OPTIMAL CONTROL AND FAULT DETECTION IN HEATING, VENTILATING, AND AIR-CONDITIONING SYSTEMS

F.L.F. Pape

J.W. Mitchell, Ph.D., P.E.
Member ASHRAE

W.A. Beckman, Ph.D., P.E.

ABSTRACT

A methodology is developed for fault detection in heating, ventilating, and air-conditioning (HVAC) systems based on optimal control. The optimal control strategy is determined using information that could come from an energy management and control system (EMCS). This information is also used to detect faults in system operation. Deviations from optimal performance are sensed by comparing the measured system power with the power predicted with the optimal strategy. Various statistical approaches to using the difference between actual and optimal power are described. Both individual measurements and sequences of data are employed to determine the existence and location of faults. The trade-off between the level at which faults can be detected and the speed of detecting faults is discussed. The methodology is tested using simulations of an HVAC system operated both with and without faults present.

INTRODUCTION

Improvements in the design, control, and maintenance of HVAC systems can result in significant savings in building energy use. Braun (1988) developed a methodology for determining the optimal control strategy for an HVAC system and showed that savings of 10% to 20% are possible through optimal control. An important aspect in the use of optimal control is the determination of equipment degradation and sensor failures that cause the performance to decrease.

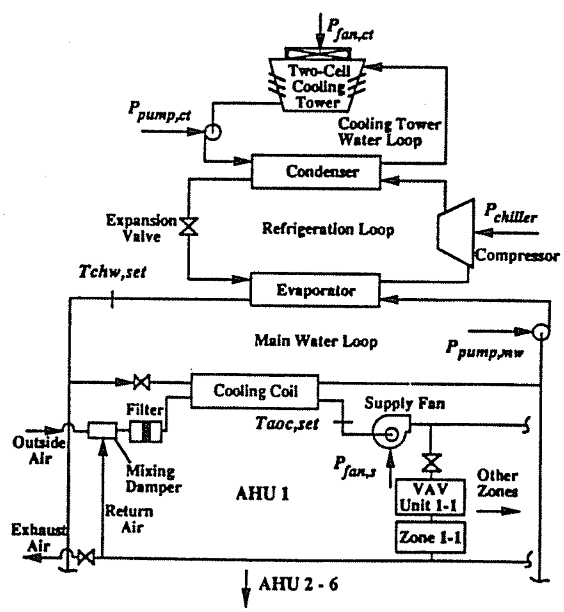
Kao (1985) showed that sensor errors in the air-handling unit of an HVAC system increase the annual energy requirements up to 50%. Urso et al. (1985a, b) developed a fault detection technique based on a model of the system without faults. The model is used to generate estimates of system performance under current operating conditions. Differences between expected and actual performance are analyzed to see whether statistically significant changes in parameters have occurred. A Kalman filter technique is used to reduce the impact of noise and uncertainty. It is feasible to detect bias errors in temperature sensors with this method. The approach requires a detailed modeling of the system to determine the expected operation under normal operating conditions.

Anderson et al. (1989) used a rule-based expert system with a statistical analysis data processor to provide an hourly diagnostic analysis of an industrial HVAC system. Monitoring and diagnostic functions were performed on data collected by the existing data acquisition system. The statistical analysis preprocessor compared incoming data with predictors and computed a normalized variance for each channel of data. The data were searched for inconsistencies as a way to find sensor failure. The variances were compared with set limits within the expert system as a check on consistent system operation. The main limitations of their study are that the predictors used in the system are not physically based.

The approach in this study will be to first develop the optimal control strategy for the system. Statistical methods will then be used to determine the deviation from optimal performance. The approach is described in detail by Pape (1989).

Figure 1 shows a schematic of a typical variable-air-volume (VAV) air-conditioning system. It is a centralized chilled-water facility, with chilled water distributed to a set of air-handling units. The HVAC system meets the air-conditioning requirements of the building zones through control of the supply fan speed, temperature settings, chilled-water and condenser water pump speed, and cooling tower fan speed. This system will be used to illustrate the tasks of optimal control and fault detection in HVAC systems.

The setting of the control variables is performed by a supervisory controller. Control variables may be both continuous, such as the temperature settings, or discrete,



$$P_{tot} = P_{fan,ct} + P_{pump,ct} + P_{chiller} + P_{fan,s} + P_{pump,mw}$$

Figure 1 *Schematic of a conventional air-conditioning system*

F.L.F Pape is a former Graduate Student and **John W. Mitchell** and W.A. **Beckman** are Professors at the University of Wisconsin, Madison.

such as fan-speed mode. The levels of the control variables change the power requirements for the components and thus for the system as a whole. Different control variable settings may produce equal building comfort but at differing power requirements.

An optimal control strategy is one in which, for a given building load and ambient condition, the total cost of operation is minimized while maintaining comfort conditions in the building. The total energy cost is the sum of the power required to operate the chillers, cooling tower fans, cooling tower water pumps, chilled-water pumps, and air-handling unit fans.

Equipment, sensor, and control component failures result in deviations from optimal control. Even though comfort may be maintained, the cost of operation will be higher than with optimal control. These failures can be detected by the increase in energy requirements compared with operation with optimal control. The methodology to detect these increases is the subject of this paper.

OPTIMAL CONTROL STRATEGY

The determination of the optimal control strategy for an HVAC system requires detailed knowledge of the performance characteristics of each of the components. Braun et al. (1989) have also developed a methodology for determining optimal control that utilizes component performance information obtained from measurements, manufacturer's data, or mechanistic models. This information is combined into a representation of the performance of the entire system. Determining the optimal strategy then involves nonlinear optimization techniques. The method yields the optimal set of control variables that minimizes power consumption at any time.

Using detailed performance data for each component is feasible, but the information requirements are large. Braun et al. (1989) also developed a methodology for determining a "near-optimal" control strategy based on the power consumption of the system as a whole rather than that of each component. The near-optimal strategy is simpler to determine and yields a strategy very close to the true optimal one.

The basis for this near-optimal strategy is the use of quadratic relations for power consumption. Braun (1988) showed that the power consumption of a chiller may adequately be represented as a quadratic function of the load and the temperature difference between leaving condenser and evaporator water temperatures. He also demonstrated that the power of the continuously adjustable pumps and fans may be accurately represented with a quadratic function of the control variables and flow rates.

This concept was then extended to the system as a whole. The power consumption of the entire HVAC system is represented by a quadratic relation for the total power in terms of the control variables, loads, and ambient conditions. The approach is used as the basis for the fault-detection methodology and is summarized below.

The total system power is represented by the following equation:

$$P(f,M,u) = u^T A u + b u + f^T C f + d f + f^T E u + g \qquad (1)$$

where

P = total system power

u = a vector of continuous control variables (e.g., temperature setpoints)

f = a vector of uncontrolled variables (e.g., loads, ambient temperatures)

M = a vector of discrete control variables (e.g., fan-speed mode)

A, C, E = coefficient matrices

b, d = coefficient vectors

g = a scalar.

The coefficients have to be determined empirically for each of the operating modes, and a separate formula is then developed for every feasible combination of discrete variables. The equations are constrained in that criteria such as temperature and humidity limits in the occupied space are met.

The optimal control is determined by equating the first derivative of the power with respect to each control variable to zero. Solving for the optimal values of continuous control variables yields linear relations between the control variables and the forcing functions:

$$u = -\frac{1}{2} A^{-1} b - \frac{1}{2} A^{-1} E f . \qquad (2)$$

If the system involves discrete variables (modes), the optimal control for each feasible combination of discrete variables has to be determined. The combination that yields the lowest system power is then the optimal operating mode.

The system shown in Figure 1 has two control variables: the chilled-water set temperature and the supply air temperature. The total power is the sum of that for the chiller, cooling tower and air-handler fans, and the chilled-water and condenser pumps. The significant uncontrolled variables are the wet-bulb temperature, cooling load, and sensible heat ratio. The ambient dry-bulb has essentially no influence (Braun 1988). The system was modeled and simulated using the TRNSYS (Klein et al. 1988) program.

An example of the total power as a function of the two control variables is shown in Figure 2. The load and ambient conditions are fixed, and each curve represents one chilled-water set temperature. For this operating condition, there exists a minimum power consumption for control settings of about 44°F (7°C) for the chilled water and 51°F (10°C) for the supply air temperature. These values will be different for other loads and ambient conditions.

The system's power consumption over a wide range of operating conditions and control variables was determined. These power values were then fit to the quadratic form of

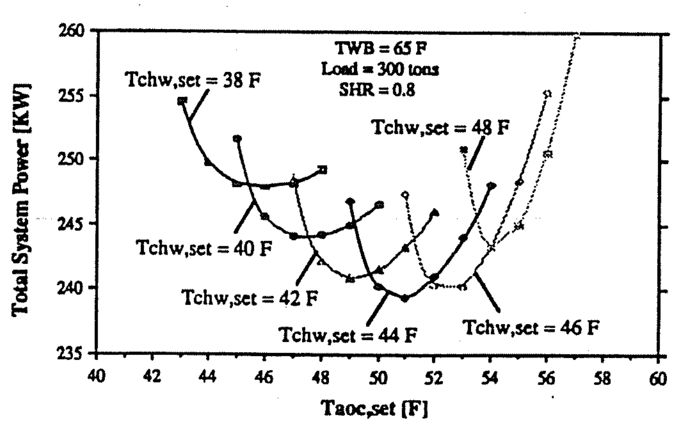


Figure 2 Total system power as a function of the set-point temperatures

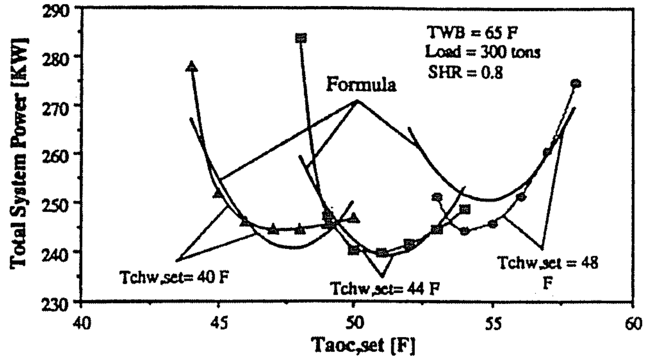


Figure 3 *Power from simulation and formula for a medium cooling load*

Equation 1 using linear least-squares regression techniques to yield the values of the coefficients. In this example system, only one set of discrete control variables (modes) exists, and thus only one formula for system power was developed. Equation 2 was solved for the optimal values of the chilled-water temperature and the air outlet temperature. The resulting equations for the two setpoint temperatures are linear relations:

$$T_{aoc,set,opt} = a_1\ T_{wb} + a_2\ Load + a_3\ \mathrm{SHR}$$
$$T_{chw,set,opt} = b_1\ T_w + b_2\ Load + b_3\ \mathrm{SHR} \qquad (3)$$

where the coefficients a_i and b_i are determined from the coefficients of the quadratic relation following Equation 2.

The quadratic relation quite accurately determines the optimal control strategy. In Figure 3, the system's power consumption from the simulation and that from Equation 1 is presented for the same conditions as for Figure 2. For each chilled-water set temperature, the optimum air temperature as predicted by Equation 1 is in good agreement with the detailed simulation results. However, as also shown in Figure 3, the energy computed from the quadratic relationship does not match the detailed simulation results exactly. For some situations, the equation predicts lower power than the simulation, and for other situations it predicts higher. This difference between simulation and equation values will be a factor in the use of the optimal control results to determine faults.

FAULT DETECTION

The term "fault" used in this study is a qualitative expression used for changes that cause nonoptimal system behavior. An error is the quantitative value of the fault. For example, a fault could occur in the chilled-water temperature sensor. The fault could be quantitatively described by saying that the error in the chilled-water temperature is 5°F (2.8°C).

The methodology for fault detection proposed in this study is based on the use of a relation for the system power consumption under the optimal control strategy. Operating variables, such as loads and ambient conditions, change over time, and equipment performance may degrade over time. Historical records may not be a sufficient reference, since they will not give a value of the power consumption for every future operating condition. The power relation for optimal control provides a method for estimating what the power consumption should be under the current operating conditions, whether it has previously been experienced or not.

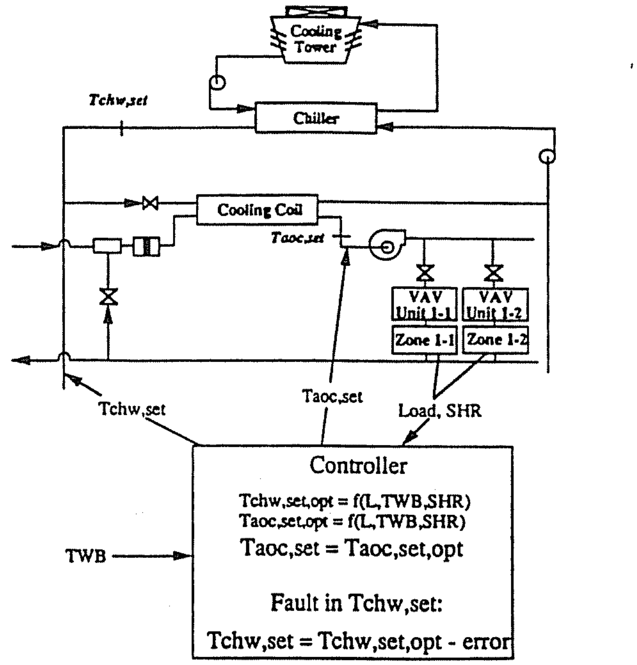


Figure 4 *Introduction of a fault in the simulation of the system*

Figure 4 demonstrates by example how faults are introduced into the simulation model. The supervisory controller bases the settings of the control variables on measures of the load, the ambient wet-bulb temperature, and the sensible heat ratio. The optimal control setpoints for chilled-water temperature and supply air temperature are then determined using the optimal strategy. An error is then introduced so that the actual setting differs from the optimal setting.

To illustrate the fault detection methodology, an error in the chilled-water set temperature is introduced. This represents a situation in which the temperature sensor is not operating properly, and the temperature indicated by a sensor deviates from the actual temperature by the amount of the error. This is modeled by introducing the error into the set temperature output from the controller. The system is then simulated with the nonoptimal value of the chilled-water temperature and the optimal value of the supply air temperature. With nonoptimal setpoints, more power will be required to meet the same load at the same comfort conditions.

The power evaluated from Equation 1 is used to predict the power for the current set of forcing functions and control variables and is called the predicted value. The power that is output from the simulation is called the measured value. The difference between the measured and the predicted power is the basis for fault detection.

The quadratic formula (Equation 1) is the best fit to the data. The predicted and measured power data are not identical for the same set of forcing functions even with no error present. In Figure 5, the residuals for 50 randomly selected sets of forcing functions are presented. The residuals have values near zero, but are rarely zero, and the mean value is approximately zero.

The determination of faults is based on whether the differences between the predicted and measured values are statistically significant. In order to perform this test, the standard deviation for the measurements without error about the predicted relation is needed. This was done by deter-

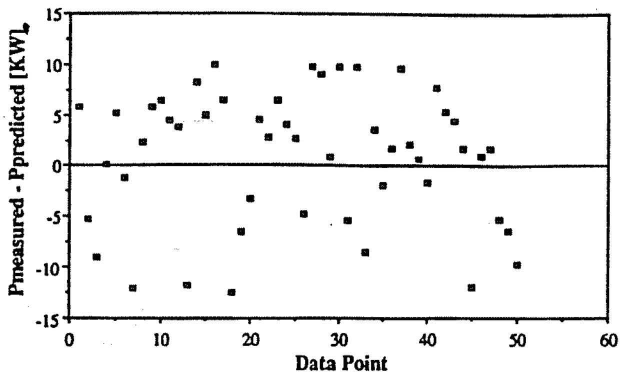


Figure 5 *Residuals for 50 randomly selected points*

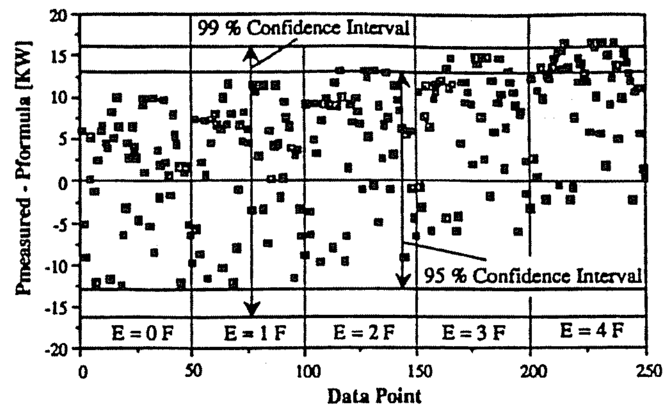


Figure 6 *95% and 99% confidence intervals for residuals with bias error*

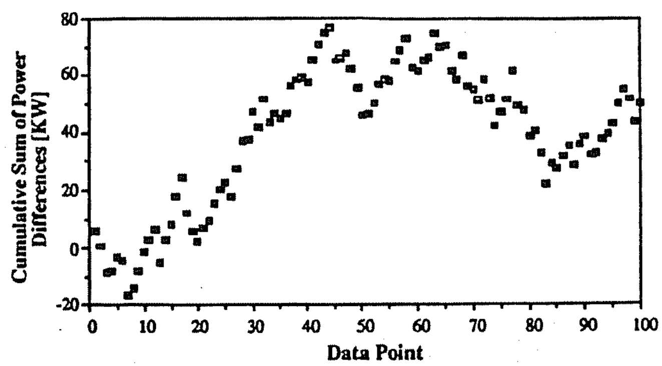


Figure 7 *Cumulative sum of power differences for 100 random data points without error*

mining the standard deviation of the differences in predicted and measured power using all of the simulation points.

Several methods were then investigated for their ability to detect faults. The first method considers the deviation of a single measurement from the optimal value. The second method uses a cumulative sum of differences. The third method considers a sequence of measurements. The method for locating faults uses measurements of the power from individual components. These approaches are described in the following sections.

SINGLE-MEASUREMENT APPROACH

In the first method, single measurements are used. For a fault due to error in the sensor of the chilled-water temperature to be detected, the measured power value must be significantly higher than the predicted value, i.e., the residual must be located outside the confidence interval. The confidence interval for a single measurement of the system power can be calculated as follows:

$$P_{pred} \pm t_{\nu,\alpha/2} \frac{s}{\sqrt{n}} \tag{4}$$

where

P_{pred} = power predicted from the formula
$t_{\nu,\alpha/2}$ = tabulated t-value for degrees of freedom and a tail area of $\alpha/2$ (two-sided) (Box et al. 1979)
ν = degrees of freedom for evaluating the standard deviation = n_{tot} - 1
$\alpha/2$ = two-sided tail area
s = standard deviation for the predicted power about the regression equation
n = number of observations for the predicted value = 1.

There is only one formula for the predicted value and only one prediction for an observation, and thus n is equal to one. Many data were employed for developing the regression formula, and thus the number of degrees of freedom is essentially infinite.

The performance of the system for the same 50 conditions as shown in Figure 5 was determined with constant bias errors in the chilled-water sensor of 1°F, 2°F, 3°F, and 4°F. The difference between predicted and measured power for these situations, along with the differences for no error, are shown in Figure 6. The 95% and 99% confidence intervals for the data are also shown.

For errors of 1°F and 2°F, the increases in power lie within the 95% confidence interval. Even though the mean of the data has increased, it is not statistically certain that this is due to the presence of an error. For an error of 3°F, all of the residuals lie within the 95% confidence limits even though the mean of the residuals is about 3% of the total power. Thus a 3% greater power consumption than optimal would not be indicated using only single measurements.

It is only when the error is 4°F that many of the residuals lie outside the 95% confidence interval. Thus, using only a single measurement, a fault can be detected only if a relatively large error in the chilled-water temperature occurs.

SEQUENCE OF MEASUREMENT APPROACHES

The second method is based on using a sequence of power measurements. During operation of the air-conditioning system, continuous measurements are made and stored by the building energy management and control system (EMCS). In practice, the EMCS could produce sums of the differences between predicted and measured power.

One aspect of this second approach is to use a cumulative sum of all power differences. The cumulative sum is shown in Figure 7 for 100 randomly selected data points without error as a function of the ith data point; in an actual system, the sequence of data points would occur over time. The cumulative sum starts out at zero but deviates thereafter. If the sum of the residuals for all of the data that were

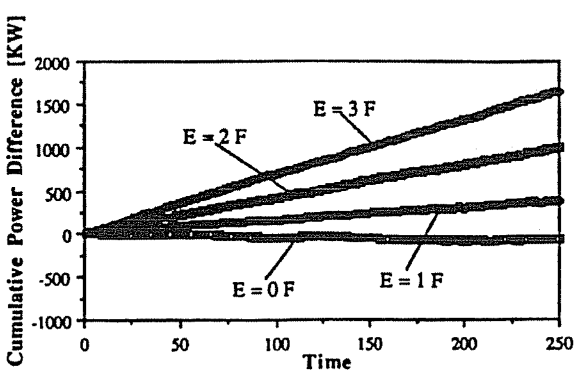


Figure 8 *Cumulative sum of residuals for errors in* $T_{chw,set}$ *of 0°F to 3°F*

employed for creating the regression formula is taken, the cumulative sum would be zero. Data sets that do not include all data yield cumulative sums that are not zero. In the case shown, the cumulative sum of residuals is approximately 50 W for 100 data points but reaches nearly 80 W around data point 40.

The sequence of measurements gives a qualitative indication that a system fault has occurred if the slope of the line exceeds a certain value. The seriousness of the fault in terms of increased energy use can be estimated. If the slope increases for a short time period and then decreases again, as in the first part of Figure 7, there is probably no error or the error is small.

In Figure 8, cumulative power differences are presented over a range of 250 time steps. Four different curves for errors in chilled-water temperature of 0°F to 3°F are shown. The 250 data points for each of the curves are randomly chosen and are different for each of the curves. A clear difference between the lines can be recognized. Thus, over a long time period, it is possible to detect a significant difference for small errors.

In the third method, an average of a sequence of several data points is used to make statistically significant comparisons. Two sets of data are needed, with the first representing data from operation without error, which is used as the reference. The reference set was 50 data points selected randomly from the entire set. Residuals for data taken during system operation when errors may be present are compared with this reference set.

The t-test is used to make the comparisons. The term $\bar{y}_A$ is average power difference (P_{meas} - P_{pred}) in the first sequence without error and is defined as

$$\bar{y}_A = \frac{\sum_i (P_{meas,i} - P_{pred,i})}{n_A} \quad . \tag{5}$$

The term $\bar{y}_B$ is the average power difference in the second sequence with error and is defined as

$$\bar{y}_B = \frac{\sum_i (P_{meas,i} - P_{pred,i})}{n_B} \quad . \tag{6}$$

A significant difference between sequence A (sequence without error) and sequence B (with error) is detected if

$$t = \frac{\bar{y}_B - \bar{y}_A}{s_p \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} > t_{\nu,\alpha} \tag{7}$$

where

$$s_p = \frac{\nu_A s_A + \nu_B s_B}{\nu_A + \nu_B} \tag{8}$$

s_p = pooled variance as given by Equation 8
n_A = number of data points with treatment A
n_B = number of data points with treatment B
$t_{\nu,\alpha}$ = t-value for ν degrees of freedom and an area of α
ν = total number of degrees of freedom = $n_A + n_B$ - 2
α = tail area probability
ν_A = degrees of freedom for $A = n_A - 1$
ν_B = degrees of freedom for $B = n_B - 1$
s_A = standard deviation for treatment A
s_B = standard deviation for treatment B.

A pooled variance is calculated because it is assumed that the system powers for both sequences have the same population variance. Therefore, all available data are utilized to estimate the variance.

The method is illustrated in Figure 9. The average residual for the reference sequence A, $\bar{y}_A$, is zero because the residuals from the regression are utilized. During the first 50 data points, no error is present, while from data point 51 to 100, a bias error of 2°F in the chilled-water temperature is present. T-tests are performed as every new data point is added to the sequence. Two of the tests are indicated in the figure. One of the tests was performed at point 55 for points 45 to 55 and the other at point 86 for points 76 to 86.

In the first of the tests shown at data point 55, five measurements without error and five measurements with error are included in the test. The average residual, $\bar{y}_B$, is close to zero and the fault is not detected using the t-test. In the second test, shown at time 86, the average of the 10 last residuals, $\bar{y}_B$, is well above zero and the fault is detected using the t-test.

The calculated t-value increases with the number of data points, n_B, included in the test. Therefore, the choice of n_B is important. A higher value of n_B increases the t-

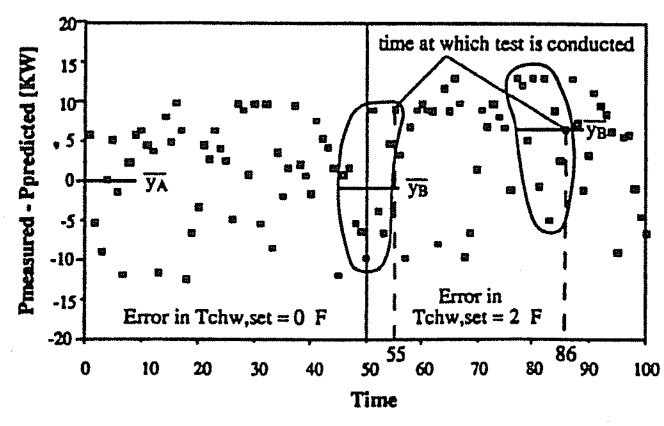


Figure 9 *Residuals without error and with an error in* $T_{chw,set}$ *of 2°F*

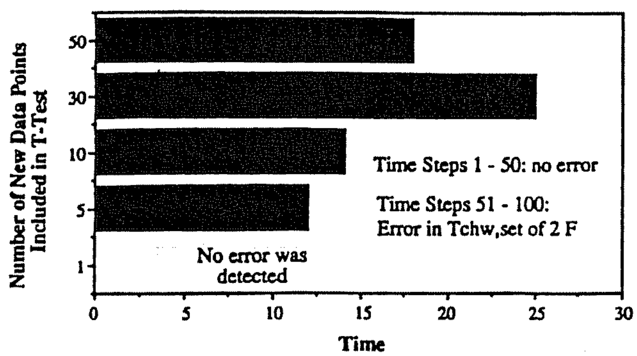


Figure 10 *First time an error in $T_{chw,set}$ of 2°F was detected for different numbers of data points included in the test*

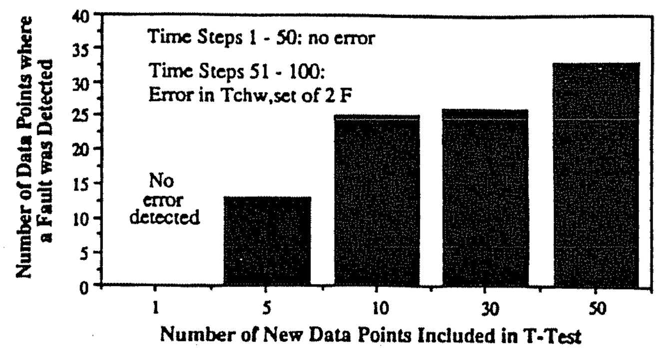


Figure 11 *Number of times a fault was detected in a 50-hour time interval with an error in $T_{chw,set}$ for different numbers of data points included in the test*

value even if the average residual ($\overline{y}_B$) is small. Thus, even small errors can be detected if many data points are included in the test. On the other hand, a high average residual ($\overline{y}_B$) is necessary to detect an error if n_B is small.

The first time a residual representing a fault is part of a t-test that includes many data points, it will not have a large influence on the average residual even if the error is relatively large. Only when the fault is present over several data points does it have a large enough influence to be detected. Hence, including less data in the test will increase the speed at which a fault is detected.

A trade-off between speed of detection and sensitivity of detection exists. In a test run, 50 sets of forcing functions were input to a simulation, first without error (data points 1 through 50) and then with an error of 2°F in the chilled-water temperature (data points 51 through 100). The total system power for the data was obtained.

During the first 50 data points, representing operation without error, no fault was indicated for any of the tests. In Figure 10, the first data point at which a fault was detected is shown for different numbers of data included in the test. An indication of a fault at one data point does not necessarily mean that a fault was also indicated during all further time steps.

As shown in Figure 10, no fault was detected at any of the 50 time steps of operation with error when only one data point was included in the test. By including five or more data points in the test, a fault was detected at least once during the 50 data points of operation with a fault. In general, the time before a fault is detected generally increases as the number of data points included in the test increases.

If a fault was indicated once and after that no further indication of a fault was given, it is possible that the fault is not real. On the other hand, if a fault is indicated at every time step, it is relatively certain that a fault occurred. In Figure 11, the number of data points for which a fault is indicated out of the 50 points with an error in $T_{chw,set}$ is shown for different numbers of data points included in the t-test. For low numbers of data points, the fault is detected only a few times. When 50 data points are included in the test, the fault is detected for every data point after the first detection.

Due to the trade-off between speed in the detection of faults and the ability to detect small faults and the certainty of having detected a fault, t-tests should be carried out with different numbers of points included in the test. By doing so, large and serious errors can be detected quickly and small errors, representing only small energy losses, can be recognized after a longer time period.

LOCATION OF A FAULT IN THE SYSTEM

The single formula for total system power allows determination of faults in the system as a whole. Individual relations for each power-consuming component allow determination of the location of faults. Through the combination of fault messages for the single component powers, the fault can be located in the system.

The component powers are determined from the simulation program similar to the measurements in an actual system. After collecting data under near-optimal conditions without error, formulas are determined for each component's power consumption. In this HVAC system, quadratic formulas for the air-handling unit fan power, the chiller power, and the main water loop pump power are determined using linear regression techniques. The power necessary to operate the cooling tower fans and pump is constant for all conditions.

When a fault is present, it may affect the power consumption of each component. The residuals could indicate a significantly higher power consumption, a significantly lower power consumption, or no significant difference between the measured and the predicted power, depending on the location of the fault. To locate the fault, the statistical tests are performed for the component power residuals following the procedure for the entire system as described earlier.

In the algorithm that was developed, a flag was used to signify the presence of a fault. If a significantly higher power consumption was detected, a value of unity was assigned to the fault indicator, while a value of -1 was assigned if a significantly lower power consumption was detected. A fault indicator value of zero indicated that no significantly higher or lower power occurred for that component. Fault messages from the single component powers are then combined to locate the fault.

As an example of such a test, a fault representing a positive error in the chilled-water temperature sensor was investigated. In Figures 12, 13, and 14, the cumulative sum of component power residuals is presented for the chiller, pump, and fan, respectively. A positive error in the chilled-water temperature, i.e., a lower chilled-water temperature, reduces the efficiency of the chiller. Thus, the chiller requires more power. Because the chilled-water temperature is lower than for optimal conditions, the pump does not have to provide as much flow to obtain the same supply air temperature as for the optimal control. Hence, the pump power requirement is lower than for optimal conditions.

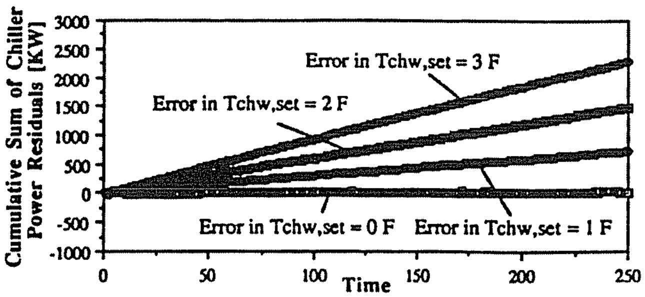


Figure 12 *Cumulative sum of chiller power residuals for errors in* $T_{chw,set}$

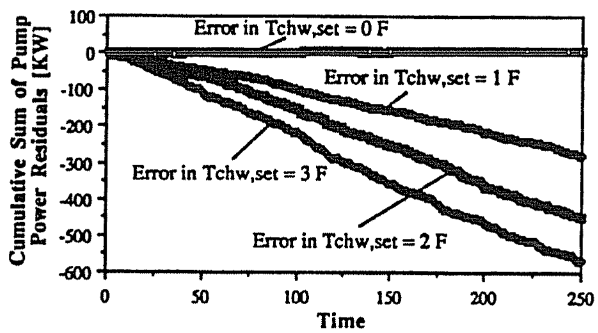


Figure 13 *Cumulative sum of pump power residuals for errors in* $T_{chw,set}$

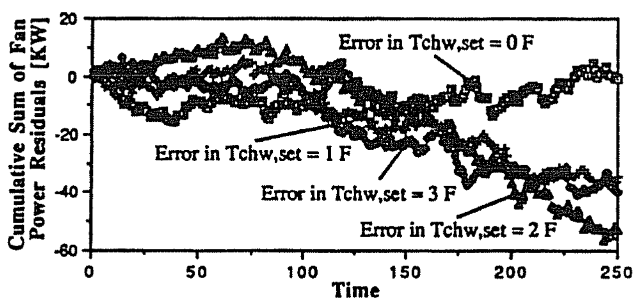


Figure 14 *Cumulative sum of fan power residuals for errors in* $T_{chw,set}$

The supply air temperature remains at its optimal value, and, therefore, no difference in the optimal behavior is indicated for the air-handling unit fan power.

The lines for the cumulative sum of pump power residuals in Figure 13 have a negative slope, while the lines for the cumulative sum of chiller power residuals have a positive slope. The t-tests applied to these residuals indicate a statistically significant error for errors of 1°F to 3°F. The lines for the air handler power have no regular pattern, and no significant difference in the residuals can be found.

For this example, the fault is readily determined to be in the chilled-water temperature. The lack of a change in the fan power from optimal indicates that the supply air temperature is correct and the fan is operating under optimal conditions. The lower-than-expected pump power indicates that the chilled-water temperature is lower than expected under optimal conditions. Thus, the fault is determined to be an error in the chilled-water sensor.

CONCLUSION

It is feasible to use an optimal control methodology in fault detection. The optimal control strategy yields a relation for predicting the system power under optimal conditions as a function of the control variables, the loads, and the ambient conditions. Increases in power consumption above that for optimal control indicate the presence of faults in the system. Statistical tests establish the confidence with which faults may be detected.

The method may be extended to locate faults in a system. The power under optimal control for each of the components is determined in a manner similar to that for the system as a whole. Statistically significant deviations in power from the optimal values are detected for some components when a fault is present. An examination of the physical relations between the components establishes the location of the faults.

REFERENCES

Anderson, D., L. Graves, W. Reinert, J.F. Kreider, J. Dow, and H. Wubbena. 1989. "A quasi-realtime expert system for commercial building HVAC diagnosis." *ASHRAE Transactions*, Vol. 95, Part 2.

Box, G.E.P., W.G. Hunter, and J.S. Hunter. 1979. *Statistics for experimenters, an introduction to design, data analysis, and model building*. New York: John Wiley & Sons.

Braun, J.E. 1988. "Methodologies for the design and control of central cooling plants." Ph.D. thesis, University of Wisconsin-Madison.

Braun, J.E., J.W. Mitchell, S.A. Klein, and W.A. Beckman. 1987. "Performance and control characteristics of a large central cooling system." *ASHRAE Transactions*, Vol. 93, Part 1.

Braun, J.E., S.A. Klein, J.W. Mitchell, and W.A. Beckman. 1989. "Applications of optimal control to chilled water systems without storage." *ASHRAE Transactions*, Vol. 95, Part 1.

Kao, J.Y., and E.T. Pierce. 1983. "Sensor errors; their effect on building energy consumption." *ASHRAE Journal*, Vol. 25, No. 12, p. 42.

Pape, F.L.F. 1989. "Optimal control and fault detection in heating, ventilating and air-conditioning systems." M.S. thesis, University of Wisconsin-Madison.

Urso, P.B., and I.C. Schick. 1985. "A hierarchical approach to HVAC system fault detection and identification." ASME DSC-1, p. 285.

Urso, P.B., I.C. Schick, and S. Negahdaripour. 1985. "HVAC system fault detection and diagnosis." American Control Conference, Vol. 1.