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

TDL Office Note 77-16

EXPERIMENTS TO IMPROVE THE CONDITIONAL PROBABILITY OF FROZEN PRECIPITATION (PoF) FORECASTS FOR THE EASTERN REGION OF THE NATIONAL WEATHER SERVICE

Joseph R. Bocchieri

August 1977

EXPERIMENTS TO IMPROVE THE CONDITIONAL PROBABILITY OF FROZEN PRECIPITATION (PoF) FORECASTS FOR THE EASTERN REGION OF THE NATIONAL WEATHER SERVICE

Joseph R. Bocchieri

I. INTRODUCTION

During the 1976-77 winter season, the PoF forecasts were generally too low (well below 50%) on many occasions when snow occurred at stations on the coastal plain east of the Appalachian Mountains. We call this area, shown in Fig. 1, the "East of the Appalachian Region" (EOA Region). In this paper, we'll show verification statistics that confirm this undesirable characteristic of the PoF forecasts in the EOA region and describe our attempts to improve the system.

To review briefly, PoF forecasts are available in both the early and final FOUS12 guidance packages (National Weather Service, 1976; Bocchieri and Glahn, 1976). The PoF forecasts in the early FOUS12 are computed solely from the Limited-area Fine Mesh (LFM) (National Weather Service, 1971) model output; the PoF forecasts in the final FOUS12 are computed solely from the Primitive Equation (PE) (Shuman and Hovermale, 1968) model output. The basic predictors in PoF are PE and LFM forecasts of 850-mb temperature (850 T), 1000-500 mb thickness (10-50 TH), 1000-850 mb thickness (10-85 TH), and boundary-layer potential temperature (BLPT). Actually, the predictors are included in the forecast equations as deviations from 50% values. The 50% value of a variable is that value below which the probability of frozen precipitation (snow) is greater than 50%. We determined 50% values for each of the above variables, for each MOS station and for both the LFM and PE models. The 50% values vary from station to station mainly due to differences in station elevation. For the LFM model, we used data from three winter seasons to develop the 50% values--September through April, 1972-73 through 1974-75. For the PE model, we used data from five seasons--1969-70 through 1973-74.

In a previous paper (Bocchieri, 1977), we analyzed two cases from the 1976-77 winter when snow occurred in the EOA region with low PoF forecasts. Our analysis seemed to indicate that the early guidance PoF underforecasted snow more than the final guidnace PoF; also, poor LFM BLPT forecasts seemed to contribute most to the inaccuracy of the forecasts. The verification shown below supports those findings.

II. EXPERIMENTAL EARLY GUIDANCE PoF SYSTEMS

Table 1 shows the bias[1] of the early and final PoF guidance for the 12- and 24-hr forecast projections from the 0000 GMT cycle time. The sample consisted of independent data from the period October 1976 through March 1977. We used a threshold probability of 50% to obtain a categorical snow forecast. The results indicated that the early and, to a lesser extent, the final PoF guidance underforecasted the snow event in the EOA region during the 1976-77 winter season. This confirmed our subjective suspicions. However, for the rest of the Eastern Region the bias was about perfect (near 1.00)

---

[1] The bias is the number of snow forecasts divided by the number of observed snow events.

for both systems.

To improve the bias in the early PoF guidance, we experimented with three new PoF systems--a generalized operator system (EXP GEN), a 10-predictor regionalized system (EXP REG10), and a 4-predictor regionalized system (EXP REG4). The EXP GEN system had the same predictors as the operational system, which is also generalized operator, except that we substituted the 10-85 TH for BLPT in the equations. This was based on Bocchieri's (1977) findings that the LFM BLPT forecasts contributed most to the inaccuracy of the early guidance forecasts for two case studies. In the regionalized systems, we developed prediction equations only for the EOA region by combining data for 21 stations. We used basically the same predictors as those in the EXP GEN system. We developed two regionalized systems because the 4-predictor system might be more stable than the 10-predictor system on independent data. We developed all the experimental systems with the same data sample used for the operational system.

We did a comparative verification for the EOA region between the operational early guidance (OPER), EXP GEN, and EXP REG10 PoF systems. We used three data samples--the developmental sample and two independent samples, (the 1975-76 and 1976-77 winter seasons). We computed the P-Score (Brier, 1950)[1], bias, and threat score[2] for the 12- and 24-hr projections from the 0000 GMT LFM cycle time. We don't show the scores for the EXP REG4 system since it didn't do as well as the EXP REG10 system. A threshold probability of 50% was used for all systems to obtain categorical snow forecasts. The results, shown in Table 2, indicated that the EXP GEN system was generally the best system for the independent data samples. As expected, the EXP REG10 system did best for the developmental data sample. Also, note that the OPER system underforecasted the snow event not only for the 1976-77 season (see Table 1) but also for the 1975-76 season and the developmental sample. In fact, all systems underforecasted snow for all the samples. This was especially surprising for the developmental sample, for which we expected a bias near 1.00.

Amazing!

In Table 2, note that the bias for the EXP GEN system was generally between .70 and .80. Therefore, while the EXP GEN system was the most accurate, it wasn't able to improve upon the bias of the OPER system to a satisfactory degree. We next tried to improve the bias by lowering the threshold probability value. We realized that a lower threshold value would produce more, but not necessarily correct, snow forecasts. Table 3 shows various verification scores for forecasts made from a number of threshold probabilities for the OPER and EXP GEN PoF systems. We used data combined from 21 stations in the EOA region from the independent samples of 1975-76 and 1976-77. Note that as the threshold probability decreased, the bias improved; the reason was that more snow forecasts were made. However, the post-agreement decreased because relatively fewer of the new snow forecasts

---

[1] The score shown is actually 1/2 the value of the score defined by Brier.

[2] The threat score is defined as: threat score = $\frac{C}{F + O - C}$,

where C is the number of correct snow forecasts, F is the number of snow forecasts, and O is the number of observed snow events.

were correct. Also, the threat-score increased; this was expected with more snow forecasts. Further examination of Table 3 for the 12-hr projection shows that, for the OPER system, a threshold probability near 20% gave a bias near 1.00. For the EXP GEN system, a 25% threshold probability gave a bias of 1.00. So, at the "minimum bias" level, the threshold probabilities for the two systems were similar. Also, for these minimum bias threshold probabilities, the other scores were about the same for the two systems. A similar result was found for the 24-hr projection. Therefore, based on this analysis, we decided not to change the operational early guidance PoF system for the EOA region for the 1977-78 winter season because we couldn't improve the bias to a satisfactory degree. However, we do recommend that forecasters in the EOA region use the threshold probabilities in Table 4 to transform the PoF forecast into a categorical forecast. We emphasize that there is no guarantee that these new threshold probabilities will give a better bias for the 1977-78 winter season. However, this is the best suggestion we could make based on results for previous data samples.

A lot better than using 50% as T.H.

III. SUMMARY

During the 1976-77 winter season, the early PoF guidance and, to a lesser extent, the final PoF guidance underforecasted snow in the EOA region, but not in the rest of the Eastern Region. We found that, by substituting 10-85 TH for BLPT in the early guidance equations, we were able to improve the accuracy of the PoF forecasts. However, the new equations still underforecasted the snow event. Attempts to improve the bias by developing PoF equations specifically for the EOA regions also failed.

We found that we could improve the bias in the PoF equations by lowering the threshold probability values. In so doing, there was little difference between the operational and experimental PoF systems. Therefore, we decided not to change the operational PoF equations for the 1977-78 winter season. However, we do recommend that forecasters in the EOA region use the PoF guidance with the threshold probabilities shown in Table 4.

Note that there are plans to implement new numerical models at the National Meteorological Center before or during the 1977-78 winter season. When that happens, the values given in Table 4 will only be approximately true. Even with the new models, we feel that lower threshold probabilities will be necessary for the EOA region.

We plan to revamp the PoF system in time for the 1978-79 winter season so that the probability of freezing rain will also be given. Hopefully, we'll be able to improve the reliability of the forecasts for the EOA region.

REFERENCES

Bocchieri, J. R., and H. R. Glahn, 1976: Verification and further development of an operational model for forecasting the probability of frozen precipitation. Mon. Wea. Rev., 104, 691-701.

_____, 1977: Two cases of inaccurate probability of frozen precipitation (PoF) forecasts. TDL Office Note 77-3, National Weather Service, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, 14 pp.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. Mon. Wea. Rev., 78, 1-3.

National Weather Service, 1971: The Limited-area Fine Mesh (LFM) model. NWS Tech. Proc. Bull., No. 67, National Weather Service, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, 11 pp.

_____, 1976: Operational probability of frozen precipitation (PoF) forecasts based on model output statistics (MOS). NWS Tech. Proc. Bull., No. 170, National Weather Service, National Oceanic and Atmospheric Administration, U.S. Dept. of Commerce, 8 pp.

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

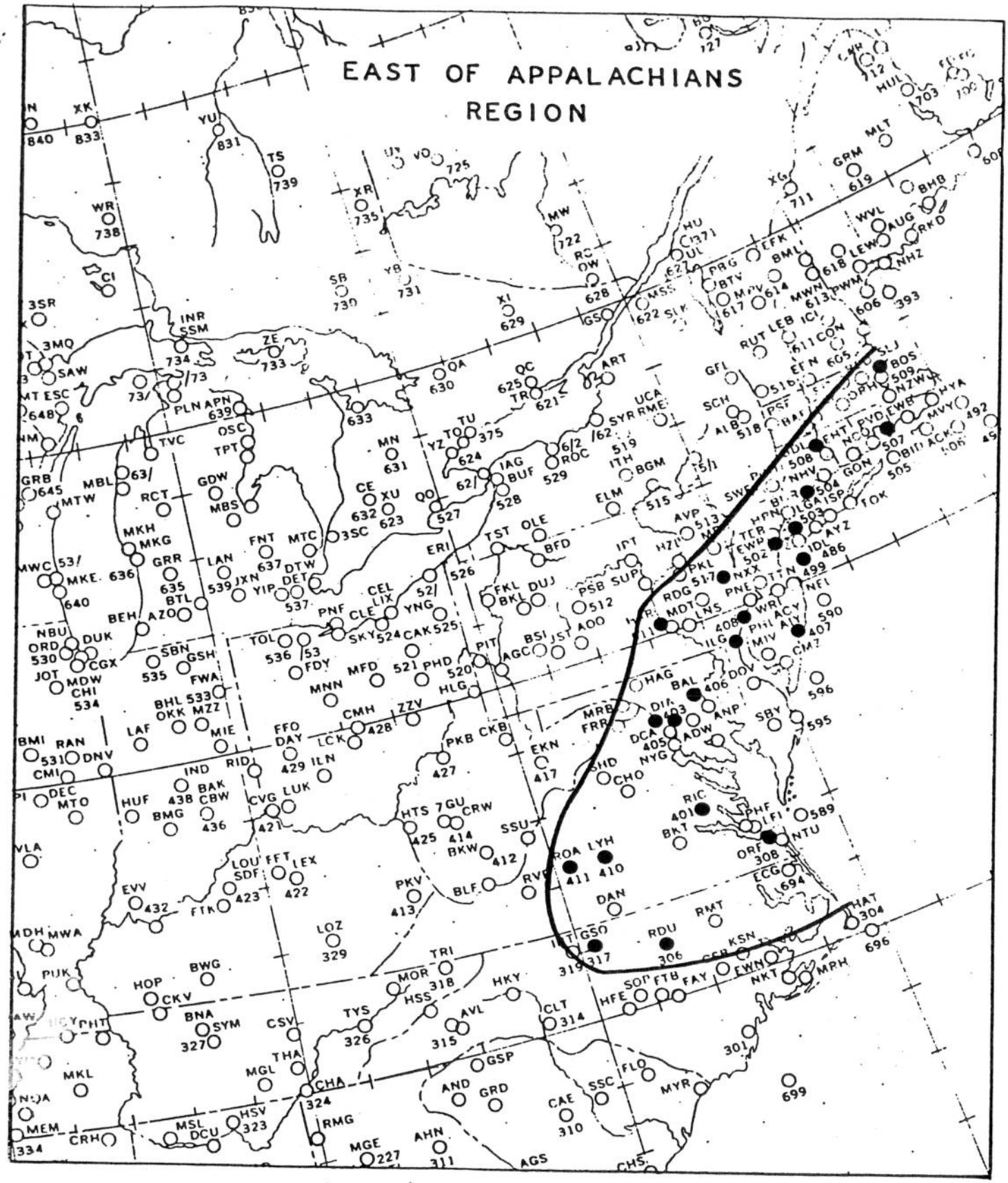

Figure 1. A depiction of the "East of Appalachians Region" (EOA Region). The dots show the 21 stations within the region that routinely receive PoF guidance.

Table 1. The bias of the early and final PoF guidance for the 12-and 24-hr forecast projections from the 0000 GMT cycle time, based on independent data for the period October 1976 through March 1977. Threshold probability was 50%.

Region	PoF System	Bias 12-hr Fcst.	Bias 24-hr Fcst.	Number of Cases
EOA	Early	.69	.70	400
	Final	.78	.80	
Rest of Eastern Region	Early	.99	1.00	1100
	Final	1.00	.99	
Whole Eastern Region	Early	.96	.96	1500
	Final	.98	.97	

Table 2. Comparative verification for the OPER, EXP GEN, and EXP REG10 PoF systems. Data were combined for 21 stations in the EOA region. The verification is shown for the developmental sample (≈1400 cases) and two independent data samples--1975-76 (≈550 cases) and 1976-77 (≈400 cases). Scores were computed for the 12- and 24-hr projections from the 0000 GMT cycle time. The threshold probability value was 50%. * Indicates best score.

PoF System	Developmental Data (1972-1975)			Independent Data (1975-1976)			Independent Data (1976-1977)		
	P Score	Bias	Threat Score	P Score	Bias	Threat Score	P Score	Bias	Threat Score
12-hr Forecast:									
OPER	.073	.74	.62	.083	.73	.63	.119	.69	.67
EXP GEN	.070	.77	.62	.079*	.79*	.67*	.092	.73	.70
EXP REG10	.062*	.86*	.71*	.081	.76	.63	.085*	.79*	.76*
24-hr Forecast:									
OPER	.088	.90	.55	.111	.69	.58	.150	.70	.54
EXP GEN	.092	.96*	.56	.104*	.79*	.59*	.136*	.82*	.63*
EXP REG10	.086*	.84	.57*	.110	.69	.53	.151	.73	.56

Table 3. Verification of the OPER and EXP GEN PoF systems for various threshold probabilities for the EOA region. Independent data, were combined for 21 stations from the 1975-76 and 1976-77 winter seasons. The projections are 12- and 24-hr from the 0000 GMT LFM cycle time.

Projection	Threshold Probability	Bias		Post-Agreement[1]		Threat Score	
		OPER	EXP GEN	OPER	EXP GEN	OPER	EXP GEN
12-hr	.50	.70	.75	.95	.95	.65	.69
	.45	.74	.80	.93	.94	.66	.72
	.40	.79	.85	.92	.92	.69	.74
	.35	.83	.91	.92	.91	.71	.77
	.30	.88	.94	.90	.90	.73	.77
	.25	.92	1.00	.90	.88	.77	.79
	.20	.98	1.05	.88	.86	.78	.79
	.15	1.04	1.14	.86	.82	.78	.78
24-hr	.50	.72	.83	.88	.85	.56	.63
	.45	.78	.88	.86	.84	.60	.65
	.40	.82	.92	.86	.83	.63	.66
	.35	.88	.98	.84	.83	.63	.69
	.30	.92	1.02	.83	.82	.66	.71
	.25	.99	1.09	.82	.80	.69	.71
	.20	1.10	1.17	.79	.78	.70	.73
	.15	1.18	1.25	.77	.74	.71	.69

[1] The post-agreement is the number of correct snow forecasts divided by the total number of snow forecasts.

Table 4. Recommended threshold probabilities for the early and final PoF guidance for the EOA region for the 1977-78 winter season. The values apply to both the 0000 GMT and 1200 GMT cycle times.

PoF SYSTEM	PROJECTIONS	THRESHOLD PROBABILITY (%)
Early Guidance	$\leq$ 12-hr	20
	> 12-hr	25
	$\leq$ 24-hr	25
Final Guidance	> 24-hr	30