The Use of COADS Ship Observations in Cloud Climatologies

Carol J. Hahn
Cooperative Institute for Research in Environmental Sciences
Boulder, CO 80309

Stephen G. Warren
Department of Atmospheric Sciences, University of Washington
Seattle, WA 98195

Julius London
Department of Astro Planetary and Atmospheric Sciences
University of Colorado, Boulder, Co 80309

1. Introduction

We have used the surface weather reports from COADS to prepare a climatology of total cloud cover and cloud type amounts over the oceans, for the years 1952-81 [Warren et al., 1988 (hereafter referred to as “the ocean atlas”)]. For that atlas we analyzed about 50 million individual synoptic observations of the “condensed marine reports” (CMR; see Woodruff et al. [1987] for a definition of this and other COADS related terms). The results contained in the ocean atlas, as well as the results in a complementary atlas for land [Warren et al., 1986 (hereafter referred to as “the land atlas”)] are available on magnetic tape [Hahn et al., 1988 (hereafter referred to as “the archive tape”)]. Included in this analysis are long term (30-year) seasonal and monthly averages, interannual variability and trends, diurnal variations, and the annual cycle for 5°x5° (latitude x longitude) grid boxes. (Some ocean results are given at 10x20° only.) Further analysis for the years 1982-91 is currently under way and utilizes the COADS Interim Product CMR5 Reports.

One of the COADS products (MSTG Group 5) includes a monthly summary of average total cloud amount (but not cloud type amounts) for 2°x2° grid boxes for the ocean areas of the globe over the period 1854-1979. One would expect that this total cloud analysis would give results similar to those in our ocean atlas since both climatologies used essentially the same data base. While this is generally true, differences are found in foggy and rainy regions of the globe because of the difference in the way the individual weather reports are processed. The differences and their causes are the subject of Section 2 of this report.

The reliability of cloud observations made by surface observers at night has often come into question. A “night- detection bias” has been documented in several regional studies by comparing the average cloud cover reported at night near the times of the full moon with that for times of little moonlight [Fig. 107 of Sverdrup, 1933; Riehl, 1947; Schneider et al., 1989]. High thin clouds are commonly missed by observers at the surface. Because of this bias, Warren et al. [1986; 19881 used only daytime observations in the analyses for middle and high clouds, and for clear sky frequency. As part of our ongoing study to analyze cloud data for 1982-91, we have undertaken the analysis of the night-detection bias on a global scale. Preliminary results from this analysis are presented in Section 3 of this report.

2. Climatology of total cloud cover: comparison of COADS to Warren et al.

In the preparation of our climatologies, each surface report is put through a series of quality control checks (Figure 1 of the land atlas). Total cloud cover is coded as a number (N) from 0 to 8, indicating eights (octas) of sky cover, or N=9, meaning “sky obscured” due to fog, rain, or other obscuring phenomena. The major difference between our processing and that of COADS is in the treatment of sky obscured observations. In the COADS processing, such reports are rejected for cloud analysis because no specific cloud information is given. We found that this occurs in over 6% of the ship observations. Since this is a fairly large fraction of the reports and since an observation of sky-obscured is often associated with overcast cloud, we check the present weather category (ww) in the synoptic code for the cause of the obscuration. If there is no indication that the obscuration is related to cloudiness, we too discard the report. This occurs about 2.6% of the time. However, if ww indicates rain or snow (1%), thunderstorms (0.2%), or fog (2.4%), we conclude that the sky is overcast with nimbostratus, cumulonimbus, or fog, respectively. Thus we should expect to obtain a greater average total cloud amount than does COADS, especially in regions where fog is common.

To test this, we compared average total cloud cover for the period 1970-79 in the two climatologies. For COADS, the 2-degree monthly summaries of total cloud were summed over all months for the 10-year period and averaged into 10x20° boxes. {The 10x20° grid used here has box sizes of 10x20° between 50°N and 50°S, 10x40 for latitudes 50-70, 10x60 for latitudes 70-80, and 10x360 for latitudes 80-90; we call this the “10c” grid, where the “c” signifies “condensed” boxes in the high latitudes. A l0°x10° box is the smallest into which both 2-degree (COADS) and 5-degree (the smallest available on the archive tape) boxes can fit evenly. However, the 10c grid size, already available on the archive tape, was most convenient for this comparison}. 1 Similarly, seasonal averages (already in 10c boxes) were taken from the archive tape and averaged over the same period. We will call this data set WHL. The COADS total cloud value was then subtracted from the WHL total cloud value in each grid box and plotted on maps.

Figure 1 shows the global (ocean) distribution of the difference between WHL and COADS for annual averages of total cloud cover during the period 1970-79. (Actual cloud amounts are given in the ocean atlas.) There is little or no difference in the equatorial regions, but differences are generally positive in the higher latitudes. The largest differences occur in the western North Pacific.

Figure 2 shows the distribution of the frequency of occurrence of sky-obscured-due-to-fog for June-July-August (JJA) 1952-81. In this season large amounts of obscuring fog (>20%) are seen in the North Pacific above 40°N. In December-January-February (DJF) there is much less fog in this region (<5%), and values in the southern hemisphere between 40-70°S are increased [Warren et al., 1988]. If the WHL - COADS differences in Figure 1 were due solely to our designating fog-obscuration as cloud while COADS does not, then we would expect to see large seasonal variations in the WHL - COADS differences. Difference values in the North Pacific for JJA (not shown) are indeed larger than those shown in Figure 1 and are greatest above 40°N. Difference values for DJF (not shown), however, are still relatively large in the western North Pacific with the large values extending southward to 30°N. These difference values in DJF are not accounted for by the occurrence of obscuring fog.

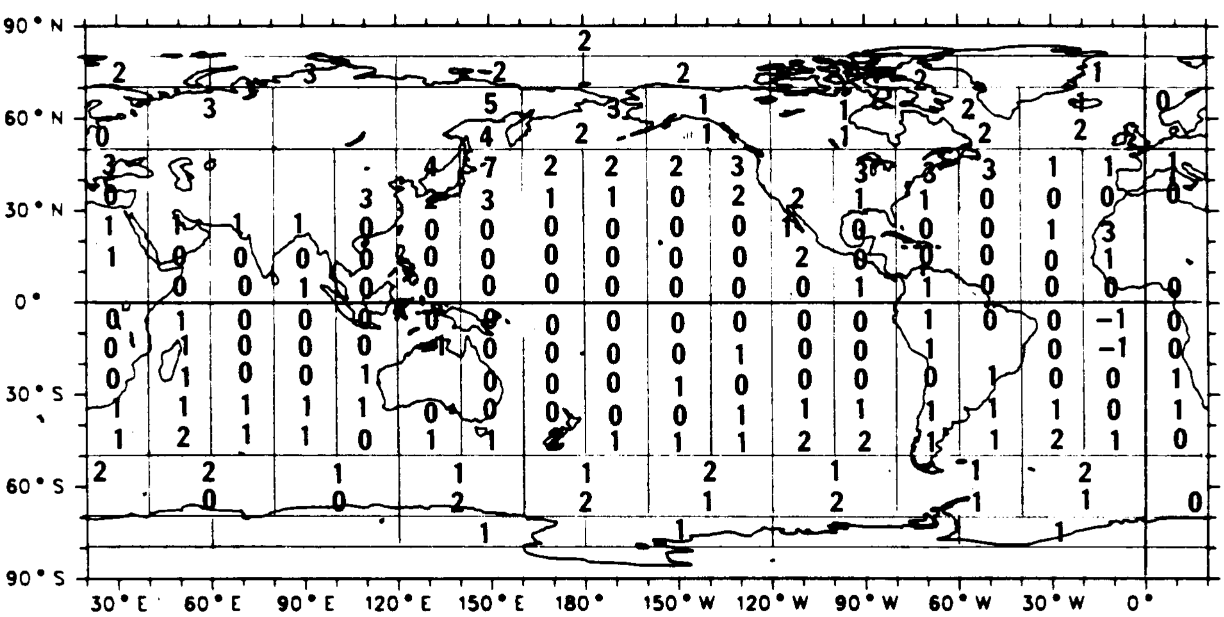

Figure 1. Difference in total cloud cover calculated by the method of WHL and that of COADS (WHL-COADS). Values are in percent of sky cover for the years 1970-1979 (annual average) and are plotted on a 10° grid. The average of the 212 filled 10° boxes is 1%.

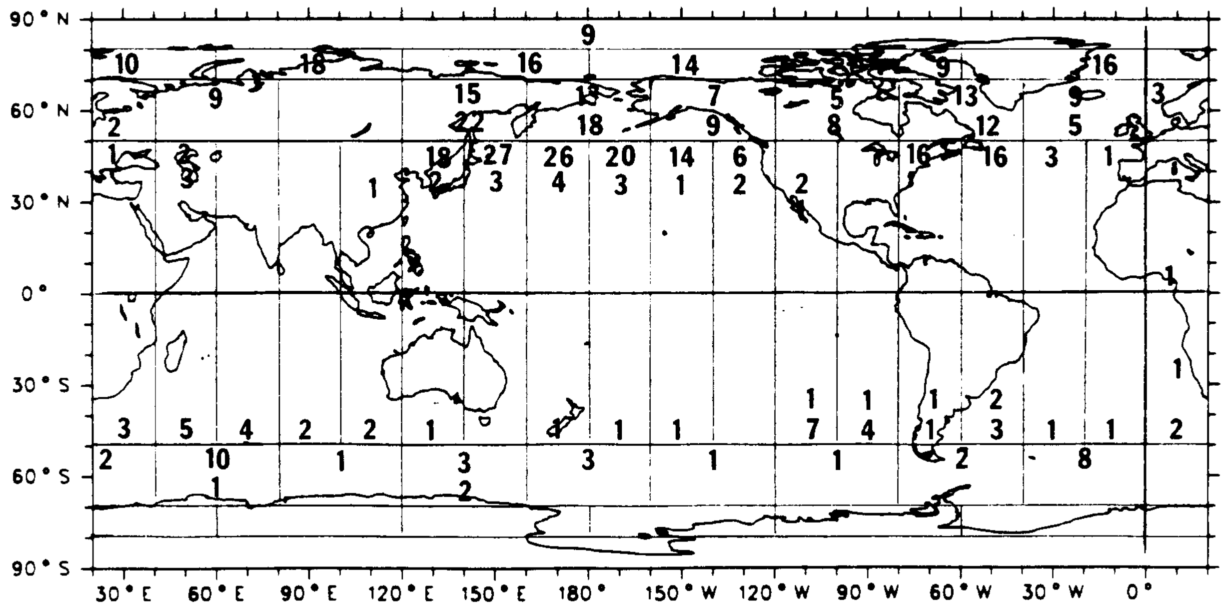

Figure 2. Frequency (%) of sky-obscured-due-to-fog, over the oceans, for JJA, 1952-1981 (from Warren et al. [1988]). Only non-zero values are plotted. Global average (ocean only) is 1%.

In our current analysis of 1982-91, we find that the frequency of occurrence of precipitation (ww codes 50-99, excluding 76 and 78) has a seasonal peak opposite to that of fog. In the North Pacific, precipitation frequencies exceed 20% in January but are less than 15% in July (globally, about 15% of the observations of precipitation also have sky obscured). Furthermore, the zone of high precipitation frequencies extends southward to 30°N. These characteristics are in accordance with the distribution of difference values in the North Pacific in DJF. Thus the major differences between the WHL and COADS total cloud analyses are consistent with the spatial and temporal distributions of sky-obscured due to fog and precipitation.[1]

3. Effect of moonlight on reported cloud amounts at night

The initial analysis of the night-detection bias shown here was performed on data within 10-degree latitude zones between the equator and 50°N (where there is a large volume of data and a latitudinal gradient of cloud amounts), for the month of December (when the moon in the northern hemisphere reaches its highest altitudes at night), and over the years 1983-88. The analysis was done for both land and ocean.

An analysis of nighttime data requires a definition of “night”. For purpose of this study, the definition of night will be related to the solar altitude (during twilight) below which the sky is too dark for human observers to distinguish clouds adequately. An initial estimate, based on our own visual observations, was that this occurred when the sun was approximately 8° below the horizon. This is confirmed by the analysis which follows.

The flux of light from the night sky due to moonlight is proportional to the product of the phase of the moon and the sine of the lunar altitude (represented below as P•sin(LA)). The phase varies from 0.0 for new moon to 1.0 for full moon. With the use of solar and lunar ephemerides (program obtained from R.F. Stephenson, University of Durham, U.K.), a value representing sky brightness was determined for each weather report, given the latitude, longitude, time and date of the report. For each latitude zone, the data were placed into 13 bins based on sky brightness. These bins, numbered 0-12, are defined in Table 1, which also gives the total number of observations that went into each bin for both ocean and land. Bins 11 and 12 represent the sunlit sky. Bin 12 contains the cases with the sun above the horizon while bin 11 contains the twilight cases. Bins 0-10 represent the night sky. The case of the moon below the horizon is contained in bin 0. The case of a full moon overhead corresponds to the upper boundary of bin 10, while a full

---

[1] Since the ship reports are not evenly distributed in either space or time, the method of averaging smaller box values into larger box values can affect the results somewhat due to various sampling biases. The method used here was simply to average all observations within a given 10c box for the 10-year period. This method, though it may not give the best geographical average in some regions, provides the most direct way of comparing the handling of the observations themselves, which is what we want here. But even this is not ideal for the present comparison because the “observations” in WHL are so-called “compressed” observations. A “compressed observation” is the average of all ship reports made within a 5c box within a single 3-hour time slot as discussed in the ocean atlas. It may be made up of one ship report or many of them. While this has advantages in forming averages for a climatology, it makes the present comparison less direct. For example, the negative values seen in Figure 1 between 0-20°S and 0-20°W are a consequence of a sampling bias.

moon at an altitude of 45° corresponds to the lower boundary of this bin. The nighttime bins are not equal in width because finer resolution is needed at the lower end of the brightness scale.

Average total cloud cover was computed for each bin of sky brightness within each 10-degree zone and plotted in Figure 3 (for ocean) and Figure 4 (for land). Generally, these curves show a fairly rapid increase in reported cloud amount as night sky brightness increases up to some threshold, after which there is little or no further change. This implies that when the sky is brighter than that threshold, total cloud cover can be adequately determined. Similar curves (not shown) were also prepared for the frequency of occurrence of low, middle, and high cloud levels and for the frequency of occurrence of clear sky. High and middle clouds show a pattern similar to that shown for total cloud, but the effect is small for low clouds, especially over land. (This supports our previous policy of using night observations for low clouds but not for middle and high clouds [Warren et al., 1986; 1988].) The reported frequency of clear sky decreases with sky brightness, as expected. These analyses were repeated for the months of June and March and similar patterns emerged.

Examination of all these results lead to the selection of P•sin(LA) = 0.20 as the threshold of night sky brightness above which clouds can be adequately detected. (This can be achieved by a full moon at an altitude of 11.5° or a partial moon at higher altitude.) This value corresponds to a sky flux of 0.6 mW/m^{-2} which, in turn, is the same as the flux from the twilight sky at a solar altitude of -8° (Fig. 4-4 of Meinel and Meinel [1983]). This supports the choice we made earlier to define “night” as the time when the sun is more than 8° below the horizon. We therefore, recommend these conditions, solar altitude > -8° or P•sin(LA) > 0.2, as the sky brightness criteria for the adequate detection of clouds by observers at the surface. (These are average values which we will use in our global climatology. For regional studies, a lower threshold might be appropriate over snow-covered surfaces.)

In Figures 3 and 4, the portions of the nighttime curves that lie beyond the brightness threshold can now be considered to represent the true nighttime cloudiness. Comparison of these values with the daytime values shown on the right of the figures gives an indication of the diurnal variation of cloud cover. On average over the five zones, the cloud cover is greater during the day than at night, by about 2% over ocean and nearly 4% over land. Diurnal variations are also evident in the curves (riot shown) for the frequency of clear sky and the frequencies of the low, middle and high level clouds, except that the day-night difference in the frequency of middle clouds was small over land.

Table 3. Sky Brightness Bins and the Number of Cloud Observations in the Zone 0-50N for December, 1983-1988.

Bin Boundaries			Ocean		Land	
	Lower	Upper	Number of Obs	%	Number of Obs.	%
Sun						
12 Altitude	-0.25°	90.00°	241701	53.5	1514930	44.4
11 Altitude	-8.00°	-0.25°	22854	5.1	204515	6.0
Moon (Sun <-8°)						
10 P•sin(LA)	0.70	1.00	26215	5.8	209813	6.1
9 P•sin(LA)	0.50	0.70	19364	4.3	183062	5.4
8 P•sin(LA)	0.35	0.50	13200	2.9	121546	3.4
7 P•sin(LA)	0.25	0.35	8453	1.9	77484	2.3
6 P•sin(LA)	0.20	0.25	4304	0.9	39530	1.2
5 P•sin(LA)	0.15	0.20	4578	1.0	41755	1.2
4 P•sin(LA)	0.10	0.15	4750	1.0	45147	1.3
3 P•sin(LA)	0.05	0.10	5770	1.3	52724	1.5
2 P•sin(LA)	0.02	0.05	4316	1.0	40499	1.2
1 P•sin(LA)	0.00	0.02	4329	1.0	41166	1.2
0 P•sin(LA)	-1.00	0.00	91991	20.3	841730	24.6
Total			451825	100.0	3413901	100.0
Sunlight (bins 11-12)			264555	58.6	1719445	50.4
Light Night (bins 6-10)			71536	15.8	631435	18.5
Dark Night (bins 0-5)			115734	25.6	1063021	31.1

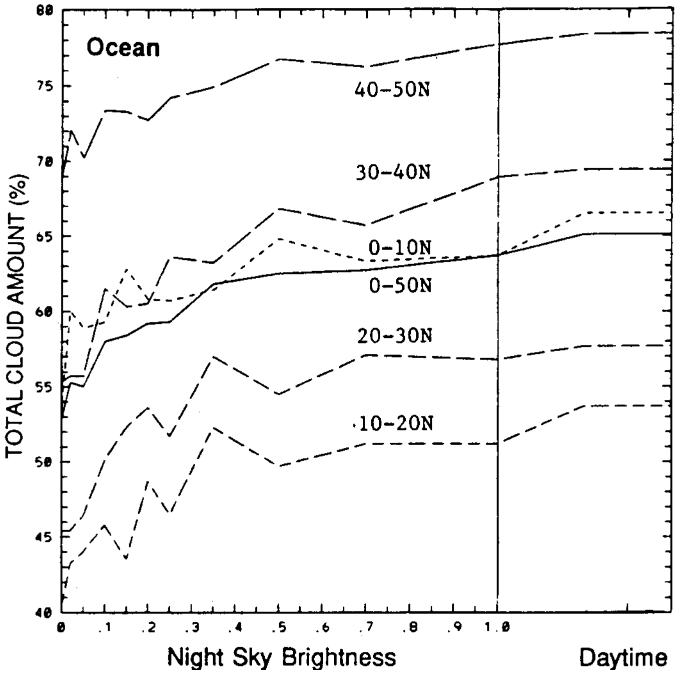

Figure 3. Reported total cloud amount as a function of night sky brightness for 10-degree zones (ocean only) and their average (0-50°N) for December, 1983-1988. Night sky brightness is represented as: luna Phase x sin(lunar Altitude). Brightness values are shown for the upper boundaries of the bins (bin sizes are defined in Table 1). For comparison, the right side of the figure shows the average cloud amount with the sun above the horizon

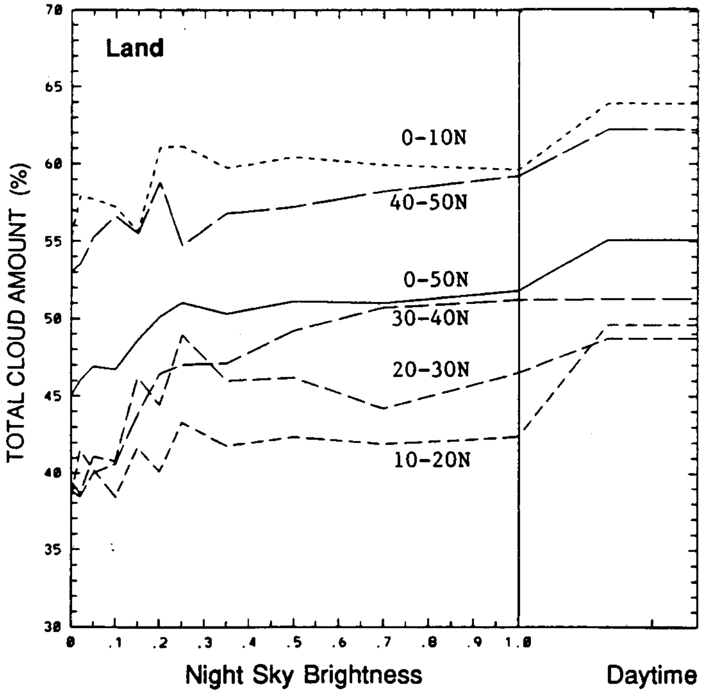

Figure 4. Same as Fig. 3, but for land.

The lower portion of Table 1 shows that the percentage of observations made during the sunlit periods in December was 58.6% over ocean and 50.4% over land. The nighttime observations are divided into the cases in which the night sky brightness criterion is met (light night) and in which it is not (dark night). Of the nighttime observations, 62% will be discarded by the application of the sky brightness criteria. For June (not shown), when the sun is above the horizon in the Northern Hemisphere for a larger portion of the day, only about 20% of the total number of observations will be discarded although these make up 72% of the nighttime observations for 0-50°N.

The sky brightness criteria determined in this study will be applied in our cloud analysis for 1982-91. Since this study shows that the average total cloud computed for night under the light night conditions is about 4% greater than that computed from all night observations, our computed diurnal average total cloud amounts will probably be about 2% greater (and clear sky frequency about 1-2% less) than that computed in our previous work for which we used all observations. Climatologies based on daytime observations only would be a percent or so higher than we would now estimate for the diurnal average. Upon completion of the analysis for 1982-91, we will be able to make more definitive statements about the diurnal variations of not only total cloud cover, but also cloud type amounts, for all land and ocean regions on the earth.

4. Recommendations for COADS

(1) Reports of sky-obscured due to precipitation or fog should be interpreted as overcast cloud. (For immediate use in weather prediction, there may be reasons not to classify fog as a cloud, but for climate studies the radiative effects of fog are those of a cloud.) This is a fairly straightforward procedure to implement as shown in Figure 1 of the land atlas. If this is to be done in future COADS processing, spurious positive trends in affected regions will appear if only new data are treated this way and added to an existing archive. Therefore, the entire data set (MSTG Group 5) would have to be re-done, perhaps during a proposed 5-year update.

(2) Application of the sky brightness criteria for the selection of observations made at night should improve our estimates of cloud cover over the globe. Perhaps the most important aspect of this improvement will be our ability to obtain more accurate diurnal cycles (and annual cycles near the poles) from surface observations. However, this is a more complicated and expensive procedure to implement, and several consequences should be considered:

> (a) Since about 2 weeks of data (often consecutive) are automatically eliminated from the nighttime averages in any single month, the resulting averages cannot be considered to be representative of an individual month, although in multi-year averages the sampling biases due to synoptic-scale variability should be averaged out. Daytime-only averages would not be affected.
>
> (b) Again, such a change in procedure would produce a discontinuity in the time series unless the procedure were applied to all the years within an archive.

We intend to prepare and archive a data file of surface cloud reports (from the COADS CMR data for 1982-91) in which “dark night” observations are flagged so that future users will not have to use an ephemeris on each report.

(3) It would be useful to have the CMR (or LMR) data sorted by time. The 2-degree box sort is often not desirable, especially since 2-degree boxes do not fit evenly into the 5x5° box size often used in climate analysis. Even a time sort within 10° boxes would be helpful.

(4) A brief response to a question, raised at the meeting, about the 1982 change in the synoptic code:

> Prior to 1982, ww was coded as either 0-99 or as "/", which indicated that present weather was not reported. With the code change, "/" can now mean either "not reported" or "no significant weather to report". To distinguish these possibilities, a new parameter, I_x (the weather indicator), must be checked [WMO, 1988]. We performed a survey of the National Meteorological Center (NMC) data set (archived at NCAR) and found that reports from land stations reflected this code change almost immediately in 1982. We can use I_x, for example, to exclude automatic weather stations which could contaminate our cloud analyses. (A code of N=/ would do as well were it not for cases of miscoding [Warren et al., 1986].) Ship reports, however, did not reliably incorporate appropriate I_x coding until 1985, even though they rapidly displayed the associated convention of coding ww=/. Since I_x is available in ship reports only for part of the decade 1982-91, we could just as well use the COADS data set (Interim Product CMR5 Reports), which does not even include I_x, for our current analysis. The consequence is that a few reports that should be discarded on the basis of I_x will be kept. However, the error caused by this should be small. (Based on analysis of December 1981, the upper limit to the number of these is less than 2% of the reports).
>
> The other code change that relates to clouds affects the analysis of cloud types. When N=0, observers are now instructed to code the cloud types, C_L, C_M, and C_H, as / rather than as 0. In the past, only stations that *never* reported cloud types would report the types as / when N=0, and these could be identified so that we could discard them in our cloud type analysis. However, beginning in 1982, stations that never report cloud types *will* contribute to the cloud type analysis, but *only* when the sky is clear, producing a clear-sky bias! This, too, is probably a small bias because most land stations, where the frequency of clear sky is relatively high (see the land atlas) normally report cloud types, and clear sky is relatively infrequent over the oceans (see the ocean atlas). A more definitive statement on this matter awaits further (somewhat complicated) analysis. A fix for this code change might be to add yet another parameter which would indicate whether a station reports cloud types. Such a parameter would also help eliminate the "sky-obscured bias" discussed by Warren et al. [1988].

(5) A comment about data volume for cloud analyses: While the addition of buoys increases the data volume of the COADS data set, it does not, of course, help cloud analyses. In the 1952-81 cloud analysis [Warren et al., 1988], about 9% of the ship reports contained no cloud information. Most of these were buoy reports. In the 1982-91 decade this number is more like 25%. It is important to be able to eliminate these reports so that spurious data do not contaminate a cloud analysis. A similar comment could be made with regard to automatic weather stations on land.

Acknowledgement

This study was supported by a grant from NASA NAG-1-998.

References

Hahn, C.J., Warren, S.G., London, J., Jenne, R.L. and Chervin, R.M., 1988: *Climatological Data for Clouds Over the Globe from Surface Observations*. NDP-026, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN.

Meinel, A. and Meinel, M., 1983: *Sunsets, twilights, and evening skies*. Cambridge University Press, 163pp.

Riehl, H., 1947: Diurnal variation of cloudiness over the subtropical Atlantic Ocean. *Bulletin of the American Meteorological Society*, **28**, 37-40.

Schneider, G., Paluzzi, P., and Oliver, J.P., 1989: Systematic error in the synoptic sky cover record of the South Pole. *Journal of Climate*, **2**, 295-302.

Sverdrup, H.U., 1933: *The Norwegian North Polar Expedition with the "Maud", 1918-1925, Scientific Results*, Vol. II: Meteorology, Part I, Discussion, pub. John Griegs, Bergen, 331pp.

Warren, S.G., Hahn, C.J., London, J., Chervin, R.M. and Jenne, R.L., 1986: *Global Distribution of Total Cloud Cover and Cloud Type Amounts Over Land*. NCAR Technical Note TN-273+STR, Boulder, Colorado, 29 pp. + 200 maps (also DOE/ER/60085-Hl).

Warren, S.G., Hahn, C.J., London, J., Chervin, R.M. and Jenne, R.L., 1988: *Global Distribution of Total Cloud Cover and Cloud Type Amounts over the Ocean*. NCAR Technical Note TN-317+STR, Boulder, Colorado, 42pp.+170 maps (also DOE/ER-0406).

Woodruff, S.D., Slutz, R.J., Jenne, R.L. and Steurer, P.M., 1987: A comprehensive ocean-atmosphere data set. *Bulletin of the American Meteorological Society*, **68**, 1239-1250.

World Meteorological Organization, 1988: *Manual on Codes, Volume 1*. (WMO Publ. No. 306), WMO, Geneva.