stationarity, since the autoregressive and moving average processes are both linear models being fit to stationary processes. Data differencing would be of the form

$$x_t - x_{t-1} = (1 - B)x_t \tag{3}$$

The full text of the paper describes in detail: 1) the identification of the time series model from autocorrelation and partial correlation of the data; 2) the estimation of the ϕ, θ, ramp and random walk parameters using maximum likelihood and nonlinear least squares; and 3) the means by which one would deduce model adequacy through autocorrelation of the white noise residuals and confidence limit theory.

As an example of the theory, consider Fig. 1, which shows a sample of normalized long term gyro drift rate. Since the process is nonstationary, the data is differenced as is shown in Fig. 2. The analysis indicated that the math model for this gyro drift rate sample is

$$x_t = x_{t-1} - \theta_1 z_{t-1} + z_t + b(1 - \theta_1) \tag{4}$$

where θ_1 = first-order moving average parameter, $b(1 - \theta_1)$ = mean of the differenced data indicating a ramp in the original process, and z_t = white noise residual indicating a random walk process in the original data. The mean squared value of Eq. (4) is given as

$$\overline{x_t{}^2} = n\sigma_z{}^2(1 - \theta_1)^2 + n^2b^2(1 - \theta_1)^2 + 2\theta^2\sigma_z{}^2 \tag{5}$$

where $E[z_t{}^2] = \sigma_z{}^2$. Note that this equation consists of random walk, ramp, and stationary noise processes. A picture of model adequacy is shown in Fig. 3. The data of Fig. 1 was broken into 64-hr segments and ensemble averaged. Using the appropriate values of θ_1, b, and $\sigma_z{}^2$ determined from the analysis, various combinations of Eq. (5) were superimposed on Fig. 3. It can be seen that model adequacy is attained only by using the combination of two nonstationary processes and the one stationary process. A good fit of the data is not attained with only one of the nonstationary processes.

It is important to note that the models and parameter values obtained were not from the ensemble mean squared gyro drift rate waveform, but from the single sample long term gyro drift rate process.

Stationary and Nonstationary Characteristics of Gyro Drift Rate

Albert S. Oravetz*
Naval Strategic Systems Navigation Facility, Brooklyn, N. Y.

AND

Herbert J. Sandberg†
Dynamics Research Corporation, Wilmington, Mass.

Procedures for the mathematical modeling of gyro drift rate based on modeling techniques for stationary and nonstationary time series analysis are presented. The essence of these new time series analysis methods is the reduction of a random process to white noise (uncorrelated) residuals. The math model of the time series is established by reducing the wave form to white noise while identifying the correlated portion of the time series. This approach differs from other known techniques in that no deterministic models of the time series are initially assumed, e.g., ramps or sinusoids. The data is treated as a random process on which systematic application of the time series techniques will determine the exact gyro drift rate math model. If deterministic phenomena exist in the data, these will be directly evident from the analysis without making a priori assumptions as to their existence. Once the time series math model has been identified, parameter estimation techniques for the numerical evaluation of the math model are employed. Application of the single sample gyro drift rate model to an ensemble average leads to model verification.

1. Introduction

OVER the past several years the investigation of a mathematical model for gyro drift rate has assumed increased importance. The drift rate of a gyro is a major error source of inertial navigation systems that are required to operate over long time intervals, and as such, must be mathematically modeled whenever an error analysis or optimization study is performed on the navigation system.

Of particular importance as far as long term navigation is concerned, is the nonstationary behavior of gyro drift rate. In order to minimize the buildup of navigation errors by use of a Kalman filter, an accurate math model of the nonstationary gyro drift rate in the filter is essential in order to provide as near optimal control to the system as possible.

The gyros used in long term inertial navigation systems must remain in operation over a period of many months. In order to gain insight into the behavior of gyros in an inertial system over prolonged time periods, gyros were tested for time periods of 500–1000 hr. It is to this aspect of gyro drift rate modeling that this paper is addressed.

Presented at the AIAA Guidance, Control, and Flight Mechanics Conference, Princeton, N.J., August 18–20, 1969; submitted September 2, 1969; revision received January 28, 1970.

* Branch Head.

† Project Engineer, Systems Analysis and Design Department. Associate AIAA.

The complete theory of stationary modeling of single sample long term gyro drift rate has been presented in Ref. 1.‡ That theory has recently been advanced to include the capability of determining the presence of two different nonstationary processes in the gyro drift rate data, again, without making a priori assumptions as to their existence. In the past, single nonstationary models have had to be assumed and then least squares fit to the data. Thus, while it will be necessary to discuss the stationary aspects of mathematical modeling, the main emphasis of this paper will be on the nonstationary characteristics of the math model for long term gyro drift rate.

2. Mathematics of Single Sample Time Series Analysis

Description of Procedures

Determining a math model from long term gyro drift rate data is achieved by a three-stage iterative procedure based on identification, estimation, and diagnostic checking. Identification means using the data in the light of the information on how the series was generated to suggest a class of models that should be considered. Estimation means using the data to make inferences about parameters conditional on the adequacy of the model chosen. Diagnostic checking means checking the fitted model in relation to the data with the hope of revealing model inadequacies and thus proceeding to model improvement.

Identification—Time Series Models

Discrete linear random processes are described by linear difference equations relating an output random process x_t to an input random process, z_t, where x_t is considered as the time series data and z_t is considered as a white noise process. If

$$x_t = z_t \tag{2.1}$$

then the data is considered as purely white noise. If x_t is of the form

$$x_t = z_t - \theta_1 z_{t-1} - \theta_2 z_{t-2} \dots \theta_q z_{t-q} \tag{2.2}$$

then the data is considered to be a moving average (MA) process of order q. A first order MA process, MA(1), is

$$x_t = z_t - \theta_1 z_{t-1} \tag{2.3}$$

The interpretation of Eq. (2.3) is that the present value of the process, x_t, is dependent on additive white noise plus the effect of the previous white noise input at time $t - 1$ weighted by $-\theta_1$. Another interpretation for the MA process is that the system receives a "shock" before the effect of the previous "shock" has worn off.

Another linear process is the autoregressive (ar) process of order p given as

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + z_t \tag{2.4}$$

A first-order ar process, ar(1) is

$$x_t = \phi_1 x_{t-1} + z_t \tag{2.5}$$

The interpretation of Eq. (2.5) is that the present value of the process, x_t, is dependent on the previous value of the process, x_{t-1} weighted by ϕ_1 plus additive white noise, z_t.

A general mixed ar MA process of order p and q is given as

$$x_t = \phi_1 x_{t-1} + \dots \phi_p x_{t-p} + z_t - \theta_1 z_{t-1} \dots - \theta_q z_{t-q} \tag{2.6}$$

‡ Though the entire stationary theory of gyro drift rate modeling was presented in Ref. 1, it is felt that its main features should be presented in this paper also, since that theory is germane to the newly developed nonstationary theory.

If we introduce the notation $Bx_t = x_{t-1}$ and $Bz_t = z_{t-1}$ where the term B means delay one unit in time, Eq. (2.6) can be written as

$$x_t = \phi_1 B x_t + \dots \phi_p B^p x_t + z_t - \theta_1 B z_t - \dots - \theta_q B^q z_t \tag{2.7}$$

Bringing terms with x_t onto the left-hand side results in

$$(1 - \phi_1 B - \dots \phi_p B^p) x_t = (1 - \theta_1 B - \dots - \theta_q B^q) z_t$$

or

$$\phi(B) x_t = \theta(B) z_t \tag{2.8}$$

This is the general form of an ar MA process for linear stationary processes. If the data were comprised of a nonstationary process (ramp and/or random walk) in addition to a stationary process, it would be necessary to difference the data to induce stationarity since the ar and MA processes are both linear models being fit to stationary processes. Data differencing would be of the form

$$x_t - x_{t-1} = (1 - B) x_t \tag{2.9}$$

Double data differencing (which is highly unlikely for the purposes of long term gyro drift rate modeling) would mean a second-order nonstationary process in the data and would be of the form

$$(x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = x_t - 2x_{t-1} + x_{t-2} = (1 - B)^2 x_t \tag{2.10}$$

Thus d differences of the data would be of the form

$$(1 - B)^d x_t \tag{2.11}$$

and the most general form of the integrated mixed ar MA process would be from Eq. (2.8)

$$(1 - B)^d \phi(B) x_t = \theta(B) z_t \tag{2.12}$$

which will henceforth be referred to as a general (p,d,q) model referring to the general (autoregressive, differencing, moving average) process.

Examples of Nonstationary Processes

Random walk

Consider Eq. (2.5) in which $\phi_1 = 1$ such that

$$x_t = x_{t-1} + z_t \tag{2.13}$$

In closed form, Eq. (2.13) becomes

$$x_t = x_0 + \sum_{t=1}^{n} z_t \tag{2.14}$$

which is the summation of white noise or a nonstationary random walk process. Equation (2.13) differenced, i.e.,

$$x_t - x_{t-1} = z_t \tag{2.15}$$

is white noise which is a stationary random process.

Random walk and random ramp

Let the white noise term z_t contain a bias b such that

$$z_t = \omega_t + b \tag{2.16}$$

Inserting Eq. (2.16) into Eq. (2.13) yields

$$x_t = x_{t-1} + \omega_t + b \tag{2.17}$$

In closed form, Eq. (2.17) becomes

$$x_t = x_0 + \sum_{t=1}^{n} \omega_t + nb \tag{2.18}$$

which is a combination of the nonstationary random walk and nonstationary ramp processes. Equation (2.17) differenced, i.e.,

$$x_t - x_{t-1} = \omega_t + b \tag{2.19}$$

is stationary white noise plus a bias b which could arise if the random "shocks" to the system had some deterministic phenomenon.

As a final example of the combination of stationary and nonstationary processes, consider a (1,1,1) process formed from Eq. (2.12). In addition, let there be a bias in the white noise. The form of the equation would be

$$(1 - B)(1 - \phi_1 B)x_t = (1 - \theta_1 B)z_t \tag{2.20}$$

clearing yields

$$\begin{aligned}(1 - B - \phi_1 B + \phi_1 B^2)x_t &= (1 - \theta_1 B)z_t \\ x_t - x_{t-1} - \phi_1 x_{t-1} + \phi_1 x_{t-2} &= z_t - \theta_1 z_{t-1} \qquad (2.21)\\ x_t - x_{t-1} &= \phi_1(x_{t-1} - x_{t-2}) + z_t - \theta_1 z_{t-1}\end{aligned}$$

now let

$$z_t = \omega_t + b,\ z_{t-1} = \omega_{t-1} + b \tag{2.22}$$

Inserting Eq. (2.22) into Eq. (2.21) and clearing yields

$$\begin{aligned}x_t - x_{t-1} &= \phi_1(x_{t-1} - x_{t-2}) + \omega_t + b - \theta_1(\omega_{t-1} + b)\\ x_t - x_{t-1} &= \phi_1(x_{t-1} - x_{t-2}) + \omega_t - \theta_1\omega_{t-1} + b(1 - \theta_1)\end{aligned} \tag{2.23}$$

where the term $b(1 - \theta_1)$ is the bias in the differenced data while b is the bias in the white noise. Call the differenced data

$$x_t - x_{t-1} = y_t \tag{2.24}$$

Substituting Eq. (2.24) into Eq. (2.23) yields

$$y_t = \phi_1 y_{t-1} + \omega_t - \theta_1\omega_{t-1} + b(1 - \theta_1) \tag{2.25}$$

and it is seen how differencing induces a stationary (1,0,1) process from a nonstationary (1,1,1) process. It can also be seen that the residual bias, $b(1 - \theta_1)$, remaining on the differenced data, y_t, is the indication of the nonstationary ramp in the original data. Estimation of the parameters ϕ_1, θ_1, and b and the significance of the term $b(1 - \theta_1)$ from zero will be considered shortly.

Identification—Autocovariance Functions

The autocovariance function (acvf) or autocorrelation function (acf) of a time series or random process is the most important tool in mathematical modeling of long term gyro drift rate for the following four reasons.

1) It is an indication of a nonstationarity in the gyro drift rate process.

2) It aids in determining the stationary part of the gyro drift rate math model.

3) Once the stationary part of the math model is determined, the white noise residuals are tested for a bias to see if the original nonstationarity contained a ramp.

4) Once the math model of the drift rate has been identified, its autocovariance function is used in computer programs for the statistical evaluation of inertial navigation systems.

Before proceeding further let us define exactly what is meant by autocovariance and autocorrelation functions.

Let x_t be the time series data. The autocovariance function, $\gamma_{xx}(k)$, is defined as (assuming zero mean)

$$\gamma_{xx}(k) = \frac{1}{T-k}\sum_{t=0}^{T-k} x_t x_{t+k} = \operatorname{cov}[x_t x_{t+k}] = E[x_t x_{t+k}]$$

for $k = 0$ (2.26)

$$\gamma_{xx}(0) = \frac{1}{T}\sum_{t=0}^{T} x_t^2 = \operatorname{cov}[x_t x_t] = E[x_t x_t] = \overline{x_t^2} \tag{2.27}$$

The autocorrelation function, $\rho_{xx}(k)$, is defined as

$$\rho_{xx}(k) = \gamma_{xx}(k)/\gamma_{xx}(0) \tag{2.28}$$

In other words, the autocovariance function normalized by the mean squared value equals the autocorrelation function. Applying the definition of autocovariance in Eq. (2.26) to Eq. (2.5), the acvf function of an ar(1) process is given as

$$\gamma_{xx}(k) = \frac{\phi_1|k|\sigma^2}{1 - \phi_1^2} \tag{2.29}$$

where

$$\operatorname{cov}[z_t z_{t+k}] = \begin{cases}\sigma^2 & k = 0\\ 0 & k \neq 0\end{cases}$$

The acf function is given as $\rho_{xx}(k) = \phi_1|k|$. Since $-1 < \phi_1 < 1$, the autocovariance function for the ar(1) process tails off as k increases.

By using Eq. (2.3) in conjunction with Eq. (2.26), the acvf and acf functions for an MA(1) process are

$$\gamma_{xx}(k) = \begin{Bmatrix}(1+\theta_1^2)\sigma^2 & k = 0\\ -\theta_1\sigma^2 & k = \pm 1\\ 0 & |k| \geq 2\end{Bmatrix} \tag{2.30a}$$

$$\rho_{xx}(k) = \begin{Bmatrix}1 & k = 0\\ \dfrac{-\theta_1}{1+\theta_1^2} & k = \pm 1\\ 0 & |k| \geq 2\end{Bmatrix} \tag{2.30b}$$

Notice that the acvf function for the MA(1) process truncates after the first lag ($k = 1$). This is an important indicator of an MA process where in general the number of lags in which the acvf function is not equal to zero is an indication of the order of the MA process.

Another means of identifying the math model of a time series is now considered, namely partial correlation functions.

Identification—Panel Correlation Functions

It was shown previously that the theoretical acf for a pure MA process of order q truncates after lag q while the acf for a pure ar process is infinite in extent [see Eqs. (2.29) and (2.30)]. Moving average processes are thus characterized by truncation of the acf whereas ar processes are characterized by attenuation of the acf. Thus the acf is a powerful tool for deciding whether the process is pure ar or pure MA. However, in most instances it is not clear whether the acf truncates or attenuates. Thus the concept of partial correlation functions becomes an additional tool in the model identification of a random time sequence.

The partial correlation function can be thought of as being generated as follows. Fit an ar model of order 1 and estimate ϕ_1. Then fit an ar model of order 2 and estimate ϕ_2. Then fit an ar model of order 3 and estimate ϕ_3 and so on. In other words, we fit models of higher- and higher-order of k and estimate the last coefficient ϕ_k. Then we plot ϕ_k (the kth ϕ) as a function of k. When we have fitted models of order higher than the order of the given data series, we expect to see the ϕ's drop to "near zero." Figure 1a shows an example of a partial correlation function ϕ_k vs k for an ar process. If we obtained a partial correlation function like that shown in Fig. 1a, we might conclude that an ar(3) model was a good model to choose. Consider what is meant by near values of a correlation function.

A value for an acvf at some lag value k can be considered not significantly different from zero if it is less than $2/(n)^{1/2}$

times the mean squared value, i.e.,[2]

$$\gamma_{xx}(k) \leq 2\gamma_{xx}(0)/(n)^{1/2} \approx 0 \tag{2.31}$$

where n = number of data points. When using the acvf and partial correlation functions as identification tools, we would normalize the acvf to the acf by dividing through by the mean squared value. Thus Eq. (2.31)can be written as

$$\rho_{xx}(k) \leq 2\rho_{xx}(0)/(n)^{1/2} \approx 0 \tag{2.32}$$

where $\rho_{xx}(0) = 1$ and $2/(n)^{1/2}$ is the 95% confidence limit on $\rho_{xx}(k)$ which says that $\rho_{xx}(k)$ is not significantly different from zero. If $\rho_{xx}(k) \geq 2/(n)^{1/2}$ then $\rho_{xx}(k)$ cannot be considered zero.

Figure 1b shows the 95% confidence limits of Fig. 1a. Notice that values of ϕ_k for $k > 4$ are within the 95% confidence limits and thus can be considered to be zero. This figure would thus indicate an ar(3) process.[2]

It was previously shown that the acf of an ar process trails off as a function of k and the acf of an MA process truncates at the number of lags k equal to the order of the MA process. Conversely, it was seen that the partial correlation function of an ar process truncates at the number of lags k equal to the order of the ar process.

By using the B notation [as in Eq. (2.7)] in Eqs. (2.2) and (2.4), it can be shown that the ma process and the ar process are duals of one another, i.e., a finite ar process is equivalent to an infinite MA process and a finite MA process is equivalent to an infinite ar process.

The use of the partial correlation function in conjunction with the acf as a time series identification tool can be summarized as follows. Whereas the acf of an ar process tails off, its partial correlation function has a cut-off. Conversely, the acf of an MA process has a cut-off while its partial correlation function tails off. Hence if the acf and partial correlation function exhibit opposite behavior, we assume that the process is either pure ar or pure MA. If both the auto and partial correlation functions tail off or are finite, it can be concluded that a mixed ar MA process is called for.

This point is illustrated in Fig. 2 which shows the acf and partial correlation functions for pure ar and MA processes. The dashed lines indicate the 95% confidence limits for zero correlation.

Fig. 1a Example of partial correlation function for a third order autoregressive process.

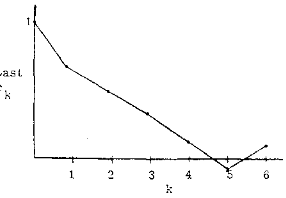
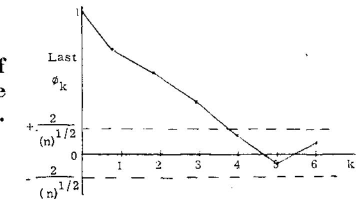


Fig. 1b Example of 95% confidence limits applied to Fig. 1a.



Taking logs of both sides, using Eq. (3.1) in conjunction with Eq. (3.2), and minimizing with respect to the ϕ's, results in the maximum likelihood estimates of the ϕ parameters.

Having established the ϕ's, Eq. (3.1) is cast into the time series $z_t = x_t - \phi_0 - \phi_1 x_{t-1} \ldots - \phi_p x_{t-p}$. The acf of the z_t series should be white noise if the order of the ar process is correct. The minimum $E[z_t^2]$ determines the best order of the ar process.

3. Estimation of Parameters

Having preliminarily identified the model of the time series by use of the acf and partial correlation functions, we turn now to the problem of estimation, which consists of using the data to estimate and make inferences about values of the parameters (the ϕ's and θ's) conditional on the tentatively identified model. Following the estimate, we perform a diagnostic check which involves the examination of the white noise residuals, z_t, from the fitted models which can result in either a) no indication of model inadequacy, or b) model inadequacy, together with information on how to better describe the series. Thus, the residuals z_t, which should be white noise, would be examined for any lack of randomness, and if the residuals are correlated, this information would be used to modify the model. The modified model would then be fitted and subjected to diagnostic checking.

Maximum Likelihood Estimation of Autoregressive Parameters

Suppose that it is required to fit the autoregressive model

$$x_t = \phi_0 + \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + z_t \tag{3.1}$$

to an observed time series $x_1, x_2, \ldots, x_n$. Under the assumption that z_t is $N(0,\sigma_z^2)$, the joint pdf of z may be written

$$f_{p+1}, \ldots, {}_n(z_{p+1}, z_{p+2}, \ldots, z_n) = \frac{1}{[2\pi\sigma_z]^{(n-p)/2}} \exp - \frac{1}{2\sigma_z^2} \sum_{t=p+1}^{n} z_t^2 \tag{3.2}$$

Estimation of the Parameters of a Moving Average Process

The first question which has to be decided in order to fit a moving average process

$$x_t = \theta_0 + z_t - \theta_1 z_{t-1} - \ldots - \theta_q z_{t-q} \tag{3.3}$$

is the appropriate order q for the model. The method of analysis is not as straight forward as that for the autoregressive process because it is difficult to write down an explicit form for the likelihood function of the process. However, a numerical technique due to Box and Jenkins at the University of Wisconsin may be used to build up the log-likelihood function recursively.

To illustrate their approach, consider the first order moving average process

$$x_t = z_t - \theta_1 z_{t-1} \tag{3.4}$$

For specified values of θ_1, Eq. (3.4) can be used to generate a set of z_t's from the observed x_t's. Since $E[z_t] = 0$, a good starting value is $z_0 = 0$ and hence $z_1 = x_1$ and $z_2 = x_2 + \theta_1 z_1$ and so on. Hence the sum of squares

$$S(\theta_0,\theta_1) = \sum_{t=1}^{n} z_t^2$$

corresponding to a particular choice of θ_1 may be readily obtained. The sum of squares can then be plotted for values of θ_1. The minimum sum of squares of the z_t determines the best choice of the θ_1 parameter.

In some cases it may be difficult to distinguish whether or not a process is ar or MA. In that case both models are fit to the data, which is usually the case, and the minimum mean squared error, $\{z_t^2\}$ dictates the best model. Sometimes discrimination between the models should be done on the basis of physical reasoning as to which form is more appropriate.

If a mixed ar MA model is called for through examination of the acf and partial correlation function, a combination of the above two parameter estimation techniques is used.

Consider the mixed model

$$x_t = \theta_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + z_t - \theta_1 z_{t-1} \tag{3.5}$$

which may be written $\phi(B)x_t = \theta(B)z_t$ and $\theta^{-1}(B)\phi(B)x_t = z_t$. Hence, by filtering a series x_t, $t = 1, 2, \ldots, n$ by fitting

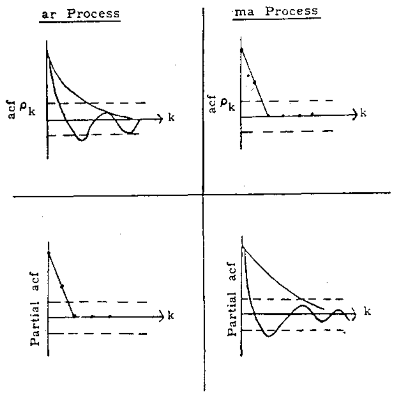


Fig. 2 Autocorrelation and partial correlation functions for pure autoregressive and moving average processes.

an MA process, it is then appropriate to do analytic least squares estimation of the ϕ's. Essentially, the mixed ar MA parameter estimates are based on the minimum mean squared z_t as a function of the θ parameters, and the appropriate ϕ parameters corresponding to that minimum. Having established the moving average and/or autoregressive parameters for the stationary model of the differenced data, the white noise residuals, z_t, are tested for the significance of a bias term, b, in the white noise [See Eq. (2.16)]. Essentially we are testing for the bias in the white noise, which can be considered the "shock" or input to the system, and we want to see if any deterministic features exist in the "shocks" such that when summed up result in a nonstationary ramp in the original process.

The test is a modified version of the "Students t test" which says that the 95% confidence limits on b,

$$b \pm (2\sigma_z/(n)^{1/2}) \tag{3.6}$$

should include zero for there to be no ramp in the original nonstationary process, where b = bias in the white noise and σ_z = rms value of the white noise.

4. Application of Time Series Analysis Techniques to Long Term Gyro Drift Rate

Figure 3 shows a sample of long term gyro drift rate where the numbers were scaled to eliminate any problem arising due to classification. The drift rate waveform exhibits nonstationary behavior, thus requiring data differencing to induce stationarity. The differenced drift rate versus time is shown in Fig. 4. It can be seen that the differenced drift rate exhibits stationary behavior.

Figures 5 and 6 show the acf and partial correlation functions of the gyro drift rate sample, while Table 1 shows the

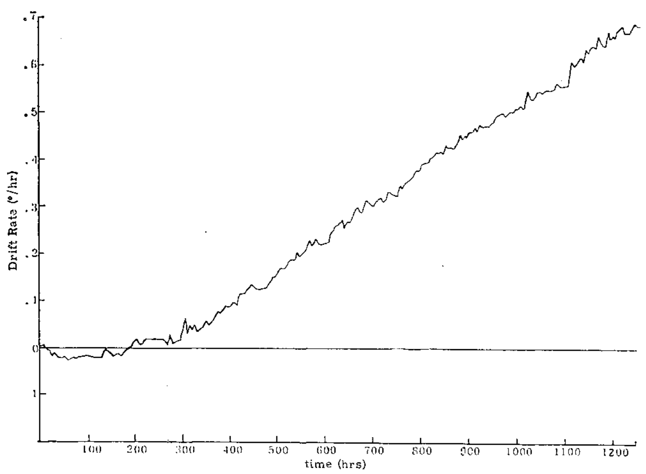


Fig. 3 Normalized gyro drift rate vs time—1250-hr drift rate test.

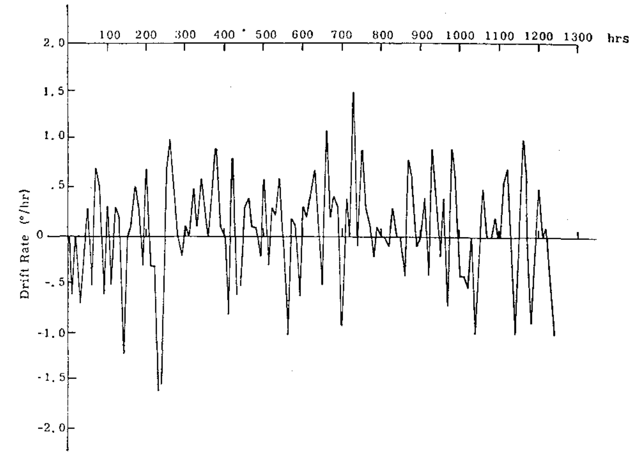


Fig. 4 Differenced gyro drift rate vs time.

results of the analysis to determine the math model parameters of the gyro. The curves and the associated analysis will now be discussed in detail.

Figures 5 and 6 show the acf and partial correlation functions along with the 95% confidence limits band which shows that values within this band are not significantly different from zero.

Figure 5 shows that the acf of the original data does not decay in a stationary manner and thus differencing of the data is necessary. The acf of the first difference, denoted by F points on the curve, does decay in a stationary manner with only one lag value outside the 95% confidence limits. Figure 6 shows the partial correlation function of the first difference has many lag values outside the 95% confidence limits. Examination of Figs. 5 and 6 to determine a possible MA or ar model seems to indicate that a (0,1,1) model would be a good starting point. We determined this because one difference was adequate, there was one lag value in the acf outside the 95% confidence limits, and some sort of decaying process exists in the partial correlation function outside the 95% confidence limits. Table 1 shows the model fit parameters for the gyro drift rate sample. It can be seen that for the (0,1,1) process, the mean squared value of the residuals, $E\{z_t^2\}$ is equal to 0.262 (this number has no actual meaning since the data has been scaled). Recall that z_t is formed by using the estimate of θ_1 and forming z_t from the (0,0,1) equation, i.e.,

$$z_t = x_t + \theta_1 z_{t-1} \tag{4.1}$$

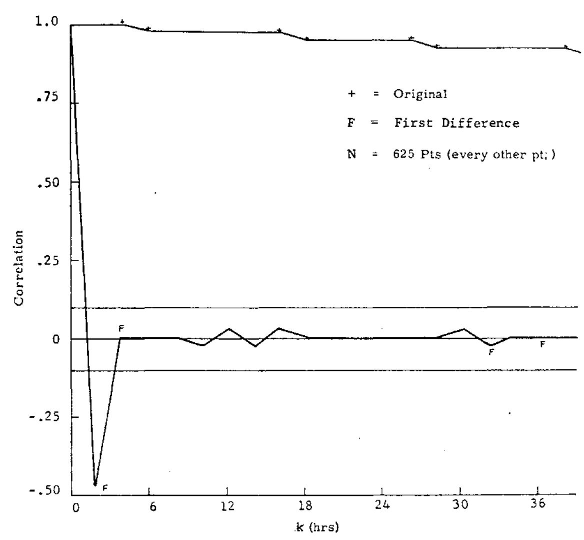


Fig. 5 Autocorrelation of differenced data.

Table 1 Table of various models fit to normalized long term gyro drift rate

95% Confidence limits	Fit	θ_1	θ_2	$\overline{z_t^2}$	ϕ_1	ϕ_2	ϕ_3	acf of Residuals, z_t $K=0$	$K=1$	$K=2$	$K=3$	$K=4$	$K=5$
$\pm \frac{2}{625^{1/2}} = \pm 0.08$	(0,1,1)	0.37	0	0.262	0	0	0	1	−0.055	−0.045	−0.029	−0.022	−0.027
	(1,1,1)	0.37	0	0.261	−0.015	0	0	1	−0.04	−0.046	−0.03	−0.023	−0.028
	(2,1,1)	0.34	0	0.261	−0.066	−0.039	0	1	−0.037	−0.037	−0.052	−0.035	−0.035
	(3,1,1)	0.31	0	0.260	−0.12	−0.07	−0.04	1	−0.036	−0.037	−0.037	−0.05	−0.042

Also shown in Table 1 is the acf of the z_t process. It is seen that all correlated values at lags not equal to zero are not significantly different from zero which shows that z_t is white noise, as it should be for an acceptable model. Also shown in Table 1 are other model fits. It is seen that a (1,1,1) process lowers the mean squared value of the residuals slightly while introducing an ar value of $\phi_1 = -0.02$. The residuals are again white noise. It can be seen in the parameter estimates that by increasing the order of the ar process, while holding the order of the MA process fixed, we are trading off MA for ar and only slightly reducing the mean squared value of the residuals. Thus for purposes of statistical analysis, we would probably go with the simpler model, the (0,1,1) process, and choose the MA(1) model as the stationary gyro drift rate math model with $\theta_1 = 0.37$.

Having then established the stationary model of the gyro drift rate process, it is now necessary to determine the form of the nonstationary model by testing for a bias in the residual white noise. For normalized values of $b = 0.085$ and $\sigma_z = (E[z_t^2])^{1/2} = 0.51$ and for $n = 625$, application of Eq. (3.6) yields a significant bias in the white noise indicating a nonstationary ramp in the original data. The math model of the long term gyro drift rate sample shown in Fig. 3 can now be stated to be of the form

$$x_t = x_{t-1} - \theta_1 z_{t-1} + z_t + b(1 - \theta_1) \tag{4.2}$$

which is the same as Eq. (2.23) with $\phi_1 = 0$ and consists of a stationary MA(1) process and the nonstationary random walk and ramp processes.

The validity of the gyro drift rate math model and the resulting parameter estimates will now be examined by considering the single sample drift rate waveform in an ensemble sense. First let us show the exact form of the equation for the mean squared gyro drift rate using the model of long term gyro drift rate shown in Eq. (4.2). Equation (4.2) can be written in closed form as

$$x_t = x_0 + \sum_{t=1}^{n} z_t - \theta_1 \sum_{t=0}^{n-1} z_t + nb(1 - \theta_1) \tag{4.3}$$

Now consider the mean squared value of Eq. (4.3),

$$E[x_t^2] = E\left[x_0 + \sum_{t=1}^{n} z_t - \theta_1 \sum_{t=0}^{n-1} z_t + nb(1 - \theta_1)\right]^2 =$$
$$E\{x_0^2\} + n\sigma_z^2(1 - \theta_1^2) - 2(n-1)\theta_1\sigma_z^2 + n^2b^2(1 - \theta_1)^2 \tag{4.4}$$

where $E[z_t^2] = \sigma_z^2$ and x_0, z_t, and b are assumed to be independent random variables. For large n and $x_0 = 0$

$$\{x_t^2\} = n\sigma_z^2(1 - \theta_1)^2 + n^2b^2(1 - \theta_1)^2 \tag{4.5}$$

Let us now discuss the two terms in Eq. (4.5). The first term is the mean squared value of the random walk coefficient which is modified by the moving average parameter θ_1. Note from Eqs. (2.14), (2.15), and (4.4) that the variance of the white noise process is also the coefficient of the random walk process.

The second term is the bias in the differenced data squared which increases as n^2. It can be seen that for even a small value of b, the term n^2b^2 grows large quite rapidly.

To cast the single sample time series in the form of an ensemble for model and parameter verification, the gyro drift rate sample was broken into 64 hr segments, the first point removed from each segment, and the 64 hr ensemble mean squared gyro drift rate obtained. The curve is shown in Fig. 7. Superimposed on the empirical mean squared gyro drift rate is its theoretical mean squared value using Eq. (4.5) and the appropriate model parameters. Also shown in Fig. 7 is the theoretical mean squared gyro drift rate for

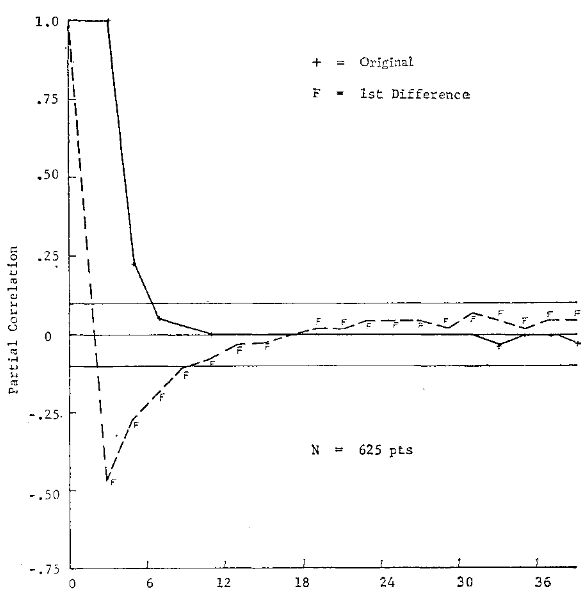


Fig. 6 Partial correlation of differenced data.

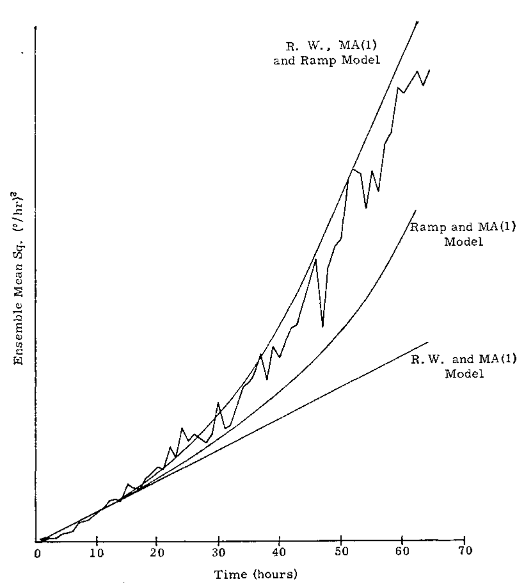


Fig. 7 Normalized ensemble mean squared gyro drift rate vs time.

only the ramp and MA(1) model, and for only the random walk and MA(1) model. Figure 7 shows that model adequacy is attained only by using the combination of two nonstationary processes and one stationary process. Note that a good fit of the data is not attained with only one nonstationary process. It can readily be seen from Eq. (4.5) that the stationary θ_1 parameter plays an important part in determining the fit of the theoretical to the empirical ensemble mean squared gyro drift rate.

It is important to note at this time that the "goodness of fit" of the theoretical mean squared gyro drift rate to the empirical ensemble mean squared gyro drift rate is based on the models and parameter values determined from the *single sample* long term gyro drift rate. The single sample time series analysis techniques give us time estimates of ensemble parameters, where the particular time series analyzed can be considered as one member of the ensemble.

5. Conclusions

The detailed methods by which one would analyze long term gyro drift rate have been presented in this paper. The motivation behind this study was to present a uniform approach for determining a mathematical model for the drift rate of a gyro that must operate in an inertial system over a prolonged period of time.

It has been demonstrated in this paper that: 1) the single sample time series analysis techniques are capable of handling random walks and ramps appearing together in time series data and 2) the technique of ensemble averaging is a useful tool in determining model adequacy, even though the model is determined from the single sample process.

References

[1] Oravetz, A. S., Jakimczyk, S. J., and Sandberg, H. J., "Mathematical Modeling of Long Term Gyro Drift Rate," *EASCON '68 Conference Proceedings*, Sept. 1968.

[2] Box, G. E. P. and Jenkins, G. M., *Statistical Models for Forecasting and Control*, Holden-Day, to be published.

[3] Jenkins, G. M. and Watts, D. G., *Spectral Analysis and Its Applications*, Holden-Day, 1968.

OCTOBER 1970 AIAA JOURNAL VOL. 8, NO. 10

Minimum-Fuel Thrust-Limited Transfer Trajectories between Coplanar Elliptic Orbits

E. A. Kern*
TRW Systems Group, Houston, Texas

AND

D. T. Greenwood†
The University of Michigan, Ann Arbor, Mich.

A method is developed for the computation of minimum-fuel transfer trajectories between coplanar elliptic orbits with a thrust-limited variable-mass rocket moving in a central gravitation force field. Each orbit is defined through the eccentricity, semilatus rectum, and argument of pericenter. Transfer time is left open. The minimum-fuel trajectory is assumed to consist of two thrusting phases separated by a coasting phase. Computation of the minimum-fuel transfer trajectory is accomplished by a direct integration of the rocket equations of motion and the associated adjoint equations. This direct approach is made possible by a transformation of the adjoint equations into a set of equations which provide a much better understanding of the general behavior of minimum-fuel transfer trajectories. An IBM 7094 digital computer program with primarily single-precision arithmetic is used for the computation. Rapid convergence is obtained over a broad class of transfer trajectories and rocket thrust levels.

Nomenclature

A = $ze \cos\phi$
a = ratio of maximum rocket thrust to initial rocket weight
B = $ze \sin\phi$
c = rocket effective exhaust velocity
e = orbital eccentricity
f = true anomaly
H = Hamiltonian
H_0 = a portion of the Hamiltonian
H_1 = a portion of the Hamiltonian
h = $1/r$
m = instantaneous rocket mass
p = semilatus rectum
r = distance from center of force to the rocket
s = switching function (s positive implies thrust is on)
t = time
u = radial velocity/r
v = transverse velocity/r
w = reciprocal of the instantaneous mass
z = reciprocal of the per-unit-mass angular momentum
β = $\tan^{-1}(\lambda_\theta/\nu)$

Presented as Paper 69-914 at the AIAA/AAS Astrodynamics Conference, Princeton, N.J., August 20–22, 1969; submitted August 13, 1969; revision received April 20, 1970. This paper represents a portion of E. Kern's Ph.D. thesis for the Aerospace Engineering Department, The University of Michigan. It was supported by The Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio.

* Formerly with the U.S. Air Force Institute of Technology; now member of the Technical Staff. Member AIAA.

† Professor, Aerospace Engineering Department. Associate Fellow AIAA.