KARL POPPER AND THE ACCOUNTABLILITY OF NUMERICAL FORECASTING

H. Tennekes
Royal Netherlands Meteorological Institute
De Bilt, Netherlands

1. INTRODUCTION

On occasion one wonders what certain famous people would have said about a subject one is vitally interested in, provided they had been in a position to study the subject in depth. This thought occurred to me when I read Karl Popper's 1965 essay "Of Clouds and Clocks" a few years ago. Karl Popper (born 1902), the outspoken philosopher of science, the originator of the sobering idea that science progresses by falsification, not verification (the latter being ultimately impossible), introduced clouds in his essay to represent physical systems "which are highly irregular, disorderly, and more or less unpredictable", and clocks to represent systems "which are regular, orderly and highly predictable in their behavior".

Popper's interest in predictability was fueled by his opposition to the views held by the proponents of determinism. The central thesis of determinism is the "staggering proposition" that all clouds are clocks; in this view the distinction between clocks and clouds is not based on their intrinsic nature, but on our lack of knowledge. If only we knew as much about clouds as we do about clocks, clouds would be just as predictable as clocks. Or, in meteorological terms, a perfect model of the atmosphere, initialized with perfect data from an observation network of infinite resolution, and run on an infinitely powerful computer, should in principle produce a perfect forecast with an unlimited range of validity.

If, at the time he wrote his essay, Popper had been living in 1990 and therefore had been aware of the explosive growth of our knowledge about the chaotic behavior of nonlinear systems since the publication of Ed Lorenz' 1963 paper on deterministic, nonperiodic flow, it would have been a lot easier for him to demolish the determinist position. We now know that the effective forecast range for a complex physical system is finite and limited to the typical life span of its most energetic phenomena. The "prediction horizon" of midlatitude weather is comparable to the average

life span of extratropical cyclones: a week, no more. Even satellites, the most clocklike of human artifacts, are not immune: the small deviations in their orbits caused by irregularities in the gravitational field also can be predicted just a few days ahead. In other words: all clocks are clouds, if only one looks closely enough.

My own interest in predictability dates from 1976, when I attempted to reconcile my experience in turbulence research with the error cascade calculations published by Lorenz, Leith and Kraichnan. By 1986 my thoughts had evolved to the point where I realized that skill forecasting should be one of the central goals of predictability research. At a workshop organized by ECMWF in April 1986 I introduced the slogan: "No forecast is complete without a forecast of forecast skill". The admittedly vague motivation for coining this phrase was that it will enhance the credibility of meteorology if we are honest about the intrinsic limitations to forecast skill. Since it is unlikely that the average performance of numerical weather forecasts can be substantially improved, it is prudent to invest in research that might lead to tolerably skillful skill forecasts. Toward the end of the forecast range the quality of a forecast will always be poor, no matter how much that range is extended. A little bit of guidance on the anticipated reliability of weather forecasts is a whole lot better than no guidance at all.

Little did I know at the time that I was dabbling in the periphery of a fundamental problem in the methodology of scientific inquiry. Fortunately, in 1990 I chanced upon a copy of the second volume of Popper's "Postscript to the Logic of Scientific Discovery" and read the astonishing paragraphs on scientific accountability that I shall quote from <u>in extenso</u> below. Those texts, based on two lectures Popper gave as early as 1950, made it painfully clear that I had merely scratched the surface of the problem I had pronounced upon, and that I had plenty of reason to be embarassed about the superficial way in which I had treated the subject at the ECMWF workshop.
In this essay, I am trying to make up for this defect. Better late than never.

2. ACCOUNTABILITY

Popper begins his quest for honesty in scientific inquiry by pointing out that it is necessary to lay down certain ground rules on the precision required of scientific calculations:

> "The fundamental idea underlying scientific determinism is that the structure of the world is such that every future event can in principle be rationally calculated in advance, if only we know the laws of nature and the present state of the world. But if every event is to be predictable, it must be predictable with any desired degree of precision: for even the most minute difference in measurement may be claimed to distinguish between different events"(The Open Universe, 1982, p.6).

In other words, we have to agree in advance on the accuracy we demand. We should not allow ourselves an escape route if things go wrong. A few pages later, Popper explains the issue in more detail:

> "Scientific determinism requires the ability to predict every event with any desired degree of precision, provided we are given sufficiently, precise initial conditions. But what does 'sufficiently' mean here? Clearly, we have to explain 'sufficiently' in such a way that we deprive ourselves of the right to plead - every time we fail in our predictions - that we were given initial conditions which were not sufficiently precise. In other words, our theory will have to account for the imprecision of the prediction: given the degree of precision which we require of the prediction, the theory will have to enable us to calculate the degree of precision in the initial conditions that would suffice to give us a prediction of the required degree of precision. I call this demand the principle of accountability" (The Open Universe, p.11).

The thrust of this argument is evident. Whenever they fail in their predictions, scientists tend to blame the poor accuracy of the observations, the lack of computer power and the inadequate parameterization in their numerical models, rather than their own lack of skill in computing the accuracy that can be obtained with present resources. Sloppy reasoning of this kind is responsible for much of the thoughtless expansion and escalation numerical modelers in all branches of science indulge in.

Popper holds researchers accountable for their own work. He feels, and so do I, that it is besides the point to complain about available resources being finite. They will always be. Instead, one should focus one's modeling skills on the creation of algorithms that compute the forecast skill obtainable at the present state of the art. To put it bluntly, a calculation that does not include a calculation of its predictive skill is not a legitimate scientific product.

If calculations of forecast skill were easy, there would be no reason to worry about accountability. Does one satisfy the principle of accountability if one collects a skill climatology by repeated experimentation? Popper quickly plugs that hole:

> "We have to demand that we must be able to find out, before we test the result of our predictions, whether or not the initial conditions are sufficiently precise. To put it more fully, we must be able to account in advance for any failure to predict an event with the desired degree of precision, by pointing out that our initial conditions are not precise enough, and by stating how precise they would have to be for the particular prediction task at hand" (The Open Universe, p.12).

My own hesitation to rely on skill climatology was rather more pragmatic. There are large fluctuations in forecast skill from day to day; the average forecast skill is often not a reliable indication of the skill of an individual forecast. In the same way, the effective forecast range or

prediction horizon, which currently is about five days on average, may be as little as two or as much as ten days from one event to the next. With that much variability, quoting an average value just does not fit the bill. Popper is a lot tougher: he insists that, since we are interested in predicting individual events (single realizations, in the language of statistics), we have to make advance calculations of the predictive skill for each event. In other words, if we wish to make a deterministic forecast, it will not do to make anything less than a deterministic skill forecast. Don't retreat into statistics when the going gets tough, that is what Popper says.

3. INFINITE REGRESS

It is evident that Popper allows himself little leeway when he thinks about the methodology of research. But once he has defined the core of the problem, he can permit himself to be considerate. He realizes that it is tougher to meet nonnegotiable demands than to formulate them. The production of skill forecasts is likely to be much more difficult than the production of the forecasts themselves. In fact, it is not at all clear whether skill forecasts of sufficient quality to meet the principle of accountability are possible, even in principle. The skill of a skill forecast is likely to be less than the skill of the original forecast, and the entire process demanded by accountability may not even converge. Popper starts his investigation into this problem by pointing out that accountability is concerned not only with the accuracy of the initial conditions, but also with the limitations inherent in the models we employ to describe the evolution of a physical system:

> "The method of science depends on our attempts to describe the world with simple theories. Theories that are complex may become untestable, even if they happen to be true (numerical modelers, please note - H.T.). Science may be described as the art of systematic over-simplification: the art of discerning what we may with advantage omit" (The Open Universe, p.44).

Since every prediction task that may be conceivably accountable operates with a simplifying model, a probably unresolvable paradox arises: if the model contains too many simplifications, it may not be possible to perform adequate calculations of forecast skill. Models have to be simple to be testable (another Popper favourite!), but simplicity and accountability may turn out to be mutually exclusive categories. In Popper's own words:

> "The question arises: how good does the model have to be in order to allow us to calculate the approximation required by accountability? Since the goodness of the model is its degree of approximation or precision, we are threatened by an infinite regress, and this threat will be very serious for systems which are complex (such as the atmosphere - H.T.). But the complexity of the system can also be assessed only if an approximate model is at hand; a consideration which again indicated that we are threatened by an infinite regress" (The Open Universe, p.50).

The matter of "infinite regress" was pointed out to me within days after launching my slogan on forecasts of forecast skill. What about the skill of the skill forecast, my friends would tease me, how do you forecast that? And what about having to repeat the process ad infinitum? Is is not true that these are progressively harder questions, and is there any chance the process will converge? Popper would have been sympathetic:

> "It is clear that we must, in advance, limit either the permissible number of demands for an improved model, or else the goodness, i.e. the degree of precision which we may require of the model. But the task of calculating either of these may lead us merely to a problem of accountability of a higher order. And with this, we may be well on our way to an infinite regress. For there is no reason whatever to believe that the higher-order problem is easier to solve than is the lower-order problem, or that a less good model is needed for its solution than for the solution of the lower-order problem. Nor is there any reason to believe that methods of approximation are always capable of improving results indefinitely" (The Open Universe,p.51).

With all of these problems piling up, there is, Popper says, "no reason whatever to believe that Newtonian mechanics is accountable". The results obtained by chaos theory bear him out. Determinism is an untenable proposition, even in the branch of physics in which it seemed most secure.

4. MARCHING ORDERS

Obviously, science is not an occupation for the weak at heart. If one is afraid of tackling problems that seem insoluble, one would do well to seek another profession. Fortunately, the human spirit is such that the hardest questions often attract the best minds. That bodes well for skill prediction research.

Because of the problem of "infinite regress", I do not think it advisable to hold skill forecasts accountable in the strict sense of the word. For the time being, statistical verifications of the skill of skill forecasts will have to suffice. The conceptual problems associated with higher-order accountability will have to wait. This puts skill forecasting midway between the presumed determinism of numerical weather prediction and the ad-hoc statistics of second-order skill. Skill forecasting deals with the statistical dynamics of the atmosphere; it is in that general area that one may anticipate the best prospects for advancement.

5. References

E.N. Lorenz, 1963: Deterministic, nonperiodic Flow. J. Atmos, Sci., 20, 130-141.

K.R. Popper, 1972: Of Clouds and Clocks. In: Objective knowledge, Oxford Univ. Press, London.

K.R. Popper, 1982: The Open Universe, Hutchinson, London.

H. Tennekes, A.P.M. Baede and J.D. Opsteegh, 1987: Forecasting Forecast Skill. In: Proc., ECMWF Workshop on Predictability, Reading, UK, April 1986.