THE UNIVERSITY OF MICHIGAN

COLLEGE OF ENGINEERING
Department of Meteorology and Oceanography

Technical Report

THE ROLE OF INITIAL UNCERTAINTIES IN PREDICTIONS

Edward S Epstein

ORA Project 07584

supported by:

NATIONAL SCIENCE FOUNDATION
GRANT NO. GP-4385
WASHINGTON, D.C.

administered through

OFFICE OF RESEARCH ADMINISTRATION ANN ARBOR

August 1968

1. Introduction

While it has long been recognized that the initial atmospheric conditions upon which meteorological forecasts are based are subject to considerable error, little if any explicit use of this fact has been made. Operational forecasting consists of applying deterministic hydrodynamic equations to a single "best" initial condition and producing a single forecast value for each parameter. While there are always ready admissions that a "perfect" or errorless forecast is not possible, no effort is made under ordinary circumstances to ascertain the manner in which the errors in the initial conditions propagate through the forecast procedure.

Recently, Gleeson (1966, 1967, 1968) has indicated how we can begin to study the influence of initial error on the error in the prediction. He has examined a few simple prognostic equations, assumed there were errors in the parameters and initial conditions required for the forecast, and examined how the uncertainty of the prediction varies with time.

One of the questions which has been entirely ignored by the forecasters, and one which Gleeson, too, chose to neglect, is whether or not one gets the "best" forecast by applying the deterministic equations to the "best" values of the initial conditions and relevant parameters. For the purpose of this discussion we may define "best" as the mean of all possible values consistent with the observations. As Gleeson (1968) clearly points out, one cannot know a uniquely valid starting point for each forecast. There is instead an ensemble of possible starting points, but the identification of the one and only one of these which represents the "true" atmosphere is not possible. The deterministic forecast equations apply equally to each member of the ensemble. The question we will pose is whether applying such equations to any single member is the proper forecast procedure under realistic conditions.

Gleeson sought to examine how the spread of the points that make up the ensemble changes with time. He points out, for example, that the variance of any parameter increases or decreases in time according to whether the correlation between the current value of the parameter and its time derivative is positive or negative. By examining several simple mathematical-physical systems, Gleeson (1968) was able to show that a wide variety of possibilities existed, and that therefore the more complex meteorological problems would require considerable study before the time behavior of errors introduced by uncertainties in the initial conditions and parameters will be understood. The significance of this is that it gets to the heart of the problem of the predictability of the atmosphere. One is able to generate useful predictions, even given perfect mathematical-physical relationships, only to the extent that errors introduced by uncertainties in initial conditions, parameters, and numerical approximations remain within acceptable bounds.

In the present study I have re-examined the same four simple examples used by Gleeson (1968). A method of greater generality is used making it possible to evaluate critically some of the assumptions which Gleeson found convenient. In addition to studying the spread of points, however, I have specifically examined the behavior of the mean of the ensemble, as opposed the behavior of the individual member of the ensemble whose location may happen to coincide with the mean. It turns out that there are, in certain cases, distinct differences between these two. This means that in general the "best" forecast will not be obtained by applying deterministic equations to the "best" initial conditions.

2. First example: constant rate of change.

The first example to be treated is represented by the simple differential equation

$$\frac{dq}{dt} = C = \text{constant} \qquad (1)$$

which has the solution

$$q = Q + Ct \qquad (2)$$

Both Q and C are subject to uncertainty. It is not unreasonable to expect the errors in the estimation of both of these parameters to be Gaussian. We may follow Gleeson in assuming that Q and C are normally distributed with means (unknown) of $\overline{Q}$ and $\overline{C}$ and standard deviations σ_Q and σ_C.

While Gleeson assumed Q and C were independent, let us assume their joint distribution is bivariate normal with covariance $\rho\sigma_Q\sigma_C$.

The mean value of q at any time t may be evaluated by taking the expected value of (2):

$$E(q) = \bar{q} = E(Q + Ct) = \bar{Q} + \bar{C}t$$

The variance of q is

$$\begin{aligned}\text{var}(q) &= E(q^2) - [E(q)]^2 \\ &= E(Q^2 + 2QCt + C^2t^2) - (\bar{Q} + \bar{C}t)^2 \\ &= \text{var}\ (Q) + t^2\ \text{var}(C) + 2t\ \text{cov}\ (Q,C)\end{aligned}$$

or

$$\sigma_q^2 = \sigma_Q^2 + t^2\sigma_C^2 + 2t\rho\sigma_Q\sigma_C$$

Taking $\rho=0$ gives the situation Gleeson considered, and the same result, i.e. that $d\sigma_q^2/dt=2t\sigma_C^2$. In that case the variance (or the standard deviation) of q increases with time at a rate governed solely by the uncertainty in C. In the more general case, however, one can only place bounds on σ_q corresponding to $-1< \rho<1$. These bounds are illustrated in Figure 1. Note that for $\rho<0$ and $t< -\rho\sigma_Q\sigma_C$, the rate of change of σ_q will be negative. Also, for $\rho<0$, σ_q remains smaller than σ_Q for $t<-2\rho\sigma_Q\sigma_C$.

3. Nocturnal Cooling

Gleeson's second example concerns the equation

$$T = T_o - At^{1/2} \tag{3}$$

which has on occasion been used to represent radiational cooling at night. A is an empirical parameter and T_o is the initial temperature; both are subject to uncertainty. If one makes the same assumptions about their joint distribution as in the previous example, namely that they are bivariate normal with parameters $\overline{T}_o$, σ_{T_o}, $\overline{A}$, σ_A, ρ, then

$$E(T) = \overline{T} = \overline{T}_o - \overline{A}t^{1/2}$$

$$\mathrm{var}(T) = \sigma_{T_o}^{2} + t\sigma_A^{2} - 2t^{1/2}\sigma_{T_o}\sigma_A\rho$$

Gleeson's result, namely that

$$\frac{\partial\sigma_T^{2}}{\partial t} = 2E[\ (T - \overline{T})\ (\frac{\partial T}{\partial t} - \frac{\partial\overline{T}}{\partial t})\ ] = \sigma_A^{2}$$

is confirmed for the case $\rho=0$. This special case, plus the bounds on σ_T^2 in the more general case, are shown in Figure 2.

The same comments apply here as did in the first example. The variance may decrease initially if the covariance between the two parameters is of the proper sign. In this case, if $\rho>o$, then σ_T is less than its initial value for $t<(2\rho\sigma_{T_o}/\sigma_A)^2$. Also the equation applied to the initial mean value gives the same results as the expected value of the ensemble.

The similarity of these two first cases lies in the linearity of (2) and (3) as concerns the statistically variable components. The next two cases will be non-linear and therefore one may expect different results. Note also, in these first two cases, that the assumptions of bivariate normality were excessively stringent. Any joint distributions with the same first and second moments would have given identical results.

4. Frictional Retardation.

Let us now examine the simplest type of friction, expressed by the differential equation

$$\frac{dv}{dt} = -kv$$

with the solution

$$v = v_o \exp(-kt). \tag{4}$$

Here v is a velocity and k is a coefficient of friction. The two parameters subject to error, in this case, are v_o, the initial velocity, and k. For this example Gleeson noted that when t=0,

$$\frac{1}{2}\frac{d\sigma_v^{\ 2}}{dt} = E[v - \bar{v}) \left(\frac{dv}{dt} - \frac{d\bar{v}}{dt}\right)] = - E[k(v_o - \bar{v}_o)^2]$$

which is certainly less than zero. He concluded that as time passes the uncertainty of v would continue to decrease, and at the same time the mean value of v would approach ever closer to zero. In time, Gleeson concluded, a very precise forecast of zero velocity is possible.

Let us now examine the problem more formally, assuming here that k and v_o are independently bivariate normal. Then

$$\bar{v} = E(v) = E(v_o\, e^{-kt}) = E(v_o)\ E(e^{-kt})$$

$$= \bar{v}_o\ E(e^{-kt})$$

$$\sigma_v^{\ 2} = E(v^2) - [E(v)]^2$$

$$= E(v_o^{\ 2})\ E\ (e^{-2kt}) - \bar{v}^2$$

Since k is normal with mean $\bar{k}$ and variance $\sigma_k^{\ 2}$,

TABLE 1

Calculated values of $\bar{v}$ and σ_v for $\bar{v}_o=10$, $\sigma_{v_o}=1$, $\bar{k}=10^{-4}\mathrm{sec}^{-1}$, $\sigma_k=3\times10^{-5}\mathrm{sec}^{-1}$

t	$\bar{v}$				σ_v			
(sec)	$\sigma_k=\infty$	Normal	Gamma	Log-Normal	$\sigma_k=\infty$	Normal	Gamma	Log-Normal
1600	8.521	8.531	8.531	8.536	.8521	.9473	.9448	.9596
3200	7.262	7.295	7.295	7.302	.7262	1.015	1.001	1.016
6400	5.273	5.371	5.367	5.378	.5273	1.176	1.124	1.125
12800	2.780	2.993	2.978	2.988	.2780	1.236	1.106	1.072
25600	0.773	1.038	0.999	0.997	.0773	.941	0.7070	.6456

Note: The columns headed $\sigma_k=\infty$, Normal and Gamma refer to the distributions of K and the calculations are based on formulae given in the text. The column headed Log-Normal is based on an experimental random sample of size 1000.

$$E(e^{-kt}) = \int_{-\infty}^{\infty} \frac{e^{-kt}}{(2\pi)^{1/2}\sigma_k} \exp[-\frac{(k-\bar{k})^2}{2\sigma_k^2}]\,dk$$

$$= \exp(-\bar{k}t + \tfrac{1}{2}t^2\sigma_k^2) \qquad (5)$$

Therefore

$$\bar{v} = \bar{v}_o\, e^{-\bar{k}t}\, e^{\frac{1}{2}t^2\sigma_k^2} \qquad (6)$$

$$\sigma_v^2 = (\bar{v}_o^2 + \sigma_{v_o}^2)\, e^{-2\bar{k}t} e^{2t^2\sigma_k^2} - \bar{v}_o^2 e^{-2\bar{k}t} e^{t^2\sigma_k^2} \qquad (7)$$

By differentation, we find that

$$\frac{d\sigma_v^2}{dt} = e^{-2\bar{k}t}\{-2\bar{k}[(\bar{v}_o^2 + \sigma_{v_o}^2)e^{2t^2\sigma_k^2} - \bar{v}_o^2 e^{t^2\sigma_k^2}]$$

$$+ 2t\sigma_k^2[2(\bar{v}_o^2 + \sigma_v^2)e^{2t^2\sigma_k^2} - \bar{v}_o^2 e^{t^2\sigma_k^2}]\}$$

Only at t=0 do we confirm Gleeson's result that $\frac{d}{dt}\sigma_v^2 = 2k\sigma_{v_o}^2$. The time rate of change of variance is negative initially, but may change sign several times as t gets large.

Eq (6), which describes the behavior of the ensemble mean, is plotted as the uppermost curve in Figure 3. The bottom curve is the estimate of $\bar{v}$ obtained by applying (4) directly to the initial mean conditions. For these plots parameter values of $\bar{v}_o$ =10 m/sec, σ_{v_o} = 1 m/sec, $\bar{k} = 10^{-4}$ sec^{-1}, σ_k = 3 x 10^{-5}sec^{-1} have been used. Not only is the ensemble mean always greater than the value obtained deterministically, but in this particular case, after about 10^5 sec $\bar{v}$ increases very rapidly toward infinity.

This peculiar behavior is a manifestation of a peculiar condition that we imposed on the distribution of k. We assumed it had a normal distribution, implying that for some members of the ensemble $k<0$. On physical grounds k can never be negative. In the example shown, only .04% of the members of the ensemble would have $k<0$, but in these few cases v increases exponentially with time, and after a sufficient period they dominate the mean.

We conclude, then, that it is improper to use a normal distribution to describe the uncertainty in k, since k must always be positive. (Note that the same argument could be applied to the distribution of A in the previous example. In that case, however, only the first two moments of the distribution of A were pertinent. In the present example all the moments of k enter into the calculation of v.) Let us instead choose a gamma distribution for k (again with $E(k)=\bar{k}$ and $\text{var}(k)=\sigma_k^2$). The functional form of the gamma distribution is

$$g(k) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} k^{\alpha+1} e^{-k\beta}, \quad k>0$$

where $\alpha=\bar{k}^2/\sigma_k^2$ and $\beta=\sigma_k^2/\bar{k}$. For $k<0$, $g(k)=0$.

Another possible form for the distribution of k, $k>0$, is the log-normal distribution

$$h(k) = (2\pi)^{-1/2}(sk)^{-1} \exp[-(\ell n k - m)^2/2s^2] \quad (8)$$

where $m=\frac{1}{2}\ell n \left(\frac{\bar{k}^4}{\bar{k}^2+\sigma_k^2}\right)$ and $s^2 = \ell n \left(1 + \frac{\sigma_k^2}{\bar{k}^2}\right)$

to maintain the same first two moments of the distribution. Figure 4 is a plot of the normal, gamma, and log-normal distributions for $\bar{k} = 1$ and $\sigma_k^2=0.1$ (or, with a simple scale change for $\bar{k} = 10^{-4}$ and $\sigma_k^2=10^{-9}$).

For the present we choose the gamma distribution because its moment generating function is known in closed form. Specifically

$$E(e^{-kt}) = (1 + \beta t)^{-\alpha}$$

Substituting (9) in (6) and (7) gives

$$\bar{v} = \bar{v}_o \left[1 + \frac{\sigma_k^2 t}{\bar{k}}\right]^{-\bar{k}^2/\sigma_k^2} \tag{10}$$

$$\sigma_v^2 = (\bar{v}_o^2 + \sigma_{v_o}^2) \left[1 + \frac{2\sigma_k^2 t}{\bar{k}}\right]^{-k^2/\sigma_k^2} - \bar{v}_o^2 \left[1 + \frac{\sigma_k^2 t}{\bar{k}}\right]^{-2\bar{k}^2/\sigma_k^2} \tag{11}$$

Again, only at $t=0$ do we find $d\sigma_v^2/dt = 2k\sigma_{v_o}^2$. But we now obtain a more reasonable behavior for $\bar{v}$. Eq. (10) is plotted as the middle curve in Figure 3. While the limit of $\bar{v}$ as t gets large is now zero we still observe that the direct application of the deterministic eq. (4) will always underestimate the ensemble mean.

Additional effects of the statistical interactions between the uncertainties in v_o and in k are shown in Figure 5. The uppermost curve, based on (7), is a plot of the standard deviation of v when k is assumed to be normal. Since this assumption has already been discredited, the curve is shown for comparison only. The middle curve is based on (11); it shows σ_v when k has a gamma distribution. (Parameters have the same values as for Figure 3.) The bottom curve in Figure 5 is a plot of $\sigma_{v_o} e^{-\bar{k}t}$, the limiting case of both (7) and (11) as σ_k approaches 0, i.e. if k is known with certainty to be equal to $\bar{k}$.

Note in particular the middle curve, i.e. for k having a gamma distribution. The variance decreases for small t, as suggested by Gleeson's result; but this decrease is very small and is temporary. In the example shown, after t∼

25,000 sec the standard deviation is as large as it was originally. A maximum of σ_v is reached near 10^5 sec, and only after this time does σ_v decrease toward zero.

It is possible to explain this behavior in terms of the changes which must occur in the distribution of v, if σ_k is sufficiently large. At time t = 0, the distribution of v is normal and the initial decrease in variance is due to the general shrinkage of all members of the ensemble toward smaller v. However, as this proceeds, the cases of small k (especially if linked with a large v_o) will lag behind, and soon will be even further from the modal value of v than they were initially; ensemble members with large k will emphasize the left hand tail of the distribution. This will tend to increase the variance. Finally, as t continues to increase the bulk of the ensemble members are near v=0 having caught up with those on the left hand tail, and are no longer decreasing, while the ones on the right-hand tail of the distribution are cases of small k and approach the origin, slowly but consistently. Thus the variance eventually decreases toward zero. (Note that this description fits the case of the normal distribution of k until large t when the very few cases of $k<0$ make their presence felt.)

To evaluate and test the ideas behind the sequence described in the preceding paragraph, the following experiment was carried out. Two sets of pairs of random numbers (1000 pairs in each set) were generated such that one member of each pair was normal (v_o: mean 10, standard error 1) and one was independently log-normal(k: $m = -\frac{1}{2}\ln(1.09\times10^{-8})$, $s^2 = \ln(1.09)$; cf. eq (8)). Eq. (4) was applied separately to each pair of values, i.e. to each member of the ensemble, and means and standard deviations for the entire ensemble were calculated for several values of t. These values are the points plotted in Figures 3 and 5. In addition, it was possible to examine the frequency distributions of v at different values of t. Some of these are plotted in Figure 6.[1] Just as the gamma

[1] The same end could have been accomplished by solving numerically Gleeson's continuity equation for probability density. If $\Psi(k,v,t)$ is the time dependent joint probability density of k and v, the continuity equation here takes the form $\partial\Psi/\partial t = k(\Psi + v\partial\Psi/\partial v)$.

distribution was used earlier because it provided analytic simplicity, the log-normal distribution was used in this experiment because of computational simplicity. Comparison of the results allows some added insight into the significance of the shape of the distributions (Figure 4).

The points plotted in Figures 3 and 5 show that the behavior of the gamma and log-normal distributions are remarkably similar. This is emphasized by the tabulation of these results in Table 1.

The distributions plotted in Figure 6 substantiate the argument given earlier. At t=1600 sec the distribution of v is still relatively normal but σ_v is just beginning to increase. The cases on the tails of the distribution are still mostly cases of small and large v_o, but the influence of the variability in k is beginning to be felt. At t=6400 sec. the standard deviation is near its maximum, but the distribution is still relatively symmetrical. The total range of values of v, 6.00 m/sec at t=1600 sec is now 6.71 m/sec. Clearly cases on both tails must be there primarily because of large or small values of k.

At t=25600 sec. the distribution is very asymmetrical. The values of v_o are of relatively little weight in determining where an individual case may lie. Note that the extreme values of v_o are in the approximate ratio 13/7, or near 2, while at t=25600 sec the extreme values of v are .028 and 4.11, a ratio near 150. Thus for large t the distribution of k becomes increasingly important.

In the foregoing the assumption was made, for simplicity, that k and v_o were independent. One can relax that assumption by selecting a particular form of dependence. It is relatively simple to examine the case where the marginal distribution of k is gamma, as previously, and the distribution of v_o, given k, is normal with mean ak+b and variance s_{v_o}. One then finds that

$$\bar{v}_o = E(v_o) = b + a\,\bar{k}$$

$$\sigma_{v_o}^{\;2} = \mathrm{var}(v_o) = s_{v_o}^{\;2} + a^2\ \mathrm{var}(k)$$

$$\mathrm{cov}(k, v_o) = a^2\ \mathrm{var}\ (k) = \rho\sigma_k\sigma_{v_o}$$

For a suitable comparison, we have calculated v and σ_v, for various values of ρ, choosing in each case values for a, b, and s_v such that $v_o = 10$ m/sec and $\sigma_{v_o} = 1$ m/sec as before. The differences in E(v) from those values shown as the middle curve in Figure 3 amount to variations of only about $\pm$ 5% as ρ varies between -1 and +1, for $<10^5$ sec. The lack of independence has a much more dramatic influence on σ_v, as illustrated in Figure 7. As one could anticipate, a negative correlation between v_o and k implies that large initial velocities decrease slowly while slower ones decrease rapidly, implying an increasing standard deviation at small t. For large positive correlation, the large initial values of v_o decrease rapidly, catch up with the ones that are small initially but decreasing slowly, and at this time a pronounced minimum of the variance occurs.

The important lesson to be learned from the above discussion is that the statistical interaction between two parameters can play a very significant role. In the present case we find that although the velocity of every member of the ensemble is known to decrease (with an average time constant of about 3 hours), 3 hours later the mean velocity has indeed decreased (by an amount we might tend, if not careful, to overestimate) but the absolute error in our specification of that smaller mean wind may be greater than the error in the original values of v_o. This result is clearly dependent on our choice of distributions and parameters, but there is the suggestion that if the distribution is reasonable then the results will be reliable. Unreasonable distributions may lead to wholly falacious results. The degree of dependence among the various parameter can play a very significant role in the variation of variance with time. Finally, since the distributions of dependent variables generally will change, the special case of Gleeson's (1966, 1968) continuity equation for probability, developed on the assumption of normality, will have little general applicability. The more general statement of that principal which Gleeson gives is of course still valid.

5. Cyclical Changes.

The last example given by Gleeson is represented by

$$q = Q + B \cos(\nu t - B) \tag{12}$$

where Q is a time-mean and B the amplitude of some parameter q. The frequency ν, and the phase angle β, as well as Q and B, are all subject to error. This type of expression has particular significance in meteorology since it is so often found convenient to express time series as sums of terms of various amplitudes and phases. Gleeson makes the point that the standard deviation of q will show a behavior that combines cyclical increases and decreases with a superimposed gradual increase. He states that the "true" value of q will be entirely cyclical in its behavior, but he mistakenly identifies this "true" value with $\overline{q}$. He misses the point that the ensemble mean (indeed $\overline{q}$) will have a somewhat different behavior.

To begin, let us assume (with some trepidation now in view of the preceding results) that the several parameters have independent normal distributions with means Q, B, ν, β, and standard deviations $\sigma_Q, \sigma_B, \sigma_\nu, \sigma_\beta$. Then by using (5) plus the identities

$$\sin\theta = -\tfrac{1}{2} i \, (e^{i\theta} - e^{-i\theta})$$

$$\cos\theta = \tfrac{1}{2} (e^{i\theta} + e^{-i\theta})$$

where $i = \sqrt{-1}$, we find that if x is normal with mean $\overline{x}$ and standard deviation σ,

$$E(\sin x) = e^{-\frac{1}{2}\sigma^2} \sin \overline{x}$$

$$E(\cos x) = e^{-\frac{1}{2}\sigma^2} \cos \overline{x}$$

$$E(\sin^2 x) = \tfrac{1}{2}(1 - e^{-2\sigma^2} \cos 2\overline{x})$$

$$E(\cos^2 x) = \tfrac{1}{2}(1 + e^{-2\sigma^2} \cos 2\overline{x})$$

$$E(\cos x \sin x) = e^{-2\sigma^2} \cos\overline{x} \sin\overline{x}$$

It then follows that

$$\overline{q} = \overline{Q} + Be^{-\frac{1}{2}(\sigma_\beta^2 + t^2\sigma_\nu^2)} \cos (\overline{\nu}t-\overline{\beta}) \qquad (13)$$

$$\sigma_q^2=\sigma_Q^2 + \tfrac{1}{2}(\overline{B}^2 + \sigma_B^2)[1+e^{-2(\sigma_\beta^2+t^2\sigma_\nu^2)} \cos 2 (\overline{\nu}t-\overline{\beta})]$$

$$-\tfrac{1}{2}\overline{B}^2 e^{-(\sigma_\beta^2+t^2\sigma_\nu^2)}[1 + \cos 2(\overline{\nu}t - \overline{\beta})] \qquad (14)$$

Note in particular that the amplitude of the ensemble mean decreases with time. In the limit, as t gets large, the best estimate for $\overline{q}$ is $\overline{Q}$. This behavior is entirely due to the uncertainty in ν, the frequency. The uncertainty of the phase angle, β, causes the amplitude of the mean of the ensemble to be less than the mean amplitude ($\overline{B}$) by a factor $e^{-\frac{1}{2}\sigma_\beta^2}$, but this effect does not change with time. In other words, even at t=0, the average departure of q from its time mean, $\overline{q} - \overline{Q}$, is less than the average of the amplitudes.

The variance of the ensemble values of q is an involved function whose specific behavior depends on the particular parameters chosen. So long as $\sigma_\nu > 0$, the variance will approach

$\sigma_Q^2 + \frac{1}{2}(\bar{B}^2 + \sigma_B^2)$ as $t \to \infty$. If $\sigma_\nu = 0$, the variance will be cyclical for all time, with a mean value of

$$\sigma_Q^2 + \tfrac{1}{2}(\bar{B}^2 + \sigma_B^2) - \tfrac{1}{2}\bar{B}^2 \, e^{-\sigma_B^2}$$

and an amplitude $\frac{1}{2}(\bar{B}^2 + \sigma_B^2) \quad e^{-2\sigma_B^2} - \frac{1}{2}\bar{B}^2 \; e^{-\sigma_B^2}$

The effect of uncertainty in ν is thus to increases the time - average uncertainty and to damp out the cyclical fluctuations in σ_q^2.

We should note here that the assumptions of a normal distribution for a phase angle is very artificial, since angles are equivalent modulus 2π. This is especially problematic since the development of (12) and (13) requires all the moments of the distribution of β be known. The same is true of ν. On the other hand, only the first two moments of B and Q enter into the calcualtion of $\bar{q}$ and σ_q. Thus we should exercise the greatest care in specifying the distributions of β and ν. In the present instant no a priori objection to specifying a normal distribution for ν is apparent. Phase angles such as β would seem to require special treatment.

As an alternative to the assumption of normality the Von Mises distribution (Downs, 1964)

$$f(\alpha) = \frac{e^{K \cos(\alpha - \hat{\alpha})}}{2\pi \, I_o(K)} \qquad 0 \leqslant \alpha, \ \hat{\alpha} \leqslant 2\pi, \ K \geqslant 0.$$

might be used for β. Here K and $\hat{\alpha}$ are parameters of the distribution and $I_o(K)$ is the Bessel function of order 0. K is called the "invariance"; as K gets larger the observations tend to cluster ever more strongly about the single modal point $\hat{\alpha}$. Indeed, for large K, $(\alpha - \hat{\alpha})$ remains small so that one may write

$$\cos(\alpha - \hat{\alpha}) = 1 - \tfrac{1}{2}(\alpha - \hat{\alpha})^2$$

and $$f(\alpha) = \frac{e^K}{2\pi\, I_o(K)} \exp[-\frac{K}{2}(\alpha - \hat{\alpha})^2]$$

Since this is the functional form of the error function, the Von Mises distribution approaches a normal distribution with mean $\hat{\alpha}$ and variance $1/K$ for large K. For K=0, $f(\alpha)$ is the uniform distribution, $0 \le \alpha < 2\pi$.

This distribution has the properties that

$$E(\cos\alpha) = \delta \cos\hat{\alpha}$$

$$E(\sin\alpha) = \delta \sin\hat{\alpha}$$

$$E(\cos^2\alpha) = \frac{\delta}{K} + (1 - \frac{2\delta}{K}) \cos^2\hat{\alpha}$$

$$E(\sin^2\alpha) = (1 - \frac{2\delta}{K}) \sin^2\hat{\alpha} + \frac{\delta}{K}$$

$$E(\sin\alpha \cos\alpha) = (1 - \frac{2\delta}{K}) \sin\hat{\alpha} \cos\hat{\alpha}$$

where $\delta = I_1(K)/I_o(K)$.

If we write $x = \cos\alpha$, $y = \sin\alpha$, then δ is the distance from the origin to the point $(E(x)\ E(y))$.

Returning to the examination of (12), we will now assume that β has the Von Mises distribution with parameters β and K_β, while the other parameters are independently normal. Then if $\delta = I_1(K_\beta)/I_o(K_\beta)$,

$$\bar{q} = \bar{Q} + \bar{B}\,\delta\, e^{-\frac{1}{2}\sigma_\nu^2 t^2} \cos(\bar{\nu}t - \hat{\beta}) \tag{15}$$

$$\sigma_q^2 = \sigma_Q^2 + \tfrac{1}{2}(\bar{B}^2 + \sigma_B^2)\left[1 + e^{-2\sigma_\nu^2 t^2}\left(1 - \frac{2\delta}{K_\beta}\right)\cos 2(\bar{\nu}t - \hat{\beta})\right]$$

$$- \bar{B}^2\,\delta^2\, e^{-\sigma_\nu^2 t^2} \cos^2(\bar{\nu}t - \hat{\beta}) \tag{16}$$

There is great similarity between these equations and (13) and (14). The δ in (15) plays the role of $e^{-\frac{1}{2}\sigma_\beta^2}$ in (13), while the factor $(1 - 2\delta/K_\beta)$ in (16) plays the role of $e^{-2\sigma_\beta^2}$ in (12). For purposes of comparison we have chosen to select the parameters for the two distributions such that $\hat{\beta} = \bar{\beta}$ and $\delta = e^{-\frac{1}{2}\sigma_\beta^2}$. A comparison of the values of the pertinent parameters as given in Table 2. Figure 8 compares the Von Mises and normal distributions under these conditions.

Except for small k there is little to distinguish one distribution from the other. Note that for large K, δ may be approximated by $1 - (2K)^{-1} \cong 1 - \frac{1}{2}\sigma_\beta^{-2} \cong e^{-\frac{1}{2}\sigma_\beta^2}$. Similarly, also for large K, $2\delta/K \cong 2/K$, and $1 - 2\delta/K \cong e^{-2\sigma_\beta^2}$.

Figures 9, 10, and 11 show plots based on (14), (15) and (16) for several values of the appropriate parameters. In the calculations for these figures the mean amplitude, $\bar{B}$, was taken as unity Consequently the the results are essentially normalized with respect to $\bar{B}$. In Figure 9, for example, $(\bar{q} - \bar{Q})/\bar{B}$ can never exceed 1.0. Only in the very special case when $K_B = \infty$ (i.e. $\sigma_\beta = 0$), and for t=0, can that value be achieved. (For simplicity we have taken $\beta = 0$. The effect of varying σ_ν is to alter the rate at which the cyclical fluctuations in $\bar{q}$ approach zero. The value of σ_B has no influence on $\bar{q}$. When the uncertainty in β is large (i.e. for small k_β) the fluctuations in $\bar{q}$ are reduced.

The influences of the several parameters on the ensemble standard deviation of q is more involved. It is easier to examine the quantity $\Sigma = (\sigma_q^2 - \sigma_Q^2)^{\frac{1}{2}}/\bar{B}$, the standard deviation of the departure of q from its time-average value, normalized to unit mean amplitude. The limit of this quantity, at

TABLE 2

Comparison of parameters related to Von Mises and normal distributions

$\underline{K}$	δ	σ_β^2	σ_β	$1-2\delta/K$	$e^{-\sigma_\beta^2}$
0.4	.1961	3.4582	1.8596	.0195	.0010
1.0	.4464	1.6138	1.2704	.1136	.0397
2.0	.6978	.8484	.8484	.3022	.2371
4.0	.8635	.2936	.5418	.5682	.5559
8.0	.9352	.1340	.3661	.7650	.7649

$t \to \infty$, is $[\frac{1}{2} + \frac{1}{2}(\sigma_B/\bar{B})^2]^{\frac{1}{2}}$ This is also the absolute maximum of Σ. At t=0, for the special case $K_\beta=\infty$, $\Sigma=\sigma_B$ This influence on uncertainty in the amplitude is shown in the plots of Figures 10 and 11. One can also readily discern the influence of K_β in Figure 10. For small K_β the initial value of Σ is large, but the cyclical variations are reduced. Note also that the values of Σ calculated on the basis of (16) and (14) are almost identical for $K \geq 4$, but even for K_β=1,2 the differences between the Von Mises and normal distributions is not very significant (cf. Figure 8). It also appears, from the plots in Figure 11, that smaller values of σ_ν, while permitting the fluctuations in Σ to persist, do not retard substantially the growth of the time-average standard deviation However the ratio of the amplitude of the fluctuations of $\bar{q}$ to the time average standard deviation of q does remain considerably larger if σ_ν is small.

6. Conclusions

Uncertainties in the initial conditions and/or parameters of time dependent processes manifest themselves in a variety of manners. By examining only a few simple cases we have been able to demonstrate that these effects must not be neglected. We have also shown that care must be taken that assumptions pertaining to the nature of the errors be reasonable in terms of any physical constraints on the problem.

Any uncertainties at all in the parameters of a time dependent problem may influence all the statistical properties of the ensemble of cases which constitute the population of possible "true" solutions. We have only examined the mean and the variance of a few simple time dependent variables. The history of the variance is clearly a complex phenomena since the variance is always non-linear, even which the basic prognostic equation is linear, and will thus be influenced by any statistical interactions that may be present.

Our most significant conclusion concerns the behavior of the ensemble mean. In general, the ensemble mean value of a variable will follow a different course than that of any single member of the ensemble. For this reason, it is clearly not an optimum procedure to forecast the atmosphere by applying deterministics hydrodynamic equations to any single initial condition, no matter how well it fits the available, but nevertheless finite and fallible, set of observations.

One cannot draw any conclusions, from the examples shown, of the relative importance of various effects in the much more complex hydrodynamic equations which describe the atmosphere. Certainly, though, the effects we have described will be present. One cannot help but note, for example, that in the examples shown correlations among the various parameters played a particular large role in determining eventual ensemble variance. Doesn't this suggest that one should design objective analysis schemes so as to introduce the right kinds of interdependences among the initial conditions? We may not want statistically independent or uncorrelated sets of initial values.

This being the case, how should we forecast the weather? To answer this question we will have to learn a great deal more about how uncertain our meteorological parameters and observations are, and how these uncertainties are propogated through the differential equations and numerical procedures of physical prediction.

References

Downs, Thomas D., 1964: The Von Mises Distribution: Derived Distributions, Regression Theory, and Some Applications to Meteorological Problems. Ph.D. Dissertation, University of Michigan, Ann Arbor, Michigan, 128 pp.

Gleeson, Thomas A., 1966: A Causal Relation for Probabilities in Synoptic Meteorology. Jour. Appl. Meteor., 5, 365-368.

_____, 1967: On Theoretical Limits of Predictability. Jour. Appl. Meteor., 6, 213-215.

_____, 1968: A Modern Physical Basis for Meteorological Prediction. Proceedings of the First National Conference on Statistical Meteorology, Amer. Meteor. Soc., Boston, Mass., 1-10.

Legends for figures

Figure 1. Ensemble standard deviation of q for various values of the coefficient of correlation between the initial value (Q) and the constant rate of change (C).

Figure 2. Ensemble standard deviation of T for various values of the coefficient of correlation between initial temperature (T_o) and the cooling coefficient (A).

Figure 3. Expected value of u as a function of time. Curves a, b, and c are cases for which the friction coefficient, k, has a normal distribution, a gamma distribution, and is known with certainty, respectively. In each case $\bar{v}_o=10$ and $\sigma_{v_o}=1$. For curves a and b $\bar{k}=10^{-4}$ sec^{-1} and $\sigma_k=3\text{x}10^{-5}$ sec^{-1}. Plotted points are based on samples of 1000, using a log-normal distribution for k.

Figure 4. Normal, log-normal and gamma probability distribution functions for E(k)=1, var(k)=0.1, if the abscissa is interpreted as k. Taking the absissa as units of $k\text{x}10^{-4}$ m the curves correspond to $E(k)=10^{-4}$ $var(k)=10^{-5}$.

Figure 5. Standard deviation of σ as a function of time. The curves are labeled as in Figure 3 and the points are again based on samples of 1000 pairs with k having a log-normal distribution.

Figure 6. Frequency distributions of v. The curve labeled t=0 is the theoretical initial distribution. Other curves are based on experimental sample of 1000 pairs of numbers for which v_o in normal with mean 10.0 and standard deviation 1.0, and k is log-normal with $\bar{k}=10^{-4}$ sec^{-1} and $\sigma_k=3\text{x}10^{-5}$ sec^{-1}.

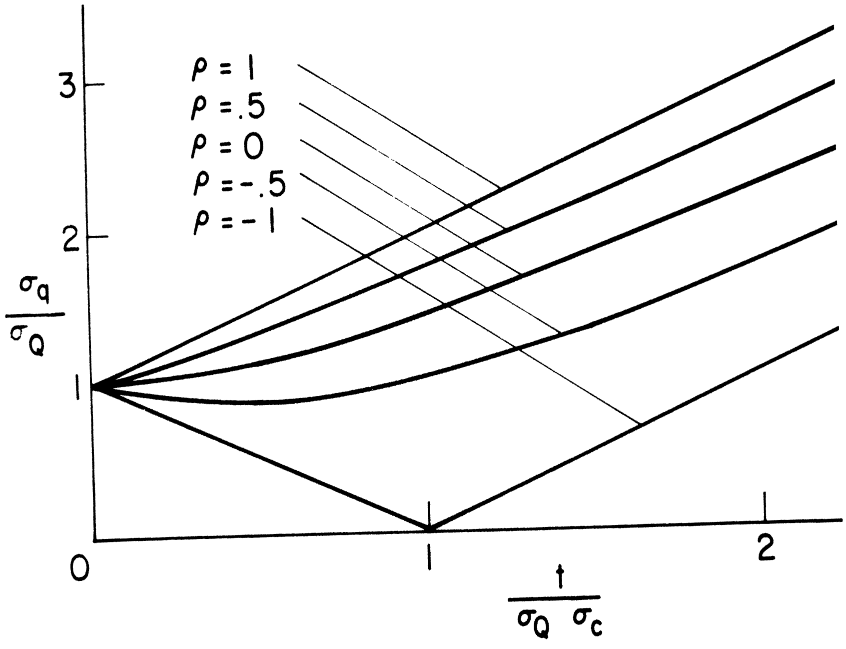
Figure 7. Standard deviation of u according to coefficient of correlation between v_o and k. The curve labeled $\rho=0$ is identical to curve b in Figure 5.

Figure 8. Normal and Von Mises distributions for comparable sets of parameters.

Figure 9. The ensemble mean of the cyclical variable. The ordinate is $(\bar{q}-\bar{Q})$ for $\bar{B}=1$, and the abscissa is $\bar{\nu}t$. (See equation 15)

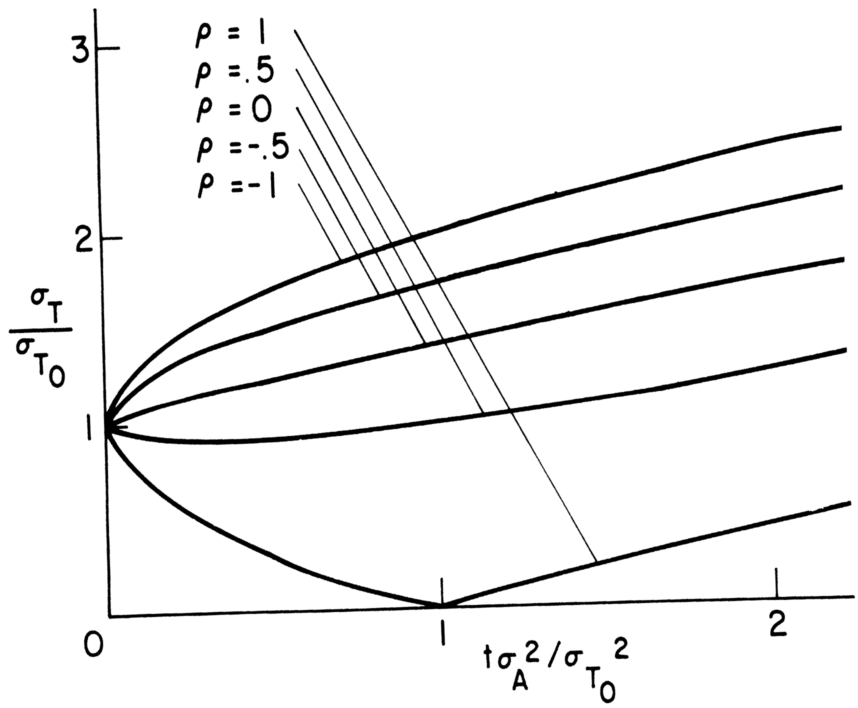
Figure 10. Plots of the quantity $\Sigma=(\sigma_q^2 - \sigma_Q^2)^{1/2}/\bar{B}$. The abscissa is $\bar{\nu}t$. All curves were calculated for $\sigma_\nu=.1$ and $\sigma_B=.1$. For curves labeled a, b, c and d, K= 1, 2, 4, ∞. respectively. The subscripts 1 and 2 refer, respectively, to the use of eq (16) and (14).

Figure 11. Plots of the quantity $\Sigma=(\sigma_q^2 -\sigma_Q^2)^{1/2}/\bar{B}$. The abscissa is $\bar{\nu}t$. The parameters used in the calculations were a) $\sigma_\nu=.1$, $\sigma_B=.3$, K=∞; b) σ_ν=Q3, $\sigma_B=.1$, K=∞; c) σ_ν=Q3, $\sigma_B=.3$, K=1; d) $\sigma_\nu=.03$, $\sigma_B = .3$, K = 4. Again the subscripts 1 and 2 refer to the use of eq(16) or (14) respectively.



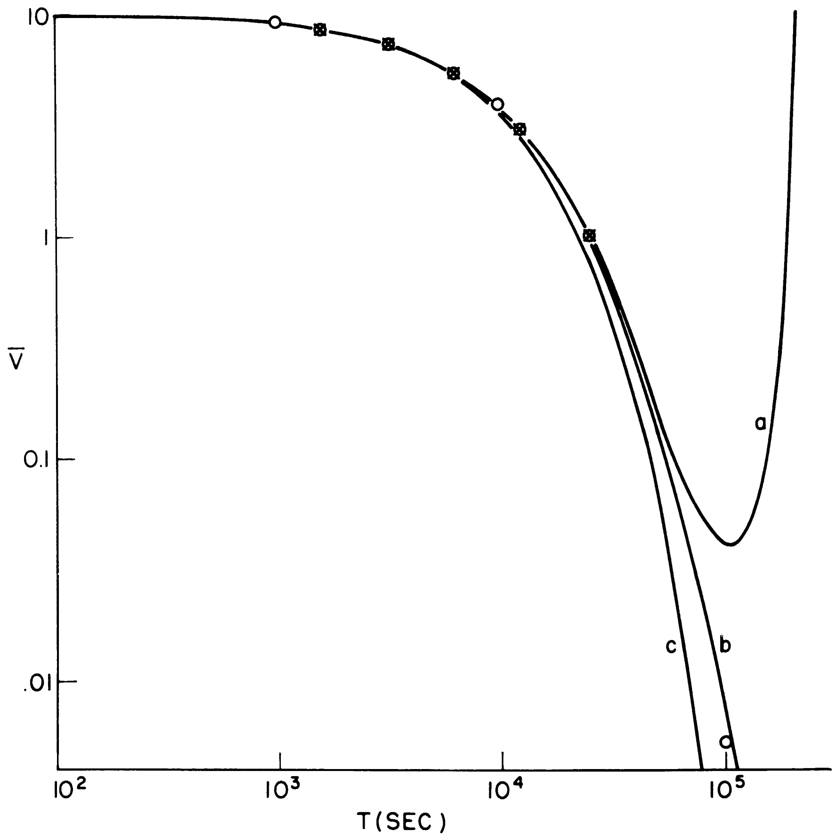
ρ = 1
ρ = .5
ρ = 0
ρ = -.5
ρ = -1
3
2
1
0
1
2
σ_q / σ_Q
$\frac{t}{\sigma_Q \sigma_c}$

Fig. 1



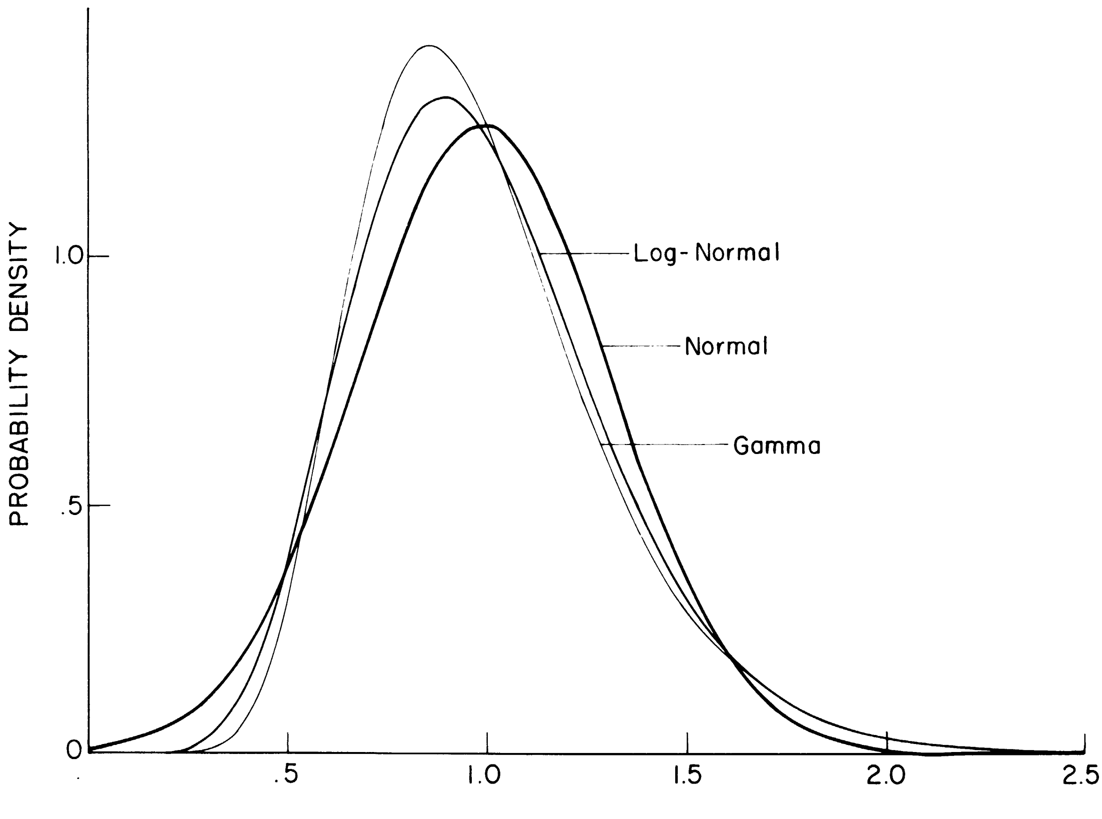
ρ = 1
ρ = .5
ρ = 0
ρ = -.5
ρ = -1
3
2
1
0
1
2
σ_T / σ_{T_0}
$t\sigma_A^2/\sigma_{T_0}^2$

Fig. 2



10
1
0.1
.01
$\bar{V}$
10^2
10^3
10^4
10^5
T(SEC)
a
b
c

Fig. 3



PROBABILITY DENSITY
1.0
.5
0
.5
1.0
1.5
2.0
2.5
Log-Normal
Normal
Gamma

Fig. 4

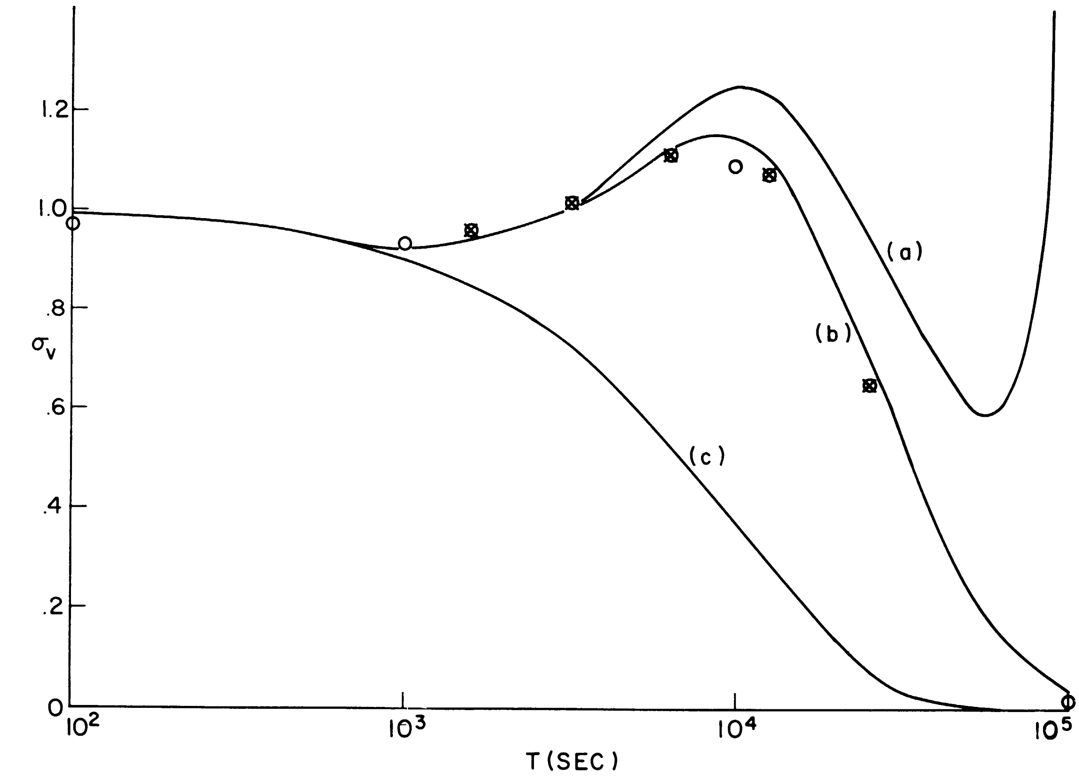

1.2
1.0
.8
σ_v
.6
.4
.2
0
10^2
10^3
10^4
10^5
T (SEC)
(a)
(b)
(c)

Fig. 5

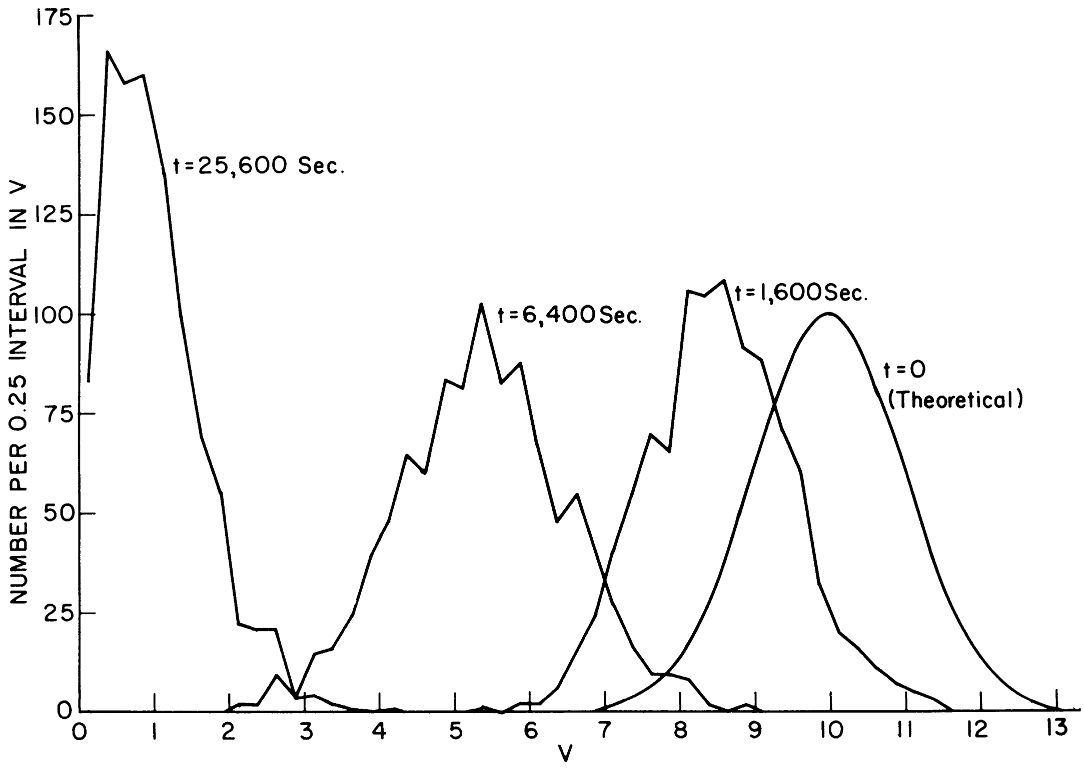
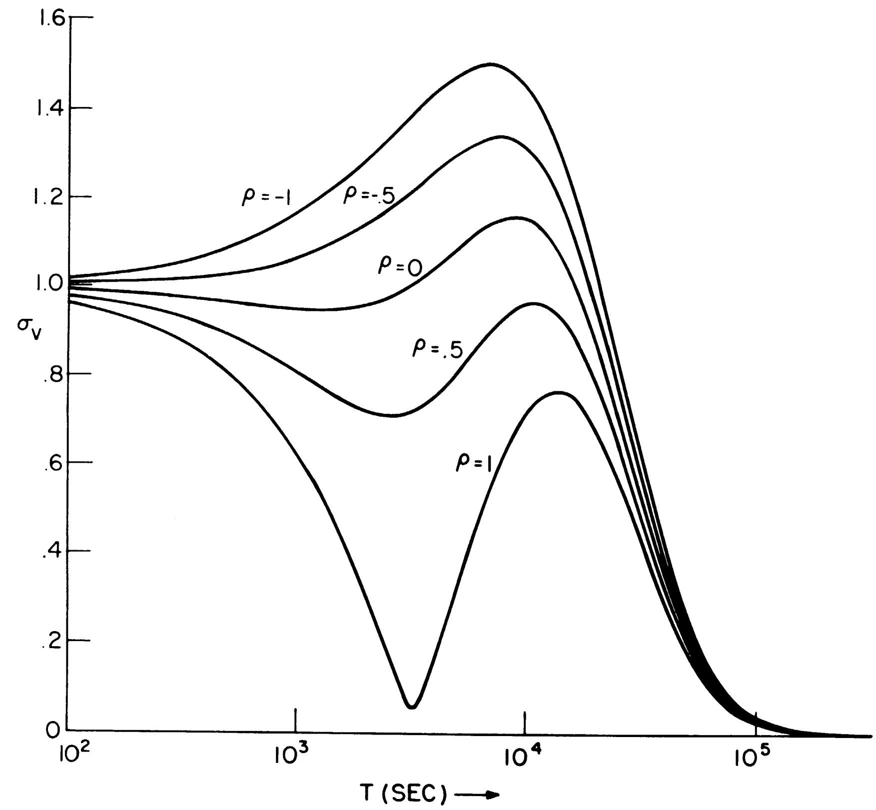
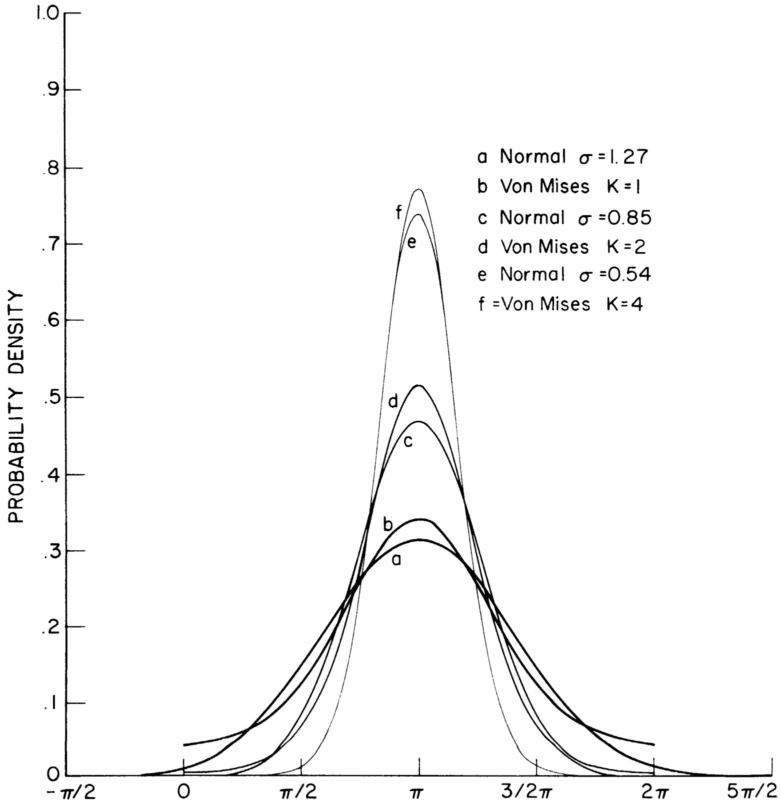

175
150
125
100
75
50
25
0
NUMBER PER 0.25 INTERVAL IN V
t=25,600 Sec.
t=6,400 Sec.
t=1,600 Sec.
t=0
(Theoretical)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
V

Fig. 6



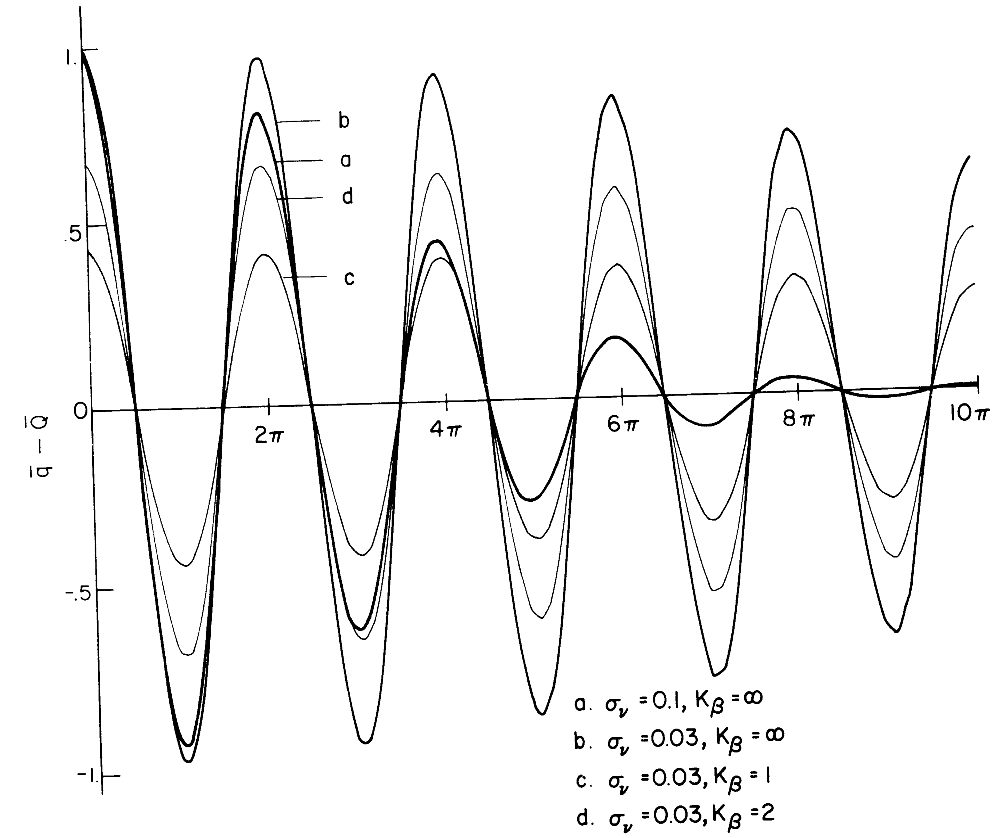
1.6
1.4
1.2
1.0
σ_V
.8
.6
.4
.2
0
ρ = -1
ρ =-.5
ρ = 0
ρ = .5
ρ = 1
10^2
10^3
10^4
10^5
T (SEC) ⟶

Fig. 7



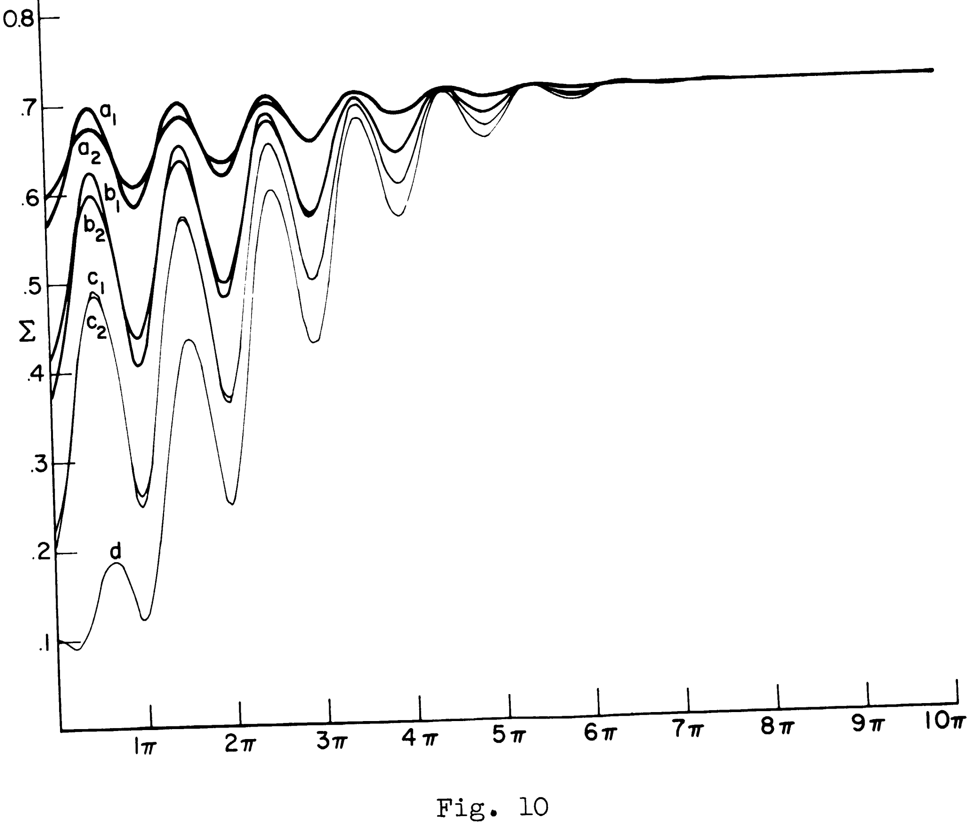
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
PROBABILITY DENSITY
a Normal σ = 1.27
b Von Mises K = 1
c Normal σ = 0.85
d Von Mises K = 2
e Normal σ = 0.54
f = Von Mises K = 4
f
e
d
c
b
a
$-\pi/2$
0
$\pi/2$
π
$3/2\pi$
2π
$5\pi/2$

Fig. 8



1.
.5
0
-.5
-1.
$\bar{q}-\bar{Q}$
2π
4π
6π
8π
10π
b
a
d
c
a. $\sigma_\nu = 0.1$, $K_\beta = \infty$
b. $\sigma_\nu = 0.03$, $K_\beta = \infty$
c. $\sigma_\nu = 0.03$, $K_\beta = 1$
d. $\sigma_\nu = 0.03$, $K_\beta = 2$

Fig. 9



0.8
.7
.6
.5
.4
.3
.2
.1
Σ
a_1
a_2
b_1
b_2
c_1
c_2
d
1π
2π
3π
4π
5π
6π
7π
8π
9π
10π

Fig. 10

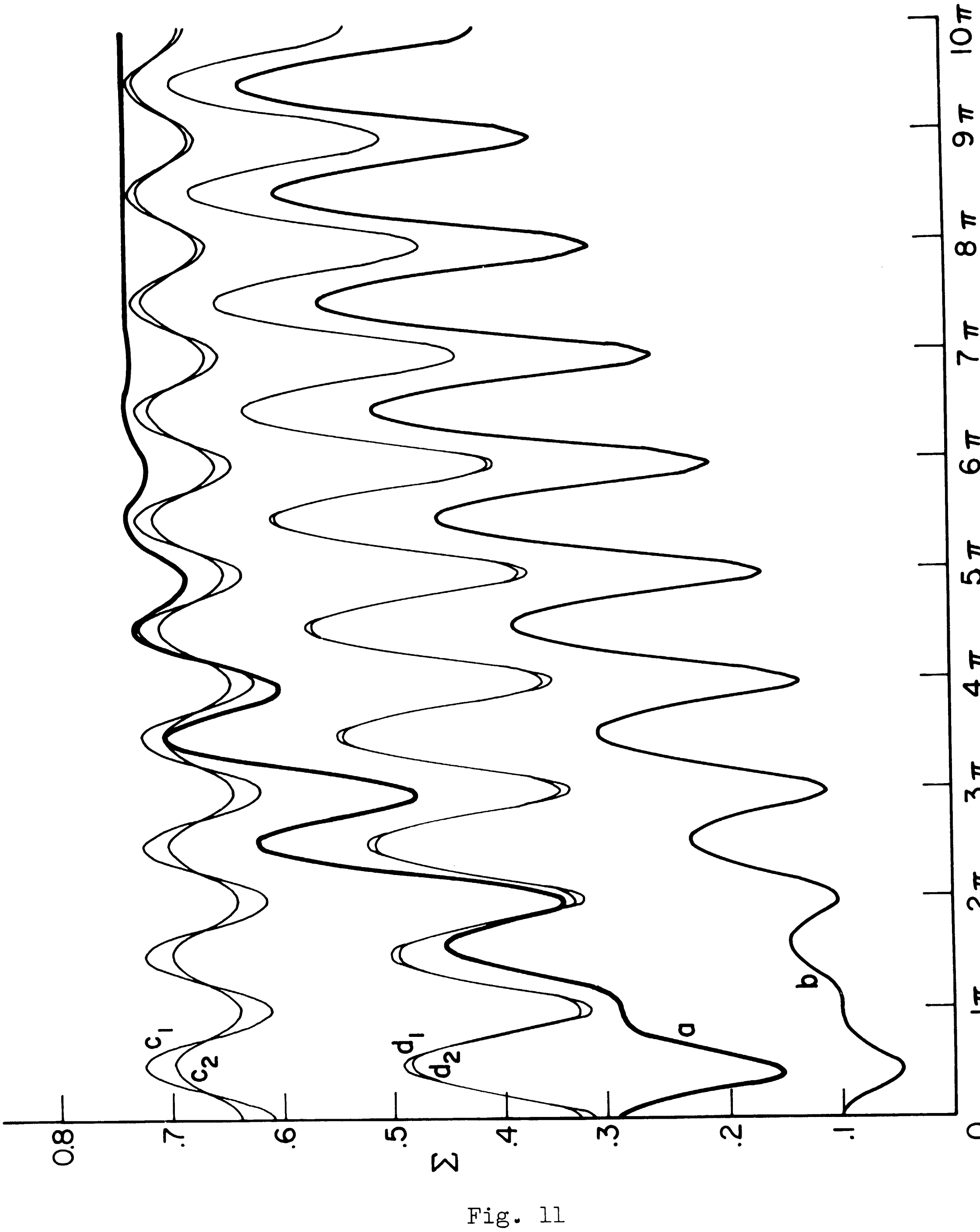


Fig. 11