CLIMATOLOGICAL DATA FOR CLOUDS OVER THE GLOBE FROM SURFACE OBSERVATIONS

CONTENTS: Total cloud cover and cloud type amounts over both land and ocean, usually at 5° latitude - longitude resolution.

This dataset includes multi-year seasonal (and some monthly) grids of total cloud cover and the amount of each cloud type, as well as grids for individual years (11 years for land, 30 years over the ocean). Average cloud data for eight individual synoptic hours and for only the daytime are also given. Seasonal average amounts of low, middle, and high clouds are included. Frequency of occurrence and amount when present are given for each cloud type. Low cloud base heights, analysis of annual and diurnal cycles, interannual variations and trends and cloud type co-occurrences are also given. Data for FGGE are given at 2.5° latitude -longitude resolution for land. The tape has all the data from five cloud atlases, (NCAR tech. notes 201, 241 and 273, and two in preparation). Information about what fields are available is given in Tables 0 and 5 of the documentation. The documentation includes references to the atlases. To reference the dataset, the documentation supplied with the tape (Hahn et al 1987) should be cited as well as one or more of the atlases, if appropriate.

This study is intended to update, extend and supersede earlier published ground-based cloud climatologies.

DATA VOLUME: 69.75 Mbytes, in 12 files on one magnetic tape.


THE TAPE: The full global grid at a nominal 5° grid resolution has 1,820 grid cells. The data are given on land and ocean subsets of the grid. For the land cloud data there are 863 cells; there are 1,493 cells for each ocean grid at this resolution. The sum of these is more than 1,820 cells because 536 cells have both land and water data so there are separate statistics for each. A map group for land (for example) starts with a 24 character logical header record followed by 863 logical records (also 24 char each) that have data for each of the grid cells. These 864 logical records are in 3 blocks (physical records) on the tape, each having 300, 300, and 264 logical records. Several different statistics are included in one 24 character data record. Tables 4, 5, 6, 7 in the documentation summarize format information.
CLIMATOLOGICAL DATA FOR CLOUDS OVER THE GLOBE

FROM SURFACE OBSERVATIONS:

DATA TAPE DOCUMENTATION

January 1988

Carole J. Hahn
Cooperative Institute for
Research in Environmental Sciences
Campus Box 449
University of Colorado
Boulder, CO 80309

Stephen G. Warren
Department of Atmospheric Sciences AK-40
University of Washington
Seattle, Washington 98195

Julius London
Department of Astrophysical, Planetary and
Atmospheric Sciences
University of Colorado
Boulder, CO 80309

Roy L. Jenne
Robert M. Chervin
National Center for Atmospheric Research
Boulder, CO 80303
# Table of Contents

1. INTRODUCTION ........................................................................................................ 4  

2. SOURCE AND ANALYSIS OF DATA ........................................................................ 4  
   A. Data Sources ............................................................................................................. 4  
   B. Data Analysis .......................................................................................................... 5  
   C. Grid Sizes ................................................................................................................ 6  

3. CONTENTS AND ORGANIZATION OF DATA TAPE ............................................. 6  
   A. General ..................................................................................................................... 6  
   B. Details of Organization ............................................................................................ 7  
      1) Map Groups and Formats. .................................................................................... 7  
      2) Group Headers. ...................................................................................................... 8  
   C. Details of Contents ................................................................................................... 9  
      1) File 1: Latitude, Longitude, Land-fractions and Number of land stations;  
         formats 10,11. ........................................................................................................ 10  
      2) Files 2 & 6: Total cloud amount; formats 20,21. ................................................ 10  
      3) Files 3 & 7: Cloud types; formats 30,31,35. ......................................................... 12  
      4) Files 4 & 8: Harmonic analyses; formats 40,41,42,43. ....................................... 13  
      5) Files 5 & 9: Interannual variations and Trends; formats 51,52,53. .................. 14  
      6) File 10: FGGE year cloud analysis; formats 20,21,31,35. ................................. 14  
      7) File 11: Cloud type co-occurrences; format 60. .................................................. 15  

4. READING THE TAPE .................................................................................................. 16  
   A. Sample Program and Output ................................................................................... 16  
   B. Tape-Reading Strategy ............................................................................................ 16  

Acknowledgements ......................................................................................................... 17  
References ....................................................................................................................... 18  
Tables ............................................................................................................................... 19  
Appendices ....................................................................................................................... 32  
   A1. Maps of Box Numbers for 5c and 10c Grids ......................................................... 32  
   A2. Conversion of Latitude, Longitude to 5x5c Box Numbers ..................................... 35  
   A3. Conversion of 5x5c Box Numbers to Box Center Latitude, Longitude .................. 36  
   A4. Given Nimbostratus or Stratus, Probability that No Other  
       Cloud is Present ....................................................................................................... 37  
   A5. Sample Tape Reading Program and Output .......................................................... 42
List of Tables

Table 0. Outline of Quantities Archived on Magnetic Tape for Land and Ocean Grid Boxes

Table 1. Cloud Information Contained in Synoptic Weather Reports

Table 2. Cloud Type Definitions Used

Table 3. Grid Box Sizes

Table 4. Tape Information

Table 5. Tape Organization

Table 6. List of Data Formats Used

Table 7. Map Group Header Format and Codes

Table 8. Description of Terms and Abbreviations Used

Table 9. Map Group Headers, Example from File 2

Table 10. List of Data Records for Group 2013
CLIMATOLOGICAL DATA FOR CLOUDS OVER THE GLOBE
FROM SURFACE OBSERVATIONS

1. INTRODUCTION

Analyses of surface-based cloud observations discussed in this report are
described in several publications (Warren et al, 1985, 1986, 1988; Hahn et al, 1982,
1984, 1988). [These publications will subsequently be referred to by the initial of
the first author and the date of publication, e.g. W85.] The atlases contain global
maps (with typically a 5 degree latitude by 5 degree longitude grid size) of long
term monthly and/or seasonal total cloud cover, cloud type amounts and frequen-
cies of occurrence, low cloud base heights, harmonic analyses of annual and diurnal
cycles, interannual variations and trends, and cloud type co-occurrences. All the
numeric results included on the maps in the published atlases, as well as consider-
able additional information such as cloud amounts for individual years and individ-
ual times of day, are archived on magnetic tape. This report describes the con-
tents of the taped archive, which are outlined in Table 0. The procedures used to
process the individual cloud observations have been described in detail in the above
publications and are only briefly outlined here.

Copies of the archive tape may be obtained from the Data Support Section,
National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307 or
from the Carbon dioxide Information Center, Oak Ridge National Laboratory, Oak
Ridge, TN 37831.

2. SOURCE AND ANALYSIS OF DATA

A. Data Sources

Synoptic weather reports were obtained for land stations from the "SPOT"
archive of the Fleet Numerical Oceanography Center (FNOC) in Monterey, Californi-
a. This archive contains reports only from stations which have been assigned offi-
cial station numbers by WMO. Approximately 116 million reports for the eleven
year period from January 1971 through December 1981 were analyzed. Data for
erlier years were either not available with global coverage or were of lesser quality
and so were not used (W86). Synoptic reports are recorded at most at 8 times per
day: 00,03,06,09,12,15,18,21 GMT. However, many stations report only every 6
hours, some less often, and some only during the daytime.

Ship observations were obtained from Release 1 of the Comprehensive Ocean-
Atmosphere Data Set (COADS) (Slutz et al, 1985) for the years 1930-1979 and
from FNOC for the years 1980-1983. Approximately 54 million reports were
analyzed for the 54 year period January 1930 through November 1983. However,
the system for coding cloud types was changed in 1982 and cloud types were not
reported consistently in the 1930's and 1940's (W88). Furthermore, there were
fewer observations recorded globally in the early years and in the 1980's. For
example, there were less than 150 thousand reports per year over the globe during
the early 1940's and about 1 million reports per year during the early 1980's com-
pared to about 2 million reports per year in the late 1960's and in 1979. Therefore
only total cloud cover was analyzed for the whole 54 year period. Both total cloud and cloud types were analyzed for the 30 year period 1952-1981 which contained approximately 43 million reports. Although ships usually report on the GMT hours divisible by 3, they may report at any hour of the day. Any reports that occurred within a 3-hour time span centered on one of the standard synoptic times were averaged and assigned to that time (see Section 2B).

B. Data Analysis

The synoptic weather reports are coded according to a system given by the World Meteorological Organization (WMO, 1956; WMO, 1974). The cloud information in these reports is summarized in Table 1. This system distinguishes 27 cloud types (9 types at each of 3 levels). For most of the present analyses these 27 types were consolidated into 6 types as summarized in Table 2. The occurrence of clear sky and sky obscured due to fog are also included in the analyses so they are listed in Table 2 and will be referred to as “types” as well. In addition, the frequencies and amounts of the low cloud types Cu, St and Cb were added together to give long-term seasonal averages for the sum of the low-level clouds (LOW) while As and Ns were combined similarly to give the mid-level clouds (MID). High-level clouds are equivalent to Ci. In a special study for the FGGE year (December 1978 through November 1979), the 9 individual low cloud types ($C_L = 1-9$) were analyzed without further grouping.

The term “number of observations” (NOBS) associated with an averaged quantity has two possible meanings as used here. The first is the number of “raw observations” (RAWOB) which is simply the number of synoptic reports made within the grid box. The second is the number of “compressed observations” (COB). This results from a pre-averaging of raw observations within a 5° x 5° box within a 3-hour time span. The concept of compressed observation was considered useful for analysis of ocean data since a 5° x 5° box may be sampled by several ships at a time (or by one ship at three successive hours) on one day but only one ship (at one hour) on the next. In this example each case would be counted as one COB and therefore be weighted equally in the seasonal average. Thus the number of COBs is always equal to or less than the number of RAWOBs. Compressed observations were also used in some of the land analyses of total cloud. Advantages and disadvantages of various averaging systems are discussed in W86 and W88. In general, it makes little difference to the final average whether preaveraging was done, but the computed standard deviations will be affected. Standard deviations of COBs are generally smaller than those of RAWOBs since the preaveraging involved in the COB effectively increases the field of view of the “observer”. The two will be identical if there is never more than one report in a grid box in each 3-hour time span. The difference will be more noticeable over land where there is an average of 7 stations per box (from 1 to more than 100). Over the ocean there is an average of about 2 reports per 3-hour time span per box (less in the southern and more in the northern hemisphere).

Further details concerning data processing can be found in H82, H84, W86 and W88.
C. Grid Sizes

The globe was divided into grid boxes for which the various cloud quantities were computed. The grid box size used for these analyses is most commonly 5° x 5° of latitude and longitude although 10° x 20° and 15° x 30° are used for some ocean analyses while 2.5° x 2.5° is used for the land analysis for the FGGE year. Because the area contained within a 5° x 5° box decreases with latitude, boxes poleward of 50° latitude were made to encompass a wider longitude range such as 5° x 10° or 5° x 20°, etc. (The convention used for ordering the coordinates will always be latitude x longitude.) A "c" is used to symbolize this condensation or contraction. Thus 5x5c (or 5c for short) means 5x5 between 50N and 50S but a larger longitude width poleward of 50° latitude.

The four grid sizes used are described in Table 3. The 10x20c (or 10c) grid is used for several ocean analyses because relatively sparse ocean data would make some analyses at the smaller grid size unreliable. The 2.5 x 2.5 (or 2c) grid is used only for FGGE analyses over land where finer resolution is practical. The "m" in 15x30m (or 15m) symbolizes "modified" since this grid does not fit the pattern of the others. This is used only for the ocean co-occurrence maps because at the time of that analysis (H82) data were available only for latitudes between 80N and 70S.

Each grid box is assigned a number. The numbering goes from north to south and west to east (beginning with the Greenwich Meridian). Box numbers are shown in map format for the 5c and 10c grids in Appendix A1. Thus on the 5c grid, the 85° -90° N box is 1 and the 85° -90° S box is 1820 while, for example, the box at 45° -50° N and 0° -5° E is 191. The west and south borders of a box are considered to be within the box while the north and east borders are within adjacent boxes (except that 90° N is considered to be within box 1). The sample subprogram in Appendix A2 converts any latitude, longitude to 5c box numbers.

3. CONTENTS AND ORGANIZATION OF THE DATA TAPE

A. General

The data are divided into 11 files on the tape as shown in Table 4. (A 12th file contains a Fortran program which will be discussed in Section 4A.) The organization of data into files is based on similarity of content and format. Thus total cloud data are stored in a separate file from cloud type data or harmonic analyses, etc. Land and ocean data are also stored separately. There are many grid boxes which contain both land and ocean and two separate values are retained in this way. The user can combine the two if desired although the years analyzed for land and ocean only partially overlap and, for some analyses, land and ocean values are given at different grid scales.

Table 4 shows the number of logical records, physical records (blocks) and characters in each file. A more detailed breakdown of the contents of each data file is given in Table 5. Each file contains a series of "map groups". A map group consists of the data records for a number of grid boxes over the globe and a header record which identifies the group. There is a separate map group for a particular season, year, time of day and cloud type. The organization within a map group is:
Header record identifying map group
Data record for first box
Data record for second box
etc. for number of boxes given in header.

This pattern is repeated for each season, year, time of day, and cloud type in the order indicated in Table 5. Data record formats are shown in Table 6 and the format of the header record is described in Table 7. The number of data records within a map group depends on the grid size. Thus in files 2-5 each land 5c map group contains 1 header record plus 863 data records for a total of 864 logical records of 24 characters each, while in file 6 the ocean 5c map groups contain 1494 records and the 10c map groups contain 231 records. Each map group begins and ends on a tape block boundary. In files 1-10, there is a maximum of 300 records per block. Thus a land 5c map group is made up of three blocks of 300, 300 and 264 logical records, respectively (864 total). The last block is not padded but is shorter than the others.

B. Details of Organization

A detailed breakdown of the contents of the data files of the tape is given in Table 5. Use of this table, along with Tables 6 and 7, will enable the user to find any desired quantity, once a few conventions are understood. Abbreviations used are listed alphabetically in Table 8.

1) Map Groups and Formats.

Referring to Table 5, each file is a series of map groups which are numbered consecutively from the beginning to the end of the tape. Each group contains the data relevant to the cloud quantities listed under the contents heading. These data are given for each box in the indicated format (described in Table 6). For example, group 6 in file 2 contains data relevant to the mean seasonal (DJF) total cloud cover over land. These data are organized according to format 20. Table 6 shows that format 20 specifies the box number, the number of raw observations, the average amount of total cloud (given to tenths of a percent), the standard deviation of the raw observations contributing (also given to tenths of a percent), the number of synoptic hours per day at which observations were made, an indicator telling whether the observations were made during the daytime only, and the number of seasons contributing to the average (not relevant for seasonal averages). This format is repeated for all the boxes for the season.

The order in which the groups follow each other, with respect to season, time of day, year or cloud type, can also be determined from Table 5. Where there are simply 4 seasons (or 12 months) as with groups 6-9 (or 90-101), all the boxes for the first season (or month) are followed by the next group header and all the boxes for the second season (or month) and so on. (The order of the seasons is DJF, MAM, JJA, SON.) In cases such as for groups 14-45 or 46-89 where more than one time or year is given for each season, the convention adopted is to increment the parameter listed first while holding the parameter listed second constant. Thus the order for groups 14-45 would be 8 synoptic hours (in order of increasing hour) for DJF, then 8 synoptic hours for MAM, etc. The order for groups 46-89 would be 4
seasons for 1971, 4 seasons for 1972, etc. Thus the individual seasons follow each other chronologically. (Note that since our land data set begins in January 1971, the season DJF 1971 contains observations only for January and February. Similarly, the ocean data set begins January 1930.) The order in which cloud type groups (as in file 3) follow each other is the numerical order of the numeric codes shown in Table 7. Using the convention of incrementing the leftmost group parameter first, it can be determined, for example, that the group number for As cloud over land for MAM 1971-81 at 03 GMT is 323. When the term "types" is used unqualified, reference is being made to the first 8 numeric codes in Table 7 (defined in Table 2). When codes 10-19 are referred to, the term "individual low types" is applied (Table 1), while the term "levels" is used in reference to codes 21(LOW) and 22(MID) (Table 2).

Data format numbers are given as 2-digit integers. The tens digit distinguishes 6 data classes as shown in Table 6. The units digit is used to emphasize some small difference in the meaning of a variable represented. Thus the only difference between formats 20 and 21 (or similarly formats 30 and 31) is that the second variable, which refers to the number of observations, represents raw observations in format 20 and compressed observations in format 21.

The differences among formats 41-43 (as well as 51-53) are simply that the data variables refer to amount, frequency, or amount-when-present, respectively. This distinction is necessary for reporting cloud types. Format 40 is used to distinguish the fact that the phase of the annual harmonic is given in units of month as opposed to hour for the diurnal harmonic. The annual harmonic was computed only for total cloud so no further distinction is needed.

Note that format 35 differs significantly from the other formats in class 3 in that it refers to heights of the cloud types.

2) Group Headers.

The first record in each map group is a header record which identifies the group. The format of this header record is shown in Table 7. The first 4 parameters of the header record give the group number (group numbers run consecutively from the beginning to the end of the tape), the number of boxes reported in the group, the grid size of the boxes, and whether the data are for land or ocean. Except in file 1, the number of boxes in a group may be less than the number of boxes in the entire grid. Of the 1820 5x5c boxes, 934 have at least some land and 1493 have at least some ocean (607 have both). Of the 934 land boxes, only 863 had data at some time during the 1971-81 period of record. Therefore only 863 boxes are included in the land 5c groups. To keep these groups uniform in size, all 863 boxes are always included in all the land 5c groups even though for some seasons or some years there may have been no data for a particular box (this will show up as a 0 in the NOBS portion of the data record and/or as a -9 for the data values, see Section 3C). Similarly, all ocean 5c groups contain 1493 boxes and all land 2c groups contain 2134 boxes. For the 10c and 15m grid sizes, all boxes are always recorded since there are so few of them. The box numbers can be converted to latitude and longitude by use of the information in file 1, which gives box center latitude and longitude for all grids, or by use of conversion routines such as that given in Appendix A3 which converts 5c grid numbers to box center latitude,
longitude. The last 5 parameters in the group header indicate the year, season (or month), time of day, and cloud type to which the data refer, and the data format. Even without the map group number, the other 8 header parameters together uniquely define each map group.

As an example of the contents and use of the map group header, Table 9 gives a list of the first 150 group headers for file 2 and shows the program that produced it. The headers are printed with the 24-character record length of the header format and the parameter names are abbreviated across the top of each column as an aid in identification. Using group 6 for an example and reading its header on Table 9 from left to right (referring to Table 7 if necessary), it is seen that 863 records of data follow this header, the boxes refer to the 5c grid and they are for land. The period of record is 1971-82 for DJF (the last year for DJF is listed as 1982 because Dec 1981 is defined as part of DJF 1982). Data were counted for all available times of day and they are for total cloud cover. The format number 20 indicates (Table 6) that the data variables reported (and skipped here) are box number, number of raw observations, average and standard deviation, number of synoptic hours contributing, a daytime bias indicator, and the number of seasons contributing. Group 10 also reports total cloud for DJF 1971-82 but the average is over daytime reports only, while group 14 gives the total cloud amount for 00 GMT. Groups 46-89 give the individual season mean total cloud, consecutively in time, though the NOBS used for these averages were COBs rather than RAWOBs as indicated by the format number 21. Groups 90-101 give the long-term mean monthly total cloud while the remainder of the groups on the list give the individual monthly means.

Thus group headers can be decoded and used to find a desired map group. Data quantities can be read by replacing the READ statement 90 in the example in Table 9 with a READ of the appropriate format. Examples are given below.

C. Details of Contents

All data on this tape were written as integers. Floating point quantities were multiplied by a power of 10 and rounded off to give an integer. When read using the formats shown in Table 6, the original units (defined in Table 8) are recovered directly. For example, the AMT in format 30, when read as F5.2, is the amount of a cloud type in percent. This quantity could also be read as I5 (or F5.0) and then divided by 100 to give percent.

Any data variable for which no value is reported is assigned the missing value code. The missing value code was chosen to be the integer -9 so that a 2-character value could be applied uniformly (since the smallest field for any data variable is 2 characters). Thus, when reading in the various floating point notations, the actual value obtained may vary but will always be less than 0. No valid data have values less than 0 except trends. However, no ambiguity is possible since the missing value code is used in the trend records only when the number of years (NYRS in formats 51-53) is 0. Therefore, when reading the trend groups (files 5 and 9), the missing value code should be considered to be NYRS=0. For the other files, when using the floating point values, the most consistent way to check for the missing value code is to check whether the value is less than zero.
As an illustration, a listing of all the records of one map group is given in Table 10. The first record, the group header, indicates that this group contains data for cumulus cloud over the ocean and is given for 230 boxes on the 10c grid. The data are for MAM averaged over the years 1952-81 for all times of day. The data format number is 31 so in box 2, for example, there were 19742 COBs giving an average Cu AMT of 6.09%, an average FQ of 13.06% and an AWP of 46.7%. The missing value code appears for box 10 because there were fewer than the minimum required 100 observations (see below). Although in box 226 there were sufficient observations (549) to report a reliable AMT and FQ, there were insufficient observations (549x.0316=17) to report a reliable AWP. [If a low cloud type is present, its amount is always known because it is a reported quantity (Nh). Therefore FQ x NOBS gives the number of occurrences of the cloud type that went into the computation of AWP.]

The methods of averaging, minima applied, and any peculiarities of the data will now be discussed for each file. Refer to Tables 5, 6 and 7 for references to file contents, format numbers and header coding, respectively, in the following discussions. Any non-standard terms not defined in this section can be found in Tables 1, 2, 3 or 8.

1) File 1: Latitude, Longitude, Land-fractions and Number of land stations; formats 10,11.

The four groups in this file give data concerning the grids used for dividing the globe (see Table 3). The latitude (90 to -90 for N to S) and longitude (0 to 360E) of the box center and the fraction of land in the box are given for each box in each grid. (Here "land" means "not ocean". Lakes and ice shelves are counted as land. See W86.) In addition, the average number of land stations contributing in 5x5c boxes is included with the data for the 5c grid. (This corresponds to Map 1 in W86.) Inclusion of the box number in the data format is redundant since here all the boxes are reported and the data records are in order of box number. Information in these groups may be used in processing other files. Alternatively, conversions between box number and latitude and longitude can be accomplished by use of subprograms such as shown in appendices A2 and A3 for 5x5c conversions.

2) Files 2 & 6: Total cloud amount; formats 20,21.

The seasonal and monthly averages for total cloud cover on these files were computed by three different series of steps (1-3 below) which is partially reflected by the two format numbers (Table 6). The differences between RAWOBs and COBs were discussed in Section 2B and the consequences of various methods of averaging are discussed in W86 and W88. For the quantities governed by format 21, COBs were simply summed over the period of record and averaged (1). The standard deviations were computed from these values. For the mean seasonal values over land (format 20), RAWOBs were accumulated for each synoptic hour then averaged to give a mean value for each of the eight hours. Then if 2 or more synoptic hours had a minimum of 200 RAWOBs, those averages were averaged together to give the mean seasonal cloud cover (2). If there were less than 2 hours with the minimum requirement, then all raw observations were averaged together irrespectively of time of day to form the mean cloud cover (3). In either case the standard deviation was computed from all raw observations irrespective of time of
day. NRAW for these averages is the total number of RAWOBs that contributed to the average and is the value plotted on Maps 2-5 in W86. In computing the daytime averages for land, raw observations were weighted equally (3).

For the land, NTP is the number of synoptic hours that had at least the minimum of 200 raw observations and is the number plotted on maps 6-9 in W86. A 1 is entered in the IDY variable as a code signifying that the only synoptic hours with 200 RAWOBs were all daytime hours, indicating that a sampling bias may exist in the computed average. This is coded as an asterisk on the atlas maps and occurs in only a few boxes. For the ocean, NTP is the number of synoptic hours that had at least 100 COBs (with a 1 in IDY if these were only daytime hours, though this is rare over the ocean). Note that it is possible for NTP to be 0 while NOBS is not. NTP is assigned a meaningful value only with the mean seasonal groups listed in Table 5 and is assigned the missing value code (-9) elsewhere. NSN is given only for the mean annual groups and is assigned the missing value code elsewhere.

The mean annual total cloud was computed by averaging the mean seasonal values of the seasons with 100 or more observations. The number of seasons averaged was entered in the NSN variable. If NSN was 0, the AMT was assigned the missing value code. Note that if NSN is 1, as may be true for ocean boxes near the poles, the reported value may not be representative of the true annual value. Standard deviations were not computed. NOBS (NRAW or NCOB) is the sum of the NOBS in the four seasons. For some boxes this value contains more than the six digits allowed in the format (16). In these cases, the value 999999 was entered.

The user should check the value of NOBS (and NSN for annual averages) before using a reported average in order to check the statistical reliability of that average. In addition, the user must check that the amount or standard deviation is not less than 0 because the missing value code was inserted for AMT and SD while NOBS was left unaltered for some boxes with erroneous reports (W86). All values of cloud amount, as well as standard deviation, are given in percent of sky cover if the floating point formats in Table 6 are used.

Individual season means of total cloud cover over the ocean are given for the years 1930-1983 (Table 5). However, after writing the tape we determined that, due to the change in coding procedures in 1982, our analysis scheme overestimates total cloud cover by about 1% on average for the years 1982 and 1983 (see W88 for discussion). Therefore the user is warned against using the data for these two years until they are reanalyzed. (In a future release of the tape, reanalyzed data will be signaled by the value 1 in the NSN variable which currently contains the value -9.) Note that the daytime means and the means by synoptic hour for the ocean were averaged over the years 1954-83. These values had been computed before it was decided to use the 1952-81 period for long term averages (Section 2). This should have no significan effect on the diurnal variations computed from these data and little effect on the 30-year daytime and synoptic hour means whose global averages are increased by less than 0.1%.

The long-term means, means by synoptic hour and the individual season means for the ocean are given at the 10c grid size as well as at the 5c size. The 10c grid size gives a better global representation for the ocean (because data are too
sparse to form averages at 5x5 degrees in some parts of the ocean) and serves as the reference mean for the diurnal harmonics and the trends and interannual variations which are given only on the 10c grid. (A COB in a 10x20c box still results from the pre-averaging of all reports within a 5x5c box over a three hour time span as defined in Section 2B.)

5) Files 8 & 7: Cloud types; formats 30,31,35.

The basic computed variables for cloud types are frequency of occurrence (FQ) and amount-when-present (AWP). The amount is then AMT=FQxAWP as discussed in W86. These variables are all assigned the missing value code if NOBS is less than some minimum. That minimum is 100 COBs for ocean data, 200 RAWOBs for long-term mean land data and 50 RAWOBs for land individual seasons. The NOBS reported refers to the number of observations used in computing FQ. In addition, since AWP is computed from a smaller data set than FQ, AWP may be assigned the missing value code when FQ is given (to have an observation for AWP, the cloud type must be present and its amount computable). The absolute minimum number of observations for reporting a computed AWP is 10 for land individual seasons and 30 otherwise. AMT may be reported even though AWP is not if the uncertainty in AWP did not lead to excessive uncertainty in the computed amount, as discussed in W86. The user must therefore always check that values are not less than 0 before using them. The value of NOBS may also be checked in order to apply a stricter criterion than we set. Only FQ is given for clear-sky and for sky-obscured-due-to-fog; AWP and AMT are assigned the missing value code. This is because AWP is meaningless for clear-sky and is always 100 percent for sky-obscured-due-to-fog.

The averages for FQ, AWP and AMT were computed in a manner similar to that for total cloud amount. For the ocean (format 31), the COBs were accumulated over the period of record and then averaged to give FQ and AWP; AMT was then computed as FQxAWP. For the land (format 30), values of FQ and AMT were formed for each synoptic hour which had the minimum number of observations, and these values were averaged if there were at least two such hours. AWP, however, was computed as the average of all the RAWOBs contributing, irrespective of time. Because of this method of averaging it is possible that, for the land, the AMT reported does not exactly equal FQxAWP. W86 gives the reasons for this choice of methods.

Average base heights (format 35) of low clouds (Cu, St, Cb) are given in meters and are simply the averages of all contributing RAWOBs. The missing value code was assigned if NOBS was less than 100. For the ocean, the period of record for heights is 1954-81 (W87).

Only seasonal (not monthly) values were computed for the cloud types. For the upper cloud types (Ci and As) and for clear sky, the averages reported were computed from daytime observations only, but averages are archived also for each of the eight synoptic hours. For the ocean, the individual season means, means by synoptic hour and low cloud heights are reported on the 10c grid only. The long-term means are given at the 10c and the 5c grid sizes.
The mean annual values for AMT and FQ were computed by averaging seasonal values if there were two or more seasons with at least 100 observations. AWP was then computed as AMT/FQ. (This method automatically gives the proper weighting factor to AWP from each season. The minimum of two seasons was required to ensure better representativeness of the reported value since no NSN parameter appears in formats 30-31 as it did in formats 20-21.)

Values for the low-level clouds (LOW) were obtained by combining Cu, St and Cb, while values for the mid-level clouds (MID) were obtained from As and Ns. Amounts and frequencies were added together and AWP was then computed as AWP = AMT/FQ. The MID level was computed only if values for both As and Ns were available; the NOBS reported is that for As. (NOBS is smaller for As than Ns because As was computed from daytime observations only.) The parameter TIME in the header was assigned the value -2 for the MID level. The definitions of St and Ns (Table 2) allow for some low clouds to be defined as part of the middle level. This results in a slight underestimate of the total amount of cloud reported as LOW and a slight overestimate reported as MID. However, the average difference in only about 0.25% for land and 0.5% for ocean (compared to the computed AMT for LOW which is, on average, 26% for land and 51% for ocean and to the computed AMT for MID which averages 26% for land and 29% for ocean).

4) Files 4 & 8: Harmonic analyses; formats 40, 41, 42, 43.

The phase and amplitude of the annual harmonic of total cloud cover were computed from the mean monthly values if all 12 months had 100 or more COBs (NT = 12 in format 40). The amplitude reported is the absolute amplitude so that, for example, if the mean value for a particular cloud amount is 25% and the maximum of the fitted cycle is 30% then the amplitude is reported as 5% (rather than 20% which would be the normalized amplitude). The phase corresponds to a numeric value for each month so that 1.0 is the middle of January, 2.0 is the middle of February, etc. Phase values reported range from 0.5 to 12.4. VAF is the percent variance accounted for by the annual harmonic. AVG is the average of the 12 months used in the analysis and may differ somewhat from the annual values reported in groups 5 and 1071. The missing value code was inserted for all parameters if the minima were not met.

The first diurnal harmonic (formats 41-43) was computed for the cloud types as well as for total cloud cover if 8 or 4 evenly space synoptic hours (NT) contained the minimum NOBS for a season. Otherwise the missing value code was inserted for the harmonic parameters (and NT was set to zero). The minimum NOBS required per synoptic hour was 100 COBs for the ocean, 300 RAWOBs for land total cloud and 200 RAWOBs for land cloud types. Additional special criteria were applied to the land analysis to minimize geographical biases (W86). The amplitude reported is, again, the absolute amplitude. The phase is the hour of day (mean solar time of box center) and may range from 0 to 24. VAF is the percent variance accounted for by the first harmonic. The AVG reported here is the average of the 8 or 4 synoptic hours that were used in the harmonic analysis and may differ somewhat from the averages reported in files 2-3 or 6-7.

Since there are three variables (AMT, FQ and AWP) which relate to cloud types, diurnal analysis was performed on all three. The distinction between these
is indicated in the format numbers (Table 6). Note that there are eight types, including Clr and Fog, for which FQ is reported. Since Clr and Fog do not have an associated AMT or AWP, only six types are reported for AMT and AWP. Diurnal cycles for the ocean were analyzed for the 10c grid only.

5c) Files 5 & 9: Interannual Variations and Trends; formats 51,52,53.

The trend (TRND) is the slope of the straight line which was fit to the data points by least-squares analysis, in units of AMT, FQ or AWP per year. UNC is the uncertainty of the slope (Bevington, 1969, page 118) in the same units. The interannual variation (IAV) for the land analyses is the standard deviation of seasonal values from individual years about the multi-year mean. For the ocean, however, it is the standard deviation about the trend line.

Individual years contributed to the computation of IAV, TRND and UNC if they had a minimum of 100 COBs (for ocean total cloud, ocean cloud types and land total cloud) or 150 RAWOBs (for land cloud types). The missing value code (-9) was inserted if no years met these criteria (NYRS=0) and zeros were inserted if NYRS was 1. The user must check NYRS (and SPAN if desired) to decide whether or not to use the results.

Trends and interannual variations in ocean data are given on the 10c grid only. To get the trend values published in W88, TRND must be multiplied by 29 (the number of annual intervals in the total data record). Since the period of record for the land was only 11 years the computed trends were not considered to be sufficiently meaningful for inclusion in the published atlas, but they are included on the tape.

6) File 10: FGGE year cloud analysis; formats 20,21,81,35.

This file contains cloud data for the time covered by the First GARP Global Experiment (FGGE; Dec 1978 through Nov 1979) for both land and ocean. The quantities presented are similar to those presented in the other files but with the addition of individual low cloud types (CLR=0-9) for two seasons. Also, land data are presented at a higher spatial resolution. Table 5 outlines the quantities presented, the order of presentation, the grid sizes used, and the data formats. On the 2c grid, only boxes with at least 2% land were analyzed. As before, no minima were applied to total cloud data (values are assigned -9 only if NOBS was 0) so the user should check NOBS before using a value. For the cloud types, the minimum NOBS for FQ is 80 for land and 50 for ocean. In addition, the absolute minimum (see section on cloud types above) for AWP is 30 for land and 20 for ocean. The user should check to see that the values AMT, FQ and AWP are not less than zero.

Reporting of the individual low types provides a finer breakdown of low clouds than that used for the main climatology (Table 2). The frequency of occurrence of CLR=0 (no low clouds) is included in this list for the individual cloud types only.

Note: SD was not computed for the daytime or the synoptic hour averages of total cloud cover over land so these SDs were assigned the missing value code.
7) *File 11: Cloud Type Co-occurrences; format 60.*

The cloud type co-occurrences (contingency probabilities) are defined as: the probability (percent frequency) of occurrence of one cloud type, B, given that another particular cloud type, A, is present. This is symbolized as \( P(A \rightarrow B) \). The "given" cloud A is coded in the group header (TYPE) and the co-occurrence probabilities of the other clouds with that cloud are archived (PB in format 60) for each box. The details of computation of these co-occurrences are discussed in H82 and H84. The values of \( P(St \rightarrow NO) \) and of \( P(Ns \rightarrow NO) \) (where NO means that no other cloud is present with the given cloud type) archived on tape for the ocean differ from those published in H82. The values on the tape were obtained by use of the improved procedure described in W85, which is also the procedure used in the land cloud analysis (H84). The revised values of \( P(St \rightarrow NO) \) on the tape average about 31\% (compared to 36\% in H82), while the values of \( P(Ns \rightarrow NO) \) average about 4\% (compared to 17\% in H82). Since the revised values differ so greatly from those published in H82, maps of these revised values are printed in Appendix A4.

The overall frequency of occurrence of the given cloud type is also reported (FQ). These frequencies may differ from those reported in files 3 and 7 because: a) time-spans of the data sets differ, b) nighttime observations were included for Ci, As and Clr, c) some type definitions differ slightly (Table 2), d) reports of upper clouds are not limited to cases of \( N_h \leq 6 \) (W86), and e) fog is ignored in computing these frequencies (except in FQ for Fog or Clr). [Modifications were made in the cloud type analysis procedure (W86, W87) after publication of the co-occurrence probabilities (H82, H84).]

In order for a co-occurrence probability to be reported, it was required that the given cloud type have at least 50 occurrences (NTY) for the ocean and 200 for the land; otherwise the missing value code was inserted for the values of PB. For FQ these minima were applied to NRAW.

The types Cu, St, Cb are mutually exclusive by definition (Table 2) so co-occurrence probabilities of these types with each other are assigned the value zero. All co-occurrences involving Clr are also assigned the value zero while co-occurrences with Fog are assigned the missing value code. (Our "Fog" means "sky obscured due to fog", so the reports have no information about what other cloud may be present above the fog.) Clr and Fog are included here merely to show their frequencies of occurrence. The co-occurrence of the given cloud type with itself is assigned the value 100\%.

The variable FC is defined as the percent of raw reports (raw reports in which at least some cloud was present, i.e., \( N>0 \)) that contributed to the statistics of the given cloud type. It is 100\% for the lower cloud types (Cu, St, Cb and Ns) which are always visible to the ground observer, but is less than 100\% for the upper types (Ci and As) which are sometimes obscured by a low overcast.

For the land co-occurrences, only boxes with at least 5\% land were analyzed.
4. READING THE TAPE

If a "map" is defined as a representation of the global distribution of a particular data parameter then there are approximately 12,000 unique maps archived on this tape (slightly less than 4000 groups times an average of slightly more than 3 data quantities per group). With the use of Tables 5, 6 and 7 (described in detail in Section 3), the user can design programs to read the desired quantities. Pertinent tape characteristics are shown in Table 4 and the various grid sizes employed are described in Table 3.

A. Sample Program and Output

A sample Fortran program (RDTAPE), given in Appendix A5, is provided as an example for reading the archived data. A copy of the program was written on file 12 of the tape. (See Table 4. The logical records were blocked so that the subroutine MAPG begins on a new block.) This program is necessarily selective in the output it produces and is not intended to suit all needs. Users will want to design their own programs to suit their own needs, though this program may serve as a useful guide. RDTAPE, as written, will print 32 maps from the entire tape, though it can easily be modified to print others. Six of these maps (one from each data class) are reproduced in Appendix A4 for user comparison. Other maps generated by the user from this tape may be compared with maps published in the atlases.

RDTAPE uses the self-documenting nature of the map group headers to select and label output, though, alternatively, map groups could be selected on the basis of the map group number, which can be determined from Table 5, and labeling could be accomplished by simply printing the header. The program checks NOBS (or NYRS) and checks for the missing value code before assigning a value to a map. Latitude and longitude data from file 1 are used for positioning the boxes on maps, so file 1 must be read first and the data written to unit 2 (line 74 of RDTAPE) must be saved before using the program for maps from other files.

B. Tape-Reading Strategy

Using RDTAPE to read every record on the tape to produce 32 maps required more than one half-hour computer time on the CDC 840. Files can be skipped to position the tape at a desired file and it may also be possible to skip blocks of data within a file to get to a desired group more efficiently. For example, it took 374 cp seconds on the CDC 840 to read all of file 3 to print two maps (St AMT and HGT for DJF 1971-82). Using Tables 5 and 7 it can be determined that St AMT for DJF 1971-82 is in the 33rd group in file 3 (group number 266). Since each group begins and ends on a block division with 300 records per block and short blocks to complete a group when necessary and since there are 864 records in a land 5c group (863 data records and 1 header record), there are 3 blocks for each of the 32 land 5c groups that precede map group 266 on file 3. Therefore, skipping to file 3 and then skipping 3x32=96 blocks would position the tape at map group 266. The CDC control cards for this are:
SKIPF (TAPE1,2)
SKIPR (TAPE1,96).

Using this method to get to group 266 and terminating the program after printing the map for St AMT required only 3.5 cp seconds.

ACKNOWLEDGEMENTS

This work was supported by NOAA grant NA80AA-D-00030 and NASA grant 2206-CL-229 to the University of Colorado, and by DOE through the NOAA grant.
REFERENCES


