NOTES AND CORRESPONDENCE

Comments on "Application of Stochastic Dynamic Prediction to Real Data"

THOMAS A. GLEESON

Department of Meteorology, The Florida State University, Tallahassee 32306

21 December 1976 and 14 February 1977

Dr. Pitcher and some other investigators [see Pitcher, (1977) and associated references] make the claim that a stochastic dynamic prediction is the best estimate of the true state in the least mean-square-error sense. No specific demonstration that this is indeed true has been offered, perhaps because it seems to be obviously valid.

However, there is reason to doubt the claim. According to Aitken (1944), the principle of least squares "can be justified under the assumption (i) that the measures are ***normally distributed*** and (ii) that the best value has ***maximum probability density*** . . . A more comprehensive derivation postulates that the best value is (i) a consistent or *unbiassed linear combination* of the observations and (ii) has *minimum variance*." In general, distributions of meteorological predictions from which stochastic dynamic predictions are obtained are not normal. Nor are the predictions linear combinations of initial state values. Therefore, a theoretical basis is lacking for the assertion that stochastic dynamic predictions are best estimates in the sense indicated.

Further study by use of a propagation-of-error procedure shows that a stochastic dynamic prediction has a systematic error if the combination is nonlinear. By contrast, a conventional deterministic prediction has no such induced error, which makes it an attractive alternative.

To demonstrate these results, an example with a single predicted variable is sufficient. The procedure to be followed below can be extended to examples of arbitrarily many variables with essentially the same outcome.

Let the governing differential equation for the variable X be represented generally by

$$dX/dt=G(X), \tag{1}$$

where t is time and $G(X)$ is a function of X; and the solution by

$$X=F(X_0,t), \tag{2}$$

where $F(X_0,t)$ is a function of time and the initial value X_0 at $t=0$. We shall regard this model as being "perfect," in that a single prediction X may be in error because of an initial-state error in the value of X_0, but not because of physical errors in the model. Then the true value $\tilde{X}$ at time t is given by

$$\tilde{X}=F(\tilde{X}_0,t), \tag{3}$$

where $\tilde{X}_0$ is the true initial value of X_0.

A stochastic dynamic prediction in this case is simply the mean μ of a large ensemble of hypothetical predictions. Therefore, from (2)

$$\mu\equiv E(X)=E[F(X_0,t)], \tag{4}$$

where E is used here to indicate an expected value of the term following it.

An expression for the prediction error $\mu-\tilde{X}$ at time t can be obtained as follows. First, expand $F(X_0,t)$ in a Taylor series to yield an equation for the error of a single prediction X. Thus,

$$X-\tilde{X}=\sum_{i=1}^{m}\frac{1}{i!}F^{(i)}(X_0-\tilde{X}_0)^i, \tag{5}$$

where (!) symbolizes a factorial, $F^{(i)}$ is the ith derivative of $F(X_0,t)$ with respect to X_0 taken at the point $(\tilde{X}_0,t)$, and $X_0-\tilde{X}_0$ is the initial-state error of X_0. Next, obtain expected values of all terms on both sides of (5). In the process it is important to recognize that there can be no more than one true value $\tilde{X}$ in the real world at a specified time, whether that value happens to be known to us or not. [In error analysis this is accepted as self-evident: see Margenau and Murphy (1949).] Therefore, the expected value $E(\tilde{X})$ must reduce to $\tilde{X}$ and consequently,

$$\mu-\tilde{X}=\sum_{i=1}^{m}\frac{1}{i!}F^{(i)}E[(X_0-\tilde{X}_0)^i]. \tag{6}$$

For simplicity, assume that systematic errors are not present in initial observations X_0 and that the latter are normally distributed. Then the mean μ_0 is a very close approximation to the true value $\tilde{X}_0$ according to the law of large numbers. Consider the rare but possible case where the two are equal:

$$\mu_0 \equiv E(X_0) = \tilde{X}_0. \tag{7}$$

If the stochastic dynamic prediction μ is just a linear combination of the observations X_0, then $m=1$ in (5) and (6). Because of (7), Eq. (6) reduces to

$$\mu = \tilde{X}, \tag{8}$$

so there is no error in this prediction. Furthermore, due to the reproductive property of the normal distribution, individual predictions X are also normally distributed. The best value μ has maximum probability density, consistent with the principle of least squares.

However, in the more complex case where (6) is nonlinear (i.e., $m>1$), the right side of the equation does not vanish, except in very special circumstances, so μ remains in error. The error is not an outgrowth of systematic errors in the initial observations nor in the physical model because these were eliminated by previous assumptions. It is, therefore, a result of the method of prediction.

A conventional deterministic prediction X^* avoids this type of error. In the present example, X^* can be represented by

$$X^* = F(\mu_0, t). \tag{9}$$

Its prediction error is given by the Taylor series expansion

$$X^* - \tilde{X} = \sum_{i=1}^{m} (1/i!) F^{(i)} (\mu_0 - \tilde{X}_0)^i. \tag{10}$$

Because of (7), (10) reduces to

$$X^* = \tilde{X}, \tag{11}$$

an error-free result for both the linear and nonlinear cases.

A more specific example might be useful at this point to illustrate the contrast between stochastic dynamic and conventional deterministic predictions. Consider the predicted displacement of a Rossby wave according to the differential equation

$$\frac{dX}{dt} = U - \frac{\beta L^2}{4\pi^2}, \tag{12}$$

where X can be defined as a predicted location of a trough line along an east-west coordinate, U and L are constant values of the west-wind speed and wavelength, respectively, and β is the rate of change of the Coriolis parameter with latitude at a specified latitude. As before, suppose there are no systematic errors in the physics of the model. Then it is readily verified that error equations for X^* and μ can be written as

$$X^* - \tilde{X} = \mu_0 - \tilde{X}_0 + t[\mu_1 - \tilde{U} - (\beta \tilde{L}/2\pi^2)(\mu_2 - \tilde{L}) - (\beta/4\pi^2)(\mu_2 - \tilde{L})^2], \tag{13}$$

$$\mu - \tilde{X} = X^* - \tilde{X} - (\beta t/4\pi^2) E[(L - \tilde{L})^2], \tag{14}$$

where $\tilde{U}$ and $\tilde{L}$ are true values and μ_1 and μ_2 are means of initial values of U and L, respectively. It is seen that errors of both predictions are functions of errors of initial means, but that μ has an additional error represented by the last term in (14).

In the limiting case where errors of initial means are zero, (13) and (14) simplify to

$$X^* - \tilde{X} = 0, \tag{15}$$

$$\mu - \tilde{X} = -(\beta t/4\pi^2) E[(L - \tilde{L})^2], \tag{16}$$

and μ still has an error. It is negative, systematically tending to make μ an underestimate of $\tilde{X}$. Of course, in the more general case represented by (14) this tendency would be counteracted or enhanced by errors in initial means.

Although $\tilde{L}$ is inherently unknown it is still possible to say something about the size of the error in (16). In the identity

$$E[(L - \tilde{L})^2] = \sigma_2^2 + (\mu_2 - \tilde{L})^2, \tag{17}$$

the only term not a function of $\tilde{L}$, and therefore in principle determinable, is σ_2^2, the variance of L. A computed value of the quantity $\beta t \sigma_2^2/(4\pi^2)$ then provides a lower bound to the absolute value of the error.

Direct comparison tests can be made for more realistic and complicated prediction models. As before, Taylor series expansions of solutions (implicit and explicit) are useful for the purpose, although they are not necessary in every instance. If the model is nonlinear, it is found that the stochastic dynamic prediction has one or more errors, systematic in some fashion, that are not present in the conventional deterministic prediction. This is a qualitative conclusion that does not depend on specific knowledge of true values of initial states. The reader can check such comparisons for himself.

It should be noted that none of the foregoing comments is intended to deny that stochastic dynamic predictions are interesting in other respects and are useful in studies of prediction variability. For example, the mean-square deviation of individual predictions X from the ensemble mean μ is the minimum possible value of the unknown mean square error of those predictions. This is always the case, independent of how well μ itself might estimate the true value.

REFERENCES

Aitken, A. C., 1944: *Statistical Mathematics*. London, Oliver and Boyd, 153 pp. (see p. 107).

Margenau, H., and G. Murphy, 1949: *The Mathematics of Physics and Chemistry*. Van Nostrand, 581 pp. (see pp. 487–493).

Pitcher, E. J., 1977: Application of stochastic dynamic prediction to real data. *J. Atmos. Sci.*, **34**, 3–21.