STATISTICAL TOOLS FOR SYSTEM DYNAMICS

by

David W. Peterson
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA

ABSTRACT

For questions of parameter choice and validity, the system dynamicist has usually relied on "manual" examination of the detailed structure of the model. Numerical data may be used in the process, but only where the implications are obvious by inspection.

This paper describes practical "automatic" tools to aid both the builder and the evaluator of a system dynamics model. The tools relate the model to available data; they are helpful in answering such questions as:

1. What are the most likely values of unknown parameters, given available data?
2. Which structural formulations are most likely?
3. Is the model consistent with all available data?
4. Which data points are likely to be wrong?
5. What is the most likely state of the system at a given time?
6. To what degree of accuracy can model-computed forecasts be trusted?

The tools are based on full-information, maximum-likelihood via optimal filtering. They operate correctly in an environment of noisy data, missing data points, unmeasured variables, and unknown inputs.

Table of Contents

I. New Capabilitios for System Dynamics

For questions of parameter choice and validity, the system dynamicist has usually relied on "manual" examination of the detailed structure of the model. The realism of both parameter values and model structure is assessed and improved by repeated simulation experiments. If a simulation experiment reveals something surprising or wrong, the modeler asks why, seeks the answer by examining the model structure, and tests the answer by new simulations. This informal procedure of model inspection and simulation is one of the great strengths of the system dynamics methodology. If the modeler proceeds with diligence and thoroughness, the model is greatly improved over "first cut" form, and the modeler gains a deep understanding of the system being modeled.

Numerical data contribute to the process, but usually only when the implications are obvious by inspection. While econometric methods are sometimes employed by system dynamicists (see Hamilton 1969 and Runge 1976 for examples), such use is infrequent. The methods often restrict the form and content of a model according to whatever data is available. Furthermore, system dynamicists have sometimes doubted the validity of standard econometric tools. For example, Senge (1974) shows that generalized least-squares estimation (GLS) may give results which are both wrong and misleading when used in the context of a system dynamics model.

This paper describes a relatively new set of statistical tools which permit the system dynamicist to make full use of numerical data. The tools relate the model to all relevant data, even if the data is incomplete, noisy, and marred by occasional "bad data" points. The tools work with models in which most of the variables are unmeasured, and for cases of unmeasured exogenous inputs. The tools work correctly under the circumstances examined by Senge, and are helpful in answering such questions as:

1. What are the most likely values of unknown parameters, given available data?
2. Which structural formulations are most likely?
3. Is the model consistent with all available data?
4. Which data points are likely to be wrong?
5. What is the most likely state of the system at a given time?
6. To what degree of accuracy can model-computed forecasts be trusted?

The tools are based on the method of full-information, maximum-likelihood via optimal filtering (FIMLOF), as discussed in Schweppe (1973) and Peterson (1975). The mathematics of FIMLOF are stated and referenced in this paper, but the formal proofs will not be repeated here. Instead, the dynamic ideas behind the methods are emphasized here, and related both to alternate methods and to practices in system dynamics. The tools have also been implemented in a computer program, the General Purpose System Identifier and Evaluator (GPSIE), as described in Peterson (1974).

A. Parameter Estimation

The FIMLOF method of parameter estimation is best understood as an optimum compromise between two less satisfactory extremes. One extreme might be called "naive simulation" (NS), and the other extreme is the standard econometric tool, ordinary least-squares (OLS). To make the presentation clear, this discussion will focus on an extremely simple system; but the method also succeeds with nonlinear, multi-state-variable systems.

Consider the system:

$$X(t) = A * X(t - 1) + W(t)$$

$$Z(t) = X(t) + V(t)$$

where X is the state of the system, Z are measurements of X, A is an unknown parameter, W(t) is driving noise ("equation error"), and V(t) is measurement noise ("errors in the variables").

First, take the case where $W(t) = 0$ and $V(t) = 0$, which is equivalent to perfect measurement of a deterministic system. From the data given in Figure 1, in this noise-free case, the parameter A can be estimated by simply taking the ratio between two successive values of Z(t). However, the simple example can help to illustrate more indirect methods, which succeed not only in the noise-free case, but also in more complicated situations. The essence of the more indirect methods, which lead to FIMLOF, is to (1) guess a value of A and simulate the system, (2) measure the error between the simulated data $\hat{Z}(t)$ and the actual data Z(t), and (3) repeat the process, making new guesses of A in an orderly fashion, until a value of A is found to minimize the error. The estimate of A is then chosen as the value which minimizes the error between actual and simulated data.

The differences among NS, OLS, and FIMLOF lie mainly in how the simulation is performed, and in the measure of the error.

Naive Simulation (NS)

In the naive simulation method, the model is initialized at the first data points and simulated without further reference to the data:

$$\hat{X}(t) = A * \hat{X}(n-1)$$
$$\hat{Z}(t) = \hat{X}(t) \quad .$$

In general, the simulated values will not coincide with the data; the differences are called residuals:

$$r(t) = Z(t) - \hat{Z}(t) \quad .$$

The NS sum of squared residuals, also called the loss function, is denoted as J:

$$J = \sum_{t=1}^{N} r^2(t) \quad .$$

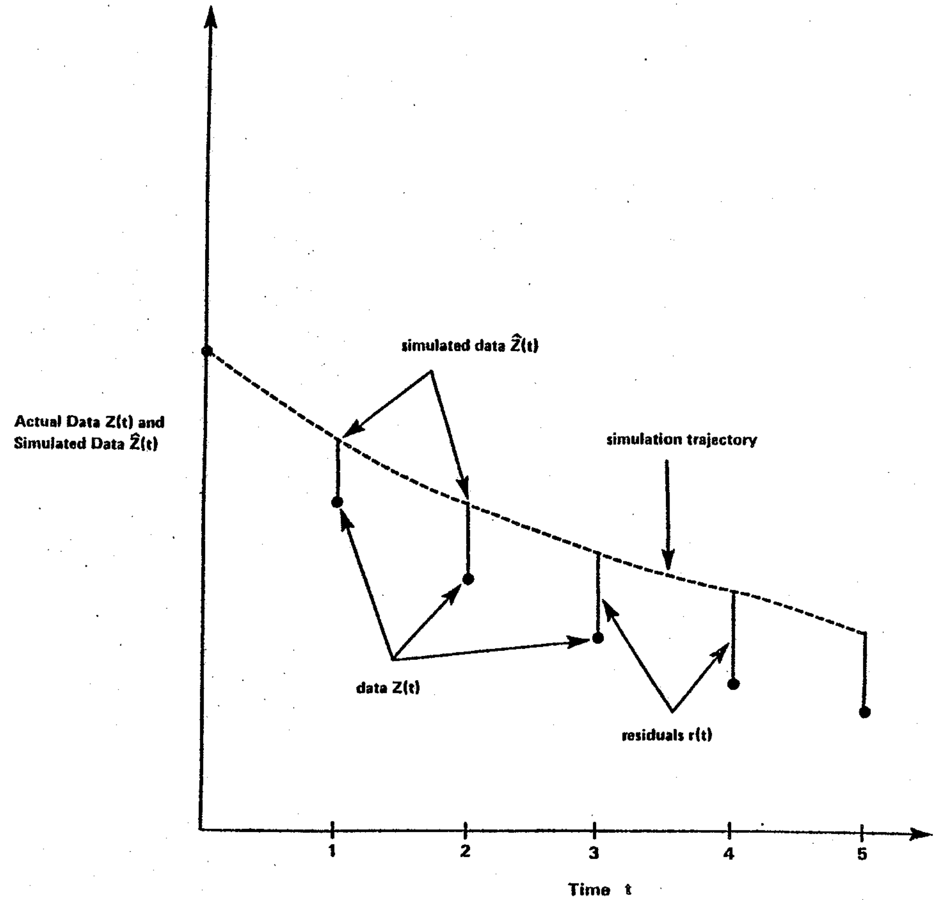

Figure 1: Completely Deterministic Data, with Naive-Simulation Estimation (NS).

In the nonpathological case of Figure 1, the error will be zero if A is guessed correctly. The modeler may guess close to the right A value more efficiently and methodically by using a good "hill climbing" algorithm[1], but the essential idea is to adjust the guess of A until no smaller error can be found.

However, if the system being modeled has equation noise $W(t) \neq 0$, then the naive simulation method may give minimum error J for a completely wrong value of A, since the real system may "drift" away from the deterministic trajectory, as shown, for example, by Forrester (Appendix K, 1961). In fact, Peterson (1976) has shown that, for noise-driven systems, the naive simulation method in effect ignores most of the data.

Ordinary Least Squares (OLS)

When driving noise W(t) is present (but V(t) is absent), the modeler can in general obtain better estimates of A by reinitializing the system at each data point, and then applying the same squared-residual error function J as in the naive simulation method. The new iteration is called ordinary least-squares (OLS):

$$\hat{X}(t) = A * Z(t-1)$$

$$\hat{Z}(t) = \hat{X}(t)$$

as illustrated in Figure 2. The dots are the data Z(t), the dashed lines are the simulation trajectories between data points, and the vertical bars are the residuals. Whenever the simulation reaches a data-sample time, the system is reinitialized, so that each segment of the simulation begins on a data point.

[1]Also called unconstrained optimization. See Murray (1972) for examples.

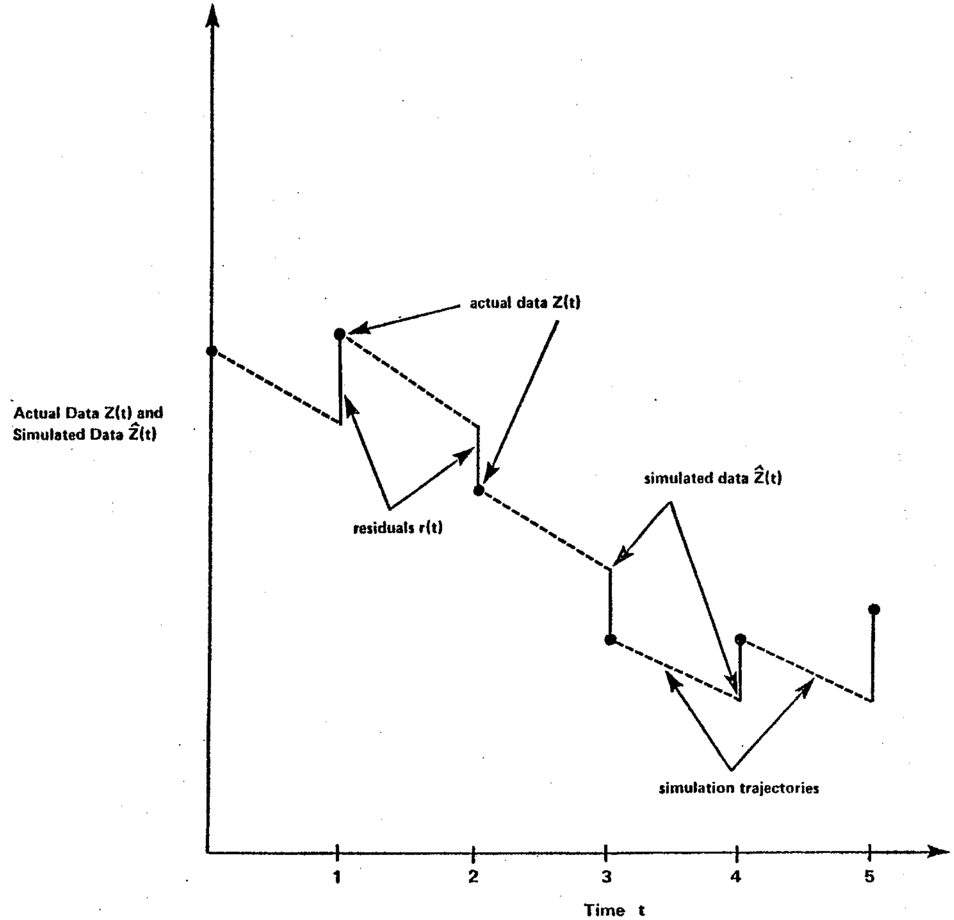

Figure 2: Noise-Driven System and Accurate Data, with Ordinary Least Squares Estimation (OLS).

Wonnacott (1970) has shown that the OLS method tends toward a good estimate of A, so long as V(t) = 0. But the method breaks down when measurement errors can lead to grossly inaccurate estimates of parameters in a typical system dynamics model. The intuitive reason for the failure is as follows. The motivation for reinitializing the model at each data point is to keep the simulation close to the true state of the system, so that any divergence in behavior, as measured by the residuals, would be meaningful. But reinitailiztion in the presence of noisy data is unlikely to keep the simulated state close to the real state. Large residuals may emerge from even a perfect model under OLS.

Full-Information Maximum Likelihood (FIMLOF)

The essence of FIMLOF is to reinitialize the system at each data point, at the value of X(t) where the system is most likely to be given all available data. To do so, the simulation must compute not only the predicted state at each data point, but also the size of the expected error (standard deviation) of the predicted state. Therefore, FIMLOF uses the iteration:

$$\hat{X}(t/t-1) = A * \hat{X}(t-1/t-1)$$

$$\hat{Z}(t/t-1) = \hat{X}(t/t-1),$$

where $\hat{X}(t/t-1)$ is the most likely value of X(t), given all information through time t-1, $\hat{X}(t-1/t-1)$ is the most likely value of X(t-1), given the same information, and $\hat{Z}(t/t-1)$ is the best guess of the next measurement Z(t), given all the previous data Z(0) .... Z(t-1).

The simulation is then updated at $\hat{X}(t/t)$, which is defined as the most likely value of X(t), given all information through time t. This information is embodied in Z(t), in $\hat{X}(t/t-1)$. and in the variances of these two quantities. The variance of Z(t) is simply the variance of the process V(t). The variance of $\hat{X}(t/t-1)$ is automatically derived from the variances

of the processes V(t) and W(t), from the variance of the guess of the initial-conditions, $\hat{X}(0/0)$, and from the structure of the model. The computation of the variances is not detailed here; the equations which compute them constitute an "optimal filter," which is documented by Schweppe (1973) and Kalman (1960), and presented in Appendix A of this paper.

Variances are employed in FIMLOF to avoid the pitfalls observed in OLS and NS. The process is illustrated in Figure 3. As in NS and OLS, the initial conditions of the model are based on the first data point. The system is simulated to the time of the first data point, and the first residual is computed as

$$r(1/0) = Z(1) - \hat{Z}(1/0).$$

So far, the process has been the same as for NS and OLS. The difference lies in how the model is reinitialized at the data point Z(1). Instead of leaving the model state at $\hat{Z}(1/0)$, as in NS, or adjusting the model state completely to match Z(1), as in OLS, FIMLOF reinitializes the model to a compromise point somewhere <u>between</u> $\hat{Z}(1/0)$ and Z(1). The compromise is based on the variances of $\hat{Z}(1/0)$ and Z(1). If the variance of Z(1) is large (noisy measurements), but the variance of $\hat{Z}(1/0)$ is small (little driving noise W(t)), then the reinitialization will be close to $\hat{Z}(1/0)$, as in NS. If, at the other extreme, the variance of Z(1) is small, but the variance of $\hat{Z}(1/0)$ is large (accurate data, but highly uncertain model), then the model will be reinitialized close to Z(1), as in OLS. In general, the reinitialization in FIMLOF will be somewhere between the two extremes, at a point always chosen as the most likely value of the model state, given the available data. Schweppe (1973) has shown that, with this optimal reinitialization (called "optimal filtering" in control theory) and a loss function based on the residuals r(t/t-1), parameter estimates will be those that are most likely, given all the information contained in the data and

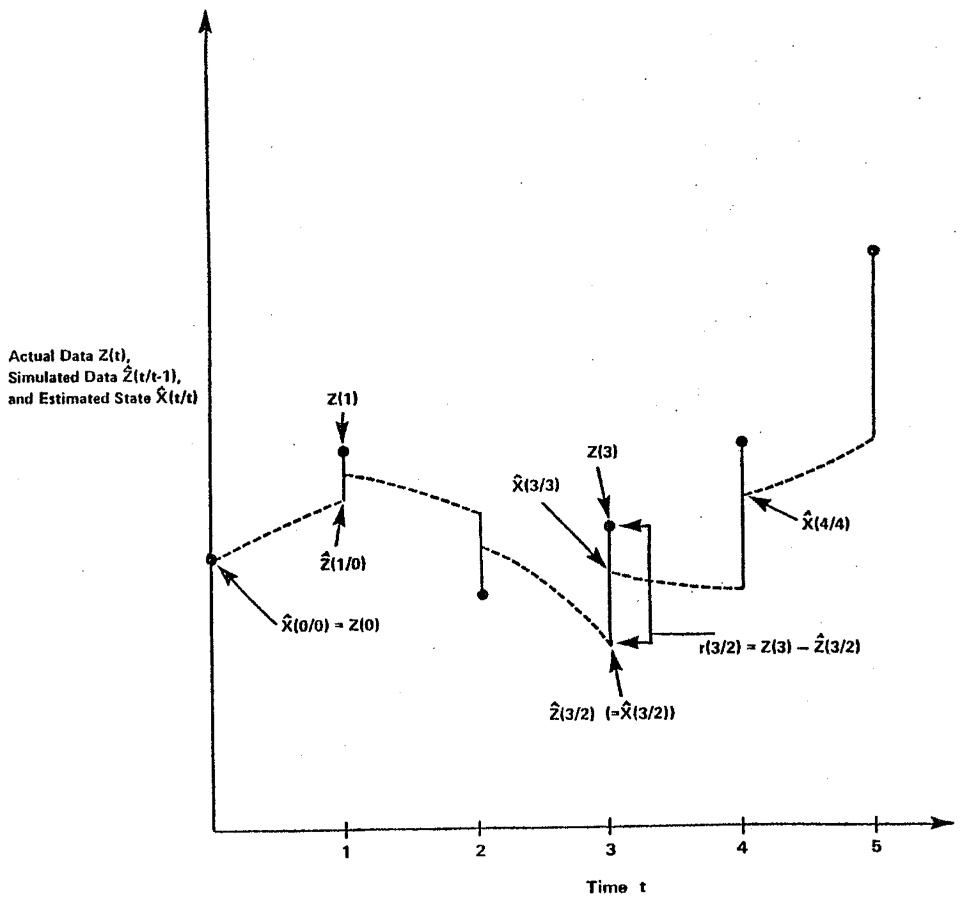

Figure 3: Noise-Driven System and Noisy Data, with Full Information Maximum-Likelihood Estimation (FIMLOF).

model structure. In fact, the FIMLOF loss function (See Appendix A) is the negative logarithm of the likelihood (probability) that the observed data could have been produced by the assumed model (including the guesses of the parameters). By making astute guesses of the parameters (again, automatically, according to a "hill climbing" algorithm, as in Murray 1972), the modeler arrives at the desired "maximum likelihood" estimated.

Feasibility: Comparison with Senge's Results

The likelihood computation is exact only in the case of linear systems with Gaussian noise W(t) and V(t). But the approximate likelihoods computed by nonlinear filters, as in the GPSIE software package, are remarkably accurate. Experiments on nonlinear engineering models by Moore (1972), Mehra (1973), and Arthur (1976) have indicated that nonlinearities are not a serious problem.

FIMLOF (as implemented in GPSIE) has also been tested on system dynamics models of social systems. Figure 4 summarizes the results of one such experiment, chosen to facilitate comparison with Senge (1974). Senge estimated parameters from noisy simulation data generated by a nonlinear, dynamic model of market growth, as published by Forrester (1968). The model consists of nine dynamic difference equations (defining nine "level" variables). Senge simulated the model, using random-number generators to introduce both equation errors and measurement errors in different amounts. In the particular experiment compared below, Senge introduced equation errors ranging from 6% to 60% of the mean of the endogenous variables, and obtained excellent estimates of the 13 system parameters, using OLS and GLS. However, when Senge introduced 10% measurement error, he obtained large errors in the parameter estimates. Figure 4 shows the results Senge obtained, compared with the estimates obtained with the FIMLOF software package GPSIE under the same conditions. The results indicate that FIMLOF techniques, as implemented in GPSIE, may yield

Parameter Name	True Value	GPSIE Estimate	OLS Estimate
SEM	400	392	4349
SED1	-.0281	-.029	-.430
SED2	-.0295	-.0295	.096
SED3	.00228	.00228	-.0074
PCF1	.61782	.615	3.7117
PCF2	-.13244	-.132	-.74891
CEF1	-.0698	-.0693	.03966
CEF2	.12442	.1245	-.14609
CEF3	-.08138	-.0813	.13953
CEF4	.027704	.02704	-.03144
DRAT	1	.97	1.3
SAT	20	19.85	18.5

Figure 4:

Comparison of Estimation Techniques for Ninth-Order Nonlinear System with Errors in Variable.

accurate results, even in the presence of system nonlinearities and measurement error which may cause difficulties with simpler estimation techniques.

In general, for nonlinear models of the system dynamics type, the approximations involved in FIMLOF cause errors no more serious than those due to other approximations in computer simulation, such as numerical-integration error and round-off error.

Although seemingly and indirect, the repeated guess-and-simulate iteration of FIMLOF allows great flexibility. For example, nowhere does the method require that all the state variables be measured, or the availability of data at each time step, or even that the data be distributed at constant intervals. If a data point is missing, for example, the simulation simply continues to the time of the next valid data point. Residuals are computed only at data points, not necessarily at all model time-steps.

Similarly, the mathematics of FIMLOF can deal with unknown (unmeasured) exogenous inputs, cross-sectional data, short data series, non-white noise, and such indirect measurements as yearly summations, averages, and functions of multiple state variables.

The alert reader may object that FIMLOF computations require the variances of measurement errors V(t) and equation errors W(t), which are seldom known with confidence. Variances can be dealt with by including them in the list of parameters to be estimated. In fact, in FIMLOF, parameters inherently may enter the model in any nonlinear or indirect way.

The features of the FIMLOF method are summarized in Figure 5. FIMLOF is compatible with the characteristics commonly found in system dynamics models.

B. Structural Estimation

The preceding observations on parameters and nonlinear systems raise other interesting possibilities. In fact, for nonlinear systems, we must

FIMLOF (as implemented in GPSIE) Operates Under Conditions of:

1. Nonlinearities in model dynamics
2. Nonlinear measurement functions
3. Measurement error (errors in variables)
4. Mixed sampling intervals (can, for example, estimate a weekly model, using monthly and yearly data)
5. Missing data (without sacrificing other data at the same sample time)
6. Models with unmeasured endogenous variables
7. Cross-sectional, time series mixed data
8. Unknown characteristics of equation errors and measurement noise.

Figure 5:

Applicability of FIMLOF and GPSIE.

redefine "structure," as opposed to "parameter." For time-invariant linear systems, the distinction is clear: the system must (by definition) take the form

$$\underline{x}(n) = \underline{A}\underline{x}(n-1) + \underline{B}\underline{u}(n) + \underline{w}(n)$$

where $\underline{A}$ and $\underline{B}$ are constant matrices, $\underline{x}(n)$ is the state vector at time n, $\underline{u}(n)$ is a vector of known inputs, and $\underline{w}(n)$ is a white, normal process of mean 0 and covariance $\underline{Q}$. In this case, the parameters are simply the constant coefficients of the matrices $\underline{A}$,$\underline{B}$, and $\underline{Q}$. A similar definition can be made for systems which are linear in the parameters. For example, in the system

$$\underline{y} = \underline{Xb}$$

where $\underline{y}$ is a vector of outputs, X is a matrix of variables which may be functions of exogenous inputs and of lagged values of $\underline{y}$; $\underline{b}$ is defined as the vector or parameters.

General nonlinear systems, however, will require a more general definition of parameters:

> A <u>parameter</u> in a nonlinear system is a constant exogenous input to the system.

By this definition, a parameter in a nonlinear system may enter the system in any nonlinear fashion. The parameter may be known or unknown, but it is always a constant whose value is not determined by the rest of the system.

This seemingly straightforward definition has the following implication: <u>In nonlinear systems, parameters may take on qualities usually associated with structure</u>. For example, consider the system

$$\underline{x}(t) = \theta\underline{f}(\underline{x}(t-1)) + (1-\theta)\underline{g}(\underline{x}(t-1)) \quad .$$

By any reasonable definition, θ would be considered a parameter in the equation. But θ determines the structure of the system. If θ = 0, then the system has the structure determined by the function $\underline{f}$; if θ = 1, then the system structure is determined by the function g.

Therefore, care must be taken in applying the usual connotations to the terms "structure" and "parameter" when dealing with nonlinear systems. By model building such as the example just illustrated, the modeler may estimate parameters which de facto result in the estimation of structure.

The estimation of structure (as in estimating θ in the above system) may be thought of as a kind of continuous hypothesis test. The maximum-likelihood value of θ may be thought of as selecting the most likely structure from the range of structures implied in the equation. In addition, completely separate models may be compared by computing the likelihood of each with respect to the same data base.

Neither parameters nor structure can be usefully estimated from mere numerical data and thin air. Estimation always entails a choice from a range of alternatives. A well-hypothesized model defines a range of plausible alternatives consistent with the purposes of the study at hand. The ability to estimate structure increases, rather than decreases, the role of experience, logic, and theory in model building.

C. Validity and Consistency Tests

The essence of validity is that the model be consistent with all available information, in the context of the purpose of the model. The automatic consistency tests related to FIMLOF represent a subset of this general notion of validity. The FIMLOF-based tests measure the consistency of the model with available numerical data. However, qualitative knowledge can usually be quantified to some approximation, and FIMLOF can make use of approximate data.

The consistency tests of FIMLOF come in two kinds. First, the likelihood evaluations for each set of parameter guesses provide information, and, second, the optimal filter itself provides several internal consistency

measures. Most of these consistency tests are based on mathematical derivations of proprieties which the likelihood evaluations and filter outputs must have if the model is an accurate representation of the real system which actually produced the data. If the properties are not observed in the FIMLOF output, then there is some inconsistency between model and data.

Use of the Likelihood Surface

The "loss function" computed in FIMLOF is the negative natural logarithm of the likelihood that the data could have been generated by the model. For each different model or set of parameter guesses, the same data will yield, in general, a different "log likelihood." Therefore, the data and model define a surface over the space of all possible parameter values. One property of the log likelihood surface is especially useful in interpreting parameter estimates. At the global maximum of the log likelihood surface, the curvature of the surface is a measure of the quantity of information about the unknown parameters contained in the data. In the extreme, if the likelihood surface is flat (so that all parameter values are equally likely), then there is no information in the data with respect to the model and parameters being estimated. Similarly, if the surface surves sharply downward from the maximum, then the estimates are highly precise, and the data (however noisy) contain a great deal of information about the model. More precisely, the second derivative of the log likelihood surface with respect to parameter A is the variance of the uncertainty in the estimate of A.

Confidence Tests from the Optimal Filter

The optimal filter computes not only the residuals r(t/t-1), but also what the standard deviation of the residuals should be, if the model is correct. The residuals, when normalized by their theoretical standard deviations, can be shown (Schweppe, 1973) to have two properties. First, the normalized residuals should have a constant variance of one; second, the sequence of residuals should be a white process. Since these properties of the residuals process are not employed directly in maximizing the log likelihood, they provide an independent test of model validity. Furthermore, experience shows that residual-based tests are sensitive to small errors in model specification.

The two theoretical properties of the residuals, whiteness and unit variance, provide two kinds of consistency tests:

First, the whiteness may be tested by computing correlation measures of the normalized residuals. Each residual (in the case of multiple-dimensional measurements) should have a correlation coefficient of one with respect to itself, zero serial correlation with respect to lagged values of itself, and zero cross correlation with all other residual processes. (See Appendix A for mathematical definitions.) The correlation test of the residuals not only indicates whether the model is consistent with the data; if the test fails, it may also reveal what is wrong. For example, the pattern of serial correlation coefficients may reveal the first-order time constant associated with a delay missing from the model structure.

Second, the theoretical unit variance of the residual processes provides a test of internal consistency. The log-likelihood function consists of two terms: (1) a sum-of-squared residuals term (SUMSQ), which is analogous to the OLS loss function; and (2) a term which is independent of the size of the residuals. The expected magnitude and standard deviation of the SUMSQ term at the true parameter values can be predicted ahead of time. If the actual SUMSQ differs from the predictdd range, then either the model is inconsistent with the data, or the global maximum of the log likelihood function has not yet been found.

D. Detection of Bad Data

The preceding discussion explained how noisy, approximate data can be used to help estimate and evaluate system structures. But most collections of data contain some points so much in error as to be best deleted. Such "bad data" points may arise from typographical errors, improper accounting procedures, and other gross malfunctions of data collection. Instead of containing useful information about the system, the bad data points serve only to mislead.

The FIMLOF methodology includes convenient techniques for detecting and isolating bad data. The basic idea behind the techniques is to look for residuals which are clearly too large. The obvious difficulty is to define "too big." The answer is provided by the optimal filter, which computes the expected standard deviation of each residual, under the assumption that the model and data are consistent. As a matter of practicality, a bad data point will create large residuals even if the model is still somewhat approximate. Therefore, even if the model has been estimated using data which contains bad data points, a residual more than four or five standard deviations away from its expected value can be taken to indicate a bad data point.

A complication arises when, as is usual, more than one varible is measured at the same time. In such a case, a large residual still reveals the presence of a bad data point, but does not necessarily indicate which of the several measurements is at fault. However, the optimal filter of the FIMLOF method provides the information required to decide which component is in error. From the variances and covariances by the filter, the evaluation can compute "normalized updated" residuals (see Peterson, 1975) which pinpoint the bad data points in both time and space.

Although bad data can often be spotted by visually scanning graphed data, the FIMLOF techniques are useful for two reasons. First, the techniques can be completely automated, allowing the thorough checking of large data files. Second, bad data is not always readily apparent, even when the data is presented in graphical form. The "wrongness" of a data point is often seen only in the context of the dynamic structure behind the data, as well as measurements of other variables. Preliminary experience (Peterson, 1975) indicates that bad data may be uncovered quickly using the FIMLOF methods, even in data files which have been manually inspected for errors.

E. Estimation of the System State

A useful by-product of the optimal filter is the computation of the most likely state of the system at a given time. The estimated state can be used to initialize the system for forecasting. The estimated state may also yield insight, or aid in decision-making. For example, a decision-maker would like to know which inputs are limiting a production process, or in which region of a nonlinear function the system is operating. The filter provides not only an estimate of the true state of the system, but also gives confidence bounds, by way of the variances and covariances of the state-variable estimates. Such computations can also be validly continued into the future, as discussed below.

F. Confidence Bounds for Forecasts

Simulation models are often used to predict the future evolution of a system. Usually, the model is intialized at some approximation of present conditions, and the (deterministic) simulated trajectory is taken as a best guess of the future. Such a forecast may be inaccurate for three reasons. First, of course, the model structure may be inaccurate. But even if the model is "perfect," two sources of uncertainty may bring abcut inaccurate forecasting. To the extent that the real system is driven by uncertain processes (events modeled as random), the future evolution of the real system is likely to drift away from any computed trajectory, thereby expand-

ing the frequency and magnitude of errors in the forecast. Finally, there is the difficulty of deciding what initial conditions to use for the forecasting simulation. Since many state variables in a model may be unmeasured, the most likely present state of the system may not be obvious from a casual inspection of the data.

Forrester (1961, Appendix K) discusses these ideas qualitatively, illustrating the interaction among model accuracy, knowledge of the present state of the real system, and forecast accuracy. The FIMLOF techniques allow these factors to be assessed quantitatively. Forecasts computed by the filter include not only the "expected" future trajectory of the system, but also confidence bounds on the trajectory. The "initial conditions" of confidence bounds are derived from the computed variance of the previously explained present state estimate. The filter than computes the <u>a priori</u> variance of future state estimates as a function of the initial variance, the model structure, and the variances of the random system inputs.

The confidence bounds on the forecast clearly show to what extent, and how far into the future, a given model may be expected to yield accurate predictions. In many social systems, the confidence bounds diverge rapidly as the simulation extends farther ahead in time. The timing and severity of the divergence will depend on the state of the model, its structural accuracy, the nature of the structure, the accuracy of the initial conditions, and the severity of random inputs to the system.

II. Implementation: GPSIE

The various FIMLOF-based techniques discussed here have been implemented in a computer program called the General Purpose System Identifier and Evaluator (GPSIE). GPSIE is a large precompiled program which links with a user-written program describing the particular model of interest. The resulting package can be used to load data, compute likelihoods via optimal filtering, search for maxima in the likelihood function, compute the validity statistics discussed here, and plot the results. GPSIE embodies a large number of options for dealing with special cases and for maintaining efficient computation in various circumstances. (For more details on GPSIE, see Peterson, 1974 and 1975).

An obvious concern with iterative methods and statistical analyses is the threat of high computation costs. GPSIE, for example, imposes no inherent limits on the model size or data base, but requirements of computer time or storage may obviously become extravagant for large systems. For example, the cost of parameter estimation and validity tests for a tenth-order system with 1000 data points would typically fall between $200 and $300 on a large time-sharing system.

The computational costs of some FIMLOF computations vary with the cube of the system dimensions. Therefore, it often pays to break the model into sectors for individual analysis. The sectors may then be recombined for final validity testing, requiring but a single filtering computation with the entire system intact.

III. Conclusions

There are two implications of the FIMLOF techniques for the field of system dynamics. First, the techniques may increase the efficiency and quality of system dynamics modeling, by complementing the manual simulate-analyze-correct techniques. Parameter estimation, consistency tests, and confidence bounds may efficiently indicate areas of sensitivity or inconsistency which might otherwise be found with difficulty. A failed consistency test or unreasonable estimated parameter, furthermore, may not only reveal the presence of trouble, but also may suggest an appropriate

remedy in model structure.

Second, the FIMLOF techniques may extend the practice of system dynamics into additional fields and disciplines. As techniques such as FIMLOF become more widely understood and available, system-dynamics type models may be employed more often for data-related activities, including forecasting, data validation, and performance monitoring.

Appendix A:
Mathematical Definitions

This appendix is for the benefit of the mathematically inclined reader. It summarizes the equations of the FIMLOF techniques discussed qualitatively in the text. Section A.1 defines the notation of the model, its relationship to data, its uncertainty, and the linearization of the model. Section A.2 gives the equations of the optimal filter and the accompanying likelihood calculation. Section A.3 gives the equations for two of the confidence tests discussed in the text. Finally, Section A.4 summarizes the techniques of bad-data detection. For more details on bad-data detection, See Peterson (1975).

A.1 System Dynamics and Linearization

State Equations: $\underline{x}(n) = \underline{f}[\underline{x}(n-1), \underline{u}(n), \underline{w}(n), n]$

Measurement Equations: $\underline{z}(n) = \underline{h}[\underline{x}(n), \underline{v}(n), n]$

Index of Data Samples: $n = 1, 2, \ldots N$

Initial Conditions: $\underline{x}(0) = N[\underline{x}_0, \underline{\Psi}]^*$

Equation Errors (Driving Noise): $\underline{w}(n) = N[\underline{0}, \underline{Q}(n)]^*$

Measurement Errors: $\underline{v}(n) = N[\underline{0}, \underline{R}(n)]^*$

* N[m,c] denotes a normal, white process with mean m and covariance matric c.

Linearization about Estimated State:

$$\tilde{\underline{F}}(n) = \left.\frac{\partial \underline{f}}{\partial \underline{x}}\right|_{\underline{x} = \hat{\underline{x}}(n-1|n-1)}$$

$$\tilde{\underline{H}}(n) = \left.\frac{\partial \underline{h}}{\partial \underline{x}}\right|_{\underline{x} = \hat{\underline{x}}(n|n-1)}$$

$$\tilde{\underline{Q}}(n) = \left.\left(\frac{\partial \underline{f}}{\partial \underline{w}}\right) \underline{Q}(n) \left(\frac{\partial \underline{f}}{\partial \underline{w}}\right)'\right|_{\underline{w} = \underline{0}}$$

$$\tilde{\underline{R}}(n) = \left.\left(\frac{\partial \underline{h}}{\partial \underline{v}}\right) \underline{R}(n) \left(\frac{\partial \underline{h}}{\partial \underline{v}}\right)'\right|_{\underline{v} = \underline{0}}$$

A.2 Filter Equations

Predicted State: $\underline{\hat{x}}(n|n-1) = \underline{f}[\underline{\hat{x}}(n-1|n-1), \underline{u}(n), \underline{\theta}, n]$

Predicted Measurement: $\underline{\hat{z}}(n|n-1) = \underline{h}[\underline{\hat{x}}(n|n-1), \underline{\theta}, n]$

Residuals: $\underline{\delta}_z(n|n-1) = \underline{z}(n) - \underline{\hat{z}}(n|n-1)$

Predicted State Covariance: $\underline{\Sigma}_x(n|n-1) = \underline{\tilde{F}}(n)\,\underline{\Sigma}_x(n-1|n-1)\,\underline{\tilde{F}}'(n) + \underline{\tilde{Q}}(n)$

Predicted Measurement Covariance: $\underline{\Sigma}_z(n|n-1) = \underline{\tilde{H}}(n)\,\underline{\Sigma}_x(n|n-1)\,\underline{\tilde{H}}'(n) + \underline{\tilde{R}}(n)$

Normalized Predicted Measurement Residuals: $\underline{\tilde{\delta}}_z(n|n-1) = \sqrt{\underline{\Sigma}_z(n|n-1)}^{\,-1}\,\underline{\delta}_z(n|n-1)$

Updated State Covariance: $\underline{\Sigma}_x(n|n) = \left[\underline{\Sigma}_x^{-1}(n|n-1) + \underline{\tilde{H}}'(n)\,\underline{\tilde{R}}^{-1}(n)\,\underline{\tilde{H}}(n)\right]^{-1}$

Filter Gain: $\underline{K}(n) = \underline{\Sigma}_x(n|n-1)\,\underline{\tilde{H}}'(n)\,\underline{\Sigma}_z^{-1}(n|n-1)$

Updated State Estimate: $\underline{\hat{x}}(n|n) = \underline{\hat{x}}(n|n-1) + \underline{K}(n)\,\underline{\delta}_z(n|n-1)$

Log Likelihood: $\xi(n) = \xi(n-1) - \frac{1}{2}\,\underline{\tilde{\delta}}_z'(n|n-1)\,\underline{\tilde{\delta}}_z(n|n-1) - \frac{1}{2}\ln\left\{\left|\underline{\Sigma}_z(n|n-1)\right|\right\}$

Initial Conditions: $\underline{\hat{x}}(0|0) = \underline{X}_0, \quad \underline{\Sigma}_x(0|0) = \underline{\Psi}, \quad \xi(0) = 0$

A.3 Confidence Tests

Residual Correlation Matrices:

$$\underline{R}(j) = \frac{1}{N-j}\sum_{n=1}^{N-j} \underline{\tilde{\delta}}_z(n|n-1)\,\underline{\tilde{\delta}}_z'(n+j|n+j-1)$$

If model and data are consistent, then $\underline{R}(0) \approx \underline{I}$ (identity matrix), and $\underline{R}(j) \approx \underline{0}$, $j \neq 0$.

SUMSQ Statistic:

$$\text{SUMSQ} = \sum_{n=1}^{N} \underline{\tilde{\delta}}_z'(n|n-1)\,\underline{\tilde{\delta}}_z(n|n-1)$$

The expected value of SUMSQ is equal to the total number of scalar data points, minus the number of unknown parameters. The standard deviation of SUMSQ is equal to the square root of twice its expected value.

A.4 Bad-Data Detection

Normalized Updated Measurement Residuals (NUMR):

$$\underline{r}_z(n|n) = \left[diag \left\{ \underline{\Sigma}_z^{-1}(n|n-1) \right\} \right]^{-1/2} \underline{\tilde{R}}^{-1}(n) \left[\underline{z}(n) - \hat{\underline{z}}(n|n) \right]$$

Normalized Updated State Residuals (NUSR):

$$\underline{r}_x(n|n) = \left[diag \left\{ \underline{\tilde{H}}'(n) \underline{\Sigma}_z^{-1}(n|n-1) \underline{\tilde{H}}(n) \right\} \right]^{-1/2} \underline{\Sigma}_x^{-1}(n|n-1) \left[\hat{\underline{x}}(n|n-1) - \hat{\underline{x}}(n|n) \right]$$

The NUMR and NUSR processes interact with each other, and must be considered togather. They have two useful properties:

1. First, both NUMR and NUSR have constant unit variance. That is, each component of NUMR and NUSR has a constant standard deviation of one, under all circumstances, as long as the model is valid and the data conforms to the model. Therefore, one or more components of NUMR and NUSR with absolute value greater than 3 or 4 is a reliable sign of a bad data point somewhere at the corresponding sample time.
2. In addition, Peterson (1975) has shown that, at a sample time involving a bad data point, the component of NUMR or NUSR with the maximum absolute value corresponds to the component of $\underline{z}(n)$ or $\underline{x}(n)$ which is in error. For example, a typographical error in the first component of $\underline{z}(3)$ may, depending on the model structure, cause several components of both $\underline{r}_z(n/n)$ and $\underline{r}_x(n/n)$ to exceed the acceptable limit (for example, 4). But the component with the largest absolute value identifies the specific component of $\underline{z}(3)$ or $\underline{x}(3)$ which contains the typographical error. (In the case of $\underline{x}(n)$, the error might be in an exogenous input to the equation determining the component of $\underline{x}(n)$). These properties

of NUMR and NUSR permit the creation of computer programs which automatically identify and delete bad data points, and the efficient screening of data sets for questionable entries. As shown by Peterson (1975), they can also be helpful in model validation and model improvement.

BIBLIOGRAPHY

Arthur, William B., 1976, private communication.

Forrester, Jay W., 1961, Industrial Dynamics (M.I.T. Press, Cambridge, Mass.).

Forrester, Jay W., 1968, "Market Growth As Influenced By Capital Investment", Industrial Management Review, Vol. 9, No. 2, pp 83-105, Winter.

Hamilton, J. R., et al., 1969, System Simulation For Regional Analysis (M.I.T. Press, Cambridge, Mass.)

Kalman, R. E., 1960, "A New Approach To Linear Filtering and Prediction Problems", Journal of Basic Engineering, Series D., Vo. 82., No. 3, pp 35-45, March.

Mehra, R. K., and Tyler, J. S., 1973, "Case Studies in Aircraft Parameter Identification", from Identification and System Parameter Estimation ed. P. Eykhoff, American Elsevier Publishing Co., New York).

Moore, R. and Schweppe, F. C., 1972, "Adaptive Control for Nuclear Power Plant Loan Changes", Proceedings, Fifth World Congress of the International Federation of Automatic Control, Paris.

Murray, W., 1972, Numerical Methods For Unconstrained Optimization (Academic Press, New York).

Peterson, D. W., and Schweppe, F. C., 1974, "Code For a General Purpose System Identifier and Evaluator: GPSIE", IEEE Transactions on Automatic Control, Vol. Ac-19, No. 6, pp 852-854, December.

Peterson, D. W., 1975, "Hypothesis, Estimation, and Validation of Dynamic Social Models" (unpublished Ph.D. thesis, Department of Electrical Engineering, M.I.T., Cambridge, Mass.), June.

Peterson, D. W., 1976, "Parameter Estimation for System Dynamics Models", Proceedings, Summer Computer Simulation Conference, Washington, D. C.

Runge, D., 1976, "Labor-Market Dynamics: An Analysis of Mobility and Wages" (Ph.D. thesis, Alfred P. Sloan School of Management, M.I.T., Cambridge, Mass.).

Schweepe, F. C., 1973, Uncertain Dynamic Systems (Prentice-Hall, Englewood Cliffs, N. J.).

Senge, P. M., 1974, "Evaluating the Validity of Econometric Methods for Estimation and Testing of Dynamic Systems", System Dynamics Group Memo D-1944-2, Alfred P. Sloan School of Management, M.I.T., February 12.

Senge, P. M., 1974, "An Experimental Evaluation of Generalized Least Squares Estimation", System Dynamics Group Memo D-1944-6, Alfred P. Sloan School of Management, M.I.T., November 14.

Wonnacott, P. J., and Wonnacott, T. H., 1970, Econometrics (John Wiley and Sons).