Cloud Predictions Diagnosed from Global Weather Model Forecasts

DONALD C. NORQUIST

Air Force Research Laboratory, Hanscom AFB, Massachusetts

(Manuscript received 12 October 1999, in final form 7 March 2000)

ABSTRACT

The U.S. Air Force has a long history of investment in cloud analysis and prediction operations. Their need for accurate cloud cover information has resulted in routine production of global cloud analyses (from their RTNEPH analysis model) and forecasts (using their ADVCLD cloud trajectory forecast model) over many years.

With the advancement of global numerical weather prediction technology and resulting forecast accuracy of noncloud meteorological quantities, it is of interest to determine if such technology could be used to benefit cloud cover forecasting. In this paper, a model output statistics approach to diagnosing cloud cover from forecast fields generated by a global numerical weather prediction model is presented. Cloud characteristics information obtained from the RTNEPH cloud analysis supplied the cloud predictands, and forecast fields from the U.S. Navy Operational Global Atmospheric Prediction System global weather prediction model provided the weather variable predictors. RTNEPH layer cloud cover was assigned to three cloud decks (high, middle, and low) based on reported cloud-base altitude, and RTNEPH total cloud cover was used as a separate predictand. Multiple discriminant analysis (MDA) was used to develop the predictand–predictor relationships for each cloud deck and total cloud using 5 days of twice-daily cloud analyses and corresponding forecasts for 30° latitude zones. The consequent relationships were applied to the forecasts fields from the forecast initialized on the day following each 5-day development period to diagnose cloud cover forecasts for the Northern or Southern Hemisphere.

In this study, cloud cover forecasts were diagnosed from global NWP model forecasts on hemispheric polar stereographic map projections with a grid spacing of 96 km. The diagnosed cloud forecasts (DCFs) were verified against the RTNEPH analyses for forecast durations of 12–72 h at 12-h intervals. Also verified were 12–48-h cloud cover forecasts (deck and total) from the ADVCLD cloud trajectory model, and from persistence (RTNEPH at initial forecast time). Biases from all three methods were generally small. The DCFs were significantly better than ADVCLD and persistence in all decks and total cloud, at almost all forecast durations in rmse and 20/20 score. ADVCLD scored better in these measures only at 12 h in total cloud, suggesting the possibility of a crossover in superior prediction skill from trajectory to diagnostic method somewhere between 12 and 24 h. DCF better preserved the cloud cover frequency distribution than did ADVCLD. ADVCLD displayed a greater degree of spatial variation inherent in RTNEPH cloud cover than did DCF. Both ADVCLD and DCF visual depictions of hemispheric total cloud cover appeared to provide useful cloud cover forecast information when compared with RTNEPH depictions. The advantages of the diagnosed cloud forecast algorithm presented in this study make it an excellent candidate for operational cloud cover prediction. It is expected that as cloud cover analyses are improved, the trajectory and diagnostic methods will prove complementary with the former more skillful at short-term predictions, and the latter better at long-term forecasts.

1. Introduction

Predictions of the cloud cover (CC) over large-scale grid areas (on the order of 100 km on a side) have long been of interest to the U.S. Air Force. Intercontinental flights common in military mobilizations and exercises and a need for mission planning for ground surveillance activities are just two of the reasons for such an interest. The air force has made a significant investment in the analysis and prediction of large-scale cloudiness. Past air force focus has primarily been on techniques based on satellite imagery for nowcasting and on the trajectory of the imagery-depicted clouds into the future for short-term forecasting purposes.

The mainstay of the air force cloud analysis and prediction program for many years has been the real-time nephanalysis model used operationally at the Air Force Weather Agency and referred to as RTNEPH (Hamill et al. 1992). This model combines imagery obtained from polar orbiting satellites (principally, the Defense Meteorological Satellite Program satellites, and secondarily the National Oceanic and Atmospheric Administration satellites) with surface observations of CC to produce global analyses of CC every 3 h. The grid areas over which the CC values are computed from the imagery data (approximately 48 km true at 60° latitude on a polar stereographic grid) provide ample

Corresponding author address: Mr. Donald C. Norquist, AFRL/VSBL WTU, % AFWA/DNX, 106 Peacekeeper Drive, Offutt AFB, NE 68113-4039.
E-mail: Donald.Norquist@afwa.af.mil

spatial resolution for many air force global-scale applications.

Total CC is the fractional surface obscuration by all clouds when looking down from above the highest cloud layer. Layer cloud cover is defined as the fraction of a specified area of the earth's surface that is obscured by any single cloud layer when looking down on that layer from above. The RTNEPH provides estimates of total CC for each grid area, and layer CC for each grid area for as many as four cloud layers. Also included for each cloud layer is a specification of cloud type, base altitude, and top altitude for each grid area occupied by cloud.

Other cloud analysis products have been developed in the 16-yr lifetime of RTNEPH. One of the more well known is the suite of cloud analyses produced by the International Satellite Cloud Climatology Project (ISCCP; Rossow and Schiffer 1991; Rossow et al. 1996). The ''gridded cloud product'' datasets produced by the ISCCP are based entirely on satellite imagery. They contain CC for equal-area grid areas of 280 km on a side. The grid-area averages that constitute the CC values are computed from 30-km pixel imagery data. The coarser resolution of the ISCCP product lends itself well to climatological applications (for which it was intended) but is not of sufficient spatial resolution to meet most military application requirements.

While satellite analysis products are useful to the military as nowcasting aids, there is a need to know future cloud coverage distribution for mission planning. In cloud prediction as well as cloud analysis, the air force has a history of extensive operational experience. Crum (1987) documents the history and design of several trajectory-based cloud forecast models used at the Air Force Weather Agency (AFWA; formerly, Air Force Global Weather Central) to produce global-scale cloud predictions. Kopp et al. (1997) described the updated Air Force Weather Agency cloud forecast model known as ADVCLD, which is also a trajectory-based model. ADVCLD produces global, 3-hourly CC predictions of total cloud and cloud at five pressure levels for grid areas approximately 48 km on a side (same grid as RTNEPH) out to 12 h and for grid areas of about 96 km on a side out to 48 h. ADVCLD is initialized from RTNEPH analyses and uses wind data for the trajectories from the navy Operational Global Atmospheric Prediction System (NOGAPS) numerical weather prediction model (Hogan and Rosmond 1991).

Diagnostic cloud prediction methods have primarily used variants of the model output statistics (MOS) technique, which involve the development of statistical relationships between the numerical weather prediction (NWP) model forecast variables (and others derived from them) as ''predictors'' and the desired cloud variables (such as CC) as ''predictands.'' These relationships are then applied to future forecasts from the same model to ''diagnose'' the cloud variables. MOS techniques use the physics of the model only implicitly in the ''prediction'' of cloud variables, through the statistical relationships.

The U.S. National Weather Service has had a long history of the use of MOS in weather forecasting at specific sites. Glahn and Lowry (1972) applied the MOS technique to the prediction of a number of surface weather conditions as well as CC. They developed and applied the MOS equations for individual locations using extensive samples of NWP variables and observed variables. They concluded that the MOS predictions can provide useful forecast guidance. Carter and Glahn (1976) used predictions from NWP models and surface observations of CC to develop and apply MOS equations to forecast CC at more the 200 U.S. stations. Verification scores indicated that the forecast method demonstrated skill comparable to subjective forecasts.

Several investigators have attempted to use the MOS technique for CC prediction in global-scale contexts. Trapnell (1992) applied the procedure of Mitchell and Hahn (1989) to the Air Force Global Weather Central (AFGWC) global spectral NWP model forecasts (Stobie 1986) and RTNEPH cloud analyses. The method relied primarily on statistical relationships between model-predicted relative humidity and the cloud analyses. Cianciolo (1993) applied a MOS approach to a 1-yr sample of AFGWC global spectral NWP model forecasts and RTNEPH cloud analyses. Total CC was the sole predictand, but a large number of predicted and derived variables were used as predictors. Norquist et al. (1994) used a similar methodology with forecasts from the Phillips Laboratory global spectral model (Norquist et al. 1992; Norquist and Chang 1994) and RTNEPH cloud analyses. In this case, relationships between CC and predictors were developed and applied for total cloud and for clouds in three discrete vertical segments of the troposphere, called ''decks.'' Finally, Nehrkorn and Zivkovic (1996) compared several methods of diagnosing CC using the Phillips Laboratory model and RTNEPH analyses. A consistent finding among these studies was that the schemes with the lowest root-mean-square errors (rmse's) had unrealistic frequency distributions, while those that sought to retain the proper frequency distribution of CC had higher rmse's.

Prognostic cloud prediction formulations within NWP models have explicit predictive equations that are solved numerically to estimate the future cloud state. They attempt to account for what are considered the more important atmospheric processes involved in cloud development, maintenance, and decay. Tiedtke (1993) described a prognostic cloud scheme that has been incorporated into the European Centre for Medium-Range Weather Forecasts's (ECMWF) operational global NWP model. The scheme uses a prognostic equation for both cloud condensate and cloud cover. Jakob (1999) reported that the ECMWF prognostic cloud scheme improved the model's cloud forecasts and its climatology.

A possible disadvantage of the use of a prognostic scheme is the requirement for initial conditions to be

specified for each prognostic variable. Research and development efforts in the area of specifying initial conditions for a prognostic cloud prediction are not fully mature. However, it is not known at this time how important the initial specification of cloud conditions is in the quality of the consequent cloud prediction using a prognostic method.

Norquist et al. (1997) attempted to develop a diagnostic technique that accurately predicted large-scale CC while preserving frequency distribution and producing realistic spatial gradients. Using Phillips Laboratory model forecasts and RTNEPH CC analyses in the Northern Hemisphere, they experimented with several statistical methods that relate the forecast predictors to the CC. Multiple linear regression (MLR) repeated the results of earlier studies—low rmse but poor reproduction of the frequency distribution. Gradients were unrealistically flat as MLR produced too many partly cloudy grid points and not enough clear or overcast ones. Regression estimation of event probabilities (REEP) was used in conjunction with MLR in an attempt to force production of a realistic number of clear and overcast grid points. Separate REEP regressions were performed on clear versus not clear categories on one hand, and nonovercast versus overcast on the other. MLR was used for all points judged by REEP to be not clear or not overcast. While the clear and overcast categories retained their contributions to the CC frequency distributions, it was at the expense of neighboring categories. Finally, the CC domain (0%–100%) was divided into six discrete categories, and multiple discriminant analysis (MDA) was applied to the predictor and CC data. Coefficients of the discriminant functions produced by the MDA algorithm were applied to subsequent forecast predictor data. This generated a probability for each of the six CC categories as each grid point. Experimentation with category selection methods designed to preserve CC frequency distribution demonstrated that the combination of MDA and one of the methods achieved the project goals. Descriptions for MLR, REEP, and MDA can be found in standard statistical forecast references, for example, Glahn et al. (1991).

The purpose of this study was to assess the performance of the MDA-based diagnostic CC prediction technique. The algorithm was designed to diagnose the following cloud variables: total CC, ceiling altitude, and cloud layer CC, type, base altitude, and top altitude for up to five layers. These are the cloud variables required for air force mission support. Norquist (1999) described a similar algorithm designed to be applied to theater-scale NWP models. In this article only the validation of CC from the large-scale algorithm is presented, in comparison to ADVCLD and persistence CC forecasts. Beyond space limitations in this article, it was felt that the accuracy of the specification of the other cloud layer characteristics in the RTNEPH (Hamill et al. 1992) did not justify the validation of their prediction using RTNEPH as a standard.

An important issue in air force cloud forecast requirements is the need for medium-range cloud predictions. While a strength of the trajectory forecast method is the direct initialization from the cloud analyses, a disadvantage is the rather rapid degradation of skill with forecast time. Medium-range (3–5 days) prediction of noncloud meteorological state variables by global NWP models has experienced significant improvement over the past few decades. The diagnostic cloud forecast approach as applied to a medium-range global NWP model seeks to exploit this success in producing cloud forecasts.

In the balance of this article, the methodology of the large-scale diagnostic cloud prediction algorithm is discussed first. Next, the testing of this algorithm on the 30-day dataset is described. Then results of the verification of the CC forecasts from the diagnostic algorithm, ADVCLD, and persistence are presented. The article concludes with a discussion of the results and their implications for operational large-scale cloud prediction.

2. Method

The following methodology description is intentionally general with respect to the characteristics of the sources of data used in the diagnostic cloud forecast algorithm. This approach emphasizes the adaptability of the algorithm to any suitable data sources. The details of the specific datasets used in this study are given in section 3.

a. Data preparation

The first stage of the diagnostic cloud prediction algorithm is the formulation of the predictors from the NWP model forecasts. Predictors were chosen based on their availability from the NWP model forecast source being used (either directly or derived) and the likelihood of their relationship to clouds. For the latter consideration, the types of predictors chosen can be placed in four categories: humidity variables, water vapor or degree of saturation; dynamic variables, influencing the distribution of atmospheric water; turbulence variables, in recognition of the high degree of spatial variability of clouds not resolved by the model; and geographic variables, model-independent physical variables that may influence cloud distribution. A list of the ''pool'' of possible predictors computed for and available to the diagnostic cloud prediction algorithm is given in Table 1. Note that certain hydrological predictors (e.g., rainfall and evaporation rate) considered in earlier studies (Norquist et al. 1997) were not included in this study. Their exclusion was due to their absence in the global NWP model data available to Air Force Weather Agency at the time of this study.

It is next necessary to represent both forecast pre-

TABLE 1. Predictors included in hemispheric cloud diagnosis algorithm for potential use in developing statistical relationships with any cloud variable. Predictor values at forecast time 6 h before development forecast time are denoted ''*t*-6,'' at development forecast times ''*t*-0.'' Data were not available in this study for those predictors marked by an asterisk. (θ = potential temperature, θ_e.)

Nbr.	Potential Predictors
1	Vorticity, predictand deck avg, *t*-6
2	Divergence, predictand deck avg, *t*-6
3	Temperature, predictand deck avg, *t*-6
4	Precipitable water, predictand deck avg, *t*-6
5	Relative humidity (RH), predictand deck avg, *t*-6
6	Vertical velocity, predictand deck avg, *t*-6
7	$d(\theta)/d(z)$, predictand deck avg, *t*-6
8	Wind speed, predictand deck avg, *t*-6
9	Wind shear, predictand deck avg, *t*-6
10	Vorticity advection, predictand deck avg, *t*-6
11	Temperature advection, predictand deck avg, *t*-6
12	RH advection, predictand deck avg, *t*-6
13	Condens. press. deficit, predictand deck avg, *t*-6
14	$d(\theta_e/d(z)$, predictand deck avg, *t*-6
15	West wind component, predictand deck avg, *t*-6
16	South wind component, predictand deck avg, *t*-6
17	Maximum RH within predictand deck, *t*-6
18	RH at layer above maximum RH (see 17), *t*-6
19	Temperature at maximum RH (see 17), *t*-6
20	$d(\theta)/d(z)$ at maximum RH (see 17), *t*-6
21	Mean sea level pressure, *t*-6
22	6-h surface precipitation rate, *t*-6*
23	Latent heat flux, *t*-6*
24	Sensible heat flux, *t*-6*
25	Surface-layer wind speed, *t*-6
26	Departure from 500-hPa zonal mean height, *t*-6
27	Vorticity, high deck avg, *t*-0
28	Vorticity, middle deck avg, *t*-0
29	Vorticity, low deck avg, *t*-0
30	Divergence, high deck avg, *t*-0
31	Divergence, middle deck avg, *t*-0
32	Divergence, low deck avg, *t*-0
33	Relative humidity, high deck avg, *t*-0
34	Relative humidity, middle deck avg, *t*-0
35	Relative humidity, low deck avg, *t*-0
36	Vertical velocity, high deck avg, *t*-0
37	Vertical velocity, middle deck avg, *t*-0
38	Vertical velocity, low deck avg, *t*-0
39	$d(\theta)/d(z)$, high deck avg, *t*-0
40	$d(\theta)/d(z)$, middle deck avg, *t*-0
41	$d(\theta)/d(z)$, middle deck avg, *t*-0
42	Wind speed, high deck avg, *t*-0
43	Wind speed, middle deck avg, *t*-0
44	Wind speed, low deck avg, *t*-0
45	Wind shear, high deck avg, *t*-0
46	Wind shear, middle deck avg, *t*-0
47	Wind shear, low deck avg, *t*-0
48	Maximum RH within high deck , *t*-0
49	Maximum RH within high deck , *t*-0
50	Maximum RH within low deck, *t*-0
51	Temperature, predictand deck avg, *t*-0
52	Precipitable water, predictand deck avg, *t*-0
53	Vorticity advection, predictand deck avg, *t*-0
54	Temperature advection, predictand deck avg, *t*-0
55	RH advection, predictand deck avg, *t*-0
56	Condens. press. deficit, predictand deck avg, *t*-0
57	$d(\theta_e)/d(z)$, predictand deck avg, *t*-0
58	West wind component, predictand deck avg, *t*-0
59	South wind component, predictand deck avg, *t*-0
60	RH for level above RH maximum, predictand deck, *t*-0
61	Temperature at maximum RH level, *t*-0
62	$d(\theta)/d(z)$ at maximum RH level, *t*-0
63	Smoothed vorticity, predictand deck avg, *t*-0
64	Smoothed vert. velocity, predictand deck avg, *t*-0
65	Mean sea level pressure, *t*-0
66	6-h surface precipitation rate, *t*-0*
67	Latent heat flux, *t*-0*
68	Sensible heat flux, *t*-0*
69	Surface-layer wind speed, *t*-0
70	Departure from 500-hPa zonal mean height, *t*-0
71	Maximum RH squared within predictand deck, *t*-0
72	Maximum RH fourth within predictand deck, *t*-0
73	RH (ice) at RH maximum, predictand deck, *t*-0
74	Lifted-cond.-dist. at RH maximum, pred. deck, *t*-0
75	Ln (Ri no.), pred. deck (maximum RH deck for total), *t*-0
76	Sine of latitude
77	Cosine of latitude
78	Sine of longitude
79	Cosine of longitude
80	Solar zenith angle, *t*-0
81	Cosine of solar zenith angle, *t*-0
82	Hours of sunshine before *t*-0
83	Hours of darkness before *t*-0
84	Surface terrain height (nine-point avg, 1/8 mesh data)
85	Standard deviation of surface terrain height
86	Percent of surface that is water
87	$3 \times 3 \times 3$ minimum of Ln (Ri no.), *t*-0
88	$3 \times 3 \times 3$ minimum of $d(\theta)/d(z)$, deck avg, *t*-0
89	$3 \times 3 \times 3$ maximum of vertical shear, deck avg, *t*-0
90	$3 \times 3 \times 3$ maximum of wind speed, deck avg, *t*-0
91	3×3 maximum of 6-h surface precip. rate, *t*-0*
92	3×3 maximum of surface layer wind speed, *t*-0
93	3×3 maximum of surface speed *x* terrain-var., *t*-0
94	Surface wind times terrain gradient, avg of *t*-6, *t*-0
95	Minimum of terrain var., wind/stability height, *t*-0
96	RH squared, predictand deck avg, *t*-0
97	RH fourth, predictand deck avg, *t*-0
98	RH squared, predictand deck avg, *t*-6
99	RH fourth, predictand deck avg, *t*-6
100	Diagnosed total cloud cover, *t*-0 (decks only)
101	Diagnosed total no. of cloud layers, *t*-0 (decks only)
102	Predictand, RTNEPH cloud variable

dictors and cloud analysis predictands on a mutual horizontal grid to develop statistical relationships. A number of considerations may influence the choice of such a grid (hereafter called the diagnosis grid). The spatial resolution of the source predictor and predictand data dictates the diagnosis grid's resolution. The constraint is the coarser of the two source grids. Cloud cover data should be interpreted differently from predictor gridpoint data. Each CC grid value actually represents the fractional cover over the grid area. Predictor gridpoint data are considered discrete values of a continuous meteorological field. Thus it is appropriate to interpolate (if necessary) the predictor data to the cloud predictand grid, but not vice versa. However, if the CC grid has a finer spatial resolution than the predictor grid, it must first be regridded to a resolution compatible with the coarser predictor grid. Because the CC grid values are a spatial average of smaller cloud–no cloud subareas, the regridding may be done by averaging all original CC grid values that cover each coarser-resolution grid

TABLE 2. Cloud variable category values. Category indices are used as predictands in the MDA diagnostic scheme. Diagnosed category values are converted to variable values using the mean of the range indicated for each category when verifying variable values. RTNEPH and ADVCLD values are converted to category indices for category verification. Cloud cover is rounded to the nearest 5% before conversion.

Variable	1	2	3	4	5	6
Cover (%)	0	05–20	25–40	45–60	65–80	85–100
Type						
low	Cb	St	Sc	Cu		
middle	As	Ns	Ac			
high	Cs	Cc	Ci			
Base altitude (m AGL)						
low	0–396	397–792	793–1188	1189–1584	1585–1980	
middle	1980–2590	2591–3199	3200–3809	3810–4418	4419–5028	
high	5028–6022	6023–7017	7018–8011	8012–9006	9007–10 000	
Thickness (m)						
low	0–1600	1601–3200	3201–4800	4801–6400	6401–8000	
middle	0–800	801–1600	1601–2400	2401–3200	3201–4000	
high	0–800	801–1600	1601–2400	2401–3200	3201–4000	
No. of layers	0	1	2	3	4	
Ceiling altitude (m)	0–2160	2161–4320	4321–6480	6481–8640	8641–10 800	

element. These are the principles that guided the construction of the diagnosis grid in this study.

Vertical compatibility of the predictors and predictands requires similar consideration. Predictor data represents the discrete values of the variation of each variable in the vertical. Cloud layers are distributed in the vertical as represented by the predictand values of base and top altitude for each reported layer. Since only a single value of cloud cover is available in each layer, one must decide how best to use the vertically discretized gridpoint predictors to represent the predictor variable value in each layer. Because cloud analyses are largely based on satellite observations, the accuracy of the reported cloud-layer base and top in the cloud analyses are subject to significant uncertainty. While cloud analysis algorithms now being developed promise a reduction of this uncertainty, current cloud analyses do no justify ''pinpointing'' the exact predictor grid points that fall within the altitude range of the reported cloud-layer base and top.

In this study, the total vertical domain within which the source cloud analysis dataset reported cloud-base altitudes was divided into three fixed subdomains, or decks (high, middle, low). The interfaces of the decks were set by the altitude ranges of the cloud types corresponding to the base altitudes of the reported cloud layers. A single value of each predictor was assigned to each deck by a weighted average (by forecast model layer depth) of the predictor grid points that fell within each deck. The cloud predictand values of CC, type, base altitude, and thickness from the reported layer having the greatest CC whose base fell within a deck were assigned as that deck's ''primary'' predictands. Another report layer's (or if none, zero) predictand values were considered the ''secondary'' layer within the deck. In this study, the high deck was assigned only a primary layer.

While it is likely that cloud-top altitudes are more accurate in a satellite-based cloud analysis, cloud-base altitudes were used to place the reported cloud layers in the decks. This is because the cloud type is more commonly associated with cloud-base altitude, and the deck domains were set by the altitude range of the cloud types. The uncertainty of cloud-base altitude is compensated for by the large vertical extent of the decks.

Because the statistical method used (MDA) requires categorical predictands, total CC, ceiling altitude, and layer cloud CC, base altitude and thickness were converted into category values (see Table 2). Categories chosen for CC were the result of a study to find the best compromise between small category range and fewest number of categories (Norquist et al. 1997). For base altitude and thickness, a simple choice of five equal-sized categories was used. The categorical predictands for each grid point of the diagnosis grid were then combined with the respective deck or total cloud predictor vector that had a forecast valid time coincident with that cloud analysis time. Total CC, total number of cloud layers, and ceiling altitude were the predictands combined with the total cloud predictor vectors. Layer CC, type, base altitude, and cloud thickness (for up to two layers for low and middle decks) were the predictands combined with the deck cloud predictor vectors. The vectors for all grid points on the diagnostic grid were grouped by the NWP model forecast duration valid at a particular cloud analysis time.

b. MDA process

An *N*-day (where *N* is set by the user) period of these predictor–predictand vector sets was then presented to the MDA statistical development module of the algorithm. However, only a subset of about one-fourth of the available vectors (representing nonmissing points on

the diagnostic grid) selected randomly were used as input to the statistical development process. This data thinning reduced the computational burden of the MDA processing while still ensuring representation of the entire hemisphere. The MDA process was invoked separately for each predictand, for each forecast duration, for each deck and total cloud, and, in the case of middle and low decks, for each of the two layers within the decks. In addition, the MDA process was executed separately for three latitude bands: 0°–30°, 30°–60°, and 60°–90°. Because of the operational setting in which the algorithm must perform, it was designed to handle occasional missing NWP forecasts or RTNEPH cloud analyses in the *N*-day development period. In this case, the sample of predictor–predictand vectors is proportionately smaller.

In MDA, the predictor values are used to discriminate between categories of the predictand. Predictor values are grouped according to the predictand category, and the mean and variance of the predictors in each group is determined. MDA performs best when the difference between group means is greater and the in-group variance is smaller. Thus, an appropriate predictor selection method for MDA would involve a screening of predictors on the basis of their collectively maximizing the ratio of these two quantities. MDA generates coefficients of a discriminant function for each category. When the coefficients are applied to the predictor values, a probability for each category is obtained. The actual category value assigned to the diagnosis must then be chosen on the basis of the diagnosed probabilities.

The diagnostic algorithm used the MDA routine DSCRM provided in the IMSL Fortran Numerical Libraries version 3.010, a product of Visual Numerics, Inc. (Visual Numerics, Inc. 1997) DSCRM does not have the facility to select the most discriminating predictors from among a longer list of available predictors. Predictors selected in a forward stepwise selection process (IMSL routine DRSTEP), where the cloud variable predictands were used in their original form, were used as input to the MDA method. This identified the leading predictors from the larger pool of predictors as shown in Table 1. For CC the leading 16 predictors were identified in this way, while for the other predictands only the leading 10 predictors were selected for MDA use. Then the reduced vectors of predictors and their respective categorical predictand values were subjected to the MDA algorithm.

The MDA process commonly could not use all of the predictors submitted to it when processing the data in an *N*-day development period for a given predictand. In this case, the DSCRM algorithm would fail because of a singularity found in calculating the determinant of the predictor matrix. When this happened, the number of predictors considered was reduced by one, and DSCRM was called iteratively with one less predictor, eliminating successively more highly correlated predictors one at a time as ordered by the forward stepwise selection process. If the matrix still could not be solved, the least highly correlated predictor was removed and the iterative process was repeated. If a solution could not be found with at least three predictors, the attempt was terminated and no MDA coefficients were obtained. This was a rare event, and appeared to occur only for predictands that were less likely to be physically related to the weather variables from the NWP model. It was not observed to occur at all when processing cloud cover.

Separate MDA coefficients were generated for each forecast duration because of the tendency of model forecast systematic error to grow in magnitude with increasing forecast duration. The values of the coefficients of the predictors change significantly to compensate for the growing model forecast error. However, the ensemble of cloud patterns present in the hemisphere changes only gradually with time, so the cloud scenes corresponding to the various forecast times are very similar.

MDA coefficients sets were developed separately for the three latitude bands because of the expected difference in the character of cloud type and CC distribution in the Tropics, midlatitudes, and high latitudes. The number of latitude bands to use is flexible in the diagnostic cloud algorithm. Using three bands gave slightly better CC diagnosis results that using just one for the hemisphere. After the cloud variables were diagnosed, they were blended over the 10° zone centered on each latitude band boundary. The diagnostic relationships from both neighboring bands were applied to all points in the 10° zone. The blending then weighted the two diagnosed values on either side of the boundary by the sine of the latitude departure from the boundary.

c. Cloud cover diagnosis

Once the discriminant function coefficients were obtained for all predictands in all decks and total cloud, for all forecast durations, they were applied to subsequent NWP model forecasts. Model forecasts from initial conditions on the day after the completion of *N*-day development period were processed in the same way as was done to get the development predictors. The discriminant function coefficients were applied to each model forecast vector of MDA-utilized predictor values on the cloud diagnosis grid for the respective cloud deck or total cloud. In middle and low cloud decks, the distinct coefficients for each of the primary and secondary cloud layers were applied to the deck's forecast predictor vectors separately. In this way, the probability of each category of the predictand is diagnosed at each cloud diagnosis grid point.

It was then necessary to select the predictand category to be assigned to each grid point based on these category probabilities. Upon experimentation with a number of methods to do this, a method developed for the diagnostic algorithm called the iterative maximum probability method was found to work best (Norquist et al.

1997). This method was designed to produce the best combination of cloud variable diagnosis accuracy and preservation of the predictand's category frequency distribution over the domain of diagnosis.

It is the design of the category selection method, not the diagnosis of the probabilities for each category, that can ensure preservation of the category frequency distribution. A major reason for choosing MDA as the statistical method was that it allows the application of a separate category selection technique. One can use the information given by the MDA-generated category probabilities to select the predictand category at each grid point in such a way as to ensure frequency distribution realism over the entire grid.

In the iterative maximum probability method, we first compute the frequency of occurrence of the predictand over the *N*-day development period for each predictand category. Next, the frequency of occurrence proportion of the number of domain grid points in each category is computed (for each category, NTOT = frequency of occurrence times the number of domain points). Then the categories for each grid point are ordered from highest probability to lowest probability. On the first pass through the grid points, we identify the highest probability category for each grid point, compute the difference between that probability and the next highest probability, sort these probability differences over all grid points that have the same highest probability category, and select for that category the NTOT grid points that have the largest differences. Once this is done for all categories, we take a second pass through the unassigned grid points, and considering only those categories that are not full (i.e., do not have NTOT grid points assigned to them), identify the grid points that have an unfilled category as their second highest probability. Within these unfilled categories, we compute the difference between that probability and the next largest category's probability. We sort these differences, and select for that category the up to NTOT grid points with the highest differences. We repeat this process until at least all but one category is filled, then assign any unassigned grid points to the remaining unfilled category.

This method has the property of imposing the frequency distribution of the predictand categories from the *N*-day development period upon the application diagnoses of the predictand. We found in our earlier study (Norquist et al. 1997) that this dramatically improves the match between the frequency distribution of the diagnosed predictand and the frequency distribution of the verification of the diagnoses, when compared to simply selecting the category with the highest diagnosed probability. This comes at the cost of only slightly higher rmse for the iterative maximum probability selection method. Furthermore, forcing the diagnosed predictands to have the development period's frequency distribution does not risk creating a hemispheric-wide bias error in the diagnoses, because over the hemisphere there is little day-to-day difference in the CC distribution. Thus the CC category frequency distribution over the *N*-day period is going to be similar to the distribution on any given subsequent day. This is in marked contrast to the application of the method in the theater-scale study (Norquist 1999) in which the frequency distribution of CC in a theater (1000–2000 km on a side) can vary significantly over several days time.

Cloud decks had two extra predictors to choose from that were actually derived from the total CC and number of cloud-layer predictands. The last two predictors listed in Table 1 were computed by diagnosing total cloud cover and number of cloud layers from the development predictor data before employing the MDA technique on the deck predictors.

3. Data

The datasets used in this study were obtained in real time from the Air Force Weather Agency. The study was the result of a test and evaluation project conducted for the agency. Part of the purpose of the test was to determine the feasibility of executing the hemispheric cloud diagnosis algorithm in real time using agency computer assets. This was clearly feasible when just the Northern Hemisphere diagnosis was executed, despite competition for computational resources from other agency projects. There is no doubt that it would be feasible to execute both hemispheres simultaneously given the agency's present and near-future computational configuration.

Global weather forecasts produced by the NOGAPS model were used to supply the forecast predictors. The NOGAPS forecasts initialized at 0000 and 1200 UTC each day were provided in near–real time to the Air Force Weather Agency by the U.S. Navy's Fleet Numerical Meteorological and Oceanographic Center (FNMOC) in gridpoint form. The forecast fields were available at 3-h intervals to 96 h. The diagnostic algorithm can diagnose clouds at forecast intervals as finely spaced as the source forecast data. However, this study used only 6-h forecast intervals in order to diagnose clouds for 12-h intervals to 72 h. Temperature, wind components, geopotential height, and dewpoint depression were available on a $1^\circ \times 1^\circ$ latitude–longitude grid. All of these upper-air variables except dewpoint depression were available on 16 pressure levels from 1000 to 10 hPa. Dewpoint depression was available only up to 300 hPa in the forecasts. Relative humidity computed from temperature and dewpoint depression was extrapolated linearly upward from the highest two available levels (usually 400 and 300 hPa) to fill in the levels up to 10 hPa. The diagnostic cloud algorithm uses predictors only as high as about 200 hPa. Surface-level variables available in the NOGAPS forecasts provided to AFWA include mean sea level pressure and surface wind components that were used in this study. However, no surface hydrological variables were available for use at the time of the study.

Each pressure level or surface data field at each forecast time interval was spectrally transformed into triangular 119 wave truncation spectral coefficients. Divergence, absolute vorticity, geopotential height, and relative humidity were the upper-air spectral coefficient fields, while mean sea level pressure, divergence, and absolute vorticity were the surface spectral coefficient fields. The spectral coefficients were then evaluated on the $1^\circ \times 1^\circ$ latitude–longitude grid. This resulted in latitude–longitude grids of temperature, wind components, divergence, relative vorticity, geopotential height, and relative humidity on 16 pressure levels and mean sea level pressure and wind components at the surface. This ''spectral filtering'' (back to spectral form, then back to grid points) was really done for two reasons. First, it allows for a more exact calculation of horizontal spatial derivatives, such as those in the divergence and vorticity. Consequently, it allows a more precise computation of the vertical motion field through the vertical integral of the divergence. Second, the assumptions used in arriving at the $1^\circ \times 1^\circ$ data obtained from FNMOC were not known. Therefore it seemed safest to ''filter'' the fields at a spectral resolution that was known. Indeed, very good agreement was found between the fields before and after the filtering, with the latter being slightly smoother.

For compatibility with reported RTNEPH altitudes, surface pressure was diagnosed from geopotential height and temperature on the RTNEPH terrain (Schaaf et al. 1990) at the latitude–longitude grid points. Then all upper-level fields were interpolated to 16 sigma (=pressure–surface pressure) levels distributed in the vertical roughly paralleling the pressure level distribution. A weighted average of the sigma layers falling within each cloud deck was used to compute the deck-averaged predictors. Deck interfaces used in RTNEPH (USAFETAC 1991) are given in meters above ground level (5028 for middle–high deck interface, 1980 for low–middle deck interface). Total cloud predictors were computed from deck-weighted averages of the deck-weighted predictors.

The deck-average and total cloud gridpoint predictor values were bilinearly interpolated from the $1^\circ \times 1^\circ$ latitude–longitude grid to the RTNEPH polar stereographic grid. The algorithm allows the user to specify the latter grid resolution in multiples of one, two, four, and eight times the grid spacing of the source RTNEPH grid resolution. The $1^\circ \times 1^\circ$ NOGAPS forecasts provided to AFWA dictated an RTNEPH grid with a grid spacing twice the source RTNEPH grid of 48 km. The resulting 96-km polar stereographic RTNEPH grid was thus used as the cloud diagnosis grid.

The RTNEPH cloud analysis was available at 3-h intervals in each hemisphere on the 48-km grid. Only the 0000 and 1200 UTC analysis for each day was matched with the valid-time 12-hourly interval NOGAPS forecasts initialized at these same two times daily. The only processing necessary for the RTNEPH data was the averaging from the 48-km grid to the chosen cloud diagnosis grid of 96 km. The RTNEPH processing used only 48-km grid values based on observed data that were no more than 2 h old. Grid points with older data were considered ''missing'' in this study. If half or more of the 48-km points used in the spatial averaging were missing, the averaged RTNEPH point was assigned as missing. Cloud cover, base altitude, and thickness were so averaged. Cloud type assigned was the most common type among the 48-km values.

All four variables were assigned in each deck as described in section 2. RTNEPH layer cloud-base altitudes are reported with reference to sea level. The RTNEPH terrain elevation had to be subtracted from each reported layer's base altitude in order to determine the deck within which the layer fell. The resulting deck predictand values were combined with the predictor vector for the corresponding deck for each nonmissing grid point on the 96-km cloud diagnosis grid for each forecast duration.

Because of the time lag between when a NOGAPS forecast is available (shortly after its initial time) and when the coinciding valid time RTNEPH is available, a different set of dates was involved in the *N*-day development period for the various forecast durations. The last day of the *N*-day development period was always set as the initial date of the NOGAPS forecasts valid on the day before the application date. This study used a 5-day development period. So, NWP forecasts initialized from 6 to 2 days prior to the application day made up the development predictors for 12- and 24-h forecasts. NWP forecasts initialized from 7 to 3 days prior to the application day made up the development predictors for 36- and 48-h forecasts. NWP forecasts initialized from 8 to 4 days prior to the application day made up the development predictors for 60- and 72-h forecasts. The MDA coefficients derived from each of these respective development periods were applied to the NWP forecasts initialized on the application day to diagnose the cloud variable ''forecast'' on the cloud diagnosis grid.

An appreciable data storage requirement is incurred to generate the MDA coefficients each day. For example, this study required using two hemispheric RTNEPH files and twice-daily NOGAPS at six forecast times to form that days' predictand–predictor files for each cloud deck. Then 5 day $\times$ 4 cloud deck $\times$ 12 NOGAPS times = 240 predicand–predictor files are needed daily to generate the MDA coefficients. The storage requirements for two RTNEPH files, 12 NOGAPS files, and 240 predictand–predictor files was 1.2 gigabytes. This has implications for the operational implementation of the algorithm.

The test and evaluation period used in this study covered the period 21 February 1999 to 4 April 1999 for the Northern Hemisphere. A short study period of 1–15 April 1999 was conducted in the Southern Hemisphere just to test the diagnostic cloud algorithm in both hemi-

spheres. The first 5-day development period started on 21 February, which required that both NOGAPS forecast and RTNEPH analysis be available beginning on that day. Predictor–predictand vector grids were generated for each forecast duration available at each RTNEPH 0000 and 1200 UTC time over 5 days to compute the MDA coefficients. The first day that had coefficients for all forecast durations that could be applied to the current day's NOGAPS forecasts was 1 March 1999. During the first four days of March, the number of predictors allowed into the MDA process for each predictand was varied to determine an optimum number. This was the number of predictors to be selected by the forward stepwise regression that would produce the best results from MDA while minimizing the computational burden. The 0000 and 1200 UTC 5 March 1999 forecasts were the first cloud diagnoses forecasts that have been included in the verification dataset for the Northern Hemisphere test. The cloud diagnoses from the 0000 and 1200 UTC 3 April 1999 forecasts are the last included in the test, making a total of 30 days in the test period. Six of the individual twice-daily NOGAPS forecasts were not successfully obtained in the test period. Two more were only obtained out to 48 h. So, fifty-four 12–48-h cloud diagnosis forecasts and fifty-two 60–72-h cloud diagnosis forecasts were included in the verification discussed in section 4.

Also archived were the 0000 and 1200 UTC ADVCLD CC forecasts for 12–36 h for the period 5 March 1999 to 3 April 1999 and 48-h ADVCLD for 15 March to 3 April in the Northern Hemisphere. The data were already available on the 96-km grid identical to the cloud diagnosis grid, so no further processing was needed to make the grids compatible. As stated earlier, ADVCLD produces CC forecasts for total cloud and five pressure levels (1000, 850, 700, 500, 300 hPa). It was only necessary to assign these cloud layers to the cloud decks used in the diagnostic cloud algorithm in order to have compatible verification. This was done by obtaining the geopotential height of these five pressure levels from the 12–48-h NOGAPS forecasts on the 96-km grid. These were provided to the verification algorithm discussed in section 4 so that the ADVCLD layers would be placed in the appropriate cloud deck for verification.

4. Results

The results of the predictor selection process in the diagnostic clouds algorithm are presented first. As mentioned in section 2, the MDA algorithm used did not have a predictor selection capability. Rather than incur the prohibitive computational expense of using the entire pool of predictors (Table 1) in the MDA, a forward stepwise regression technique was used to select a subset. The stepwise regression finds the predictors whose linear combination is most highly correlated with the predictand (in this case, CC). Up to 16 predictors could be chosen in this way. For CC, rarely were fewer than 16 chosen in any of the three latitude bands. This indicates that there were enough predictors correlated with CC (either directly or in combination with each other) to ensure statistical relationship development.

Table 3 lists the leading predictors selected by the forward stepwise regression selection process for each of the three latitude bands for the CC statistical relationship development in each deck and total cloud. The predictors listed in the table as ''leading'' resulted from assigning points according to the order in which the predictors were selected in each development period, then adding these points over all 30 development periods. This was done separately for each latitude band and forecast duration, though only the results for the 12-h forecasts are shown in Table 3.

It is clear from Table 3 that relative humidity (and its variations) is a variable prominent in its association with CC. In the low latitudes, some function of relative humidity was considered the overall most highly selected predictor in the forward stepwise regression selection process. There is at least one occurrence of relative humidity as a top 10 selected predictor in each cloud deck–latitude band category in the table. With just one exception (low deck, middle latitudes), relative humidity is the only measure of degree of saturation included in the prominently selected predictors. The close association of relative humidity with CC found here supports the use of relative humidity as a sole predictor of CC in many diagnostic methods.

Table 3 indicates that besides relative humidity, a number of other types of variables can make important contributions to the diagnosis of CC. First, the absolute humidity variable precipitable water appears many times in the table. Another category of predictor that is highly selected are the turbulence variables, such as Richardson number, vertical potential temperature gradient (dry and moist), and wind speed and wind shear. Especially prominent in the high deck are the geographic predictors, especially percent of the surface covered by water and the sine of latitude and longitude. Of the mass variables that might be associated with cloudiness, only temperature or temperature advection appears significantly in the table. Mean sea level pressure and departure from 500-hPa zonal mean height are virtually absent. Though vertical velocity has commonly been associated with cloud diagnosis, it appears infrequently in Table 3. Finally, the diagnosed total CC or total number of cloud layers that are available as predictors to the decks are selected in the deck cloud regressions, especially the low cloud.

Before the objective verification scores of the diagnosed cloud forecasts are presented, a look at the nature of the RTNEPH CC is necessary to set the context. Figure 1 shows the categorical CC distribution for the RTNEPH averaged to the 96-km grid, using only timely (no more than 2 h old) data. Figure 1 shows that 75% and 68% of the timely grid points were cloud free in

TABLE 3. Leading cloud amount predictors for each cloud deck for the thirty 5-day development periods from 27 Feb–3 Mar to 28 Mar–1 Apr 1999, for the 12-h forecast (*t*-6 is value 6 h before 12-h forecast time). Deck name is given in description when predictor selected is not from the same deck: RH = relative humidity, $d\theta/dz$ = vertical potential temperature gradient, ∇Z_* = terrain gradient, ↑ = above.

Rank	High	Middle	Low	Total
0°–30°N				
1	RH, middle*	RH, high*	RH, high*	RH^2
2	Percent sfc. water*	Percent sfc. water	RH^2, *t*-6	RH, *t*-6
3	RH^2	RH^2, *t*-6	Sin (lat)	Sin (lat)
4	Ln (Ri)	RH^2	Temp (max RH), *t*-6	Temp
5	RH^2, *t*-6	Sin (long)	Sin (long)	RH, low
6	Sin (long)	3 × 3 × 3 min ln (Ri)	Max RH, middle	$d\theta_E/dz$, *t*-6
7	3 × 3 max. sfc. wind	Sin (lat)	RH^4, *t*-6	Sin (long)
8	Precip. water, *t*-6	Max RH^4	Temp	Max RH, *t*-6
9	RH^4, *t*-6	Max RH^2	Diag. total CC	$d\theta/dz$, high
10	Sin (lat)	RH advection, *t*-6	RH, middle	RH^4
30°–60°N				
1	RH^2*	RH, high*	Diag. total CC*	Percent sfc. water
2	Sin (long)	3 × 3 × 3 max wind shear	Percent sfc. water	RH^2
3	Smth. vert. velocity	Max RH^2	Surface terrain height	Max RH^2
4	Percent sfc. water	Temp, *t*-6	H dark before fcst.	Max RH, *t*-6
5	Precip. water, *t*-6	RH ↑ max RH level	3 × 3 max sfc. wind	3 × 3 × 3 max wind speed
6	3 × 3 × 3 min ln (Ri)	Sin (long)	Cond. press. deficit	3 × 3 × 3 min ln (Ri)
7	3 × 3 × 3 max wind shear	3 × 3 × 3 min ln (Ri)	Diag. total no. cloud layers	Max RH^4
8	RH^4, *t*-6	wind shear, high	RH^2	RH^4
9	Min (∇Z_*, wind/stab.)	Precip. water, *t*-6	Max RH, high	Surface terrain height
10	Cos (long)	Min (∇Z_*, wind/stab.)	Max RH^4	H dark before fcst.
60°–90°N				
1	Percent sfc. water*	RH, high*	Diag. total no. cloud layers	$d\theta/dz$, low
2	Precipitable water	Percent sfc. water	$d\theta/dz$	Precipitable water
3	Sin (long)	Precipitable water	3 × 3 max sfc. wind	Cos (lat)
4	Temp advection	Vert. velocity, high	Precipitable water	3 × 3 × 3 min ln (Ri)
5	Max RH^4	Cos (lat)	H dark before fcst.	$d\theta/dz$, *t*-6
6	RH^4	RH^4	MSL pressure, *t*-6	RH^2
7	3 × 3 × 3 min ln (Ri)	$d\theta/dz$, low	Precip. water, *t*-6	3 × 3 max sfc. wind
8	Diag. total no. cloud layers	Temp advection	3 × 3 × 3 min $d\theta/dz$	Surface terrain height
9	Surface terrain height	3 × 3 × 3 max wind shear	RH^4	Temp advection
10	South wind, *t*-6	South wind	Cos (zenith angle)	MSL pressure, *t*-6

* On average, in the top four predictors on a daily basis.

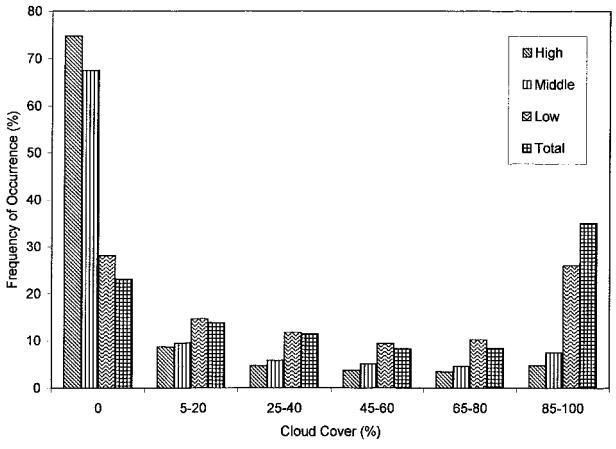

FIG. 1. Cloud cover (percent) frequency distribution for the RTNEPH averaged to the 96-km grid, using only timely (no more than 2 h old) data. Statistics shown are based on 0000 and 1200 UTC analyses during the 30-day period 5 Mar–3 Apr 1999 over the entire Northern Hemisphere.

the high and middle decks, respectively. In these two decks, a CC forecast of clear is likely to be correct. We expect that the MDA scheme will predict a clear category for a majority of the points in these decks, so we anticipate favorable prediction scores. By contrast, 72% and 77% of grid points are not cloud free in the low deck and for total cloud, respectively. In this case, most of the points have CC distributed over five categories, so the prediction is considerably more difficult. We expect poorer scores for low and total cloud than for high and middle cloud.

An objective verification of the diagnosed cloud forecasts was conducted for each 12-h interval forecast in the 30-day period 5 March–3 April 1999 over the entire Northern Hemisphere. In all cases, CC diagnosed on the 96-km grid was verified against 96-km RTNEPH CC at forecast time at all nonmissing grid points. All CC representations were rounded to the nearest 5% before verification, as RTNEPH layer CC is reported to this degree of accuracy. Graphs of the verification scores for the entire Northern Hemisphere are shown in Figs. 2 and 3. Separate verification statistics were computed for

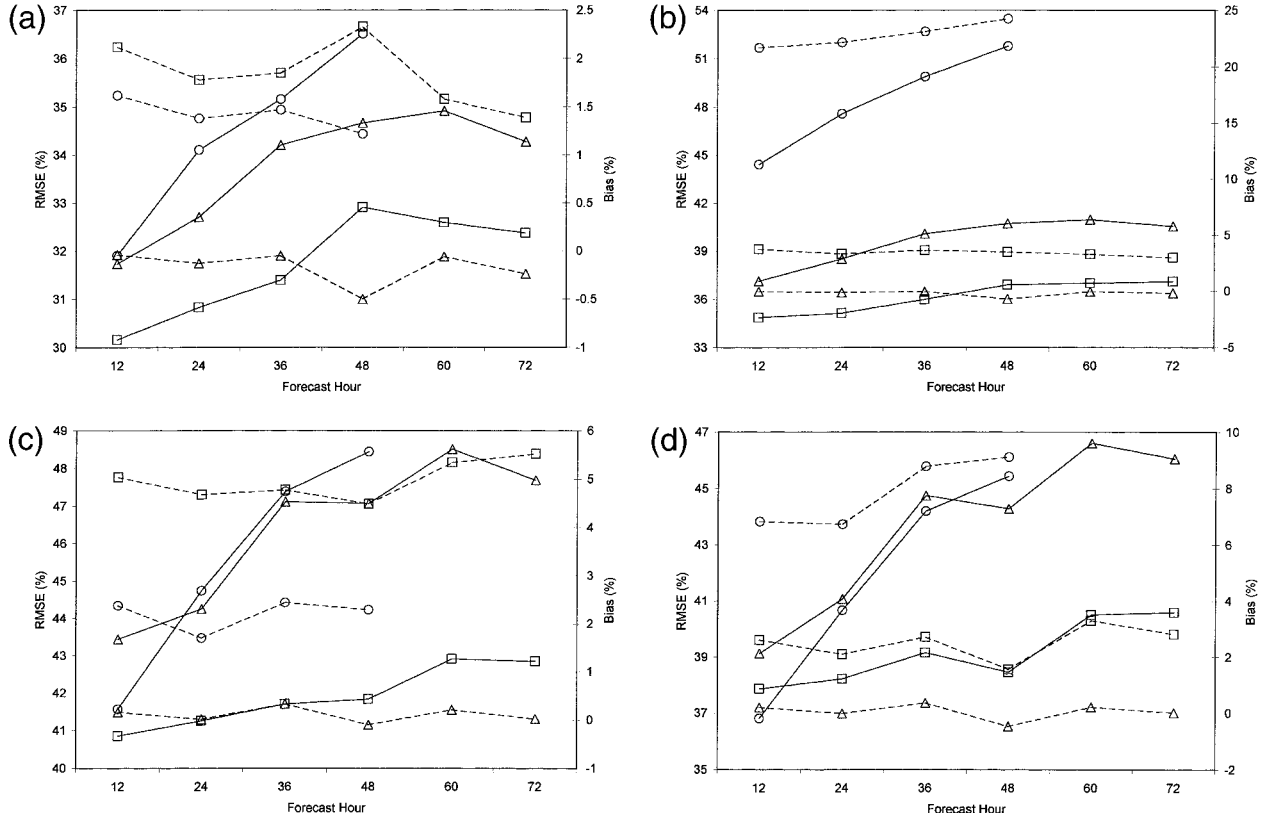

FIG. 2. Rmse (solid lines) and bias (dashed lines) in % for cloud cover forecasts from DCF (squares), ADVCLD (circles), and persistence (triangles). Verification statistics for (a) high, (b) middle, (c) low decks, and (d) total cloud are based on forecasts from 0000 and 1200 UTC initial times during the 30-day period 5 Mar–3 Apr 1999 over the entire Northern Hemisphere.

each 30° latitude band used in the development and application of the diagnosis algorithm, but are not shown.

The CC verification statistics from the diagnostic cloud forecasts (DCFs), ADVCLD trajectory model forecasts, and persistence (RTNEPH at the forecast initial time) forecasts were computed to ensure comparison on a common basis. The verifications of all three methods took place over just the grid points at which all three plus the verifying RTNEPH had nonmissing data. DCF and ADVCLD could have been verified at many more grid points (all the grid points where just the verifying RTNEPH had nonmissing data) than this more limited set. However, this approach ensured that the three forecast methods used the same basis for verification so that the resulting statistics could be compared between the methods. Over the 30-day verification period, about 35% of the hemisphere grid points were common to DCF, ADVCLD, persistence, and verifying RTNEPH.

Figure 2 shows the bias and rmse for the CC forecasts from the three CC forecast methods. It is clear from Fig. 2 that bias is small and is within acceptable limits for DCF in all three decks and total cloud. There is no discernable trend to the DCF bias with increasing forecast time. The largest values (biases of about 5%) are in the low cloud deck. In all decks, the sign of the DCF bias is positive. This can be attributed in part to the CC assignment of the DCF-produced CC categories. As mentioned, we used the category average to represent the percent CC in each category. When this was used in the verification process, it was rounded upward to the next 5% CC. Without this rounding, the small positive DCF biases would be further reduced.

The ADVCLD forecast bias also appears acceptable except in the middle cloud deck, where a positive bias of more than 20% is apparent. In the assignment of the ADVCLD data to decks, we used the 300-hPa layer to represent the high deck, the greater of the 500- and 700-hPa CC to represent the middle deck, and the greater of the 850- and 1000-hPa CC to represent the low deck. We found what appeared to be excessive CC forecast values in both the 500- and 700-hPa ADVCLD layers as compared with RTNEPH. Because the RTNEPH hemispheric cloud cover changes little from day to day, we observe trivially small biases in the persistence forecasts.

Figure 2 also shows that the DCF CC forecasts have smaller rmse than either ADVCLD or persistence forecasts for virtually all cloud decks and forecast times. The only exception is the 12-h forecast in total CC, where ADVCLD is slightly better than DCF. Another

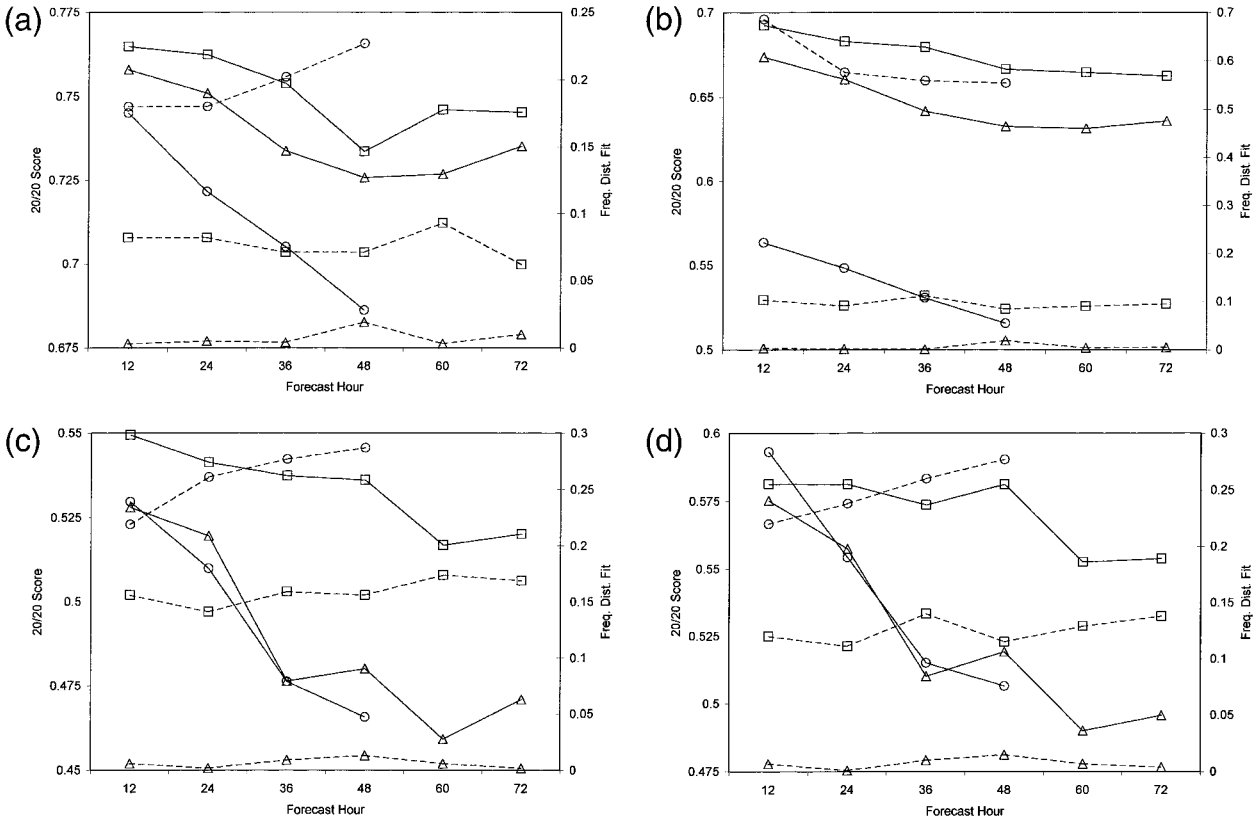

FIG. 3. Same as Fig. 2 but for 20/20 score (solid lines) and frequency distribution fit (dashed lines).

observation from the graphs is that rmse does not increase as greatly with forecast duration for DCF as it does for the other two forecast methods. Especially in low and total cloud, the difference in rmse is greater at greater forecast times. Improvement in forecast skill over persistence varies from 5% to 10% of the DCF rmse in the high and middle cloud decks to 10%–15% in the low deck and total cloud. The excessive rmse scores for ADVCLD in the middle deck are due in large part to the large positive biases found there.

It is also of interest to examine the skill of the forecast methods with regard to the approximate correctness of the prediction. In particular, we may wish to know if the forecast is within a certain percentage CC of being correct. This avoids the heavy penalty imposed by rmse if, for example, the position of a mass of clouds is not located quite correctly in the forecast. The score we use for approximate correctness of the forecast is 20/20 score, which is the fraction of the grid points predicted that have CC within 20% of being correct. For the diagnosed CC categories, this represents a prediction within one category of being correct.

The 20/20 scores for the cloud decks and total cloud are shown in Fig. 3. The scores shown indicate that, as in rmse, the DCF forecasts are an improvement upon persistence and ADVCLD at virtually every forecast hour and cloud deck. However, the improvement over persistence is not as great for 20/20 score as for rmse, ranging from 2% of DCF values in high and middle decks to 11% in the low deck. As with rmse, the only exception is the better ADVCLD total cloud forecast at 12 h. The 20/20 scores for persistence remain at about the same interval below DCF for all forecast hours in the high and middle deck, while dropping in time with respect to DCF for low and total clouds. High and middle deck scores for DCF and persistence seem to cease decreasing and even rise a bit at later forecast times. By contrast, 20/20 scores for ADVCLD decrease monotonically for the 48-h duration of the forecast.

How well the CC diagnoses fit the frequency distribution of the RTNEPH is important. In order to ensure a realistic visual depiction of CC, the proper proportion of each CC category must be present. The frequency distribution fit (FDF) metric used in the validation of the cloud predictions is given by

$$\mathrm{FDF} = (1/N) \sum_{g=1}^{G} |(m_g/n_g) - 1| n_g,$$

where g is the category index of G categories, m_g is the number of grid points diagnosed to be in category g, n_g is the number of grid points observed in that category (the verification points), and N is the total number of verification points. An FDF value of zero represents the

perfect representation of the verification's categorical frequency distribution.

FDF scores for the cloud decks and total cloud are shown in Fig. 3. It is immediately clear that persistence forecasts retain the frequency distribution of cloud cover out to at least 3 days. This confirms the earlier statement that the hemispheric cloud distribution changes little over several days and suggests that the hemisphere's cloud cover remains fairly constant for such time periods. The FDF values shown for DCF are no greater than 0.1 for high and middle clouds, less than 0.2 for low clouds, and less than 0.15 for total clouds. The corresponding ADVCLD scores are 0.2 and above for high, 0.6 and above for middle, and 0.2 and above for low and total clouds. As with the other verification scores, DCF FDF values depict little change with increasing forecast duration, while ADVCLD values show degradation with time.

A further explanation of what is happening to the distribution can be obtained by considering the actual CC category frequency distribution. Total CC distribution was compared for DCF, ADVCLD, and verifying RTNEPH at 12- and 48-h forecast times. This comparison showed that DCF underpredicted the number of clear points (12% clear compared to 20% clear for RTNEPH), but assigned the remaining points primarily to the cloud cover categories just greater than clear. By contrast, ADVCLD similarly underpredicted the occurrence of clear grid points (just 10%–11% clear), but most of the remaining grid points were placed in the overcast category. For ADVCLD, the partly cloudy categories matched fairly well with RTNEPH frequency distribution. The problem of ADVCLD overpredicting the number of overcast points worsened with forecast duration, growing from 43% overcast at 12 h to 53% overcast at 48 h (compared with about 38% for RTNEPH). At the same time, the number of DCF-diagnosed overcast points changed little (40%–42%). Because FDF weights each category's contribution by the percentage of verification grid points in each category, the overcast category receives about twice as much weight in the FDF score as does clear. Therefore, in the situation described here ADVCLD receives worse FDF scores than does DCF.

To examine the day-to-day variation in the verification statistics, DCF and ADVCLD forecasts were verified separately for each of 8 days at 4-day intervals in the 30-day verification period. Allowing for the fact that the single-day verifications were based on a much smaller sample that the 30-day period statistics, the single-day statistics were quite representative of those of the entire period. The statistical values changed somewhat from day to day, DCF being slightly worse on some days and slightly better on others. The relative order of DCF, PER, and ADVCLD was maintained in all four statistical measures for virtually all single-days verified.

Cloud predictions become useful to most military users when they are displayed in a user-friendly visual depiction. These visualizations allow mission planners to quickly spot areas to avoid or direct aircraft through, depending on the mission. Because of space considerations, only depictions of total cloud cover are presented here. These are shown for two individual cases during the 30-day verification period in Figs. 4 and 5. These two cases were among the single days that were verified separately, so their representativeness of the entire 30-day period is confirmed.

Figure 4 shows the depiction of the DCF and ADVCLD 24-h total cloud cover forecasts valid 0000 UTC 11 March 1999, along with the verifying RTNEPH analysis. In the latter, portions of the hemisphere that appear black are the missing (more than 2 h old) grid points. The first impression one gets from the images is that the DCF forecast has less spatial CC variation than is depicted in ADVCLD and RTNEPH. The normalized (with respect to RTNEPH) average spatial standard deviation was computed over the 30-day verification period by computing the standard deviation of CC in each 9×9 grid point array (of nonmissing points), averaging these over the hemisphere, and dividing by the same computed from the verifying RTNEPH. Normalized standard deviation values of approximately 0.6 were obtained for DCF, compared with about 0.8 for ADVCLD and 1.0 for persistence. In Figs. 4 and 5 we see that the DCF depiction tends to be patches of a single CC category, compared with similar areas in ADVCLD and RTNEPH that can have several categories included.

The major features that are visible in the RTNEPH image in Fig. 4 are fairly well reproduced by the forecasts. Cloud cover voids over the northern Atlantic Ocean and northern Africa are common to DCF and ADVCLD, but unfortunately cannot be clearly validated against the missing RTNEPH data. On the other hand, areas of overcast CC in east equatorial Africa, south of India, and off the west coast of North America are well represented by both DCF and ADVCLD. The latter does a better job of predicting the overcast in the Bay of Bengal, while DCF better produces the overcast in the southeastern United States and Gulf of Mexico. ADVCLD generally captures the clear areas over the Saudi Arabian peninsula while DCF renders these as partly cloudy. ADVCLD is too heavy handed with the overcast over southwest Asia (north of the Arabian Sea) while DCF get the CC more correct there. Both forecast methods produce too much cloudiness west of Central America, but get the relatively clear area over the Caribbean Sea correct.

Figure 5 shows the depiction of the DCF and ADVCLD 48-h total cloud cover forecasts valid 1200 UTC 20 March 1999, along with the verifying RTNEPH analysis. DCF does somewhat better than ADVCLD in predicting the generally cloud-free areas over western Canada, north of the West Indies, and west of the Red Sea. The band of lesser cloud cover extending west from Central America in the RTNEPH analysis is better represented by the DCF prediction. Partly cloudy skies are predominant in southwest Asia, India, and southeast

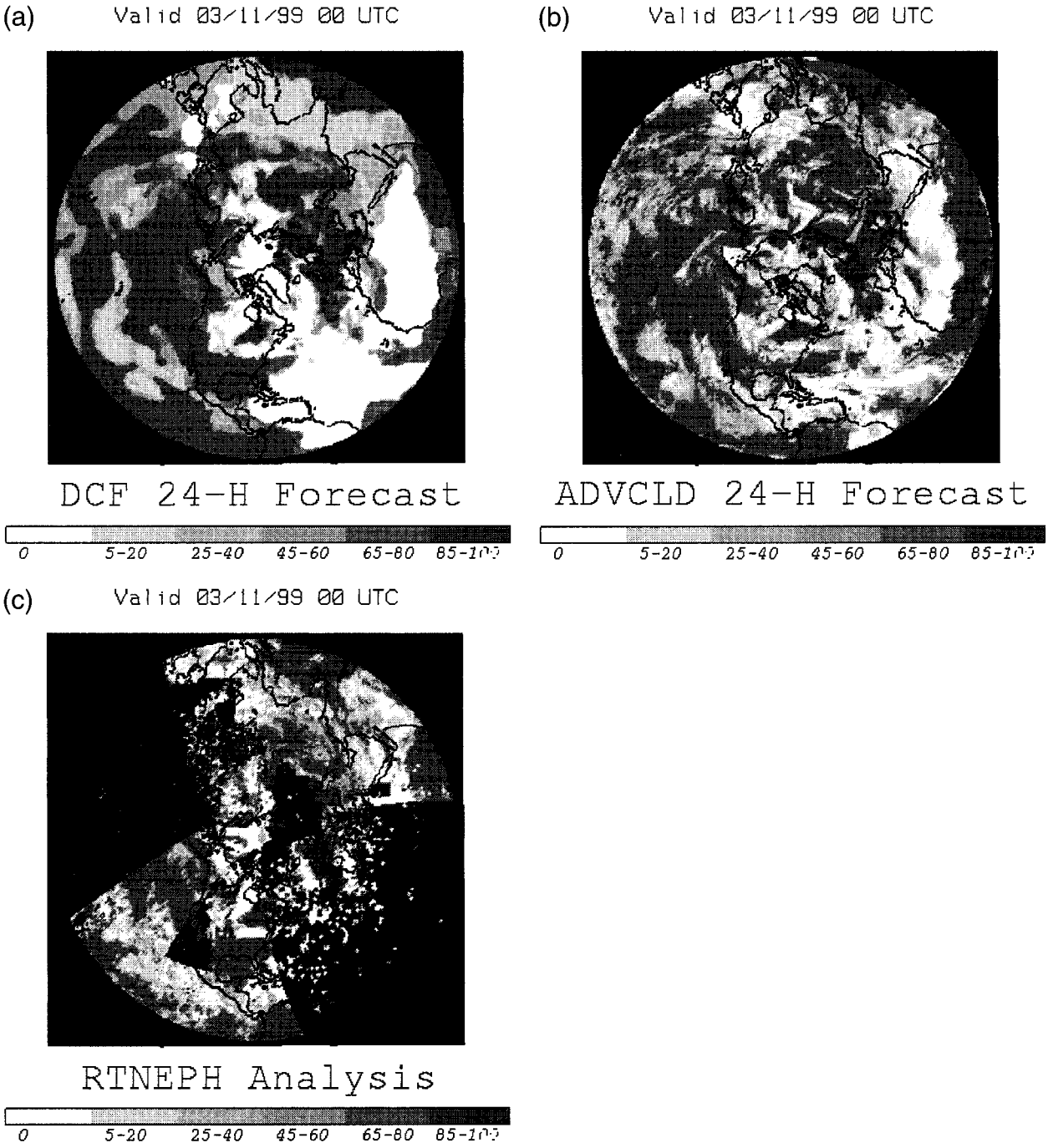

FIG. 4. Total cloud cover (%) for 11 Mar 1999 from (a) 24-h DCF forecast, (b) 24-h ADVCLD forecast, and (c) RTNEPH analysis. Black areas in RTNEPH depiction indicate missing (more than 2 h old) data.

Asia in the RTNEPH. These are depicted well in the DCF forecast, but are mixed with large patches of overcast in the ADVCLD forecast. Both forecasts capture the heavy clouds in the Tropics extending east from south of the Hawaiian Islands and over most of northern South America, though DCF reduces the CC a bit too much over the northeastern South American coast. Both forecast methods correctly predict the large overcast area off the west coast of North America. The overcast over the south-central United States and northern Mexico is better depicted by DCF.

5. Discussion

The diagnostic cloud forecast algorithm described and demonstrated in this study was developed to meet the

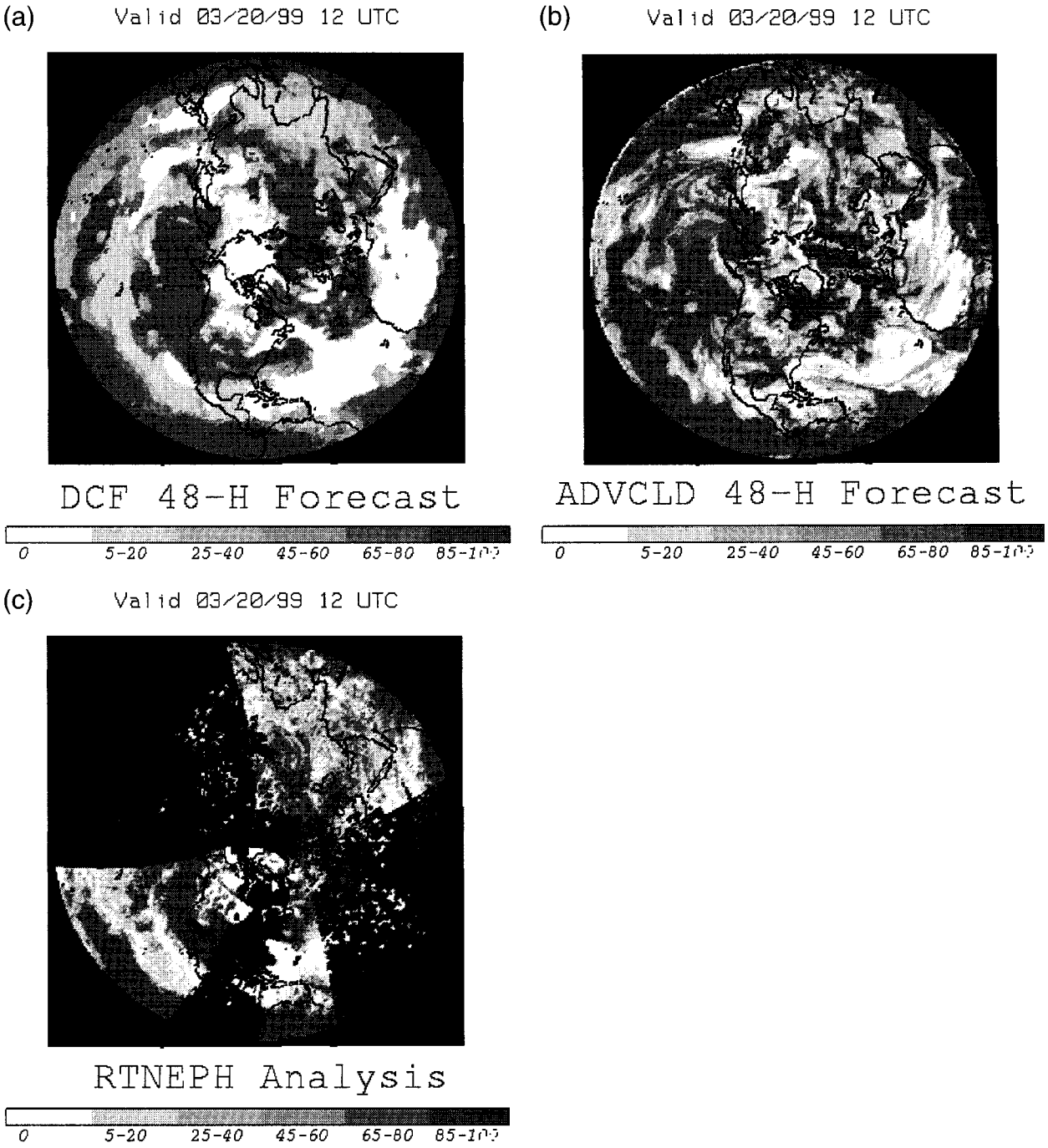

FIG. 5. Same as Fig. 4 but for 20 Mar 1999 from (a) 48-h DCF forecast, (b) 48-h ADVCLD forecast, and (c) RTNEPH analysis.

immediate needs of the operational air force for cloud cover predictions. While cloud water prognostic schemes are becoming more common in mesoscale NWP models, they cannot be converted into CC in a straightforward way. The ambiguity in the relationship between cloud water concentration and what we consider to be fractional obscuration of the ground by clouds also makes it difficult to use satellite observations of the latter to initialize and verify model predictions of the former. Explicit prediction of cloud matter in some form (probably cloud water content as a component of total atmospheric water in all phases) is no doubt the most promising path toward cloud forecasting. Progress is being made in that direction in the atmospheric research community. But a recognition of our current limited status in that progression makes it nec-

essary to find short-term solutions to respond to immediate needs. It was this recognition that motivated the development and testing of the algorithm that is the subject of this study.

Another shortfall in our ability to predict three-dimensional distributions of clouds accurately is the limitation of our cloud observing systems. For all practical purposes, certainly on a worldwide scale, satellite imagery is our only current source of cloud observations. While research in multispectral techniques (e.g., d'Entremont et al. 1997) is attempting to derive more information from the clouds that can be seen by satellite, lower clouds that are not observed and the base altitudes of those that are remain in question. Cloud analysis models like RTNEPH, and more recent NWP model-related cloud analysis schemes (Albers et al. 1996; Zhang et al. 1998; Zhao et al. 1998), use non-cloud-related information to make estimates of the unseen cloud properties like low CC or cloud-base altitude. However, we have no way of knowing how accurate these estimates are. Considerable uncertainty is evident when the products of several cloud analysis methodologies are compared (see, e.g., Masutani et al. 1999). Ground-based cloud radar (e.g., Mace et al. 1998) offers some hope of using active-sensor technology to measure some of these cloud properties accurately, but only over very limited geographic areas. Perhaps even these limited measurements can provide information needed to make advancements in explicit cloud prediction. However, they cannot be expected to provide the cloud information on large spatial scales needed to initialize or verify cloud prediction models.

Because of the uncertainty of the actual three-dimensional cloud distribution on global scales, no information on the performance of the current algorithm's performance in diagnosing cloud type, base altitude, or thickness is included in this paper. While these performance statistics were computed with respect to RTNEPH as a reference, they are considered too uncertain to be presented here. It could be stated that the algorithm could be used to simply project into the future the limited information on cloud type, base altitude, and thickness available in the analyses. However, even this ability is compromised by the consequent lack of physical relationship between these estimated cloud properties and the corresponding NWP model noncloud data that are being used as predictors. In fact, it may be considered a stretch to try to estimate lower CC in a spatially comprehensive manner when it is known that not all of the CC information used to do so was actually observed. However, the design of the diagnostic cloud forecast algorithm to make such estimates for predicted states at least offers a computational means for such predictions that could be used if and when more extensive three-dimensional cloud observations are available.

A model output statistics approach was used in this study in light of the relatively good forecast performance of global models in noncloud predictions. It was felt that perhaps they could provide considerable future-state information that could be related to cloud information. Also, the MOS approach allows us to compensate for the systematic errors of the NWP model. By developing a separate relationship between the forecast model's predictors and the cloud variables at each forecast duration, the MOS technique can account for the growth of such errors in the derived relationship coefficients. By redeveloping the relationships on an ongoing basis using a recent sample of model forecasts and cloud analyses, any changes in the NWP model or cloud analysis model can be quickly absorbed in the MOS technique.

Past studies of CC diagnosis using MOS methods showed that statistical methods that improved rmse skill did so at the expense of compromising the CC frequency distribution. Alternatively, methods designed to preserve the CC frequency distribution of the reference analysis often produced poor skill scores. Previous work (Norquist et al. 1997) suggested that, at least on global scales, multiple discriminant analysis could be coupled with a predictand category selection technique designed to preserve frequency distribution without a significant cost in skill. Consequently, the visual depictions of the resulting cloud diagnoses were much more realistic in appearance than those resulting from methods that sought only to minimize rmse.

The diagnostic cloud forecast algorithm was applied daily to 5-day sets of NOGAPS NWP model forecasts and corresponding forecast valid time RTNEPH CC analysis during the period 5 March–3 April 1999 over the entire Northern Hemisphere. Resulting MDA coefficients were applied forecasts subsequent to each 5-day set to diagnose CC on the 96-km RTNEPH hemispheric grid. The diagnosed CC for total cloud and three vertical decks (high, middle, and low, determined by base altitude of diagnosed cloud layers falling within fixed above-ground-level boundaries) were compared with RTNEPH as a reference for generating verification statistics. Cloud cover predictions from the Air Force Weather Agency ADVCLD cloud trajectory forecast model were also placed in decks and verified against RTNEPH. Persistence forecasts (RTNEPH at forecast initial time) were also verified against RTNEPH as a baseline indicator of forecast value.

The diagnostic cloud forecast algorithm used a forward stepwise regression scheme to select a subset of predictors from a larger pool of candidate predictors. The stepwise regression scheme identified those predictors most closely correlated with CC in each cloud deck and total cloud. Relative humidity was prominent among predictors selected, which substantiates the use of this variable in many simple univariate cloud cover schemes. However, several other predictor variable types were commonly selected as well, including absolute humidity, turbulence variables (such as static stability, and wind shear and wind speed), and nonpredicted geographic variables. Evaporation and precipi-

tation rates were not available in the forecast data used in this study, though past studies (Norquist et al. 1997) have shown that they can contribute as predictors as well.

The diagnostic cloud forecast algorithm, ADVCLD, and persistence all had small and nongrowing CC biases except for ADVCLD in the middle cloud deck. Those larger biases resulted in considerably larger ADVCLD CC rmse in the middle cloud deck that those of DCF or persistence. In the other decks and total cloud, DCF had lower rmse than ADVCLD or persistence at virtually all forecast durations. In addition, the rate of rmse growth with forecast duration was less for DCF than for ADVCLD or persistence. DCF had better 20/20 scores than did ADVCLD or persistence, though the level of improvement was not as great as in rmse. While for 12–72-h forecasts evaluated in this study the CC frequency distribution over the hemisphere persists, DCF clearly outperformed ADVCLD in frequency distribution fit to RTNEPH. ADVCLD better preserves the degree of spatial variability of RTNEPH. In limited visual comparisons of DCF and ADVCLD total CC predictions with RTNEPH, it appears that the DCF depictions provide considerable information for air force mission planning guidance. The greater rate of prediction error growth with forecast time in ADVCLD seems to be reflected in the visualizations; some useful information was present in the 24-h forecast depiction but was very limited in the 48-h forecast graphic.

In evaluating the quality of the diagnosed cloud forecasts, it must be kept in mind that the best a diagnostic algorithm can do is to accurately predict future states of the source predictand. The realism of the predictions can only be as good as the realism of the source of the predictand data used to develop the algorithm. If one wishes to evaluate the predictions against real-world cloud representations (like visual comparisons with satellite imagery), the deficiencies of the source predictand data must be taken into account.

If current research efforts result in substantially improved satellite-based, three-dimensional cloud analyses, it is likely that trajectory, diagnostic, and prognostic cloud prediction techniques would improve. Before this study was conducted, it was expected that a ''crossover'' point in forecast time would be identified, where diagnostic predictions would become better than trajectory predictions. It was thought that since the trajectory model (ADVCLD) was initialized directly from the RTNEPH, its very short duration forecasts would be superior. In total cloud cover only, a crossover point appeared to be present somewhere between 12 and 24 h of forecast time. While such a crossover was not very apparent in this study when RTNEPH was used, it may be more evident if a superior cloud analysis were used. Thus, in the next-generation cloud analysis era, ADVCLD and the diagnostic algorithm are expected to be complementary—ADVCLD better at smaller forecast times, the diagnostic algorithm better at the greater times.

Perhaps the most common question asked during the course of this and past evaluations of the diagnostic cloud forecast algorithm was that of the impact of the duration of the development period on quality of the diagnosed clouds. A 10-day development period was used in the first evaluation of the algorithm (Norquist et al. 1997). In between that study and this, we compared results from the use of a 5- and a 10-day development period using the same predictor and predictand data. Differences in the quality of the diagnosed clouds were negligible as measured by statistical verifications. Thus, it seems that longer development periods do not provide any appreciable benefit in the global context, and actually increase the data storage burden. As for shorter duration periods, we expect that with the smaller development sample (particularly if latitude zones are used) the statistical method (here, MDA) would become less stable. The resulting predictand–predictor relationships would likely yield lower quality cloud diagnoses.

Current cloud analysis techniques such as RTNEPH and ISCCP make different assumptions when placing the various observational data in the three-dimensional depiction. For example, whether or not high clouds are treated as blackbody emitters can make several kilometers difference in the cloud-top altitude assigned to them. Unfortunately, at this relative early stage in the development of objective cloud analysis techniques, there are too many differences in the available analyses to extend the results of such a study as this from one cloud analysis to another. Ironically, it may be easier to do so with the NWP model, because of the similarity of assumptions in their formulations between models. Therefore, the reader is warned against transferring the results of this study to other pairs of NWP models and objective cloud analyses.

In summary, the subject cloud diagnosis forecast algorithm offers many features that lend themselves to operational utility. It uses available NWP model and cloud analysis products and can adapt quickly to changes in both. It can quickly and efficiently produce CC estimates immediately after the hemispheric NWP forecast is available. It produces no significant systematic error in CC forecasts, shows appreciable skill improvement over the current operational cloud forecast model and persistence, and well preserves the CC frequency distribution of the cloud analysis. As more accurate air force cloud analysis products are available (Gustafson et al. 1997), it is expected that the skill and usefulness of the diagnostic cloud forecast algorithm's products (CC as well as cloud type, base altitude, and thickness) will improve correspondingly.

Acknowledgments. The author acknowledges the contribution of several people who were helpful in the production of this paper. Captain Maggie Wonsick of Air Force Weather Agency helped in gaining access to and

reading the RTNEPH data. Chris Franks of Harris Corporation provided similar assistance with the NOGAPS data. Dan DeBenedictus of SenCom Corporation produced the visual depictions shown. Don Chisholm of Air Force Research Laboratory gave supervisory support and personal encouragement without which this project would not have been possible. Visual Numerics, Inc., provided the IMSL software libraries for the computer used in this project. This project was funded through U.S. Air Force Program Element 62101F.

REFERENCES

Albers, S. C., J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analysis of clouds, precipitation, and temperature. *Wea. Forecasting,* **11,** 273–287.

Carter, G. M., and H. R. Glahn, 1976: Objective prediction of cloud amount based on model output statistics. *Mon. Wea. Rev.,* **104,** 1565–1572.

Cianciolo, M. E., 1993: Short-range cloud amount forecasting with model output statistics. PL-TR-93-2205, Phillips Laboratory (AFMC), Hanscom AFB, MA, 166 pp. [NTIS ADA 274769.]

Crum, T. D., 1987: AFGWC cloud forecast models. AFGWC/TN-87-001, Air Force Global Weather Central, Offutt AFB, NE, 76 pp. [Available from AFWA, 106 Peacekeeper Dr., Offutt AFB, NE 68113-4039.]

d'Entremont, R. P., G. B. Gustafson, and T. R. Caudill, 1997: Comparison of satellite-based cirrus retrievals with coincident ground-based radar and lidar observations. Preprints, *Cloud Impacts on DOD Operations and Systems Conf.,* Newport, RI, Department of Defense, 145–148. [NTIS ADA 330020.]

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statisics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

——, A. H. Murphy, L. J. Wilson, and J. S. Jensenius, 1991: Lectures presented at the WMO Training Workshop on the interpretation of NWP products in terms of local weather phenomena and their verification. WMO/TD-No. 421, World Meteorological Organization.

Gustafson, G. B., R. G. Isaacs, and B. McDonald, 1997: The cloud depiction and forecast system II nephanalysis model. Preprints, *Cloud Impacts on DOD Operations and Systems Conf.,* Newport, RI, Department of Defense, 119. [Available from AER, 840 Memorial Drive, Cambridge, MA 02139.]

Hamill, T. M., R. P. d'Entremont, and J. T. Bunting, 1992: A description of the air force real-time nephanalysis model. *Wea. Forecasting,* **7,** 288–306.

Hogan, T. F., and T. E. Rosmond, 1991: A description of the Navy Operational Global Atmospheric Prediction System's spectral forecast model. *Mon. Wea. Rev.,* **119,** 1786–1815.

Jakob, C., 1999: Cloud cover in the ECMWF reanalysis. *J. Climate,* **12,** 947–959.

Kopp, T. J., M. M. Wonsick, L. E. Cantrell, and F. D. Bieker, 1997: ADVCLD: Air Force Global Weather Center's updated cloud forecast model. Preprints, *Cloud Impacts on DOD Operations and Systems Conf.,* Newport, RI, Department of Defense, 107–110. [NTIS ADA 330020.]

Mace, G. G., C. Jakob, and K. P. Moran, 1998: Validation of hydrometeor occurrence predicted by the ECMWF model using millimeter wave radar data. *Geophys. Res. Lett.,* **25,** 1645–1648.

Masutani, M., and Coauthors, 1999: Observing system simulation experiments for NPOESS. Preprints, *13th Conf. on Numerical Weather Prediction,* Denver, CO, Amer. Meteor. Soc., 1–6.

Mitchell, K. E., and D. C. Hahn, 1989: Objective development of diagnostic cloud forecast schemes in global and regional models. Preprints, *Seventh Conf. on Atmospheric Radiation,* San Francisco, CA, Amer. Meteor. Soc., J138–J145.

Nehrkorn, T., and M. Zivkovic, 1996: A comparison of diagnostic cloud cover schemes. *Mon. Wea. Rev.,* **124,** 1732–1745.

Norquist, D. C., 1999: Cloud predictions diagnosed from mesoscale weather model forecasts. *Mon. Wea. Rev.,* **127,** 2465–2483.

——, and S. Chang, 1994: Diagnosis and correction of systematic humidity error in a global numerical weather prediction model. *Mon. Wea. Rev.,* **122,** 2442–2460.

——, C.-H. Yang, S. Chang, and D. C. Hahn, 1992: Phillips Laboratory global spectral numerical weather prediction model. PL-TR-92-2225, Phillips Laboratory (AFMC), Hanscom AFB, MA, 166 pp. [NTIS ADA 267293.]

——, H. S. Muench, D. L. Aiken, and D. C. Hahn, 1994: Diagnosing cloudiness from global numerical weather prediction model forecasts. PL-TR-94-2211, Phillips Laboratory (AFMC), Hanscom AFB, MA, 152 pp. [NTIS ADA 289456.]

——, D. L. Aiken, and D. A. DeBenedictus, 1997: Cloud cover predictions diagnosed from global numerical weather prediction model forecasts. PL-TR-97-2015, Phillips Laboratory (AFMC), Hanscom AFB, MA, 150 pp. [NTIS ADA 331790.]

Rossow, W. B., and R. A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 2–20.

——, A. W. Walker, D. E. Beuschel, and M. D. Roiter, 1996: International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud datasets. WMO/TD-No. 737, World Meteorological Organization, 115 pp.

Schaaf, C. B., J. M. Ward, H. S. Muench, R. P. d'Entremont, M. K. Griffin, G. B. Gustafson, and J. T. Bunting, 1990: The hemispheric eighth mesh terrain elevation and geography data sets. GL-TR-90-0223, Geophysics Laboratory (AFSC), Hanscom AFB, MA, 16 pp. [Available from AFRL/VSBL, 29 Randolph Rd, Hanscom AFB, MA 01731-3010.]

Stobie, J. G., 1986: AFGWC's Advanced Weather Analysis and Prediction System (AWAPS). AFGWC/TN-86-001, Air Force Global Weather Central, Offutt AFB, NE, 98 pp. [Available from AFWA, 106 Peacekeeper Dr., Offutt AFB, NE 68113-4039.]

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Wea. Rev.,* **121,** 3040–3061.

Trapnell, R. N., 1992: Cloud curve algorithm test program. PL-TR-92-2052, Phillips Laboratory (AFMC), Hanscom AFB, MA, 170 pp. [NTIS ADA 253918.]

USAFETAC, 1991: RTNEPH USAFETAC climatic database users handbook no. 1. USAFETAC/UH—86/001, USAFETAC, Asheville, NC, 18 pp. [Available from AFCCC, 151 Patton Ave., Rm. 120, Asheville, NC 28806-5002.]

Visual Numerics, Inc., 1997: IMSL Fortran subroutines for statistical applications, Stat/Library Vols. 1 and 2. [Available online from http://www.vni.com/products/imsl/TechSpecs/FNLStatDescrip.html.]

Zhang, J., F. H. Carr, and K. Brewster, 1998: ADAS cloud analysis. Preprints, *12th Conf. on Numerical Weather Prediction,* Phoenix, AZ, Amer. Meteor. Soc., 185–188.

Zhao, Q., T. L. Black, G. J. DiMego, R. Aune, Y. Lin, M. E. Baldwin, E. Rogers, and K. A. Campana, 1998: Assimilating cloud and precipitation observations into the Eta Model to improve cloud and precipitation forecasts. Preprints, *12th Conf. on Numerical Weather Prediction,* Phoenix, AZ, Amer. Meteor. Soc., 179–180.