

MATHEMATICAL INVESTIGATION OF THE TRACKING SIGNAL
IN THEORY OF FORECASTING

by 6408

PRASAD PATHARE

B. Eng. (Mech.) (Hons.), University of Bombay, 1965

B. Eng. (Elec.) (Hons.), University of Bombay, 1966

—

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1971

Approved by:

L. E. Grosh

Major Professor

LD
2668
T4
1971
P3
C.2

TO Papa and Mummy
who spelled LIFE for me.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	iv
Chapter 1 : INTRODUCTION	1
Chapter 2 : MEASUREMENT AND ANALYSIS OF ERRORS	5
Chapter 3 : TRACKING SIGNAL	20
Chapter 4 : EXPERIMENTAL RESULTS AND CONCLUSION	33
REFERENCES	47
APPENDICES :		
APPENDIX A - NORMALITY OF FORECAST ERRORS	49
APPENDIX B - FOLDED NORMAL DISTRIBUTION	54
APPENDIX C - COMPUTER PROGRAMS AND SUBROUTINES	58
APPENDIX D - CHI-SQUARE TEST FOR RANDOM NUMBER GENERATOR	69

ACKNOWLEDGEMENT

The author wishes to express his deep appreciation and sincere gratitude to Dr. L. E. Grosh for his valuable counsel and guidance in this research.

The author is also very grateful to Mrs. Jeanne Carpenter for her prompt and excellent typing work.

**THIS BOOK
CONTAINS
NUMEROUS PAGES
WITH THE ORIGINAL
PRINTING BEING
SKEWED
DIFFERENTLY FROM
THE TOP OF THE
PAGE TO THE
BOTTOM.**

**THIS IS AS RECEIVED
FROM THE
CUSTOMER.**

CHAPTER 1

INTRODUCTION

Forecasting occupies a unique position in the modern industrial structure in that it helps to plan the future development. Actual forecasts are only a part of some control system. A specific action will be based on the forecast made some prior period and hence the success of that action, and the subsequent profit, will be a function of the forecast accuracy. In practical situations the difference between the actual observation and its prior forecast is almost inevitable because any time series can always be decomposed into the following components:

- a. a constant entity or level,
- b. a trend, which can, generally, be expressed by a low degree polynomial model,
- c. a cyclic or seasonal variation, a periodic term,
- and d. a noise, a component that cannot be categorized as any of the above, it is unaccounted for and varies randomly, but is always present in practical situations.

These factors form the basics of the forecasting models.

Noise can be measured in statistical terms only; hence, its prediction necessitates probabilistic models rather than simple mathematical models (which may have algebraic, trigonometric or exponential form) designed to represent the first three factors.

Apparently, a rigid control system cannot be designed to restrain the magnitude or occurrence of the noise factor, which makes it more difficult to generate a prior forecast with a rigid level of precision.

To deal with different types of forecasting problems, there exist many kinds of forecasting systems that can be broadly classified under Static and Dynamic methods. The first classification includes subjective estimates, graphical curve fitting and regression analysis, whereas the second has simulation, moving average, adaptive smoothing, exponential smoothing, etc. In all cases, the efficiency of the system is decided by the magnitudes of the forecast errors.

A forecasting error is defined as the difference between the actual observation and its prior forecast. A good forecasting system will care for,

1. a satisfactory and feasible forecasting method,
- and 2. an allowance for the forecast errors.

The adequacy of the forecasting system will depend upon the extent to which these factors are taken care of.

The first factor implies the establishment of a practical forecast technique. The second urges us to devise a method to increase the accuracy of the forecasts. The present research deals with a part of the problem involved in this second aspect.

STATEMENT OF THE PROBLEM

Since improvement in system accuracy involves the reduction in error magnitudes, the measurement of system errors and setting the control limits for their variational range plays an important role. There are different criteria for error measurement, the most widely applicable and versatile of them being a statistic called a 'tracking signal.' These will be discussed in detail in the next chapter. However, it needs to be mentioned that the tracking signal statistic smooths the error statistic and maps it to a smaller range of variation.

The present research deals with the statistical distribution of this tracking signal statistic. The latter being a ratio of two other error statistics, it is merely a series of numbers, the magnitudes of which will depend upon the magnitudes of system errors. Thus, if the statistical distribution of the errors is known, the statistical distribution of the tracking signal can be found. The present thesis is an attempt to find the tracking signal distribution, making suitable simplifying assumptions.

The knowledge of the distribution of this tracking signal statistic will facilitate the determination of suitable confidence range to check on the magnitude of the system errors of any individual forecasting system.

Although the theory developed here may be extended and modified to suit other dynamic systems, the exponential smoothing technique is the one chosen for illustration. The entire treatment uses exponential smoothing as a basis. The basic reason for this is that the exponential smoothing provides for updating the coefficients of the forecasting model at each new sampling interval to correct for the past errors; also, the ability of this technique to react to changes in the observed process is another desirable feature. The later part of this work includes a simulation experiment to test the practical agreement of the developed theory with the real-world-situation.

CHAPTER 2

ERROR MEASUREMENT AND ANALYSIS

The forecasting error is defined as the difference between the actual observation and the forecast model fitted. It is similar to the residuals in the regression technique, but that the forecast is forward in time and hence the observation is not the one used in building the forecast.

The mean of the distribution of the forecast errors will be zero only if the model of the underlying process continues to hold throughout the forecast lead time. The actual observation is a modulation of this process with a noise component. The adaptive (or dynamic) forecasting system amplifies the noise. That is, the forecasts have a statistical distribution caused by the past noise. Since the forecast is a linear combination of the past observations, the distribution of the forecasts will tend to be normal. (1,2)

The distribution of the forecast errors will be the convolution of the noise distribution and the forecast distribution; the noise samples may or may not be random, the forecasts, being a linear combination of the past observations, are definitely correlated.

Due to the nature of dynamic forecasting, "...whenever any such system is set up, it is very desirable to incorporate some form of automatic monitoring to ensure that the system remains in control." (5)

NORMALITY OF FORECAST ERRORS

In order to effect an improvement in forecast accuracy, it is necessary to represent the data on forecast errors by a probability distribution. This is due to the statistical nature of the system noise. The data also need a meaningful statistical interpretation to make a close allowance for these errors. Different statistics and parameters have already been introduced to effectively deal with this aspect. In the present research, a statistic called "tracking signal" is given further consideration in this respect.

Three types of information are basically needed:

1. the form of the error distribution,
2. the parameters of the error distribution,
3. the serial correlation between different samples of the distribution.

Simulation has proved that in a great number of cases, the error distribution is approximately normal. (1,2) The analytical approximation to this effect is derived in Appendix A. In most of the cases this approximate normality of errors can be assumed without question. If the noise in the past data is random and is around a zero mean, the error distribution will have a zero mean as long as the underlying process does not change, and the error variance can be estimated from the response characteristics of the forecasting system.

**THIS BOOK
CONTAINS
NUMEROUS PAGES
WITH DIAGRAMS
THAT ARE CROOKED
COMPARED TO THE
REST OF THE
INFORMATION ON
THE PAGE.**

**THIS IS AS
RECEIVED FROM
CUSTOMER.**

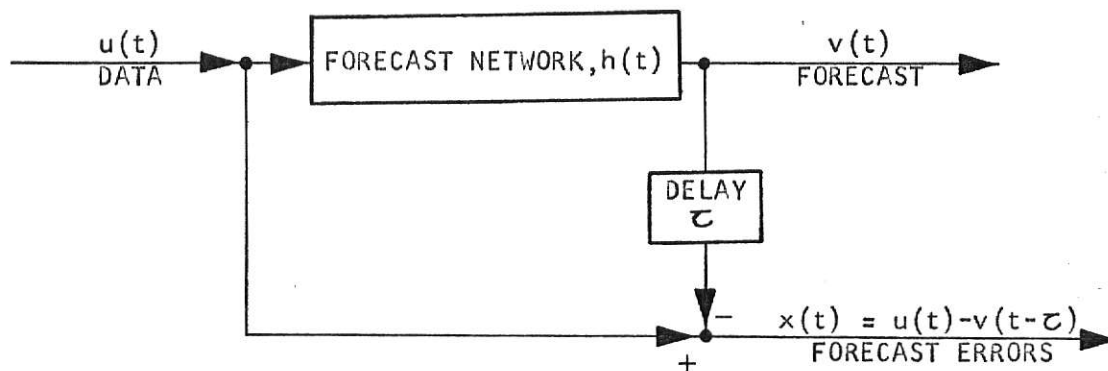
ERROR GENERATOR:

Fig. 1

The block diagram in the Fig. 1 above represents the analogous error generator. The time series of data is fed into the forecasting system represented by the impulse response function, $h(t)$. The output is the forecasts of the observations. This is delayed over the lead time and forms a negative feedback to the time series of data, the resultant interaction of which generates the forecast errors.

SERIAL CORRELATION OF FORECAST ERRORS:

If τ represents the forecast lead time, the forecast error $x(t)$ at time t is given as,

$$x(t) = u(t) - \hat{u}_{\tau}(t - \tau) \quad \dots (1)$$

where $u(t)$ is the actual observation at time t , and $\hat{u}_{\tau}(t - \tau)$ is

its forecast made τ periods ago. In terms of impulse response function $h(t)$ of the forecast system, the forecast can be represented as the convolution equation of the form,

$$\hat{u}_{\tau}(t) = \sum_{n=0}^{\infty} h(n) \cdot u(t-n) \quad \dots (2)$$

which changes the equation (1) to the form,

$$x_{\tau}(t) = u(t) - \sum_{n=0}^{\infty} h(n) \cdot u(t-\tau-n) \quad \dots (3)$$

If the forecast system is accurate enough, all the forecast errors would be distributed around a zero mean value. The errors may or may not be serially correlated. The relevance in this respect may be evaluated from the autocorrelation coefficients of the errors.

If F_{τ} denotes the sum of cross products of the errors τ periods apart in the error time series such that

$$F_{\tau} = \sum_t x(t) \cdot x(t+\tau), \quad \dots (4)$$

the mean value of F_{τ} , viz. $E(F_{\tau})$, is called its average lagged product. If the forecast errors are distributed about a zero mean, that is if $E(x(t)) = 0$, the average lagged product is the autocovariance of the errors; and is denoted by $R_{xx}(\tau)$.

It is, thus, obvious from equation (4) that the variance, σ_x^2 , of the error distribution will simply be,

$$\sigma_x^2 = R_{xx}(0).$$

Also, that the autocovariance function is symmetrical, that is

$$R_{xx}(\tau) = R_{xx}(-\tau).$$

The set of values of autocovariance over the entire range of τ is called the autocorrelation function, most commonly expressed as a ratio,

$$\rho(\tau) = \frac{R_{xx}(\tau)}{R_{xx}(0)},$$

called the autocorrelation coefficient. This ratio has a finite domain of $(-1,1)$. In the case of noisy data (uncorrelated errors), the value of this ratio is quite close to zero. The errors in such cases are quite independent of each other. If the value, on the other hand, lies quite close to ± 1 , high correlation exists between the errors and the knowledge of any error may be effectively used to forecast the subsequent errors. From the equation (3) above,

$$R_{xx}(\tau) = R_{uu}(\tau) - \sum_{n=0}^{\infty} h(n) \left[R_{uu}(\tau + n) + R_{uu}(\tau - n) \right]$$

$$+ \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} h(m) \cdot h(n) \cdot R_{uu}(\tau + n - m), \quad \dots (5)$$

where $R_{uu}(\tau)$ is the autocovariance of the observations sampled τ periods apart in the given time series. If the observations were random and uncorrelated,

$$R_{uu}(0) = \sigma_u^2$$

and $R_{uu}(\tau) = 0$ for $\tau \neq 0$.

$$\begin{aligned} \text{This yields, } R_{xx}(0) &= \sigma_u^2 \cdot \left[-1 + \sum_{n=0}^{\infty} h^2(n) \right] \\ \text{and } R_{xx}(\tau) &= \sigma_u^2 \cdot \sum_{n=0}^{\infty} h(n) \cdot h(\tau + n) \end{aligned} \quad \dots (6)$$

MEASUREMENT AND ANALYSIS OF ERRORS:

The following criteria can be used in evaluating errors:

1. cumulative error
2. mean error
3. variance of errors
4. mean absolute error
5. smoothed error
6. smoothed mean absolute deviation
7. tracking signal.

Cumulative error is the sum of forecast errors. Mean error and cumulative error are two different criteria of expressing the same statistic. For normally distributed errors with zero mean, the cumulative error will fluctuate around zero.

But even if this condition is satisfied, it does not necessarily testify to the forecast accuracy. The precision of the forecasts will, in such cases, be judged by the standard deviation (or variance, in other words) of the errors.

In analytical problems it is most convenient to deal with the variance σ^2 as a measure of scatter of the data around the mean. However, the computation of the standard deviation (square-root of variance) from a large number of observations is an involved task. Hence, a simpler and more meaningful statistic called 'mean absolute error' is introduced. If $x(t)$ represents the error having a statistical distribution $p(x)$ about the mean μ , then the mean absolute error is defined as,

$$\Delta = \int_{\Omega} |x - \mu| \cdot p(x) \cdot dx$$

for a continuous density function. For a discrete distribution this takes the following form:

$$\Delta = \sum_{t=0}^{\infty} |x(t) - \mu| \cdot p(x(t)) \cdot$$

If $\mu = 0$, as in most cases, the computation is considerably reduced. In the general case, where the error distribution may or may not be normal, a normalized variable z may be defined such that

$$z = \frac{x - \mu}{\sigma_x} \quad ,$$

where σ_x is the standard deviation of errors. The mean absolute error is then

$$\begin{aligned}\Delta &= \int_{\Omega} |z\sigma_x| \cdot p(x) \cdot dx \\ &= \sigma_x \int_{\Omega} |z| \cdot p(z) \cdot dz \quad \dots (7)\end{aligned}$$

the integral being over the entire sample space Ω . Hence, the mean absolute error is proportional to the standard deviation of the errors. For normally distributed errors, the constant of proportionality is 0.7979 (1).

All the criteria introduced above are, thus, more or less related to the same common statistic. If $y(t)$ represents the cumulative error, then

$$y(t) = \sum_{i=0}^t x(i) \cdot$$

For the forecasting system using exponential smoothing, the impulse response function, $h(t)$, producing the cumulative error is given as,

$$h_Y(t) = \beta^t,$$

where $\beta = 1 - \alpha$, and α is the smoothing statistic of the system. (Also, $0 < \alpha, \beta < 1$).

Hence, if σ_u^2 is the variance of the input data noise, the variance of the cumulative error is,

$$\sigma_Y^2 = \sigma_u^2 \cdot \sum_{t=0}^{\infty} h_Y^2(t) = \frac{\sigma_u^2}{1 - \beta^2} \quad \dots (8)$$

If the errors are normally distributed, and the noise samples were serially independent, the mean absolute error is (1),

$$\Delta = \sqrt{2/\pi} \cdot \sqrt{2/(2-\alpha)} \cdot \sigma_u$$

$$\text{or} \quad \sigma_u = \Delta \cdot \sqrt{\pi/2} \cdot \sqrt{2 - \alpha/2} \quad \dots (9)$$

Equations (8) and (9) above give the relation between σ_Y and Δ . The standard deviation of the cumulative error is, thus, proportional to the mean absolute error.

All the criteria discussed above are quite adequate as statistical measures of error analysis as long as the actual process level does not change. However, if a sudden change in the level occurs, the errors grow larger in magnitude. As the forecasts 'catch up' with the changed process level, the errors again tend to settle down about zero mean. Yet the forecast catch-up period does not match with the settling period of the error statistic described above.

For example, let us consider a practically noiseless process infinitely settled at a level 'D.'

The forecasts in such cases are practically equal to the observation values, and the mean error lies on the zero level. If now there is a sudden rise to another noiseless level 'G,' a big error of magnitude $(G-D)$ is observed in the forecast. As the forecasts catch up with this change, the magnitude of the error decreases. In the figure below, the nature of impulse response of the forecasts is shown.

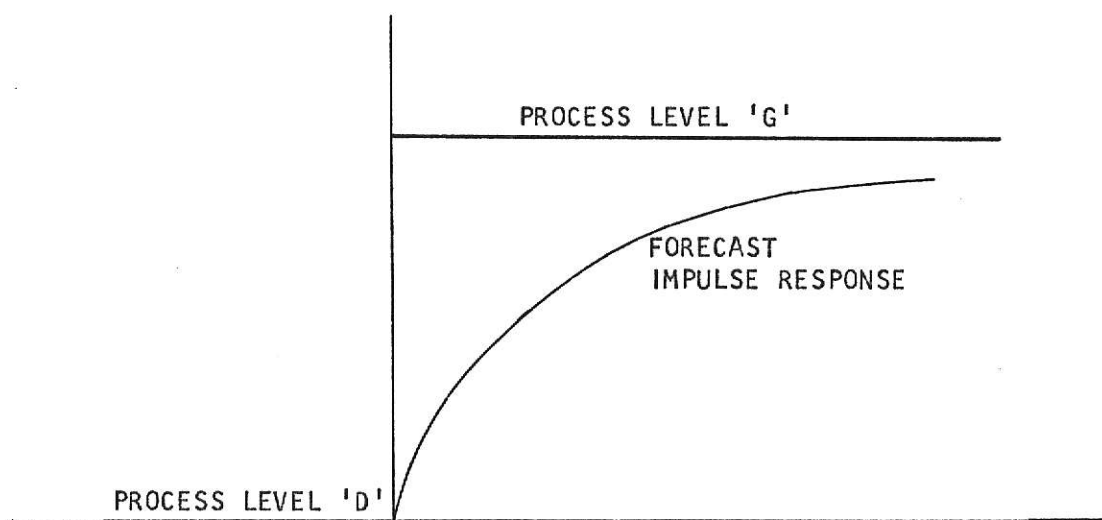


Fig. 2

However, during this catching up period, the errors are all of the same sign and the mean error drifts away from zero. The period for the mean error to drop down to practically zero is substantially different, though, from the forecast catching-up period. This means that during such transition stage the above error statistics are inadequate to judge the forecast precision. This necessitates the definition of some other modified statistic that can effectively track the forecast precision.

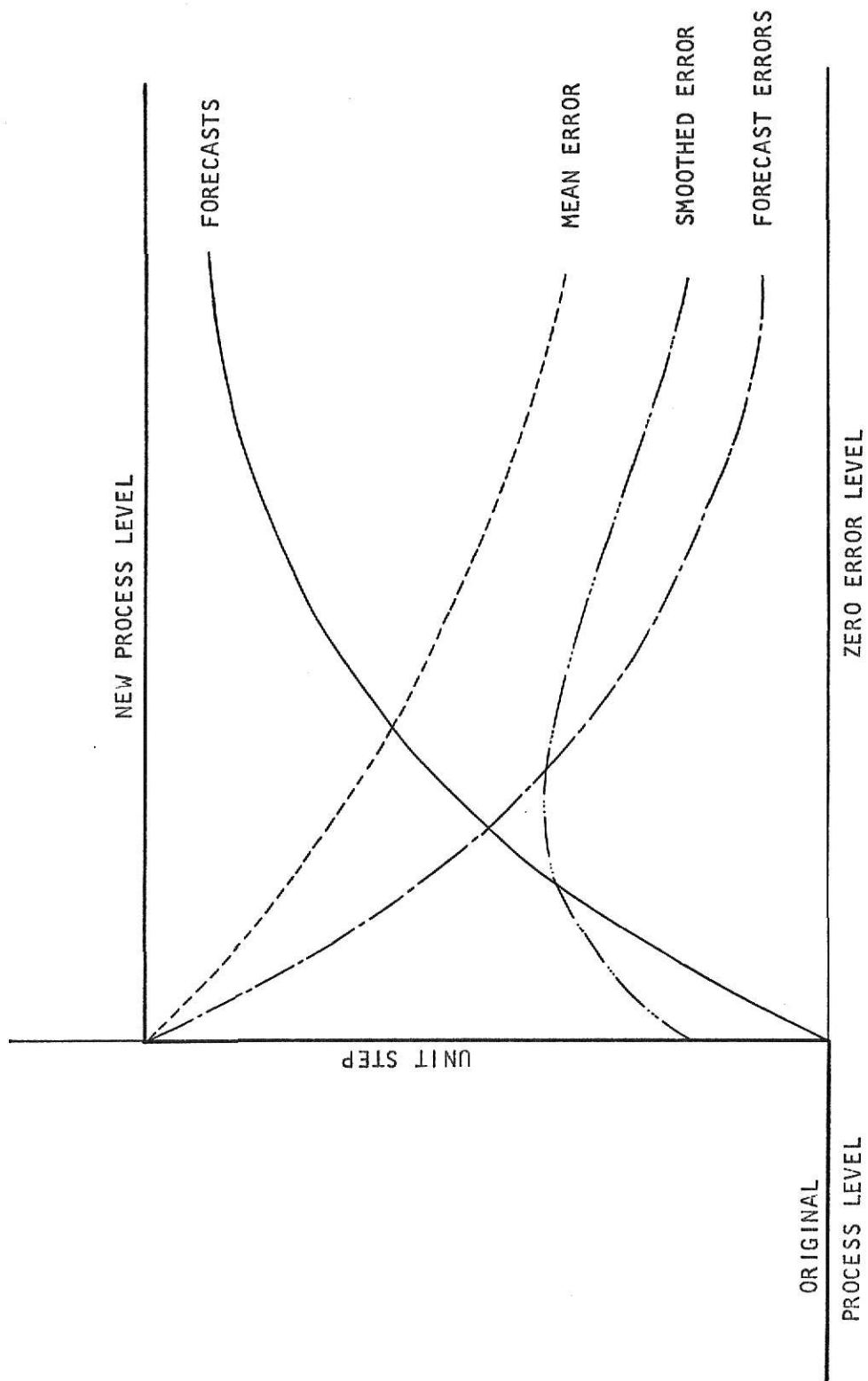


Fig. 3 Impulse Response of the Forecast System to a Unit Step

Thus, smoothed statistics are introduced.

Smoothed error: This is defined as,

$$\text{smoothed error} = (1 - \alpha) \text{ previous smoothed error} + \alpha (\text{latest error}),$$

where α is the smoothing statistic. However, this α need not be the same as the smoothing statistic of the forecast system, though the same value in both the cases may have a better synchronization of the 'catching-up' and 'settling-down' periods.

Figure 3 compares the forecast catch-up process with the settling-down process of the mean error and the smoothed error. The smoothing statistic is the same for both the forecasts and the smoothed error. It could be observed that the smoothed error and the forecasts reach the desired levels almost simultaneously. Therefore, this error statistic can be effectively used as a measure of forecast precision at any instance.

$$\text{By definition, smoothed error} = \sum_{i=0}^n \alpha \beta^i x(n-i),$$

and for a stabilized system (that is for sufficiently large n) variance of the smoothed error is,

$$\begin{aligned} \sigma_p^2 &= \sum_{i=0}^{\infty} \alpha \beta^i \sigma_x^2 = \frac{\alpha \cdot \sigma_x^2}{1 - \beta^2} \\ &= \frac{\alpha}{2 - \alpha} \cdot \sigma_x^2 \end{aligned} \quad \dots (10)$$

That is, the standard deviation of the smoothed error is proportional to the standard deviation of the errors.

Smoothed mean absolute deviation (SMAD) : This is another smoothed statistic defined as,

$$\text{smoothed mean absolute deviation} = (1 - \alpha) \text{ previous SMAD} + \alpha (\text{latest absolute error}).$$

This is simply a convenient measure of the noise in the system.

In fact, its mean is proportional to the standard deviation of errors and is simpler to compute - the facts that make this statistic quite important and meaningful.

Of all these criteria the most important and the most widely applicable is a statistic called the 'tracking signal.' R. G. Brown (1,2) first defined this as,

$$\text{Tracking signal} = \frac{\text{Cumulative error}}{\text{Mean absolute error}}$$

He used the mean absolute error as a measure of scatter around the mean. The error tracking signal is, thus, a normalized measure of whether the sum of forecast errors is reasonably close to zero or not. The rationale of this approach is that if the forecast is good, the cumulative (and also the mean error) will be approximately zero. As discussed before, if the actual process changes, the cumulative error would tend to drift away from this zero level.

This, in turn, makes the tracking signal to drift away considerably. By setting the control limits, the tracking signal can be checked for its variability. If the limits were exceeded for more than two consecutive forecasts, re-evaluation of the forecasting model may be necessary.

However, D. W. Trigg (5), of Kodak, Ltd., pointed out two serious faults with Brown's definition of this statistic even though he agreed with the approach. To quote Trigg,

"---- two serious difficulties arise:

1. Once the tracking signal has gone out of control, it will not necessarily return within limits even though the forecasting system itself comes back in control. Consequently, intervention is necessary to set the cumulative error back to zero if future false alarms are to be avoided. Such interventions can be tedious and expensive when several hundred items are involved.
2. Ironically, if the system starts to give exceptionally accurate forecasts, the tracking signal may go out of limits. If perfect forecasts begin to occur, the mean absolute error will tend to zero while the sum of errors will remain unaltered; this leads to the tracking signal tending to infinity." (5).

As remedial measures he suggested that the tracking signal be defined as,

$$\text{Tracking signal} = \frac{\text{Smoothed error}}{\text{Smoothed mean abs. deviation}} .$$

Both the above drawbacks were thus eliminated. Hereafter, in this research, this modified definition of tracking signal is used.

CHAPTER 3

THE TRACKING SIGNAL

As stated in the previous chapter, in most of the cases, but not always, the distribution of the forecast errors is normal. One desirable feature, though, is that the mean of the error distribution should be zero. The concept of the tracking signal statistic can be used to determine whether it is so or not. Also, the significance of the computed value of the high order coefficients in a polynomial model can be checked. The distribution of this statistic, thus, assumes a significant measure in determining an allowance for forecast errors in actual forecasting problems. From the definition, it is quite obvious that this statistic should be centered about zero if the error distribution has a zero mean. In what follows Trigg's (5) definition of the tracking signal is used.

SMOOTHED ERROR:

The errors occurring in the forecasting system are assumed to be serially uncorrelated (purely random noise) and normally distributed with a zero mean and a constant variance, σ^2 . If α is the smoothing statistic of the forecasting system, such that

$\beta = 1 - \alpha$, and $(x_n, x_{n-1}, x_{n-2}, \dots, x_1, x_0)$ are the errors respectively dating backwards, then the smoothed error is,

$$\begin{aligned}
 P_n &= \alpha x_n + \alpha \beta x_{n-1} + \alpha \beta^2 x_{n-2} + \dots + \alpha \beta^n x_0 \\
 &= \alpha x_n + \beta P_{n-1}
 \end{aligned}$$

That is, smoothed error = α (latest error) + β (previous smoothed error) .

Since, $x_i \sim N(0, \sigma^2)$, $\alpha x_i \sim N(0, \alpha^2 \sigma^2)$ *.

And similarly, $P_n \sim N \left[0, \left(\sum_{i=0}^n \alpha^2 \beta^{2i} \right) \sigma^2 \right]$

i.e. the smoothed error P_n is distributed normally with zero mean and variance of,

$$\sigma_{P_n}^2 = \left(\sum_{i=0}^n \alpha^2 \beta^{2i} \right) \sigma^2 .$$

Considering the series,

$$\begin{aligned}
 &\alpha^2 + \alpha^2 \beta^2 + \alpha^2 \beta^4 + \dots + \alpha^2 \beta^{2n} , \\
 &\quad \quad \quad (n+1 \text{ terms})
 \end{aligned}$$

$$\begin{aligned}
 \text{we have, } \sum_{i=0}^n \alpha^2 \beta^{2i} &= \alpha^2 \sum_{i=0}^n \beta^{2i} \\
 &= \alpha^2 \cdot \frac{1 - \beta^{2n+2}}{1 - \beta^2} \\
 &= \frac{\alpha (1 - \beta^{2n+2})}{2 - \alpha}
 \end{aligned}$$

* $x_i \sim N(0, \sigma^2)$ denotes that x_i has Normal distribution.

For a stabilized system, i.e., for a practically infinite data series, the above series approaches the limit of

$$\gamma^2 = \frac{\alpha}{2 - \alpha} \quad , \quad \text{since } \alpha, \beta < 1.$$

Thus, for a stabilized forecast system,

$$P \sim N \left[0, \gamma^2 \sigma^2 \right]$$

where γ^2 is a factor as defined above, and the probability density function for $P \leq p$ can be written as,

$$f_1(p) = \frac{1}{\sqrt{2\pi} \cdot \gamma \sigma} \exp \left(- \frac{p^2}{2\gamma^2 \sigma^2} \right) \quad \dots (11)$$

SMOOTHED MEAN ABSOLUTE DEVIATION:

This statistic plays an important role in computing the average allowance for error in the used forecasting system. The reasons being that it is easier to compute and also that its expectation is proportional to the standard deviation of the system errors to a first degree of approximation. In most applications, the smoothed mean absolute deviation changes slowly with time, and it also has a statistical distribution; but at any particular instance the value may be regarded as constant for all practical purposes. Normality of errors is assumed here, and this may be an essential factor to be considered.

Since $SMAD = \alpha(\text{latest absolute error}) + (1-\alpha) \text{ previous SMAD}$,

$$\begin{aligned} Q_n &= \alpha |x_n| + \beta Q_{n-1} \\ &= \alpha |x_n| + \alpha \beta |x_{n-1}| + \dots \dots \dots + \alpha \beta^n |x_0| . \end{aligned}$$

Since $x_i \sim N(0, \sigma^2)$, $|x_i| \sim N_f(0, \sigma^2)$

where N_f represents the folded normal distribution (4). Thus

in terms of σ^2 , the probability density function of $|x_i|$ may be written as,

$$f_2(|x_i|) = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sigma} \cdot \exp\left(-\frac{x_i^2}{2\sigma^2}\right), \quad -\infty < x_i < \infty$$

If μ_1 and σ_1^2 are the mean and variance values of the random variable $w = |x|$, where $x \sim N(0, \sigma^2)$, it can be shown that,

(4)

$$\begin{aligned} \mu_1 &= \sqrt{\frac{2}{\pi}} \cdot \sigma = 0.7979 \sigma \\ \text{and } \sigma_1^2 &= \frac{\pi - 2}{\pi} \sigma^2 = 0.3634 \sigma^2 \end{aligned} \quad \left| \dots (12) \right.$$

and the probability density function of w , in terms of μ_1 and σ_1^2 , can be expressed as, (see Appendix B for the proof)

$$f_2(w) = \frac{\pi \mu_1^3}{2 (\mu_1^2 + \sigma_1^2)^2} \exp\left(-\frac{\mu_1^2 w^2}{4 (\mu_1^2 + \sigma_1^2)^2}\right),$$

$$0 \leq w < \infty \quad \dots (13)$$

At this stage if we define a constant V such that ,

$$V = \frac{\sigma_1}{\mu_1} = 0.755$$

the above equation can be written as,

$$f_2(w) = \frac{\pi}{2 \mu_1 (1 + V^2)^2} \exp \left(- \frac{\pi w^2}{4 \mu_1^2 (1 + V^2)^2} \right) \dots (14)$$

This constant V in general represents the dispersion of the given or associated distribution in relation to the standard normal distribution.

If $K^2 = \frac{2(1 + V^2)^2}{\pi}$, we get

$$f_2(w) = \frac{1}{\mu_1 K^2} \exp \left(- \frac{w^2}{2 K^2 \mu_1^2} \right) , 0 \leq w < \infty \dots (15)$$

The moment generating function of this distribution is,

$$\begin{aligned} \phi_w(t) &= E(e^{tw}) = \int_0^\infty e^{tw} f_2(w) dw \\ &= \sqrt{\frac{\pi}{2K}} \exp \left(- \frac{K^2 \mu_1^2 t^2}{2} \right) \dots (16) \end{aligned}$$

Expanding the term, $\exp \left(- \frac{K^2 \mu_1^2 t^2}{2} \right)$, with a Taylor's series the moments of the distribution can be obtained.

Returning back to the statistic, smoothed mean absolute dev., (SMAD), the above function can be modified to represent the probability density function for $Q \leq q$ and large value of n .

If $y = aw$, the moment generating function of y will be,

$$\begin{aligned}\phi_y(t) &= E(e^{ty}) \\ &= E(e^{taw}) = \phi_w(at) = \frac{\sqrt{\pi}}{2K} \exp\left(\frac{1}{2} K^2 a^2 \mu_1^2 t^2\right)\end{aligned}$$

from equation (16). Thus the mean of the distribution of the variable $y = aw$ is $a\mu_1$ where μ_1 is the corresponding mean of w alone. And the moment generating function for q_n will be,

$$\begin{aligned}\phi_{q_n}(t) &= E(e^{q_n t}) \\ &= E(\exp(\alpha w_n + \alpha \beta w_{n-1} + \dots + \alpha \beta^n w_0)) \\ &= E(e^{\alpha w_n t} e^{\alpha \beta w_{n-1} t} \dots e^{\alpha \beta^n w_0 t})\end{aligned}$$

and since w_0, w_1, \dots, w_n are all serially uncorrelated, that is, independent and identically distributed,

$$\phi_{q_n}(t) = E(e^{\alpha w_n t}) \cdot E(e^{\alpha \beta w_{n-1} t}) \dots E(e^{\alpha \beta^n w_0 t})$$

Hence,
$$\phi_{q_n}(t) = \phi_w(\alpha t) \cdot \phi_w(\alpha \beta t) \dots \phi_w(\alpha \beta^n t)$$

$$= \left(\frac{\sqrt{\pi}}{2K}\right)^{n+1} \exp \frac{1}{2} \sum_{i=0}^n K^2 (\alpha \beta^i)^2 \mu_1^2 t^2$$

$$= \left(\frac{\sqrt{\pi}}{2K}\right)^{n+1} \exp \frac{1}{2} K^2 t^2 \alpha^2 \left(\sum_{i=0}^n \beta^{2i}\right) \mu_1^2$$

Thus the series $Q_n \leq q_n$ is identically distributed with a mean,

$$\mu_n = \sqrt{\left(\sum_{i=0}^n \alpha^2 \beta^{2i} \right)} \mu_1$$

For a stabilized system,

$$\mu_q = \sqrt{\frac{\alpha}{2-\alpha}} \mu_1 = \nu \mu_1$$

The statistical distribution function of SMAD for $Q \leq q$ can be, thus, written as

$$f_2(q) = \frac{1}{K^2 \nu \mu_1} \exp \left(- \frac{q^2}{2K^2 \nu^2 \mu_1^2} \right), \text{ for } 0 \leq q < \infty.$$

And since $K^2 = \frac{\pi}{2}$ and $\frac{\pi}{2} \mu_1^2 = \sigma^2$ (from eqⁿ (2)) above eqⁿ reduces to,

$$f_2(q) = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\nu \sigma} \exp \left(- \frac{q^2}{2 \nu^2 \sigma^2} \right) \dots (17)$$

TRACKING SIGNAL:

Let us denote the tracking signal by R, and its general value by r.

Then we have $R = P / Q$.

It could be noted that we need to have only previous smoothed statistics and the latest error in order to be able to compute the tracking signal. To find the statistical distribution of $R \leq r$, the following procedure is adopted:

Let us make a transformation from the statistics (p,q) to the statistics (r,s) defined by,

$$r = \frac{p}{q} \quad \text{and} \quad s = q .$$

This allows us to express the joint probability function of (r,s) in terms of that of (p,q) . Integration of this joint distribution function with respect to s over its entire domain will, then, yield the statistical distribution function of $R \leq r$.

Thus, if $T(p,q)$ represents the transformation giving (r,s) , we have,

$$(r,s) = T(p,q)$$

where T is a vector transformation function having two components,

$$r = T_1(p,q) = p / q \quad \text{and} \quad s = T_2(p,q) = q .$$

If $g(r,s)$ represents the joint density function of (r,s) and $f(p,q)$ that of (p,q) , then we have,

$$g(r,s) = \frac{1}{|J(p,q)|} \cdot f(p,q)$$

where the Jacobian $J(p,q)$ is given by,

$$J(p,q) = \begin{vmatrix} \frac{\partial T_1}{\partial p} & \frac{\partial T_2}{\partial p} \\ \frac{\partial T_1}{\partial q} & \frac{\partial T_2}{\partial q} \end{vmatrix}$$

Trigg (5) assumed that the smoothed error and the smoothed mean absolute deviation are independent and uncorrelated. With the same assumption, the joint statistical distribution for (p,q) can be written from equations (11) and (17) as,

$$f(p,q) = \frac{1}{\pi v^2 \sigma^2} \cdot \exp \left(- \frac{1}{2 v^2 \sigma^2} (p^2 + q^2) \right).$$

and the Jacobian is,

$$J(p,q) = \begin{vmatrix} \frac{1}{q} & 0 \\ -\frac{p}{q} & 1 \end{vmatrix} = \frac{1}{q} = \frac{1}{s},$$

which yields, $g(r,s) = s \cdot f(p,q)$.

With a transformation multiplier A incorporated to account for any possible change in the mapping domain, we get,

$$g(r,s) = \frac{A \cdot s}{\pi v^2 \sigma^2} \exp \left(- \frac{1}{2 v^2 \sigma^2} (r^2 + 1) s^2 \right) \dots (18)$$

To obtain the statistical distribution of $R \leq r$ from $g(r,s)$, equation (18) will be integrated over the entire domain of s .

Hence,

$$g(r) = \frac{A}{\pi v^2 \sigma^2} \int_0^{\infty} s \cdot \exp \left(- \frac{1}{2 v^2 \sigma^2} (r^2 + 1) s^2 \right) ds$$

**THIS BOOK WAS
BOUND WITH TWO
PAGES NUMBERED
29. THESE PAGES
ARE THE SAME.**

**THIS IS AS
RECEIVED FROM
CUSTOMER.**

$$= -\frac{A}{\pi(r^2 + 1)} \left[\exp \left(-\frac{1}{2\sigma^2} (r^2 + 1) s^2 \right) \right]_0^\infty$$

$$\text{Hence, } g(r) = \frac{A}{\pi} \frac{1}{1 + r^2}, \quad -\infty < r < \infty \quad \dots (19)$$

The probability density function for $R \leq r$, from equation (19), has a Cauchy form of distribution. From the definition of the tracking signal it should be noted that the range of variation of r will be the closed interval $(-1,1)$. Thus, the multiplier A can be evaluated for this mapping space of the tracking signal, R .

The following points to be noted:

1. The cumulative probability at any point inside the domain should be less than unity,
- and 2. The probability over the entire sample space is one.

These conditions respectively give,

$$G(r) = \int_{-1}^r g(r) dr < 1$$

$$\text{and } \int_{-1}^1 g(r) dr = 1.$$

From the second equation of the above,

$$\frac{A}{\pi} \cdot \left[\arctan r \right]_{-1}^1 = 1$$

$$= -\frac{A}{\pi(r^2 + 1)} \left[\exp \left(-\frac{1}{2\sigma^2} (r^2 + 1) s^2 \right) \right]_0^\infty$$

$$\text{Hence, } g(r) = \frac{A}{\pi} \frac{1}{1 + r^2}, \quad -\infty < r < \infty \quad \dots (19)$$

The probability density function for $R \leq r$, from equation (19), has a Cauchy form of distribution. From the definition of the tracking signal it should be noted that the range of variation of r will be the closed interval $(-1, 1)$. Thus, the multiplier A can be evaluated for this mapping space of the tracking signal, R .

The following points to be noted:

1. The cumulative probability at any point inside the domain should be less than unity,
- and 2. The probability over the entire sample space is one.

These conditions respectively give,

$$G(r) = \int_{-1}^r g(r) dr < 1$$

$$\text{and } \int_{-1}^1 g(r) dr = 1.$$

From the second equation of the above,

$$\frac{A}{\pi} \cdot \left[\arctan r \right]_{-1}^1 = 1$$

and hence, $A = 2$.

Thus, the statistical distribution function of the tracking signal is given by the modified Cauchy distribution function as,

$$g(r) = \frac{2}{\pi} \cdot \frac{1}{1 + r^2}, \quad |r| \leq 1 \quad \dots (20)$$

The statistics of the distribution can readily be found:

$$(a) \text{ Mean } = E(r) = \int_{-1}^1 r \cdot g(r) dr = 0$$

$$\text{and (b) Variance } = E(r^2) = \int_{-1}^1 r^2 g(r) dr = 0.273$$

It can be concluded from this that if the errors are distributed normally around zero mean with a fixed variance, the tracking signal is also distributed symmetrically about zero. However, the distribution is independent of the forecasting system, and has a fixed variance. In other words, the tracking represents the behavior of the errors, irrespective of the system they occur in, and in doing so, scales down the error distribution to a small range.

PHYSICAL SIGNIFICANCE OF THE RESULT:

The above result presents numerous advantages for adoption of tracking signal as a measure of allowance of errors in the forecasting system.

First, the irregularities in the errors are reduced and smoothed to a smaller range of variation without distorting the basic nature of their behavior. Thus, the close confidence limits will serve to judge any erratic behavior or intensity of the system errors.

Second, once the error distribution is decided from the available data, there will not be any further necessity to store and accumulate the successive changes in the data, since, once the relative measure is fixed between the original error magnitudes and the tracking signal magnitudes, the same measure will remain effective in judging the accuracy of the forecasting system. This advantage may not be quite apparent as long as the occurrence of the errors remains random and within the decided limits. However, whenever any assignable cause exists, the statistic immediately undergoes the corresponding variation through the comparative measures without actually having to re-evaluate the change in the distribution of the errors.

And third, it is possible to compare the different prediction systems on the same common level and domain of their precision. This facilitates the comparative approach of the accuracy of different systems and the steps towards the betterment.

At this stage, one fact needs to be mentioned. In computing the smoothed statistics for the tracking signal it is assumed that the

same value of the smoothing constant is used in both the cases. This is not absolutely necessary, but it simplifies the matter a great deal. However, it is quite desirable that α used for the smoothed error should be equal to or larger than that used for the smoothed mean absolute deviation, as it is possible otherwise for the tracking signal to exceed the set control limits for very accurate forecasts, which is quite undesirable.

CHAPTER 4

EXPERIMENTAL RESULTS AND CONCLUSIONS

Testing the suitability of the developed theory in practice assumes an important aspect of the research. Accordingly, a computer simulation was programmed to probe the validity of the applications of the research.

The preceding theory is based on the following two major assumptions:

1. All errors are independent and normally distributed with a zero mean,
- and 2. The smoothed error and the smoothed mean absolute deviation statistics are uncorrelated.

Accordingly, the experiment consisted of three phases. The computer subroutines and programs are in Appendix C.

PHASE I:

The forecast system was simulated assuming independent and normally distributed observations having constant mean and variance values. In other words a constant process and normality of the superimposed process noise were assumed. Random normal deviates were generated as observations.

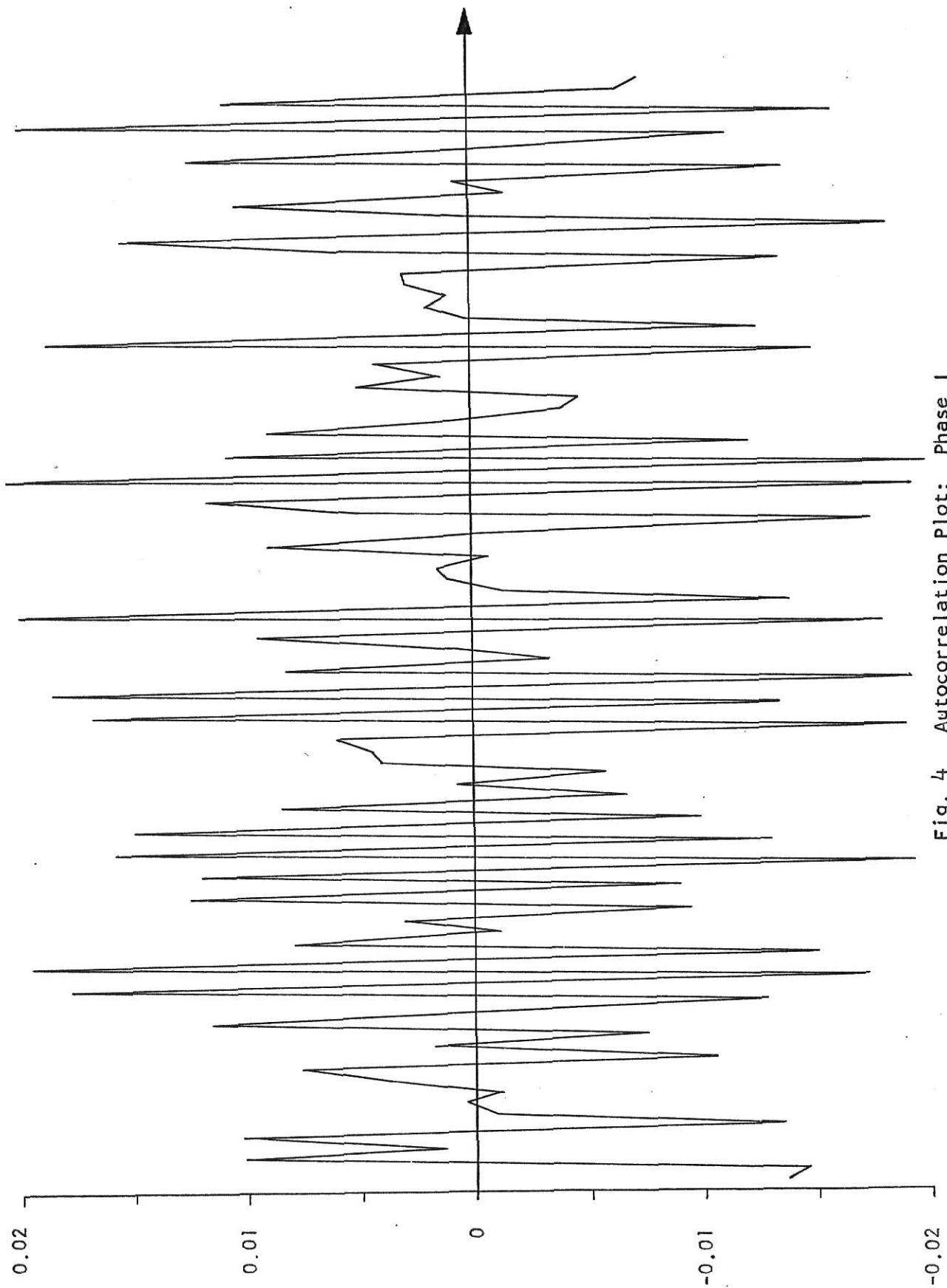


Fig. 4 Autocorrelation Plot: Phase I
No. of lags = 100

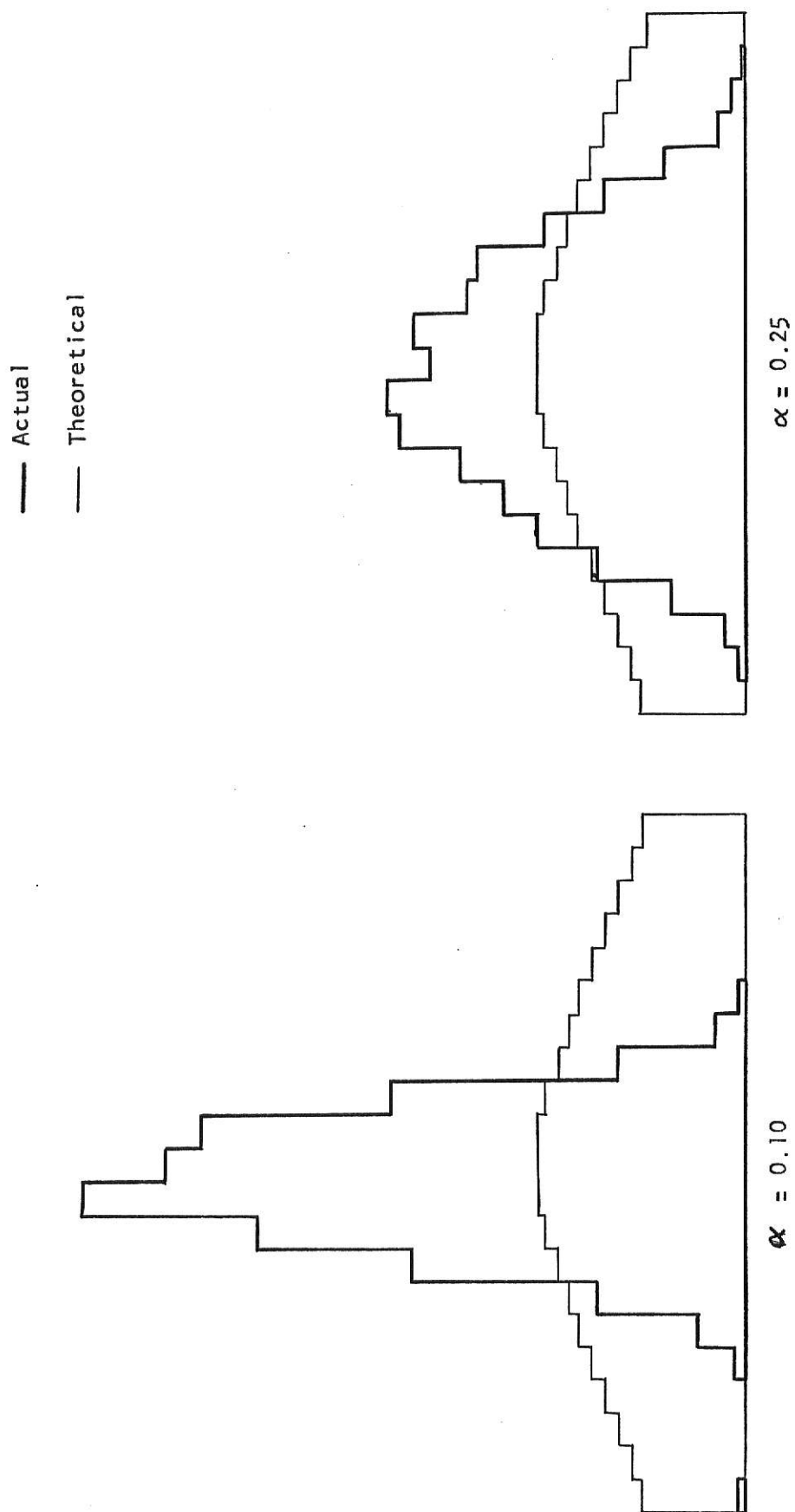


Fig. 5 Histogram: Phase I

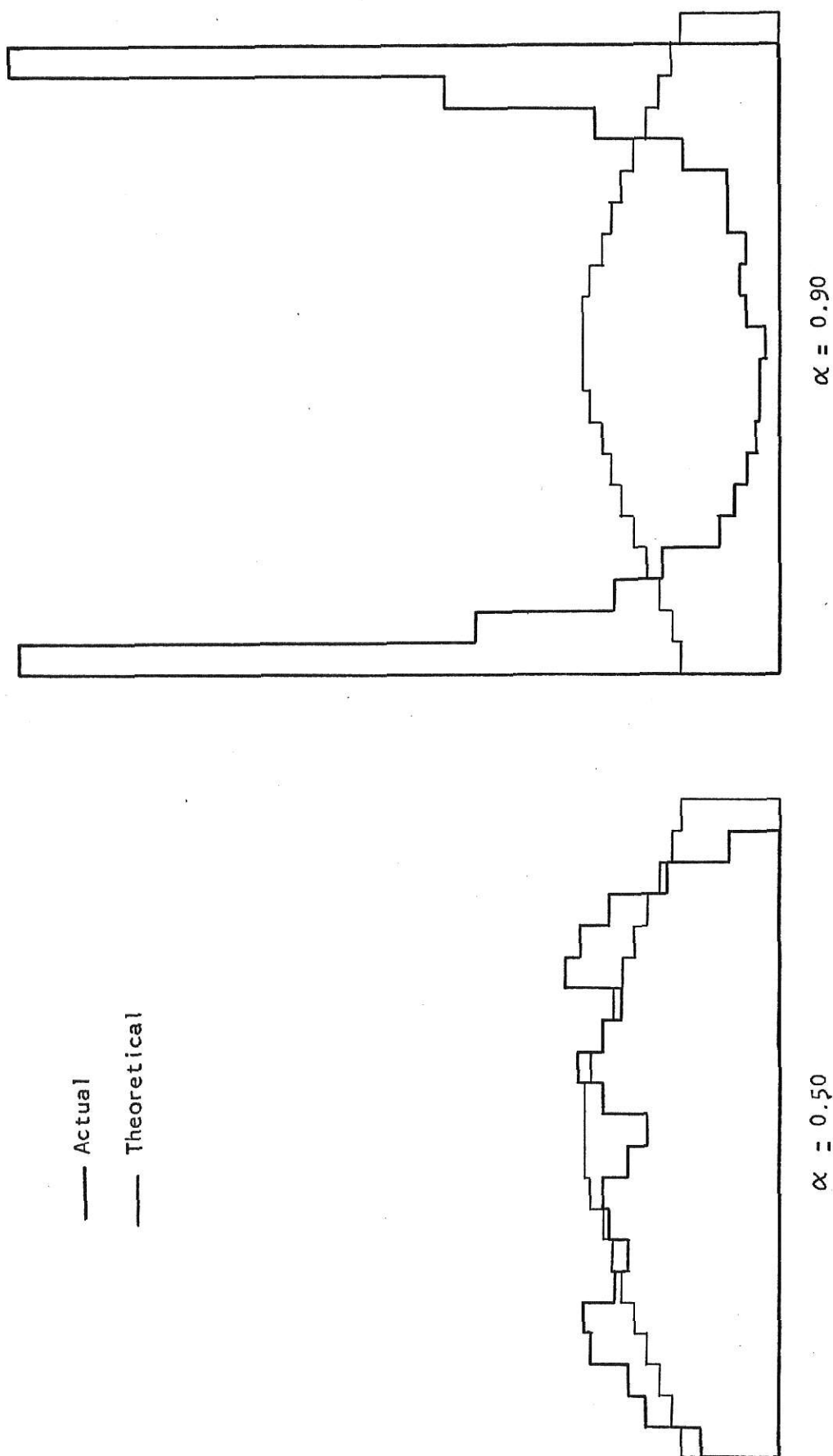


Fig. 6 Histogram: Phase I

Forecasts were generated over a unit lead time using single exponential smoothing. The process was simulated over 2000 observations (steady state) for different values of α . Errors were the difference between the observations and their prior forecasts. The tracking signal statistic was computed for each forecast. An error autocorrelogram and a tracking signal frequency histogram were obtained in each case. Following results were apparent:

1. Forecast errors were serially uncorrelated (i.e., independent and identically distributed). Fig. 4 represents a sample autocorrelogram for this phase of the experiment.
2. Frequency histograms revealed that the tracking signal has a symmetrical distribution with a peaked concentration about its mean for small values of α . For higher values of α , the peak flattens and the distribution spreads over the variational range of the tracking signal. Also, a practically symmetrical bimodality of the distribution is exhibited. For very high value of α ($\alpha = 0.90$), the distribution becomes U-shaped.

Fig. 5,6 illustrate the results of this phase of the experiment in comparison with the theoretical distribution. The results were not par expectation.

— Actual

--- Theoretical

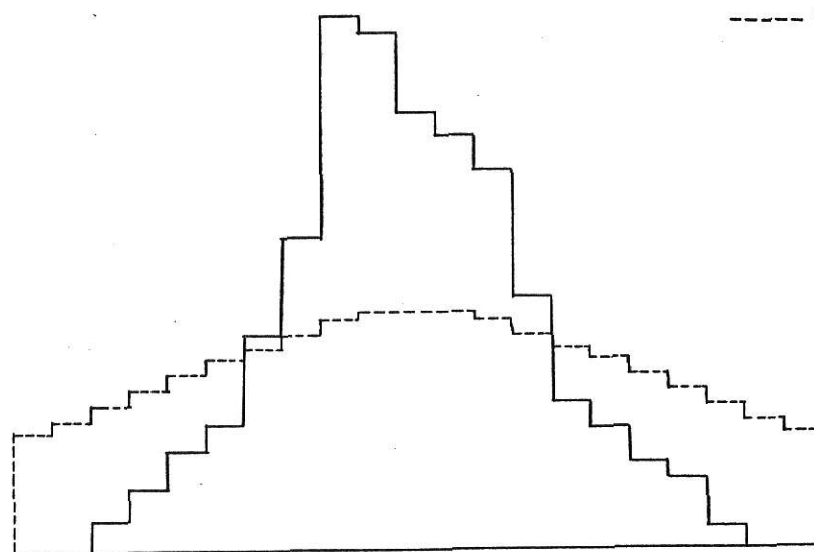
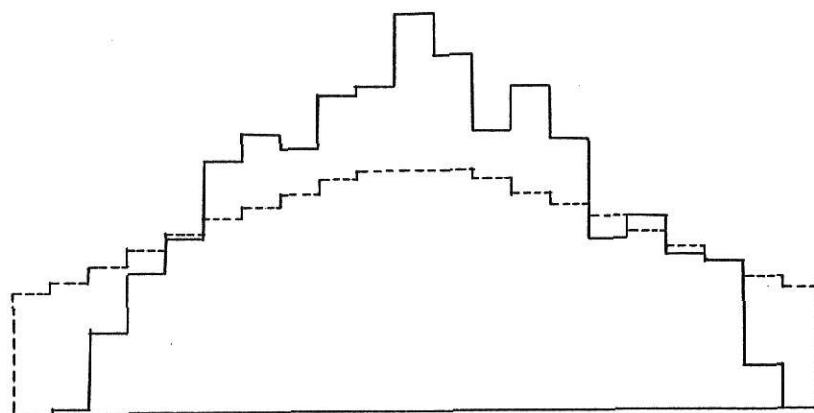
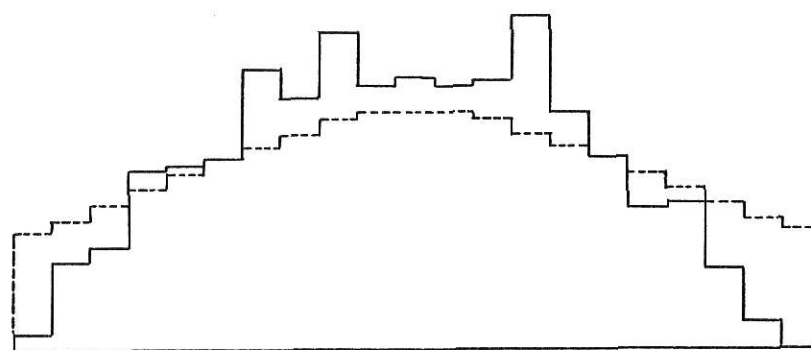
 $\alpha = .10$  $\alpha = 0.15$  $\alpha = 0.20$

Fig. 7 Histogram: Phase II

— Actual

----- Theoretical

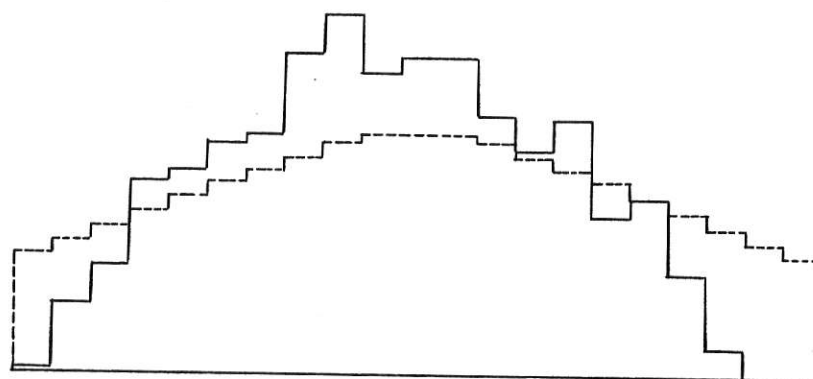
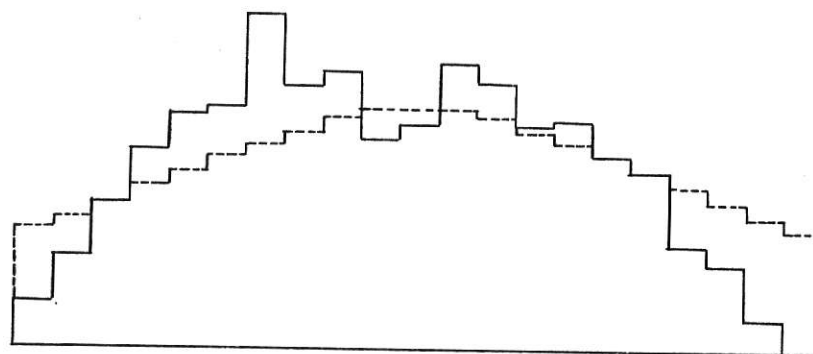
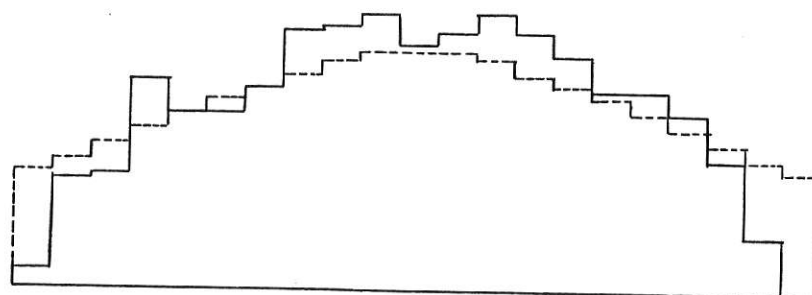
 $\alpha = 0.25$  $\alpha = 0.30$  $\alpha = 0.35$

Fig. 8 Histogram: Phase II

PHASE II:

In this phase, the first assumption was incorporated. Instead of generating the normally distributed observations, normally distributed errors were generated. The process was again simulated over 2000 values for different values of α . Tracking signal frequency histogram was obtained for each value of α . Following results were apparent:

1. Error autocorrelogram showed that the errors were serially uncorrelated as desired.
2. Frequency histograms revealed that the tracking signal has a peaked symmetrical distribution about its mean for small values of α . As α takes higher values, the distribution tends to flatten, its spread widens and a tendency towards bimodality becomes quite apparent.

Fig. 7, 8 illustrate the results of this phase of the experiment in comparison with the theoretical distribution. Results not as expected.

PHASE III:

In this phase, both the theoretical assumptions were incorporated. Two different seeds were provided to the uniform random number generator, and two distinct and independent series of random normal deviates were generated.

— Actual
— Theoretical

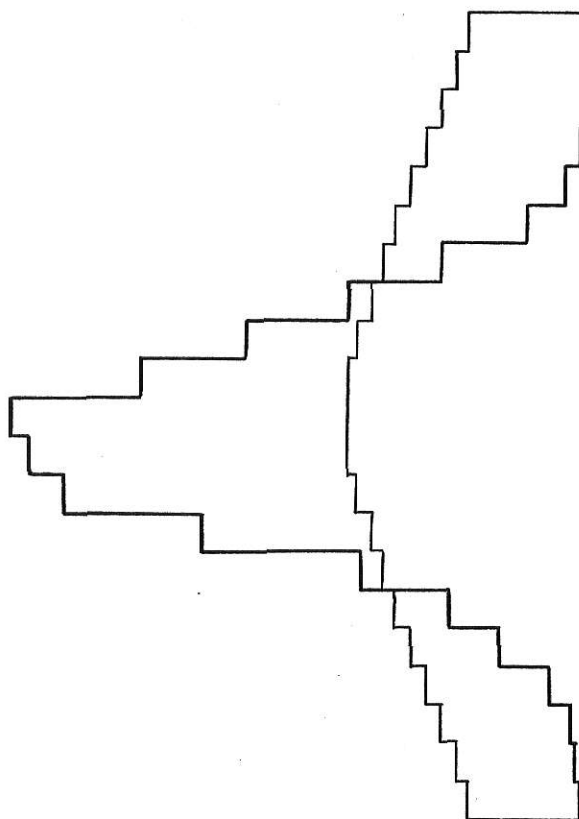
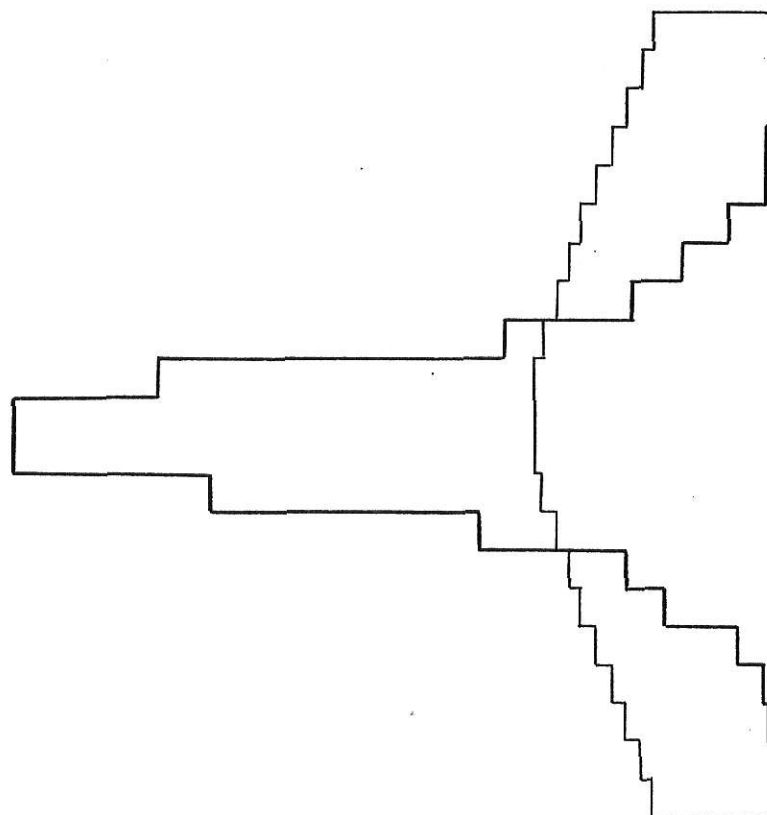
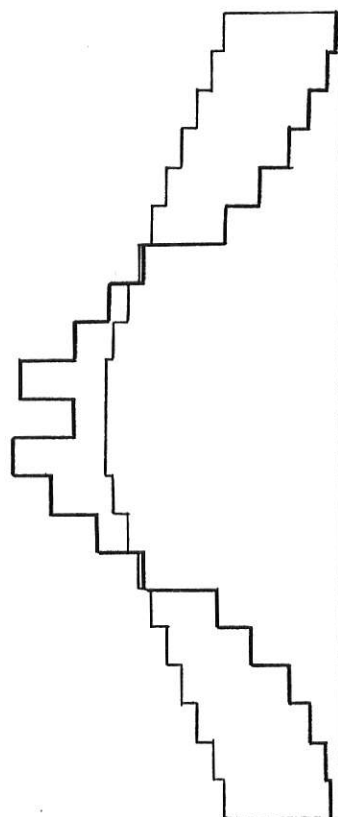
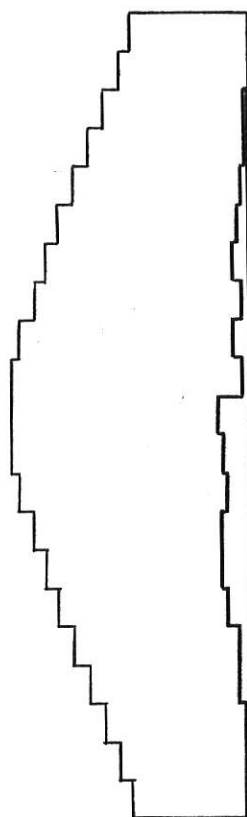


Fig. 9 Histogram: Phase III

— Actual
— Theoretical



$\alpha = 0.15$



$\alpha = 0.20$

Fig. 10 Histogram: Phase III

Both these series of numbers had a zero mean and the same variance value. Each series was simulated to 2000 numbers and was used for obtaining independent smoothed error and smoothed mean absolute deviation statistics respectively. Serial autocorrelogram for each series and the tracking signal frequency histogram were obtained for different values of α . Following are the conclusions of this phase of the experiment.

1. Autocorrelogram of either series showed that the generated numbers were serially uncorrelated as desired.
2. Two series being independent and distinct, the smoothed error statistic and the smoothed mean absolute deviation statistic were uncorrelated.
3. Frequency histograms of the tracking signal had peak values about the mean of the distribution for small values of α . For increased values of α , the frequency histograms showed the flattening and wide-spreading tendency of the distribution. Also, the distribution became bimodal. For higher value of α , the tracking signal statistic exceeded the theoretical limits of $(-1,1)$. This is because the two smooth statistics were not the outcome of the same series of random numbers, and hence the above limits will not be valid.

Fig. 9, 10 present the results of this phase of the experiment in comparison with the theoretical distribution.

The results were not par expectation.

This gap between the theoretical results and the simulation results can be explained as follows:

For simplicity of the procedure and ease of derivation it is assumed that the successive error values are uncorrelated. However, they are correlated in actual practice. Due to the nature of the adaptive forecasting techniques all the forecasts are correlated. Since the forecast errors are the interactions between the observations and the delayed forecasts, the errors also are serially correlated. Hence, the results of the first phase of the simulation did not agree with the theoretically expected values. Also, the theory assumes the two smoothed statistics to be independent whereas the simulation experiment did not account for this.

The second phase of the experiment dealt with the assumed independence of the forecast errors. But the smoothed statistics were yet correlated. The results obtained were, hence, not as expected.

The third phase of the experiment tried to deal with both the theoretical assumptions. However, the two smoothed statistics generated did not result from the same error values as in the theory. Hence, this phase failed to obtain the expected results.

Uniform Random Number Generator was tested for uniformity and randomness. The results of the test were satisfactory. (See Appendix D.)

Cross-correlation between the smoothed error and the smoothed mean absolute deviation was also checked for various values of ρ in phases I and II. Following table shows the computer results in this respect.

Table: Cross-correlation coefficients between the Smoothed error and the smoothed mean absolute deviation.

Value of ρ	Correlation Coefficients	
	Phase I	Phase II
0.05	0.0811	0.2490
0.10	0.1334	0.2523
0.20	0.1307	0.2530
0.25	0.1332	0.2589
0.40	0.0930	0.2498
0.60	-0.0466	0.2560

There exists practically a constant correlation between the two smoothed statistics p and q when the normal errors were generated.

CONCLUSIONS:

The results of this experiment are similar to the conclusions Trigg (5) arrived at.

1. The distribution of the Tracking Signal depends upon the smoothing parameter α ,
2. For higher values of α , the distribution tends to be bimodal.

However, Trigg's conclusions resulted from the simulation outputs although he decided the confidence limits for the Tracking Signal theoretically. Whereas the smoothed mean absolute deviation was theoretically either approximated as a constant, or assumed to be independent of the smoothed error, the experiment simulated the practical situation wherein these two smoothed statistics are highly correlated. Hence, the confidence limits decided by Trigg may not be valid in their practical applicability.

The present research also brought out the same fact. Further theoretical work in this respect will be applicable if it incorporates the correlation between the smoothed statistics.

REFERENCES:

1. BROWN, R. G., "Smoothing, forecasting and prediction of the discrete time series," Prentice-Hall. 1963.
2. BROWN, R. G., "Decision rules for Inventory Management," Holt, Rinehart & Winston. 1967.
3. CURTISS, J. H., "On the distribution of the quotient of two chance variables," Annals of Mathematical Statistics, vol. 12, 1941.
4. ELANDT, Regina C., "The Folded Normal Distribution: Two methods of estimating parameters from moments," Technometrics, vol. 3, No. 4, Nov. 1961.
5. TRIGG, D. W., "Monitoring a Forecasting System," Operations Research Quarterly, vol. 15, No. 3, 1964.

APPENDICES

APPENDIX A

NORMALITY OF FORECAST ERRORS

The following proof is due to Ronald Howard.

("Smoothing, forecasting and prediction of the discrete time series," R. G. Brown. (1).)

Consider a linear, discrete, time invariant forecasting system. Any such system can be represented by an impulse response function $h(n)$. The input to the system is a time series of observations $u(t)$, the output will be various forecasts $\hat{u}(t)$. Let us denote the output by $v(t)$, which can be represented by the convolution equation,

$$v(t) = \sum_{n=0}^{\infty} h(n) \cdot u(t-n) .$$

The output and the corresponding input are statistically related. An infinite set of joint probability density functions is required to describe the statistics involved; and even if the inputs were statistically independent, the output may exhibit a high order of serial correlations.

Let $f_1(u)$ be the density function of the input amplitude $u(t)$ and $f_2(v)$ be the first order density function of the output $v(t)$. The characteristic function of the input $\phi_u(w)$ is the Fourier transform of the density function.

$$\phi_u(w) = E(e^{-wu}) = \int_{-\infty}^{\infty} e^{-wu} f_1(u) du$$

The first order characteristic function of the output is similarly given by,

$$\phi_v(w) = E(e^{-wv}) = \int_{-\infty}^{\infty} e^{-wv} f_2(v) dv$$

The output characteristic function can be written as,

$$\phi_v(w) = E \left[\prod_{n=0}^{\infty} \exp -(w.h(n).u(t-n)) \right]$$

Since the successive values of $u(t)$ are independent, the expected value of the product may be written as the product of the expected values. Therefore,

$$\phi_v(w) = \prod_{n=0}^{\infty} \phi_u(w.h(n))$$

The first order characteristic function of the output can be obtained by multiplying together the characteristic functions of the input with w replaced by w times the successive values of the impulse response.

The k th moment of v , viz. \underline{v}^k , can be expressed in terms of the derivatives of the characteristic function of v at the origin by,

$$\underline{v}^k = \left. \frac{\partial^k \phi_v(w)}{\partial w^k} \right|_{w=0}$$

The mean of the output function is

$$\underline{v} = \bar{u} \sum_{n=0}^{\infty} h(n)$$

When the input values are independent, the average value of the output is the average value of the input multiplied by the sum of the values of the impulse response. The variance is,

$$\sigma_v^2 = \sigma_u^2 \sum_{n=0}^{\infty} h^2(n)$$

Thus, under the same conditions of independent input values, the variance of the output is proportional to the input variance, the constant of proportionality being the sum of impulse response squares.

Now, let us suppose that $u(t)$ is normally distributed with mean m_u and variance of σ^2 . Then ,

$$\phi_u(w) = \exp - \left(\frac{w^2 \sigma^2}{2} + w.m_u \right)$$

$$\text{and } \phi_v(w) = \exp - \left[\frac{w^2 \sigma_u^2}{2} \sum_{n=0}^{\infty} h^2(n) + w \cdot m_u \sum_{n=0}^{\infty} h(n) \right]$$

Therefore, by comparison of the characteristic functions, it is apparent that the output $v(t)$ is normally distributed with mean,

$$\bar{v} = m_u \sum_{n=0}^{\infty} h(n)$$

$$\text{and variance of } \sigma_v^2 = \sigma_u^2 \sum_{n=0}^{\infty} h^2(n) .$$

Since the output $v(t)$ has been shown to be the sum of independent samples from $f_1(u)$ weighted by the values of the impulse response, one should expect from the central limit theorem that, for a very large class of input distributions $f_1(u)$, the output $v(t)$ would be approximately normally distributed.

The mean of the noise is zero and the forecasting processes are linear, so that the superposition theorem holds. Hence, the mean of the forecasts will be the expected value of the 'true' process. For simplicity if a case be considered when the input signal is the noise, the mean will be zero and the variance will be the variance of the noise, and no serial correlations. The variance of the forecasts will be proportional to the noise variance, the constant of proportionality being the sum of squares of the impulse

responses. The error variance will be the sum of the noise variance and the forecast variance, since, for the moment, one may reasonably stipulate that the noise samples have no serial correlations.

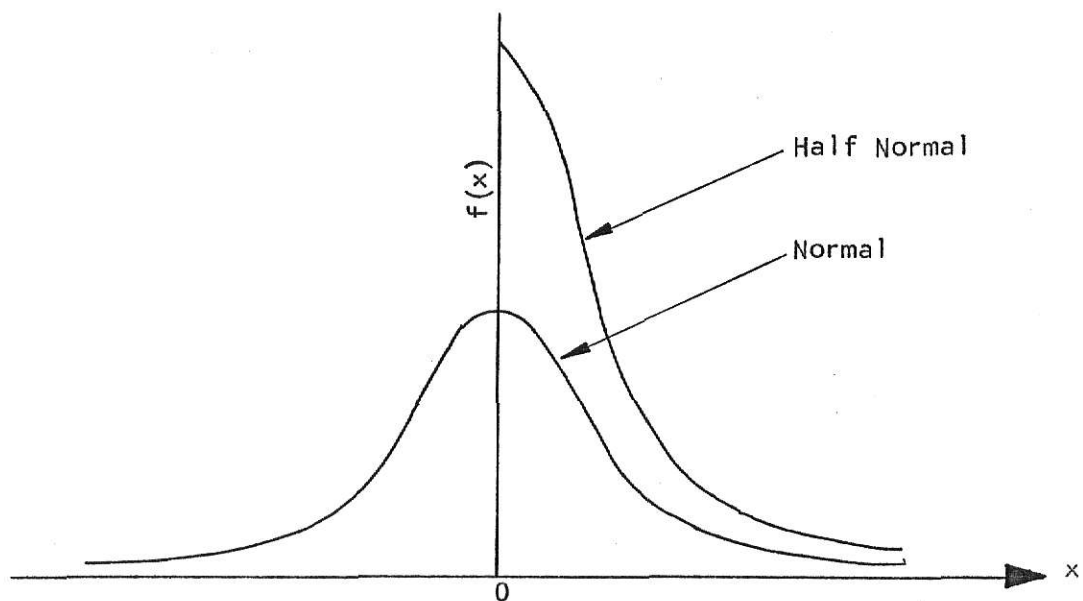
APPENDIX B

FOLDED NORMAL DISTRIBUTION

The following derivation is due to the author.

For other details of the distribution, refer to "Folded Normal Distribution" by R. C. Elandt. (4)

If a random variable X is distributed normally with zero mean and variance of σ^2 , a variable W defined by $W = |X|$, has a folded normal distribution. In fact, any normal variable gives rise to a folded normal variable if its domain be truncated. To explain mathematically, if $X \sim N(\mu, \sigma^2)$, $-a \leq x \leq b$; another random variable defined as $W = |X - k|$ has a folded normal distribution, the original distribution being folded along $x = k$, $-a \leq k \leq b$. The case presently under consideration is a special case called 'Half Normal Distribution.'



A general equation to this type of distribution curve can be written as,

$$f(w) = C \cdot \exp(-D \cdot w^2) ,$$

where $W \leq w$. C and D are constants in terms of distribution statistics. To decide the exact form of the distribution, C and D need be determined.

If μ_1 is the mean of the half normal distribution and σ_1^2 the variance,

$$\begin{aligned} \mu_1 &= \int_0^{\infty} wC \exp(-Dw^2) dw \\ &= C \left[-\frac{1}{2D} \exp(-Dw^2) \right]_0^{\infty} \\ &= \frac{C}{2D} \end{aligned} \quad \dots (A)$$

$$\text{Also, } \sigma_1^2 = \int_0^{\infty} w^2 C \exp(-Dw^2) dw - \mu_1^2$$

$$\text{Thus, } \mu_1^2 + \sigma_1^2 = E(w^2) = \int_0^{\infty} w^2 C \exp(-Dw^2) dw \dots (B)$$

To solve the above equation, consider the following integral,

$$I = \int_0^{\infty} \exp(-Dx^2) dx .$$

Since this is independent of the integration variable, we can write

$$I^2 = \int_0^{\infty} \int_0^{\infty} \exp(-D(x^2 + y^2)) dx dy$$

Changing the rectangular coordinates to the polar coordinates

$$\begin{array}{l} x = r \cos \theta \\ y = r \sin \theta \end{array} \quad \left| \begin{array}{l} \text{---} \\ \text{---} \end{array} \right. \quad 0 \leq r < \infty, 0 \leq \theta \leq \frac{\pi}{2} .$$

And the double integration can be written as,

$$\begin{aligned} I^2 &= \int_0^{\frac{\pi}{2}} \int_0^{\infty} \exp(-Dr^2) r dr d\theta \\ &= \frac{\pi}{4D} . \end{aligned}$$

Hence, the integral, $I = \int_0^{\infty} \exp(-Dx^2) dx = \sqrt{\frac{\pi}{4D}} \dots (C)$

Now differentiating both sides of the equation (C) above with respect to D we get,

$$\begin{aligned} \frac{d}{dD} \int_0^{\infty} \exp(-Dx^2) dx &= \frac{d}{dD} \sqrt{\frac{\pi}{4D}} , \\ \text{or } \int_0^{\infty} -x^2 \exp(-Dx^2) dx &= -\frac{1}{4D} \sqrt{\frac{\pi}{D}} . \end{aligned}$$

And, $E(x^2) = \frac{C}{4D} \sqrt{\frac{\pi}{D}}$, which yields ,

$$\sigma_1^2 = \frac{c}{4D} \sqrt{\frac{\pi}{D}} - \frac{c^2}{4D^2} \quad \dots (D)$$

Solving equations (A) and (D) above for C and D we get,

$$C = \frac{\pi \mu_1^3}{2 (\mu_1^2 + \sigma_1^2)^2} \quad .$$

$$D = \frac{\pi \mu_1^2}{4 (\mu_1^2 + \sigma_1^2)^2} \quad .$$

Thus, the probability density function of $w = |x|$, in terms of its own statistics μ_1 and σ_1^2 , is given as

$$f(w) = \frac{\pi \mu_1^3}{2 (\mu_1^2 + \sigma_1^2)^2} \exp \left(- \frac{\pi \mu_1^2 w^2}{4 (\mu_1^2 + \sigma_1^2)^2} \right) \quad .$$

APPENDIX C

COMPUTER PROGRAMS AND SUBROUTINES

1. Main Program : Phase I
2. Main Program : Phase II
3. Main Program : Phase III
4. Subroutine PARZN
Parzen routine to compute autocorrelation coefficients
5. Subroutine SPLOT
Autocorrelogram plot routine
6. Subroutine RANDU
Uniform random number generator
7. Subroutine RTNORM
Random Normal variate generator
8. Subroutine RANORD
Random Normal variate generator
9. Program for testing the Uniform random number generator
Main and subroutine INT
10. Program for crosscorrelation coefficient between p and q

ILLEGIBLE

**THE FOLLOWING
DOCUMENT (S) IS
ILLEGIBLE DUE
TO THE
PRINTING ON
THE ORIGINAL
BEING CUT OFF**

ILLEGIBLE

```
0001      DIMENSION FREQ(30),X(2000),XHAT(2000),ERR(2000)
0002      DIMENSION SSE(2000),SMAD(2000),TRSIG(2000)
0003      READ 9,IY
0004      9 FORMAT(I9)
0005      READ 10, MEAN,STDDEV,NCOUNT
0006      10 FORMAT(15,F7.1,18)
0007      DO 202 M = 1,6
0008      READ 11,ALPHA
0009      11 FORMAT (F4.2)
0010      BETA = 1. - ALPHA
0011      SSE(1) = 0.
0012      SMAD(1) = 0.7979*STDDEV
0013      TRSIG(1) = 0.
0014      XHAT(1) = MEAN
0015      SW = 0.
0016      WRITE (3,18)ALPHA,STDDEV
0017      18 FORMAT (1H1,' ALPHA = ' F4.2,10X ' STD DEV = 'F6.1)
0018      DO 14 J = 1,25
0019      14 FREQ(J) = 0.
0020      DO 201 I = 1,NCOUNT
0021      X(I) = RTNORM(STDDEV,MEAN,IY,SW)
0022      ERR(I) = X(I) - XHAT(I)
0023      Z = ABS(ERR(I))
0024      SSE(I+1) = ALPHA*ERR(I) + BETA*SSE(I)
0025      SMAD (I+1) = ALPHA*Z + BETA*SMAD(I)
0026      TRSIG(I+1) = SSE(I+1)/SMAD(I+1)
0027      XHAT(I+1) = ALPHA*X(I) + BETA*XHAT(I)
0028      IM = 11 + 10*TRSIG(I+1)
0029      201 FREQ(IM) = FREQ(IM) + 1.
0030      DO 15 J = 1,21
0031      16 FORMAT (' FREQUENCY(' I2,' ) = ' F6.0)
0032      15 WRITE (3,16) J,FREQ(J)
0033      202 CALL PARZNI(ERR,NCOUNT,NCOUNT,100)
0034      STOP
0035      END
```



```
CC01      DIMENSION FREQ(30),ERR(2000)
CC02      DIMENSION SSE(2000),SMAD(2000),TRSIG(2000)
CC03      READ 9,IY,IZ
CC04      9 FORMAT(2I9)
CC05      READ 10, MEAN,STDDEV,NCOUNT
CC06      10 FORMAT(I5,F7.1,I8)
CC07      DO 202 M = 1,6
CC08      READ 11,ALPHA
CC09      11 FORMAT (F4.2)
CC10      BETA = 1. - ALPHA
CC11      SSE(1) = 0.
CC12      SMAD(1) = 0.7979*STDDEV
CC13      TRSIG(1) = 0.
CC14      SW = 0.
CC15      WRITE (3,18)ALPHA,STDDEV
CC16      18 FORMAT (1H1,' ALPHA = ' F4.2,10X ' STD DEV = 'F6.1)
CC17      DO 14 J = 1,25
CC18      14 FREQ(J) = 0.
CC19      DO 201 I = 1,NCOUNT
CC20      ERR(I) = RTNORM(STDDEV,MEAN,IY,SW)
CC21      Z = ABS(ERR(I))
CC22      SSE(I+1) = ALPHA*ERR(I) + BETA*SSE(I)
CC23      SMAD (I+1) = ALPHA*Z + BETA*SMAD(I)
CC24      TRSIG(I+1) = SSE(I+1)/SMAD(I+1)
CC25      IM = 11 + 10*TRSIG(I+1)
CC26      201 FREQ(IM) = FREQ(IM) + 1.
CC27      DO 15 J = 1,21
CC28      16 FORMAT (' FREQUENCY(' I2,' ) = ' F6.0)
CC29      15 WRITE (3,16) J,FREQ(J)
CC30      202 CALL PARZNI(ERR,NCOUNT,NCOUNT,100)
CC31      STOP
CC32      END
```

```
0C01      DIMENSION FREQ(30),ERR(2000),GAP(2000)
0C02      DIMENSION SSE(2000),SMAD(2000),TRSIG(2000)
0C03      READ 9,IY,IZ
0C04      9 FORMAT(2I9)
0C05      READ 10, MEAN,STDDEV,NCOUNT
0C06      10 FORMAT(I5,F7.1,I8)
0C07      DO 202 M = 1,6
0C08      READ 11, ALPHA
0C09      11 FORMAT (F4.2)
0C10      BETA = 1. - ALPHA
0C11      SSE(1) = 0.
0C12      SMAD(1) = 0.7979*STDDEV
0C13      TRSIG(1) = 0.
0C14      SW = 0.
0C15      WRITE (3,18)ALPHA,STDDEV
0C16      18 FORMAT (1H1,' ALPHA = ' F4.2,10X ' STD DEV = 'F6.1)
0C17      DO 14 J = 1,25
0C18      14 FREQ(J) = 0.
0C19      DO 201 I = 1,NCOUNT
0C20      GAP(I) = RANORD(STDDEV,MEAN,IZ,SW)
0C21      ERR(I) = RTNORM(STDDEV,MEAN,IY)
0C22      Z = ABS(ERR(I))
0C23      SSE(I+1) = ALPHA*GAP(I)+BETA*SSE(I)
0C24      SMAD (I+1) = ALPHA*Z + BETA*SMAD(I)
0C25      TRSIG(I+1) = SSE(I+1)/SMAD(I+1)
0C26      IM = 11 + 10*TRSIG(I+1)
0C27      201 FREQ(IM) = FREQ(IM) + 1.
0C28      DO 15 J = 1,21
0C29      16 FORMAT (' FREQUENCY(' I2,' ) = ' F6.0)
0C30      15 WRITE (3,16) J,FREQ(J)
0C31      CALL PARZN1 (GAP,NCOUNT,NCOUNT,200)
0C32      202 CALL PARZN1 (ERR,NCOUNT,NCOUNT,200)
0C33      STOP
0C34      END
```

```

0001      SUBROUTINE PARZN1(X,N,NQ,M)
0002      1 FORMAT(2I4)
0003      2 FORMAT(4X,F5.0,4X,F5.0)
0004      3 FORMAT(F12.5,12X,F12.5)
0005      DIMENSION X(1)
0006      REAL R1(400)/400*0.0/
0007      DATA D1,D3/0.0,0.0/
0008      MM=0
0009      DO 5 I=1,N
0010      D1=D1+X(I)**2
0011      5 CONTINUE
0012      MM=M+1
0013      DO 7 KK=1,MM
0014      KK1=KK-1
0015      NM=N+KK-1
0016      NK=N-KK+1
0017      SUM1=0.0
0018      DO 6 JL=1,NK
0019      SUM1=SUM1+X(JL)*X(JL+KK1)
0020      6 CONTINUE
0021      R1(KK)=SUM1/D1
0022      7 CONTINUE
0023      CALL SPLOT (M,R1,1)
0024      RETURN
0025      END
    
```

```

C
0001      SUBROUTINE SPLOT(N,X,NC)
C
C          THIS ROUTINE IS USED TO MODIFY THE RESULTS FROM THE PARZE
C          SO THAT THEY CAN BE GRAPHED BY SUBROUTINE PLOT.
C
0002      DIMENSION  X(2000),OUT(501),YPR(50),ANG(10),PLT(4000),Y(2000)
0003      102 FORMAT(1H ,' AUTOCORRELATION PLOT ',(/8F14.5))
C
C          INITIALIZE VALUES
0004      PI=3.1415927
0005      K=0
0006      M=2
0007      IF(NO.NE.3)      GO TO 10
0008      DO 106 I=1,N
0009      IF(X(I))100,100,99
0010      99 X(I)=ALOG(X(I))
0011      GO TO 106
0012      100 X(I)=-10.0
0013      106 CONTINUE
0014      DO 8 I=1,N
0015      PLT(I)=I*PI/N
0016      PLT(I+N)=X(I)
0017      8 CONTINUE
0018      10 DO 11 I=1,N
0019      PLT(I)=I
0020      PLT(I+N)=X(I)
0021      11 CONTINUE
0022      WRITE(3,102)  (X(I),I=1,N)
0023      RETURN
0024      END

```

```
0001      FUNCTION RANDU(IY)
0002      IY = IY*97119
0003      IF(IY.LT.0) IY = IY + 2147483647 + 1
0004      RANDU = IY * .4656613E-9
0005      RETURN
0006      END
```

**THIS BOOK
CONTAINS
NUMEROUS
PAGES WITH
MULTIPLE
PENCIL MARKS
THROUGHOUT
THE TEXT.**

**THIS IS THE
BEST IMAGE
AVAILABLE.**

```
0001      FUNCTION RTNORM(STDDEV,MEAN,IB,SW)
0002      IF (SW) 1,1,2
0003      1 U = SQRT(-ALOG(RANDU(IB)))
0004      U2 = 6.2832 * RANDU(IB)
0005      X1 = U*COS(U2)
0006      X2 = U*SIN(U2)
0007      SW = 1.
0008      3 RTNORM = X1*STDDEV + MEAN
0009      RETURN
0010      2 X1 = X2
0011      SW = 0.
0012      GO TO 3
0013      4 RETURN
0014      END
```

using !

```
0C01      FUNCTION RANORD(STDDEV,MEAN,IA,SW)
0C02      IF (SW) 1,1,2
0C03      1 V = SQRT(-ALOG(RANDU(IA)))
0C04      W = 6.2832 * RANDU(IA)
0C05      Y1 = V * COS(W)
0C06      Y2 = V * SIN(W)
0C07      SW = 1.
0C08      3 RANORD = Y1*STDDEV + MEAN
0C09      RETURN
0C10      2 Y1 = Y2
0C11      SW = 0.
0C12      GO TO 3
0C13      4 RETURN
0C14      END
```

works


```
0001      DIMENSION ICEL(100),LCEL(20,20),KCEL(50),IEXT(50)
0002      READ(1,2)NNO,ICLS,NCEL,KCLS,IX,NEXP,IRNJ
0003      2 FORMAT(7I5)
0004      CHI1=0.0
0005      CHI2=0.0
0006      CALL INT(ICEL,ICLS)
0007      CALL INT(KCEL,KCLS)
0008      CALL INT(IEXT,KCLS)
0009      DO 1 K=1,NCEL
0010      DO 1 N=1,NCEL
0011      LCEL(K,N)=0
0012      1 CONTINUE
0013      RJ=FLOAT(KCLS)/FLOAT(IRNJ)
0014      DO 15 IN=1,NNO
0015      R=RANDU(IX)
0016      JK=R*ICLS+1
0017      ICEL(JK)=ICEL(JK)+1
0018      IF(IN-1)20,20,30
0019      20 K=R*NCEL+1
0020      30 J=R*NCEL+1
0021      LCEL(K,J)=LCEL(K,J)+1
0022      15 K=J
0023      TRK=FLOAT(NNO)/FLOAT(ICLS)
0024      DO3J=1,ICLS
0025      CHI1=CHI1+(ICEL(J)-TRK)**2
0026      3 CONTINUE
0027      CHI1=CHI1/TRK
0028      TRIK=FLOAT(NNO-1)/FLOAT(NCEL**2)
0029      DO 5 M=1,NCEL
0030      DO 5 N=1,NCEL
0031      5 CHI2 = CHI2 + (LCEL(M,N) - TRIK)**2
0032      CHI2=CHI2/TRIK
0033      WRITE(3,8)
0034      8 FORMAT(' UNIFORM DISTRIBUTION IN (0,1)',//)
0035      DO6M=1,5
0036      MM=20*(M-1)+1
0037      ML=MM+19
0038      6 WRITE(3,9) (ICEL(J),J=MM,ML)
0039      9 FORMAT(20I6)
0040      WRITE(3,10)CHI1
0041      10 FORMAT(' -COMPUTED CHI-SQ VALUE =',F10.3)
0042      WRITE(3,11)
0043      11 FORMAT('12-D CHI-SQ TEST FOR SERIAL CORRELATION',//)
0044      DO7I=1,NCEL
0045      WRITE(3,9)(LCEL(I,J),J=1,NCEL)
0046      7 CONTINUE
0047      WRITE(3,10)CHI2
0048      STOP
0049      END
```

```
0001      SUBROUTINE INT(K,L)
0002      DIMENSION K(1)
0003      DO1I=1,L
0004      K(I)=0
0005 1 CONTINUE
0006      RETURN
0007      END
```

APPENDIX D

CHI - SQUARE TEST FOR RANDOM NUMBER
GENERATOR.UNIFORMITY TEST :

Since the generated random numbers are uniformly distributed, each number in the generated random series should have an equal probability of occurrence. The theoretical frequency histogram will be a rectangle. The actual histogram is compared with the theoretical and Chi-square statistic is computed for the goodness of fit. The whole uniform interval is divided in n subintervals and the Chi-square statistic is,

$$\chi^2 = \sum_{i=1}^n \frac{(f_{oi} - f_{ei})^2}{f_{ei}}$$

where, f_{oi} = observed frequency for the i^{th} subinterval,
 f_{ei} = expected (theoretical) frequency for the i^{th} subinterval.

If the chi-square value thus obtained is less than the critical Chi-square value obtained from the tables, the uniformity of the generated random numbers is satisfactory.

The numbers generated by the subroutine (RANDU) were divided into 100 equal subintervals. The uniformity was tested for 99% confidence. Thus, for this test the degrees of freedom = 99, and the critical Chi-square value = 123.230. The computed Chi-square goodness-of-fit statistic = 72.800. Hence, the uniformity of the numbers generated by the Random Number Generator routine was satisfactory.

RANDOMNESS TEST :

In essence, this test checks for uniformity of the successive lagged pairs of random numbers. Also, it checks only high order decimal digits of the random numbers for pairwise uniformity.

As in the uniformity test the interval is subdivided into n subintervals. A k -by- k matrix is formed by determining the frequency of numbers $f_{oi,j}$ that are in the i^{th} subinterval followed by a number in the j^{th} subinterval (i and j referring to row number and column number of the matrix). If N is the length of the random sequence, and ν is a power of 2, the number of pairs (x_i, x_{i+1}) , $i = 1, 2, \dots, N-1$, is determined for which the first $(\log \nu)$ bits of x_i had the value m and the first $(\log \nu)$ bits of x_{i+1} had the value n ; n, m assume values from 0 to $\nu-1$. Chi-square statistic is,

$$\chi^2 = \sum_{m,n=0}^{\nu-1} \left(f_{n,m} - \frac{N}{\nu^2} \right)^2 / \left(\frac{N}{\nu^2} \right)$$

The degrees of freedom for the critical Chi-square value is $\nu^2 - 1$. For the test, 20 subintervals were formed, $\nu = 19$. Degrees of freedom = 360. 99% confidence range was used. Critical Chi-square value = 432.00. Computed Chi-square value = 349.768. Hence, the numbers are satisfactorily random. This showed that the Random Number Generator was satisfactory.

For details, refer to P. A. W. Lewis, A. S. Goodman and J. M. Miller, "A Pseudo-random Number Generator for System/360," IBM Systems Journal, Vol. 8, No. 2, 1969.

MATHEMATICAL INVESTIGATION OF THE TRACKING SIGNAL
IN THEORY OF FORECASTING

by

PRASAD PATHARE

B. Eng. (Mech.) (Hons.), University of Bombay, 1965

B. Eng. (Elec.) (Hons.), University of Bombay, 1966

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1971

ABSTRACT

The thesis dealt with the statistical distribution of the Tracking Signal which was conceptualized to smooth the forecast errors and map them to a fixed and a smaller variational range. It was proposed to investigate the statistical behavior of the Tracking Signal. This, in turn, would facilitate determination of appropriate control limits to check on the magnitude of forecast error.

The procedure assumed,

- a. normality and independence of the forecast errors,
- b. independence of the smoothed error and smoothed mean absolute deviation.

The theoretical results indicated that,

1. the Tracking Signal statistic has a modified Cauchy distribution,
2. the distribution was independent of the smoothing parameter.

Simulation experiment was carried out to test the practical validity of the theoretical development. The experiment simulated a steady state forecasting system. The conclusions were:

- a. The Tracking Signal distribution is dependent upon the smoothing parameter.
- b. The distribution tends to be bimodal as α increases.
- c. The theoretical assumptions were not justifiably valid in practical situations.