NOAA Technial Memorandum NWS SR-108

EVALUATION OF POINT PRECIPITATION PROBABILITY FORECASTS USING RADAR ESTIMATES OF RAINFALL AREAL COVERAGE

Pamela S. Naber and Daniel L. Smith
National Weather Service Southern Region
Fort Worth, Texas

Scientific Services Division
Southern Region
Fort Worth, Texas
November 1983

UNITED STATES
DEPARTMENT OF COMMERCE
Malcolm Baldrige, Secretary

National Oceanic and
Atmospheric Administration
John V. Byrne, Administrator

National Weather
Service
Richard E. Hallgren, Director

Evaluation of Point Precipitation Probability Forecasts Using Radar Estimates of Rainfall Areal Coverage

Pamela S. Naber and Daniel L. Smith

NOAA, National Weather Service, Southern Region
Fort Worth, Texas

1. Introduction

This study continues earlier investigations into the relationships between first-period probability of precipitation (PoP) forecasts prepared at WSFO Birmingham, Alabama, and radar estimates of areal coverage of rainfall in Southern Alabama and Northwest Florida. Forecast and observed data from the summer (June through August) of 1982 have been analyzed and results are compared with similar analyses from the summers of 1976 and 1977 (see Smith, 1977 and Smith and Smith, 1978; hereafter referred to as S77 and S78, respectively).

Evaluation of NWS precipitation probability forecasts using observations (or estimates) of precipitation areal coverage is a procedure not without some controversy. At issue is whether or not it is proper to attempt to verify a single probability forecast. We argue the answer is "Yes", and we base our approach on an aspect of NWS PoP forecasts which seems to be often overlooked, namely, that the PoPs apply to any given point within a zone, not just a single specific point. In a manner of speaking a single PoP forecast might be thought of as a number of forecasts (all the same) for many points in the zone. Thus, verifying that (those) forecasts with estimates of areal coverage amounts to verifying, in effect, several forecasts.

No significant changes were made in the PoP program between the time of the earlier studies and the present one, but there have been several changes in the forecasting staff at the WSFO. Repeating earlier analyses using data from the summer of 1982 provides an opportunity to compare current forecasts with those made several years ago and may reflect the extent to which knowledge gained from the earlier studies has been incorporated into operational procedures at the WSFO. For example, it was obvious in 1976 and again in 1977 that forecasters grossly overused probability values between 20% and about 50%, when frequencies of use of such PoPs were compared with frequencies of occurrence of similar rainfall coverages for zones. Was there any evidence in 1982 that a tendency to do so was checked?

An additional investigation was made using the 1982 radar data. Studies by Lopez, et al., (1983), using high resolution digital radar data acquired as part of FACE (Florida Area Cumulus Experiment) in 1978-80, show that 12-hour cumulative rainfall volume can be estimated for large areas with considerable accuracy using radar echo maps (PPI displays) separated by intervals greater than one hour. To test whether similar results might hold for areal coverage estimates from manually derived data we produced 12-hour composite echo maps using overlays at 2-, 4-, and 6 hour intervals and compared the results to composites based on hourly data. Results have significance in terms of the effort involved in carrying out a PoP coverage comparison in real time.

It will be useful to examine once again in Section 2 the data and analysis procedures used in this (and earlier) studies. Readers are encouraged to review S77 for an in-depth discussion of NWS PoP forecasts, point and areal probability concepts, and relationships between areal coverage of rainfall and PoP forecasts. Similar discussions are given by Winkler and Murphy (1976), Murphy (1978), and Smith (1979). While a brief review is given in Section 3, it will be assumed readers are knowledgable of these basic concepts regarding PoP forecasts. In Section 4 we consider summer long averages of PoP forecasts and areal coverage estimates. An analysis of averages reveals a number of similarities with earlier data and confirm earlier conclusions. In Section 5 we compare frequencies of use of individual PoP forecast values with frequencies of occurrence of corresponding coverage amounts. Significant differences in the frequencies are obvious. Section 6 presents results of specific comparisons of PoP forecasts and coincident coverage. This allows an evaluation of PoP forecast resolution and reliability, as judged against areal coverage. Finally, in Section 7 we briefly discuss results of coverage estimates based on fewer overlays.

2. Data

Fig. 1 shows the eight forecast zones, four in Alabama and four in the Florida Panhandle, which were used in this study. First period PoP forecasts for the 12-hour night (0000-1200 GMT) and day (1200-0000 GMT) periods were tabulated. The forecasts were issued at roughly 4 pm and 4 am, CST, a few hours before the start of the valid periods, by forecasters at WSFO Birmingham. Areal coverage of rainfall in each zone was estimated for the same 12-hour periods by compositing hand-drawn hourly radar overlays (Fig. 2) prepared by the staff at WSO Pensacola. Observations were made using a WSR-57 radar and the study area was restricted to roughly the 125 mile range of the radar. As in earlier studies the investigation was limited to a three month summer (June-August) period to emphasize convective activity and minimize effects of organized precipitation systems. One of the goals of the initial study (S77) was to develop improved procedures for forecasting PoPs for the summer "scattered shower" regime of the study area.

Echo coverage from the composite hourly overlays was determined visually to the nearest 10% with an estimated accuracy of ±10%. It should be kept in mind when comparing results from S77, S78 and the present study that coverage estimates were made by different individuals in each case. This introduces a possibility of some bias from year to year. Indeed, evidence for systematic differences is found in this study. In addition, accuracy of overlays, movement and changes of echoes between hourly overlays, correlation between echoes and measurable rainfall, and other factors all affect the accuracy of rainfall coverage estimates. These problems are treated in some detail in S77 and S78 and we will let it suffice here to note that earlier results argue in favor of using this simple approach to analyzing PoP forecasts.

PoPs were extracted from public forecasts issued for individual zones or, more commonly, combinations of zones. When zones are grouped it is assumed the same PoP applies to each zone. Because NWS procedures specify that no PoP, nor mention of precipitation, is to be included in a zone for which the PoP is less than 20%, we are forced to group all such forecasts under the heading "less than 20%". This includes forecasts which might have been logged (for verification purposes) as 0% or 10%. For comparison purposes, estimated coverages of 0% and 10% are also grouped as "less than 20%".

3. PoP Forecasts and Areal Coverage

In the following sections PoP forecasts will be evaluated by comparison with estimates of the areal coverage of precipitation. This implies that PoPs are forecasts of areal coverage rather than forecasts of point probability. Since this is a matter of some confusion to many forecasters, an explanation is in order. PoP forecasts, as used in the NWS, are defined to be average point forecasts of the probability of precipitation for specified areas (zones) and time periods (usually 12 hours). A forecast of the expected coverage of precipitation in an area, given that it rains at all, may have some value, but it is difficult to determine who the user might be. If we tell someone, for example, "If it rains today, the rain will cover 50% of the area", the listener's response is sure to be, "Fine, so what's the chance it will rain?" Presumably the listener still wants to know the likelihood it will rain on him! That is precisely what is provided by the point probability (P).

The coverage forecast described above is the expected, or conditional, areal coverage--call it A--and is related to the point probability (P) by the simple expression

$$P = C \times A$$

where C is the probability that it will rain somewhere in the forecast area (the areal probability). The product of the areal probability and the conditional areal coverage (C x A) is the unconditional areal coverage, which is the same as the point probability (P). (In addition to the references cited earlier, Hughes (1980) also discusses this relationship in considerable detail.) Before the fact, the forecaster may visualize an areal distribution of rain (A) he thinks likely from a given weather pattern, but he must always wrestle with the areal probability (C) to resolve his final PoP forecast. After the fact, regardless of whether it rained in the forecast area or not, the chance that a given point in the forecast area received rain is just the same as the areal coverage of rain in the area. In other words, for a given forecast period, Nature reveals to the forecaster (after the fact) the best forecast he could have made of the probability of rain at any given point in the area--the areal coverage.

We should emphasize that a PoP which fails to match the ensuing areal coverage cannot be termed automatically a "bad" forecast, any more than a forecast of 50% for "heads" on any given coin flip can be termed "bad" when the coin comes up "tails" (or "heads", for that matter). All things considered, a 50% forecast for either face of the coin is a reasonable forecast. In fact, any other probability would rely on sheer luck--unless the forecaster knew of something amiss with regard to the coin or the flipping process! We can certainly argue, however, that consistently calling "heads" or "tails"--correctly--results in the best outcome...especially if a wager is involved!!

Evaluating point probability forecasts by comparing them with areal coverage estimates has the advantage of returning useful information about individual forecasts. This is not possible with the use of conventional statistical measures such as the Brier Score, for which a set (usually a season or longer) of forecasts must be evaluated. It was one of the initial purposes of S77 to develop procedures by which evaluation of today's PoP forecast could be integrated into an improvement in tomorrow's forecast. To some extent we succeeded, as discussed in S78, but more significant results came from the broader perspective on PoP forecasts provided by the new evaluation procedure.

4. Summer Averages of PoP Forecasts and Radar Coverage

Before making specific comparisons between PoP forecasts and corresponding rainfall coverages, it's useful to examine summer averages. For such purposes we've considered the average of all PoPs grouped as "less than 20%" to be 5%. This is reasonable since, from an examination of PoPs recorded for verification, there appear to be about as many 0% forecasts as 5% forecasts. Table 1 shows results from 1982 data along with corresponding results provided in similar tables in S77 and S78. The table shows averages--for the four Florida zones, the four Alabama zones, and all eight zones combined--of:

a. Forecast probability (PoPs).

b. Areal echo coverage on all days (and nights).

c. Areal echo coverage on days (and nights) when an echo occurred somewhere in a zone, regardless of coverage.

d. Frequency of occurrence of an echo somewhere in a zone.

e. Observed rainfall frequency (averaged for available raingages in the zones).

A cursory examination of Table 1 reveals three areas of notable correspondence between this and earlier studies. First, for both night- and daytime periods the observed rainfall frequencies (e.) were about the same in 1976, '77, and '82. A curious exception was the Florida area during the day. Rainfall frequencies appeared low (∿25%) in previous study years when compared with radar estimates of coverage. This was attributed to the fact that afternoon showers were typically inland, while coastal locations provided the observed rainfall frequencies. Daytime frequencies were much higher (∿40%) for similar stations in 1982. A simple answer is that showers were more homogeneously distributed within the coastal zones during the daytime periods of the latest study. No attempt was made to confirm this by close examination of radar data from individual days.

Second, averages of forecast PoPs (a.) were consistent for the zone groups in all three summers. In fact, 1982 averages were midway between 1976 and '77 averages. If rainfall is homogeneously distributed in the zones, and if the PoP forecasts are reliable, the average of PoP forecasts should equal the average of observed rainfall frequencies. Once again, in the current study, they do--to an excellent degree--for the daytime periods. Nighttime PoP forecast averages (as in previous years) exceed observed rainfall frequencies by about 5%. We continue to attribute this to a tendency on the part of forecasters to overestimate the persistence and coverage of showers after nightfall. Additional evidence for this is presented later. A subjective evaluation of hourly rainfall observations, made while determining rainfall frequencies, suggests that most of the time when rain occurred during nighttime periods, it fell in the first few hours after 0000 GMT. Rain during the period--depending on coverage--calls for a PoP value in the forecast, but the rainfall observations seem to be consistent with a fairly high incidence of lingering showers which diminish rapidly in coverage soon after the start of the nighttime period. They might easily lead a forecaster to "leave a mention of rain in the forecast...just in case."

A third area of similarity between the three summers is the high incidence of rain <u>somewhere</u> in the zones (d.). As in previous years it rained somewhere in the zones on about 75-85% of the days and 65% of the nights, based on radar estimates. A moment's reflection will show that average occurrence of rain in a zone, without regard to coverage, coincides with the average areal probability (C) discussed in Section 3. This statistic is particularly significant for what it reveals about daytime periods. If, in effect, it rains somewhere in the zone almost every day ($C \rightarrow 1$), the forecaster's job reduces to one of simply (!) forecasting areal coverage (i.e., $P \rightarrow A$). Information about areal coverage is contained in (b.) and (c.) in Table 1.

When we consider these radar estimates of areal coverage we see significant differences beween the 1982 data and that from previous studies. Coverage estimates, from both periods and both zone groups, are about 5-10% lower in the present study than they were in 1976 and '77. This is true for days on which there were echoes in the zones (c.) and all days (b.). The most obvious explanation of this difference is sytematic bias in estimation of coverage. We can get some idea of the accuracy of the echo coverage estimates by comparing averages with an independent measure...rainfall frequency. It is obvious that averaged over all days the frequency of occurrence of rain at any point (e.) must coincide with the average areal coverage of rain, which we estimate from radar in (b.). Notice that in previous study years nighttime coverage estimates coincided almost perfectly with rainfall frequencies (15-17%). Rain frequencies were similar in 1982, but coverage estimates were about 6% lower, suggesting radar coverages were, in fact, estimated systematically low by at least a few percent. Similar problems are seen in the daytime estimates, particularly for the Alabama zones. The large discrepancy between average coverage (25%) and rainfall frequency (39%) may reflect a particular difficulty in making estimates in the typically higher coverage situations during the day. In any event, this possible bias should be kept in mind in later sections.

Since only periods in which an echo occurred somewhere in a zone were used in forming the average coverages in (c)., these averages represent conditional areal coverages (A). As we have seen, the values in (d.) are just the average areal probabilities (C). The product of these two should be the unconditional areal coverage (or P). In fact, for both groups of zones, night- and daytime periods, products of the values in (c.) and (d.) equal values in (b.), the average echo coverages on all days (or unconditional coverage). Do the average forecast PoP values agree with these estimated unconditional coverages? As in earlier years, the answer is "yes and no".

For daytime periods average PoPs are about 5% higher than coverage estimates. As already noted, coverage estimates might be slightly low, so correspondence appears to be reasonably good. Even better correspondence was seen in earlier summers. Notice also that average PoPs compare very well with average rainfall frequencies, as expected. A significant difference is seen in nighttime periods, however, as average PoPs are significantly higher than coverages estimated from all days, even if the coverage estimates are inflated to account for low estimates. Similar results can be seen in data from earlier years and this is probably another indication of a "high bias" in nighttime PoPs. It's worth noting that the average nighttime PoPs compare more favorably with the average conditional areal coverages than with the unconditional coverages. Are forecasters indicating with their PoPs (perhaps subconsciously) the chance of rain, given that rain continues somewhere in the zone?

5. Comparison of Frequencies: PoP Use and Observed Coverages

The frequencies with which forecasters used various PoP values for night- and daytime periods in each area during the summer of 1982 are shown by the broken lines in Fig. 3. Also shown are frequencies of occurrence of various areal coverages, which will be discussed below. When the PoP frequencies are compared with similar data from earlier studies the differences are minimal; in fact, differences do not warrant inclusion of combined graphs as shown in S78. In each of the three study years frequency of use statistics show a broad hump during daytime periods between 30% and about 50% (extending to 60% for Florida zones in 1982), with a sharp decline in use of PoPs above 60%. PoPs of 90% or higher were almost never used. For nighttime periods, frequency of use of PoP values shows a more-or-less monotonic decline from 20% to 70 or 80%, with PoPs higher than 80% used only a few times in the three study years. The effect of the climatological rain frequency for the study area (about 20% at night, 35-40% during the day) on frequency of use statistics is obvious. Forecasters tend to cluster PoP forecasts around these values.

As will be discussed in greater detail below, the frequencies of occurrence of higher areal coverages indicate that similarly high PoPs are appropriate for the study area. This was first pointed out in S77. It was tempting to speculate (S78), when data were analyzed for the second summer, that slight increases in use of higher PoPs over the previous summer reflected forecasters' appreciation of this. Data from 1982 do not support that assumption, however. Most likely, differences from year to year are not significant and there still exists a bias toward lower PoPs.

When the areal coverage frequency of occurrence graphs are examined (solid lines in Fig. 3) two features are obvious for both areas and both periods...from 20% to 100% each decile coverage is about equally likely, especially during the daytime; and coverages of less than 20% are by far most common. (Since we've already seen that two-thirds of the nights and three-fourths or more of the days have at least a single shower somewhere in the zone, we may conclude that "isolated showers" is probably the single most appropriate forecast for the study area.) As with the PoP frequency graphs, frequency of occurrence graphs for the latest study period are not significantly different from those presented in S77 or S78, except that they are noticably flatter (i.e., there is less variation of frequencies from decile to decile).

The real significance of Fig. 3 lies in comparison of the sets of curves on each graph. It is obvious at a glance that forecasters use PoPs below 20% much less frequently than similar coverages occur; use PoPs in the mid-ranges (20%-60%) too often; and essentially fail to use highest PoPs at all, even though high coverages are as common as low coverages. Since individual forecasts are not compared with coincident coverage we cannot judge the extent to which over- or underforecast bias exists. One might suspect from Fig. 2, however, that both reliability and resolution of the PoPs examined could be improved. Similar results from 1976 and '77 were analyzed exhaustively in the earlier studies and conclusions will not be repeated here except to call attention to either end of the frequency of use curves.

At the upper end, it might not be surprising that PoPs above 80% are almost non-existent if one considers the forecaster must be virtually certain rain will occur in the zone ($C \rightarrow 1$) and expect almost total coverage ($A \rightarrow 1$) to issue a 90% PoP. The first may not be too difficult (in light of what we've seen about the average areal probability), but a conception of the showery nature of summertime rainfall (and mental image of all the rain-free areas on a radar display, even when numerous showers are in evidence) may make it difficult for the comprehensive forecaster to convince himself rain will cover the entire forecast area! Indeed, many forecasters have expressed much surprise at the results revealed by the radar coverage studies. At the lower end of the graph, to some extent the underuse of 0% and 10% PoPs may actually arise from overestimation of the area affected by isolated showers. We continue to believe a more likely explanation is the "20% threshold" which precludes mention of rainfall (or PoP) in the public forecast unless the PoP is at least 20%. For several reasons a forecaster may feel the need to "inflate" a lower PoP to 20% in order to include a mention of rain in the forecast. For instance, there seems to be a common feeling among forecasters that the public expects to hear some mention of rain in the forecast if they hear a radar summary which describes "isolated showers". To avoid the need for a later update, forecasters may accommodate them by adjusting the PoP in their forecast (released in the morning) upward accordingly. (After 1982 PoP forecasting procedures were changed slightly in an attempt to address this problem. Under certain circumstances forecasters are allowed to mention "isolated showers" and include PoPs of less than 20% in their public forecasts. A very preliminary examination of Birmingham forecasts from the summer of 1983 indicates that, indeed, frequency of use of PoPs below 20% doubled over that for 1982.)

6. Radar Indications of PoP Resolution and Reliability

We have examined average characteristics of PoP forecasts and areal coverages and seen that patterns of over- and underuse of certain PoP values can be inferred from coverage estimates. How did individual PoP forecasts compare with coincident estimates of coverage? Fig. 4, similar to corresponding figures in S77 and S78, shows the average of coverages which were observed for all instances of a given PoP forecast (solid lines) and the average of PoP forecasts which were made for each decile of coverage observed (broken lines). For example, considering all 60% daytime PoPs for the Alabama zones (upper left graph), the average of observed coverages for those forecasts was about 53%. For all periods in which 60% coverage was observed, the average PoP forecast was about 44%.

The graphs in Fig. 4 provide a significant amount of diagnostic information about PoP forecasts. The broken lines graphically show prefigurance of the forecasts--indicating at a glance how well given degrees of areal coverage were anticipated by the forecasts. They also reveal something about resolution of the forecasts--showing how well forecasters were able to resolve various areal coverages. The solid lines graphically depict post-agreement by showing how well given PoP values verified against subsequent areal coverage. The degree to which the plotted points fall along the diagonal line also reveals reliability of the forecasts. Even though we plot average areal coverage against PoPs the graphs are actually the same as those normally used which show frequency of occurrence of precipitation at a point plotted against PoP values. This is because, as shown in Section 4, frequency of precipitation at a point and average areal coverage are the same as long as rainfall is homogeneously distributed in the zones.

First, let's examine resolution of the PoP forecasts. An obvious feature of the daytime graphs (broken lines) is the fact that, as coverages range from 20% to 100%, the average PoPs vary only from 40% to 60%. PoPs do show a definite (but slight) tendency to increase with increasing coverages, but overall ability to resolve coverage is limited. Resolution of coverage at night is almost totally lacking. For Florida zones nighttime PoPs seem to average 20% to 40% throughout the full range of coverage, and for Alabama zones the trend is actually for average PoPs to decrease as coverage increases. These nighttime results comprise one of only a couple of significant differences beween this study and those of earlier summers. In 1976 and '77 both zone groups showed a positive--but very slight--trend for average PoPs versus coverage during the night period. The other difference is seen in average PoPs for highest coverages (90-100%) in Alabama zones during the day. Averages were about 15% higher in the present study than in earlier years--a positive sign which suggests a slightly better resolution for highest coverage days.

The failure of PoPs to average more than 50-60%, even on days when rain covers the entire forecast zone, is consistent with results obtained by Murphy (1977) who used a dense network of raingages to evaluate PoP forecasts for Rapid City, South Dakota, and St. Louis, Missouri. Even though there is considerable room for improvement, the trends shown by the averages (for daytime periods, at least) are encouraging. As pointed out earlier, it's very difficult to assign a high PoP in the convective regime of the study area during the summertime. At the lower end of the PoP scale, operational considerations sometimes argue for a PoP higher than might otherwise be appropriate. Nevertheless, it is hoped this and earlier studies, by calling to attention these areas, might lead to improvements at either end of the PoP range.

Another aspect of resolution is provided by analysis of variability of coverage over the study area and comparison with corresponding PoP forecasts. To what extent do forecasters attempt (successfully or otherwise) to apply different PoPs to zones in the forecast area? How often is coverage significantly different from zone-to-zone? Table 2 provides answers to these questions by showing the frequencies with which 1-, 2-, 3-, or 4 significantly different coverages (PoPs) were observed (used) in the study areas. We should note that forecasters often avoid slight (10%) differences in PoPs between adjacent zones because of operational considerations--they might thus combine zones and reduce overall workload as well as minimize already heavy communications traffic. Thus, we considered a PoP difference of 20% or more to be significant. Similarly, we considered a 20% difference in coverage estimates from one zone to another to be significant (also acknowledging the accuracy of coverage estimates). Table 2 reveals that during summer days in 1982 each of the four-zone Alabama and Florida areas experienced <u>three</u> significantly different coverages on about one-fourth of the days. Correspondingly, forecasters never assigned three significantly different PoPs to zones in either group. In fact, the same PoP or PoPs differing by no more than 10% were applied to all four zones over 90% of the time, but zones had essentially the same coverage only one-fourth to one-third of the time! Observed variability among the zones was less during nighttime periods, but two- and three different coverages were still observed a significant number of times. Forecasters, on the other hand, even less often deviated from a single PoP.

Two problems are indicated by these statistics. First, it's obvious that PoP forecasts failed to indicate the variability of coverage which, most of the time, occurred in the forecast area--certainly during daytime periods. A major reason for this is non-meteorological and involves procedures, as indicated above. It's easy to see, however, that to whatever extent forecasters apply the same PoP to a four-zone area, it must be that much harder to utilize very high PoPs with anywhere near the frequency with which such coverages occur in the zones.

The second problem appears when data from this study are compared with those from earlier years (S77 and S78). Observed variability of coverage was very similar in each of the three summers studied. For PoPs, however, a significant change was seen from the first to second study years in that frequencies of use of two and three different PoPs increased markedly...approaching the frequencies of occurrence. We attributed this to a realization, from the first year's study, that such variability was common, thus an attempt to reflect it in the forecasts the second year. The summer of 1982, however, was the most conservative of all years in terms of variability of PoP use! Over 90% of the time forecasters either a) used a single PoP for all zones, or b) expressed anticipated variability over a four-zone area by using PoPs which differed by no more than 10%. This is an unfortunate turn of events and probably contributes in no small way to limits on skill reflected in the graphs representing resolution in Fig. 4.

Now, what of radar indications of reliability of the PoP forecasts? Those accustomed to ordinary reliability diagrams will notice from a glance at Fig. 4 that daytime PoP forecasts for both groups of zones were quite reliable during the summer of 1982. The solid lines in the top graphs of the figure closely follow the diagonal for most PoP values. The obvious diversion for 80% PoP forecasts for Florida zones is worth mentioning. For one thing, we concluded, based on results of the earlier studies, that forecasters seemed to be able to make more reliable forecasts at high PoPs for Florida zones than for Alabama zones. While there were no 80% or higher PoPs for Alabama zones in 1982, those for Florida were not very reliable! On closer inspection, there were only two 80% forecasts in 1982, so care should be taken not to put too much weight on the average coverage. Overall, PoPs for the Alabama zones consistently overestimated coverage by about 5-10%, but some of this might be attributable to a possible systematic bias in underestimating coverage from radar overlays (discussed earlier). The plot of data from the Florida zones, on the other hand, shows an "S" shaped curve which is often seen for PoP forecasts. Below roughly the climatological rainfall frequency the PoPs are too high, above they are too low. Daytime reliability curves from the latest study compare very favorably with similar curves from each of the earlier studies.

Reliability curves for the nighttime period--particularly for the Alabama zones--confirm earlier indications that PoPs were not well chosen for antecedent coverage conditions. Interestingly, the data for 1982 are very similar to those from 1977, but differ markedly from 1976, during which nighttime PoPs were seen to be just as reliable as those from the daytime period. In 1977, and again in the present study, nighttime forecasts overestimated coverage by about 10-15% for PoPs up to about 40%, then grossly overestimated from 50% to 60%. This result is difficult to explain, but it may stem simply from overestimation of the longevity and areal extent of daytime activity which persists into the nighttime period. This cannot be the entire answer, however, because an inspection of individual forecasts for the Alabama zones showed that there were thirty-six PoP forecasts of 50% to 70% and these were preceded by daytime coverages of from less than 10% to near 100% in the subject zones. Coverage during the day, which might have influenced the forecaster's high PoP for the nighttime period, exceeded 50% on only about half of the thirty-six days. We know from Fig. 3 that forecasts of 50% to 70% were made for the nighttime periods with about the same frequencies with which such coverages were observed. It's clear from Fig. 4, however, that the two failed to coincide with anything like the success rate shown for similar PoPs during the day.

7. Dependence of Areal Coverage Estimates on Time Interval of Sampling

Coverage estimates for this study were determined from archived overlays long after the fact. The compositing procedure is not a complicated one and can be accomplished soon after the close of the 12-hour period to improve utility of the data, as demonstrated in S78. For operational considerations it would be useful to minimize the effort even further (short of fully automating the process...easily accomplished at those very few sites equipped with appropriate hardware). With this in mind we compiled areal coverages for intervals of 2-, 4-, 6- and 12 hours for day and night periods. In other words, we compared 12-hour coverage estimates from overlays separated by these intervals with those obtained from overlays at 1 hour intervals. For example, an estimate based on overlays at 4 hour intervals (three) would have combined echoes from 1530Z, 1930Z and 2330Z for the daytime period. Obviously, estimates made from fewer overlays cannot show more total coverage than those from the full 12 overlays, so the real question is, how many overlays can be eliminated--for summer rainfall regimes--before a significant loss of coverage is seen? Not too surprisingly, the answer is, very few!

12-hour coverage estimates based on every other hour's overlay comprised only about 80% of the coverage derived from hourly data for the daytime periods. For nighttime periods the percentage dropped to about 70%. Estimates based on three overlays (4-hour separation) failed completely to show a reasonable representation of the "true" coverage. Since spatial and temporal distribution of summertime convection in the study area is known to vary considerably--depending primarily, it seems, on prevailing low level winds--we could easily anticipate these results. Smith and Henderson (1977) used data from the initial year's study to show that distinct rainfall regimes existed. With this in mind, we might expect that fewer overlays could more accurately depict the total coverage *if* they were chosen at proper times during the period (during the afternoon, for example).

We conclude that even though rainfall volumes might be accurately estimated using data at intervals greater than one hour (Lopez, et al, 1983)--for areas and seasons climatologically similar to our study area--the same is not true for rainfall coverage estimates.

8. Summary

An initial study in 1976 showed that radar estimates of summertime rainfall coverage--derived from hand-drawn hourly overlays--could provide useful information about both true rainfall distribution and PoP forecasts. In 1977 we tested the possibility of obtaining coverage estimates in near real-time, i.e., estimates of daytime coverage prior to formulation of forecasts for the subsequent daytime period, in order to increase the utility of the coverage estimates for forecast purposes. The present study, six years after the first, was primarily for benchmark purposes to evaluate differences, if any, in characteristics of coverage (or estimates of coverage) and PoP forecasts. We also tested the possiblity of making useful estimates by an even simpler procedure of using fewer overlays. Results showed this to be not useful. Composites of radar overlays at intervals of less than one hour--at least for summertime rainfall--do not yield sufficient information about coverage for 12-hour periods.

Results of analyses showed major differences in forecasts for the summer of 1982, compared to earlier study periods. On the whole, the forecasts did not appear to have either resolution or reliability equal to that demonstrated earlier, although for some zones and periods the reverse was true. Because the possibility exists some bias may be reflected in areal coverage estimates from year to year, results are not strictly comparable and caution is advised. In general, however, the following results appear to be well-founded after three years of PoP/coverage comparisons:

--Inter-zonal variability of coverage (differences of 20% or more) is common...as many as three significantly different coverage values characterize each four-zone group studied on as many as 25% of the daytime periods. PoP forecasts--particularly in 1982--failed entirely to reflect this.

--It rains somewhere in a zone-sized area in southern Alabama and northwest Florida almost every summer day (frequency 80% or more) and it rains on most of the nights (frequency about 60%). For daytime periods PoP forecasts can be thought of as estimates of expected areal coverage <u>or</u> point probabilities.

--Even though summer rainfall is predominately convective in the study area, coverage in the zones is just as likely to 100% as 20%. The likelihood of occurrence of any decile coverage between 20% and 100% averages about 5%. Coverages less than 20% dominate and occur more than half the time.

--PoP forecasts of 30% and less are used much more frequently than corresponding areal coverges are observed, while PoPs of 80% or higher are virtually never used in the summertime--even though such high coverages are not uncommon, even at night. This results in poor resolution of summer forecasts. As coverages increase from 20% to 100% the averages of coincident PoP forecasts remain between 40% and 60%.

References---

Hughes, Lawrence A., 1980: Probability forecasting - Resources, procedures, problems. NOAA Tech. Memo. NWS RCST 24, 84 pp.

Lopez, Raul E., Jack Thomas, David O. Blanchard, and Ronald L. Holle, 1983: Estimation of rainfall over an extended region using only measurements of the area covered by radar echoes. Proc. 21st Radar Meteorology Conf., Boston: American Meteorological Society.

Murphy, A. H., 1978: On the evaluation of point precipitation probability forecasts in terms of areal coverage. Mon. Wea. Rev., 106, 1680-1686.

Murphy, A. H., 1977: Personal communication.

Smith, D. L., 1979: Comments "On the evaluation of point precipitation probability forecasts in terms of areal coverage". Mon. Wea. Rev., 107, 1224-1225.

Smith, D. L., 1977: An examination of probability of precipitation forecasts in light of rainfall areal coverage. Nat. Wea. Dig., 2, No. 2, 15-26.

Smith, D. L., and M. Smith, 1978: A comparison of probability of precipitation forecasts and radar estimates of rainfall areal coverage. NOAA Tech. Memo. NWS SR-96, 17 pp. [NTIS PB279987/AS].

Smith, D. L., and K. Henderson, 1977: Summertime rainfall regimes in southern Alabama and northwest Florida as deduced from wind-stratified radar data: preliminary study 1976. NOAA Tech. Memo. NWS SR-92, 15 pp. [NTIS PB275183/AS].

Winkler, Robert L., and Allan H. Murphy, 1976: Point and area precipitation probability forecasts: some experimental results. Mon. Wea. Rev., 104, 86-95.

Table 1. Averages derived from forecast probabilities and observed echo coverages. Data are for years 1976/1977/1982.

	NIGHT Fla	NIGHT Ala	NIGHT Comb	DAY Fla	DAY Ala	DAY Comb
a. Avg forecast PoP	20/26/23	20/25/23	20/25/23	32/40/37	29/36/34	30/38/36
b. Avg covg echoes, all days	16/17/12	17/17/11	16/17/12	35/45/36	30/38/25	32/41/30
c. Avg covg echoes, rain days*	31/27/19	29/24/17	29/25/18	46/51/43	40/45/32	42/48/37
d. Avg freq occrnc of echo	52/63/64	59/71/67	56/67/66	77/88/84	76/85/77	76/87/81
e. Avg rain freq	13/16/18	16/18/18	15/17/18	25/22/40	31/37/39	28/30/39

* Note: Rain days are periods during which an echo occurred somewhere in a zone, regardless of coverage.

Table 2. Frequencies of observation (use) of areal coverage (PoP) values separated by 20% or more. See text for details.

	Values observed / PoPs used 1	2	3	4
Night				
AL Zones	66 (96)	24 (4)	10 (0)	0 (0)
FL Zones	60 (96)	31 (4)	8 (0)	1 (0)
Day				
AL Zones	36 (92)	38 (8)	25 (0)	1 (0)
FL Zones	26 (95)	48 (5)	24 (0)	1 (0)

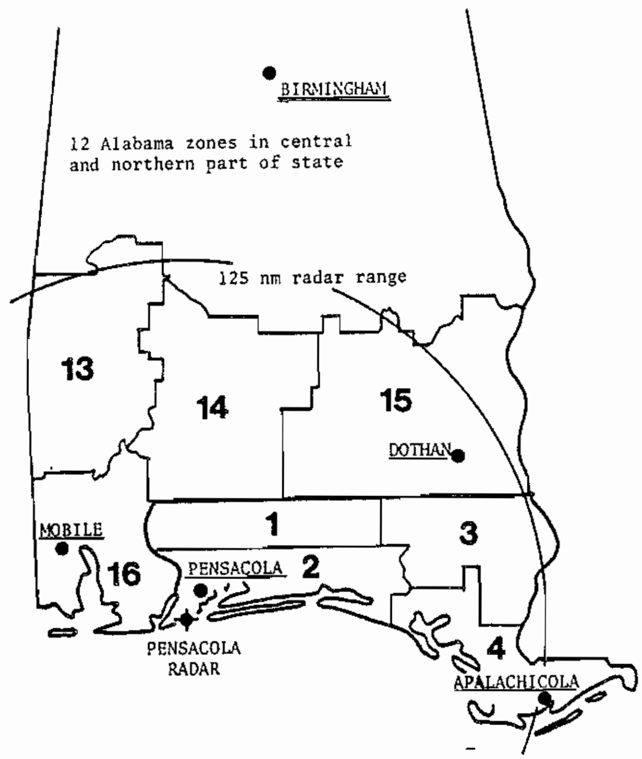

Figure 1. Florida and Alabama forecast zones used in this and earlier studies.

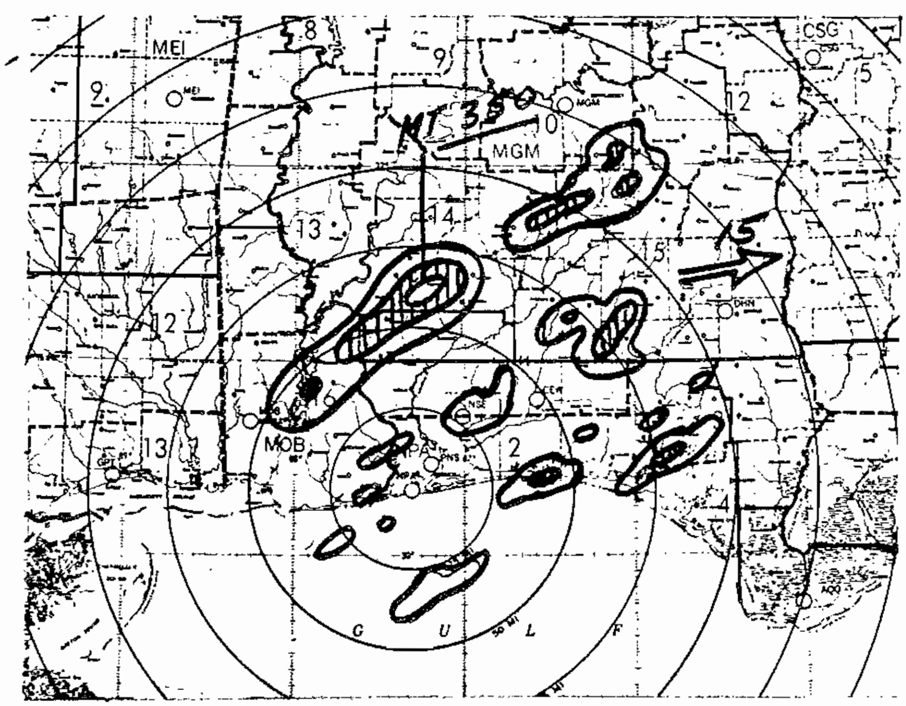

Figure 2. Manual analysis (radar overlay) typical of those produced hourly by operators of network radars.

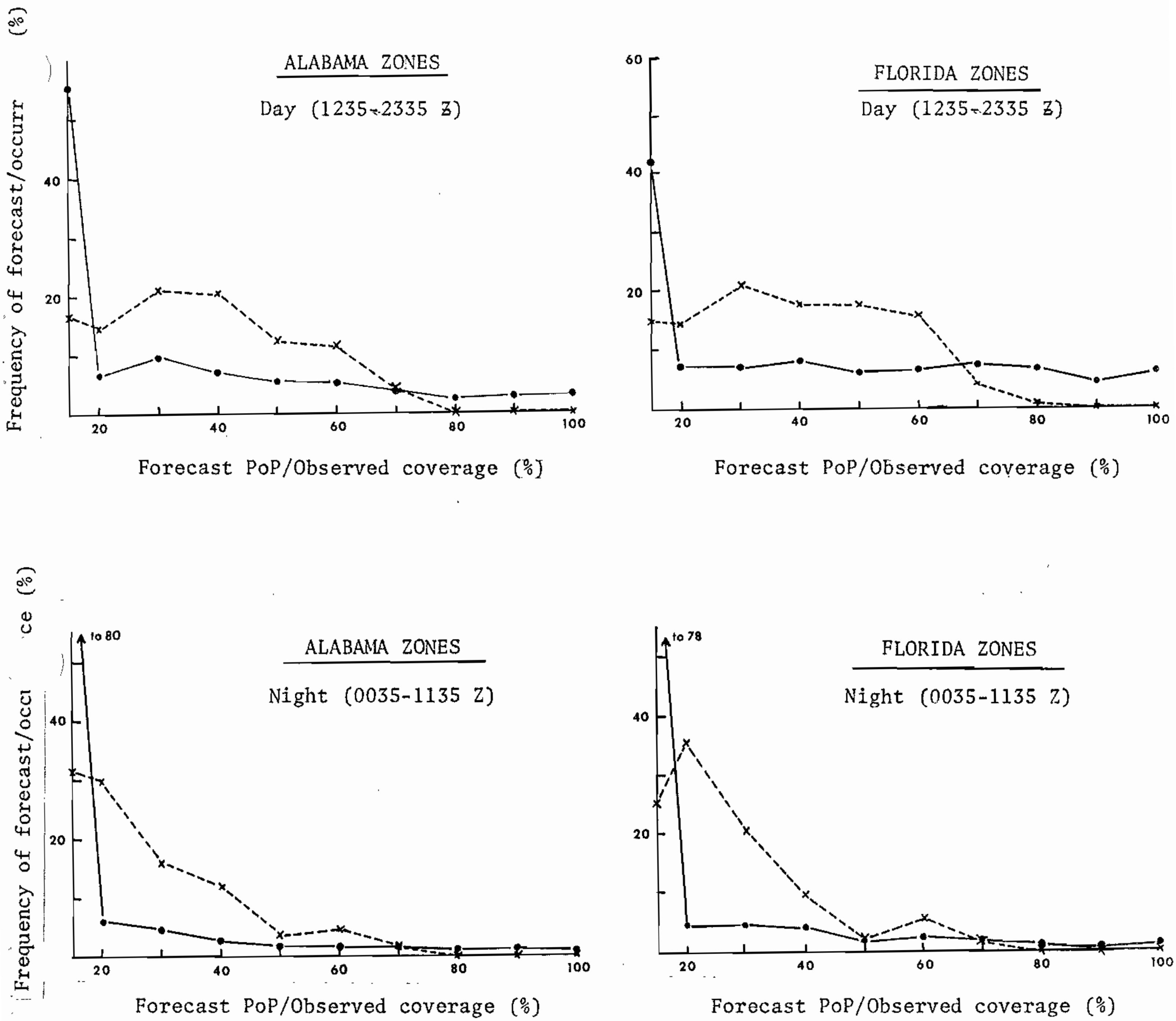

Figure 3. Frequencies of use of various PoP forecast values (broken lines) and frequencies of observation of various estimated values of rainfall coverage (solid lines) for summer (June-August) 1982.

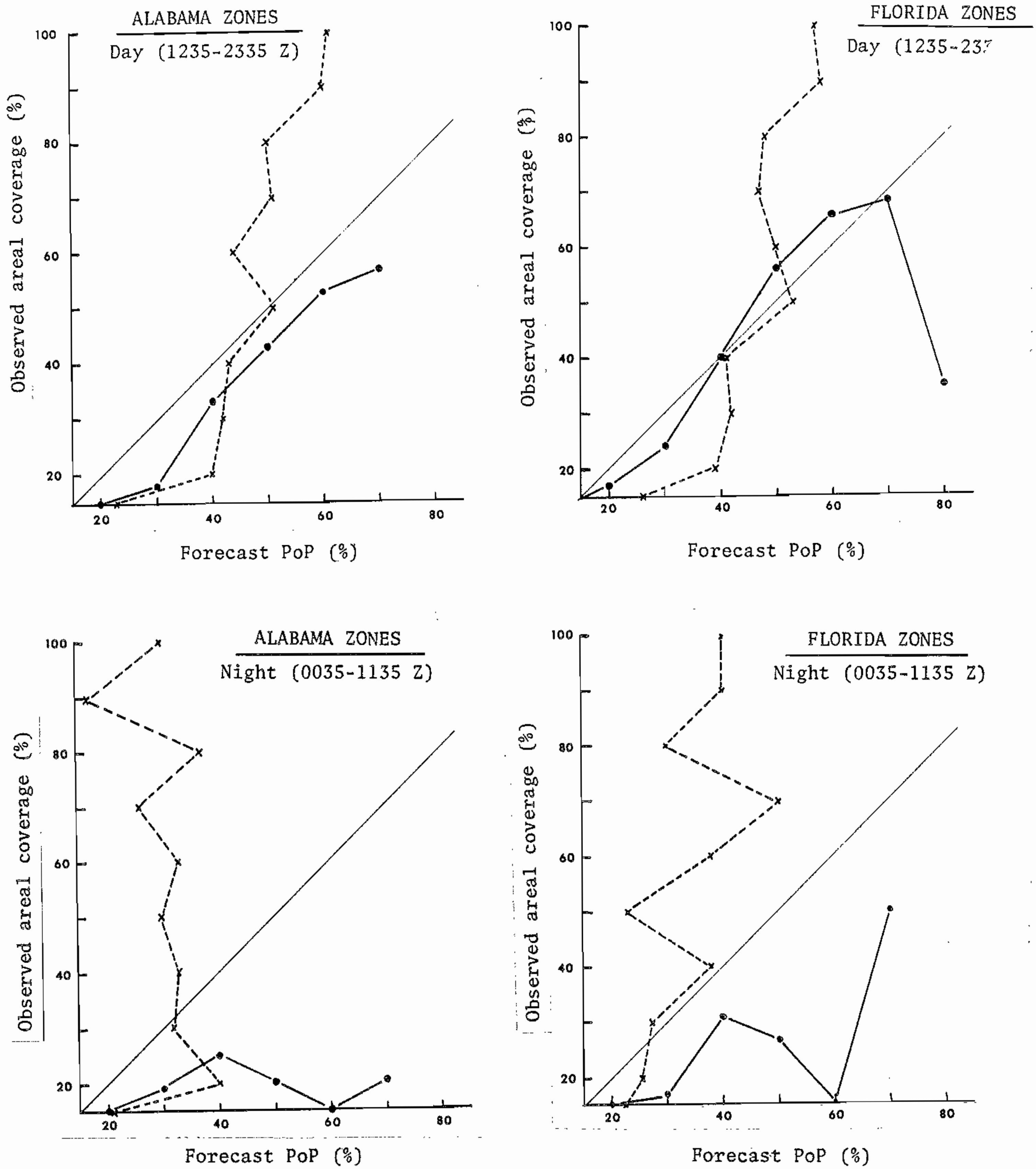

Figure 4. Average forecast PoP for various values of observed areal coverage (broken lines); average observed coverage for various values of forecast PoP (soild lines) for summer (June-August) 1982.