scale in Fig. 1, the regression lines yield 10-year increases in skill in the NWP 36 h sea level and 500 mb circulation forecasts of 11.5 and 13.5%, respectively. The corresponding increases in skill of the second and third period PoP forecasts (Fig. 5) are 6.9 and 7.4%, respectively.

While it is tempting to use these results as a quantitative comparison of skill trends for circulation and PoP forecasts, such a comparison may be inappropriate due to the different nature of these forecasts (atmospheric pressure or height versus precipitation) and the different scoring rules used (S_1 versus Brier scores). Suffice it to say that one would expect the improvement in large-scale circulation forecasts to be greater than that for precipitation forecasts—because of the difference in scales involved—and the comparison qualitatively supports this hypothesis.

As noted previously, most of the improvement in PoP skill has been achieved since the early 1970s. Two additions to the NWS centralized guidance system at NMC are believed to be significant contributors to this nationwide upturn. One was the introduction of the LFM model in 1971, with subsequent extensions and refinements as noted earlier. The other was the introduction of model output statistics (MOS) forecasts of PoP in 1970 (Glahn and Lowry, 1972; Lowry and Glahn, 1976). The MOS forecasts gradually improved in subsequent years as the NWP output on which these statistical forecasts are based became more accurate and longer data samples became available for development. Another change that may have contributed to forecast improvement during the 1970s was a reorganization within the NWS, increasing the number of WSFOs from 26 in 1968 to 50 in 1972 (Cressman, 1979, personal communication). It should be noted that other factors such as new or improved data (e.g., satellite measurements), new understanding or insights, and increased forecaster experience with PoP also may have contributed to these forecast improvements.

Snellman (1977) and Klein (1978) have shown that local forecasters in the Western Region of the NWS improve more on centralized MOS PoP guidance than forecasters in other parts of the country. Figure 5 shows the nationwide annual skill (in terms of Brier Score improvement over climatology) of the MOS PoP guidance for comparison with the local subjective scores. We see that the MOS and local curves generally are similar to one another. For instance, the linear correlation coefficients between the two products in the second and third periods are 0.84 and 0.90, respectively. In the first period, local forecasters clearly improved on the guidance, evidently by making effective use of observed data several hours more recent than those used in the preparation of the guidance. Nevertheless, the two curves retain a good deal of year-to-year similarity with a linear correlation of 0.69.

Close inspection of the comparative curves in Fig. 5 suggests that in recent years local forecasters have been achieving a margin of improvement over guidance in the second and third periods, whereas in the first period, they apparently have been losing some of their earlier advantage. In fact, these recent tendencies are reflected in the overall trends in the local and guidance scores. For instance, overall trend lines in the first period yield 10-year improvements of 8.9 and 2.0% for the guidance and locals, respectively; in the third period this ranking of trends is reversed with 10-year im-

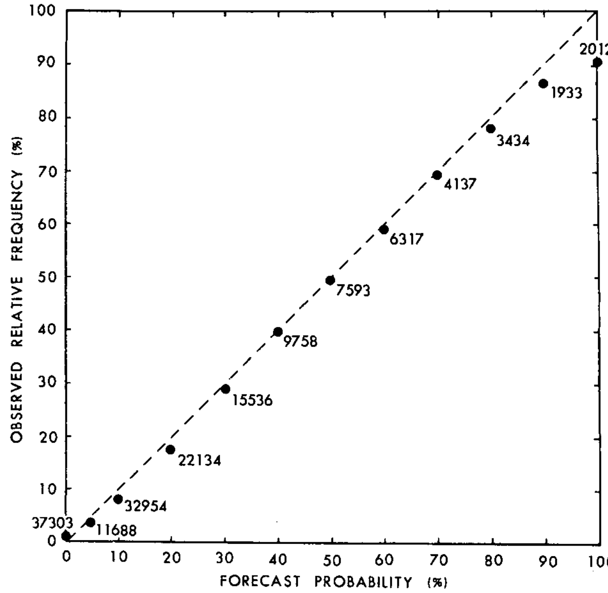

FIG. 6. Reliability plot of local PoP forecasts for 87 United States stations. The sample includes forecasts issued once daily (from 0000 GMT) for the three periods and only the last two years in Fig. 5. The probability values used in local PoPs are 0, 5, 10, 20, 30, . . ., 90, 100%. Numbers next to the plotted points are the sample sizes.

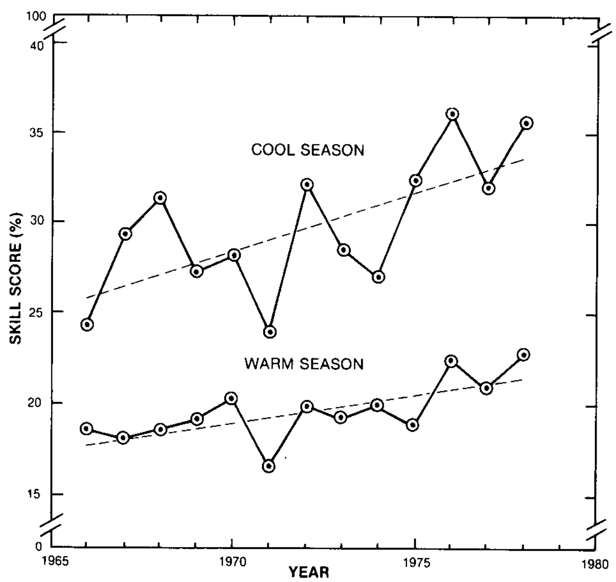

FIG. 7. Warm (April–September) and cool (October–March) season local PoP skill scores. The plotting convention for the cool season scores is as in Fig. 4 and trend lines are dashed as before.