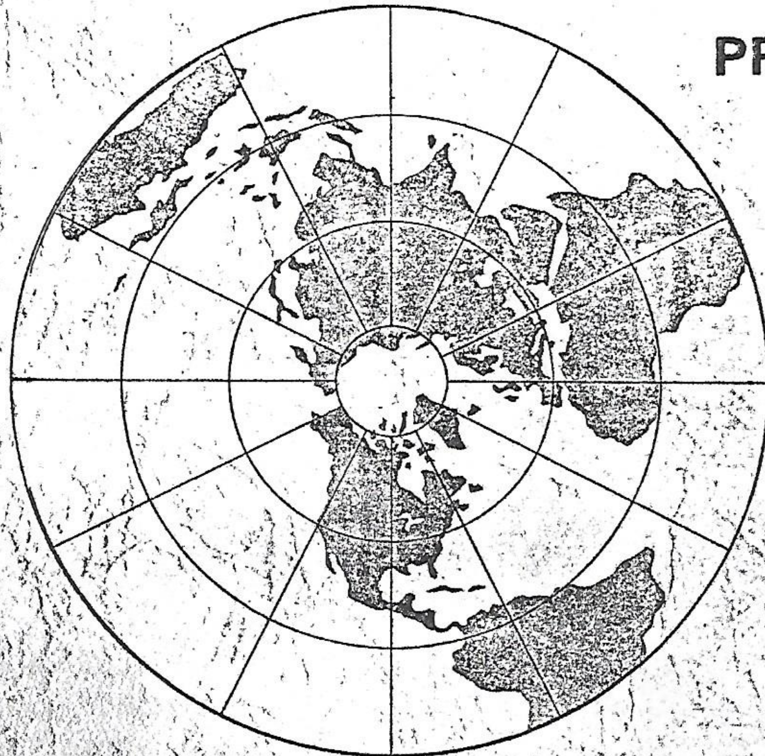
PREPRINTS
OF
PAPERS
PRESENTED
AT THE



INTERNATIONAL SYMPOSIUM ON PROBABILITY AND STATISTICS IN THE ATMOSPHERIC SCIENCES

JUNE 1-4, 1971
HONOLULU, HAWAII

STATISTICS OF NUMERICAL PREDICTION MODELS

Harry R. Glahn, Dale A. Lowry, George W. Hollenbaugh, and John R. Annett

Techniques Development Laboratory
National Weather Service, NOAA
Silver Spring, Maryland

1. INTRODUCTION

Until rather recently, objective forecasting methods have been considered as falling into one of two categories--dynamical and statistical. Now, the relatively new field of stochastic-dynamic prediction is being explored and is beginning to show promise for operational use sometime in the future. However, until stochastic-dynamic prediction is developed much further and more powerful computers are available, we must use some combination of dynamical and statistical methods for practical forecasting.

There has been little success in the prediction of such variables as surface wind, probability and form of precipitation, maximum and minimum temperature, cloudiness, ceiling, and visibility with dynamic models, and indeed, most models do not even forecast these variables directly. There are two general ways in which statistics can be used and the results applied to predictions from numerical models to yield estimates of those elements not successfully forecast directly by the numerical models.

The first is usually called the perfect prog method. A concurrent statistical relationship is developed between the variable to be estimated and selected variables forecast by a dynamic model. Both predictand and predictors are observed quantities in the developmental sample. In application, this relationship is applied to numerical model output at, say, a projection of 36 hours to get an estimate of the predictand 36 hours after the data input time for the numerical model.

The other method, which we call Model Output Statistics (MOS), consists of determining a statistical relationship between the predictand and variables from the numerical model at the desired projection time. Application is made in exactly the same way as with the perfect prog method.

The MOS technique is, in effect, the determination of the "weather-related" statistics of a numerical model. For instance, we may want to know what percent of the time it rains when the model predicts 80 percent relative humidity, or, what the best estimate is of the surface wind at an airport, when a model predicts a particular 1000-mb geostrophic wind at that point in time and space. In the following sections we will describe some of our results with this technique.

2. PROBABILITY OF PRECIPITATION

Perhaps the first major use of MOS was in the estimation of probability of precipitation (PoP) (Glahn and Lowry, 1969). To do this, we used a combination of the variables predicted by two models being run operationally at the National Meteorological Center (NMC): the Primitive Equation (PE) model (Shuman and Hovermale, 1968) and the Subsynoptic Advection Model (SAM) (Glahn, Lowry, and Hollenbaugh, 1969). These PoP forecasts are made about 3:00 A.M. EST for the "today" period 7:00 A.M. to 7:00 P.M. The area covered is generally the United States east of the Mississippi River.

Screening Regression was used to select predictors which contributed most to the reduction of variance of the binary predictand. Data for nearly 100 stations were grouped together and the equations developed and updated twice a year (Summer: April-September; Winter: October-March). Predictors were all binary and were computed from PE relative humidity, SAM saturation deficit, PE precipitation amount, and SAM sea level pressure.

Figure 1 shows a portion of a facsimile chart transmitted January 18, 1971. The isopleths indicate PoP forecasts for the 12-hr period 1200 GMT January 18 to 0000 GMT January 19. In this particular chart, there is a variation from < 5% to > 55% over a distance of about 150 miles.

We have compared the MOS forecasts with those made at local stations since July 1968 and with those made subjectively at NMC since July 1969; Figure 2 gives the monthly and yearly Brier scores.* This figure indicates:

a) The MOS forecasts were better than the locals for the first 15 months of comparison. The reverse is true for the last 15 months. (The locals <u>are</u> made with 1 or 2 hours later data available, and many stations receive the MOS forecasts before the locals are finalized.)

b) The MOS forecasts have been better, on the average, than the NMC guidance for the 18-month period of comparison.

---

*Brier Score $= \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2$, where N = sample size, F = forecast probability, and O = observed (O = no precipitation, 1 = precipitation).

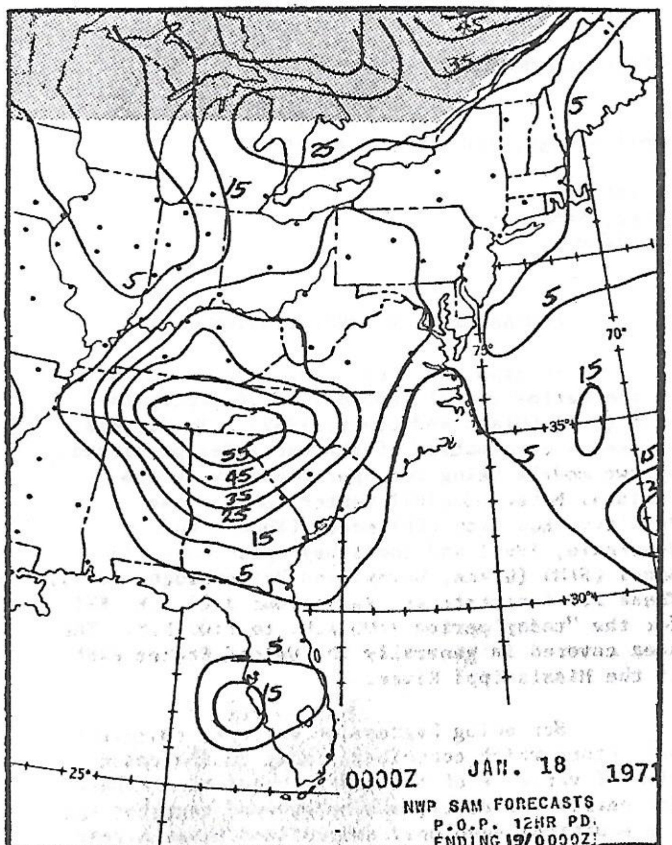


Figure 1. A portion of a facsimile chart transmitted January 18, 1971. The isopleths, labeled in percent, indicate PoP forecasts for the 12-hr period 1200 GMT January 18 to 0000 GMT January 19.

The sharp decrease in skill of the MOS forecasts relative to the locals was mainly due to a major change in the moisture portion of the PE model in late October 1969. Because of this change, PE predictors were not used from December 1969 through August 1970. Since that time PE predictors have been included, following another PE change in September 1970.

The effort in PoP forecasting is being extended to all three forecast periods by combining the output from the TDL trajectory model (Reap, 1968), which is also being run operationally at NMC, with that from the PE model. These Primitive Equation and Trajectory (PEAT) MOS forecasts have been compared with those made by the perfect prog technique (Klein, 1971) as well as with the official local office forecasts and the NMC guidance.

Table 1 gives the Brier Scores for October, November, and December for 86 stations over the 48 states. Figure 3 shows the bias in the four sets of forecasts. From this somewhat limited verification, we conclude:

a) The PEATMOS forecasts are competitive with NMC and, except for the first period, with the local offices.

b) The PEATMOS forecasts are too low for all forecast categories except 6-15 percent. Since a large number of forecasts are in this category, the overall bias is small. All other forecasts are too high on the average, particularly the perfect prog system.

Table 1. Brier Scores for local, NMC, PEATMOS, and perfect prog PoP forecasts for 86 stations in the 48 states. Climatology is long-term relative frequency by month and by station

			LOCAL	NMC	MOS	PERFECT PROG	CLIMATOLOGY	CASES
PERIOD I (TODAY)	BRIER SCORE	OCTOBER	.0804	.1020	.0957	.1134	.1273	750
		NOVEMBER	.0836	.1012	.1013	.1199	.1452	2316
		DECEMBER	.0891	.1082	.1024	.1189	.1522	2220
		AVERAGE	.0854	.1042	.1009	.1185	.1456	(5286)
	IMPROVEMENT OVER CLIMATOLOGY	AVERAGE	.413	.284	.307	.186		
	BIAS (PERCENT)	AVERAGE	2.6	4.6	- .2	6.9		
PERIOD 2 (TONIGHT)	BRIER SCORE	OCTOBER	.1013	.1073	.0981	.1256	.1229	752
		NOVEMBER	.1104	.1167	.1175	.1290	.1503	2282
		DECEMBER	.1196	.1247	.1221	.1275	.1529	2218
		AVERAGE	.1129	.1187	.1166	.1278	.1474	(5252)
	IMPROVEMENT OVER CLIMATOLOGY	AVERAGE	.234	.195	.209	.133		
	BIAS (PERCENT)	AVERAGE	2.2	4.4	- .7	6.6		
PERIOD 3 (TOMORROW)	BRIER SCORE	OCTOBER	.1219	.1225	.1212	.1312	.1342	750
		NOVEMBER	.1208	.1226	.1316	.1373	.1468	2289
		DECEMBER	.1373	.1341	.1375	.1341	.1546	2215
		AVERAGE	.1279	.1274	.1326	.1350	.1482	(5254)
	IMPROVEMENT OVER CLIMATOLOGY	AVERAGE	.137	.140	.105	.089		
	BIAS (PERCENT)	AVERAGE	1.2	3.1	- 2.8	6.5		

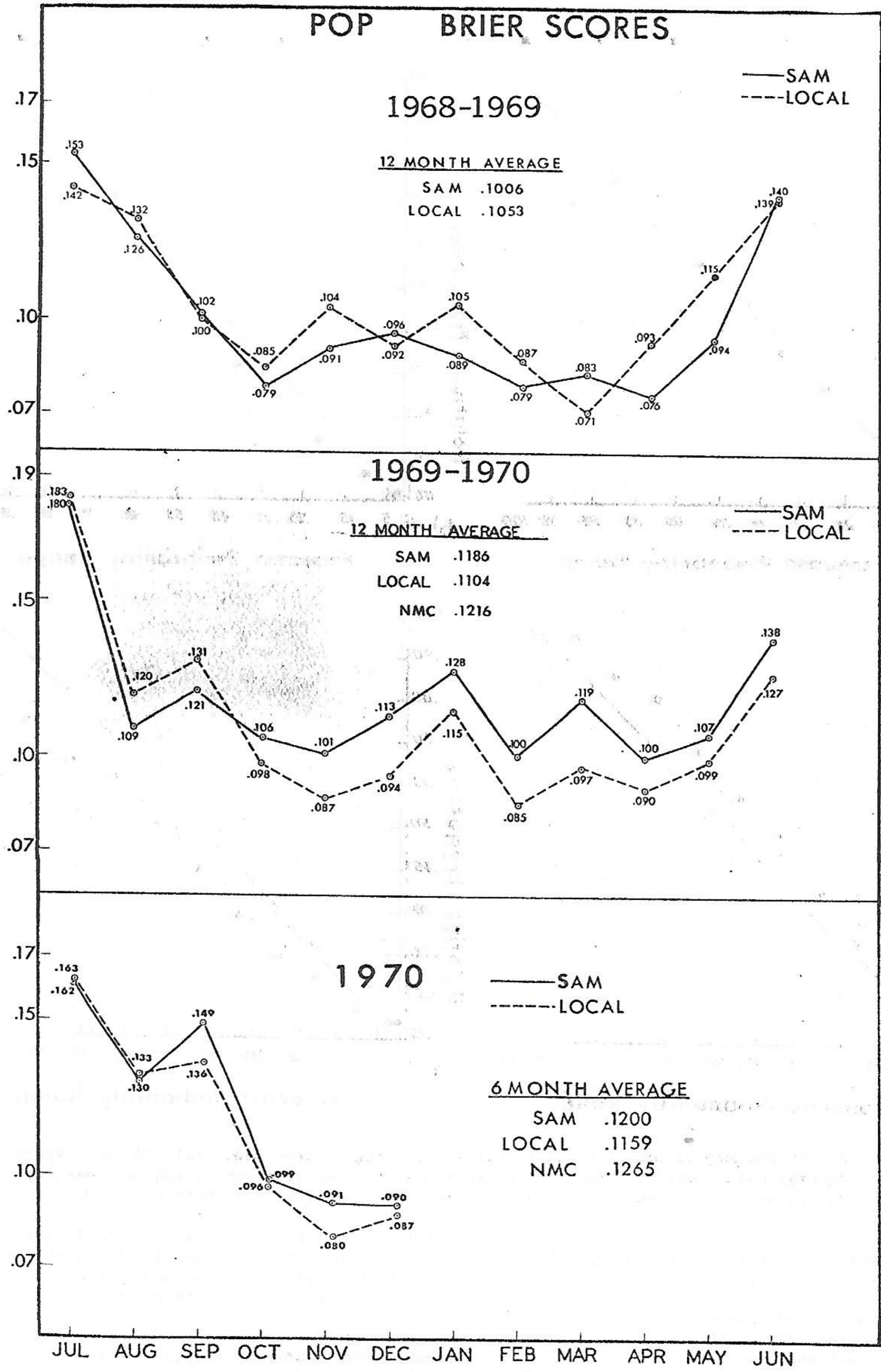


Figure 2. Monthly and average Brier Scores for local forecasts and MOS forecasts from the SAM teletype and facsimile products for July 1968 through December 1970. NMC Scores are given as averages only. Nineteen stations in the eastern U. S. are included.

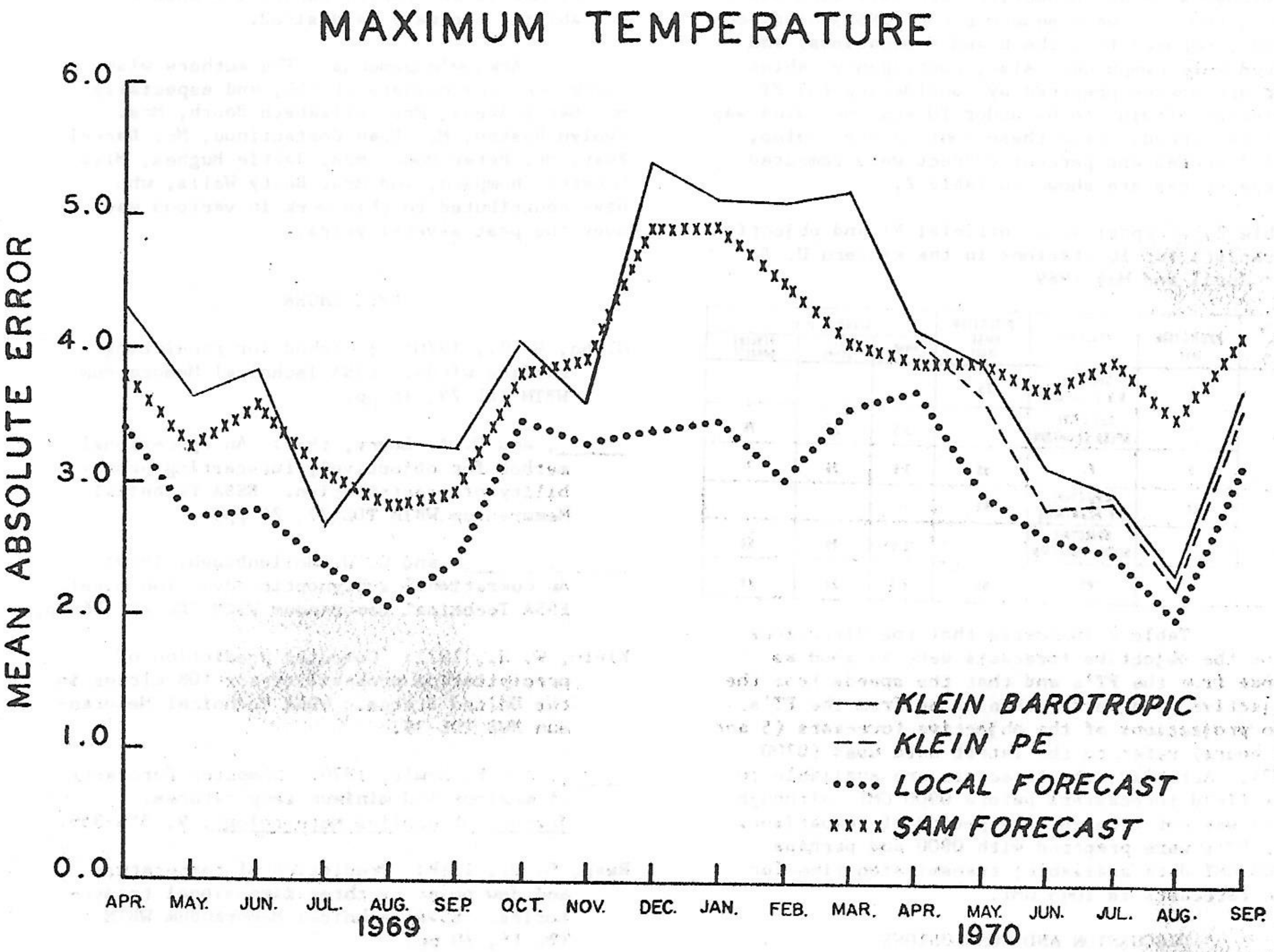


Figure 4. Mean absolute errors for local, MOS (SAM), and perfect prog (Klein) forecasts. Sixteen stations in the eastern U. S. are included.

b) The perfect prog and MOS techniques gave forecasts of about equal skill.

c) The perfect prog system applied to PE forecasts gave slightly better results than the same system applied to barotropic forecasts in all 6 months for which the comparison was made.

It is clear that for this short range forecast, the forecasts made at local stations have less error, on the average, than our objective forecasts. Quite likely, to equal the local skill, we will have to include more information about the initial soundings and, perhaps, stratify according to synoptic situation. A complicating factor is that the maximum temperature does not occur at the same time each day. The poor relative performance of MOS during the 1970 summer was probably due to some change in the PE model that affected the PE temperature forecasts.

4. SURFACE WIND

Separate regression equations were developed for estimating the U and V wind components and the wind speed valid at 1200 and 1800 GMT for each of 10 stations in the eastern U. S.--Albany, Atlanta, Baltimore, Cleveland, Cincinnati, Washington, New York, New Orleans, Chicago, and St. Louis (Glahn, 1970). Data for April through September 1967 and 1968 were used. The most important predictors were geostrophic winds from SAM, although other variables such as 1000- and 500-mb winds and temperatures and initial observed winds also played a part.

The equations were evaluated for each day in April and May 1969 for which SAM data tapes were available. The wind forecasts in the terminal forecasts (FT's) made at the Weather Service Offices were used for comparison. Since the FT's do not mention wind if the speed is expected to be less than 10 kts, the comparison was made in two ways.

For all those cases were the FT's included wind and objective forecasts were available, the root mean square error (RMSE) of direction (computed from the U and V equations) and speed were computed. Also, contingency tables for speed were prepared by considering the FT forecast of wind to be under 10 kts when wind was not mentioned. From these contingency tables, skill scores and percent correct were computed. These scores are shown in table 2.

Table 2. Comparison of official FT and objective forecasts for 10 stations in the eastern U. S. for April and May 1969.

VALID TIME (GMT)	PROJECTION (HR)	FORECAST	DIRECTION RMSE (DEG)	SPEED (kts)		
				RMSE	SKILL SCORE	PERCENT CORRECT
12	5	OBJECTIVE U,V EQUATIONS	35			
	5	OBJECTIVE SPEED EQUATION		3.5	.37	76
	3	FT	33	3.6	.36	71
18	11	OBJECTIVE U,V EQUATIONS	47			
	11	OBJECTIVE SPEED EQUATION		3.5	.29	54
	9	FT	50	4.3	.24	49

Table 2 indicates that the directions from the objective forecasts were as good as those from the FT's and that the speeds from the objective were better than those from the FT's. The projections of the objective forecasts (5 and 11 hours) refer to the latest data used (0700 GMT). Actually, the forecasts are available to the field forecasters before 0900 GMT, although this was not true for the period of comparison. The FT's were prepared with 0900 and perhaps 1000 GMT data available; transmission time for the forecasts is 1045 GMT.

5. DISCUSSION AND CONCLUSIONS

In this paper we have described the MOS technique and some of our uses of it within TDL. The verifications completed to date are very encouraging in that the MOS forecasts are competitive with both local forecasts and NMC guidance, although in most cases the local forecasts are slightly better. Only within the last few years have objective forecasts of actual "weather" variables been mass-produced with an accuracy that can compete with trained forecasters. We expect this trend to continue so that objective forecasts can be used unchanged a large portion of the time and modified only in special and difficult synoptic situations.

The main problems encountered with MOS are:

a) Only relatively small data samples are available.

b) Frequent updates of the system are necessary, or, at least, desirable.

c) Numerical model changes may render a system unusable or necessitate a change in application.

Even with the problems listed above, we feel the use of MOS is desirable, especially when a probability estimate is desired.

Acknowledgements. The authors wish to thank the many members of TDL, and especially Mr. Harry Akens, Mrs. Elizabeth Booth, Mrs. Evelyn Boston, Mr. Dean Costantinou, Mr. Darrel Foat, Mr. Peter Howe, Mrs. Jackie Hughes, Miss Loretta Thompson, and Mrs. Betty Wells, who have contributed to this work in various ways over the past several years.

REFERENCES

Glahn, H. R., 1970: A method for predicting surface winds. ESSA Technical Memorandum WBTM TDL 29, 18 pp.

______, and D. A. Lowry, 1969: An operational method for objectively forecasting probability of precipitation. ESSA Technical Memorandum WBTM TDL 27, 24 pp.

______, ______, and G. W. Hollenbaugh, 1969: An operational subsynoptic advection model. ESSA Technical Memorandum WBTM TDL 23, 26 pp.

Klein, W. H., 1971: Computer prediction of precipitation probability for 108 cities in the United States. NOAA Technical Memorandum NWS TDL-39.

______, and F. Lewis, 1970: Computer forecasts of maximum and minimum temperatures. Journal of Applied Meteorology, 9, 350-359.

Reap, R. M., 1968: Prediction of temperature and dew point by three-dimensional trajectories. ESSA Technical Memorandum WBTM TDL 15, 20 pp.

Shuman, F. G., and J. B. Hovermale, 1968: An operational six-layer primitive equation model. Journal of Applied Meteorology, 7, 525-547.