Methods for Validating Computer Simulation Models of Missile Systems

Douglas C. Montgomery*
Georgia Institute of Technology, Atlanta, Georgia
and
Latricha Greene†
Army Missile Laboratory, U.S. Army Missile Command, Huntsville, Alabama

An overview of simulation model validation methodology is presented, with emphasis on how these techniques may be applied in the missile simulation environment. Because many missile simulations produce an output time series that must be matched to corresponding flight or other laboratory test data, the primary focus is on time series methods. Univariate spectral analysis and cross-spectral analysis techniques useful in simulation model validation are described. Two examples of the use of spectral analysis in validation computer simulation models are given.

Nomenclature

$C_{xy}(\omega)$ = coincident spectral density (co-spectrum) of x_t and y_t
$\hat{C}_{xy}(\omega)$ = sample co-spectrum of x_t and y_t
$F_{p,a,b}$ = pth percentage point of the F distribution with a numerator and b denominator degrees of freedom
$f_{xx}(\omega)$ = sample spectrum of the time series x_t
$f_{xy}(\omega)$ = sample cross-spectrum of x_t and y_t
$K^2_{xy}(\omega)$ = squared coherency between x_t and y_t
$\hat{K}^2_{xy}(\omega)$ = sample squared coherency between x_t and y_t
$q_{xy}(\omega)$ = quadrature spectral density of x_t and y_t
$\hat{q}_{xy}(\omega)$ = sample quadrature spectral density of x_t and y_t
$t_{p,\nu}$ = pth percentage point of t distribution with ν degrees of freedom
x_t = time series observed in flight test
y_t = time series observed in computer simulation
$\phi_{xx}(\omega)$ = spectrum of the time series $x_t\pi$
$\phi_{xy}(k)$ = cross-spectrum of x_t and y_t at lag k
$\phi_{yy}(\omega)$ = spectrum of the time series y_t
γ_k = autocovariance function of a time series at lag $k=0,1,2,\ldots$
$\gamma_{xy}(k)$ = autocovariance function for x_t and y_t at lag k
λ_k = weights used in estimating the sample spectrum
$\theta_{xy}(\omega)$ = phase spectrum of x_t and y_t
$\hat{\theta}_{xy}(\omega)$ = sample phase spectrum of x_t and y_t
μ_x = mean of the time series x_t
μ_y = mean of the time series y_t
ω = frequency $0\leq\omega\leq\pi$

Introduction

MISSILE flight and performance characteristics are investigated using several different types of testing, including actual flight tests, laboratory tests of major system components, and computer simulation studies. For example, in addition to flight tests, extensive laboratory tests of the electronic subsystems and wind tunnel tests of the airframe are conducted. Computer simulation models may be used to gain further information about missile performance at a particular set of flight conditions (such as target type, speed, altitude, and maneuver; environmental factors including windspeed, temperature, and visibility, and so on), or they may be used to explore sets of conditions where little or no flight testing has been performed. Various types of computer simulation models have been used in analyzing missile systems, including pure digital simulators, hybrid digital-analog simulators, and hardware-in-the-loop simulators.

Model validation is an important aspect of using computer simulation in the test program for a missile system. By validation we mean an investigation of the consistency of the simulation model with the real missile system. Successful validation provides a basis for confidence in the model results, and is a necessary step if the model is to be used to draw valid inferences about the behavior of the real missile. A reasonable definition of validity is that a set of input conditions to the model should produce output similar to that produced by the real missile system when it was exposed to the same input. Consequently, methods for comparing computer simulation model output to data generated during actual flight tests (or laboratory tests of major components or subsystems) are typically used for model validation.

This paper is a review of methods useful for validation of computer simulation models of missile systems. The variables of interest in analyzing missile performance are often time dependent and thus constitute a *time series*. Consequently, this paper focuses on time series analysis methods, emphasizing the usefulness of spectral and cross-spectral methods. For other more general discussions of simulation model validation, see Refs. 1-3.

Review of Validation Methodology

Missile performance characteristics are either static or dynamic. Static performance characteristics do not vary over the time of flight and are usually represented by a single number or a vector of values at the end of flight. Examples of static performance characteristics are kill probability and terminal miss distance. Dynamic performance characteristics are output phenomena that vary continuously during missile flight, such as roll position, roll rate, wing deflection, lateral accelerations, and various guidance system parameters. These characteristics are usually observed as discrete time series.

In this section we review some of the more important validation techniques. References are provided that describe the procedures in more detail.

Validation with Static Performance Measures

Many published discussions of simulation model validation focus on the analysis of static performance measures.

Received Aug. 19, 1981; revision received July 8, 1982. This paper is declared a work of the U.S. Government and therefore is in the public domain.

*Professor, School of Industrial & Systems Engineering.
†Statistician, Systems Simulation and Development Directorate.

Relatively standard statistical procedures, such as hypothesis testing methods, confidence intervals, regression analysis, and the design of experiments can be used in this context. Statistical methods are particularly useful in validating Monte Carlo simulations, when different input streams of random numbers can produce different realizations of the output variable of interest.

For example, suppose that replication using n different random number streams results in n observations $x_1, x_2, \ldots, x_n$ on terminal miss distance. Now we may view these observations as a sample from some population with mean μ and variance σ^2. The sample mean $\bar{x}$ and sample variance s^2 are unbiased estimators of μ and σ^2, and, if the distribution of the population of terminal miss distances is not too different from the normal distribution, then an approximate 100 (1 $-\alpha$) % confidence interval on mean terminal miss distance is

$$\bar{x} - t_{\alpha/2,n-1} s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1} s/\sqrt{n} \tag{1}$$

This confidence interval can be useful in assessing the validity of the simulation model. For example, if μ_0 represents the value of terminal miss distance observed in actual flight, then if μ_0 lies between the lower and upper $100(1-\alpha)\%$ confidence limits in Eq. (1) ($-t_{\alpha/2,n-1}s/\sqrt{n}$ and $t_{\alpha/2,n-1}s/\sqrt{n}$, respectively), then at the $100(1-\alpha)\%$ level of confidence we may conclude that the simulation model is consistent with the results observed during flight test. Note that this is equivalent to testing the statistical hypothesis that the mean terminal miss distance generated by the Monte Carlo simulation model is identical to the value of terminal miss distance observed in the flight test.

Other basic statistical methods can be adapted directly for use in analyzing static performance measures for Monte Carlo simulation. For an introduction to statistical methods in data analysis, see Refs. 4-7. Fishman[8] gives a good general description of the use of these methods in simulation model validation. For an excellent example of validation using basic statistical methods, see Ref. 9.

Validation with Dynamic Performance Measures

Dynamic performance measures vary continuously during the time of missile flight. These characteristics are usually observed as discrete time series. Let x_t be the time series of interest observed in the actual flight test and y_t be the corresponding time series generated by the computer simulation model, $t=1,2,\ldots,T$. These time series are usually highly autocorrelated, and may exhibit other internal structure (such as nonstationarity, the presence of deterministic components, and so on). To validate the computer simulation model, we must compare the two times series x_t and y_t to determine if they are statistically equivalent; that is, statistical equivalence refers to the degree of confidence that observed differences in the time series are statistically significant.

A variety of methods can be used to compare the time series x_t and y_t. In the validation of missile systems, nonstatistical methods are often used. The most common method of nonstatistical comparison involves plotting the time series x_t and y_t, overlaying the plots, and sliding them along until as close a match as possible is obtained. Then the analyst determines subjectively whether or not the output time series from the simulator agrees with the flight test results. A major difficulty with this approach is that it does not quantify the risk associated with any decision, and it is entirely possible that different analysts will arrive at different conclusions.

Another nonstatistical procedure sometimes used in validating computer simulation models is Theil's inequality coefficient (TIC)[10-12]:

$$\mathrm{TIC} = \sqrt{\sum_{t=1}^{T} (x_t - y_t)^2} \Bigg/ \left(\sqrt{\sum_{t=1}^{T} x_t^2} + \sqrt{\sum_{t=1}^{T} y_t^2} \right) \tag{2}$$

This coefficient is an index that measures the degree of conformance of one time series with another. Clearly, $0 \leq \mathrm{TIC} \leq 1$. If TIC is near zero, then the time series are very different, while if TIC is near unity they are nearly identical. Values of TIC that exceed 0.7 are usually considered evidence of good agreement between the time series.

Theil's inequality coefficient has been extensively used to validate computer simulation models of missile systems (for example, see Ref. 3). While this procedure is more quantitative than simple visual comparison of time series, there is no standard distribution theory for Theil's inequality coefficient, and so no statistical statements relative to the conformance of the two time series can be made.

Goodness-of-fit testing has been applied in simulation model validation. This procedure involves testing the hypothesis that the entire sample of data generated by a computer simulation model has the same probability distribution as the sample of data observed in the flight test. Thus our attention is now focused on the conformance of the entire distribution of sample data from the simulation with the flight, and not just on the point-by-point comparison of the two time series. Goodness-of-fit testing is usually performed by forming frequency distributions or histograms of the two time series and then comparing the two histograms. The two-sample Kolmogorov-Smirnov test is a distribution-free test that is highly useful in this regard. For further methods of goodness-of-fit testing, see Ref. 13.

Goodness-of-fit testing may often be viewed as an improvement over static data analysis methods for validation. It is entirely possible that two time series look visually similar when x_t and y_t are plotted vs time, but their probability distributions may differ considerably. Goodness-of-fit testing is designed to help detect such a situation. Its only weakness is that most goodness-of-fit tests require independent observations (random samples) and many computer simulations produce output time series that are highly autocorrelated. Consequently, this autocorrelation may render the goodness-of-fit testing approach less useful.

Several statistical methods may be useful in comparing time series. One approach is to fit an appropriate stochastic model to x_t and y_t, usually an autoregressive integrated moving average model (see Ref. 14), and then compare the two models. If the two models are the same, the inference is that the two time series are the same. For example, we might model x_t and y_t with pure first-order autoregressive models, say

$$x_t = \xi_x + \psi_x x_{t-1} + a_t$$

and

$$y_t = \xi_y + \psi_y y_{t-1} + b_t$$

where ξ_x, ψ_x, ξ_y, and ψ_y are unknown parameters to be estimated from the sample data and a_t and b_t are white-noise error components. Now if the parameters $\xi_x = \xi_y$ and $\psi_x = \psi_y$, and if the variances of the two noise processes are the same, the two time series x_t and y_t are statistically equivalent.

A test for the equivalence of two time series models is described by Hsu and Hunter,[15] who also illustrate the use of the procedure in validating a computer simulation model of an airport. Unfortunately, the two time series could have been generated by the same underlying stochastic model and still differ significantly in certain characteristics, particularly over the relatively short records typically associated with time series obtained from missile systems. For example, the two series could be significantly out of phase, and yet both could have been generated, for example, from the same autoregressive model. Furthermore, differences in phase angle, gain, and frequency usually have specific interpretations to the missile designer. Therefore, he would like to know if such differences are present. Methods of comparing time series based on the spectrum are particularly useful in analyzing these types of characteristics. We now turn to a discussion of these procedures.

Spectral Methods for Simulation Model Validation

Spectral methods have been suggested by many authors for validation of computer simulation models.[9,16,17] The general approach consists of comparing the sample spectra of the simulation model output and the corresponding flight test data to infer how well the simulation matches the flight. This is usually done using univariate spectral analysis methods. Cross-spectral methods are also useful in investigating the interrelationships between two time series.

Univariate Spectral Methods

The spectrum $\phi_{xx}(\omega)$ of a time series x_t is a decomposition of the total variance of the series by frequency over the interval $0 \leq \omega \leq \pi$. Thus $\phi_{xx}(\omega)$ measures the variance contribution to x_t at frequency ω. The spectrum is related to the autocovariance function $\{\gamma_k\}$ of a wide-sense stationary time series by the relationship

$$\phi_{xx}(\omega) = \gamma_0 + 2\sum_{k=1}^{\infty} \gamma_k \cos(\omega k) \qquad (3)$$

Thus the spectrum is the Fourier cosine transform of the autocovariance function. The spectrum, Eq. (3), is estimated by the *sample spectrum*

$$f_{xx}(\omega_j) = \gamma_0 c_0 + 2\sum_{k=1}^{m} \lambda_k c_k \cos(\omega_j k) \qquad (4)$$

where $f_{xx}(\omega_j)$ is an estimate of the spectrum averaged over a band of frequencies centered at $\omega_j = \pi j/m$, $j=0,1,\ldots,m$; m is the number of frequency bands estimated (often called the truncation point); λ_k, $k=0,1,\ldots,m$ are a set of constants or weights, and

$$c_k = \frac{1}{T-k}\sum_{t=1}^{T-k}(x_t - \bar{x})(x_{t+k} - \bar{x}) \qquad k=1,2,\ldots,m$$

is the *sample* autocovariance function. The weights $\{\lambda_k\}$ in Eq. (4) depend on the type of spectral window used in the estimation process (see Refs. 18 and 19). Spectral windows are employed to give a smoother estimate of the spectrum.

Let $f_{xx}(\omega_j)$ denote the sample spectrum of the flight test data and $f_{yy}(\omega_j)$ denote the sample spectrum of the simulation output. To compare these spectra at a specific frequency, say ω_j, construct a $100(1-\alpha)\%$ confidence interval on the ratio of the true spectra, say $\phi_{xx}(\omega_j)/\phi_{yy}(\omega_j)$, using

$$\frac{f_{xx}(\omega_j)/f_{yy}(\omega_j)}{F_{\alpha/2,u,u}} \leq \phi_{xx}(\omega_j)/\phi_{yy}(\omega_j) \leq \frac{f_{xx}(\omega_j)/f_{yy}(\omega_j)}{F_{1-\alpha/2,u,u}} \qquad (5)$$

where $F_{p,u,u}$ is the pth percentage point of the F distribution with $u=2T/m$ degrees of freedom in the numerator and denominator. If the upper and lower confidence limits in Eq. (5) contain the value $\phi_{xx}(\omega_j)/\phi_{yy}(\omega_j)=1$, then we conclude that at that frequency the two time series are identical, in that they have statistically equivalent spectra. The set of all confidence intervals at the frequency points ω_j, $j=0,1,\ldots,m$ is called a confidence band. While the confidence level at any one specific frequency is $1-\alpha$, the confidence level is *not* $1-\alpha$ for the entire confidence band. That is, the statement in Eq. (5) is a one-at-a-time confidence interval.

For the time series to be identical, their spectra must be equal at all frequencies ω_j, $j=0,1,\ldots,m$. The *simultaneous* confidence band allows us to state with probability at least $1-\alpha$ that *all* $m+1$ confidence intervals are simultaneously true. The $100(1-\alpha)\%$ simultaneous confidence band is computed from

$$\frac{f_{xx}(\omega_j)/f_{yy}(\omega_j)}{F_{\alpha/[2(n+1)],u,u}} \leq \phi_{xx}(\omega_j)/\phi_{yy}(\omega_j) \leq \frac{f_{xx}(\omega_j)/f_{yy}(\omega_j)}{F_{1-\alpha/[2(m+1)]u,u}}$$

$$j=0,1,2,\ldots,m \qquad (6)$$

For example, if we wished to make the statement that all $m+1$ frequencies are simultaneously equal with an error probability of at most 0.05 (that is, a 95% simultaneous confidence interval), then the probability level associated with the F-statistic at each frequency is $0.05/2(m+1) = 0.025/(m+1)$. Thus if there are 16 bands in the sample spectrum, then $0.026/16 = 0.00156$ probability is assigned to each tail of the F distribution. Note that the simultaneous confidence interval, Eq. (6), is wider than the individual confidence interval, Eq. (5), at any frequency. For an application of this methodology to the validation of a computer simulation model of a missile system, see Ref. 17.

In constructing either individual or simultaneous confidence intervals, the analyst must specify the level of confidence $100(1-\alpha)$. We may interpret α as the probability that we will conclude that the two spectra differ when they really do not. Typical reasonable values of α are in the interval $0.1 \leq \alpha \leq 0.01$. Generally, as α gets smaller the confidence band widens making it easier to conclude that the two time series match when they really do not, while as α increases the confidence band becomes more narrow, making it easier to conclude that the time series differ when they really do not. What is desired is a reasonably narrow interval at a high level of confidence.

Spectral methods can be applied only to a stationary series. A time series is considered to be stationary if the probability distribution of $\{x_t\}$ is identical to the probability distribution of $\{x_{t+\tau}\}$. That is, if the form of the distribution of x_t does not depend on the period of time in which the time series is observed, then x_t is stationary. If the series is nonstationary, then the nonstationary part of the process must be removed, either by successive differencing or by fitting a polynomial model (or other appropriate function) to the data and analyzing the residuals. Piecewise polynomial fitting may be necessary when the time series exhibits different behavior in different local segments of time. This may be conveniently done using splines. References 2 and 7 are very useful. Indications of nonstationarity are usually observed in either the sample autocorrelation function or the sample spectrum. If the autocorrelation function does not die down even at very long lags or if the power is concentrated at the lowest frequency in the spectrum, then the series is probably nonstationary. Note that if two nonstationary series x_t and y_t are compared, they are equivalent in a frequency sense if both their stationary representations have the same spectrum and if the same level of differencing (or the same order polynomial) is required to reduce both of them to stationarity.

Internal Validity Checking

While simulation-to-flight test comparisons usually form the basis of simulation model validation, it is also necessary to validate the internal logic of the simulation model. Spectral methods can be useful in this aspect of validation also. These techniques can be used to investigate the interrelationships between two time series generated by the simulator. For example, suppose that the computer simulation model produces time series output of fin deflection and airframe lateral acceleration. The lateral accelerations are a physical result of the fin deflections and the nonlinear aerodynamic response of the airframe. Casual or correlative structure between these two time series should be reflected in the analysis.

Interrelationships between two time series x_t and y_t may be investigated using the cross-spectrum. The cross-covariance function is

$$\gamma_{xy}(k) = E(x_t - \mu_x)(y_{t+k} - \mu_y) \tag{7}$$

where μ_x and μ_y are the means of x_t and y_t, respectively. Unlike the autocovariance function for a single time series, the cross-covariance function $\gamma_{xy}(k)$ may not be symmetric about zero. The cross-spectrum is defined as

$$\phi_{xy}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_{xy}(k) e^{-i\omega k} \tag{8}$$

Note that $\phi_{xy}(\omega)$ is a continuous periodic function of ω (the frequency). Since $\gamma_{xy}(k)$ may not be symmetric about zero, the cross-spectrum is in general a complex function, say

$$\phi_{xy}(\omega) = C_{xy}(\omega) - iq_{xy}(\omega) \tag{9}$$

where $C_{xy}(\omega)$ is the coincident spectral density (co-spectrum) and $q_{xy}(\omega)$ is the quadrature spectral density. Both $C_{xy}(\omega)$ and $q_{xy}(\omega)$ are real-valued functions of ω. $C_{xy}(\omega)$ is the cosine portion of the transform and is an even function of ω, while $q_{xy}(\omega)$ is the sine portion of the transform and is an odd function of ω. Let $f_{xy}(\omega)$ denote an estimate of the cross-spectrum and let $\hat{C}_{xy}(\omega)$ and $\hat{q}_{xy}(\omega)$ denote the estimates of $C_{xy}(\omega)$ and $q_{xy}(\omega)$, respectively. For a detailed treatment of the estimation procedure see Refs. 18 and 19.

The squared coherency is defined as

$$K_{xy}^2(\omega) = \frac{|\phi_{xy}(\omega)|^2}{\phi_{xx}(\omega) \cdot \phi_{yy}(\omega)} \tag{10}$$

where

$$|\phi_{xy}(\omega)|^2 = C_{xy}^2(\omega) + q_{xy}^2(\omega)$$

and $\phi_{xx}(\omega)$ and $\phi_{yy}(\omega)$ are the spectra of x_t and y_t, respectively. The coherency is analogous to the coefficient of multiple determination R^2 in multiple regression. Thus coherency is a measure of independence of x_t and y_t at frequency ω. If coherency is 0, then the series are independent (unrelated), while if coherency is 1, then the series are perfectly dependent (related). Coherency is a nondimensional measure of the correlation between two time series as a function of frequency. Another way to interpret coherency is in terms of the predictability of one series from the other. If coherency is 0, then one series cannot be predicted from the other, while if coherency is 1, one series can be perfectly predicted from the other. We may also think of coherency as the proportion of the total power (by frequency) in one time series that can be explained by the other time series.

An F-test for zero coherency is given by

$$F_0 = \frac{4d\hat{K}_{xy}^2(\omega)}{2[1 - \hat{K}_{xy}^2(\omega)]} \tag{11}$$

which if $K_{xy}^2(\omega) = 0$ is distributed as $F_{2,4d}$, where d is the number of points at which the spectrum is estimated, and

$$\hat{K}_{xy}^2(\omega) = \frac{|f_{xy}(\omega)|^2}{f_{xx}(\omega) f_{yy}(\omega)} \tag{12}$$

is an estimate of coherency. If $F_0 > F_{\alpha,2,4d}$, the hypothesis of zero coherency is rejected. While the limiting values of coherency, 0 and 1, are of obvious interest, intermediate values are also of interest because of the natural interpretation of coherency as "percent variability explained." If the coherency function is greater than zero but less than unity, one or more of three possibilities exist:

1) extraneous noise is present in the measurements;
2) the system relating x_t and y_t is not linear; or
3) y_t is an output related to the input x_t as well as to other inputs.

The phase spectrum is defined as

$$\theta_{xy}(\omega) = \tan^{-1}[-q_{xy}(\omega)/C_{xy}(\omega)] \tag{13}$$

and estimated by

$$\hat{\theta}_{xy}(\omega) = \tan^{-1}[-\hat{q}_{xy}(\omega)/\hat{C}_{xy}(\omega)] \tag{14}$$

The phase spectrum shows whether the frequency components in one series lead or lag the components at the same frequency in the other series. If the coherency is zero at frequency ω, that implies that a $100(1-\alpha)\%$ confidence interval on the phase angle is $(-\pi/2, \pi/2)$. That is, the average phase difference between the two processes is zero, but the phase difference is equally likely to lie anywhere in the range $(-\pi/2, \pi/2)$. If coherency is not zero, then a $100(1-\alpha)\%$ confidence interval is

$$\hat{\theta}_{xy}(\omega) - \delta \leq \theta_{xy}(\omega) \leq \hat{\theta}_{xy}(\omega) + \delta \tag{15}$$

where

$$\delta = \sin^{-1}\left[t_{\alpha,4d}^2 \frac{1 - \hat{K}_{xy}^2(\omega)}{4d\hat{K}_{xy}^2(\omega)}\right]^{1/2} \tag{16}$$

Generally speaking, the cross-correlation structure between two time series can be adequately described by their squared coherency and phase spectra. Therefore, it is recommended that internal validity checking concentrate on these measures.

Screening and Preparation of Data

A potentially frustrating problem for the data analyst is dealing with wild or unusual observations in either the observed flight data or the simulation data. These wild or aberrant observations may severely distort the sample spectrum or estimates of the parameters of the underlying distributions. Editing or preliminary screening of the data is often found to be necessary.

Two approaches are useful in this regard. The first of these is to smooth the data with a nonlinear robust filter to eliminate spikey noise. The more popular nonlinear smoothers are usually based on running medians (see Refs. 20 and 21). A second approach is to fit a model to the data that describes the smooth portion of the signal. Fitting methods more robust than least squares are recommended in this approach, as it is well known that least squares is severely distorted by outliers. Robust fitting procedures are reviewed by Hogg.[22] Agee and Turner[23] describe the use of these methods in preprocessing of missile trajectory data.

Examples

In this section, we apply the spectral methods described previously to two examples. The first example illustrates univariate spectral analysis while the second involves cross-spectral methods. Computations were done using an interactive FORTRAN IV program written for the CDC CYBER 74, utilizing a Textronix 4014 Graphics terminal. Details of the computer programs are available from the authors.

Univariate Spectral Analysis

Figure 1 presents 800 realizations of digital simulation model output data (dashed line) and the corresponding flight test data (solid line) for a control variable for a typical missile. The flight test data were derived from range telemetry records made at White Sands Missile Range during a tactically relevant intercept. The simulation data were generated from a digital simulation model exercised using target trajectory data and other inputs similar to those for the flight test. The time

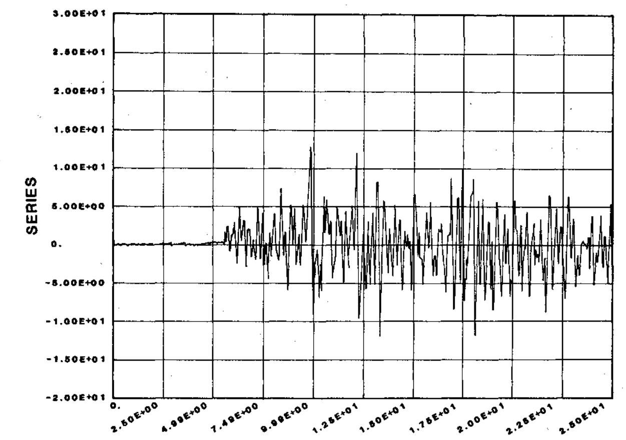

Fig. 1 Simulation and flight test data.

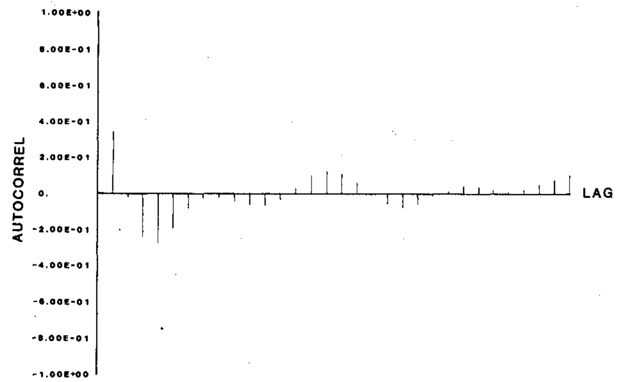

Fig. 2 Sample ACF, simulation data.

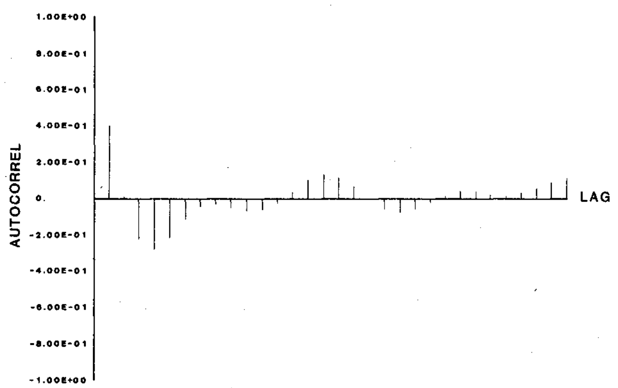

Fig. 3 Sample ACF, flight test data.

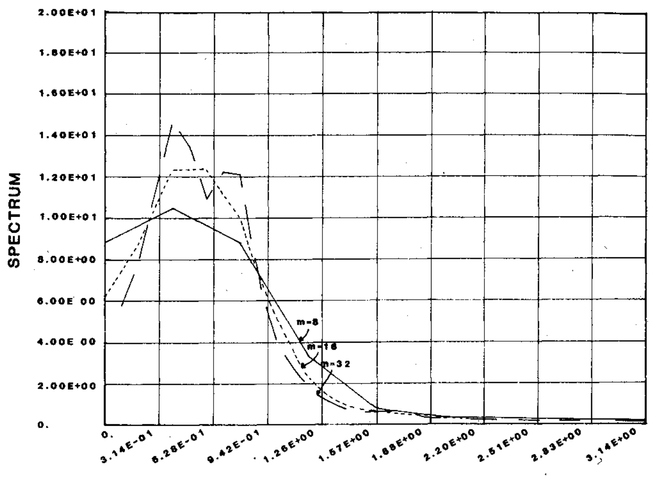

Fig. 4 Sample spectral estimates, simulation data.

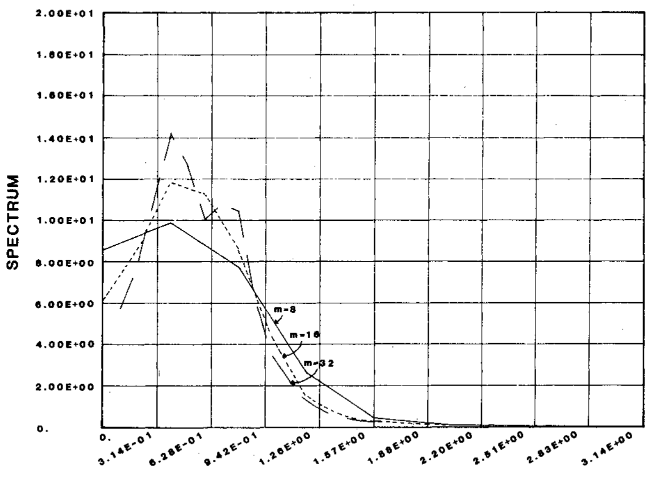

Fig. 5 Smooth spectral estimates, flight test data.

of flight is approximately 25 s. From examining Fig. 1 the visual impression is that the two time series match closely.

The sample autocorrelation functions and spectra, shown in Figs. 2-5, are typical of those associated with second-order autoregressive processes. Each spectra is computed using truncation points of $m = 8$, 16, and 32. There is evidence of a narrow two-sided peak in the low-frequency range. Reasonable estimates of the spectra are obtainable at truncation points of either $m = 16$ or $m = 32$.

Table 1 presents the statistical comparison of the spectra for a truncation point of $m = 16$. Columns b and c of this table list the sample spectra by frequency. Column d is the observed ratio of the simulation and flight test spectra. The 95% one-at-a-time confidence limits, computed from Eq. (5), are shown in columns e and f. These confidence limits would be of interest if the analysts were interested in a specific frequency. Usually we are concerned with all frequencies containing significant power, so the simultaneous confidence limits are most useful in validation.

The 95% simultaneous confidence limits, shown in columns g and h, indicate that at all points in the spectrum containing significant power there is good agreement between the simulation and flight test data. The only simultaneous confidence limits that do not include one are at higher frequencies where the power is low. From this analysis, conclude that at the 95% level of confidence the simulation model and the flight test data are substantially equivalent.

Cross-Spectral Analysis

The second example illustrates how cross-spectral methods can be used to assist the analyst in verifying the internal logic of a computer simulation model. The simulation model investigated is a real-time, hybrid, six-degree-of-freedom, hardware-in-the-loop model of a ground-to-air missile. This simulation model was providing target trajectory data and other inputs similar to those observed during an actual flight test at the White Sands Missile Range.

The two variables chosen for demonstrating the verification procedure are fin deflection and lateral acceleration. These are critical variables in the analysis of missile guidance. In the missile, the fin deflections are produced by a nonlinear pneumatic servomechanism. The nonlinearity of this subsystem is preserved in the mathematical model and the implementation of that model in the hybrid simulation. Similarly, the lateral accelerations are a physical result of the fin deflections and the nonlinear response of the airframe by which those deflections induce turning moment, angle of attack, and side-force. This is explicitly incorporated in the

Table 1 Comparison of spectra

				Individual confidence limits		Simultaneous confidence limits	
Frequency a	x_t Spectrum b	y_t Spectrum c	Ratio d	Lower e	Upper f	Lower g	Upper h
0.000000	6.207218	6.145827	1.009989	0.680929	1.498067	0.553924	1.841548
0.031250	8.667697	8.487142	1.021274	0.688538	1.514806	0.560113	1.862125
0.062500	12.335819	11.797968	1.045588	0.704930	1.550870	0.573448	1.906458
0.093750	12.327129	11.283451	1.092496	0.736555	1.620446	0.599175	1.991987
0.125000	9.982977	8.631579	1.156564	0.779750	1.715475	0.634313	2.108805
0.156250	5.309353	4.301735	1.234235	0.832115	1.830681	0.676911	2.250424
0.187500	2.098018	1.513495	1.386207	0.934574	2.056093	0.760259	2.527520
0.218750	0.963606	0.597802	1.611913	1.086744	2.390872	0.884047	2.939058
0.250000	0.655426	0.329599	1.988556	1.340674	2.949528	1.090615	3.625804
0.281250	0.511515	0.212382	2.408470	1.623778	3.572365	1.320915	4.391447
0.312500	0.319251	0.103718	3.078053	2.075207	4.565525	1.688144	5.612322
0.343750	0.223154	0.047651	4.683142	3.157349	6.946274	2.568448	8.538936
0.375000	0.185645	0.027379	6.780629	4.571464	10.057374	3.718806	12.363358
0.406250	0.160495	0.011932	13.450970	9.068573	19.951165	7.377125	24.525625
0.437500	0.141352	0.005733	24.655609	16.622682	36.570456	13.522259	44.955435
0.468750[a]	0.119460	0.000000					
0.500000[a]	0.108827	0.000000					

[a] No ratio of spectra or confidence limits calculated at these frequencies because the smoothed estimate of the spectrum of the y_t series is zero.

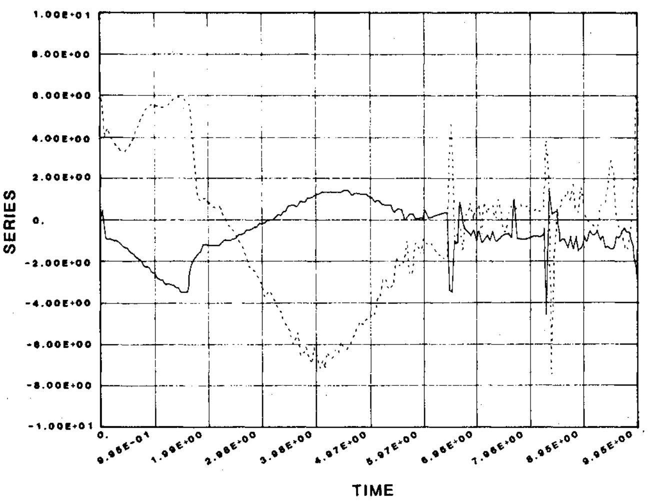

Fig. 6 Plot of fin deflection (–) and lateral acceleration (--) from a hybrid simulation model.

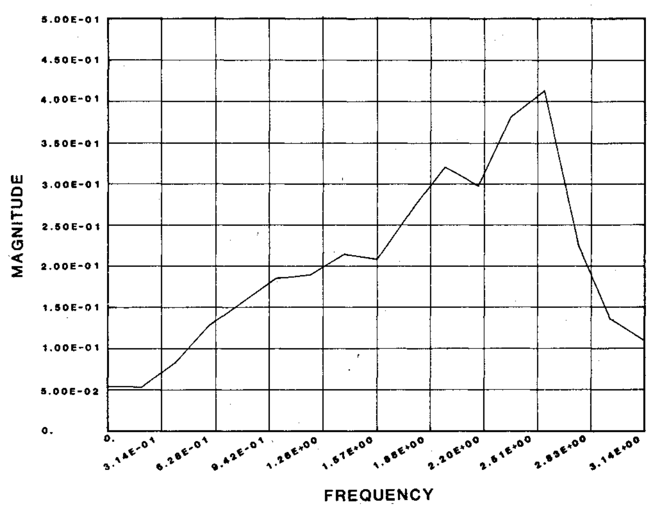

Fig. 8 Smoothed cross-spectrum.

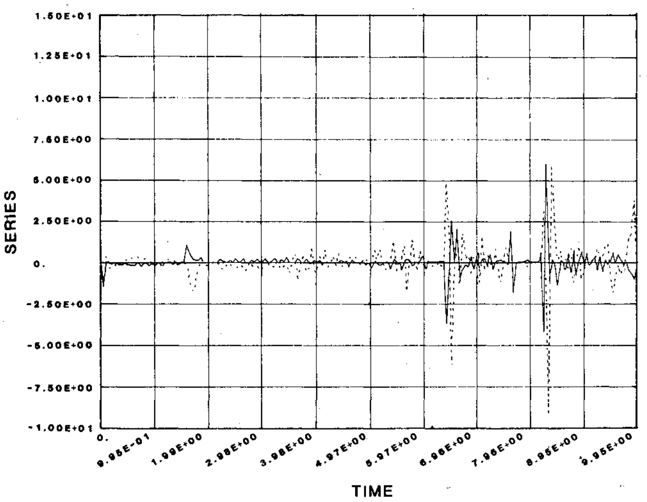

Fig. 7 Plot of first differences of fin deflection (–) and lateral acceleration (--).

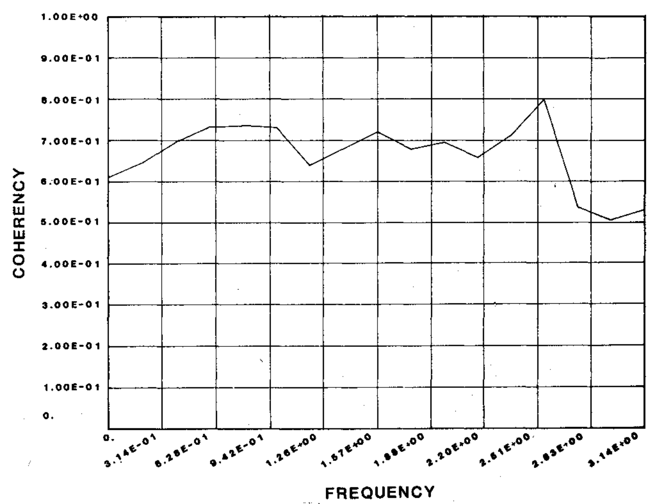

Fig. 9 Squared coherency.

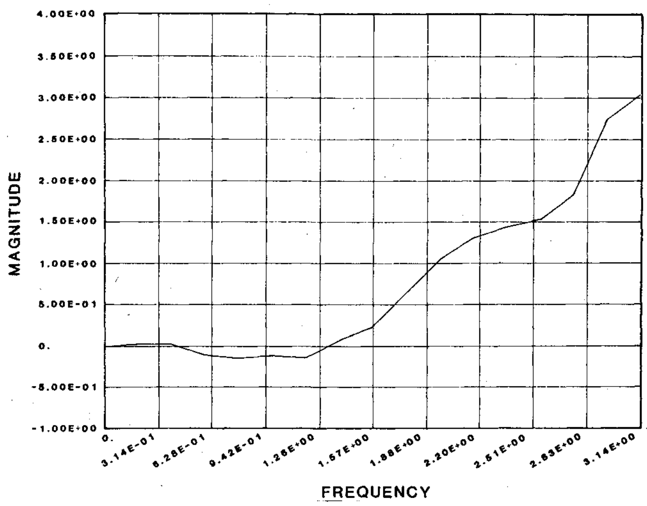

Fig. 10 Phase spectrum.

hybrid simulation model. Consequently, we would expect to find significant coherency between fin deflection and lateral acceleration. The lack of such coherency could be viewed as evidence that either the mathematical modeling or its incorporation in the simulator is incorrect.

Figure 6 shows a plot of the time series of fin deflection and the corresponding lateral acceleration (200 observations). Because these variables are integral to missile translational dynamics, nonstationary behavior is observed in the two time series. The time series of first differences, shown in Fig. 7, exhibit stationarity, although there is still some evidence of internal stochastic structure for both series.

The smoothed cross-spectrum based on a truncation point of $m = 16$ for the series of first differences is shown in Fig. 8. Note that there is significant power in the higher frequencies, and a strong two-sided peak. A plot of coherency is shown in Fig. 9. Coherency varies between 0.5 and 0.8 over the entire frequency range, and the F-test for zero coherency [Eq. (11)] is significant at all frequencies. Recalling that coherency is a measure of linear correlation (or predictability) between fin deflection and lateral acceleration, we conclude that there is a strong linear relationship between the two series. Clearly, the relationship is not entirely linear, and this accounts for the fact that coherency is not extremely close to unity. The phase spectrum is shown in Fig. 10. The low-frequency components of fin deflection and lateral acceleration are in-phase, but at higher frequency fin deflection leads lateral acceleration changes. Again, this behavior is anticipated in light of the physical nature of the variables analyzed. Both the squared coherency and phase spectrum imply that the modeling of the relationship between fin deflection and lateral acceleration is reasonable, and that this relationship has been satisfactorily represented in the simulator.

Conclusion

Development of a valid simulation model is a necessary part of the evaluation of a missile system, if such evaluations are to be performed in an efficient and effective manner. This paper has reviewed several methods for validating computer simulation models of missile systems. Such validation efforts usually take the form of comparing time series that represent some performance measure of interest in both the simulation and the flight test. Univariate spectral analysis methods are a useful way to make these comparisons on a statistical basis. Interval validity checking is also a necessary part of model validation. Cross-spectral analysis is helpful in determining if reasonable relationships exist between physical parameters in the simulation model.

Examples were presented that illustrate the application of these methods in practice. We believe that spectral methods are a useful supplement to the other validation techniques often employed by the missile system designer.

References

[1] Naylor, T.H. and Finger, J.M., "Verification of Computer Simulation Model," *Management Science,* Vol. 14, Oct. 1967, pp. 92-101.

[2] Van Horn, R., "Validation," in *The Design of Computer Simulation Experiments,* T.H. Naylor, ed., Duke University Press, Durham, N.C., 1969.

[3] Kheir, N.A. and Holmes, W.M., "On Validation Simulation Models of Missile Systems," *Simulation,* April 1978, pp. 117-128.

[4] Hines, W.W. and Montgomery, D.C., *Probability and Statistics in Engineering and Management Science,* 2nd ed., John Wiley & Sons, New York, 1980.

[5] Montgomery, D.C., *Design and Analysis of Experiments,* John Wiley & Sons, New York, 1976.

[6] Montgomery, D.C. and Peck, E.A., *Introduction to Linear Regression Analysis,* John Wiley & Sons, New York, 1982.

[7] Draper, N.R. and Smith, H., *Applied Regression Analysis,* 2nd ed., John Wiley & Sons, New York, 1981.

[8] Fishman, G.S., *Concepts and Methods in Discrete Event Digital Simulation,* John Wiley & Sons, New York, 1973.

[9] Naylor, T.H., Wallace, W.H., and Sasser, W.E., "A Computer Simulation Model of the Textile Industry," *Journal of the American Statistical Association,* Vol. 62, Dec. 1967, pp. 1338-1364.

[10] Theil, H., *Economic Forecasts and Policy,* North-Holland Publishing Co., Amsterdam, 1961.

[11] Theil, H., *Optical Decision Rules for Government and Industry,* Rand McNally, Chicago, 1964.

[12] Theil, H., *Applied Economic Forecasting,* Rand McNally, Chicago, 1966.

[13] Conover, W.J., *Practical Nonparametric Statistics,* 2nd ed., John Wiley & Sons, New York, 1980.

[14] Box, G.E.P. and Jenkins, G.M., *Time Series Analysis, Forecasting and Control,* Holden-Day, San Francisco, 1970.

[15] Hsu, D.A. and Hunter, J.S., "Analysis of Simulation-Generated Responses Using Autoregressive Models," *Management Science,* Vol. 24, Oct. 1977, pp. 181-190.

[16] Fishman, G.S. and Kiviat, P.J., "The Analysis of Simulation-Generated Time Series," *Management Science,* Vol. 13, 1967, pp. 525-557.

[17] Montgomery, D.C. and Conard, R.G., "Comparison of Simulation and Flight-Test Data for Missile Systems," *Simulation,* Feb. 1980, pp. 63-72.

[18] Fuller, W.A., *Introduction to Statistical Time Series,* John Wiley & Sons, New York, 1976.

[19] Jenkins, G.M. and Watts, D.G., *Spectral Analysis and Its Application,* Holden-Day, San Francisco, 1968.

[20] Tukey, J.W., *Exploratory Data Analysis,* Addison-Wesley Publishing Co., Reading, Mass., 1977.

[21] Velleman, P.F., "Definition and Comparison of Robust Nonlinear Data Smoothing Algorithms," *Journal of the American Statistical Association,* Vol. 75, Sept. 1980, pp. 609-615.

[22] Hogg, R.V., "Statistical Robustness: One View of Its Use in Application Today," *The American Statistician,* Vol. 33, No. 2, Aug. 1979, pp. 108-115.

[23] Agee, W.S. and Turner, R.H., *Application of Robust Statistical Methods to Data Reduction,* Technical Report 65, U.S. Army White Sands Missile Range, N. Mex., 1978.