Demonstration of Fault Detection and Diagnosis Methods for Air-Handling Units (ASHRAE 1020-RP)

L.K. Norford, Ph.D. Member ASHRAE
J.A. Wright, Ph.D., C.Eng Member ASHRAE
R.A. Buswell
D. Luo, Ph.D. Student Member ASHRAE
C.J. Klaassen Member ASHRAE
A. Suby Member ASHRAE

Results are presented from controlled field tests of two methods for detecting and diagnosing faults in HVAC equipment. The tests were conducted in a unique research building that featured two air-handling units serving matched sets of unoccupied rooms with adjustable internal loads. Tests were also conducted in the same building on a third air handler serving areas used for instruction and by building staff. One of the two fault detection and diagnosis (FDD) methods used first-principles-based models of system components. The data used by this approach were obtained from sensors typically installed for control purposes. The second method was based on semiempirical correlations of submetered electrical power with flow rates or process control signals.

Faults were introduced into the air-mixing, filter-coil, and fan sections of each of the three air-handling units. In the matched air-handling units, faults were implemented over three blind test periods (summer, winter, and spring operating conditions). In each test period, the precise timing of the implementation of the fault conditions was unknown to the researchers. The faults were, however, selected from an agreed set of conditions and magnitudes, established for each season. This was necessary to ensure that at least some magnitudes of the faults could be detected by the FDD methods during the limited test period. Six faults were used for a single summer test period involving the third air-handling unit. These fault conditions were completely unknown to the researchers and the test period was truly blind.

The two FDD methods were evaluated on the basis of their sensitivity, robustness, the number of sensors required, and ease of implementation. Both methods detected nearly all of the faults in the two matched air-handling units but fewer of the unknown faults in the third air-handling unit. Fault diagnosis was more difficult than detection. The first-principles-based method misdiagnosed several faults. The electrical power correlation method demonstrated greater success in diagnosis, although the limited number of faults addressed in the tests contributed to this success. The first-principles-based models require a larger number of sensors than the electrical power correlation models, although the latter method requires power meters that are not typically installed. The first-principles-based models require training data for each subsystem model to tune the respective parameters so that the model predictions more precisely represent the target system. This is obtained by an open-loop test procedure. The electrical power correlation method uses polynomial models generated from data collected from "normal" system operation, under closed-loop control.

Leslie K. Norford is an associate professor at the Massachusetts Institute of Technology, Cambridge, MA. **Jonathan A. Wright** is a senior lecturer and **Richard A. Buswell** is a research associate with Loughborough University, Department of Civil and Building Engineering, Loughborough, U.K. **Dong Luo** is a senior engineer with United Technologies Corporation. **Curtis J. Klaassen** is manager of Iowa Energy Center's Energy Resource Station. **Andy Suby** is project engineer at the Center for Sustainable Environmental Technologies, Iowa State University.

Both methods were found to require further work in three principal areas: to reduce the number of parameters to be identified; to assess the impact of less expensive or fewer sensors; and to further automate their implementation. The first-principles-based models also require further work to improve the robustness of predictions.

INTRODUCTION

In the last decade, a considerable amount of research has been carried out in the field of fault detection and diagnosis in HVAC systems. Hyvarinen and Karki (1996) summarized the efforts of an international collaboration [International Energy Agency (IEA) Annex 25] that listed typical faults in heating systems ranging from oil burners to district heating distribution systems; vapor-compression and absorption refrigeration machines; variable air volume (VAV) air-handling units (AHUs); and thermal storage systems. This work also produced a number of fault detection and diagnosis (FDD) methods:

Innovation approaches
- Physical models
- Time-series models
- State-estimation methods

Parameter-estimation approaches
- Methods based on physical models
- Characteristic curves
- Characteristic parameters

Classification approaches

Topological case-based modeling
- Artificial neural networks

Expert-system approaches
- Rule-based methods
- Associative networks

Qualitative approaches
- Formal qualitative approaches
- Fuzzy models

Many participants in this effort described their own methods; those of a generic nature or that focused on AHUs include Dexter and Benouarets (1996), Haves et al. (1996), Lee et al. (1996a, 1996b), Salsbury (1996), Yoshida et al. (1996), and Peitsman and Bakker (1996). These methods were developed and tested with simulations and laboratory test rigs where a high degree of experimental control can be applied. Such issues as interfaces with commercially available control systems, identification of the intended users and their needs, and methods for testing and evaluating the performance of FDD systems in systems installed in real buildings were not addressed. IEA Annex 34 followed Annex 25 and focused on the practical application of FDD techniques in real buildings (Dexter and Pakanen 2001).

ASHRAE sponsored the research described in this paper as a contribution to the global effort to demonstrate FDD methods in real buildings. This research focused on demonstrating FDD methods applied to systems installed in real buildings, encompassing the FDD methods, results from the trials, and the evaluation of the FDD method performance. The three objectives of the research were to

1. Demonstrate the operation of FDD methods for HVAC systems in a realistic building environment

2. Compare the performance of different FDD methods for different types of faults in AHUs and to assess the costs of their implementation
3. Archive and document the test data so they can be used to develop and test other FDD methods

Each of these objectives carried equal weight and each was largely met. This paper reports the completed work with respect to the first two objectives. A more substantial account and archived data are available in the final report to ASHRAE (Norford et al. 2000).

This research considered VAV AHUs only, although the methods can be applied to other types of HVAC systems. Both methods were based on a reference-model approach, where measurements from the system are compared to model predictions. A significant difference between the model predictions and the observations indicates that the system has deviated from the expected operating condition, which is taken to indicate the presence of a fault. The first-principles-based approach modeled the subsystem components and used fluid (air/water) quantity and property measurements and control-signal observations. The detection of faults focused on the effect of the fault condition on the output of the system process. The electrical power correlation method related electrical power measurements to fluid quantity and property measurements and control-signal observations; changes in the correlations were considered to be faults.

A description is given of the test building and the HVAC systems. The faults, their implementation and the FDD methods are described. The results from the four test periods are presented. The methods are evaluated on the basis of the accuracy, calibration, and cost of the required sensors; the ease of implementation of the methods, including selection and estimation of the model parameters and thresholds; the sensitivity and robustness of the methods in fault detection; and their success in fault diagnosis.

BUILDING, SYSTEMS, AND FAULTS

The fault-test program was conducted in a unique building that combined laboratory testing capability with real building characteristics and was capable of simultaneously testing two full-scale, commercial building systems side-by-side with identical thermal loads. The building was equipped with three VAV AHUs. Two were nominally identical (AHU-A and AHU-B), each serving four test rooms (Figure 1). The building had a true north-south solar alignment so that the pairs of test rooms had nearly identical exposures to the external thermal loads. The test rooms were unoccupied but were equipped with two-stage electric baseboard heaters to simulate thermal loads and with two-stage room lighting, both scheduled to represent various usage patterns. The third AHU (AHU-1) served the general areas of the facility including offices, reception space, a classroom, a computer center, a display room, service spaces, and a media center. A second classroom (not shown in Figure 1) was added to the east side of the building during the later stages of this project. AHU-1 was subject to variable occupant, lighting, external, and internal loads.

The test rooms, heating and cooling loops, and AHUs were well instrumented, including watt transducers for all components of interest. The A and B test rooms were individually controlled by a single commercial energy-management and control system (EMCS) and the general areas were controlled by a second EMCS. The building had a structural steel frame with internally insulated, pre-cast concrete panels, a flat roof, slab-on-grade flooring, and a floor area of 862 m^2 (9272 ft^2), including the new classroom. The east, south, and west test rooms each had 6.9 m^2 (74 ft^2) windows with double-layer, clear glass.

The heating plant consisted of a gas-fired boiler, circulation pumps, and the necessary control valves. Heating operation of the HVAC systems was not required as part of the tests conducted in this research, other than the preheating of the outside air during winter operation to simulate

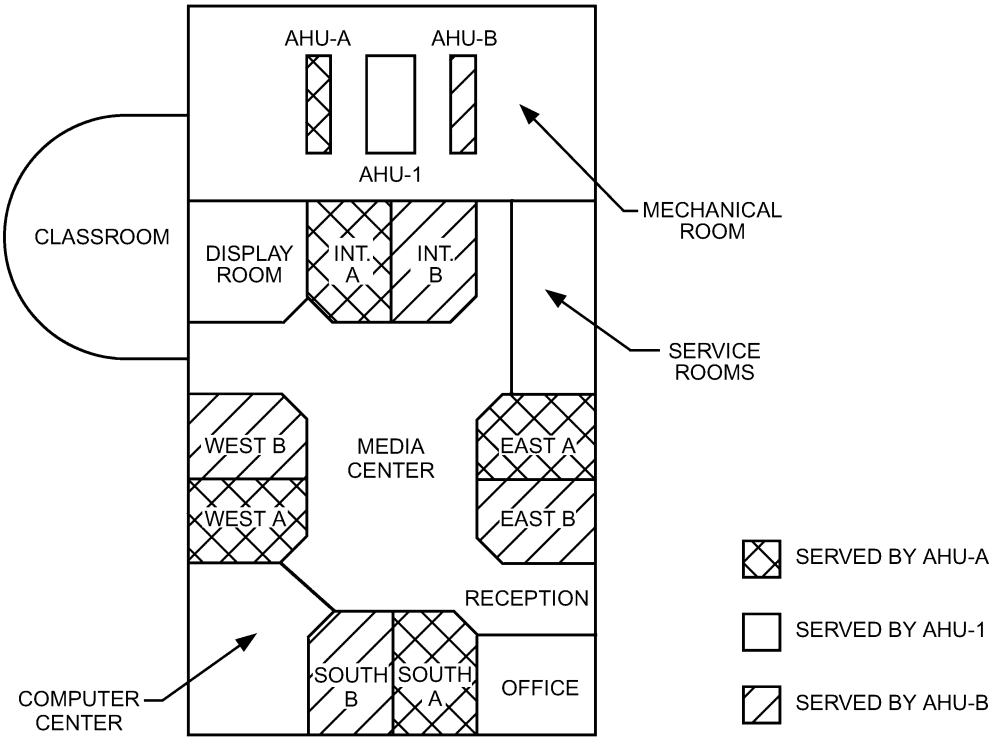

Figure 1. Plan of Test Building

higher outside temperatures and force the HVAC systems into economizer mode. The cooling plant (Figure 2), consisted of a nominal 35 $kW_{thermal}$ (10 ton) two-stage, reciprocating air-cooled chiller; a 525 $kWh_{thermal}$ (149 ton·h) thermal energy storage (TES) unit that was isolated from the cooling system for this research; and chilled water supplied by a central facility, with pumps, valves, and piping to circulate chilled water through the HVAC components.

The major components of the AHUs were the recirculated air, exhaust air, and outdoor air dampers; cooling and heating coils with control valves; and the supply and return fans (Figure 3). Ducts transferred the air to and from the conditioned spaces. Both the supply and return fans were controlled with variable-frequency drives. An additional heating coil was installed for this research on AHU-A and B, between the outside air (OA) inlet and the flow and temperature sensors. This coil was employed during the winter test period to preheat the outside air so as to force the control system into the free-cooling mode. AHU-A and B were identical, while AHU-1 was similar but larger to accommodate higher thermal loads. Air from the AHUs was supplied to VAV box units, each having electric or hydronic reheat.

The supply fan speed for all three AHUs was controlled to regulate supply duct static pressure. The AHU-A and AHU-B return fans were controlled to maintain a constant percentage of supply airflow; in AHU-1, the return fan control signal was a constant percentage of the supply fan control signal. The chilled water flow rate through the cooling coils in AHU-A and AHU-B was controlled by three-port mixing valves in a diverting application. A two-port valve was used in AHU-1.

The requirements for this project stipulated that a minimum of six faults be investigated, with at least two degradation faults and at least one fault in each of the three AHU subsystems (air-

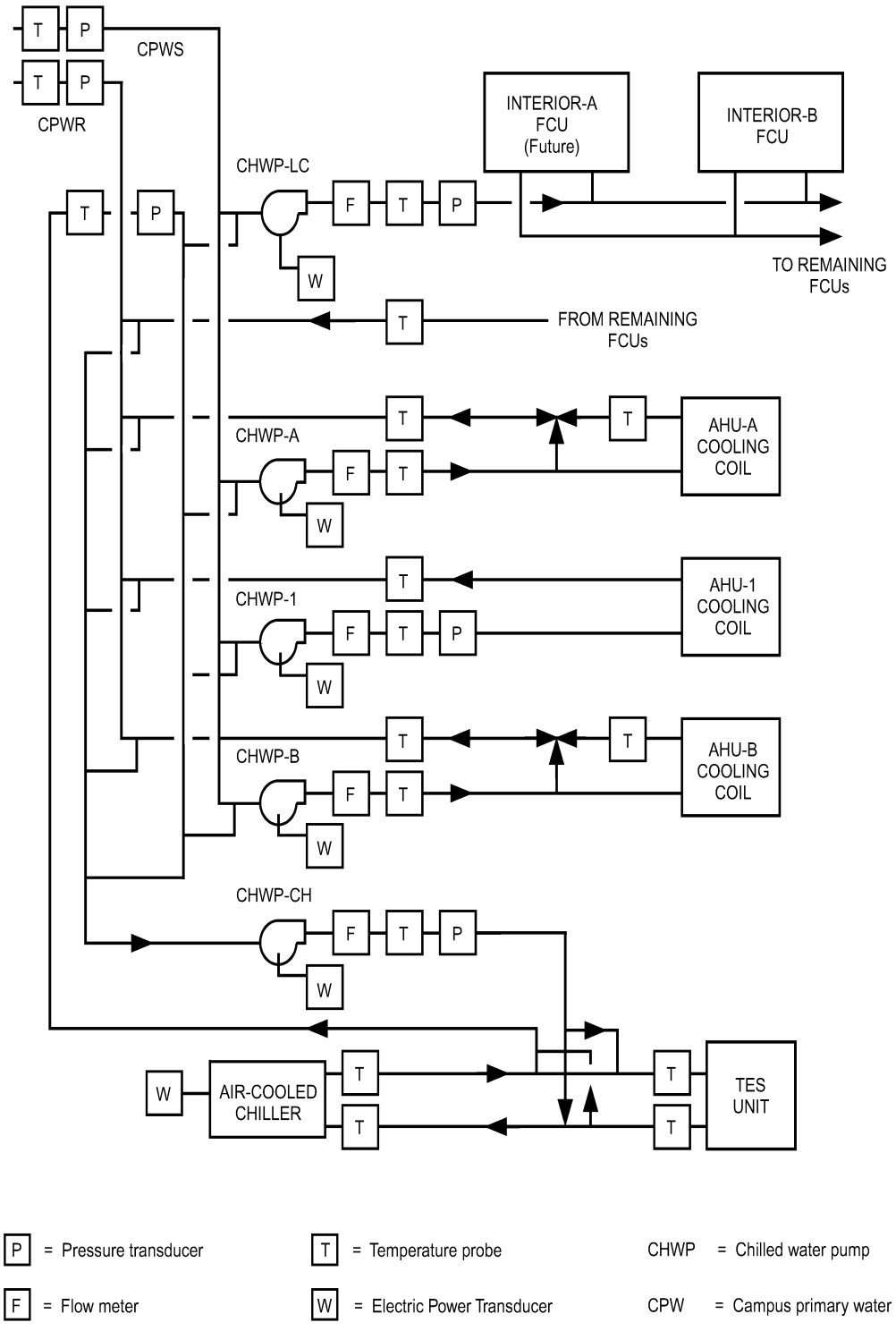

Figure 2. Chilled Water Flow Circuit in Test Building

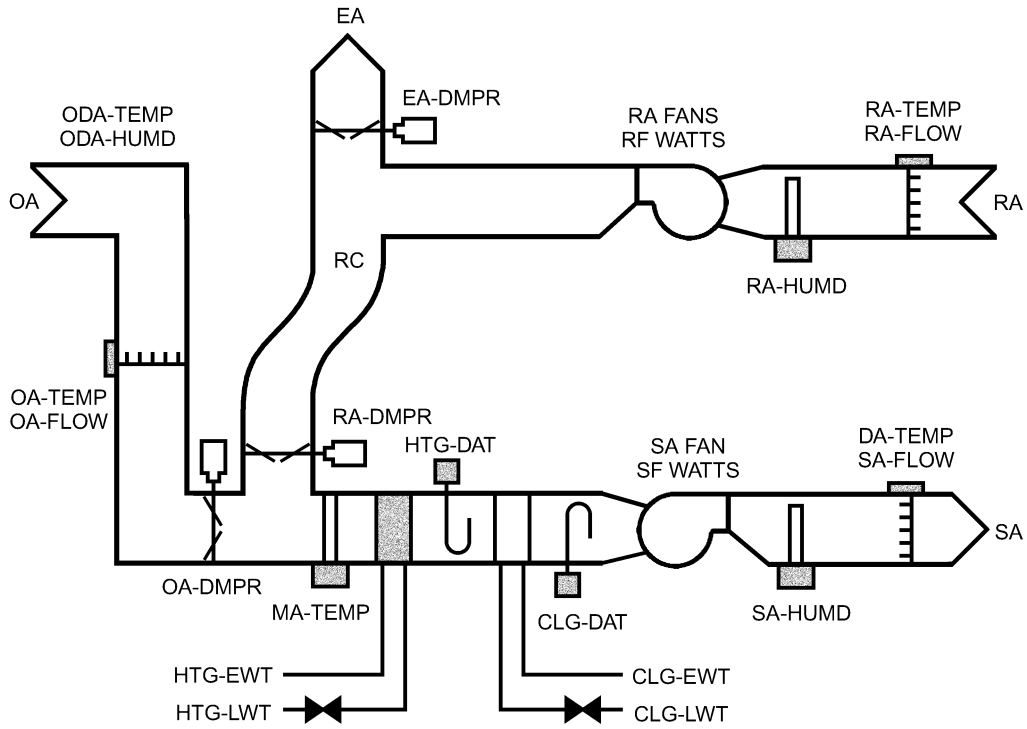

Figure 3. Air-Handling Unit in Test Building

mixing, filter-coil, and fan). Table 1 shows the seven selected faults and their method of implementation for AHU-A and AHU-B. Although faults implemented through software or easily disconnected hardware (such as actuator linkages) were readily introduced, others required substantial system modification, including the installation of bypass piping and additional valves.

It was necessary, in the context of the research, to ensure that both FDD methods were (in principle) capable of detecting all of the faults. Clearly, little would be learned from a series of null results. This criterion eliminated some faults, such as temperature sensor faults that the electrical power method would have difficulty detecting. This criterion was relaxed for tests with AHU-1, discussed later.

Table 2 indicates that each fault was implemented in at least two of the three test periods held during summer, winter, and spring seasons. Each test period consisted of a week of controlled tests, when the research staff of the test building introduced faults known to the investigators, a short analysis period, and a week of blind tests. For each blind test period, the list of possible faults was made known, but not the order of implementation or whether they were implemented at all. The lists for each season excluded faults that would not be seen in that season. For example, the recirculation damper would normally be fully open in hot weather (minimum outside air) and a damper leak could not be detected. Abrupt faults were typically implemented over a 24 h period while most degradation faults required three days, one for each of three stages. The three stages of the drifting pressure sensor fault were introduced over a single day.

Fault magnitudes were established during an initial period when the FDD methods were commissioned and the procedures for introducing faults and the HVAC systems were developed. The magnitudes of the degradation faults were selected such that it was anticipated that the two FDD methods *would* be able to detect the largest level, *should* be able to detect the middle level

Table 1. Method of Implementation of Faults

Fault	Type	Implementation
Air-Mixing Section		
Stuck-closed recirculation damper	Abrupt	Application of a control voltage from an independent source to maintain the damper in the closed position.
Leaking recirculation damper	Degradation	Removal of the recirculation damper seals, with one seal removed for the first fault stage, two for the second, and all seals for the third stage.
Filter-Coil Section		
Leaking cooling coil valve	Degradation	Manual opening of a coil bypass valve.
Reduced coil capacity (water-side)	Degradation	Manual throttling of the cooling coil balancing valve, to 70%, 42%, and 27% of the maximum coil flow of 1.7 L/s (27.5 gpm) for the three fault stages.
Fan		
Drifting pressure sensor	Degradation	Introduction of a controlled leak in the pneumatic signal tube from the supply duct static pressure sensor to the transducer, to a maximum reduction of 225 Pa (0.9 in. of water).
Unstable supply fan controller	Abrupt	Introduction of alternative gains for the PID controller that adjusts fan speed to regulate static pressure.
Slipping supply fan belt	Degradation	Adjustment of fan belt tension to reduce maximum fan speed by 15% at 100% control signal for the first stage and 20% for the second stage. The third stage had an extremely loose belt with variable fan speed.

and *could possibly* detect the lowest level. Fault magnitudes were consistent in each of the three test periods. Constant-magnitude faults provided a firmer basis for evaluating the FDD methods and were implemented with less difficulty than the variable-magnitude faults (a change of fault magnitude over different occurrences at different times) that would likely occur in practice. HVAC system commissioning consisted primarily of sensor calibration and establishing standard system operating configurations; the latter was required because the research facility altered the systems between test periods to meet the needs of other research programs. The configuration setup, which proved to be a major task for the test-building staff, encompassed fan control algorithms, isolation of the thermal storage tank (which provided a thermal capacitance that interfered with analysis of chiller cycling periods), and operating schedules for both HVAC equipment and false loads in the test rooms.

A more realistic set of blind tests was conducted with AHU-1, the air handler serving areas of the building occupied by research staff and classroom visitors. Four days of normal operation for training FDD methods and 17 days for fault introduction were included in a summer period of about six weeks. Building loads were not controlled and four of the six faults (listed in Table 3) had not been implemented in AHU-A and AHU-B and were completely unknown to the investigators. This test period was considerably longer than each of the three test periods on the matched AHUs, and increased the possibility of naturally occurring faults.

Two of the faults in this program produced signatures different from the naturally occurring faults they were intended to represent: the leaky cooling coil valve and the coil capacity fault. The leaking valve was implemented with a specially installed bypass valve that generated the same thermal effect as a leakage past a closed control port, but changed the flow resistance. The coil capacity fault mimicked the impact of water-side fouling on heat transfer across the cooling coil to some extent. A simpler alternative to replacing the existing coil with an older coil with tubes fouled with calcium carbonate was to close a valve in the inlet leg to the coil, thus increas-

Table 2. Faults Introduced into AHU-A and AHU-B During Three Blind Test Periods

Fault	Summer	Winter	Spring
Air-Mixing Section			
Stuck-closed recirculation damper	X	X	
Leaking recirculation damper		X	X
Filter-Coil Section			
Leaking cooling coil valve		X	X
Reduced coil capacity (water-side)	X		X
Fan			
Drifting pressure sensor	X	X	X
Unstable supply fan controller	X	X	X
Slipping supply fan belt	X	X	

Table 3. Faults Introduced into AHU-1 During Blind Test Period and Their Method of Implementation

Fault	Type	Implementation
Air-Mixing Section		
Stuck-closed recirculation damper	Abrupt	Application of a control voltage from an independent source to maintain the damper in the closed position for about 24 h
Stuck-open outside air damper	Abrupt	Application of a control voltage from an independent source to maintain the damper in the open position for 24 h
Filter-Coil Section		
Leaking heating coil valve	Abrupt	Adjustment of output voltage to the heating coil valve, causing it to unseat and leak for about 29 h
Fouled cooling coil	Degradation	Blockage of the cooling coil with a curtain drawn from the bottom to cover 25%, 50%, and 75% of the 61 cm (24 in.) coil in the three fault stages
Fan		
Drifting pressure sensor	Degradation	Introduction of a controlled leak in the pneumatic signal tube from the supply duct static pressure sensor to the transducer, with pressure reduced by 50, 100, and 150 Pa in the three fault stages (0.2, 0.4, and 0.6 in. of water) and each stage implemented for at least 6 h
Loss of control of supply fan	Abrupt	Supply fan VFD isolated from EMCS and operated at a constant speed for about 23 h

ing the resistance to water flow. This change in flow resistance became the basis for its detection with the electrical power FDD method.

Daily data sets for normal and faulty operation were assembled by the test-building staff from logs made by the EMCSs and were posted for electronic transfer to the investigators' home sites.

DETECTION AND DIAGNOSIS

The two fault detection methods compare the differences between the observed system behavior and a reference model of the system operation. The approaches differ significantly in how the fault effects are observed. The first-principles-based method considers the performance of the monitored system in terms of the system output useful to the air-conditioning process. In this case, the model predicts the temperature of the air or the static air pressure at the outlet of the component. A fault can be described in these terms as a degradation in the expected system performance.

The electrical power correlation method uses models derived from the system characteristics that relate electrical load to certain variables. This method predicts the expected power consumption. In this case a fault can be described as a change in the expected system energy consumption.

Both methods can take advantage of certain fault characteristics and not of others. The first-principles-based approach will always detect a degradation in the performance of the thermo-fluid system (as long as it is significant), regardless of the cause. The electrical power correlation method will not detect a fault that affects performance but has no effect on the electrical load. It is, however, predisposed to generating an operating cost associated with the faulty behavior and it can, in principle, detect faults associated with motors and drive trains that the first-principles-based approach cannot detect.

Fault diagnosis is also based on the information available to each method. The sensors required to implement FDD are one principal difference between the two methods. The first-principles-based approach uses measurements typically installed for control (temperatures, humidities, flow rates, etc.). One disadvantage with this approach is that in general terms there can be less control over the quality of these measurements in any given installation. The electrical power correlation method uses sensors over and above those normally installed, but these are more focused for the intended application and are not as susceptible to poor installation and maintenance.

First-Principles-Based Models with Thermo-Fluid Measurements

First-principles-based (or analytical) models can be used as a reference for the "correct" or expected operation of a HVAC system. The approach used in this research relied on the sensors typically installed in most VAV systems for control. Three subsystem models [described more fully in Norford et al. (2000)] were used to implement the FDD scheme: a fan/duct model of the air system, an economizer model, and a model of the coiling-coil subsystem. Figure 4 demonstrates the modeling arrangement. The black dots indicate where the comparisons to the observations from the real system were made, and hence where the fault detection for each subsystem was focused. Simple, steady-state simulations of the subsystems are formed by the models, which are based on the following principles:

- The fan/duct model is based on the fan laws and simple quadratic expressions for the change in system resistance and predicts the supply air static pressure.
- The economizer model is based on the analytical representation of the mixed-air condition as a function of damper position and the inlet temperatures and humidities. The model also includes an actuator model.
- The cooling coil model is based on the SHR method effectiveness-NTU heat and mass transfer calculation method [similar to the ASHRAE 3-line method; a review of both methods is given by Stephan (1994)]. The subsystem model also includes fan temperature rise and models for the control valve and the actuator.
- The fan-temperature-rise model is a simple addition to the air temperature, linearly dependent on the fan-control signal.
- The valve model is based on a first-principles analysis of the water-circuit resistance with respect to the control valve. The model predicts the mass flow rate of water through the coil, a typically unmeasured variable.
- The actuator model is an analytical representation of the movement of the actuator in response to a control signal. This models the dead-bands at either end of the operating range and any hysteresis (slack in the linkage) that may be present in the system.

The models have parameters for which values must be estimated for a specific system. The parameters are designed, as far as possible, to represent some tangible system characteristic and give greater precision in prediction. An example of this is the actuator "low activation point"

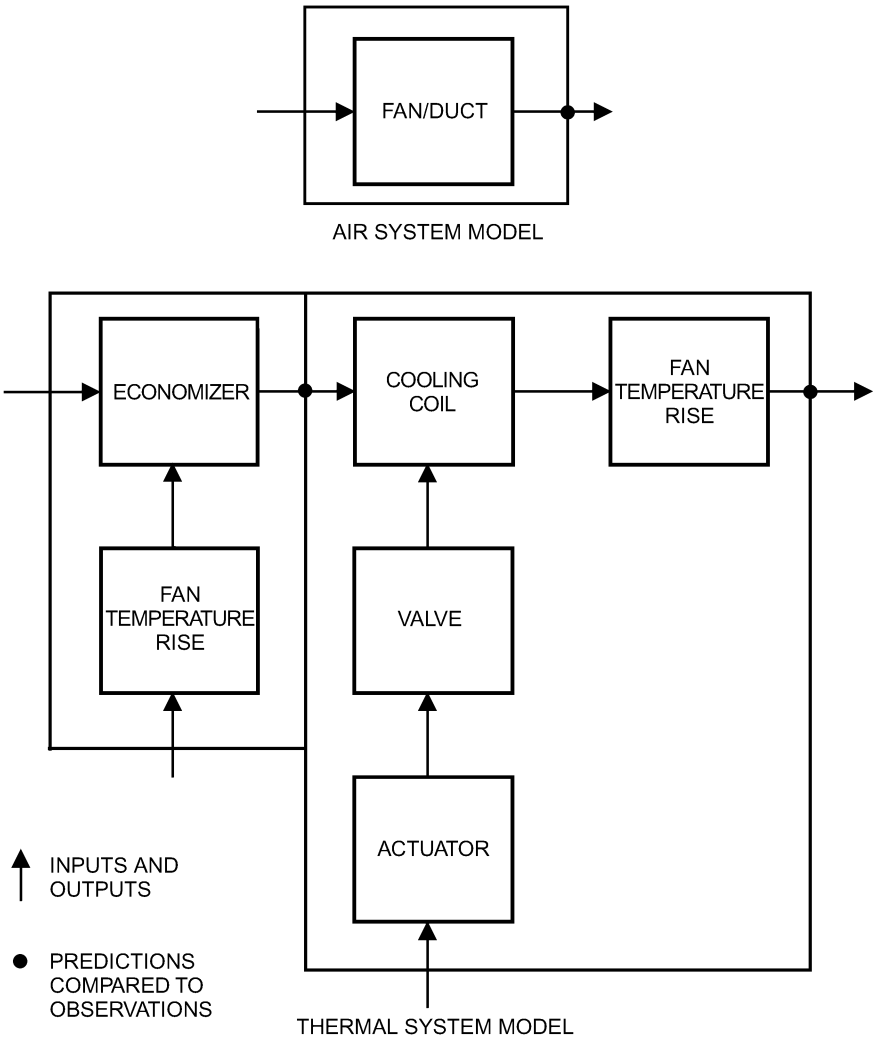

Figure 4. First-Principles-Based Model Functionality

parameter, which describes the value of the control signal required before the valve stem starts to move as it is opened from the closed position (i.e., dead-band). Some of the model parameters can be obtained from design information or inspection of the installed system. The face area and number of rows and circuits in the cooling coil are examples. The remaining parameters are identified simultaneously for each subsystem model. The model parameters are estimated by applying a nonlinear optimization technique, minimizing the model prediction errors in a least-squares sense. The data used for this procedure were generated by applying a sequence of open-loop "steps" in the inputs to capture the system characteristics when the system was in a normal (fault-free) condition.

With the system models characterized, model predictions can be applied to the observations to generate the "prediction error" (demonstrated in the top halves of Figures 5 and 6). The models only apply to observations that are close to steady state and a steady-state filter removes data containing transients. A lack of steady-state data was used to identify the presence of oscillatory, or unstable, control.

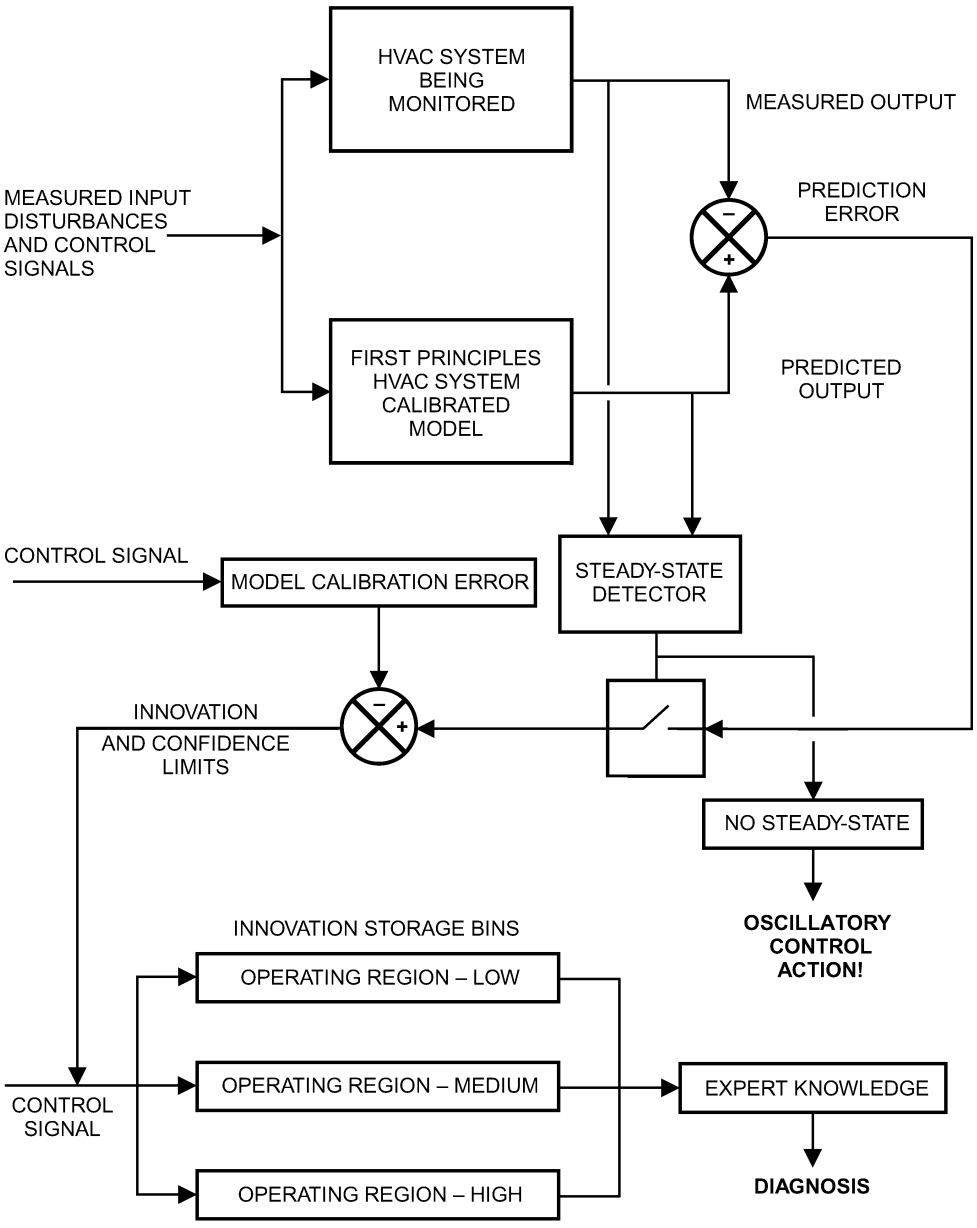

Figure 5. Method for Fault Detection and Fault Diagnosis Using Expert Rules

Some prediction error will always exist due to uncertainties in the measurements and unmodeled system disturbances. Statistically based thresholds are applied to the prediction error, such that a certain magnitude of error is required before triggering an alarm.

Once an alarm has been identified, the cause is diagnosed. Two methods were investigated in this work, fault diagnosis by (1) expert rules and by (2) recursive parameter estimation. The schemes are shown in Figures 5 and 6, respectively. Figure 5 shows that the "innovations" (the magnitude of the prediction error over and above the thresholds) were split into three "bins."

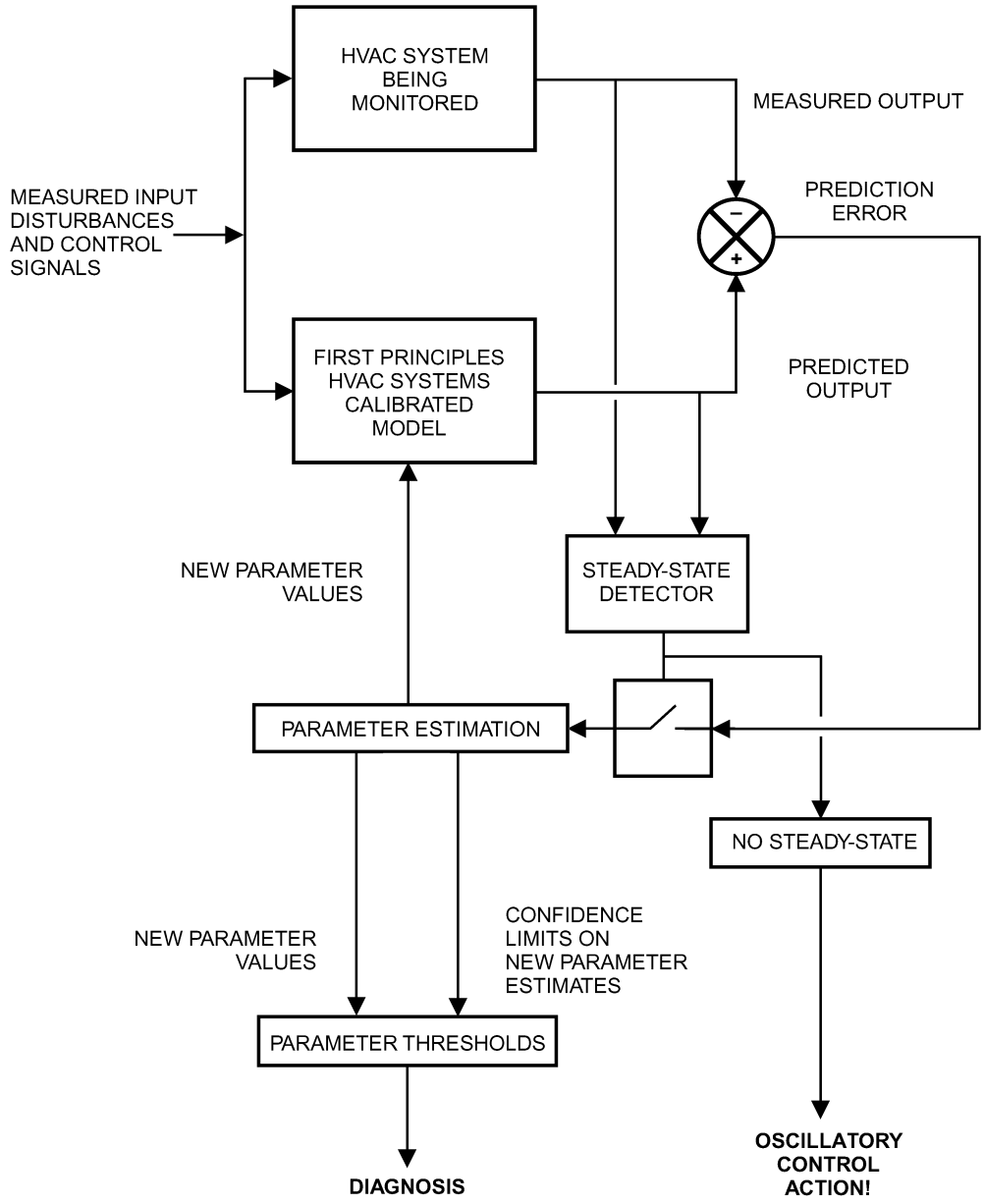

Figure 6. Method for Fault Diagnosis by Recursive Parameter Estimation

The bins contain the average magnitudes of the innovations, exponentially weighted with age. Each bin represents a portion of the operating space of the monitored process. Crisp expert rules were then applied to the average values in these bins to determine the cause of the fault.

In the recursive parameter estimation scheme, some of the parameters are designed to represent the effects of the faults on the system output. These parameters are recursively re-estimated to track the developing fault. The algorithm minimizes the prediction error and uses the sensitiv-

ity coefficients of the fault parameters with respect to the model output to drive the estimation procedure (Salsbury 1996). As the fault develops, the current fault parameter values implicitly describe the state of the system and hence, diagnose the state of the system.

Gray-Box Correlations with Electrical Measurements

The electrical power correlation FDD method produced prediction errors in electrical power. The method also made use of statistically derived confidence intervals for predictions of performance under normal operation. The method is a semiempirical approach that correlates measured fan or pump power with such exogenous variables as airflow, motor speed control signals and actuator position control signals. Power correlations were third-order polynomials; confidence intervals reflected the influence of disturbances during training periods, such as those due to normal variation in damper positions. Oscillatory power data, indicative of unstable local-control loops, were detected via a calculation of signal variance over a sliding window of data points; this calculation effectively acted as a steady-state filter by excluding oscillatory data from comparison with power correlations established during the training phase.

Analysis of chiller power (associated with the economizer and cooling coil valve leakage faults) was more difficult than for fans (air system faults) and pumps (cooling coil undercapacity fault), for two reasons. First, the chiller in the test building was a two-stage reciprocating unit with discrete power levels (0, 5, and 10 kW). In principle, it was possible to time-average the power to obtain a continuous power variable suitable for the same sort of power correlations used for fans and pumps. In practice, the cycling frequency was often long (i.e., on the order of 30 minutes), making it impossible to calculate a short-term power average needed for reasonable correlation with driving variables. Second, chiller power was strongly influenced by environmental conditions (expressed by dry-bulb and wet-bulb temperatures and solar radiation) and by building internal loads. These variables are not all easily measured. Even those that are directly and simply measurable require sensors that have a cost associated with them and are subject to errors. It is necessary to either include these influences in a model of chiller power or exclude them and limit the analysis of chiller power to narrow and known operating conditions.

The FDD method developed and applied in the test building assessed chiller cycling periods under two low-load conditions where it was, in principle, possible to discern a change in chiller loading due to damper and valve leaks.

The method relied heavily on one-minute-average data from installed power transducers to assess the benefits of such data and set the stage for a future cost-benefit analysis. A detailed discussion of the method is presented in a companion paper (Shaw et al. 2002), which includes examples of power correlations, detection of chiller cycling, and analysis of power oscillations.

Table 4 summarizes the types of electrical power analyses used in this method, along with a list of possible faults that each analysis can detect. The list of faults is not exhaustive but is long enough to indicate the difficulties in distinguishing a particular fault from other possible causes of the same deviation between predicted and measured electrical power.

Expert rules were used for limited fault diagnosis. Table 4 indicates how rules can distinguish a slipping fan belt from a fault caused by a change in flow resistance: the former leads to power measurements that differ from the predicted value for a given motor speed while the latter does not. (For the slipping fan belt, the reduction in power due to reduced airflow was a stronger effect than an increase in thermal dissipation from the fan belt itself, which became quite hot when slipping.) Although careful analysis of fan curves indicates that this statement is not entirely true, the impact of a change in duct pressure on the power-speed correlation is sufficiently minimal to be of no concern. As a second example, the leaky recirculation damper or cooling coil valve will affect chiller power but not fan power for a given airflow.

Table 4. Nonexhaustive Listing of Faults Associated with a Given Electrical Power Signature

Type of Electrical Power Analysis	Possible Faults Causing a Deviation Between Predicted and Measured Electrical Power
Polynomial correlation of supply fan power with supply airflow	Change in airflow resistance, possibly due to stuck air-handler dampers or fouling of heating or cooling coils
	Static pressure sensor error (affects portion of fan power due to static pressure)
	Flow sensor error
	Power transducer error
	Change in fan efficiency, caused by change in blade type or pitch, or use of VFD in lieu of inlet vanes
	Change in motor efficiency
Polynomial correlation of supply fan power with supply fan speed control signal	Slipping fan belt
	Disconnected control loop (fan speed differs from control signal)
	Power transducer error
	Change in fan efficiency
	Change in motor efficiency
Polynomial correlation of chilled water pump power with cooling coil control valve position control signal	Change in water flow resistance, possibly due to constricted cooling coil tubes or piping
	Disconnected control loop
	Power transducer error
	Change in pump efficiency
	Change in motor efficiency
Detection of change in cycling frequency for two-stage reciprocating chiller	Leaky cooling coil valve
	Leaky recirculation damper
Detection of power oscillations	Unstable local-loop controller

Ideally, a given power signature would be associated with nonoverlapping lists of faults, providing a high level of "orthogonality" useful for fault diagnosis. As can be seen in Table 4, this ideal was not achieved. Errors in power transducers and changes in fan or motor performance will affect both types of fan power correlations. In the blind tests, it was possible in some cases to distinguish faults associated with a given power correlation by limiting the analysis to a small range of the correlation or to a narrow band of another variable:

- The stuck-closed recirculation damper could be distinguished from the pressure sensor offset in the test building via analysis of power at low airflows. The impact of the stuck-closed recirculation damper fault was exacerbated in the early evening, when the air handlers were still running but, in cold or hot weather, the outdoor air damper was fully closed. The supply fan then drew air across two closed dampers. There is substantial variation across buildings in control strategies at the beginning and end of the working day and it is difficult to generalize such an approach. Even in the test building, this strategy could not be generalized to AHU-1, which operated continuously.
- The leaky cooling coil valve and the leaky recirculation damper could be distinguished with the help of measurements of the valve position control signal and outdoor temperature (Shaw et al. 2002).

This approach to fault diagnosis, while unable to distinguish a large set of possible faults, was easily implemented in rules. For the blind tests, where the number of possible faults was limited,

one rule was "if the electrical power exceeds the confidence interval and the airflow is less than a threshold, then the fault is a stuck-closed recirculation damper." More generally, the "if-then" statement could be modified to include a larger list of possible causes of the detected fault.

RESULTS

The results from the FDD trials on AHU-A/B and AHU-1 are shown in Tables 5 through 8. The original intent of the first-principles-based approach was to not use the mixed-air temperature sensor because this is not commonly available. The predictions of the mixed-air humidity ratio and temperature by the economizer model were used as inputs to the cooling coil model. The magnitudes of prediction errors associated with normal operation in the mixing box and cooling coil (due to the influence of unmodeled disturbances) led to a reduction in sensitivity to fault detection. There was also a significant reduction in the isolation of the cause of faults (faults possibly being in one of two subsystems), which led to ambiguous diagnoses. The model parameters were re-estimated using the mixed-air temperature measurement and the summer tests were rerun with the addition of this measurement. The analyses for the other season were carried out using the same models.

Table 5 describes the first blind test period, conducted in summer conditions on AHU-A and AHU-B. Tables 6 and 7 describe the winter and spring blind test periods. The tests on AHU-1 are summarized in Table 8; for these tests the first-principles-based FDD approach gave a single diagnosis through application of the RPE method and expert rules.

For both the first-principles-based models and the electrical power correlation method, the results are discussed first for AHU-A and AHU-B and then AHU-1. Where individual faults are highlighted, the order is the same as that presented in Tables 2 and 3.

First-Principles Models with Thermo-Fluid Measurements

In general, the first-principles-model-based FDD method proved to be effective in the detection of the faults implemented on AHU-A and B. All faults were detected in each season they were implemented, with the exception of the leaking recirculation damper and leaking cooling coil valve. Diagnosis was less reliable, in that no single method of diagnosis (expert rules or recursive parameter estimation) was able to provide a diagnosis for all fault conditions.

A leaking recirculation damper can be expected to produce small differences between the expected and observed mixed-air temperature. This is, however, a function of the temperature difference between the ambient and return air streams (when these temperatures are equal, no faults can be detected using temperature measurements) and the size of the leakage. In order to detect these changes, the model of the economizer must be quite precise in its prediction of the mixed-air temperature. The principal factors affecting the precision of the model were

- Stratification at the locations of the temperature sensors (return and mixed), which affects the calibration of the model parameters and subsequent calculations of prediction error.
- Localized offsets caused by differences between the measured ambient-air temperature and the temperature of the ambient air entering the inlet duct, some distance away. The effect was similar to the above point.
- Unmodeled effects, of which there were two primary sources:
    - The pressure/resistance characteristics for the fan/duct system were not constant as a result of the fan control strategy (return fan runs at a fixed percentage of the supply air volumetric flow rate). This meant that at different fan speeds, different proportions of the ambient air and return air were mixed for a given damper position. Prediction errors are generated because the model assumes that these proportions are unchanging (i.e., that the system is well balanced).

Table 5. Detection and Diagnosis of Faults During Summer Blind Test Period for AHU-A and AHU-B

Test Day	AHU A	AHU B	Physical Models: Detect from Innovations	Physical Models: Diagnose from Expert Rules	Physical Models: Diagnose from Recursive Parameter Estimation	Electrical Power Models: Detect	Electrical Power Models: Diagnose
1	Slipping fan belt (Stage 1)		No	—	—	Yes	Yes
		Reduced cooling coil capacity (Stage 1)	Yes	Yes	No	Yes	Yes
2	Slipping fan belt (Stage 2)		No	—	—	Yes	Yes
		Reduced cooling coil capacity (Stage 2)	Yes	Yes	No	Yes	Yes
3	Slipping fan belt (Stage 3)		Yes	No	Yes	Yes	Yes
		Reduced cooling coil capacity (Stage 3)	Yes	Yes	No	Yes	Yes
4	No fault		No fault	—	—	No fault	—
		Unstable pressure control	Yes*	—	—	Yes	Yes
5	Unstable pressure control		Yes*	—	—	Yes	Yes
		No fault	No fault	—	—	No fault	—
6	Static-pressure sensor offset (Stages 1-3)		Yes	No	Yes	Yes	Yes
		Stuck-closed recirculation damper	Yes	No	Yes	Yes	Yes
7	Stuck-closed recirculation damper		Yes	Yes	Yes	Yes	Yes
		Slipping fan belt (Stage 1)	No	—	—	No	—

*The unstable pressure controller was detected via the steady-state filter, which indicated that the measured pressure was in a dynamic state for a prolonged period. No detection or diagnosis method was applied, because the measurements did not pass the filter.

- At certain conditions, the fan/duct system imbalance also caused ambient air to flow in through the exhaust grille, resulting in a higher proportion of the ambient air in the mixed air than expected by the model. This problem was exacerbated in the system under investigation because of the increased resistance to airflow through the ambient air inlet duct due to the installation of the preheat coil.

These problems can be considered to be system faults and not problems with the FDD method. The last point highlights the disadvantages of an analytical modeling approach applied to observations of the result of the process, rather than modeling the process itself. Given a more detailed fan/duct system model, it may be possible to predict the airflow rates within the system (within a tolerable degree of uncertainty), which would eliminate the need for the mixing box model as presented in this work.

Table 6. Detection and Diagnosis of Faults During Winter Blind Test Period for AHU-A and AHU-B

Test Day	AHU A	AHU B	Physical Models: Detect from Innovations	Physical Models: Diagnose from Expert Rules	Physical Models: Diagnose from Recursive Parameter Estimation	Electrical Power Models: Detect	Electrical Power Models: Diagnose
1	No fault		No fault	—	—	No fault	—
		Stuck-closed recirculation damper	Yes[1]	No	Yes[2]	Yes	Yes
2	Leaking cooling coil valve (Stages 1-3)		No	—	—	Yes	Yes
		Slipping fan belt (Stage 1)	No	—	—	No	—
3	Leaking recirculation damper (Stage 1)		No	—	—	Yes	Yes
		Slipping fan belt (Stage 2)	No	—	—	No	—
4	Leaking recirculation damper (Stage 2)		No	—	—	Yes	Yes
		Slipping fan belt (Stage 3)	Yes	Yes[3]	Yes	Yes	Yes
5	Leaking recirculation damper (Stage 3)		No	—	—	Yes	Yes
		Static pressure sensor offset (Stages 1- 3)	Yes	Yes	Yes	Yes	Yes
6	Static pressure sensor offset (Stages 1- 3)		Yes	Yes[4]	Yes[5]	Yes	Yes
		Unstable pressure control	Yes	Yes	Yes	Yes	Yes
7	Unstable pressure control		Yes[6]	—	—	Yes	Yes
		No fault	No fault	—	—	No fault	—

[1]Alternative diagnosis of static pressure sensor drift.
[2]Alternative diagnosis of slipping fan belt.
[3]Alternative diagnosis of unknown mixing box fault.
[4]Alternative diagnoses of excessive control dynamics and excessive outside air (due to flow of outdoor air into the exhaust damper, an actual—not artificial—system fault).
[5]Alternative diagnosis of slipping fan belt.
[6]The unstable pressure controller was detected via the steady-state filter, which indicated that the measured pressure was in a dynamic state for a prolonged period.

Estimating the parameters for the economizer model means that either the system has to be balanced correctly or the training data (ideally) need to encompass the complete range of expected airflow/resistance characteristics. If the model does not describe its output in these terms, then robustness comes through increased uncertainty in the model output (wider confidence limits or thresholds) and reduced sensitivity in fault detection. Practically, the model training data have to be collected at one or two airflow rates. An improvement would be to retune the thresholds as new regions of operation are encountered, to maintain maximum sensitivity.

Table 7. Detection and Diagnosis of Faults During Spring Blind Test Period for AHU-A and AHU-B

Test Day	AHU A	AHU B	Physical Models: Detect from Innovations	Physical Models: Diagnose from Expert Rules	Physical Models: Diagnose from Recursive Parameter Estimation	Electrical Power Models: Detect	Electrical Power Models: Diagnose
1	Static pressure sensor offset (Stages 1-3)		Yes	No (unknown fault)	Yes	Yes	Yes
		Leaking cooling coil valve (Stage 1)	No	—	—	Yes	Yes
2	Unstable pressure control[1]		No fault[2]	—	—	No fault[2]	—
		Leaking cooling coil valve (Stage 2)	Yes	Yes	No[3]	Yes	Yes
3	Normal operation		No fault	—	—	No fault	—
		Leaking cooling coil valve (Stage 3)	Yes	Yes	Yes		—
4	Reduced cooling coil capacity (Stage 1)		Yes	No[4]	No[5]	Yes	Yes
		Normal operation	Yes	No[4]	—	No fault	—
5	Reduced cooling coil capacity (Stage 2)		No	—	No[6]	No	—
		Leaking recirculation damper (Stage 1)	No	—	—	No	—
6	Reduced cooling coil capacity (Stage 3)		Yes	Yes	Yes[7]	Yes	Yes
		Leaking recirculation damper (Stage 2)	No	—	No[6]	No	—
7	Unstable pressure control[1]		Yes	No[4,8]	No [8,9]	Yes	Yes
		Leaking recirculation damper (Stage 3)	Yes	Yes	No[6]	No	—
8	Unstable pressure control		Yes[10]	—	—	Yes	Yes
		Leaking recirculation damper (Stage 3)	Yes	No[11]	No[6]	No	—

[1]Implementation of unstable pressure control was considered by test implementers not to have taken effect.

[2]Although the day was labeled as normal by the physical model and electrical power FDD methods, periods of control instability were detected by both methods.

[3]Diagnosed as increased coil capacity (discharge-air temperature lower than predicted) or an unknown fault. An increase in coil capacity would not represent a degradation fault but could indicate that the cooling coil had been cleaned after the model calibration period.

[4]Diagnosed as a leaking cooling coil valve.

[5]Diagnosed as increased coil capacity.

[6]Diagnosed as an unknown fault in the mixing box.

[7]Alternative diagnosis as an unknown fault in the mixing box.

[8]The physical model FDD method detected brief periods of unstable control.

[9]Diagnosed as increased coil capacity.

[10]The unstable pressure controller was detected via the steady-state filter, which indicated that the measured pressure was in a dynamic state for a prolonged period.

[11]Diagnosed as an unknown fault.

Table 8. Detection and Diagnosis of Faults for AHU-1 Blind Test Period

Test Day	Fault	Physical Models: Detect from Innovations	Physical Models: Diagnose from Expert Rules and Recursive Parameter Estimation	Electrical Power Models: Detect	Electrical Power Models: Diagnose
July					
12	No fault	False alarm	Coil undercapacity	No fault	—
14	No fault	False alarm	False alarms from supply air and cooling coil models; design fault from mixing box model[1]	No fault	—
16	No fault	False alarm	False alarm from cooling coil model[2]	No fault	—
17	No fault	False alarm	False alarm from cooling coil model[2]	No fault	—
27	Aborted fault	No fault		No fault	
28	Static pressure sensor offset (Stages 1-2)	Yes	Yes	Yes	Yes
29	Static pressure sensor offset (Stages 2-3)	Yes	Yes	Yes	Yes
30	Stuck outdoor air damper	Yes[3]	Inconclusive diagnosis from mixing box model	No	—
31	Stuck outdoor air damper	No	—	No	—
August					
1	Normal operation	No fault	—	No fault	—
9	Cooling coil air-side fouling (Stage 1)	Yes[4]	False alarm from cooling coil model[2]	No	—
10	Cooling coil air-side fouling (Stage 2)	Yes[4]	False alarm from cooling coil model[2]	No	—
10-11	Cooling coil air-side fouling (Stage 3)	Yes[4]	False alarm from cooling coil model[2]	No	—
11-12	Leaking heating coil	Yes[5]	False alarm from supply air model; cooling coil capacity fault	No	—
13	Stuck-closed recirculation damper	Yes[5]	Yes—correct diagnosis from the supply air model; a false alarm from cooling coil model[2]	Yes	No
14, 15, 16	Normal operation	No fault	—	No fault	—
17	Loss of supply fan control, aborted after 2 h	Yes	No	Yes	No
18	Loss of supply fan control	No	—	No	—
19	Normal operation	No fault	—	No fault	—

[1]A fault was detected from innovations in each of the three models: supply air, mixing box, and cooling coil. The diagnosis from the supply air model was a false alarm due to a poor representation of the air distribution system. The diagnosis from the mixing box model was a design fault because the return fan control for AHU-1, a fixed percentage of the supply fan speed, influenced the proportion of outside air in the supply air. The diagnosis from the cooling coil model was a false alarm because the prediction of mixed air humidity needed in the model was influenced by the return fan control strategy.

[2]The diagnosis from the cooling coil model was a false alarm because the prediction of mixed air humidity needed in the model was influenced by the return fan control strategy.

[3]A fault was detected from innovations in the mixing box model.

[4]A fault was detected from innovations in the cooling coil model but the diagnosis was a false alarm.

[5]A fault was detected from innovations in the supply air and cooling coil models.

The lower levels of leakage in the cooling coil valve were difficult to detect. A significant factor that resulted in this insensitivity was the modeling of the valve. The nonlinear characteristics associated with the heat exchanger process combined with the poorly balanced chilled water circuit resulted in very high gain as the valve opens. This is difficult to model and results in a high degree of uncertainty in the low region of operation. It was found that leakage could only be detected when the valve was closed.

The leakage fault was not detected during the winter test period, which should have been the period when the fault was most visible. The steady-state detector deemed that an extremely high proportion of the data were transient and hence there were almost no data with which to monitor the system. This excessive dynamic activity in the coil control system was due to the cycling of the chilled water inlet temperature, due to the close coupling of the chiller to the coil and to the two-stage control of the reciprocating chiller. Although this could be considered to be a design fault, these effects were considered to represent "acceptable operation." The steady-state threshold was reset during the spring test period to allow more data through the steady-state detector and hence permit the fault monitoring function. The relaxed steady-state criterion, however, resulted in larger model prediction errors. The fault threshold were increased accordingly, again reducing the sensitivity of the method to detection.

The reduced coil capacity fault implemented on AHU-A and AHU-B was easily detected with the exception of spring operation, when only the highest level of the fault was detected. This was attributed to the relatively low load on the coil, which limited the effect of the fault on the coil performance.

The fan-duct reference model for the static pressure prediction was sufficiently accurate to allow the detection of the offset in static pressure, at least for the second and third magnitudes of the fault. The first two stages of the slipping fan belt were not detectable, although the observations from the data showed the effects on the performance to be very small. Oscillatory supply duct pressure control was detected by a prolonged period of dynamic activity (as indicated by a lack of steady-state data classified by the steady-state detector). This approach proved to be reliable: the fault was detected each time it was implemented.

The stuck-open outside air damper fault implemented on AHU-1 was not detected. The return fan overloaded the supply fan to such an extent that the relative proportion of the recirculation airflow increased from normal operation. The effect of the stuck-*open* outside damper was masked to the extent that the system appeared to have a fault similar to a stuck-*closed* outside air or exhaust air damper. The first-principles-based method indicated that a fault condition was present in the economizer on the day that the stuck-open damper fault was implemented, although no firm diagnosis could be made.

Neither the leaking heating coil valve nor the fouled cooling coil surface faults were detected in AHU-1. Both faults should have been detected by observation of the prediction error at the supply air temperature point, although it would have been unlikely that the faults could have been distinguished because they both result in a reduction in cooling coil capacity. Failure to detect both faults was due to the high degree of uncertainty in the model predictions. Factors contributing to this uncertainty that the point temperature sensors on the air-side are more susceptible to airflow-related temperature offsets such as stratification, and that the estimation of the mixed-air humidity ratio (inlet humidity to the cooling coil) was poor, partly because of the airflow imbalance problems discussed for AHU-A and B, and partly due to the estimation of the parameters of the economizer model from data taken from point temperature measurements. The cooling coil model was dehumidifying during the summer test period and model predictions were sensitive to uncertainty in the estimate of inlet humidity. These problems led to several false alarms during the AHU-1 test period. After completion of the test period, it was clear that these false alarms could have been avoided by a marginal increase in the level of the thresholds.

Conclusive diagnosis with both the expert-rules and the recursive parameter estimation was limited by the need for data to be available across the range of operation (low, high and mid ranges). For example, a leaking cooling coil valve could only be distinguished from a sensor offset if the fault was apparent only when the control valve was closed or nearly closed. However, during most of the tests implemented in this study, the systems remained in a narrow region of operation.

The problems associated with unmodeled disturbances, the lack of independence in the parameters and tests carried out over nominally one operating condition (season) led to difficulties in generating reliable performance from this method.

Gray-Box Correlations with Electrical Measurements

Results with submetered power data were very satisfactory for the three blind test periods for AHU-A and AHU-B. Almost all faults were detected. Careful maintenance and control of the HVAC systems and a limited pallet of faults to choose from made fault diagnosis possible, whereas it would be substantially more difficult or impossible in a less-controlled setting.

The stuck-closed recirculation damper was detected and diagnosed in the two test periods in which it was implemented. The leaky recirculation damper was the most difficult to detect. Analysis of chiller cycling frequency was limited to a narrow range of outdoor temperatures, to block the influence of outside temperature on chiller loading. Suitable conditions were present in the late-winter test and the fault was successfully detected and diagnosed. Temperatures were milder in the spring test and the fault was not found. A less restrictive temperature band, not evaluated, might have made it possible to find the fault in spring at the expense of possible false alarms.

The leaky cooling coil valve was detected and diagnosed in the two test periods in which it was implemented. The coil capacity fault was detected and diagnosed successfully in the summer test period and was also found on two of the three implementation days in the spring test period. It was not detected during the second of the three degradation stages in spring because the cooling loads were relatively low and the cooling coil valve did not open to an extent sufficient to reveal the fault.

The pressure sensor offset fault was detected and diagnosed successfully in all three test periods and the unstable fan controller was detected and diagnosed in the two periods in which it was implemented. All three degradation stages of the slipping fan belt were detected and diagnosed in the summer test period but only the most severe stage was found in the winter tests. At that time, the detection algorithm required that the fan-speed control signal be 100%, an unduly severe restriction that was met only on the last day, when there were large loads on the fan. The detection algorithm was changed for AHU-1, to rely on confidence intervals above and below the normal-operation correlation of fan power with speed, with no restriction on the speed signal as a prerequisite for detection of a fault.

As noted earlier, four of the six AHU-1 faults were entirely unknown to the investigators and had not been studied on AHU-A and AHU-B. The electrical power method successfully detected three of the six faults (stuck-closed recirculation damper, pressure sensor error, and loss of control of the supply fan), successfully diagnosed only one (pressure sensor error), and did not find the three remaining faults. Balancing this mixed performance, it is worth noting that one of the detected faults, the loss of control of the supply fan, was not among those for which the method had been commissioned. Further, the method did not generate any false alarms.

After the AHU-1 faults were revealed to the investigators, the electrical power FDD method was extended and applied with more care to data recorded during days when the undetected faults were implemented. The three faults still defied detection. Neither the stuck-open outside air damper nor the fouling on the cooling coil affected the supply fan power for a given airflow.

The impact of fouling on cooling coil capacity was not investigated because chiller cycling at high loads is strongly affected by unmeasured variables (internal and solar loads, for example). The leaking heating coil valve could not be detected via a change in power consumption of the source of hot water because the boiler was not monitored. An analogous method was successful in finding the leaking cooling coil valve, as already noted. While the leaking heating coil valve did introduce a heating load on the (downstream) cooling coil that affected the chiller cycling period, the change was not sufficiently conclusive to warrant flagging it as a fault.

DISCUSSION

All sensors required to implement the first-principles-based methods (Table 9) are typically installed in VAV systems for controlling the HVAC processes; the sole exception is the supply airflow sensor, which is not used when the return fan is controlled on the basis of supply fan speed. For effective fault isolation it is desirable to generate prediction errors at the outlet of each modeled subsystem. The attendant sensor is required to make this possible. Improvements to the performance of the cooling coil FDD scheme could have been realized if the coil outlet air temperature measurement was used, rather than the supply air temperature (the fan temperature rise model would not have been necessary).

Results from the test periods demonstrated that the accuracy of the cooling coil model predictions under dehumidifying conditions might have been improved if a better estimate of the inlet humidity ratio were available. A possible solution is the installation of an additional sensor if the cooling coil is designed for latent duty, although the air local to the sensor location needs to be well mixed. An alternative would be to measure the ambient airflow rate as well as the supply airflow rate. The proportions of ambient and return air in the supply air could then be calculated directly.

Ideally, the outside-air temperature and humidity sensor would be located in the inlet duct to the system rather than being external to the building to reduce the effects of improper representation of the air properties. This would inevitably increase the cost of the implementation of the FDD scheme.

The electrical power FDD method required substantially fewer measurements, which reduced the sensor maintenance requirements. Sensors for this method are listed in Table 9. The electrical power data that are at the heart of the method are not available in typical HVAC plants. The method as implemented in the test building required a power meter for the supply fan for each AHU. Another power meter was required for the single onsite chiller. The apparent economy in having just one chiller and one chiller power meter was more than outweighed by difficulties in ascribing changes in chiller cycling to faults in a particular air handler. To provide an informative test of the electrical power FDD method, it was necessary to couple the chiller to a single air handler and use district chilled water for the others, even when cooling loads were low and the single chiller would have had adequate capacity. In a high-rise office building where a single chiller serves separate air handlers on each floor, the electrical power method would not be able to detect faults on the basis of chiller power.

The sensors available for use on this project were generally instrument-grade devices having a higher accuracy, more stability and less drift than standard HVAC-grade devices. Exceptions include the return air humidity sensor and the supply duct static pressure transducers, which were standard HVAC-grade devices. Equipment costs for instrument-grade sensors are typically 3 to 5 times the cost of standard, mass-produced HVAC-grade sensors. Costs noted below exclude installation, setup, and any onsite calibration, which is estimated to average $75 per sensor.

Sensors that require periodic calibration, such as those measuring temperature, pressure, and flow, were within their appropriate calibration dates to comply with certification requirements.

Table 9. Sensors and Control Signals Required for Implementation of Each FDD Method

Sensor Type or Control Signal	Condition	First-Principles-Model FDD	Electrical Power FDD
Temperature	Return air	X	
	Outside (ambient) air	X	X
	Mixed air	X	
	Supply air	X	
	Chilled water flow to coil	X	
Humidity	Return air	X	
	Outside (ambient) air	X	
Flow	Supply air	X	X
Pressure	Supply duct static pressure	X	X (training only)
Electrical power	Chiller		X
	Supply fan		X
	Secondary chilled water pump		X
Control Signal	Return fan	X	
	Economizer	X	
	Cooling coil control valve	X	X
	Supply fan	X	X

Note: All sensors and control signals are required for fault detection; there would no reduction in sensor count if monitoring were limited to detection and excluded diagnosis.

Calibration represents the most significant maintenance requirement for the sensors. Typically, the listed sensors are calibrated annually, and more frequently than sensors in a standard HVAC system:

- The return, supply, and outside airflow rates and temperatures were measured by electronic airflow measuring stations. Each measuring station costs $1400 and has a stated accuracy of ±2% for airflow greater than 2.54 m/s (500 fpm) and ±0.2°C (0.36°F) for temperature, with zero long-term drift.
- Mixed-air temperatures were sensed using instrument-grade 1000-ohm platinum RTDs arranged in a multipoint array. The array has a listed device accuracy of ±0.14°C (0.25°F) and an average cost of $2500. Return air humidity is measured with a standard HVAC-grade humidity sensor with an accuracy of ±3% and an equipment cost of $50.
- Supply duct static pressure transducers were standard HVAC-grade devices with a stated accuracy of ±1% of full scale [±0.75 Pa (0.03 in. of water)]. Stability was listed as ±1.0% of full-scale deviation from original calibration for one year under normal operating conditions. These transducers cost $225 each, compared with $485 for sensors with an accuracy of ±0.25% of full scale, which were installed after this project was completed.
- Water temperatures were sensed with single point, direct immersion, instrumentation-grade, 1000-ohm 2-wire and 100-ohm 4-wire, platinum RTD sensors. The sensors have a stated accuracy of ±0.14°C (0.25°F) and are very stable, with little long-term drift. These temperature sensors and related equipment have an average cost of $125 each.
- Electrical power to the supply fans, pumps and chiller was measured with precision AC watt transducers with a stated accuracy of ±0.2%. These devices have NIST-traceable calibration. The watt transducers cost $400 each, including $60 for calibration; the transducer for the chiller required additional current transducers and cost $550. HVAC-grade watt transducers,

not used at this site, are available for about $250.

- Outside air conditions were measured by an instrument-grade temperature and humidity measuring weather station. A 100-ohm platinum RTD with a stated accuracy of ±0.2°C (0.36°F) was used for temperature and a polymer sensor with a stated accuracy of ±1% (at 0 to 90%rh) was used for humidity measurement. The measuring station has a total cost of approximately $1000.

The overall accuracy of the instrumentation system depends on factors in addition to the stated accuracy of the sensing device: transducers, stability of power supplies, wiring type, lead length, A/D converters, processor resolution, scaling values and, in particular, measurement representation of average fluid properties and quantities. The impact of these uncertainties on the performance of FDD methods was not evaluated as part of this work. A full assessment of the uncertainties associated with the first-principles-based FDD methods in HVAC systems is presented in Buswell (2001).

The parameters of the first-principles models need to be calibrated for each test system. Some parameters are identified directly from design information and/or inspection. The remaining parameters are simultaneously identified using test data from the target system. These data were obtained by increasing the control signal to each sub-subsystem in a series of steps from 0% to 100% and back to 0%. Each step is held until steady state is considered to exist. For the tests conducted in this research, the total time taken to commission three subsystems in one AHU (economizer, cooling coil, and fan-duct system) was 23 h. This was controlled largely by the time constants associated with the system and the need for the observation of a number of consecutive points at each step to decide whether steady-state conditions exist. To decrease this overhead, simultaneous commissioning of all three subsystems was investigated on AHU-1 (Norford et al. 2000). The simultaneous test took 14 h to complete, saving 9 h. Alternative schemes to generate training data may provide a quicker method of capturing the system characteristics.

The model parameters are described fully in Norford et al. (2000) and are listed in Table 10. The commissioning tests were designed to provide data to capture the principal system characteristics. It was not always possible to identify all parameters simultaneously such that each parameter represented its prescribed system characteristic. This was caused by a lack of parameter independence, which increased as the number of parameters increased (i.e., one parameter estimate may in part be offset by the value in another and hence the subsequent parameter estimates become gray).

This phenomenon, particularly apparent in the economizer model, led to an approach in which subsets of the parameters were estimated from the commissioning data that most related to their effect on the model. Leakage parameters were identified from data for which the control elements were closed (0% or 100% control signal). With the leakage parameters fixed, the model "gain" parameters were then identified from the data for the opening movement of the control signal (0% control signal to 100% control signal). Finally, with the leakage and gain parameters fixed, the control actuator hysteresis parameters were identified from all of the commissioning data (0% control signal to 100% control signal and then reverse from 100% to 0%).

The steady-state detector also required two parameters for each monitored subsystem. One describes the dominant subsystem time constant in relation to system dynamics and the other controls the amount of data that is considered to be transient. As the stringency of the parameter is increased, the amount of data passed on to the FDD methods is reduced. Trials on a number of HVAC systems, including those systems tested here, have revealed that the values of these remain quite consistent for similar subsystems.

Table 10. Parameters of First-Principles Models

Subsystem	Design Parameters	Calibrated Parameters
Fan-duct	Rotational speed of the fan (estimated from the control signal) (–)	*Static pressure sensor offset (Pa)*
	Static pressure at zero mass flow rate (Pa)	Total fan/duct resistance to airflow (sPa/kg)
	Fan minimum rotational speed (–)	
	Fan maximum rotational speed (–)	
	Control signal relating to minimum rotational speed (–)	
	Speed at which fan belt slippage occurs (–)	
Air temperature rise due to fan	Control signal relating to the minimum fan speed (–)	Minimum temperature rise (K)
		Maximum temperature rise (K)
Economizer (dampers)		Mixed air temperature offset (K)
		Parameter that describes the degree of curvature in the process relationship (–)
		Parameter that defines the asymmetry of the process (–)
		Leakage through the return damper (–)
		Leakage through the outside air damper (–)
(actuator)		High activation point (–)
		Low activation point (–)
		Hysteresis (–)
Cooling Coil (coil)	Coil face area (m^2)	*Heat-transfer scaling factor (–)*
	Number of rows (–)	Supply air temperature sensor offset (K)
	Number of circuits for parallel flow (–)	*Fractional flow leakage (–)*
(three port control valve)	Maximum chilled water mass flow rate (kg/s)	Curvature coefficient (–)
		Authority (–)
(actuator)		High activation point (–)
		Low activation point (–)
		Hysteresis (–)

Note: The parameters in italics are those used for fault diagnosis.

In concept, the methods do not require the selection of fault thresholds. A fault is simply detected when a prediction error is considered to be significant against some statistical measure. Similarly, a statistically significant change in the value of a recursively re-estimated fault parameter could also be used to indicate the presence of a fault. In practice, the scheme did not formally account for the uncertainty with respect to the sensitivity of the model output to unmodeled phenomena. The addition of the thresholds, above which a prediction error generated an alarm, was a subjective and somewhat ad hoc attempt to account for this uncertainty. Better model representation of the process and a formal methodology to account for the uncertainties

from all sources has been shown to produce a more sensitive and a more robust first-principles-based fault detection scheme than that used in this research (Buswell 2001).

The expert-rule method that employs the bins requires a number of parameters to locate the bins within the operating space for each subsystem (i.e., what value of control signal is attributed to low, mid and high operation). The expert rules are generic and principally describe what evidence is expected to exist with respect to the fault and the bin categories.

The sensitivity of the electrical power FDD method to faults and robustness with respect to false alarms are functions of the extent to which the semiempirical power correlations capture the process characteristics and the calculation of a large number of thresholds and other parameters, listed in Table 11. These thresholds and parameters can be usefully grouped as follows: (1) normal equipment power levels, from measurement or manufacturer's data; (2) statistical confidence intervals; (3) fault detection thresholds based on commissioning the FDD method with known faults; or (4) data analysis regions, based on commissioning the FDD method with known faults and designed to improve the robustness of the method in both detection and diagnosis.

The list of parameters does not include the parameters in the third-order polynomial curve fits that express the correlation of electrical power with flow or rotational speed. Values for these parameters were determined from training data for individual fans and pumps.

Required values for the first two groups are easily obtained or assigned. One goal of ongoing research is to reduce the number of parameters and thresholds in the last two groups, to simplify and eventually automate the process of commissioning this FDD method. To that end, the AHU-1 test period in this research provided an opportunity to replace the effective but hand-tuned analyses of fan power as a function of speed and pump power as a function of valve position with more straightforward polynomial power correlations, for which the statistics are rigorous and the only need is to supply a confidence interval.

Calibration of the power correlations for normal operation was entirely a passive procedure and demanded only a reasonable range of operating conditions. For the test building, about 10 h of data from power meters and supporting sensors were required. At the test building, these data were collected in a single summer season. These data were sufficient for fault detection but not diagnosis, where knowledge of fault signatures was required to distinguish faults that reveal themselves as a deviation in a power correlation for a single component (i.e., the supply fan). Fault diagnosis rules developed for equipment in the test building relied on observations of faulty performance. Further tests in different buildings are necessary to determine the extent to which these rules are general or can be easily adjusted. It is clearly not practical to commission an FDD method with onsite faults.

The first-principles-based methods are formulated to represent what is normally considered as "ideal" system operation. In this respect, they are sensitive to any nonideal system behavior, which could represent a design fault. Two forms of nonideal behavior affected the sensitivity and robustness of the first-principles methods during these tests. Both forms concerned changes in the relative proportion of outside and recirculation airflow rates through the mixing box, as the supply fan speed varied. One form of this phenomenon is reported in Seem et al. (1998). The nonideal behavior could cause innovations in both the mixed-air temperature and supply air temperature, suggesting faults in the economizer and cooling coil respectively.

In addition to the effect of the nonideal system behavior, the first-principles-based methods were sensitive to the changeover in system configuration necessary at the start of each test period (the test systems were used for other project work in between the seasonal tests conducted in this study). The changeover in configuration concerned control strategies and reinstallation of some sensors. Further, some physical disruption to the systems, such as repairs to

Table 11. Thresholds and Other Parameters Required by Electrical Power FDD Method

Description of Thresholds and Parameters	Value
Fan-power correlations with airflow and speed control signal	
Maximum deviation of static pressure from set point for training data	25 Pa (0.1 in. of water)
Confidence level to establish boundary between normal and faulty data	90%
Airflow boundary to distinguish stuck-closed recirculation damper from static pressure offset/drift[1]	500 cfm
Fan power at 100% speed below which a slipping-fan-belt fault was flagged, subject to a minimum time duration[2]	1 kW
Time duration for low fan power at 100% speed, above which a slipping-fan-belt fault was flagged	3 one-min. power samples
Pump-power correlation with cooling coil valve position control signal	
Valve position control signal above which pump-power data were analyzed for a cooling coil capacity fault[3]	40%
Measured normal-operation power level of the secondary chilled water pump	400 W
Minimum decrease of pump power below normal-operation value, in excess of which a coil capacity fault was flagged[4]	10 W
Confidence level to establish boundary between normal and faulty data (used for AHU-1)	90%
Chiller-cycling analysis	
Power level above which the chiller is considered to be operating in the low-power stage[5]	4 kW
Cycling interval when the cooling coil valve control signal is at 0%, below which a leaky-valve fault is flagged[4]	30 min.
Normalized outdoor air temperature, below which chiller cycling is analyzed to detect a leaky recirculation damper[6]	0.2
Power-oscillation analysis	
Size of sliding window for averaging one-minute power data from submeters	5 samples
Standard deviation of power signal above which a fault is flagged, as a percentage of average power	15%

[1]This parameter was used solely for fault diagnosis.

[2]Fan-power analysis at 100% speed was used in AHU-A and B to detect the slipping fan belt. For AHU-1 this approach was replaced by the more rigorous and sensitive polynomial correlation of fan power with speed control signal.

[3]Pump-power analysis relative to a measured and near-constant normal-operation value was used in AHU-A and B to detect the coil capacity fault. For AHU-1 this approach was replaced with a polynomial correlation of pump power with valve position control signal.

[4]The parameters for the decrease in pump power and the change in chiller cycling interval were used for a single-stage detection and diagnosis in the test building, where the number of faults was limited.

[5]The chiller's high-power stage was not of concern, because the chiller-cycling analysis was limited to low-load conditions when the chiller was either off or in the low-power state.

[6]The normalized outdoor air temperature, defined in Shaw et al. (2002), is the difference between the outdoor air temperature and the supply air temperature set point, normalized by the difference between the supply and room air temperature set points.

damper linkages, was necessary between test periods. This could have caused a change in the system characteristics and reduced the validity of the model calibration.

The nonideal system characteristics and the necessary disruption to the state of the system that occurred between test periods resulted in the setting of fault thresholds for each test period. This was done in a subjective manner to eliminate false alarms during periods of

known normal operation (Norford et al. 2000). This approach was not completely effective for the AHU-1 tests, during which the fault thresholds were adjusted for two consecutive days of operation. Following the setting of the thresholds, the nonideal system behavior became more predominant, which led to several false alarms for the test period. Although it was recognized that the alarms were false and that they could be eliminated with very minor changes to the thresholds (~0.25°C), the thresholds were not readjusted during the fault testing, because this would not have been consistent with the FDD methodology as proposed for this project. In practice, the thresholds could easily be made more robust by setting them over a longer period of operation.

In some instances, the nonideal system behavior resulted in the setting of relatively wide fault thresholds, which necessarily reduced the sensitivity of the methods, particularly for the more subtle leakage faults. Considering that the systems were subject to change between test periods and that they exhibited nonideal behavior, the results of this study make it clear, however, that the first-principles methods remained robust in fault detection. The models were not recalibrated to account for disruption to the systems that occurred between test periods.

The diagnosis of faults by the first-principles methods was less robust than the fault detection. A number of faults detected during the blind test periods were misdiagnosed. Robust diagnosis by expert rules requires the system to have operated over its complete range during the occurrence of a fault. Considering that some faults can force the system to move to, and remain at, one operating point, this requirement is impractical. However, this restriction could be eliminated by developing a methodology that includes the injection of test signals to exercise the system across its range of operation once a fault has been detected.

Fault diagnosis by the recursive re-estimation of the first-principles-based model parameters was sensitive to the unmodeled disturbances, the limited excitation in terms of operating condition, and the lack of independence of the parameters. It is unlikely that it would be possible to include parameters to specifically represent all fault conditions. Evidence from this work suggests that typical HVAC system data could probably support two recursively re-estimated parameters; one describing the under/over capacity at the "high duty end" of operation and one describing under/over capacity at the "low duty end" of operation.

The electrical power FDD method was effective in detecting faults for AHU-A and AHU-B. The detection methodology was straightforward for some faults, including those that affected fan power for a given airflow (the stuck-closed recirculation damper and the pressure sensor offset), those that impacted the cycling of the reciprocating chiller, and the oscillating controller. The detection methodologies for other faults were developed in response to the configuration and operation of the equipment in the test building. Notably, the cooling coil capacity fault was detected by identifying changes in the electrical power drawn by the secondary chilled water pump. This method was effective only because the fault was introduced in a way that significantly obstructed the water flow. Water-side tube fouling would likely not have been detected. Detection of the slipping fan belt was effective but done with a threshold developed from observations of the impact of the fault, rather than by relying strictly on a fan power correlation. Detection of this fault for AHU-A and AHU-B was also limited to 100% speed control signal, an apparently unnecessary restriction on the use of the power-speed correlation that was eliminated for AHU-1. This same restriction was employed in the first-principles model to distinguish this fault from the pressure sensor offset.

Several of the faults introduced into AHU-1 did not produce detectable changes in electrical power, including the stuck-open outside air damper, the obstruction of the cooling coil, and the leaky heating coil valve. Detection failure in these cases is acceptable, because the FDD method developed for other faults was rationally applied and simply did not reveal significant deviations from normal operation. It is possible that further tuning of the method could have reliably

revealed the leaky heating coil valve, in the test building and others where the heating coil is upstream of the cooling coil. This will be pursued in the future.

As with the first-principles method, fault diagnosis via the electrical power FDD method was less reliable than fault detection. Diagnosis techniques also required considerably more effort to develop and commission. For example, it was relatively easy to measure the cycling periods of the reciprocating chiller, but considerable care was required to establish operating regions where the leaky recirculation damper could be diagnosed. Inadequate care in this step not only confounds the diagnosis process but leads to false alarms, as would occur if a change in chiller cycling were flagged but that change were due solely to variations in the thermal load on the chiller. Diagnosis of the stuck-closed recirculation damper and the pressure sensor offset required careful observation of the HVAC plant and a knowledge of the normal control of the mixing box dampers at different times of day and different seasons.

The electrical power method proved to be robust in terms of false alarm generation. None were generated in the AHU-1 test period. In part, this is due to the choice of confidence intervals for the power correlation; intervals of lower confidence would be tighter and would tend to reveal more faults and generate false alarms.

CONCLUSIONS

The relatively rare opportunity to thoroughly test FDD methods in a building operated as a research facility is an invaluable bridge between simulation and lab testing and field deployment. Such "real-world" issues as sensor placement and calibration, fault magnitudes, and imperfectly understood equipment performance under normal and faulty operation make it unlikely that a first-generation FDD method will successfully leap from lab to commercial use. Controlled field tests such as were required for this project are not so much a proof of performance of fully mature methods but a means of revealing flaws in the methods and subsequently refining them to the point where their performance is substantially improved.

The first-principles-model-based methods can be implemented without installation of special sensors. However, further work is required to shorten the time required to gather the training data for model calibration. The electrical power correlation method requires the use of additional electrical power sensors, but models are calibrated primarily from data collected during normal system operation. The advantage of a single testing procedure and models that can extrapolate is that the FDD scheme is operational immediately after installation (and calibration) of the FDD software. Normal operation data could be used to calibrate the first-principles-based models in a similar manner to the electrical power correlation approach. The disadvantage, however, is that data from across the subsystem operating season are required before the models are fully calibrated.

As shown in Tables 5 through 8, both FDD methods performed reasonably well in detecting faults. The electrical power FDD method was less developed and may not apply to other sites. Power correlations can be used as an effective method for detecting faults, but it is not clear how much work will be required to adapt the basic approach to different plants. The first-principles-based scheme is more mature than the electrical power correlation scheme as an FDD approach. The methods have been applied to other systems installed in real buildings. The reduction of false alarms and increase in sensitivity in the detection of faults has been addressed in subsequent work by Buswell (2001). In the future, developers of the electrical power method will consider extending that method to use additional sensors, such as temperature sensors typically found in control systems. Table 12 summarizes the evaluation of the two methods.

Both FDD approaches require some additional information to diagnose faults. This effort was simplified at the test site because of the limited number of introduced faults. Implementation of

Table 12. Comparison of First-Principles Physical Models FDD Method and Gray-Box Electrical Power Method

Feature	First-Principles Physical Models	Gray-Box Electrical Power Models
Operates only on steady-state data	Yes	Yes
Calibration time	15 to 23 h	10 h
Active or passive training data	Active; a single training period is sufficient	Passive; need seasonal data
Training methodology	Well defined	Less well defined
Thresholds	Statistical confidence intervals, arbitrarily selected thresholds	Statistical confidence intervals, arbitrarily selected thresholds
Critical sensors	Supply, return, mixed-air and outdoor air temperature sensors, return air and outdoor air humidity, supply airflow, control signals	Electrical power meters, supply airflow, control signals for valves and fan variable-speed drives
Fault detection	Excellent results for faults which the methods were tuned to find	Excellent results for faults which the methods were tuned to find
Fault diagnosis	Moderate number of misdiagnoses	Good results when list of potential faults is small
False alarms	Moderate	None
Extension of method to larger list of faults	Possible	Some faults cannot be detected with electrical power measurements
Other issues	Excessive number of fault parameters in models	Effects of faults on electrical power must be carefully defined for each component (fan, pump and chiller) in each system (constant volume, VAV, etc.)

either method on a system with no prior knowledge of the causal faults will almost certainly result in ambiguous diagnosis.

The first-principles-based methods are sensitive to the occurrence of nonideal system behavior. The detection of nonideal system behavior by the method is an advantage where nonideal behavior is considered to be a design fault. However, where the nonideal behavior must be accepted as part of the system characteristic, the uncertainty in the prediction error increases, reducing the sensitivity of the method to fault detection. Better modeling of the system behavior results in a reduction in the uncertainty in the model predictions and improved fault detection rate. The gray-box electrical power models are less sensitive to nonideal system behavior because the correlations model the system behavior under closed-loop control. The gray-box modeling methods were, therefore, more robust than the first-principles-based methods in that they generated no false alarms.

Robust fault diagnosis using first-principles-based models and expert rules is limited by the need for the system to have moved across its range of operation during the occurrence of a fault condition. A way of generating this information "out of season" would be to inject test signals

that excite the system across the deficient region. Fault diagnosis by the recursive re-estimation of model parameters did not prove to be reliable in the test environment in this research.

The apparent complexity of both methods reflects their degree of development. Much of the identification and data gathering processes described could be fully automated. As the methods mature, the selection of some of the required parameters will become better understood and easier. There are still issues surrounding the sensitivity and robustness of FDD methods (Dexter and Pakanen 2001). A balance between sensitivity in detecting faults and robustness in minimizing false alarms is needed. More testing with data sets from real buildings is required, and in particular a concerted effort is required to generate reliable, streamlined, and automated commissioning processes for FDD methods. Both FDD methods investigated here need to be simpler in terms of application and the analysis of the data and need to be more "transparent" to the end user.

ACKNOWLEDGMENT

The authors warmly acknowledge the financial support of ASHRAE, the California Energy Commission (via subcontracts from Lawrence Berkeley National Laboratory and Architectural Energy Corporation), and Honeywell, Inc.; the vision of Phil Haves, who played a leading role in conceiving this project; and the technical advice of the ASHRAE Project Monitoring Committee, chaired by John House.

REFERENCES

Buswell, R.A. 2001. Uncertainty in the First Principle Model Based Condition Monitoring of HVAC Systems. Ph.D. thesis, Loughborough University, U.K.

Dexter, A.L. and M. Benouarets. 1996. A Generic Approach to Identifying Faults in HVAC Plant. *ASHRAE Transactions* 102(1):550-556.

Dexter, A.L. and J. Pakanen, eds. 2001. *Demonstrating Automated Fault Detection and Diagnosis Methods in Real Buildings*. IEA Annex 34 Final Report. VTT Building Technology, Espoo, Finland.

Haves, P., T.I. Salsbury, and J.A. Wright. 1996. Condition Monitoring in HVAC Subsystems Using First-Principles Models. *ASHRAE Transactions* 102(1):519-527.

Hyvarinen, J. and S. Karki, eds. 1996. *Building Optimization and Fault Diagnosis Source Book*. Technical Research Center (VTT), Espoo, Finland.

Lee, W.Y., C. Park, and G.E. Kelly. 1996a. Fault Detection of an Air-Handling Unit Using Residual and Parameter Identification Methods. *ASHRAE Transactions* 102(1):528-539.

Lee, W.Y., J.M. House, C. Park, and G.E. Kelly. 1996b. Fault Diagnosis of an Air-Handling Unit Using Artificial Neural Networks. *ASHRAE Transactions* 102(1):540-549.

Norford, L.K., J.A. Wright, R. Buswell, and D. Luo. 2000. Demonstration of Fault Detection and Diagnosis Methods in a Real Building (ASHRAE 1020-RP) ASHRAE 1020-RP Final Report.

Peitsman, H.C. and V.E. Bakker. 1996. Application of Black-Box Models to HVAC Systems for Fault Detection. *ASHRAE Transactions* 102(1):628-640.

Salsbury, T.I. 1996. Fault Detection and Diagnosis in HVAC Systems Using Analytical Models. Ph.D. thesis, Loughborough University, U.K.

Seem, J.E., J.M. House, and C.J. Klaassen. 1998. Volume Matching Control: Leave the Outdoor Air Damper Wide Open. *ASHRAE Journal* 40(2):58-60.

Shaw, S.R., D. Luo, L.K. Norford, and S.B. Leeb. 2002. Detection of HVAC Faults via Electrical Load Monitoring. *International Journal of HVAC&R Research* 8(1).

Stephan, W. 1994. Comparison of Different Models for Cooling Coils Under Wet Conditions. *Proceedings of the Fourth International Conference on System Simulation in Buildings, Liege,* p. 291.

Yoshida, H., T. Iwami, H. Yuzawa, and M. Susuki. 1996. Typical Faults of Air-Conditioning Systems and Fault Detection by ARX Model and Extended Kalman Filter. *ASHRAE Transactions* 102(1):557-564.