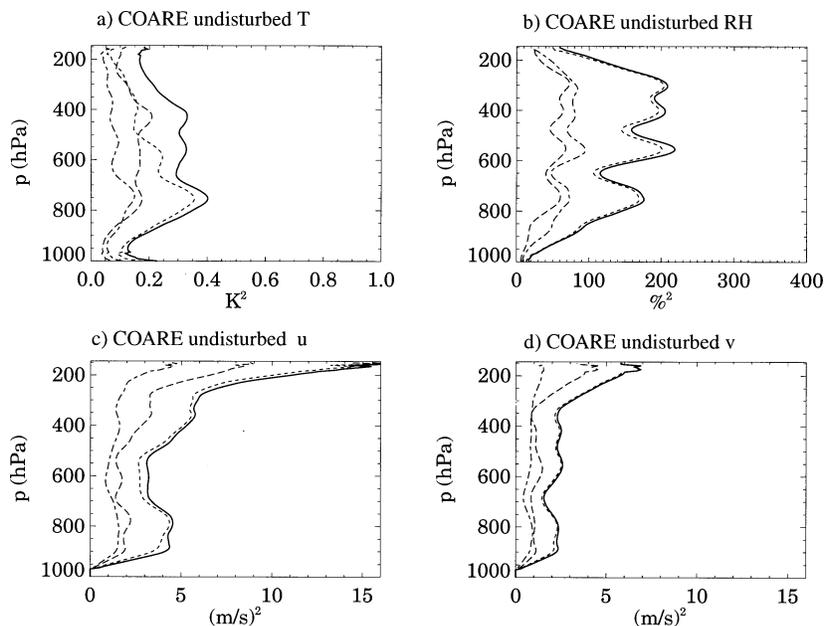


FIG. 5. As in Fig. 3, but for undisturbed conditions.

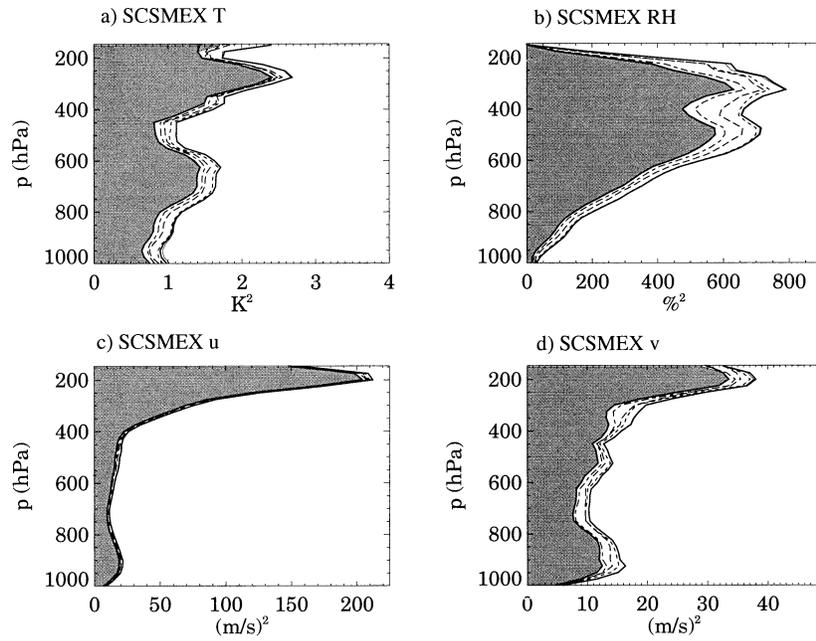


FIG. 6. As in Fig. 2, but for SCSMEX stations Dongsha and *Shiyan 3*.

difference in SCSMEX is that the unresolved variance of lower-tropospheric zonal wind (Fig. 7c) is considerably smaller even than that in COARE undisturbed conditions, and up to half of this smaller amount is apparently due to the mean diurnal cycle below the 650-hPa level (as the dotted line in Fig. 7c is far from the solid line). Otherwise, the mean diurnal cycle is a small fraction of the unresolved variance except in upper-tropospheric temperature, as in COARE. Overall, unre-

solved variance is quite similar to the COARE values: the main difference is that unresolved wind standard deviation is about 1.5 m s^{-1} rather than 2 m s^{-1} in the lower troposphere.

The other feature to note about Fig. 7 is that the spatially unresolved variance (dashed-dotted line) is especially small, much smaller than the temporally unresolved variance (dashed line). This reflects the fact that the stations are quite close together, about 40 km.

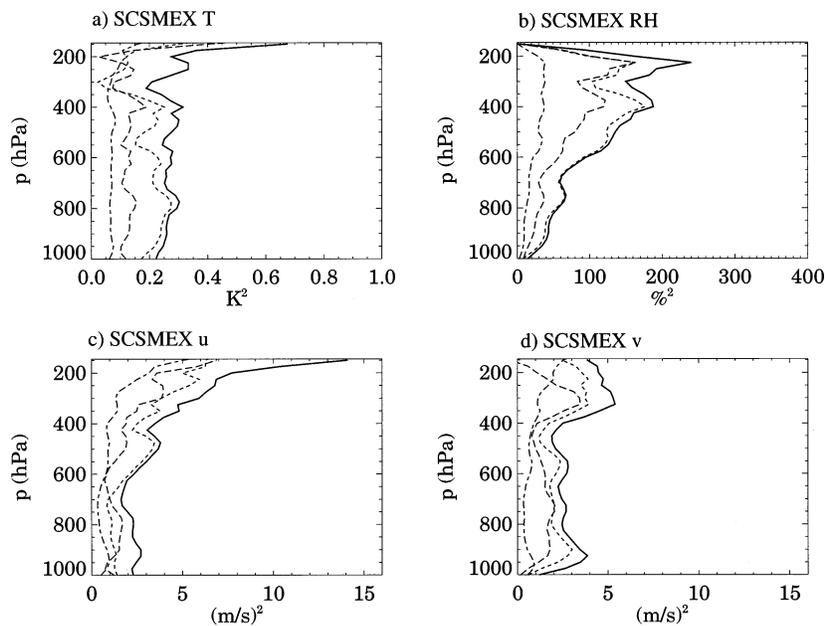


FIG. 7. As in Fig. 3, but for SCSMEX stations Dongsha and *Shiyan 3*.

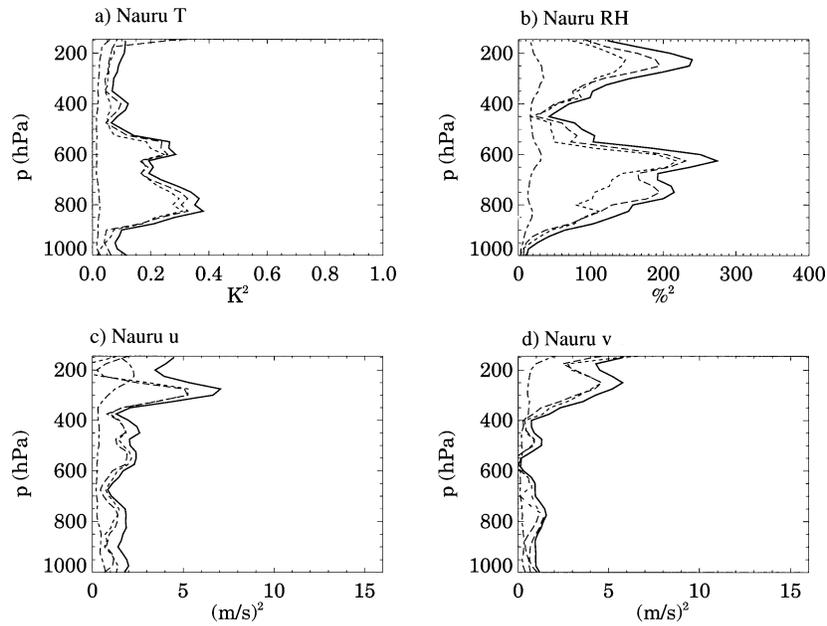


FIG. 8. As in Fig. 3, but for island and ship soundings (~ 10 km apart) in Nauru99.

An even closer pair of sounding stations is available from the Nauru99 experiment, when the National Oceanic and Atmospheric Administration (NOAA) research vessel *Ronald H. Brown* launched radiosondes within 10 km of land-based soundings from Nauru for 12 days in July 1999 (Westwater et al. 2003). The unresolved variances from this pair of stations are plotted in Fig.

8. The spatially unresolved variance (dashed-dotted curve) is very small in this case, with square root magnitudes near 0.2 K in temperature; $< 5\%$ in RH; and < 1 m s^{-1} in winds throughout the lower half of the troposphere. The smallness of these statistical measures of random instrument error indicates that we can interpret the large majority of the unresolved variance in typical

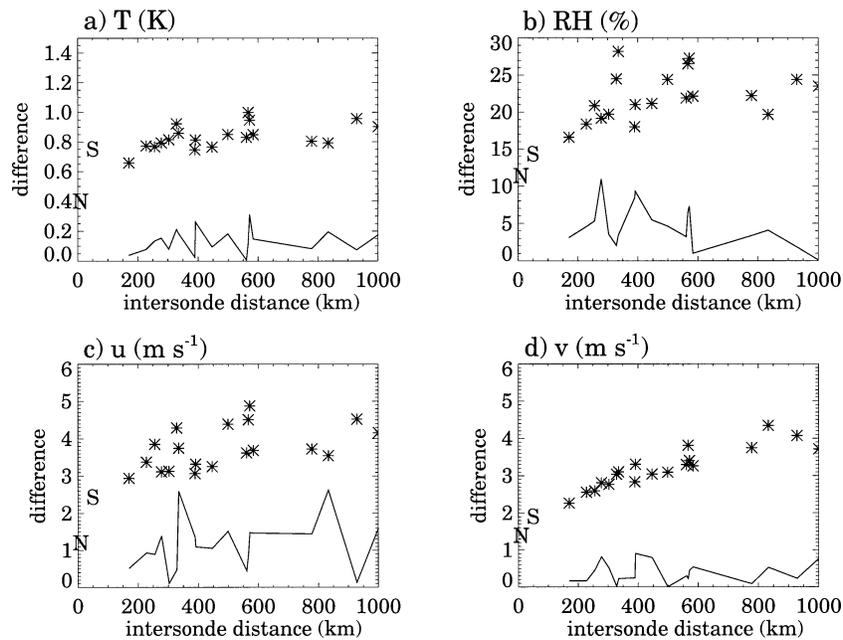


FIG. 9. Statistics of instantaneous pairwise differences between rawinsondes launched at various distances apart. Solid line: absolute value of time-mean difference. Symbols: rms of residual differences, from COARE (*) and closer station pairs in SCSMEX (S), and Nauru99 (N). (a) Temperature, (b) relative humidity, (c) zonal wind, and (d) meridional wind.

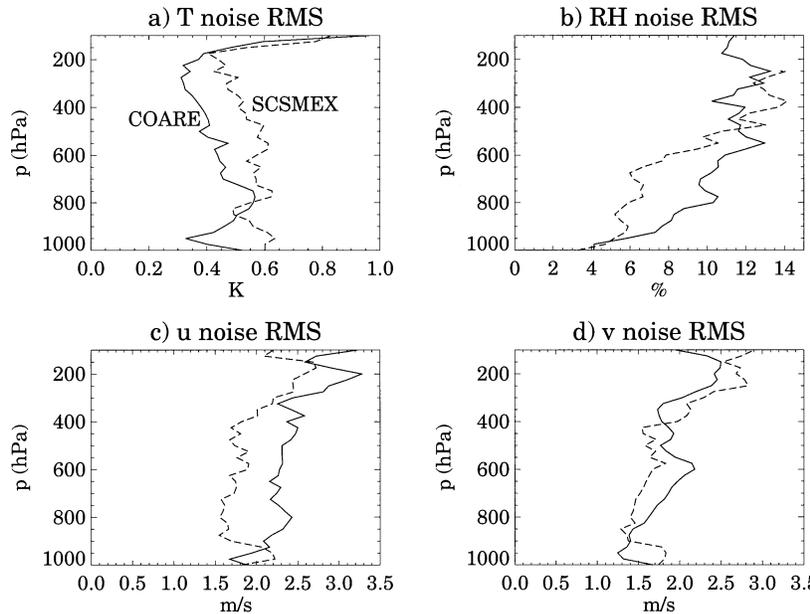


FIG. 10. Standard deviation profiles for the unresolved variability dataset used to perturb the observations entering the large-scale analysis system. Solid line: COARE. Dashed line: SCSMEX. (a) Temperature, (b) relative humidity, (c) zonal wind, and (d) meridional wind.

rawinsonde arrays as real meteorological variability, with space scales and timescales smaller than the intersonde distances and time intervals.

To isolate even more clearly the behavior of error with interstation distance, Fig. 9 shows statistics of differences between 500-hPa data from individual rawinsondes launched at the same nominal hour. All possible pairs were analyzed from the set of six COARE IFA stations plus Manus (Fig. 1a), as well as the two closer station pairs discussed above from SCSMEX (“S” on Fig. 9) and Nauru99 (“N” on Fig. 9). The solid curve indicates the absolute value of the bias (time mean) differences, while plotting symbols indicate the root-mean-square (rms) of each set of differences after its mean is removed. In general, the time-mean biases are much smaller than the rms differences, which increase rapidly with distance in the first 50 km, then increase modestly with station distances >100 km.

At the limit of zero station spacing, the rms difference should asymptote to $2^{1/2}$ times the rms of the random instrument error, because the error contributes equally to both measurements. For similar reasons, the pairwise rms differences in Fig. 9 are about $2^{1/2}$ times the single-sounding sampling variability shown in Sobel et al. (2004), which is based on the deviations of individual soundings from the mean of five soundings in a region ~ 200 km in size. The SUV in Figs. 3, 5, 7, and 8 is also smaller than the square of the values in Fig. 9, because SUV is derived from two-point time averages rather than single instantaneous measurements.

4. Random realizations of unresolved variability

Section 3 explored the *magnitude* of unresolved variability, but to estimate the effects on budget calculations more information is needed. In particular, realizations of random noise that statistically resemble unresolved variability (plus instrumental error) are needed for input to the ensemble of data analyses. For simplicity, the random noise will be assumed to be uncorrelated in time or among radiosonde stations, with no weather regime dependence. It is, however, important to realistically capture the vertical correlation properties of unresolved variability, especially in wind. The reason is that sampling error in wind at a point affects the calculation of divergence, which is integrated vertically to give vertical velocity. This vertical integration process is sensitive to the depth over which wind and hence divergence errors tend to be coherent.

To extract realizations of unresolved variability from the data, large-scale or resolved fields are first constructed and subtracted from the actual sounding data at a point. For the resolved values, we take the output of the CSU multiquadric large-scale analyses, averaged over some space–time region centered roughly on some individual soundings. For COARE, we subtract the IFA mean time–height section, which necessarily averages information coming from several sounding sites, from the individual soundings at the *Moana Wave* and *Xn5* ships, which sat within the interior of the array (Fig. 1a). Two hundred twenty-three profiles with good-quality thermodynamic and wind data spanning the entire 1000–100-hPa layer were used. The (small) mean over

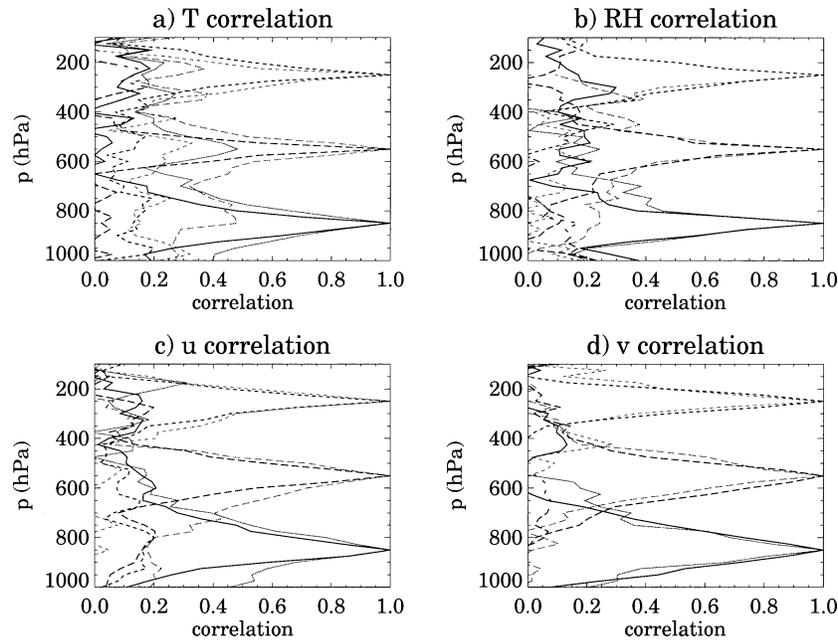


FIG. 11. Vertical correlation structure of unresolved variability. (a) Temperature, (b) relative humidity, (c) zonal wind, and (d) meridional wind. (a)–(d) Display profiles of correlation coefficient between the data at each level and data at three base levels: 850 hPa (solid), 550 hPa (dashed), and 250 hPa (dotted). Black curves: COARE. Gray curves: SCSMEX.

this set of the sonde-analysis difference at each level was subtracted, so that the resulting set of unresolved fluctuation profiles introduces zero bias. Histograms of these fluctuations look very close to Gaussian (not shown). These profiles form a “noise” dataset, from which profiles are sampled at random and added to individual soundings entering the ensemble of analysis and budget computations. For SCSMEX, a 4×4 box in the 1° analyses (19° – 22° N, 115° – 118° E) was chosen, centered roughly on the ship at 21° N, 117° E, where 122 profiles without errors in 1000–100 hPa were available. Because of the data sparsity in that region, the analysis fields too closely resemble the ship data, so a three-point running mean in time was performed on the analysis time–height section, to bring unresolved variances up to values comparable to the COARE noise dataset.

The standard deviation profiles for these noise datasets are shown in Fig. 10 for COARE (solid) and SCSMEX (dashed). Values are similar to, but appropriately smaller than (for reasons discussed above), those discussed in section 3 for station-pair statistics. Values in Fig. 10 are more directly comparable to, and are quite similar to, values shown in Sobel et al. (2004). Figure 11 shows the vertical-lag autocorrelation structure of unresolved variability for base points at three altitudes, for both COARE (black) and SCSMEX (gray). The e -folding distance for correlation decay tends to be a bit less than 100 hPa, implying a characteristic thickness of roughly 200 hPa to unresolved features in all fields. There are also weaker long-range correlations,

which can be important to vertical integrals in the computation of major budget terms. This vertically extended correlation could come from slight individual sonde instrument biases, or may be meteorological in origin. The most prominent COARE–SCSMEX differences are that SCSMEX temperature and zonal wind fluctuations tend to have a larger correlation thickness (Figs. 11a,c).

Before displaying the ensemble results for various computed budget quantities, it is instructive to anticipate the error characteristics based on simple scale analysis from Figs. 10 and 11. In the moisture budget, used for rainfall computations, a dominant term is vertical advection. We may estimate the sampling error in this term as follows. Middle-tropospheric vertical velocity is obtained from a vertical integral of wind divergence. The wind sampling error of $\sim 2 \text{ m s}^{-1}$ (Figs. 10c,d), divided by the scale of the IFA ($L \sim 500 \text{ km}$) becomes a divergence error of $\sim 4 \times 10^{-6} \text{ s}^{-1}$. If that error has a 50% coherence over a 200-hPa layer, the resulting vertical velocity error is $\sim 4 \times 10^{-4} \text{ hPa s}^{-1}$ or 40 hPa day^{-1} . Multiplying by a typical mean vertical gradient of specific humidity ($20 \text{ g kg}^{-1}/800 \text{ hPa}$) yields $1 \text{ g kg}^{-1} \text{ day}^{-1}$. Rainfall rate is a vertical integral of this moisture tendency over mass. We thus multiply by the mass of the troposphere, since vertical velocity error tends to be coherent throughout the troposphere, yielding $\sim 10 \text{ kg m}^{-2} \text{ day}^{-1}$ or 10 mm day^{-1} sampling error in diagnosed rainfall rate. This is in reasonable agreement with the results of section 5. Since it is uncorrelated in time, this rainfall rate error should decrease with time averaging

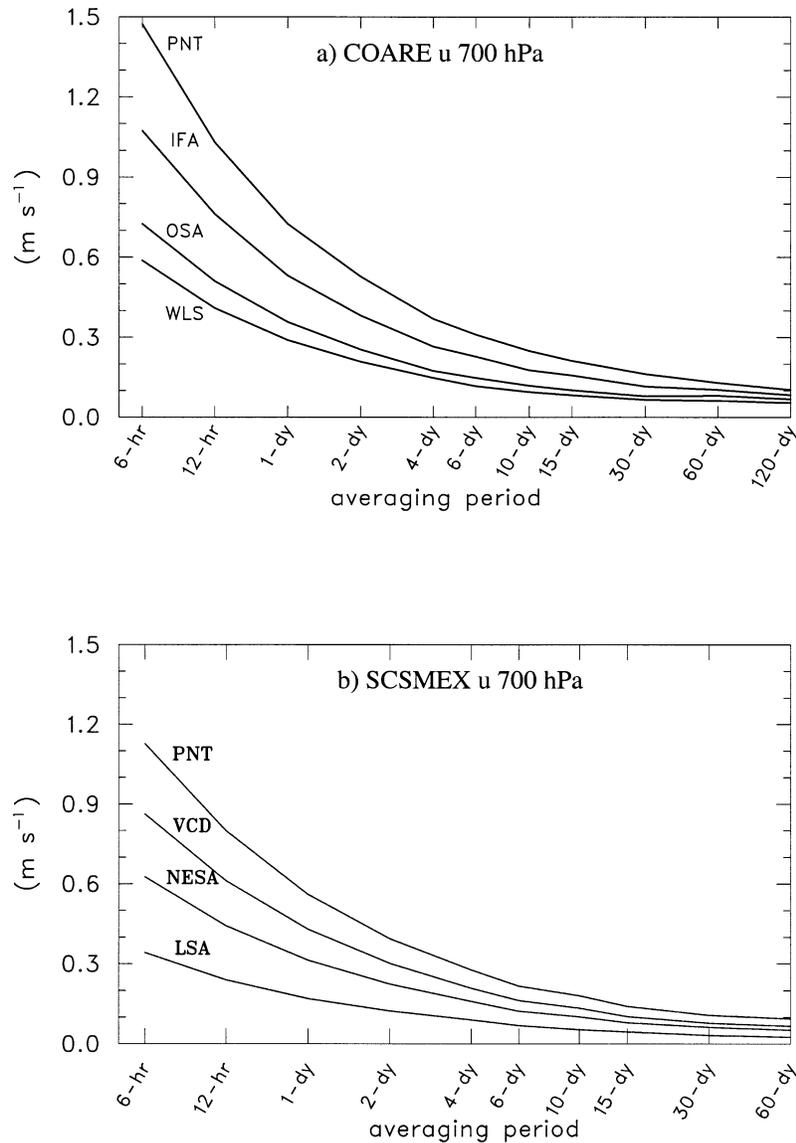


FIG. 12. Ensemble standard deviation of zonal wind at 700 hPa as a function of averaging time, averaged over the indicated geographical regions. (a) COARE and (b) SCSMEX.

approximately as $N^{-1/2}$, where N is the number of time samples averaged. It should decrease roughly linearly with the length scale characterizing the geographical region in question, or perhaps a bit faster because wind information from more soundings gets averaged along the perimeter of larger geographical areas. The contribution of humidity sampling error to rainfall rate, via the vertical advection term, turns out to be smaller than the contribution of wind sampling error. Since a humidity anomaly within the troposphere implies a dipole of vertical gradient anomalies, there is some tendency for the resulting vertical advection error to cancel in the final vertical integration.

Another error term worth noting is in geopotential

thickness, whose horizontal gradients vex momentum budgets (Carr and Bretherton 2001; Tung and Yanai 2002a,b). Temperature sampling errors of ~ 0.5 K (Fig. 10a), coherent over ~ 100 -hPa layers, yield geopotential height (thickness) random sampling errors of roughly 25 m. Height gradient errors across 500 km are therefore of order 5×10^{-5} , implying wind acceleration errors of $40 \text{ m s}^{-1} \text{ day}^{-1}$. Such large errors completely overshadow typical values of vertical momentum flux convergence in tropical convection (e.g., Mapes and Wu 2001). As a result, geopotential heights from global model analyses must be used in momentum budgets, introducing unknown model dependence into the results.

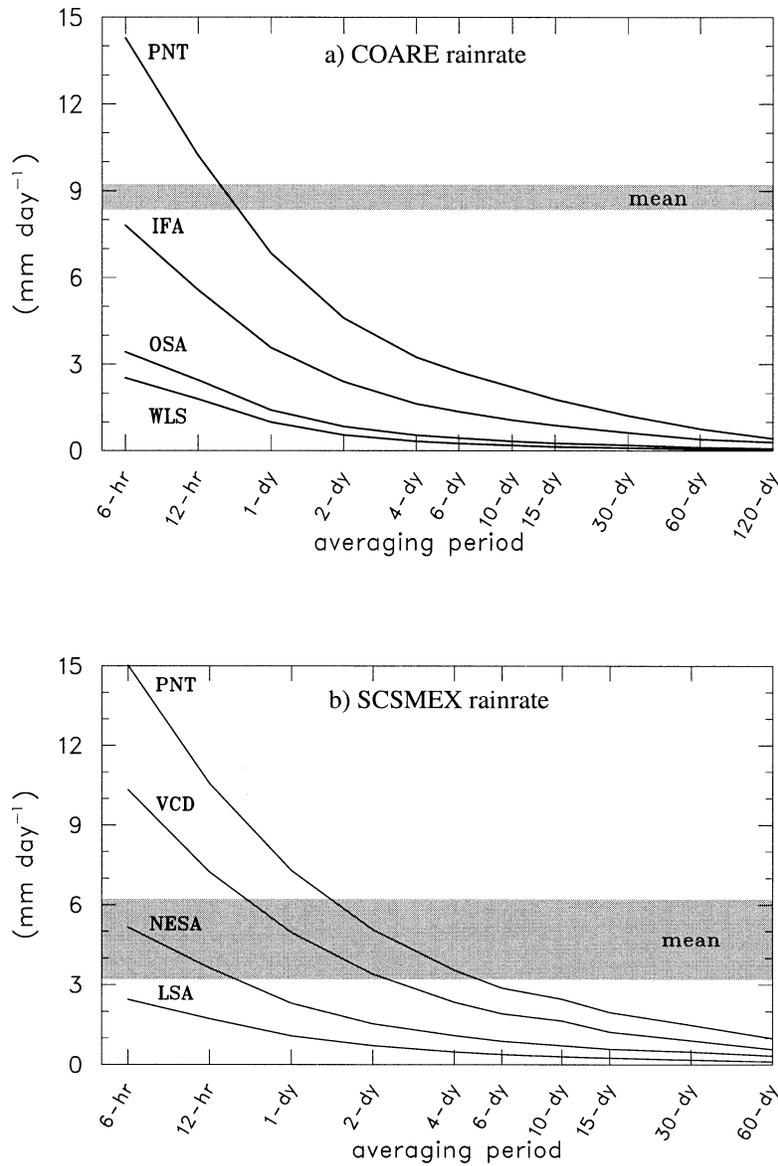


FIG. 13. Ensemble standard deviation of rainfall rate as a function of averaging time, averaged over the indicated geographical regions. (a) COARE; (b) SCSMEX. Shaded bar represents the range of experiment-mean values for the various geographical regions.

5. Sampling errors in analyses and budgets

The sampling errors in directly measured variables will propagate through analysis and budget computations into more indirectly derived variables, as anticipated above. The uncertainty of results is here displayed as the standard deviation among an ensemble of 20 analysis and budget computations performed with perturbed observations. As a simplest example (a directly measured variable), Fig. 12 shows the 700-hPa zonal wind error as a function of averaging time for results averaged over various-sized regions in Fig. 1. For instantaneous (labeled 6 hr) point values, the errors are essentially equal to the point error standard deviations seen in Figs.

3c, 5c, or 10. These errors decrease with averaging scale in a way that must depend on the station density, and with time in a way that is roughly consistent with elementary $N^{-1/2}$ sampling statistics, where N is the number of observation times. For example, errors are roughly halved for 1-day (4 time level) averages, and halved again for 4-day averages. Other directly analyzed fields have similarly simple error characteristics (not shown).

One major reason for undertaking this project was to derive error estimates for more elaborately constructed quantities like rain rate and net radiation (Johnson and Ciesielski 2000). Figure 13 shows rainfall rate sampling errors, again as a function of averaging time and for

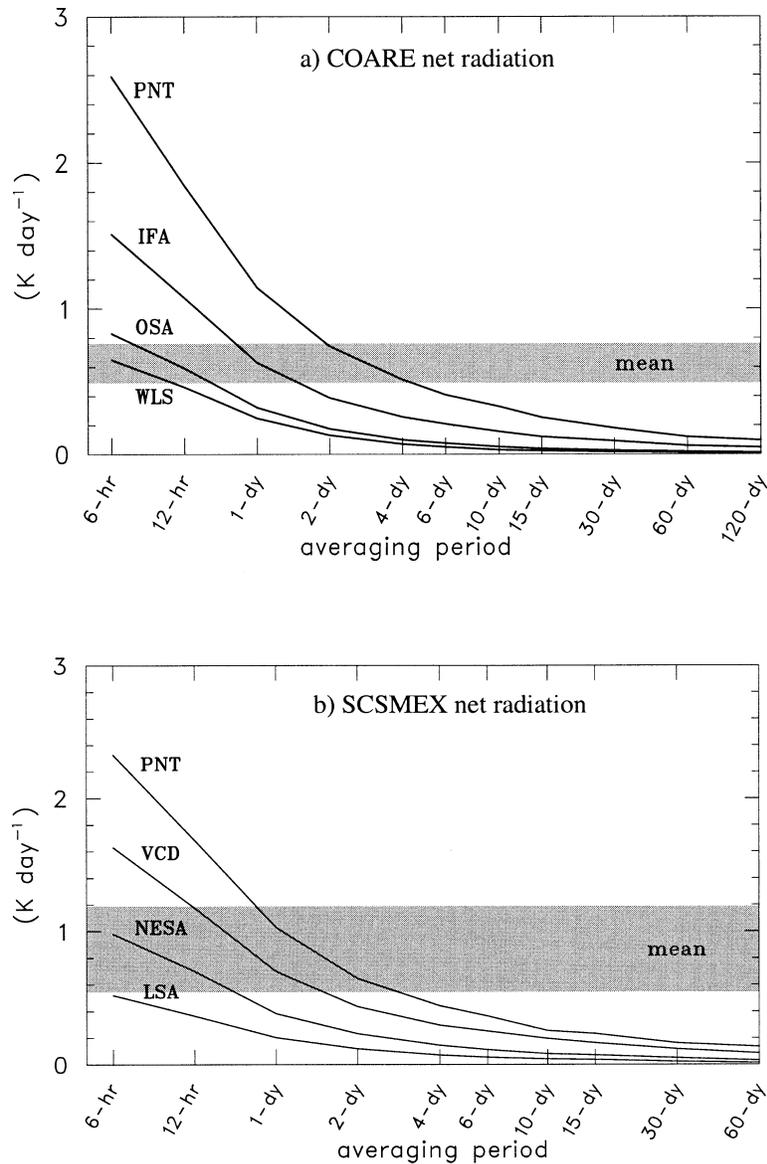


FIG. 14. Ensemble standard deviation of net radiation as a function of averaging time, averaged over the indicated geographical regions. (a) COARE and (b) SCSMEX. Shaded bar represents the range of experiment-mean values for the various geographical regions.

areas of different spatial scales. The decay of error with space scales and timescales resembles that seen in Fig. 12. The absolute values are of interest, especially in relation to the long-term mean rainfall rate (horizontal bar). For point or ~ 500 km (IFA) scales, instantaneous sampling error rivals this mean value, being ~ 10 mm day^{-1} for the IFA as anticipated from the scale analysis above. The errors are comparable in both SCSMEX and COARE. Net radiation is an even more delicate computation, so instantaneous sampling errors exceed the long-term mean even for very large spatial scales (Fig. 14). Multiday averages are necessary to bring the rawinsonde sampling error down to a small proportion of

the mean. As an example, Fig. 15 shows an ensemble of time series of 5-day mean rain rate and net radiation for the COARE IFA.

Figure 16 shows the ensemble standard deviation of the altitude of the peak of the profile of heating diagnosed in the temperature budget. This statistic is more difficult to foresee from simple scaling analysis, but begins (for instantaneous estimates) near the ~ 100 hPa correlation scale of sampling errors in the basic variables, even for quite large area averages. Error drops below 50 hPa for averages longer than 3 days.

A common use of rawinsonde array analyses is to compute estimates of the quantities of the "apparent

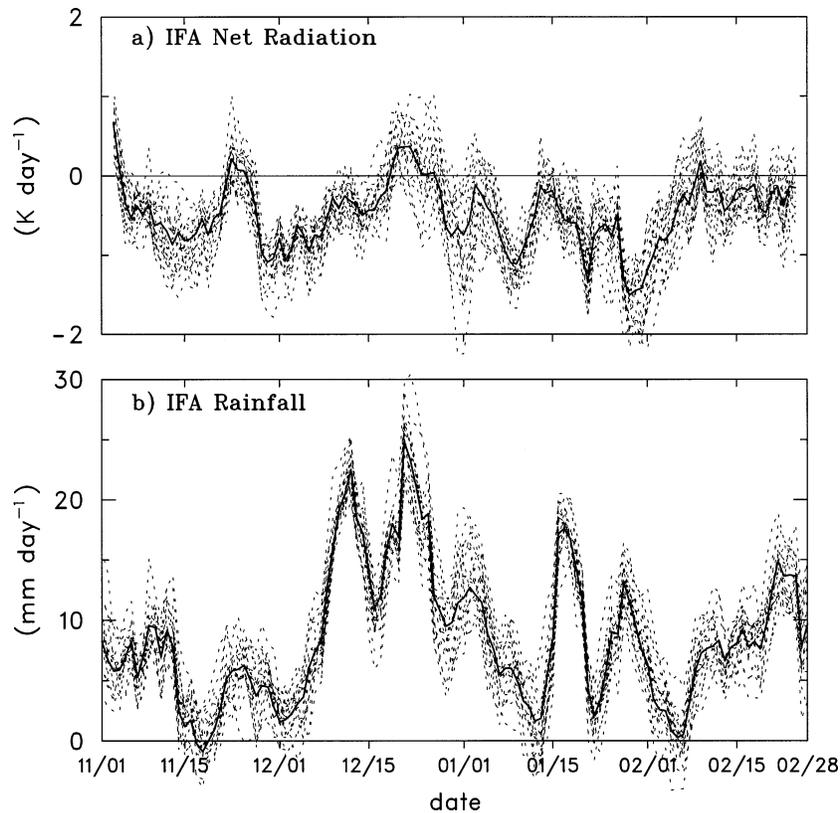


FIG. 15. Time series of (a) COARE IFA net radiation and (b) rainfall. Solid curve is the unperturbed value, lighter broken curves are members of the perturbed ensemble. All curves have been smoothed with a 5-day running mean.

heat source” Q1 and “apparent moisture sink” Q2 (see Yanai and Johnson 1993 for background). Here Q1 and Q2 contain the source and sink terms associated with condensation, as well as terms expressing the convergence of fluxes by circulations of smaller dimensions than the averaging scale, which appears in the Q1 and Q2 definitions. In short, Q1 and Q2 contain (and are, in the wet Tropics, dominated by) the thermodynamic impacts of convection.

For example, Fig. 17 shows vertical-mean Q1 and Q2 estimates (solid curves) from the COARE IFA region in a period leading up to the deep convective phase of the Madden–Julian oscillation in December 1992 (Lin and Johnson 1996a,b). The curves gradually rise over days 30–39, then suddenly spike up on day 40, when organized deep convection developed within the IFA. The information content of these instantaneous Q1 and Q2 estimates is seen to be utterly dominated by the vertical advection term (ensemble of dotted lines in Fig. 17) to such a degree that the uncertainty in the latter (as reflected in the ensemble) is substantially larger than the total magnitude of all other terms. In other words, instantaneous vertical-mean IFA-scale Q1 and Q2 are observationally indistinguishable from vertical advection alone.

Because the physical “signals” in Q1 and Q2 have

low-frequency components, while sampling “noise” is uncorrelated, time smoothing reduces noise faster than it reduces signal amplitude. As a result, for example, a 5-day smoothed plot like Fig. 17b (not shown) has the Q2 curve falling systematically below the envelope of vertical advection curves. This gap indicates the statistical significance of the observation that horizontal advection contributed a systematic drying tendency to the IFA humidity budget, especially during humid (convective) periods. In contrast, however, the Q1 curve in corresponding smoothed versions of Fig. 17a never emerges from the ensemble, with either spatial (OSA versus IFA) or 5-day time averaging. Appropriate caution is required regarding statements using evocative, application-specific labels for sonde–array-derived quantities. For example, the apparent heat source, often viewed as a convective quantity, is not significantly different observationally from “large-scale forcing,” as advective tendencies are called when used to drive cloud-resolving models (CRMs) or single-column models (SCMs).

In Fig. 17, Q1 and Q2 look very similar, but the small difference between them has a special importance in SCM and CRM applications. Thermodynamic biases in long runs of such forced models led Emanuel and Zivkovic-Rothman (1999) and Wu et al. (2000) to analyze

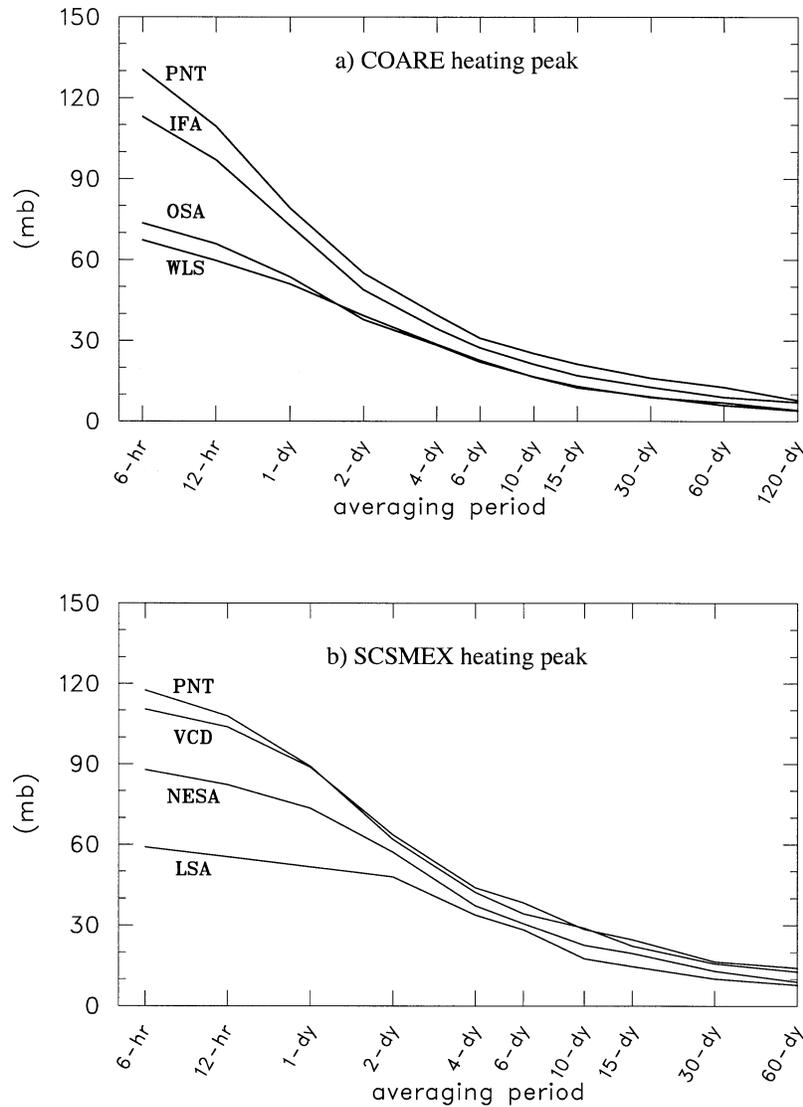


FIG. 16. Ensemble standard deviation of the altitude of the peak of the heating profile, as a function of averaging time, averaged over the indicated geographical regions. (a) COARE and (b) SCSMEX.

forcing datasets in terms of their implied mass-averaged tendencies of enthalpy $K_o = (c_{pd} + c_l q)T + Lq$ (where c_{pd} is the specific heat capacity for dry air, and c_l is the specific heat capacity of liquid water). Since enthalpy is conserved during condensation and evaporation processes, moist convection drops out of the governing equation for its mass-weighted vertical average, whose changes are then dominated by the imposed forcing tendencies. If the forcing data have an imbalance between cooling and moistening, then forced CRM and SCM simulations will tend to develop temperature and/or moisture biases, regardless of their representation of convection. In these authors' early experiments, differences between observed and predicted enthalpy at the end of a 120-day integration using COARE forcing data

were found to be equivalent to a 25-K temperature error integrated over the troposphere.

Ciesielski et al. (2003) have shown that the difference between observed enthalpy changes and the 120-day time integral of COARE data-derived enthalpy forcing was reduced substantially (from 25 to 5 K) by use of the latest humidity-corrected sonde data, in which a low-level dry bias in the original Vaisala sonde data was reduced. To see how sampling error compares to this humidity correction in terms of enthalpy tendencies, Fig. 18 shows observed (heavy curve) and predicted (lighter curves) changes of enthalpy during COARE. Figure 18 was constructed following the procedure outlined in Wu et al. (2000), and using his radiative flux data. The spread of predicted moist enthalpy after 120 days in the

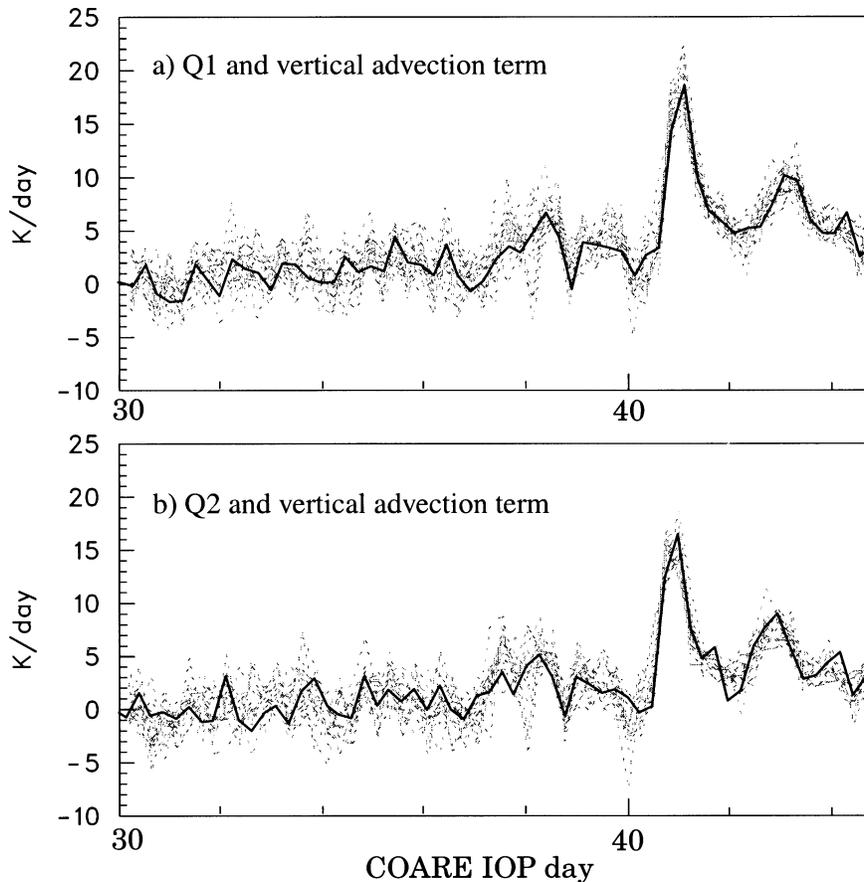


FIG. 17. Line plots connecting instantaneous estimates of the vertical-mean apparent (a) heat source Q1 and (b) moisture sink Q2, averaged over the COARE IFA region (solid curve), and the ensemble of estimates of the contribution solely by the Q1 and Q2 vertical advection term (gray dashed curves). Days are measured within the COARE IOP, which began on 1 Nov 1992.

ensemble of perturbed analyses is almost 25 K, equal to the impact of humidity corrections. Given this fundamental uncertainty, data adjustments (Zhang and Lin 1997) are probably necessary to avoid the problem of enthalpy drift in such forced column model integrations.

6. Summary and conclusions

The sampling error inherent in constructing large-scale analyses by interpolating among rawinsonde point measurements has been examined. Analysis of data from station pairs with spacings typical of special field research programs (a few hundred kilometers) and a 6-hourly launch interval indicates that these errors have typical standard deviations of roughly 0.5 K in temperature; in relative humidity, about 5% near the surface, increasing with altitude to nearly 15% near 400 hPa; and in winds, about 2 m s^{-1} increasing to about 3 m s^{-1} high in the upper troposphere. The main source of error is not instrument accuracy, but rather the aliasing associated with undersampled variations which are too small-scale or high-frequency to be resolved by the ar-

ray spacing. Sampling error is undoubtedly more severe in sparser arrays.

In undisturbed weather conditions during COARE, wind variance (both resolved and unresolved) was found to be smaller than in disturbed conditions. Consistent with reduced homogenization by the reduced wind variability, the unresolved variance in relative humidity is actually greater in undisturbed conditions. Well-resolved (large scale, low frequency) variance is greater in temperature and in upper-level zonal wind during SCSMEX (at 20°N in the summer monsoon) than in COARE. However, unresolved variances are similar for the two regions, suggesting similar gross statistics of the smaller-scale processes such as convection. The mean diurnal cycle is a small fraction of unresolved variance, except in upper-tropospheric temperature, where it approaches 50%.

Spatially and temporally unresolved variances are comparable for COARE 6-hourly soundings spaced ~ 200 km apart. This suggests that a space-time ratio of 200 km/6h $\sim 9 \text{ m s}^{-1}$ is approximately optimal for the spacing and timing of rawinsonde launches in future

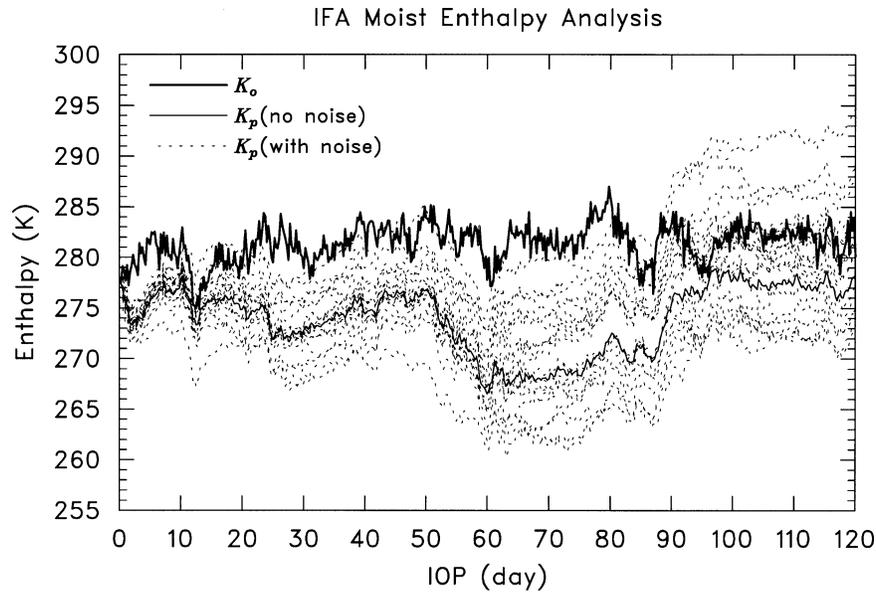


FIG. 18. Moist enthalpy during COARE, observed (thick solid curve) and computed with cumulative (from day 0) IFA average tendencies of temperature and moisture (thin curves). Thin solid curve uses advective tendencies from unperturbed analysis; broken curves are from the ensemble of 20 randomly perturbed analyses.

tropical experiments. This ratio is a characteristic of the unresolved variability, which is probably generated by convective processes in these regions. The ratio may be expected to vary with wind speed, since wind advects unresolved atmospheric structure across the array. Typical thicknesses for unresolved fluctuations (twice the vertical scale of correlation e -folding) are a bit less than 200 hPa in all variables.

Instantaneous large-scale analyses of rawinsonde array data have been performed multiple times to form an ensemble, perturbing the observations with typical realizations of sampling error. The standard deviation of the analyzed quantities, and of budget terms computed from them, allow error bars to be attached to rawinsonde array results. These error estimates were derived from the densest part of the COARE rawinsonde array and are conservative (i.e., too small) for sparser arrays. Sampling errors are also presumably weather dependent, although the datasets analyzed here are not large enough to subdivide meaningfully beyond the disturbed/undisturbed COARE division employed in section 3. Some highly derived budget residuals of special scientific interest include rainfall rate and net radiative heating, whose error characteristics (shown in Figs. 13 and 14) suggest that averages of at least several days are likely necessary to bring random errors down to a small fraction of long-term mean values. Errors in vertical advection estimates make instantaneous estimates of large-scale forcing and apparent heating and drying by convection indistinguishable. Cumulative enthalpy tendency errors due to sampling errors can reach up to 25 K over a 120-day period, producing temperature and/

or moisture biases of that magnitude in cloud-resolving or single-column models forced with rawinsonde-array-derived advective tendencies.

Acknowledgments. This material is based on work supported by the National Science Foundation under Grants ATM-0073206 and ATM-0112715. PEC and RHJ were supported by NOAA Project NA67RJ0152 and NASA NAG5-9665. The helpful input of Jeff Whitaker, Tom Hamill, John Molinari, and an anonymous reviewer are gratefully acknowledged.

REFERENCES

- Brown, R. G., and C. Zhang, 1997: Variability of midtropospheric moisture and its effects on cloud-top height distribution during TOGA COARE. *J. Atmos. Sci.*, **54**, 2760–2774.
- Carr, M. T., and C. S. Bretherton, 2001: Convective momentum transport over the tropical Pacific: Budget estimates. *J. Atmos. Sci.*, **58**, 1673–1693.
- Ciesielski, P. E., L. M. Hartten, and R. H. Johnson, 1997: Impacts of merging profiler and rawinsonde winds on TOGA COARE analyses. *J. Atmos. Oceanic Technol.*, **14**, 1264–1279.
- , R. H. Johnson, P. T. Haertel, and J. Wang, 2003: Corrected TOGA COARE sounding humidity data: Impact on diagnosed properties of convection and climate over the warm pool. *J. Climate*, **16**, 2370–2384.
- Emanuel, K. A., and M. Zivkovic-Rothman, 1999: Development and evaluation of a convection scheme for use in climate models. *J. Atmos. Sci.*, **56**, 1766–1782.
- Haertel, P., 2002: Spurious divergence within objective analyses, with application to TOGA COARE heat and moisture budgets. Preprints, *25th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 192–193.
- Johnson, R. H., and P. E. Ciesielski, 2000: Rainfall and radiative

- heating rates from TOGA COARE atmospheric budgets. *J. Atmos. Sci.*, **57**, 1497–1514.
- , and —, 2002: Characteristics of the 1998 summer monsoon onset over the northern South China Sea. *J. Meteor. Soc. Japan*, **80**, 561–578.
- Lin, X., and R. H. Johnson, 1996a: Kinematic and thermodynamic characteristics of the flow over the western Pacific warm pool during TOGA COARE. *J. Atmos. Sci.*, **53**, 695–715.
- , and —, 1996b: Heating, moistening and rainfall over the western Pacific warm pool during TOGA COARE. *J. Atmos. Sci.*, **53**, 3367–3383.
- Mapes, B. E., and P. Zuidema, 1996: Radiative-dynamical consequences of dry tongues in the tropical troposphere. *J. Atmos. Sci.*, **53**, 620–638.
- , and X. Wu, 2001: Convective eddy momentum tendencies in long cloud-resolving model simulations. *J. Atmos. Sci.*, **58**, 517–526.
- Molinari, J., and S. Skubis, 1988: Calculation of consistent flux and advective terms from adjusted vertical profiles of divergence. *Mon. Wea. Rev.*, **116**, 1829–1837.
- Neter, J., W. Wasserman, and M. H. Kutner, 1990: *Applied Linear Statistical Models*. Irwin, 1181 pp.
- Nuss, W. A., and D. W. Titley, 1994: Use of multiquadric interpolation for meteorological objective analysis. *Mon. Wea. Rev.*, **122**, 1611–1631.
- Randall, D. A., Harshvardan, and D. A. Dazlich, 1991: Diurnal variability of the hydrologic cycle in a general circulation model. *J. Atmos. Sci.*, **48**, 40–62.
- Sobel, A. H., S. E. Yuter, C. S. Bretherton, and G. N. Kiladis, 2004: Large-scale meteorology and deep convection during TRMM KWAJEX. *Mon. Wea. Rev.*, in press.
- Tung, W.-W., and M. Yanai, 2002a: Convective momentum transport observed during the TOGA COARE IOP. Part I: General features. *J. Atmos. Sci.*, **59**, 1857–1871.
- , and —, 2002b: Convective momentum transport observed during the TOGA COARE IOP. Part II: Case studies. *J. Atmos. Sci.*, **59**, 2535–2549.
- Wang, J., H. L. Cole, D. J. Carlson, E. R. Miller, K. Beierle, A. Pakkunen, and T. K. Laine, 2002: Corrections of humidity measurement errors from the Vaisala RS80 radiosonde—Application to TOGA COARE data. *J. Atmos. Oceanic Technol.*, **19**, 981–1002.
- Webster, P. J., and R. Lukas, 1992: TOGA COARE: The Coupled Ocean–Atmosphere Response Experiment. *Bull. Amer. Meteor. Soc.*, **73**, 1377–1416.
- Westwater, E. R., B. B. Stankov, D. Cimini, Y. Han, J. A. Shaw, B. M. Lesht, and C. N. Long, 2003: Radiosonde humidity soundings and microwave radiometers during Nauru99. *J. Atmos. Oceanic Technol.*, **20**, 953–971.
- Wu, X., M. W. Moncrieff, and K. A. Emanuel, 2000: Evaluation of large-scale forcing during TOGA COARE for cloud-resolving models and single-column models. *J. Atmos. Sci.*, **57**, 2977–2985.
- Yanai, M., and R. H. Johnson, 1993: Impacts of cumulus convection on thermodynamic fields. *The Representation of Cumulus Convection in Numerical Models of the Atmosphere*, Meteor. Monogro, No. 46, Amer. Meteor. Soc., 39–62.
- Zhang, M., and J. Lin, 1997: Constrained variational analysis of sounding data based on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to ARM measurements. *J. Atmos. Sci.*, **54**, 1503–1524.
- Zuidema, P., 1998: The 600–800-mb minimum in tropical cloudiness observed during TOGA COARE. *J. Atmos. Sci.*, **55**, 2220–2228.