NOAA TECHNICAL MEMORANDUM NWS CR-82

SOME PROPOSALS FOR MODIFYING THE PROBABILITY OF PRECIPITATION PROGRAM OF THE NATIONAL WEATHER SERVICE

Wayne E. Sangster and Michael D. Manker
National Weather Service Central Region
Scientific Services Division
Kansas City, Missouri

July 1986

UNITED STATES DEPARTMENT OF COMMERCE
Malcolm Baldrige, Secretary /
National Oceanic and Atmospheric Administration
Anthony Calio, Administrator /
National Weather Service
Richard E. Hallgren, Assistant Administrator

FOREWORD

This document is intended to stimulate discussion on the part of interested people, both within and outside of the National Weather Service. It is not official policy and represents the feelings of the authors who think that with the greater capabilities of the new numerical models we need to pass information on to the users in a more sophisticated manner than has been done in the past. The Probability of Precipitation program in the past has had its problems and this paper is an attempt to suggest how a smoother running system of more utility to the public could be attained.

TABLE OF CONTENTS

		Page No.
1.	INTRODUCTION	1
2.	REDUCING THE NUMBER OF PROBABILITY CLASSES	1
3.	SOME NEW STATISTICS	13
4.	A PROPOSED NEW SET OF EVENT-DEFINING THRESHOLDS	16
5.	WORDING OF FORECASTS IN THE NEW SYSTEM	25
6.	SUMMARY	28
7.	ACKNOWLEDGEMENTS	29
8.	REFERENCES	29

SOME PROPOSALS FOR MODIFYING THE PROBABILITY OF PRECIPITATION PROGRAM OF THE NATIONAL WEATHER SERVICE

Wayne E. Sangster and Michael D. Manker
National Weather Service Central Region
Scientific Services Division
Kansas City, Missouri 64106

1. INTRODUCTION

The senior author has been closely associated with the Probability of Precipitation (PoP) program of the National Weather Service (NWS) since its inception two decades ago. The ideas expressed and experiments discussed herein are an outgrowth of those years of experience. During these 20 years the basics of the program have remained unchanged, although variations in the methods of incorporating the PoP numbers into the worded forecasts have occurred. We think it is time to think about changing some of our ways of doing things in order for the program to be more effective and pertinent.

Verification of PoP forecasts is especially important since feedback on how well a forecaster is doing is needed so that future forecasts can be based on past experience. For example, a 40 percent PoP is both right and wrong regardless of whether or not .01 inch or more of precipitation fell in the gauge. It takes many 40 percent PoP's by the same forecaster for the same period, etc., for a verification system to show how things are going.

In this technical memorandum we first propose reducing the number of probability classes from 12 (or 13) to six in order to allow a new kind of verification display (and also to allow more meaningful verification). Next some new scores designed to put more emphasis on catching the larger amounts with high PoP's will be presented. Later we will explore the possibilities of changing the threshold for the probability event from the present .01 inch to three thresholds of Trace, .04 inch, and .40 inch.

An experiment was conducted to illustrate the validity of these concepts. It consisted of making subjective probability forecasts of several sorts based on the RAFS (Regional Analysis and Forecast System) model output and comparing them with objective MOS forecasts based on the LFM (Limited Fine Mesh) model.

2. REDUCING THE NUMBER OF PROBABILITY CLASSES

Typical PoP verification systems present "reliability" numbers giving the observed "rain" frequency for each PoP value for both guidance and the public forecast. This allows some judgment to be made as to how to modify forecasts in the future. Brier scores are also given to assess the relative quality of the two sets of forecasts. What is missing, however, is the joint performance of the two sets. What, for example, was the relative frequency of precipitation when guidance was 30 percent and the forecast was 50 percent? Also, who had the better Brier score for these forecasts?

It is straightforward to construct tables showing this information when there are 12 classes of PoP's, but the number of cells becomes quite large (144). Clemen and Murphy (1985) actually have done this for 13 classes of objective forecasts and 11 classes of subjective forecasts in spite of the unwieldy nature of the tables. They refer to them as joint calibration functions. Believing that simplicity has virtues, we chose to use only six classes of probabilities for both types of forecasts — namely 0, 10, 30, 50, 70, and 90 percent. This reduces the number of cells to 36,[1] but as will be shown, does not significantly affect the information content of the forecasts.

As a feasibility study we ran an experiment in which only the six classes were used. The forecasts were for 30 Central Region stations divided into seven subareas as shown in Fig. 1. Two sets of numbers were generated as follows:

(1) Model Output Statistics (MOS) PoP's from the Limited Fine Mesh (LFM or Early) model.

(2) Subjective PoP's made by the senior author based on the Regional Analysis and Forecast System (RAFS) output.

We shall refer to these as MOS and RAFS forecasts (M and R).

A fair question to ask is how MOS PoP forecasts for six classes were obtained, since there are 13 in the data received from the National Meteorological Center, and there is no obvious way to divide 40 percent forecasts, for example, into 30 and 50 percent classes from the digital data. This problem was solved simply using the graphical output for PoP forecasts sent on AFOS and facsimile. The graphical output is not analyzed perfectly for the station data, so occasional numbers had to be changed to be consistent with the digital output. The subjective PoP's were, of course, made using only six classes.

In Table 1 are shown results from a sample of forecasts made on 62 days from the 1200 GMT runs for the first period (TONIGHT) during the cool season of 1985-86. The numbers in each cell are the number of "rains," the number of forecasts, the relative "rain" frequency, the difference (MOS minus RAFS) in "Total Brier Score" (before dividing by the number of forecasts), the Total Brier Score for MOS, and the Total Brier Score for RAFS, according to the key on the bottom of the table. Although they are not shown here, each projection and each cycle (0000 or 1200 GMT) would have its own table in actual practice.

It is of interest to note that 13 cells had no forecasts, so the number of cells with forecasts is only 23. Additionally, the six-cells on the

[1] In actuality fewer than 36 cells will be likely to have entries in them since, for example, a guidance of zero percent and a forecast of 90 percent is unlikely.

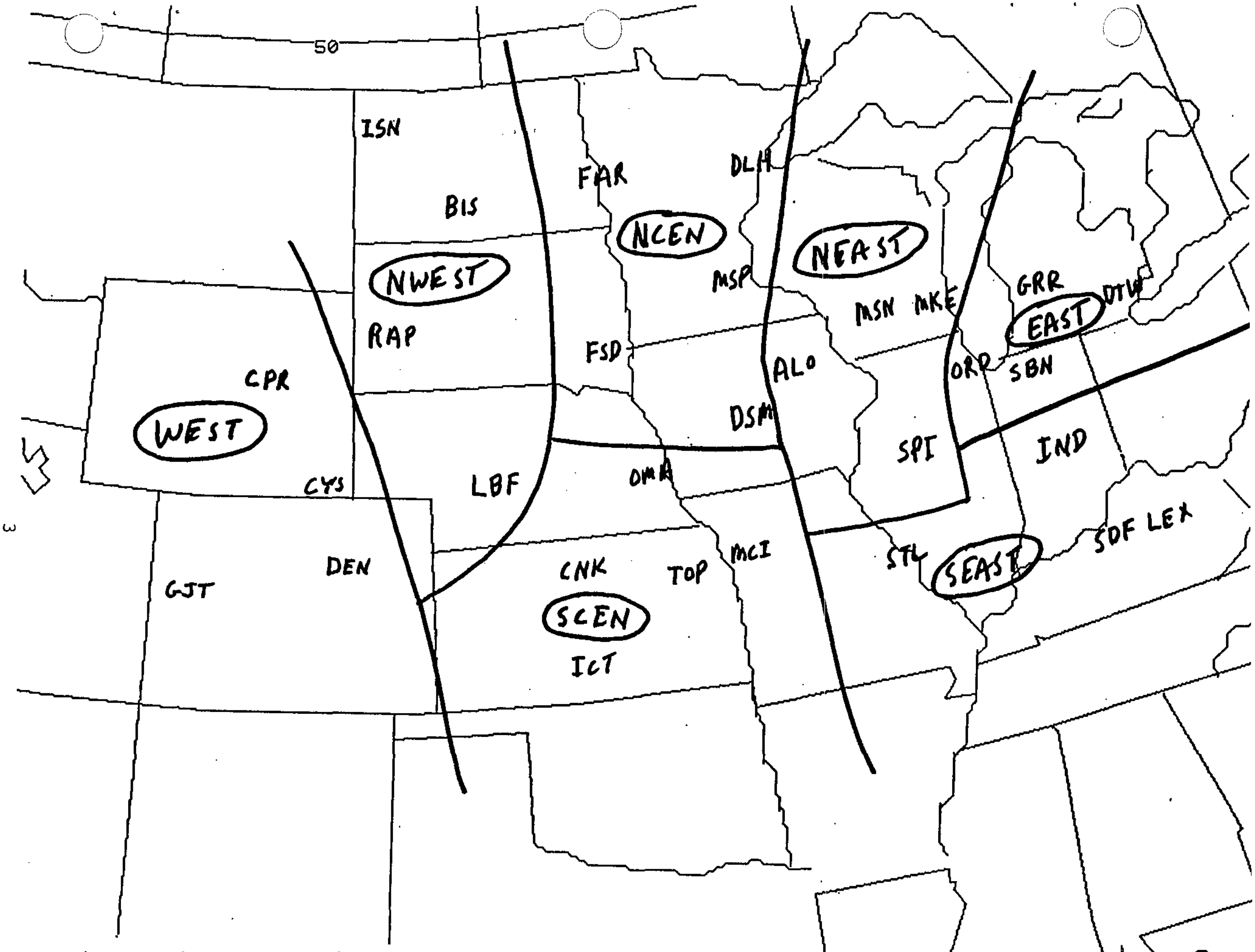

Fig. 1. Subareas and stations used in the 30-station study.

Table 1. RAFS vs. MOS verification statistics for the cool season 1985-86. Key to numbers is shown at the bottom.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		00	10	30	50	70	90	ALL
	00	7/ 496	3/ 77	3/ 21	1/ 1	0/ 0	0/ 0	14/ 595
		1.4/ 0.	3.9/ -17.	14.3/ -9.	100.0/ 75.	.0/ 0.	.0/ 0.	2.4/ 49.
		700./ 700.	300./ 317.	300./ 309.	100./ 25.	0./ 0.	0./ 0.	1400./ 1351.
	10	6/ 29	27/ 388	29/ 140	3/ 8	0/ 0	0/ 0	65/ 565
		20.7/ -91.	7.0/ 0.	20.7/ 40.	37.5/ 48.	.0/ 0.	.0/ 0.	11.5/ -3.
M		509./ 600.	2548./ 2548.	2460./ 2420.	248./ 200.	0./ 0.	0./ 0.	5765./ 5768.
	30	0/ 0	5/ 35	72/ 197	20/ 38	2/ 3	0/ 0	99/ 273
		.0/ 0.	14.3/ 80.	36.5/ 0.	52.6/ 192.	66.7/ 40.	.0/ 0.	36.3/ 312.
		0./ 0.	515./ 435.	4653./ 4653.	1142./ 950.	107./ 67.	0./ 0.	6417./ 6105.
	50	0/ 0	0/ 0	6/ 24	89/ 130	16/ 20	2/ 2	113/ 176
		.0/ 0.	.0/ 0.	25.0/ 144.	68.5/ 0.	80.0/ 160.	100.0/ 48.	64.2/ 352.
O		0./ 0.	0./ 0.	600./ 456.	3250./ 3250.	500./ 340.	50./ 2.	4400./ 4048.
	70	0/ 0	0/ 0	3/ 6	13/ 25	69/ 80	19/ 20	104/ 131
		.0/ 0.	.0/ 0.	50.0/ 0.	52.0/ 80.	86.2/ 0.	95.0/ 120.	79.4/ 200.
		0./ 0.	0./ 0.	174./ 174.	705./ 625.	1160./ 1160.	220./ 100.	2259./ 2059.
	90	0/ 0	0/ 0	0/ 0	1/ 5	16/ 18	79/ 88	96/ 111
		.0/ 0.	.0/ 0.	.0/ 0.	20.0/ 200.	88.9/ -64.	89.8/ 0.	86.5/ 136.
S		0./ 0.	0./ 0.	0./ 0.	325./ 125.	178./ 242.	808./ 808.	1311./ 1175.
	ALL	13/ 525	35/ 500	113/ 388	127/ 207	103/ 121	100/ 110	491/ 1851
		2.5/ -91.	7.0/ 63.	29.1/ 175.	61.4/ 595.	85.1/ 136.	90.9/ 168.	26.5/ 1046.
		1209./ 1300.	3363./ 3300.	8187./ 8012.	5770./ 5175.	1945./ 1809.	1078./ 910.	21552./ 20506.

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.

RR / NN
RF / ITBS
MTBS / RTBS

diagonal have identical scores for MOS and RAFS, so only 17 cells are of interest in this case with regard to relative performance of the two forecast systems.

The largest gain by RAFS over MOS was in the cell (M = 90, R = 50) where the rain frequency was 20 percent (one out of five). It is obvious that the lower PoP would be superior and it is, by 200 points (these numbers are 100 times the usual Brier Score used in the NWS). At the other extreme, the cell (M = 10, R = 0) had an observed rain frequency of 20.7 percent so RAFS had a loss of 91 points.

Other places where RAFS improved are in cell (M = 30, R = 50) (192 points) and cell (M = 50, R = 70) (160 points). In both of these cases RAFS improved the reliability of the forecasts by using a higher PoP. A case where RAFS used a lower PoP and thereby got a better score is cell (M = 50, R = 30) (144 points). On cell (M = 90, R = 70) RAFS should not have used a lower PoP since the relative frequency was 88.9 percent. It lost 64 points here. As an aside, note that for cell (M = 90, R = 90) the relative frequency was 89.8 percent — virtual perfection.

Examination of cell (M = A, R = 50) reveals that the 50 percent RAFS PoP contributed 595 points to the total of 1046 points gained by RAFS overall (M = A, R = A). All RAFS forecasts of 50 percent (except of course on the diagonal) showed improvement over the MOS forecasts at all MOS values when taken in toto. Cell (M = 50, R = A) showed a total improvement for RAFS of 352 points, about a third of the total, so RAFS improved things when MOS was on the fence. It is apparent that tables such as these can provide much more insight into problems and strengths than can the usual separate reliability numbers or diagrams.

In Tables 2a-g are breakdowns of the forecasts by seven subareas shown in Fig. 1. To further examine the large gain of RAFS over MOS for (M = 90, R = 50) we look at Table 2f and see that more than half (112 points) of the 200-point total improvement by RAFS were in the SCEN subarea. The NCEN and NEAST subareas also contributed positively to the improvement. Since these three cells only account for three forecasts, a great amount of significance cannot be placed in the results. Looking at Table 2b for (M = 10, R = 0), however, there were nine forecasts in the SCEN subarea which accounted for much of the loss of points by RAFS. Five of nine forecasts of zero having "rain" is clearly out of line.

In Table 2c the (M = 30, R = 50) column shows that the WEST, NWEST, NCEN, and NEAST subareas all had RF's of 50 percent or greater, leading to an improvement by RAFS. The NCEN and SCEN subareas in Table 2d for column (M = 50, R = 30) had a combined record of two "rains" out of 12 forecasts, so the lowering of the PoP by RAFS helped account for most of the 144 points gained by RAFS overall. Table 2f under the (M = 90, R = 70) column shows that four of the seven subareas lost points. The combined RF for these subareas was 15 of 16 (93.8 percent), so the lowering of MOS 90 percent values was clearly a mistake in these locations.

Table 2a. RAFS vs. MOS verification statistics for the cool season 1985-86. For forecasts when MOS was zero percent.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		M=00,R=00		M=00,R=10		M=00,R=30		M=00,R=50		M=00,R=70		M=00,R=90		M=00,R=A		
	WEST	0/	77	0/	28	1/	10	1/	1	0/	0	0/	0	2/	116	
		.0/	0.	.0/	-28.	10.0/	-30.	100.0/	75.	.0/	0.	.0/	0.	1.7/	17.	5.4
		0./	0.	0./	28.	100./	130.	100./	25.	0./	0.	0./	0.	200./	183.	
	NWEST	2/	85	0/	10	1/	5	0/	0	0/	0	0/	0	3/	100	
		2.4/	0.	.0/	-10.	20.0/	15.	.0/	0.	.0/	0.	.0/	0.	3.0/	5.	2.5
M		200./	200.	0./	10.	100./	85.	0./	0.	0./	0.	0./	0.	300./	295.	
	NCEN	0/	81	0/	10	1/	2	0/	0	0/	0	0/	0	1/	93	
		.0/	0.	.0/	-10.	50.0/	42.	.0/	0.	.0/	0.	.0/	0.	1.1/	32.	1.7
		0./	0.	0./	10.	100./	58.	0./	0.	0./	0.	0./	0.	100./	68.	
	SCEN	0/	88	3/	11	0/	0	0/	0	0/	0	0/	0	3/	99	
		.0/	0.	27.3/	49.	.0/	0.	.0/	0.	.0/	0.	.0/	0.	3.0/	49.	1.1
O		0./	0.	300./	251.	0./	0.	0./	0.	0./	0.	0./	0.	300./	251.	
	NEAST	2/	59	0/	11	0/	0	0/	0	0/	0	0/	0	2/	70	
		3.4/	0.	.0/	-11.	.0/	0.	.0/	0.	.0/	0.	.0/	0.	2.9/	-11.	1.6
		200./	200.	0./	11.	0./	0.	0./	0.	0./	0.	0./	0.	200./	211.	
	EAST	3/	55	0/	3	0/	2	0/	0	0/	0	0/	0	3/	60	
		5.5/	0.	.0/	-3.	.0/	-18.	.0/	0.	.0/	0.	.0/	0.	5.0/	-21.	1.5
S		300./	300.	0./	3.	0./	18.	0./	0.	0./	0.	0./	0.	300./	321.	
	SEAST	0/	51	0/	4	0/	2	0/	0	0/	0	0/	0	0/	57	
		.0/	0.	.0/	-4.	.0/	-18.	.0/	0.	.0/	0.	.0/	0.	.0/	-22.	1.8
		0./	0.	0./	4.	0./	18.	0./	0.	0./	0.	0./	0.	0./	22.	
	ALL	7/	496	3/	77	3/	21	1/	1	0/	0	0/	0	14/	595	
		1.4/	0.	3.9/	-17.	14.3/	-9.	100.0/	75.	.0/	0.	.0/	0.	2.4/	49.	2.4
		700./	700.	300./	317.	300./	309.	100./	25.	0./	0.	0./	0.	1400./	1351.	

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.
MRNF = MOS FCST. RAIN FREQ.

RR / NN
RF / ITBS MRNF
MTBS / RTBS

Table 2b. Same as for Table 2a, except for MOS forecasts of 10 percent.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		M=10,R=00	M=10,R=10	M=10,R=30	M=10,R=50	M=10,R=70	M=10,R=90	M=10,R=A	
	WEST	0/ 0	4/ 51	2/ 16	0/ 1	0/ 0	0/ 0	6/ 68	
		.0/ 0.	7.8/ 0.	12.5/ -48.	.0/ -24.	.0/ 0.	.0/ 0.	8.8/ -72.	15.3
		0./ 0.	371./ 371.	176./ 224.	1./ 25.	0./ 0.	0./ 0.	548./ 620.	
	NWEST	0/ 2	3/ 50	6/ 18	0/ 1	0/ 0	0/ 0	9/ 71	
		.0/ 2.	6.0/ 0.	33.3/ 96.	.0/ -24.	.0/ 0.	.0/ 0.	12.7/ 74.	15.4
M		2./ 0.	290./ 290.	498./ 402.	1./ 25.	0./ 0.	0./ 0.	791./ 717.	
	NCEN	1/ 7	5/ 76	10/ 34	1/ 1	0/ 0	0/ 0	17/ 118	
		14.3/ -13.	6.6/ 0.	29.4/ 128.	100.0/ 56.	.0/ 0.	.0/ 0.	14.4/ 171.	15.5
		87./ 100.	476./ 476.	834./ 706.	81./ 25.	0./ 0.	0./ 0.	1478./ 1307.	
	SCEN	0/ 7	3/ 84	4/ 25	0/ 0	0/ 0	0/ 0	7/ 116	
		.0/ 7.	3.6/ 0.	16.0/ -40.	.0/ 0.	.0/ 0.	.0/ 0.	6.0/ -33.	13.7
O		7./ 0.	324./ 324.	345./ 385.	0./ 0.	0./ 0.	0./ 0.	676./ 709.	
	NEAST	0/ 2	5/ 43	6/ 15	0/ 0	0/ 0	0/ 0	11/ 60	
		.0/ 2.	11.6/ 0.	40.0/ 120.	.0/ 0.	.0/ 0.	.0/ 0.	18.3/ 122.	14.7
		2./ 0.	443./ 443.	495./ 375.	0./ 0.	0./ 0.	0./ 0.	940./ 818.	
	EAST	5/ 9	5/ 37	1/ 11	1/ 2	0/ 0	0/ 0	12/ 59	
		55.6/ -91.	13.5/ 0.	9.1/ -48.	50.0/ 32.	.0/ 0.	.0/ 0.	20.3/ -107.	13.6
S		409./ 500.	437./ 437.	91./ 139.	82./ 50.	0./ 0.	0./ 0.	1019./ 1126.	
	SEAST	0/ 2	2/ 47	0/ 21	1/ 3	0/ 0	0/ 0	3/ 73	
		.0/ 2.	4.3/ 0.	.0/ -168.	33.3/ 8.	.0/ 0.	.0/ 0.	4.1/ -158.	17.1
		2./ 0.	207./ 207.	21./ 189.	83./ 75.	0./ 0.	0./ 0.	313./ 471.	
	ALL	6/ 29	27/ 388	29/ 140	3/ 8	0/ 0	0/ 0	65/ 565	
		20.7/ -91.	7.0/ 0.	20.7/ 40.	37.5/ 48.	.0/ 0.	.0/ 0.	11.5/ -3.	15.0
		509./ 600.	2548./ 2548.	2460./ 2420.	248./ 200.	0./ 0.	0./ 0.	5765./ 5768.	

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.
MRNF = MOS FCST. RAIN FREQ.

RR / NN
RF / ITBS MRNF
MTBS / RTBS

Table 2c. Same as for Table 2a, except for MOS forecasts of 30 percent.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		M=30,R=00		M=30,R=10		M=30,R=30		M=30,R=50		M=30,R=70		M=30,R=90		M=30,R=A		
	WEST	0/	0	0/	3	7/	17	3/	4	0/	0	0/	0	10/	24	
		.0/	0.	.0/	24.	41.2/	0.	75.0/	56.	.0/	0.	.0/	0.	41.7/	80.	30.8
		0./	0.	27./	3.	433./	433.	156./	100.	0./	0.	0./	0.	616./	536.	
	NWEST	0/	0	0/	4	12/	30	5/	10	1/	1	0/	0	18/	45	
		.0/	0.	.0/	32.	40.0/	0.	50.0/	40.	100.0/	40.	.0/	0.	40.0/	112.	33.6
M		0./	0.	36./	4.	750./	750.	290./	250.	49./	9.	0./	0.	1125./	1013.	
	NCEN	0/	0	3/	7	8/	34	2/	3	0/	0	0/	0	13/	44	
		.0/	0.	42.9/	-64.	23.5/	0.	66.7/	32.	.0/	0.	.0/	0.	29.5/	-32.	28.2
		0./	0.	183./	247.	626./	626.	107./	75.	0./	0.	0./	0.	916./	948.	
	SCEN	0/	0	1/	4	6/	27	0/	1	0/	0	0/	0	7/	32	
		.0/	0.	25.0/	-8.	22.2/	0.	.0/	-16.	.0/	0.	.0/	0.	21.9/	-24.	28.1
O		0./	0.	76./	84.	483./	483.	9./	25.	0./	0.	0./	0.	568./	592.	
	NEAST	0/	0	0/	2	10/	25	7/	10	0/	1	0/	0	17/	38	
		.0/	0.	.0/	16.	40.0/	0.	70.0/	120.	.0/	-40.	.0/	0.	44.7/	96.	35.3
		0./	0.	18./	2.	625./	625.	370./	250.	9./	49.	0./	0.	1022./	926.	
	EAST	0/	0	0/	3	19/	33	2/	6	0/	0	0/	0	21/	42	
		.0/	0.	.0/	24.	57.6/	0.	33.3/	-16.	.0/	0.	.0/	0.	50.0/	8.	31.4
S		0./	0.	27./	3.	1057./	1057.	134./	150.	0./	0.	0./	0.	1218./	1210.	
	SEAST	0/	0	1/	12	10/	31	1/	4	1/	1	0/	0	13/	48	
		.0/	0.	8.3/	56.	32.3/	0.	25.0/	-24.	100.0/	40.	.0/	0.	27.1/	72.	27.5
		0./	0.	148./	92.	679./	679.	76./	100.	49./	9.	0./	0.	952./	880.	
	ALL	0/	0	5/	35	72/	197	20/	38	2/	3	0/	0	99/	273	
		.0/	0.	14.3/	80.	36.5/	0.	52.6/	192.	66.7/	40.	.0/	0.	36.3/	312.	30.7
		0./	0.	515./	435.	4653./	4653.	1142./	950.	107./	67.	0./	0.	6417./	6105.	

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.
MRNF = MOS FCST. RAIN FREQ.

RR / NN		
RF / ITBS		MRNF
MTBS / RTBS		

Table 2d. Same as for Table 2a, except for MOS forecasts of 50 percent.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		M=50,R=00		M=50,R=10		M=50,R=30		M=50,R=50		M=50,R=70		M=50,R=90		M=50,R=A		
	WEST	0/	0	0/	0	1/	3	13/	20	2/	2	0/	0	16/	25	
		.0/	0.	.0/	0.	33.3/	8.	65.0/	0.	100.0/	32.	.0/	0.	64.0/	40.	49.2
		0./	0.	0./	0.	75./	67.	500./	500.	50./	18.	0./	0.	625./	585.	
	NWEST	0/	0	0/	0	0/	1	11/	16	0/	0	0/	0	11/	17	
		.0/	0.	.0/	0.	.0/	16.	68.7/	0.	.0/	0.	.0/	0.	64.7/	16.	48.8
M		0./	0.	0./	0.	25./	9.	400./	400.	0./	0.	0./	0.	425./	409.	
	NCEN	0/	0	0/	0	0/	4	13/	18	3/	3	1/	1	17/	26	
		.0/	0.	.0/	0.	.0/	64.	72.2/	0.	100.0/	48.	100.0/	24.	65.4/	136.	50.8
		0./	0.	0./	0.	100./	36.	450./	450.	75./	27.	25./	1.	650./	514.	
	SCEN	0/	0	0/	0	2/	8	4/	10	1/	3	0/	0	7/	21	
		.0/	0.	.0/	0.	25.0/	48.	40.0/	0.	33.3/	-32.	.0/	0.	33.3/	16.	45.2
O		0./	0.	0./	0.	200./	152.	250./	250.	75./	107.	0./	0.	525./	509.	
	NEAST	0/	0	0/	0	2/	6	18/	23	5/	6	0/	0	25/	35	
		.0/	0.	.0/	0.	33.3/	16.	78.3/	0.	83.3/	56.	.0/	0.	71.4/	72.	50.0
		0./	0.	0./	0.	150./	134.	575./	575.	150./	94.	0./	0.	875./	803.	
	EAST	0/	0	0/	0	1/	1	19/	25	3/	4	0/	0	23/	30	
		.0/	0.	.0/	0.	100.0/	-24.	76.0/	0.	75.0/	24.	.0/	0.	76.7/	0.	52.0
S		0./	0.	0./	0.	25./	49.	625./	625.	100./	76.	0./	0.	750./	750.	
	SEAST	0/	0	0/	0	0/	1	11/	18	2/	2	1/	1	14/	22	
		.0/	0.	.0/	0.	.0/	16.	61.1/	0.	100.0/	32.	100.0/	24.	63.6/	72.	52.7
		0./	0.	0./	0.	25./	9.	450./	450.	50./	18.	25./	1.	550./	478.	
	ALL	0/	0	0/	0	6/	24	89/	130	16/	20	2/	2	113/	176	
		.0/	0.	.0/	0.	25.0/	144.	68.5/	0.	80.0/	160.	100.0/	48.	64.2/	352.	50.0
		0./	0.	0./	0.	600./	456.	3250./	3250.	500./	340.	50./	2.	4400./	4048.	

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.
MRNF = MOS FCST. RAIN FREQ.

RR / NN
RF / ITBS MRNF
MTBS / RTBS

Table 2e. Same as for Table 2a, except for MOS forecasts of 70 percent.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		M=70,R=00		M=70,R=10		M=70,R=30		M=70,R=50		M=70,R=70		M=70,R=90		M=70,R=A		
	WEST	0/	0	0/	0	0/	0	1/	1	8/	10	0/	0	9/	11	
		.0/	0.	.0/	0.	.0/	0.	100.0/	-16.	80.0/	0.	.0/	0.	81.8/	-16.	68.2
		0./	0.	0./	0.	0./	0.	9./	25.	170./	170.	0./	0.	179./	195.	
	NWEST	0/	0	0/	0	0/	0	2/	3	8/	9	0/	0	10/	12	
		.0/	0.	.0/	0.	.0/	0.	66.7/	-8.	88.9/	0.	.0/	0.	83.3/	-8.	65.0
M		0./	0.	0./	0.	0./	0.	67./	75.	121./	121.	0./	0.	188./	196.	
	NCEN	0/	0	0/	0	0/	0	2/	4	10/	11	4/	4	16/	19	
		.0/	0.	.0/	0.	.0/	0.	50.0/	16.	90.9/	0.	100.0/	32.	84.2/	48.	70.0
		0./	0.	0./	0.	0./	0.	116./	100.	139./	139.	36./	4.	291./	243.	
	SCEN	0/	0	0/	0	1/	2	1/	7	7/	7	3/	4	12/	20	
		.0/	0.	.0/	0.	50.0/	0.	14.3/	128.	100.0/	0.	75.0/	-8.	60.0/	120.	63.0
O		0./	0.	0./	0.	58./	58.	303./	175.	63./	63.	76./	84.	500./	380.	
	NEAST	0/	0	0/	0	1/	1	3/	4	9/	11	6/	6	19/	22	
		.0/	0.	.0/	0.	100.0/	-40.	75.0/	-24.	81.8/	0.	100.0/	48.	86.4/	-16.	70.0
		0./	0.	0./	0.	9./	49.	76./	100.	179./	179.	54./	6.	318./	334.	
	EAST	0/	0	0/	0	1/	1	2/	3	15/	16	3/	3	21/	23	
		.0/	0.	.0/	0.	100.0/	-40.	66.7/	-8.	93.7/	0.	100.0/	24.	91.3/	-24.	68.3
S		0./	0.	0./	0.	9./	49.	67./	75.	184./	184.	27./	3.	287./	311.	
	SEAST	0/	0	0/	0	0/	2	2/	3	12/	16	3/	3	17/	24	
		.0/	0.	.0/	0.	.0/	80.	66.7/	-8.	75.0/	0.	100.0/	24.	70.8/	96.	66.7
		0./	0.	0./	0.	98./	18.	67./	75.	304./	304.	27./	3.	496./	400.	
	ALL	0/	0	0/	0	3/	6	13/	25	69/	80	19/	20	104/	131	
		.0/	0.	.0/	0.	50.0/	0.	52.0/	80.	86.2/	0.	95.0/	120.	79.4/	200.	67.4
		0./	0.	0./	0.	174./	174.	705./	625.	1160./	1160.	220./	100.	2259./	2059.	

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.
MRNF = MOS FCST. RAIN FREQ.

RR / NN
RF / ITBS MRNF
MTBS / RTBS

Table 2f. Same as for Table 2a, except for MOS forecasts of 90 percent.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		M=90,R=00		M=90,R=10		M=90,R=30		M=90,R=50		M=90,R=70		M=90,R=90		M=90,R=A		
	WEST	0/	0	0/	0	0/	0	0/	0	0/	0	0/	0	0/	0	
		.0/	0.	.0/	0.	.0/	0.	.0/	0.	.0/	0.	.0/	0.	.0/	0.	.0
		0./	0.	0./	0.	0./	0.	0./	0.	0./	0.	0./	0.	0./	0.	
	NWEST	0/	0	0/	0	0/	0	0/	0	0/	0	2/	3	2/	3	
		.0/	0.	.0/	0.	.0/	0.	.0/	0.	.0/	0.	66.7/	0.	66.7/	0.	90.0
M		0./	0.	0./	0.	0./	0.	0./	0.	0./	0.	83./	83.	83./	83.	
	NCEN	0/	0	0/	0	0/	0	0/	1	2/	2	6/	7	8/	10	
		.0/	0.	.0/	0.	.0/	0.	.0/	56.	100.0/	-16.	85.7/	0.	80.0/	40.	82.0
		0./	0.	0./	0.	0./	0.	81./	25.	2./	18.	87./	87.	170./	130.	
	SCEN	0/	0	0/	0	0/	0	0/	2	3/	3	12/	15	15/	20	
		.0/	0.	.0/	0.	.0/	0.	.0/	112.	100.0/	-24.	80.0/	0.	75.0/	88.	83.0
O		0./	0.	0./	0.	0./	0.	162./	50.	3./	27.	255./	255.	420./	332.	
	NEAST	0/	0	0/	0	0/	0	0/	1	4/	4	15/	17	19/	22	
		.0/	0.	.0/	0.	.0/	0.	.0/	56.	100.0/	-32.	88.2/	0.	86.4/	24.	84.5
		0./	0.	0./	0.	0./	0.	81./	25.	4./	36.	177./	177.	262./	238.	
	EAST	0/	0	0/	0	0/	0	0/	0	6/	7	25/	25	31/	32	
		.0/	0.	.0/	0.	.0/	0.	.0/	0.	85.7/	-16.	100.0/	0.	96.9/	-16.	85.6
S		0./	0.	0./	0.	0./	0.	0./	0.	87./	103.	25./	25.	112./	128.	
	SEAST	0/	0	0/	0	0/	0	1/	1	1/	2	19/	21	21/	24	
		.0/	0.	.0/	0.	.0/	0.	100.0/	-24.	50.0/	24.	90.5/	0.	87.5/	0.	86.7
		0./	0.	0./	0.	0./	0.	1./	25.	82./	58.	181./	181.	264./	264.	
	ALL	0/	0	0/	0	0/	0	1/	5	16/	18	79/	88	96/	111	
		.0/	0.	.0/	0.	.0/	0.	20.0/	200.	88.9/	-64.	89.8/	0.	86.5/	136.	85.0
		0./	0.	0./	0.	0./	0.	325./	125.	178./	242.	808./	808.	1311./	1175.	

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.
MRNF = MOS FCST. RAIN FREQ.

RR / NN
RF / ITBS MRNF
MTBS / RTBS

Table 2g. Same as for Table 2a, except for all MOS forecasts.

RAFS VS. MOS POP VERIFICATION STATISTICS FOR COOL SEASON 1985-86
TONIGHT
R A F S

		M=A ,R=00	M=A ,R=10	M=A ,R=30	M=A ,R=50	M=A ,R=70	M=A ,R=90	M=A ,R=A		
	WEST	0/ 77	4/ 82	11/ 46	18/ 27	10/ 12	0/ 0	43/ 244	38.8	
		.0/ 0.	4.9/ -4.	23.9/ -70.	66.7/ 91.	83.3/ 32.	.0/ 0.	17.6/ 49.	2.3	18.0
		0./ 0.	398./ 402.	784./ 854.	766./ 675.	220./ 188.	0./ 0.	2168./ 2119.	.00201	
		.0	7.3	17.8	44.4	66.7	.0	14.0		
	NWEST	2/ 87	3/ 64	19/ 54	18/ 30	9/ 10	2/ 3	53/ 248	30.1	
		2.3/ 2.	4.7/ 22.	35.2/ 127.	60.0/ 8.	90.0/ 40.	66.7/ 0.	21.4/ 199.	6.8	19.1
M		202./ 200.	326./ 304.	1373./ 1246.	758./ 750.	170./ 130.	83./ 83.	2912./ 2713.	.00803	
		.2	9.7	20.9	44.0	66.0	90.0	16.2		
	NCEN	1/ 88	8/ 93	19/ 74	18/ 27	15/ 16	11/ 12	72/ 310	34.8	
		1.1/ -13.	8.6/ -74.	25.7/ 234.	66.7/ 160.	93.7/ 32.	91.7/ 56.	23.2/ 395.	11.0	21.6
		87./ 100.	659./ 733.	1660./ 1426.	835./ 675.	216./ 184.	148./ 92.	3605./ 3210.	.01275	
		.8	10.4	21.1	50.7	68.7	80.0	19.5		
	SCEN	0/ 95	7/ 99	13/ 62	5/ 20	11/ 13	15/ 19	51/ 308	29.8	
		.0/ 7.	7.1/ 41.	21.0/ 8.	25.0/ 224.	84.6/ -56.	78.9/ -8.	16.6/ 216.	7.2	21.0
O		7./ 0.	700./ 659.	1086./ 1078.	724./ 500.	141./ 197.	331./ 339.	2989./ 2773.	.00702	
		.7	9.7	25.8	60.0	70.0	85.8	20.7		
	NEAST	2/ 61	5/ 56	19/ 47	28/ 38	18/ 22	21/ 23	93/ 247	37.6	
		3.3/ 2.	8.9/ 5.	40.4/ 96.	73.7/ 152.	81.8/ -16.	91.3/ 48.	37.7/ 287.	7.9	30.3
		202./ 200.	461./ 456.	1279./ 1183.	1102./ 950.	342./ 358.	231./ 183.	3617./ 3330.	.01162	
		.3	8.8	27.0	47.9	66.4	84.8	28.4		
	EAST	8/ 64	5/ 43	22/ 48	24/ 36	24/ 27	28/ 28	111/ 246	39.5	
		12.5/ -91.	11.6/ 21.	45.8/ -130.	66.7/ 8.	88.9/ 8.	100.0/ 24.	45.1/ -160.	-4.3	32.8
S		709./ 800.	464./ 443.	1182./ 1312.	908./ 900.	371./ 363.	52./ 28.	3686./ 3846.	-.00650	
		1.4	10.7	25.4	46.1	72.2	87.9	31.9		
	SEAST	0/ 53	3/ 63	10/ 57	16/ 29	16/ 21	23/ 25	68/ 248	47.8	
		.0/ 2.	4.8/ 52.	17.5/ -90.	55.2/ -48.	76.2/ 96.	92.0/ 48.	27.4/ 60.	2.3	30.3
		2./ 0.	355./ 303.	823./ 913.	677./ 725.	485./ 389.	233./ 185.	2575./ 2515.	.00242	
		.4	13.2	23.3	46.6	68.1	86.0	28.7		
	ALL	13/ 525	35/ 500	113/ 388	127/ 207	103/ 121	100/ 110	491/ 1851	40.3	
		2.5/ -91.	7.0/ 63.	29.1/ 175.	61.4/ 595.	85.1/ 136.	90.9/ 168.	26.5/ 1046.	4.9	24.5
		1209./ 1300.	3363./ 3300.	8187./ 8012.	5770./ 5175.	1945./ 1809.	1078./ 910.	21552./ 20506.	.00566	
		.6	9.9	23.0	47.9	68.7	85.6	22.6		

KEY

RR = PRECIP. FREQ.
NN = TOTAL NO. OF FORECASTS.
RF = REL. FREQ. (PCT) OF PRECIP.
ITBS = IMPR. IN TOTAL BRIER SCORE.
MTBS = MOS TOTAL BRIER SCORE.
RTBS = RAFS TOTAL BRIER SCORE.

PIMSF = PCT. IMPR. OF MOS OVER SAMPLE FREQ. BRIER SCORE.
PIRM = PCT. IMPR. OF RAFS OVER MOS BRIER SCORE.
BSIRM = IMPR. OF RAFS OVER MOS BRIER SCORE.
MRNF = MOS FCST. RAIN FREQ.
RRNF = RAFS FCST. RAIN FREQ.

RR / NN	PIMSF	
RF / ITBS	PIRM	MRNF
MTBS / RTBS	BSIRM	
RRNF		

Table 2g (for MOS = ALL) shows that except for the SEAST all subareas showed an improvement when RAFS was 50 percent. Table 2d shows that all cells for (M = 50, R = A) except the EAST subarea showed positive improvements.

This sort of presentation in another setting (a breakdown by WSFO) can pinpoint an office which may be contributing to good (or poor) scores and where its strengths (problems) lie. A breakdown by forecaster within an office might be too much subdivision to make tables like these worthwhile, however.

The reduction of the number of PoP classes to six makes these tables wieldy, but the question arises as to how much skill is lost by not allowing a 40 percent PoP between 30 and 50 percent, for example. We included in this project an experiment to find out what loss there was. The six-class (from graphics) MOS PoP's (G) were compared with the 12-class (MOS PoP of 2 percent was set to zero) MOS digital PoP's (D) to determine the differences in Brier scores.

The D system improvements over the G system are shown in Table 3. Much of the superiority of the D system forecasts was found to occur when the G system forecasts were 10 percent and the D system forecasts were 20 percent. For this combination the observed "rain" frequency was over 20 percent in all three lead times.

Table 3. Percent improvement of Brier score of 12-class system (D) over 6-class system (G) for various lead times.

TONIGHT	0.9
TOMORROW	0.4
TOMORROW NIGHT	1.6

On the other hand, for TONIGHT and TOMORROW the "rain" frequency for G forecasts of 30 percent and D forecasts of 20 percent was 30 percent. Thus the G forecasts actually got a better score than did the D forecasts, in spite of the latter's freedom to use more classes.

So there were conflicting results for the probabilities which were on the "climatological fence." (We later will propose adding the 20 percent class back into the G system, making a seven class system for future work.)

3. SOME NEW STATISTICS

As far as the usual verification system for PoP's is concerned, .01 inch of rain is treated the same as six inches. Strictly speaking, this is the way the game is played. Realistically, however, a forecaster who gets a hundredth on his 10 percent forecast will not feel like he/she "blew it" as much as one who got a two-inch downpour on the same PoP value. We propose some new statistics in this section designed to lend a quantitative flavor to the verification.

Table 4 has the total precipitation for each class for MOS and RAFS for each projection. For the first period it is seen that the total precipitation for the 90 percent class is by far the largest of any class. Almost half of the precipitation was in this class for MOS, and over half of all precipitation was with a RAFS PoP of 90 percent. Certainly this speaks well for the quality of the forecasts. Note that only a very small amount (.28") fell on zero RAFS PoP's in the first period. A preponderance of the precipitation was on either a 70 or 90 percent for both MOS and RAFS.

Table 4. Total observed precipitation amount (inches) for each PoP class for MOS and RAFS by lead time.

		0	10	30	50	70	90	ALL
1st PD	MOS	1.00	3.19	6.35	15.66	21.93	41.55	89.68
TONIGHT	RAFS	.28	1.84	7.92	13.24	19.48	46.92	89.68
2nd PD	MOS	.51	7.43	16.37	15.91	27.08	12.97	80.27
TMRW	RAFS	.20	3.04	11.38	23.09	25.60	16.96	80..27
3rd PD	MOS	.23	7.05	16.94	18.87	22.37	15.69	81.15
TMRW NT	RAFS	.24	8.30	15.95	20.98	18.58	17.10	81.15

For the second period the class amounts maximized at 70 percent for both MOS and RAFS, though 50 percent was close behind for RAFS. Very little precipitation fell on zero PoP's. For the third period MOS PoP's the largest total was still on 70 percent, but for RAFS the maximum was on 50 percent. For the fourth period the largest totals were on 30 percent PoP's, while the totals for the zero class were larger than for the shorter projections.

Table 5 shows the <u>average</u> amounts of precipitation <u>when it precipitated</u> (.01" or more). For all periods the average amounts were largest on 90 percent for both MOS and RAFS. Note that the average amounts were only a few hundredths for PoP's of zero, 10 and 30 percent.

Table 5. Average observed precipitation amount (inches) <u>when it precipitated</u>.

		0	10	30	50	70	90	ALL
1st PD	MOS	.07	.05	.06	.14	.21	.43	.18
TONIGHT	RAFS	.02	.05	.07	.10	.19	.47	.18
2nd PD	MOS	.06	.08	.16	.16	.29	.36	.19
TMRW	RAFS	.03	.06	.09	.21	.30	.31	.19
3rd PD	MOS	.03	.08	.13	.20	.31	.39	.19
TMRW NT	RAFS	.03	.10	.11	.23	.26	.49	.19

Table 6 shows the average amounts for all forecasts of that class, regardless of whether it precipitated. It will be seen that for both MOS and RAFS the largest averages were on 90 percent on all three periods. In all cases the averages for zero PoP's were less than .005".

Table 6. Average observed precipitation amount (inches) for all cases.

		0	10	30	50	70	90	ALL
1st PD	MOS	.00	.01	.02	.09	.17	.37	.05
TONIGHT	RAFS	.00	.00	.02	.06	.16	.43	.05
2nd PD	MOS	.00	.01	.05	.09	.22	.30	.04
TMRW	RAFS	.00	.01	.03	.12	.23	.28	.04
3rd PD	MOS	.00	.01	.04	.10	.18	.32	.04
TMRW NT	RAFS	.00	.01	.03	.12	.16	.45	.04

A new score which tells how well the forecasts caught the larger amounts on high probabilities is the weighted precipitation ratio (WPR).

It is defined by

$$WPR = \frac{\sum_{i=1}^{N} o_i \sum_{k=1}^{6} f_k A_k}{\sum_{i=1}^{N} f_i \sum_{k=1}^{6} A_k},$$

where o is the observed event (0 or 1), f is the forecast probability, and A is the amount of precipitation observed. The summations over o_i in the numerator and f_i in the denominator are designed to discourage overforecasting to get a better score. The results from this experiment are shown in Table 7. This score is positively oriented — that is, the higher the better.

Table 7. WPR scores (defined in text) for MOS and RAFS by lead time.

1st PD	MOS	75.7
TONIGHT	RAFS	85.0
2nd PD	MOS	57.6
TMRW	RAFS	71.9
3rd PD	MOS	58.4
TMRW NT	RAFS	56.8

4. A PROPOSED NEW SET OF EVENT-DEFINING THRESHOLDS

a. The Present System and Its Faults

The instigators of PoP forecasting in the 1960's (or earlier) decided that a precipitation event would be defined as .01 inch of precipitation or more in six or twelve hours. Was this a wise decision? The phrase "Probability of Precipitation," if taken literally, would mean the event is any precipitation, however slight, meaning Trace or more. The .01 inch threshold we use causes problems in wording our forecasts for drizzle and snow flurry events. To get around this we sometimes say "Probability of Measurable Precipitation." This really isn't good, either, because it all depends on what one is using to measure the precipitation. Not everyone has an eight-inch gauge. Many plastic gauges cannot measure very small amounts. On the other hand one could count the number of rain drops which fall on the sidewalk in one square during a sprinkle and call this a "measure" of the precipitation. It would be a fair guess that of the general public not one person in ten (maybe not one in a hundred) knows what we mean by "Measurable Precipitation." If we were in the metric system we certainly would use something different from .01 inch as the threshold.

Besides this there is the question of whether .01 inch is significant enough to warrant all the hoop-la being made over it. The senior author visited a northern WSFO one February, and it snowed most of the time he was there, but not much. He was told that much of the time they had little skill in separating Traces from .01 inch or so, and in such cases they simply did not try to beat MOS guidance and went along with it.

Perhaps the most serious drawback to using .01 inch is that smaller amounts occur much more frequently than larger ones, so the number of borderline events is bound to be greater than for a higher threshold. Also, are we so sanguine as to believe that observers are above being gently pushed by the forecaster into converting a Trace into .01 inch, or vice versa? There is bound to be a considerable demoralizing effect on the forecaster when a marginal call goes the wrong way when he/she has departed from guidance. If the decision is more clear cut the forecaster would lick his/her wounds and accept it. Lastly, the use of a higher threshold would mean that we are tending to reserve the probability numbers for the more important events. Since the event definition is more restrictive, a larger percentage of the events will have substantial amounts.

A very important drawback of the present PoP program is that we tend not to inform the public about something for which we have information —the amount of precipitation to occur. We have considerable skill now in this area which is not getting to the public and we should have even more skill in the future as models become better. So our proposal will deal with this aspect of the situation.

b. Proposed System

We propose replacing the present system of a limit at .01 inch with one using three thresholds — Trace, .04 inch, and .40 inch. Trace or more takes care of drizzle-snow flurry problems. The limit at .04 inch is large enough to reduce significantly the borderline case problem, while not so large as to eliminate as events cases which will certainly get the streets slick with rain, snow, or ice, for example. The .40 inch limit was chosen so that rains or snows which are rather major events would be separated from minor events.

The .04 and .40 inch limits are not cast in concrete but use by the authors confirms the belief that they are in the right ball park. It so happens that .04 inch is about a millimeter and .40 inch is about a centimeter, so that these limits would fit right in if the pendulum swings the other way and the U. S. goes to metric full scale some day.

It is not proposed that the Probability of a Trace (PoT) numbers would be issued to the public — one has to be careful in loading up the forecasts with numbers. Probability of Four (Po4) numbers would always be issued, while only occasionally would Probability of Forty (Po40) numbers be issued. In this way the confusion which could exist in the forecast would be minimized.

As an aside we note that other investigators have used several limits (e.g., Murphy et al., 1985) going up to as large as 2 inches. We fear that this complicates the problem to the point where it is not practical to incorporate them into the public forecasts. This is not to preclude the internal use of some higher limit or limits for the purpose of issuing flash flood watches, however.

c. The Effect of Probability of Four on the Number of Borderline Cases

We have tabulated observations to show the effects of raising the threshold from .01 inch to .04 inch. We first will discuss a tabulation done for the winter (December, January, February) of 1978-79 for the Central Region as a whole (using each WSFO location data). From a sample of 2520 twelve-hour total precipitation amounts the breakdown was as follows:

Trace	538
.01 inch	132
.02 inch	64
.03 inch	53
.04 inch	42

From these data we have estimated the number of borderline cases for each threshold by assuming that one eighth of the Traces were close to being .01 inch, and half of the .01 inch amounts were close to being Traces. This is arbitrary, but it provides some way of proceeding. Similarly, half of each of the .03 and .04 inch cases were considered borderline. This gives totals of

133.3 and 47.5 borderline cases for the .01 inch and .04 inch thresholds, respectively. Thus, higher limit would result in 36 percent as many fenceline situations, a rather substantial reduction.

The same statistics for the summer (June, July, August) of 1978 for the same 14 stations are 70.9 and 21.5 for the lower and higher limits, respectively. This gives only 30 percent as many events on the line.

One additional source of information is a file of hourly precipitation data for Omaha, Nebraska for July and August of a 30 year (1955-84) period. For a sample size of 3660 twelve-hour periods there were 89.4 and 29 borderline cases for the .01 inch and .04 inch limits, respectively, so the higher limit has only 32 percent as many cases. All in all there is then about a two-thirds reduction in the number of borderline situations with the higher threshold.

Additionally, we calculated the number of borderline cases at still a higher threshold, .10 inch, and found only 14 such cases, about half as many as at .04 inch. However, with this higher limit some winter events which could be considered important might be missed. Freezing rain is an example where the .10 limit would definitely be too high.

In Figs. 2a-b are given the relative frequencies of events at and above four thresholds — Trace, .01 inch, .04 inch, and .10 inch plotted at the WSFO's for winter 1978-79 and summer 1978. One item of note is the frequency of a Trace or more at Bismarck in the winter — 71 percent. This contrasts dramatically with the 9 percent for .04 inch or more, and only 4 percent for .10 inch or more. Thus, we find that 87 percent of the precipitation events (Trace or more) in Bismarck during this winter consisted of less than .04 inch in 12 hours.

For the Central Region as a whole raising the threshold to our new value would result in only 58 percent (14/24) as many events in winter and 74 percent (14/19) as many events in the summer. The forecast probabilities would be lower than the PoP's we have now so the public forecasts would take on a more optimistic (if you don't want rain or snow) tone while at the same time not inordinately raising the hopes of a farmer who is looking for rain for his crops during a dry August (a few hundredths won't help).

d. Experimental Results

Concurrently with the RAFS vs. MOS experiment described in Section 1, we made forecasts for the three thresholds, Trace, .04", and .40" for the cool season 1985-86. In order to have a comparison from MOS for the Po40 forecasts, we used the QPF12 numbers on the FPC's as guidance. The first two digits are the probabilities of .25" or more and .50" or more in 12 hours, in tens of percent. The average of these two was used as a MOS Po40, giving probabilities at 5 percent intervals. This assumption doesn't seem to cause any serious problems since Po40 is certainly bracketed by Po25 and Po50.

NOAA NATIONAL WEATHER SERVICE
CENTRAL REGION

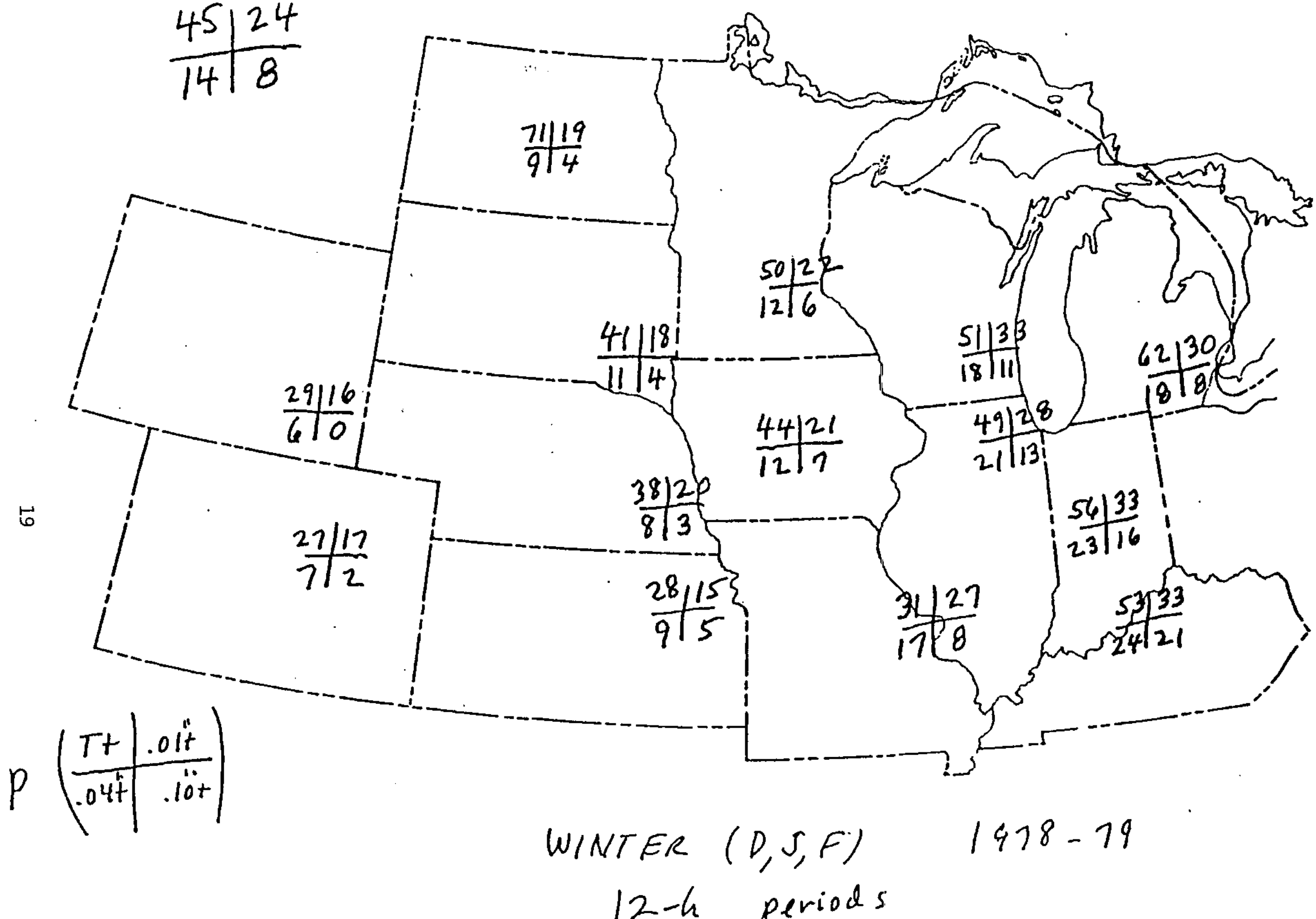

Fig. 2a. Relative frequencies (in percent) of precipitation events for thresholds of Trace, .01", .04", .10" as shown by key in lower left. For the winter (December, January, and February of 1978-79. Twelve-hour periods.

NOAA NATIONAL WEATHER SERVICE
CENTRAL REGION

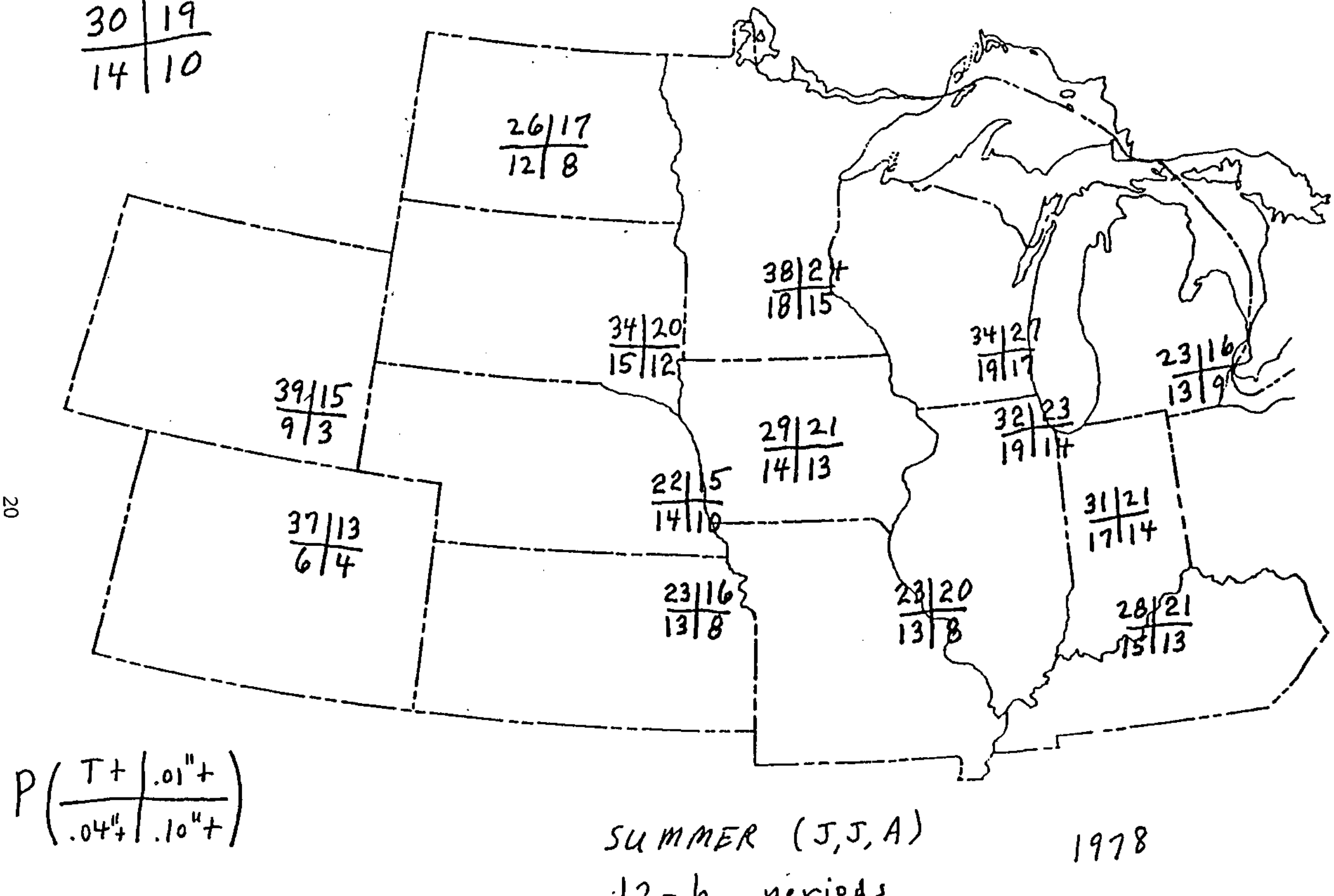

Fig. 2b. Same as for Fig. 2a, except for summer (June, July, and August) of 1978.

Tables 8a-c show reliability numbers and scores for four sets of forecasts for three projections. RPoT is the probability of a Trace or more based on RAFS, RPo4 is the probability of .04" or more based on RAFS, MP040 is the probability of .40" or more based on MOS (LFM), and RPo40 is the probability of .40" or more based on RAFS. PISF is the percentage improvement over the Brier Score obtained from using a constant forecast equal to the sample frequency. PIRM is the percentage improvement of RAFS over MOS in terms of the Brier Score.

A glaring systematic error for the RPoT's for all three periods is evident. Except for 100 percent, where all but three of 211 forecasts had a Trace or more, there is an underforecasting in all classes. The forecaster had no prior experience in making PoT's, and typically the PoT's were assigned values 20 percent higher than the Po4's. That this is too low in general is evident. This type of feedback would no doubt improve the reliability of future forecasts, and better PISF scores would result. Note that Trace or more events are quite frequent, almost double the frequency of .01" or more events.

The RPo4 forecasts show reasonable reliability. At least the relative frequencies (RF's) increase monotonically as the Po4 number increases from left to right. The poorest reliability is on the 60 percent value in the third period, where there were events only 44 percent of the time. The highest Po4's (80 percent) held up quite well, with events occurring more than 80 percent of the time on all three periods. PISF numbers are respectable.

We had a 12-class system for MPo40 due to the way they were determined, and the agreement between the RF's and the probabilities is erratic, to say the least. This is an indication that for a sample of this size 12 classes are too many. MPo40 had useful skill at all projections, however.

The reliabilities for RPo40 were better, but still had some problems in spite of the fact only six classes were used. Looking first at the Tonight period, none of the 81 five percent values had an event, and the 40 percent values represented a significant underforecast. These forecasts had a rather good score, with a PISF of 25.9. This represented a slight (2.1 percent) improvement over MOS.

For the second period the RF dropped between 20 percent and 30 percent, and 40 percents were again an underforecast. The improvement over MOS (PIRM) went up to 4.1 percent. For the third period the reliability went haywire, jumping from an RF of 19.4 on 30 percent to 76.9 on 40 percent. It is not obvious why this should have happened. The third period forecasts still have skill (PISF = 17.7) and were an improvement over MOS by a margin of 5.9 percent, the best of all three periods.

The RF of the larger amount events is 3 to 4 percent, so that they are not really rare, occurring often enough to make the statistics meaningful.

Table 8a. Verification statistics for RAFS PoT forecasts (RPoT), RAFS Po4 forecasts (RPo4), MOS Po40 forecasts (MPo40), and RAFS Po40 forecasts (RPo40) for the first period (TONIGHT).

				RPoT			
	0	20	40	60	80	100	ALL
R/M	43/512	142/511	232/386	170/210	116/120	110/111	813/1850
Rel. Freq.	8.4	27.8	60.1	81.0	96.7	99.1	43.9
						PISF = 32.7	

				RPo4			
	0	10	20	40	60	80	ALL
R/M	7/654	20/485	45/284	76/200	77/117	92/110	317/1850
Rel. Freq.	1.1	4.1	15.8	38.0	65.8	83.6	17.1
						PISF = 40.9	

				MPo40			
	0,5	10,15	20,25	30,35	40,45	50,55	ALL
R/M	1/1435	4/67	10/27	8/21	4/16	4/6	60/1850
Rel. Freq.	0.1	6.0	37.0	38.1	25.0	66.7	3.2
R/M	1/151	12/68	3/26	3/11	4/10	6/12	
Rel. Freq.	0.7	17.6	11.5	27.3	40.0	50.0	
						PISF = 24.3	

				RPo40			
	0	5	10	20	30	40	ALL
R/M	2/1448	0/81	10/161	15/82	18/54	15/24	60/1850
Rel. Freq.	0.1	0	6.2	18.3	33.3	62.5	3.2
						PISF = 25.9	
						PIRM = 2.1	

R = Number of "rains"   M = Number of forecasts

PISF = Percent Improvement over Sample Frequency Brier score.

PIRM = Percent Improvement of RAFS over MOS Brier score.

Table 8b. Same as for Table 8a, except for the second period (TOMORROW).

	RPoT						
	0	20	40	60	80	100	ALL
R/M	55/512	185/507	253/435	164/204	97/109	59/61	813/1828
Rel. Freq.	10.7	36.5	58.2	80.4	89.0	96.7	44.5
						PISF = 20.1	
	RPo4						
	0	10	20	40	60	80	ALL
R/M	5/629	34/495	61/345	75/192	70/106	50/61	295/1828
Rel. Freq.	0.8	6.9	17.7	39.1	66.0	82.0	16.1
						PISF = 32.7	
	MPo40						
	0,5	10,15	20,25	30,35	40,45	50,55	ALL
R/M	7/1432	6/40	9/49	10/39	5/8	0/2	67/1828
Rel. Freq.	0.5	15.0	18.4	25.6	62.5	0	3.7
R/M	15/159	2/53	7/24	4/12	1/7	1/3	
Rel. Freq.	9.4	3.8	29.2	33.3	14.3	33.3	
						PISF = 12.7	
	RPo40						
	0	5	10	20	30	40	ALL
R/M	5/1453	6/80	17/146	22/93	9/39	8/17	67/1828
Rel. Freq.	0.3	7.5	11.6	23.7	23.1	47.1	3.7
						PISF = 16.3	
						PIRM = 4.1	

Table 8c. Same as Table 8a, except for the third period (TOMORROW NIGHT).

RPoT	0	20	40	60	80	100	ALL
R/M	38/400	195/625	285/492	127/179	96/116	39/39	780/1851
Rel. Freq.	9.5	31.2	57.9	70.9	82.8	100.0	42.1
						PISF = 17.7	

RPo4	0	10	20	40	60	80	ALL
R/M	9/529	47/620	77/377	63/174	50/113	33/38	279/1851
Rel. Freq.	1.7	7.6	20.4	36.2	44.2	86.8	15.1
						PISF = 19.9	

MPo40	0,5	10,15	20,25	30,35	40,45	50,55	ALL
R/M	4/1313	4/61	13/61	11/39	1/17	0/3	59/1851
Rel. Freq.	0.3	6.6	21.3	28.2	5.9	0.0	3.2
R/M	5/220	6/89	4/25	5/13	3/6	3/4	
Rel. Freq.	2.3	6.7	16.0	38.5	50.0	75.0	
						PISF = 12.5	

RPo40	0	5	10	20	30	40	ALL
R/M	8/1484	2/63	8/134	25/126	6/31	10/13	59/1851
Rel. Freq.	0.5	3.2	6.0	19.8	19.4	76.9	3.2
						PISF = 17.7	
						PIRM = 5.9	

5. WORDING OF FORECASTS IN THE NEW SYSTEM

Having shown that there is no great problem in making forecasts with the new limits, and that they have skill, we now turn to the problem of incorporating them into the public forecasts.

First of all, though, this experiment has led us to the conclusion that seven classes of probabilities for each threshold, instead of six, should be used. For PoP's and Po4's the lack of another class close to the climatological value was restricting. For Po40 there may be skill in using a class higher than the 40 percent we used. Also, the utility of the 5 percent class for Po40 is questionable. For these reasons we propose that in the future the following classes be used:

PoT	(0, 10, 20, 40, 60, 80, 100)
Po4	(0, 10, 20, 30, 50, 70, 90)
Po40	(0, 10, 20, 30, 40, 60, 80)

The basic rules we propose are:

(1) All probability numbers will be at the end of the forecast for each period.

(2) PoT numbers are used as a guide to the modifier usage (CHANCE, LIKELY, ETC.), but never included in the forecast.

(3) Po4 numbers are always included (even near zero, 10, and 20 percent) in the first three periods.

(4) Po40 numbers are sometimes included in either or both of the first two periods.

(5) When the first period is twelve hours in length (4 AM forecast — TODAY; 4 PM forecast — TONIGHT), two Po4's and sometimes one or two Po40's are provided for the six-hour subperiods. If both Po4's are zero, a zero twelve-hour Po4 is given. If the Po4's are (0, 10) or (10, 0) for verification, a single 10 percent Po4 for a twelve hour period is given. For Po4's of (10, 10) a 10 or 20 percent Po4 for twelve hours is used.

We propose that the requirement to link words strictly to the numbers be abolished. The numbers speak for themselves, and the use of two Po4's in the first twelve-hour period presents complications. Besides, when we say SLIGHT CHANCE OF SHOWERS it does little good to link SLIGHT CHANCE to a probability number because the event SHOWERS is not defined. Are we talking about having showers somewhere in the forecast area, or are we talking about a point? If the latter is the answer, how much rain constitutes an event? Another serious problem with the present system is in a forecast which reads "30 PERCENT CHANCE OF THUNDERSTORMS." Thunderstorms often occur with no precipitation (or only a Trace), while on the other hand sometimes when they are forecast only showers occur. In contrast, with some public education the event associated with Po4 and Po40 numbers becomes well defined to the user so that the use of a number is justified.

Forecaster judgement on what qualifiers to use is proposed, with only a few guidelines. Precipitation would never be mentioned on a zero PoT. Precipitation need not be mentioned on PoT's of 10 and 20 percent at the discretion of the forecaster. It is envisioned that more often than not no mention would be made on these PoT's. Tying the qualifiers to PoT's, rather than Po4's, allows them to apply to small-amount events such as snow flurries or drizzle.

The point and areal qualifiers would still be ranked from low to high probability without strict linkage as follows:

POINT QUALIFIER	AREAL QUALIFIER
SLIGHT CHANCE	ISOLATED
CHANCE	WIDELY SCATTERED
GOOD CHANCE	SCATTERED
LIKELY	NUMEROUS
(Unqualified mention of precipitation)	

The Po4 and Po40 statements consist of:

CHANCE OF FOUR IS XX PERCENT

CHANCE OF FORTY IS XX PERCENT

The use of FOUR instead of PRECIPITATION saves four syllables; likewise for CHANCE instead of PROBABILITY.

We should include near zero Po4's because they are among the best we make — good weather is easier to forecast than is bad weather. We feel the 10 percent values should be included because, as shown in Tables 8a-c, we do have the ability to distinguish them from zero and 20 percent values. Quite a few "rains" fall on 10 percent. Indeed, Table 8c shows that 17 percent (47/279) of all "rains" in the third period were on 10 percent values. We are now making 10 percent PoP's — why not use them? The inclusion of zero's and 10's is no problem because the forecasts are already short in most instances. Why make a short forecast shorter?

We propose that Po40 numbers be included for all values except zero. If Po40 is 30 percent or higher the possibility of heavy precipitation should be mentioned in the text.

The above rules are it—no exceptions, few options, clear and simple (the KISS Principle—Keep It Simple Stupid). By not binding the words strictly to the numbers more independent information can be transmitted to the user. For example, by using DRIZZLE LIKELY with a Po4 of 10 percent we are saying that small amounts are very probable, but larger ones are unlikely.

f. Some Examples of Public Forecasts in the New System

In this section we will present some examples from a variety of weather situations using the new system.

(1) .TODAY...RAIN AND THUNDERSTORMS...HEAVY AT TIMES...WITH EAST WINDS 10 MPH BECOMING STRONG SOUTHEAST 20 MPH WITH GUSTS TO 30 BY EVENING. HIGH IN THE LOWER 60S. CHANCE OF FOUR IS 70 PERCENT THIS MORNING AND 90 PERCENT THIS AFTERNOON. CHANCE OF FORTY IS 40 PERCENT THIS MORNING AND 40 PERCENT THIS AFTERNOON.

.TONIGHT...SHOWERS AND THUNDERSTORMS LIKELY...HEAVY AT TIMES. LOW IN THE LOWERS 50S. SOUTHEAST WINDS 15 TO 20 MPH. CHANCE OF FOUR IS 70 PERCENT. CHANCE OF FORTY IS 30 PERCENT.

.TOMORROW...BREEZY WITH SCATTERED SHOWERS. CLOUDY...HIGH IN THE LOWERS 60S. CHANCE OF FOUR IS 30 PERCENT.

(2) ...WINTER STORM WARNING FOR THIS MORNING...
.TODAY...OCCASIONAL SNOW...POSSIBLY MIXED WITH RAIN OR FREEZING RAIN AT TIMES. SNOW ACCUMULATING 1 TO 2 INCHES BEFORE TAPERING OFF TO FLURRIES THIS AFTERNOON. HIGH IN THE MID 30S. VARIABLE WINDS 20 TO 30 MPH. CHANCE OF FOUR IS 70 PERCENT THIS MORNING AND 30 PERCENT THIS AFTERNOON. CHANCE OF FORTY IS 20 PERCENT THIS MORNING.

.TONIGHT...FLURRIES LIKELY. LOW IN THE MID 20'S. WEST WIND 10 TO 20 MPH. CHANCE OF FOUR IS 10 PERCENT.

.SATURDAY...CLOUDY WITH A GOOD CHANCE OF FLURRIES. HIGH NEAR 30. CHANCE OF FOUR IS 10 PERCENT.

(3) .TONIGHT...INCREASING CLOUDINESS WITH SNOW LIKELY LATE TONIGHT...ACCUMULATION AROUND AN INCH OVERNIGHT. LOW 10 TO 15 EARLY THEN TEMPERATURES RISING TO NEAR 25 BY MORNING. SOUTH WINDS INCREASING TO 15 TO 20 MPH. CHANCE OF FOUR IS 10 PERCENT BEFORE MIDNIGHT AND 50 PERCENT AFTER MIDNIGHT. CHANCE OF FORTY IS 10 PERCENT AFTER MIDNIGHT.

.FRIDAY...PERIODS OF SNOW WITH ADDITIONAL ACCUMULATIONS LIKELY...NOT AS COLD. HIGH AROUND 30. SOUTH WINDS 15 TO 25 MPH. CHANCE OF FOUR IS 70 PERCENT. CHANCE OF FORTY IS 20 PERCENT.

.FRIDAY NIGHT...MOSTLY CLOUDY WITH A CHANCE OF FLURRIES. LOW NEAR 20. CHANCE OF FOUR IS 20 PERCENT.

.SATURDAY...MOSTLY SUNNY. HIGH 30 TO 35.

(4) .THIS AFTERNOON...WARM WITH A RECORD HIGH IN THE MID 80S. PARTLY CLOUDY WITH A GOOD CHANCE OF SHOWERS AND THUNDERSTORMS. SOUTHWEST WINDS 10 TO 20 MPH. CHANCE OF FOUR IS 50 PERCENT. CHANCE OF FORTY IS 10 PERCENT.

.TONIGHT...CLOUDY WITH SHOWERS AND THUNDERSTORMS LIKELY. LOW 55 TO 60. LIGHT SOUTH WINDS. CHANCE OF FOUR IS 70 PERCENT. CHANCE OF FORTY IS 30 PERCENT.

.SATURDAY...SHOWERS AND THUNDERSTORMS LIKELY. HIGH IN THE MID 70S. SOUTHEAST WINDS 10 TO 15 MPH. CHANCE OF FOUR IS 50 PERCENT.

(5) .TONIGHT...BECOMING MOSTLY CLEAR. LOW 40 TO 45. SOUTH TO SOUTHWEST WINDS DIMINISHING TO 10 TO 15 MPH. CHANCE OF FOUR IS NEAR ZERO.

.SATURDAY...MOSTLY SUNNY. HIGH 65 TO 70. SOUTHWEST WINDS 10 TO 15 MPH. CHANCE OF FOUR IS NEAR ZERO.

.SATURDAY NIGHT...MOSTLY CLEAR. LOW SATURDAY NIGHT 40 TO 45. CHANCE OF FOUR IS NEAR ZERO.

.SUNDAY...MOSTLY SUNNY. HIGH AROUND 70.

(6) .TODAY...LINGERING SHOWERS THIS MORNING BECOMING MOSTLY SUNNY BY AFTERNOON. HIGH NEAR 90. SOUTH WINDS 10 TO 20 MPH. CHANCE OF FOUR IS 30 PERCENT THIS MORNING AND 10 PERCENT THIS AFTERNOON.

.TONIGHT...A GOOD CHANCE OF SHOWERS AND THUNDERSTORMS DEVELOPING AGAIN AFTER MIDNIGHT. LOW NEAR 70. SOUTH WINDS 10 TO 15 MPH. CHANCE OF FOUR IS 50 PERCENT. CHANCE OF FORTY IS 20 PERCENT.

.SATURDAY...SLIGHT CHANCE OF SHOWERS IN THE MORNING. HIGH NEAR 90. CHANCE OF FOUR IS 30 PERCENT.

6. SUMMARY

Three separate ideas — reducing the number of probability classes, new statistics, and introducing three new thresholds to replace .01 inch, are presented here. They do not depend upon one another for implementation. The number of classes could be reduced relatively easily. A first step would be to present MOS PoP's to the nearest percent in the FPC's, making it easier to change classes without going to the graphical products.

The new statistics could be used with the present system. The third proposal has more far reaching implications and couldn't be undertaken without high level approval, no doubt. It has the promise of providing more information to the users in a palatable form. It would make the probability of precipitation program more pertinent and stir interest among forecasters. Some present hassles such as the drizzle-snow flurry problem would be eliminated. Numbers would be given to the chance for a major rain or snow event. All of this would mean more service to the public by the National Weather Service.

7. ACKNOWLEDGEMENTS

I (WES) wish to give Gregg Hooker, formerly at the Cheyenne WSFO (now a lead forecaster at the Seattle WSFO), credit for stirring my initial interest in the possibility of raising the threshold for precipitation probabilities. Our boss, Joe Schaefer, has been unflagging in his support of the ideas contained herein, for which we are most grateful. Bev Lambert's unsurpassed ability with the word processor made working with her to get the finished product a pleasure.

8. REFERENCES

Clemen, R. T., and A. H. Murphy, 1985: Objective and subjective precipitation probability forecasts: Statistical analysis of some interrelationships. Manuscript submitted to Weather and Forecasting, 29 pp.

Murphy, A. H., W. Hsu, R. L. Winkler, and D. S. Wilks, 1985: The use of probabilities in subjective quantitative precipitation forecasts: Some experimental results. Mon. Wea. Rev., 113, 2075-2089.