Precipitation Forecast Quality of the NMC Nested Grid Model According to Synoptic Situation: Case Study for the Upper Midwest

DAVID D. HOUGHTON AND EVAN A. RUBIN*

Department of Meteorology, University of Wisconsin-Madison, Madison, Wisconsin

(Manuscript received 26 July 1989, in final form 17 April 1990)

ABSTRACT

Twelve-hour quantitative precipitation forecasts from the National Weather Service National Meteorological Center's nested-grid model are evaluated for 12 stations in the upper Midwest for the period April–July 1987. Statistics for threat score, bias, post agreement, over and under forecasts, and "quantitative correctness" are determined from frequency distributions for precipitation forecasts and verifications categorized into four quantitative amount levels: none, low, medium, and high. The analysis is performed for all cases as a group and for subpopulations representing six different categories for the associated synoptic situation. The synoptic situation descriptors involve proximity to surface frontal or trough positions or lack thereof. It was found that the warm and occluded frontal situations had better forecast performance than the other synoptic situations reflecting the better handling of grid-scale in contrast to convective-scale precipitation by the model. Results provide an example of the aid that can be given to forecasters by suggesting relative levels of reliance to be assigned to specific model forecasts.

1. Introduction

The forecast for precipitation, including quantitative amounts, is one of the most important operational weather prediction products. Such forecasts are very difficult to produce because of the dominant mesoscale structure of precipitation and the dependency on the smaller resolvable scales and convective scale parameterization in the numerical models. It has been generally noted that the forecast improvement for smaller scale phenomena such as precipitation has been more gradual and debatable than for the larger scale circulations (Charba and Klein 1980; Glahn 1985; Sanders 1986). A general study by Murphy and Sabin (1986) has shown overall improvement in both subjective and objective precipitation probability forecasts by the U.S. National Weather Service (NWS) from the mid-1960s to the mid-1980s.

The forecaster needs all the information available about overall model performance, including specific indicators of forecast quality, to meet the challenge of precipitation forecasting. Historically, manually drawn quantitative precipitation forecasts, model output statistics products, and direct numerical model output have been used to assist the forecaster. The statistical methods have been used extensively at the National Meteorological Center (NMC) to provide general quantitative guidance from numerical model forecasts and to compensate for model deficiencies (Glahn and Lowry 1972). Beyond such general guidance, information on forecast quality according to the specific weather situation should also be of value to the forecaster.

The nested grid model (NGM) has served a key role in providing regional forecasts for the NWS since 1985. It is a 16-layer primitive equation model with 80-km horizontal resolution over North America. Qualitative precipitation forecasts are produced according to a prognostic grid-scale moisture equation and Kuo-type parameterization for sub-grid scale convective processes. See Phillips (1979), NWS (1986), and Hoke et al. (1989) for a complete description of the model.

The NGM along with the older limited-area finemesh model (LFM) (NWS 1978) provide the primary numerical model precipitation forecasts for the NWS. The NGM performance is considered to be generally better than the LFM for precipitation forecasts (Jensenius 1988, 1989). However, this is not uniformly true for all quantitative levels as shown by reports in the quarterly publication, NMC Seasonal Performance Summary, published by the NMC Development Division.

The purpose of this case study is to make an exploratory examination of how numerical model forecast performance for precipitation varies according to

---

* Present affiliation: WWLP Channel 22, P.O. Box 2210, Springfield, MA 01102.

Corresponding author address: David D. Houghton, Department of Meteorology, University of Wisconsin, 1225 West Dayton Street, Madison, WI 53706.

the synoptic situation. Forecasts from the NMC's NGM are examined.

The case study is not intended to provide a comprehensive and statistically robust result. Many factors of the synoptic situation, such as the upper-level jet flow, are not considered. The data sample is limited to four months and to the North Central region of the United States. Results are most relevant to regions in the United States with propagating and developing extratropical systems where topographic effects are not a major factor. The study is intended to exemplify a type of analysis that could facilitate the use of numerical model products and to provide a benchmark for NGM performance.

Section 2 describes the data analyzed in this case study. Section 3 presents the method of analysis. Results and conclusions follow in Sections 4 and 5, respectively.

2. Data

Model and verification data were analyzed for twelve first-order stations in the upper Midwest shown in Fig. 1 for the period 1 April 1987–31 July 1987. This four-month period includes both spring and summer months and thus is representative of large-scale and convective precipitation events. It follows the time when several significant modifications were made in the physics of the NGM (National Weather Service 1986) including the adoption of a diurnal cycle and a reduction in evaporation effects for convective precipitation. However, since this case study period, other changes have been made. For instance, the nonlinear normal mode initialization for the NGM has been changed from an 8-mode to a 2-mode procedure which did impact the precipitation forecasts especially for heavy precipitation (Carr et al. 1989).

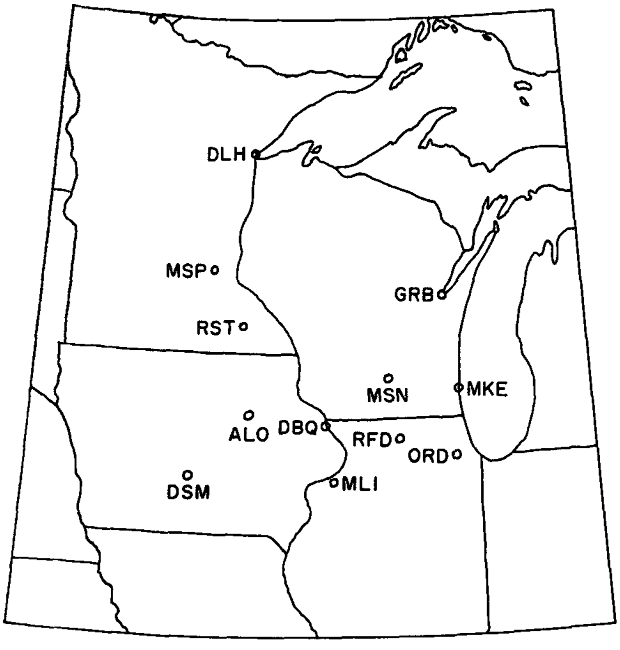

FIG. 1. The twelve upper Midwest stations used to obtain the data for the study.

Hoke et al. (1989) provide the most recent overall description of the NGM including changes since the case-study period. Junker et al. (1989) highlight the effects of the NGM model changes after July 1987 on precipitation forecast scores including tests for this case study period. They found that the model changes tended to reduce bias and threat scores for 0.50-in. threshold values. Jensenius (1989) summarizes the precipitation prediction performance of the post-July 1987 version of the model.

Despite the fact that operational models are altered as seen with the NGM here, it is felt that this case study suggests relationships relevant to more than a specific model formulation especially for relative differences in forecast performance. In addition, the results serve as a benchmark for quantitative precipitation forecasts by the NGM.

The case-study period was characterized by conditions generally warmer and dryer than normal. Mean temperature for the 12 stations was about 2°C above normal for each of the four months. This was related to a tendency for excess ridging in the upper-air flow over the western United States together with northward displacement of the upper tropospheric jet stream. Rainfall was below normal averaging about 80% of normal for the twelve stations for the four-month period. April and June were particularly dry with rainfall amounts averaging barely over 50% of normal. July was the wettest month with rainfall being more than 30% above normal. Despite the dry weather, there were a reasonable number of synoptic systems that moved through the study area during the four month period.

Twelve-hour quantitative precipitation amounts were analyzed for the times 1200–0000 UTC (daylight in the upper Midwest) and 0000–1200 UTC (night in the upper Midwest). The observation amounts were obtained from hourly quantitative rainfall records (local climate data publications). Forecast values were obtained from NGM run cycles initialized only at 1200 UTC and for the three forecast periods, 0–12-, 12–24-, and 24–36 h. Model grid-point values were interpolated to the station positions using procedures established at the NWS Techniques Development Laboratory (Dallavalle, pers. comm.).

The dataset had a relatively satisfactory total sample size; however, it was not sufficient for some of the subcategories. The total sample size consisted of three forecasts for each of the 122 days in the period and each of the 12 stations, yielding a total of 4392 forecasts. These data were not independent so the effective sample size was much smaller.

3. Analysis

a. Basic compilation

Frequency of occurrence statistics were determined for the 4392 forecasts and verifications according to four categories for precipitation amount in the 12-h period:

1) None
2) Low: Trace up to 0.1 in. (2.54 mm)
3) Medium: 0.1 in. (2.54 mm) up to 0.5 in. (12.7 mm)
4) High: 0.5 in. (12.7 mm) or more.

The low group was defined including trace precipitation amounts rather than using 0.01 in. as the lower limit. It was reasoned that to an operational forecaster sprinkles or snow flurries that can occur for several hours without amounting to 0.01 in. are in the view of the general public indeed precipitation.

The cutoff between low and medium precipitation is consistent with previous studies. Balling (1985) looked at nocturnal precipitation in the Great Plains and considered the "larger events" to be those where at least 2.54 mm (one-tenth of an inch) of water fell. Englehart and Douglass (1985) analyzed precipitation in the United States and also used 2.54 mm as a criterion to define a day where precipitation frequency could be evaluated.

Similarly the cutoff between medium and high was also made consistent with previous work. Bohm (1979) used "at least 30 mm (1.18 in.) of daily precipitation" to characterize heavy rainfall events in Vienna. The study by Klaus et al. (1983) for Mexico City contained various classes for rainfall, of which the heaviest was defined as having "30 mm per twenty-four hour period" (equivalent to 15 mm (0.59 in.) for the 12-h period and close to the half-inch criterion used here).

Although there is some advantage in having similar populations for the three quantitative amount categories, it was considered more useful to have cutoff criteria consistent with commonly accepted terminology. Of course, the smallest interval range of the three precipitation categories, "Low," had the greatest frequency of occurrence, representing nearly two-thirds of all precipitation cases. The "High" category had a much smaller frequency of occurrence than the others accounting for only 5% of the cases when precipitation was observed (Fig. 2).

The basic data assembly thus consists of a 16-category frequency distribution contingency table as shown in Fig. 2. The 16 categories are identified by letters for later reference. Actual frequency values of the total dataset for the observations and forecasts in the 0–12-h forecast period are also shown for reference. Frequency values are qualitatively similar for the other forecast periods.

	OBSERVED			
FORECAST	NONE (69.3%)	LOW (20.4%)	MEDIUM (7.9%)	HIGH (2.5%)
NONE (62.0%)	A 778	B 111	C 15	D 3
LOW (19.3%)	E 136	F 103	G 40	H 3
MEDIUM (17.3%)	I 90	J 81	K 56	L 26
HIGH (1.5%)	M 10	N 4	O 4	P 4

FIG. 2. The 16-category frequency distribution contingency table used for data compilation. Each category is identified by a letter in the upper left hand corner for use in the discussion. The percent of occurrence of each of the four quantitative amount levels is shown in parenthesis for both the observed and forecast amounts for the 0–12-h forecast period. The actual frequency count for each of the 16 categories for the 0–12-h period is shown in the center of each category box. The frequency distributions are similar for the other two 12-h periods.

b. Classification of associated synoptic conditions

For each forecast, the associated synoptic conditions were determined by inspecting the NMC surface analyses for 0000 UTC and 1200 UTC and noting the relationship of analyzed surface weather fronts and trough positions to each station. Six categories were defined for the associated conditions according to an estimate of the minimum distance of the fronts and troughs from the station during the 12-h period:

1) Cold Front (COLD): Station within 100 km of front on either side
2) Warm Front (WARM): Station within 500 km on the leading (cold) side
3) Stationary Front (STNY): Station within 500 km on the cold side or 100 km on the warm side and the front moves less than 200 km in 12 h
4) Trough (TROF): Station within 100 km of trough line on either side
5) Occluded Front (OCCL): Station within 500 km on the leading side or 100 km on the back side.
6) No front or trough (NONE): Station not within any of the distances to frontal or trough weather systems as noted above.

The distances used above were intended to correspond to realistic dimensions for the precipitation area associated with the weather systems. Spatial scales for frontal influences on precipitation were specified according to "ideal" front models. Values chosen corresponded to those frequently mentioned in the literature and text books.

For the cold frontal cases, it is recognized that precipitation commonly falls in a 75–100-km zone on either side of the surface front position (Miller 1971). This figure does not take into account squall lines that may occur in advance of the surface front.

Two figures, 300- and 500 km, were seen most often in the literature for the maximum horizontal extent of associated precipitation in advance of a warm front. There is likely to be considerable variability in the distance of precipitation from the warm front due to advective effects in the baroclinic zone. Several authors (Navarra 1979; Miller 1971; Ahrens 1988) agreed that cirrus clouds can be found 1000 km from the front although they disagreed as to where precipitation was most likely to be found. In order to minimize the specification of actual warm frontal precipitation as in the NONE situation 500 km was used in this study.

In the case of the stationary front, no definition for extent of precipitation could be found in the literature. It was acknowledged, however, that on some occasions overrunning may occur and precipitation of the warm frontal type could be expected on the cool side of the stationary front. On the warm side, the definition for the cold front model was used (100 km) to account for the cold frontal characteristics manifested by frontal undulations where colder air would locally replace warmer air.

The trough condition was considered analogous to the cold front in terms of horizontal scaling. It is assumed that the influence distance for precipitation is the same (100 km).

The definition given by Navarra (1979) was used for occluded front cases. He points out that the weather associated with an occluded front is similar to that of a warm front and cold front combined. Tables 1 and 2 summarize the climatology for actual occurrence of precipitation associated with the weather systems considered in this study. Note that the percent of occurrence is considerably larger (roughly double) for the weather system domains than for regions outside those domains (the NONE category). As shown in Table 2, the occurrence was not close to 100% for any individual type of weather system. The larger percent rates were for the WARM, STNY, and OCCL systems.

TABLE 1. Observed percent occurrence of precipitation sample climatology for the weather system influence domains defined for this study. Values are the average for the 0–12- and 12–24-h forecast periods which covers essentially the entire time for the case-study period.

Data category	Percent occurrence of precipitation
Total population	30.5
Groups associated with weather systems	39.7
Group not associated with weather systems (NONE)	17.9

TABLE 2. Observed percent occurrence of precipitation sample climatology for the individual weather system influence domains. Values obtained as described in Table 1.

Individual weather system category	Percent occurrence of precipitation
COLD	33.4
WARM	49.4
STNY	39.3
TROF	34.3
OCCL	62.0

The relative and actual frequencies of occurrence of the various synoptic categories are shown in Fig. 3 for the 0–12-h forecasts (and 24–36-h forecasts) i.e. corresponding to the daylight period. The figures for the night period, relevant to the 12–24-h forecasts, are very similar. The NONE group is the largest and consists of approximately 40% of all the cases. It is expected that its large data sample size will provide better score reliability than for the smaller sets such as the TROF (8%) and OCCL (2.4%). However, the OCCL category was so small (35 cases in all) that the results could be considered nearly meaningless.

c. *Evaluation statistics*

A number of statistical parameters were determined. Several of these are comparable to the standard scores used by NMC such as the Threat Score and Bias and others are introduced to define further forecast characteristics. All of the parameters are defined using frequency of occurrence as the counterpart to "area inside an isopleth" which is used when evaluating a precipitation forecast over an area at a given time. These are listed and briefly described in Table 3.

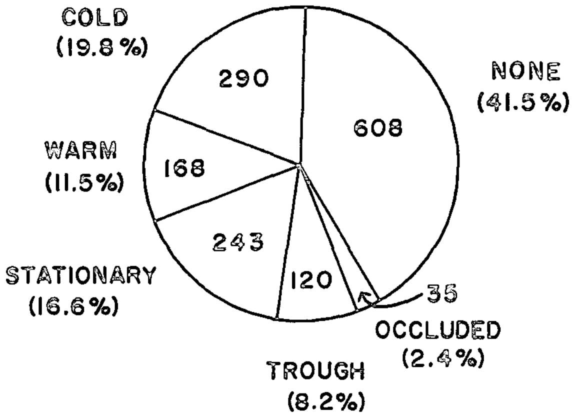

FIG. 3. The percent of occurrence and actual population count for each of the six synoptic categories used in this study. Figures are for the daylight period.

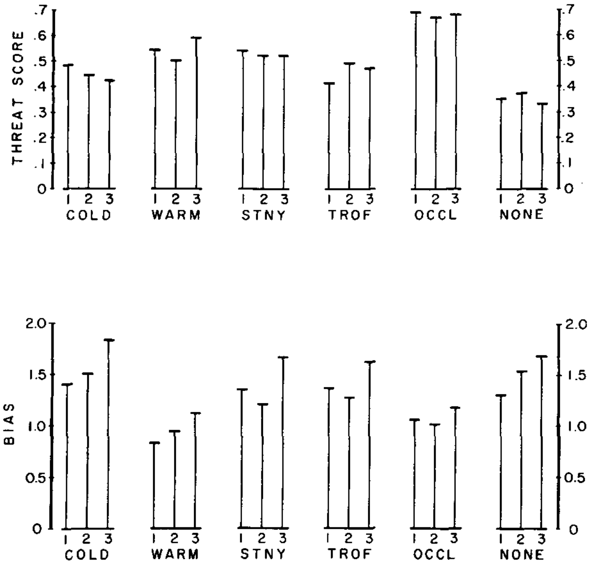

FIG. 7. Threat scores (THR; top) and bias scores (BIAS; bottom) for the three forecast intervals, labeled as before, for each of the six synoptic situation categories (dimensionless ratios).

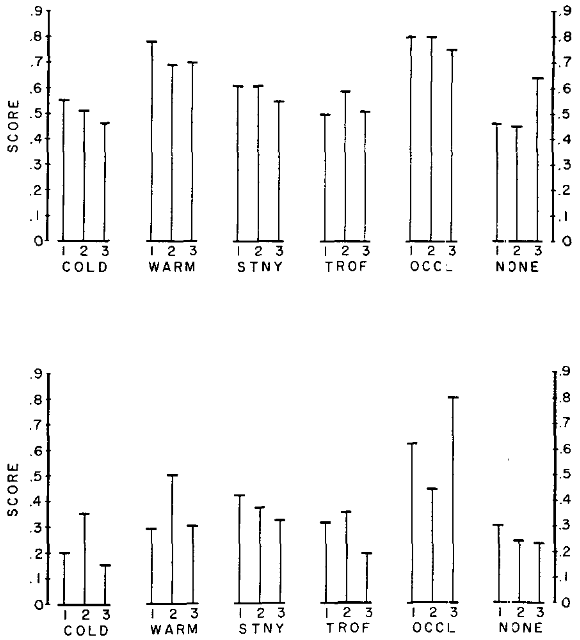

FIG. 8. Trace (top) and medium level (bottom) threshold post agreement scores (POAG-T and POAG-M, respectively) for the three forecast intervals, labeled as earlier, for each synoptic category (percent).

Clearly a model prediction of no precipitation should be taken seriously. As suggested in Fig. 6, scores for low, medium, and high amount categories are so low that it was not considered useful to look at details of the score variation as a function of the synoptic situation.

c. Seasonal variations

The case study spans two seasons, spring (April/May) and summer (June/July) and thus mixes seasonal differences in model performance that have long been recognized. Charba and Klein (1980) and more recently Junker et al. (1989) document that the winter season tends to be the best and the summer season the worst. In this study only part of the seasonal cycle is included; nevertheless, model precipitation forecasts measured by the statistics used in this study did show better performance in the spring compared to the summer period. This reflects the increased importance of convective precipitation activity in the summer which is more difficult both to model and to observe. This seasonal variation did not change the qualitative relationship among the synoptic situations for model performance, the key focus of the study.

5. Conclusions

Forecast performance for precipitation was found to be best for warm and occluded frontal situations in terms of threat, bias, and post agreement scores. Other synoptic situations such as cold frontal, trough, or the absence of any synoptic-scale, weather producing systems had consistently lower performance. In the latter cases it is expected that precipitation structures would be generally smaller scale with more emphasis on convective systems which have to be represented by parameterization techniques in the model and which

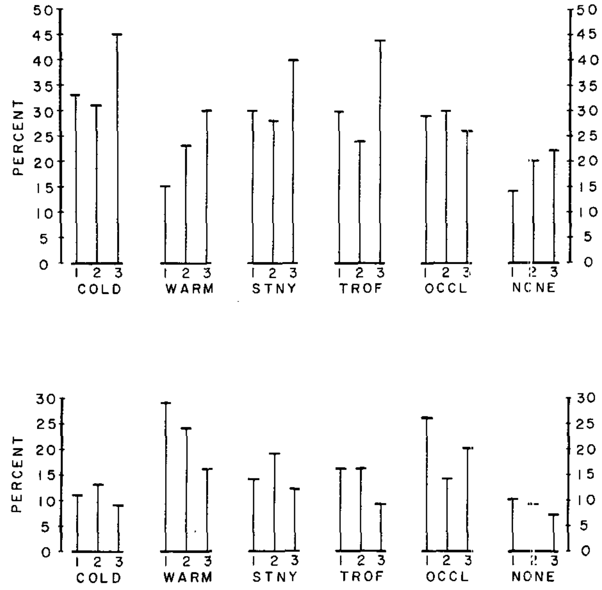

FIG. 9. Overforecast (top) and underforecast (bottom) scores (OVER and UNDER, respectively) for the three forecast intervals, noted as before, for each of the synoptic situation categories (percent).

would be less reliably depicted by the rain gauge observing network.

For most cases, forecast guidance for precipitation occurrence as measured by post agreement scores with trace threshold was barely above the 50% level. The important exception was for the warm frontal situation where correct forecasts occurred about 75% of the time. Success rates for the occurrence of more significant rainfall amounts (medium and high) were generally much less than 50%. It is worth noting that post agreement trace threshold scores for all six synoptic categories are well above the observed climatological percent occurrence of precipitation for the six categories respectively.

Forecasts for no precipitation were found to be correct for more of the time than for precipitation in any of the amount categories. The success rate was in the middle 80% range for all synoptic categories combined, which is higher than the 69.5% occurrence of no precipitation or the percent correct obtained using probability based on the observed frequency of no precipitation. The percent correct for no-precipitation forecasts was lowest for the warm and occluded frontal situations.

Results suggest that there is a differential in model performance based on the synoptic situation which could be explored further by forecasters seeking guidance directly from model output. Results also confirm that the success rates for numerical model precipitation forecasting, while small, remain better than the climatological percent occurrence rates for precipitation (and thus the success rate that would be obtained by chance forecasting).

As noted before, limitations in sample size precluded making meaningful statistical significance tests for the differentials in model performance. Results here are intended to provide ideas and suggest specific contingency categories for which a comprehensive statistical analysis could be useful.

TABLE 4. Categories summed for each statistical parameter ratio factor.

Statistical score acronym	Numerator	Denominator
THR	F, G, H, J, K, L, N, O, P	B–P (all categories except A)
BIAS	E, F, G, H, I, J, K, L, M, N, O, P	B, C, D, F, G, H, J, K, L, N, O, P
POAG-T	F, G, H, J, K, L, N, O, P	E, F, G, H, I, J, K, L, M, N, O, P
POAG-M	K, L, O, P	I, J, K, L, M, N, O, P
OVER	B, C, D, G, H, L	A–P (all categories)
UNDER	E, I, J, M, N, O	A–P (all categories)
QCORR		
None (zero) level	A	A, B, C, D
Low level	F	E, F, G, H
Medium level	K	I, J, K, L
High level	P	M, N, O, P

Acknowledgments. Many thanks are due Linda Keller who did the drafting work and gave editorial advice for the paper. John Jensenius, National Weather Service Techniques Development Laboratory, very kindly supplied the quantitative precipitation forecast data. Pam Naber, State Climatologist Office, helped to find the climatological data for the case study period. Enormous help was given by three comprehensive and constructive anonymous reviews. Partial support was given by the National Science Foundation through grant ATM-8514730.

APPENDIX

Specific Definitions for the Statistical Quantities Used in this Study

The statistical parameters were defined by ratios involving the summation of populations from the 16-category contingency table. Refer to Fig. 2 for the description of the contingency table and the definition of the specific categories denoted by capital letters. The sums used in the numerator and denominator for each statistical score ratio are defined in Table 4.

REFERENCES

Ahrens, D. C., 1988: *Meterology Today.* West Publishing Co., St. Paul, Minnesota, 514 pp.

Balling, R. C., 1985: Warm season nocturnal precipitation in the Great Plains of the United States. *J. Clim. and Appl. Meteor.*, **24**, 1383–1387.

Bohm, R., 1979: Heavy precipitation in the regions of Vienna, Austria. *Zentralanstalt fur Meteor. und Geodynamik,* **31**, 207–230.

Carr, F. H., R. L. Wobus and R. A. Petersen, 1989: Synoptic evaluation of normal mode initialization experiments with the NMC Nested Grid Model. *Mon. Wea. Rev.*, **117**, 2753–2771.

Charba, J. P., and W. H. Klein, 1980: Skill in precipitation forecasting in the National Weather Service. *Bull. Amer. Meteor. Soc.*, **96**, 1546–1555.

Englehart, P. J., and A. V. Douglass, 1985: Statistical analysis of precipitation frequency in the conterminous United States, including comparisons with precipitation totals. *J. Clim. and Appl. Meteor.*, **24**, 350–362.

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

——, and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.

Hoke, J. E., N. A. Phillips, G. J. DiMego, J. J. Tuccillo and J. G. Sela, 1989: The regional analysis and forecast system of the National Meteorological Center. *Weather and Forecasting*, **4**, 323–334.

Jensenius, J. S., Jr., 1988: Statistical characteristics of the National Meteorological Center's regional and global weather prediction models. *Preprints Eighth Conference on Numerical Weather*

Prediction, Baltimore, American Meteorological Society, pp. 550–557.

——, 1989: A statistical comparison of the forecasts produced by the NGM and LFM for the cool season of 1987–88. TDL Office Note 89-4, Techniques Development Laboratory, National Weather Service, Department of Commerce, Silver Spring, MD, 14 pp.

Junker, N. W., J. E. Hoke and R. H. Grumm, 1989. Performance of NMC's regional models. *Weather and Forecasting,* **4,** 368–390.

Klaus, D., E. Jaurequi and W. Lauer, 1983: Urban-induced precipitation effects in the Mexico City region. *Arch. Meteorol. Geophys. und Bioklimatol.,* **33,** 275–288.

Miller, A., 1971: *Meteorology.* Merrill Publishing Co., Columbus, OH, 154 pp.

Murphy, A. H., and T. E. Sabin, 1986: Trends in the quality of National Weather Service forecasts. *Weather and Forecasting,* **1,** 42–55.

National Meteorological Center, 1987a: NMC Monthly Performance Summary. NOAA, U.S. Dept. of Commerce, Washington, D.C., Vol. 14, Nos. 5–6.

——, 1987b: NMC Quarterly Performance Summary. NOAA, U.S. Dept. of Commerce, Washington, D.C., Vol. 1, No. 1.

National Weather Service, 1978: The limited-area fine mesh model (LFM). *NWS Technical Procedures Bulletin,* No. 232, NOAA, U.S. Dept. of Commerce, Washington, D.C., 11 pp.

——, 1986: Modelling of physical processes in the nested grid model. *NWS Technical Procedures Bulletin,* No. 363, NWS/NOAA, Silver Spring, MD, 24 pp.

Navarra, J. G., 1979: *Atmosphere, Weather and Climate.* W.B. Saunders Co., Philadelphia, PA, 519 pp.

Phillips, N. A., 1979: The nested grid model. *NOAA Technical Report,* NWS 22, U.S. Dept. of Commerce, Silver Spring, MD, 80 pp.

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