System and method providing surveillance of an asset such as a process and/or apparatus by providing training and surveillance procedures that numerically fit a probability density function to an observed residual error signal distribution that is correlative to normal asset operation and then utilizes the fitted probability density function in a dynamic statistical hypothesis test for providing improved asset surveillance.
OTHER PUBLICATIONS


* cited by examiner
Figure 1

Asset and/or Apparatus 12

Data Acquisition, Signal Processing, and Digitization Means 20

Communication Means 26

Operator Console, Alarm and Data Display 80

Asset Control Means 82

Computer 22
- Training Procedure 30
- Surveillance Procedure 60
  - Parameter Estimation Procedure 64
  - Fault Detection Procedure 66
    - Compute Data Residuals 76
    - Perform ASP Test(s) 78

Memory Means 24
- Training Data 34
- Process Model 50
  - Parameter Estimation Model(s) 52
  - Fault Detection Model(s) 54

Figure 1
Figure 3

Training Procedure

1. Data Acquisition and Digitization Means
2. Store Historical Operating Data
3. Data
4. Training Data
5. Calibrate Parameter Estimator(s)
6. Calibrate Fault Detector(s)
7. Process Model
8. Parameter Estimation Model(s)
9. Fault Detection Model(s)

Asset
Figure 4

1. Specify Parameter Estimator Settings
2. Model Nominal Signal Behavior
3. Training Data
4. Parameter Estimation Model(s)
5. Process Model
6. Estimate Process Parameters
7. Compute Training Data Residuals
8. Specify Fault Detector Settings
9. Compute Nominal Residual PDF
10. Fault Detection Model(s)

Designer
Figure 7

Sensor 1
Acquire Signal 1
Value = S1

Compute Residual = S1 - E1

Perform ASP Test(s) for S1

Estimate S1
E1 = f(S2)

Sensor 2
Acquire Signal 2
Value = S2
Acquire Signal 1 Value = S1

\[ E_1 = f(S_1, \ldots, S_n) \]

Compute Residual = S1 - E1

Perform ASP Test(s) for S1

Acquire Signal n Value = Sn

\[ E_n = f(S_1, \ldots, S_n) \]

Compute Residual = Sn - En

Perform ASP Test(s) for Sn

Figure 8
Acquire and Store Asset Training Data

Calibrate Parameter Estimator(s) (Procedure (36)) Using Asset Training Data And Training Procedure (e.g., MSET Training Procedure) For Producing Parameter Estimation Model(s)

Calibrate Fault Detector(s) (Procedure (38)) Including Fitting A General PDF To Computed Training Data Residual Distribution, For Example By Fitting A Standard Gaussian PDF To The Computed Training Data Residuals And Then Adding Successive Higher Order Terms Of A Remainder Function To The Standard Gaussian PDF Until An Adequate Fit Between The General PDF And Computed Training Data Residual Distribution Is Produced

Acquire and Digitize Current Asset Data

Estimate Process Parameters (Procedure (64)) Using Acquired and Digitized Current Asset Data And Parameter Estimation Model(s)

Perform Fault Detection (Procedure (66)) By Computing Data Residuals (Procedure (76)) And Using The Fitted General PDF In Performing The Adaptive Sequential Probability (ASP) Test(s) (Procedure (78)) In Accordance With The Present Invention

Fault Found? No → Surveillance Complete? No

Yes → Alarm and/or Control Action

Yes → End Surveillance

Figure 9
Figure 10

- $H_0$: Mean = 0, Variance = 1
- $H_1$: Mean = 2, Variance = 1
- $H_2$: Mean = -2, Variance = 1
Figure 11

- H0: Mean = 0, Variance = 1
- H3: Mean = 0, Variance = 2
- H4: Mean = 0, Variance = 0.5
Figure 12
Figure 13
Figure 14
Figure 15
Figure 16
<table>
<thead>
<tr>
<th>Sensor Number</th>
<th>Gaussian</th>
<th>One-Term Series Expansion</th>
<th>Two-Term Series Expansion</th>
<th>Three-Term Series Expansion</th>
<th>Four-Term Series Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.414</td>
<td>0.409</td>
<td>0.122</td>
<td>0.151</td>
<td>2.06</td>
</tr>
<tr>
<td>2</td>
<td>0.389</td>
<td>0.396</td>
<td>0.147</td>
<td>0.223</td>
<td>2.13</td>
</tr>
<tr>
<td>3</td>
<td>0.233</td>
<td>0.234</td>
<td>0.0605</td>
<td>0.0651</td>
<td>0.171</td>
</tr>
<tr>
<td>4</td>
<td>0.339</td>
<td>0.320</td>
<td>0.0964</td>
<td>0.146</td>
<td>0.256</td>
</tr>
<tr>
<td>5</td>
<td>0.156</td>
<td>0.167</td>
<td>0.121</td>
<td>0.0985</td>
<td>0.0983</td>
</tr>
<tr>
<td>6</td>
<td>0.180</td>
<td>0.134</td>
<td>0.107</td>
<td>0.326</td>
<td>0.656</td>
</tr>
</tbody>
</table>

Figure 17
<table>
<thead>
<tr>
<th>Sensor Number</th>
<th>Gaussian</th>
<th>One-Term Series Expansion</th>
<th>Two-Term Series Expansion</th>
<th>Three-Term Series Expansion</th>
<th>Four-Term Series Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.414</td>
<td>0.176</td>
<td>0.0826</td>
<td>0.0755</td>
<td>0.0899</td>
</tr>
<tr>
<td>2</td>
<td>0.389</td>
<td>0.174</td>
<td>0.0811</td>
<td>0.0816</td>
<td>0.0659</td>
</tr>
<tr>
<td>3</td>
<td>0.233</td>
<td>0.122</td>
<td>0.0522</td>
<td>0.0573</td>
<td>0.0522</td>
</tr>
<tr>
<td>4</td>
<td>0.339</td>
<td>0.166</td>
<td>0.0706</td>
<td>0.0726</td>
<td>0.0713</td>
</tr>
<tr>
<td>5</td>
<td>0.156</td>
<td>0.0699</td>
<td>0.0394</td>
<td>0.0352</td>
<td>0.0373</td>
</tr>
<tr>
<td>6</td>
<td>0.180</td>
<td>0.0710</td>
<td>0.0346</td>
<td>0.0352</td>
<td>0.0369</td>
</tr>
</tbody>
</table>

Figure 18
Figure 19
Theoretical Upper Bound

Figure 20
SURVEILLANCE SYSTEM AND METHOD HAVING AN ADAPTIVE SEQUENTIAL PROBABILITY FAULT DETECTION TEST

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation patent application of U.S. application Ser. No. 10/095,835, filed Mar. 8, 2002, now U.S. Pat. No. 6,892,163.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

The invention described herein was made in the performance of work under NASA Small Business Technology Transfer Research (STTR) Contract NAS8-98027, and is subject to the provisions of Public Law 96-517 (35 USC 202) and the Code of Federal Regulations 48 CFR 52.227-11 as modified by 48 CFR 18.227-11, in which the contractor has elected to retain title.

The United States Government has rights in this invention pursuant to Contract No. W-31-109-ENG-38 between the United States Government and the University of Chicago representing Argonne National Laboratory.

FIELD OF THE INVENTION

The instant invention relates generally to a system and method for process fault detection using a statistically based decision test and, in particular, to a system and method for performing high sensitivity surveillance of an asset such as a process and/or apparatus wherein the performance of the surveillance is performed using an adaptive sequential probability (ASP) fault detection test comprised of a probability density function model empirically derived from a numerical analysis of asset operating data.

BACKGROUND OF THE INVENTION

Conventional process surveillance schemes are sensitive only to gross changes in the mean value of a process signal or to large steps or spikes that exceed some threshold limit value. These conventional methods suffer from either a large number of false alarms (if thresholds are set too close to normal operating levels) or from a large number of missed (or delayed) alarms (if the thresholds are set too expensively). Moreover, most conventional methods cannot perceive the onset of a process disturbance or sensor signal error that gives rise to a signal below the threshold level or an alarm condition. Most conventional methods also do not account for the relationship between measurements made by one sensor relative to another redundant sensor or between measurements made by one sensor relative to predicted values for the sensor.

Recently, improved methods for process surveillance have developed from the application of certain aspects of artificial intelligence technology. Specifically, parameter estimation methods have been developed using either statistical, mathematical or neural network techniques to learn a model of the normal patterns present in a system of process signals. After learning these patterns, the learned model is used as a parameter estimator to create one or more predicted (virtual) signals given a new observation of the actual process signals. Further, high sensitivity surveillance methods have been developed for detecting process and signal faults by analysis of a mathematical comparison between an actual process signal and its virtual signal counterpart. In particular, such a mathematical comparison is most often performed on a residual error signal computed as, for example, the difference between an actual process signal and its virtual signal counterpart.

Parameter estimation based surveillance schemes have been shown to provide improved surveillance relative to conventional schemes for a wide variety of assets including industrial, utility, business, medical, transportation, financial, and biological systems. However, parameter estimation based surveillance schemes have in general shown limited success when applied to complex processes. Applicants recognize and believe that this is because the parameter estimation model for a complex process will, in general, produce residual error signals having a non-Gaussian probability density function. Moreover, a review of the known prior-art discloses that virtually all such surveillance systems developed to date utilize or assume a Gaussian model of the residual error signal probability density function for fault detection. Hence, a significant shortcoming of the known prior-art is that, inter alia, parameter estimation based surveillance schemes will produce numerous false alarms due to the modeling error introduced by the assumption of a Gaussian residual error signal probability density function. The implication for parameter estimation based surveillance schemes is that the fault detection sensitivity must be significantly reduced to prevent false alarms thereby limiting the utility of the method for process surveillance. An alternative for statistically derived fault detection models is to mathematically pre-process the residual error signals to remove non-Gaussian elements prior to using the residual error signals in the fault detection model; however this approach requires an excess of additional processing and also limits the sensitivity of the surveillance method. Therefore, the implication of assuming a Gaussian residual error signal probability density function for a parameter estimation based surveillance scheme is simply that the system becomes less accurate thereby degrading the sensitivity and utility of the surveillance method.

Many attempts to apply statistical fault detection techniques to surveillance of assets such as industrial, utility, business, medical, transportation, financial, and biological processes have met with poor results in part because the fault detection models used or assumed a Gaussian residual error signal probability density function.

In one specific example, a multivariate state estimation technique based surveillance system for the Space Shuttle Main Engine’s telemetry data was found to produce numerous false alarms when a Gaussian residual error fault detection model was used for surveillance. In this case, the surveillance system’s fault detection threshold parameters were desensitized to reduce the false alarm rate; however, the missed alarm rate then became too high for practical use in the telemetry data monitoring application.

Moreover, current fault detection techniques for surveillance of assets such as industrial, utility, business, medical, transportation, financial, and biological processes either fail to recognize the surveillance performance limitations that occur when a Gaussian residual error model is used or, recognizing such limitations, attempt to artificially conform the observed residual error data to fit the Gaussian model. This may be attributed, in part, to the relative immaturity of the field of artificial intelligence and computer-assisted surveillance with regard to real-world process control applications. Additionally, a general failure to recognize the specific limitations that a Gaussian residual error model
imposes on fault decision accuracy for computer-assisted surveillance is punctuated by an apparent lack of known prior art teachings that address potential methods to overcome this limitation. In general, the known prior-art teaches computer-assisted surveillance solutions that either ignore the limitations of the Gaussian model for reasons of mathematical convenience or attempt to conform the actual residual error data to the artificial Gaussian model, for example, by using frequency domain filtering and signal whitening techniques.

For the foregoing reasons, there is a need for a surveillance system and method that overcomes the significant shortcomings of the known prior-art as delineated herein-above.

**BRIEF SUMMARY OF THE INVENTION**

The instant invention is distinguished over the known prior art in a multiplicity of ways. For one thing, an embodiment of the instant invention provides a surveillance system and method having a fault detection model of unconstrained probability density function form and having a procedure suitable for overcoming a performance limiting trade-off between probability density function modeling complexity and decision accuracy that has been unrecognized by the known prior-art. Specifically, an embodiment of the instant invention can employ any one of a plurality of residual error probability density function model forms, including but not limited to a Gaussian form, thereby allowing a surveillance system to utilize the model form best suited for optimizing surveillance system performance.

Moreover, an embodiment of the instant invention provides a surveillance system and method that uses a computer-assisted learning procedure to automatically derive the most suitable form of the residual error probability density function model by observation and analysis of a time sequence of process signal data and by a combination of a plurality of techniques. This ability enables surveillance to be performed by the instant invention with lower false alarm rates and lower missed alarm rates than can be achieved by the known prior-art systems and methods.

Thus, an embodiment of the instant invention provides a surveillance system and method that performs its intended function much more effectively by enabling higher decision accuracy. Further, an embodiment of the instant invention is suitable for use with a plurality of parameter estimation methods thereby providing a capability to improve the decision performance of a wide variety of surveillance systems.

In one embodiment, the instant invention provides a surveillance system and method that creates and uses, for the purpose of process surveillance, a multivariate state estimation technique parameter estimation method in combination with a statistical hypothesis test fault detection method having a probability density function model empirically derived from a numerical analysis of asset operating data.

Particularly, and in one embodiment, the instant invention provides a surveillance system and method for providing surveillance of an asset such as a process and/or apparatus by providing training and surveillance procedures.

In accordance with one embodiment, the instant invention provides a training procedure comprised of a calibrate parameter estimator procedure and a calibrate fault detector procedure.

The calibrate parameter estimator procedure creates a parameter estimation model(s) and trains a parameter estimation model by utilizing training data correlative to expected asset operation and, for example, utilizing a multivariate state estimation technique (MSET) procedure. The calibrate parameter estimator procedure further stores this model as a new parameter estimation model in a process model. The process model may contain one or more parameter estimation models depending upon design requirements. Additionally, each parameter estimation model contained in the process model may be created to implement any one of a plurality of parameter estimation techniques.

The calibrate fault detector procedure makes use of the parameter estimation models to provide estimated values for at least one signal parameter contained in the training data. Generally, the calibrate fault detector procedure will create a separate and distinct fault detection model for each sensor or data signal associated with the asset being monitored for the presence of fault conditions during the surveillance procedure.

The calibrate fault detector procedure includes, for example, a method of fitting a standard Gaussian probability density function (PDF) to a training data residual distribution (computed as a function of estimated process parameters and the training data) and then adding successive higher order terms of a remainder function to the standard Gaussian PDF for the purpose of defining a general PDF that better fits the computed training data residual distribution. Other techniques for fitting a general PDF to the training data are similarly feasible and useful in accordance with the instant invention and include, for example, a technique for fitting a polynomial function to the training data.

The training procedure is completed when all training data has been used to calibrate the process model. At this point, the process model preferably includes parameter estimation models and fault detection models for each sensor or data signal associated with the asset being monitored for the presence of fault conditions during the surveillance procedure. The process model is thereafter used for performing surveillance of the asset.

The surveillance procedure is performed using an adaptive sequential probability (ASP) fault detection test comprised of the general probability density function model empirically derived from a numerical analysis of the asset training data.

The surveillance procedure acquires and digitizes current asset data and then estimates process parameters as a function of the acquired digitized current asset data and the parameter estimation model(s) obtained from the calibrate parameter estimator procedure. Then, fault detection is determined by first computing data residuals as a function of the estimated process parameters and the acquired digitized current asset data and then performing the ASP test(s) as a function of the fault detection models and thus, as a function of the fitted general PDF obtained in the calibrate fault detector procedure. Each ASP test returns one of three possible states: a null state which rejects the probability that a null hypothesis is true and excepts an alternative hypothesis correlative to unexpected operation of the asset; a null state which accepts the probability that a null hypothesis is true and excepts the null hypothesis correlative to expected operation of the asset; and an in-between state which excepts neither the null hypothesis nor the alternative hypothesis as being true and requires more data to reach a conclusion.

The results of the fault detection procedure are then analyzed according to the instant invention such that an alarm and/or a control action are taken when the analysis determines that the results indicate unexpected operation of the asset.
Hence, the instant invention is distinguished over the known prior-art by providing a surveillance system and method for performing surveillance of an asset by acquiring residual data correlative to expected asset operation; fitting a mathematical model to the acquired residual data; storing the mathematical model in a memory means; collecting a current set of observed signal data values from the asset; using the mathematical model in a sequential hypotheses test to determine if the current set of observed signal data values indicate unexpected operation of the asset indicative a fault condition; and outputting a signal correlative to each detected fault condition for providing asset surveillance.

Accordingly, the instant invention provides a new, novel and useful surveillance system and method having an adaptive sequential probability fault detection test.

In one embodiment of the instant invention the surveillance system and method includes an unconstrained form of a residual error probability density function model used in said surveillance system’s fault detection method.

In one embodiment of the instant invention the surveillance system and method can perform high sensitivity surveillance for a wide variety of industries including industrial, utility, business, medical, transportation, financial and biological processes and apparatuses wherein such process and/or apparatus asset preferably has at least one pair of redundant actual and/or virtual signals.

In one embodiment of the instant invention the surveillance system and method includes a statistical hypothesis test surveillance decision procedure that uses a fault detection model comprised of a probability density function model of a residual error signal that is of an unconstrained form.

In one embodiment of the instant invention the surveillance system and method creates an improved fault detection model for a process surveillance scheme using recorded operating data for an asset to train a fault detection model.

In one embodiment of the instant invention the surveillance system and method provides an improved system and method for surveillance of on-line, real-time signals, or off-line accumulated signal data.

In one embodiment of the instant invention the surveillance system and method provides an improved system and method for surveillance of signal sources and detecting a fault or error state of the signal sources enabling responsive action thereto.

In one embodiment of the instant invention the surveillance system and method provides an improved system and method for surveillance of signal sources and detecting a fault or error state of the asset processes and apparatuses enabling responsive action thereto.

In one embodiment of the instant invention the surveillance system and method provides an improved system and method for ultra-sensitive detection of a fault or error state of signal sources and/or asset processes and apparatuses wherein the parameter estimation technique used for the generation of at least one virtual signal parameter is a neural network having any one of a plurality of structures, training procedures, and operating procedures.

In one embodiment of the instant invention the surveillance system and method provides an improved system and method for ultra-sensitive detection of a fault or error state of signal sources and/or asset processes and apparatuses wherein the parameter estimation technique used for the generation of at least one virtual signal parameter is a mathematical process model having any one of a plurality of structures, training procedures, and operating procedures.

In one embodiment of the instant invention the surveillance system and method provides an improved system and method for ultra-sensitive detection of a fault or error state of signal sources and/or asset processes and apparatuses wherein the parameter estimation technique used for the generation of at least one virtual signal parameter is an autoregressive moving average (ARMA) model having any one of a plurality of structures, training procedures, and operating procedures.

In one embodiment of the instant invention the surveillance system and method provides an improved system and method for ultra-sensitive detection of a fault or error state of signal sources and/or asset processes and apparatuses wherein the parameter estimation technique used for the generation of at least one virtual signal parameter is an autoregressive moving average (ARMA) model having any one of a plurality of structures, training procedures, and operating procedures.

In one embodiment of the instant invention the surveillance system and method provides a novel system and method to classify the state of a residual error signal wherein said classification is made to distinguish between a normal signal and an abnormal signal.

In one embodiment of the instant invention the surveillance system and method provides a novel system and method to classify the state of a residual error signal wherein said classification is performed using a statistical hypothesis test having any one of a plurality of probability density function models, training procedures, and operating procedures.

In one embodiment of the instant invention the surveillance system and method provides a novel system and method to classify the state of a residual error signal wherein said classification is performed using a probability density function model having any one of a plurality of structures, training procedures, and operating procedures.

Accordingly, having thus summarized the invention, it should be apparent that numerous modifications and adaptations may be resorted to without departing from the scope and fair meaning of the present invention as set forth hereinbelow by the claims.

**BRIEF DESCRIPTION OF THE DRAWINGS**

FIG. 1 is a block diagram of a surveillance system of an embodiment in accordance with the instant invention.

FIG. 2 is a schematic functional flow diagram of an embodiment in accordance with the instant invention.
FIG. 3 is a schematic functional flow diagram of a method and system for training a process model consisting of at least one parameter estimation model and at least one fault detection model using recorded observations of the actual process signals in accordance with the instant invention.

FIG. 4 is a schematic functional flow diagram of a method and system for the fault detection model training procedure in accordance with the instant invention.

FIG. 5 is a schematic functional flow diagram of a preferred method and system for performing surveillance of an asset using at least one parameter estimation model and at least one fault detection model in accordance with the instant invention.

FIG. 6 is a schematic functional flow diagram of a method and system for the fault detection surveillance procedure in accordance with the instant invention.

FIG. 7 is a schematic functional flow diagram of a method and system for performing parameter estimation and fault detection using a redundant sensor.

FIG. 8 is a schematic functional flow diagram of a method and system for performing parameter estimation and fault detection using a generalized parameter estimation model, such as a multivariate state estimation model, a neural network model, an analytical model, or a Kalman filter model.

FIG. 9 is a flow diagram of a training and surveillance procedure in accordance with the instant invention.

FIG. 10 illustrates characteristics of the null and alternate hypotheses for the prior-art sequential probability ratio test (SPRT) mean tests.

FIG. 11 illustrates characteristics of the null and alternate hypotheses for the prior-art SPRT variance tests.

FIG. 12 illustrates acquired operating data, estimated parameter data, and residual error data for a typical Space Shuttle Main Engine accelerometer.

FIG. 13 illustrates a probability density function of the residual error data for a typical Space Shuttle Main Engine accelerometer with comparison to a Gaussian probability density function.

FIG. 14 illustrates an un-optimized one-term expansion probability density function model of the residual error data for a typical Space Shuttle Main Engine accelerometer with comparison to the actual residual error data and a Gaussian probability density function.

FIG. 15 illustrates an un-optimized two-term expansion probability density function model of the residual error data for a typical Space Shuttle Main Engine accelerometer with comparison to the actual residual error data and a Gaussian probability density function.

FIG. 16 illustrates an un-optimized three-term expansion probability density function model of the residual error data for a typical Space Shuttle Main Engine accelerometer with comparison to the actual residual error data and a Gaussian probability density function.

FIG. 17 lists the root mean square errors for five different un-optimized probability density function models computed for each of six Space Shuttle Main Engine accelerometers used for feasibility testing of a preferred embodiment in accordance with the instant invention.

FIG. 18 lists the root mean square errors for five different optimized probability density function models computed for each of six Space Shuttle Main Engine accelerometers used for feasibility testing of a preferred embodiment in accordance with the instant invention.

FIG. 19 illustrates the optimized two-term expansion probability density function model of the residual error data for a typical Space Shuttle Main Engine accelerometer with comparison to the actual residual error data and a Gaussian probability density function.

FIG. 20 illustrates the empirical false alarm rates for feasibility testing of a preferred embodiment in accordance with the prior art SPRT method.

DETAILED DESCRIPTION OF THE INVENTION

Considering the drawings, wherein like reference numerals denote like parts throughout the various drawing figures, reference numeral 10 is directed to the system according to the instant invention.

In its essence, and referring to FIGS. 1 and 2, the system 10 is generally comprised of a method and apparatus for performing high sensitivity surveillance of a wide variety of assets including industrial, utility, business, medical, transportation, financial, and biological processes and apparatus wherein such process and/or apparatus asset preferably has at least one distinct measured or observed signal or sequence comprised of characteristic data values which are processed by the system 10 described herein for providing ultra-sensitive detection of the onset of sensor or data signal degradation, component performance degradation, and process operating anomalies. The system 10 includes a training procedure 30 carried out on a computer 22 such that a process model 50 of an asset 12 (e.g., a process and/or apparatus) is stored in an associated memory means 24 after being learned from historical operating data using at least one of a plurality of computer-assisted procedures in accordance with the instant invention. The historical operating data includes a set or range of observations from expected or typical operation of the asset 12 that are acquired and digitized by a data acquisition means 20 and stored in a memory means 24 as training data 34 by using any combination of electronic data acquisition hardware and signal processing software 20 known to those having ordinary skill in the art, and informed by the present disclosure. As will be delineated in detail infra, one hallmark of the instant invention is the process model 50 for the asset 12 that is derived during the training procedure 30.

The system 10 further includes a surveillance procedure 60 wherein the stored process model 50 is used for high sensitivity computer-assisted surveillance of the asset 12 for the purpose of determining whether a process fault or failure necessitates an alarm or control action. The process model 50 is comprised of a parameter estimation model 52 or collection of parameter estimation models 52 as necessary to provide an estimated value for each sensor or data signal 14 of asset 12 to be monitored for the presence of fault conditions during the surveillance procedure 60. The process model 50 is further comprised of a fault detection model 54 or collection of fault detection models 54 such that at least one fault detection model 54 is provided for each sensor or data signal 14 of asset 12 to be monitored for the presence of fault conditions during the surveillance procedure 60. The fault detection model 54 is another hallmark of the instant invention and will be delineated in further detail hereinbelow.

Process Model Training Procedure:

More specifically, and referring to FIGS. 1 and 3, the training procedure 30 of the system 10 includes a method and apparatus for training or preparing the process model 50 using historical operating data from the asset 12 that has been acquired by the data acquisition means 20 using any
A combination of conventional electronic data acquisition hardware and signal processing software is well known in the art. The historical operating data is acquired in digital format and stored in memory means using a data storage procedure to create a training data set. The training data set includes at least N discrete observations of the asset wherein each single observation, herein denoted Xobs, is comprised of a vector of data values for each signal parameter to be included in the process model for the purposes of the training procedure, the number of observations, N, acquired is at least great enough to adequately bound the operating state space of the asset. Thus, the training data set provides a representative sample of the signals produced by the asset during all normal modes of operation.

Upon acquiring the training data set, the unique training procedure can be implemented in accordance with instant invention. The unique training procedure is comprised of a calibrate parameter estimator procedure and a calibrate fault detector procedure. The calibrate parameter estimator procedure creates the parameter estimation model and trains the parameter estimation model using the training data. The calibrate parameter estimator procedure further stores this model as a new parameter estimation model in the process model.

The process model may contain one or more parameter estimation models depending upon the requirements of the approach taken by a designer. Continuing to refer to FIG. 3, the training procedure may be, in general, performed using any parameter estimation method suitable for defining a parameter estimation model useful for estimating the values of one or more process signals. Methods suitable for the calibrate parameter estimator procedure include, but are not limited to, a plurality of redundant sensor techniques, a plurality of multivariate state estimation techniques, a plurality of neural network techniques, and a plurality of Kalman filter techniques. Each parameter estimation model contained in the process model may be created to implement any one of a plurality of parameter estimation techniques. Further, the parameter estimation technique implemented for an individual parameter estimation model is not constrained to be the same as the parameter estimation technique implemented for any other parameter estimation model contained in the process model.

One example of the calibrate parameter estimator procedure would be the computation of a bias term between two redundant sensors wherein the parameter estimation model used for estimating the value of one sensor during the surveillance procedure consisted of summing the observed value of a second redundant sensor with a bias term computed during the training procedure. This procedure is the mean difference between the two sensor values over the training data set. More sophisticated examples of the training procedure using multivariate state estimation techniques will be described herein below.

Still referring to FIG. 3, the calibrate fault detector procedure makes use of the parameter estimation models to provide estimated values for at least one signal parameter contained in the training data. Generally, the calibrate fault detector procedure will create a separate and distinct fault detection model for each sensor or data signal of the asset to be monitored for the presence of fault conditions during the surveillance procedure. As delineated infra, one hallmark of the instant invention is the fault detection model 54 element of the process model 50 for the asset that is derived during the training procedure. In particular, the instant invention encompasses a statistical hypothesis test type of fault detection model having novel and unique characteristics and calibration procedures described herein below including but not limited to having a probability density function model empirically derived from a numerical analysis of asset operating data.

Continuing to refer to FIG. 3, the training procedure is completed when all training data has been used to calibrate the process model. At this point, the process model contains parameter estimation models and fault detection models for each sensor or data signal of the asset to be monitored for the presence of fault conditions during the surveillance procedure. The process model array is thereafter useful for performing surveillance of the asset.

Referring to FIG. 4, the training procedure is illustrated in additional detail. A designer initializes the calibrate parameter estimators procedure by specifying a set of parameter estimator methods and settings. The parameter estimator methods and settings are then used to operate on the training data via a nominal signal behavior modeling procedure, for example using an MSET training procedure as described herein below, in order to create the parameter estimation models, which are stored in the process model.

Still referring to FIG. 4, the training procedure next proceeds to the calibrate fault detectors procedure wherein the parameter estimation models are an input to the procedure. The designer initializes the calibrate fault detectors procedure by specifying a set of fault detector methods and settings. Next, the estimate process parameters procedure operates the parameter estimation models over the training data to generate an estimated value for each monitored signal value contained in the training data. It is important that the estimate process parameters procedure used in the calibrate fault detectors procedure be the same estimate process parameters procedure that will later be used in the surveillance procedure.

Next, A compute training data residuals procedure calculates the training data residuals for each monitored signal. The training data residuals are next used by a compute nominal residual probability density function (PDF) procedure to create a fault detection model for each monitored signal. In accordance with the instant invention, the fault detection models are typically comprised of mathematical descriptions of the probability density function that best characterizes or best fits the training data residual for the monitored signal. The training data is presumed to accurately characterize the expected normal operating states of the asset. Therefore, the training data residuals are characteristic of the expected normal deviations between the observed signal values and the values estimated using the parameter estimation models. The fault detection models are stored in the process model thereby completing the training procedure.

One hallmark of the instant invention is the method and system for computing the fault detection models by the means of the compute nominal residual probability density function (PDF) procedure. As will be described mathematically herein below, the compute nominal residual probability density function (PDF) procedure fits a general open-ended probability function to the training data residu-
als and employs this fitted function when implementing a herein named Adaptive Sequential Probability (ASP) method and system for computing the fault detection model 54 and thereafter employing said fault detection model 54 for the purpose of performing a fault detection procedure 66 of the surveillance procedure 60.

Surveillance Procedure:
Referring to FIG. 5, the surveillance procedure 60 is comprised of acquiring successive vectors of current operating data and determining for each such observation vector whether the current operating data is indicative of a fault or failure of the asset 12. The surveillance procedure 60 further includes implementing an alarm or control action 70 for the purpose of notifying an operator and/or taking a corrective action in response to a detected fault or failure of the asset 12. The surveillance procedure 60 is in general an open-ended data acquisition and analysis loop that continues until such time as the operator chooses to terminate the surveillance 74.

More specifically, and referring to FIG. 5, the surveillance procedure 60 begins 58 with an acquire current operating data procedure 62 that employs the data acquisition and digitization means 20 (FIG. 1) to acquire a current set of signal data from the monitored asset 12. The current set of signal data is provided to the estimate process parameters procedure 64 that uses the parameter estimation models 52 to estimate values for one or more of the current signal data values. The observed and estimated data are next processed by a perform ASP tests procedure 66 that uses the fault detection models 54 to determine whether a fault is found 68 in the current operating data. If a fault is found 68 is true, an alarm and/or control action 70 is taken. Upon completing the fault found procedure 68, the surveillance procedure 60 then repeats for the next available set of signal data for as long as a surveillance complete decision procedure 72 determines that additional surveillance data are available or terminates at surveillance complete step 74 when no more surveillance data are available or when terminated by an operator.

Referring now to FIG. 6, the perform fault detection procedure 66 of surveillance procedure 60 is illustrated in additional detail. For each current set of signal data values acquired the estimate process parameters procedure 64 uses the parameter estimation models 52 to estimate values for one or more of the current signal data values. The compute data residuals procedure 76 performs a mathematical transformation on the acquired and estimated values to produce a current set of residual data values. Said mathematical transformation is most typically a simple mathematical difference, however, any appropriate transformation may be used including transformations that remove correlated and uncorrelated noise from the residual data values. The residuals produced and transformed in the compute data residuals procedure 76 are next processed by a perform ASP tests procedure 78 that uses the fault detection models 54 to produce an ASP fault indication that is a hallmark of the method and system of the instant invention. Next, the fault found decision procedure 68 is performed on the basis of the ASP fault indication results produced by the perform ASP tests procedure 78. The fault found decision procedure 68 may have any one of a plurality of structures and procedures, including but not limited to methods and systems to perform false alarm filtering by means of observing a time series of ASP fault indication results for the purposes of determining the actual presence of a fault. In one preferred embodiment of the instant invention, a conditional probability fault found decision procedure 68 is used to perform said false alarm filtering.

Continuing to refer to FIG. 6, the estimate process parameters procedure 64 uses the parameter estimation models 52 to estimate values for one or more of the current signal data values wherein the parameter estimation method used may have any one of a plurality of structures and procedures, including but not limited to, a plurality of redundant sensor techniques, a plurality of multivariate state estimation techniques, a plurality of neural network techniques, a plurality of mathematical model techniques, a plurality of autoregressive moving average techniques, and a plurality of Kalman filter techniques.

Referring to FIG. 7, one possible redundant sensor technique for the estimate process parameters procedure 64 is illustrated. The acquire current operating data procedure 62 is used to acquire current signal data values from signals 14 monitored from asset 12 via sensors 18. The estimated value for a first redundant sensor signal is computed using a mathematical transformation on the acquired value of a second redundant sensor signal. Said mathematical transformation is the estimate process parameters procedure 64 that in this case may be a simple equivalence or may include biasing, de-noising or other signal processing. The compute data residuals procedure 76 is then performed followed by the perform ASP tests procedure 78 as described hereinabove and further delineated hereinafter.

Referring to FIG. 8, one possible multivariable parameter estimation technique for the estimate process parameters procedure 64 is illustrated. The acquire current operating data procedure 62 is used to acquire current signal data values from signals 14 monitored from asset 12 via sensors 18. The estimated value for one or more sensor signals is computed using a mathematical transformation on the acquired values of one or more sensor signals. Said mathematical transformation is the estimate process parameters procedure 64 that in this case may implement any feasible parameter estimation technique or procedure, including but not limited to a plurality of multivariate state estimation techniques, a plurality of neural network techniques, a plurality of mathematical model techniques, and a plurality of Kalman filter techniques. The compute data residuals procedure 76 is then performed followed by the perform ASP tests procedure 78 as described hereinabove and further delineated hereinafter.

Referring again to FIG. 6, the usefulness of the instant invention is, inter alia, the improvement achieved in the accuracy of the fault decision procedure 68 that results from the improvement achieved in the accuracy of perform fault detection procedure 66 made possible by the novel perform ASP tests procedure 78 that is a hallmark of the instant invention. Improving the accuracy of the fault decision procedure 68 accomplishes a reduction in the number of false alarms sent to a process operator or control system that can in turn result in an erroneous alarm or control action by the alarm or control action procedure 70. Further, improving the accuracy of the fault decision procedure 68 accomplishes a reduction in the number of missed alarms thereby accomplishing more timely alarm or control action by the alarm or control action procedure 70. The instant invention thereby enables improved operating safety, improved efficiency and performance, and reduced maintenance costs for a wide variety of industrial, utility, business, medical, transportation, financial, and biological processes and apparatuses wherein such process and/or apparatus asset 12 has at least one characteristic data signal suitable for surveillance.

In use and operation, and in one preferred form, FIGS. 1 and 9 outline a general surveillance procedure of the system 10 when employing the novel fault detection model 54
contained in the process model 50 and the accompanying novel fault detection procedure 66 having the perform ASP tests procedure 78. In a typical surveillance procedure, the asset 12 is the source of at least one process signal 14 that is acquired and digitized using conventional data acquisition means 20 for providing the data acquisition procedure for the purpose of computer-assisted surveillance. The digitized signal data is generally evaluated using a computer 22 having computer software modules implementing the estimate process parameters procedure 64, and the perform fault detection procedure 66. The estimate process parameters procedure 64 is used to produce an estimated signal value for at least one process signal 14 emanating from the asset 12. The estimate process parameters procedure 64 in general makes use of the process model 50 stored in a memory means 24 associated with the computer 22 to produce the estimated signal values. The estimated signal values are then generally evaluated using the perform fault detection procedure 66 to identify faults or operating anomalies in the asset 12. The results of the fault detection evaluation are thereafter communicated by a conventional communications means 26 (as is known to those having ordinary skill in the art, and informed by the present disclosure) to an operator console 80 and/or automated process control system 82 for possible alarm and/or control action.

The computer 22 along with the associated memory means 24 can also be employed to perform the training and surveillance procedures 30, 60 as delineated hereinabove and to produce and store all the data associated with these procedures, for example, the historical training data, designer defined settings, and the process model.

Multivariate State Estimation Technique (MSET) for Estimate Process Parameters Procedure:

In one embodiment of the instant invention, the estimate process parameters procedure 64 uses a multivariate state estimation technique (MSET) procedure having an MSET parameter estimation model 52 structure. The U.S. Department of Energy’s Argonne National Laboratory originally developed the implementation of MSET described herein for surveillance of sensors and components in nuclear power plant applications. However, other implementations of a multivariate state estimation technique are possible and useful in conjunction with the instant invention. MSET is in general a statistically derived parameter estimation algorithm that uses advanced pattern recognition techniques to measure the similarity or overlap between signals within a defined domain of asset operation (set of asset operating states). MSET “learns” patterns among the signals by numerical analysis of historical asset operating data. These learned patterns or relationships among the signals are then used to identify the learned state that most closely corresponds with a new signal data observation. By quantifying the relationship between the current and learned states, MSET estimates the current expected response of the asset signals. MSET then uses threshold comparisons or a form of statistical hypothesis testing, such as a Sequential Probability Ratio Test (SPRT) as disclosed in U.S. Pat. No. 5,459,675 which is hereby incorporated by reference in its entirety or the Adaptive Sequential Probability (ASP) procedure that is a hallmark of the instant invention to compare the current estimated value of a signal with its observed value. The MSET procedure provides an accurate and widely applicable method to estimate the operating signal values for an asset. However, other implementations of the parameter estimation procedure are possible and useful in conjunction with the instant invention.

An MSET model is created for the asset 12 using the MSET training algorithms to learn the inherent data relationships within a set of historical asset operating data. The trained MSET model is then used with the MSET parameter estimation and fault detection algorithms to perform the process surveillance function when presented with a new observation of signal data values. The following sections will first provide a mathematical overview of the MSET algorithms and procedures useful for training a parameter estimation model and for using this trained model for process surveillance. The description is followed by a detailed description of a preferred embodiment of the instant invention using a novel Adaptive Sequential Probability (ASP) procedure for fault detection during process surveillance.

MSET Training and Parameter Estimation Procedures:

The MSET methods are generally described in the following two U.S. Government documents produced and maintained by the U.S. Department of Energy’s Argonne National Laboratory, Argonne, Ill., disclosure of which is incorporated in its entirety herein by reference.


The MSET algorithm uses pattern recognition with historical operating data from an asset to generate a parameter estimation model. If data is collected from a process over a range of operating states, this data can be arranged in matrix form, where each column vector (a total of m) in the matrix represents the measurements made at a particular state. Thus, this matrix will have the number of columns equal to the number of states at which observations were made and the number of rows equal to the number of measurements (a total of n signal data values) that were available at each observation. We begin by defining the set of measurements taken at a given time t_j as an observation vector X(t_j),

\[ X(t_j) = [x_{1}(t_j), x_{2}(t_j), \ldots, x_{n}(t_j)]^T \]  

(1)

where x_{i}(t_j) is the measurement from signal i at time t_j. We then define the data collection matrix as the process memory matrix D:

\[
\bar{D} = \begin{bmatrix}
    d_{1,1} & d_{1,2} & \cdots & d_{1,n} \\
    d_{2,1} & d_{2,2} & \cdots & d_{2,n} \\
    \vdots & \vdots & \ddots & \vdots \\
    d_{m,1} & d_{m,2} & \cdots & d_{m,n}
\end{bmatrix} = [\bar{x}_1(\tau), \bar{x}_2(\tau), \ldots, \bar{x}_n(\tau)]
\]  

(2)

Each of the column vectors (X(t_j)) in the process memory matrix represents an operating state of the asset. Any number of observation vectors can be assigned to the process memory matrix. Training an MSET model includes collecting enough unique observation vectors from historical operation of the asset during normal conditions such that the process memory matrix encompasses the full dynamic operating range of the asset. Computation of the D matrix is the first of three steps in the method for training an MSET model based on historical operating data.

One of at least two algorithms is used by MSET to select the vectors in the D matrix. The MinMax algorithm extracts vectors that bound the vector space defined by the training
data and returns the smallest process memory matrix that will produce an effective system model (see also U.S. Pat. No. 5,764,509 which is hereby incorporated by reference in its entirety). A vector ordering algorithm selects and includes representative vectors from the inner regions of the vector space producing a more accurate system model.

Once the process memory matrix has been constructed, MSET is used to model the dynamic behavior of the system. For each current observation of the system (Xobs), MSET compares the observation vector to the stored operating states to calculate an estimate of the process parameter values. The parameter estimate of the current process state (Xest) is an n-element vector that is given by the product of the process memory matrix and a weight vector, W:

$$X_{est} = D \cdot W$$  \hspace{1cm} (E3)

The weight vector represents a measure of similarity between the estimate of the current state and the process memory matrix. To obtain the weight vector, we minimize the error vector, R, where:

$$R = X_{obs} - X_{est}$$ \hspace{1cm} (E4)

The error is minimized for a given state when:

$$W = (D^T \cdot D)^{-1} \cdot (D^T \cdot X_{obs})$$ \hspace{1cm} (E5)

This equation represents a “least squares” minimization when the pattern recognition operator \( \theta \) is the matrix dot product. Several advanced pattern recognition operators have been defined that provide excellent parameter estimation performance. Pattern recognition operators used by MSET include, but are not limited to, the System State Analyzer (SSA) method (see also U.S. Pat. No. 4,937,763 which is hereby incorporated by reference in its entirety), the Bounded Angle Ratio Test (BART) method (see also U.S. Pat. No. 5,987,399 which is hereby incorporated by reference in its entirety), the Vector Pattern Recognizer (VPR) method, the Vector Similarity Evaluation Technique (VSET) method, and the Probabilistic State Estimation Method (PSEM).

Once the weight vector is found, the resulting current state estimate of the system (i.e., the parameter estimate vector) is given by:

$$X_{est} = D \cdot (D^T \cdot W)^{-1} \cdot (D^T \cdot X_{obs})$$ \hspace{1cm} (E6)

The first application of the pattern recognition operator in equation E6 \((D^T \theta D)\) involves a comparison between the row vectors in the \(D^T\) matrix and each of the column vectors in the \(D\) matrix. If we define \(G=D^T \theta D\), then \(G\), the similarity matrix, is an \(m \times m\) matrix. The element in the \(i\)-th row and \(j\)-th column of the matrix \(g_{ij}\) represents a measure of the similarity between the \(i\)-th and \(j\)-th column vectors (i.e., memorized states) in the process memory matrix. The second application of the pattern recognition operator in equation E6 \((D^T \theta X_{obs})\) involves a comparison between the row vectors in the \(D^T\) matrix and each of the elements in the observation vector \(Xobs\). If we define \(A=D^T \theta X_{obs}\), then \(A\), the similarity vector, is an \(m \times 1\) vector. Each element in the similarity vector is a measure of the similarity between the observation vector and the \(i\)-th column vector (i.e., memorized state) in the process memory matrix.

Note that the similarity matrix is a function of the process memory matrix only. Thus, the similarity matrix and its inverse \(G^{-1}\) can be calculated as soon as the process memory matrix has been derived thereby making the application of MSET to an on-line surveillance system more computationally efficient. Computation of the \(G\) matrix initializes the parameter estimation model and completes the second of three steps in the procedure for training an MSET model based on historical operating data.

The third and final step in the MSET training procedure includes analyzing the historical training data using equation E4 to produce a residual error vector, \(R\), for each observation vector in the training data. The collection of residual error vectors comprises the training data residuals necessary for training the fault detection model using any one of a plurality of techniques, including but not limited to the SPRT technique, and the novel ASP technique that is a hallmark of the instant invention.

The Sequential Probability Ratio Test (SPRT) technique is a statistical hypothesis test fault detection algorithm historically used for MSET process surveillance. The SPRT technique is described in U.S. Pat. No. 5,459,675, which is incorporated herein by reference in its entirety. The SPRT analyzes a sequence of discrete residual error values from a sensor to determine whether the sequence is consistent with normal signal behavior or with some other abnormal behavior. When the SPRT reaches a decision about the current signal behavior, e.g., that the signal is behaving normally or abnormally, the decision is reported and the test continues analyzing the signal data. For any SPRT, signal behavior is defined to be normal when the signal data adheres to a Gaussian probability density function (PDF) with mean 0 and variance \(\sigma^2\). Normal signal behavior is referred to as the null hypothesis, \(H_0\). MSET employs four specific SPRT hypothesis tests. Each test determines whether current signal behavior is consistent with the null hypothesis or one of four alternative hypotheses. The four tests are known as the positive mean test, the negative mean test, the nominal variance test, and the inverse variance test. For the positive mean test, the corresponding alternative hypothesis, \(H_1\), is that the signal data adheres to a Gaussian PDF with mean \(+M\) and variance \(\sigma^2\). For the negative mean test, the corresponding alternative hypothesis, \(H_2\), is that the signal data adheres to a Gaussian PDF with mean \(-M\) and variance \(\sigma^2\). For the nominal variance test, the corresponding alternative hypothesis, \(H_3\), is that the signal data adheres to a Gaussian PDF with mean 0 and variance \(\sigma^2\). For the inverse variance test, the corresponding alternative hypothesis, \(H_4\), is that the signal data adheres to a Gaussian PDF with mean 0 and variance \(\sigma^2/V\). The user-assigned constants \(M\) and \(V\) control the sensitivity of the tests.

Limitations of the SPRT Fault Detector Training and Surveillance Method and System:

One significant shortcoming of the SPRT technique is found in the assumptions underlying its mathematical formulation. Specifically, the SPRT technique presumes that the residual error signals adhere to a Gaussian probability density function. Residual error signals that are non-Gaussian, the fault detector false alarm rates and/or missed alarm rates specified by the designer are not accomplished by the SPRT procedure thereby degrading the fault decision accuracy of the asset control and/or surveillance system. The novel ASP technique of the instant invention specifically removes the assumption that the residual error signals
adhere to a Gaussian probability density function. The ASP technique implements any one of a plurality of methods to numerically fit a probability density function to the observed residual error signal distribution that is characteristic of normal asset operation. The derived probability density function is then used to perform a dynamic statistical hypothesis test thereby achieving the fault detector false residual error signal distribution that is characteristic of binary hypothesis testing problems. The possible states of a the fact that in causal reasoning the relationship \( P(e|H) \) is known as the posterior probability, numerical fit a probability density function to the observed

technique implements any one of a plurality of methods to

and/or surveillance system.

dynamic statistical hypothesis test for non-Gaussian residual

Fault Detection Using Statistical Hypothesis Test Procedures:
The general theory underlying the statistical hypothesis test will now be delineated below. Next, the SPRT implementation of a dynamic statistical hypothesis test will be described. Finally, the novel ASP implementation of a dynamic statistical hypothesis test for non-Gaussian residual error signals will be delineated in detail along with a delineation of its reduction to practice.

Bayes’ Rule for a Single Observation:
Statistical decision problems in which there are just two possible outcomes constitute an important class called binary hypothesis testing problems. The possible states of a system are called hypotheses and each individual state of the system is termed a simple hypothesis. A simple hypothesis is a complete specification of a probability distribution for the system (i.e., the distribution of possible observations or samples from the system). The “hypothesis” being tested is that the particular distribution is the correct one.

The basic operation in a binary hypothesis test is to evaluate the veracity of a hypothesis, \( H \), given a piece of evidence or observation, \( e \), from the system being studied. Because of the unpredictability or element of chance inherent in the system, the test deals with the probabilities that events occur or that hypotheses are true. The probability that a hypothesis is true given a piece of evidence is written as \( P(H|e) \). The notation identifies a conditional probability—namely the probability that the hypothesis is true under the condition that the event has occurred with absolute certainty.

Bayes’ well known inversion rule, as described in Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, by Judea Pearl, Morgan Kaufmann Publishers, Inc., San Mateo, Calif., 1988, Second Edition at pages 29 through 39, provides a means for evaluating the conditional probability of a hypothesis, \( P(H|e) \), which is also known as the posterior probability,

\[
P(H|e) = \frac{P(e|H)P(H)}{P(e)} \quad \text{(E7)}
\]

\( P(e|H) \) is the probability that the observation would occur if the hypothesis is true. \( P(H) \) is the probability that the hypothesis is true before any observations of the system have been made, also known as the prior probability. The denominator, \( P(e) \), is the unconditional probability that the observation occurs.

Basic axioms of set theory can be used to prove the following identity for two events \( R \) and \( S \): \( R-(R\cap S) \cup (R^c\cap S^c) \) (\( R^c \) is the converse of event \( S \). In probability theory, the analog of this identity is

\[
P(R)-P(R,S)+P(R,^cS^c), \quad \text{(E8)}
\]

where the notation \( P(R,S) \) is used represent the probability of the joint event \( R \cap S \). The multiplication law states that

the probability of two events occurring jointly can be expressed as a function of the conditional probability of one event based on the other,

\[
P(R,S)=P(R|S)P(S) \quad \text{(E9)}
\]

If \( R \) and \( S \) are independent events, then \( P(R,S)=P(R) \) and \( P(S) \), and the multiplication law simplifies to \( P(R,S)=P(R) \cdot P(S) \).

Bayes’ rule can be simplified by eliminating the denominator in equation E7. Combining equations E8 and E9, and substituting \( e \) for \( R \) and \( H \) for \( S \), the denominator can be written as

\[
P(e)=P(e|H)P(H)+P(e|\overline{H})P(\overline{H}) \quad \text{(E10)}
\]

Therefore, Bayes’ rule becomes

\[
P(H|e) = \frac{P(e|H)P(H)}{P(e|H)P(H)+P(e|\overline{H})P(\overline{H})} \quad \text{(E11)}
\]

The power of Bayesian techniques comes primarily from the fact that in causal reasoning the relationship \( P(e|H) \) is local, namely, given that \( H \) is true, the probability of \( e \) can be estimated naturally and is not dependent on many other propositions. For instance, given that the measurements from an experiment adhere to a particular PDF, the probability that any single measurement will occur is easily computed.

The complementary form of Bayes’ rule provides the posterior probability for the converse of the hypothesis. It is evaluated by substituting \( \overline{H} \) for each instance of \( H \) in equation E11 and noting that \( \overline{(\overline{H})} = H \):

\[
P(\overline{H}|e) = \frac{P(e|\overline{H})P(\overline{H})}{P(e|H)P(H)+P(e|\overline{H})P(\overline{H})} \quad \text{(E12)}
\]

An alternate form of Bayes’ rule is produced by dividing it (i.e., equation E11) with its complementary form (i.e., equation E12) to obtain

\[
\frac{P(H|e)}{P(\overline{H}|e)} = \frac{P(e|H)P(H)}{P(e|\overline{H})P(\overline{H})} \quad \text{(E13)}
\]

This form of Bayes’ rule is further manipulated by first defining the prior odds on hypothesis \( H \) as

\[
O(H) = \frac{P(H)}{P(\overline{H})} \quad \text{(E14)}
\]

the likelihood ratio as

\[
L e(H) = \frac{P(e|H)}{P(e|\overline{H})} \quad \text{(E15)}
\]
and the posterior odds on $H$ as

$$O(H) = \frac{P(H|e)}{P(\neg H|e)}.$$  

Bayes’ rule then specifies the posterior odds as the product of the likelihood ratio and the prior odds,

$$O(H|e) = L(e|H)O(H).$$  

Bayes’ Rule for a Time Series:

The formulation above specifies Bayes’ rule for a single observation and a binary hypothesis. For the application of binary hypothesis tests to real world signals, the formulation must be able to handle a sequence of discrete observations. This is accomplished by beginning with a single observation and successively updating Bayes’ rule for each successive observation. Let the sequence $\{Y_n\}$ be an ordered set of $n$ elements, $\{Y_1, Y_2, \ldots, Y_n\}$, in which the elements are observations of the signal made at $n$ discrete moments in time such that $t_1 < t_2 < \ldots < t_n$. Bayes’ rule for the first observation ($Y_1$) in the time series is

$$P(H|Y_1) = \frac{P(Y_1|H)P(H)}{P(Y_1)}.$$  

Adding the second observation from the time series, Bayes’ rule for the joint event $Y_1, Y_2$ is

$$P(H|Y_1, Y_2) = \frac{P(Y_2|Y_1, H)P(Y_1)P(H)}{P(Y_2|Y_1)P(Y_1)}.$$  

With the aid of the multiplication law (equation E9), the joint event probabilities are converted to conditional probabilities so that the right hand side of this equation can be rewritten as

$$P(H|Y_1, Y_2) = \frac{P(Y_2|Y_1, H)P(Y_1|H)P(H)}{P(Y_2|Y_1)P(Y_1)}.$$  

Note that the probability of the joint event $Y_1, Y_2$ is written as $P(Y_1, Y_2)=P(Y_1|Y_2)P(Y_2)$ instead of the equivalent $P(Y_1, Y_2)=P(Y_2|Y_1)P(Y_1)$ because of the temporal dependency of the data. The second form of the multiplication law reduces to $P(Y_2|Y_1)=P(Y_2)$ because earlier events (e.g., $Y_1$) in a time series cannot be dependent on later events (e.g., $Y_2$).

The multiplication law is used for each successive observation in the time series to derive the form of Bayes’ rule for the joint event ($Y_1, Y_2, \ldots, Y_n$):

$$P(H|Y_1, Y_2, \ldots, Y_n) = \frac{P(Y_n|Y_{n-1}, \ldots, Y_1, H)P(Y_{n-1}|Y_{n-2}, \ldots, Y_1, H)\ldots P(Y_2|Y_1, H)P(Y_1|H)P(H)}{P(Y_n|Y_{n-1}, \ldots, Y_1)P(Y_{n-1}|Y_{n-2}, \ldots, Y_1)\ldots P(Y_2|Y_1)P(Y_1)}.$$  

Since the probabilities for each of the later observations in the time series are conditioned on earlier observations, Bayes’ rule is difficult to solve in the general case. But if the observations in the series are independent of each other (i.e., random), then the probabilities will be dependent on the hypothesis only. In this case the conditional probability for the $i$th observation, $P(Y_i|Y_{i-1}, \ldots, Y_1, H)$, is simply $P(Y_i|H)$. Thus, Bayes’ rule for independent observations in a time series is

$$P(H|Y_1, Y_2, \ldots, Y_n) = \frac{P(Y_n|H)\cdots P(Y_2|H)P(H)}{P(Y_n|Y_{n-1}, \ldots, Y_1)P(Y_{n-1}|Y_{n-2}, \ldots, Y_1)\cdots P(Y_2|Y_1)P(Y_1)}.$$  

If an explicit time dependency can be established for the observations in the time series (i.e., a function is found that relates earlier events to later events), then the general form of Bayes’ rule (equation E21) can be used to develop failure detection models for serially-correlated signals. However, the residual signals formed by the difference between the observed and estimated signal values are in general random signals; thus, Bayes’ rule for random time series is used as the basis for the fault detection models.

Dividing by the complementary form of Bayes’ rule for random time series and utilizing the definition of the posterior odds, prior odds, and likelihood ratio from above, an alternate form of Bayes’ rule for a time series is developed:

$$O(H|y_1, y_2, \ldots, y_n) = O(H)\prod_{i=1}^{n} \frac{P(y_i|H)}{P(y_i|\neg H)}.$$  

If we take the logarithm of this equation, the incremental nature of Bayesian formulation becomes more apparent. Equation E24 shows the log of the likelihood ratio as a weight, carried by each observation in the sequence, which additively sways the belief in the hypothesis one way or the other.

$$\ln(O(H|y_1, y_2, \ldots, y_n)) = \ln(O(H)) + \sum_{i=1}^{n} \ln(P(y_i|H)).$$  

Sequential Hypothesis Tests:

Wald first presented and studied the following sequential test of a simple hypothesis against a simple alternative, as described in Sequential Analysis, by A. Wald, John Wiley and Sons, Inc., New York, 1947. Let $H_0$ be a specific probability density function called the null hypothesis. Then the probability that the time series $\{Y_n\}$ contains samples drawn from $H_0$ is $P(Y_1, Y_2, \ldots, Y_n|H_0)$. Let $H_1$ be a different probability density function called the alternative hypothesis. Then the probability that the time series $\{Y_n\}$ contains samples drawn from $H_1$ is $P(Y_1, Y_2, \ldots, Y_n|H_1)$. Two threshold limits $A$ and $B$ are chosen, with $A < B$, and after each observation in the series the following statistic $(L_{AB})$ is calculated:

$$L_{AB} = \frac{P(Y_1, Y_2, \ldots, Y_n|H_1)}{P(Y_1, Y_2, \ldots, Y_n|H_0)}.$$  

The test procedure is then as follows. If the statistic is greater than or equal to the upper threshold limit (i.e., $L_{AB} \geq B$), then a decision is made to accept hypothesis $H_1$ as true. If the statistic is less than or equal to the lower
threshold limit (i.e., \( \Lambda_{\text{w}} \leq A \)), then a decision is made to accept hypothesis \( H_0 \) as true. If the statistic falls between the two limits (i.e., \( A < \Lambda_{\text{w}} < B \)), then neither hypothesis can yet be accepted to be true and sampling continues. If the observations in the series are independent random variables, then the test statistic reduces to a product of conditional probabilities:

\[
\Lambda_{\text{w}} = \prod_{i=1}^{n} \frac{P(y_i|H_1)}{P(y_i|H_0)}
\]  
(E26)

In the sequential hypothesis tests, the logarithm of the test statistic is often easier to work with:

\[
\ln \Lambda_{\text{w}} = \sum_{i=1}^{n} [\ln \frac{P(y_i|H_1)}{P(y_i|H_0)}]
\]  
(E27)\]

The sequential hypothesis test consists of calculating the logarithm of the test statistic for each observation in the series and comparing the result to the logarithms of the lower and upper threshold limits.

The statistic in the sequential hypothesis test is a product of a sequence of likelihood ratios. Each term in the product is the ratio of a probability conditioned on one hypothesis to a probability conditioned on a second hypothesis. The difference between the likelihood ratios in the sequential hypothesis tests and those in Bayes’ rule (see equations E15 and E23), is that in Bayes’ rule the probabilities are conditioned on a hypothesis and its converse, whereas in the sequential hypothesis tests the probabilities are conditioned on two hypotheses from a set of related hypotheses. In principle, the two hypotheses in the sequential hypothesis tests could be the converse of each other (i.e., the set contains two elements), but in practice the hypotheses are selected from an infinite set of exhaustive and mutually exclusive hypotheses. For instance, suppose the null hypothesis is a Gaussian PDF with a mean of 1 and a variance 10. Then a sequential hypothesis test can be defined in which the alternate hypothesis is any Gaussian PDF in which the mean is not 1 and/or the variance is not 10.

It is informative to compare the sequential hypothesis tests to Bayes’ rule for a time series. Given a null hypothesis \( H_0 \) and an alternative hypothesis \( H_1 \), the likelihood ratio for an observation \( e \) conditioned on the two hypotheses is defined as

\[
L_1(e|H_0) = \frac{P(e|H_1)}{P(e|H_0)}
\]  
(E28)

The subscripts emphasize the fact that the two hypotheses are selected from an infinite set of related hypotheses. The prior odds for the two hypotheses are defined as

\[
O_j(H_0) = \frac{P(H_j)}{P(H_0)}
\]  
(E29)

while the posterior odds are defined as

\[
O_j(H_0|e) = \frac{P(H_j|e)}{P(H_0|e)}
\]  
(E30)

Assuming the observations in the time series are independent, then Bayes’ rule conditioned on hypothesis \( H_j \) (i.e., equation E22 with the symbol \( \Pi \) replaced by \( \Pi \)) can be divided by Bayes’ rule conditioned on hypothesis \( H_0 \) to produce Bayes’ rule for a sequential hypothesis test:

\[
O_j(H_0|y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} L_1(y_i|H_0).
\]  
(E31)

Dividing this equation through by the prior odds, it becomes apparent that the statistic in the sequential hypothesis test is just the ratio of the posterior odds to the prior odds:

\[
\frac{O_j(H_0|y_1, y_2, \ldots, y_n)}{O_j(H_0)} = \prod_{i=1}^{n} L_1(y_i|H_0) = \Lambda_{\text{w}}.
\]  
(E32)

The prior odds are the ratio of the probability that the alternative hypothesis is true to the probability that the null hypothesis is true, before any data have been collected from the system. In many cases, these probabilities are not known—no prior information about the system is known. In the absence of prior information about the system, these probabilities are taken to be 1/2, making the prior odds equal to 1. Thus in the absence of prior information, the test statistic \( \Lambda_{\text{w}} \) is equal to the odds that the system adheres to the alternative PDF as opposed to the null PDF.

Error Probabilities for the Sequential Hypothesis Tests:

Because the sequential hypothesis tests make decisions based on probabilities, there is always a finite probability that a decision reached by the test is erroneous. If a sequential hypothesis test makes a decision to accept the alternative hypothesis \( H_1 \) as true when the null hypothesis \( H_0 \) is true, then an error of type I is said to have occurred. If a sequential hypothesis test makes a decision to accept the null hypothesis \( H_0 \) as true when the alternative hypothesis \( H_1 \) is true, then an error of type II is said to have occurred. Although the designation is arbitrary, it stems from situations in which one kind of error is more serious than the other. Since the normal usage of the sequential hypothesis tests is to detect a change in signal response from its normal behavior (i.e., hypothesis \( H_0 \)) to some abnormal behavior (i.e., hypothesis \( H_1 \)), the error of accepting \( H_1 \) when \( H_0 \) is true is the more serious error. The probability that a decision to accept hypothesis \( H_j \) is erroneous is denoted by \( \alpha \). A type I decision error is also called a false alarm and the probability of a type I error is called the false alarm probability. The probability that a decision to accept hypothesis \( H_j \) is erroneous is denoted by \( \beta \). A type II decision error is also called a missed alarm and the probability of a type II error is called the missed alarm probability.

The sequential hypothesis tests are open-ended. The tests will continue to collect observations from the system and update the test statistic until the test statistic satisfies one of the two decision conditions. In principle, the number of observations needed to reach a decision can be any positive
integer, although it can be shown that a decision will be reached in a finite number of observations. Since the number of observations needed to make a decision is indeterminate, the probability that a decision is erroneous is found by summing the probability of an erroneous decision being made after 1 observation, 2 observations, and so on. Formally, in terms of the threshold limits A and B that define the test, the false alarm probability is given by:

$$\alpha = P(A_{exp} \leq a, B_{exp} \geq B) \ldots$$ (E33)

The first term in the sum is the probability that the test statistic drops below the lower threshold limit after only one observation given that the alternative hypothesis is true. The second term is the probability that the test statistic drops below the lower threshold limit after two observations given that the alternative hypothesis is true. Similarly, the missed alarm probability is given by:

$$\beta = P(A_{exp} \geq A, B_{exp} \leq B) \ldots$$ (E34)

These expressions are by no means easily computed. Moreover, one could not hope to solve these equations for A and B in terms of given a and b, despite the desirability of being able to do so in setting up a test to provide a specified protection. Although these equations cannot be solved, it can be shown that the error probabilities and the threshold limits are related by the following inequalities:

$$A \geq \frac{a}{1-\beta} \quad \text{and} \quad B \geq \frac{\beta}{1-\alpha}.$$ (E35)

The error probabilities and the threshold limits are related by inequalities because the test statistic $A_{exp}$ does not usually attain exactly the value A or the value B when the test is completed. But since a decision is declared as soon as an observation drives the test statistic past either threshold, the inequalities are almost equalities. Indeed, in practice, A and B are taken to be equal to $\alpha(1-\beta)$ and $1-\alpha)/\beta$, respectively. Doing so, of course, means that the sequential hypothesis test actually carried out has error probabilities that are somewhat different than those specified. Let $\alpha'$ and $\beta'$ denote the empirical error probabilities actually attained by a test using specified threshold limits of $A = \alpha/(1-\beta)$ and $B = (1-\alpha)/\beta$. Then according to the inequalities in equation E35, the empirical error probabilities (i.e., $\alpha'$ and $\beta'$) are related to the preassigned error probabilities (i.e., the values of $\alpha$ and $\beta$ used to specify the threshold limits) by

$$\frac{\alpha'}{1-\beta'} \leq A = \frac{\alpha}{1-\beta} \quad \text{and} \quad \frac{\beta'}{1-\alpha'} \geq B = \frac{\beta}{1-\alpha}.$$ (E36)

Multiplying these through to eliminate denominators, one obtains

$$\alpha' \leq \alpha \quad \text{and} \quad \beta' \geq \beta.$$ (E37)

Adding these two equations together, one obtains an inequality relating the empirical error probabilities to the preassigned error probabilities:

$$\alpha' + \beta' \geq \alpha + \beta.$$ (E38)

One of the key features of the hypothesis test technique is that the designer can specify the error probabilities. This is particularly important for type I errors, because false alarms can cause an operator to make an incorrect decision. Type II errors typically do not lead to incorrect decisions. This is because in the event that a real failure does occur, missed alarms may delay the time to detect the failure but not the ability to detect the failure. The result above (equation E38) shows that the preassigned false alarm probability is not a strict upper limit for the empirical false alarm probability. Similarly, the preassigned missed alarm probability is not a strict upper limit for the empirical missed alarm probability, even when the hypothesis test is applied to purely random data. It is the sum of the preassigned error probabilities that is an upper limit for the sum of the empirical error probabilities. Thus it is possible with purely random data for one of the empirical error probabilities to exceed its corresponding preassigned error probability, but both empirical error probabilities cannot be greater than their corresponding preassigned error probabilities.

True upper bounds for the empirical error probabilities can be determined from the inequalities in equation E36,

$$\alpha' \leq \frac{\alpha}{1-\beta} \quad \text{and} \quad \beta' \leq \frac{\beta}{1-\alpha}.$$ (E39)

For small preassigned error probabilities, the true upper bounds are only slightly greater than the preassigned error probabilities. For instance, if both of the preassigned error probabilities are 0.01 then the empirical error probabilities of the test will not exceed 0.0101.

The Sequential Probability Ratio Test (SPRT):

The sequential hypothesis tests described herein above are general statistical tests valid for any pair of related hypotheses. MSET has historically employed four specific sequential hypothesis tests to detect signal faults. These four tests are called the Sequential Probability Ratio Tests, or SPRTs. The SPRTs monitor for changes in the statistical characteristics of the residual signals. A residual signal is the difference between an actual signal and MSET’s estimate of that signal. The SPRTs continually monitor a residual signal, generating sequences of decisions. A decision in which the null hypothesis is accepted (i.e., $A_{exp} \geq B$) is called a normal decision and implies that the residual signal is behaving as anticipated. A decision in which the alternative hypothesis is accepted (i.e., $A_{exp} \leq B$) is called a fault decision and implies that the residual signal is behaving abnormally.

The null hypothesis upon which the SPRTs are based specifies that the residual signal consists of Gaussian data that have a sample mean of 0 and a sample variance of $\sigma^2$. A training procedure, during which the system is operating normally, is used to verify that the mean of the signal is 0 and to evaluate the variance of the signal. Note that if the residual signal does not have a mean of 0, the calculated mean from the training phase is used to normalize the residual signal for the surveillance procedure using the model. Thus the null hypothesis, $H_0$, for the SPRTs is that the signal being analyzed adheres to Gaussian PDF ($\mathcal{N}(y; \mu, \sigma^2)$),

$$\mathcal{N}(y; \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{1}{2\sigma^2} (y - \mu)^2 \right).$$ (E40)

for which the mean ($\mu$) is 0 and the variance is $\sigma^2$. Thus, the conditional probability that a discrete observation ($y_i$) occurs given the null hypothesis is expressed as

$$P(y_i|H_0) = \mathcal{N}(y; 0, \sigma^2)$$ (E41)
The four SPRT tests historically used with MSET are the positive mean test, the negative mean test, the nominal variance test, and the inverse variance test. For the positive mean test, the corresponding alternative hypothesis, \( H_1 \), is that the signal data adhere to a Gaussian PDF with mean \( +\mu \) and variance \( \sigma^2 \). For the negative mean test, the corresponding alternative hypothesis, \( H_2 \), is that the signal data adhere to a Gaussian PDF with mean \( -\mu \) and variance \( \sigma^2 \). For the nominal variance test, the alternative hypothesis, \( H_3 \), is that the signal data adhere to a Gaussian PDF with mean 0 and variance \( \sigma^2 \). For the inverse variance test, the corresponding alternative hypothesis, \( H_4 \), is that the signal data adhere to a Gaussian PDF with mean 0 and variance \( \sigma^2/V \), where \( V \) is the determined system disturbance magnitude for the variance test. For the inverse variance test, the corresponding alternative hypothesis, \( H_4 \), is that the signal data adheres to a Gaussian PDF with mean 0 and variance \( \sigma^2 \).

The SPRT index for the positive mean test (SPRT\(_{+\mu} \)) is given by

\[
\ln \Lambda_{SPRT+} = \sum_{i=1}^{n} \ln \left( \frac{N(y_i; +\mu, \sigma^2)}{N(y_i; 0, \sigma^2)} \right)
\]

The SPRT index for the negative mean test (SPRT\(_{-\mu} \)) is given by

\[
\ln \Lambda_{SPRT-} = \sum_{i=1}^{n} \ln \left( \frac{N(y_i; -\mu, \sigma^2)}{N(y_i; 0, \sigma^2)} \right)
\]

The SPRT index for the nominal variance test (SPRT\(_{nom} \)) is given by

\[
\ln \Lambda_{SPRT_{nom}} = \sum_{i=1}^{n} \ln \left( \frac{N(y_i; 0, \sigma^2)}{N(y_i; 0, \sigma^2/V)} \right)
\]

The SPRT index for the inverse variance test (SPRT\(_{inv} \)) is given by

\[
\ln \Lambda_{SPRT_{inv}} = \sum_{i=1}^{n} \ln \left( \frac{N(y_i; 0, \sigma^2/V)}{N(y_i; 0, \sigma^2)} \right)
\]

The logarithm of the test statistic for the four SPRT tests can be evaluated by substituting the conditional probabilities from equations E41 and E42 into the general formula (equation E27) and simplifying. Thus, for the positive mean test,

\[
\ln \Lambda_{SPRT+} = \frac{1}{2\sigma^2} \sum_{i=1}^{n} y_i^2 - \frac{\mu^2}{\sigma^2} = SPRT_{+\mu}
\]

The Adaptive Sequential Probability (ASP) method defines four new sequential hypothesis tests. The ASP technique is an advanced failure detection technique that broadens the domain of applicability of the SPRT technique to non-Gaussian PDFs. In the ASP method, the problem associated with non-Gaussian data distributions that are typical for residual error signals produced by parameter estimation based techniques, such as MSET. The ASP method uses binary hypothesis tests that are numerically tuned to better accommodate the non-Gaussian data distributions that are typical for residual error signals produced by parameter estimation based techniques, such as MSET. Mathematical Foundations of the ASP Test:

The Adaptive Sequential Probability (ASP) method defines four new sequential hypothesis tests. The Adaptive Sequential Probability method is an advanced failure detection technique that broadens the domain of applicability of the SPRT technique to non-Gaussian PDFs. In the ASP method, the problem associated with non-Gaussian data distributions that are typical for residual error signals produced by parameter estimation based techniques, such as MSET.
method, the assumption that the data fit a Gaussian PDF is relaxed and the test statistic is evaluated for any arbitrary data distribution. In the ASP method, the signal is assumed to consist of random observations that adhere to a specific PDF that is a function of the sample mean, variance, and possibly higher order terms. The PDF is denoted by the general function \( f(y; \mu, \sigma^2, \ldots) \). The parameter list of the function is open-ended to indicate that additional terms, such as the sample skewness, kurtosis, or width of the distribution at half-maximum, may be required to characterize the function.

The null hypothesis upon which the ASP tests are based specifies that the data distribution has a sample mean of 0 and a sample variance of \( \sigma^2 \). A training phase, during which the system is operating normally, is used to verify that the mean of the signal is 0 and to evaluate the sample variance. If the PDF is dependent upon any other additional terms, they are also evaluated during the training phase for numerically tuning one or more probability functions obtained by fitting an open-ended general probability functions to one or more data distributions obtained during typical or normal operating conditions of the asset under surveillance. Thus, the conditional probability that a discrete observation \( (y_i) \) occurs given the null hypothesis is expressed as

\[
P(y_i | H_0) = \mathcal{N}(y_i; 0, \sigma^2, \ldots)
\]  

Typically, four sequential hypothesis tests (i.e., a positive mean test, a negative mean test, a nominal variance test, and an inverse variance test) are utilized by the ASP method. For the positive mean test, the corresponding alternative hypothesis, \( H_1 \), is that the signal data adhere to the specified PDF with mean \( +M \) and variance \( \sigma^2 \) where \( M \) is a preassigned system disturbance magnitude for the mean test. For the negative mean test, the corresponding alternative hypothesis, \( H_2 \), is that the signal data adheres to the specified PDF with mean \( -M \) and variance \( \sigma^2 \). For the nominal variance test, the corresponding alternative hypothesis, \( H_3 \), is that the signal data adhere to the specified PDF with mean 0 and variance \( \sigma^2 \). For the inverse variance test, the corresponding alternative hypothesis, \( H_4 \), is that the signal data adhere to the specified PDF with mean 0 and variance \( \sigma^2/N \). The conditional probabilities that a discrete observation \( (y_i) \) occurs given one of the four alternative hypotheses are expressed as

\[
P(y_i | H_1) = \mathcal{N}(y_i; M, \sigma^2, \ldots)
\]
\[
P(y_i | H_2) = \mathcal{N}(y_i; -M, \sigma^2, \ldots)
\]
\[
P(y_i | H_3) = \mathcal{N}(y_i; 0, V \sigma^2, \ldots)
\]
\[
P(y_i | H_4) = \mathcal{N}(y_i; 0, \sigma^2/N, \ldots)
\]

The logarithm of the test statistic for the four ASP tests can be evaluated by substituting the conditional probabilities from equations E47 and E48 into the general formula (equation E27). The logarithm of the test statistic for a given test is defined to be the ASP index for that test. The ASP index for the positive mean test (\( \text{ASP}_{\text{pos}} \)) is given by

\[
\ln \Lambda_{1n} = \sum_{i=1}^{n} \ln \frac{\mathcal{N}(y_i; M, \sigma^2, \ldots)}{\mathcal{N}(y_i; 0, \sigma^2, \ldots)} = \text{ASP}_{\text{pos}}.
\]

The ASP index for the negative mean test (\( \text{ASP}_{\text{neg}} \)) is given by

\[
\ln \Lambda_{2n} = \sum_{i=1}^{n} \ln \frac{\mathcal{N}(y_i; -M, \sigma^2, \ldots)}{\mathcal{N}(y_i; 0, \sigma^2, \ldots)} = \text{ASP}_{\text{neg}}.
\]

The ASP index for the nominal variance test (\( \text{ASP}_{\text{nom}} \)) is given by

\[
\ln \Lambda_{3n} = \sum_{i=1}^{n} \ln \frac{\mathcal{N}(y_i; 0, V \sigma^2, \ldots)}{\mathcal{N}(y_i; 0, \sigma^2, \ldots)} = \text{ASP}_{\text{nom}}.
\]

The ASP index for the inverse variance test (\( \text{ASP}_{\text{inv}} \)) is given by

\[
\ln \Lambda_{4n} = \sum_{i=1}^{n} \ln \frac{\mathcal{N}(y_i; 0, \sigma^2/N, \ldots)}{\mathcal{N}(y_i; 0, \sigma^2, \ldots)} = \text{ASP}_{\text{inv}}.
\]

The ASP tests are then implemented in the same manner as the SPRT tests. Namely, for each time step in the calculation, the four ASP indices are calculated (equations E49 through E52). Each ASP index is compared to the upper and lower log-thresholds and the status of the test is evaluated (i.e., faulted, normal, or continuing).

ASP Method for Near-Gaussian Distributions:

The Adaptive Sequential Probability method consists of four specific sequential hypothesis tests applied to non-Gaussian data distributions. In order to use the method, a general PDF, \( f(y; \mu, \sigma^2, \ldots, \lambda) \), must be defined for the target signal. In this section, the ASP method is derived for data distributions that are nearly Gaussian.

In applications of MSET to the Space Shuttle Main Engine accelerometer signals, applicants discovered that the residual signals produced by the system model have a nearly, but not truly, Gaussian behavior, as described in Dynamics Sensor Data Validation Phase I Final Report, by applicants: Randall L. Bickford and James P. Herzog, NASA Contract NASX-40874, Jul. 1, 1997. When plotted as a histogram, the residual data will appear to have the same bell-curve shape as a Gaussian distribution. But when a Gaussian PDF of the same mean, standard deviation, and area as the data distribution is superimposed on it, the histogram was found by applicants to be non-Gaussian. Typically, applicants found that the histogram has thicker tails than the Gaussian curve, which corresponds to a sample kurtosis that is greater than 3.5.

A PDF can be written as a sum of a standard normal distribution, \( \mathcal{N}(x) \), and a remainder term, \( R(x) \), as described in Mathematical Methods of Statistics, by H. Cramer, Princeton University Press, Princeton, N.J., 1946.

\[
f(y) = \mathcal{N}(y) + R(y)
\]
where the standard normal PDF

\[ Z(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \]  

(E54)

is a function of the dimensionless variable \( x = (y - \mu)/\sigma \). The standard normal PDF is related to the general Gaussian PDF (equation E40) through the normalization condition:

\[ \int_{-\infty}^{\infty} Z(x) \, dx = \int_{-\infty}^{\infty} N(y, \mu, \sigma^2) \, dy = 1. \]  

(E55)

The remainder term can be expressed as an infinite series of orthogonal polynomials whose terms are functions of derivatives of the standard normal PDF. The constant coefficients of the polynomial terms are dependent on the central moments of the target distribution. As described in Handbook of Mathematical Functions, by M. Abramowitz and I. A. Stegun, Dover Publications, Inc., New York, 1972, the first four terms in the series expansion of the remainder function are:

\[
R(x) = \left[ \frac{1}{6!} \frac{\mu_6}{\sigma^6} Z^{(6)}(x) \right] + \left[ \frac{1}{4!} \frac{\mu_4}{\sigma^4} Z^{(4)}(x) + \frac{1}{6!} \frac{\mu_6}{\sigma^6} Z^{(6)}(x) \right] + \left[ \frac{1}{2!} \frac{\mu_2}{\sigma^2} Z^{(2)}(x) \right] + \left[ \frac{1}{4!} \frac{\mu_4}{\sigma^4} Z^{(4)}(x) \right] + \left[ \frac{1}{6!} \frac{\mu_6}{\sigma^6} Z^{(6)}(x) \right] + \ldots 
\]

(E56)

The \( \mu_i \) factors are central moments of the discrete data sequence \( \{Y_n\} = \{y_1, y_2, \ldots, y_n\} \),

\[ \mu = \frac{1}{n} \sum_{j=1}^{n} (y_j - \mu)^2 \]  

(E57)

The \( Z^{(n)}(x) \) functions are derivatives of the standard normal PDF. The \( n \)th derivative of the standard normal PDF is given by:

\[ Z^{(n)}(x) = (-1)^n \frac{d^n}{dx^n} Z(x) = (-1)^n \frac{n!}{6^n} \frac{\mu_6}{\sigma^6} Z^{(6)}(x), \]  

(E58)

where \( H_n(x) \) are the Hermite polynomials. The first twelve Hermite polynomials are

\[ H_0(x) = 1, \]
\[ H_1(x) = 2x, \]
\[ H_2(x) = 4x^2 - 2, \]
\[ H_3(x) = 8x^3 - 12x, \]
\[ H_4(x) = 16x^4 - 48x^2 + 18, \]
\[ H_5(x) = 32x^5 - 120x^3 + 120x, \]
\[ H_6(x) = 64x^6 - 384x^4 + 576x^2 - 240, \]
\[ H_7(x) = 128x^7 - 1344x^5 + 3360x^3 - 1440x, \]
\[ H_8(x) = 256x^8 - 4096x^6 + 11520x^4 - 11520x^2 + 1792, \]
\[ H_9(x) = 512x^9 - 12288x^7 + 43560x^5 - 71680x^3 + 30240x, \]
\[ H_{10}(x) = 1024x^{10} - 32768x^8 + 111488x^6 - 179200x^4 + 80640x^2 - 6400, \]
\[ H_{11}(x) = 2048x^{11} - 81920x^9 + 376784x^7 - 716800x^5 + 675840x^3 - 201600x, \]
\[ H_{12}(x) = 4096x^{12} - 2097152x^{10} + 10737418x^8 - 21810240x^6 + 22118400x^4 - 9322400x^2 + 1193920, \]

If the data distribution that is to be approximated with equation E53 is nearly Gaussian, a good approximation will be achieved with only a few of the terms in the series. In the ASP method for near-Gaussian distributions, the data distribution is first approximated with the standard normal PDF, \( Z(x) \). The approximation is refined by adding successively higher derivatives of the remainder function until an adequate fit between the general PDF, \( S(x) \), and the data distribution is produced. In practice it is usually not advisable to go beyond the fourth term in the remainder function because the tail regions of the general PDF become unstable.

Using the equations for the derivatives of the standard normal PDF (equations E58 and E59), the remainder function can be written as a product of the standard normal PDF and a polynomial whose coefficients depend on the central moments of the data distribution.

\[ R(x) = Z(x) \sum_{n=0}^{\infty} r_n(x), \]  

(E60)

where the first four terms in the series are

\[ r_1(x) = \frac{1}{6} \frac{\mu_2}{\sigma^2} H_2(x), \]  

(E61)
residual signal data values that fall into each bin. The count number is a discrete function of the sequence \( \{y_i\} \), \( y_1 \), \( y_2 \), \( \ldots \), \( y_{n'} \), where the \( i \)th element in the sequence (\( y_i \)) is given by

\[
y_i = y_{\min} + \frac{(y_{\max} - y_{\min})}{m}.
\]

5

and \( y_{\min} \) is the minimum datum and \( y_{\max} \) is the maximum datum in the residual signal. A normalized residual signal PDF is produced by dividing each element in the count number function by the total number of elements in the residual signal. The number of bins used to produce the residual signal PDF's was 1000. As shown in FIG. 13, a PDF for a typical residual signal has a nearly Gaussian shape. Superimposed on top of residual signal PDF is a Gaussian PDF of the same mean, standard deviation, and area as the residual signal PDF. The first four moments of the residual signal data are as follows: mean = 1.1 \times 10^{-11} \text{g}, standard deviation = 8.2 \times 10^{-12} \text{g}, skewness = 0.32, and kurtosis = 5.7, where the skewness and kurtosis are related to the third and fourth central moments of the data by

\[
\text{skewness} = \frac{\mu_3}{\sigma^3} \quad \text{and} \quad \text{kurtosis} = \frac{\mu_4}{\sigma^4}.
\]

15

Because the kurtosis of the distribution is greater than 3, the residual signal PDF has thicker tail regions than a true Gaussian distribution. This is confirmed by a visual examination of FIG. 13.

The ASP method for near-Gaussian distributions was applied to the residual signals from the MSET model of the SSME accelerometers. The residual signal PDFs were approximated with the series expansion formula of a general PDF (equations E60 through E65). Four calculations were performed for each residual signal in the model. In the first calculation, the residual signals were approximated with the one-term series expansion formula. In the one-term formula, the remainder function \( R(x) \) is given by the product of the standard normal PDF and the first term of the series, \( r_1(x) \). Subsequent calculations introduced additional terms from the series, culminating with the four-term formula in which the remainder function is given by the product of the standard normal PDF and the four-term series, \( r_1(x) + r_2(x) + r_3(x) + r_4(x) \).

For each calculation, the quality of the approximation is measured by the root mean-squared (rms) error of the calculation. The rms error \( (E_{rms}) \) is a function of the difference between the calculated PDF \( (\psi(x)) \) from equation E65 and the residual signal PDF \( (\Phi(x)) \).

\[
E_{rms} = \left( \frac{1}{m} \sum_{i=1}^{m} \left[ \psi(x_i) - \Phi(x_i) \right]^2 \right)^{1/2}.
\]

60

The calculated and residual signal PDFs are discrete functions of the dimensionless variable \( x_i = (y_i - \mu) / \sigma \), where \( y_i \) is an element in the sequence that spans the range of data in the residual signal distribution (see equation E66).

In general terms, each additional term in the series expansion formula improves the approximation of the residual signal PDF. Referring to FIG. 14, FIG. 15 and FIG. 16, adding terms from the series expansion formula improves the fit of
the residual signal PDF from the first sensor. The approxi-
mation of the residual signal PDF generated by the one-term
series expansion formula is compared to a Gaussian approxi-
mation and the residual signal PDF in FIG. 14. The one-term
approximation shows little improvement over the Gaussian
approximation. The rms error for the Gaussian PDF is 0.414,
while the rms error for the one-term series expansion PDF
is slightly smaller at 0.409. The PDF generated by the
two-term series expansion formula is shown in FIG. 15. The
fit of the two-term approximation is significantly better than
that of the Gaussian approximation. The two-term formula
provides an excellent fit of the residual signal PDF, espe-
cially near the peak in the distribution. The greatest source
of error in the approximation occurs in the transition regions
of the distribution, between 1 and 4 standard deviations from
the mean. The rms error for the two-term series expansion
PDF is 0.122, which is nearly a factor of four smaller than
the rms error of the Gaussian PDF. The PDF generated by
the three-term series expansion formula is shown in FIG. 16.
The three-term formula provides a better fit of the peak in
the distribution than does the two-term formula. The rms
error for the three-term approximation is 0.151, which is
slightly larger than that of the two-term approximation
because its fit of the lower transition region is less accurate
than that of the two-term approximation.

The rms errors of the four calculations for each of the
sensors in the model are listed in FIG. 17. Also included in
FIG. 17 are the rms errors for Gaussian PDFs of the same
mean, standard deviation, and area as the residual signal
PDFs. As indicated by the data in the first two columns of
FIG. 17, the one-term series expansion formula produces
PDFs that are slightly more accurate than the corresponding
Gaussian PDFs. Significant improvement is exhibited by the
two-term series expansion formula, which produces PDFs
whose rms errors are as much as a factor of four smaller than
those from the corresponding Gaussian and one-term series
expansion approximations. The three-term series expansion
formula generally produces larger rms errors than does the
two-term formula because the fit of the transitions regions of
the distribution is less accurate. This trend is further exhib-
ited by the four-term series expansion formula. For three
of the calculations (i.e., sensor numbers 1, 2, and 6), the
four-term approximations are unstable in the transition
regions, resulting in rms errors that are much larger than
those of the corresponding Gaussian approximations. The
four-term series expansion formula produces an improved
approximation for the residual signal from sensor number 5
only. The residual signal for this accelerometer is the most
nearly Gaussian of the sensors to begin with, as indicated by
the rms errors for the Gaussian approximations for all six
sensors. These results suggest that the higher order terms in
the series expansion should be used with caution: only those
PDFs that are nearly Gaussian to begin with should be
approximated with terms higher than the second term in the
series expansion.

In the ASP method for near-Gaussian distributions, the
central moments of the distribution to be approximated are
used to evaluate the coefficients in the series expansion. A
second approach, known as the optimized ASP method for
near-Gaussian distributions, was also evaluated. In the sec-
dond approach, the higher-order central moments (i.e., \( \mu_3 \), \( \mu_4 \), \( \mu_5 \), and \( \mu_6 \)) are treated as free parameters in the model and
the approximation is optimized by searching for values of
the central moments that provide the best fit of the distri-
bution. In this approach, the simplex optimization method is
used to minimize the rms error of the calculation, as de-
scribed in A Simplex Method for Function Minimization,
by J. A. Nelder and R. Mead, Computer Journal, Vol. 7,
1965, at pages 308 through 313. The simplex method is
particularly useful for minimizing complicated transcendent-
al functions of two or more variables because it minimizes
a function by utilizing evaluations of the function only—no
evaluations of function derivatives are required.

In the optimized ASP method for near-Gaussian distri-
butions, the first step is to select the approximation formula
to be optimized, either the one, two, three, or four-term series
expansion formula. The number of free parameters in the
model is determined by the approximation formula used. For
instance, the one-term formula requires only one free parame-
ter (i.e., \( \mu_3 \)), whereas the four-term formula requires four
free parameters (i.e., \( \mu_3 \), \( \mu_4 \), \( \mu_5 \), and \( \mu_6 \)). The function that is
minimized by the simplex algorithm is the rms error for
the approximation, given by equation E68. The higher-order
central moments of the distribution are used as initial values
of the free parameters in the model. The simplex algorithm
iterates on the free parameters until a minimum in the
function is found, thereby producing the best fit (i.e., small-
est rms error) between the calculated PDF and the residual
signal PDF for a given approximation formula.

The rms errors of the calculations using the optimized
ASP method for near-Gaussian distributions are listed in
FIG. 18. Also included in FIG. 18 are the rms errors for
Gaussian PDFs of the same mean, standard deviation, and
area as the residual signal PDFs. Comparing the rms errors
of the Gaussian approximations to those from the optimized
calculations shows that all optimized series expansion for-
mulas produce significant reductions in the rms error.
Comparing the data in FIG. 18 to those of FIG. 17 shows that
the optimized ASP method produces more accurate approxima-
tions than does the ASP method for each series expansion
formula. For a given sensor, the optimized two-term series
expansion formula produces rms errors that are roughly a
central factor of two smaller than those of the optimized one-term
formula. The optimized three-term and four-term series
expansion formulas, though, produce rms errors that are
roughly equivalent to those of the optimized two-term
formula. Since the three-term and four-term models incor-
porate more complicated functions with more free param-
eters than the two-term model, these results indicate that the
optimized two-term series expansion method provides the
best balance between accuracy and efficiency.

The PDF generated by the optimized two-term series
expansion formula for the first sensor is shown in FIG. 19.
The optimized two-term formula provides an excellent fit of
the residual signal PDF across all regions of the distribution,
including the transition regions. Comparing the results of the
optimized ASP calculation in FIG. 19 to those from corre-
sponding ASP calculations in FIG. 15 and FIG. 16, reveals
that the main effect of the optimization is to improve the
accuracy of the approximation through the transition regions
of the distribution.

The optimized ASP method for near-Gaussian distribu-
tions produces an excellent fit of the residual signal distri-
butions from the MSET model of SSME accelerometer
signals. Because the optimized ASP method is tuned to the
specific data distributions analyzed, the empirical error
probabilities for the method will better adhere to the theo-
retical limits for the fault detection procedure.

A parametric study of the empirical false alarm rate was
performed to compare the false alarm of the SPRT method
to the new ASP method for the SSME residual signals. In the
study, the residual signals generated by applying the six
sensor MSET model to the accelerometer signals from flight
STS057 engine 1 were analyzed with the SPRT and ASP
The residual signals were large enough to produce accurate results but small enough to permit multiple calculations in a reasonable time span. Typically, the sequential hypothesis tests make a decision for every 1 to 10 data points analyzed. Thus for the parametric calculations of the SSME residual signals, hundreds of thousands of decisions were made for each signal. By dividing the number of fault decisions from a SPRT analysis of a residual signal into the total number of fault and normal decisions made by the test, the empirical false alarm rate could be accurately evaluated.

Referring to FIG. 20, the results from the application of the positive mean SPRT and ASP tests to the residual signal from the first sensor are shown. The calculations shown in FIG. 20 were performed with a constant preassigned missed alarm probability ($\beta$) of 0.01, a constant system disturbance magnitude (M) of 6, and a variable preassigned false alarm probability ($\alpha$). The solid diagonal line in the figure represents the theoretical upper bound for the empirical false alarm probability, as defined by equation E39. For the SPRT calculation, the empirical false alarm rate satisfies the theoretical upper bound only for preassigned false alarm probabilities greater than 0.003. For smaller values of the preassigned false alarm probability, the SPRT model reaches more fault decisions than one would expect based on theoretical arguments. The theoretical upper bound for the empirical false alarm probability is not met for all values of the preassigned false alarm probability because the residual signals are not purely Gaussian, as confirmed in FIG. 13. The residual signal distribution has more data populating the tail regions than a Gaussian distribution of the same mean, standard deviation, and area as the residual signal distribution. The data from the tail regions of a distribution are a major source of false alarms in the SPRT calculations, because they are more representative of the alternative hypothesis than they are of the normal hypothesis. Because the residual signals exhibit more heavily populated tail regions than true Gaussian distributions, they have a tendency to trigger more false alarms than anticipated, especially at high levels of sensitivity (i.e., small values of the false alarm probability). Because the optimized ASP method is tuned to the specific data distributions exhibited by the residual signals, the empirical false alarm rate for the ASP calculation satisfies the theoretical upper bound for all values of the preassigned false alarm probability. The ASP method produces fewer false alarms than the SPRT method for all values of the preassigned false alarm probability.

ASP Benefits and Applications:

The Adaptive Sequential Probability (ASP) technique was developed as the fault detection element of a software program that reliably detects signal data faults for an asset, such as a process and/or apparatus. These signal validation modules improve safety, reduce maintenance cost, and enable optimal performance for a wide range of aeronautical, industrial, chemical, power generating, medical, biological, financial, and military assets. Signal validation is required in all types of process critical control and safety systems where unexpected process interruptions due to sensor or control component failures or false alarms are unsafes or uneconomical. Signal validation assures the safe, reliable operation of a process or apparatus and reduces the manpower, schedule and uncertainty associated with sensor and component failure detection. Signal validation prevents a facility safety or control system from making critical decisions, such as the decision to shut down a process or abort a mission, on the basis of bad sensor data. Signal validation improves process quality and efficiency by ensuring that closed loop control or data analysis is performed using good data. Finally, signal validation increases system up time and decreases system maintenance cost by enabling sensor calibration using on-condition criteria rather than time-in-service criteria.

Moreover, having thus described the invention, it should be apparent that numerous structural modifications and adaptations may be resorted to without departing from the scope and fair meaning of the instant invention as set forth hereinabove and as described herein below by the claims.

We claim:

1. A method for performing surveillance of an asset, the steps including:

   acquiring residual data correlative to expected asset operation;
   fitting a mathematical model to the acquired residual data and storing the mathematical model in a memory means;
   collecting a current set of observed signal data values from the asset;
   using the mathematical model in a sequential hypotheses test to determine if the current set of observed signal data indicates unexpected operation of the asset that is indicative of a fault condition;
   outputting a signal correlative to a detected fault condition for providing asset surveillance.

2. The method of claim 1 wherein the step of acquiring residual data includes a step of acquiring residual data correlative to at least one mode of asset operation.

3. The method of claim 2 wherein the step of acquiring residual data correlative to at least one mode of asset operation includes a step of acquiring past observations of signal data values from the asset and obtaining deviations between the past observations of signal data values and expected signal values obtained by using a model of expected asset operation.

4. The method of claim 1 wherein the step of fitting the mathematical model to the acquired residual data and storing the mathematical model in the memory means includes a step of fitting at least one mathematical function correlative to a frequency distribution to the acquired residual data.

5. The method of claim 4 wherein the step of fitting at least one mathematical function correlative to a frequency distribution to the acquired residual data includes a step of fitting at least one function defining a frequency curve to the acquired residual data.

6. The method of claim 4 wherein the step of fitting at least one mathematical function correlative to a frequency distribution to the acquired residual data includes a step of fitting at least one function defining a frequency histogram to the acquired residual data.

7. The method of claim 4 wherein the step of fitting at least one mathematical function correlative to a frequency distribution to the acquired residual data includes a step of fitting at least one function defining a frequency polygon to the acquired residual data.

8. The method of claim 1 wherein the step of using the mathematical model in the sequential hypotheses test to determine if the current set of observed signal data indicates unexpected operation of the asset that is indicative of a fault
condition further includes the step of performing at least one of a group of sequential hypothesis tests comprised of a positive mean test, a negative mean test, a nominal variance test, and an inverse variance test all as a function of the mathematical model.

9. The method of claim 8 wherein the step of performing at least one of the group of sequential hypothesis tests comprised of the positive mean test, the negative mean test, the nominal variance test, and the inverse variance test all as a function of the mathematical model includes a step of processing the at least one of the group of sequential hypothesis tests and comparing an outcome of each processed test to upper and lower threshold limits for determining unexpected operation of the asset that is indicative of a fault condition.

10. The method of claim 1 further including a step of determining if an outputted signal correlative to a detected fault condition necessitates a control action.

11. The method of claim 1 further including a step of determining if an outputted signal correlative to a detected fault condition necessitates an alarm action.

12. The method of claim 2 where the step of acquiring residual data correlative to at least one mode of asset operation includes a step of transforming the residual data values to remove noise from the residual data values.

13. The method of claim 12 wherein the step of transforming the residual data values to remove noise from the residual data values includes a step of transforming the residual data values to remove correlated noise from the residual data values.

14. The method of claim 12 wherein the step of transforming the residual data values to remove noise from the residual data values includes a step of transforming the residual data values to remove uncorrelated noise from the residual data values.

15. The method of claim 1 wherein the asset is a rocket engine.

16. The method of claim 1 wherein the asset is a rocket engine turbo pump.

17. The method of claim 1 wherein the step of acquiring residual data correlative to expected asset operation includes a step of acquiring observed signal data values from an accelerometer.

18. The method of claim 1 wherein the step of using the mathematical model in the sequential hypotheses test further includes a step of determining if a determined unexpected operation of the asset that is indicative of a fault condition is a degraded signal.

19. The method of claim 1 wherein the step of using the mathematical model in the sequential hypotheses test further includes a step of determining if a determined unexpected operation of the asset that is indicative of a fault condition is a degraded sensor.

20. The method of claim 1 wherein the step of using the mathematical model in the sequential hypotheses test further includes a step of determining if a determined unexpected operation of the asset that is indicative of a fault condition is a degraded component.

21. The method of claim 1 where the step of using the mathematical model in the sequential hypotheses test further includes a step of determining if a determined unexpected operation of the asset that is indicative of a fault condition is an operating anomaly.

22. The method of claim 1 where the step of fitting the mathematical model to the acquired residual data and storing the mathematical model in the memory means includes a step of selecting the parameters in the mathematical model that provide a best fit to a frequency distribution of the acquired residual data.

23. The method of claim 22 where the step of selecting the parameters in the mathematical model that provide the best fit to the frequency distribution of the acquired residual data includes a step of performing an optimization procedure to minimize the error between the mathematical model and the frequency distribution of the acquired residual data.

24. The method of claim 23 wherein the optimization procedure used to minimize the error between the mathematical model and the frequency distribution of the acquired residual data is a simplex optimization procedure.

25. A method for performing surveillance and control of an asset, the steps including:

- acquiring residual data correlative to expected asset operation;
- fitting a mathematical model to the acquired residual data and storing the mathematical model in a memory means;
- collecting a current set of observed signal data values from the asset;
- using the mathematical model in a sequential hypotheses test to determine if the current set of observed signal data indicates an unexpected operation of the asset that is indicative of a fault condition; and determining if a determined unexpected operation of the asset necessitates a control action.

26. The method of claim 25 further including a step of performing a control action when a control action is determined to be necessitated by the determining step.

27. The method of claim 25 further including a step of determining if a determined unexpected operation of the asset necessitates an alarm action.

28. The method of claim 27 further including a step of sounding an alarm when an alarm action is determined to be necessitated by the determining step.

29. A system for performing surveillance of an asset, said system comprising in combination:

- means for acquiring residual data correlative to expected asset operation;
- means for fitting a mathematical model to said acquired residual data and storing said mathematical model in a memory means;
- means for collecting a current set of observed signal data values from the asset;
- means for using said mathematical model in a sequential hypotheses test to determine if said current set of observed signal data indicates unexpected operation of the asset that is indicative of a fault condition;
- a communication means for communicating detected fault conditions to a remote location for providing asset surveillance.

30. The combination of claim 29 further including means for determining if an outputted signal correlative to a detected fault condition necessitates a control action.

31. The combination of claim 29 further including means for determining if an outputted signal correlative to a detected fault condition necessitates an alarm action.

32. The combination of claim 29 wherein said means for fitting said mathematical model to said acquired residual data and storing said mathematical model in said memory means includes means for selecting parameters in said mathematical model that provide a best fit to a frequency distribution of said acquired residual data.

33. The combination of claim 29 wherein said means for using said mathematical model in said sequential hypotheses
test to determine if said current set of observed signal data indicates unexpected operation of the asset that is indicative of a fault condition further includes means for performing at least one of a group of sequential hypothesis tests comprised of a positive mean test, a negative mean test, a nominal variance test, and a inverse variance test all as a function of said mathematical model.

34. The combination of claim 33 wherein said means for performing at least one of said group of sequential hypothesis tests comprised of said positive mean test, said negative mean test, said nominal variance test, and said inverse variance test all as a function of each processed test to upper and lower threshold limits for determining unexpected operation of the asset that is indicative of a fault condition.