PACIFIC REGION TECHNICAL NOTES

83-019
May 6, 1983

Verification of POP over the South Coast of British Columbia

D. Grimes, Supervising Meteorologist
Pacific Weather Centre, Vancouver, B.C.

INTRODUCTION

Probability of Precipitation (POP) forecasts were introduced into the regional public forecasts in July of 1982. A verification procedure was devised from the beginning to assess the reliability and credibility of such a program in British Columbia.

The verification procedure generates Brier and skill scores for a number of forecast sites, but the focus of this report is on Victoria and Vancouver, located in the South Coast forecast area of British Columbia.

Some forethoughts have also been included in an attempt to put a perspective on POP forecasting in this region.

COMMENTS ON POP FORECASTS

The POP forecast is for a point, not necessarily any point, but the user's point of interest. In reality, this represents a very large number of points in any one forecast region, and the forecaster is presented with an improbable task of predicting a POP forecast to be "true" for all. Before embarking on such a task, one must first discover who the significant users are as well as defining what POP really means.

It is inconceivable to assume all "receivers" of the POP forecasts are significant users. It is more likely to say, that POP adds a new dimension to the Public Forecast by providing a trend (...an increasing/decreasing chance of precipitation) which appeals to the majority of the "receivers". However, the real value in POP is the added information it provides for decision makers. These users attach more importance to POP (and the forecasters' reliability) when confronted with making weather sensitive decisions. In Pacific Region, attempts should be made to determine the demographic characteristics of this group of users before considering future changes to the program.

POP is officially defined as "the probability of an occurrence of measurable precipitation at the point of interest located in the forecast area, sometime during the forecast period". There are two terms in this statement which need further clarification. One of the key words is 'probability' which can have only one of two viable interpretations. The first is a (pseudo) relative frequency of occurrence - an objective measure based on historical data, of similar synoptic patterns; while the second is subjective - the forecaster's degree or measure of belief that measurable precipitation will occur.

The latter is a more realistic interpretation, since synoptic situations are never identical. However, the former has some validity as a starting point for evaluating similar circulation patterns.

The second term in the definition warranting some discussion is 'the point of interest'. The users should perceive this forecast to apply specifically to their location. From the forecasters' viewpoint, one representative forecast to apply to all user points may not be possible, even in relatively homogeneous forecast districts. One could make a case when any one point in the forecast district could have the same value as any other if: (1) showers are occurring randomly (independent of terrain characteristics), (2) no precipitation will occur anywhere (0%); or (3) every point will have measurable precipitation (100%). Under this premise, certainty forecasts should be infrequent and the climatology should reflect similar precipitation records for all sites. A more realistic approach is to allow the forecasters to predict a POP for always the same point - indicating their degree of belief that precipitation will occur at that point. If precipitation occurrences are not uniform over the forecast area, the users who have a vested interest in such forecasts should evaluate the reliability of this new information and make adjustments for the forecasters' bias.

VERIFICATION

1) Scoring Method

The method chosen for this verification is the simplified Brier score which represents the mean square error over a number of forecasts. The score varies between 0 and 1, with 0 being the best possible score. The score is represented mathematically as:

$$P = \frac{1}{N} \sum^{N} (F_i - e_i)^2$$

where e = 0 (if no precipitation occurs)
= 1 (if precipitation does occur)

if (as in the case for POP) probability forecasts (F) can be divided into 10% categories, the Brier score can be further divided into two components:

$$P = \frac{1}{N} \sum^{11} n_i \, (F_i - O_i)^2 + \frac{1}{N} \sum^{11} n_i \, (O_i \, (1 - O_i))$$

where F = 0, 10, 20, ..., 90, 100%
O_i = observed precipitation frequency in each forecast category
= (# rain observations/total # forecasts)

The first term is referred to as the reliability, while the second is called the resolution. This latter way of considering the Brier score offers additional insight into the explanation of the error arising in a number of probability forecasts.

Finally, to add meaning to the score, it needs to be compared to some standard. In this case, two have been chosen: the first is the toss of a coin, representing the uneducated guess (i.e. F = 50% always); and the second is climatology. The results are presented in the form of a skill score which varies from 1 to negative infinity. In this case, 1 would be the best score.

2) Results

Two sets of forecasts are being examined: the objective POPA guidance from CMC (Probability of Precipitation Amount) and the subjective forecasts issued by PWC. The forecasts for Vancouver and Victoria have been categorized into 2 and 3 month groups (see Table 1).

Table I provides a summary of the Brier scores for Today, Tonight, Tomorrow, and the third day based on climatology, the subjective forecasts issued at 5 a.m. and 4 p.m., and the objective forecasts based on the 00Z and 12Z data.

Three major observations are evident from the Table, the first is illustrated by the scores based on climatology and the other forecasts. Except for the July-August period, the Brier scores of both the objective/subjective forecasts are showing skill over climatology (also over the toss of the coin which yields a Brier score of always .25). The zero or negative skill generated in July-August, may be a result of two things: firstly the program was new and forecasters required a "breaking in" period; and secondly, during the summer season, precipitation is generally more widely scattered and more often related to meso-scale events.

The second point to be emphasized, is the variability in the Brier scores between Victoria and Vancouver, although the forecast frequency distributions were nearly the same. This just re-affirms the necessity of producing different POP forecasts for these locations.

The third significant observation is that the subjective forecasts for Tonight and Tomorrow issued at 4 p.m. are, for most cases, better than the 5 a.m. forecasts. This is the result one would expect, given the additional information assimilated during the day. However, in the case of the objective guidance (POPA) this result is not observed. As well, the Brier score does not exhibit the same pattern as the subjective forecasts. Instead of the scores getting progressively larger as the time period moves further into the future, the scores for the POPA forecasts are much more irregular.

Figures 1 through 4 compare the subjective (5 a.m.) and objective verification results for the Today and Tomorrow periods. All the figures are similar with the subjective results at the top and the objective results at the bottom. The circle to the left splits the Brier score into its resolution and reliability components and indicates the level of skill over a toss (50% POP) and climatology. The circle on the right depicts the forecast distribution and the reliabilty error ($F_i - O_i$) in each

category (-·- shading means an overforecast and a —— shading means an underforecast). The number in the brackets to the bottom-right of the left hand circle indicates what the resolution error (and Brier score) would have been if the forecasts were perfectly reliable. In other words, with the forecast probability (F) in each category matched to observed frequency (O) in each category.

Figures 1a and 1b compare Vancouver's subjective and objective forecast results for Today during the Sept.-Nov. and Dec.-Feb. periods, while figures 2a and 2b depict a similar comparison for Victoria. The most striking feature which is common to all four figures is that nearly 50% of the subjective forecasts predicted certainty (0 or 100%), while the objective forecasts predicted near certainty up to 12% of the time for Vancouver and over 17% for Victoria. Except for Victoria (Figure 2b), the guidance appeared very reliable. The subjective forecasts had small reliability errors with Vancouver (Figure 1b) being the best. Focussing on the left hand side of these figures, it appears that the objective guidance is at the least as reliable as the subjective forecasts. Surprisingly, the objective guidance also shows comparable resolution errors to the subjective forecasts. However, the subjective scores showed more skill during the Dec.-Feb. period (a reversal from the Sept.-Nov. period). One could conclude that the forecasters should use this guidance especially if reliability is what one is striving for. However, there is one more interesting result. Suppose that both the subjective and objective forecasts were perfectly reliable ($F_i = O_i$) in each of the 11 (0, 10, ...90, 100%) different categories. Under such reliable circumstances, the resolution errors (numbers in brackets) do not significantly change for the subjective forecasts, but increase to at least .16 for the objective forecasts providing a much less skillful product.

Considering Figures 3a, 3b, 4a, and 4b for Tomorrow, similar observations can be made. The forecast frequency diagrams indicate that certainty predictions are made generally 25% of the time by the forecaster and only about 5% of the time by the guidance. The objective guidance was very reliable while the subjective forecasts were a little less so. The majority of the objective guidance forecasts ranged between 10% and 60% but the resolution errors were similar in magnitude to the subjective forecasts. In all circumstances the skill of the objective forecasts (Tomorrow) were better than the skill levels of the subjective forecasts, but all of the forecasts demonstrated some skill. However, if both the objective and subjective forecasts were perfectly reliable in the tomorrow period, then the subjective forecasts would have been more skillful (comparing the numbers in bracket).

CONCLUSIONS

Overall, the POP forecasts issued by PWC for the South Coast are demonstrating skill, with the best forecasts produced in the first 12 hours. The objective POPA guidance is also demonstrating skill over climatology for

all forecast periods, generally demonstrating better skill than the subjective forecasts after the first 12 hours. Up until now, POP forecasts have been produced with little reliance on the POPA guidance. These verification results suggest using POPA as the initial step, and subjectively adjusting them according to the evolution of the weather patterns and forecaster's experience. This method would likely yield a very skillful POP forecast for Vancouver and Victoria.

REFERENCES

Grimes, D., 1982, A Preliminary Verification of the Pacific Weather Centre's Probability of Precipitation Forecasts, PRTN 82-013.

Richardson, H. & Grimes, D., 1982, A Comparison of the POP Forecasts between CMC and PWC during August 1982, PRTN 82-021.

TABLE I

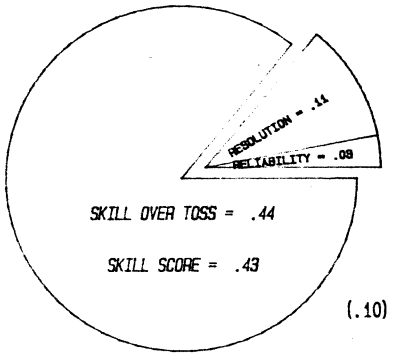
PoP Brier Scores

		Today	Tonight	Tomorrow	Third Day
July-Aug '82					
Vancouver	Climate	.19	.14	.18	.18
	5am fcst	.12	.14	.20	
	4pm fcst		.11	.20	.22
Victoria	Climate	.12	.14	.13	.13
	5am fcst	.08	.13	.13	
	4pm fcst		.10	.17	.14
Sept-Nov '82					
Vancouver	Climate	.25	.23	.25	.25
	5am fcst	.14	.13	.18	
	4pm fcst		.10	.16	.21
	PoPA (00z)	.12	.11	.15	
	PoPA (12z)		.12	.12	.14
Victoria	Climate	.21	.19	.24	.24
	5am fcst	.14	.16	.16	
	4pm fcst		.14	.16	.15
	PoPA (00z)	.12	.14	.15	
	PoPA (12z)		.16	.12	.15
Dec-Feb '83					
Vancouver	Climate	.26	.26	.23	.23
	5am fcst	.09	.15	.15	
	3pm fcst		.13	.15	.18
	PoPA (00z)	.12	.17	.15	
	PoPA (12z)		.14	.12	.16
Victoria	Climate	.25	.25	.25	.25
	5am fcst	.10	.17	.20	
	3pm fcst		.14	.18	.22
	PoPA (00z)	.15	.16	.19	
	PoPA (12z)		.13	.17	.17

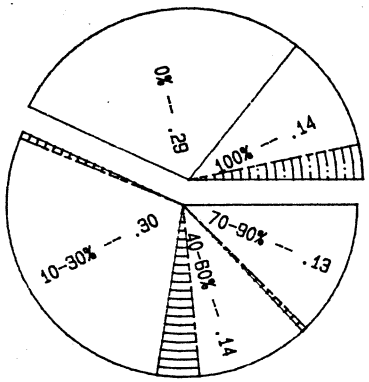
Vancouver's Today Forecast from Sept to Nov '82

Figure 1a

PNC Forecast Scores

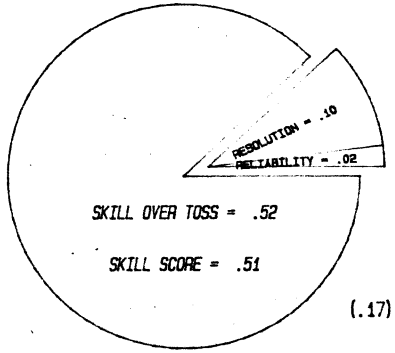


Forecast Frequency/Reliability

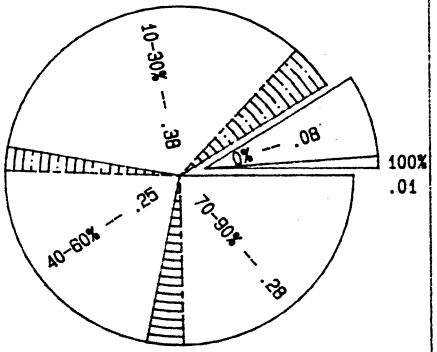


TOTAL # FORECASTS = 91

Guidance Forecast Scores



Guidance Frequency/Reliability

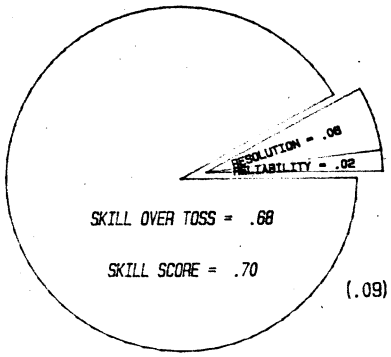


TOTAL # FORECASTS = 89

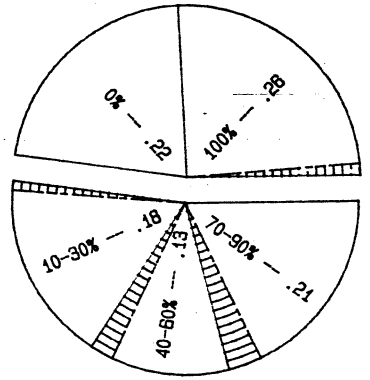
Vancouver's Today Forecast from Dec to Feb '83

Figure 1b

PNC Forecast Scores

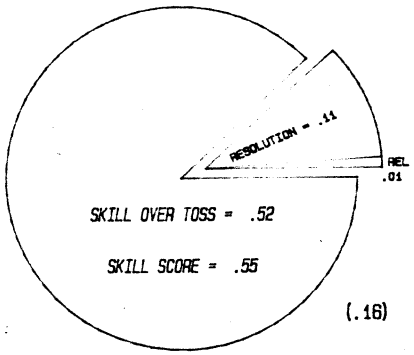


Forecast Frequency/Reliability

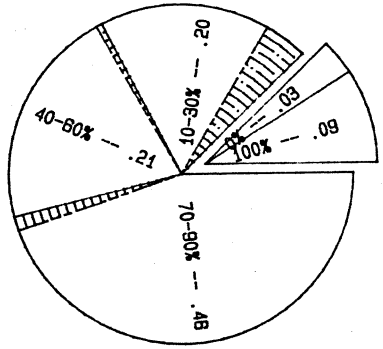


TOTAL # FORECASTS = 90

Guidance Forecast Scores



Guidance Frequency/Reliability

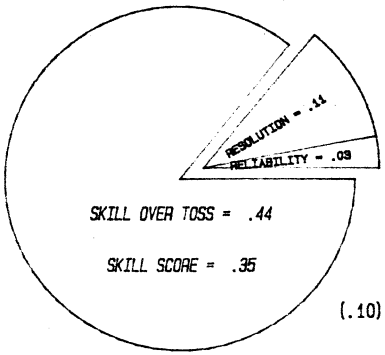


TOTAL # FORECASTS = 89

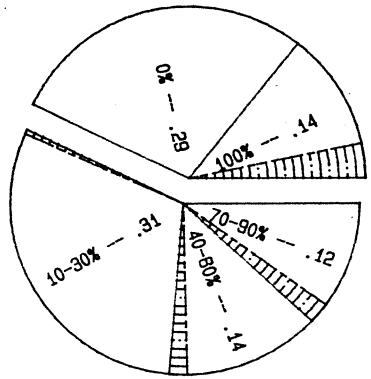
Victoria's Today Forecast from Sept to Nov '82

Figure 2a

PMC Forecast Scores

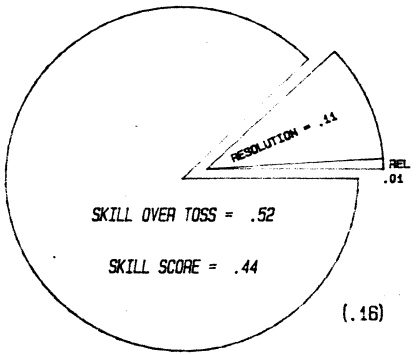


Forecast Frequency/Reliability

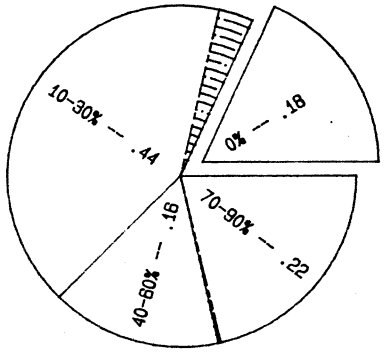


TOTAL # FORECASTS = 91

Guidance Forecast Scores



Guidance Frequency/Reliability

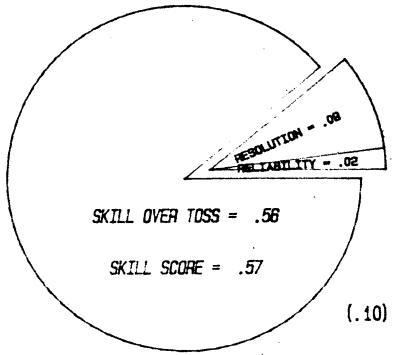


TOTAL # FORECASTS = 88

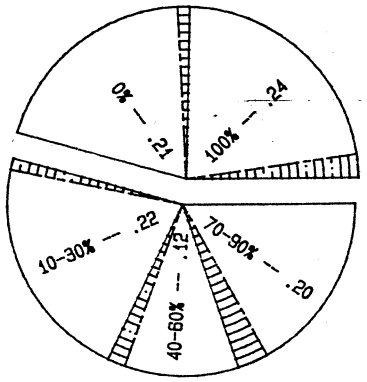
Victoria's Today Forecast from Dec to Feb '83

Figure 2b

PMC Forecast Scores

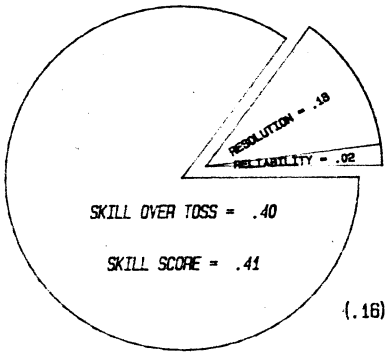


Forecast Frequency/Reliability

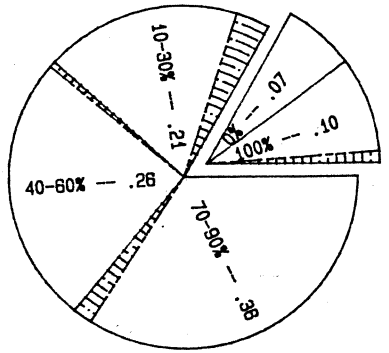


TOTAL # FORECASTS = 90

Guidance Forecast Scores



Guidance Frequency/Reliability

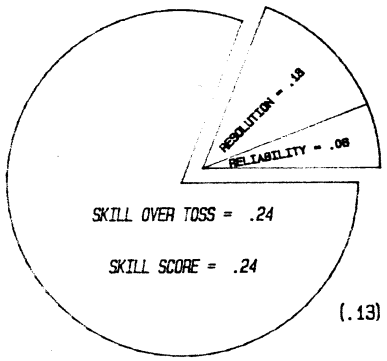


TOTAL # FORECASTS = 89

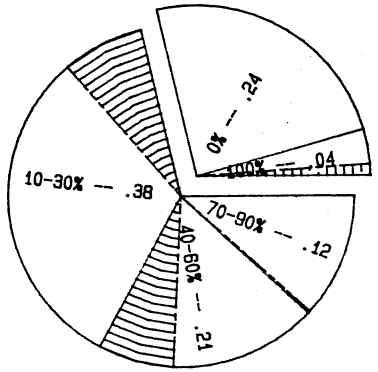
Vancouver's Tomorrow Forecast from Sept to Nov '82

Figure 3a

PWC Forecast Scores

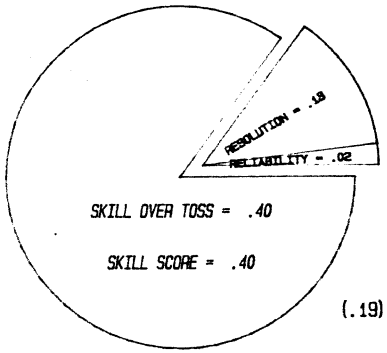


Forecast Frequency/Reliability

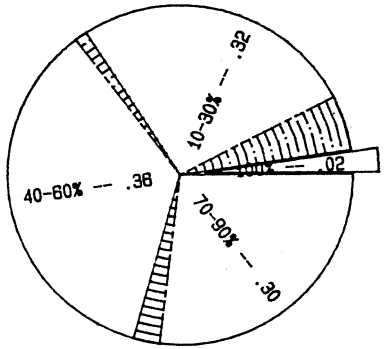


TOTAL # FORECASTS = 91

Guidance Forecast Scores



Guidance Frequency/Reliability

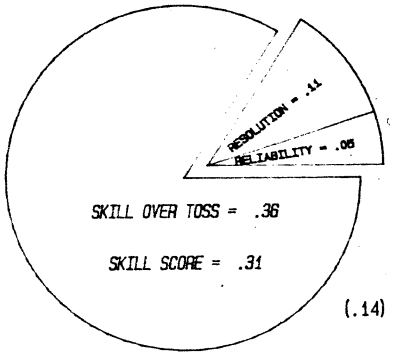


TOTAL # FORECASTS = 88

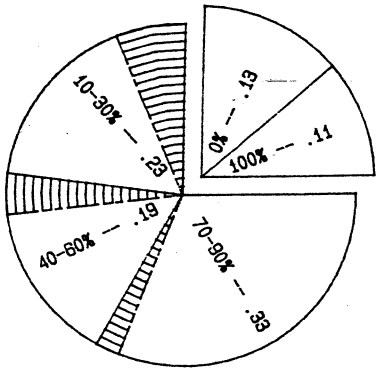
Vancouver's Tomorrow Forecast from Dec to Feb '83

Figure 3b

PWC Forecast Scores

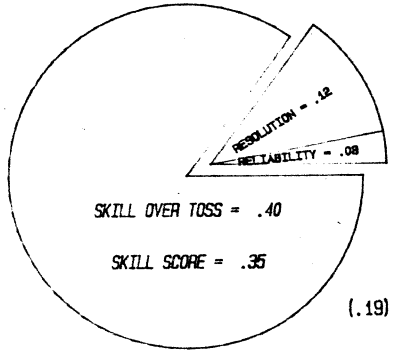


Forecast Frequency/Reliability

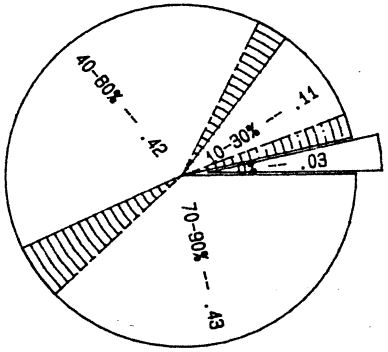


TOTAL # FORECASTS = 90

Guidance Forecast Scores



Guidance Frequency/Reliability

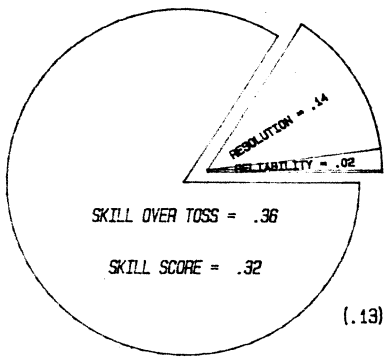


TOTAL # FORECASTS = 88

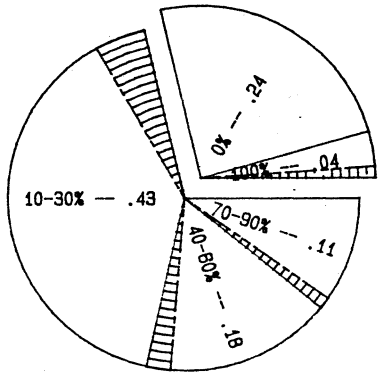
Victoria's Tomorrow Forecast from Sept to Nov '82

Figure 4a

PWC Forecast Scores

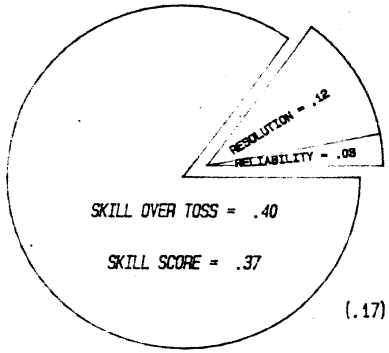


Forecast Frequency/Reliability

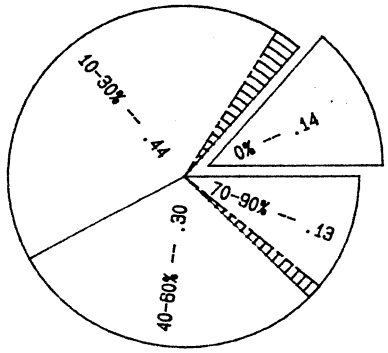


TOTAL # FORECASTS = 91

Guidance Forecast Scores



Guidance Frequency/Reliability

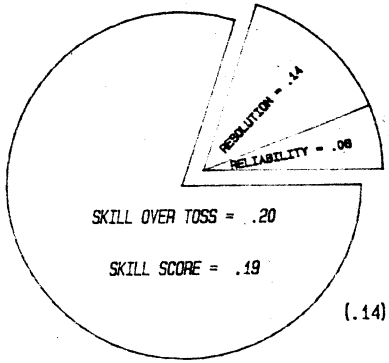


TOTAL # FORECASTS = 88

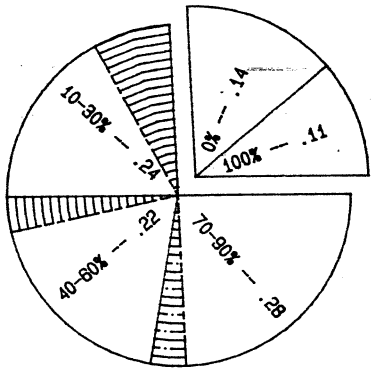
Victoria's Tomorrow Forecast from Dec to Feb '83

Figure 4b

PWC Forecast Scores

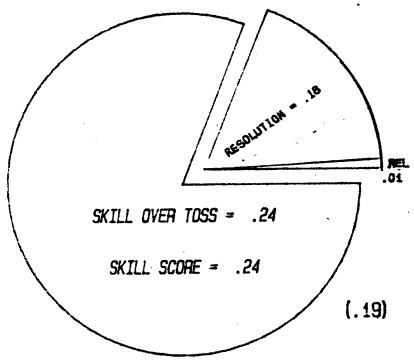


Forecast Frequency/Reliability

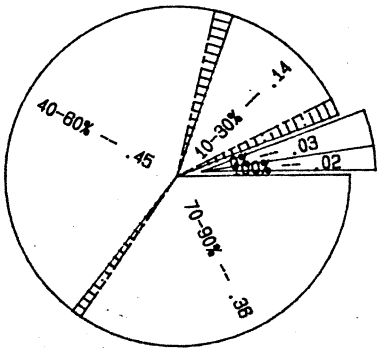


TOTAL # FORECASTS = 90

Guidance Forecast Scores



Guidance Frequency/Reliability



TOTAL # FORECASTS = 87