JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 94, NO. D7, PAGES 9925-9935, JULY 20, 1989

Cloud Overlap Statistics

LIN TIAN[1] AND JUDITH A. CURRY[2]

Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana

The U.S. Air Force Three-Dimensional Nephanalysis (3DNEPH) has been employed to analyze the vertical distribution of clouds and cloud overlap statistics during January 1979 over the north Atlantic Ocean (40° to 60°N). The 3DNEPH integrates both satellite and conventional observations and gives what is probably the best available information on the vertical distribution of clouds. This region was chosen because of the predominance of layered clouds and the high density of conventional observations. The grid size of the 3DNEPH is ~45 km, and the data set was additionally averaged over three different grid sizes: 90, 221, and 442 km. Three cloud overlap assumptions that have been commonly used were tested, namely the maximum, minimum, and random overlap assumptions. For adjacent layers of the 3DNEPH that contained cloud the maximum overlap assumption performed the best, accurately determining the total cloud cover to within the roundoff error for the layered cloud amounts in 82% of the cases. For two or three cloud layers that were separated by clear interstices the random overlap assumption performed the best overall for all resolutions, although there were systematic biases in the predicted cloud fraction which depend on the cloud fraction of the layers and on grid size. For grid sizes ≥90 km, random overlap resulted in a systematic underestimation of total cloud cover, ~5%. Random overlap performed worst when there were two layers with intermediate cloud fractions (30-70%). The results of this study indicate that while random overlap performs reasonably well on average, the systematic bias (which depends on grid resolution) and random discrepancies could result in significant errors when this approximation is used in general circulation modeling and cloud climatologies.

INTRODUCTION

Clouds play an undeniably important role in weather and climate by changing the radiative heating of the planet. The presence of clouds greatly increases the solar radiation reflected back to space and also reduces the thermal radiation emitted to space. The net effect of these two processes is not clear, indicating the importance of establishing the cloud/radiation interaction. The resolution of this issue requires an accurate cloud climatology to input into the models that require prescribed cloudiness and to validate the models that predict cloudiness.

An issue that has not been adequately addressed is the clarification of cloud overlap in the vertical. Partial cloudiness is frequently observed to occur in multiple vertical layers in the atmosphere. Specification of the overlap that occurs between cloud layers is required for calculations of radiative transfer in general circulation models (GCMs), since only cloud fraction for each model level is determined by the model. The cloud overlap assumptions that have been made in GCMs include the assumptions of maximum, minimum, and random overlap between cloud layers. These assumptions are derived from simple geometric considerations and do not have an observational basis. Most of the radiation codes presently used in GCMs assume random overlap of partially cloudy layers [see *Stephens,* 1984, Table 6], including *Ramanathan et al.* [1983]. *Geleyn and Hollingsworth* [1979] have proposed an alternative which contains maximum overlapping of adjacent cloudy layers but random overlap of cloud layers separated by a cloud-free layer. Systematic differences in the radiative forcings of GCMs have been shown by *Morcrette and Fouquart* [1986] to occur because of different cloud overlap assumptions. Additional application of cloud overlap assumptions may be made in interpreting clouds from surface or satellite observations [e.g., *Warren et al.,* 1986].

The requirements for an observational data set to test these overlap assumptions, and to possibly indicate an alternative overlap scheme, are not easily met using either surface or satellite observations: upper level clouds are frequently obscured from surface observations if lower clouds are present, and lower level clouds are obscured from the satellite observations if high clouds are present. The U.S. Air Force Three-Dimensional Nephanalysis (3DNEPH) in principle would meet the requirements for such a data set: the 3DNEPH combines satellite, surface, and aircraft observations on a horizontal grid of approximately 45 km and a vertical grid consisting of 15 layers.

For this study we have examined the 3DNEPH over the north Atlantic Ocean from 40°-60°N during January 1979. The region that we have chosen to analyze is one of the most persistently cloudy regions in the northern hemisphere. The cloudiness originates mainly from the large-scale convergence associated with the frequent and intense synoptic activity and consists of a complicated mixture of layered and convective clouds. The period we have chosen was dominated by layered clouds, with cumuliform cloud types being reported for less than 10% of the observations and contributing little to the total cloud cover. A comparison of the 3DNEPH data set with independent satellite cloud data and with climatology lends credence to the accuracy of this data set. We presume that the accuracy of the 3DNEPH, particularly in terms of the vertical layering of the clouds, is enhanced in this region because of the high density of ship observations and aircraft reports. The validity of the different cloud overlap assumptions is examined using the 3DNEPH data by comparing the difference between the observed total cloud amount and the computed total cloud amount by using the different overlap assumptions.

DATA DESCRIPTION

The U.S. Air Force Three-Dimensional Nephanalysis began in 1972 to provide a real-time cloud data set and is described fully by

[1]Current affiliation Department of Meteorology, The Pennsylvania State University, University Park.

[2]Current affiliation Department of the Geophysical Sciences, The University of Chicago.

Paper number 89JD00810.
0148-0227/89/89JD-00810$05.00

Fye [1978]. The 3DNEPH provides total cloud cover, cloud amount at 15 levels, heights of the lowest bases and highest tops, and types of low, middle, and high clouds. The 3DNEPH grid is based on a polar stereographic projection, which is true at a latitude of 60°. Actual grid spacing is variable with latitude due to the distortion of the projection, with the horizontal resolution of the 3DNEPH approximately 46 km at a latitude of 60°. The temporal resolution of the analysis is 3 hours.

One of the major problems encountered with the 3DNEPH data base is the tremendous volume of data that must be processed. To reduce the data to a manageable size, only 0000 and 1200 UT data are used in the present study. The study period is January and February 1979 (corresponding to FGGE Special Observing Period I) for general cloud statistics and January 1-14 for the cloud overlap study. The region analyzed is confined to 40-60°N and 10-50°W. The horizontal resolution for this region is about 45 km, with a total number of 393,530 observations for 2 months and 93,380 for the 14-day period.

Visible (0.4-1.1 μm) and infrared (10.5-12.5 μm) data from the polar-orbiting Defense Meteorological Satellite Program (DMSP) provide the basic input to the 3DNEPH, using a threshold technique. When visible and infrared estimates of total cloud amount differ, the highest value is retained. This method ensures that high clouds, which are more easily detected by infrared processing, and low clouds which are more easily detected by visible processing, are most likely to be retained in the final analysis. The appropriate height of cloud top is determined from a model temperature which is assumed to represent the temperature of a cloud field emitted as a blackbody. Cloud thickness is determined as a function of the layer cloud amount and layer cloud top. The conventional data used in the 3DNEPH consists of surface data, upper air data, and aircraft reports. These three data sources are built into a "best reports" file, with surface cloud observations given priority over upper air sounding data. Visible and infrared data are merged and then combined with the best reports file. The strength of the 3DNEPH for this study lies in the merging of conventional and satellite data; in a region with sufficient conventional observations, lower cloud amounts are determined from surface observations, with the higher cloud amounts determined from the satellite observations, supplemented by aircraft observations (Hamill, personal communication 1988).

The 3DNEPH data has been used by a number of authors in studies of global cloud cover [e.g., *Stowe,* 1984; *Hughes and Henderson-Sellers,* 1985], cloud-radiation interaction [e.g., *Koenig and Liou,* 1983; *Gordon et.al.,* 1984], and regional energy budgets [*Curry and Herman,* 1985]. Most notably, *Hughes and Henderson-Sellers* [1985] examined 3DNEPH data for the entire FGGE year (December 1978-November 1979). They found that 3DNEPH produced an excellent global picture of cloudiness, which agrees well with the known features of the general circulation and other widely used cloud climatologies, such as *London* [1957].

Several problems with the 3DNEPH have been identified. These are described below, with a brief discussion of their potential impact on the present study.

Lack of discrimination between clouds and snow/ice surfaces. The results of *Hughes and Henderson-Sellers* [1985] illustrated a discrepancy between the 3DNEPH results and climatology in the polar regions. *Curry and Herman* [1985] found that the 3DNEPH seriously underestimated low cloud cover over the Arctic Ocean. Since the present study is over an ice-free oceanic region, it will not be affected by this problem.

Satellite underestimation of cirrus height. If the emissivity of a cloud is less than one, which is most likely to occur for cirrus clouds, the infrared satellite processor will overestimate the cloud temperature. This will result in an underestimation of cloud height and possible misidentification of cloud type. The magnitude of this problem in the present data set is uncertain, although the results of section 3 (Table 3) indicate that the frequency of cirrus determined from the 3DNEPH is reasonable. Accurate cloud heights for the present study are significant only to the extent that they allow the discrimination of different cloud layers, for example, an underestimation of cirrus height may result in the merging of a distinct cirrus layer with a lower altostratus cloud layer, yielding one layer where in reality two layers exist.

Timeliness of the data. The 3DNEPH is produced every 3 hours. If there are no new data for an update, the previous analysis or a cloud forecast is used for the new analysis. We have minimized this problem in the present study by choosing a region and period with a large number of conventional observations and where the cloud situations are not changing rapidly.

Bias toward random overlap in the conventional data processing. If more than one cloud layer is observed in a surface observation, the report does not contain sufficient information to define each layer unless a surface supplementary cloud data record is included. The random overlap assumption may be used in the following instances: (1) if the cloud height categories indicate that two cloud layers are present, and no surface supplementary cloud data record is included, and (2) if the cloud layers are described in a supplementary cloud data record but total cloud cover is missing. Independent analysis of the Level II-b surface observations for this period indicate that instance 2 is not a factor in this data set, while instance 1 probably is, since only about 10% of the surface observations have the supplementary cloud data.

Bias toward random overlap in the infrared data processor. If two modes are identified by the infrared processor, the layer amount for the warmer (lower cloud) mode is determined by using the random overlap assumption.

For purposes of the present study the only problems that are likely to bias our results are the utilization of random overlap by the 3DNEPH in processing conventional surface observations and the infrared satellite data, although it is shown in section 3 that this problem does not dominate the results. The reason for this is due to the manner in which 3DNEPH merges the satellite with conventional observations, using the satellite observation for the upper cloud amounts and the surface observation for the lower cloud amounts. The cloud amount for one or more middle layers is determined by a weighting of any derived middle cloud amounts, which may be supplemented by aircraft data. In fact, the only way that 3DNEPH can have three or more cloud layers in a given box is by combining conventional and satellite observations; satellite alone, or surface alone, can give at most two layers. In the presence of both satellite and conventional observations the layer cloud amounts will have no bias toward random overlap for two cloud layers. For three cloud layers the middle layer may have been determined using the random overlap assumption from the satellite and/or surface observation, in the absence of a relevant aircraft report.

In the presence of both conventional and satellite observations there is no reason to suspect a large bias toward random overlap. The issue then becomes whether 3DNEPH has sufficient conventional observations, concurrent with satellite observations, to provide meaningful information about cloud layers. The satellite observations are derived from two DMSP satellites (polar

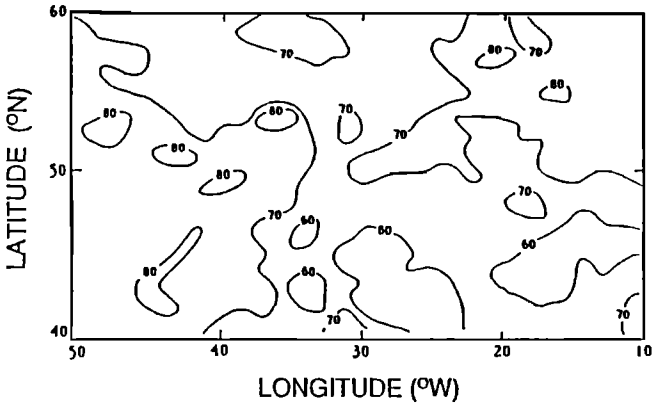

Fig.1. The area distribution of mean total cloud amount for January-February 1979.

orbiters) that cross at approximately 0600 and 1800 UT and 1100 and 2300 UT, for a total of four passes per day (Hamill, personal communication, 1988). The region that we have chosen to examine probably has the best conventional data coverage of any oceanic region in the world [*Bjorheim et al.*, 1981], which is one of the reasons why this particular region was chosen for analysis. Independent analysis of the Level II-b surface observations gives an average for the region of 82 and 110 surface observations at 0000 UT and 1200 UT, respectively. Aircraft reports are abundant but appear to vary with the day of the week [*Bjorheim et al.*, 1981]. Since there are 3335 $(45\ km)^2$ boxes in the analysis region, there are insufficient conventional observations to provide direct information for each box. Before the conventional data is merged with the satellite data, conventional data in the best reports file for a single grid point are spread to influence surrounding grid points as a prelude to actual integration with the satellite data [*Fye*,1978]. The conventional data in this region appears to be sufficiently dense so that most boxes will be meaningfully influenced by one or more conventional observations. Given the relatively large horizontal extent of the cloud decks, as determined from the Advanced Very High Resolution Radiometer (AVHRR) images, the amount of conventional data spreading should provide a reasonable nephanalysis and when combined with the satellite data should provide a reasonable analysis of layer cloud amounts.

General Cloud Statistics

The general cloud statistics are presented here to describe the characteristics of the data set and also to establish the reliability of the data set by comparing it with other cloud climatologies. Two months of data are analyzed to get the general cloud statistics of the 3DNEPH for the region, which include the total cloud amount and cloud types.

TABLE 1. Mean Total Cloud Amount Over North Atlantic Region During January Obtained From Different Cloud Climatologies

Source	Cloud cover,%
Haurwitz and Austin (1944)	60
London (1957)	60
Clapp (1964)	75
P.E. Sherr et al. (1968)	79
Miller and Feddes (1971)	60
Becker (1979)	70
Beryland and Strokina (1980)	70

All sources given by *Hughes* [1984]

Total Cloud Amount

Figure 1 shows the area distribution of mean cloudiness for the 2-month period. In most of the region the cloud amount averaged over 2 months ranges from 50 to 80%, resulting in an areal average of 69% for the 2-month period. Analysis of the day-to-day variation of cloudiness for the 2 months indicates that the areally averaged cloudiness is persistently high for all days, ranging from a low of 49% to a high of 83%.

The comparison of these results with other cloud climatologies is ambiguous due to year-to-year variations in cloudiness, that is, a 10-year climatology is almost certainly bound to differ from an individual season. However, such a comparison can give an indication of the reliability of the data. Table 1 gives the average cloud amount (estimated to the nearest 5%) over the north Atlantic Ocean, derived from several data sets provided by *Hughes* [1984]. The values range from 60-75%, which is comparable with the mean cloud amount of 69% derived from this study.

The fact that the 3DNEPH provides a reasonable average seasonal cloud cover in this region is not sufficient to prove the reliability of the 3DNEPH total cloud cover on a day-to-day grid point basis. This aspect of the 3DNEPH was validated by comparing the 3DNEPH total cloud cover with a visual analysis of AVHRR (channel 4) satellite pictures from the TIROS N. The region that was analyzed corresponded to one swath of the TIROS N that generally covered the northeast quadrant of the region shown in Figure 1. Local crossing times for the TIROS N were generally between 1300 and 1430 UT, which was compared with the 1200 UT 3DNEPH. Comparisons were made for grid boxes that were 2° latitude by 5° longitude, after averaging the 3DNEPH total cloud cover on this grid. Five days from January 1979 were chosen at random for the comparison. The comparison was thus made for approximately 30 grid boxes for each of the 5 days, for a total of about 150 grid boxes. The comparison showed that the 3DNEPH provided accurate total cloud amount on a day-to-day grid point basis. Although this was not an exhaustive comparison, it was probably a sufficient one, given the numerous other studies that have validated the 3DNEPH total cloud cover over oceanic surfaces.

Cloud Types

For purposes of the present study we would like to validate the vertical structure of the cloud layers that are determined by the 3DNEPH. The only published information on the vertical structure of cloud layers has been presented by *Hahn et al.* [1982], in the form of probabilities of the occurrence and cooccurrence of different cloud types as determined from surface observations. Although the probability of occurrence and cooccurrence of different cloud types is not the ideal way to validate the vertical structure of the 3DNEPH cloud layers, a comparison of these statistics between the 3DNEPH and the surface climatology of Hahn et al. will provide useful information regarding the reliability of the vertical layering of the different cloud types in the

TABLE 2. Classification of Cloud Types

Cloud Types Used in This Paper	Shorthand Notation	3DNEPH Cloud Type Codes
Ci/Cs/Cc	CI	hct 1-7
As/Ac	AS	mct 1, 2, 4, 5, 6, 7
Ns	NS	mct 3, 5, 6, 7
Cu	CU	lct 3, 6, 8, 10, 11, 13, 14, 15
St/Sc	ST	lct 1, 2, 5-9, 11-15
Cb	CB	lct 4, 7, 9, 10, 12-14

TABLE 3. Frequency of Occurrence and Cooccurrence of the Different Cloud Types from the 3DNEPH and from *Hahn et al.* [1982]

Cloud Type	Frequency of Occurence	Probability That These Types are also present							Data Source
		CI	ST	AS	NS	CU	CB	No	
CI	22	...	61	49	26	15	8	7	3DNEPH
	29	...	53	63	19	17	11	8	Hahn et al. [1982]
ST	66	20	...	46	7	2	0	41	3DNEPH
	59	29	...	47	18	...	...	35	Hahn et al. [1982]
As	45	23	67	...	0	12	0	18	3DNEPH
	41	46	63	...	1	16	13	7	Hahn et al. [1982]
NS	10	54	45	1	...	19	16	24	3DNEPH
	15	38	70	6	...	2	3	21	Hahn et al. [1982]
CU	9	35	12	56	21	...	0	19	3DNEPH
	20	26	...	33	1	...	...	54	Hahn et al. [1982]
CB	2	96	15	9	91	0	...	1	3DNEPH
	11	28	...	44	4	...	...	43	Hahn et al. [1982]
Clear	10	...	...	...	...	...	...	...	3DNEPH
	4	...	...	...	...	...	...	...	Hahn et al. [1982]

See Table 2 for definitions of cloud types.

3DNEPH, which is of course a prerequisite to obtaining accurate layer cloud amounts.

For the comparison the cloud types from the 3DNEPH are grouped into six categories (Table 2), which is essentially the same as the classification adopted by *Hahn et al.* [1982]. The conditional probability of given type A, that type B also occurs, is given by [*Hahn et al.*, 1982]

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{1}$$

The probability that a cloud type occurs by itself is defined by

$$P(No \mid A) = \frac{P(A \cap No)}{P(A)} \tag{2}$$

where "No" means that no other cloud type exists.

The most frequent cloud type (Table 3) occurring in this region is ST, with a frequency occurrence of 66%. The frequency of ST occurring by itself (indicated by No in Table 3) is about 41%. The frequency of cumuliform clouds is low, 9.3% for CU and 1.8% for CB. The sum of the frequencies of all cloud types is greater than 100, which implies the cooccurrence of different cloud types. Compared with the results of *Hahn et al.* [1982], which is compiled from 12 years (1965-1976) of ship data, the differences in frequency of occurrence of individual cloud types are less than 12%, with the largest differences associated with convective clouds. It seems reasonable to suspect that the 3DNEPH is typing as stratocumulus what a surface observer would call a shallow cumulus cloud; this would account for the 3DNEPH overestimation of ST and underestimation of CU when compared with *Hahn et al.* [1982]. The large discrepancy in the frequency of CB may result either from the misidentification by satellite or from less convective cloud being present during our study period. The occurrence of middle cloud type from both data sources agrees well with occurrence of AS/NS of 55% for 3DNEPH and 56% for Hahn et al. The 3DNEPH shows a higher frequency of clear cases (10%) than Hahn et al. (4%), which may have resulted from the anticyclonic domination of the southern half of this region during the period of 3DNEPH data used in this study.

The contingency probabilities for the 3DNEPH and the surface observations described by Hahn et al. are also shown in Table 3. Overall, a comparison of the contingency probabilities for the two data sets shows fairly good agreement. ST, AS, and CI are seen to have high probabilities of cooccurrence, as would be expected from clouds associated with frontal systems. A principle used by *Hahn et al.* [1982] when computing the contingency probabilities was that the probability $P(U \mid L)$ of an upper cloud U, given a lower cloud L, is assumed to be the same when U cannot be seen (because L is overcast) as when it can be seen (when L is present but not overcast). For 3DNEPH, since satellite and surface data are combined, there is no need to use such an assumption. However, by comparing both data sets the validity of this assumption can be examined. Large discrepancies (20%) are seen in Table 3 for the contingency probabilities $P(\text{AS} \mid \text{CI})$, $P(\text{NS} \mid \text{ST})$, $P(\text{CU} \mid \text{AS})$, $P(\text{CU} \mid \text{No})$, $P(\text{CB} \mid \text{CI})$, $P(\text{CB} \mid \text{AS})$, $P(\text{CB} \mid \text{NS})$ and $P(\text{CB} \mid \text{No})$. The largest discrepancies are associated with CB. The 3DNEPH has a very high cooccurrence of CB with CI, which seems reasonable, since mid-latitude CB are frequently associated with cirrus anvils. In contrast, Hahn et al. show a low $P(\text{CB} \mid \text{CI})$ and a relatively high $P(\text{CB} \mid \text{AS})$. This difference seems especially surprising considering the possibility that 3DNEPH may be misclassifying CI as AS if the CI emissivity is less than one (this may also be a factor in the discrepancy for $P(\text{AS} \mid \text{CI})$). The relatively high $P(\text{CB} \mid \text{AS})$ in the Hahn et al. data set, when compared with the relatively high $P(\text{CB} \mid \text{NS})$ in the 3DNEPH data set, may simply indicate a difference in classification of $\text{AS} \mid \text{NS}$. Hahn et al. show CU occurring primarily by itself, while the 3DNEPH shows CU occurring primarily with AS. It is not clear

which of these situations is more reasonable. However, for the cloud types that predominate in this region, namely ST and AS, the contingency probabilities show fairly good consistency between the two data sets.

Cloud Overlap Statistics

Three different overlap assumptions for multiple layers that are partially cloudy are used in this analysis, namely the minimum, maximum and random overlap assumptions. Assuming $A_1, A_2, ..., A_n$ are the cloud amounts for each layer, the total cloud cover (TC) computed from the minimum, maximum, and random overlap assumptions are

$$TC_{MIN} = \text{Min}\left(\sum_{i=1}^{n} A_i, 1\right)$$

$$TC_{MAX} = \text{Max}\,(A_1, A_2, ..., A_n) \qquad (3)$$

$$TC_{RAN} = 1 - (1 - A_1)(1 - A_2)...(1 - A_n)$$

The methodology of the cloud overlap analysis is to compare the difference between the observed total cloud cover determined by the 3DNEPH (TC) and total cloud cover obtained by the different overlap assumptions (TC_{MIN}, TC_{MAX} or TC_{RAN}), using the cloud amounts determined by 3DNEPH for each of the 15 vertical layers. Since layered cloud amounts in the 3DNEPH are rounded to the nearest 5%, the roundoff error for each layer may be as large as 2.5%. Assuming that in reality cloud fractions are distributed equally in a 5% interval, the average roundoff error is 1.25%. In computing total cloud amount using the different overlap assumptions this roundoff error will propagate. Suppose y is a function of $x_1, x_2, ..., x_n$, that is, $y = f(x_1, x_2, ..., x_n)$, and the errors in $x_1, x_2, ..., x_n$ are $\Delta x_1, \Delta x_2, ..., \Delta x_n$. The error for y can be

$$\Delta y = \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 + ... + \frac{\partial f}{\partial x_n}\Delta x_n \qquad (4)$$

Using (4), the errors for the computed total cloud amounts are

$$\Delta TC_{MIN} = \sum_{i=1}^{n} \Delta A_i$$

$$\Delta TC_{MAX} = \text{Max}\,(\Delta A_1, \Delta A_2, ..., \Delta A_n) \qquad (5)$$

$$\Delta TC_{RAN} = \sum_{j=1}^{n} \prod_{\substack{i=1 \\ (i \neq j)}} (1 - A_i)\,\Delta A_j$$

where A_i is the cloud fraction and ΔA_i is the roundoff error. Table 4 gives the values of the maximum roundoff errors incurred for the different overlap assumptions for n=2 and n=3 cases, where n is the number of overlapping cloud layers and ΔA_i=2.5% is used. The average roundoff errors were computed using ΔA_i=1.25%, which assumes that in reality the cloud fraction amounts are equally distributed over the 5% interval. Note that the average roundoff error is simply one half of the maximum roundoff error for TC_{MAX} and TC_{MIN}, while the roundoff error for TC_{RAN} varies with the cloud fraction of each of the layers. ΔTC_{RAN} is largest if all of the cloud layers have small cloud fractions and smallest if all of the cloud layers have large cloud fractions.

TABLE 4. Maximum (ΔA_i =2.5%)and Average (ΔA_i =1.25%) Roundoff Errors for the Computed Total Cloud Cover for n=2 and n=3 Cases

Error	Maximum Error n=2	Maximum Error n=3	Average Error n=2	Average Error n=3
$\lvert\Delta TCM_{MAX}\rvert$	2.5	2.5	1.25	1.25
$\lvert\Delta TC_{MIN}\rvert$	5.0	7.5	2.50	3.75
$\lvert\Delta TC_{RAN}\rvert$	4.8	7.5	2.46	3.00

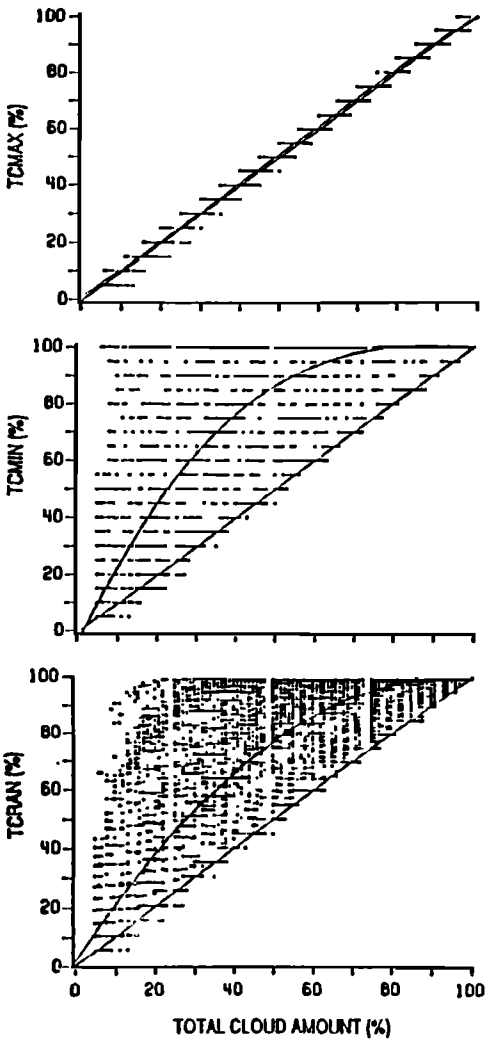

Fig. 2. Scatter plots of observed total cloud amounts (TC) versus computed total cloud amounts (TC_{MIN}, TC_{MAX} and TC_{RAN}) for cases which have cloud in adjacent layers that are not separated by a clear interstice for (45 km)^2 grid.

Overlap of Adjacent Cloud Layers

The different cloud overlap assumptions are tested in order to choose an appropriate one to combine the adjacent vertical levels in the 3DNEPH that are cloudy. For simplicity, only cases which have cloud in adjacent layers that are not separated by a clear interstice (52,649 observations) are used for this testing. Figure 2 gives scatter plots which show the relations between observed (TC) and computed (TC_{MIN}, TC_{MAX}, TC_{RAN}) total cloud cover using the different overlap assumptions. Two lines are drawn on these diagrams, one with a 45° slope which represents the line corresponding to a perfect fit between the observed and calculated cloud cover, and the other line representing a cubic least squares fit to the data points. Both TC_{MIN} and TC_{RAN} are seen to

systematically overestimate the total cloud cover. For the maximum overlap assumption (TC_{MAX}, Figure 2*a*), 81.7% of the cases approximate TC to within ±2.5%, which is the maximum roundoff error for TC_{MAX}. TC_{MAX} results in a small systematic overestimation of total cloud cover by 0.5%. Clearly the maximum overlap assumption is appropriate to use for adjacent levels in the 3DNEPH that contain cloud.

Overlap of Discrete Cloud Layers

For the bulk of the cloud overlap analyses the original 15 levels of the 3DNEPH are reduced to 1-5 cloud layers (n=1 to 5) by combining adjacent 3DNEPH layers that have cloud, using the maximum overlap assumption. This has the obvious advantage of simplifying the overlap calculations but, more importantly, yields discrete cloud layers that correspond to observed cloud layers. Among the 93,380 observations analyzed in this data set, 56.4% are one layer (n=1), 24.3% have two layers (n=2), and 7.9% have three layers (n=3). Only 0.5% of the cases have more than 3 layers. The n=2 and n=3 cases where none of the individual cloud layers had 100% cloud cover (a total of 24,000 observations) are used to examine the overlap of discrete cloud layers.

Figure 3 is the same as Figure 2, except that it describes the n=2 and n=3 cases. The maximum overlap assumption is seen to

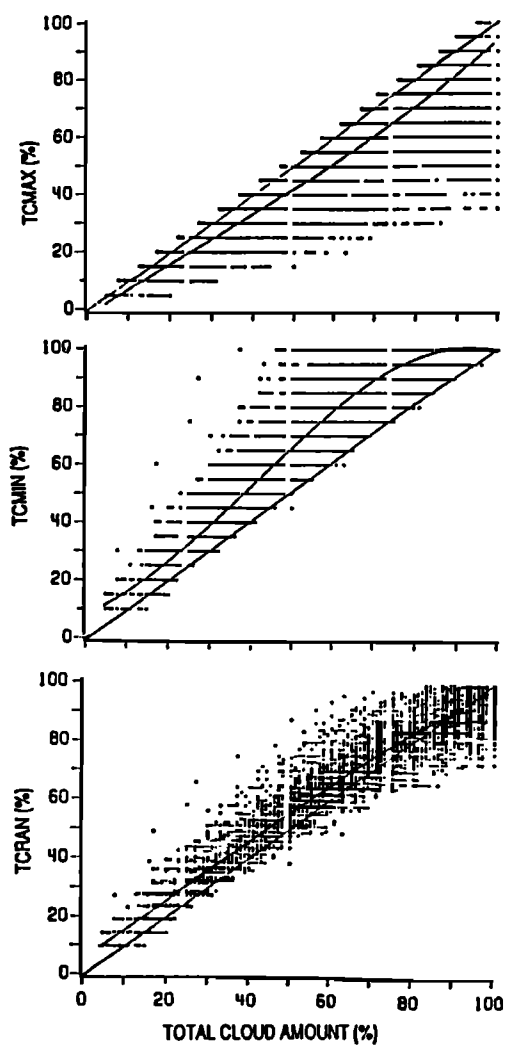

Fig. 3. Scatter plots of observed total cloud amounts versus computed total cloud amounts for n=2 and n=3 cases for (45 km)^2 grid.

TABLE 5. Errors in Total Cloud Cover Derived by the Overlap Assumptions When Compared With the Observed Total Cover (TC)

	Systematic Error $TC_{calculated}$-TC		Random Error $\lvert TC_{calculated}\text{-}TC \rvert$		Maximum Error $MAX[\lvert _{calculated}\text{-}TC \rvert]$	
	n=2	n=3	n=2	n=3	n=2	n=3
TC_{MAX}	-7.4	-14.6	7.8	14.7	47	65
TC_{MIN}	9.7	10.7	9.7	10.7	47	63
TC_{RAN}	1.3	1.1	4.3	5.3	28	30

systematically underestimate the total cloud cover and the minimum overlap assumption to overestimate the total cloud cover. The overestimation for the minimum overlap assumption is quite large for total cloud cover, that is, 50-90%. There is a slight bias toward overestimation for the random overlap assumption, which is greatest for the lower values of total cloud amount. Table 5 shows the average systematic and random errors associated with the different overlap assumptions. The magnitude of the systematic underestimation by the random overlap assumption is 1.2%, while the systematic biases for maximum and minimum overlap are considerably larger. For the maximum overlap assumption the errors are significantly larger for the n=3 cases than for the n=2 cases. The average magnitude of the error for the random overlap assumption is 4.3% for n=2 and 5.3% for n=3, with the magnitude of the largest error being 39%. As expected, the random overlap assumption performs the best at describing the overlap of discrete cloud layers. However, there is a slight but systematic overestimation of total cloud cover by the random overlap assumption, which is greatest for total cloud amounts <60% (Figure 3), and the magnitude of errors in individual cases can be large.

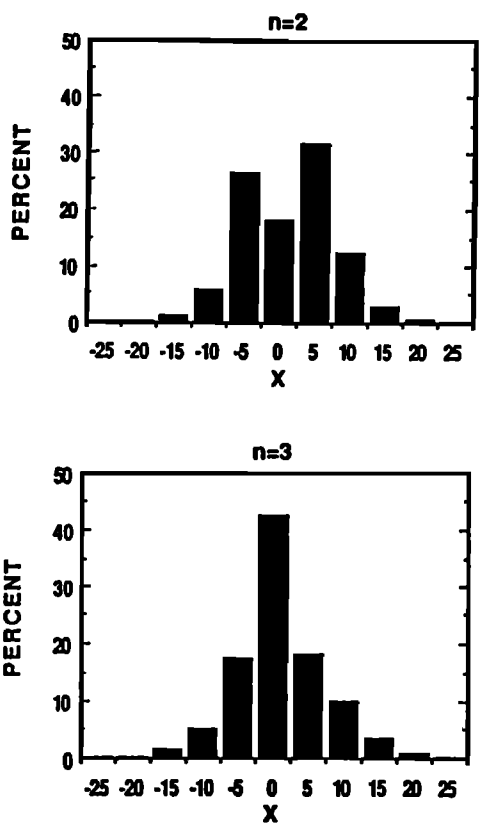

Fig. 4. Frequency distribution of the difference between the random overlap approximation and the observed total cloud cover that exceeds the roundoff error, $\chi = TC_{RAN} - TC \pm \Delta TC_{RAN}$.

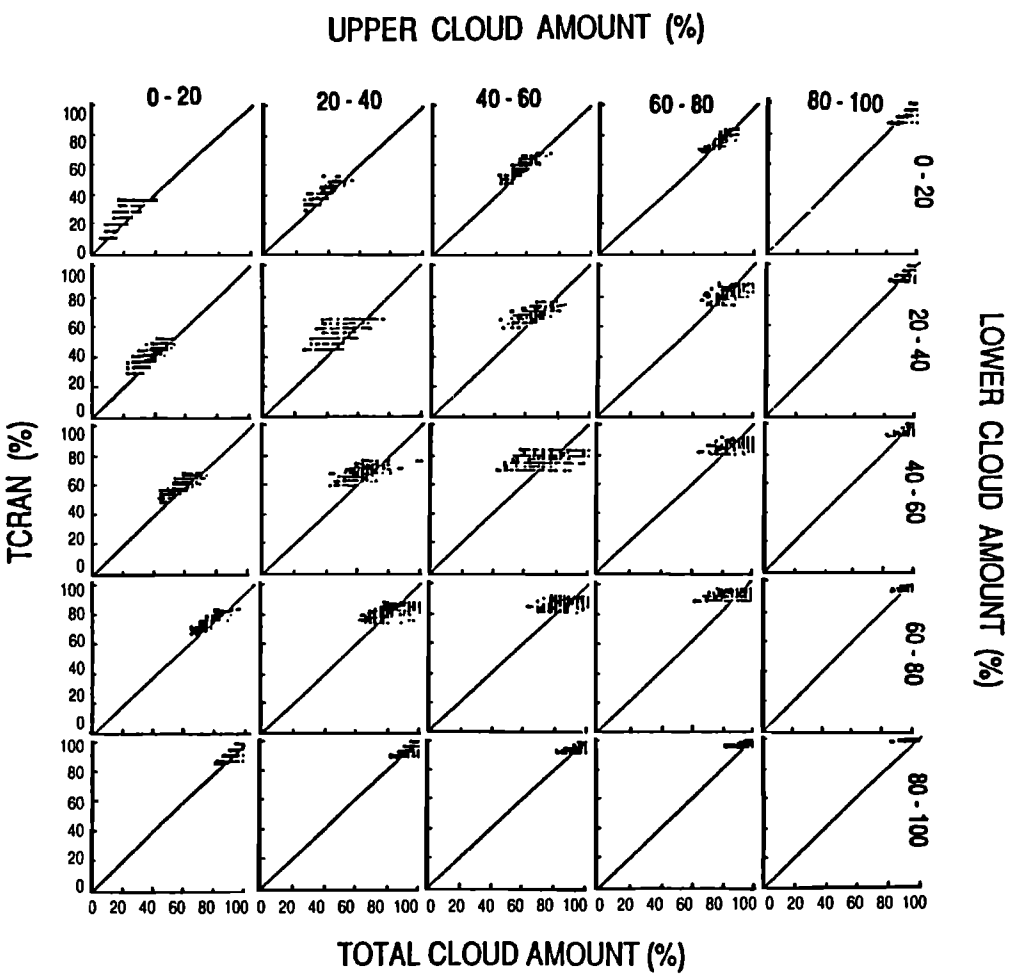

Fig. 5. Scatter plots of observed total cloud amounts versus cloud amounts computed using the random overlap assumption for n=2 cases for the (45 km)^2 grid.

Figure 4 gives the frequency distribution of [X=TC_{RAN} - TC $\pm$ ΔTC_{RAN}] (minus used for overestimation by TC_{RAN} and plus used for underestimation by TC_{RAN}), where the average roundoff A_f=1.25% was used. The bar X=0 gives the percent of the cases that TC_{RAN} minus TC is within the roundoff error (the roundoff error for each individual case was used here, rather than the average roundoff error given in Table 4). The bar X=5 gives the percent of the cases that are overestimated by TC_{RAN}, where the overestimation exceeded the roundoff error by 5% or less. For two layers (n=2) it is seen in Figure 4 that 18% of the cases are estimated by TC_{RAN} to within the roundoff error, and 76% of the cases are estimated to within 5% of the roundoff error. For 3 layers, 43% of the cases are estimated by TC_{RAN} to within the roundoff error, and 77% of the cases are estimated to within 5% of the roundoff error. These results indicate that any bias toward random overlap that is built into the 3DNEPH (as discussed in section 2) appears to be quite small.

The error associated with the random overlap assumption is further investigated by considering only the two-layer cases (n=2) and categorizing the observations according to the cloud amount in each layer. The latter is accomplished by considering the layer cloud amounts in 20% cloud cover intervals, for example, 0-20%, 20-40%, etc., defining each category by its midpoint. Figure 5 shows scatter diagrams of the 25 resulting categories of the overlapping of two cloud layers. Both systematic and random errors are evident, with a variation seen across the different categories. Note that the results for two symmetric categories (e.g., 10% lower layer and 70% upper layer versus 70% lower layer and 10% upper layer) are not necessarily the same. Positive bias associated with the random overlap assumption is largest when both layers have cloud amounts less than 50% and also when the lower layer has a cloud fraction of 70%. Negative bias occurs when the upper cloud amount is 70-90% and the lower cloud amount is less than or equal to 50%. The largest random errors are seen in Figure 5 to be associated with cases where both layers have cloud amounts between 30 and 70%.

Influence of Horizontal Resolution

The appropriate overlap assumption to use may depend on the spatial scale under consideration. The use of an overlap assumption in a numerical model of the atmosphere may therefore depend on the horizontal grid spacing of the model. A mesoscale model with grid spacing less than 100 km may require a different cloud overlap assumption than a general circulation model with a grid spacing exceeding 200 km. For example, when the spatial scale is large enough to contain parts of more than one synoptic situation in a grid box, minimum overlap may become more prominent, since, for example, the cumulus forming in one corner of a 500-km grid box will not overlap the cirrus formed in a distant part of the box.

To study the effects of changing horizontal resolution on cloud overlap properties, three additional data sets with different

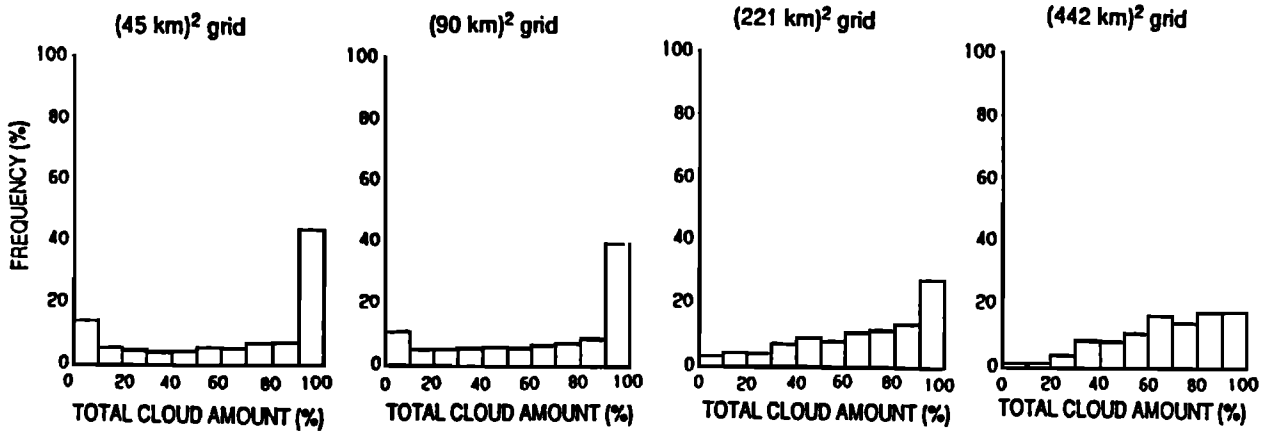

Fig. 6. Frequency distribution of total cloud amount for each horizontal resolution.

horizontal resolutions, (90 km)^2, (221 km)^2, and (442 km)^2, are created from the original one. Horizontal averaging to obtain these data sets poses problems if the number of overlapping cloud layers is to be kept to a manageable number. The procedure used for the horizontal averaging is described as follows. First, the 1-5 layers (which were reduced from the original 15 layers using maximum overlap for the adjacent levels containing clouds) are classified as low, middle, or high cloud layers according to their cloud bases in each of the (45 km)^2 boxes. Low clouds have bases below 2000 m, middle clouds have bases between 2000-6000 m, and high clouds have bases above 6000 m. Clouds of vertical development are classified as low clouds, since the cloud bases exist in the low cloud height interval. In only 8% of the (45 km)^2 boxes, more than one cloud layer occurred in a single-layer category (e.g., two stratus layers occurring in the low cloud category). In these cases the appropriate layer cloud amounts are obtained by combining the cloud layers using random overlap assumption. As shown in Figure 4, this procedure of reducing vertical resolution will cause at most 24% of the (8%) cases to have an error that exceeds the roundoff error by more than 5%. The low, middle, and high cloud amounts are then averaged over (90 km)^2, (221 km)^2, and (442 km)^2 grids.

As pointed out by *Falls* [1974], observations of cloudiness over small areas may be expected to lead to a U-shape frequency distribution, since the probability of observing either completely clear sky or completely cloudy sky increases as the observation area decreases. Increasing the horizontal scale under consideration finally leads to a bell-shaped distribution, while transitional shapes of the frequency curves may be anticipated at intermediate resolution scale. Figure 6 shows the frequency distribution of the total cloud amounts for four different horizontal resolutions using the 3DNEPH data. For the 45 and 90 km grids a predominance of clear and overcast cases are seen, while for the 442 km grid a more nearly monomodal distribution of cloud amounts is seen, centered near the mean cloud amount.

The effect of different spatial scales on the validity of the different cloud overlap assumptions is considered by examining the overlap assumptions for the three data sets created and comparing them with those derived for the original grid size, ~ (45 km)^2. Coarsening the resolution reduces the percent of clear (n=0) and one-layer (n=1) cases and hence increases the percent of n=2 and n=3 cases, which are included in the overlap analysis. Table 6 gives the percent of the observations for the different number of layers as a function of horizontal resolution. At the lowest resolution (442 km)^2, clear and single layer cases are virtually nonexistent, while at the highest resolution these cases comprise more than half of the observations.

Table 7 compares the systematic differences between the observed total cloud cover and the cloud cover determined from the different overlap assumptions, for the different horizontal resolutions. Since the random errors show little variation with the horizontal resolution, they are not shown here. It is seen in Table 7 that the systematic error for the random overlap approximation goes from slightly positive (overestimation) at 45 km resolution to a systematic underestimation of ~5% at the coarser resolutions. The systematic underestimation for maximum overlap increases in magnitude as resolution decreases, while that for the minimum overlap shows little variation as the resolution changes. A comparison of the results for n=2 and n=3 cases shows that the errors are slightly larger for the n=3 cases compared with the n=2 cases.

The scatter diagrams of TC versus TC_{MIN}, TC_{MAX}, and TC_{RAN} for the 221-km resolution are shown in Figure 7 (the 90 km and 442-km resolution plots show similar features and are not included for brevity). The plots for 221-km resolution in Figure 7 show basically the same features as those in Figure 3 for the 45-km resolution. A notable difference, however, is the increasingly large systematic error shown by TC_{RAN} for TC >70%.

TABLE 6. Number of Observations and Cloud Layer Distribution for Different Horizontal Resolutions

	(45 km)^2	(90 km)^2	(221 km)^2	(442a km)^2
Observations	93,380	23,352	1,792	448
Number of grid boxes	3,336	834	64	16
Percent n=0	11	5	0	0
Percent n=1	56	35	5	1
Percent n=2	24	46	33	7
Percent n=3	8	13	61	93

TABLE 7. Systematic Errors in Overlap Assumptions [$TC_{calculated}$-TC] for Different Horizontal Resolutions

	45 km	90 km	221 km	442 km
TC_{MAX}	-74	-14.4	-14.9	-16.9
TC_{MIN}	8.1	5.1	6.2	7.9
TC_{RAN}	1.1	-4.9	-4.7	-4.8

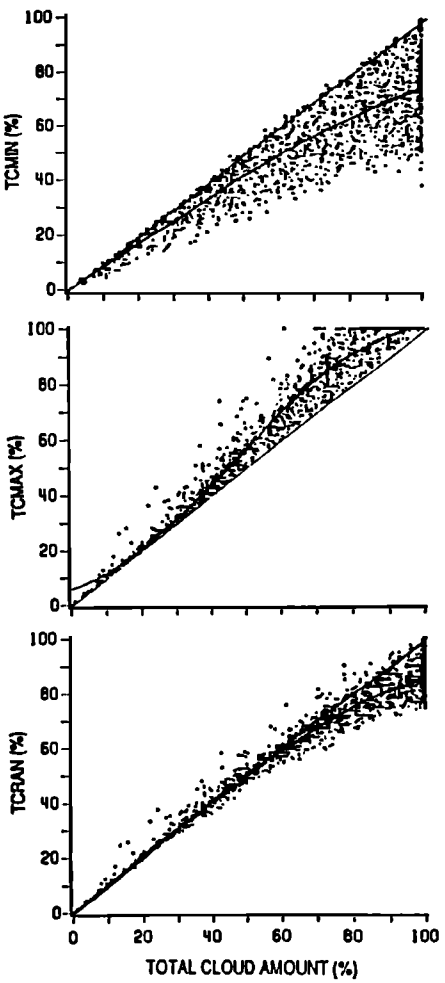

Fig. 7. Scatter plots of observed total cloud amounts versus computed total cloud amounts for n=2 and n=3 cases for (221 km)^2 grid.

Although random overlap performed reasonably well with little systematic error for the grid size of 45 km, at larger grid sizes typical of those found in general circulation models the random overlap assumption results in a systematic underestimate of total cloud fraction by about 5%. The underestimation by the random overlap assumption at larger scales can be explained as follows: For the original 45-km grid, 56% of the cases had only one cloud layer and thus were not included in the overlap statistics. As several of these grid boxes are combined to determine the overlap for a larger region, many of the n=1 cases will be incorporated into the overlap statistics, promoting a tendency toward minimum overlap. Consider the following extreme example, whereby four boxes which contain one cloud layer each are averaged. If all of the boxes contain only low cloud, then by definition there will be no overlap. However, if two of the boxes have low cloud and two of the boxes have middle cloud, the resulting overlap for the average of the four boxes will of course be minimum overlap. In the more complicated case when many grid boxes are averaged, some of which have one layer and some of which have several layers, the resulting overlap will be some sort of weighted average according to the number of grid boxes with more than one layer (random overlap) and the number of grid boxes with one layer (minimum overlap). Examination of Table 6 shows that for 45-km horizontal resolution, approximately half of the cases have only one cloud layer. This effect seems to have the most significance when the total cloud cover exceeds 70%, which is approximately the value of the mean cloud cover for the region. In light of these results the following algorithm is suggested for grid sizes exceeding 45 km:

$$TC_{overlap} = 0.5\,TC_{RAN} + 0.5\,TC_{MIN} \quad \text{if } TC_{RAN} > \overline{TC}$$

$$TC_{overlap} = TC_{RAN} \quad \text{if } TCRAN < \overline{TC} \qquad (6)$$

where $\overline{TC}$ is the mean cloud cover for the region. The coefficients used in (6) (i.e., 0.5) were chosen to represent the likely proportion of n=1 versus n>1 cases for $TC_{RAN} > TC$. A scatter plot of (6) is shown in Figure 8, along with the plot of TC_{RAN}, for horizontal resolution of 442 km. The plot of $TC_{overlap}$ using (6) shows that there is no overall systematic bias for this approximation, but there is a slight bias toward underestimation for TC > 90% and a slight overestimation for 70% < TC < 90%.

Summary and Conclusions

The U.S. Air Force 3DNEPH has been used to examine cloud overlap statistics during January 1979 over the north Atlantic Ocean. Efforts were made to determine the reliability of the 3DNEPH total cloud cover by comparison with other January climatologies and also with a visual analysis of AVHRR satellite pictures for the same period. The layered structure of the 3DNEPH was examined by comparing the occurrence and cooccurrence of different cloud types with the surface cloud climatology of *Hahn et al.* [1982]. The results of these comparisons indicate that the 3DNEPH is providing reliable total cloud cover and vertical layering of cloud types although the layer cloud amounts must be regarded as somewhat uncertain because of the lack of any independent method to verify them. As far as

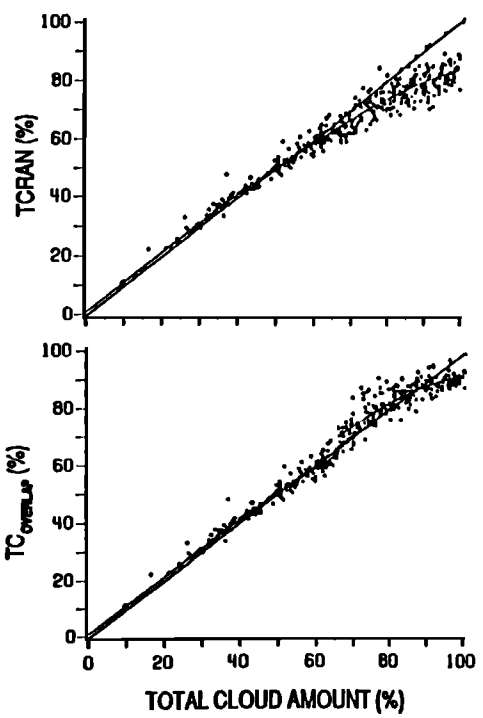

Fig. 8. Scatter plots of observed total cloud amounts versus cloud amounts computed using the random overlap assumption and using (6) for (422 km)^2 grid.

we can tell, the 3DNEPH in this particular region provides an adequate data set to examine cloud overlap; the reliability of the data set is believed to be enhanced in the regions analyzed due to a combination of ocean background (for reliable satellite retrievals) and the large number of conventional observations.

Analysis of the overlap of adjacent 3DNEPH levels containing cloud showed that these layers could be combined to a high degree of accuracy using the maximum overlap assumption. Using the maximum overlap assumption for adjacent 3DNEPH levels containing cloud, the 15 vertical levels of the 3DNEPH were reduced to 1-5 discrete cloud layers. The maximum, minimum, and random overlap assumptions were applied to the discrete cloud layers, and the results were compared with the observed total cloud cover. It was shown that maximum overlap assumption results in a systematic underestimation of total cloud cover, while minimum overlap assumption results in an overestimation. The random overlap results in a small systematic overestimation of the total cloud cover, of approximately 1%. The average magnitude of the difference between the total cloud cover estimated by the random overlap approximation and the observed cloud cover was approximately 4%, which exceeds the roundoff error associated with the overlap assumption by slightly more than 1%. Although the random overlap assumption has been shown to perform fairly well for the original 3DNEPH grid size (~ 45 km), it must be kept in mind that there may be an inherent bias toward random overlap in the 3DNEPH. The magnitude of this bias is uncertain, but it is easily seen from the results that this bias does not dominate the results due to the large scatter observed.

Although the overlap results for the original 45-km resolution may be somewhat biased toward random overlap, averaging the 3DNEPH data over larger regions will reduce the role of any such bias, since the averaging will incorporate the cases containing only one cloud layer (~56% of the total number of 3DNEPH grid cases). After averaging a number of 3DNEPH grid boxes to result in larger grid boxes with horizontal resolutions of $(90\ km)^2$, $(221\ km)^2$, and $(442\ km)^2$, it was shown that the random overlap assumption systematically underestimated the total cloud cover by about ~5% for the coarser resolutions. The underestimation was most marked for cases where the total cloud cover was observed to exceed 70%. This underestimation was attributed to the increasing influence of areas with a single layer of cloud, which will bring the situation toward minimum overlap. The bias toward underestimation for cases with large total cloud cover was eliminated by using the mean of the random and minimum overlap approximations if the random overlap assumption indicated a total cloud cover exceeding 70%, which is the mean total cloud cover for this region.

These results were obtained using a limited data set; for different regions the results may be different in a quantitative sense, although the qualitative assessments presented here are reasonable in terms of simple geometric considerations and probably have a more widespread validity. Regional factors which may affect the nature of the overlap include the horizontal extent of the cloud decks and the average total cloud cover.

The implications of these results for including cloud overlap assumptions for radiative transfer calculations in atmospheric models are as follows: Adjacent levels containing cloud may be accurately combined by using maximum overlap to form a discrete cloud layer. For two or three discrete cloud layers, use of the minimum overlap assumption results in systematic overestimation of total cloud cover, and the use of maximum overlap results in a substantial underestimation of total cloud cover, especially for the large grid sizes employed in general circulation models. Random overlap of discrete cloud layers, which is the most widely used overlap assumption, performed the best, although this assumption results in a systematic underestimation of the total cloud cover for the horizontal resolutions used in general circulation models. The underestimation by the random overlap assumption for large grid sizes that was obtained here was ~5%; the significance of an error of this size in total cloud cover is illustrated by the result of *Randall et al.* [1984] that an increase of global total cloud cover by 4% would be sufficient to offset the warming associated with a doubling of CO_2.

The present study has raised a number of issues regarding the appropriate cloud overlap assumption to use for radiative transfer calculations in numerical models that predict or diagnose layer cloud amounts. These issues need to be addressed using a larger data set to examine possible regional differences. The 3DNEPH data set appears to be reliable for this type of study, provided that there are sufficient conventional data in the analysis region, although there may be a slight bias toward random overlap in this data set. Additionally, there may be other aspects of the 3DNEPH data processing that we are unaware of that affect the overlap statistics. Although the 3DNEPH layer cloud amounts seem reasonable, they must be considered as somewhat uncertain because of the lack of any method to verify their accuracy.

In spite of the uncertainties in the present analysis it has been shown that the overlap assumptions currently employed deserve more careful consideration, and future efforts should be made to examine this subject.

Acknowledgments. This research has benefitted from several discussions with S. Warren. Comments from Harshvardhan and E. Agee are also gratefully acknowledged. Discussions with L. Hamill, AFGWC, and P. Giese, USAFETAC, helped clarify the 3DNEPH merge processor. Comments from the two anonymous reviewers helped to improve the manuscript. We would also like to thank Chris Ardeel for his assistance. This research was supported by NSF grant ATM-850527.

References

Bjorheim, K., P. Julian, M. Kanamitsu, P. Kallbert, P. Price, S. Tracton, and S. Uppala, FGGE III-B daily global analyses, 1, December 1978-February 1979, ECMWF Report, Eur. Cent. for Medium Range Weather Forecasts, Reading, England, 1981.

Curry, J.A., and Herman, G.F., Relationship between large-scale heat and moisture budgets and the occurrence of Arctic stratus clouds, *Mon. Weather Rev., 113*, 1441-1457, 1985.

Falls, L.W., The beta distribution, A statistical model for world cloud cover, *J. Geophys. Res., 79*, 1261-1264, 1974.

Fye, F. K., The AFGWC automated cloud analysis model, *Tech. Memo. 78-002*, U.S. Air Force Global Weather Central, Offutt Air Force Base, Nebr., 1978.

Geleyn, J.F., and A. Hollingsworth, An economical analytical method for the computation of the interaction between scattering and line absorption of radiation, *Contrib. Atmos. Phys., 52*, 1-16, 1979.

Gordon, C.T., R.D. Hovanec, and W.F. Stern, Analyses of monthly mean cloudiness and their influence upon model-diagnosed radiative fluxes, *J. Geophys. Res., 89*, 4713-4738, 1984.

Hahn, C.J., S.G. Warren, J. London R.M. Chervin, and R. Jenne, Atlas of simultaneous occurrence of different cloud types over the ocean, *NCAR Technical Note TN-201+STR NCAR*, Natl. Cent. for Atmos. Res., Boulder, Colo., 1982.

Hughes, N.A., Global cloud climatologies: A historical review, *J. Clim. Appl. Meteorol., 23*, 724-751, 1984.

Hughes, N.A., and A. Henderson-Sellers, Global 3D Nephanalysis of total cloud amount climatology for 1979, *J. Clim. Appl. Meteorol., 24*, 669-686, 1985.

Koenig, G., and K.-N. Liou, Cloud climatology and radiative budget studies from Air Force 3D-NEPH data base, *J. Geophys. Res., 92*, 5540-5554, 1987.

London, J., A study of the atmospheric heat balance. *Final Rep. AFCRC-TR-57-287*, 99pp., New York University, N.Y., 1957.

Morcrette, J.J., and Y. Fouquart, The overlapping of cloud layers in shortwave radiation parameterization, *J. Atmos. Sci., 43*, 321-328, 1986.

Ramanathan, V., E.J. Pitcher, R.C. Malone, and M.L. Blackmon, The response of a spectral general circulation model to refinement in radiative process, *J. Atmos. Sci., 40*, 605-630, 1983.

Randall, D.A., J.A. Coakley, C.W. Fairall, R.A. Kropfli, and D.H. Lenschow, Outlook for research on subtropical marine stratiform clouds, *Bull. Am. Meteorol. Soc., 65*, 1290-1301, 1984.

Stephens, G.L., The parameterization of radiation for numerical weather prediction and climate models, *Mon. Weather Rev., 112*, 826-867, 1984.

Stowe, L.L., Evaluations of NIMBUS 7 THIR/CLE and Air Force Three-Dimensional Nephanalysis estimates of total cloud amount, *J. Geophys. Res., 90*, 5370-5380, 1984.

Warren, S.G., C.J. Hahn, J. London, R.M. Chervin, and R.L. Jenne, Global distribution of total cloud cover and cloud type amounts over land, *Tech. Note NCAR/TN-273 + STR. Natl. Cent. for Atmos. Res.*, Boulder, Colo., 1986.

J.A. Curry Department of Meteorology, The Pennsylvania State University, University Park, PA 16802.

L. Tian Department of the Geophysical Sciences, The University of Chicago, Chicago, IL 60637.

(Received August 17, 1988;
revised December 10, 1988;
accepted March 17, 1989.)