A BRIEF REVIEW OF FREQUENCY STABILITY MEASURES

Gernot M. R. Winkler
U. S. Naval Observatory, Washington, D. C.

ABSTRACT

A standard statistical treatment of the measurements of frequencies of clocks can be done simply by computing the mean μ and variance σ^2 of these measurements. However, drifts (aging of crystal clocks) pose a problem because they cause a troublesome dependency of μ and σ^2 on N, the number of samples. While such measures in the time domain can be made more meaningful by using the mean square successive difference instead of σ^2 as a measure of clock stability, RADAR, microwave spectroscopy and other applications require measures which give sideband to carrier power ratios (frequency domain measures). The principles of such measures and their various advantages and disadvantages will be discussed.

INTRODUCTION

This paper has been prepared at the request of the program committee which wished to have the subject of frequency stability measures emphasized as one of the issues of this year's conference. It is therefore proposed to review the subject as an introduction to more specific papers and to the panel discussion which is to follow these presentations.

There is a considerable amount of written material available which can serve as introduction to our subject. One of the best introductory coverages is given in Rovera's (IEN, 1974) lecture notes. Appended is a selected list of original contributions. Two excellent and complementary reviews are the NBS Technical Notes No. 669 (D. Allan, 1975) and No. 679 (D. Howe, 1976). One cannot expect to improve upon such lucid and exemplary expositions of the subject. One can only attempt to present it in a different style, context and accent.

We can divide our subject according to various lines of which the division short-term vs. long-term stability is only a more superficial division. By measurement techniques one can distinguish measurements in the time domain from frequency domain measurements. One can characterize stability in the time domain and in the frequency domain. Here we would like to emphasize yet another dichotomy:

A. Trends, Systematic or Deterministic Variations of Frequency,

vs.

B. Random Variations as Part of a Random Process.

A failure to separate the systematic variations from the random part will, in most cases, affect any characterization of a clock in a very misleading way (Barnes and Allan, 1964). On the other hand, the random part of the observed frequency variations of a clock is of the same type as the observed variables which arise from many other random processes. Such data form a time series which can be investigated and characterized with the same methods as have proved useful in other sciences. We therefore do not have to re-invent the wheel but can simply apply what we can find in the literature of Time Series Analysis etc. (Box and Jenkins, 1970, Jenkins and Watts, 1968, Wiener, 1949).

CLOCK SAMPLING

Table 1 shows the general scheme which we can follow if we sample our clock error $x(t)$ at regular intervals h. Such sampling is the basic measurement in the time domain. We use a clock output marker (which could be a pulse or a sine wave zero crossing) and a reference clock "R" which we assume to be perfect. The sampling of the clock error of the test clock "T" is done with a time interval counter at the time t.

$$\text{START(T)} - \text{STOP(R)} = x(t) \qquad (1)$$

The complement of $x(t)$ is called the clock correction, $C(t) = -x(t)$. These measurements constitute our basic time series in which we designate the number of each measurement as the index k.

Table 1

Index k	Time $t=t_o+kh$	Clock Error $x(k)$	First Diff. $\nabla x(k)$	Second Diff. $\nabla^2 x(k)$	Third Diff. $\nabla^3 x(k)$
0	$t_o + 0h$	$x(0)$	-	-	-
1	$t_o + 1h$	$x(1)$	$\nabla x(1)$	-	-
2	$t_o + 2h$	$x(2)$	$\nabla x(2)$	$\nabla^2 x(2)$	-
3	$t_o + 3h$	$x(3)$	$\nabla x(3)$	$\nabla^2 x(3)$	$\nabla^3 x(3)$
4	$t_o + 4h$	$x(4)$	$\nabla x(4)$	$\nabla^2 x(4)$	$\nabla^3 x(4)$
5	$t_o + 5h$	$x(5)$	$\nabla x(5)$	$\nabla^2 x(5)$	$\nabla^3 x(5)$

We assume n + 1 measurements (last k = n). We use "backward" differences ∇ because that is what we can compute in real time immediately after a measurement. The symbol ∇ may also be understood as a linear operator to be applied to a table entry z_k (cf. Hamming, 1962):

$$\nabla(z_k) = z_k - z_{k-1} \tag{2}$$

$$\nabla(\nabla^m(z_k)) = \nabla^m(z_k) - \nabla^m(z_{k-1}) \tag{3}$$

A similar operator is the backshift operator B:

$$B(z_k) = z_{k-1} \tag{4}$$

We have obviously the following operator equation which is true for whatever argument is chosen:

$$\nabla = 1 - B \tag{5}$$

(1 is the identify operator). A formalism such as (5) will be used in the discussions of the Box-Jenkins (1970) approach to statistical modeling. Let us apply these concepts to table 1 where we interpret the m^{th} difference as a symbolic m^{th} power of ∇. We use the binomial theorem:

$$\nabla^m = (1 - B)^m = \sum_{r=0}^{m} (-1)^r \binom{m}{r} B^r \tag{6}$$

B^r means that we apply B r times. As an example we can write the general expression for the 3^{rd} difference in our table 1 by multiplying equation (6) with x(k):

$$\nabla^3 x(k) = \sum_{r=0}^{3} (-1)^r \binom{3}{r} x(k-r)$$

$$= x(k) - 3x(k-1) + 3x(k-2) - x(k-3)$$

FREQUENCY

The average normalized (relative) frequency departure (error) during the basic interval h is

$$\bar{y}_h(k) = \frac{\nabla x(k)}{h} = \frac{\Delta x}{\Delta t} \tag{7}$$

Its dimension is time error per unit time, i.e. it is a dimensionless (relative) quantity. It is therefore given as parts per billion etc., whereas the time error is given as time, e.g. as microseconds (μs). Frequency F, as distinguished from relative frequency error $\bar{y}(t)$, is given in cycles (of 2π radians each) per second (Hertz, Hz). The period P is the duration of one cycle: P = 1/F. Angular frequency Ω is given in radians per second (rad/s) and we have $\Omega = 2\pi F$ and

$$\bar{y} = \frac{\Delta F}{F} = \frac{1}{\Omega}\cdot\frac{\Delta\phi}{\Delta t}, \quad x = \frac{\phi}{\Omega} \tag{8}$$

where ϕ is the phase error at the frequency F.

Our timing waveform is always bandlimited since it comes from mateial circuits. Our concept of average frequency may therefore be brought to the limit for a smaller and smaller sampling interval Δt and we obtain the concept of the instantaneous relative frequency departure

$$y(t) = \lim_{\Delta t \to o} \frac{\Delta x}{\Delta t} = \frac{dx}{dt} = \frac{1}{\Omega}\frac{d\phi}{dt} = \frac{1}{\Omega}\dot{\phi} \tag{9}$$

For practical reasons our basic sampling interval h must be chosen sufficiently long so that the Δx can be resolved above the noise level of measurement. Once the measurements are recorded it is of course easy to choose any multiple of h as the averaging or integration time τ. In the general case we have

$$\tau = kh \tag{10}$$

where $k \geq 1$ is a positive integer.

Now, if our clock errors would be purely deterministic and if they followed an m^{th} degree polynomial

$$x(t) = a'_0 + a'_1 t + a'_2 t^2 + a'_3 t^3 + \cdots\cdots + a'_m t^m \tag{11}$$

which we could consider e.g. as the beginning of a Taylor series, then, according to the fundamental theorem of the difference calculus, the m^{th} difference would be a constant. By applying the operator ∇ m times one can verify that

$$\nabla^m x(t) = a'_m m! h^m \tag{12}$$

If we use the index k as the argument with h = 1, then the polynomial is now in powers of k

$$x(k) = a_0 + a_1 k + a_2 k^2 + a_3 k^3 + \cdots\cdots + a_m k^m \tag{11'}$$

with

$$\nabla^m x(k) = a_m m! \tag{12'}$$

The (m + 1)st differences are zero. Since (12) and (12') refer to the same quantity we also can see that

$$a_m = a'_m h^m \tag{13}$$

where a'_m has the dimension of time $^{m+1}$. But in addition to measurement noise, our clocks are always subject to random disturbances and therefore the ∇^{m+1} will not be zero but will reflect the clock erratics. These erratics will be magnified as higher differences are taken according to (6).

These properties of the higher differences offer a possibility to obtain insight into the clock's performance. We could try to form the average difference

$$\overline{\nabla^\ell x(k)} = \frac{1}{n - \ell + 1} \sum_{k=\ell}^{n} \nabla^\ell x(k) \quad (\ell \geq 1) \tag{14}$$

for successive $\ell = 1, 2, 3$ etc. until we find a $\overline{\nabla^m x(k)}$ which is not only close to zero but also shows no trends if we take different samples, i.e. which is stationary. Our clock performance would then contain a systematic part in the form of a polynomial of degree m-1 and the coefficient of the highest power term would be

$$a_{m-1} = \frac{1}{(m-1)!} \cdot \nabla^{m-1} x(k) \tag{15}$$

As an example consider a quartz-crystal clock which can be modeled typically with a second degree polynomial for its clock errors

$$x(k) = a_0 + a_1 k + a_2 k^2 + x'(k) \tag{16}$$

where x'(k) are the residuals. The third differences will average (show no trend) close to zero and we find for the aging coefficient

$$a_2 = \frac{\overline{\nabla^2 x(k)}}{2} \tag{17}$$

Furthermore in this case one can use as a measure for the clock erratics (the random part) the average third absolute difference as it once was actually used by H. M. Smith in the early 50's at the RGO. Such

a measure would be written in our terms

$$V = \left|\overline{\nabla^3 x(k)}\right|$$

By using the root mean square we obtain a sort of higher variance (since the average is close to zero) as it has been discussed by Kramer at the PTB (1975) who called it curvature variance. Expressed as sigma (variance $^{-1/2}$) we have

$$\hat{\sigma}_c(h) = \sqrt{\frac{\overline{[\nabla^3 x(k)]^2}}{3!\ h^2}} \qquad (18)$$

This measure is insensitive to frequency drifts and in that regard would be superior to the pair variance (two sample Allan variance). The pair sigma, which for good reasons is the most popular frequency stability measure, is in our terms

$$\hat{\sigma}_y(h) = \sqrt{\frac{\overline{[\nabla^2 x(k)]^2}}{2!\ h^2}} \qquad (19)$$

Formulae (18) and (19) are, in fact, just special cases of the discussion given by Audoin & Lesage (1975). A sample variance $\sigma_{\bar{y}}^2$, finally, can simply be obtained by averaging the $\nabla x(k)$ with a naive determination of the standard deviation and squaring it. But again, this is usually a misleading procedure if we do not first take off any trends in the record (in another way of looking at the problem this amounts to forcing the averages to zero in which case, if we have a white noise process, all of the above discussed mean square measures become variances of the same magnitude).

The simplest case in principle, and a very important one for conceptual clarification, is the case of purely random $\bar{y}(k)$'s. We call that white FM. A cesium beam atomic clock would be an example (typically for $\tau > 10s$) since the servo loop error signal coming from the atomic beam tube is shot noise fundamentally due to the arrival of individual atoms at the detector. We should have no trends in this case but the random additions of noise produce a random walk in $x(t)$ as is shown in fig. 1. This random walk (RW) has neither a stationary mean nor a stationary variance in $x(t)$ but of course, the differences constitute a normal process by definition. This is a major reason why the $\bar{y}(t)$ play such a large role in comparison with the $x(t)$ which are the values one measures directly. By using differences one gets rid of not only systematics but a RW as well.

An actual clock performance as shown in fig. 2 (which gives the $\overline{y}(1^d)$ of Cs 571) exhibits always some systematic disturbances in addition and the clock errors are therefore much larger. In addition to a linear trend of $+1.5 \times 10^{-13}$ p.a. which has been removed, we see that there is some obvious correlation in the frequency residuals. It is this tendency of frequency variations to persist for a while which produces a dispersion of the time error which is greater than in the case of purely random FM.

It will be clear at this point that

(1) The measures of clock performance based on sampling (table 1) become less sensitive to long term trends as higher differences are used for the analysis.

(2) It would be most desirable to agree on a standard way to reduce the data to a clock performance measure which is related to the application of clocks rather than to adopt procedures such as those which led to formulae (14) and (15).

The statements (1) and (2) are in mutual conflict only at first sight because in almost all cases we do not need to go to higher than second differences if we take off any long-term trends first. This allows us to satisfy the requirement (2) in the most straightforward way.

A. Elimination of Trends

There are three main classes of approximating functions which we can consider viz.

(1) Polynomials in t (or k) of degree m,

(2) Fourier Series,

(3) Exponential Functions, $\sum_i e^{-a_i t}$

All of these transform into another function of the same kind (they are class conservative) under a transformation $t \to t + k$, i.e. we can assume that there is no special origin inherent to the problem. If any singularity is present in the record then one can't use these functions, at least not for the whole range. In the case of clock records we are not so much concerned with genuine singularities. But often we see breaks in clock rate due to a specific but usually unknown disturbance. The most natural way to handle such cases is to break the record at these points and to use a particular function for the approximation of each undisturbed part. The simplest case would be the approximation of a clock with pieces of straight lines. This can usually be done with the $\overline{y}(k)$.

Only class (1) is class conservative also under the transformation $t \rightarrow kt$. This class is therefore of major importance if no natural scale is inherent in the problem. For an extensive discussion of approximation with these mathematical functions cf. Hamming (1962). For these reasons clocks have been approximated with polynomials in the overwhelming number of cases. Quadratic and cubic polynomials are usual for the $\overline{y}(k)$ and the $x(k)$ of crystal clocks (the cubic is to account for a changing frequency drift due to aging). For the mathematics of polynomial approximation cf. Jordan (1965) if one does not have a "canned" computer or calculator program available.

It is clear that by increasing the degree of the polynomial one can reduce the variance of the residuals but there is an absolutely essential point to be kept in mind. Any attempt to absorb the random variations in a mathematical model with a complexity greater than what is necessary for the elimination of the genuine long-term trend will lead to completely illusory gains. An apparent slight decrease of the variance of the residuals will have been obtained at the price of a representation of the trends which is useless and even dangerous for any amount of extrapolation. The same danger exists if a clear break due to a real clock disturbance occurs and is not recognized as such but incorporated into a greater complexity of the clock model. Eighth and higher degree polynomials have been used for long term clock modeling instead of breaking up the record into pieces which can be fitted with a low order polynomial. This is fundamentally wrong and dangerous. Similar is the principle of parsimony (Box & Jenkins, 1970).

However, the elimination of trends is usually extremely simple. Of greatest importance is the recognition of breaks or steps in the record. In most cases we will then find a more or less linear drift of frequency which can easily be subtracted from the record. In some cases the elimination of trends is a tricky problem and it is wise to be conservative and to take off only the most obvious overall trend instead of trying to be sophisticated. In that case one is likely to end up with a mathematical model for the low Fourier frequency components of the random part of the frequency variations.

Of course, the problem, if there is any, can always be solved by going to higher differences. Such a filtering (1^{st} order differences) is already being done automatically by using frequency instead of phase records for evaluation. This brings us one step closer to the goal of stationarity which produces stable measures. But this is of major importance mainly in clock modeling for purposes of prediction and much less so for the characterization of performance as needed for specifications, testing etc. It is a main benefit of the pair variance that it uses second differences thereby avoiding most of the problems mentioned while still giving an objective reproducible measure which is closely related to at least one important class of applications.

B. Random Process Characterization

Assume that we have freed our record from obvious trends. A natural question will be whether the residuals x'(t) will now be purely random and/or stationary. Particularly the stationarity question has plagued the discussions because of a number of frequent misunderstandings. First we must realize that no natural process can be assumed stationary in the sense that its statistical measures are independent of time. The universe does not allow any process to go on indefinitely, therefore life time limitations and aging phenomena are commonplace. Stationarity can only be a property of models as it was pointed out so emphatically by Barnes (1976) and Barnes et al (1971). The question must then be asked whether and how one can apply stationary models.

Let us consider an example: Fig. 3 gives the daily frequencies of clock 837 after removal of a linear frequency drift of -7.2×10^{-13} p.a., in a record of more than 900 days. We can see two types of non-stationarity quite clearly:

(i) a residual long-term frequency variation of systematic character (non-stationarity of μ) and

(ii) an increase in the frequency variations from day to day (non-stationarity in σ^2).

If we really want to characterize fully the behavior of this clock then we must display the record as it is given in fig. 3. Any other measure of frequency stability characterization can only be done by compacting data. But this is a euphemism, we actually must throw data away. If I say that after drift removal the sample variance is 4.4×10^{-26} ($\sigma = 2.1 \times 10^{-13}$) then this will give only a very general idea of the overall variations and it will be too pessimistic for most purposes. The day to day variations are really much smaller. In contrast to the sample variance which is rather useless in such a case, the pair sigma (corresponding to the two sample Allan variance) is more realistically tuned to what is useful in a time keeping application. It is only $\sigma_y(1^d) = 4.5 \times 10^{-14}$. This datum will be even more useful if we add the explanation that this is an average value, that the sigma is better for the first two years of operation and is getting worse now. Therefore we must also realize that the stability measures have a time variability. This adds a significant complication to the problem of stability measures. We interpret our samples as pieces of stationary time series models and we state the model parameters as they change with the samples which come from later dates (when the clock ages). Since the samples are necessarily limited in time, our confidence in the statistics must be limited also (cf. Lesage and Audoin, 1973). But there is no point at all in insisting on hairsplitting perfection which is of no concern in practice since we don't have perennial clocks.

THE AUTOCOVARIANCE AND THE SPECTRAL DENSITY

As we could see in the last example, the frequency residuals $\bar{y}'(t)$ can be correlated in time (fig. 3). A purely random (uncorrelated, normal) process is easy to characterize with its mean μ and the variance σ^2. A process with correlated distrubances $z(k)$ can be characterized by its sample mean $\hat{\mu}$ (we use the "hat" to distinguish the estimates from the ideal population parameters):

$$\hat{\mu} = \frac{1}{n} \cdot \sum_{k=1}^{n} z(k) = \bar{z} \tag{20}$$

and the autocovariance (acvf) for lag u:

$$c(u) = \frac{1}{n} \cdot \sum_{i=u+1}^{n} (z_i - \bar{z})(z_{i-u} - \bar{z}) = \hat{\gamma}(u) \tag{21}$$

and we see immediately that

$$c(0) = \hat{\sigma}^2 \tag{22}$$

The acvf is the average lagged product of the deviations from the mean and can therefore be interpreted as a quantity which originates from the variance but which is "spread out" into the lag axis. For any precision in the estimates it is clear that n will need to be large and in general will be $n > 50$. The normalized acvf is known as the autocorrelation function (acf) $\rho(u)$:

$$\rho(u) = \frac{\gamma(u)}{\gamma(o)} = \frac{1}{\sigma^2}\gamma(u) \tag{23}$$

Obviously the $c(u)$ or the $r(u) = \hat{\rho}(u)$ cannot be estimated confidently for greater lags than a fraction of n and in practice one should stay within $u < n/2$. (Jenkins and Watts recommend $u < \frac{n}{10}$). Completely random uncorrelated disturbances will have an acf which drops to zero for any $u \geq 1$. In contrast we see that actual clocks show a $r_{\bar{y}}(u)$ which indicates significant correlation for lags of many days (figs. 4 and 5). Indeed the acvf of $\bar{y}(\tau)$ gives most of the significant information of interest but it does not give it in a form which is best suited for further analysis. That is available in the Fourier transform of the acvf, the (one-sided) spectral density of the $\bar{y}$:

$$S_{\bar{y}}(f) = 4 \int_0^{\infty} \gamma(u)\cos(2\pi f \cdot u)du \tag{24}$$

Since the S(f) and the acvf are a Fourier transform pair we also can go backwards to the acvf once we know the S(f):

$$\gamma_{\overline{y}}(u) = \int_0^\infty S_{\overline{y}}(f)\cos(2\pi f \cdot u)df \tag{25}$$

A consequence of this last formula is

$$\gamma_{\overline{y}}(0) = \sigma^2_{\overline{y}} = \int_0^\infty S_{\overline{y}}(f)df \tag{26}$$

which explains S(f) as a variance density function of Fourier frequency. Its dimension is therefore variance per Hz. One must not be misled by these simple and transparent relationships. S(f) is not directly available but must be computed from a finite sample. This can be a tricky process and Jenkins and Watts (1968) or similar references must be consulted for guidance on details. Figs. 6 and 7 show examples of $\hat{S}_{\overline{y}}(f)$ computed, however, from the acf which normalizes the plots to a $\sigma^2 = 1$.

A special caveat concerns the role of τ. For each value chosen, $\gamma_{\overline{y}}(u)$ and $S_{\overline{y}}(f)$ will be different as can be seen from (26) but also because of the measurement resolution which changes with τ. The essential point is, however, that sampling the frequency for an interval τ corresponds to a convolution with a rectangular time window. This is transformed into the frequency domain as a factor

$$K = \frac{\sin^2(\pi f\tau)}{(\pi f\tau)^2} \tag{27}$$

so that the spectral density of $\overline{y}_\tau(K)$ is really obtained from the spectral density of y, the instantaneous relative frequency departure, by multiplication with K:

$$S_{\overline{y}}(f) = S_y(f)\cdot K \tag{28}$$

A second factor K' is needed if we want to compute an estimated variance for a sample of finite data length. As explained by Cutler and Searle (1966) in great detail (their paper is indispensable for an understanding of details, unfortunately it is full of annoying misprints), sampling for a total time n·h, cf. table 1, and removing the drift and average (i.e. our elimination of systematics) corresponds to a high pass filtering with two zeros and a cutoff frequency

$$f_L = 1/nh\pi \tag{29}$$

Therefore the variance in such a sample must be expected to be smaller than the true variance. The computation achieves this with the factor

$$K' = 1 - (\sin^2 fnh\pi)/(fnh\pi)^2 \tag{30}$$

and we obtain

$$\sigma^2_{\bar{y}}(h) = \int_0^\infty S_y(f)\cdot K\cdot K'\cdot df \tag{31}$$

The high pass filter action of K' can usually (if we have removed the systematics) be approximated by a low frequency cutoff on the integral

$$\sigma^2_{\bar{y}}(h) \sim \int_{1/nh\pi}^\infty S_y(f)\cdot K\cdot df \tag{32}$$

A method of obtaining $\hat{\sigma}$ from the frequency domain without sampling and digital analysis is actually based on this formula (Rutman, 1974, Rutman and Sauvage 1974).

On the other hand if we consider how we actually obtain our estimate of $S_{\bar{y}}(f)$, the $\hat{S}_{\bar{y}}(f)$, then the situation is somehow the reverse of what we have just discussed. Data acquisition for a total time h·n corresponds to a convolution of the true S(f) with a filter function $Q_c(f)$ which is the Fourier transform of the time window (the Hanning or Hamming etc. which we perform on the acvf). Therefore the true S(f) differs from the estimate $\hat{S}(f)$ which we compute:

$$E[\hat{S}] = S \circledast Q_c(f) \qquad (\circledast \text{ signifies convolution}) \tag{33}$$

The effect of this is a blurring of the details of S(f), making it impossible to resolve details finer than $\Delta f \sim 1/nh\pi$.

Since $S_y(f)$ is of such fundamental importance for frequency stability considerations, the accepted terminology concerning types of clock erratics often refers to it. As an important example we mention white FM, completely uncorrelated frequency disturbances, which is characterized by a $S_y(f)$ = constant between the practical limits f_L and f_h given by record length and sample interval. We have

$$f_h = 1/2h \tag{34}$$

This is related to the sampling theorem which one must also keep in mind in regard to the danger of aliasing. If the sampled process contains substantial noise at or above f_h then special filtering is advisable (cf. Baugh 1971).

GENERAL COMMENTS ON THE USE OF STATIONARY MODELS:

If z(t) (which could be a model for our x, y or any of the ∇'s) is stationary then $\gamma(u)$ is an even function and the probability distribution function P(z) (pdf) is time invariant.

If z(t) is "wide sense" stationary then μ is constant and $\gamma(u)$ is only a function of u.

If z(t) is stationary and ergodic then the ensemble (stochastic) average μ is equal to the time average $\overline{z(t)}$:

$$\mu = \overline{z(t)} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} z(t)dt \qquad (35)$$

If we have done a sufficient job with the removal or filtering of the systematics then our averages will remain stable regardless of which part of the record we select for estimating the average:

$$\hat{\mu}(N,i) = \frac{1}{N+1} \sum_{k=i}^{i+N} z(k) \qquad (36)$$

will be nearly constant if N is sufficiently large.

A large power of S(f) at low frequencies f is typical for trends which have not been removed (cf. figs. 3 and 7). Barnes (1976) gives an excellent discussion that in such cases the experimenter cannot make believable statistical estimates. Many learned arguments about what happens at f = 0 have simply overlooked the fact that as one runs out of data, one also loses the justification for the use of statistics. This is the real meaning of the distinction of trends from the higher frequency noise which alone is the genuine raw material for statistics. But there is also this other aspect of the separation. Most applications are concerned only with clock properties for time intervals which are shorter than a certain limit of usefulness given by the nature of the problem (e.g. re-synchronization interval, or time constant in a feedback loop etc.). Therefore measures, in order to be useful, must have been freed from long term effects (Barnes & Allan, 1964).

DISTRIBUTION OF THE $\bar{y}(k)$

It was obvious from some of the previous examples that the actual $\bar{y}(k)$ may not be Gaussian distributed. If z be again a representative random variable, p(z) the probability density function and P(z) the probability distribution function then we have, of course,

$$P(z) = \int_{-\infty}^{z} p(z)dz \quad \text{and} \quad \int_{-\infty}^{\infty} p(z)dz = 1 \tag{37}$$

One can plot the P(z) as a function of the magnitude of the deviation preferably expressed in units of σ. Figs. 8 and 9 give an estimate of $P(\bar{y}_{1d})$ of our two representative cesium clocks 571 and 837. One notes the non-normal nature of the erratics of 837. Such plots seem to have only very limited usefulness compared with the plot of $\bar{y}(t)$ itself.

THE PAIR VARIANCE (TWO SAMPLE ALLAN VARIANCE)

In our examples which showed actual clock behavior (Cs 571 and 837) we could see that even in the absence of long-term trends, the determiniation of the estimates of variance ($\hat{\sigma}^2$) depends on the number of samples taken. It was therefore proposed by Barnes and Allan (1964) to use N = 2 samples for a variance determination and to average over many such groups to improve the estimate. Such grouping produces a variance which is the limiting case of a sample variance and one idea would be to simply standardize the procedure to N = 2. However, this pair variance has additional benefits. The general sample variance depends not only on the number N but also on any deadtime T between frequency measurements and Allan (1966) introduced the notation $\sigma_y^2(N,T,\tau)$ for this general sample variance. In our case of x(t) sampling we have, of course, no deadtime between frequency measurements because we can compute the $\bar{y}_{\tau=kh}$, cf. our equation (10), for any selected sequence of measurements. Since the pair variance is developed by Allan as the limit for N = 2 of the general sample variance, it contains a factor of 2 in the denominator which is the only difference from the mean square successive difference of $\bar{y}$. The mean square successive difference was used already by von Neumann and others (1941) for the characterization of processes with a drift such as it was encountered in ballistic testing.

Allan has also introduced a variance ratio χ (we assume zero deadtime:

$$\chi(N,\tau) = \frac{\sigma^2(N,\tau)}{\sigma^2(2,\tau)} \quad \text{(later also denoted } B_1(N,\mu)\text{)} \tag{38}$$

For purely random noise $\chi = 1$, and this allows an easy check for the noise type present in one's data. For details one should consult Allan (1966) but it must be kept in mind that the utility of χ is primarily limited in usefulness to cases where "power law" noises predominate. We can see this from the following relationships:

$$\hat{\sigma}^2_{\bar{y}}(2,\tau) = \frac{\sum_{k=2}^{k=n} (\bar{y}_k - \bar{y}_{k-1})^2}{2(n-1)} \tag{39}$$

which for large n becomes approximately

$$\hat{\sigma}^2_{\bar{y}}(2,\tau) \simeq \frac{1}{n}\sum_{k=1}^{n} \bar{y}_k^2 - \frac{1}{n-1}\sum_{k=2}^{n} \bar{y}_k \bar{y}_{k-1} \tag{40}$$

If $\mu(y)$ is assumed to be zero then the first term will be the estimated variance and the second the estimated acvf for a lag of 1 (which is τ, if expressed in s):

$$\sigma^2_{\bar{y}}(2,\tau) = \sigma^2[\bar{y}(\tau)] - c_{\bar{y}}(1) \tag{41}$$

since $c(u) = r(u)\hat{\sigma}^2$ we obtain as a useful approximation for large samples

$$\chi(N,\tau) \sim \frac{1}{1 - r_{\bar{y}}(1)} \tag{42}$$

We can see that Allan's variance ratio can be related easily to the beginning of the autocorrelation function. Considering the great simplicity of the concept and the computation of $\chi(N,\tau)$, if we restrict ourselves to the case N=n and no deadtime, it is a very useful tool in the set of the various stability measures. The pair variance $\hat{\sigma}^2_{\bar{y}}(\tau)$ which is a short hand notation for $\hat{\sigma}^2_{\bar{y}}(2,\tau)$ and much more so the pair sigma $\hat{\sigma}_y(\tau)$ have become a standard for frequency stability specifications and measurements for averaging times greater than about 1s because it is the simplest quantity to measure and it is relatively insensitive to the choice of measurement circumstances. It provides an objective measure even in the presence of some systematics in the record because it uses second differences of phase as we saw in our equation (19). In practice, the incentive to go to more sophisticated aggregates of the $y^m(k)$ for $m \geq 2$ has been rarely given in view of what we said before about the necessary removal of systematics.

An excellent characterization of oscillator performance can be given in the form of sigma-tau double logarithmic plots as shown for our two Cesium clocks 571 and 837 in figs. 10 and 11. In the case of Cs 837 we can see the effect of the systematic variations in an upswing of the graph for large tau's.

MEASURES IN THE FREQUENCY DOMAIN

Up to now we have used the $x(t)$ measures as basis for our performance characterization; we operated in the time domain and did or did not transform our results also into the frequency domain to obtain $S_y(f)$. However, for about $f > 1$ Hz, such measures can also be obtained directly. Since we are now dealing with relatively fast phenomena, the problem of removal of the systematics can be easily circumvented by phaselocking the oscillators together. This assures that $\phi(t)$ and $\dot{\phi}(t)$ will only vary around zero. The phaselock loop (PLL) can act as a frequency or as a phasedetector. If the time constant of the PLL is τ_L then the error signal will be proportional to the phase error $\phi(t)$ for times $\tau << \tau_L$ and proportional to $\dot{\phi}(t)$ for $\tau >> \tau_L$. We can assume that the phase detector is operated in its linear range. A narrow-band low frequency spectrum analyzer or wave analyzer is used to scan the PLL error signal for the AC power which is contained as a function of (Fourier) frequency f. We obtain directly the sideband power of the phase or frequency variations depending on the choice of the PLL time constant. Of course, the various measures are closely interrelated (cf. table 2). However, in contrast to the long-term measures where one prefers $S_y(f)$ for the reasons discussed, here $S_\phi(f)$ is more popular because most oscillators are dominated by white phase noise in a large part of the spectrum and this gives a horizontal line in S_ϕ. Also in the applications, phase noise is the more fundamental concept and $S_\phi(f)$ is more closely related to what is measurable in the laboratory. The unit of $S_\phi(f)$ is again a variance (now in radians^2) per Hz. The sideband power can also be directly expressed as a ratio in terms of the (Suppressed) carrier in db.

Therefore, we see three different phase-time spectral densities in use today: $S_x(f)$ is the time error spectral density which is independent of the signal frequency F. $S_\phi(f)$ is the phase error spectral density which increases with the square of F. Finally, $\mathcal{L}(f)$ is defined variously as the single sideband to carrier power ratio per Hz in the RF spectrum assuming negligible AM. The ratio is also often given in respect to total power with little practical difference for high performance oscillators. In the latter case one can speak of a normalized density measure since the integral over the total RF spectrum (which in f-measure goes from -F to $+\infty$, since f is centered on the carrier) must be one:

$$\int_{-F}^{\infty} \mathcal{L}(f)df = 1 \qquad (43)$$

$\mathcal{L}(f)$ was used by NBS for some time but there is general agreement today that $S_\phi(f)$ is more clearly defined and more directly related to what can be measured with the usual test set-up. For small phase deviations (ϕ << 1 rad) we may use as an excellent approximation

$$\mathcal{L}(f) = \frac{1}{2} S_\phi(f) \qquad (44)$$

A PLL phase detector output voltage V_p (we assume $\tau_L >> \frac{1}{f}$) can be converted into the phase error (assuming the signal is kept in quadrature with the reference which is assumed noise free)

$$\phi = \frac{V_p}{V/rad} \qquad (45)$$

where V/rad is the phase detector sensitivity to phase errors. The variance of ϕ at the wave analyzer setting f is per Hz:

$$\hat{S}_\phi(f) = \left(\frac{V_{p,rms}}{V/rad}\right)^2 /Hz \qquad (46)$$

If the signal frequency is measured after a frequency multiplication of m times then $S_\phi(f)$ will be increased by m^2 (its db measure will increase by 20 log m).

Example: Two equal quartz crystal oscillators with a 5 MHz output are being measured at 25 MHz (m = 5) and we assume a phase detector sensitivity of 2V/rad. We measure with a wave analyzer of 1 Hz bandwidth a rms voltage of 200 nV at 100 Hz. This will give for one oscillator at 5 MHz a

$$\hat{S}_\phi \text{ (100 Hz)} = -140 \text{ db} - 14 \text{ db} - 3 \text{ db (2 oscillators)}$$
$$= -157 \text{ db/Hz}$$

For more details cf. Howe (1976) and Shoaf et al (1973). Practical questions are discussed by M. Fischer (this volume).

Note: We have used only one-sided spectral densities S (with $0 \leq f \leq \infty$). Two sided spectral densities $\underline{S}$ are more popular in theoretical work (where f goes from $-\infty$ to $+\infty$). We have

$$\underline{S} = \frac{1}{2} S \quad \text{and} \tag{47}$$

$$\int_{-\infty}^{\infty} \underline{S}(f)df = \int_{0}^{\infty} S(f)df$$

RANDOM PROCESS MODELING AND FORECASTING

Given a white noise process a_k with zero mean and (constant) variance σ_a^2 we can ask how a more complicated process Z_k such as observed in the $\overline{y}(k)$ of clocks can be simulated on a digital computer. Box and Jenkins (1970) discuss several classes of models:

a) The Moving Average (MA) of order q is given by

$$Z_k = a_k + \Theta_1 a_{k-1} + \Theta_2 a_{k-2} + \ldots + \Theta_q a_{k-q} \tag{48}$$

which can be written with the aid of our operator notation

$$Z_k = \Theta(B)a_k \quad \text{with } \Theta(B) = 1 + \Theta_1 B + \ldots + \Theta_q B^q \tag{49}$$

where $\Theta(B)$ is the MA operator.

For finite q this process is always stationary.

b) The Autoregressive Process of order p (AR) is

$$Z_k = \Phi Z_{k-1} + \ldots + \Phi_p Z_{k-p} + a_k \tag{50}$$

which in our short hand becomes

$$(1 - \Phi_1 B - \ldots - \Phi_p B^p)Z_k = a_k \text{ or } \Phi(B)Z_k = a_k \tag{51}$$

where $\Phi(B)$ is the AR operator. The magnitude and sign of the Φ_i determines the degree of internal correlation of the process and it is seen that a large variety of processes can be obtained since e.g. for large i and large positive Φ_i a low frequency component of S(f)

results, whereas large negative coefficients at small i must produce high frequency components.

c) A mixed model can combine the AR with a MA to represent an ARMA model

$$\phi(B)Z_k = \theta(B)a_k \tag{52}$$

With such a model almost any spectrum can be obtained.

d) This process can be further modified to include integration (which produces a random walk type behavior in the Z_k) by modeling a difference of the process. This could be written as the ARIMA model of order (p, d, q)

$$\phi(B)\nabla^d Z_k = \theta(B)a_k \tag{53}$$

We note that we have a stochastic trend of order d if $\nabla^d Z_k$ is stationary. In the presence of such trends, differencing reduces $\hat{\sigma}^2(\nabla^i Z)$ to a minimum for $i = d$ then it increases again (cf our formulae 6 and 12). This leads to our discussion around equations 14 and 15.

An equivalent formulation would be to model a random variable w_k

$$\phi(B)w_k = \theta(B)a_k \tag{54}$$

and then to sum the w_kd times to obtain the Z_k.

Now assume that we have found a reasonably simple model which produces clock noise of the same kind (with the same measures such as the acvf, $\hat{S}$, etc) as observed in the clock under test. Assume that a sufficiently long series of clock data is available and has been represented by our model Z_k up to time k. We call the $\hat{Z}_k(\ell)$ the forecast at k for a lead of ℓ. Under the assumption that the errors in the estimated model parameters are small and that the model remains unchanged in the future, a so-called optimum forecast can be obtained (optimum on a minimum mean square error basis):

$$E\left[Z(k+\ell) - \hat{Z}_k(\ell)\right]^2 = \min. \tag{55}$$

However, in practice these assumptions are rarely valid. Box and Jenkins (1970) emphasize these points:

a) It is essential to separate trends before modeling.

b) When the model is not adequate, simple visual extrapolation is the best method of forecasting.

One may also note that the above approach is completely equivalent to the "filter" approach as sketched in fig. 12. The coefficients of the digital filter (as well as the operators Φ & Θ) contain the same information as the spectral density S(f). However, as Barnes (1976) emphasizes, it is easy to go from the filter and the model to S(f) and to a time domain measure such as $\sigma_y(\tau)$ but the opposite is not possible, at least in the general case.

In summary it must be stressed that any theoretical basis available must be utilized for the removal of systematics instead of building models which are solely based on purely statistical fits. This is true even if the theoretical ideas are most general which may still be sufficient, e.g., to explain the puzzling flicker noise (cf. Percival, 1976).

Models have an important place in diagnostics, simulation and systems optimization. In this regard, Box and Jenkins (1970) also discuss the estimation of a system's transfer function from an available model.

CONCLUSIONS AND RECOMMENDATIONS

In order to obtain sound measures of clock performance it is necessary to:

a) Specify the conditions of measurement such as f_h, the systems bandwidth; n, the number of measurements; F, the frequency of the signal; T, dead time if any, in the case of frequency measurements; the experimenters trust in the results if not expressed as confidence based on n; the environmental conditions, etc.

b) Remove any obvious systematics such as drifts and state them separately before the random part of the clock performance (the erratics) is analyzed. Failure to do so leads to unnecessary complications and often to erroneous results.

c) Determine and state, if possible, the environmental sensitivity in coefficients of sensitivity to pressure, temperature, vibration, acceleration, magnetic field, etc.

d) State the clock erratics in the same language (time or frequency domain) in which the needs can be identified. A conversion is possible but problematic, particularly from time domain to frequency domain.

For the important case of power law spectra, conversion can be accomplished by means of Table 2. State any observed time dependency of the statistics.

1) Time Domain. The two sample Allan variance $\sigma_y^2(\tau)$ (or rather the "pair" sigma $\sigma_y(\tau)$) has become a de facto standard. A double log plot of $\sigma_y(\tau)$ contains most of the information of possible interest in timekeeping. The practical range is for $\tau > 1s$.

2) Frequency Domain. For $\tau < 1s$ it is generally easier and more reliable to determine $S_\phi(f)$ directly with a phase detector and wave analyzer. $\mathcal{L}(f)$ which is often used, is a practically equivalent measure; it "looks better" for a given oscillator by 3 db. $S_\phi(f)$ is the recommended measure.

ACKNOWLEDGEMENT

I am obliged to Mr. Dave Allan of National Bureau of Standards, Dr. Len Cutler of Hewlett Packard Company, and Dr. U. Hübner of PTB for valuable comments and suggestions.

REFERENCES

1. Allan, D. W. (1966) "Statistics of Atomic Frequency Standards", Proc. IEEE, Vol. 54, pgs 221-230.

2. Allan, D. W. (1974) "The Measurement of Frequency and Frequency Stability of Precision Oscillators", Proc. 6th PTTI Planning Meeting, pgs 109-142 and NBS TN 669.

3. Audoin, C. and Lesage, P. (1975) "Instabilité de frequence des oscillateurs", L'Onde Electrique, Vol. 55/2, pgs 82-89.

4. Barber, R. E. (1971) "Short-Term Frequency Stability of Precision Oscillators", BSTJ, Vol. 50/3. pgs 881-915.

5. Barnes, J. A., and Allan, D. W. (1964) "Effects of Long-Term Stability etc.", Proc. IEEE-NASA Symposium on Short-Term Frequency Stability.

6. Barnes, J. A. et al (1971) "Characterization of Frequency Stability", IEEE Trans. I&M, Vol. 20/2, pgs 105-120.

7. Barnes, J. A. (1976) "Models for the Interpretation of Frequency Stability Measurements", NBS TN 683.

8. Baugh, R. A. (1971) "Frequency Modulation Analysis with the Hadamard Variance", Proc. 25th ASFC, pgs 222-225.

9. Box, G. E. P. and Jenkins, G. M. (1970) "Time Series Analysis, Forecasting and Control", Holden-Day, San Francisco.

10. CCIR Study Group VII Report 580, ITU, Geneva.

11. Cutler, L. S. and Searle, C. L. (1966) "Some Aspects of the Theory and Measurement etc.", Proc. IEEE, Vol. 54/2, pgs 136-154.

12. Fischer, M. C. (1976) "Frequency Stability Measurement Procedures", Proc. 8th PTTI Planning Meeting.

13. Jordan, Ch. (1965) "Calculus of Finite Differences", Chelsea, NY, (CH. VIII).

14. Hamming, R. W. (1962) "Numerical Methods for Scientists and Engineers", McGraw-Hill.

15. Howe, D. A. (1976) "Frequency Domain Stability Measurements", NBS TN 679.

16. Jenkins, G. M. and Watts, D. G. (1968) "Spectral Analysis and its Applications", Holden-Day, San Francisco.

17. Kramer, G. (1975) "Digitale Bestimmung von Schwankungsspectren", PTB Jahresbericht Teil II, p. 157.

18. Lesage, P. and Audoin, C. (1973) "Uncertainty Due to the Finite Number etc.", IEEE Trans. I&M, Vol. 22, pgs 157-161.

19. Lindsey, W. C. and Chie, C. M. (1976) "Theory of Oscillator Instability Based Upon Structure Functions", Proc. IEEE, Dec.

20. Loeb, H. W. (1972) "Efficient Data Transformation in Time Domain Spectrometry", IEEE Trans. I&M, Vol. 21/2, pgs 166-158.

21. Neumann, J. V., Kent, R. H., Bellison, H. R. and Hart, B. I. (1941) "The Mean Square Successive Difference", Annals of Mathematical Statistics, Vol. 12, pgs 153-162.

22. Percival, D. B. (1976) "A Heuristic Model of Atomic Clock Behavior", Proc. 30th ASFC.

23. Rovera, G. (1974) "Stabilità di Frequenza", Editrice Tecnico Scientifica, Pisa, Italy (Obtained from IEN, Torino).

24. Rutman, J. (1974) "Relations Between Spectral Purity and Frequency Stability", Proc. 28th ASFC, pgs 160-165.

25. Rutman, J. (1974) "Characterization of Frequency Stability", IEEE Trans. I&M, Vol. 23/1, pgs 40-48.

26. Rutman, J. and Sauvage, G. (1974) "Measurement of Frequency Stability etc.", IEEE Trans. I&M, Vol. 23/4, pgs 515-518.

27. Shoaf, J. H., Halford, D. and Risley, A. S. (1973) "Frequency Stability Specification and Measurement", NBS TN 632.

28. Vessot, R., Mueller, L. and Vanier, J. (1966) "The Specification of Oscillator Characteristics etc.", Proc. IEEE, Vol. 54/2, pgs 199-207.

29. Wiener, N. (1949) "Time Series", MIT Press (ISBN 0 262 73005 7).

APPENDIX

The Various Uses of Spectrum Analyzers

The output signal of a standard frequency generator can be written as a time varying voltage:

$$V(t) = A(t) \cdot \sin [\Omega \cdot t + \phi(t)] \rightarrow = \int_0^t \dot{\phi} dt$$

↓	↓	↓	↓
RF Spectrum A.	AM Detector	Phase Detector	FM Detector
	↓	↓	↓
	Wave Analyzer	Wave Analyzer	Wave Analyzer
↓	↓	↓	↓
$S_{RF}(f)$	$S_A(f)$	$S_\phi(f)$	$S_{\dot{\phi}}(f)$

Notes: The center of $S_{RF}(f)$ will be at $f = F$. $(2\pi F = \Omega)$

The wave analyzer scans from f = 1Hz up to about 100 kHz in most practical applications.

Table 2

Power Law Noise Models With Their Respective Stability Measures

Noise Type	α	Spectrum $S_y(f)$	Slope S_x	Slope S_y	Time Domain $\sigma_y^2(\tau)$	Slope $\sigma_x(\tau)$	Slope $\sigma_y(\tau)$	μ
White Phase	2	$h_2 f^2$	0	+2	$\frac{3f_h \cdot h_2}{4\pi^2 \cdot \tau^2}$	0	- 1	-2
Flicker Phase	1	$h_1 f$	-1	+1	$[3\ln(2\pi f_h \tau) + 1.038] \frac{h_1}{(2\pi\tau)^2}$			
White FM	0	h_0	-2	0	$\frac{h_0}{2 \cdot \tau}$	1/2	-1/2	-1
Flicker FM	-1	h_{-1}/f	-3	-1	$2\ln 2 \cdot h_{-1}$	1	0	0
Random Walk FM	-2	h_{-2}/f^2	-4	-2	$\frac{4\pi^2}{6} \cdot h_{-2} \cdot \tau$	3/2	1/2	1

Notes:

1. Slope refers to double logarithmic plots of S(f) or $\sigma(\tau)$ in which the noise types are distinguished as pieces of straight lines with the slope given.

2. Allan's (1965) μ corresponds with the parameter μ as it is used here. However, his α is our α - 2 because we refer to S_y instead of S_x. This is the new convention.

Table 2 (Continued)

3. $S_y(f) = (2\pi f)^2 S_x(f) = \frac{1}{\Omega^2} S_{\dot{\phi}}(f) = (\frac{f}{F})^2 S_\phi(f)$

4. $S_\phi(f) = (\frac{F}{f})^2 S_y(f) = \frac{1}{(2\pi f)^2} S_{\dot{\phi}}(f) = \Omega^2 S_x(f)$, See also equation (8).

Table 3

Long Term Performance of Ten Cesium Clocks at the USNO

Oscillator	Days	Linear Frequency Drift Removed (Per Year in 10^{-13})	Sample Variance (After Drift Removed) (in 10^{-26})
Cs 346/1C	1200	+ 3.5	4.0
Cs 532/1C	1200	-	2.9
Cs 549/1	1200	-	2.9
Cs 571/1C-2	1185	+ 1.5	0.7
Cs 591/1	1138	+ 4.1	3.1
Cs 654/1C-2	972	- 4.9	0.8
Cs 660/1C-2	952	-	6.4
Cs 783/1C-2	845	3)	14.2
Cs 834/1C-2	727	-	1.5
Cs 837/1C-2	726	- 7.2	4.4

Notes:

1. The data refer to the $\bar{y}(1^d)$ which are measured in reference to A.1(USNO, MEAN).

2. The units with a "-2" designation are high performance units (004).

3. The large variance of Cs 783/1C-2 is caused by a large non-linear frequency drift which was not removed.

4. Dashes indicate that no significant drift was found in a straight line fit.

(Data courtesy of D. Percival, USNO)

DAILY CLOCK ERRORS FOR WHITE FREQUENCY NOISE

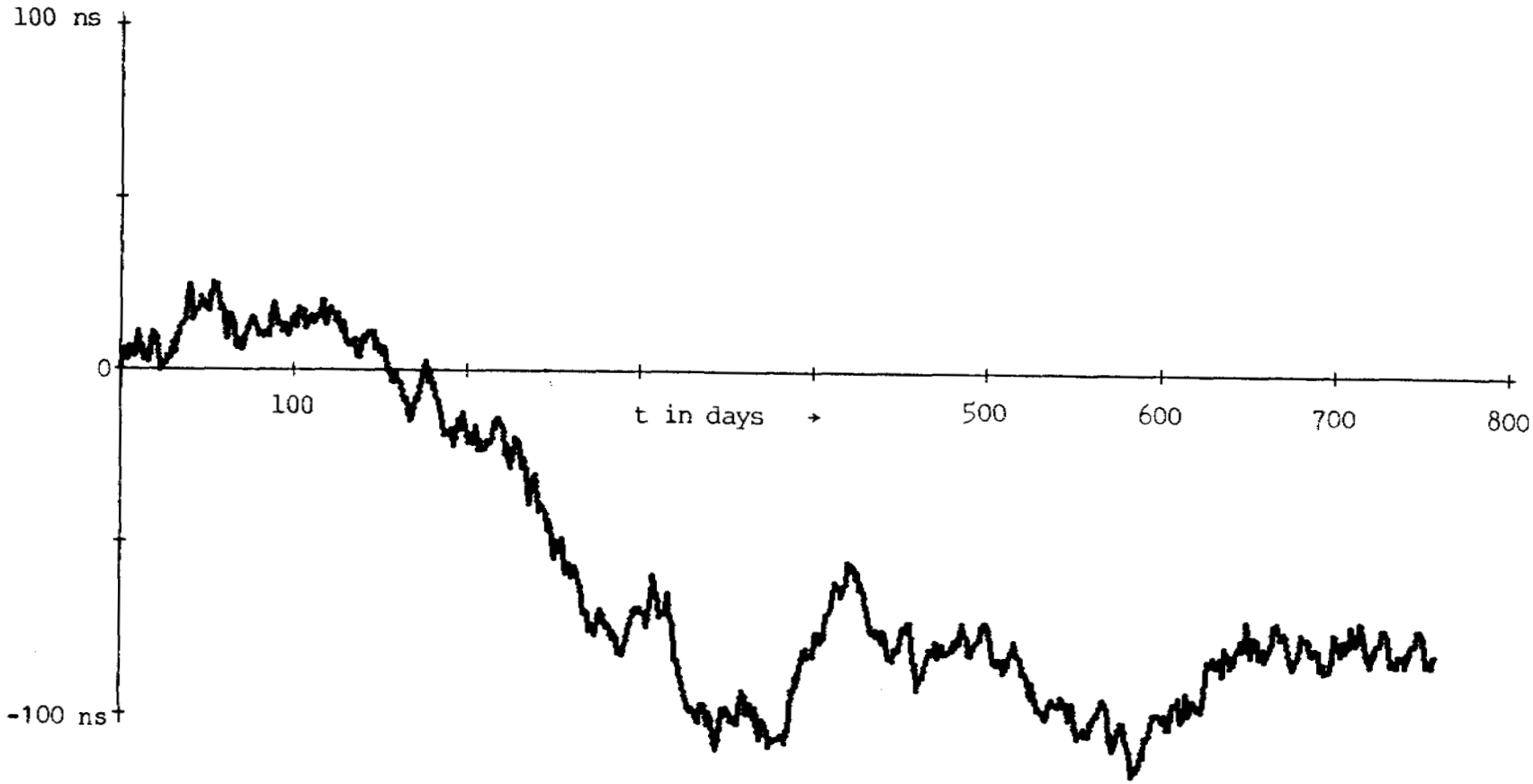


Fig. 1
516

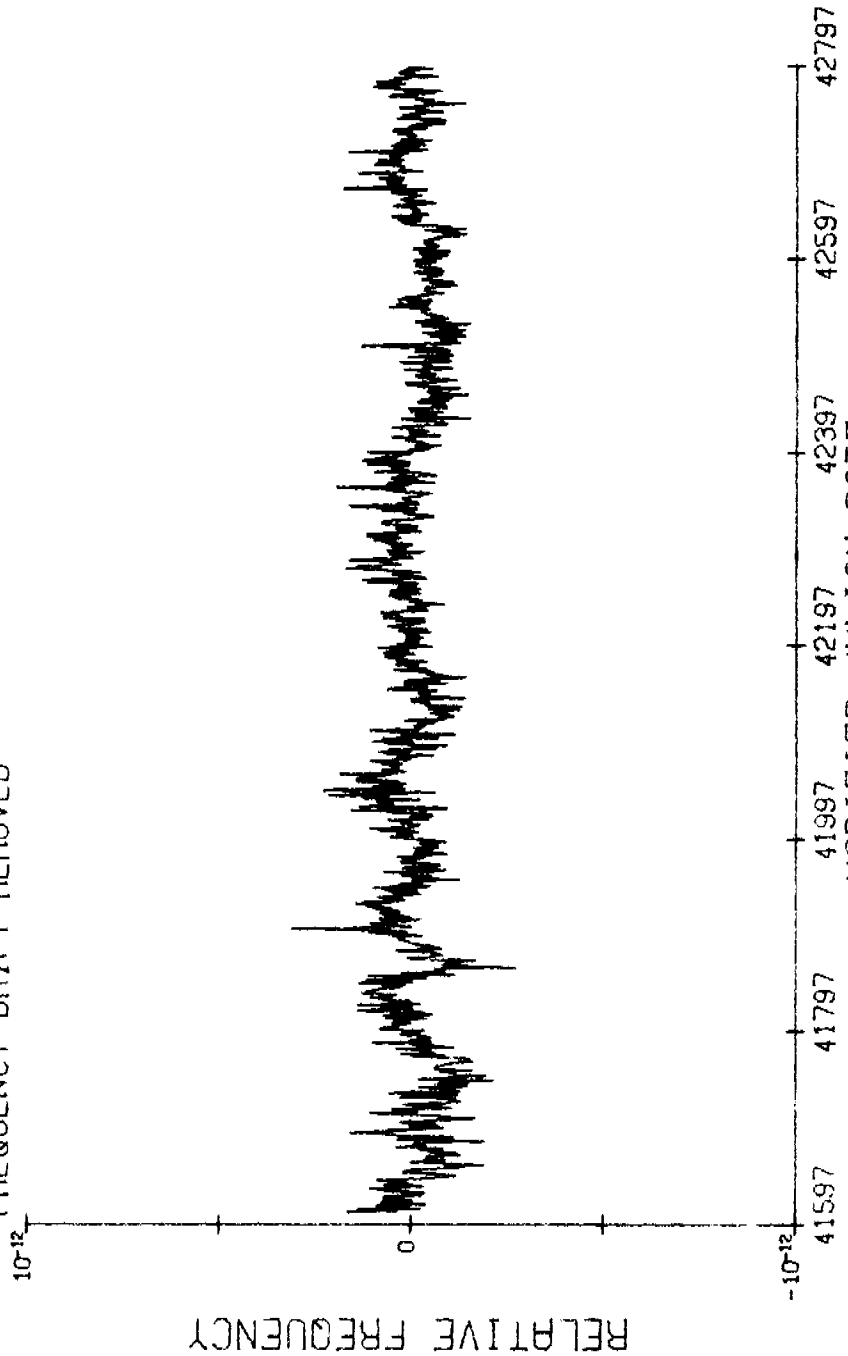

CS 571/1C-2 VS. A.1(USNO.MEAN)
FREQUENCY DRIFT REMOVED
10⁻¹²
0
-10⁻¹²
RELATIVE FREQUENCY
41597
41797
41997
42197
42397
42597
42797
MODIFIED JULIAN DATE

Fig. 2

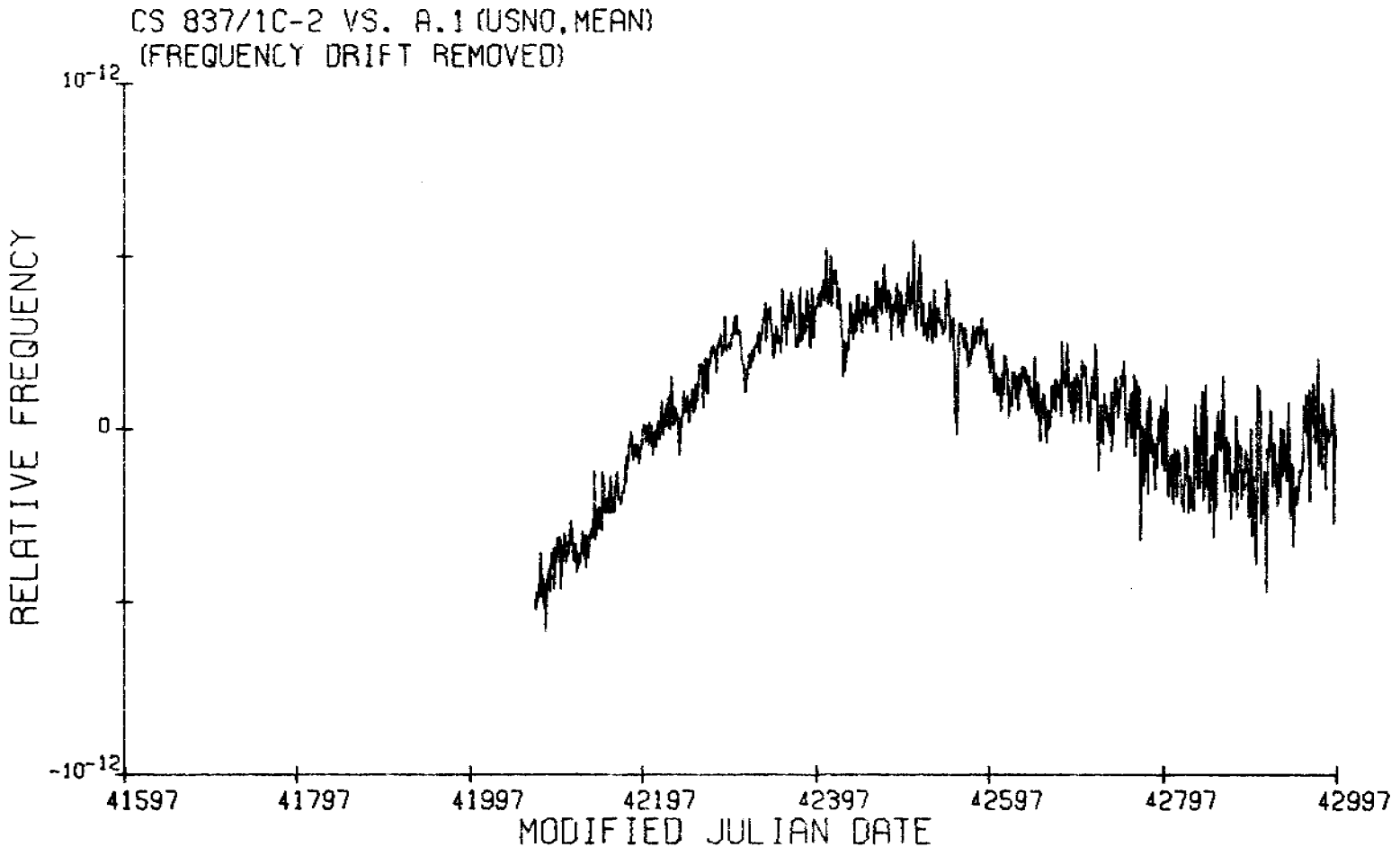

CS 837/1C-2 VS. A.1 (USNO,MEAN)
(FREQUENCY DRIFT REMOVED)
10^-12
0
-10^-12
RELATIVE FREQUENCY
41597
41797
41997
42197
42397
42597
42797
42997
MODIFIED JULIAN DATE

Fig. 3

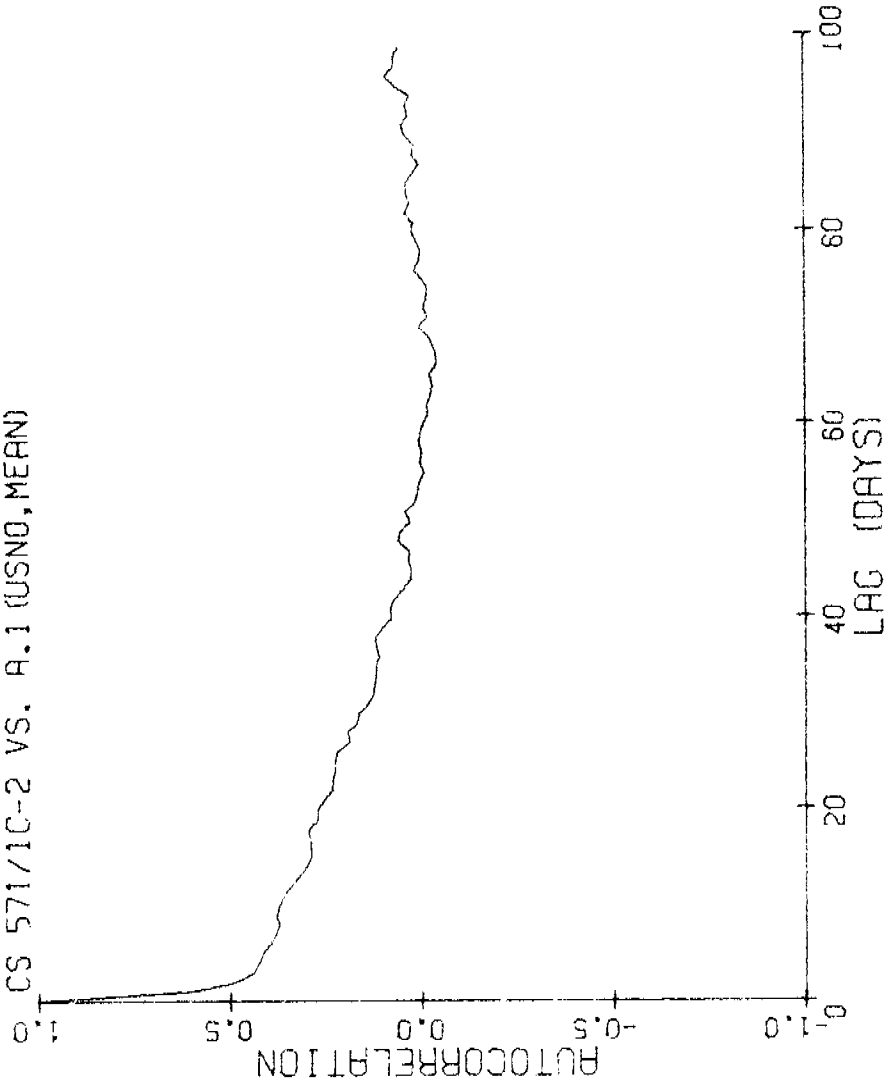

CS 571/1C-2 VS. A.1 (USNO,MEAN)
AUTOCORRELATION
1.0
0.5
0.0
-0.5
-1.0
0
20
40
60
80
100
LAG (DAYS)

Fig. 4

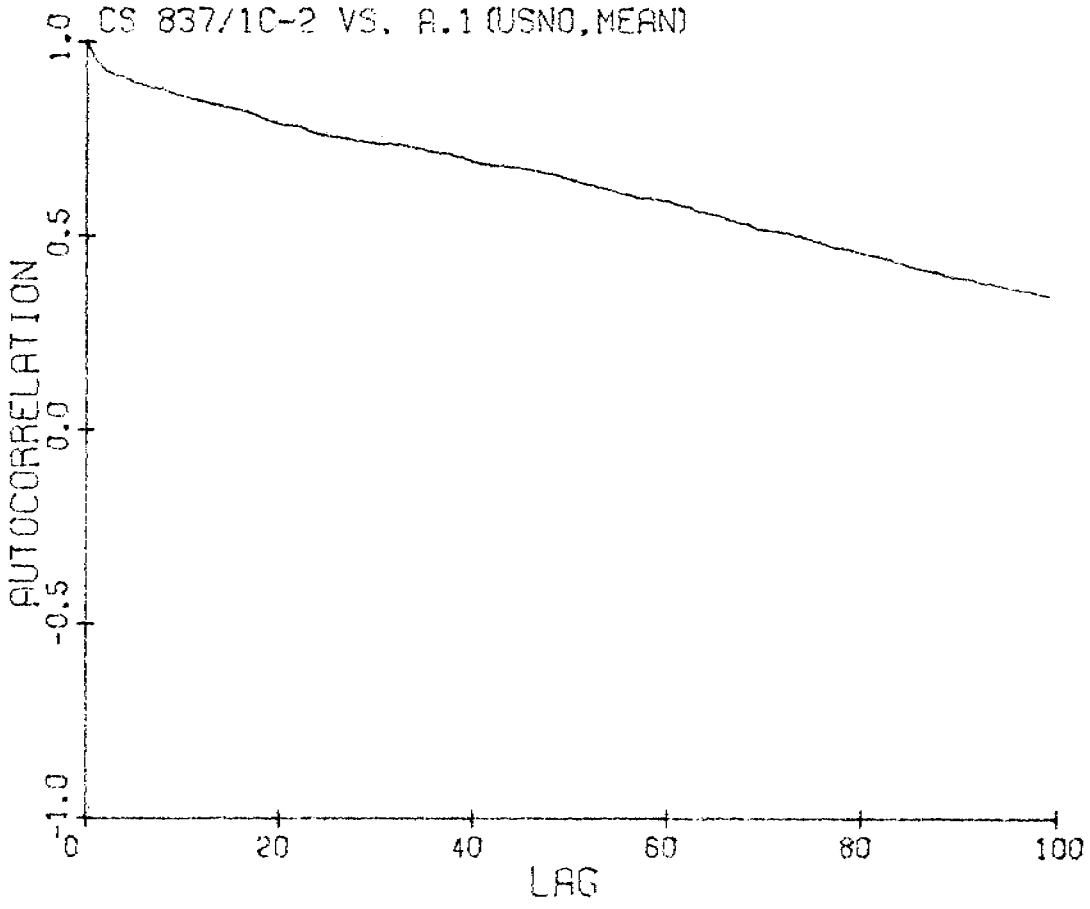

CS 837/1C-2 VS. A.1 (USNO,MEAN)
AUTOCORRELATION
1.0
0.5
0.0
-0.5
-1.0
0
20
40
60
80
100
LAG

Fig. 5

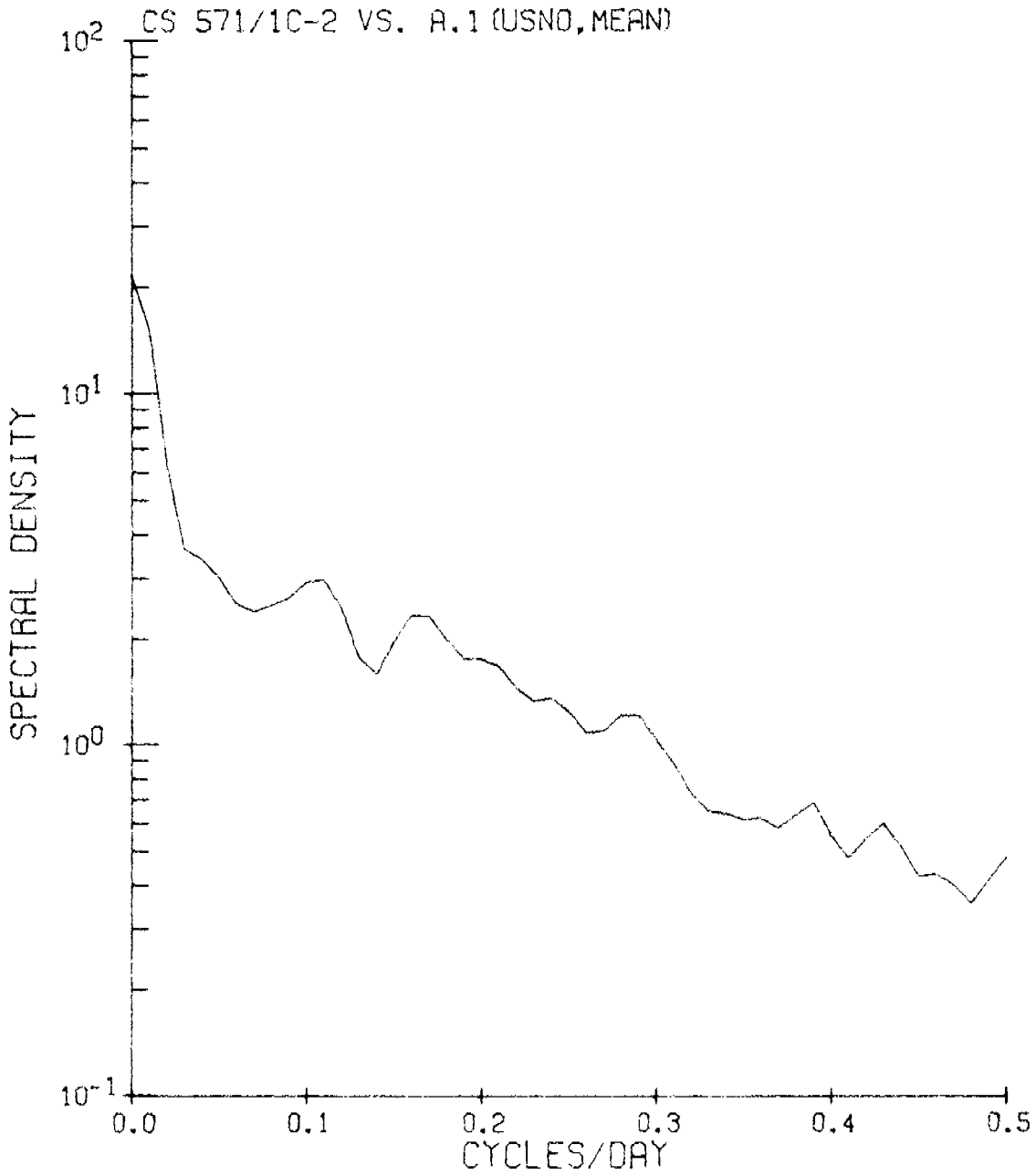

CS 571/1C-2 VS. A.1 (USNO,MEAN)
SPECTRAL DENSITY
10^2
10^1
10^0
10^{-1}
0.0
0.1
0.2
0.3
0.4
0.5
CYCLES/DAY

Fig. 6

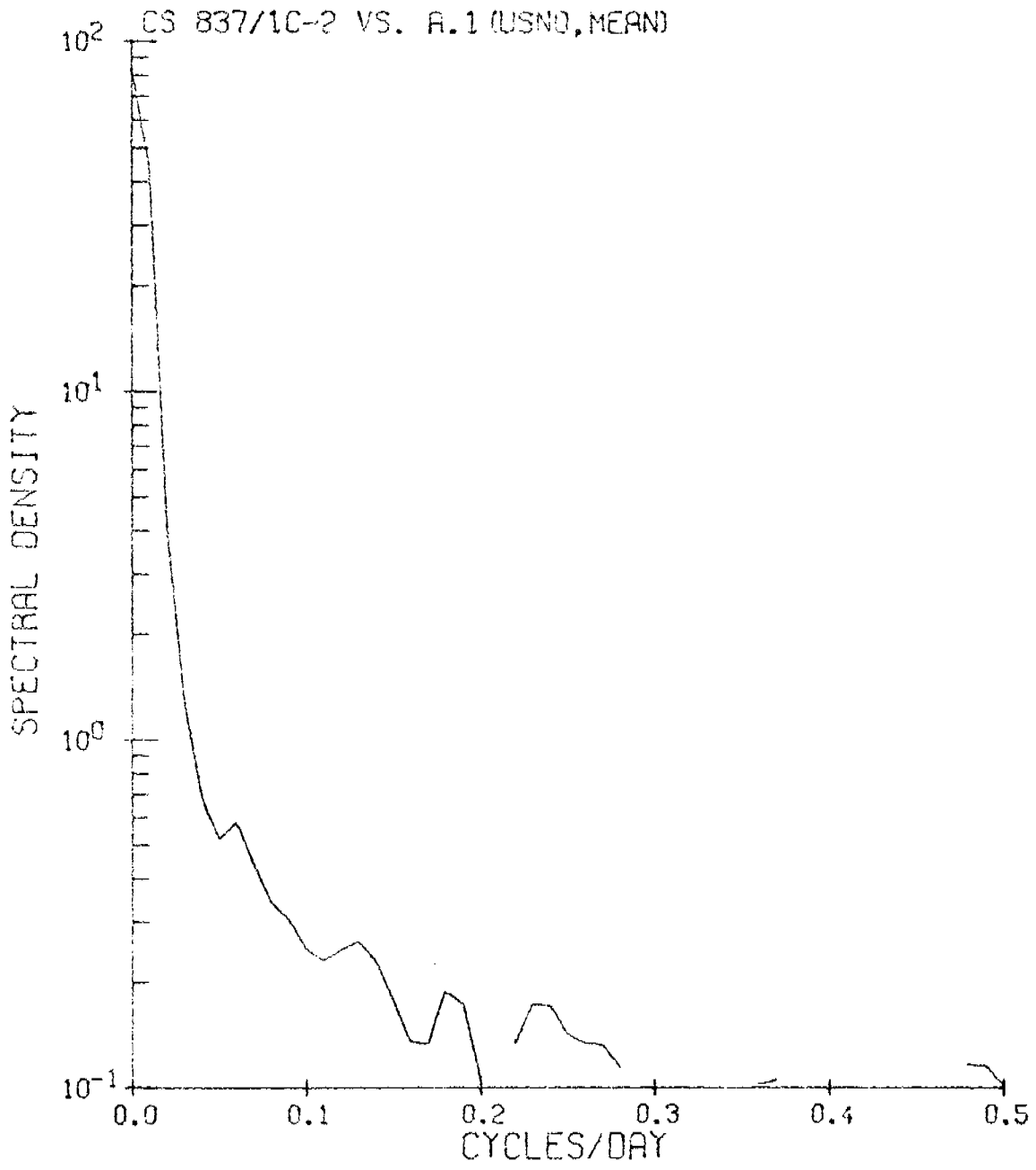

CS 837/1C-2 VS. A.1 (USNO, MEAN)
SPECTRAL DENSITY
10^2
10^1
10^0
10^{-1}
0.0
0.1
0.2
0.3
0.4
0.5
CYCLES/DAY

Fig. 7

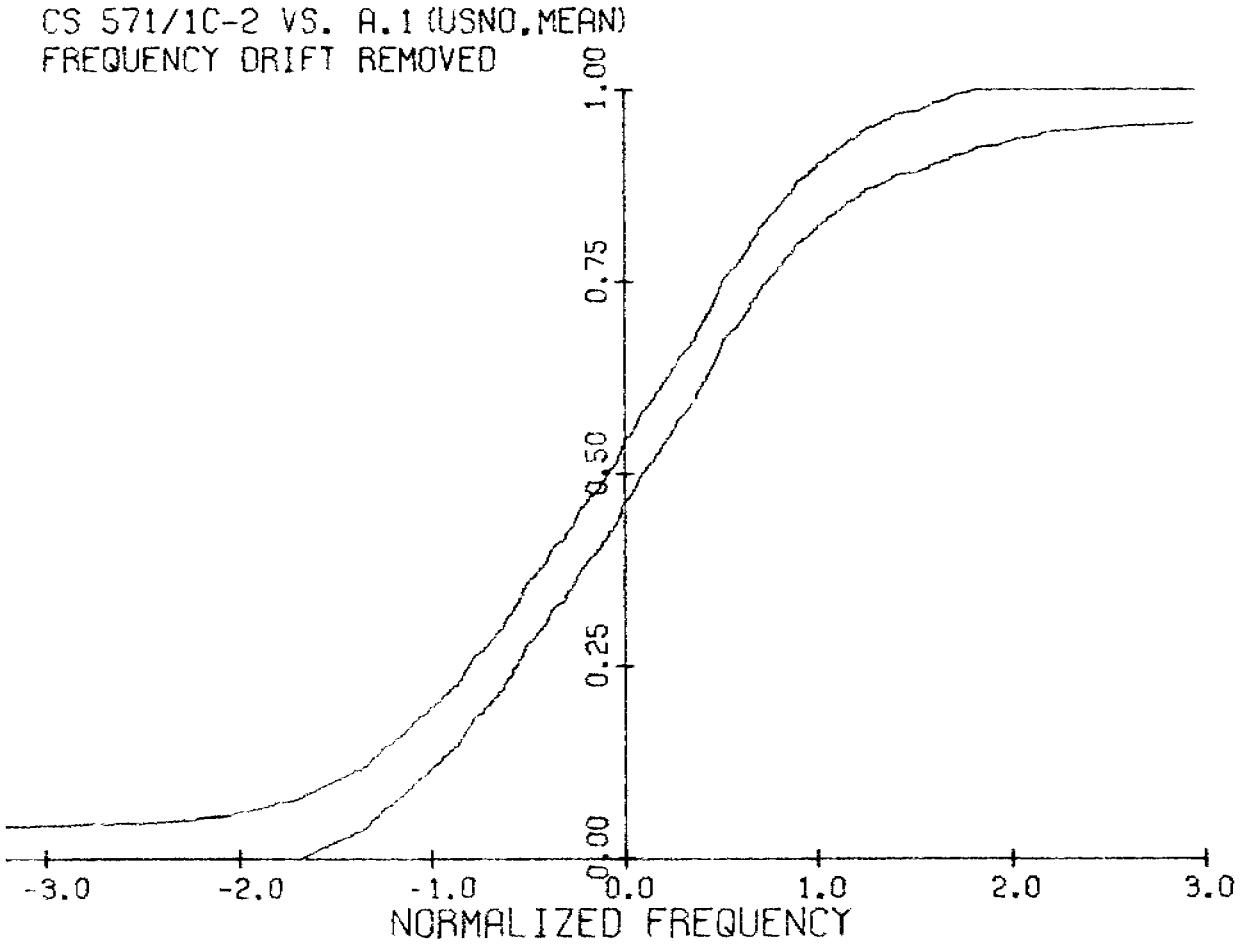

CS 571/1C-2 VS. A.1 (USNO.MEAN)
FREQUENCY DRIFT REMOVED
1.00
0.75
0.50
0.25
0.00
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
NORMALIZED FREQUENCY

Fig. 8

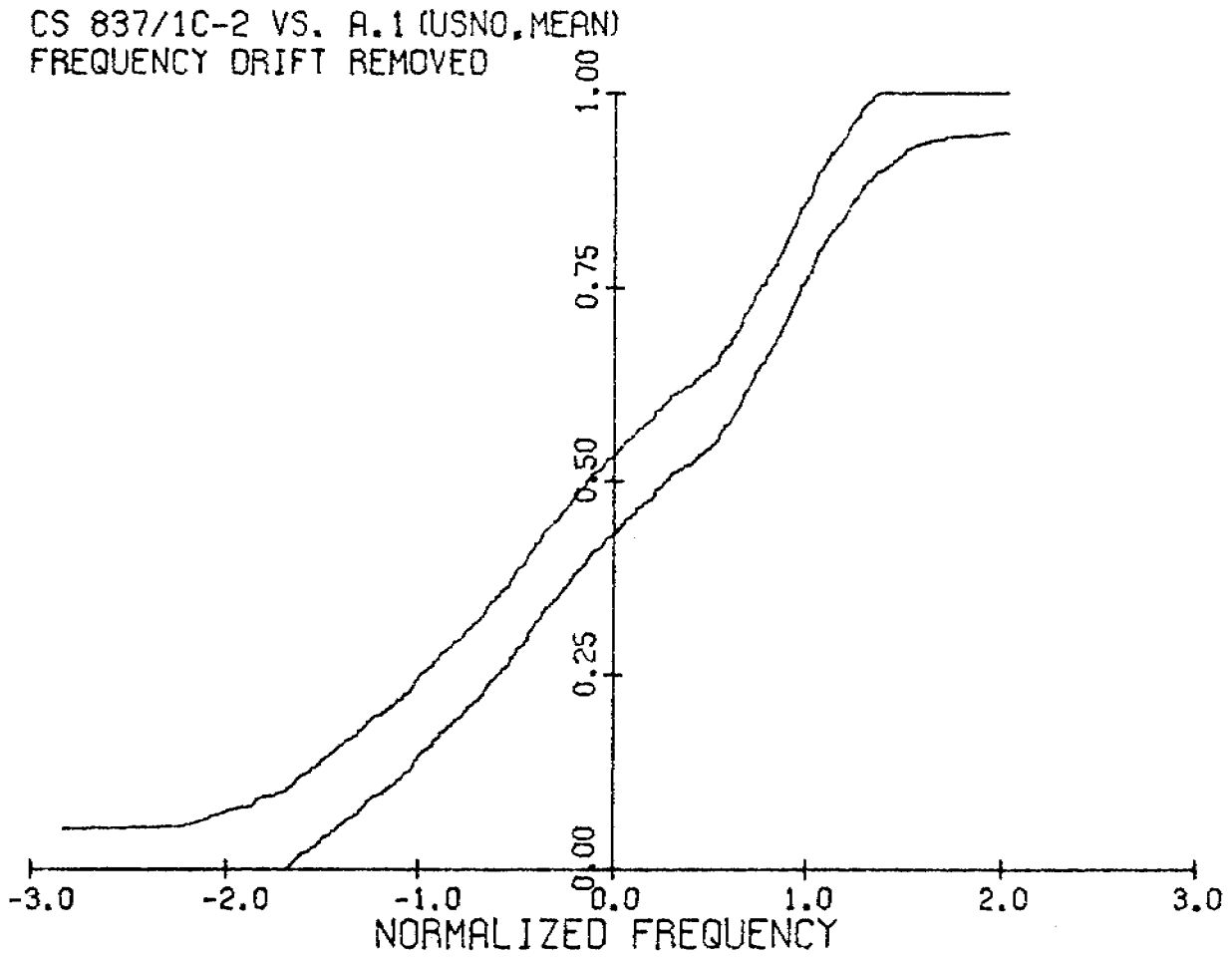

CS 837/1C-2 VS. A.1 (USNO,MEAN)
FREQUENCY DRIFT REMOVED
1.00
0.75
0.50
0.25
0.00
-3.0
-2.0
-1.0
0.0
1.0
2.0
3.0
NORMALIZED FREQUENCY

Fig. 9

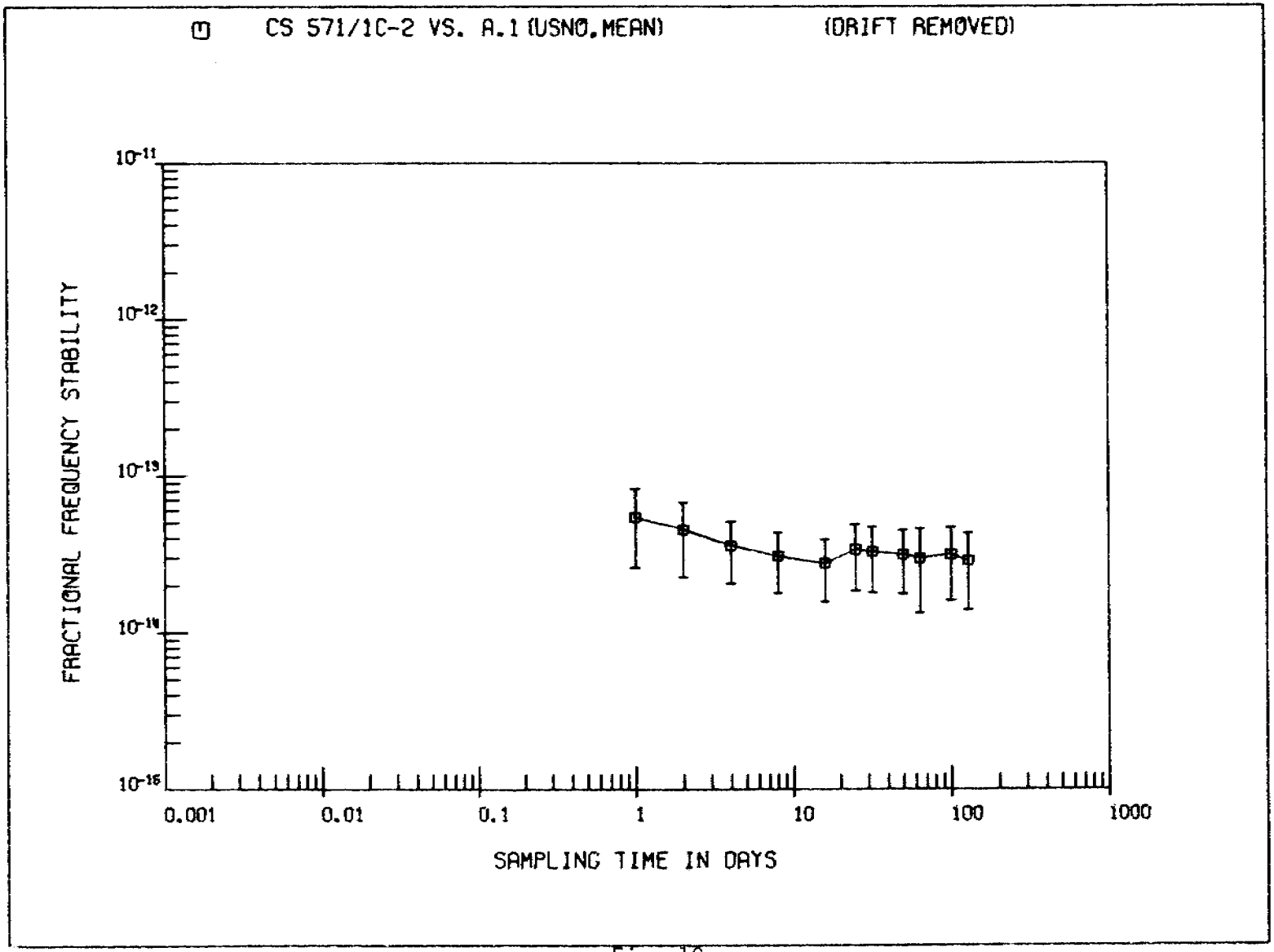

CS 571/1C-2 VS. A.1 (USNO, MEAN)
(DRIFT REMOVED)
FRACTIONAL FREQUENCY STABILITY
10^{-11}
10^{-12}
10^{-13}
10^{-14}
10^{-15}
0.001
0.01
0.1
1
10
100
1000
SAMPLING TIME IN DAYS

Fig. 10

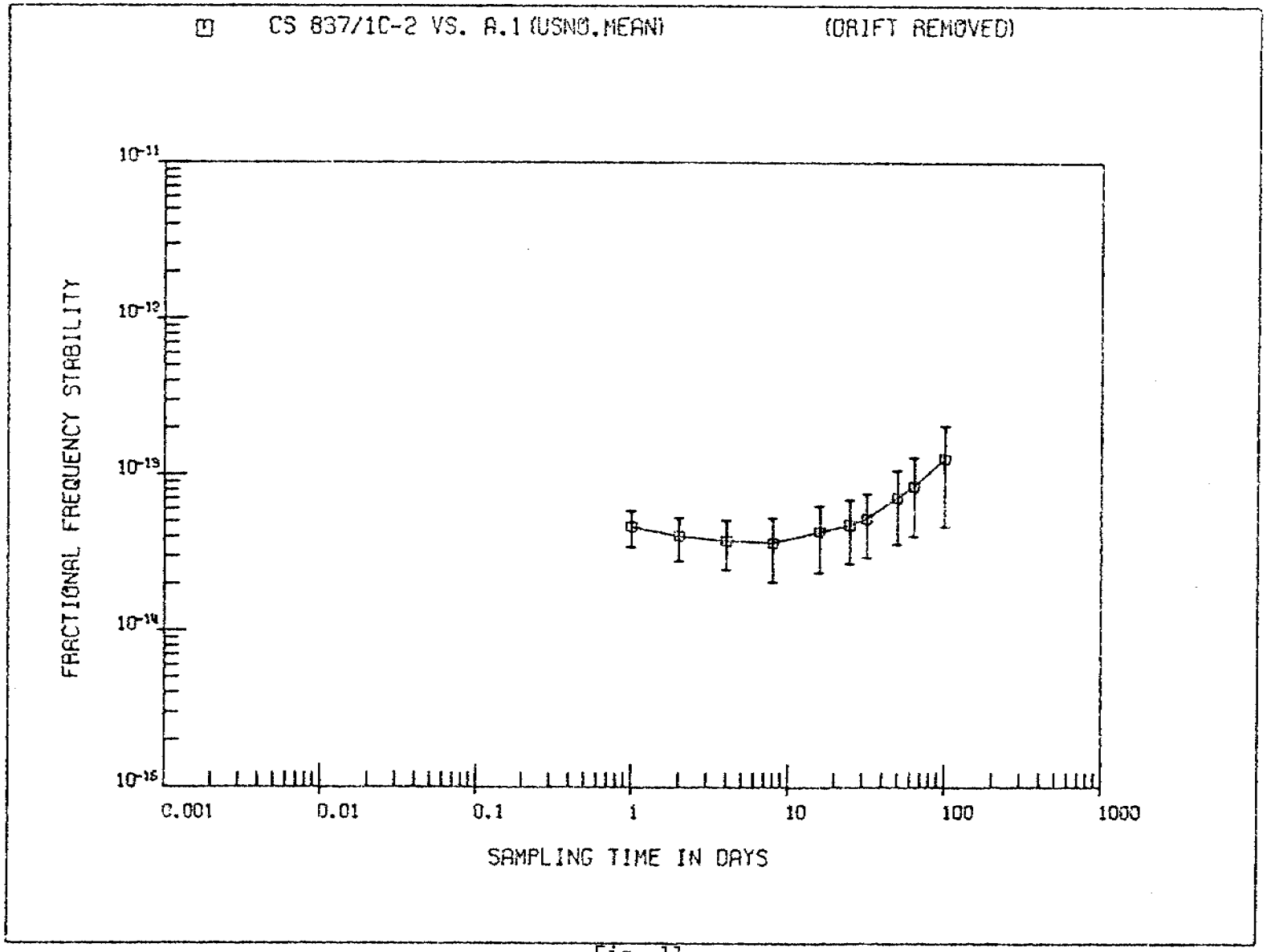
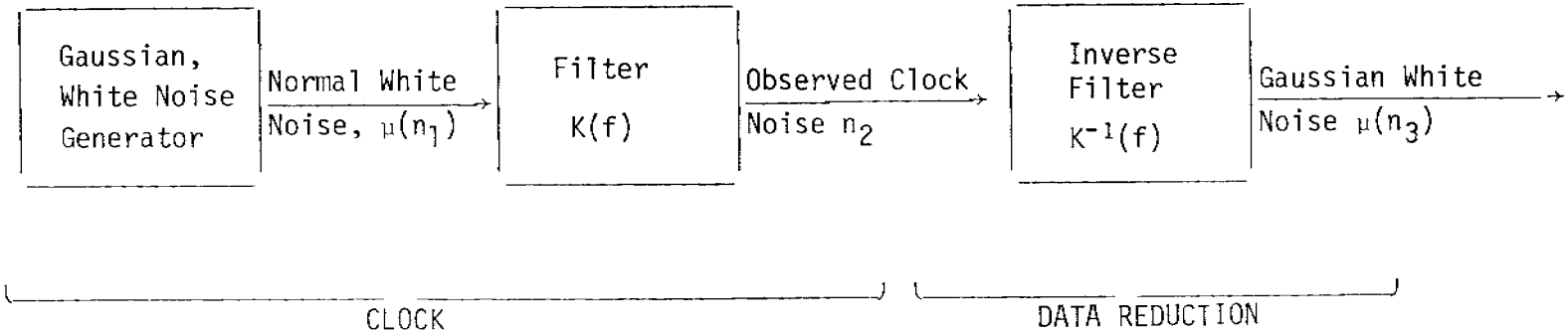

CS 837/1C-2 VS. A.1 (USNO,MEAN)
(DRIFT REMOVED)
FRACTIONAL FREQUENCY STABILITY
10^-11
10^-12
10^-13
10^-14
10^-15
0.001
0.01
0.1
1
10
100
1000
SAMPLING TIME IN DAYS

Fig. 11



Clock Modeling; The Filter Approach

Procedure: Determine $K^{-1}(f)$ from the data ("whitening")

Invert K^{-1} into K(f)

Use final state of K^{-1} for the initial conditions of K(f)

Use the mean $\mu(n_3)$ as input to K(f) to obtain an optimum prediction.

Fig. 12