U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
SYSTEMS DEVELOPMENT OFFICE
TECHNIQUES DEVELOPMENT LABORATORY

TDL Office Note 76-18

TESTING LFM- AND PE-BASED STATISTICAL CLOUD PREDICTION FOR THE COOL SEASON

Gary M. Carter

December 1976

Testing LFM- and PE-Based Statistical Cloud Prediction for the Cool Season

Gary M. Carter

1. INTRODUCTION

Twice daily the Techniques Development Laboratory's (TDL's) "early" and "final" guidance forecasts of cloud amount (clear, scattered, broken, and overcast) are distributed to National Weather Service (NWS) forecasters via facsimile and/or teletypewriter. Probability estimates and categorical forecasts are available for approximately 230 stations throughout the conterminous United States.

Initially, this operational system to forecast cloudiness relied entirely on prediction equations derived separately for each station. In October 1976 we implemented a set of early guidance cloud prediction equations developed for the 21 regions shown in Fig. 1. We had developed these equations simultaneously with a new set of ceiling prediction equations. Our objective was to improve the overall consistency between TDL's automated forecasts of these two weather elements.

In the near future we also plan to use regionalized equations in the final guidance forecast system. Consequently, we have been studying the bias characteristics (i.e., the number of forecasts in a given category of cloudiness divided by the total number of observations of that category) associated with our newly developed regionalized cloud prediction equations.

2. OPERATIONAL AND TEST EQUATIONS

We conducted several verification experiments for the cool season of October through March using both operational and test equations. All the equations were derived using the Model Output Statistics (MOS) technique (Glahn and Lowry, 1972).

Our early guidance developmental sample consisted primarily of forecasts from the Limited-area Fine Mesh (LFM) model (NWS, 1971) during the cool seasons of 1972-73, 1973-74, 1974-75, and 1975-76. Weather elements from surface observations taken 3 hr after LFM input time were also screened. We derived the operational equations on this entire period, but we used only the first three seasons to develop our test equations.

We applied this same approach in deriving the final guidance operational and test equations. Forecast fields from the Primitive Equation (PE) model (Shuman and Hovermale, 1968) were also included in the data sample, and surface observations 6 hr after the model input times were screened. The same developmental seasons were used as for our early guidance equations.

3. TEST PROCEDURES

As part of this test we devised a new technique for converting our four-category probability forecasts into single "best" category predictions. This new procedure called "modified inflation" increases the likelihood that our categorical forecasts will be consistent with the original forecast probabilities.

The first step in this new procedure involves an analysis of the predicted probability distribution. Either clear and scattered (categories one and two) or broken and overcast (categories three and four) are removed from consideration depending on where the majority of the forecast probability is centered. We make this decision by summing the probability estimates for clear and scattered and then comparing this value with the combined probability for broken and overcast. The two categories with the smallest combined probability are then eliminated as possibilities for our selection of the best category.

Based on the outcome of the preliminary test, a modified form of the inflation transformation is applied as follows:

$$\hat{P}'_j = \frac{\hat{P}_j - \overline{P}_j}{R_j} \times F + \overline{P}_j$$

where $\hat{P}'_j$ is the "inflated" forecast for the j^{th} category of cloud amount ($j=1,2$ or $j=3,4$), $\hat{P}_j$ the original objective probability estimate, $\overline{P}_j$ the average frequency of the cloud amount predictand from the developmental sample, R_j the multiple correlation coefficient of the predictand with the predictors in the forecasting equation, and F an "inflation factor" used to adjust the overall bias characteristics.

The category with the largest value for $\hat{P}'_j$ is selected as the best category forecast.

4. TEST RESULTS

a. Early guidance operational equations

We used developmental data from the cool seasons of 1974-75 and 1975-76 to study the bias characteristics of our early guidance operational equations. Forecasts were produced for all the stations (approximately 230) for projections of 6, 12, 18, and 24 hr from 0000 GMT. Inflation factors of 0.75, 1.0, and 1.5 were applied depending on the length of the forecast projection. Table 1 shows the overall verification scores.

The scores and bias values in Table 1 indicate that a factor of 1.0 works reasonably well for the 6-hr projection, while 1.5 is most desirable for the other three periods. This difference appears to be related to the strong influence of the 0300 GMT surface observations on the bias of the forecasts valid at 0600 GMT.

b. Early guidance test equations

We also generated 12- and 24-hr cloud forecasts using the regionalized early guidance test equations. This time, only the 40 stations shown in Fig. 1 were involved. Our independent data were from the cool season of 1975-76 (0000 GMT cycle forecasts).

One aspect of this test was to compare bias values when both regional and individual station means were used for $\overline{P}_j$ in the modified inflation equation. We conducted this test for the 12-hr projection only. Figs. 2 and 3 show the scattered and broken category bias values for each station.

Also, for both projections we compared regionalized equation forecasts with another set based on our traditional single station approach (see Carter and Hebenstreit, 1976). We determined the single station best categories using a "standard inflation" technique which does not involve preliminary elimination of any categories, and uses a constant inflation factor of 1.0. The results of all our tests are summarized in Table 2.

Table 2 shows that the verification scores are quite similar for each procedure we tested. The bias values differ the most for the difficult to forecast categories of scattered and broken.

As noted before, 12-hr forecast regional mean versus individual station mean scattered and broken category bias values for each of the 40 test stations are presented in Figs. 2 and 3. For most stations it appears as though the use of individual means offers little advantage except near the Great Lakes where the scattered bias is improved (i.e., closer to unity).

c. Final guidance operational equations

Using developmental data from the cool seasons of 1974-75 and 1975-76 we produced seven sets of final guidance cloud forecasts for projections of 12- to 48-hr after 0000 GMT. These predictions for all 230 stations were based on our newly developed regionalized prediction equations. We experimented with inflation factors of 1.0, 1.5, and 2.0 depending on the length of the forecast projection. Table 3 gives the overall verification scores.

The results in Table 3 indicate that an inflation factor of 1.5 produces reasonable bias values for the first three periods. A factor of 2.0 appears to work well for the latter four projections.

d. Final guidance test equations

We derived final guidance test equations for only the 36-hr projection from 0000 GMT. We generated forecasts for all 230 stations using independent data from the cool season of 1975-76. Here we compared our new regionalized equations with the traditional single station equations described by Carter and Glahn (1976). Modified inflation was applied to the regional equation forecasts, while the single station best categories were based on a combination of standard inflation and a minimum bias matrix. The comparative results are given in Table 4.

The overall verification scores and bias values in Table 4 show that the regional equation forecasts using an inflation factor of 2.0 are similar to the predictions from the single station technique.

5. CONCLUSIONS

The results of several tests using both early and final guidance cloud prediction equations indicates that our regionalized equations can produce forecasts with the same level of accuracy and overall bias characteristics as those associated with our traditional single station equations. However, the individual station-by-station bias values often differ considerably for these two schemes.

Our modified inflation technique produces relatively accurate best category forecasts. Depending on the length of projection, inflation factors of 1.0 and 1.5 result in acceptable bias values for the early guidance forecasts. Factors of 1.5 and 2.0 work better for the final guidance predictions. Using station means instead of regional means as part of the modified inflation procedure contributes very little towards improving the bias values for individual stations.

ACKNOWLEDGMENTS

I am grateful to Tom Grayson and Bob Glahn for the valuable suggestions they contributed towards developing the test procedures used in this study. I also appreciate Anna Booth's assistance in preparing the figures, and Mercedes Bakon's help with typing the manuscript and tables. Lt. Col. Karl Hebenstreit, USAF Liaison Officer assigned to TDL, and Rich Crisci assisted in preparing the developmental data used in this study.

REFERENCES

Carter, G. M., and K. F. Hebenstreit, 1976: Testing LFM-based statistical cloud prediction equations for the cool season. TDL Office Note 76-12, Techniques Development Laboratory, Silver Spring, 8 pp.

______, and H. R. Glahn, 1976: Objective prediction of cloud amount based on model output statistics. Mon. Wea. Rev., 104, 1565-1572.

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

National Weather Service, 1971: The limited area fine mesh (LFM) model. NWS Tech. Proc. Bull., No. 67, 11 pp.

Shuman, F. G., and J. B. Hovermale, 1968: An operational six-layer primitive equation model. J. Appl. Meteor., 7, 525-547.

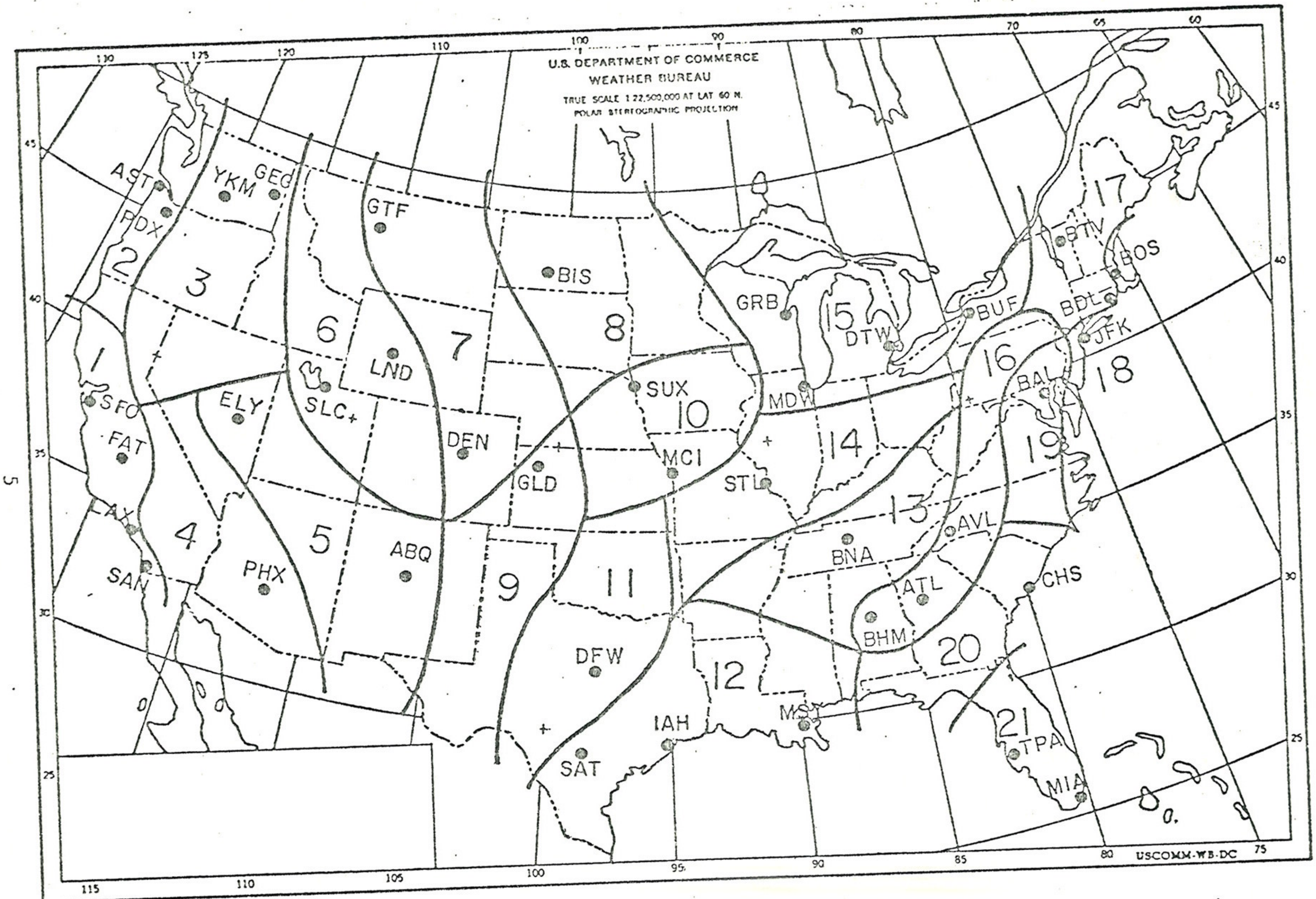

Figure 1. The 21 developmental regions and 40 stations used for testing the regional cloud prediction equations.

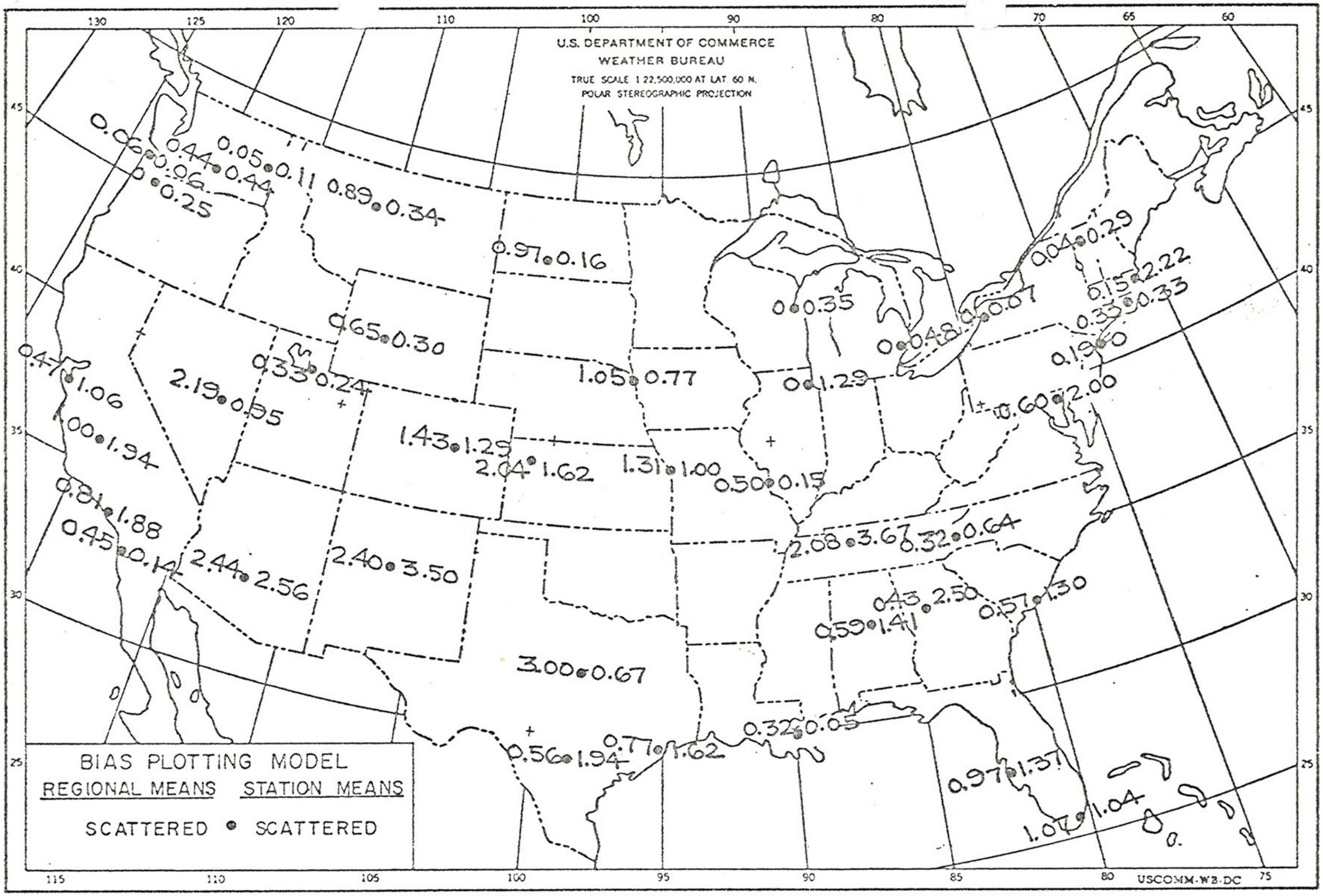

Figure 2. Bias values for 12-hr forecasts of scattered cloudiness at 40 U.S. stations during October 1975 through March 1976 using regional and individual station means in the best category selection process.

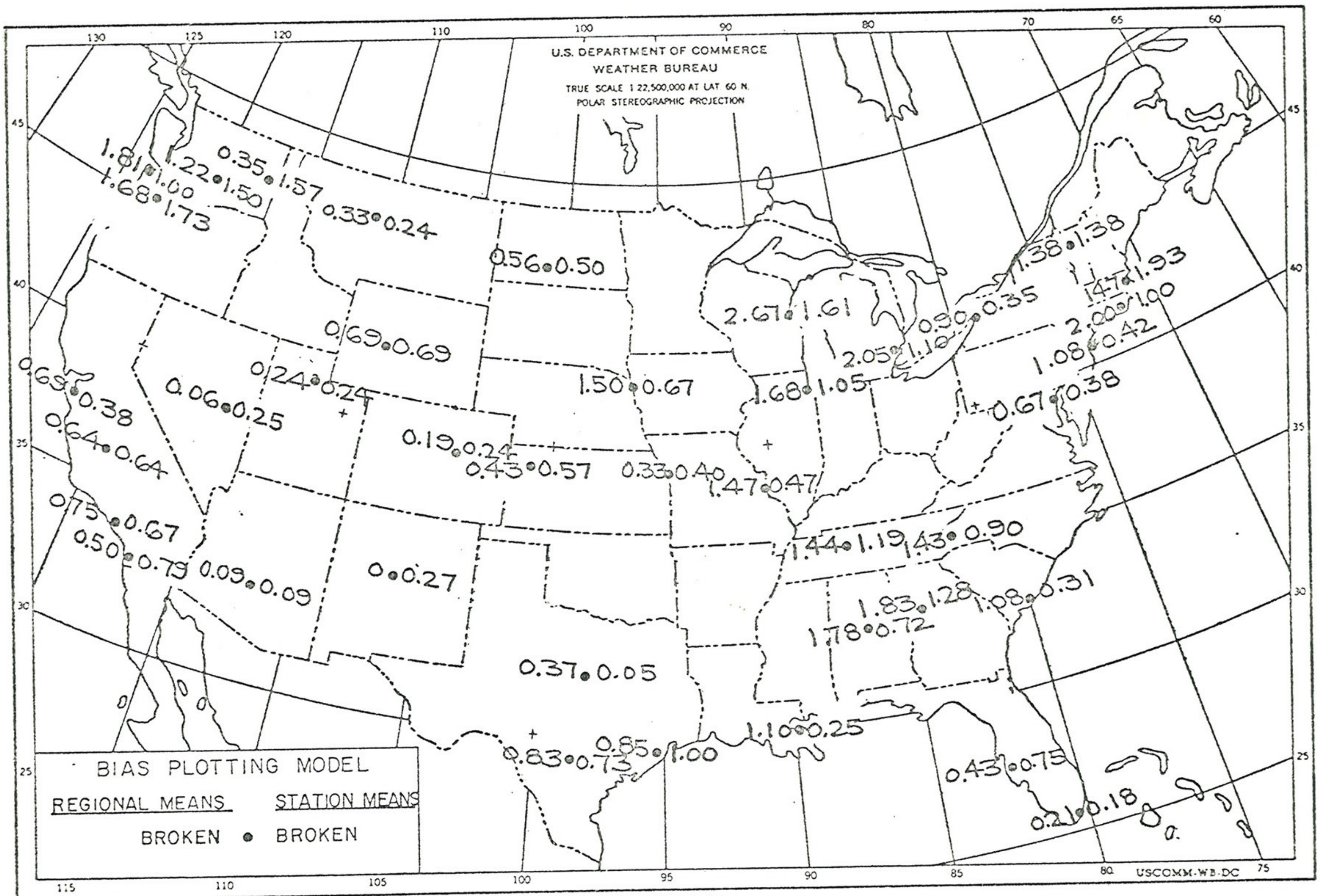

Figure 3. Same as Fig. 2 except for 12-hr forecasts of broken cloudiness.

Table 1. Combined verification scores for approximately 230 U.S. stations during October through March of 1974-75 and 1975-76 for early guidance forecasts of four categories of cloud amount using the modified inflation best category selection technique and regional prediction equations.

PROJECTION (HOURS FROM 0000 GMT)	TYPE OF FORECAST	BIAS = NO. FCST/NO. OBS				PERCENT CORRECT	SKILL SCORE	NO. OF CASES
		CLEAR (No. Obs.)	SCTD (No. Obs.)	BROKEN (No. Obs.)	OVERCAST (No. Obs.)			
6	Inflation Factor = 0.75 Inflation Factor = 1.0	0.99 0.93 (29043)	1.16 1.35 (9090)	0.56 0.68 (6834)	1.07 1.04 (24297)	69.4 68.6	0.543 0.537	69264
12	Inflation Factor = 1.0 Inflation Factor = 1.5	1.20 1.14 (25573)	0.64 0.78 (10197)	0.64 0.87 (8138)	1.06 0.99 (26914)	63.2 62.0	0.450 0.442	70822
18	Inflation Factor = 1.0 Inflation Factor = 1.5	1.18 1.14 (21347)	0.80 0.85 (14196	0.86 0.98 (12641)	1.04 0.97 (22953)	55.3 54.7	0.385 0.380	71137
24	Inflation Factor = 1.0 Inflation Factor = 1.5	1.16 1.12 (24074)	0.79 0.87 (13963)	0.75 0.88 (10837)	1.08 1.01 (22353)	55.0 54.2	0.370 0.363	71227

Table 2. Combined verification scores for 40 U.S. stations during October 1975 through March 1976 for early guidance predictions of four categories of cloud amount based on regional and single station equations.

PROJECTION (HOURS FROM 0000 GMT)	TYPE OF FORECAST	BIAS = NO. FCST/NO. OBS				PERCENT CORRECT	SKILL SCORE	NO. OF CASES
		CLEAR (No. Obs.)	SCTD (No. Obs.)	BROKEN (No. Obs.)	OVERCAST (No. Obs.)			
12	Regional Equations (Regional Means, Inflation Factor = 1.5)	1.17	0.71	0.98	0.94	58.8	0.397	6238
	Regional Equations (Station Means, Inflation Factor = 1.5)	1.07	0.90	0.73	1.06	59.2	0.404	
	Single Station Eqns. (Standard Inflation)	1.07 (2450)	0.85 (926)	0.87 (769)	1.03 (2093)	59.1	0.405	
24	Regional Equations (Regional Means, Inflation Factor = 1.5)	1.07	0.86	1.05	0.98	51.3	0.324	6239
	Single Station Eqns. (Standard Inflation)	1.05 (2304)	0.90 (1264)	0.92 (978)	1.05 (1693)	50.9	0.317	

Table 3. Same as Table 1 except for final guidance cloud forecasts.

PROJECTION (HOURS FROM 0000 GMT)	TYPE OF FORECAST	BIAS = NO. FCST/NO. OBS				PERCENT CORRECT	SKILL SCORE	NO. OF CASES
		CLEAR (No. Obs.)	SCTD (No. Obs.)	BROKEN (No. Obs.)	OVERCAST (No. Obs.)			
12	Inflation Factor = 1.0	1.13	0.75	0.71	1.05	64.4	0.471	69596
	Inflation Factor = 1.5	1.09	0.86	0.90	1.00	63.3	0.463	
		(25148)	(10001)	(7966)	(26481)			
18	Inflation Factor = 1.0	1.20	0.77	0.86	1.04	55.6	0.389	69805
	Inflation Factor = 1.5	1.16	0.82	0.97	0.98	55.0	0.384	
		(20937)	(13956)	(12424)	(22488)			
24	Inflation Factor = 1.0	1.18	0.77	0.73	1.08	55.6	0.377	69802
	Inflation Factor = 1.5	1.14	0.84	0.86	1.01	55.0	0.373	
		(23560)	(13700)	(10632)	(21910)			
30	Inflation Factor = 1.0	1.16	0.64	0.44	1.10	61.2	0.400	69464
	Inflation Factor = 2.0	1.08	0.91	0.82	0.99	59.0	0.385	
		(29019)	(9172)	(6930)	(24343)			
36	Inflation Factor = 1.0	1.17	0.59	0.57	1.12	58.1	0.371	69906
	Inflation Factor = 2.0	1.10	0.77	0.98	1.00	56.0	0.357	
		(21125)	(10086)	(8025)	(26670)			
42	Inflation Factor = 1.0	1.17	0.72	0.78	1.14	50.4	0.315	70415
	Inflation Factor = 2.0	1.12	0.79	0.95	1.05	49.4	0.307	
		(20949)	(14252)	(12534)	(22680)			
48	Inflation Factor = 1.0	1.19	0.73	0.61	1.15	50.3	0.298	70114
	Inflation Factor = 2.0	1.12	0.85	0.79	1.07	49.0	0.289	
		(23542)	(13776)	(10695)	(22101)			

Table 4. Same as Table 2 except for 230 stations and for final guidance cloud forecasts.

PROJECTION (HOURS FROM 0000 GMT)	TYPE OF FORECAST	BIAS = NO. FCST/NO. OBS				PERCENT CORRECT	SKILL SCORE	NO. OF CASES
		CLEAR (No. Obs.)	SCTD (No. Obs.)	BROKEN (No. Obs.)	OVERCAST (No. Obs.)			
36	Regional Equations (Inflation Factor = 1.0)	1.12	0.52	0.52	1.22	56.8	0.349	33981
	Regional Equations (Inflation Factor = 2.0)	1.06	0.69	0.95	1.08	55.0	0.339	
	Single Station Eqns. (Standard Inflation and Minimum Bias Matrix)	1.04 (13109)	0.80 (4982)	0.81 (3935)	1.11 (11955)	53.9	0.323	