

probability forecasts: a survey of National Weather Service forecasters¹

RP-75-9
Reprint

Allan H. Murphy²

National Center for Atmospheric Research³
Boulder, Colorado 80303

and Robert L. Winkler⁴

Graduate School of Business
Indiana University
Bloomington, Indiana 47401

Abstract

Some results of a nationwide survey of National Weather Service forecasters with regard to probability forecasting in general and precipitation probability forecasting in particular are summarized. Specifically, the questionnaire which was used in the survey, the participants in the survey (i.e., the forecasters), and the nature of the results are briefly described, and some recommendations based upon these results are presented.

1. Introduction

This paper summarizes some of the results of a nationwide survey of National Weather Service (NWS) forecasters relative to precipitation probability forecasting in particular and probability forecasting in general.⁵ The survey, which was conducted in May and June, 1972, was inspired by an earlier survey of the forecasters of the Travelers Weather Service (TWS) (Murphy and Winkler, 1971a, b) and by research related to probability assessment and evaluation undertaken by the authors of this paper (e.g., Murphy, 1972; Murphy and Winkler, 1970; Peterson, Snapper, and Murphy, 1972; and Winkler and Murphy, 1968a, b).

The objective of the survey was to obtain information about probability forecasting in an operational setting from a large sample of forecasters who make such forecasts on a regular basis. Specifically, we wanted to obtain information about the process by which precipitation probability forecasts are formulated; about the judgments of the forecasters relative to the definition, interpretation, and evaluation of these forecasts; about the

effects of feedback and experience upon such forecasts; and about the forecasters' opinions concerning the need for and use of probability forecasts by the general public and by specific user groups. This information was expected to provide a reasonably accurate description of 1) current practices and beliefs relative to probability forecasting in meteorology and 2) the nature and extent of any problems related to these practices and beliefs. With regard to the latter, the results of the TWS survey (Murphy and Winkler, 1971a, b) and of other studies (see, for example, Murphy, 1972, and Murphy and Winkler, 1974) have suggested that a number of important theoretical and practical problems do indeed exist. We believe that such problems must be resolved if probability forecasting in meteorology is to realize its full potential.

In this paper we briefly describe the questionnaire (Section 2), the participants (i.e., the forecasters) (Section 3), and the nature of the principal results (Section 4) and present some recommendations based upon these results (Section 5).

2. The questionnaire

a. Questionnaire topics

The questionnaire was designed to elicit information from the NWS forecasters with regard to the following topics related to probability forecasting:

- 1) the process by which probability of precipitation (PoP) forecasts are formulated, including the information sources examined, their relative importance, and their order of examination;
- 2) the relationships between a forecaster's judgments (expressed in probabilistic terms) and that forecaster's official PoP forecasts, and the factors which lead to differences between judgments and forecasts;
- 3) the definition of the event "precipitation";
- 4) the definition and interpretation of PoP forecasts;
- 5) the availability to forecasters of information in the form of feedback regarding the accuracy, reliability, etc., of their PoP forecasts and those of other forecasters, and the effects of such information upon their forecasts;
- 6) the evaluation of PoP forecasts and the reasons for any improvements in probability forecasting ability;
- 7) the effects of forecasting experience upon PoP forecasts;

¹ Supported in part by the National Science Foundation (Atmospheric Sciences Section) under Grants GA-31735 and GA-41232.

² On leave and visiting the International Institute for Applied Systems Analysis, Laxenburg, Austria (September 1974-March 1975).

³ The National Center for Atmospheric Research is sponsored by the National Science Foundation.

⁴ On leave and visiting the International Institute for Applied Systems Analysis, Laxenburg, Austria (July 1973 to January 1974), and the Department of Statistics, Stanford University, Stanford, Calif. (February-August 1974).

⁵ This paper is based upon a lengthy report (Murphy and Winkler, 1973) that describes these results in some detail. Copies of the report can be obtained from the authors of this paper.

- 8) the present and future need for and usefulness of forecasts expressed in probabilistic terms; and
- 9) the forecasters' opinions regarding the interpretation and use of PoP forecasts by members of the general public and specific user groups, and the effects of these opinions upon their forecasts.

These topics are similar to the topics considered in the TWS survey. However, the NWS questionnaire contained approximately twice as many questions as the TWS questionnaire and, as a result, the NWS questionnaire was able to treat these topics in considerably more detail.

b. Development and administration of questionnaire

A preliminary version of the questionnaire was developed in March 1972. This version of the questionnaire was pretested by a small group of NWS forecasters in early April. The results of the pretest, as well as the comments and suggestions received from individuals who were asked to review the preliminary version of the questionnaire, led to the preparation of a revised questionnaire during the latter half of April. This version of the questionnaire was mailed to all Weather Service Forecast Offices (WSFOs) and Weather Service Offices (WSOs) involved in the survey (see Section 3) in May. Almost all of the completed questionnaires were received prior to 1 July 1972.

The questionnaire was prepared in two formats, and the only difference between these formats was that the order in which the alternative answers to the questions were presented was reversed. Otherwise, questionnaires in the two formats were indistinguishable. The set of questionnaires that was sent to each WSFO and WSO consisted of approximately an equal number of questionnaires in each format. Thus, the probability that a particular forecaster in a specific office would receive a questionnaire of a given format was approximately one-half.

3. The participants

The survey was designed to obtain responses from all of the NWS meteorologists who were experienced in making PoP forecasts, as well as from a sample of the NWS meteorological technicians who also were experienced in making such forecasts. The regional contacts (see *Acknowledgments*) provided estimates of the number of meteorologists who were involved (at that time) in making precipitation probability forecasts at each WSFO and WSO in their respective regions. In addition, they selected a sample of at least 30 meteorological technicians at various WSFOs and WSOs in their respective regions who were making these forecasts. An appropriate number of questionnaires was then sent to each WSFO and WSO.

The forecasters who participated in the survey were asked to provide the following personal information:

- 1) classification (meteorologist, meteorological technician, etc.);

- 2) age;
- 3) education and training (general education and meteorological education and training);
- 4) forecasting experience (total and station experience in both general forecasting and probability forecasting);
- 5) nature of participation in PoP forecasting program (regular or occasional participant);⁶ and
- 6) attendance at the NWS Forecasters' Training Course.

Nationwide and regional tabulations of the forecasters' responses to the personal information questions are presented in Murphy and Winkler (1973). This information permits the results of the survey to be stratified by the forecasters' classification, age, education, and so forth.

4. The results

a. Response rate

The overall response of the forecasters to the survey is presented in Table 1. Nationwide, the response rate was almost 68%, while the regional response rates varied from 89.1% in the Central Region to 58.3% in the Western Region. These differences are largely the result of somewhat different procedures used by the regional contacts in identifying the survey participants, primarily meteorologists, in their respective regions. Moreover, the "real" response rates were actually higher than the "apparent" response rates indicated in Table 1, because some WSFOs and WSOs are believed to have received too many questionnaires. In any case, a response rate exceeding 50%, which was obtained in each region, is considered excellent for a survey administered by mail.

b. Response bias

In order to investigate the existence of any response bias that might have been produced by the order in

TABLE 1. Number of questionnaires sent out and returned and the response rates for each region and for the nation as a whole.

Region	Number of questionnaires sent out	Number of questionnaires returned	Response rate (%)
Eastern	293	173	59.0
Southern	253	189	74.7
Central	174	155	89.1
Western	295	172	58.3
Nation	1015	689	67.9

⁶ These processes were made somewhat more difficult by the fact that, just prior to and during the period in which the survey was conducted, the NWS policy according to which forecasters were concerned with either aviation forecasting or public weather forecasting was undergoing a significant change. PoP forecasts were, and still are, exclusively part of the public weather forecasting program, but present NWS policy is such that the forecasters are each involved in both programs.

which the answers to the questions were presented, we have compared the responses of the forecasters who completed questionnaires in the two formats. A preliminary analysis suggests that, except for the questions that required the forecasters to rank the alternative answers, the response bias was minimal. We are presently investigating the magnitude of the response bias for the ranking questions in more detail.

c. Nature of results

The results summarized in this paper consist of nationwide and regional tabulations of the forecasters' responses to the individual questions, and these tabulations together with a brief discussion of the results for each question are contained in Murphy and Winkler (1973).

Our conclusions, based upon an examination of these tabulations, can be summarized as follows:

- 1) For the most part, forecasters do not formulate their forecasts directly in probabilistic terms (i.e., they do not "think" in probabilistic terms). Further, they generally interpret probabilities in terms of relative frequencies.
- 2) Forecasters prefer to obtain a general, overall picture of weather situations, past and present, before making their PoP forecasts. In this regard, when other information sources are available, the forecasters do not appear to consider the guidance forecasts⁷ to be particularly important sources of information (vis-à-vis these other sources). On the other hand, if they can examine only one information source, then many forecasters would prefer to depend upon the guidance forecasts.
- 3) Almost all of the forecasters hedge at least occasionally. However, a distinction should be made between real hedging and apparent hedging. In the case of the former, the forecaster's judgment and forecast differ although both the judgment and the forecast refer to the same event (e.g., measurable precipitation at a particular point), while, in the case of the latter, the forecaster's judgment and forecast differ but the forecast and the judgment do not refer to the same event (e.g., measurable precipitation at two different points). In the presence of a strictly proper scoring rule such as the Brier score (Brier, 1950), real hedging is clearly inappropriate (see, for example, Murphy and Winkler, 1971b). While apparent hedging is not inappropriate in this sense, such hedging appears to be related to the use, by some forecasters, of different definitions of a precipitation event on different occasions [see 5) below], and this practice presents potential users of these forecasts with a difficult problem.
- 4) A serious problem appears to exist with regard to the forecasters' understanding of the relationships between precipitation probabilities for different periods.
- 5) Different forecasters prefer different definitions of a precipitation event and of a precipitation probability, and, as a result, they often use different definitions in connection with their PoP forecasts. As previously indicated, this situation presents potential users of these forecasts with the difficult, if not impossible task of determining their meaning on each occasion.
- 6) The forecasters in the Central Region receive more feedback relative to their own forecasts and those of their colleagues than do the forecasters in the other regions, and this information is perceived as having several important, and apparently beneficial, effects upon their probability forecasting processes and their PoP forecasts. In this regard, while some uses of feedback have been identified (see, for example, Hughes, 1965, 1967), further studies should be undertaken to investigate these and other possible ways of using such feedback to improve PoP forecasts.
- 7) Forecasters, in general, do not believe that probability forecasts can be evaluated by a single quantitative measure, but those who do prefer a relative measure to an absolute measure (e.g., they prefer the skill score to the Brier score).
- 8) Most forecasters believe that their PoP forecasts have improved during the last three years, primarily as a result of better probability guidance material and additional general forecasting experience. Further, they indicate that their probability forecasting processes have changed over time, and that these changes are primarily related to i) considering certain information sources either more or less important, ii) thinking of certain probabilities in particular situations, and/or iii) depending more upon guidance forecasts.
- 9) Approximately one-half of the forecasters receive requests to express forecasts of parameters other than precipitation occurrence in probabilistic terms. These requests relate most often to precipitation amount, precipitation type, and temperature, indicating a need for probabilistic forecasts of these parameters. In this regard, most forecasters believe that forecasts will be expressed in probabilistic terms at least as frequently, if not more frequently, in the future than they are at the present time.
- 10) Forecasters believe that categorical forecasts, accompanied by verbal modifiers when appropriate, are most useful for members of the general public and that point probability forecasts, perhaps accompanied by categorical forecasts, are most useful for members of specific user groups. In this regard, more than one-half of the forecasters believe that the general public interprets PoP forecasts properly, but uses them improperly.

⁷ The guidance forecasts of concern are the precipitation probability forecasts prepared by "objective" procedures (see Klein and Glahn, 1974) at the NWS's National Meteorological Center and routinely provided to forecasters in the WSFOs.

- 11) Approximately one-half of the forecasters are affected by their impressions of the general public's interpretation of their PoP forecasts, and these impressions cause some forecasters to modify their forecasts in the direction of categorical forecasts and other forecasters to modify their forecasts in the direction of the climatological probabilities.

5. Recommendations

The conclusions presented in Section 4 lead to the following recommendations:

- 1) A need exists to train forecasters in probability and statistics as these topics relate to probability forecasting in meteorology. This training could take several forms: i) self-training by the forecasters themselves in their WSFOs and WSOs through the use of programmed learning material; ii) short courses conducted at several locations (e.g., at each regional headquarters for the forecasters in that region); and iii) lectures and written material presented during the NWS Forecasters' Training Course. We believe that the approach involving programmed learning material offers certain advantages vis-a-vis the other approaches in that this approach would accomplish the objective of training all of the forecasters in a relatively short period of time and at a relatively modest cost. Thus, we recommend that a programmed learning manual be developed for use by the NWS forecasters which would contain detailed explanations and illustrations of those concepts in probability and statistics that are related to the theory and practice of probability forecasting in meteorology.
- 2) The fact that the forecasters perceive some confusion on the part of the public with regard to probability forecasts and the fact that the forecasters themselves exhibit some confusion both suggest that some confusion undoubtedly exists among members of the general public. Thus, we believe that a need also exists to initiate a program to educate the general public and members of specific user groups with regard to the purposes, meaning, and use of probability forecasts. Such a program should be designed to reach potential users of these forecasts in a variety of ways (e.g., directly by means of brochures and pamphlets, as well as indirectly through articles in newspapers, magazines, and other publications). Furthermore, this program should be conducted on a continuing basis. As a first step in initiating such a program, we recommend that a brochure be prepared which can be distributed free of charge to potential users and to others, such as members of the news media, who have direct contact with users. The brochure prepared by Bennett *et al.* (1969) would represent an excellent starting point for this effort.
- 3) With regard to forecasting practices, we believe that the ability of forecasters to converse in the "language of uncertainty" (i.e., in terms of probabilities) could be significantly enhanced if they began to think

and speak in these terms when formulating their forecasts on a day-to-day basis. Thus, we recommend that forecasters be encouraged to formulate all (or at least more) of their forecasts in probabilistic terms,⁸ and then, if necessary, they could translate these forecasts into categorical terms for dissemination to the public. In addition, we recommend that WSFOs and WSOs be encouraged to undertake probability forecasting experiments designed to provide forecasters with actual operational (or pseudo-operational) experience in expressing forecasts of a variety of meteorological variables in probabilistic terms.

Acknowledgments. We would like to express our appreciation to all of the NWS forecasters who participated in the survey and to those individuals in the regional and local offices of the NWS who assisted in the administration of the questionnaire. In particular, we would like to thank Dr. Duane S. Cooley, Technical Procedures Branch, NWS, and our regional contacts, Messrs. Carlos R. Dunn (Eastern Region), Paul L. Moore (Southern Region), Lawrence A. Hughes (Central Region), and Leonard W. Snellman (Western Region), without whose interest and cooperation this survey could not have been conducted.

⁸The fact that "objective" probability forecasts of a variety of parameters are routinely provided as guidance to forecasters in the WSFOs (see Klein and Glahn, 1974) should facilitate this practice.

References

- Bennett, E. S., J. E. Newman, P. W. Dailey, Jr., C. E. Petzhold, and L. H. Smith, 1969: Precipitation probabilities in weather forecasts. Lafayette, Ind., Purdue University, Cooperative Extension Service, Publication AY-174, 15 pp.
- Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1-3.
- Hughes, L. A., 1965: On the probability forecasting of the occurrence of precipitation. Washington, D.C., Environmental Science Services Administration, Weather Bureau, Central Region Technical Note 20-CR-3, 36 pp.
- , 1967: Improving precipitation probability forecasts using the Central Region verification printout. Kansas City, Mo., Environmental Science Services Administration, Weather Bureau, Central Region Technical Memorandum 15, 18 pp.
- Klein, W. H., and H. R. Glahn, 1974: Forecasting local weather by means of model output statistics. *Bull. Amer. Meteor. Soc.*, **55**, 1217-1227.
- Murphy, A. H., 1972: Probability forecasting in meteorology: a review of recent developments. Boulder, Colo., National Center for Atmospheric Research, unpublished manuscript, 19 pp.
- , and R. L. Winkler, 1970: Scoring rules in probability assessment and evaluation. *Acta Psychologica*, **34**, 273-286.
- , and —, 1971a: Forecasters and probability forecasts: the responses to a questionnaire. *Bull. Amer. Meteor. Soc.*, **52**, 158-165.
- , and —, 1971b: Forecasters and probability forecasts: some current problems. *Bull. Amer. Meteor. Soc.*, **52**, 239-247.

- , and —, 1973: National Weather Service forecasters and probability forecasts: preliminary results of a nationwide survey. Boulder, Colo., National Center for Atmospheric Research and Bloomington, Ind., Indiana University, unpublished report, 146 pp.
- , and —, 1974: Subjective probability forecasting experiments in meteorology: some preliminary results. *Bull. Amer. Meteor. Soc.*, 55, 1206-1216.
- Peterson, C. R., K. J. Snapper, and A. H. Murphy, 1972: Credible interval temperature forecasting. *Bull. Amer. Meteor. Soc.*, 53, 966-970.
- Winkler, R. L., and A. H. Murphy, 1968a: Evaluation of subjective precipitation probability forecasts. *Proc. First National Conference on Statistical Meteorology* (Hartford, 1968). Boston, Mass., American Meteorological Society, 148-157.
- , and —, 1968b: "Good" probability assessors. *J. Appl. Meteor.*, 7, 751-758.