LONG RANGE FORECASTING PRACTICES IN THE UNITED STATES: STATUS, OUTLOOK, AND ETHICS

Robert P. Harnack

Department of Meteorology and Physical Oceanography
Cook College - New Jersey Agricultural Experiment Station
Rutgers, The State University of New Jersey
New Brunswick, NJ 08903

1. INTRODUCTION

The Committee on Long Range Forecasting wishes to inform other members of the National Weather Association (NWA) and its council on the nature of long range forecasting, the role that the NWA should play vis-a-vis long range forecasting activities, and the future prospects for long range forecasting improvement. This is written by the Chairman of the Committee on Long Range Forecasting after consultation with committee members on the role that the committee should play in recommending actions which the NWA could take regarding long range forecasting practices, and on how to inform the NWA membership about the nature and problems of long range forecasting. The Committee on Long Range Forecasting has the following individuals as members: Chairman, Robert Harnack (Rutgers University), James F. Andrews (Intercon Weather Consultants), Anthony Broccoli (Rutgers University), Douglas LeComte (Environmental Data Service), Robert E. Livezey (National Meteorological Center), Joseph A. Shipps, Jr. (National Meteorological Center), and A. James Wagner (Climate Analysis Center).

2. CURRENT LEVEL OF SKILL IN LONG RANGE FORECASTING

Long range forecasting is defined here to include forecasting for periods in excess of five days from forecast day. The National Weather Service (NWS) currently issues forecasts for 6 to 10 day periods, 30-day periods (monthly outlooks), and 90-day periods (seasonal outlooks) to the public. The first type is issued by the Medium Range Forecast Group, while the latter two types are issued by the recently formed Climate Analysis Center, formed partially from the Long Range Prediction Group.

In addition to the NWS, various researchers issue experimental forecasts for portions of the long range to a selected audience. Table 1 summarizes long range forecasters, their main methods, and verification statistics as they have been reported in the literature or obtained from the forecaster directly. Only forecasts for the United States are included in this table. It should be noted that active long range forecast programs are evident in many other countries, especially in Great Britain, Canada, Germany, Japan, and the Soviet Union. For detailed information on each of the listed forecast methods, the interested reader should refer to the references given for each. No attempt has been made to include forecasters who have not published their methodologies or provided verification statistics of some sort. In addition, the given forecast method must be in use currently for making real time forecasts for inclusion in this table. The reader should not draw conclusions from the table with regard to comparative quality or skill of forecast method since sizes, format of forecasts, and main purpose of forecast (i.e., operational vs. research) varies considerably. This information, edited by the author, is provided only for the purpose of informing the membership on the range of forecast methods employed and the kinds of skill reported by the forecasters. It can be seen that at least some skill relative to chance has been achieved for these forecast methods.

3. LIMITS ON PREDICTABILITY

Various theoretical studies have been conducted in order to assess the greatest number of days in advance that skillful, dynamic predictions of the atmosphere state (e.g., isobaric height field) can be made. The work and reviews by Lorenz (1969) address this question. The consensus is that daily predicitions cannot be made with any skill greater than a few weeks in advance. This is believed, then, to be the THEORETICAL upper limit of predictability for DAILY predictions. No definitive work has been completed yet on the theoretical limits on prediction of the mean period (e.g., monthly, seasonal, etc.) state of the atmosphere. The operational and experimental forecast methods cited earlier all pertain to mean period forecasts, not daily predictions in the long range. The current practical limit of predictability (about five days) is obviously shorter than the theoretical limit when daily, dynamic predictions of the atmosphere state are being considered. The gap between the current practical and theoretical limits is due to the inability to accurately define the initial state of the atmosphere (e.g., insufficient data coverage and resolution, observational errors); numerical errors due to the methods used to solve the governing differential equations; and errors due to imcomplete understanding of all the important atmospheric processes. The essential point to be emphasized here is that skillful, daily forecasts beyond a few weeks in advance are probably impossible and that daily forecasts beyond a week are CURRENTLY unattainable. Therefore, forecast claims that are contrary to the above must be viewed with great skepticism, and it must be added that despite a few such forecast claims to the contrary, no evidence has been put forward to substantiate them.

The rest of this paper represents a drawing together of opinions, solicited by the Chairman, from each member of the Long Range Forecasting Committee. The degree of unanimity and some minority opinions are also included.

4. PROSPECTS FOR FORECAST IMPROVEMENT

All but one member was optimistic about continued improvement of mean period forecasts for the medium range (i.e., 6 days to 3 weeks). Table 1 indicates some current operational skill for forecasts in the 6 to 10-day period. Members were generally in agreement that at least marginal skill will be attained for forecast periods extended out to 2-3 weeks in advance, within a decade, without a major breakthrough in physical understanding. Factors cited for this optimistic viewpoint include the belief in continued improvement of numerical weather prediction models, including extension to global domains and more complex physics, improved data bases resulting from the GARP experiments, and improved understanding of processes and phenomena on extended period time scales. Research encouraged by the passage of the National Climate Act was also cited as a basis for optimism.

Expectations for forecast improvement on the monthly and seasonal time scale was also generally optimistic, although great amounts of skill are not anticipated. Some members believe that 30-day forecasting will improve due to extension of skill of numerical prediction models out to several weeks, while others believe improvement will come from the application of sophisticated statistical methods, which have been used mainly by researchers, and from increased empirically-derived understanding of the climatic system. For seasonal periods, and even longer, statistical methods will have to be relied on. Seasonal temperature predictions are likely to continue to improve slowly, especially as the data base lengthens, while seasonal precipitation forecasts are likely to lag behind those of temperature in terms of skill and improvement. Extension of seasonal or longer period predictions to several seasons in advance is possible in the next decade.

5. Role of the NWA regarding ethical practices and guidelines for long range forecasting.

The committee feels strongly that the NWA should issue guidelines that outline what is acceptable and what is not for the issuance of long range forecasts. These guidelines should include the following:

(1) All forecasts should be given in a format which allows them to be verified objectively and quantitatively. Acceptable formats include category forecasts for SPECIFIC periods and locations, in which the category limits are clearly defined; departure from normal forecasts by period and location; and forecasts which state specific, actual values of parameters for a defined period and location. These formats can be accommodated in map, table, or narrative form. Unacceptable formats include forecasts that involve the use of undefined or vague terminology or do not specify exactly the period or area that the forecast applies to. For example, a forecast of "a cold winter for the Midwest" is unacceptable, since it cannot be verified objectively and quantitatively. "Cold" is undefined and not quantitative, and "Midwest" is similarly undefined.

(2) All forecasts should be accompanied by some realistic indication of the probability of success, preferably in terms of or based on previous verification scores. The basis for success claims should be given. If such an accompanying statement does not append each forecast issued, then it either should be given periodically or the information should be referenced. In any case, potential users should have this type of information available to them, in case they desire to use it, along with the forecast itself.

(3) All forecasts should include or be in a format such that the user can easily compute verification scores for climatology versus those of the forecaster. If the forecast is given in terms of departure from normal or in categories, then this point is taken care of. If the forecast is essentially climatology, then this should be made clear to the user.

The members of the committee also feel strongly that the NWA should include in a statement on long range forecasting practices, the ever increasing growth of unethical practices in long range forecasting from portions of the private sector. While the committee does not feel that individuals or groups should be mentioned by

Table 1. Operational and experimental long range forecasts issued in the United States

Forecaster	Forecast Period	Format of Forecast	Verification Statistics	Forecast Method	References
Medium Range Forecast Group of NWS	6-10 day forecasts from 12/77 through 2/79. Forecasts produced 3 times weekly.	Five temperature classes and 3 precipitation classes used for most of U.S. Mean period type forecast made.	Temperature: 32% of U.S. in correct class (21% expected by chance). Precipitation: 41% of U.S. in correct class (35% expected by chance).	NWP guidance modified by climatology, recent trends, and taking account of recent errors to produce 6-10 day subjective 500 mb hts. Temperature forecast produced statistically from ht. forecast, plus some subjective modifications. Precipitation forecast mostly subjective using model output, climatology, and synoptic reasoning.	National Weather Service (1979), Andrews (1978), NMC (1979).
Climate Analysis Center of NWS	30-day forecasts from Fall 1974 to Spring 1979. Forecasts produced every 2 weeks. (Total of about 110 forecasts.)	Three temperature classes and 2 precipitation classes used for period.	Temperature: 40% of U.S. in correct class (~33% expected by chance). Precipitation: 52% of U.S. in correct class (~50% expected by chance).	Mainly statistical, except short range NWP output used for days 1-5. Subjective 30 day 700 mb ht. forecast made first using NWP, climatology, and past observed data plus synoptic reasoning. Temperature forecast produced mainly statistically from ht. forecast. Some manual modification employed. Precipitation forecast produced from ht. forecast mainly subjectively with the use of various climatological aids.	National Weather Service (1979), Namias (1953).
Climate Analysis Center of NWS	90-day forecasts from Winter 1959 to Spring 1979. Forecasts produced each season. (Total of about 82 forecasts.)	Two temperature classes and 2 precipitation classes used for period. "Indeterminant" also used in selected regions for temperature forecasts.	Temperature: 60% of U.S. in correct class, where forecast was made (~50% expected by chance). Precipitation: 54% of U.S. in correct class (~50% expected by chance).	Statistical methods (e.g. autocorrelation maps of season to season 700 mb hts. at various lags) used to infer important circulation features. Temperature and precipitation classes subjectively forecast from projected circulation features using synoptic reasoning and various climatological aids.	National Weather Service (1979), Namias (1964).

Table 1. Continued

Forecaster	Forecast Period	Format of Forecast	Verification Statistics	Forecast Method	References
R. Harnack	90-day forecasts for winter temperature only in period 1973-1979. (Total of 7 forecasts for eastern 2/3 of U.S. only.	Three temperature classes used for period. Various models tested on independent sample of 7 winters.	SST model: 50% Circulation model: 28% Circulation - SST model: 39% (33% expected by chance). Note: These statistics pertain to independent test forecasts, not real-time forecasts.	Various statistical models were developed, then tested on a recent independent sample of years. All forecasts included here are objective. Models differed from one another with respect to predictor types used. Pacific & Atlantic SST are used in one, 700 mb hts are used in another, and SST plus hts were used in a third model	Harnack (1979)
T.P. Barnett and R.W. Preisendorfer	90-day forecasts for all seasons in period 1952-75.	Three temperature classes used.	Verifications presented in maps, graphs, & tables in terms of "local" and "global" skill scores stratified by season and method. Percent correct not given or derivable in reference. Positive skill was evident for various seasons, lags, and methods.	Analogue methods used, in which a temperature class is predicted for a station on the basis of evolution of climate for a "best match" between the current observed conditions and a previous set of conditions. Note: See reference for further details and for verification scores. All forecasts in paper were not real-time forecasts.	Barnett and Preisendorfer (1978)
G.C. Henricksen	120-day forecasts for winter temperature only in period 1971-75. (Total of 5 forecasts for eastern 2/3 of U.S. only.	Maps of temperature departure from normal.	Root mean square error for 16 stations of 1.63°F which compares favorably to a standard deviation of 2.5°F.	Statistical methods used in which various predictors (SST mostly) have been related to winter temperatures at each of 16 stations.	Henricksen (1979)
J. Namias	90-day forecasts from Winter 1974 to Winter 1979. Forecasts produced each season. (Total of about 21 forecasts)	Three temperature and 3 precipitation classes used for period.	Temperature: 39% of U.S. in correct class (33% expected by chance). Precipitation: 38% of U.S. in correct class (33% expected by chance).	Statistical methods used, in which obs. sea surface temperature of N. Pacific is used to predict SST for the following period, then from predicted SST, a 700 mb ht. forecast is made. Finally, temperature and precipitation are forecast by relating predicted 700 mb heights to each forecast parameter. Subjective judgment is employed at each step.	Namias (1976)

name, it does feel that various types of practices should be specifically cited as unethical and unprofessional. These practices include making forecasts for specific days beyond a few weeks, making unsubstantiated claims of accuracy about forecast methods, and using language in forecasts which does not allow them to be verified objectively and quantitatively.

All but one member of the committee feels that the NWA should, at some point, revoke the NWA membership of any member who willfully violates ethical guidelines which the NWA sets up. Backing of the entire membership is needed to set up a mechanism which can result in membership revocation when circumstances warrant such action. Perhaps a referendum of sorts is needed prior to setting up such a mechanism. We think that the NWA should squarely face up to this issue, unlike another meterogical organization which we know! A void exists to be filled, and the NWA should fill it. Some specifics include setting up a committee to: review long range forecasting practices of its members, persuade said members involved in long range forecasting to follow NWA guidelines (after they are written), warn offenders that are not willing to follow the guidelines, and recommend membership revocation to the NWA Council in the worst cases. Perhaps the Long Range Forecasting Committee could play this role, since it is composed of members who know and understand the problems inherent in long range forecasting. In any case, enforcement of guidelines is essential.

Finally, the committee believes that despite the marginal skill of many long range forecasts, as well as the necessarily small samples on which they are based and/or tested, long range forecasts should still be issued as long as certain guidelines, like those discussed earlier, are followed. The committee is anxious to hear from the NWA Council and other NWA members on the recommendations contained here.

Author's note: This paper was submitted for publication in October 1979. Due to unfortunate circumstances, the editors were unable to publish it until now. It is recognized that some of the information contained herein may be out of date, for which the editors apologize to the author.

Editor's note: We believe delays of this nature will not be the rule in the future, as we endeavor to get the Digest out on time. However, individuals considering submitting papers must be aware of some delay even in the best of circumstances due to the large number of incoming articles and the quarterly publication schedule of the magazine. We thank everyone for their patience.

References

Andrews, J.F., 1978: Medium-Range Forecasting in the National Meteorological Center. Preprints Conf. on Weather Forecasting and Analysis, Silver Spring, Amer. Meteor. Soc., 13-19.

Barnett, T.P. and R.W. Preisendorfer, 1978: Multifield Analog Prediction of Short Term Climate Fluctuations Using a Climate State Vector, J. Atmos. Sci., 10, 1771-1787.

Harnack, R., 1979: A Further Assessment of Winter Temperature Predictions Using Objective Methods, Mon. Wea. Rev., 107, 250-267.

Henricksen, G.C., 1979: An Attempt to Project Temperature Departures for the Eastern United States, Nat. Wea. Digest, 4, 27-37.

Lorenz, E.N., 1969: Three Approaches to Atmospheric Predictability, Bull. Amer. Meteor. Soc., 50, 345-349.

Namias, J., 1953: 30-Day Forecasting: A 10-Year Experiment, Meteor. Monographs, 2, Amer. Meteor. Soc., 83pp.

Namias, J. 1964: A 5-Year Experiment in the Preparation of Seasonal Outlook, Mon. Wea. Rev., 92, 449-464.

Namias, J., 1976: Seasonal Forecasting Experiments Using North Pacific Air-Sea Interactions, Preprint Vol. 6th Conf. on Wea. Forecasting and Analysis, Albany, N.Y., Amer. Meteor. Soc., 13-16.

National Meteorological Center, 1979: Verification of 1-5 Day and 6-10 Day Mean Forecasts, National Meteorological Center Newsletter, April 1979, 6.

Nationl Weather Service, 1979: Forecasting Handbook No. 1 - Facsimile Products, U.S. Dept. of Commerce, Washington, D.C.