Adequacy of the Rawinsonde Network for Global Circulation Studies Tested through Numerical Model Output

ABRAHAM H. OORT

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, N. J. 08540

(Manuscript received 1 September 1977, in final form 17 October 1977)

ABSTRACT

Numerical output from a GFDL global climate model was used to determine whether the present distribution of rawinsonde stations is adequate to deduce the atmospheric structure and its variability in space and time over the globe. Spatial data gaps were found to cause typical rms wind errors averaged over a hemisphere of 2 to 3 m s^{-1}, increasing for the zonal wind component to 5 or 6 m s^{-1} at jet stream levels. In temperature the spatial data gaps led to rms errors on the order of 0.5 to 1°C in the free atmosphere, in geopotential height between 20 and 30 gpm in the upper troposphere, and in specific humidity between 1 and 2 g kg^{-1} near the surface and about 0.3 g kg^{-1} at 500 mb.

Errors due to instrumental deficiencies, unrepresentativeness of the local soundings, deficiencies in the analysis technique and gaps in the time series were found to be less important than those due to the spatial gaps, even in the Northern Hemisphere.

In the Northern Hemisphere, the rawinsonde network was found to be generally adequate to measure large-scale circulation statistics. However, in the Southern Hemisphere the incorporation of additional data sources (rawinsonde, satellite or otherwise) is necessary, especially for defining the fluxes by the mean meridional and stationary eddy circulations.

1. Introduction

Before the 1950's progress in the understanding of the atmospheric circulation was hampered by a lack of upper air data. However, since World War II aviation requirements have led to a greatly improved coverage in the Northern Hemisphere. Thus, the number of meteorological stations taking daily rawinsonde observations has grown from less than one hundred during the late 1940's to nearly one thousand in recent years. The improvement was not restricted to the continental regions but also included the North Atlantic and North Pacific Oceans through the introduction of about 15 specially equipped Weather Ships, and also many meteorological observing stations at various islands.

In the early 1950's it was found feasible to undertake for the first time general circulation studies on a hemispheric scale. The theoretical framework to put this wealth of information in a comprehensive form was laid largely by the late Victor P. Starr and his co-workers at the Massachusetts Institute of Technology (MIT). The fundamental equations describing the conservation of mass, angular momentum, water substance and energy were applied to the study of the earth's fluid envelope. Statistics were evaluated for the climate parameters by taking long-term averages along different latitude circles (so-called zonal averages), and studying the properties of both the zonal mean, and the transient and stationary eddy components.[1] This breakdown proved very useful. The observations reduced in this framework have contributed greatly to establishing our present conception of how the earth's climate is maintained (e.g., Lorenz, 1967).

However, there is always the lingering doubt whether the climate over the globe and its variability in space and time can be adequately determined from the present rawinsonde network (e.g., Walker, 1970; Stoldt, 1971; Baer and Tribbia, 1976). The problem becomes evident when one considers the distribution of rawinsonde stations used in our more recent studies of the global climate during the 1968–73 period as shown in Figs. 1 and 2. The first figure shows the location of the available rawinsonde stations over the globe, and the second one a block diagram of their meridional distribution summed over 5° latitude wide belts. From a study of these figures one notices (expressed in a qualitative manner) that the coverage over the continents is good, over the northern oceans fair, and over the southern oceans quite poor. The main purpose of the present paper is to quantify these impressions, and thus to determine to what extent general circulation statistics based on rawinsonde information (i.e., practically all presently avail-

[1] For a list of symbols and definitions see Appendix.

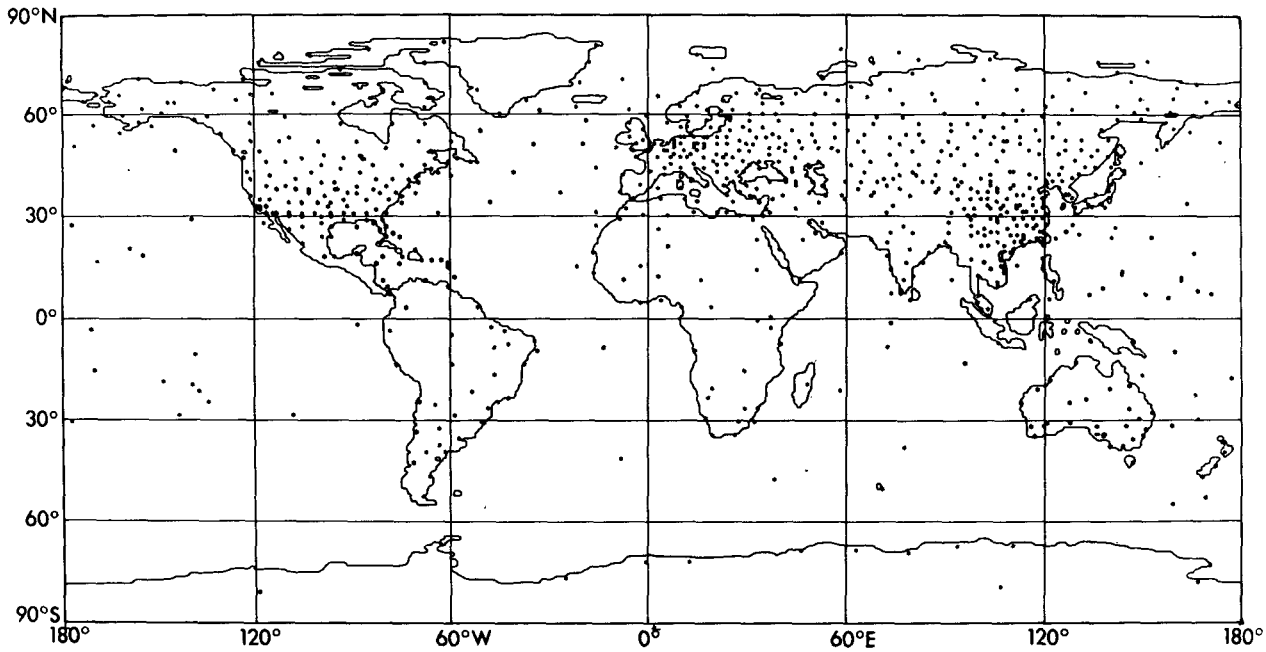

FIG. 1. Global map of the location of rawinsonde stations used in our recent general circulation studies for the period 1968–73.

able atmospheric statistics) are affected by spatial gaps in the observing network. Other types of errors also influence the results. However, it will be shown in Section 7 that data gaps over oceanic regions constitute the most serious source of error, even in case of the Northern Hemisphere.

The other types of errors are related to 1) the accuracy of the basic daily observations, and the interpretation of a local sounding as representing large-scale conditions (Section 4), 2) deficiencies in the objective analysis scheme used (Section 5), and 3) gaps in the time series at each station (Section 6).

2. Method of approach

To investigate the nature of the errors caused by temporal and spatial data gaps we have used information available on GFDL model history tapes. These tapes contain, among other things, the daily grid-point values of the wind components, temperature, pressure and humidity obtained during a two-year integration of a GFDL general circulation model. They were created originally for the purpose of diagnostic analysis of the model results. However, the information on the tapes can also be used for other purposes, such as, testing the adequacy of the rawinsonde network as will be demonstrated in the present paper. In the case of a general circulation model, the "observational" network is probably "perfect" since data are available for each time step and at each point of a regular three-dimensional grid.

In the error tests to be described in the following sections, the daily global meteorological fields from the history tapes were used to compute a variety of monthly mean general circulation statistics. By withholding information at certain times or for certain locations and by reanalyzing the basic monthly mean fields the typical data deficiencies experienced in the real atmosphere were simulated. Through an intercomparison between the model statistics obtained with a realistic data distribution and those obtained for the full model it proved possible to compute the errors in the general circulation statistics for the model. These errors will most likely also represent a good estimate of the probable errors in the general circulation statistics for the true atmosphere. However, it is necessary to explicitly state here that in generalizing these results to the true atmosphere, the basic assumption is made that the numerical model generates a model atmosphere that sufficiently closely resembles the real atmosphere. How one defines a sufficiently close resemblance is, of course, open to discussion. Whether the error estimates are insensitive to the particular approach taken, can and probably should be examined by further tests with other models. However, this is beyond the scope of the present study. It may be mentioned that the evidence presently available to the author, seems to indicate that the present error results are certainly qualitatively and probably also quantitatively representative of the true atmosphere.

3. Description of general circulation model

The particular model used here is the so-called ZODIAC model developed at GFDL by S. Manabe

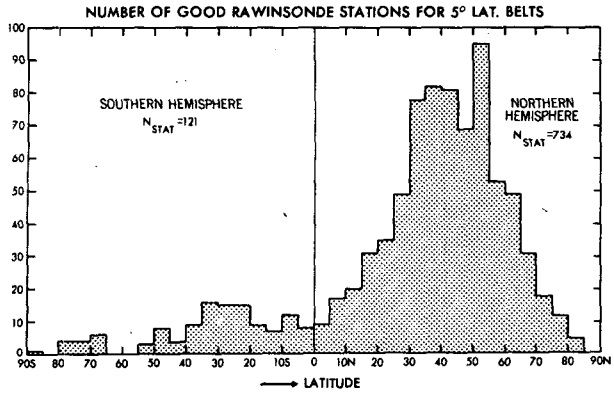

FIG. 2. Meridional distribution of the number of rawinsonde stations in 5° latitude belts, which were used in our recent global general circulation studies for the period 1968–73.

and his co-workers. It is a global, general circulation model with a seasonal cycle. The model has 11 sigma levels in the vertical and a horizontal resolution of about 250 km. Because the evaluation of the ZODIAC model is not the object of the present study, the model results will not be compared here with the results for the actual atmosphere presented in earlier publications (see e.g., Oort and Rasmusson, 1971; Oort, 1977). It should be mentioned, however, that the model in spite of certain obvious deficiencies appears to generate a surprisingly good replica of the true atmospheric circulation and its annual variation. This is shown, for example, by Manabe *et al.* (1974) for the tropics, by Manabe and Mahlman (1976) for the stratosphere, by Hayashi and Golder (1977) for the space-time character of the model disturbances, and by Holloway and Manabe (1971) and Manabe and Holloway (1975) for the global hydrological cycle.

As obvious deficiencies in the ZODIAC model let us single out two parameters, the mean zonal wind distribution and the transient eddy kinetic energy. Thus the subtropical and polar night jet streams do not appear as separate phenomena, but are connected in an unrealistic manner. As regards the general level of transient eddy kinetic energy, the ZODIAC model computes only about half of what is observed in the real atmosphere. However, these problems are not unique for the ZODIAC model but are typical for practically all general circulation models.

The ZODIAC model was run for a period of two years. Because the first year of integration was a test period during which various corrections were made in the computer code, only the output of the second year was used for the present experiments. To show the seasonal variation, results will be presented for the months of January and July.

To make the data handling exactly compatible with that used in our real observational studies it was necessary to interpolate the daily data in the vertical from sigma to pressure levels. The 11 pressure levels are at 1000, 950, 900, 850, 700, 500, 400, 300, 200, 100 and 50 mb. The next step was to calculate monthly mean statistics (mean values, variances and covariances) at each of the 7140 model grid points (or at a subset) and at each pressure level. Then, followed an objective analysis to interpolate the monthly fields to our 47×51 Northern Hemisphere (NH) and Southern Hemisphere (SH) polar stereographic analysis grids. The final step was the computation of the desired spatial variance and covariance, zonal mean, vertical mean, hemispheric mean and/or global mean statistics.

4. Accuracy and representativeness of basic reports

Soundings of wind (u,v), temperature (T), specific humidity (q) and geopotential height (z) in the free atmosphere are subject to various types of errors. First of all there are the pure instrumental errors, as well as the errors due to uncertainty in position of the rawinsonde balloon both in the vertical and horizontal directions. These effects tend to create errors that increase with height; for the upper troposphere root-mean-square (rms) values are about 5 m s^{-1} for the wind components, a few tenths of 1°C for the temperature, and a few decameters for the geopotential height (Lenhard, 1970, 1973). Relative humidity errors may amount to about 10% (Reeves *et al.*, 1976).

Another error is made when one interprets the soundings, as is frequently done, to represent the average value for an area of typically 100 km×100 km or more. Although the meteorological instruments will record the local properties of the air through which the balloon travels fairly accurately, the presence of microscale and mesoscale variations may make these recorded values unrepresentative for the larger area. Resulting errors are generally larger than the estimates given above. This is evident from the recent comparisons that were made between simultaneous reports from adjacent stations (see e.g., Bruce *et al.*, 1977), between rawinsonde reports and interpolated analyses at the same station (see e.g., Schlatter *et al.*, 1976), and between cloud motions and rawinsonde winds (Bauer, 1976). Reasonable values for the rms errors in the upper troposphere seem to lie between 5 and 10 m s^{-1} for the wind components, to be about 1°C for the temperature and about 50 m for the geopotential height.

In addition to the errors in the basic measurements, other problems may arise during transmission and processing of the data. Fortunately, gross errors can generally be detected by a careful analysis of the time series at each station. For example, doubtful data points several standard deviations away from the mean value may be checked further and recognized as obviously erroneous points.

It appears reasonable to assume that the errors discussed up to now are random in character and

uncorrelated [an exception is the possible bias for missing upper air reports during high wind rather than low wind speed conditions, see e.g., Priestley and Troup (1964) and Lorenz (1967, pp. 33 and 88)]. This would imply that the statistics evaluated for a certain station would become progressively more reliable as longer averaging periods are considered. By considering an averaging period with n (independent) data points, the errors in the linear quantities would reduce by a factor of $1/n^{\frac{1}{2}}$. Because the day-to-day variations due to microscale and mesoscale phenomena should be uncorrelated at each station, one may assume about 30 independent data points in one month. This would mean a reduction by a factor of $\frac{1}{5}$ or $\frac{1}{6}$ in the errors given above for the daily values. Thus the error $\bar{u}$ and $\bar{v}$ would be about 1 m s^{-1}, in $\bar{T}$ about 0.2°C, and in $\bar{z}$ about 10 gpm, where the bar denotes the monthly mean. In addition, one may expect a reduction of 30 to 40% in these random errors due to the analysis scheme. This will be discussed extensively in Section 5c.

5. Errors associated with objective analysis method

In order to duplicate the real data situation as well as possible, the same statistical treatment of the model daily data was adopted as the one used by us for real atmospheric data. Thus monthly mean station values were calculated from the time series at each "rawinsonde station" point. Next, the same objective analysis technique was applied to interpolate from the station values to the full analysis grids (for further information see Oort and Rasmusson, 1971).

a. Description of analysis technique

Our original goal in selecting an objective analysis scheme was to arrive at an analyzed field that would as closely as possible resemble the field as obtained by a careful hand analysis, but of course done in an objective manner and in a minimum of computer time.

In the actual procedure, the first guess field for all parameters was a zonally symmetric one, and equal to the zonal average of the available data in the appropriate latitude belt. Next a CRAM (Conditional Relaxation Analysis Method) technique as described by Harris *et al.* (1966), was applied. In this technique the interpolation is done by requiring that the values at all grid points satisfy a Poisson-type equation. The forcing function at the right-hand side of the equation is specified to be the second derivative of the smoothed first guess field. The "observations" are then used as internal boundary points in solving the Poisson equation, while the external boundary points are arbitrarily defined to equal the values of the first guess field. Thus after the rawinsonde grid points have been corrected by the difference between the initial guess at the grid point and the observed value, they become the internal boundary points. Next the Poisson equation is solved in finite-difference form through relaxation at all nonboundary grid points. After smoothing of the resulting field the entire procedure is repeated by analyzing the difference but now between the smoothed field at the grid point and the observed value. The procedure is continued until the differences at the rawinsonde points become sufficiently small.

In the actual calculations an advanced computer language "ANAL68" developed by J. G. Welsh at GFDL was used. This language was designed especially for the handling of meteorological and oceanographic problems. It contains all basic subroutines needed for a CRAM analysis. The analyses were performed on a modified National Meteorological Center (NMC) grid with a grid distance 1.5 times the NMC distance leading to arrays of 47 by 51 grid points (average grid point distance of about 430 km). The NH analyses were done on a NH polar stereographic map with the North Pole located in the center of the grid and the y-axis along the Greenwich meridian. At the left and right sides the analysis extended southward to 6°S, and at the bottom and top sides to 11°S. Similarly the SH analyses were done on a SH polar stereographic map extending northward to 6°N at the left and right sides, and to 11°N at the bottom and top sides. In the construction of the global cross sections and profiles to be presented in this paper the values were taken directly from the NH and SH maps except within a few degrees latitude of the equator where the average of the NH and SH analyses was used. In this way the possibly detrimental influence of the boundaries was largely eliminated.

b. Deficiencies of the analysis technique

Here we will mention some of the good features and also possible shortcomings of the objective analysis scheme described above.

Most relevant is probably a study by Leary and Thompson (1973) who tested this scheme over the Northern Hemisphere by comparing a known function with the objective analysis of its values at the locations of the rawinsonde stations. They found that a surface spherical harmonic of wavenumber 2 was reproduced with little spectral distortion. In contrast the power input for a very high wavenumber 12 was reduced to only 13% of its original value, compared to 87% for wavenumber 2. These results are qualitatively consistent with what one would expect on the basis of the station distribution. Of course, it should be kept in mind that most of the meteorological information of interest is contained in the low, and not in the high, wavenumbers.

Over relatively data dense regions, as one finds over most of the continents, all analysis schemes will yield very similar results (see, e.g., Gandin, 1963,

p. 67). However, this is not the case over the data sparse oceanic regions. Here, the choice of the initial guess field seems quite important. The reason is that in almost all present interpolation schemes the first guess field is hardly modified at grid points more than 1000 to 2000 km away from any rawinsonde station. This is related to the sharp drop-off in the spatial correlation functions at those distances. The local, long-term mean value would be an obvious first guess, but unfortunately this normal is not known and is just the quantity we would like to determine from our analyses. Therefore, in these data-void regions, no analysis scheme can be expected to give very realistic results. Our choice of a zonal mean of all data in the same latitude belt seems reasonable, at least for those monthly-mean parameters, such as $\bar{u}$, $\bar{T}$, $\bar{z}$ and $\bar{q}$, that tend to have the largest gradients in the meridional direction. For quantities with a more cellular structure in the east-west direction, such as $\bar{v}$, the zonal average is obviously not very helpful as a first guess.

Next in importance to the first guess is the choice of how to modify the first guess. Besides the CRAM analysis technique, the successive approximations technique as described by Cressman (1959) and the optimum interpolation technique as promoted by Gandin (1963) are frequently used in analyzing meteorological fields. The optimum interpolation technique is probably the most elegant one, but seems more suitable for the analysis of daily or monthly anomaly rather than long-term mean fields. The reason for this is that the normal conditions and the shape of the correlation functions are supposed to be known.

A Cressman-type analysis scheme modified by Eddy (1967) was used extensively by Newell *et al.* (1972, 1974). Their final cross sections of the real atmosphere resemble closely our real data cross sections based on the CRAM analysis technique. Further Starr *et al.* (1970) compared various general circulation statistics obtained by a number of objective and other schemes of analyses. They found close agreement in the results obtained by the different schemes, with the possible exception of the computation of the mean meridional velocity. On the basis of these earlier findings one may tentatively conclude that the results of gross general circulation statistics such as those presented in this paper will not depend crucially on the choice of the first guess field and the particular scheme of analysis used. Thus it seems permissible at this point to generalize the present results for the ZODIAC model and assume them to be relatively independent of the analysis technique used. Nevertheless, this point should be checked further in later studies.

c. Experiments with random number input

Another question one might ask regarding the analysis scheme is how it handles a random error field. Random errors associated with, e.g., instrumental deficiencies and mesoscale processes (see Section 4), unfortunately, do contaminate meteorological statistics. Thus random station errors should be reduced in the final grid-point values, while on the other hand, no significant meteorological information should be lost.

From various tests (not to be described here) it was found that the objective analysis scheme is practically linear so that one can analyze the random error field by itself, and does not have to study it superposed on each different data field. Therefore in the present discussion we will simply consider the response to an input of random numbers with a mean of zero and a standard deviation of one unit (e.g., 1 m s^{-1} for the wind components, 1°C for the temperature, etc.) for various spatial distributions of the input.

The resulting rms values were computed for each hemisphere, and are presented in Table 1. For greater reliability the same case was repeated 21 times with different choices of random numbers. Thus both the

TABLE 1. Influence of objective analysis scheme on a field consisting of random numbers with mean=0 and standard deviation=1.

Case	Input specified	22-sample means (±standard deviation) of the rms value computed over the Northern Hemisphere	Southern Hemisphere
1	At all 47×51 points of analysis grid	1.00(±0.02)	1.01(±0.02)
2	At "rawinsonde" grid points only (414 in NH, 110 in SH)	0.70(±0.03)	0.76(±0.03)
3	At exact rawinsonde locations only (734 in NH, 121 in SH)	0.69(±0.03)	0.59(±0.06)
4	At exact rawinsonde locations only; the final analysis field is smoothed.*	0.57(±0.03)	0.56(±0.06)
5	At all 47×51 points of analysis grid; the final analysis field is smoothed.*	0.59(±0.03)	0.59(±0.01)

* In the analysis of real data one frequently applies a final smoother to the analyzed fields which does only affect the very high wavenumbers. In our case, twice the smoother $a_{ij}=\frac{1}{2}\,a_{ij}+(a_{i+1,j}+a_{i-1,j}+a_{i,j+1}+a_{i,j-1})/8$ was applied, and once the desmoother $b_{ij}=2b_{ij}-(b_{i+1,j}+b_{i-1,j}+b_{i,j+1}+b_{i,j-1})/4$ to restore the contribution to the low wavenumbers.

TABLE 2. Errors due to temporal gaps only. Shown are rms differences between the 20-random-day mean and the 31-day mean fields averaged over the entire Northern or Southern Hemisphere for various parameters at selected pressure levels for January. Case (1) for the full ZODIAC model with information at all grid points of our 47×51 polar stereographic analysis grids (i.e., 1353 grid points in each hemisphere), and case (2) for the rawinsonde-simulated ZODIAC model with information at only rawinsonde grid points of our analysis grids (i.e., 414 grid points in NH and 110 grid points in SH).

		(1) Full ZODIAC model				(2) Rawinsonde-simulated ZODIAC model				
		850 mb	500 mb	200 mb	50 mb	850 mb	500 mb	200 mb	50 mb	Units
$\bar{u}$	NH	0.4	0.4	0.6	0.4	0.4	0.5	0.7	0.6	m s^{-1}
	SH	0.4	0.4	0.6	0.3	0.4	0.4	0.7	0.4	
$\bar{v}$	NH	0.3	0.4	0.5	0.3	0.5	0.4	0.6	0.4	m s^{-1}
	SH	0.3	0.3	0.6	0.2	0.4	0.4	0.8	0.4	
$\bar{T}$	NH	0.2	0.1	0.1	0.2	0.4	0.2	0.1	0.3	°C
	SH	0.2	0.1	0.1	0.1	0.2	0.2	0.1	0.2	
$\bar{q}$	NH	0.2	0.1	—	—	0.2	0.1	—	—	g kg^{-1}
	SH	0.2	0.1	—	—	0.2	0.1	—	—	
$\overline{u'^2}$	NH	3	4	7	4	4	5	8	6	m^2 s^{-2}
	SH	4	3	8	2	4	3	9	2	
$\overline{v'^2}$	NH	3	3	6	2	4	5	9	3	m^2 s^{-2}
	SH	3	3	7	1	4	4	11	1	
$\overline{T'^2}$	NH	1	1	0	2	2	1	1	3	°C^2
	SH	1	0	0	0	2	1	0	1	
$\overline{q'^2}$	NH	0.9	0.1	—	—	1.0	0.1	—	—	g^2 kg^{-2}
	SH	0.9	0.1	—	—	1.1	0.1	—	—	
$\overline{v'u'}$	NH	2	2	4	2	3	4	6	3	m^2 s^{-2}
	SH	2	2	5	1	3	3	7	1	
$\overline{v'T'}$	NH	1	1	1	1	2	1	1	2	m s^{-1} °C
	SH	1	1	1	0	1	1	1	1	
$\overline{v'q'}$	NH	0.8	0.2	—	—	1.0	0.3	—	—	m s^{-1} g kg^{-1}
	SH	1.1	0.3	—	—	1.6	0.3	—	—	

22 sample means as well as their standard deviations are tabulated.

One notices in Table 1 that the error field is reduced in all cases from an input value of 1.00 to values ranging between 0.56 and 0.76. Therefore, the analysis method has the desirable property of being conservative. In other words, it does not create sizeable maxima or minima away from the input data. If the input is specified at the actual rawinsonde locations (cases 3 and 4) rather than at the nearest grid points (case 2), the random error is reduced even more because of the extra interpolation needed. The effect of high wavenumber smoothing is, of course, very large in the case of random numbers specified at all grid points (case 5). The result is a reduction from 1.00 to less than 0.60. However, for more realistic data distributions (compare cases 3 and 4) the reduction by smoothing is rather small because of the sparse station network. For the same reason, the intersample variability is generally larger in the Southern Hemisphere than in the Northern Hemisphere. It can be shown analytically that the lowest rms value one may expect for a network would be about 0.5 (Simmonds, 1977, private communication).

6. Errors associated with temporal gaps

A different type of error to be discussed next is that due to missing reports in the final time series at each station available to the researcher. For global studies one has to accept stations with incomplete records, especially in data sparse regions. To minimize the detrimental effects of such data gaps in the time series we have specified in our earlier observational studies a cutoff criterion of 10 reports per month. Only stations that reported more frequently were used in further analyses. It was found that the bulk of the stations used reported with an average frequency between 60 and 90%, and that only about 5% of the stations reported less than 50% of the time.

The influence of such temporal data gaps was estimated in the ZODIAC model by recomputing the model statistics after creating artificial record gaps. In the actual procedure the same reporting frequency of about 65% was assumed for all grid points. Thus 20 days were selected at each grid point, independently and at random, from the total sample of 31 days available for the month of January. These 20 days, in general different for each point, were kept in the record, while the 11 remaining days were given a missing value. Using this 20-day sample monthly mean, variance and covariance statistics were computed for all grid points, then new horizontal analyses were obtained and new general circulation statistics were computed.

The comparison between the 20-day results and those for the full 31-day sample is given in Table 2. Shown are the rms differences between the two sets of monthly mean fields, which were averaged over the two hemispheres separately. Both the linear mean

quantities, as well as their variances and covariances in time are considered at various pressure levels. The computations were done for two cases. Case 1 is the situation that all points of the 47×51 polar stereographic grids (1353 grid points in one hemisphere) were given the appropriate 20 random day values in the test run and the 31 day values in the control run. Almost no interpolation was needed to evaluate the zonally averaged statistics in the two runs. In case 2 the 20 random day values were assigned only to those grid points that were located within a range of 200 km of a rawinsonde station (414 grid points in the Northern Hemisphere and 110 in the Southern Hemisphere; see Fig. 1 for actual station distribution). At all other grid points the values were interpolated using an objective analysis scheme. The same procedure was followed at all 11 pressure levels. The generated analyses were then compared with similar analyses made for the full 31-day samples but again based on the 524 rawinsonde grid points only.

The rms values in Table 2 for cases 1 and 2 are almost identical. This shows that the results are the same whether the full spatial coverage is used or only the rawinsonde coverage in conjunction with an objective analysis scheme. This seems to be another indication that the analysis scheme is a reasonable one. Typical error values are about 0.5 m s^{-1} for $\bar{u}$ and $\bar{v}$, 0.2°C for $\bar{T}$, 5 m^2 s^{-2} for $\overline{u'^2}$ and $\overline{v'^2}$, 1°C^2 for $\overline{T'^2}$, 5 m^2 s^{-2} for the momentum transport $\overline{v'u'}$, and 1 m s^{-1} °C for the heat transport $\overline{v'T'}$ at all levels. Because of the rapid decrease of humidity with height, the humidity errors are found to decrease with height. Below 850 mb (not presented here), the errors in the temperature and its variance are somewhat larger than in the free atmosphere. This is partly due to real variability and partly due to another (probably minor) sampling problem that occurs when the pressure levels are sometimes above and sometimes below the ground.

If one compares the previous error estimates discussed in Section 4 with the present values it is found that the previous ones are larger. Thus one can conclude that the occasional unrepresentativeness of the local soundings for a larger synoptic area could possibly present a serious problem in global observational studies, more serious than the problems caused by any of the other error sources discussed until now.

TABLE 3. Errors due to spatial gaps only. Shown are the rms differences between the full ZODIAC model fields and the rawinsonde-simulated ZODIAC model fields averaged over the entire Northern or Southern Hemisphere for various parameters at selected pressure levels for January and July.

		January					July					
		1000 mb	850 mb	500 mb	200 mb	50 mb	1000 mb	850 mb	500 mb	200 mb	50 mb	Units
$\bar{u}$	NH	1.6	2.4	3.0	5.4	2.5	1.8	2.4	2.4	4.7	2.1	m s^{-1}
	SH	2.3	3.5	3.6	5.9	2.2	2.1	2.9	3.2	4.0	3.3	
$\bar{v}$	NH	1.4	1.2	1.4	1.9	0.8	1.4	1.4	1.3	1.8	0.7	m s^{-1}
	SH	1.8	1.8	1.8	2.6	1.5	2.0	1.7	2.7	2.8	2.0	
$\bar{T}$	NH	2.9	1.0	0.8	0.6	0.6	1.7	1.0	0.6	0.5	0.4	°C
	SH	2.1	1.6	1.0	0.9	0.8	3.3	1.7	1.2	0.6	1.0	
$\bar{z}$	NH	—	6	14	25	26	—	5	12	19	14	gpm
	SH	—	7	19	25	21	—	9	23	33	39	
$\bar{q}$	NH	1.4	1.0	0.2	—	—	1.2	1.2	0.3	—	—	g kg^{-1}
	SH	2.0	1.4	0.3	—	—	1.6	1.0	0.3	—	—	
$\overline{u'^2}$	NH	—*	8	9	18	10	—*	12	8	22	2	m^2 s^{-2}
	SH	—	14	10	25	5	—	12	18	27	6	
$\overline{v'^2}$	NH	—	11	9	14	3	—	7	6	16	1	m^2 s^{-2}
	SH	—	10	11	29	4	—	11	16	22	6	
$\overline{T'^2}$	NH	—	5	1	1	3	—	2	1	1	1	°C^2
	SH	—	4	2	2	1	—	8	3	1	1	
$\overline{q'^2}$	NH	—	2.6	0.2	—	—	—	2.4	0.3	—	—	g^2 kg^{-2}
	SH	—	2.8	0.3	—	—	—	2.6	0.2	—	—	
$\overline{v'u'}$	NH	—	5	5	11	4	—	5	4	13	1	m^2 s^{-2}
	SH	—	6	7	17	3	—	8	10	18	4	
$\overline{v'T'}$	NH	—	4	2	2	2	—	2	1	3	1	m s^{-1} °C
	SH	—	3	3	4	1	—	5	4	3	2	
$\overline{v'q'}$	NH	—	2.2	0.6	—	—	—	2.5	0.8	—	—	m s^{-1} g kg^{-1}
	SH	—	3.0	0.8	—	—	—	2.5	0.6	—	—	
$v\bar{u}$	NH	7	8	17	47	21	5	9	7	25	9	m^2 s^{-2}
	SH	7	9	15	46	14	10	10	27	76	71	
$v\bar{T}$	NH	32	17	20	106	53	31	22	15	98	40	m s^{-1}°C
	SH	36	27	34	149	77	34	25	48	175	132	
$v\bar{q}$	NH	19	7	2	—	—	23	11	2	—	—	m s^{-1} g kg^{-1}
	SH	23	13	2	—	—	17	7	2	—	—	

* The transient eddy statistics at 1000 mb are not reliable.

However, it will be shown in the next section that spatial data gaps in the global network generally lead to the largest errors, and that these gaps will make certain general circulation statistics highly unreliable, especially in the Southern Hemisphere.

7. Errors associated with spatial gaps

In the tests to be described here the monthly data were deleted at grid points more than 200 km removed from any "rawinsonde grid point." Next new data were obtained at these points by interpolation from the still available information at the rawinsonde grid points. As stated before, interpolation was accomplished through the same objective analysis scheme as is currently being used by us in the analysis of the real atmosphere. A thorough comparison of these newly generated monthly-mean fields with the full-model control fields should then show the influence of the spatial data gaps on the model statistics and, by inference, also the influence on the real atmospheric statistics.

a. Comparison of horizontal fields

To characterize the differences between the full model and the rawinsonde model results it seems appropriate to discuss first a gross, overall measure. As before in the case of temporal gaps, we have chosen the rms difference value averaged over an entire hemisphere as a convenient measure. For a certain parameter and at a certain pressure level one number will now describe the difference or error field. The computed numbers for the temporal means $\bar{u}$, $\bar{v}$, $\bar{T}$, $\bar{z}$ and $\bar{q}$, the temporal variances $\overline{u'^2}$, $\overline{v'^2}$, $\overline{T'^2}$ and $\overline{q'^2}$, the temporal covariances $\overline{v'u'}$, $\overline{v'T'}$ and $\overline{v'q'}$ and the products $\bar{v}\bar{u}$, $\bar{v}\bar{T}$ and $\bar{v}\bar{q}$ are given in Table 3. Shown are the numbers at five representative pressure levels in the troposphere and lower stratosphere for the months of January and July.

Let us discuss some of the outstanding features of the results. First, one may point out that for all parameters in Table 3, errors in the Southern Hemisphere are somewhat larger than in the Northern Hemisphere, perhaps not as much as one might expect on the basis of the station distribution. The reason may be that where there are only few stations (such as in the SH) the analysis scheme gives a smooth pattern, probably uniformly poor, but with relatively low rms values. A much greater sensitivity to the station network will be found in certain zonal mean quantities to be discussed later. As a second point, regarding seasonal differences one notices a decrease in the NH rms error from winter (January) to summer (July) in agreement with the decrease in intensity of the NH circulation. However, for the Southern Hemisphere, the rms errors are practically the same in winter and summer (except for $\bar{z}$). This is perhaps not so surprising because also the intensity of the SH circulation does not change much during the year. As is well known, the occurrence of only small seasonal variations in the Southern Hemisphere is related to its surface characteristics, namely, that it is dominated by oceans in low and middle latitudes and by the permanent Antarctic cold source at high latitudes. Typical values of the error in $\bar{u}$ and $\bar{v}$ are 2 to 3 m s^{-1}, increasing for $\bar{u}$ to 5 or 6 m s^{-1} near jet stream levels.

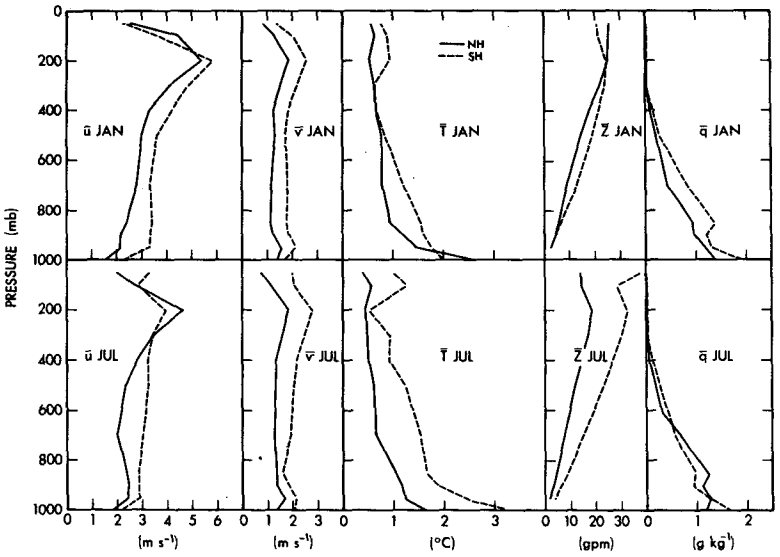

FIG. 3. The rms difference between the full ZODIAC model fields and the rawinsonde simulated ZODIAC model fields for the monthly mean wind components, temperature, geopotential height and specific humidity as a function of pressure. Shown are NH and SH averages by solid and dashed lines, respectively, for January (top) and July (bottom).

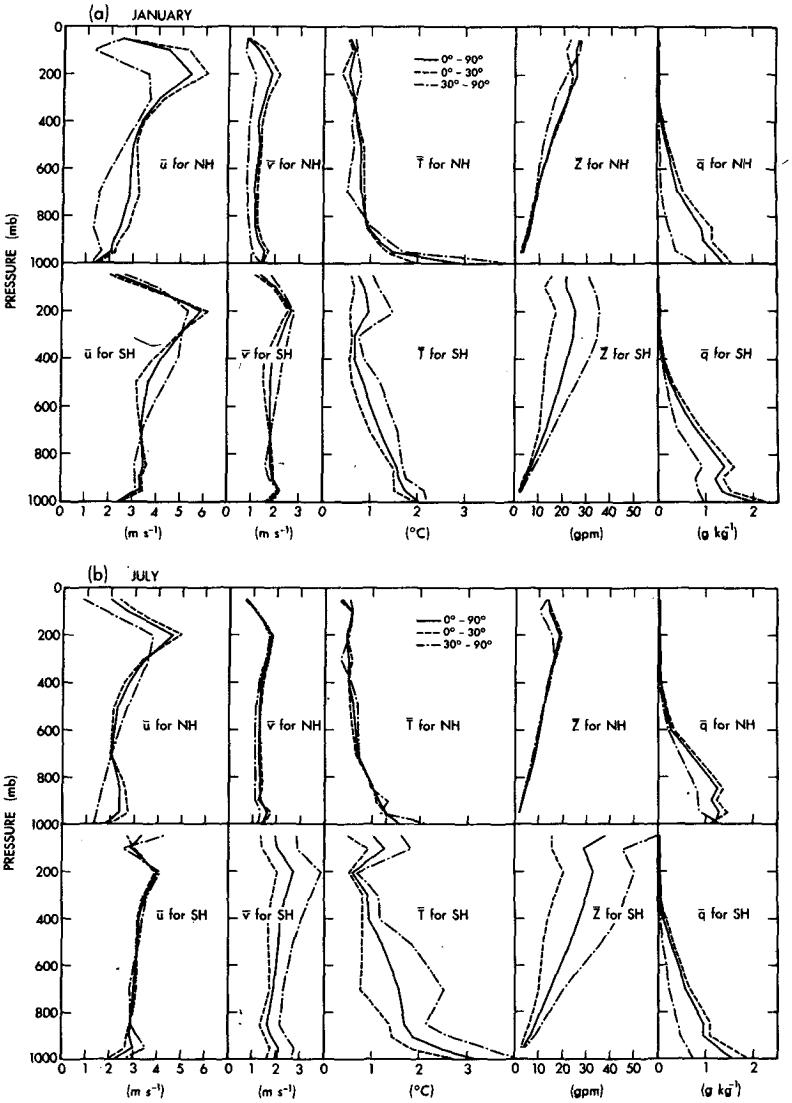

FIG. 4. The rms difference between the full ZODIAC model fields and the rawinsonde simulated ZODIAC model fields for the monthly mean wind components, temperature, geopotential height and specific humidity as a function of pressure for January (a) and July (b). Shown are tropical (0°–30°), extratropical (30°–90°) and hemispheric (0°–90°) averages for the Northern (top) and Southern Hemisphere (bottom).

Temperature errors are of the order of 0.5 to 1°C in the free atmosphere and 2 to 3°C in the surface boundary layer. Errors in the geopotential height range between 20 to 30 gpm in the upper troposphere. Errors in the humidity decrease with height from a value of 1 to 2 g kg^{-1} near the surface to 0.3 g kg^{-1} at 500 mb. For the wind variance and covariance statistics errors tend to increase from the surface to the 200 mb level and then again to decrease, in parallel with typical vertical wind profiles as found at middle latitudes. The statistics involving temperature often show an increase in the rms error near the surface. This is, of course, even more pronounced in the case of humidity.

Figs. 3 and 4 supply more detailed, visual information concerning the vertical and meridional distributions, respectively, of the rms error in $\bar{u}$, $\bar{v}$, $\bar{T}$, $\bar{z}$ and $\bar{q}$. The meridional dependence in Fig. 4 is shown by a breakdown of each hemisphere into a tropical zone (0–30°) and an extratropical zone (30–90°). The rms errors averaged over these zones and over the hemisphere as a whole are plotted in the figure as a function of pressure for January and July. In the Northern Hemisphere rms errors tend to be largest in the 0–30°N zone where the station network is relatively poor. The more dense station network in the 30–90°N zone apparently measures the larger spatial variations in the normal $\bar{u}$, $\bar{v}$, $\bar{T}$ and $\bar{z}$ fields very well. However,

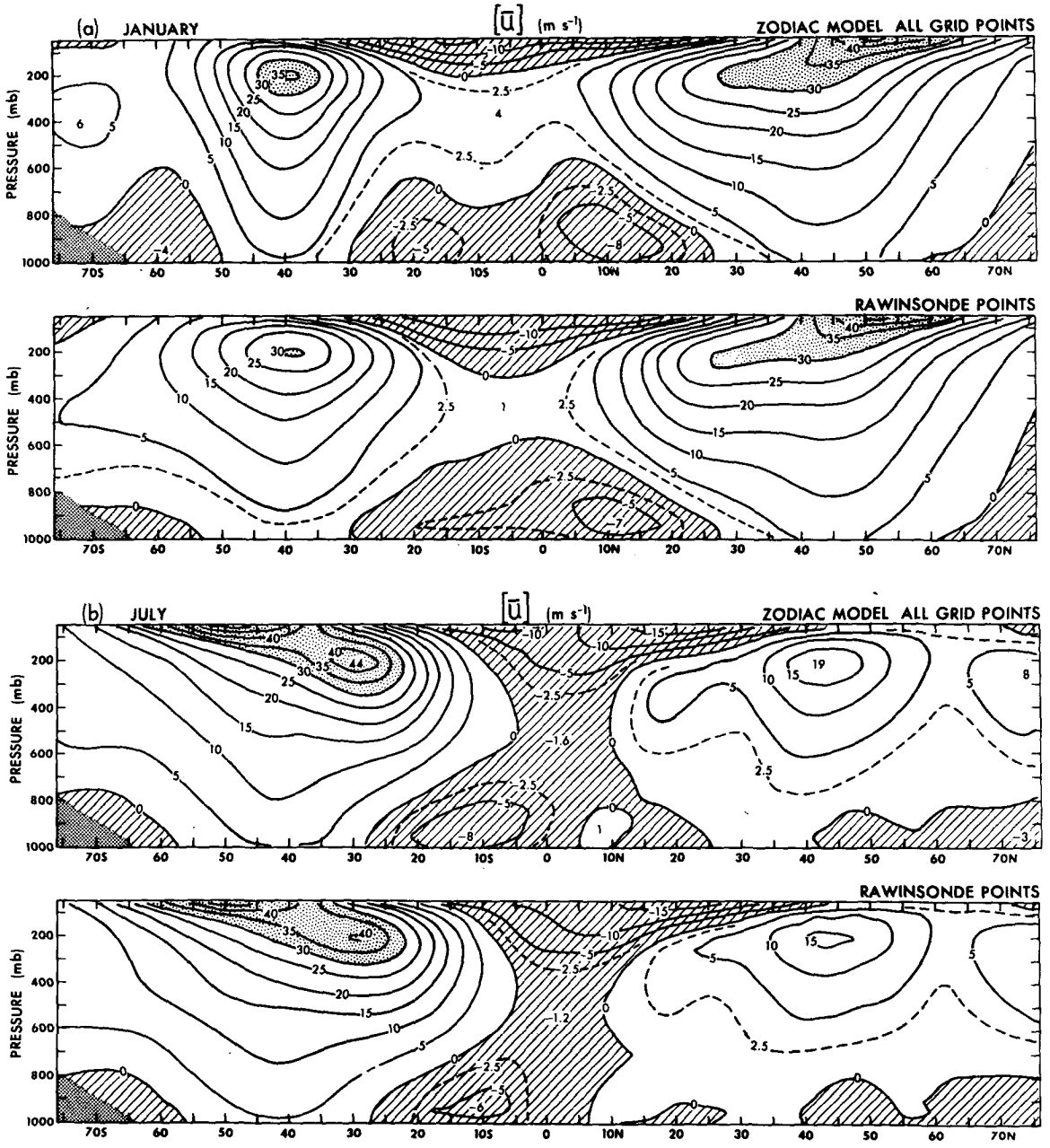

FIG. 5. Comparison between meridional cross sections of the mean zonal wind component $[\bar{u}]$ for the full ZODIAC model (top) and the rawinsonde simulated ZODIAC model (bottom) for January (a) and July (b). Units are in m s^{-1}.

in the Southern Hemisphere it is clear that the worst results are found in the extratropics and not in the tropics. This is in part related to the less dense network in the SH extratropics, and also associated with the larger spatial variations in the $\bar{u}$, $\bar{v}$, $\bar{T}$ and $\bar{z}$ fields at high latitudes.

b. Comparison of meridional cross sections

To further illustrate the type of errors associated with spatial gaps in the rawinsonde network two sets of meridional cross sections will be compared here. The first set (top diagrams) is based on the full model analyses (information at all 1353 grid points in each hemisphere) and the second set (bottom diagrams) on the rawinsonde model analyses (information at only 414 grid points in NH and 110 in SH). Shown are zonal mean pole-to-pole cross sections for the zonal wind component, the meridional streamfunction, the eddy kinetic energy and the eddy northward transport of total energy in Figs. 5–8, respectively. These quantities have been selected to illustrate some of the key differences between the two analyses. The

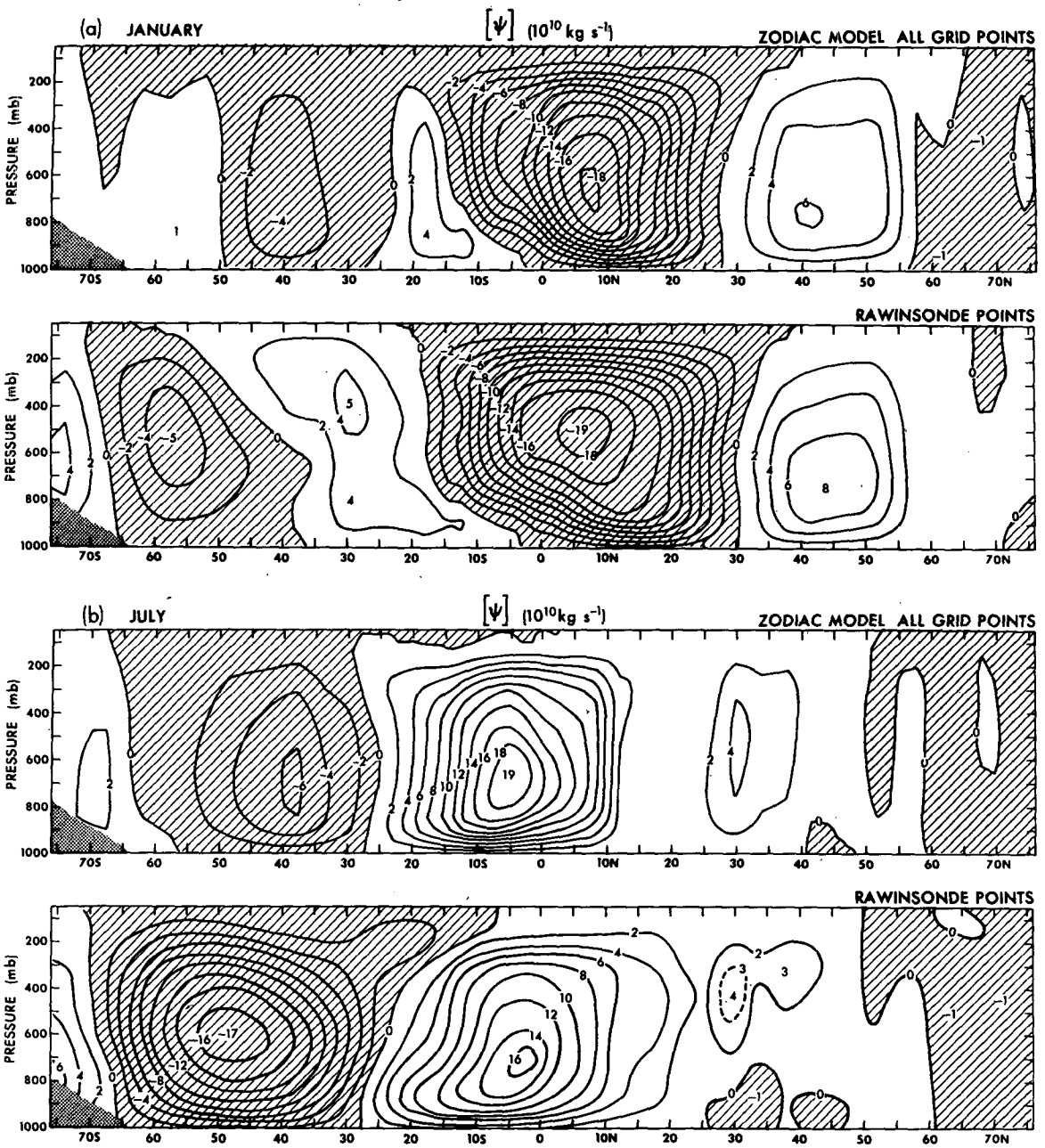

FIG. 6. Comparison between meridional cross sections of the mean meridional streamfunction $[\psi]$ for the full ZODIAC model (top) and the rawinsonde simulated ZODIAC model (bottom) for January (a) and July (b). Units are in 10^{10} kg s^{-1}.

description for a more complete list of general circulation parameters will be postponed to the next subsection.

The mean zonal wind patterns for January (Fig. 5a) and July (Fig. 5b) are remarkably similar in the Northern Hemisphere. However, south of 10°N the analyses begin to deviate. For example, for January in the inner tropics near 400 mb the rawinsonde model analyses give mean westerlies of 1 to 2 m s^{-1} while the full model analyses give stronger westerlies of 3 to 4 m s^{-1}. Also in the upper troposphere the rawinsonde analyses give too strong easterlies. The main reason for these discrepancies is probably the lack of rawinsonde stations in the tropical Eastern Pacific Ocean (see Fig. 1). Over that area at 200 mb the control data in the model atmosphere show the existence of fairly strong westerly winds in a zonal belt otherwise occupied by easterlies. However, in the rawinsonde case because of extrapolation from other longitudes weak easterlies are also found over the eastern Pacific. It is interesting to note that in real atmospheric analyses a similar bias is evident if one

compares the analyses based on rawinsonde stations alone with those based on both rawinsonde and aircraft reports over that area (Sadler, 1975). This fact already provides some justification for our earlier assertion that many of the conclusions based on the present model tests will also be valid for the actual atmosphere. In the Southern Hemisphere one finds large discrepancies in January at all latitudes. Although the location of the maximum zonal current at 40°S is the same, the isolines are too much spread out in the north-south direction in the rawinsonde case. The largest differences between the control and rawinsonde model analyses are found in the 50–70°S belt where the station network is also the poorest. Surprisingly, the July case comparison is very good, especially in the Southern Hemisphere. This is probably fortuitous since the rms values in Table 3 and Fig. 4 show similar errors for January and July.

One of the most sensitive parameters in the general circulation is probably the mean meridional velocity. The associated streamfunction in the (y,p) plane is shown in Fig. 6 for January and July. In the Northern Hemisphere the rawinsonde model seems to reproduce the essential features of the full model, such as the

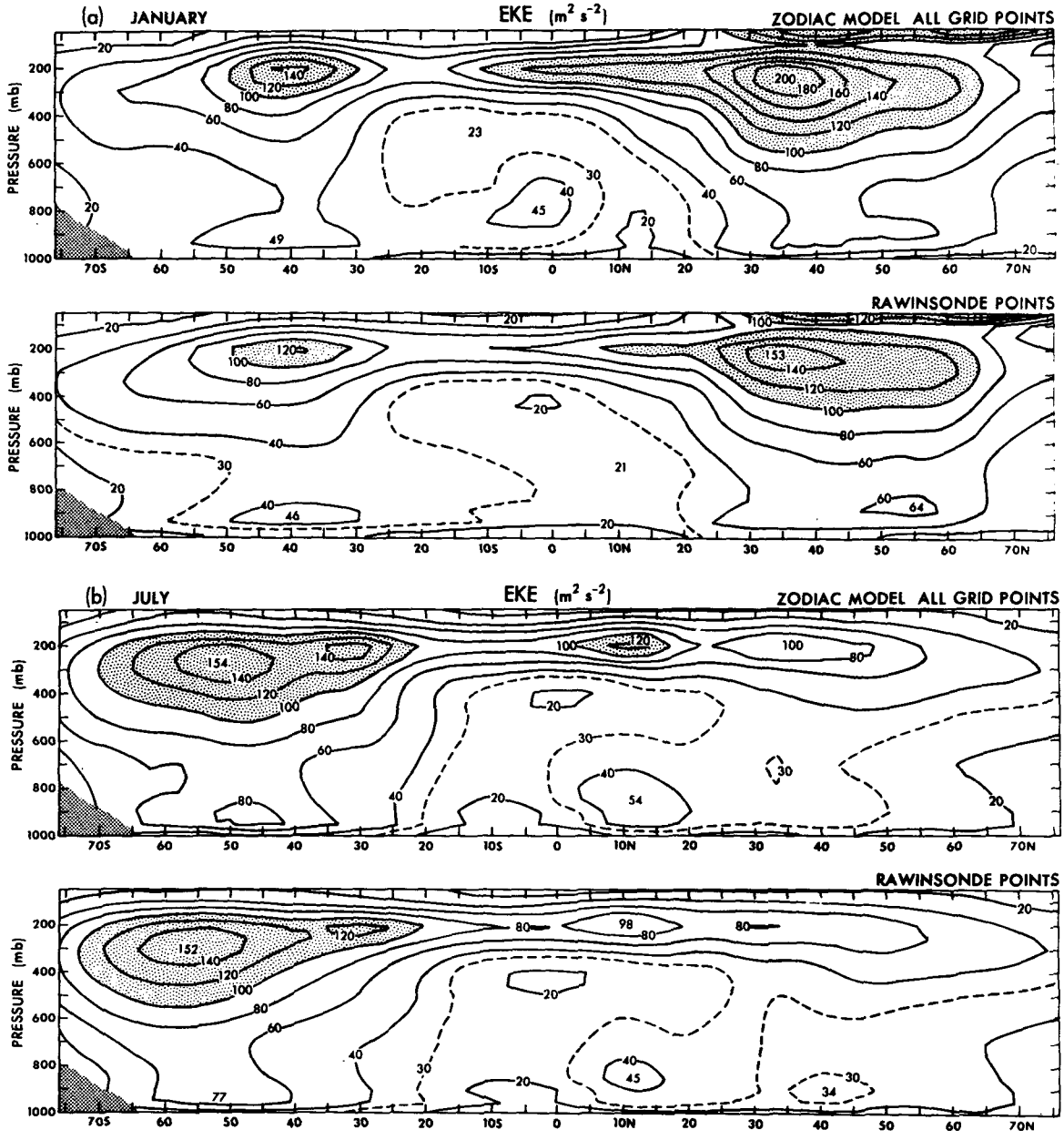

FIG. 7. Comparison between meridional cross sections of the eddy kinetic energy, $\frac{1}{2}[\overline{u'^2}+\overline{v'^2}+\bar{u}^{*2}+\bar{v}^{*2}]$, for the full ZODIAC model (top) and the rawinsonde simulated ZODIAC model (bottom) for January (a) and July (b). Units are in $m^2\ s^{-2}$.

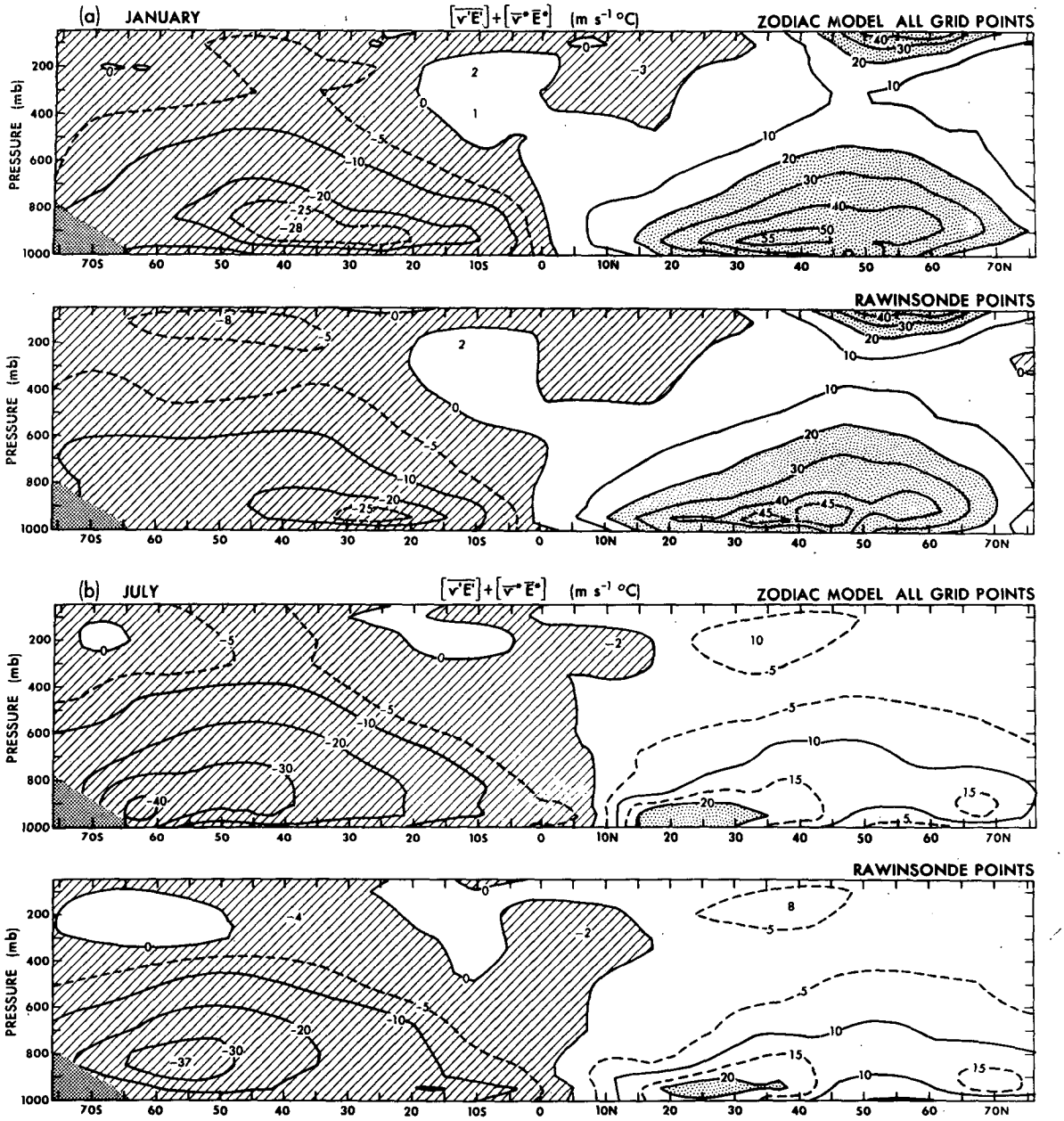

FIG. 8. Comparison between meridional cross sections of the northward eddy transport of energy, $[\overline{v'E'}+\overline{v}^*\overline{E}^*]$, for the full ZODIAC model (top) and the rawinsonde simulated ZODIAC model (bottom) for January (a) and July (b). Units are in m s^{-1} °C.

strong tropical Hadley cell, a weaker mid-latitude Ferrel cell and an ill-defined polar cell. There are slight differences in the intensity and orientation of the cells. In the Southern Hemisphere the situation is very much worse. One finds hardly any correspondence in the mean meridional cells south of about 20°S between the full and the rawinsonde model analyses. In middle latitudes an indirect Ferrel cell is in evidence in both cases but the rawinsonde analysis centers it 10° to 15° latitude too far south, and makes it a factor 2 or 3 too strong. These errors in location and intensity of the SH Ferrel cell will lead to very serious errors in the computed transports of angular momentum and heat by mean meridional overturnings. In fact, it will be shown in the next subsection that the resulting transports based on the rawinsonde network are unusable in the SH extratropics. This seems to be one of the major stumbling blocks for global budget studies. In our recent, real data analyses for the globe a similar result was found in the actual atmosphere, i.e., an unreasonably strong Ferrel cell in the Southern Hemisphere.

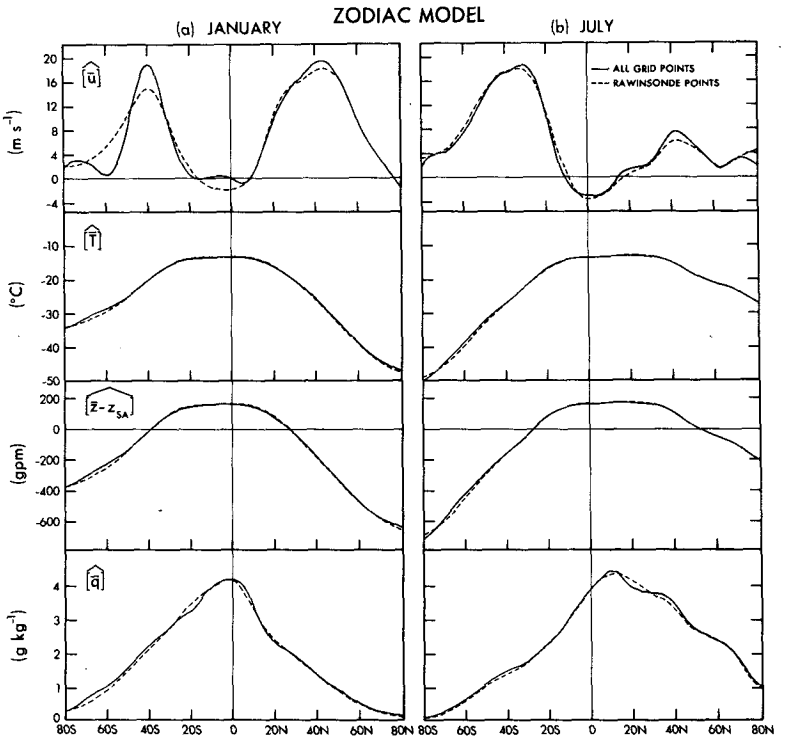

FIG. 9. Meridional profiles of the vertical- and zonal-mean values of the zonal wind component (m s^{-1}), temperature (°C), geopotential height (gpm), and specific humidity (g kg^{-1}) for the full ZODIAC model fields (solid) and for the rawinsonde simulated ZODIAC model fields (dashed) for January (a) and July (b).

The third parameter to be discussed here is the eddy kinetic energy. It consists of a transient and a stationary eddy component. These components are defined as

$$\text{EKE}=\text{TEKE}+\text{SEKE}$$
$$=([\overline{u'^2}]+[\overline{v'^2}])/2+([\bar{u}^{*2}]+[\bar{v}^{*2}])/2,$$

where the overbar and prime indicate a time mean and the departure from the time mean, and the brackets and starr indicate a zonal mean and the departure from the zonal mean, respectively (see also Appendix). Meridional cross sections of EKE are given in Fig. 7. It will be shown later that the transient eddy kinetic energy can be estimated fairly well based on the rawinsonde network. The reason is probably the predominantly zonal character of the transient eddy quantities in both hemispheres that necessitates only a fair distribution of rawinsonde points in the meridional direction. On the other hand, the stationary eddy quantities are, of course, determined by the east-west asymmetries, and require therefore in addition a relatively uniform distribution of stations along each latitude circle. The appropriate number and location of the stations required depends on the scale of the dominant east-west anomalies at that latitude. Fortunately, in the Southern Hemisphere the earth's surface is mostly covered by oceans, and thus quite uniform in the zonal direction. Because of this uniform lower boundary the stationary component is only on the order of 20% of the total eddy kinetic energy in the Southern Hemisphere. Consequently, the estimates of EKE for the control and rawinsonde cases shown in Fig. 7 agree well in the Southern Hemisphere in spite of relatively large discrepancies in the (small) absolute values of the stationary eddy component (not shown here). In the Northern Hemisphere, almost 40% of the eddy kinetic energy is in the stationary component. In spite of the better network the stationary component is about 30% too low compared to that in the full model, and a sizeable error is made in estimating the total Northern Hemisphere EKE. This is evident at all latitudes when one compares the full and rawinsonde model cross sections in Fig. 7. For example, the low-level equatorial maximum in EKE during January is even missing in the rawinsonde model case. The significant underestimation of the NH stationary eddies at all latitudes in the rawinsonde case seems surprising in view of the assumed adequacy of the network in the Northern Hemisphere. This is a new and probably significant result of the present study.

The fourth and last parameter that we will discuss presently is the northward transport of total energy

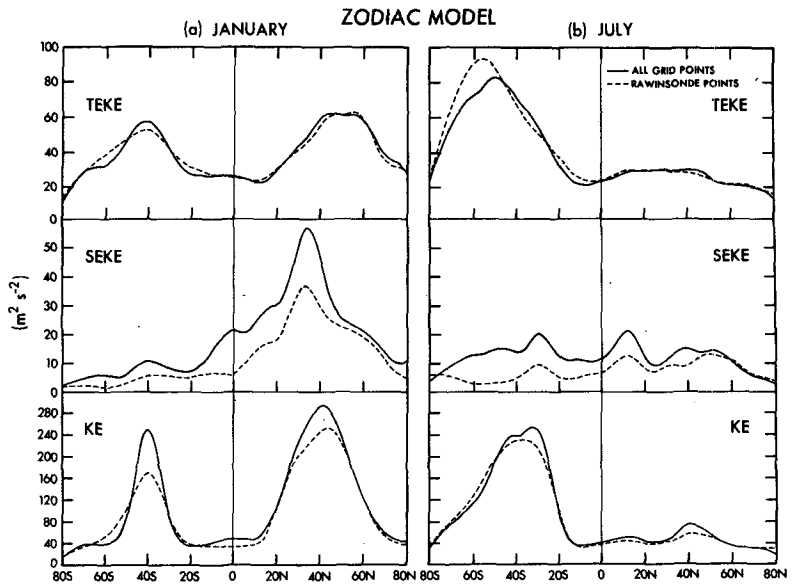

FIG. 10. Meridional profiles of the vertical- and zonal-mean transient eddy kinetic energy, standing eddy kinetic energy, and total kinetic energy for the full ZODIAC model fields (solid) and for the rawinsonde simulated ZODIAC model fields (dashed) for January (a) and July (b). Units are in $m^2 s^{-2}$.

by eddies. The total energy is defined here as the sum of sensible heat, potential energy and latent heat. Cross sections for January and July are shown in Fig. 8. Comparing again the rawinsonde model analyses with the full-model analyses one finds excellent, quantitative agreement in the Northern Hemisphere. Apparently the underestimation of the stationary eddies does not noticeably affect the total eddy trans-

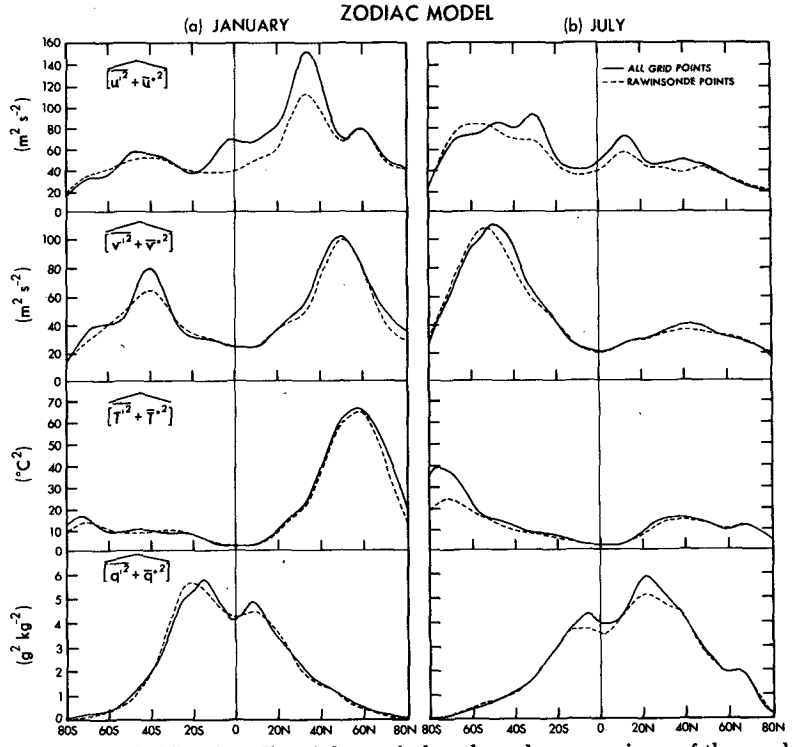

FIG. 11. Meridional profiles of the vertical- and zonal-mean variance of the zonal wind component ($m^2 s^{-2}$), the meridional wind component ($m^2 s^{-2}$), the temperature ($°C^2$), and the specific humidity ($g^2 kg^{-2}$) for the full ZODIAC model fields (solid) and for the rawinsonde simulated ZODIAC model fields (dashed) for January (a) and July (b).

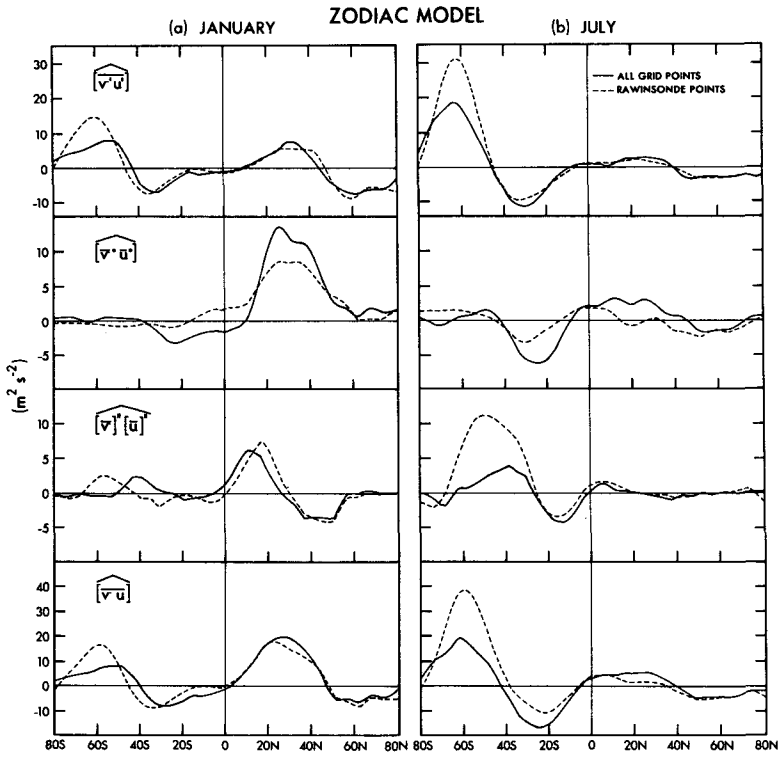

FIG. 12. Meridional profiles of the vertical- and zonal-mean momentum flux due to transient eddies, due to stationary eddies, due to mean meridional circulations and due to all types of motions for the full ZODIAC model fields (solid) and for the rawinsonde simulated ZODIAC model fields (dashed) for January (a) and July (b). Units are in $m^2 s^{-2}$.

port of energy. In the Southern Hemisphere one recognizes the same maxima in transport in the lower troposphere in the two analyses, but the rawinsonde analysis tends to lead to an underestimate of about 30% in SH transports at middle and high latitudes. Nevertheless, the overall agreement in this eddy component of the energy transport appears satisfactory. The present conclusion is in sharp contrast with our earlier statement about the contributions by mean meridional circulations which were discussed in connection with Fig. 6. The values of these last transports based on the SH rawinsonde network were found to be unusable for SH budget studies.

c. Comparison of vertical and zonal mean parameters

For an overall picture of the errors caused by spatial gaps in the station network a more complete set of atmospheric variables will be discussed now, but in a much simplified form. Thus first the three-dimensional fields have been averaged in both the zonal and vertical directions, so that only a latitudinal dependence remains. For each quantity two curves will be shown as a function of latitude, a solid curve for the full model or control case, and a dashed curve for the rawinsonde model case. The comparison of the two curves will yield a coarse measure of the adequacy of the rawinsonde network to measure one of the essential parameters of the full model atmosphere. The curves for the linear quantities, the transient plus standing eddy variances, the kinetic energy, the northward momentum transport, the northward water vapor transport and the northward energy transport are shown in Figs. 9–14, respectively. A close correspondence between the solid and dashed curves does not imply a close correspondence between meridional cross sections such as shown for example in the previous subsection, nor will it guarantee agreement between the horizontal fields as given by the rms errors in Section 7a. Nevertheless, these gross averages constitute a simple and useful measure, and supply additional information concerning the nature of errors to be expected also in real atmospheric analyses. If the present comparisons show a discrepancy, one can be fairly certain that much more serious discrepancies occur in the three-dimensional fields. In either case the reader should go back to the tables or figures in previous sections for further information.

The zonal wind component in Fig. 9 shows the same problems as discussed before in connection with Fig. 5, namely, too strong easterlies in the inner

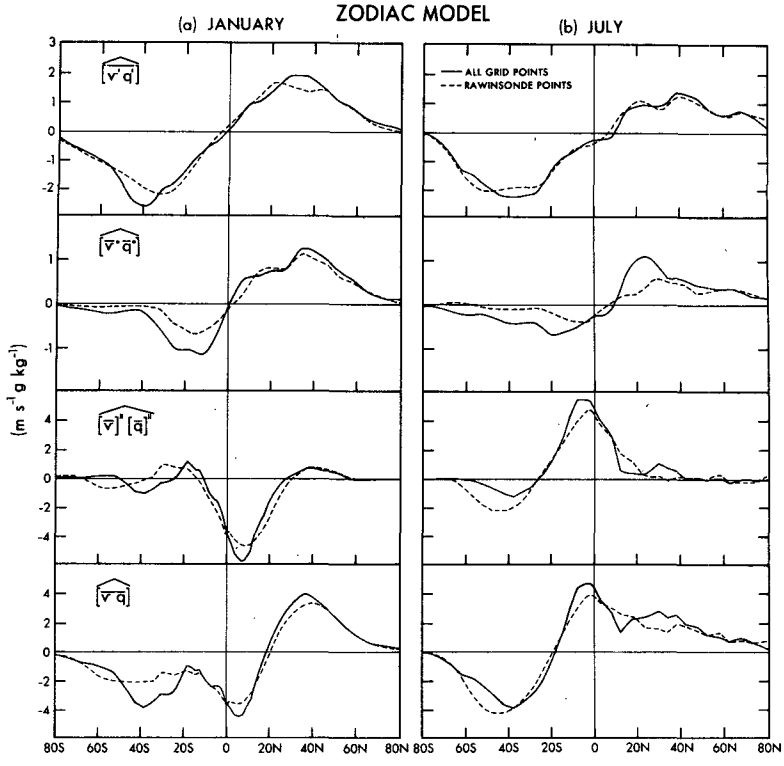

FIG. 13. Meridional profiles of the vertical- and zonal-mean water vapor flux due to transient eddies, due to stationary eddies, due to mean meridional circulations, and due to all types of motions for the full ZODIAC model fields (solid) and for the rawinsonde simulated ZODIAC model fields (dashed) for January (a) and July (b). Units are in m s^{-1} g kg^{-1}.

tropics and a too much widening of the SH jet. The curves for temperature and geopotential height and humidity agree well and seem therefore rather insensitive to the station distribution.

The curves for the kinetic energy in Fig. 10 illustrate well the difference in present ability to measure transient eddy and stationary eddy contributions. The transient eddy part is well simulated but the stationary part is grossly underestimated. This is especially true for the zonal component, as will be shown in the next figure. The errors made in the total kinetic energy include, of course, those made in the zonal mean component. Based on the errors found in the cross sections of the zonal wind itself in Fig. 5 one can explain the spreading out of the profile of total kinetic energy in the Southern Hemisphere mainly by the errors made in the mean component.

In the case of the variance of the horizontal wind components in Fig. 11, the deficiencies in the rawinsonde network to adequately represent the stationary eddies are again evident, especially in the zonal component. On the other hand, the meridional distributions of the temperature and humidity variances are reproduced fairly well.

The momentum transport curves in Fig. 12 bring out some new interesting points. Considering first the transient eddy component the discrepancies are relatively small, except near 60°S where no rawinsonde stations are available in a 10° latitude wide zone. Large differences are found in the standing eddy fluxes at all latitudes in both hemispheres. In January even the sign of the equatorial flux is in error, i.e., the flux is not from winter to summer hemisphere at it probably should be. The mean meridional circulation flux in middle latitudes of the Southern Hemisphere is greatly in error, especially in July. The reason for this is the too intense, indirect Ferrel cell in the rawinsonde case as depicted before in Fig. 6. In summary, the total transport of momentum appears reasonably well measured in the Northern Hemisphere and rather poorly in the Southern Hemisphere especially during the winter season.

The graphs of the transport of water vapor in Fig. 13 show deficiencies similar to those discussed for momentum. However, the discrepancies are somewhat smaller because the water vapor field is a much more slowly varying and more uniform field than the momentum field. The poleward transient eddy trans-

TABLE 4. Hemispheric and global mass-weighted averages of various parameters for the full ZODIAC model and in parentheses for the rawinsonde network simulation of the ZODIAC model.

Parameter	January Northern Hemisphere	January Southern Hemisphere	January globe	July Northern Hemisphere	July Southern Hemisphere	July globe	Units
$\bar{u}$	9.56 (9.32)	5.19 (5.11)	7.38 (7.22)	2.08 (1.66)	8.82 (9.13)	5.45 (5.40)	m s^{-1}
$\bar{u}\cos\phi$	7.35 (7.15)	3.92 (3.64)	5.64 (5.40)	1.33 (0.91)	6.73 (6.93)	4.03 (3.92)	m s^{-1}
$\bar{T}$	−23.73 (−23.76)	−19.11 (−19.10)	−21.42 (−21.43)	−15.41 (−15.36)	−23.69 (−23.85)	−19.55 (−19.60)	°C
$\bar{z}$†	−103 (−103)	11 (11)	−46 (−46)	100 (101)	−95 (−98)	2. (1)	gpm
$\bar{q}$	1.97 (1.95)	2.64 (2.63)	2.31 (2.29)	3.42 (3.41)	1.92 (1.89)	2.67 (2.65)	g kg^{-1}
MKE	103.6 (98.9)	48.4 (43.6)	76.0 (71.2)	13.9 (12.7)	105.9 (105.6)	59.9 (59.2)	m^2 s^{-2}
SEKE	31.0 (20.6)	9.9 (5.2)	20.4 (12.8)	13.4 (9.9)	13.4 (5.8)	13.4 (7.8)	m^2 s^{-2}
TEKE	41.0 (40.6)	35.8 (36.9)	38.4 (38.8)	26.8 (26.7)	49.2 (52.4)	38.0 (39.5)	m^2 s^{-2}
KE	175.6 (160.1)	94.1 (85.7)	134.8 (122.9)	54.1 (49.3)	168.5 (163.8)	111.3 (106.6)	m^2 s^{-2}
$\overline{u'^2}$	43.3 (42.9)	35.2 (37.5)	39.2 (40.2)	29.7 (29.4)	47.6 (51.9)	38.6 (40.6)	m^2 s^{-2}
$\overline{v'^2}$	38.7 (38.4)	36.4 (36.3)	37.6 (37.4)	23.9 (24.0)	50.9 (52.9)	37.4 (38.4)	m^2 s^{-2}
$\overline{T'^2}$	10.1 (9.8)	5.8 (6.5)	7.9 (8.2)	4.7 (4.6)	8.4 (7.8)	6.5 (6.2)	°C^2
$\overline{q'^2}$	1.62 (1.62)	2.22 (2.35)	1.92 (1.99)	2.82 (2.83)	1.53 (1.57)	2.81 (2.20)	g^2 kg^{-2}
$\bar{u}^{*2}$	47.0 (28.8)	13.3 (5.8)	30.2 (17.3)	19.8 (13.8)	18.5 (7.5)	19.2 (10.6)	m^2 s^{-2}
$\bar{v}^{*2}$	14.9 (12.4)	6.5 (4.5)	10.7 (8.4)	7.1 (6.0)	8.4 (4.1)	7.7 (5.0)	m^2 s^{-2}
$\bar{T}^{*2}$	16.9 (15.7)	2.5 (1.7)	9.7 (8.7)	5.0 (4.5)	3.2 (1.5)	4.1 (3.0)	°C^2
$\bar{z}^{*2}$	7780 (7350)	939 (527)	4360 (3938)	2070 (1870)	1190 (493)	1630 (1182)	gpm^2
$\bar{q}^{*2}$	0.85 (0.79)	0.95 (0.86)	0.90 (0.82)	1.09 (0.88)	0.60 (0.46)	0.85 (0.67)	g^2 kg^{-2}

† Departure from standard atmosphere.

ports appear well measured, while the stationary eddy transports are consistently too weak. The errors in the mean meridional transports reflect again the difficulties in measuring the mean meridional streamflow. Because the SH Ferrel cell is too strong in the rawinsonde model, its lower branch transports too much water vapor toward the equator. The shape of the total transport curves at the bottom of Fig. 13 show reasonable correspondence over the entire domain of latitude.

Finally, the energy fluxes presented in Fig. 14 are subject to errors resembling those in the momentum and water vapor fluxes. However, in this case the mean meridional circulation estimates seem to be even worse than for the earlier quantities. One can make here the general statement that reasonable global energy flux estimates are not possible based on the present rawinsonde network alone, and that additional Southern Hemisphere data are absolutely essential.

d. Comparison of hemispheric and global mean parameters

In this final subsection the errors in the parameters will be discussed after they have been area averaged over one hemisphere or over the entire globe and mass-weighted averaged in the vertical. This discussion may throw some light on the question of reliability of "observed" climatic trends in the real atmosphere. Through the various averaging processes in time and space, of course, most errors become progressively smaller going from local to global mean estimates. Nevertheless, sizeable errors still remain as will be seen next. Both the control values as well as the rawinsonde model estimates (in parentheses) are presented in Table 4 for many climatic variables for the months of January and July. The difference between the control and estimated values gives the desired error estimate for the particular variable.

In case of the zonal wind component the January NH value is underestimated by 0.24 m s^{-1}, while in July the NH estimate is 0.42 m s^{-1} too low. Going back to the meridional profiles in Fig. 9, the reason seems to be that almost all NH values at latitudes south of 55°N have been underestimated. In case of the product of the zonal wind component and $\cos\phi$, which is a measure of the relative angular momentum, the NH discrepancy in July becomes even worse because more weight is given to the contributions at low latitudes. It is interesting to note that through a compensation of errors in both hemispheres, the global underestimate is only about 1%.

In view of the recent interest in global temperature trends the error in monthly mean temperature was also computed for four additional winter months. The results are shown in Table 5. The probable error is found to be about 0.05°C or less for the Northern Hemisphere, 0.1 to 0.2°C for the Southern Hemisphere, and 0.1°C for the globe. The actual observed year-to-year "trends" are typically on the order of 0.1°C per year with superimposed random month-to-month variations of about 0.2°C (Starr and Oort, 1973; Angell and Korshover, 1975). Therefore, one may tentatively conclude that global trends in temperature

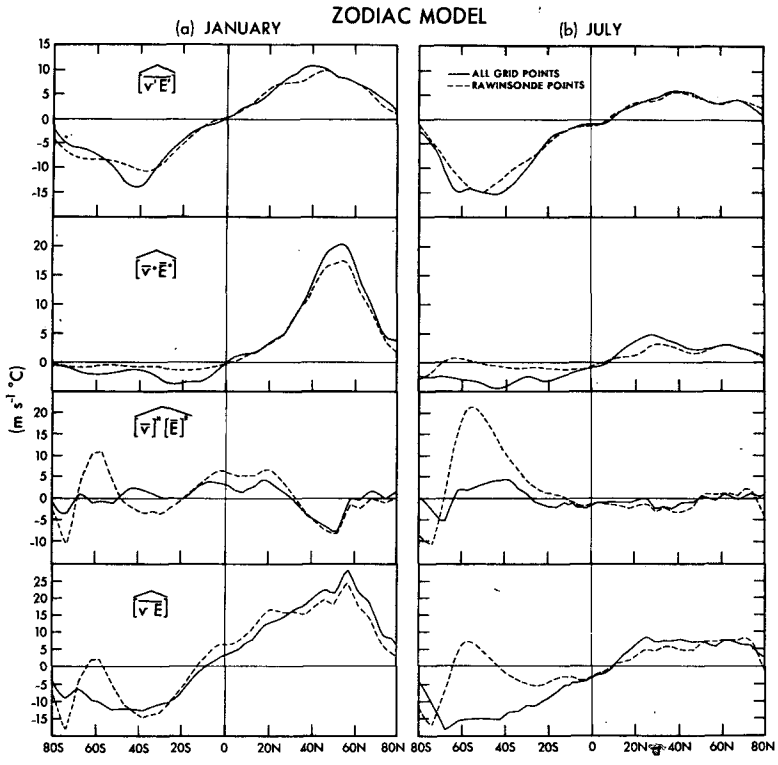

FIG. 14. Meridional profiles of the vertical- and zonal-mean energy flux due to transient eddies, due to stationary eddies, due to mean meridional circulations, and due to all types of motions for the full ZODIAC model fields (solid) and for the rawinsonde simulated ZODIAC model fields (dashed) for January (a) and July (b). Units are in m s^{-1} °C.

can indeed be determined based on the present rawinsonde network.

In case of the geopotential height the errors are about 1 m or less, and in case of the specific humidity about 0.02 g kg^{-1}.

The estimates for the kinetic energy components give percentage errors of about −5 to −10% for the mean component, −30 to −60% for the stationary eddy component, −1 to +6% for the transient eddy component, and −5 to −10% for the total energy. Thus it is found that the relative error in the stationary eddy component is by far the largest.

The transient eddy variance statistics as a group seem to show some tendency to be underestimated in the Northern and overestimated in the Southern Hemisphere. In absolute value, the percentage errors are a few per cent in the Northern and up to 10% in the Southern Hemisphere.

Finally, the stationary eddy variance statistics show again very large errors. In the zonal wind component they are of the order of −40 to −60%, in the meridional component −20 to −50%, in the temperature −10 to −50%, in the geopotential height −10 to −60%, and in the specific humidity −10 to −20%.

8. Summary of conclusions and final remarks

In this paper the problem investigated was to what extent the present distribution of rawinsonde stations is adequate to deduce the atmospheric structure and its variability in space and time over the globe. With the aid of numerical output from a two-year integration of a global climate model, it was shown how large the influence of temporal and spatial data gaps would be on atmospheric model statistics, and by inference also on real atmospheric statistics. The tests on spatial data gaps were performed by witholding information at all grid points more than 200 km away from any rawinsonde station, and by assigning new interpolated values to these grid points through an objective analysis scheme. To simulate the real atmospheric situation as well as possible, the same statistical treatment of the model data at each "rawinsonde" grid point, as well as the same objective analysis methods were used as in previous real atmospheric analyses. Various comparisons between the results of the simulated model version with those of the full model provided a direct measure of the deficiencies of the rawinsonde network over different parts of the globe.

Besides the errors related to spatial data gaps, estimates were provided of the uncertainty due to other sources of errors. Discussed were also instrumental errors in the basic daily soundings (Section 4), errors due to interpretation of the local soundings as being representative for large-scale conditions (Section 4), deficiencies of the particular objective analysis technique used (Section 5), and errors due to gaps in the time series at each station (Section 6). The uncertainties caused by the analysis technique itself are, of course, closely related to the distance between rawinsonde stations and thus to the data gaps. However, it was argued that the use of any other, reasonable analysis technique would not materially have changed the results of the spatial data gap tests. Of the other errors, the possible nonrepresentativeness of the local sounding due to microscale or mesoscale variabilty was found to be the next most important source of error.

Let us list some of the results of the spatial data gap experiments:

1) Typical rms wind errors in $\bar{u}$ and $\bar{v}$ averaged over a hemisphere are 2–3 m s^{-1}, increasing for $\bar{u}$ to 5–6 m s^{-1} near jet stream levels. Temperature errors in $\bar{T}$ are on the order of 0.5–1°C in the free atmosphere and 2–3°C in the surface boundary layer. Errors in $\bar{z}$ in the upper troposphere range between 20 and 30 gpm. Errors in mean humidity decrease with height from 1–2 g kg^{-1} near the surface to 0.3 g kg^{-1} at 500 mb.

2) Rms errors in the Southern Hemisphere are somewhat larger than in the Northern Hemisphere. The NH rms errors are noticeably smaller in summer than in winter but little seasonal variation is found in the SH.

3) Meridional cross sections showed clearly the problems in the Southern Hemisphere. The simulated SH jet stream in January was excessively spread out toward high latitudes into the almost data-void latitude belt at 50°–70°S. The most serious problems were found in the meridional streamfunction. In the Northern Hemisphere only slight differences in intensity and orientation of the cells were computed. However, in the Southern Hemisphere the indirect Ferrel cell was found about 10° to 15° too far south and a factor 2 or 3 too strong. This discrepancy makes the computation of SH mean meridional cell transports on the basis of rawinsonde stations practically impossible.

4) As regards the eddy variances and covariances, the transient components were generally simulated quite well over the entire globe probably because of their zonally uniform character. Although the stationary eddy components are highly unreliable over the Southern Hemisphere, this is not such a serious problem in the overall balances because their contributions are small compared to the transient ones. It was rather surprising to find that in the Northern Hemisphere the stationary kinetic energy was underestimated by about 30% in spite of the better network. This result may vindicate some of the model results of the stationary eddies that were thought to be too intense compared to the "real" observed values.

5) A table comparing hemispheric and global mean, estimated values with the true model values gave the first quantitative information on the extent to which various climatic indices, such as the global temperature trend, can be expected to be measurable in the actual atmosphere.

6) Additional data sources are essential to define the fluxes by the mean meridional and stationary eddy circulations in the Southern Hemisphere. New rawinsonde stations, aircraft soundings, and/or satellite derived wind statistics should be obtained over the data void regions in order that reliable global budget studies will become possible.

TABLE 5. Temperature difference (°C), rawinsonde network simulation minus full ZODIAC model, averaged over entire mass (below 75 mb) of Northern Hemisphere, Southern Hemisphere and globe.

Month	Northern Hemisphere	Southern Hemisphere	Globe
January, year 1	0.01	−0.01	0.00
February, year 1	−0.03	0.17	0.07
December, year 1	−0.03	0.09	0.03
January*, year 2	−0.03	0.01	−0.01
February, year 2	0.00	0.19	0.10
July*, year 2	0.05	−0.16	−0.05
Average±standard deviation	−0.00±0.03	0.05±0.12	0.02±0.05

* Only January and July for the second year of integration were used in the calculations described earlier in the paper.

Finally, the following points may be mentioned:

1) Tests similar to those described in Section 7, but for five different winter months of the ZODIAC model, have been performed in an earlier study (Oort, 1977). These tests made an evaluation possible of our present ability to measure, besides the long-term mean, also the range of the interannual variability. It was shown that in the NH, based on data from the actual rawinsonde network, such an evaluation is indeed possible, and that it can be expected to give realistic results for most large-scale circulation parameters. In other words, the station distribution seemed adequate in the NH to catch interannual shifts in the long-wave patterns.

2) In a recent paper it was suggested that the oceans play a more crucial role in the heat balance

of the earth than was believed previously (Oort and Vonder Haar, 1976). For example, residual calculations showed a very large seasonal variation in the oceanic heat transport reaching near 10°N a maximum amplitude of 3.6×10^{15} W in the annual component and 3.2×10^{15} W in the semiannual component. Of course, the errors in the measured individual components of the heat balance (i.e., the total energy transport required by satellite radiation measurements, the transport by the atmosphere, and the atmospheric and oceanic rates of heat storage) will all affect the residual to make it less reliable. The results presented in Fig. 14 of the present paper throw some light on one possible source of errors, namely the determination of the atmospheric energy transport. At low latitudes this transport is dominated (both in the model and in the real world) by the mean meridional Hadley circulation which is generally thought to be difficult to measure. When converted to the same units, the graphs in Fig. 14 show that the rawinsonde network tends to overestimate the northward heat transport in the atmosphere in January by about 1×10^{15} W, and underestimate it in July by about the same amount. Although these values are higher than the 95% confidence limits of 0.5×10^{15} W estimated on other grounds in the earlier paper, it is clear that the large seasonal cycle in the residual ocean transport cannot be explained by errors in the evaluation of the atmospheric transport.

3) Another extension of the present work would be to analyze the influence of data gaps on daily maps instead of on monthly mean maps. However, in that case, the present analysis technique would not be useable because of the high wavenumber components present on a daily map. Another technique, for example, a four-dimensional analysis scheme as described by Miyakoda *et al.* (1976) or Simmonds (1976) or a multivariate statistical scheme as described by Schlatter *et al.* (1976), could be used. Furthermore, the analysis could be extended to include more complicated parameters, such as energy transformation and generation rates.

Acknowledgments. The author is grateful to Syukuro Manabe for making available the output of the ZODIAC model, the basic ingredient of the present study. Ian H. Simmonds' comments and suggestions on the original manuscript are much appreciated. Finally, the author would like to acknowledge the support of Melvin Rosenstein in programming, Mrs. Hilda Philander and Ms. Laura Cilino in various calculations, Philip G. Tunison and William H. Ellis in drafting and Mrs. Betty M. Williams in typing the manuscript.

APPENDIX

List of Symbols and Definitions

c_p specific heat at constant pressure

E $c_pT+gz+Lq=$ total energy per unit mass

EKE eddy kinetic energy per unit mass

$=$TEKE$+$SEKE

$\{=([\overline{u'^2}]+[\overline{v'^2}])/2+([\bar{u}^{*2}]+[\bar{v}^{*2}])/2\}$

g acceleration due to gravity

KE kinetic energy per unit mass $[=(u^2+v^2)/2]$

L heat of condensation

MKE mean kinetic energy per unit mass $[=([\bar{u}]^2+[\bar{v}]^2)/2]$

p pressure

p_0 pressure at ground level (where there are no mountains $p_0=1012.5$ mb)

p_t top level of vertical integration ($=75$ mb)

q specific humidity (usually in units of grams of water vapor per kilogram of moist air)

t time

T temperature

u zonal wind component (positive if eastward)

v meridional wind component (positive if northward)

z geopotential height

λ geographic longitude

ϕ geographic latitude

$\psi(p)$ meridional streamfunction

$$\left[=\int_p^{p_0}\int_0^{2\pi}\bar{v}a\cos\phi d\lambda dp/g\right]$$

$\bar{A}$ time average of A

$$\left[=(t_2-t_1)^{-1}\int_{t_1}^{t_2}A dt\right]$$

A' departure from time average of A $[=A-\bar{A}]$

$[A]$ zonal average of A $\left[=(2\pi)^{-1}\int_0^{2\pi}A d\lambda\right]$

A^* departure from zonal average of A $[=A-[A]]$

$\hat{A}$ mass weighted "vertical" average of A

$$\left[=(p_0-p_t)^{-1}\int_{p_t}^{p_0}A dp\right]$$

A'' deviation from "vertical" average A $[=A-\hat{A}]$.

Examples of nomenclature follow:

1. $[\overline{A^2}]=[\overline{A'^2}]+[\bar{A}^{*2}]+[\bar{A})]^2$ where

$[\overline{A'^2}]$ = variance of A resulting from transient eddies

$[\bar{A}^{*2}]$ = variance of A resulting from standing (or stationary) eddies

2. $[\overline{vA}] = [\overline{v'A'}] + [\bar{v}^*\bar{A}^*] + [\bar{v}][\bar{A}]$ where

$[\overline{v'A'}]$ = meridional transport of A resulting from transient eddies

$[\bar{v}^*\bar{A}^*]$ = meridional transport of A resulting from standing eddies

$[\bar{v}][\bar{A}]$ = meridional transport of A resulting from (standing) mean meridional circulations.

REFERENCES

Angell, J. K., and J. Korshover, 1975: Estimate of the global change in tropospheric temperature between 1958 and 1973. *Mon. Wea. Rev.*, **103**, 1007–1012.

Baer, F., and J. J. Tribbia, 1976: Spectral fidelity of gappy data. *Tellus*, **28**, 215–227.

Bauer, K. G., 1976: A comparison of cloud motion winds with coinciding radiosonde winds. *Mon. Wea. Rev.*, **104**, 922–931.

Bruce, R. E., L. D. Duncan and J. H. Pierluissi, 1977: Experimental study of the relationship between radiosonde temperatures and satellite-derived temperatures. *Mon. Wea. Rev.*, **105**, 493–496.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374.

Eddy, A., 1967: The statistical objective analysis of scalar data fields. *J. Appl. Meteor.*, **6**, 597–609.

Gandin, L. S., 1963: Objective analysis of meteorological fields. Israel Program for scientific translation, 1965 [NTIS TT-65-50007].

Harris, R. G., A. Thomasell, Jr. and J. G. Welsh, 1966: Studies of techniques for the analysis and prediction of temperature in the ocean. Part III: Automated analysis and prediction. Interim Report, prepared by Travelers Research Center, Inc., for the U.S. Naval Oceanographic Office, Contract N62306-1675, 97 pp.

Hayashi, Y., and D. G. Golder, 1977: Space-time spectral analysis of mid-latitude disturbances appearing in a GFDL general circulation model. *J. Atmos. Sci.*, **34**, 237–262.

Holloway, J. L., Jr., and S. Manabe, 1971: Simulation of climate by a global general circulation model, Part I, Hydrological cycle and heat balance. *Mon. Wea. Rev.*, **99**, 335–370.

Leary, C., and R. O. R. Y. Thompson, 1973: Shortcomings of an objective analysis scheme. *J. Appl. Meteor.*, **12**, 589–594.

Lenhard, R. W., 1970: Accuracy of radiosonde temperature and pressure-height determination. *Bull. Amer. Meteor. Soc.*, **51**, 842–846.

——, 1973: A revised assessment of radiosonde accuracy. *Bull. Amer. Meteor. Soc.*, **54**, 691–693.

Lorenz, E. N., 1967: The nature and theory of the general circulation of the atmosphere. World Meteorological Organization, No. 218, TP 115, Geneva, Switzerland, 161 pp.

Manabe, S., and J. L. Holloway, Jr., 1975: The seasonal variation of the hydrological cycle as simulated by a global model of the atmosphere, *J. Geophys. Res.*, **80**, 1617–1649.

——, and J. D. Mahlman, 1976: Simulation of seasonal and interhemispheric variations in the stratospheric circulation. *J. Atmos. Sci.*, **33**, 2185–2217.

——, D. G. Hahn and J. L. Holloway, Jr., 1974: The seasonal variation of the tropical circulation as simulated by a global model of the atmosphere. *J. Atmos. Sci.*, **31**, 43–83.

Miyakoda, K., L. Umscheid, D. H. Lee, J. Sirutis, R. Lusen and F. Pratte, 1976: The near-real-time, global, four-dimensional analysis experiments during the GATE period, Part I. *J. Atmos. Sci.*, **33**, 561–591.

Newell, R. E., J. W. Kidson, D. G. Vincent and G. J. Boer, 1972: *The General Circulation of the Tropical Atmosphere and Interactions with Extratropical Latitudes*. Vol. 1. The MIT Press, 258 pp.

——, ——, —— and ——, 1974: *The General Circulation of the Tropical Atmosphere and Interactions with Extratropical Latitudes*. Vol. 2. The MIT Press, 371 pp.

Oort, A. H., 1977: The interannual variability of atmospheric circulation statistics. NOAA Prof. Paper No. 8, U.S. Govt. Printing Office, Washington, D. C., 76 pp.

——, and E. M. Rasmusson, 1971: Atmospheric circulation statistics. NOAA Professional Paper No. 5, U.S. Govt. Printing Office, Washington, D. C. 323 pp. [limited number of copies available from Geophysical Fluid Dynamics Laboratory/NOAA or in microfiche form from NTIS COM 7250295].

——, and T. H. Vonder Haar, 1976: On the observed annual cycle in the ocean-atmosphere heat balance over the northern hemisphere. *J. Phys. Oceanogr.*, **6**, 781–800.

Priestley, C. H. B., and A. J. Troup, 1964: Strong winds in the global transport of momentum. *J. Atmos. Sci.*, **21**, 459–460.

Reeves, R., S. Williams, E. Rasmusson, D. Acheson, T. Carpenter and J. Rasmusson, 1976: GATE convection subprogram data center-analysis of rawinsonde intercomparison data. NOAA Tech. Rep. EDS 20, Washington, D. C., 75 pp.

Sadler, J. C., 1975: The upper tropospheric circulation over the global tropics. Dept. of Meteorology, University of Hawaii, NSF Grant GA-36301, 35 pp.

Schlatter, T. W., G. W. Branstator and L. G. Thiel, 1976: Testing a global multivariate statistical objective analysis scheme with observed data. *Mon. Wea. Rev.*, **104**, 765–783.

Simmonds, I., 1976: Data assimilation with a one-level, primitive equation spectral model. *J. Atmos. Sci.*, **33**, 1155–1171.

Starr, V. P., and A. H. Oort, 1973: Five-year climatic trend for the northern hemisphere. *Nature*, **242**, 310–313.

——, J. P. Peixóto and N. E. Gaut, 1970: Momentum and kinetic energy balance of the atmosphere from five years of hemispheric data. *Tellus*, **22**, 251–274.

Stoldt, N. W., 1971: Effects of missing data on zonal kinetic energy calculations. *Pure Appl. Geophys.*, **92**, 207–218.

Walker, H. C., 1970: Unbiased station grid for study of the kinetic energy balance of the atmosphere. *Pure Appl. Geophys.*, **81**, 313–322.