Reply

ERIC J. PITCHER

Division of Atmospheric Science, University of Miami, Coral Gables, Fla. 33124

21 January 1977 and 21 March 1977

The true state of a physical system generally remains unknown to us and can only be estimated. Consequently, the estimate suffers from some uncertainty. The reader should keep in mind that by a single observation I mean the specification of *all* of the quantities in the system. It may turn out that an observation would lie very close to the unknown true state. Should they coincide, an error-free forecast in the case of a perfect model would be given by the conventional deterministic prediction—not a surprising result. Professor Gleeson considers such a possibility [his Eqs. (7) and (11)]. Although such an occurrence must be regarded as an unattainable ideal in the case of realistic physical systems, I have no quarrel with him on this point. However, to conclude on the basis of this fact that a stochastic dynamic forecast has a systematic error, in the case where the initial state is subject to observational error, is invalid. In particular, we shall presently see that for the simple example posed by Gleeson the stochastic dynamic forecast becomes the "best" one in the least mean-square-error sense, contrary to Gleeson's claim.

He asserts that the ensemble mean (to be defined later) is always a "very close approximation to the true value $\tilde{X}_0,\cdots$." But the whole point of introducing statistical considerations is to quantify the approximate nature of the initial state. The "closeness" of the ensemble mean to the truth is measured in terms of the variance (and in general covariance) information. His concept of ensemble mean, whose definition is rather crucial, is never made clear, and his appeal to the law of large numbers in order to support his contention is not entirely relevant. Of the several equivalent statements of this law the following will serve our purpose. In the event that a finite set of observations $(i=1,2,\cdots,n)$ of an unknown true state is made available, the sample mean of the n observations represents a combined observation with a mean-square error smaller than that of an individual observation by a factor of the order n^{-1}. (I am assuming, as does Gleeson, that the observations are free of any bias.) It is noteworthy that such a hypothetical situation differs from reality in the case of numerical weather prediction. In practice we only have available a *single* observation X_1 of the unknown true state.

Gleeson states that there can be no more than one true value in the real world at a specified time. Indeed, this *is* self-evident. But which of the possible infinity of states represents the truth? What is the likelihood of finding the true state $\tilde{X}_0$ a given "distance" from the observation X_1? These questions lead us to a probabilistic framework in specifying the initial state. We conclude that the true state $\tilde{X}_0$ is likely to be in the neighborhood of X_1 with a probability distribution which is related to the probability distribution for observational error. In the statistical model that I consider, and which has won traditional acceptance in treating a variety of problems (Papoulis, 1965), the true state becomes a random variable and we may consider an *ensemble* of *possible true states* distributed about X_1. The *ensemble mean* of these states is given by the observation X_1 and as such represents our best estimate of the true state in the least mean-square-error sense. In general the mean-square error of estimation must be obtained by a combination of theoretical considerations and past experience in observing the system. A stochastic dynamic model predicts the time evolution of this estimate such that it remains optimal in the least-squares sense. The stochastic dynamic forecast also provides a prediction of the covariance of the forecast ensemble which is the simplest measure of uncertainty in that estimate. The reader may consult Leith (1974) for further amplification.

Eqs. (15) and (16) given by Gleeson allegedly express the error in a deterministic and stochastic dynamic forecast, respectively, whenever $\tilde{L}$ is known in the absence of any uncertainty. He concludes that the stochastic forecast has a systematic error. We shall see that this apparent "bias" actually accounts for the optimal nature of a stochastic dynamic forecast, and that in fact such a forecast is an *unbiased* estimate of the truth whenever the latter is subject to observational error. Given $\tilde{X}_0$, $\tilde{U}$ and $\tilde{L}$, we have

$$\tilde{X}=[\tilde{U}-\alpha\tilde{L}^2]t+\tilde{X}_0, \qquad (1)$$

where, by definition, $\alpha=\beta/4\pi^2$. Because the initial state $\tilde{X}_0$ enters in a statistically linear sense, the full utility of the stochastic dynamic method cannot be demonstrated with this extremely simple model. In fact the only term in (1) which results in a nontrivial difference between the conventional deterministic and stochastic dynamic prediction of the ensemble mean of $\tilde{X}$ is $-\alpha\tilde{L}^2 t$. In the interest of brevity we treat the case where $\tilde{U}$ and $\tilde{X}_0$ may be specified exactly and where only the determination of $\tilde{L}$ is subject to error, i.e.,

$$\tilde{L}=L+\Delta,$$

where $\tilde{L}$ is an unknown true state, L a given observation and Δ a random variable such that

$$\left.\begin{aligned} E\Delta&=0\\ E\Delta^2\equiv E(\tilde{L}-L)^2&\equiv\sigma_2^2\end{aligned}\right\}, \tag{2}$$

where E denotes an ensemble-averaging operator, and by definition σ_2^2 is the variance of $\tilde{L}$ (now a random variable). In the light of (2) we have

$$E\tilde{L}\equiv\mu_2=L, \tag{3a}$$

$$E\tilde{L}^2=\mu_2^2+\sigma_2^2. \tag{3b}$$

With the substitution for $\tilde{L}$ Eq. (1) becomes

$$\tilde{X}=[\tilde{U}-\alpha(L^2+2L\Delta+\Delta^2)]t+\tilde{X}_0. \tag{4}$$

To obtain the deterministic prediction X^* we simply replace $\tilde{L}$ in (1) by its observed value and get

$$X^*=[\tilde{U}-\alpha L^2]t+\tilde{X}_0. \tag{5}$$

The error in the deterministic prediction is given by the difference between (4) and (5)

$$\tilde{X}-X^*=-\alpha(2L\Delta+\Delta^2)t. \tag{6}$$

If by chance $\Delta=0$, (6) shows that X^* would be a perfect forecast. The value of such a forecast is limited somewhat by the fact that we have no greater *a priori* confidence in it. For us to know its superiority beforehand we would have to know $\tilde{L}$ with absolute certainty. But we have presumably made use of all the finite and fallible data in arriving at our best guess for $\tilde{L}$ as well as its uncertainty given by σ_2^2. To find the expected error (i.e., bias) we take the ensemble average of (6) over all possible states, *noting that L has a fixed value:*

$$E(\tilde{X}-X^*)=-\alpha\sigma_2^2 t$$

from (2). A stochastic prediction of the ensemble mean is obtained by taking the ensemble average of (1). The reader should recall that, since $\tilde{L}$ is a random variable distributed about the observation L, $\tilde{X}$ is also a random variable whose mean μ is defined as $E\tilde{X}$. Gleeson and I differ on this point. The result is

$$\mu\equiv E\tilde{X}=[\tilde{U}-\alpha(\mu_2^2+\sigma_2^2)]t+\tilde{X}_0 \tag{7}$$

from (3b). Its expected error is then the ensemble average of the difference between (4) and (7), i.e.,

$$E(\tilde{X}-\mu)=-\alpha E(2L\Delta+\Delta^2-\sigma_2^2)t=0 \tag{8}$$

from (3a) and (2). As a direct consequence of the definition of the mean μ, it is easy to show that its mean-square error is a minimum. In fact the mean-square error of any forecast χ may be related to that of a stochastic dynamic forecast as follows;

$$\begin{aligned} E(\tilde{X}-\chi)^2&=E[(\tilde{X}-\mu)-(\chi-\mu)]^2\\ &=E(\tilde{X}-\mu)^2+(\chi-\mu)^2-2(\chi-\mu)E(\tilde{X}-\mu)\\ &=E(\tilde{X}-\mu)^2+(\chi-\mu)^2.\end{aligned} \tag{9}$$

The last line follows because of (8). Both terms on the right-hand side of (9) are clearly nonnegative. Thus $E(\tilde{X}-\mu)^2<E(\tilde{X}-\chi)^2$ unless $\chi=\mu$. Thus μ is an unbiased minimum variance estimate of the true state, as of course it should be. This does not mean that μ will be an exact forecast, only that its expected error when averaged over all possible states is zero. Parenthetically it should be pointed out that if the variance (uncertainty) of the true state L were known to be zero *a priori* (i.e., $\sigma_2^2=0$), then the stochastic dynamic forecast (7) would become identical with the conventional deterministic prediction (5).

Finally, the point which Gleeson makes in his second paragraph on the principle of least squares is an illusory one (at least to me). Estimation of the initial state in my study may be carried out quite independently of whether the prediction model is a linear or nonlinear one. (Dynamical constraints may of course be introduced.) The estimate of the initial state [Eq. (7) in Pitcher (1977)] satisfies the criteria given by Gleeson in that it is both a linear combination of the observations and a minimum variance estimate. He correctly points out that an ensemble of states though initially distributed according to a normal distribution will not remain normal in general. However, normality is preserved in the case of a nonlinear prediction model by closing the moment equations after the second moments (Pitcher, 1977), a procedure which appears to be satisfactory for short-range forecasting. Adopting the more restrictive closure of neglecting all reference to second moments would yield the conventional deterministic model.

Optimality has been defined here as minimization of mean-square error. The desirability of such a definition may not always be obvious, but this has *not* been the issue in the present exchange. *Any* definition of optimality will stipulate certain criteria with which to judge the various possible states of a system and to choose the "best" one. Rather, I have attempted to demonstrate here that, given one such definition, we may utilize the method of stochastic dynamic prediction in order that the stated error criterion be satisfied.

REFERENCES

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418.
Papoulis, A., 1965: *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 583 pp. (see p. 16 and Chap. 9).
Pitcher, E. J., 1977: Application of stochastic dynamic prediction to real data. *J. Atmos. Sci.*, **34**, 3–21.