Modeling of Atomic Clock Performance and Detection of Abnormal Clock Behavior

William A. Ganter
Department of Engineering Design and Economic Evaluation
University of Colorado

Final Report on Research Sponsored by Time and Frequency Division
Institute for Basic Standards
National Bureau of Standards
Boulder, Colorado 80302

NBS Technical notes are designed to supplement the Bureau's regular publications program. They provide a means for making available scientific data that are of transient or limited interest. Technical Notes may be listed or referred to in the open literature.

U.S. DEPARTMENT OF COMMERCE, Frederick B. Dent, Secretary
NATIONAL BUREAU OF STANDARDS, Richard W. Roberts, Director

Issued March 1973

National Bureau of Standards Technical Note 636

Nat. Bur. Stand. (U.S.), Tech. Note 636, 39 pages (March 1973)
CODEN: NBTNAE

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D. C. 20402
(Order by SD Catalog No. C13.46:636)

Contents

List of Figures

List of Figures (cont'd.)

List of Tables

Modeling of Atomic Clock Performance and Detection of Abnormal Clock Behavior

William A. Ganter

We have assumed that the nominal performance of an atomic clock can be well characterized by (1) a noise amplitude and (2) a mix of white and flicker pure noise processes. A number of specific kinds of changes are assumed that a clock might encounter. We assume that these changes can occur with either sign and with varying magnitudes. The changes considered are a noise amplitude different from nominal, a flicker component in the noise mix which is different from nominal, a step jump in the time counts for a clock and a linear (frequency jump) or quadratic (frequency drift) trend in the time counts for a clock.

Detection of a change is accomplished with a multiple sequential test having compound limits. The test is designed to respond quickly to an actual change but to make few incorrect detections (identify the wrong change) or false detections (when no change from nominal operation has occurred). When a change is detected for a clock, the laboratory time scale can be adjusted accordingly for this condition.

Key words: Atomic clock model; detection; flicker noise; sequential test; time scale.

1. Introduction

A recent paper by Allan, Gray, and Machlan [1] gives a comprehensive view of the AT(NBS) time scale. These authors discuss the influence of an individual commercial cesium beam clock on this time scale. Of special interest for this report is a section of [1] on clock modeling where the belief is expressed that a clock can experience changes in its rate due to either internal or external perturbations. In addition, the interpretation of "time" from a clock is subject to counting noise. A prominent model of clock noise is a mix of two kinds of noise: white FM and flicker FM noise processes. These noise processes are discussed in [2], [3], and [4].

In [1] the authors disclose that a weight w_i is used to incorporate measured counts from clock i into the AT(NBS) scale. They state that this weight depends upon the quality of performance that clock i is expected to give. This means that the intensity and the kinds of noise processes affect the weighting. The assumption of the quality of performance for clock i must be derived from a comparison of clock i with other clocks or from its direct calibration with a primary standard.

Under the assumption that the noise process of a clock might change or that its rate might change, this research addresses the problem of detecting changes in the performance of a clock. The noise changes considered are intensity (amplitude) changes in either the white FM or flicker FM component. The other changes are a jump in time, a drift in time, or a drift in rate. The objective is to explore detection schemes which have the following properties:

1. Make few false detections;
2. Ignore very minute changes;
3. Respond rapidly to large changes; and
4. Identify the type of change correctly.

Unfortunately these desired characteristics are in conflict in a number of ways. Very rapid detection of a change causes more false detections. Trying to detect very small changes causes more false detections. Trying to distinguish between many changes causes more detections for the wrong reason. Therefore, a good detection scheme is one which maintains a suitable balance between the goals of the detection procedure. It is also one which responds quickly to actual changes but makes only an acceptable number of false detections. A design goal for the procedures developed in this project was to make a false detection on the average only once per year with daily testing for changes.

Detection schemes are used in many other applications. Two very prominent examples are quality control and submarine warfare. Other examples are early failure detection for operating machinery and aircraft engines, medical diagnosis, and environmental control. Graphs comparing the false detection frequency and the speed of detection of various magnitudes of changes are often called operating characteristic curves. Another related notion is that machines are conceived to operate in a particular manner by their designers. Often a failure of some type occurs causing abnormal performance from the machine. A detection scheme can be devised to recognize non-nominal behavior when it is important to be aware of such conditions.

The statistical techniques used in this research are moving averages, geometric smoothing, estimation, prediction, likelihood functions, and comparisons of quantities against limits. The procedures (which are really hypothesis tests) are multiple sequential tests using compound limits to better distinguish between possible changes. Geometric smoothing is often employed to require a persistence of evidence before a detection is made.

In section 2, we discuss the noise model of a clock that we have used in this study. Section 3 shows the structure of the multiple detection schemes which resulted from the experimental development using simulation of a clock. The bulk of the effort in this study was devoted to the development of a multiple test which was consistent with the design goals. Section 4 discusses this experimental development and then discusses the results of an experimental evaluation of the procedures of section 3. Section 5 gives conclusions and recommendations for additional study. No attempt was made to specify any adjustments to be made to a time scale after a change was detected. A change might suggest a revision of the w_i weights or a one time increment to be applied to the scale.

2. Noise Model

The noise process about the rate of an atomic clock considered in this study has a white noise FM component and a flicker noise FM component. The white FM process has constant spectral density of the frequency fluctuation while the flicker process has a spectral density of the frequency fluctuation inversely proportional to the Fourier frequency. Let us consider other properties of these stochastic processes, the autocovariance function and the Allan variance. The Allan variance is specified as a sample statistic (time averaging of a process realization) in [2], [3], and [4]. It will be convenient to view this measure of frequency stability in its expectation form in this section.

Let W_t for $t = 0, 1, \ldots, \infty$ be a white noise stochastic process. Let t be a day index; however, other counting intervals are equally valid. Let us also denote a realization of the W_t process by w_t. We also assume that random variable W_t is normally distributed with mean 0 and unit variance. The process outcomes at various days are independent, thus the autocovariance function

$$A_W(\tau) = \mathrm{Cov}(W_t, W_{t+\tau}) = 0, \text{ for } \tau > 0,$$

and (1)

$$A_W(0) = \mathrm{Var}(W_t) = 1.$$

Let $\sigma^2_W(\tau)$ denote the Allan variance for the W_t noise process. We define

$$\sigma^2_W(1) = \frac{E[(W_t - W_{t+1})^2]}{2}. \tag{2}$$

This expectation expands to

$$\frac{E[W_t^2] - 2E[W_t W_{t+1}] + E[W_{t+1}{}^2]}{2}$$

$$= \frac{E[W_t^2] - 2E[W_t]E[W_{t+1}] + E[W_{t+1}{}^2]}{2}$$

$$= \frac{E[W_t{}^2] + E[W_{t+1}{}^2]}{2} \quad ,$$

since W_t and W_{t+1} are independent random variables. By definition we have that $W_{t+1}{}^2$ is distributed chi-square with 1 degree of freedom; the expected value of a chi-square random variable with 1 degree of freedom is 1. Thus,

$$\sigma_W{}^2(1) = \frac{1+1}{2} = 1 \ .$$

Next consider

$$\sigma_W{}^2(2) = \frac{E\left[\left(\frac{W_t + W_{t+1}}{2} - \frac{W_{t+2} + W_{t+3}}{2}\right)^2\right]}{2}$$

$$= \frac{1+1}{4} - \frac{2(0)}{4} + \frac{1+1}{4} = \frac{1}{2} \quad .$$

In similar fashion it is easily shown that

$$\sigma_W{}^2(\tau) = \frac{1}{\tau} \ . \tag{3}$$

Let F_t be a flicker noise stochastic process. It is shown in [4] that

$$\sigma_F{}^2(\tau) = c, \text{ a constant.} \tag{4}$$

However, the autocovariance function

$$A_F(\tau) \neq 0 \text{ for any } \tau \ . \tag{5}$$

The flicker process is thought to be positively autocorrelated for all τ; this result should be demonstrable using eq (4) and the definition of $A_F(\tau)$ since many similar terms are present. It is necessary to divide

F_t by c so that

$$\sigma_F^2(\tau) = 1 \tag{6}$$

in this model.

Let us define noise stochastic process

$$Z_t = a\ W_t + b\ F_t\ , \tag{7}$$

where a + b = 1. The Allan variance for this process is

$$\sigma_Z^2(\tau) = \frac{a^2}{\tau} + b^2\ . \tag{8}$$

For $\tau = 1$ in eq (8) we see that

$$\sigma_Z^2(1) = \frac{E[(aW_t - aW_{t+1})^2]}{2} + \frac{E[(bF_t - bF_{t+1})^2]}{2}$$

$$= a^2\ \sigma_W^2(1) + b^2\ \sigma_F^2(1)\ .$$

Henceforth, let us denote the Allan variance of the Z_t process by simply $\sigma^2(\tau)$. We will let $\hat{\sigma}^2(\tau)$ be an estimate of the Allan variance from a particular process realization. Let us also define the notation

$$\text{sigma}\ (a, b, \tau) = \sqrt{\frac{a^2}{\tau} + b^2}\ . \tag{9}$$

Table 1 gives values of this function for a and b equal to 0., .1, .2, ..., 1. and τ + 1, 2, 4 and 8.

Let us also define a particular estimate of the Allan variance for the Z_t process with outcomes z_t by

$$\hat{\sigma}_T^2(\tau) = \frac{\sum_{i=1}^{\frac{T}{\tau} - 1} (\bar{z}_i - \bar{z}_{i+1})^2}{2(\frac{T}{\tau} - 1)}\ , \tag{10}$$

where $\bar{z}_i$ is the arithmetic average of the i-th group of τ Z_t outcomes. For comparison, a common estimator for the autocovariance function is

$$\hat{A}_T(\tau) = \frac{1}{T-\tau} \sum_{t=1}^{T-\tau} (z_t - \overline{z})(z_{t+\tau} - \overline{z}), \tag{11}$$

for $\tau = 0, 1, \ldots, T-1$ and where $\overline{z}$ is the average over all T outcomes of Z_t.

TABLE 1. FUNCTION SIGMA (a, b, τ).

sigma (a, b, 1)

a \ b	0.	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0.	0.000	.100	.200	.300	.400	.500	.600	.700	.800	.900	1.000
.1	.100	.141	.224	.316	.412	.510	.608	.707	.806	.906	1.005
.2	.200	.224	.283	.361	.447	.539	.632	.728	.825	.922	1.020
.3	.300	.316	.361	.424	.500	.583	.671	.762	.854	.949	1.044
.4	.400	.412	.447	.500	.566	.640	.721	.806	.894	.985	1.077
.5	.500	.510	.539	.583	.640	.707	.781	.860	.943	1.030	1.118
.6	.600	.608	.632	.671	.721	.781	.849	.922	1.000	1.082	1.166
.7	.700	.707	.728	.762	.806	.860	.922	.990	1.063	1.140	1.221
.8	.800	.806	.825	.854	.894	.943	1.000	1.063	1.131	1.204	1.281
.9	.900	.906	.922	.949	.985	1.030	1.082	1.140	1.204	1.273	1.345
1.0	1.000	1.005	1.020	1.044	1.077	1.118	1.166	1.221	1.281	1.345	1.414

sigma (a, b, 2)

a \ b	0.	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0.	0.000	.100	.200	.300	.400	.500	.600	.700	.800	.900	1.000
.1	.071	.122	.212	.308	.406	.505	.604	.704	.803	.903	1.002
.2	.141	.173	.245	.332	.424	.520	.616	.714	.812	.911	1.010
.3	.212	.235	.292	.367	.453	.543	.636	.731	.828	.925	1.022
.4	.283	.300	.346	.412	.490	.574	.663	.755	.849	.943	1.039
.5	.354	.367	.406	.464	.534	.612	.696	.784	.875	.967	1.061
.6	.424	.436	.469	.520	.583	.656	.735	.819	.906	.995	1.086
.7	.495	.505	.534	.579	.636	.704	.778	.857	.941	1.027	1.116
.8	.566	.574	.600	.640	.693	.755	.825	.900	.980	1.063	1.149
.9	.636	.644	.667	.704	.752	.809	.875	.946	1.022	1.102	1.185
1.0	.707	.714	.735	.768	.812	.866	.927	.995	1.068	1.145	1.225

Table 1. (continued)

sigma (a, b, 4)

a \ b	0.	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0.	0.000	.100	.200	.300	.400	.500	.600	.700	.800	.900	1.000
.1	.050	.112	.206	.304	.403	.502	.602	.702	.802	.901	1.001
.2	.100	.141	.224	.316	.412	.510	.608	.707	.806	.906	1.005
.3	.150	.180	.250	.335	.427	.522	.618	.716	.814	.912	1.011
.4	.200	.224	.283	.361	.447	.539	.632	.728	.825	.922	1.020
.5	.250	.269	.320	.391	.472	.559	.650	.743	.838	.934	1.031
.6	.300	.316	.361	.424	.500	.583	.671	.762	.854	.949	1.044
.7	.350	.364	.403	.461	.532	.610	.695	.783	.873	.966	1.059
.8	.400	.412	.447	.500	.566	.640	.721	.806	.894	.985	1.077
.9	.450	.461	.492	.541	.602	.673	.750	.832	.918	1.006	1.097
1.0	.500	.510	.539	.583	.640	.707	.781	.860	.943	1.030	1.118

sigma (a, b, 8)

a \ b	0.	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0.	0.000	.100	.200	.300	.400	.500	.600	.700	.800	.900	1.000
.1	.035	.106	.203	.302	.402	.501	.601	.701	.801	.901	1.001
.2	.071	.122	.212	.308	.406	.505	.604	.704	.803	.903	1.002
.3	.106	.146	.226	.318	.414	.511	.609	.708	.807	.906	1.006
.4	.141	.173	.245	.332	.424	.520	.616	.714	.812	.911	1.010
.5	.177	.203	.267	.348	.437	.530	.625	.722	.819	.917	1.016
.6	.212	.235	.292	.367	.453	.543	.636	.731	.828	.925	1.022
.7	.247	.267	.318	.389	.470	.558	.649	.742	.837	.933	1.030
.8	.283	.300	.346	.412	.490	.574	.663	.755	.849	.943	1.039
.9	.318	.334	.376	.437	.511	.593	.679	.769	.861	.955	1.049
1.0	.354	.367	.406	.464	.534	.612	.696	.784	.875	.967	1.061

3. Structure of the Detection Procedures

In this section we merely present the detection procedures developed in this research. The rationale behind them is defended in the next section. Let T_t denote the time from a clock at day t; we update time by

$$T_t = T_{t-1} + z_t . \tag{12}$$

Let us first consider the tests for a noise intensity change from eq (7) of either of two types:

$$Z_t = a\ W_t + \beta\ F_t \tag{13}$$

or

$$Z_t = \alpha\ W_t + b\ F_t , \tag{14}$$

where α and β are changed values of a and b, respectively.

An estimate of $\sigma(\tau)$ is obtained by letting T = 16 in eq (10). Every 5 days thereafter let our estimate of $\hat{\sigma}(\tau)$ be replaced by

$$.95\ \hat{\sigma}(\tau) + .05\ \hat{\sigma}(\tau) . \tag{15}$$

At day 16 we start the estimator with $\hat{\sigma}(\tau) = \hat{\sigma}_{16}(\tau)$. The noise test is performed every 5 days.

It proved necessary to adjust the $\hat{\sigma}_{16}(\tau)$ values to make them more closely correspond to the theoretical values of the sigma (a, b, τ) function. The multiplicative factors at $\tau = 1, 2, 4, 8$ are .97, 1.02, 1.05, 1.11, respectively. The need for this correction stems from a minor imperfection in the flicker generator [5] and from a skewed sampling distribution for $\hat{\sigma}_{16}(\tau)$. The sampling skew is suggested by analogy to the following example. Let random variable X be distributed chi-square with 1 degree of freedom. The $E[X] = 1$ and the $\Pr\{X<1\}$ is approximately equal to .7. So we infer by analogy that in a small sample of X outcomes, a number of low values is very probable. Since $\hat{\sigma}_{16}(8)$ is estimated from a single sample it often assumes a value lower than sigma (a, b, 8). Different correction factors would likely be better for $T \neq 16$.

A likelihood test is used to detect a noise change. Assume that $\hat{\sigma}(1)$ has a greater value than nominal in a realization. This suggests that a noise intensity increase may have occurred such that either $\alpha > a$ or $\beta > b$; we need to determine which one or neither (the test does not explicitly look for a change in both).

If $\hat{\sigma}(1) > \min\{\text{sigma}(a + .1, b, 1); \text{sigma}(a, b + .1, 1)\}$, (16)

a threshold value, let us start the following procedure. The first day that $\hat{\sigma}(1)$ exceeds the threshold value let $x_1 = x_2 = 0$ and $x_3 = 1$. At other times we update the x_i values. We interpolate twice on $\hat{\sigma}(1)$ to obtain the best fit to both sigma $(a, b^*, 1)$ and sigma $(a^*, b, 1)$. Then we compute the absolute differences

$$d_1 = \sum_{\tau=1,2,4,8} \left| \hat{\sigma}(\tau) - \text{sigma}(a, b^*, \tau) \right|$$

$$d_2 = \sum_{\tau=1,2,4,8} \left| \hat{\sigma}(\tau) - \text{sigma}(a^*, b, \tau) \right|$$

$$d_3 = \sum_{\tau=1,2,4,8} \left| \hat{\sigma}(\tau) - \text{sigma}(a, b, \tau) \right| .$$

We update the x_i by replacement signified as

$$x_i = .75\, x_i + .25 \left(\frac{\frac{1}{d_i}}{\frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3}} \right) . \tag{17}$$

In similar manner, if

$$\hat{\sigma}(1) < \max\{\text{sigma}(a - .1, b, 1); \text{sigma}(a, b - .1, 1)\} \tag{18}$$

we start $y_1 = y_2 = 0$ and $y_3 = 1$ the first time the threshold is exceeded and update on subsequent consecutive times. We interpolate on $\hat{\sigma}(1)$ again for both possible changes and compute d_i as above and update the y_i by replacement signified as

$$y_i = .75\, y_i + .25 \left(\frac{\frac{1}{d_i}}{\frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3}} \right) . \tag{19}$$

Thus, the $\sum_{i=1}^{3} x_i = \sum_{i=1}^{3} y_i = 1$ always.

The test using compound limits is

$$
\begin{aligned}
&\text{when } x_1 > .55 \text{ and } x_3 < .20 \quad \text{detect } \beta > b \\
&\text{when } y_1 > .55 \text{ and } y_3 < .20 \quad \text{detect } \alpha > a \\
&\text{when } x_2 > .55 \text{ and } x_3 < .20 \quad \text{detect } \beta < b \\
&\text{when } y_2 > .55 \text{ and } y_3 < .20 \quad \text{detect } \alpha < a \ .
\end{aligned}
\tag{20}
$$

This test has 7 parameters. For review they are

1. T = 16 day averaging interval;
2. Smooth every 5 days into $\hat{\sigma}(\tau)$;
3. .95 smoothing factor in eq (15);
4. The .1 threshold in eqs (16) and (18);
5. .75 smoothing factor in eqs (17) and (19);
6. .55 limit in eq (20); and
7. .20 limit in eq (20) .

Two procedures are used for detection of a time jump change. This change is signified by replacing z_t with z_t plus or minus a step as shown as follows:

$$z_t = z_t \pm J \ . \tag{21}$$

The first procedure has 3 parameters and either detects or fails to detect jump J on day t. Consider the filter

$$p_t = .90p_{t-1} + .10z_{t-1} , \tag{22}$$

where $p_1 = p_2 = z_1$. We let p_t predict the value z_t and if

$$|p_t - z_t| > 3.4 + .6b \tag{23}$$

we say that a jump must have been present at day t.

The second procedure using 5 parameters is a backup to the first since if

$$|z_t| > 2.8 + .6b \tag{24}$$

and

$$|z_t| > 4.8 \left(\frac{\sum_{k=t-10}^{t-1} z_k \quad \sum_{k=t+1}^{t+10} z_k}{20} \right) \tag{25}$$

and

$$|z_k| < .8|z_t| \quad \text{for all } k = t-10, \ldots, t-1, t+1, \ldots, t+10 \tag{26}$$

we detect that a jump occurred at day t. A jump at day t can only be detected at day t or at day t+10 using these procedures.

The drift change conditions can be represented by letting

$$z_t = z_t \pm D \pm Q(t - t_c), \tag{27}$$

where t_c is the day that the drift started.

A predictor

$$r_t = .95\, r_{t-1} + .05 z_t \tag{28}$$

is compared to a 5-day moving average

$$\bar{z}_5 = \frac{\sum_{k=t-4}^{t} z_k}{5} \tag{29}$$

in order to update one of four test quantities P_H, P_L, N_H, or N_L. The initial values at day 1 are $P_H = N_H = .75$ and $P_L = N_L = .50$. The update is signified in replacement form as follows:

$$\begin{aligned}
&\text{if } r_t > 0 \text{ and } \bar{z}_5 > 0, \text{ then } P_H = .8P_H + .2\bar{z}_5 \\
&\text{or if } r_t \le 0 \text{ and } \bar{z}_5 \le 0, \text{ then } N_H = .8N_H - .2\bar{z}_5 \\
&\text{or if } r_t > 0 \text{ and } \bar{z}_5 \le 0, \text{ then } N_L = .8N_L - .2\bar{z}_5 \\
&\text{or if } r_t \le 0 \text{ and } \bar{z}_5 > 0, \text{ then } P_L = .8P_L + .2\bar{z}_5 \quad .
\end{aligned} \tag{30}$$

The drift test is if

$$\begin{aligned}
&(P_H - N_H > .5 \text{ and } P_L > N_L) \\
&\text{or } (P_H - N_H > .7) \\
&\text{or } (P_L > .9N_H \text{ and } P_H > .8)
\end{aligned} \tag{31}$$

detect a drift in the positive direction, or if

$$(P_H - N_H > .5 \text{ and } P_L > N_L) \text{ or } (P_H - N_H > .7)$$
$$\text{or } (N_L > .9P_H \text{ and } N_H > .8) \tag{32}$$

detect a drift in negative direction. This test using 9 parameters is made every 5 days.

4. Experimental Development and Evaluation

4.1. Development of the Detection Procedures

Considerable effort was required to verify and normalize the Z_t generator so that $\sigma_W^2(1)$ and $\sigma_F^2(1)$ averaged out to 1. The experimental evaluation of the noise model was by simulation. A sequence of z_t values was generated for $t = 1, 2, \ldots, 512$ days. At day t_c, usually around day 30, a change could be introduced of the type shown in eqs (13), (14), (21), or (27). If no change were made at day t_c the run was said to be nominal.

The $\hat{\sigma}(\tau)$ estimator was studied. The parameter choices made were considered to be suitable. Longer averaging times have both advantages and disadvantages. A major drawback is that the estimator would require more than 16 days to start and would be less responsive to changes in the a and b variables. In the noise detection scheme the .1 threshold parameter was chosen because it was near the outer edges of the variation experienced in $\hat{\sigma}(1)$ under nominal conditions. It is possible to detect noise changes of less than $\pm$ .1 variation around either a or b.

In the case of the jump change the limit values in eqs (23) and (24) were chosen to balance falsely detecting a jump and actually detecting ones just inside the noise level. The .6b term in these limits reflects the apparently greater variation in the F_t process than in the W_t process.

The drift condition proved to be complex and the most difficult to detect. The motivation behind the P_H, N_H, P_L, and N_L quantities (meaning positive high, negative high, positive low, and negative low, respectively) is as follows: The Z_t process is autocorrelated due to bF_t. The direction of the process is generally guided by the low frequency "energy" in the bF_t process, while the higher frequencies generally come from the aW_t process. The drifts D or Q of eq (27) can be in a positive or negative direction. (If the signs are different we must wait for $Q(t - t_c)$ to dominate.) Let us assume that the r_t predictor in eq (28) indicates a positive direction and that a positive drift was introduced at day t_c. We contend that $\overline{z}_5$ will generally be greater in magnitude when it is in the same direction as r_t and generally smaller in magnitude when it is running counter to r_t. The reason for this condition is that D and $Q(t - t_c)$ are additive in one case and subtractive in the counter case. Thus in this example explanation P_H should be getting larger and N_L should be getting smaller. By reversing signs in the above example N_H would get larger and P_L smaller.

Several other schemes were tried using this same notion. They proved less effective until a very large number of days beyond day t_c. For D only not zero in eq (27) this scheme is complicated by the predictor gradually incorporating the D term into the prediction of $\overline{z}_5$. It should also be noted that D does not affect $\hat{\sigma}_{16}(\tau)$; the Q term gradually will, however.

4.2. Evaluation of the Detection Procedures

We estimate that the time between false detections is on the average about 400 days when b = .3. (The design center for these schemes was a = .7, b = .3.) Recall that 1 year was desired overall. Very accurate estimates could not be made as many nominal runs went to 512 days without a false detection. Figures 1 to 5 display cumulative noise for 5 nominal runs that went to 512 days without detection.

An evaluation of the jump change was made in 27 trials where $|J|$ ranged from 3.0 to 4.6. We experienced 21 jumps detected and 6 jumps missed. Figures 6 to 9 display 4 of these 21 correctly detected jumps. Of the 6 trials where the jump was missed, 4 went to 512 days nominally (as they should have) while 1 drift and 1 $\alpha < a$ condition were mistakenly detected. As $|J|$ increases, the chance of successfully detecting the jump also increases toward certainty.

In an evaluation of the time drift change consisting of 12 trials with $|D|$ equal to 1.0 or 1.5 and $Q = 0$, we obtained 9 correct detections, 2 missed detections which reported nominal to 512 days, and 1 $\beta < b$ mistake. The average number of days to make a $|D| = 1.0$ detection was 83 while a $|D| = 1.5$ detection was reduced to 27 days. Figures 10 and 11 show a correct $D = 1.0$ and $D = 1.5$ detection, respectively. A $|D|$ value around 3 might easily be mistaken for a jump. Values less than 1 require on the average more than 83 days to detect while values very near zero may not be detected much faster than the interval between false detections.

The rate drift change combined with a small time drift was evaluated in 10 trials; this resulted in 8 correct detections, 1 jump by mistake on the 6th day of drift and 1 $\alpha < a$ condition on the 35th day of drift. Had these 2 improper detections not been made so quickly, it is likely that all 10 rate drifts would have been properly detected. In the trials where $|D| = .1$ and $|Q| = .02$ the average number of days until detection was 76 while with $|D| = .2$ and $|Q| = .05$ the average was reduced to only 37 days. In practice, it is likely that Q values would be near zero requiring a longer period of drift before proper detection is made. Figures 12 and 13 display correct detections of combined time and rate drifts.

An evaluation of the noise change consisted of 45 trials where $\alpha = a \pm .2$ and $\beta = b \pm .2$. The overall results were that 19 correct detections were made; in addition we had 9 cases where the direction of the intensity was correctly sensed (a β change was reported instead of an α or vice versa), while 17 mistakes or missed detections occurred. Of the 17 incorrect cases 11 were drifts, 4 were jumps, and 2 were nominal to 512 days. In the correct detection cases the average number of days observed to detect $\beta = b + .2$, $\beta = b - .2$, $\alpha = a + .2$, $\alpha = a - .2$ were about 200, 175, 150, and 110 days, respectively. Figures 14 to 18 display 5 correct noise change detections.

The most successful detection by percentage was the $\alpha = a - .2$ followed by $\beta = b - .2$, $\alpha = a + .2$, with $\beta = b + .2$ being the hardest change to detect. It appears that the reason for $\beta = b - .2$ being detected slowly (175 days) was the difficulty in distinguishing this change from the $\alpha = a - .2$ change rather than from nominal. Most of the mistaken drift detections occurred on the $\beta = b + .2$ and $\alpha = a + .2$ trials. It is easily seen why these mistakes occurred in trials where the Z_t noise process has larger than nominal outcomes as opposed to the cases where Z_t is lower than nominal. In the 9 cases where we detected the wrong reason for a noise intensity increase or decrease but the correct direction, it was because the likelihood of the change indicated was greater than the other non-nominal alternative. In the cases where $\alpha = a - .2$ was detected instead of the proper $\beta = b - .2$ condition, the detection interval averaged around 100 days. When more days were taken to make the detection, more chance of getting the proper reason was shown; logically, this represents an effectively larger sample size and thus should do better. Figures 19 to 23 display 5 noise changes that were not correctly detected by procedure in eq (20).

An alternative way to check for a noise change might be to compute $\hat{\sigma}_{128}(\tau)$ for $\tau = 1, 2, 4, 8, 16, 32, 64$, and perform an interpolation on $\hat{\sigma}_{128}(1)$ and $\hat{\sigma}_{128}(2)$. Then we could compute a likelihood of any change. However, in this method one may go a number of 128 day intervals before detecting a change since the $\hat{\sigma}_{128}(\tau)$ estimator experiences a fairly substantial amount of variation about the theoretical sigma (a, b, τ) values.

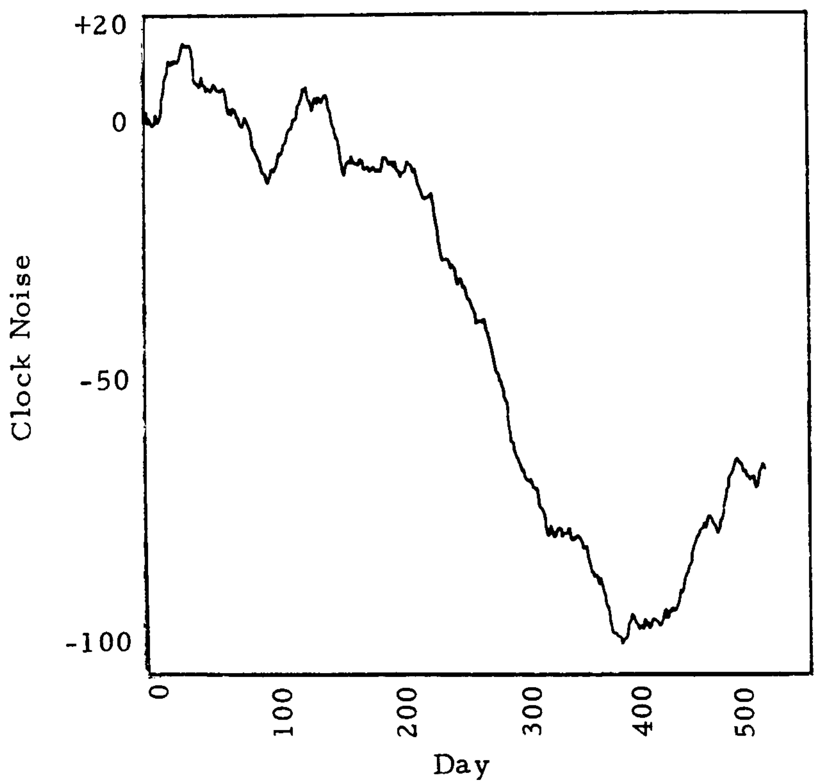

Figure 1. A nominal run with (a, b) = (.7, .3) and no detection in 512 days.

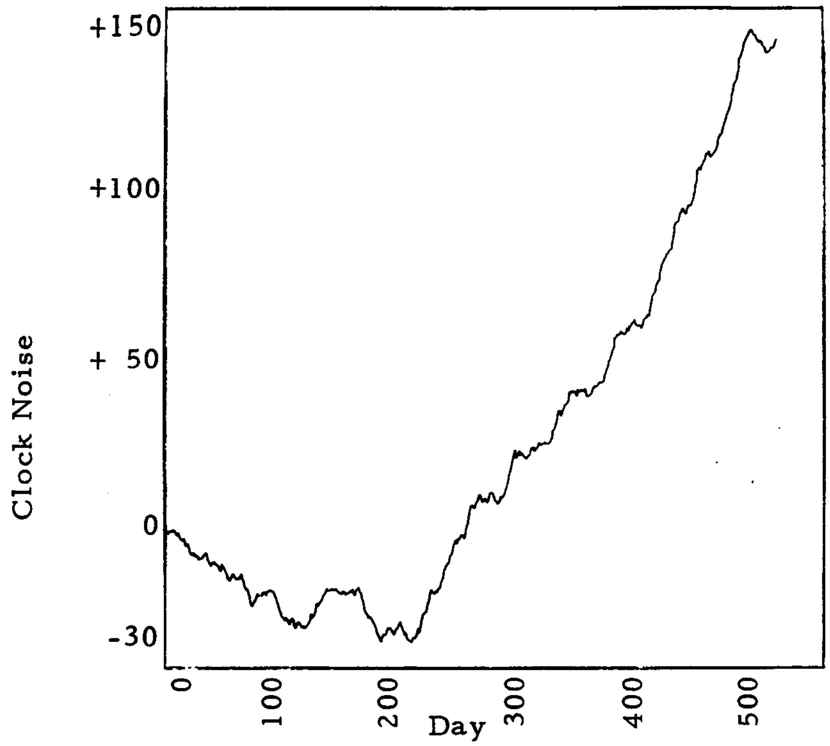

Figure 2. A nominal run with (a, b) = (.7, .3) and no detection in 512 days.

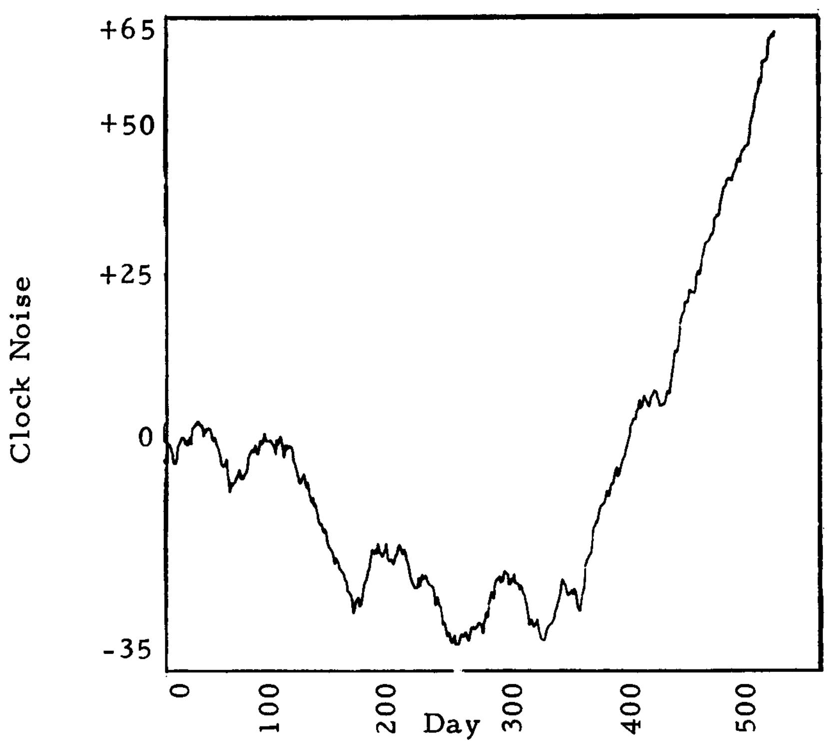

Figure 3. A nominal run with (a, b) = (.8, .2) and no detection in 512 days.

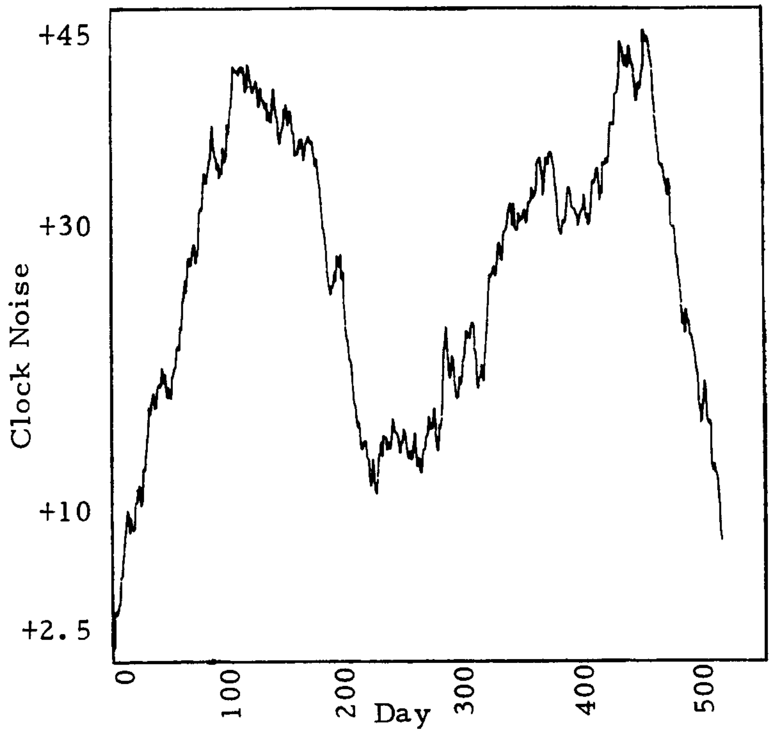

Figure 4. A nominal run with (a, b) = (.8, .2) and no detection in 512 days.

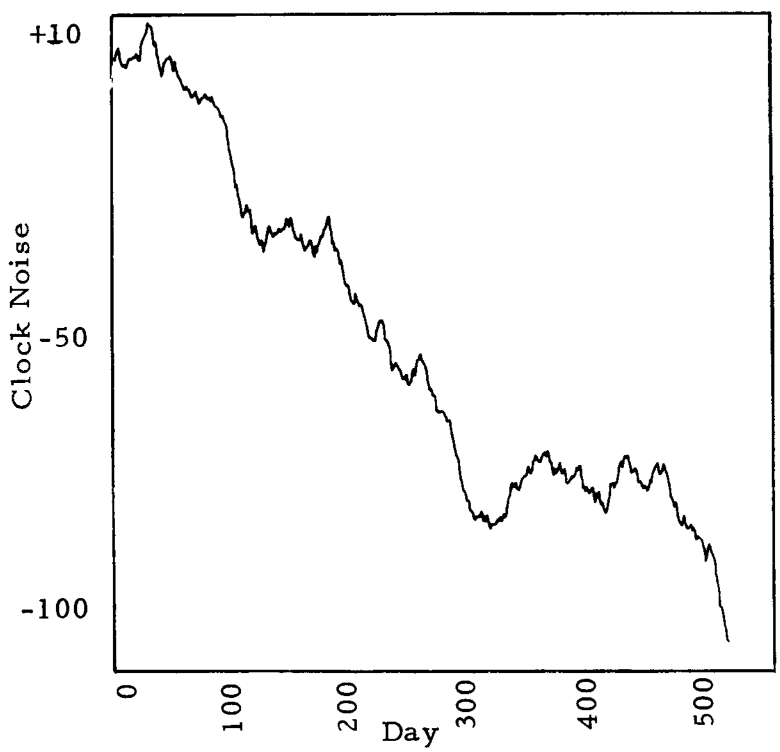

Figure 5. A nominal run with (a, b) = (.8, .2) and no detection in 512 days.

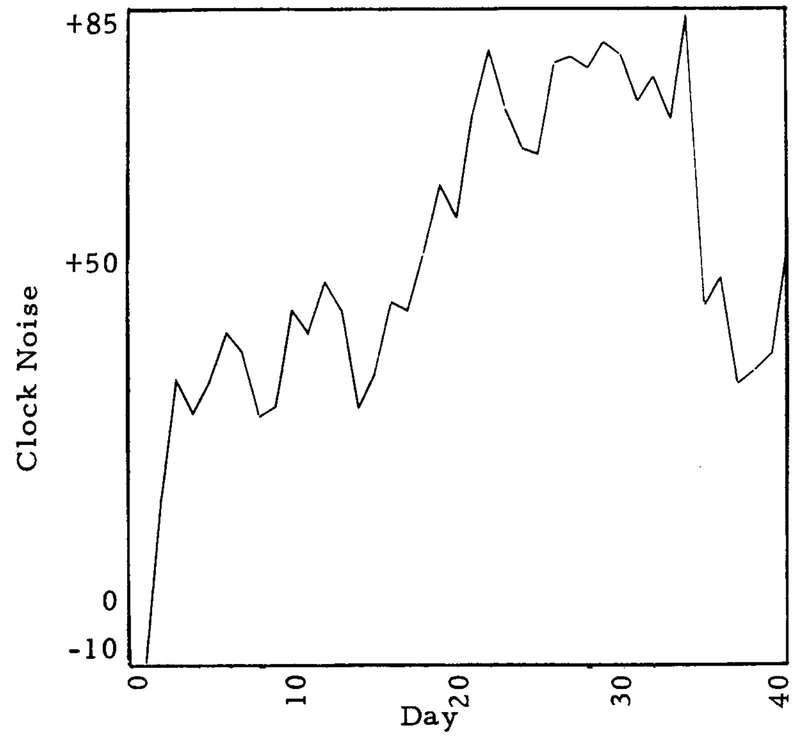

Figure 6. Detection of a J = 3.4 jump by eq (23) when (a, b) = (.8, .2) and t_c = 35.

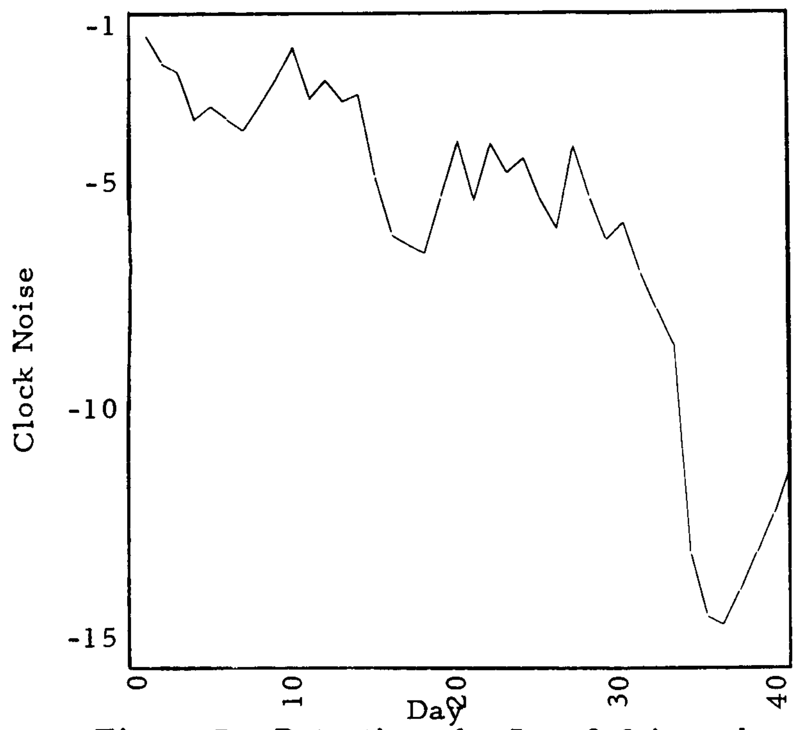

Figure 7. Detection of a J = -3.2 jump by eq (23) when (a, b) = (.8, .2) and t_c = 34.

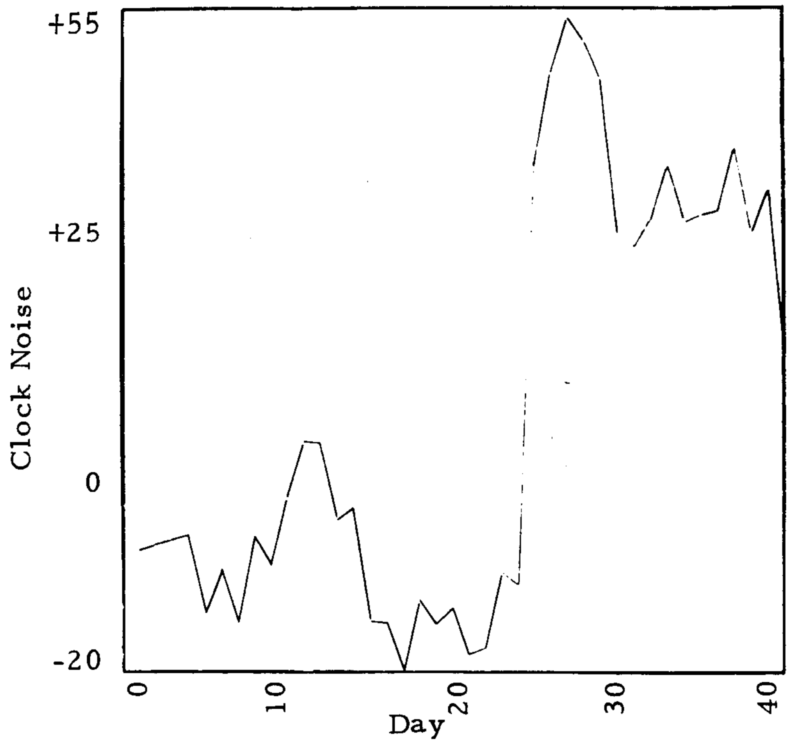

Figure 8. Detection of a J = 3.5 jump by eq (23) when (a, b) = (.7, .3) and t_c = 25.

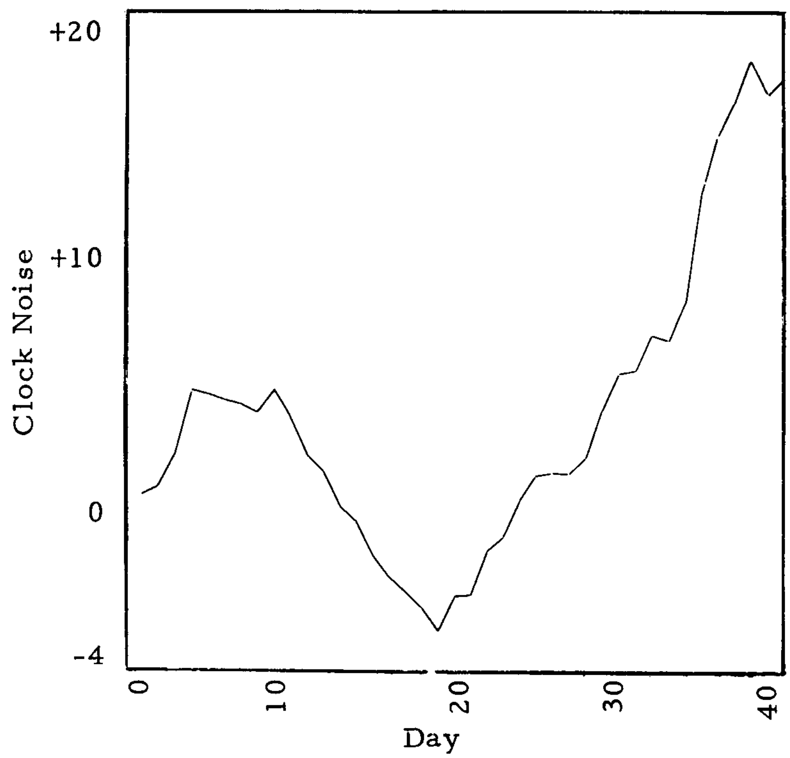

Figure 9. Detection of a J = 3.0 jump by eqs (24-26) when (a, b) - (.7, .3) and t_c = 35.

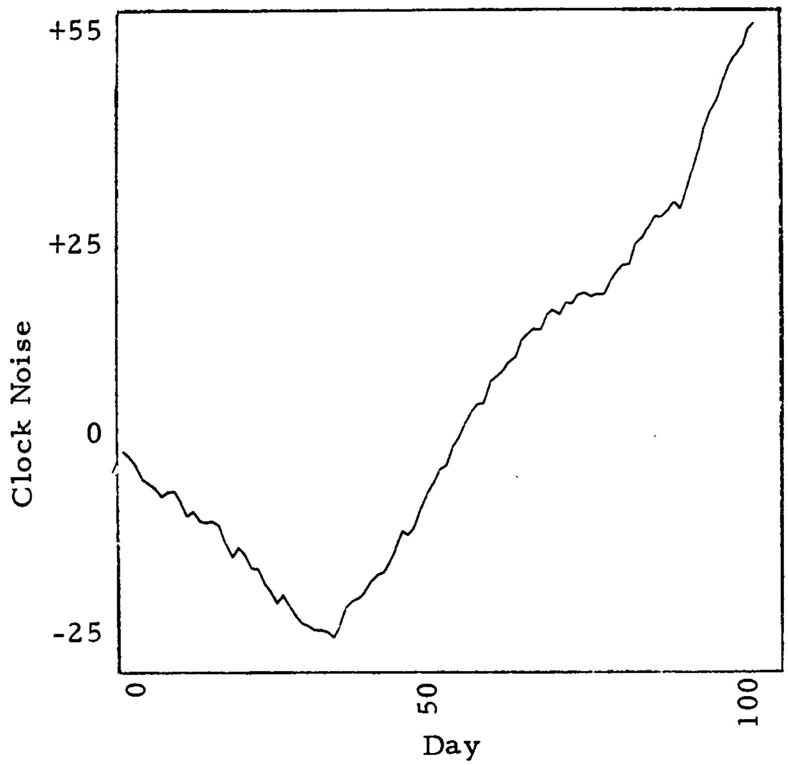

Figure 10. Detection of a D = 1.0 time drift by eq (31) when (a, b) = (.8, .2) and t_c = 35.

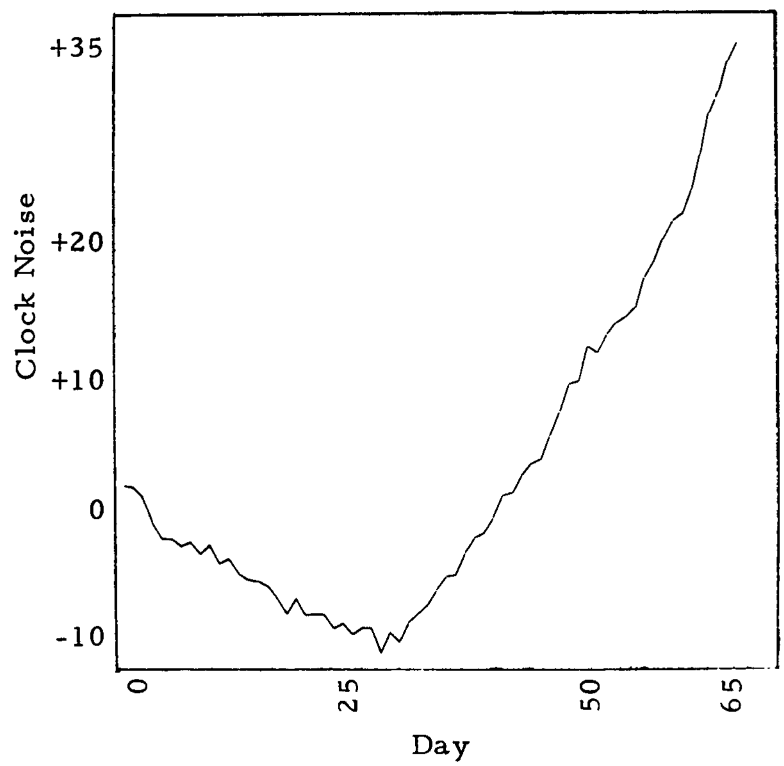

Figure 11. Detection of a D = 1.5 time drift by eq (31) when (a, b) - (.7, .3) and t_c = 31.

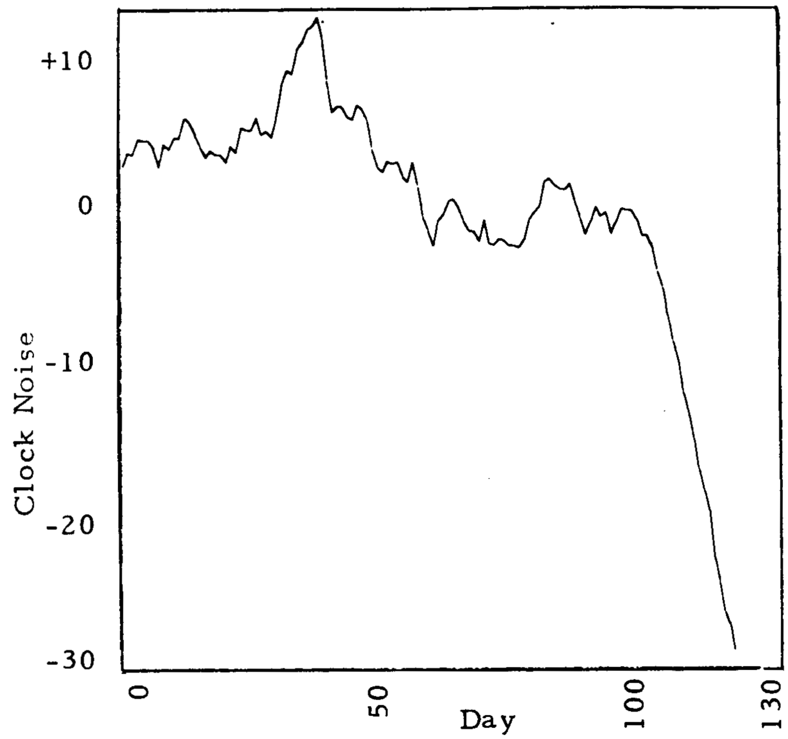

Figure 12. Detection of a D = -.1, Q = -.02 time and rate drift by eq (32) when (a, b) - (.7, .3) and t_c = 28.

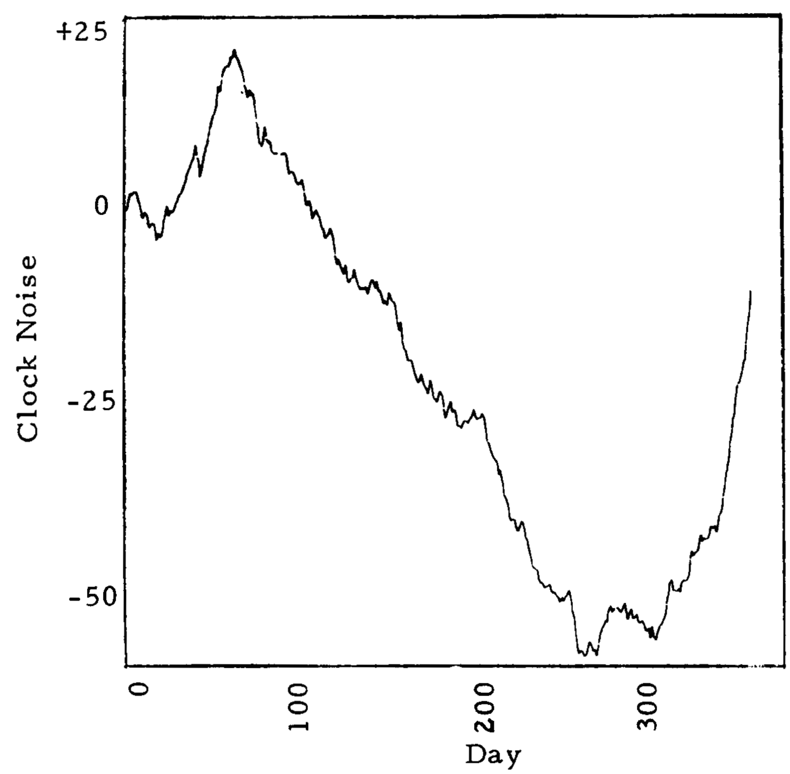

Figure 13. Detection of a D = .2, Q = .05 time and rate drift by eq (31) when (a, b) = (.7, .3) and t_c = 320.

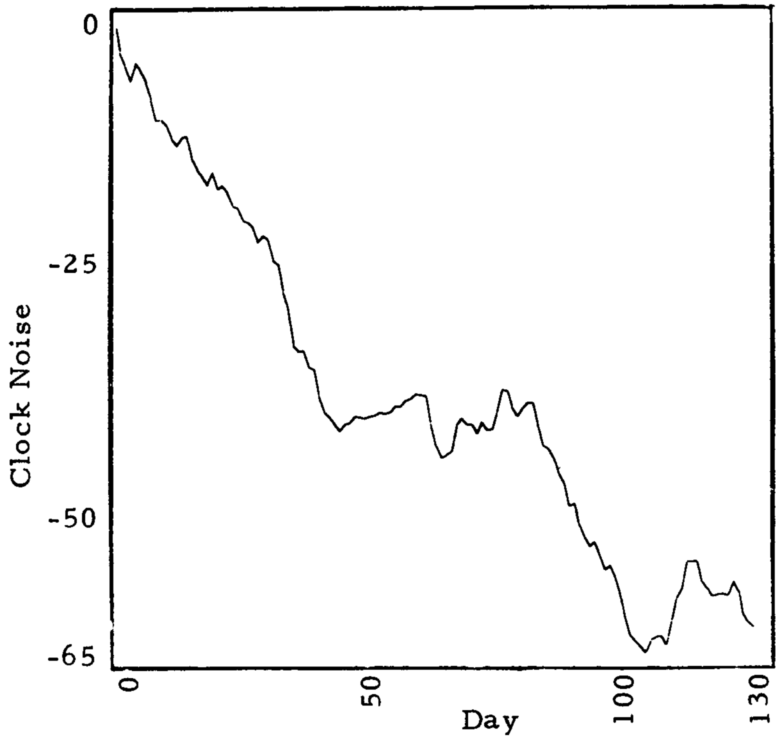

Figure 14. Detection at day 126 of noise change $\beta = .6$ when $(a, b) = (.7, .3)$ and $t_c = 28$.

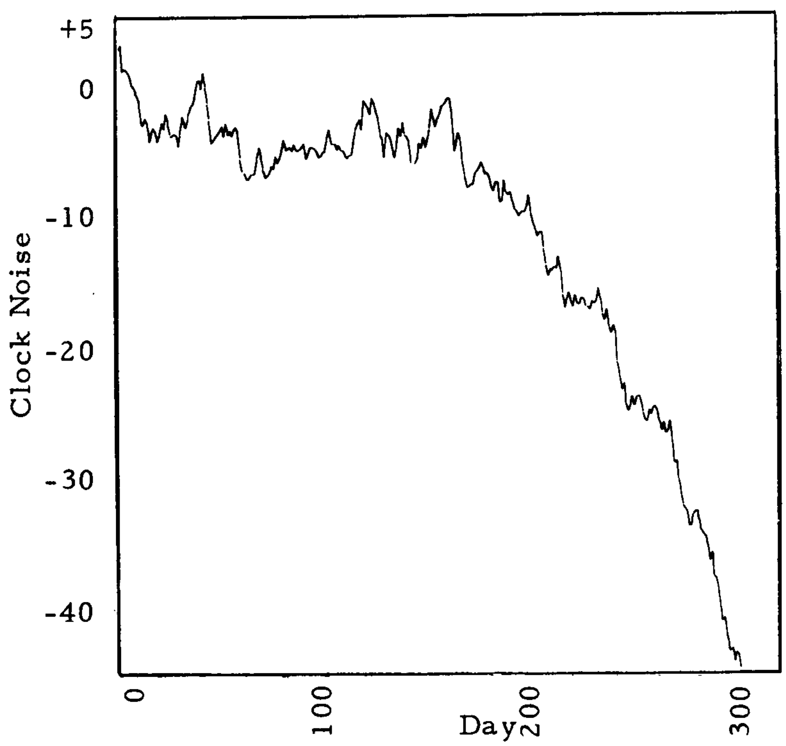

Figure 15. Detection at day 301 of noise change $\beta = .1$ when $(a, b) = (.7, .3)$ and $t_c = 26$.

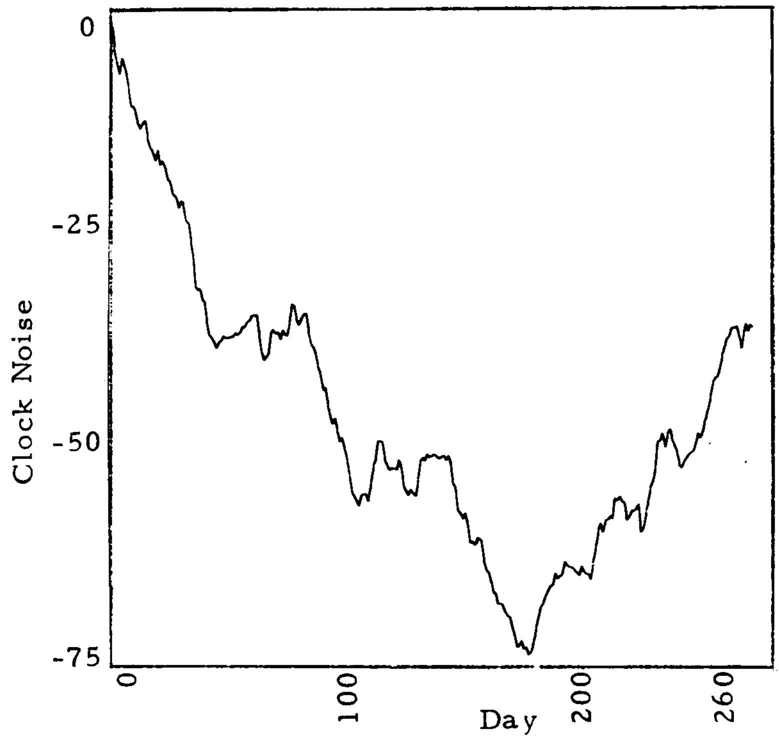

Figure 16. Detection at day 271 of noise change $\beta = .5$ when $(a, b) = (.7, .3)$ and $t_c = 28$.

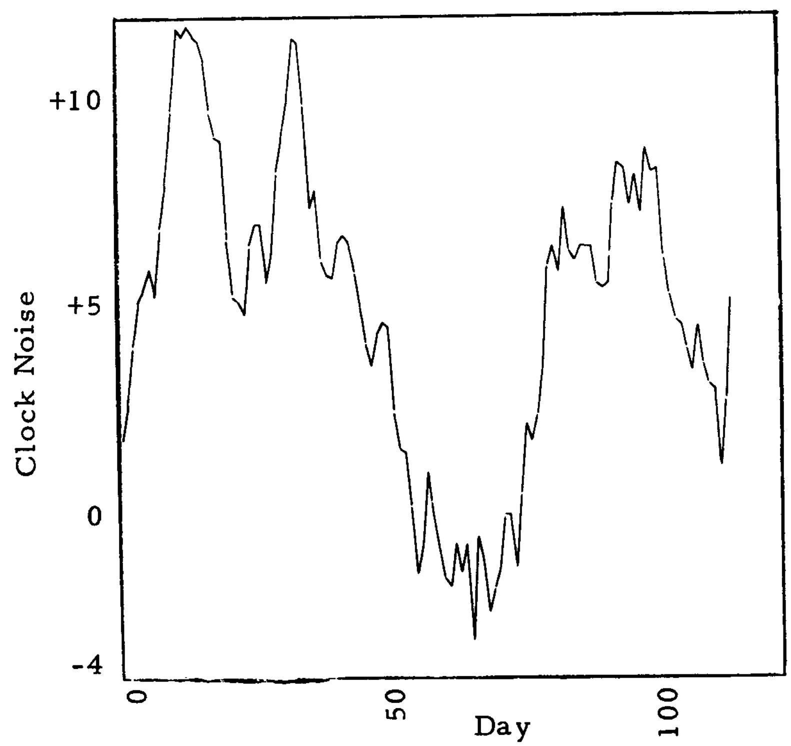

Figure 17. Detection at day 111 of noise change $\alpha = .9$ when $(a, b) = (.7, .3)$ and $t_c = 24$.

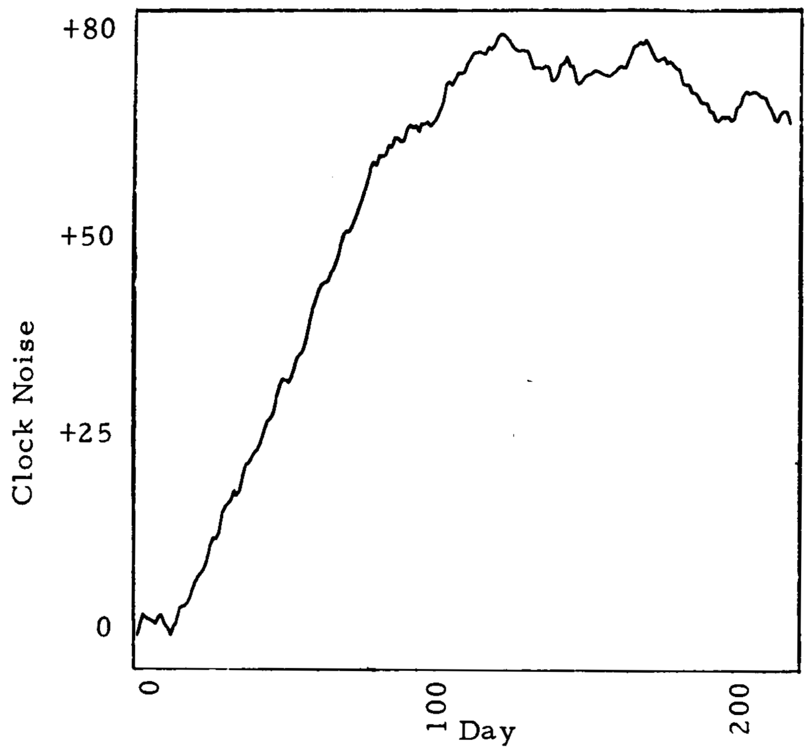

Figure 18. Detection at day 216 of noise change $\alpha = .5$ when (a, b) = (.7, .3) and $t_c = 28$.

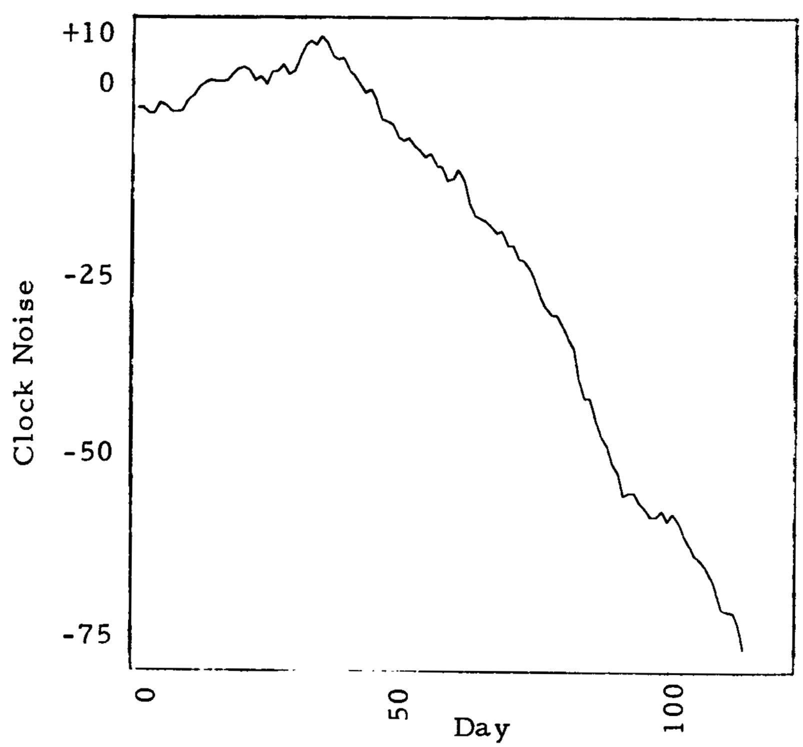

Figure 19. False inference at day 111 that $\alpha > .7$ when actually $\beta = .5$. (a, b) = (.7, .3) and t_c 26.

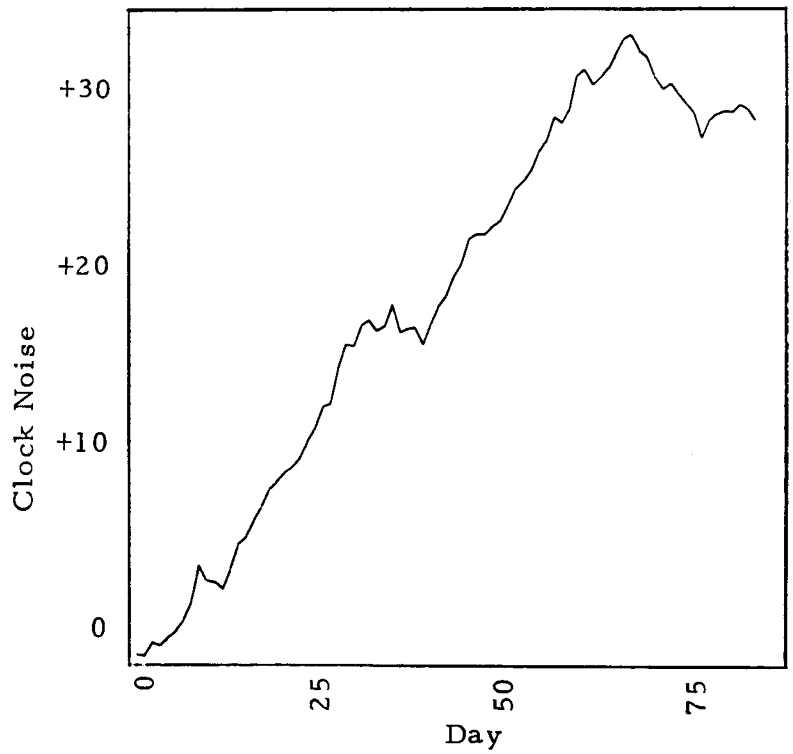

Figure 20. False inference at day 81 that $\alpha > .7$ when actually $\beta = .1$. (a, b) = (.7, .3) and $t_c = 31$.

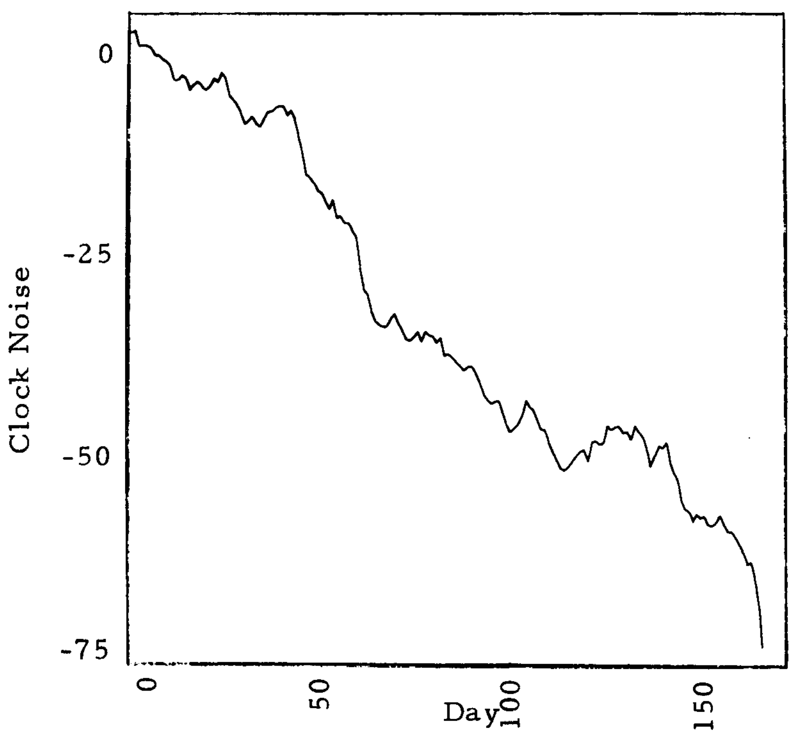

Figure 21. False detection of a jump by eq (23) at day 164 when actually $\beta = .5$. (a, b) = (.7, .3) and $t_c = 28$.

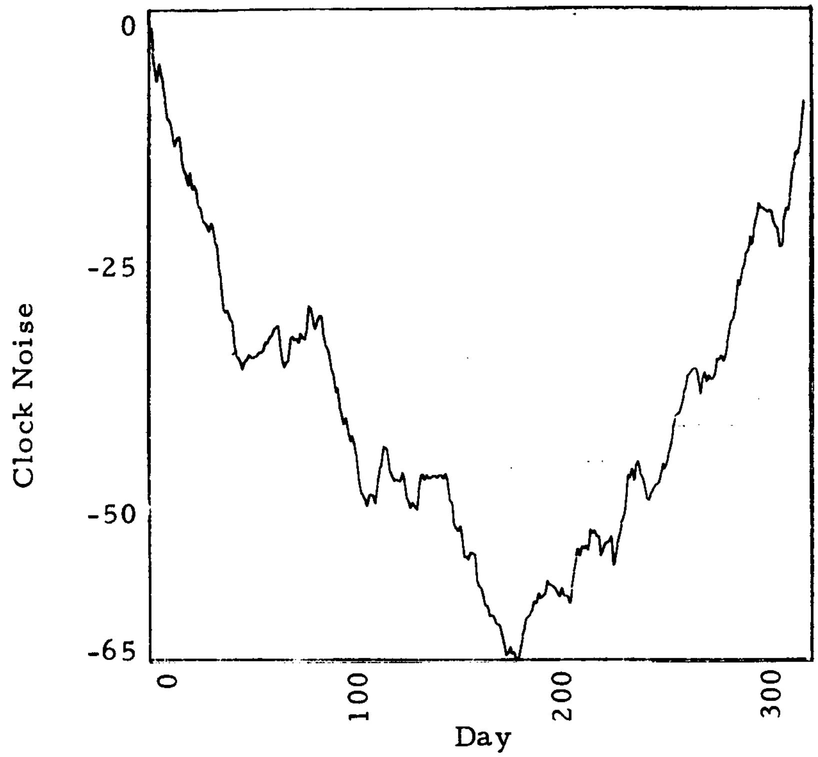

Figure 22. False detection of a drift by eq (31) at day 316 when actually β = .4. (a, b) = (.7, .3) and t_c = 28.

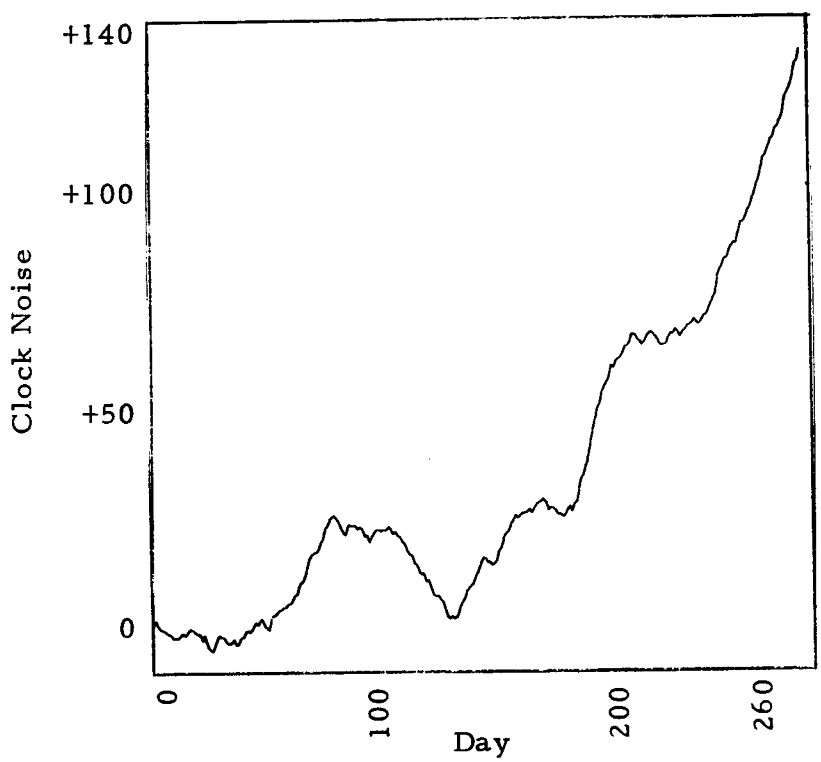

Figure 23. False detection of a drift by eq (31) at day 276 when actually β = .4. (a, b) = (.8, .2) and t_c = 29.

5. Conclusions

We can conclude that the false detection interval came out as desired. Also, nearly all of the runs where some change is introduced a detection is obtained; however, sometimes it is an incorrect detection. Rationale as to why mistakes are being made is available and enhances our insight into the difficulty of establishing the cause of any apparently non-nominal behavior of an individual clock in an ensemble. The experimental evaluation reported in section 4 was not intended to be comprehensive, but rather to give a general idea of successes and failures of these detection procedures.

The 24 parameters used is an unfortunately large number, however, the results indicate that even more sophisticated schemes might be necessary to obtain better discrimination; this could require even more parameters.

The problems of $\beta > b$ in the noise process looking like drift and $\beta < b$ appearing as though $\alpha < a$ since β cannot be allowed to go below zero when a lower than average $\hat{\sigma}(1)$ estimate is obtained can be resolved by specializing the test to the b intensity to a greater degree. We can also be comforted by the contention that jumps within the noise level should not degrade a time scale very much.

It could be very useful to know more about the conditional distribution of $Z_{t+\tau}$ given knowledge of Z_k for all $k \leq t$ as would other information about the Z_t noise process in terms of classical mathematical statistics. It would also be useful to know more about the sampling distribution for $\hat{\sigma}_T(\tau)$. In short, there are many more things to be learned about detection of changes in atomic clock performance. However, it is hoped that some of the concepts discussed in this report will aid the ongoing effort of improving atomic time scales.

6. References

[1] Allan, D. W., Gray, J. E., and Machlan, H. E., "The National Bureau of Standards atomic time scale system: Generation, dissemination, precision, and accuracy," IEEE Trans. Instrum. and Meas., IM-21, No. 4, pp. 388-391 (November 1972).

[2] Allan, D. W., "Statistics of atomic frequency standards," Proc. IEEE, 54, No. 2, pp. 221-230 (February 1966).

[3] Barnes, J. A., et al., "Characterization of frequency stability," IEEE Trans. Instrum. and Meas., IM-20, No. 2, pp. 105-120 (May 1971).

[4] Barnes, J. A., and Allan, D. W., "A statistical model of flicker noise," Proc. IEEE, 54, No. 2, pp. 199-207 (February 1966).

[5] Barnes, J. A., and Jarvis, Stephen, Jr., "Efficient numerical and analog modeling of flicker noise processes," NBS Technical Note 604 (June 1971).