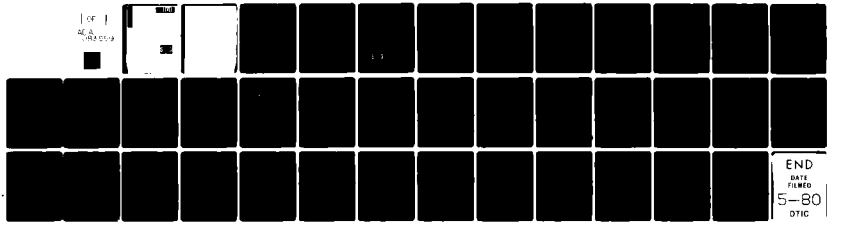
AD-A083 559 NAVAL SURFACE WEAPONS CENTER DAHLGREN VA F/G 12/1
LEAST-SQUARES POLYNOMIAL FITS TO GAUSSIAN RANDOM WALK SEGMENTS,(U)
JAN 80 P L YOUNG, P J FELL
UNCLASSIFIED NSWC/TR-79-300 NL



END
DATE FILMED
5-80
DTIC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: NSWC TR-79-300 ✓	2. GOVT ACCESSION NO.: AD-A083 559 / 3. RECIPIENT'S CATALOG NUMBER:
4. TITLE *(and Subtitle)*: LEAST-SQUARES POLYNOMIAL FITS TO GAUSSIAN RANDOM WALK SEGMENTS	5. TYPE OF REPORT & PERIOD COVERED: / 6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Phillip L. Young and Patrick J. Fell	8. CONTRACT OR GRANT NUMBER(s): 12 [illegible]
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Surface Weapons Center (K10), Dahlgren, VA 22448	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Defense Mapping Agency, Washington, DC 20360	12. REPORT DATE: Jan[illegible]1980 / 13. NUMBER OF PAGES: 40
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*	15. SECURITY CLASS. *(of this report)*: UNCLASSIFIED / 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

gaussian random sequences
least-squares polynomial approximation
one-dimensional random walks

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

This report addresses the least-squares (l.s.) polynomial approximation of gaussian random walk (r.w.) segments. The r.w.'s treated are assumed to be generated by the summation of zero mean, stationary, gaussian random sequences of real numbers. Expressions are developed that relate the statistics of these stationary underlying sequences to the statistics of the associated coefficients and residuals of fit of the approximating l.s. polynomials. When more than one segment of a given r.w. has been approximated, the statistical relationship between the residuals →

(Continues)

next page

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

411567

JCD

20. Continued

of the fits to the different segments is explicitly defined. Finally, an example using the main results is presented. It involves calculating the statistics of the fit residuals when the given r.w. is generated by the fractional frequency errors of an atomic clock.

FOREWORD

The work reported herein was conducted in the Astronautics and Geodesy Division of the Naval Surface Weapons Center under the sponsorship of the Defense Mapping Agency. Appreciation is expressed to C. Branch, J. Headley, and R. Strachen for computer programming support.

This report has been reviewed and approved by R. J. Anderle, Head, Astronautics and Geodesy Division.

Released by:

R. T. RYLAND, Head
Strategic Systems Department

CONTENTS

I. INTRODUCTION

A one-dimensional random walk (r.w.) is a sequence, $\{Z_n\}$, whose terms are given by

$$Z_n = \sum_{K=1}^{n} U_K , \tag{1.1}$$

where $\{U_K\}$ is a random sequence of real numbers. A segment of the random walk $\{Z_n\}$ is simply a set of the form, $\{Z_n\}_{n=N_1}^{N_2}$, consisting of $(N_2 - N_1 + 1)$ consecutive terms of $\{Z_n\}$. In this report we will confine our attention to r.w.'s whose underlying random sequences, $\{U_K\}$, are stationary, gaussian, and zero mean.

The usual interpretation of a one-dimensional r.w. may be stated as follows. Suppose that a particle is constrained to move along the x-axis. If the particle starts at the origin and at time $t_K = Kh$, $K = 1, 2, \ldots$, moves $|U_K|$ units to the right if U_K is positive or to the left if U_K is negative, then Z_n represents the position of the particle at time t_n. The constant, h, denotes the time interval between steps. Thus, using this interpretation, we may think of $\{U_K\}$ and $\{Z_n\}$ as time sequences; i.e., let $U_K = U(t_K)$ and $Z_n = Z(t_n)$.

In many applied problems that involve r.w.'s, it is necessary to model segments of an r.w. as simple algebraic functions. Assuming that $Z_n = Z(t_n)$, polynomials of the form,

$$P_d(t) = \sum_{j=0}^{d} a_j t^j , \quad d = \text{degree of } P_d(t) , \tag{1.2}$$

are particularly convenient models for various applications. Given an r.w. segment, a polynomial of degree d may be fit to the points of the segment, using some fitting criterion. Throughout this report, we will use the least-squares (l.s.) criterion for all polynomial fits; i.e., the coefficients of the polynomial will be determined such that the sum of the squares of the fit residuals is minimized. The fit residuals, r_i, are simply the differences between the points of the segment and the values of the fit polynomial at corresponding times. We will take these residuals in the sense,

$$r_i = Z_i - \hat{Z}_i , \tag{1.3}$$

where $\hat{Z}_i = P_d(t_i)$.

It is important to note that by fitting (in the least-squares sense) a polynomial to an r.w. segment we are not performing linear estimation. Instead, we are performing linear approximation. When you perform linear l.s. estimation to determine the coefficients of a fitting polynomial on the basis of a particular set of data, you make two assumptions. First, you assume that the data you are given consist of observations that have been corrupted by noise. Second, you implicitly assume that, if the noise on the data was not present, then the estimation algorithm would yield a perfect fit; i.e., that the polynomial is a perfect model for the process that generated the basis (uncorrupted) data. Thus, the coefficients determined by the l.s. estimation algorithm will be in error due to the effects of the data noise, in the sense that they will differ from the exact values they would have if no data noise was present.

In contrast, when you perform linear approximation using the l.s. criterion you make two very different assumptions. First, you assume that the given data set has not been corrupted by noise. Second, you assume that the form of the fitting polynomial is not equivalent to the process that generated the data but is simply an approximation to that process. Thus, the coefficients determined by l.s. approximation are not in error; they are correct in the sense that any other polynomial of the same degree would yield a larger sum of the squares of the residuals, taken with respect to the given data set. It follows that it is not meaningful to talk about the statistics of errors in fit coefficients, as we would if l.s. estimation was being performed.

We may, however, ask meaningful statistical questions about l.s. polynomial fits to r.w. segments, even though the coefficients are exact in the sense explained in the preceding paragraph. For example, what are the statistics of the fit residuals for a l.s. polynomial fit of degree d to an r.w. segment of length m? If more than one segment of a given r.w. has been approximated, how are the residuals of the fits to the different segments related statistically? For an arbitrary segment of length m of a given r.w., what is the variance of each of the coefficients in the fit polynomial? These, and other questions to be discussed later, are well posed and have definite answers once the statistics of the underlying random sequence, $\{U_K\}$, have been specified.

The primary purpose of this report is to answer the three questions just mentioned. We begin with a brief discussion of one-dimensional r.w.'s whose underlying random sequences, $\{U_K\}$, are stationary, gaussian, and zero mean. The special case of first degree l.s. polynomial fits to r.w. segments is then examined in some detail. Next, the machinery needed to treat the general case of fits of arbitrary degree is developed. Finally, we present a nontrivial example of our main results by calculating the statistics of the fit residuals when the given r.w. is generated by the fractional frequency errors of an atomic clock.

II. ONE-DIMENSIONAL (GAUSSIAN) RANDOM WALKS

One-dimensional random walks (as defined by Equation (1.1)) arise in many areas of both the "hard" and "soft" sciences. Fluctuations in stock market indices, the size of biotic populations, the fortunes of a gambler, and the errors in a clock are common examples. The motivation for the work we are reporting on comes from the need to model segments of the r.w.'s that occur in some physical systems because of random errors at the input of a summing subsystem. These random errors are often composites, composed of errors generated by several separate error mechanisms such as thermal noise, stiction (i.e., stick-and-slip resulting from nonuniform friction in mechanical assemblies), and random environmental influences on sensitive components. When digital processing is included in the system, roundoff and truncation are also sources of error within the system.

It was mentioned in section I that we will restrict our attention to r.w.'s whose underlying random sequences, $\{U_K\}$, are gaussian, stationary, and zero mean. In many physical systems, the composite errors that exist at the input of the summing subsystem may be assumed to be gaussian. The mathematical basis for this assumption is a remarkable theorem known as the central limit theorem. For our discussion, a useful form of this theorem is (from Reference 1):

Theorem. "Let $x_1, x_2, x_3, \ldots, x_m$ be a series of statistically independent random variables having arbitrary distributions for which the means $\mu_1, \mu_2, \ldots, \mu_m$ and variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_m^2$ exist. If m is sufficiently large, the sum $\Sigma_{i=1}^{m} x_i$ will be approximately normally distributed [i.e., gaussian] with mean μ and variance σ^2, whether the x_i's are normally distributed [gaussian] or not and where the mean and variance of the sum are given by $\mu = \Sigma_{i=1}^{m} \mu_i$ and $\sigma^2 = \Sigma_{i=1}^{m} \sigma_i^2$, respectively."

It follows that if each U_K (interpreted as the input to a summing subsystem at time t_K) is composed of a large number of independent random errors, $x_1(t_K), x_2(t_K), \ldots, x_m(t_K)$, then, for large m, we are justified in assuming that each U_K is gaussian. (In fact, m may not need to be very large at all before the distribution of U_K becomes essentially gaussian. An excellent example is given by Papoulis on pages 267 and 268 of Reference 2.) The assumption that $\{U_K\}$ is a stationary sequence is based on a requirement that the systems under consideration be in steady state.

We are fortunate that composite errors in many systems may be assumed to be gaussian because gaussian random variables have several properties that simplify any analysis involving them. For example, a gaussian distribution is completely specified by its first two moments (i.e., by its mean and variance). This is not the case for arbitrary distributions. Also, a linear transformation of a set of gaussian random variables yields a set of gaussian random variables. This is clearly not true for sets of random variables with arbitrary distributions, as evident from the central limit theorem. (Demonstrations of the truth of these two statements concerning gaussian random variables may be found in Reference 2.)

The final restriction concerning $\{U_K\}$ that holds throughout this report is that it be a zero mean sequence. This restriction guarantees that the terms of $\{Z_n\}$ (as defined by Equation (1.1)) represent random sums. If underlying sequences with nonzero means were allowed, the resulting r.w.'s would be composed of terms of the form,

$$Z_n = \left(\sum_{K=1}^{n} W_K \right) + nb,$$

where $\{W_K\}$ is a zero mean sequence and b is a constant bias. Clearly, the ramp sequence, $\{nb\}$, may be approximated exactly by a polynomial and so is of no interest in this report.

Now, let's examine the r.w.'s whose underlying sequences satisfy our three restrictions. We will look at two cases.

Case (1): $\{U_K\}$ Uncorrelated (White)

Let $\{U_K\}$ be a white, stationary, zero mean gaussian random sequence, with variance σ_U^2. Then, the r.w. defined by

$$Z_n = \sum_{K=1}^{n} U_K$$

is a zero mean gaussian sequence, and for this case the variance of its *n*th term is given by

$$\sigma_n^2 = E[Z_n^2] = n\sigma_U^2 , \tag{2.1}$$

where E[ ] denotes the expectation operator. The correlation between Z_I and Z_J, $I < J$, is given by the correlation coefficient,

$$\rho_Z(I, J) = E[Z_I Z_J]/\sigma_I \sigma_J = (I/J)^{1/2} . \tag{2.2}$$

In particular, if $J = I + q$,

$$\rho_Z(I, I+q) = \left(\frac{I}{I+q}\right)^{1/2} . \tag{2.3}$$

(Of course, $\rho_Z(J, I) = \rho_Z(I, J)$.)

* * *

Case (2): $\{U_K\}$ Correlated

Let $\{U_K\}$ be a stationary, zero mean gaussian random sequence with variance σ_U^2, and with the correlation between U_K and U_{K+q} given by the correlation coefficient, $\rho_U(q) \neq 0$. Then, the r.w. defined by

$$Z_n = \sum_{K=1}^{n} U_K$$

is a zero mean gaussian sequence, and for this case the variance of its nth term is given by

$$\left.\begin{aligned} \sigma_Z^2(1) &= \sigma_U^2 \\ \sigma_Z^2(n) = E[Z_n^2] &= \left[n + 2 \sum_{i=1}^{n-1} i\rho_U(n-i) \right] \sigma_U^2 , \quad n \geqslant 2 . \end{aligned}\right\} \tag{2.4}$$

The correlation between Z_I and Z_J is given by the correlation coefficient, $(I < J)$,

$$\rho_Z(I, J) = E[Z_I Z_J]/\sigma_Z(I)\sigma_Z(J) = \frac{I + S_I + S_{IJ}}{(I + S_I)^{1/2}(J + S_J)^{1/2}} \tag{2.5}$$

where

$$S_I = 2 \sum_{i=1}^{I-1}{}' \, i\rho_U(I - i),$$

$$S_J = 2 \sum_{i=1}^{J-1}{}' \, i\rho_U(J - i),$$

$$S_{IJ} = \sum_{i=1}^{I}{}' \sum_{j=I+1}^{J}{}' \, \rho_U(j - i)$$

Notice that if $\rho_U(q) = 0$ for q = 1, 2, ..., then Equations (2.4) and (2.5) reduce to Equations (2.1) and (2.2), respectively, as we would expect. Also, $\rho_Z(J, I) = \rho_Z(I, J)$, as in case (1).

* * *

It should be emphasized that when we compute the expectation of a function of terms from an r.w., the implied average is taken over the ensemble of all possible r.w.'s whose underlying sequences have the specified statistics. For example, the variance of the *n*th term of $\{Z_n\}$ in case (1) is interpreted to be the "average" value of the squares of the *n*th terms from all possible r.w.'s whose underlying sequences are white, stationary, zero mean, gaussian and have variance σ_U^2. In both case (1) and case (2), it is clear that $\{Z_n\}$ is a nonstationary sequence.

In Figure 2 and Figure 3 we have examples of the first hundred steps of an uncorrelated r.w. and a correlated r.w., respectively. These examples were prepared as follows. The first hundred steps of a white, stationary, zero mean gaussian sequence with unit variance were chosen from table XXIII of Reference 1. These numbers will be denoted by $\{U_K\}_{K=1}^{100}$, and are plotted in Figure 1a. Next, the first hundred terms of a stationary, zero mean, first-order gauss-markov sequence, with unit variance were generated. The correlation time of this gauss-markov sequence, τ, was chosen to be five times the step increment, h. This second set of numbers will be denoted by $\{V_K\}_{K=1}^{100}$, and are plotted in Figure 1b. They were generated from $\{U_K\}_{K=1}^{100}$ by the difference equation,

$$\left.\begin{aligned} V_1 &= U_1 , \\ V_K &= \rho_V(1) V_{K-1} + \left[1 - \rho_V^2(1)\right]^{1/2} U_K , \quad 2 \leqslant K \leqslant 100, \end{aligned}\right\} \tag{2.6}$$

where $\rho_V(q) = \exp(-qh/\tau) = \exp(-q/5)$. A derivation of Equation (2.6) is given in Reference 3. Finally, $\{U_K\}_{K=1}^{100}$ and $\{V_K\}_{K=1}^{100}$ were each summed to obtain the r.w.'s $\{Z_n\}$ and $\{Y_n\}$, respectively, shown in Figure 2 and Figure 3, where the points have been connected by straight lines.

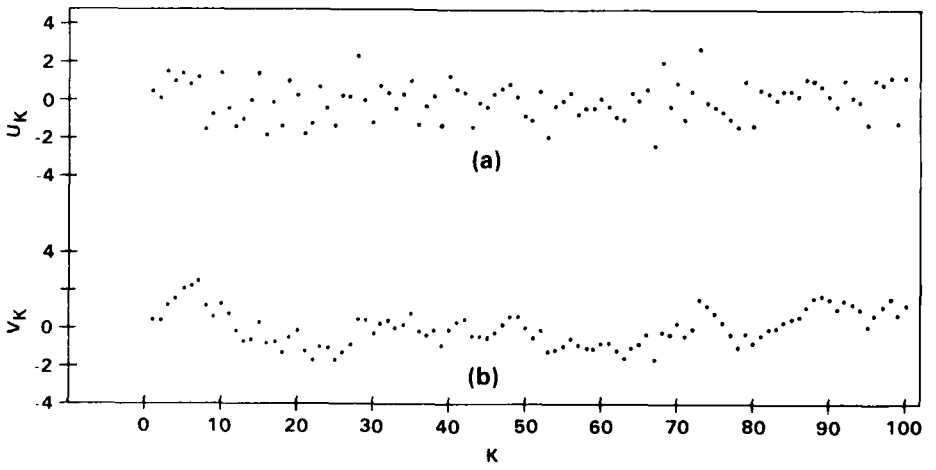


Figure 1

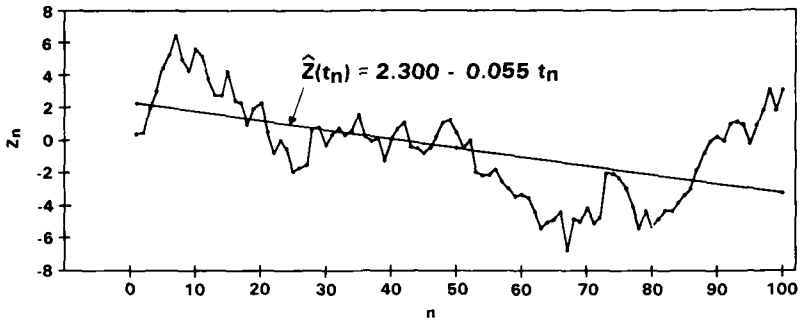


Figure 2

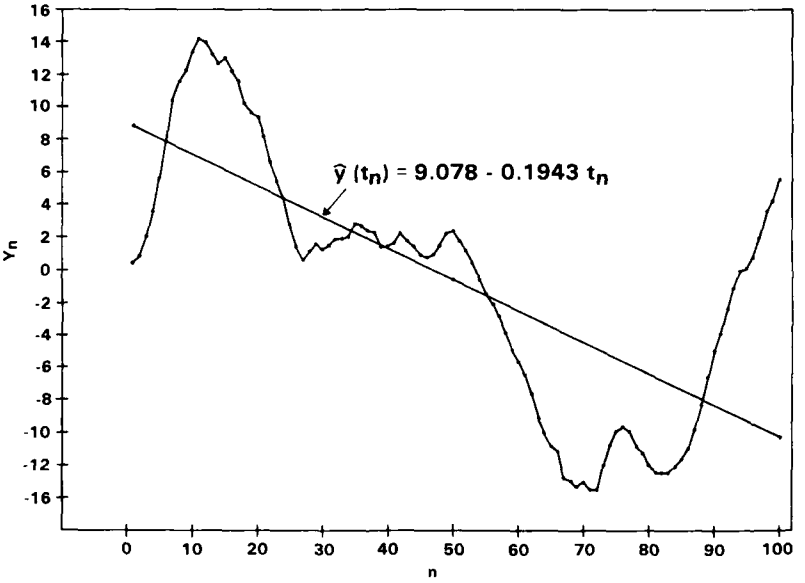


Figure 3

In addition to the r.w.'s, one other item is plotted in Figures 2 and 3. Each r.w. was fit by a first-degree l.s. polynomial, and the line segments defined by the fitted polynomials are plotted. When looking at these figures, it should be remembered that both r.w.'s were generated by underlying sequences with the same variance, i.e., $\sigma_U^2 = \sigma_V^2 = 1.0$. The differences between the two r.w.'s result from the fact that $\{V_K\}$ is correlated and $\{U_K\}$ is not. The two most noticeable differences are that $\{Y_n\}$ is much smoother than $\{Z_n\}$, and the variances of the terms of $\{Y_n\}$ are larger than those of the corresponding terms of $\{Z_n\}$.

The last two plots in this section illustrate the statistics of $\{Z_n\}$ and $\{Y_n\}$. Figure 4a shows the variance of the terms of $\{Z_n\}$ and $\{Y_n\}$, and Figure 4b gives the corresponding standard deviations. Figure 5 is a plot of $\rho_Z(1, J)$, $\rho_Z(25, J)$, $\rho_Z(50, J)$ and $\rho_Z(100, J)$.

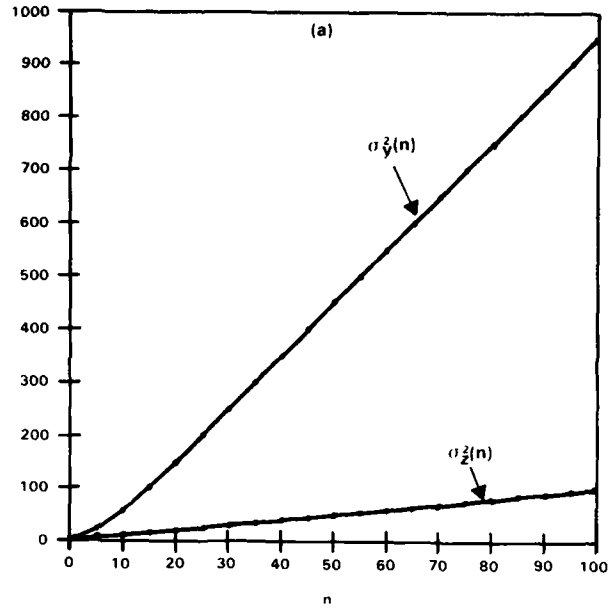


Figure 4a

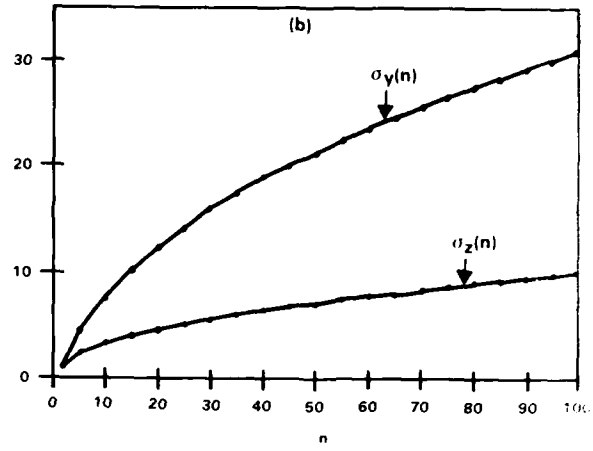


Figure 4b

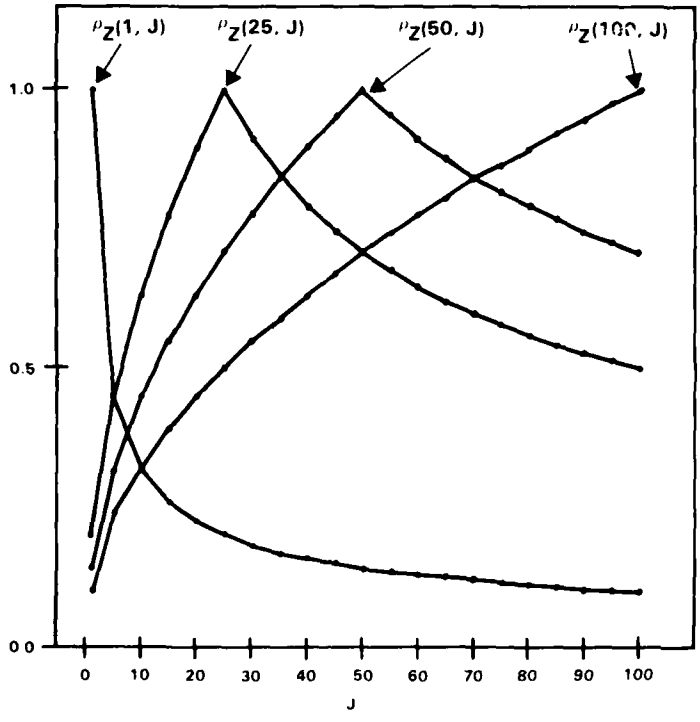


Figure 5

III. FIRST-DEGREE LEAST-SQUARES FITS

In this section, only first-degree l.s. polynomial fits to r.w. segments are considered. This special case is of importance for two reasons. First, the simplicity of first-degree fits allows the derivation of analytic results that are instructive and form the basis for the generalization to fits of arbitrary degree. Second, many practical applications use first-degree fits to r.w. segments, so the development of the fit residual statistics without resorting to matrix notation is computationally useful.

Our immediate objective may be stated as follows. Given a segment of the r.w. $\{Z_n\}$, generated by the underlying sequence $\{U_K\}$, express the residuals of the first-degree l.s. polynomial fit to this segment as linear combinations of the terms of $\{U_K\}$. Once this is accomplished, the statistics of the fit residuals may readily be found in terms of the statistics of $\{U_K\}$.

Consider the r.w. segment, $\{Z_n\}_{n=1}^{N}$, consisting of the first N terms of $\{Z_n\}$. (It will be shown later in this section that no generality is lost by treating the first segment of length N.) The first degree l.s. polynomial fit to this data will be

$$\hat{Z}_\ell = a_0 + a_1 t_\ell, \quad t_\ell = \ell h, \quad \ell = 1, 2, \ldots, N \tag{3.1}$$

where

$$a_0 = \left[\frac{2(2N+1)}{N(N-1)}\right] \sum_{n=1}^{N} Z_n + \left[-\frac{6}{hN(N-1)}\right] \sum_{n=1}^{N} t_n Z_n, \tag{3.2}$$

and

$$a_1 = \left[-\frac{6}{hN(N-1)}\right] \sum_{n=1}^{N} Z_n + \left[\frac{12}{h^2 N(N^2-1)}\right] \sum_{n=1}^{N} t_n Z_n. \tag{3.3}$$

Substituting the expressions for a_0 and a_1 into Equation (3.1) and collecting terms yields

$$\hat{Z}_\ell = \sum_{n=1}^{N} C(\ell, n) Z_n \tag{3.4}$$

where

$$C(\ell, n) = c_0[c_1 + c_2(\ell + n) + c_3 n\ell] \tag{3.5}$$

with

$$c_0 = \frac{2}{N(N-1)}, \quad c_1 = (2N+1), \quad c_2 = -3 \quad \text{and} \quad c_3 = \frac{6}{(N+1)}.$$

But, Z_n is just the sum of the first n terms of $\{U_K\}$ (by Equation (1.1)), so Equation (3.4) may be written in the form,

$$\hat{Z}_\ell = \sum_{n=1}^{N} \sum_{K=1}^{n} C(\ell, n) U_K .$$

This double sum may be written as a single sum by introducing a new set of weighting coefficients. Thus,

$$\hat{Z}_\ell = \sum_{K=1}^{N} D(\ell, K) U_K , \tag{3.6}$$

where

$$D(\ell, K) = \sum_{j=K}^{N} C(\ell, j) . \tag{3.7}$$

Substitution of Equation (3.5) into Equation (3.7) yields

$$D(\ell, K) = d_0[(d_1 + d_2\ell) + (d_3 + d_4\ell)K + (d_5 + d_6\ell)K^2] \tag{3.8}$$

where

$$d_0 = \frac{1}{N(N-1)}, \quad d_1 = (N+1)(N+2), \quad d_2 = -6$$

$$d_3 = -(4N+5), \quad d_4 = \frac{6(N+2)}{(N+1)}, \quad d_5 = 3, \quad d_6 = \frac{-6}{(N+1)} .$$

Finally, from Equation (1.3),

$$r_\ell = Z_\ell - \hat{Z}_\ell , \quad \ell = 1, 2, \ldots, N.$$

So, using Equations (1.1) and (3.6) the fit residuals for the first degree l.s. fit become

$$r_\ell = \sum_{K=1}^{\ell} U_K - \sum_{K=1}^{N} D(\ell, K) U_K$$

or

$$r_\ell = \sum_{K=1}^{N} F(\ell, K) U_K , \tag{3.9}$$

where

$$F(\ell, K) = \begin{cases} 1 - D(\ell, K) & \text{for} \quad 1 \leq K \leq \ell \\ -D(\ell, K) & \text{for} \quad \ell < K \leq N. \end{cases} \tag{3.10}$$

Thus, by introducing a third set of weighting coefficients, the F(ℓ, K)'s, we have met our immediate objective. Before using Equation (3.9) to determine the statistics of the fit residuals, several facts concerning the three sets of weighting coefficients should be pointed out.

From the definitions of C(ℓ, n), D(ℓ, K), and F(ℓ, K) given by Equations (3.5), (3.8) and (3.10), respectively, the following facts may be established:

$$C(\ell, n) = C(n, \ell) \tag{3.11}$$

$$C(\ell, n) = C(N - n + 1, N - \ell + 1) \tag{3.12}$$

$$\sum_{n=1}^{N} C(\ell, n) = 1, \quad \ell = 1, 2, \ldots, N \tag{3.13}$$

$$D(\ell, 1) = 1, \quad \ell = 1, 2, \ldots, N \tag{3.14}$$

$$\sum_{K=1}^{N} D(\ell, K) = \ell, \quad \ell = 1, 2. \ldots, N \tag{3.15}$$

$$F(\ell, 1) = 0, \quad \ell = 1, 2, \ldots, N \tag{3.16}$$

$$\sum_{K=1}^{N} F(\ell, K) = 0, \quad \ell = 1, 2, \ldots, N. \tag{3.16}$$

It might be noted that the computation of all three sets of weighting coefficients depends only on the value of N, the number of terms being fit. The facts presented in Equations (3.11)—(3.16) represent only a few of the results and relations that might prove useful in various applications, but they should serve as a basis for further derivations. For example, if we think of the C(ℓ, n)'s as the elements of an N × N matrix, C, then Equations (3.11) and (3.12) show that C is symmetric about both of its diagonals. Equation (3.13) implies that the rows of C all sum to one, and Equations (3.11) and (3.13) together imply that the columns of C also sum to one.

The C(ℓ, n)'s are particularly useful in different applications since they define the relationship between the value of $\hat{Z}_\ell = \hat{Z}(t_\ell)$ for a first-degree l.s. polynomial fit and the data values, $Z_n = Z(t_n)$, over which the fit was performed. The only assumption made during the derivation of Equations (3.4) and (3.5) was that the data be equally spaced. If we think of C(ℓ, n) as a function of n for some fixed value of ℓ, we see from Equation (3.5) that C(ℓ, 1), C(ℓ, 2), ..., C(ℓ, N) lie on a straight line whose y-axis intercept is $c_0(c_1 + c_2\ell)$, and whose slope is $c_0(c_2 + c_3\ell)$. The constants c_0, c_1, c_2, and c_3 depend only on N.

Similarly, the D(ℓ, K)'s define the relationship between the value of $\hat{Z}_\ell = \hat{Z}(t_\ell)$ for a first-degree l.s. polynomial fit and the values $\{U_K\}_{K=1}^{N}$ of the underlying sequence that generated the r.w. segment over which the fit was performed. In this instance, we see from Equation (3.8) that for fixed values of ℓ, the weighting coefficients D(ℓ, 1), D(ℓ, 2), ..., D(ℓ, N) lie on a quadratic curve. Figures 6 and 7 are plots of the C(ℓ, n)'s and D(ℓ, K)'s, respectively, for the case N = 10, and Figure 8 is a plot of the associated F(ℓ, K)'s for this case. For the sake of clarity, Figure 8 consists of two plots, Figures 8a and 8b, showing F(1, K) through F(5, K) and F(6, K) through F(10, K), respectively.

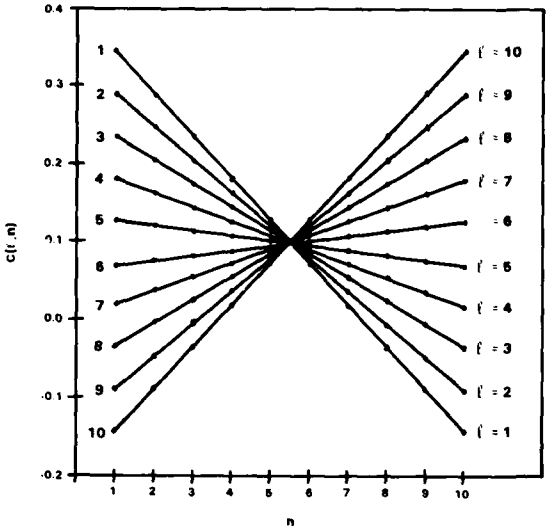


Figure 6

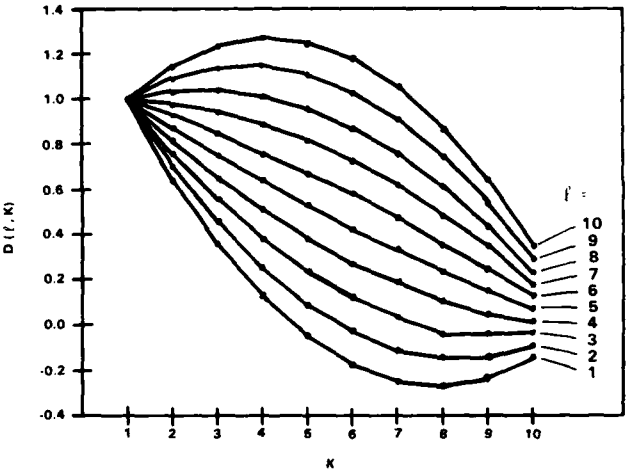


Figure 7

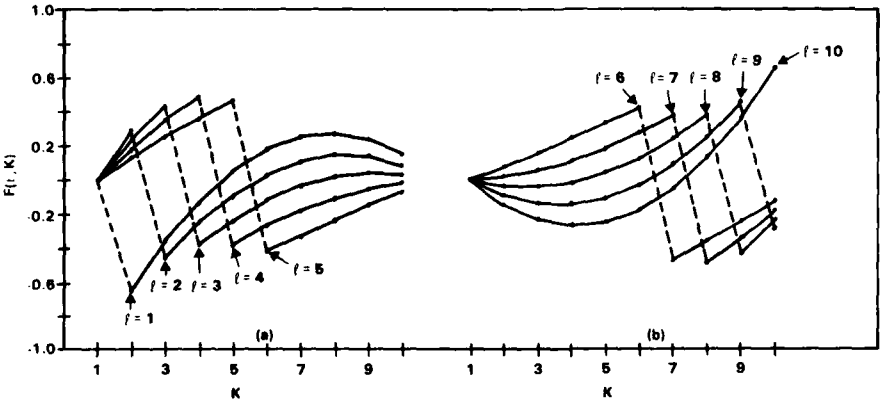


Figure 8

Returning to our main theme, we see from Equation (3.9) that the mean of r_ℓ is zero, since $\{U_K\}$ is a zero mean sequence. The variance of r_ℓ is given by

$$\sigma_r^2(\ell) = E\left[r_\ell^2\right] = \sigma_U^2\left[\sum_{K=1}^{N} F^2(\ell, K) + 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} F(\ell, i)F(\ell, j)\rho_U(j-i)\right]. \quad (3.17)$$

The correlation between r_I and r_J, $I < J$, is given by the correlation coefficient,

$$\rho_r(I, J) = E[r_I r_J]/\sigma_r(I)\sigma_r(J), \quad (3.18)$$

where

$$E[r_I r_J] = \sigma_U^2\left\{\sum_{K=1}^{N} F(I, K)F(J, K) + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N}[F(I, i)F(J, j) + F(I, j)F(J, i)]\rho_U(j-i)\right\}. \quad (3.19)$$

Since Equation (3.9) equates the fit residuals to a linear combination of the U_K's, it follows that the residuals are gaussian and so Equations (3.17), (3.18), and (3.19) completely specify their statistics. These equations represent the answer to one of the primary questions addressed by this study.

Figures 9, 10, and 11 display the variances of the fit residuals for three different values of N. Each of the plots shows the results for $\{U_K\}$ uncorrelated and for $\{U_K\}$ a first-order markov sequence with $\tau_U = 5.0$. In all cases, $\sigma_U^2 = 1.0$ and h = 1.0. The values of N used in Figures 9, 10, and 11 are 10, 15, and 50, respectively. Since only two levels of correlation of the underlying sequence are shown in these plots, it is not possible to conclude much from them concerning the effect on the fit residual statistics of increasing the correlation of $\{U_K\}$. For this reason, Figure 12 has been included. It shows the effect of various values of τ_U on $\sigma_r^2(1)$, $\sigma_r^2(4)$, and $\sigma_r^2(8)$, with N = 15, $\sigma_U^2 = 1.0$, and h = 1.0. The results for $\tau_U = 0$ (i.e., for $\{U_K\}$ uncorrelated) through $\tau_U = 10.0$ are plotted; then a break in the plot is indicated, followed by the values corresponding to $\tau_U = 99.0$ (The fact that the values of $\sigma_r^2(\ell)$ approach zero as $\tau_U \to \infty$ is easily explained. For very large values of τ_U, $\{U_K\}$ looks like a random constant sequence, so $\{Z_n\}$ is almost a random ramp sequence. The more closely $\{Z_n\}$ resembles a ramp the better it will be approximated by a first-degree polynomial.) The last plot associated with the fit residual statistics is Figure 13. It shows correlation coefficients $\rho_r(I, J)$ for I = 1, 10, 20, 25, and J = 1, 2, ..., 50, for the case N = 50, $\sigma_U^2 = 1.0$, h = 1.0, and $\{U_K\}$ uncorrelated.

In the third paragraph of this section, we claimed that no generality is lost by developing the fit residual statistics for fits to r.w. segments of length N by treating the segment $\{Z_n\}_{n=1}^{N}$. At this point, we will show that for first-degree l.s. polynomial fits the fit residual statistics are identical for fits to arbitrary segments of length N of a given r.w. (Our usual assumptions concerning $\{U_K\}$ still obtain, of course.) This result will be demonstrated for l.s. polynomial fits of arbitrary degree in the next section.

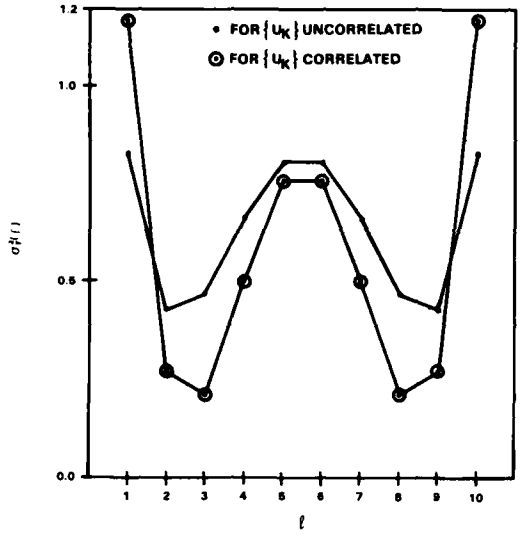


Figure 9

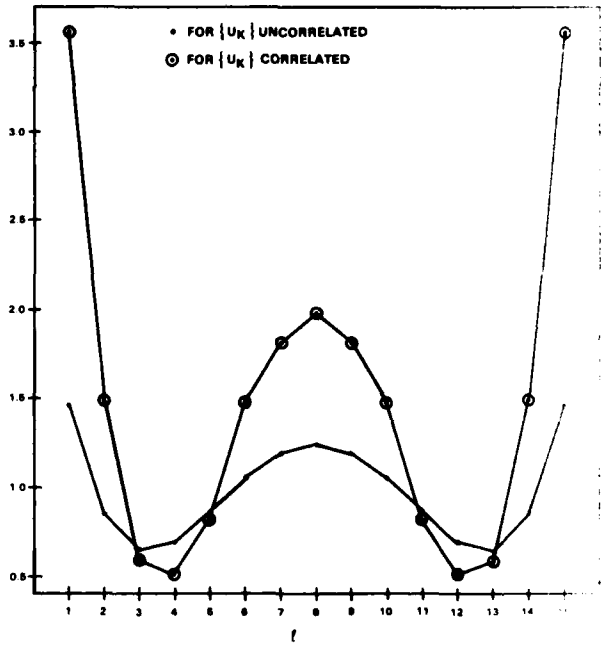


Figure 10

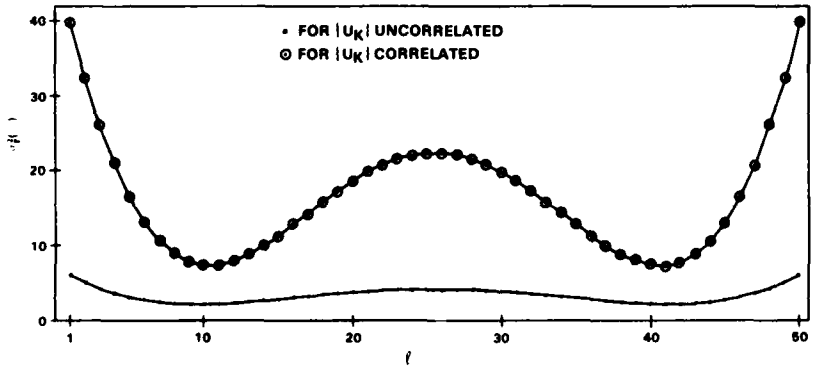


Figure 11

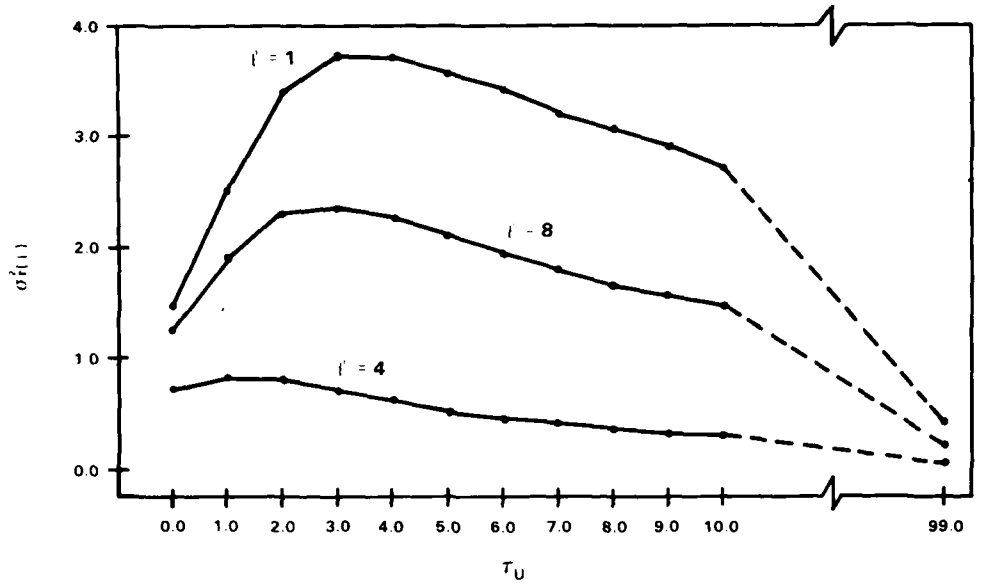


Figure 12

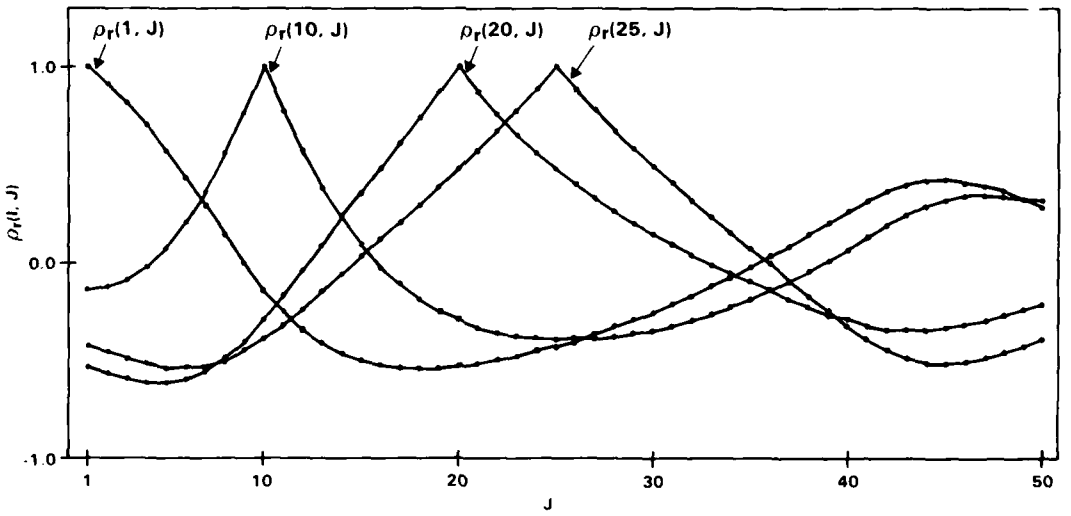


Figure 13

Let $\{Z_n\}$ be an r.w. whose underlying sequence, $\{U_K\}$, is a stationary, zero mean gaussian sequence. Suppose that an arbitrary segment of length N has been chosen from $\{Z_n\}$. We will denote this segment by $\{Z_n\}_{n=\xi+1}^{\xi+N}$. The terms of this segment are given by

$$Z_{\xi+\ell} = \sum_{K=1}^{\xi+\ell} U_K = Z_\xi + \sum_{K=1}^{\ell} U_{\xi+K}, \quad \ell = 1, 2, \ldots, N.$$

The values of the first degree l.s. polynomial fit to this segment, corresponding to the times $t_{\xi+\ell}$, are,

$$\hat{Z}_{\xi+\ell} = \sum_{n=1}^{N} C(\ell, n) Z_{\xi+n} = \left(\sum_{K=1}^{N} D(\ell, K) U_{\xi+K} \right) + Z_{\xi} \sum_{n=1}^{N} C(\ell, n).$$

But, from Equation (3.13),

$$\sum_{n=1}^{N} C(\ell, n) = 1, \quad \text{for} \quad \ell = 1, 2, \ldots, N.$$

Thus,

$$\hat{Z}_{\xi+\ell} = Z_{\xi} + \sum_{K=1}^{N} D(\ell, K) U_{\xi+K}$$

and so, the ℓth fit residual is

$$r_{\ell} = Z_{\xi+\ell} - \hat{Z}_{\xi+\ell} = \sum_{K=1}^{N} F(\ell, K) U_{\xi+K}. \tag{3.20}$$

Since $\{U_K\}$ is stationary, r_{ℓ} is zero mean and the variance and correlations are given by Equations (3.17), (3.18), and (3.19). This completes the demonstration.

Now, the second of our primary questions concerns the relationship between the fit residuals of fits to different segments of the same r.w. Given our usual restrictions on $\{U_K\}$, suppose that two nonoverlapping segments of length N of the r.w. $\{Z_n\}$ have been fit by first-degree l.s. polynomials. To be specific, let $\{Z_n\}_{n=\eta+1}^{\eta+N}$ and $\{Z_n\}_{n=\xi+1}^{\xi+N}$, with $\xi \geq \eta + N$, be the first and second segments, respectively. Then by the argument leading to Equation (3.20) we find that the ℓth residuals of the fits to the first and second segments are given by

$$\left.\begin{aligned} r_{\ell}^{(1)} &= \sum_{K=1}^{N} F(\ell, K) U_{\eta+K} \\ \text{and} \\ r_{\ell}^{(2)} &= \sum_{K=1}^{N} F(\ell, K) U_{\xi+K}, \end{aligned}\right\} \tag{3.21}$$

respectively. We have just shown that the statistics for these two sets of residuals are identical, so the only remaining question concerns the correlation between a residual of the first fit and one of the second fit.

From Equations (3.21), the correlation between $r_I^{(1)}$ and $r_J^{(2)}$ is given by the correlation coefficient, $(I < J)$,

$$\rho_r^{(1,2)}(I, J) = E\left[r_I^{(1)} r_J^{(2)}\right] \Big/ \sigma_{r_I} \sigma_{r_J}, \tag{3.22}$$

where

$$E\left[r_I^{(1)} r_J^{(2)}\right] = \sum_{i=1}^{N} \sum_{j=1}^{N} F(I, i) F(J, j) \rho_U(\xi - \eta + j - i). \tag{3.23}$$

The standard deviations in Equation (3.22) are found from Equation (3.17) and do not depend upon which set of residuals are under consideration. It is clear that if $\{U_K\}$ is a white sequence then the correlation between any residual from the first fit and any residual from the second fit will be zero.

The third, and last, of the primary questions mentioned in section I asks about the statistics of the fitted coefficients, a_0 and a_1. The statistics of a_0 represent the first occurence in this report of statistics that depend upon where the fitted segment occurs in $\{Z_n\}$. We will begin by deriving the statistics for these coefficients when the first N terms of $\{Z_n\}$ constitute the fitted segment and then will show how the statistics for a_0 are computed for an arbitrary segment.

Consider the segment, $\{Z_n\}_{n=1}^{N}$, of the r.w. $\{Z_n\}$ whose underlying sequence is zero mean, gaussian, and stationary. Using Equation (1.1) in Equations (3.2) and (3.3), we may write a_0 and a_1 in terms of $\{U_K\}$, for $N \geq 2$;

$$a_0 = \sum_{K=1}^{N} G_0(K) U_K, \tag{3.23}$$

$$a_1 = \sum_{K=1}^{N} G_1(K) U_K, \tag{3.24}$$

where

$$G_0(K) = \frac{1}{N(N-1)} \left\{ 3K^2 - (4N + 5)K + (N^2 + 3N + 2) \right\} \tag{3.25}$$

and

$$G_1(K) = \frac{-6}{hN(N^2 - 1)} \left\{ K^2 - (N + 2)K + (N + 1) \right\}. \tag{3.26}$$

It follows from Equations (3.23) and (3.24) that a_0 and a_1 are both zero mean and gaussian and that their variances are given by

$$\sigma^2_{a_q}(N) = E[a_q^2] = \sigma_U^2 \left[\sum_{K=1}^{N} G_q^2(K) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} G_q(i) G_q(j) \rho_U(j-i) \right] \tag{3.27}$$

where $q = 0, 1$. The correlation between a_0 and a_1 is given by the correlation coefficient,

$$\rho_{a_0 a_1} = E[a_0 a_1] / \sigma_{a_0} \sigma_{a_1} \tag{3.28}$$

where

$$E[a_0 a_1] = \sigma_U^2 \left\{ \sum_{K=1}^{N} G_0(K) G_1(K) + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} [G_0(i) G_1(j) + G_0(j) G_1(i)] \rho_U(j-i) \right\} . \tag{3.29}$$

From the definitions of $G_0(K)$ and $G_1(K)$ given in Equations (3.25) and (3.26), respectively, the following facts may be established. For $N = 2, 3, \ldots$,

$$G_0(1) = 1, \tag{3.30}$$

$$\sum_{K=1}^{N} G_0(K) = 0, \tag{3.31}$$

$$\sum_{K=1}^{N} G_0^2(K) = [(2N+1)(N+1)(N+2)] / [15N(N-1)], \tag{3.32}$$

$$G_1(1) = 0, \tag{3.33}$$

$$\sum_{K=1}^{N} G_1(K) = 1/h, \tag{3.34}$$

$$\sum_{K=1}^{N} G_1^2(K) = [6(N^2+1)] / [h^2 5N(N^2-1)] . \tag{3.35}$$

For white underlying sequences the double sum in Equation (3.27) vanishes, so for cases where $\{U_K\}$ is white,

$$\sigma_{a_0}^2(N) = \left[\frac{(2N+1)(N+1)(N+2)}{15N(N-1)}\right]\sigma_U^2$$

and

$$\sigma_{a_1}^2(N) = \left[\frac{6(N^2+1)}{h^2 5N(N^2-1)}\right]\sigma_U^2 .$$

In Figures 14 and 15 we have plotted $\sigma_{a_0}^2$ and $\sigma_{a_1}^2$, respectively, for the case of a white underlying sequence with unit variance and time step h = 1, for $2 \leqslant N \leqslant 25$.

Now, let's consider the arbitrary segment, $\{Z_n\}_{n=\xi+1}^{\xi+N}$, of length N. We may define a new segment, $\{X_n\}_{n=1}^{N}$, where

$$X_n = Z_{\xi+n} - Z_\xi, \quad n = 1, 2, \ldots, N; \tag{3.36}$$

so

$$X_n = \sum_{K=1}^{n} U_{\xi+K} . \tag{3.37}$$

Since $Z_{\xi+n} = Z_\xi + X_n$, it is easy to see (from Equations (3.2) and (3.3)) that the first-degree l.s. polynomial fit to the segment $\{Z_n\}_{n=\xi+1}^{\xi+N}$ is given by Z_ξ plus the first-degree fit to $\{X_n\}_{n=1}^{N}$. That is, the desired fit is given by

$$\hat{Z}_{\xi+\ell} = (Z_\xi + a_0) + a_1 t_\ell, \quad t_\ell = \ell h, \quad \ell = 1, 2, \ldots, N, \tag{3.38}$$

where

$$a_0 = \sum_{K=1}^{N} G_0(K) U_{\xi+K}$$

$$a_1 = \sum_{K=1}^{N} G_1(K) U_{\xi+K} .$$

The sequence $\{U_K\}$ is stationary so a_0 and a_1 are zero mean and their variances are given by Equation (3.27).

The constant term in Equation (3.38), i.e., the constant term in the first-degree fit to $Z_{n\,n=\xi+1}^{\xi+N}$, has a variance given by

$$\tilde{\sigma}_{a_0}^2 = E[(Z_\xi + a_0)^2] = E[Z_\xi^2] + 2E[Z_\xi a_0] + E[a_0^2] . \tag{3.39}$$

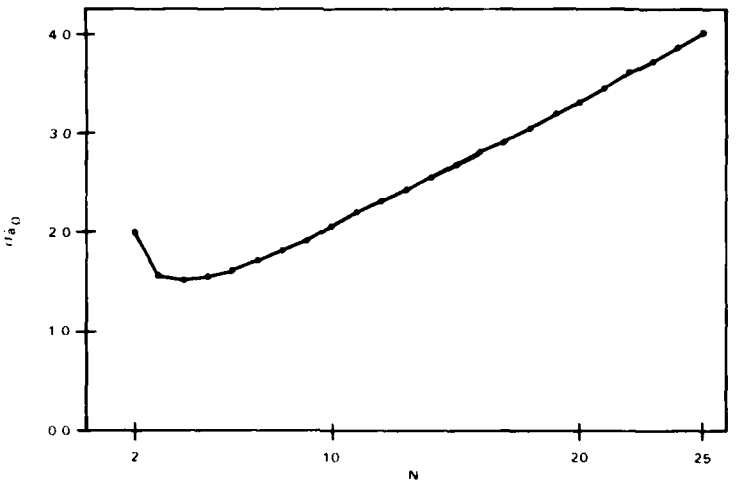


Figure 14

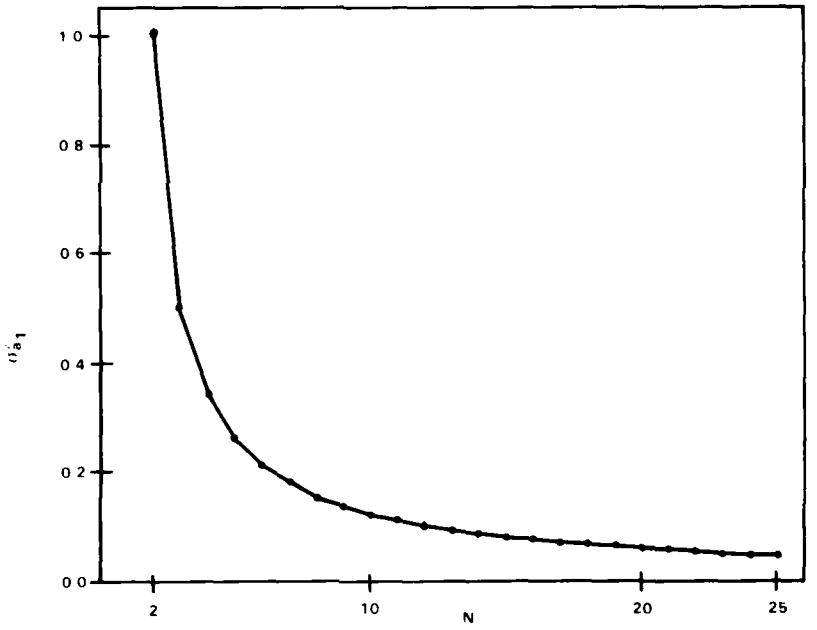


Figure 15

The expectation $E[Z_\xi^2]$ is given by Equation (2.4), $E[a_0^2]$ is given by Equation (3.27) as mentioned above, and

$$E[Z_\xi a_0] = \sigma_U^2 \left[\sum_{i=1}^{\xi} \sum_{j=1}^{N} G_0(j) \rho_U(\xi + j - i) \right]. \tag{3.40}$$

For white underlying sequences, the right-hand side of Equation (3.40) vanishes.

Similarly, the correlation between a_1 and the constant term in Equation (3.38) is given by the correlation coefficient,

$$\tilde{\rho}_{a_0 a_1} = E[(Z_\xi + a_0)a_1]/\tilde{\sigma}_{a_0}\sigma_{a_1} \tag{3.41}$$

where

$$E[(Z_\xi + a_0)a_1] = E[Z_\xi a_1] + E[a_0 a_1]. \tag{3.42}$$

Finally, the expectation $E[a_0 a_1]$ is given by Equation (3.29) and

$$E[Z_\xi a_1] = \sigma_U^2 \left[\sum_{i=1}^{\xi} \sum_{j=1}^{N} G_1(j)\rho_U(\xi + j - i) \right]. \tag{3.43}$$

IV. LEAST-SQUARES FITS OF ARBITRARY DEGREE

In this section, we generalize the results given in section III in three ways. The principal generalization is that polynomial fits of arbitrary degree are allowed. In addition, the independent variable will be $(t - t_0)$ instead of t, where t_0 is arbitrary. Thus, the fitted polynomials will be of the form,

$$P_d(t) = \sum_{j=0}^{d} a_j(t - t_0)^j. \tag{4.1}$$

Finally, the terms of $\{U_K\}$ and $\{Z_n\}$ will not be assumed to be evenly spaced in time.

These generalizations are easily accomodated by the introduction of matrix notation. As in section III, we will begin by considering fits to the r.w. segment, $\{Z_n\}_{n=1}^{N}$, consisting of the first N terms of $\{Z_n\}$, and show later that no generality is lost by this choice. As usual, we will assume that the underlying sequence, $\{U_K\}$, is a stationary, zero mean gaussian sequence. The (column) vectors **U**, **Z**, $\hat{\mathbf{Z}}$, **r** and **a** are defined by

$$\mathbf{U} = [U_1, U_2, \ldots, U_N]^T$$

$$\mathbf{Z} = [Z_1\ Z_2, \ldots, Z_N]^T$$

$$\hat{\mathbf{Z}} = [\hat{Z}_1, \hat{Z}_2, \ldots, \hat{Z}_N]^T$$

$$\mathbf{r} = [r_1, r_2, \ldots, r_N]^T$$

$$\mathbf{a} = [a_0, a_1, \ldots, a_d]^T$$

where $\hat{Z}_\varrho = P_d(t_\varrho)$, $r_\varrho = Z_\varrho - \hat{Z}_\varrho$, and a_i, i = 0, 1, ..., d are the coefficients of the l.s. polynomial $P_d(t)$. We will always assume that $N > (d + 1)$.

The dth-degree l.s. polynomial fit to $\mathbf{Z}$ is given by

$$\mathbf{a} = (A^T A)^{-1} A^T \mathbf{Z}, \tag{4.2}$$

where A is the $N \times (d + 1)$ observation matrix. For this case, A is simply

$$A = \begin{bmatrix} 1 & (t_1 - t_0) & \dots & (t_1 - t_0)^d \\ 1 & (t_2 - t_0) & \dots & (t_2 - t_0)^d \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (t_N - t_0) & \dots & (t_N - t_0)^d \end{bmatrix}. \tag{4.3}$$

The values of the fitted polynomial and the associated residuals at the times $t_1, t_2, \dots, t_N$, are given by

$$\hat{\mathbf{Z}} = C\mathbf{Z} \tag{4.4}$$

and

$$\mathbf{r} = H\mathbf{Z} \tag{4.5}$$

where C and H are the $N \times N$ matrices defined by

$$C = A(A^T A)^{-1} A^T, \tag{4.6}$$

$$H = (I - C), \tag{4.7}$$

with I denoting the $N \times N$ identity matrix.

Now, since Z_n is just the sum of the terms U_K, $K = 1, 2, \dots, N$, $\mathbf{Z}$ and $\mathbf{U}$ are related by

$$\mathbf{Z} = S\mathbf{U} \tag{4.8}$$

where S is the lower triangular $N \times N$ matrix whose elements on and below the main diagonal are all unity. Thus, using Equation (4.8) in Equations (4.4) and (4.5) we have

$$\hat{\mathbf{Z}} = D\mathbf{U} \tag{4.9}$$

and

$$\mathbf{r} = F\mathbf{U}, \tag{4.10}$$

where D and F are the $N \times N$ matrices defined by

$$D = CS, \tag{4.11}$$

$$F = HS. \tag{4.12}$$

Since $\{U_K\}$ is zero mean and gaussian, it follows that $\mathbf{r}$ is also zero mean and gaussian. Thus, the statistics of $\mathbf{r}$ are completely specified by its covariance matrix, Q_r. From Equation (4.10),

$$Q_r = E[\mathbf{r}\mathbf{r}^T] = FQ_UF^T, \tag{4.13}$$

where $Q_U = E[\mathbf{U}\mathbf{U}^T]$ is the covariance matrix of $\mathbf{U}$.

At this point, we have determined the statistics of the fit residuals for fits to the segment $\{Z_n\}_{n=1}^N$. We will now show that, for any other segment of length N, whose spacing scheme is identical to that of $\{Z_n\}_{n=1}^N$, the fit residual statistics will also be given by Equation (4.13). The proof of this claim will make use of the fact that the rows of C all sum to one. So, we need the following lemma.

Lemma (1): Let A be the $N \times (d+1)$ observation matrix, given in Equation (4.3), associated with a dth-degree l.s. polynomial fit to N points, and let C be the $N \times N$ matrix defined by $C = A(A^TA)^{-1}A^T$. Then the rows of C each sum to one.

Proof of Lemma (1): Consider the matrix $H = (I - C)$. The product A^TH vanishes because

$$A^TH = A^T(I - C) = A^T(I - A(A^TA)^{-1}A^T) = A^T - A^T = 0.$$

By Equation (4.3), the first row of A^T consists of N ones, so it follows from the fact that $A^TH = 0$ that

$$\sum_{i=1}^{N} H_{ij} = 0, \quad \text{for} \quad j = 1, 2, \ldots, N.$$

Since H is symmetric, this implies that

$$\sum_{j=1}^{N} H_{ij} = 0, \quad \text{for} \quad i = 1, 2, \ldots, N.$$

Now, from the definition of H, $C = I - H$, so

$$\sum_{j=1}^{N} C_{ij} = 1 - \sum_{j=1}^{N} H_{ij} = 1, \quad \text{for} \quad i = 1, 2, \ldots, N,$$

i.e., the sum of each row of C is one. *Q.E.D.*

It should be noted that if A is an arbitrary observation matrix associated with some l.s. procedure, the rows of $A(A^TA)^{-1}A^T$ will not, in general, sum to one. For example, consider the 4×2 matrix,

$$A = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 3 & 1 & 1 \end{bmatrix}^T .$$

For this case, $A^T A$ is nonsingular and the four rows of $A(A^T A)^{-1} A^T$ sum to 21/29, 38/29, 30/29, and 17/29, for i = 1, 2, 3 and 4, respectively.

With the preceding lemma in hand, we may now prove the following theorem.

Theorem: Let $\{Z_n\}$ be a r.w. whose underlying sequence, $\{U_K\}$, is zero mean, stationary, and gaussian. Consider the two segments of $\{Z_n\}$ denoted by

$$\mathbf{Z}_I = [Z_1, Z_2, \ldots, Z_N]^T$$

and

$$\mathbf{Z}_{II} = [Z_{\xi+1}, Z_{\xi+2}, \ldots, Z_{\xi+N}]^T ,$$

where ξ is an arbitrary integer such that $\xi > N$. If the spacing schemes for the two segments are identical; i.e., if $(t_{i+1} - t_i) = (t_{\xi+i+1} - t_{\xi+i})$ for i = 1, 2, ..., N − 1, and each segment is approximated by a *d*th-degree l.s. polynomial, then the fit residuals for the two fits will have identical statistics, given by Equation (4.13).

Proof of Theorem: Let $\mathbf{U}_I = [U_1, U_2, \ldots, U_N]^T$ and $\mathbf{U}_{II} = [U_{\xi+1}, U_{\xi+2}, \ldots, U_{\xi+N}]^T$, with associated covariance matrices Q_{U_I} and $Q_{U_{II}}$, respectively. The elements of $\mathbf{Z}_{II}$ are given by

$$Z_{\xi+\ell} = \sum_{K=1}^{\xi+\ell} U_K = Z_\xi + \sum_{K=\xi+1}^{\xi+\ell} U_K , \quad \text{for} \quad \ell = 1, 2, \ldots, N.$$

Thus, $\mathbf{Z}_{II}$ may be written in the form,

$$\mathbf{Z}_{II} = \mathbf{S}\mathbf{U}_{II} + Z_\xi [1, 1, \ldots, 1]^T .$$

The fit values, $\hat{\mathbf{Z}}_{II}$, are given by,

$$\begin{aligned} \hat{\mathbf{Z}}_{II} = C\mathbf{Z}_{II} &= C\mathbf{S}\mathbf{U}_{II} + Z_\xi C[1, 1, \ldots, 1]^T \\ &= \mathbf{D}\mathbf{U}_{II} + Z_\xi C[1, 1, \ldots, 1]^T \\ &= \mathbf{D}\mathbf{U}_{II} + Z_\xi [1, 1, \ldots, 1]^T , \end{aligned}$$

where the last equality follows because the rows of C all sum to one. The fit residuals for the fit to $\mathbf{Z}_{II}$ are then

$$\mathbf{r}_{II} = \mathbf{Z}_{II} - \hat{\mathbf{Z}}_{II} = \mathbf{S}\mathbf{U}_{II} - \mathbf{D}\mathbf{U}_{II} = \mathbf{F}\mathbf{U}_{II} .$$

Clearly, $\mathbf{r}_{II}$ is zero mean and gaussian, with covariance matrix,

$$\mathbf{Q}_{r_{II}} = \mathbf{F}\mathbf{Q}_{U_{II}}\mathbf{F}^T .$$

By Equation (4.13), the fit residuals for the fit to $\mathbf{Z}_I$ have covariance matrix,

$$\mathbf{Q}_{r_I} = \mathbf{F}\mathbf{Q}_{U_I}\mathbf{F}^T .$$

Since $\{U_K\}$ is stationary and $\mathbf{U}_I$ and $\mathbf{U}_{II}$ have identical spacing schemes, by assumption, it follows that $Q_{U_{II}} = Q_{U_I}$ and hence $Q_{r_{II}} = Q_{r_I}$. *Q.E.D.*

We have just shown that if the conditions stated in the preceding theorem are met then the individual statistics for $\mathbf{r}_I$ and $\mathbf{r}_{II}$ are identical. To complete our discussion of the relationship between the two sets of residuals we need to determine the cross-covariance matrix, $R_{I,II}$, for $\mathbf{r}_I$ and $\mathbf{r}_{II}$. Since $\mathbf{r}_I = \mathbf{F}\mathbf{U}_I$ and $\mathbf{r}_{II} = \mathbf{F}\mathbf{U}_{II}$, both are zero mean and $R_{I,II}$ is given by

$$R_{I,II} = E[\mathbf{r}_I\mathbf{r}_{II}^T] = \mathbf{F}E[\mathbf{U}_I\mathbf{U}_{II}^T]\mathbf{F}^T . \tag{4.14}$$

If $\{U_K\}$ is a white sequence, it follows from Equation (4.14) and the fact that $\{U_K\}$ is gaussian that $\mathbf{r}_I$ and $\mathbf{r}_{II}$ are statistically independent for this case.

Finally, we need to address the question of the statistics of the fit coefficients, $\mathbf{a}$. As mentioned in section III, the statistics of the constant term, a_0, depend on where the fitted r.w. segment occurs in $\{Z_n\}$. We will begin, as before, by deriving the statistics of $\mathbf{a}$ for the segment consisting of the first N terms of $\{Z_n\}$ and then will show how the statistics are computed for an arbitrary segment. The extension will require the result given in the following lemma.

Lemma (2): Let A be the observation matrix associated with a l.s. dth degree polynomial fit to the N data points, $Z_n = Z(t_n)$, n = 1, 2, ..., N, given explicitly by Equation (4.3). Let W be the (d + 1) × N matrix given by $\mathbf{W} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$, $\mathbf{V}$ be an N-dimensional column vector of ones, and $\mathbf{e}$ be the (d + 1)-dimensional column vector whose first component is one and whose remaining components are all zero; i.e., $\mathbf{V} = [1, 1, \ldots, 1]^T$ and $\mathbf{e} = [1, 0, 0, \ldots, 0]^T$. Then, $\mathbf{WV} = \mathbf{e}$.

Proof of Lemma (2): By hypothesis (see Equation (4.2)), W acts on a given set of N points to yield the coefficients of the unique dth-degree polynomial which minimizes $\mathbf{r}^T\mathbf{r}$, where $\mathbf{r}$ is the column vector of fit residuals. Thus, WV will be the coefficients of the dth-degree polynomial which best fits (in the l.s. sense) the components of $\mathbf{V}$. If $\mathbf{WV} = \mathbf{e}$, then $\mathbf{r} = \mathbf{V} - \mathbf{Ae} = 0$ and $\mathbf{r}^T\mathbf{r} = 0$. But $\mathbf{r}^T\mathbf{r} \geq 0$ for any $\mathbf{r}$, so it follows that no other (d + 1)-dimensional column vector, $\tilde{\mathbf{e}} \neq \mathbf{e}$, could yield a lower value for $\mathbf{r}^T\mathbf{r}$. Since the best fitting l.s. dth-degree polynomial is unique, WV must equal $\mathbf{e}$. *Q.E.D.*

Consider the segment $\{Z_n\}_{n=1}^N$ of the r.w. $\{Z_n\}$ whose underlying sequence, $\{U_K\}$, is zero mean, stationary, and gaussian. From Equations (4.2) and (4.8) we have

$$\mathbf{a} = \mathbf{GU}, \tag{4.15}$$

where

$$G = (A^T A)^{-1} A^T S. \tag{4.16}$$

It follows from Equation (4.15) that **a** is zero mean and gaussian, and that its covariance matrix is just

$$Q_a = G Q_U G^T. \tag{4.17}$$

Now, let's consider the arbitrary segment, $\{Z_n\}_{n=\xi+1}^{\xi+N}$, of length N. Proceeding as we did in section III, we define a new segment, $\{X_n\}_{n=1}^{N}$, where (see Equation (3.36))

$$X_n = Z_{\xi+n} - Z_\xi, \quad n = 1, 2, \ldots, N$$

or

$$X_n = \sum_{K=1}^{n} U_{\xi+K}.$$

Letting $\mathbf{Z} = [Z_{\xi+1}, Z_{\xi+2}, \ldots, Z_{\xi+N}]^T$ and

$$\mathbf{X} = [X_1, X_2, \ldots, X_N]^T$$

we see from the definition of X_n given above that these vectors are related by

$$\mathbf{X} = \mathbf{Z} - Z_\xi \mathbf{V}$$

or

$$\mathbf{Z} = \mathbf{X} + Z_\xi \mathbf{V},$$

where **V** is an N-dimensional column vector each of whose components is one, i.e., $\mathbf{V} = [1, 1, \ldots, 1]^T$.

The fits to **X** and **Z** are thus given by

$$\mathbf{a}_X = (A^T A)^{-1} A^T \mathbf{X} = W\mathbf{X} \tag{4.18}$$

and

$$\mathbf{a}_Z = (A^T A)^{-1} A^T \mathbf{Z} = W\mathbf{Z} = W\mathbf{X} + Z_\xi W\mathbf{V} = \mathbf{a}_X + Z_\xi W\mathbf{V}.$$

But, from Lemma (2) we know that $W\mathbf{V} = \mathbf{e}$, so

$$\mathbf{a}_Z = \mathbf{a}_X + Z_\xi \mathbf{e}. \tag{4.19}$$

Thus, the covariance matrix of the coefficients of the fit to $\mathbf{Z}$ is just,

$$Q_{a_Z} = E[a_Z a_Z^T] = Q_{a_X} + E[a_X e^T Z_\xi] + E[Z_\xi e a_X^T] + E[Z_\xi^2 e e^T], \qquad (4.20)$$

where Q_{a_X} is the covariance matrix of a_X.

Since $\mathbf{X} = \mathbf{S}U_X$, where $U_X = [U_{\xi+1}, U_{\xi+2}, \ldots, U_{\xi+N}]^T$ and Z_ξ is simply the sum of U_1 through U_ξ, it follows that Q_{a_X} is given by

$$Q_{a_X} = GE[U_X U_X^T]G^T = GQ_{U_X}G^T,$$

and the other terms of Equation (4.20) are also easily rewritten in terms of the statistics of $\{U_K\}$. It should be noted that the second, third, and fourth terms on the right of Equation (4.20) act to modify only the first row and first column of Q_{a_X}.

Figures 16 and 17 display the variances, $\sigma_r^2(\ell)$, of the fit residuals for fits of degree d = 1, 2, 5 and 8. In each case, N = 50, $\sigma_U^2 = 1.0$, and $(t_{K+1} - t_K) = 1.0$, for K = 1, 2, ... The results shown in Figure 16 are for the case with $\{U_K\}$ a white sequence. Figure 17 shows the corresponding results with $\{U_K\}$ a first-order markov sequence whose correlation time, τ_u, is equal to 5.0. (Samples of the underlying sequences and the associated random walks for these two cases were presented in Figures 1, 2, and 3.) It should be noted that the vertical scales for the plots with d = 1 and d = 2 differ between Figures 16 and 17, while the plots with d = 5 and d = 8 have the same vertical scales in both figures.

It is interesting to note that the values of $\sigma_r^2(\ell)$ with d = 8 are slightly higher, for most values of ℓ, when $\{U_K\}$ is white than they are for the correlated case. This is due to the fact that an r.w. generated by a white underlying sequence contains more power at higher frequences than one generated by a correlated underlying sequence. This is illustrated by a comparison of Figures 2 and 3.

V. EXAMPLE; ATOMIC CLOCK ERRORS

A clock is any device which counts the cycles of a periodic phenomenon. Among the most stable clocks in use are the atomic clocks which form the basis for atomic time scales such as International Atomic Time (TAI). Atomic time is used primarily as a measure of time interval and is based on the electromagnetic oscillations produced by quantum transitions within the atom.

If an atomic clock is used as a component of some distance measurement system, then clock errors become part of the total measurement error of the system. When data analysis is performed on the output of such a system, it is often convenient to model (approximate) the clock errors over some time interval by a polynomial. In this way, a substantial portion of the clock's contribution to the total measurement error may be removed. This is the motivation for the work presented in this example.

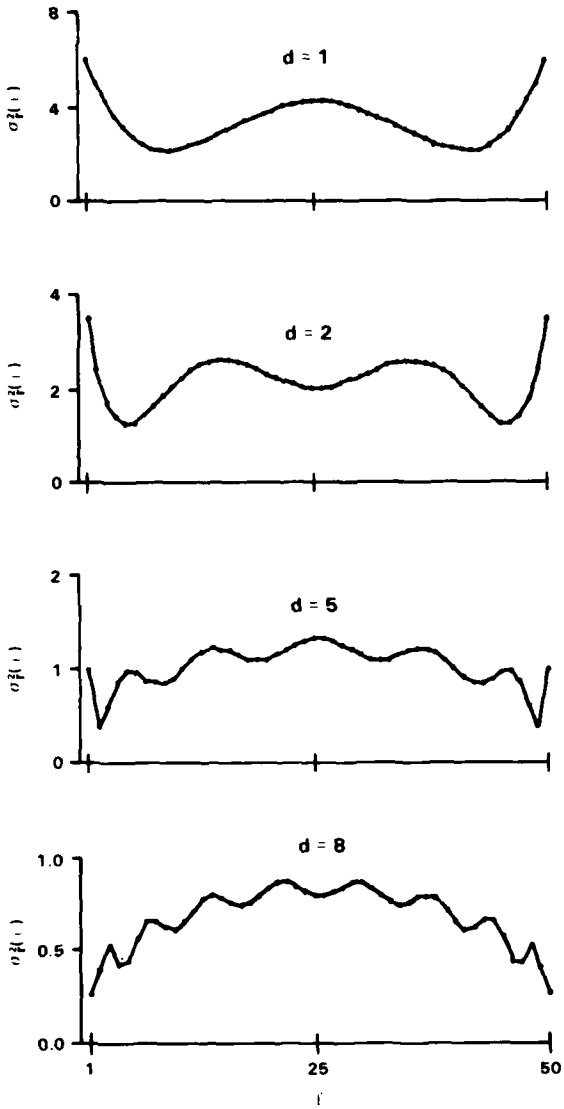


Figure 16

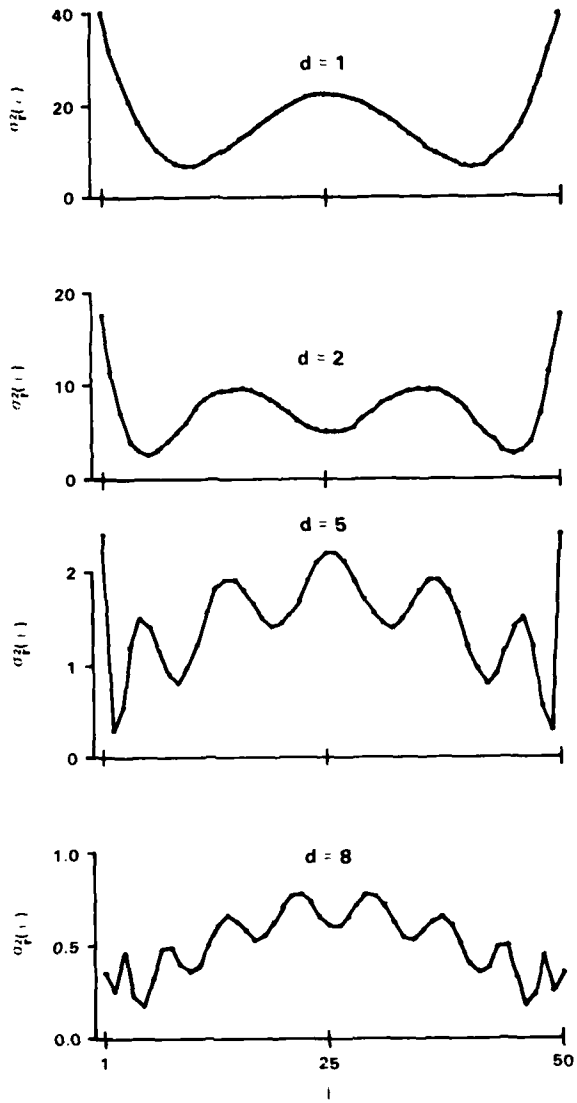


Figure 17

Atomic clocks are subject to both deterministic and random errors causing deviations from an idealized time scale. The random time error, X(t), results from the summation of all prior time changes induced by random changes in frequency, known as fractional frequency fluctuations, y(t). These are defined by

$$y(t) = \Delta F(t)/F,$$

where F is the ideal frequency of the clock's oscillator and $\Delta F(t)$ represents the departure of the actual frequency from F. Standard procedures for specifying the statistics of y(t) involve the use of the Allan variance or the use of power spectral densities (see Chapter 8 of Reference 4).

Since the random time error is the sum of all prior random time errors caused by the random frequency fluctuations, the statistics of the random time error may be developed from the statistics of y(t). In the present example, we have assumed that measurements of X(t) have been made discretely with a fixed-time interval, Δt, between measurements and that y(t) may also be approximated as a discrete sequence. Thus, in terms of our previously adopted notation, $\{y_K\}$ plays the role of $\{U_K\}$ and $\{X_n\}$ the role of $\{Z_n\}$. Given that the error in the clock at time t_0 is zero (i.e., that t_0 corresponds to a reset or calibration time of the clock) then $\{X_n\}$ takes the form of a random walk, with

$$X_n = \Delta t \sum_{K=1}^{n} y_K .$$

Figure 18 is a plot of the square root of the Allan variance for a typical cesium clock. Using this information, a sequence of random time errors was generated whose Allan variance closely approximates that shown in the figure. These time errors were converted to range or distance errors by multiplying by the speed of light. The plot of these errors is labeled "actual range error" in Figure 19. The data are given over a five-day period with the clock error assumed to be zero initially. Notice that the error appears quite systematic over the five-day span, even though it does represent a segment of a random walk. This is a result of the long correlation time of the underlying sequence, so the horizontal scale would have to be extended over a much longer time period in order to see a pattern more representative of the behavior we expect to see from a random walk process.

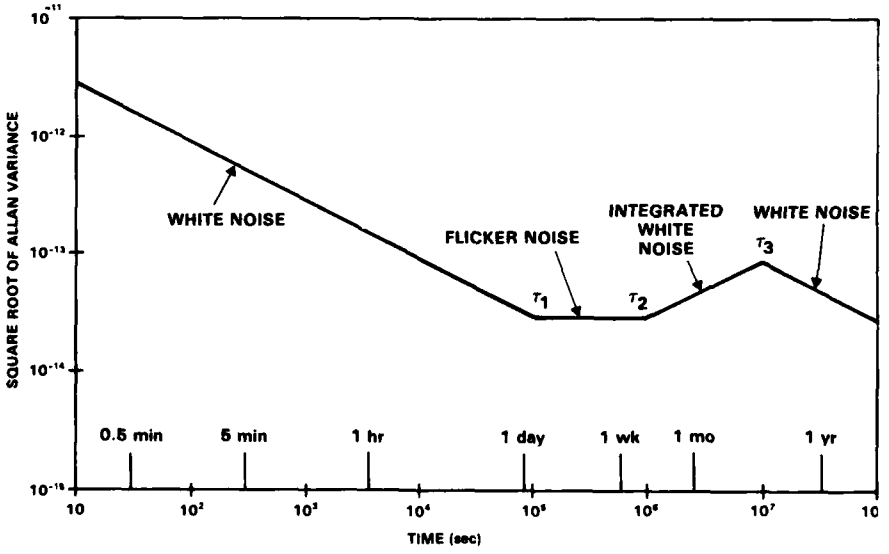


Figure 18

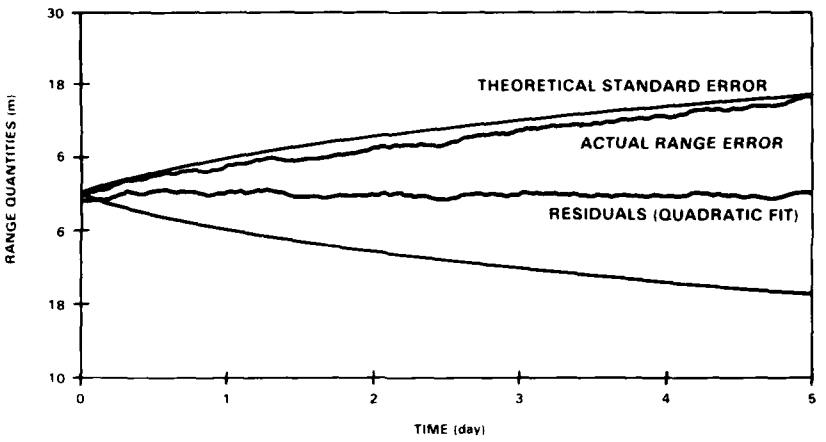


Figure 19

A second-degree polynomial of the form,

$$P_2(t) = a_0 + a_1 t + a_2 t^2 ,$$

was fit to the range errors over the five-day period. The resulting coefficient values were

$$a_0 = 1.0589 \text{ m}$$
$$a_1 = 0.39213 \times 10^{-4} \text{ m/sec}$$
$$a_2 = -0.12103 \times 10^{-10} \text{m/sec}^2 .$$

The associated fit residuals are plotted in Figure 19. The contributions of the linear and quadratic terms at $t = 0.432 \times 10^6$ sec (= 5 days) are 16.94 m and −2.26 m, respectively, so for time intervals of much shorter duration than five days a first-degree polynomial should be adequate.

For example, an eight-hour span of this simulated cesium clock error was fit with a polynomial of the form,

$$P_1(t) = a_0 + a_1 t .$$

The fit residuals for this case are shown in Figure 20. The theoretical standard deviations of the residuals for a first-degree fit over eight hours are given in Figure 21.

Finally, Figure 22 gives a contour plot of the matrix of correlation coefficients of the fit residuals for the linear model over the eight hour span. Notice that the correlation between residuals separated by a fixed time interval varies with the location of the residuals within the fit span. This, as well as the curve in Figure 21, demonstrates the non-stationarity of the statistics of the residuals.

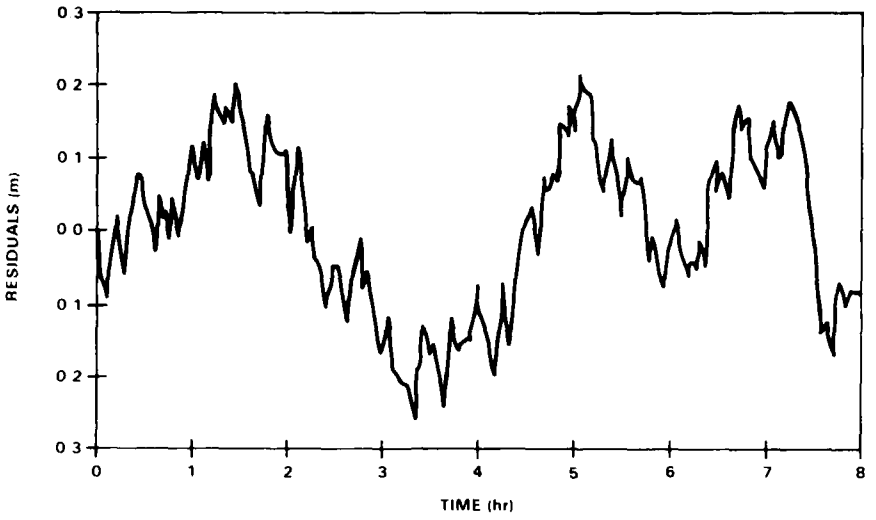


Figure 20

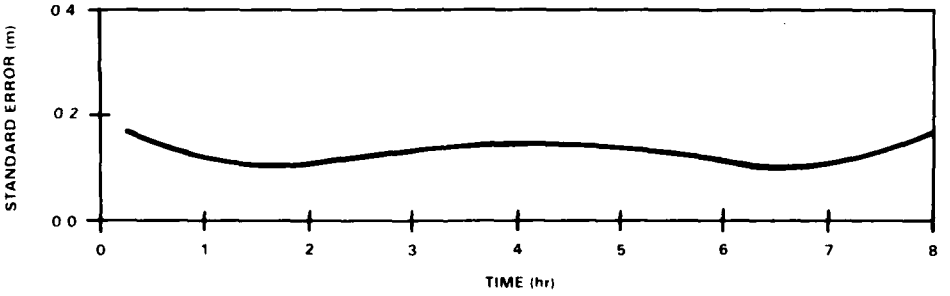


Figure 21

* * *

This concludes the present report. Clearly, the practical application of l.s. polynomial approximations to one dimensional r.w. segments raises many questions that were not addressed in this report. It is hoped, however, that the results and techniques that have been presented will serve as a basis for the derivations required to answer specific questions of interest to the reader.

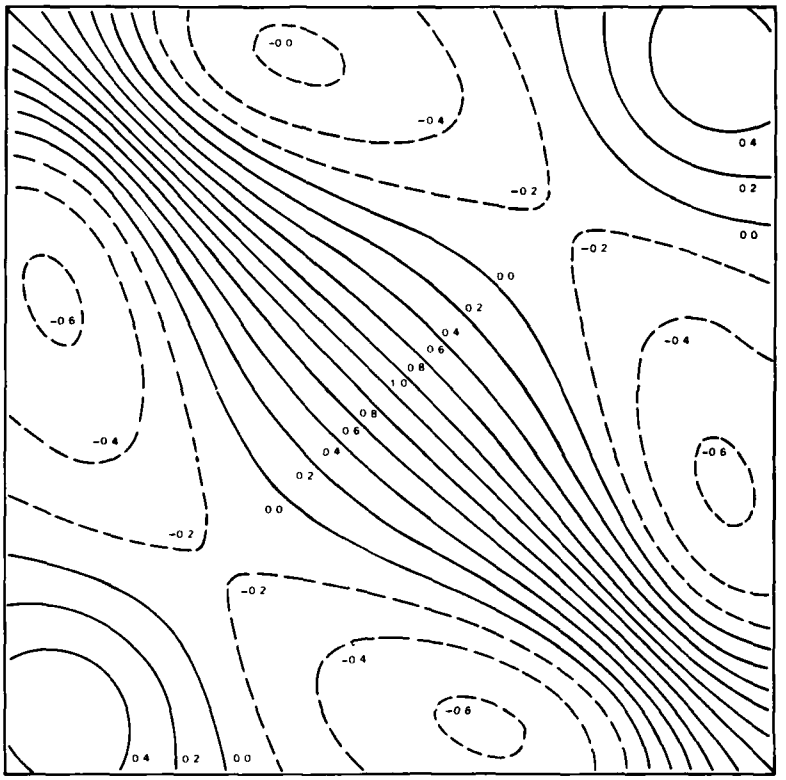


Figure 22

REFERENCES

1. R. S. Burington and D. C. May, *Handbook of Probability and Statistics with Tables*, Second Edition, McGraw-Hill, New York, 1970.
2. A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1965.
3. P. L. Young, "Some Practical Aspects of the Generation of Gaussian Random Sequences," NAVSWC Technical Report, in preparation.
4. B. E. Blair, Editor, *Time and Frequency: Theory and Fundamentals*, National Bureau of Standards Monograph No. 140, 1974, U.S. Government Printing Office, Washington, D.C.

DISTRIBUTION

Defense Technical Information Center
Cameron Station
Alexandria, VA 22314 (12)

Library of Congress
Washington, DC 20540
Attn: Gift and Exchange Division (4)

Defense Mapping Agency
Aerospace Center
St. Louis, MO 63118
Attn: R. Ballew
J. Finkland (Code GDG) (2)
M. Shultz (Code GDGA) (2)

Defense Mapping Agency
Hydrographic/Topographic Center
Washington, DC 20360
Attn: R. Smith (2)
A. Carlson (2)

U.S. Army Topographic Laboratory
Ft. Belvoir, VA 22060

U.S. Naval Academy
Annapolis, MD 21402
Attn: Technical Library

Naval Electronics Systems Command
Department of the Navy
Washington, DC 20360
Attn: Technical Library

Naval Research Laboratory
Washington, DC 20360
Attn: J. Buisson (2)
Technical Library (2)

Naval Oceanographic Office
NSTL Station
Bay St. Louis, MS 39522
Attn: T. Davis (2)

HQ SAMSO
Air Force Unit Post Office
Los Angeles, CA 90045
Attn: YE

Aerospace Corporation
2350 East El Segundo Boulevard
El Segundo, CA 90245
Attn: W. Feese

U.S. Naval Observatory
Washington, DC 20390
Attn: G. Winkler
D. Parseval

Applied Physics Laboratory
The John Hopkins University
11100 John Hopkins Road
Laurel, MD 20810
Attn: L. Levy
H. Black
G. Weiffenbach

The Charles Stark Draper Laboratory
Cambridge, MA 02139
Attn: W. Robertson

Naval Air Development Center
Warminster, PA 18974
Attn: Technical Library

The Analytic Science Corporation
6 Jacob Way
Reading, MA 01867
Attn: W. Heller

National Bureau of Standards
Time and Frequency Division
325 Broadway
Boulder, CO 80302
Attn: D. Allan

Local:

E31 (GIDEP)
E41

K10 (25)
K12 (P. Fell) (15)
K13 (P. Young) (15)

K20

DISTRIBUTION—(Continued)

K30		
K33	(M. Thomas)	
K33	(J. Crigler)	
K50		(2)
K55		(2)
K50	(T. Alexander)	
K50	(O. Schultz)	
K51	(D. Owen)	
K56	(W. Farr)	
K56	(S. Lipscomb)	
K56	(G. Sitzman)	
F106	(B. Clark)	
F20	(R. Hughey)	
N23	(P. O'Neil)	
X210		(6)
X211		(2)