Does Distance from the Forecast Site Affect Skill?

PAUL J. ROEBBER

Department of Geosciences, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin

LANCE F. BOSART

Department of Atmospheric Science, State University of New York at Albany, Albany, New York

GREGORY S. FORBES

Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

(Manuscript received 8 April 1996, in final form 15 July 1996)

ABSTRACT

An outstanding issue in the assessment of forecast skill (and value) is whether any advantage that can be obtained through regional knowledge not readily available to distant forecasters supersedes the leveling effect of information obtained by all forecasters (through the interpretation of numerical weather predictions). An analysis of 1 yr of data from the National Collegiate Weather Forecasting Contest was conducted in order to evaluate whether physical separation from the forecast site (defined by distances outside of and within 1000 km, with a minimum separation of 100 km) has a measurable effect on skill. The results indicate that regional effects (on the meso-α scale) are manifested in forecasts of both temperature (maximum and minimum) and precipitation amount (by category). Furthermore, these effects are a function of the experience level of the forecaster. Specifically, experienced forecasters are able to use regional knowledge to their advantage in forecasting temperature and precipitation amount, while their less-experienced counterparts cannot advantageously use such information for either type of forecast. The implication of these results with respect to the allocation of National Weather Service resources is also addressed.

1. Introduction

Roebber and Bosart (1996a) recently studied the effect of education and experience on the skill of local forecasts of maximum–minimum temperature and probability of precipitation (POP) and found that experience was a strong factor in determining skill. Roebber and Bosart (1996a) also documented the continued downward trend in the relative skill advantage of human forecasters with respect to the model output statistics (MOS) and found that the advantage, where it still existed, largely reflected the ability of human forecasters to recognize those instances in which the MOS approach does not account for bias that is specific to certain synoptic situations. In light of these findings, Roebber and Bosart (1996b) found that across a range of real-world user contexts these kinds of routine local meteorological forecasts possessed considerable value but that human intervention in making those forecasts has led to only minimal gains in value beyond that which is obtainable through the direct use of numerical–statistical guidance.

Given that skill and value advantages appear to be obtained through regional or local knowledge concerning model biases, a question that needs to be addressed is whether physical separation from the forecast site has a measurable affect on skill. At issue is whether any advantage that can be obtained through regional knowledge not readily available to distant forecasters supersedes the leveling effect of information obtained by all forecasters through the interpretation of numerical weather predictions. By regional knowledge, we mean that forecasters are likely to have been examining more carefully synoptic-scale weather systems within about a day's translation to their location than systems more distant. Thus, these forecasters are more likely to perceive any meso-α-scale phenomena and other peculiarities likely to arise in their region than are forecasters from more distant locations. In addition, such forecasters might also be more likely to understand the potentially complex interactions between these systems and regional physiographic features than distant forecasters. In this paper, we seek to address this question through the examination of 1 yr of data obtained from the National Collegiate Weather Forecasting Contest

Corresponding author address: Dr. Paul J. Roebber, Department of Geosciences, University of Wisconsin—Milwaukee, Lapham Hall 352, P.O. Box 413, Milwaukee, WI 53201.
E-mail: roebber@csd.uwm.edu

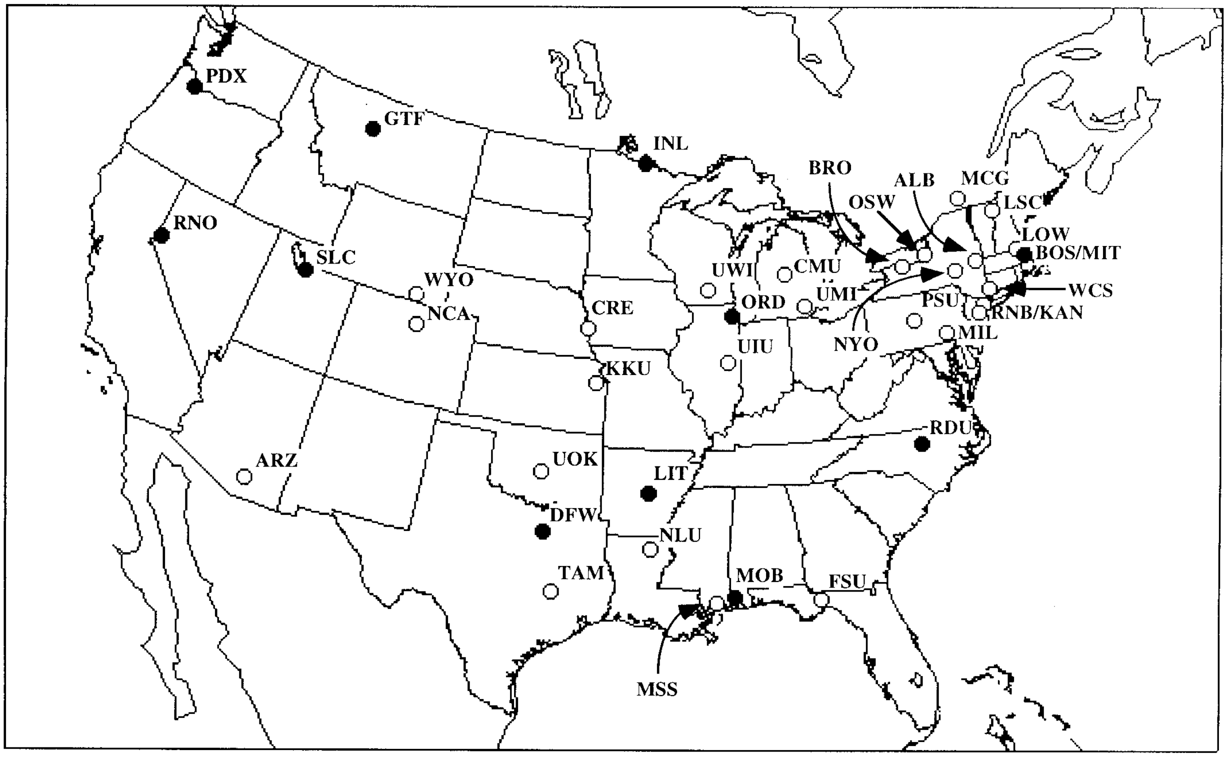

FIG. 1. Geographic distribution of the institutional teams (open circles) and forecast sites (filled circles) of the 1992–93 NCWFC used in the analysis. The identifiers for the teams are listed in Table 1, while the sites are listed in Table 2.

(NCWFC). These data are well suited to this task because the contest (described in section 2) involves broad multiuniversity participation across a range of forecast sites (Fig. 1) and thus provides a reasonable sample for examining the relationship between physical proximity and forecast skill.

However, because of the nature of the NCWFC contest rules, we are prevented from examining forecasts collocated with the verification site and thus are restricted a priori to an examination of regional rather than local variations in forecast skill. Given that we seek to address the regional question, we have selected 1000 km (the meso-α scale) to be representative of the kinds of meteorological systems that might play the most prominent role in the determination of (fall–spring) forecast skill. Thus, any forecaster whose school lies within 1000 km of the forecast site would fall within the regional classification. Further experi-

TABLE 1. Guide to institution codes for 1992–93 NCWFC participants.

Code	Participant	Code	Participant
ALB	State University of New York at Albany	NCA	National Center for Atmospheric Research (CO)
ARZ	The University of Arizona		
BRO	State University of New York at Brockport	NLU	Northeast Louisiana University
CRE	Creighton University (NE)	NYO	State University of New York at Oneonta
CMU	Central Michigan University	OSW	State University of New York at Oswego
FSU	The Florida State University	PSU	The Pennsylvania State University
KAN	Kean College (NJ)	RNB	Rutgers University (NJ)
KKU	University of Kansas	TAM	Texas A&M University
LOW	University of Massachusetts at Lowell	UIU	University of Illinois at Champaign–Urbana
LSC	Lyndon State College (VT)	UMI	University of Michigan
MCG	McGill University (Montreal, PQ, Canada)	UOK	University of Oklahoma
MIL	Millersville University (PA)	UWI	University of Wisconsin—Madison
MIT	Massachusetts Institute of Technology	WCS	Western Connecticut State University
MSS	Mississippi State University	WYO	University of Wyoming

TABLE 2. Guide to forecast site codes for 1992–93 NCWFC. Two forecast sites in the 1992–93 contest (DCA—Washington National and BTV—Burlington International) were eliminated from the analysis due to problems with the data archive.

Code	City, State
BOS	Boston, MA
DFW	Dallas–Forth Worth, TX
GTF	Great Falls, MT
INL	International Falls, MN
LIT	Little Rock, AK
MOB	Mobile, AL
ORD	Chicago, IL
PDX	Portland, OR
RDU	Raleigh–Durham, NC
RNO	Reno, NV
SLC	Salt Lake City, UT

mentation with the distance criterion assured us that our results were not very sensitive to this selection. The analysis approach and results are provided in section 3, while a discussion of those results and the limitations imposed upon them by the data is presented in section 4.

2. The National Collegiate Weather Forecasting Contest

The forecast data used for this study are based upon forecasts submitted as part of the 1992–93 NCWFC. This annual contest emanated from a competition between The Florida State University (FSU) and the University of California, Los Angeles (UCLA), held in the fall of 1965 (Peyrefitte and Mogil 1968). In 1992–93, 27 teams and 575 individuals participated in the NCWFC with team size varying from 5 to 86 members (median team size was 14). The teams included universities from most regions of the United States, McGill University in Canada, and the National Center for Atmospheric Research. The contest has been coordinated since 1987–88 by The Pennsylvania State University.[1]

The NCWFC requires that teams forecast maximum–minimum temperature and a precipitation amount category for a 24-h period beginning a minimum of 6 h after forecast submission. The contest also features individual participant competition in four divisions, based upon their level of professional attainment: 1) faculty/staff, 2) graduate students, 3) juniors and seniors, and 4) freshmen and sophomores.

The forecast site moves every 2 weeks of the contest. These sites are selected by the NCWFC chairpersons prior to the start of the annual competition. Sites within 100 km of longstanding NCWFC participant schools are excluded. Similarly, new or late entrant schools are prohibited from forecasting for sites within 100 km of their location, so that private local data and personal observations do not give them an advantage. Figure 1 shows the geographic distribution of forecast teams (Table 1) and forecast sites (Table 2) used in this analysis.

Early in the history of the NCWFC, the yearly schedule of forecasting periods evolved to a ''least common denominator'' of weeks when most schools would be in session, since each school had to take persistence (their forecast for tomorrow becomes the values of the forecast parameters that verify today) for missed forecasts. By the late 1970s, the contest was further altered, and schools were allowed to select the periods (a minimum of 8 out of the 13) in which they would participate. This change also brought about potential inequities in end-of-year standings because of the relative differences in the difficulty of the forecasts in different periods, so normalization schemes were implemented to address this problem (see below).

The 1992–93 contest was held under the following rules. Forecasts of maximum and minimum temperature and precipitation category (Table 3) were prepared on 4 days per week (Monday–Thursday) during each 2-week period assigned to a particular forecast site. Forecasts, which were prepared prior to 0000 UTC, were valid for the 24-h period beginning at 0600 UTC of the following day. On a daily basis, forecasts were assigned error points using a scheme that dates back to the early days of the competition. One error point was assigned per degree (Fahrenheit) of error in maximum and minimum temperature forecasts, and four error points were assigned per error in the precipitation category. The rules allow trace measurements of precipitation to verify both the category 0 and category 1 forecasts (with category 1 verifying for forecasts exceeding category 1). To qualify for a period, individuals must submit numerical forecasts for at least 6 of the 8 days of a period. To qualify in the annual competition, a forecaster must submit numerical forecasts for at least 75% of the days in the 2-week periods in which they participate and participate in at least the minimum number of these periods (8).

The above emphasis on ''numerical'' forecasts is to distinguish such forecasts from other contest options.

TABLE 3. NCWFC precipitation categories.

Category	Amount
0	0.00 in. or trace
1	Trace–0.05 in. (1.27 mm), inclusive
2	0.06–0.24 in. (6.10 mm), inclusive
3	0.25–0.49 in. (12.45 mm), inclusive
4	0.50–0.99 in. (25.15 mm), inclusive
5	1.00 in. (25.40 mm) or more

[1] Additional information about the NCWFC is available by writing to the NCWFC, Department of Meteorology, 503 Walker Building, University Park, PA 16802.

Because it is recognized that most participants have a scheduled absence at some time during the contest, participants are occasionally allowed to issue a forecast of ''persistence,'' ''guidance,'' ''adjusted persistence,'' or ''adjusted guidance.'' A forecaster opting for guidance is assigned a forecast based upon the available Nested Grid Model (NGM) MOS temperature forecasts and NGM MOS precipitation amount for the site. A forecaster may indicate a forecast of ''guidance + 2'' for maximum temperature, for example, or ''persistence − 3,'' etc. No more than two such forecasts may be issued per period in order for a participant to count toward a team score (with at least five qualifying individuals required for a valid team score). These types of nonnumerical forecasts have been excluded from the dataset used in preparing this paper.

Scoring of the team and individual forecasts in the annual competition must invoke a normalization procedure to account for the fact that teams and individuals participate in different periods that can vary in difficulty. The normalization scheme uses the consensus forecast as a measure of the absolute standard of excellence and uses the variance in the errors of the individual forecasters as a measure of the difficulty of the forecasts in a period:

$$\text{normalized score} = 80 + 10\,\frac{\text{individual} - \text{consensus}}{\text{standard deviation}}, \quad (1)$$

where ''consensus'' is the error point score of the daily consensus forecast (i.e., a forecast that represents the average of all the participants in the contest), ''individual'' is the error point score of an individual, and ''standard deviation'' is the standard deviation of the participants' error point scores. In this scheme, a forecaster receives ± 10 points for every standard deviation above or below the consensus error (the consensus forecaster receives a normalized score of 80). For example, if the standard deviation is 14 at a particular site, an individual forecaster with 14 fewer error points than the consensus forecaster receives a normalized score of 70.

In effect, some of the philosophy of a skill score is adopted in the normalization scheme, with the numerator of the normalized score fundamentally equivalent to the numerator of a skill score invoking consensus as the standard. The differences in scoring philosophy emerge in selection of the denominator term. If climatological values were used as a standard, an approach typical of skill scores, the denominator would represent how closely the 2-week period conformed to normal conditions once verifications were obtained, but would not necessarily represent the perceived degree of difficulty of the daily forecasts at the time they were being prepared. In that context, the normalization scheme adopted in the NCWFC uses the standard deviations of the participant forecasts as the denominator term, because variance of the errors in the prepared forecasts certainly is one measure of the difficulty of the forecast. The NCWFC normalization scheme also uses the consensus score as a standard, because it is commonly observed that a consensus forecast is very hard to beat in a contest. Thus, the consensus score itself tends to supply a stable benchmark upon which to intercompare forecaster scores from various periods. In this paper, a modified normalization approach is used to measure skill (section 3).

3. Forecast distance and the degradation of skill

We wish to examine the effect of distance from the forecast site on the skill of both temperature and precipitation forecasts. Since the results of Roebber and Bosart (1996a) indicated a clear dependence of forecast skill on experience, we will seek to examine these populations separately. As discussed in section 2, the NCWFC identifies forecasters according to four levels of professional attainment; in this study, we shall form two groups: (a) all faculty, staff, and graduate student forecasters (FS) and (b) all undergraduate forecasters (US). Although these classes do not by themselves define forecast experience (an individual faculty participant may have a limited forecast background, while an undergraduate may have relatively extensive experience), we expect that over a sufficiently large sample these classes should provide a reasonable measure.

Since we are interested in measuring the effect of distance on the skill of forecasting temperature and precipitation separately, we shall use a modified form of the scoring system defined by (1). Specifically, we will separate temperature and precipitation errors and calculate normalized scores for each category independently. Since we do not know a priori whether all the teams are equally balanced in terms of experience and skill composition (indeed, the point of the contest is to try to determine these differences in declaring a winner) and because the teams are highly variable in size, we will not use the original teams in the analysis. Instead, we will form teams on the basis of experience level (FS vs US) and distance from the forecast site (within or outside of some critical range).

As mentioned previously, we have used a critical range of 1000 km; thus, any forecaster whose school lies within 1000 km of the forecast site would fall into one of two bins, depending on professional status. In total, four teams are formed for each 2-week forecast site—namely ''FS—within 1000 km,'' ''FS—outside 1000 km,'' ''US—within 1000 km,'' and ''US—outside 1000 km''—and the normalized scores are computed for each of these teams for each of the 11 sites identified in Fig. 1 and Table 2. This procedure results in median team sizes of 50–150 forecasters for US and 30–75 forecasters for FS, ensuring a robust consensus.

In order to test the relationship between distance and forecast skill on a rigorous statistical basis, we have

TABLE 4. ANOVA-relationship between forecast temperature skill (normalized score) and experience. *P* value defines the probability that the differences in the categories result from chance variation. A *p* value of less than 0.05 (.01) implies statistical significance at the 95% (99%) levels.

Experience category	Average normalized score	*p* value
FS	79.23	<.0001
US	80.44	
ANOVA-relationship between forecast precipitation skill (normalized score) and experience.		
FS	79.46	.0004
US	80.39	

TABLE 5. ANOVA-relationship between forecast temperature skill (normalized score) and distance for faculty, staff, and graduate students (FS). *P* value defines the probability that the differences in the categories result from chance variation. A *p* value of less than 0.05 (.01) implies statistical significance at the 95% (99%) levels.

Distance category	Average normalized score	*p* value
FS—within 1000 km	78.95	.0206
FS—outside 1000 km	79.47	
ANOVA-relationship between forecast precipitation skill (normalized score) and distance for faculty, staff, and graduate students (FS).		
FS—within 1000 km	79.43	.8650
FS—outside 1000 km	79.49	

applied the one-way analysis of variance (ANOVA) technique using the normalized scores defined by (1) as the dependent variable and the regional distance categorization as the independent variable. Our data do not represent the product of a rigorously controlled experiment, because the NCWFC is a contest, not an experiment. For the same reasons, we are prevented from applying the principles of experimental design. However, these same limitations also assure us of randomization: the forecast sites have not been selected with this application in mind and are effectively random (except for the constraint that they are confined to the continental United States). Furthermore, our estimates of chance variation (experimental error) require replication, a constraint that is satisfied by repeating our team combinations over the 11 available sites. Thus, our dataset has been collected under circumstances that assure us that we will be able to reliably test our hypothesis concerning the skill–distance relationship under fairly general conditions.

As a first step, in an effort to confirm our assumptions concerning team structure, we have sought to test the results of Roebber and Bosart (1996a) concerning the role of forecast experience on skill. The results of this test, shown in Table 4, confirm the findings of Roebber and Bosart (1996a) that experience is a significant factor (well above the 99% level) in determining forecast skill. Thus, in order to test our hypothesis concerning the skill–distance relationship, we are fully justified in analyzing our populations separately. However, it is important to note that in absolute terms, the differences in error between these two groups are rather small. For example, for median temperature and precipitation category error standard deviations, the skill differences between US and FS amounts to about 4°F (2.2°C) extra error per 2 weeks (eight forecasts) in temperature and two categories per 2 weeks in precipitation.

The results of the skill–distance investigation are shown in Tables 5 and 6. The analysis indicates that there are significant (above the 95% level) differences in forecast temperature skill as a function of regional distance for experienced forecasters (Table 5). In particular, forecasters desiring to make predictions of maximum and minimum temperature for locations outside of their own region (defined as the meso-α scale) are at a distinct disadvantage with respect to forecasters within the target region (resulting in an increase in error about half again as much as that between FS and US groups). In contrast, the less experienced forecasters (US, Table 6) possess virtually identical skill in forecasting temperatures, regardless of forecast distance. Roebber and Bosart (1996a) concluded that experienced forecasters were better able to account for biases in model output specific to certain synoptic situations and adjust their temperature forecasts accordingly. Our results reinforce the findings of Roebber and Bosart (1996a) and further indicate that the advantages extend beyond the local scale to the meso-α scale.

However, these results are not repeated for forecasts of precipitation category. Little difference in skill is apparent for forecasters within or outside of the 1000-km limit, regardless of experience level. Roebber and Bosart (1996a) found that the experience advantage was sharply reduced or eliminated in the context of precipitation probability forecasts, presumably because the overall skill level of precipitation forecasts is low.

TABLE 6. ANOVA-relationship between forecast temperature skill (normalized score) and distance for undergraduate students (US). *P* value defines the probability that the differences in the categories result from chance variation. A *p* value of less than 0.05 (.01) implies statistical significance at the 95% (99%) levels.

Distance category	Average normalized score	*p* value
US—within 1000 km	80.42	.9096
US—outside 1000 km	80.45	
ANOVA-relationship between forecast precipitation skill (normalized score) and distance for undergraduate students (US).		
US—within 1000 km	80.40	.9326
US—outside 1000 km	80.38	

TABLE 7. Precipitation events and subjective classification for 1992–93 NCWFC sites. Trace measurement indicated by T. WAA denotes warm air advection.

Station	Date (forecast submitted)	Precipitation category	Description of event	Event classification
BOS	21 Sep.	2	500-hPa trough/front	Synoptic
BOS	22 Sep.	2	500-hPa trough/front	Synoptic
GTF	5 Oct.	2	500-hPa trough	Synoptic
GTF	7 Oct.	T	500-hPa trough/upslope	Regional
GTF	8 Oct.	T	500-hPa trough/upslope	Regional
PDX	27 Oct.	2	500-hPa trough/front	Synoptic
PDX	28 Oct.	4	500-hPa trough/surface trough	Synoptic
PDX	29 Oct.	2	500-hPa trough/surface trough	Synoptic
PDX	2 Nov.	T	WAA, surface cold air drainage	Regional
PDX	3 Nov.	3	WAA, surface cold air drainage	Regional
PDX	5 Nov.	T	Weak warm front	Synoptic
ORD	9 Nov.	2	500-hPa trough	Synoptic
ORD	10 Nov.	1	500-hPa trough	Synoptic
ORD	11 Nov.	4	500-hPa trough	Synoptic
ORD	12 Nov.	1	500-hPa trough	Synoptic
ORD	18 Nov.	2	500-hPa trough/WAA	Synoptic
ORD	19 Nov.	2	500-hPa trough/WAA	Synoptic
SLC	2 Dec.	1	Weak 500-hPa trough	Synoptic
SLC	7 Dec.	T	Weak 500-hPa trough	Synoptic
SLC	8 Dec.	2	Cold front	Synoptic
SLC	10 Dec.	3	Major 500-hPa trough	Synoptic
LIT	9 Feb.	4	Cutoff 500-hPa cyclone	Synoptic
LIT	10 Feb.	3	Cutoff 500-hpa cyclone	Synoptic
LIT	15 Feb.	1	500-hPa trough	Synoptic
RNO	22 Feb.	2	Lee trough/upslope	Regional
RNO	23 Feb.	2	Lee trough/upslope	Regional
RNO	25 Feb.	1	Lee trough/upslope	Regional
RNO	1 Mar.	T	Weak upslope	Regional
RNO	2 Mar.	T	Weak upslope	Regional
RDU	11 Mar.	3	Superstorm 1993	Synoptic
RDU	15 Mar.	1	WAA/cold air damming	Regional
RDU	16 Mar.	2	WAA/cold air damming	Regional
INL	25 Mar.	T	Front	Synoptic
INL	29 Mar.	1	Intensifying cyclone	Synoptic
INL	30 Mar.	1	Intensifying cyclone	Synoptic
DFW	6 Apr.	4	Thunderstorm	Regional
DFW	12 Apr.	1	Thunderstorm	Regional
DFW	13 Apr.	4	Thunderstorm	Regional

For example, within the 1992–93 NCWFC, the median consensus error was three precipitation categories per 2-week forecast period, which, given that 43% of the forecast days included at least a trace of precipitation, amounts to roughly 1 category error per precipitation day.

In concluding that there are no significant differences in precipitation forecast skill with distance for either forecast group, we have not yet considered the types of precipitation events that occurred. In particular, it may be that events in which regional controls manifest themselves (such as upslope flow, convective outbreaks, cold air damming) are more likely to result in skill–distance relationships than those that are driven largely by synoptic-scale factors [e.g., 500-hPa shortwave troughs and the associated differential cyclonic vorticity advection and the Laplacian of thermal advection; see Bluestein (1993)]. We have therefore attempted to stratify the precipitation results further, using a subjective criterion as to whether the observed precipitation was substantially modulated by such regional factors. A list of the dates and locations for which precipitation events occurred, along with our subjective event classification, is presented in Table 7. It should be noted that the subjective event classification was conducted by one of us (LFB) independent of the analysis (performed by PJR), helping to assure that no bias was introduced. The list shows that the number of event periods under consideration is relatively small, since we were able to identify substantial regional controls at only five of the sites. However, these events displayed a range of mechanisms, so we have some confidence that our results generalize reasonably well.

We have computed a precipitation skill score following (1) as a function of precipitation event class and distance from the forecast site for each forecaster category (FS and US) for each 2-week period. The median number of forecasts making up the two FS groups

TABLE 8. ANOVA-relationship between forecast precipitation skill (normalized score) and distance for faculty, staff, and graduate students (FS) for events with substantial regional controls. *P* value defines the probability that the differences in the categories result from chance variation. A *p* value of less than 0.05 (.01) implies statistical significance at the 95% (99%) levels.

Distance category	Average normalized score	*p* value
FS—within 1000 km	78.86	.0400
FS—outside 1000 km	80.10	
ANOVA-relationship between forecast precipitation skill (normalized score) and distance for undergraduate students (US) for events with substantial regional controls (subjectively determined).		
US—within 1000 km	81.06	.5359
US—outside 1000 km	80.76	

(within and outside, respectively) for these events was 96 and 287, while for the US groups it was 127 and 414, so we are again assured that our consensus is robust. The ANOVA results (Table 8) are clear: distance appears to be a significant factor in precipitation forecast skill, provided that the events are modulated by regional controls. Experienced forecasters within the region are able to take advantage of their presumed better knowledge of these conditions compared to other forecasters (those that are farther from the site or are simply lacking the necessary experience). Such skill differences are masked when we examine precipitation forecasts as a whole, because of the leveling effects of the numerical model output and the associated guidance, which capture well the synoptic-scale signal.

4. Discussion

The results from this study suggest that regional knowledge can provide a slight advantage to experienced forecasters concerned with broadscale high- and low-temperature patterns. We suggest that experienced temperature forecasters have obtained implicit knowledge as to the vagaries of maximum and minimum temperatures as a function of weather regime and time of the year in their particular regions (defined as locations within between 100 and 1000 km of the forecast site). As an example, experienced forecasters may use the model-forecast 850-hPa temperature and 1000–500-hPa thickness patterns extensively to forecast maximum temperatures the next day. Experienced forecasters with regional knowledge might be able to make subtle modifications to the model-derived large-scale temperature patterns by factoring in how such things as the prevailing wind direction (upslope or downslope), the presence or absence of snow cover and its distribution in the forecast domain, or current soil moisture characteristics and their likely impact on low-cloud formation can affect the temperature forecast.

In general, forecaster skill at distinguishing between wet and dry days (as measured by POP forecasts) is less than for maximum and minimum temperature forecasts and decays at a more rapid rate with time (Sanders 1986; Bosart 1983; Glahn 1985). However, our results suggest that the prediction from the more inherent mesoscale configuration of rainstorms is also amenable to improvements through forecaster experience. For example, we speculate that regional knowledge can prove useful for the prediction of the timing, amount, and intensity of precipitation in mountainous regions or near coastlines where differential diabatic heating and differential roughness may be important physical mechanisms that contribute to rainfall totals. We further speculate that if a dataset existed that included *local* as well as regional and distant forecasts, analysis of such data would show an even stronger effect than was found in this paper. We base this consideration on the fact that forecasters typically learn how to interpret and modify the output of numerical models in light of their knowledge of local peculiarities of the weather and that this knowledge base undoubtedly becomes degraded as one moves away from the local area. Thus, one might expect that a highly experienced forecaster's skill would trace out a rapidly declining curve as a function of distance from the forecast site (asymptotically approaching some baseline of skill approaching MOS at distances beyond the meso-α scale).

Our findings suggest that forecast managers (public and private) might want to consider whether it makes sense to ensure the retention of a small subset of forecasters on station (or who forecast for specific sites) to serve as an experienced nucleus of forecasters with considerable regional knowledge. For example, in the National Weather Service (NWS) it is common for forecasters to rotate through individual locations every 2–3 yr as they move up the career ladder. While this situation is not NWS policy, it is an operational reality imposed upon forecasters who wish to advance relatively quickly. Whether regional forecasts could be improved somewhat if a small subset of these forecasters would be allowed to remain on station for a longer period (and still be able to move up the career ladder) might need to be explored.

Acknowledgments. Partial support for this research was provided by National Science Foundation Grant ATM-9114598. Ideas for this research began to germinate through the continued involvement of PJR and LFB in the practical forecast problem, including most recently the STORM-FEST field program of 1992. Our thanks go to Dr. Bradley Colman and the reviewers for their constructive comments on an earlier version of this paper. The authors also thank The Pennsylvania State University for continuing to serve as the host institution of the NCWFC and National Chairpersons Fred Gadomski and Paul Knight for their assistance in providing the 1992–93 data used for this study.

REFERENCES

Bluestein, H. B., 1993: *Synoptic-dynamic Meteorology in Midlatitudes.* Vol. II. *Observations and Theory of Weather Systems.* Oxford University Press, 594 pp.

Bosart, L. F., 1983: An update on trends in skill of daily forecasts of temperature and precipitation at the State University of New York at Albany. *Bull. Amer. Meteor. Soc.*, **64,** 346–354.

Glahn, H. R., 1985: Yes, precipitation forecasts have improved. *Bull. Amer. Meteor. Soc.,* **66,** 820–830.

Peyrefitte, A., Jr., and H. Mogil, 1968: A national collegiate forecasting contest. *Weatherwise,* **21,** 162–165.

Roebber, P. J., and L. F. Bosart, 1996a: The contributions of education and experience to forecast skill. *Wea. Forecasting,* **11,** 21–40.

——, and ——, 1996b: The complex relationship between forecast skill and forecast value: A real world analysis. *Wea. Forecasting,* **11,** 544–559.

Sanders, F., 1986: Trends in skill of Boston forecasts made at MIT, 1966–84. *Bull. Amer. Meteor. Soc.,* **67,** 170–176.