A Real Time Approach for Revising Generation Unit Performance Characteristics

G. L. Viviani
Member

Electrical Engineering
Lamar University
Beaumont, TX 77710

C. E. Lin
Member

Electrical Engineering
National Cheng Kung University
Tainan, Taiwan

M. G. Webb

Gulf States Utilities
Beaumont, TX 77704

ABSTRACT- This paper presents a unique method for representing the performance characteristics of generation units of electric utilities. The approach utilizes digitally sampled information in an on-line environment. The resulting accuracy is superior to conventional approaches, as the true time varying nature of performance characteristics is taken into account.

1.0 Introduction

Operational costs for generation units within an electric utility have increased dramatically with the increases in fossil fuel prices. To minimize these costs, it is essential to have accurate information regarding the operational characteristics of generation units. The input-output relationship for a generation unit relates BTU's/Hr of fuel, f(t), flowing into a unit to the output MW's, p(t) produced. The greatest source of inaccuracy in operational cost functions results from inaccuracies in the true input-output relationship for each generation unit. Following convention, the relationship between f(t) and p(t) is expressed as

$$f(t) = Q(p(t)) \qquad (1)$$

where Q is an unknown function. Often, Q is approximated with a quadratic function determined from design data or measured data or both. While this formulation may be accurate for some "snapshot in time", it does not accurately characterize the actual time varying nature of Q. Hence, accuracy of conventional approaches to optimization are limited by the inaccuracy in the representation of Q.

Improvement of the accuracy of measured data associated with generation units is a topic in estimation theory. Application of modern estimation techniques in generation plants is not widespread. A recent paper, [4], describes the application of Kalman Filtering techniques for improved control. The approach of [4] is complicated. In this paper, a simpler approach is presented which the authors feel is equally as good. Improvement in the accuracy of cost curves is addressed directly. The approach of this paper is made possible because of the advancements in digital transducer design, and the recent availability of cost effective transducers and associated digital equipment.

84 SM 612-8 A paper recommended and approved by the IEEE Power System Engineering Committee of the IEEE Power Engineering Society for presentation at the IEEE/PES 1984 Summer Meeting, Seattle, Washington, July 15 - 20, 1984. Manuscript submitted June 7, 1983; made available forp printing May 29, 1984.

Sufficient technology exists to monitor both p(t) and f(t) for each generation unit (natural gas, oil, coal) in a system in real time. The cost of fuel (and other factors) make this technology cost effective. At Gulf States Utilities, both p(t) and f(t) are monitored in real time. Information is sampled digitally, resulting in a continuous stream of information p(nT), f(nT) (n=1,2,3,.....) where T is the sampling period. A value of T equal to (say) 15 seconds allows for accurate representation of power system related variations in Q.

The purpose of this paper is to present a method for representing the variations in Q to more accurately represent true operational costs. Time variations in Q result from various sources including: 1) changing operating characteristics of generation units, 2) transducer inaccuracy, 3) Digital/Analog, Analog/Digital, and roundoff error, 4) environmental and seasonal variations in ambient operating conditions, 5) mechanical failures, and 6) differences in fuel supplies. The historical practice of utilizing a fixed function, Q, to represent the various conditions above is questionable in light of the tremendous costs associated with slight inaccuracies. The method of this paper suggests an approach for on-line monitoring of Q which is practical, reliable and accurate, for gas-fired generating units.

In a pure sense, determination of Q(p(t)) is an estimation problem. Various methods exist for estimating unknown parameters of a mathematical model of known form [4]. However, a general mathematical model of a generation unit, in a relatively simple form, is not available. Hence, the most accurate information describing operational characteristics of generation units results directly from measured quantities f(nT), and p(nT). The approach of this paper is to statistically reduce measured data to a simplified form to accurately characterize the true time variations of Q.

2.0 Method

The method of this paper operates directly on the measured data which is telemetered to a central computer which performs system optimization. For simplicity and data reduction, only integer values of p(nT) (n=1,2,3,.....) are considered. Hence, raw data, p(nT), are always rounded to the nearest integer MW. This results in only slight error. As more and more data are gathered, the random nature of Q becomes evident. For a particular set of data points,

$$(p_1,f_1),\ (p_2,f_2),\ (p_3,f_3),\ldots\ldots\ldots\ldots,(p_m,f_m)$$

0018-9510/85/0002-0280$01.00©1985 IEEE

where $p_1 = p_2 = p_3 = \ldots\ldots = p_m = K_i$, then, in general $f_1 \neq f_2 \neq f_3 \neq \ldots\ldots \neq f_m$. Equivalently, the flow rate for an observed MW output fluctuates due to the random nature of the observations. A similar result is observed for other K_i. It is useful to characterize the statistics of the m observations, for each K_i, of f_i $(i=1,2,3,\ldots\ldots,m)$. Each integer K_i must be constrained with the operational range,

$$MW_{min} < K_i < MW_{max} \quad ,$$

for each unit. The sample mean, h_s, and sample variance, v_s, characterize the statistics of the observations f_i $(i=1,2,3,\ldots\ldots,m)$. These quantities are defined for a general random variable, x, as,

$$h_s = \frac{1}{m}\sum_{i=1}^{m} x_i \qquad (2)$$

$$v_s = \frac{1}{m}\sum_{i=1}^{m} (x_i - h_s)^2 \qquad (3)$$

where m is the number of observations. Since the observations are not stationary, the values of h_s and v_s will vary depending on the observation set.

For simplicity, it is assumed that the time varying $h_s(t)$ and $v_s(t)$, which change depending on when the m samples are taken, can be modeled by piecewise constant functions. Alternatively, the random process which characterizes Q is stationary for short time intervals, $[t_0,t_s]$. The assumption is made that if m samples are gathered within a time period $[t_0,t_s]$, for a particular K_i, the resulting h_{si} and v_{si} will accurately represent the statistics of Q. For this to be possible, it is necessary to assume that $p_i = K_i$ is also constant throughout the interval $[t_0,t_s]$. Given the slowly varying nature of power systems, this is not overly restrictive.

With the above assumptions, it is possible to calculate a set of statistics h_{si}, v_{si} $(i=1,2,3,\ldots,r)$ where r is the number of allowable integer output MW's for a particular unit. For a particular unit, each MW output level K_i $(i=1,2,3,\ldots,r)$ is associated with a time interval $[t_0,t_s]_i$ during which data are gathered. Once the mean and variance are determined for each K_i, the result is a set of points,

$$(K_1,h_{s1}),\ (K_2,h_{s2}),\ \ldots\ldots\ldots\ldots,\ (K_r,h_{sr}).$$

This set of points represents an accurate measure of the true performance of the generation unit based on metered data, alone.

Assuming the data are gaussianly distributed and independent for a particular K_i, confidence intervals [2] are easily determined. (Confidence intervals could also be determined if the data were not independent and gaussian, but determination would likely not be convenient.) A confidence interval provides an interval within which a certain parameter, associated with a random variable is likely to be found with a precise degree of certainty. For a confidence level, α , [2] the interval I determines the likely range of variation in each sample f_i,

$$I_i = \left[\ h_{si} - \frac{v_{si}\ z(1-.5\alpha)}{m},\ h_{si} + \frac{v_{si}\ z(1-.5\alpha)}{m}\right] \qquad (4)$$

where z ( ) denotes a zero mean and unit variance normal distribution function. Typically, α is chosen between 0.95 and 0.99 depending on statistical preferences [2].

As m approaches infinity, the length of I approaches zero, which is the most desirable condition. Practically, it is necessary to limit I to a sufficiently small interval. It is necessary for m to be sufficiently large to insure that the sample mean, h_s, is stable. The length of the interval, I, will also vary as α is varied. For most applications, a choice of α is determined and fixed a priori. By specifying the upper limit on the width of I, the lower limit on m is determined. Engineering judgement is applied to insure that the size of m is compatible with limitations imposed by $[t_0,t_s]$ and T. In general, compromises are necessary.

It is important to prevent bad data from corrupting the pairs of points (h_{si},p_i). A decision function which measures the relative "goodness" of data is necessary. It is possible to utilize the interval I for this purpose, however, a simpler decision function

$$D_i = (f_i - h_{si})^2 \quad , \qquad (5)$$

evaluates the quality of the data by measuring the square of the distance from the sample mean of existing data. To accept or reject data, an appropriate decision rule is necessary. Let b be an integer, then

$$\begin{array}{ll} D < bv_s & \text{accept data} \\ D > bv_s & \text{flag data as invalid} \end{array} \qquad (6)$$

is a practical decision rule for testing incoming data. Bad data can occur for the usual reasons, however, the non-stationary nature of the variations in Q provide an atypical source of bad data. Bad data are omitted from future calculations of the sample mean and variance.

A practical realization of the method described above requires data storage in a push-down stack. (The oldest datum is dropped as the new datum is stored.) For each $p_i = K_i$, for a particular unit, only the last m data are retained. Typically, m = 20. The values of h_{si} and v_{si} are determined based on the last 20 data. If a unit does not operate at a particular output level, K_i, for a period of time, it is possible that the decision rule (6) will flag data as invalid. This results from the non-stationary nature of the process. This problem is overcome by decreasing the sensitivity of the decision rule (increase b) or by having flagged data manually checked. In general, a single flagged datum within a stack for a particular K_i, is not a cause for concern. Flagged data are stored, but not included in the calculation of h_s and v_s. There is a finite chance that flagged data will eventually just be pushed out of the stack. However, the software must indicate a problem exists when the ratio

$$\frac{(\text{flagged data for } K_i)}{m} > E \qquad (7)$$

is greater than some value E (less than 1). This situation is characteristic of abnormal operation of the unit (which may be accurate) or the result of the non-stationary characteristics of the variations in Q. Each situation is easily identified.

3.0 Calculation of Incremental Costs

Most economic dispatch algorithms require piecewise linear incremental cost functions to determine the optimal operating point. For each unit, there is an allowable range of operation $[MW_{min}, MW_{max}] = R$. Also, c = (cost/BTU) for the fuel in a unit. The set of points,

$$(p_1, ch_{s1}),\ (p_2, ch_{s2}),\ \ldots\ldots, (p_r, ch_{sr})$$

represent the cost function for a particular unit. This is a sampled version of cQ. The incremental cost (IC) is found by evaluating the derivative,

$$\frac{d\ cQ}{d\ p} = IC\ .$$

Hence, if a quadratic least squares fit [3] is made for the points (p_i, ch_{si}) (i=1,2,3,....,r), a linear IC function results. For greater accuracy, the range R is divided in subintervals and a quadratic least squares fit is performed over each sub interval. The resulting IC function is piecewise linear which is the required form for most Economic Dispatch algorithms. Typically, the interval R is divided into 4 subintervals.

4.0 Results

Gulf States Utilities is installing a new energy management system (EMS). A test version of the algorithm described in this paper is currently being evaluated. For this system, m=20, T=15 sec, and R is divided in 5 subintervals. Data are recorded continuously. After a new data point, for a particular unit, is recorded, the associated h_{si} and v_{si} are recalculated. Revised least squares fits, for the newly stored data, are performed regularly. The resulting cost curves replace obsolete cost curves. This replacement is made on-line to insure that all economic calculations are performed with curves that reflect the true operating characteristics of the system. All incoming data are stored in push-down stacks and marked with a time stamp indicating the date and time of storage. Bad data are detected with the decision rule of (6). An error condition results when the ratio (7) is sufficiently large. For the most part, this algorithm operates with no intervention.

To test the method of this paper, it is necessary to determine the sort of variations in telemetered data likely to be encountered. At Gulf States Utilities, the capability to monitor f(t) and p(t) continuously is limited to the plant environment only (at the present time). The communications system to transmit the information back to the centralized system computer is under construction. Hence, sample data were gathered at the plant. This information was transferred (by hand) to the central computer. Figure 1 represents the input-output relationship, for a subset of the entire operational range, for a 100 MW unit. Two curves shown represent the input-output curve for ramping up and down 10MW in approximately 10 minutes. Only one point was recorded for each output level. The purpose of these curves is to illustrate the variability of the input output relationship for different operating conditions. The curve labelled "average" represents the input-output curve (for a 10 MW range) as calculated by taking the average of 20 observations, for each MW output level, under normal operating conditions. The curve labelled "actual" is a section of the actual input-output relation, in use, for this unit. One can see that the "average" curve gives a realistic representation of the unit's performance including the apparent valve point at approximately 88 MW. Since virtually all economic dispatch algorithms require a quadratic input-output relationship, it is necessary to find a quadratic least squares fit to the "average" curve. This will tend to minimize the effects of a valve point. However, one can conclude that since the initial curve included the effects of a valve point, the fitted result will be more accurate than fitting data which neglects valve points.

The statistics of the data were analyzed. Specifically, the sample variance for a particular MW output was calculated in order to have a realistic idea about the size of the variations that will result in practice (see Figure 2 and 3). While the true variations at other units, in other plants, for each MW output, will likely be different, the sample data gathered gives an adequate indication of the true variations (for test purposes).

Figure 2 represents the average BTU/Hr flows, h_{si} (i=1,2,3,....,50) where T = 15 seconds and $p(nT)_i$ = 92 MW. The 50 observations were selected from a larger contiguous set of samples which included observations for $p(nT)_i$ not equal to 92 MW. v_{si} (i=1,2,3,....50) is shown in Figure 2. The horizontal axis on both plots is the same, for easy comparison. The fluctuations are small but the real dollar costs of inaccuracies may not be small. For example, assume that a large unit in the system was one percent less efficient than presumed by conventional measurements. This could result in incorrect results from economic dispatch, which could result in significant costs to the utility.

Additional insight regarding the sorts of operational variations actually encountered is found by referring to Table 1. For the same unit, in the same day, at two different times, sample data for operation at approximately 91 MW were recorded consecutively at 15 second intervals.

TIME	MW	MBTU/hr	AVERAGE
11:40:00	90.9	959	
11:40:15	91.3	964	
11:40:30	91.1	973	
11:40:45	91.0	970	968
11:50:00	90.9	972	
11:50:15	91.0	971	
11:50:30	91.4	970	
12:30:00	91.0	936	
12:30:15	91.1	926	
12:30:30	90.9	920	
12:30:45	90.9	911	916
12:40:00	90.9	908	
12:40:15	90.7	906	
12:40:30	90.7	905	

TABLE 1

Throughout this time period, the unit was experiencing "normal" operations under the control of AGC. Approximately a 6 percent difference in efficiencies is verified. This tends to support the argument that input-output curves should be updated regularly, for greatest accuracy, as described in this paper.

5.0 Discussion

To minimize data manipulations and for clarity, the previous analyses have focused on a particular subset of the unit operational range which is believed to be characteristic of other subintervals. Exactly how to determine, R, and the location of subintervals for unit operation is based on engineering judgement and other factors. Important considerations include the input requirements of available economic dispatch algorithms, simplicity, accuracy and valve point locations. Ideally, it would be nice to split the operational range at the valve points to enhance the accuracy of the fitted curves. Independent of how the range of operation is divided into subintervals, the measured data will reflect the effects of valve points. Using these data to determine a quadratic least squares fit (or other appropriate fit) will incorporate the effects of valve points to the input-output characteristics for the unit.

6.0 Potential Improvements

The installation as described will significantly improve the accuracy of all economic optimization algorithms which require input-output relationships for each unit. However, it may be desirable to direct system operator attention to malfunctioning generation units or metering equipment. This goal is satisfied by allowing flagged data to drive alarming functions within the EMS. This type of improvement would tend to make the system operator more aware of operational problems within plants.

7.0 Conclusions

The approach outlined in this paper has at least two advantages over conventional approaches. First, the input-output relationship is modified to reflect actual operating conditions. This information is useful to many computer analyses associated with modern power system operation and planning. Secondly, this approach provides a mechanism for alarming abnormal operating conditions. The first result is a long overdue necessity. The second result may provide a mechanism for achieving a closer working relationship between plants and the system operations staff.

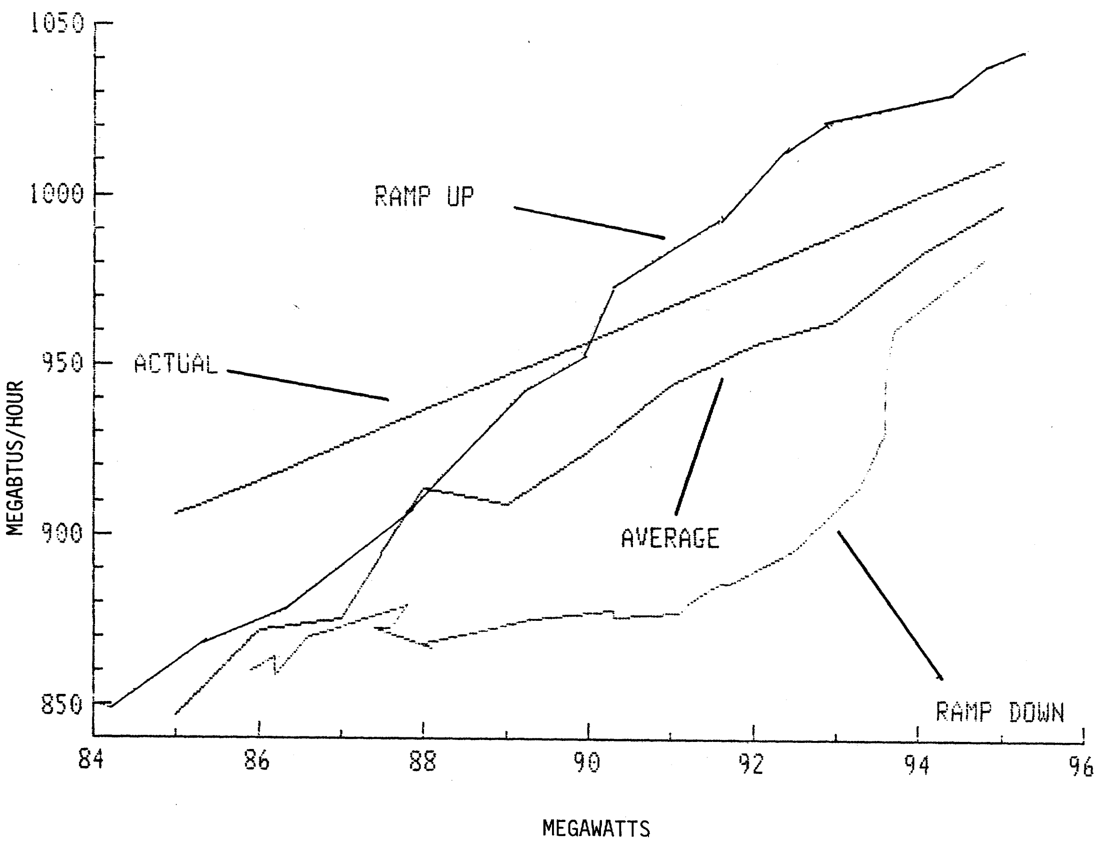

Figure 1

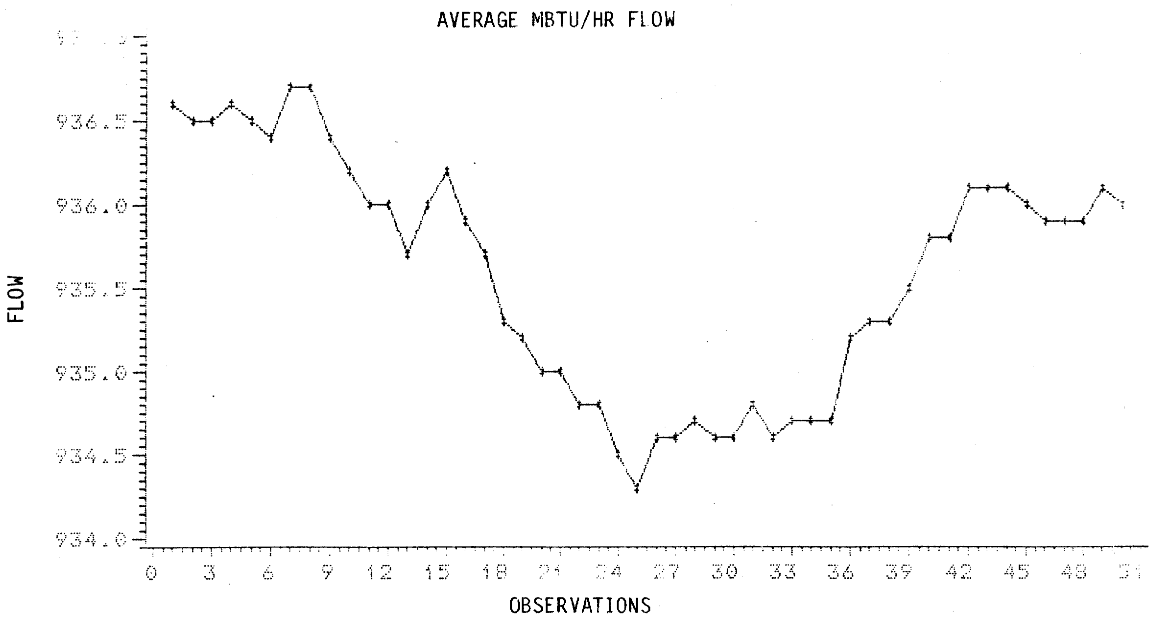

FIGURE 2

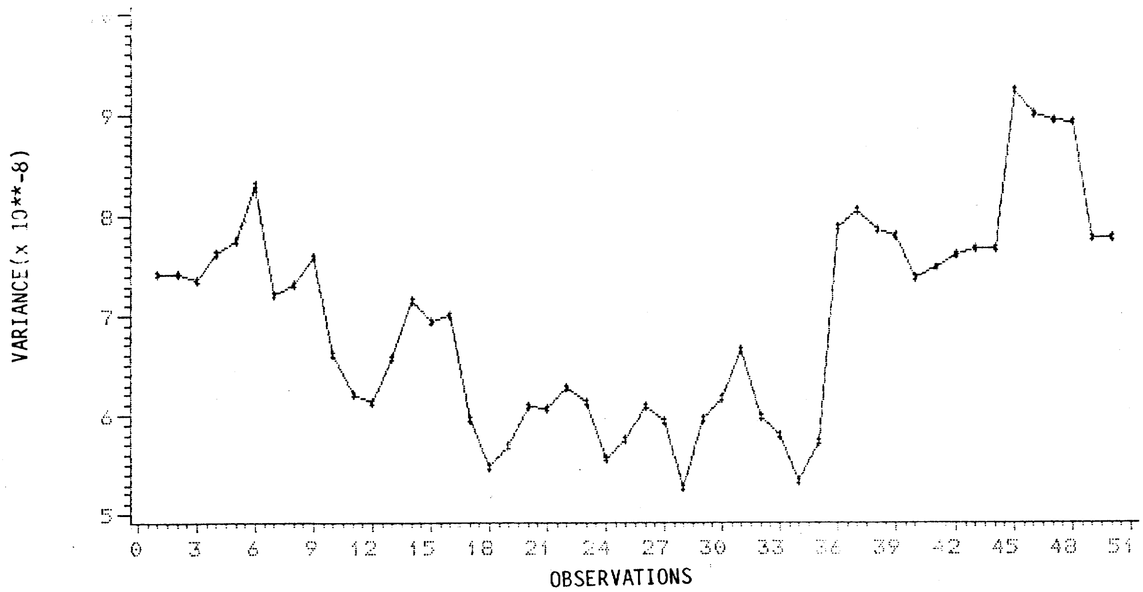

FIGURE 3

8.0 References

[1] Papoulis, A., Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1965.

[2] Bickel, P. J., Doksum, K. A., Mathematical Statistics, Holden-Day, San Fransisco, 1977.

[3] Box, G. E. P., Jenkins, G. M., Time Series Analysis - forecasting and control, Holden-Day, San Fransisco, 1976.

[4] Wallace, J. N., Clarke, R., "The Application of Kalman Filtering Estimation Techniques in Power Station Control Systems," IEEE Transactions on Automatic Control, March 1983, pp. 416-427.

Discussion

Anjan Bose (Arizona State University, Tempe, AZ): The real time updating of generator input-output characteristics has received a lot of discussion since the dramatic rise in fuel costs. Thus, it is gratifying to see an algorithm being actually implemented by Gulf States Utilities. The approach is particularly attractive because the whole generation plant is considered a black box whose input (BTU) and output (MW) are only of interest. Very little needs to be assumed about the black box itself.

The plant, however, is a very complex system and whether its input-output characteristics can be accurately calculated without considering some of these complexities is a quesiton that has to be considered quite carefully. The off-line determination of these characteristics has been the subject of recent research but the approaches have taken into account much of the combustion cycle. If, as is claimed in the paper, the black box approach produces adequate accuracy, the off-line tests can also be significantly simplified. Such simplification will enable more frequent off-line updating of the cost characteristics, a major improvement on present day practice.

It is generally assumed that the drift in the characteristics of a generator is small when there are no changes in the plant system. However, the characteristics can change drastically when there is an outage of a major piece of equipment at the plant causing the biggest errors in economic dispatch. Since equipment failures are quite common in a large plant, it has been suggested that input-output characteristics should be determined and stored for different conditions of equipment status. The plant operator can then manually select the right curve depending on the real time status. The method proposed in the paper appears to be able to handle such a scheme. It is not clear, however, whether the method can be successful without this scheme. Since the method assumes gradual changes, an abrupt change in the characteristics may produce wrong results. If the gradual updating is then aborted, an alternate curve may be necessary to restart the updating.

The authors' comments on these issues will be appreciated. This subject is of great interest because of its economic consequences.

Manuscript received August 6, 1984.

G. L. Viviani, C. E. Lin and **M. G. Webb:** The authors wish to express their appreciation for the interest shown in their paper. As indicated in the discussion, this topic maintains considerable economic interest to utilities. Our intent was to extend existing approaches, without radical change. Conventionally, the "black box" approach is the usual approach for a utility to determine the input-output characteristics of a unit. The limitation of this approach has been the frequency of updates. Therefore, the usefulness of the black box approach is well established. The authors agree that is may be possible to improve the accuracy of this approach by applying nonlinear Kalman Filtering of similar techniques. However, the increased overhead with regard to a plant model would be very significant. Except for REf. [4] in the paper and some pilot studies underway with EPRI, the authors are unaware of any efforts to apply more sophisticated estimation techniques to power plants. Our tacit conclusion is that utilities perceive application of sophisticated techniques as too expensive and/or complicated.

As shown in our Fig. 1, significant drift in operational charàcteristics does occur for our test unit. We believe this is the norm, more than the exception. Therefore, by reducing the number of retained samples, m, it is possible to allow for more rapid change of the input-output characteristics. A suitable m can be determined by trial and error. Retaining and storing numerous curves for various conditions is perceived, by the authors, as an unwieldy approach for modeling the operational characteristics for a unit.

The method of our paper is suitable for indicating abnormal operational conditions. If the likely causes of component outages could be determined in advance, it might be possible to retain a limited number of curves to rapidly reflect operational changes as suggested by Dr. Bose. However, the proposed method will eventually reflect the revised operational conditions which represents a significant improvement over present methods which are not flexible.

Manuscript received September 10, 1984.