MASTER

APPENDIX H

CONF-801049--1

A CONCEPT OF A $(1-\alpha)$ PERFORMANCE CONFIDENCE INTERVAL*

H.H. Leong, G.R. Johnson and T.N. Bechtel

Summary

A multi-input, single-output system is assumed to be represented by some model. The distribution functions of the input and the output variables are considered to be at least obtainable through experimental data. Associated with the computer response of the model corresponding to given inputs, a conditional pseudo-response set is generated. This response can be constructed via the model by using the simulated pseudo-random input variates from a neighborhood defined by a preassigned probability allowance. A pair of such pseudo-response values can then be computed via a procedure corresponding to a $(1-\alpha)$ probability for the conditional pseudo-response set. The range defined by such a pair is called a $(1-\alpha)$ performance confidence interval with respect to the model. The application of this concept can allow comparison of the merit of two models describing the same system, or, it can detect a system change when the current response is out of the performance interval with respect to the previously identified model.

Introduction

The concept of the confidence interval is well established in the theory of statistics. The confidence interval is a numerical region expected to contain the true value of a system parameter. An inference is made of estimating the interval according to a pre-established measure of expectation in probability, $(1-\alpha)$, called the confidence coefficient. If the confidence coefficient equals 0.90, we say that

* Paper to be presented at the 1980 International Conference of Cybernetics and Systems; Boston, MA, October, 1980.

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

there is a probability of 0.90 that the true value of a parameter falls within the confidence region. The confidence interval is established through the probability distribution law of some appropriate statistic function (estimator) that estimates the value of the parameter from a given sample.

The idea of the confidence region has been extended in this paper to encompass the output of a system model rather than an estimation of its parameters. The inference is made here that the true, unknown performance of the system is expected to have a designated confidence of being contained within the estimated interval -- hence the term performance confidence interval. The inference is valid, of course, under certain given conditions only if the model is an equivalent representation of the true system. In our treatment the system model plays the role of the statistical estimator which will be used to simulate the confidence interval for some prescribed condition.

The importance of the performance interval concept stems from the fact that simple agreement between the computed response estimate and the observed system response is inadequate for judging the robustness of a model. A model which is identified on the basis of observed data consisting of a finite set of discrete values can often provide an accurate estimate to the given data in the training period. Applied to a different period or to conditions perturbed among the discrete training points of the same period, the model may prove unreliable due to instability or modelling error. This is because the criteria of simple agreement between model estimate and observed data is limited to a known single input condition at each point within the training period. Also, in most cases, the observed response value is different from the true

response due to system noise or observational error. Furthermore, the inputs to the system are quite often also subjected to noise or behave as random variables. Therefore, it is meaningless to pursue exact agreement between the estimated and observed responses while the observed information is subject to errors even for identical input conditions.

What has been done in statistics literature is to establish an interval estimate such that the probability that the "true" response falls within the desired limits will have some preassigned confidence. In practice it is more desirable to know within what range the response of the real system will occur when we take into account the probable fluctuation of the input and system noise. It is apparent that the more reliable the model (as well as the system) the closer the estimated "true" response will be to the observed response and, hence, the narrower the probable performance range.

It is, however, impossible to enumerate all the probable outcomes. Thus, it is desirable to provide a range of the estimate with a prescribed confidence representing the performance of the system with respect to a perturbation of the given input. The performance confidence interval establishes the range of performance of the model on the basis of some preassigned allowance for perturbations of the input distribution pattern derived from the model training period. This allowance gives more credibility to the model in subsequent periods with similar conditions and derives threshold limits within which the model response has a specified probability to occur. This tempers the criteria of judging the model's accuracy with an indication of its stability.

The presentation of the performance confidence interval concept will be prefaced in the next section by a brief discussion of the perturbed state and variation allowance. These topics are fundamental to the development of the $(1-\alpha)$ performance confidence interval described later. An outline of a proposed procedure for establishing the performance confidence interval is given. Finally, the results of the performance confidence interval calculation for a model of the space-heating energy consumption of a community are presented.

Perturbed State And Variation Allowance

As indicated above, we are not interested in the confidence interval of performance corresponding to the input varying over the whole variation range. Rather, we shall restrict ourselves to the confidence interval conditional on the performance corresponding to the neighborhood of some predefined state. For convenience this neighborhood of perturbation is called the perturbed state.

The measure of the variation range of a given perturbed state is called the variation allowance. This is a probability measure with respect to the model response where the probability of the event that the response estimate is inside the whole unrestricted range is equal to one. A perturbed state with $1/K$ variation allowance means that the event that the model response falls within this perturbed state range has a $1/K$ probability.

The derivation of a performance confidence interval of a model response is then dependent on the behavior of the perturbed responses computed by the model corresponding to those perturbed inputs of the perturbed state. A convenient means of representing the frequency of occurrence of the perturbed fluctuation is by the probability distribution function (pdf) or, equivalently, by the corresponding cumulative

distribution function (cdf). We shall use only an empirical cdf¹ or a curve fitted pdf in approximating the necessary distribution functions, provided that a considerable number of the corresponding data are available (Bury, 1975; Kendall and Stuart, 1961).

Now we proceed to show that the perturbed state centered about a response estimate, $\hat{\theta}_x$, with 1/K variation allowance can be defined on the basis of information provided from the data during the evaluation (model training) period. Let $\{X_i\}$, $\{Y_i\}$ and $\{\hat{\theta}_i\}$, with $i=1$ to NK , be complete samples of inputs, outputs and estimated responses of the system concurrently obtained with respect to the index i . With the knowledge of $\{\hat{\theta}_i\}_{i=1}^{NK}$, an empirical cdf of the model response, denoted as $F_{NK}(\hat{\theta}_i)$, can be easily derived. Using $F_{NK}(\hat{\theta}_i)$ and the 1/K variation allowance, a pair of lower and upper bounds, B_L and B_U , of the perturbed state with respect to the model response is determined by satisfying the probability requirement such that

¹An empirical cdf of a random variable based on a sample of M data points drawn from a population with unknown cdf, $F(Z)$, is denoted as $F_M(Z)$, which is an estimate of the true cdf. The procedure for obtaining $F_M(Z)$ is given as:

1. Rearrange of sample $\{Z_i\}_{i=1}^M$ to an order sequence, $\{Z_{(i)}\}_{i=1}^M$ such that $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(i)} \leq \dots \leq Z_{(M)}$.
2. Assign a value between 0 and 1 to $F_M(Z)$ according to the value of Z as the proportion of observations not exceeding the given value Z . That is,

$$F_M(Z) = \frac{\text{Number of counts of } Z_i \text{ for all } Z_i \leq Z}{M}$$

$$= \begin{cases} 0 & \text{for } Z < Z_{(1)} \\ i/m & \text{for } Z_{(i)} \leq Z < Z_{(i+1)} \\ 1 & \text{for } Z_{(M)} \leq Z \end{cases}$$

$$\Pr[B_L \leq \hat{\theta}_i \leq \hat{\theta}_x] = F_{NK}(\hat{\theta}_x) - F_{NK}(B_L) = \frac{1}{2K}$$

and

(2.1)

$$\Pr[\hat{\theta}_x < \theta_i \leq B_U] = F_{NK}(B_U) - F_{NK}(\hat{\theta}_x) = \frac{1}{2K}.$$

A subset, $\{\hat{\theta}_k\}$, of $\{\hat{\theta}_i\}_{i=1}^{NK}$ of the model estimates in this neighborhood, $[B_L, B_U]$, can be thought as a sample drawn from the perturbed state of output population with unknown conditional cdf $F(\hat{\theta}|\text{perturbed state})$. Similarly, by reverse mapping from the observed output to the concurrent observed input the portion partitioned from the input data $\{\underline{X}_i\}_{i=1}^{NK}$ such that

$$\{\underline{X}_k | \hat{\theta}_k = \theta^*(\underline{X}_k) \in [B_L, B_U]\}_{k=1}^N \quad (2.2)$$

is described as the sample drawn from a perturbed state of the input population with unknown conditional cdf, $F(\underline{X}|\text{perturbed state})$. Then $F_N(\underline{X}|\text{perturbed state})$, an empirical cdf, can be constructed from $\{\underline{X}_k\}_{k=1}^N$.

We stress that the subsample $\{\hat{\theta}_k\}$ is being used as an intermediary of the input subset $\{\underline{X}_k\}$ to define the perturbed state input domain and retrieve its distribution characteristic via $F_N(\underline{X}_k|\text{perturbed state})$. Consequently, we are able to generate another pseudo set of model responses $\{\hat{\theta}_p|\text{perturbed state}\}$ virtually computed through the model $\theta^*(\cdot)$ corresponding to the perturbed inputs, $\{\underline{X}_p|\text{perturbed state}\}$. The perturbed input set can be simulated as desired through the use of the cdf $F_N(\underline{X}_k|\text{perturbed state})$ in conjunction with the Monte Carlo method (Chu, 1969). The distinct feature between this subsample $\{\hat{\theta}_k\}$ and a pseudo set $\{\hat{\theta}_p|\text{perturbed state}\}$ of model responses $\hat{\theta}$ is that the pattern in $\{\hat{\theta}_k\}$ represents the old discrete information from the model evaluation period while the pattern from the pseudo set reflects performance possibilities through random perturbation of input conditions which may not have existed in the evaluation period data.

Construction of a $(1-\alpha)$ Performance Confidence Interval with Respect to $\hat{\theta}_X$ with $1/K$ Variation Allowance

We begin our discussion of the methodology for constructing a $(1-\alpha)$ performance confidence interval with a brief review of the general form of a $(1-\alpha)$ confidence interval for an estimate $\hat{\theta}$, based on a random sample of $X_1, X_2, \dots, X_N$ drawn from the cdf $F(X;\theta)$. The symbol θ is a parameter, or a function of the parameter, which is to be estimated by $\hat{\theta}$ on the basis of the sample values. For each random sample $X_1, X_2, \dots, X_N$ there is an estimate $\hat{\theta}$ for the unknown θ . Because $\hat{\theta}$ is a function of a random sample, it, too, is a random function. When this sampling of fixed sample size is repeated an infinite number of times, the corresponding estimates, $\hat{\theta}$'s, will form a probability population from which we can construct the cdf of $\hat{\theta}$ denoted as $G(\hat{\theta};\theta)$ or equivalently, the pdf of $\hat{\theta}$, denoted as $g(\hat{\theta};\theta)$ or equivalently, the pdf of $\hat{\theta}$, denoted as $g(\hat{\theta};\theta)$. If $\hat{\theta}_L(X_1, X_2, \dots, X_N)$ and $\hat{\theta}_U(X_1, X_2, \dots, X_N)$ are two functions of the sampled values such that $\hat{\theta}_L < \hat{\theta}_U$ and

$$\begin{aligned} \Pr[\hat{\theta}_L < \hat{\theta} < \hat{\theta}_U] &= \int_{\hat{\theta}_L}^{\hat{\theta}_U} g(\hat{\theta};\theta) d\hat{\theta} \\ &= G(\hat{\theta}_U;\theta) - G(\hat{\theta}_L;\theta) \\ &= 1 - \alpha \end{aligned} \quad (3.1)$$

where $(1-\alpha)$ is the preassigned value called the confidence coefficient, then $\hat{\theta}_L$ and $\hat{\theta}_U$ are called the lower and upper confidence limits and $(\hat{\theta}_L, \hat{\theta}_U)$ is called a $(1-\alpha)$ confidence interval of θ .

This analysis does not result in an unique solution for $\hat{\theta}_L$ and $\hat{\theta}_U$. However, the optimum $(1-\alpha)$ confidence interval is the pair $(\hat{\theta}_L, \hat{\theta}_U)$ that

yields the narrowest interval. For example, the best 0.80 confidence interval among the three pairs of $(\hat{\theta}_{Li}, \hat{\theta}_{Ui})$, $i = 1, 2$, and 3, that are shown in Fig. F1 is $(\hat{\theta}_{L2}, \hat{\theta}_{U2})$, because the distance between $\hat{\theta}_{L2}$ and $\hat{\theta}_{U2}$ is less than that between any other pair. For simplicity in the calculation of this procedure a deterministic condition is placed on selecting the pair $(\hat{\theta}_L, \hat{\theta}_U)$ such that

$$\Pr[\hat{\theta} \leq \hat{\theta}_L] = \Pr[\hat{\theta}_U \leq \hat{\theta}] = \alpha/2. \quad (3.2)$$

This is sometimes referred to as an equal-tails or central confidence interval. In fact, such a confidence interval is at its optimum when the pdf of the estimator, $\hat{\theta}$, is symmetric with single mode (Mood et al., 1974).

It is important to notice that knowledge of the distribution $G(\hat{\theta}; \theta)$ is essential for determining the limits, $\hat{\theta}_L$ and $\hat{\theta}_U$; while the derivation of $G(\cdot)$ depends on the distribution of the random sample $X_1, X_2, \dots, X_N$ or equivalently on the knowledge of the distribution $F(X; \theta)$.

We want to extend this general form to a $(1-\alpha)$ performance confidence interval with respect to $\hat{\theta}_X$, an estimate of the system response generated by a model, $\theta^*(X)$, corresponding to an observation of input X with a $1/K$ variation allowance.

Assuming that the multi-input, single output system is represented as

$$Y_X = \theta(X) + W \quad (3.3)$$

where Y_X is the observable output corresponding to the concurrent observable random multi-input, X ; $\theta(\cdot)$ is the true but unknown system response function and W is additive noise with zero mean.

Next we assume the random input, X , and the white noise, W , are statistically independent. Now, since the true response function,

$\theta(\cdot)$, depends on $\underline{X}$ only, and $\underline{X}$ is independent of W , then by statistical theory $\theta(\cdot)$ is also statistically independent from W . This assumption allows separation of the respective fluctuations contributed by $\theta(\cdot)$ and W to the confidence interval.

Furthermore, assume there exists a model, $\theta^*(\cdot)$, which best describes the true system, $\theta(\cdot)$, under a specified criterion. This criterion assures at least that if the model $\theta^*(\cdot)$ is an equivalent representation of the true system (Box and Jenkins, 1976; Davies, 1970) then, by virtue of Eq. 3.3 the residual between the observed and estimated response is an estimate of the white noise and is therefore statistically independent of the input. This assumption can be expressed as

$$\hat{W}_X = Y_X - \hat{\theta}_X, \hat{W}_X \perp \underline{X} \text{ and } \hat{W}_X \perp \hat{\theta}_X,$$

where $\hat{\theta}_X = \theta^*(\underline{X})$, is an estimate of $\theta(\underline{X})$.

With these stipulations in mind, the construction of a performance confidence interval involves estimation of the possible perturbation range of Y_X through the simulation by combining the fluctuations of $\theta^*(\underline{X})$ due to input $\underline{X}$ and the noise, W . Simulations are carried out over the required cdf's in the neighborhood prescribed with the specific conditions, namely, the $1/K$ variation allowance and the confidence coefficient $(1-\alpha)$. In other words, construction of a performance confidence interval is a procedure to simulate the conditional distribution of Y_X over the perturbed state with respect to a given $\underline{X}$ with a $1/K$ variation allowance in $\hat{\theta}_X$, adding the influence of W , and deriving the $(1-\alpha)$ confidence interval from this resulting distribution.

As described in Section 2, on the basis of the information from the evaluation (model training) period, where data of $\{\underline{X}_i\}$, $\{\hat{\theta}_i\}$ and $\{\hat{W}_i = Y_i - \hat{\theta}_i\}$, $i=1$ to NK are available, we can obtain empirical cdf's of the

input $\underline{X}$, the model response $\hat{\theta}$, and the residual (estimated white noise) denoted as $F_{NK}(\underline{X}_i)$, $F_{NK}(\hat{\theta}_i)$ and $F_{NK}(\hat{W}_i)$, respectively. We need only be concerned with the transformation from those cdf's corresponding to the evaluation period to the specific cdfs, namely, the conditional cdfs of input and output of the perturbed state.

With the knowledge of cdf $F_{NK}(\hat{\theta}_i)$ and the prescribed $1/K$ variation allowance, a perturbed state with respect to $\hat{\theta}_X$ corresponding to $\underline{X}$ can be defined through Eq. 2.1. Using the bounds of the perturbed state we can isolate data, both the model response and concurrent input from $\{\hat{\theta}_i\}$ and $\{\underline{X}_i\}$ as samples of the perturbed state. The conditional distribution function of the perturbed state input, denoted as $F_N(\underline{X}|\text{perturbed state})$, can therefore be empirically derived on the basis of the perturbed state sample, $\{\underline{X}_k\}$, $k=1$ to N as given in Eq. 2.2.

At this point we use a statistical simulation technique known as the Monte Carlo method² for generating the desired perturbed variates. As shown in Eq. 3.3 two sources of fluctuations may occur; one from $\theta(\underline{X})$ which in turn is due to $\underline{X}$ and the other from the white noise, W . To take care of the first fluctuation, a perturbed input variate, $\underline{X}_p$, called the pseudo random input (because of its artificial nature) is generated according to the cdf $F_N(\underline{X}_k|\text{perturbed state})$ by the Monte Carlo

²A Monte Carlo method is a probabilistic approach to draw random qualities (or reproduce pseudo random samples) from an underlying probability distribution function for simulating the performance of a stochastic model or approximately solving some complex mathematical problem which is intractable by analytic methods. The systematic formulation was initiated by Von Neumann and his associates during the Second World War for investigating the neutron diffusion problems which were too difficult to be solved analytically and too expensive to be evaluated experimentally (see Hammersley et al., 1964 and Naylor et al., 1968).

method. The pseudo random output of the true but unknown system is estimated by the model and denoted such that $\hat{\theta}_p = \theta^*(\underline{x}_p)$. To account for the second source, which is the influence of white noise, a pseudo random white noise, denoted as $\hat{w}_p$, is simulated according to the empirical cdf of the residual, $F_{NK}(\hat{w}_i)$. The pseudo white noise value is then added to the pseudo response estimate to obtain the approximation of the real system response, called the pseudo response, and denoted such that $Y_p = \hat{\theta}_p + \hat{w}_p$.

The Y_p simulation steps are repeated a large number of times, NP. The resulting population of this pseudo response $\{Y_p | \text{perturbed state w.r.t. } \theta^*(\underline{x})\}_{p=1}^{NP}$ is then used to compute the empirical cdf of system response conditioned over the perturbed state denoted as $G_{NP}(Y_p | \text{perturbed state w.r.t. } \theta^*(\underline{x}))$. With this conditional cdf, $G(\cdot)$ applied to Eqs. 3.1 and 3.2, the performance confidence interval for the response estimate $\hat{\theta}_x$ for a given input $\underline{x}$ is finally computed according to the preassigned confidence, $(1-\alpha)$, such that Y_L satisfies

$$\Pr[Y_p \leq Y_L] = G_{NP}(Y_L | \text{perturbed state}) = \alpha/2$$

and Y_U satisfies

(3.2)

$$Y_U = \Pr[Y_U \leq Y_p] = 1 - G_{NP}(Y_U | \text{perturbed state}) = \alpha/2$$

where Y_L and Y_U are lower and upper confidence limits, respectively.

Outline of a Proposed Procedure

A proposed procedure for constructing a $(1-\alpha)$ performance confidence interval is shown in Fig. F2. For the system model output $\hat{\theta}_x$ with a $1/K$ variation allowance as discussed in the previous section, the procedure is as follows:

1. Specification:

- a) Assign a value for $1/K$ (e.g. $1/20$, $1/10$ or $1/5$) such that the range of variation about the estimate, $\hat{\theta}_X$, will have a $1/K$ change of occurrence in reference to the whole variation range of $\hat{\theta}$.
 - b) Assign a value for $(1-\alpha)$ (e.g. 95%, 90%, or 80%) such that the confidence interval will cover the expected response value $Y_X = \theta(X) + W$ with a $(1-\alpha)$ probability conditioned over the perturbed state specified in (1.a).
2. Compute a best estimate $\hat{\theta}_X$ of the system response corresponding to the given input X by using the system model $\theta^*(\cdot)$ such that $\hat{\theta}_X = \theta^*(X)$.
3. Establish a $1/K$ allowance perturbed state and its distribution function of input.
- a) Obtain the empirical cdf of $\hat{\theta}_i$ on the basis of a data set $\{\hat{\theta}_i\}_{i=1}^{NK}$, which is available from the last evaluation period.
 - b) Construct a perturbed state corresponding to the value of $\hat{\theta}_X$ with a preassigned $1/K$ variation allowance in $\hat{\theta}$ by selecting $[B_L, B_U]$, the lower and upper bounds as defined in Eq. 2.1.
 - c) Isolate all the data belonging to this perturbed state and call this subgroup the sample of the perturbed state. The sample of particular interest is the input data subset:

$$\{X_k \mid \hat{\theta}_k = \theta^*(X_k) \in [B_L, B_U]\}_{k=1}^N$$
 - d) Using the sample of the perturbed state of the input variables, compute the empirical conditional joint pdf, denoted as $f_N(X_k \mid \text{perturbed state})$.

4. Establish the estimated white noise distribution function by retrieving the residual data $\{\hat{W}_i | \hat{W}_i = Y_i - \hat{\theta}_i\}_{i=1}^{NK}$ to construct the required empirical cdf, denoted as $F_{NK}(\hat{W}_i)$. Note that the white noise is assumed statistically independent of $\underline{X}$ and, therefore, is statistically independent of all $\underline{X}$ in the perturbed state and $\theta^*(\underline{X})$.

5. Generate pseudo-random variates

- a) Simulate a pseudo-random input $\underline{X}_p$ according to pdf $f_N(\underline{X}_k | \text{perturbed state})$ by the Monte Carlo method.
- b) Simulate a pseudo-random noise $\hat{W}_p$ according to cdf $F_{NK}(\hat{W}_i)$ by the Monte Carlo method.
- c) Compute a pseudo-random response estimate, Y_p , which is the sum of the model estimated response corresponding to the pseudo-random input plus the pseudo-random white noise. That is, $Y_p = \theta^*(\underline{X}_p) + \hat{W}_p$.
- d) Repeat (5.a) to (5.c) a large number of times, NP , to obtain a large simulated set of pseudo-random response estimates: $\{Y_p\}_{p=1}^{NP}$. This set constitutes the population of the response corresponding to the perturbation of the input over the perturbed state.

6. Establish the distribution function of the pseudo-random response estimate to compute the required confidence limits:

- a) Construct the empirical cdf from $\{Y_p\}_{p=1}^{NP}$, denoted as $G_{NP}(Y_p | \text{perturbed state})$.
- b) Determine the two limits, Y_L and Y_U , corresponding to the preassigned confidence coefficient $(1-\alpha)$, such that Eq. 3.5 is satisfied. The interval (Y_L, Y_U) is the required performance confidence interval of $\hat{\theta}_X$ with the prescribed specification.

It is appropriate here to remark that the construction of the performance confidence interval is quite heuristic, especially in the assignment of a value for the $1/K$ variation allowance. The value $1/K$ should, in fact, be a function of the root mean square error and the number of available data points. From our experience in simulating the perturbation neighborhood of the desired pseudo-random multivariate, even when there were 20 to 23 data points in a perturbed state from 212 original data points, it was not an easy task to construct a joint pdf of 3 variates and obtain a reasonably smooth pdf surface (Rosenblatt, 1956; Watson et al., 1969). Thus, it was necessary to assume that each perturbed state has a multivariate normal distribution, which could be defined by computing the sample mean and covariance matrix.

Illustrations

The procedure described above has been applied to establish confidence in models used in a continuing Colorado State University project to simulate regional space-conditioning energy consumption. Here, we present several illustrations from the project to demonstrate the application of the performance confidence interval concept to dynamic system modelling. As part of an overall framework of adaptive system identification, a set of hypothesis testing procedures was devised to measure how well the model results compare with observed energy consumption. The performance confidence interval is an integral part of one of these hypothesis tests, the reliability test, which has been previously reported (Reiter et al., 1978, 1979 and 1980). The null hypothesis of the reliability test says that for any input the estimation made by the current physical model shall fall within the performance confidence limits simulated by a statistical reference model. Using this modelling

approach, the model is then capable, not only of detecting changes in the habit patterns prevalent in a region, but, also, of updating the current model to cope with the evolution of the region.

The first region of performance confidence interval was generated for the community of Greeley, Colorado. A statistical reference model (Leong and Johnson, 1979) was implemented to produce 212 daily energy consumption values encompassing the 1975-76 and 1976-77 winter heating seasons. From these values the empirical pdf of the estimated system response was computed. Using the empirical pdf and assigning a value of $1/10$ to $1/K$, perturbed states for each of the estimated reference points were constructed. To obtain the conditional pdf of the perturbed output, $g(\hat{\theta}|\text{perturbed state})$, a perturbation neighborhood of 200 pseudo-random input points was computed. From this pdf the performance confidence interval with $(1-\alpha)$ equal to 90% was constructed and is shown in Fig. F3 for the 1976-77 winter season.

Inspection of this figure indicates reliability of the physical model for all but the period March 10-12 where both the observed and estimated energy consumption values (with symbols o and p , respectively) fell outside the 90% confidence region with respect to the reference model response. It was later noted that this period was characterized by a storm with high velocity winds that caused an unusual increase in energy consumption, in other words, an extreme event. Thus, the performance confidence criterion was able to single out this period for special handling rather than rejecting an otherwise acceptable model because of error during this peculiar weather influence.

The ability of the performance confidence region to discern when the parameters of a model need to be updated was demonstrated with the

same Greeley model applied to a later period, the 1978-1979 winter. The results are shown in Fig. F4 including the 90% confidence interval of the reference response, the observed and the physical model estimated space-heating energy consumption denoted by θ and ψ , respectively. Clearly, from the substantial number of observed values which no longer fall within the confidence interval, the statistical reference model originally evaluated during the 1975-76 season is no longer acceptable. The parameters of this model need to be adapted to represent the conservative trend of the changing energy consumption patterns apparent in the community. By the same token, the response based on the physical model (ψ) does not agree with the statistical reference model. One of the models does not represent the real system. Indeed, both models show certain recurring patterns when compared with the observed data.

The performance confidence interval indicated the presence of a similar conservative trend in a second community, Cheyenne, Wyoming, which was modelled for space-heating energy consumption. In this case the 1975-76 winter season was used to evaluate the model and the 1976-77 winter season was used to test the prediction based on the same model. The results for the 152-day period of the first season and the 80-day period of the second season are shown in Figs. F5 and F6, respectively. The season used for evaluation demonstrates 90% confidence in the model as evinced by the correspondence of the observed and estimated values to the confidence region. It is interesting to note that even the period of the March 10-12 wind storm in Greeley falls within the 90% confidence limit for Cheyenne. Despite their relative proximity, Cheyenne experiences much higher average wind velocities than Greeley. Thus, for Cheyenne the wind storm was not signaled as an extreme event as it was

for Greeley. In other words, the presence of higher winds in Cheyenne was compensated for in the confidence interval boundaries.

The confidence interval for the 1976-77 prediction period, on the other hand, does not compensate for an apparent conservative trend indicated by the observed energy consumption in Fig. F6. This is because the confidence region was generated from a model evaluated a year before the conservative trend had gathered momentum. The tendency of the observed energy consumption to dip toward the lower confidence envelope during the latter portion of the season indicates that this model would soon need to be adapted to reflect this trend.

Conclusion

A concept of performance confidence interval applied to check the merit of an observable system model has been introduced. A procedure for constructing such a performance confidence interval on the basis of the Monte Carlo method and a perturbed state is proposed and implemented in a digital computer program. Typical results with the adaptive system identification framework for modelling space heating energy consumption are presented. The significant role played by the performance confidence interval is summarized as follows:

The ability to establish an expected range of deviation of a given predicted response gives the performance confidence interval a twofold purpose as applied to dynamic system modelling. First, by quantifying the range of the response, the interval can be used to check the accuracy of a model. Second, it distinguishes between error and change in the deviation of the model prediction with respect to the real system. Since it is impossible to know the accuracy of the predicted value

before a physical measurement is obtained, the performance confidence interval gives a statistical measure of reliability as a reasonable approximation. This approximation establishes the degree of confidence or accuracy of the model. In this capacity the possibility of modelling error or the occurrence of an extreme event can be isolated whenever the prediction is found to lie outside limits of the performance confidence interval. Alternatively, the observed system response may be found to remain within the confidence region; but tend toward one or the other of its boundaries. This trend indicates that the actual system is changing and the model parameters should be updated accordingly. Thus, the performance confidence interval provides the criteria for judging when a dynamic system model needs to be adapted to cope with the evolution of the actual system. In this way the adaptive identification framework uses the interval as a buffer (no change) zone with a prescribed confidence level to avoid frequent modification of the current model due to small irregular changes.

Acknowledgment

This research was funded by the Department of Energy under Contract DE-AS02-76EV01340.

References

- Box, G.E.P. and G.M. Jenkins, 1976: Time series analysis: forecasting and control. Holden-Day, San Francisco, 1976.
- Bury, K.V., 1975: Statistical models in applied science. John Wiley, N.Y.
- Chu, K., 1969: Quantitative method for business and economic analysis. International Textbook Co., Scranton, PA.
- Davies, W.D.T., 1970: System identification for self-adaptive control. John Wiley, N.Y.
- Hammersley, J.M. and D.C. Handscomb, 1964: Monte Carlo methods. Methue, London.
- Kendall, M.G. and A. Stuart, 1961: The advanced theory of statistics, Vol. I: Distribution theory. Hafner, N.Y.
- Leong, H.H. and G.R. Johnson, 1979: Modelling of energy consumption for space heating for community via GMDH approach. Proceedings of the International Conference of Cybernetics and Systems, 1979, pp. 268-273.
- Mood, A.M., F.A. Graybill and D.C. Boes, 1974: Introduction to the theory of statistics. McGraw-Hill, N.Y.
- Naylor, T.H., J.L. Balintf, D.S. Burdick and K. Chu, 1968: Computer simulation techniques. John Wiley, N.Y.
- Reiter, E.R., E. Dreiseitl, G.R. Johnson, H.H. Leong, B.C. MacDonald, W.L. Somervell, Jr., A.M. Starr and K. Timbre, 1978: The effects of atmospheric variability on energy utilization and conservation. Environmental Research Paper 14, Colorado State University, Ft. Collins, CO.
- , C.C. Burns, E. Dreiseitl, G.R. Johnson, H.H. Leong, B.C. MacDonald, J.D. Sheaffer, and A.M. Starr, 1979: The effects of atmospheric variability on energy utilization and conservation. Environmental Research Paper 18, Colorado State University, Ft. Collins, CO.
- , C.C. Burns, H. Cochrane, H.H. Leong, G.R. Johnson, J.R. McKean, J.D. Sheaffer, A.M. Starr and J. Webber, 1980: The effects of atmospheric variability on energy utilization and conservation. Environmental Research Paper 24, Colorado State University, Ft. Collins, CO.

Rosenblatt, M., 1956: Remarks on some nonparametric estimates of a density function. Ann. Math Statist., 27, pp. 832-837.

Watson, G.S. and M.R. Leadbetter, 1969: On estimating a probability density I. Ann. Math Statist., 40, pp. 1496-1498.

CUMULATIVE RELATIVE FREQUENCY AND RELATIVE FREQUENCY HISTOGRAM CURVES
FROM SIMULATED RESIDUALS OF CHEYENNE, WYO, 1976-77 AND 76-77.

Fig. H1 Simulated probability distribution function of residual (observed minus computed) energy consumption for Cheyenne, Wyoming during 1976-77 heating season. Also shown is the cumulative distribution function of residual values. The lower and upper confidence interval bounds $(\theta_{Li}, \theta_{Ui})$, $i = 1, 2$, and 3 are all qualified to be a $(1-\alpha) = 80\%$ confidence interval of the residual energy consumption for Cheyenne.

Flow chart of generating a $(1-\alpha)$ PERFORMANCE CONFIDENCE INTERVAL to a response of a Model $\theta^*(.)$ for a given input X with a $1/K$ variation allowance

Fig. H2 Flow chart for the generation of a $(1-\alpha)$ Performance Confidence Interval for the response of a model $\theta^*(.)$ for given input X and with a $1/K$ variation allowance.

Fig. H3 Observed and estimated energy consumption for Greeley, Colorado, 1976-77 heating season.

Fig. H4 Observed and estimated energy consumption for Greeley, Colorado, 1978-79 heating season.

Fig. H5

Observed and estimated energy consumption and weather conditions for Cheyenne, Wyoming, during the 1975-76 winter as the evaluation period. The estimated time series represents an approximation of the unknown stochastic process which generates the observed residual energy consumption time sequence. The final error is the difference between observed consumption and the computed consumption after adjustment with the estimated residual time series.

Fig. H6 Observed and estimated energy consumption and weather conditions for Cheyenne, Wyoming, during the 1976-77 winter as the prediction period. The predicted time series is computed with the estimated time series description identified from the previous 1975-76 residual and meteorological data.