Aust. Met. Mag. 29 (1981) 25-29 551.509

Aiding the weather forecaster: comments and suggestions from a decision analytic perspective

G. Allen, Head Office, Bureau of Meteorology, Melbourne

(Manuscript received November 1980; revised February 1981)

Weather forecasting is considered as a decision task. Similarities with medical diagnosis are noted and it is suggested that methods used in computer-aided medical diagnosis, particularly Bayesian probability aggregation methods, hold promise for weather forecasting. It is also suggested that the development of weather forecasting procedures has not been sufficiently innovative in the employment of human pattern recognition ability. This neglect may lead to sub-optimal weather forecasting systems.

Introduction

Within the weather forecasting and meteorological research community, weather forecasting is most usually thought of as an application of meteorological physics and mathematics. To think of it solely in these terms, however, is to overlook that weather forecasting is also a decision making task. This perspective has the advantage that it makes available an additional set of methods for improving the quality of weather forecasts. These methods focus attention on decision making behaviour, discovering how decisions are reached and how this process can be improved.

Devices or procedures which raise the quality of decisions are called decision aids (e.g. Triggs 1975). Weather forecasters use decision aids without calling them such. The oldest and most valued decision aid of the Australian weather forecaster, at present, is the synoptic weather chart. It functions as a decision aid by transforming an otherwise unintelligible mass of numerical information about the atmosphere into visible patterns. The forecaster recognises certain pattern elements and uses his knowledge and experience of their behaviour to decide the likely course of weather events during the forecast period. A related device is the so-called 'objective forecasting aid'. The term objective forecasting aid has evolved within the weather forecasting community, and refers to a device or procedure which furnishes the forecaster with some estimate (numerical, categorical or probabilistic) of a future weather variable. Such an estimate is an additional information item to be considered by the forecaster, and may simply increase his cognitive load without leading to more accurate forecasts. MacDonald (1977) reports an example of this. Thus objective forecasting aids may or may not turn out to be decision aids.

In order to design decision aids for forecasters (at some given time in the evolution of observational and information-processing technology, and meteorological science) it is necessary to know what their decision processes are. Decision analysis (Howard 1980) yields such information. A complete decision analysis would describe the information items from which decision input is selected, the form in which these are presented, the strategies, decision rules, time and other resources used in selecting, evaluating, organising and aggregating decision input to reach a decision, the decision maker's uncertainty at different stages and in different circumstances, the demand for statistical or other systematic information, the range of decision options, the objectives of the decision, the perceived penalties and rewards of various incorrect and correct decisions, and measures of decision validity and utility. Given part or all of such a description, the decision analyst, knowing something of the characteristics of human information processing in general, together with the decision maker, knowing something about the decision setting in particular, can identify opportunities for decision aiding. Depending on the objectives of the decision analysis project and the resources committed to it, changes to the decision task can range from the introduction of a simple look-up table to avoid errors in recall, to a complete restructuring of the entire decision task.

In the following, an attempt will be made to draw some useful inferences from a superficial decision analytic view of weather forecasting.

Weather forecasting and medical diagnosis

A number of similarities between weather forecasting and medical diagnosis become evident when each is considered as a decision task. For example, in both tasks the decision maker is faced with information of varying degrees of reliability and significance, and is required to reach the most reliable judgment possible in the least possible time. Both tasks involve selective gathering of data from prodigious data sources. In each task, complex physical relationships underlie the connection between available information and the state to be inferred. Both kinds of practitioners rely on education and experience, and both make fallible judgments. Until two decades ago both kinds of practitioners performed their decision tasks without computer assistance. Computers are now firmly entrenched in both work areas, but with interesting differences in their mode of use, as will emerge.

Sufficient evidence relating to human judgment in medical diagnosis accumulated between 1960 and 1970 to support the view (Goldberg 1970) that humans should not make diagnostic decisions: '. . . the proper role of the human in the decision making process is that of scientist, (a) discovering or identifying new clues which will improve predictive accuracy, and (b) constructing new sorts of systematic procedures for combining predictors in increasingly more optimal ways'. Subsequent evidence and published debate on the question of whether computer-based decision systems perform better than skilled humans (Libby 1976a, b; Goldberg 1976) has strengthened Goldberg's case.

The sorts of computer-based decision systems which outperform human diagnosticians include linear regression models of the diagnosticians themselves, Bayesian probability aggregation systems, and symptom profile matching systems (Croft 1972; Greist 1973; Gustafson et al. 1973). The reasons for under-performance by humans are several. They include fundamental limitations in human information processing ability, in particular, imperfect and biased long-term recall, short-term recall of only three or four items, and slow serial execution of logical or algebraic operations on conceptual information. These limitations lead to human information processing strategies which minimise cognitive effort. For example in tasks calling for Bayesian aggregation of prior probability and information specific to a case at hand, prior probability is usually neglected (Lyon and Slovic 1976), and in hypothesis testing, disconfirming evidence is usually ignored (Evans and Wason 1976; Mynatt et al. 1978). As well as using sub-optimal strategies, humans have a very robust predisposition to change strategies without justification, which leads to judgmental inconsistency and error (Brehmer and Kuylenstierna 1978).

Two tentative inferences might be drawn from the above account. Firstly, it is probable that Goldberg's assertion applies to weather forecasting as well as medical diagnosis. Secondly, the procedures used in medical diagnosis might be effective in weather forecasting. Indeed some progress has already been made in applying them. Mason (1977) has developed a Bayesian probability aggregation procedure for fog forecasting. Linear regression models of forecasters' judgments have not been developed, but linear regression objective forecasting aids are numerous; all but one of the ten statistical forecasting procedures presented at the Conference on Probability and Statistics in Meteorology (American Meteorological Society 1979) were based on either multiple linear regression or discriminant analysis.

That these methods resemble those used in medical diagnosis is fortuitous rather than the result of a deliberate attempt to exploit medical experience, and what appears (Croft 1972) to be the most powerful method, Bayesian probability aggregation, has so far been least used in weather forecasting. A particularly attractive property of self-instructing Bayesian diagnosis systems (Vishnevskiy et al. 1973) is their ability to refine their decision making as they accummulate experience. The ability of humans to do this in decision tasks as complex as medical diagnosis or weather forecasting is very limited (Hammond 1971; Gedzelman 1978). Development of computer diagnosis has proceeded vigorously in the last decade. In a recent review, Rogers et al. (1979) cite 48 computer-based diagnostic systems. This appears to be a rich field for forecasters to prospect.

The human role in weather forecasting

The development of numerical weather prediction has also proceeded vigorously since establishment of the first operational system in 1955 (JNWPU 1957). And despite the current practice in Australian forecasting offices of treating the products of numerical weather prediction as 'forecast guidance information' to be rejected, accepted, modified, weighted or otherwise dealt with at the forecaster's discretion, it is widely assumed that numerical weather prediction, based on fluid dynamics and atmospheric physics,is the 'proper', and will eventually be the best, means to generate forecasts. Numerical weather prediction based on physics has already outperformed man at a number of traditonal forecasting tasks. Sixty per cent (37 out of 62) of the charts transmitted daily by the Australian National Meteorological Analysis Centre are now generated by computer. Judging by American experience (e.g. MacDonald 1977), weather forecasting procedures based on physical models plus model output statistics can be expected to outperform man at rain forecasting. It is a simple step to assume that these demonstrations of the superiority of computer products vindicate numerical weather prediction and the progressive transfer of forecasting tasks from man to computer.

But comparing the skill of any forecasting procedure with that of a human operating in the traditional forecaster role is hardly a stringent test. Even comparatively simple objective forecasting aids such as Maine's (1958) maximum temperature prediction scheme manage to pass this test. It is an easy test to pass because traditional forecasting procedure, i.e. synoptic analysis coupled with careful deliberation, evolved to suit man's limited logical processing ability. Synoptic analysis is not only a means of organising information, it is also a simplifying process, to reduce the forecaster's decision input to manageable proportions. And the decision strategies which operate on this simplified decision input also employ simple decision rules. Moreover, traditional forecasting procedure is sub-optimal not only because of simplification; errors are also introduced by strategy changing, risk preference, imperfect recall, confirmation bias, probability misjudgment and penalty avoidance.

It is surprising then that the merit of traditional forecasting procedure has not been called into question in connection with the introduction of computers into forecasting. The earliest application of computers to the forecasting task in Australia (Maine 1968) was to produce meteorological charts, both current and prognostic, of the same kind that forecasters traditionally prepared. In 1980 this is still the main contribution of the computer to forecasting in Australia. Research is in progress (e.g. Woodcock 1980) to link computer prognoses to future weather variables. This also has traditionally been a task for human judgment. As long as human performance in these traditional tasks is taken as the standard against which machine-based methods are tested, machine-based methods will eventually overtake man and replace him. The process is in motion, and the consensus view in papers at the 1978 Forecasting Conference of the American Meteorological Society (see for example Klein and Haggard 1978) is that it will continue. According to this view, short-range mesoscale forecasting and now-casting are the only human forecasting roles likely to survive. Other than that man's roles will become peripheral to computer-based forecasting systems; interfacing with the community, solving special forecast needs, monitoring system performance and carrying out system development.

An alternative course is possible. Although traditional forecasting procedure evolved to suit man's limited logical processing capability, this limitation is now irrelevant. Computers can handle logical and algebraic tasks of immense complexity. It is now possible to consider using man's fast and efficient (see for example visual fault detection in the NASA Apollo program (Bowen 1971, Figs 1 and 2)) visual pattern and motion recognition ability as a source of input for forecasting systems based on statistical relationships (for which physical explanations may or may not be available). Information displays presented for view need not be limited to traditional linear space frames. Predictive relationships have emerged using novel frames, for example, rain in humidity/stability space (Doneaud et al. 1979), and rain in pressure/time space (Lynch et al. 1970). What systematic phenomena would appear if conditions at one site were represented by a point moving in pressure/temperature/humidity space, and rain and frontal passages were marked on these trajectories? Clearly this is speculation, not respectable science, but it might yield decision aids. At any rate it would be hasty to discard man on the basis of his performance as a traditional forecaster, when his performance at that task is limited by his logical capability. His visual information processing capability is superb, and its potential in alternative approaches to weather forecasting is untried.

Conclusion

From a decision-oriented examination of weather forecasting it appears that (a) forecasting systems might benefit from the application of computer-based techniques currently used in medical diagnosis, particularly Bayesian methods, and (b) designers of forecasting systems should not hastily abandon humans as information processors on the basis of their limited success in traditional forecasting roles. Man/machine systems employing visual information processing may have advantages over traditional and pure machine systems.

It is tempting to seek reasons why both these proposals are novel. Medical diagnostic procedures have probably escaped notice because the content and setting of medical decisions are so grossly different from those of weather forecasting decisions. Innovative ways of representing meteorological data to expose hitherto unrecognised regularities have not been sought, because synoptic meteorology has already provided a perceptual/conceptual framework which custom has placed beyond challenge. In fact, this framework has become such an integral part of our thinking about the atmosphere that even the visible products of numerical weather prediction are expressed in its terms. Furthermore, numerical weather prediction has been met with such high and universal optimism that relatively little research effort has been directed to alternative weather prediction procedures during the last 25 years. Finally, because meteorology is a physical science, recognition that decision processes and human performance viewpoints are relevant to the forecasting task (American Meteorological Society 1977), has been late in coming.

References

American Meteorological Society. 1977. Summary of the Sessions on Problems in Forecasting, AMS 57th Annual Meeting (Recommendation 34). *Bull. Am. Met. Soc.*, *58*, 404-407.

American Meteorological Society. 1979. *Papers from the 6th Conference on Probability and Statistics in Meteorology*, Banff.

Bowen, H. M. 1971. The imp in the system. *The Human Operator in Complex Systems*, Singleton, W. T., Easterly, R. S. and Whitfield, D.C., Eds., London: Taylor and Francis, 12-19.

Brehmer, B. and Kuylenstierna, J. 1978. Task information and performance in probabilistic information tasks. *Organizational Behaviour and Human Performance*, *22*, 445-64.

Croft, D. J. 1972. Is computerized diagnosis possible? *Comp. Biomed. Res.*, *5*, 351-67.

Doneaud, A. A., Sumedha Sengupta, Smith, P. L. and Dennis, A. S. 1979. A combined synoptic and statistical method for forecasting rain volume over small areas. *Papers from the 6th Conference on Probability and Statistics in Meteorology*, Banff, American Meteorological Society, 39-45.

Evans, J. St. B. T. and Wason, P. C. 1976. Rationalization in a reasoning task. *Br. J. Psych.*, *67*, 479-86.

Gedzelman, S. D. 1978. Forecasting skill of beginners. *Bull. Am. Met. Soc.*, *59*, 1305-09.

Goldberg, L. R. 1970. Man versus model of man: A rationale, plus some evidence, for a method of improving a clinical inference. *Psych. Bull.*, *73*, 422-32.

Goldberg, L. R. 1976. Man versus model of man: Just how conflicting is that evidence? *Organizational Behaviour and Human Performance*, *16*, 13-22.

Griest, J. H., Gustafson, D. H., Stauss, F. F., Rouse, G. J., Laughren, J. P. and Chiles, J. A. 1973. A computer interview for suicide risk prediction. *Am. J. Psych.*, *130*, 1327-32.

Gustafson, D. H., Kestly, J. J., Ludke, R. J. and Larson, F. 1973. Probabilistic information processing: Implementation and evaluation of a semi PIP diagnostic system. *Comp. Biomed. Res.*, *6*, 355-70.

Hammond, K. R. 1971. Computer graphics as an aid to learning. *Science*, *172*, 903-08.

Howard, R. A. 1980. An assessment of decision analysis. *Operations Research*, *28*, 4-27.

JNWPU. 1957. One year of operational numerical weather prediction. (Report by Staff Members, Joint Numerical Weather Prediction Unit) *Bull. Am. Met. Soc.*, *38*, 263-68.

Klein, W. H. and Haggard, W. G. 1978. The future roles of people and machines in weather forecasting. *Preprints, Conference on Weather Forecasting and Analysis and Aviation Meteorology*, Silver Spring, Md., American Meteorological Society, 265-70.

Libby, R. 1976a. Man versus model of man: Some conflicting evidence. *Organizational Behaviour and Human Performance, 16*, 1-12.

Libby, R. 1976b. Man versus model of man: The need for a nonlinear model. *Organizational Behaviour and Human Performance, 16*, 23-26.

Lynch, K., Thorp, D. and Horry, S. 1970. The high correlation between pressure and the commencement of rain at Perth during the winter months. *Met. Note 36*, Australian Bureau of Meteorology.

Lyon, D. and Slovic, P. 1976. Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psychologia, 40*, 287-98.

MacDonald, A. E. 1977. The MAN-MOS program. *NOAA Technical Memorandum NWS WR-112*, United States Department of Commerce, Washington, D.C.

Maine, R. 1958. Maximum temperature prediction. *Aust. Met. Mag., 22*, 50-63.

Maine, R. 1968. A real time experiment in operational meteorological data processing. *Aust. Met. Mag., 16*, 75-99.

Mason, I. B. 1977. Objective probabilities for fog at Canberra Airport using Bayesian aggregation and principal components. *Aust. Met. Mag., 25*, 121-34.

Mynatt, C. R., Doherty, M. E. and Tweney, R. D. 1978. Consequences of confirmation and disconfirmation in a simulated research environment. *Q. Jl. Exp. Psychol., 30*, 395-406.

Rogers, W., Ryack, B. and Moeller, G. 1979. Computer aided medical diagnosis: literature review. *International Journal of Biomedical Computing, 10*, 267-89.

Triggs, T. J. 1975. Decision aids to the individual medical practitioner. *Human Factors in Health Care*, Pickett, R. M. and Triggs, T. J., Eds., Massachusetts: D. C. Heath and Co., 153-68.

Vishnevskiy, A. A., Artobolevskiy, J. J. and Bykhovskiy, M. L. 1973. *Machine Diagnosis and Information Retrieval in Medicine in the USSR*. Translation published by U.S. Department of Health, Education and Welfare. DHEW Publications (NIH) 73-424.

Woodcock, F. 1980. On the use of analogues to improve regression forecasts. *Mon. Weath. Rev., 108*, 292-97.