LB-07-036 (RP-1275)

# Formulation of a Generic Methodology for Assessing FDD Methods and Its Specific Adoption to Large Chillers

**T. Agami Reddy, PhD, PE**
*Member ASHRAE*

## ABSTRACT

*Just as standard methods have been developed for evaluating different building energy simulation programs, there is growing realization of the imperative need to develop methods to evaluate the performance of different fault detection and diagnosis (FDD) methods and tools for different heating, ventilating, air-conditioning, and refrigerating equipment. This paper describes research done in this regard by first reviewing past pertinent literature on FDD evaluation and identifying various elements of a general conceptual evaluation framework. In order to make the tool practical for an impartial evaluation, a small set of quantitative criteria are identified based on which an objective function formulation is described where one minimizes the sum of the costs associated with false positives and those of missed opportunities. Subsequently, we describe how this function can be modified and simplified so as to obtain expressions for normalized rating of any FDD tool in terms of fault detection only as well as in terms of the combined FDD process. These general expressions are then tailored to the evaluation of FDD methods and tools specific to large chillers, and specific numerical values of several of the quantities appearing in the normalized rating expression are proposed based on discussions with chiller manufacturers and service companies as well as analysis of fault-free and faulty chiller performance data gathered in a laboratory from a previous research study. Another paper will describe its application and usefulness to assess four chiller FDD methods customized to the same performance data sets.*

## BACKGROUND AND OBJECTIVE

Research and development into fault detection and diagnosis (FDD) methods and tools applicable to heating, ventilating, air-conditioning, and refrigerating (HVAC&R) equipment has been ongoing for close to 15 years. There is a need to develop a general testing methodology against which different FDD methods and tools could be evaluated. The ultimate objective is to develop a test standard for the industry akin to those available for testing the normal performance of different types of HVAC&R equipment and also evaluating building energy analysis computer programs (ASHRAE 2001). This paper reviews the engineering literature in order to first learn what has already been published and whether any attempt or consensus has been reached in this regard. The various issues to be considered within a general conceptual evaluation framework are described with the objective of identifying a small set of quantitative criteria based on which an analytical generic expression for the normalized measure to rate FDD methods or tools can be developed. How the general expression can be tailored toward specific HVAC&R equipment, namely, large chillers, has been investigated. The research results reported in this paper are based on research sponsored by ASHRAE in Research Project 1275, and the interested reader can refer to the complete final report for further details (Reddy 2006).

## PAST WORK ON FDD EVALUATION

A good way to compare FDD tools is side-by-side comparisons. Research by Norford et al. (2000) was meant to demonstrate FDD methods for air-handler units (AHU) in a real building, to assess the strengths and weaknesses of the methods investigated and to provide guidance for future research in this area that will accelerate the development of FDD technology. The comparison included data from a research test facility for seven different faults (abrupt and degradation) collected during multiple seasons of the

**T. Agami Reddy** is a professor in the Civil, Architectural and Environmental Engineering Department, Drexel University, Philadelphia, PA.

year. The test procedure consisted of the following three steps: (a) preliminary commissioning tests; (b) one-week of control tests during different seasons in which faults were implemented and the researchers were told what faults were implemented (including the severity level), at what time they were implemented, and the duration they were sustained; and (c) one-week blind tests, also during different seasons, in which the researchers knew only that the faults considered during the control tests would be implemented at some time during that week. Both FDD methods proved capable of consistently detecting the faults, with a small number of exceptions. Fault diagnosis procedures were improved over the course of the tests, and the conclusion reached was that these procedures were generally effective. However, diagnosis was made considerably easier than what is likely to occur in typical conditions due to the limited number of known faults, the known magnitude of the faults, and the excellent maintenance of building equipment and sensors. The test procedure was then altered to evaluate the performance of the methods without the benefit of the control test data. The new test procedure was carried out on a different AHU, and the researchers were not told what faults were implemented. The performance of the methods suffered with the removal of step (b). In particular, the ability to diagnose the implemented faults was poor. This research illustrated how challenging it is to detect and diagnose faults in real buildings and, furthermore, how challenging it is to evaluate FDD tools.

Section D of the report written by House et al. (2001) and compiled by the International Energy Agency's Annex 34 working group (Dexter and Pakanen 2001) has also addressed the issue of evaluating FDD tools. The need for performance criteria and an evaluation approach is stated, along with the consideration that application and cost be traded off versus the benefit. Issues that are critical in other industries may not be so in the HVAC&R industry. For example, in power-producing gas turbines, the fault detection time delay is crucial, but it is much less so in the HVAC&R industry, where failures do not usually result in injury to occupants. A checklist of characteristics to consider while assessing the capability of a particular tool is also provided by House et al. (2001) along with summaries of previous work done to compare and evaluate FDD tools in general.

Gruber (2001) describes the feedback from industrial partners regarding issues affecting commercialization. Fifty industrial partners were consulted in various building-related activities, from building automation, to HVAC equipment, to maintenance. The following criteria were flagged in response to a question seeking to identify the essential properties of a good FDD method:

a. Simplicity in terms of understanding it, commissioning it, using it, integrating it to existing building automation systems, and ease in changing it.
b. No or very few false alarms, with the user being allowed to decide on the threshold level.
c. No disruption of normal operation.
d. Detection and diagnosis can be separate.
e. Robustness is a crucial factor in that changing operating conditions should be accommodated.
f. Cost-effective with no human assistance during operation (i.e., FDD tool to be automated).
g. Impact on savings of energy and comfort at different sites.

Issues affecting commercialization, such as how thoroughly it has been tested, how easy it is to use, additional cost, etc., as well as what type of topics should supporting documentation cover in order to convince management of the advantages of purchasing such an FDD tool are also described.

Any methodology to evaluate different FDD tools should explicitly consider the characteristics or capabilities of what these tools are supposed to do. The basic requirements of an FDD system according to Patel and Kamrani are (1996) are as follows:

1. Identify abnormal conditions accurately.
2. Do not give false alarms of abnormal conditions (i.e., be robust).
3. Report the level of confidence associated with each diagnosis.
4. Rank the conclusions.

One could add other requirements as well:

5. Be able to detect the necessary faults at the stipulated sensitivity level (i.e., at the pre-specified fault severity level).
6. As far as possible use sensors already installed by the equipment manufacturer.
7. Select sensors that provide the needed level of accuracy while minimizing cost.
8. Be able to handle insufficient data and uncertain situations.

The success of the FDD system depends on proper selection of methods for both detection and diagnosis (Katipamula et al. 2001). Often methods are selected because of the interest of the developer or the availability of an existing tool. While such methods may yield satisfactory results under small-scale laboratory tests, several professionals and researchers have shown that such an evaluation is often inadequate for satisfactory field operation (for example, Norford et al. [2000], House et al. [2001]). For some FDD applications, fault diagnosis may not be needed because detection isolates the fault. On the other hand, fault diagnosis may not be possible because data resolution may be inadequate.

Requirements (1) and (2) above need some discussion. An important consideration in the design of fault detection systems is to have a system that responds rapidly when a fault occurs but that is not too sensitive to noise so as to generate too many false alarms. The relative magnitude of this trade-off is specific to the mathematical model used and is an indication of the robustness of the model. For non-critical applications, the methods used for detection and diagnosis should minimize the number of false positives (false alarms). If a number of false positive faults are incorrectly detected and diagnosed, the operators may disable the FDD system completely. FDD methods applied to critical systems are tuned to be sensitive to fault detection; therefore, these applications may generate false alarms more often. On the other hand, FDD methods applied to noncritical systems (most building energy systems fall in this category) are likely to be tuned to generate fewer false alarms even if this results in a certain missed opportunity or missed alarms.

## ELEMENTS OF A GENERAL EVALUATION METHODOLOGY

From the above discussion of past work and requirements, we suggest that any FDD evaluation methodology consider the following four distinct elements:

a. **Site-specific criteria** must explicitly consider the type of equipment/system (say, an AHU or a chiller) for which the FDD tool has been developed. One needs to gather the following information: (1) list of faults to be identified in order of importance, based on past cited surveys; (2) list of sensors available and their accuracy, which will allow determination of any additional cost involved with additional sensors or replacing existing sensors; (3) available building automation system that will allow any additional setup and system integration costs to be deduced; (4) range and frequency of different operating states (i.e., external driving conditions) that the equipment is likely to face during a year, which will allow preferential evaluation of more frequently occurring states as compared to less frequent ones; (5) annual energy operating costs and maintenance costs, which will allow deducing the cost penalty of the FDD tool associated with its minimum detection sensitivity level; and (6) operator labor costs in order to assign cost to false positives of the FDD tool.

b. **Performance criteria of FDD tool** (some of the criteria will depend on specific faults and severity levels) are divided into qualitative and quantitative criteria:
   - *Qualitative criteria*: simplicity in understanding and using it, comprehensive in terms of being able to identify all important faults, ability to rank possible faults if they cannot be uniquely identified, ability to learn online and gradually increase its sensitivity and robustness, ability to adapt to changing building and system operation over time, and ability to handle simultaneous faults.
   - *Quantitative criteria* (at different fault severity levels): correct detection fraction during fault-free operation (or low false positives or false alarms), correct detection fraction during faulty operation (or low false negatives), correct diagnosis fraction, non-unique diagnosis fraction (diagnosis rules are not able to distinguish between more than one possible fault), unable to diagnose fraction (observed fault patterns do not correspond to any rule within the diagnosis rules), minimum detectable fault severity level (or sensitivity level), rapidity (i.e., minimum time required to identify a fault from its onset—applies to abrupt faults only), fault evaluation and prognostic capability in suggesting course of action.

c. **Implementation costs, training and maintenance criteria of FDD tool** are divided into qualitative and quantitative criteria:
   - *Qualitative criteria*: ease in modifying capabilities of tool or flexibility if system operation changes, ease of integration with existing automation system or controllers, ease of transportability, ease of routine maintenance either by FDD developer or site engineer, basis of fault-free and faulty data (manufacturer's performance data or a lumped chiller model that needs to be calibrated with limited in-situ performance data).
   - *Quantitative criteria*: initial cost of tool (includes licensing cost), cost of additional sensors if needed, cost of more accurate sensors than existing ones if needed, cost of implementation and commissioning (fault detection thresholds may have to be tuned manually), cost of training technician and site engineer to use tool, cost due to delayed benefit of tool in case the tool needs a training period, cost in operator time and repairs to maintain the tool over time, operator intervention cost due to false positives, savings due to reduced energy use, savings due to reduced maintenance cost, benefit due to improved comfort and enhanced productivity.

d. **Testing sequence and cost benefit analysis.** This last step is assumed to occur on a separate test bed or on an emulator where different operating states can be simulated with different fault severities of specific faults prior to the client purchasing the FDD tool. This would be similar to a certified testing process currently adopted before purchasing large equipment such as chillers (Corcoran and Reddy 2003). How to specify these fault severity levels and the test sequence also needs to be determined. How to introduce different types of faults (abrupt or gradual or cyclic) is also a relevant issue.

## FDD EVALUATION PROCEDURE

The list of categories and criteria described in the previous section is a first attempt at compiling a comprehensive list. Many of these pertain to site-specific issues, which are beyond the purview of, say, an independent FDD tool evaluation agency performing the tests on a laboratory test stand. Consequently, it would be useful to select a small sub-set of *quantitative* criteria against which to evaluate different FDD methods. Consider the FDD tool as one that when in operation first sorts or flags incoming system performance data into either fault-free or faulty categories (Himmelblau 1978) to which we have added further subdivisions as described below (see Figure 1).

a. **False negative rate** denotes the probability of calling a faulty process good, i.e., missed opportunity loss (Type II error).
b. **Correct fault-free detection rate** denotes the probability of calling a good process good.
c. **False positive rate** denotes the probability of calling a good process faulty, i.e., false alarm (Type I error).
d. **Correct faulty detection rate** denotes the probability of calling a faulty process faulty.

Note that in case only one crisp fault detection threshold value is used, probabilities (a) and (d) should add to unity, and so should (b) and (c). Thus, if the correct fault-free detection rate is signaled by the FDD tool as 95%, then the corresponding false alarm rate is 5%. Once a fault has been detected correctly, there are four possibilities, each of which has a cost implication (in terms of technician's time to deal with it (see Figure 1):

- **Correct and unique diagnosis,** where the fault is correctly and unambiguously identified.
- **Correct but not unique,** where the diagnosis rules are not able to distinguish between more than one possible fault.
- **Unable to diagnose,** where the observed fault patterns do not correspond to any rule within the diagnosis rules.
- **Incorrect diagnosis,** where the fault diagnosis is done improperly.

The evaluation of different FDD methods will consist of two distinct aspects:

a. *How well is fault detection done*? Given the rudimentary state of practical implementation of FDD methods in chillers, it would suffice for many service companies merely to know whether a fault has occurred; they would then send a service technician to diagnose and fix the problem.
b. *How well does the FDD methodology perform overall*? Each of the four diagnosis outcomes as shown in the flowchart of Figure1 will affect the time needed for the service technician to verify the suggested diagnosis of the fault (or to diagnose it).

### Evaluation of Different Methods Based on Fault Detection Consideration

Varying the fault detection threshold affects the sensitivity of detection, i.e., the "Correct Fault-Free Detection Rate" and the "False Positive Rate" are affected in compensatory and opposite ways. Since these rates have different cost implications, one cannot simply optimize the total error rate. Instead, FDD evaluation can be stated as a minimization problem

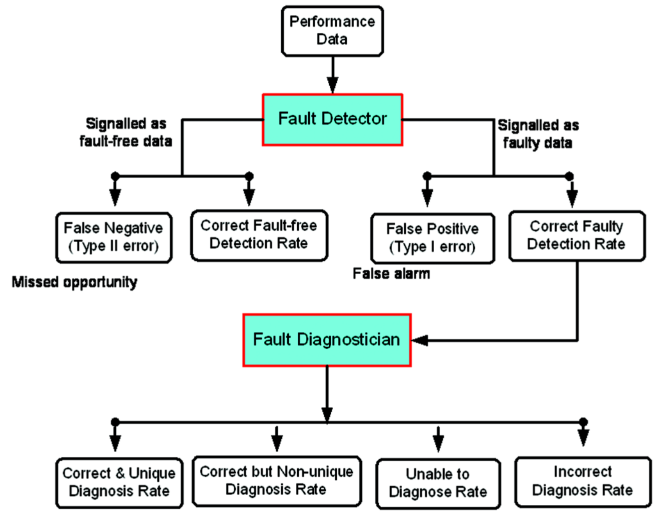


***Figure 1*** *Evaluation procedure for detecting and diagnosing faults.*

where one tries to minimize the sum of penalties associated with false positives (which require a service technician to needlessly respond to the service call) against the extra operating cost incurred due to excess energy use arising from the onset of different faults. The frequency of occurrence of different faults over a long enough period (say, one year of operation) needs to be explicitly considered as well as their energy cost penalties. Such considerations lead to the formulation of the following objective function, which is applicable to any HVAC&R system:

$$J = \min \{J_1 + J_2 + \dots\}$$
$$\text{or } J = \min \left\{ P_0 \cdot F_P \cdot C_{FP} + c_e \sum_{f=1}^{N_F} (P_f \cdot \Delta E_f \cdot F_{N,f} \cdot \Delta t_f) \right\} \quad (1)$$

where

$J_1$ = cost penalty due to false positives

$J_2$ = cost due to increased energy use as a result of missed detection of faults (due to false negatives); other cost penalty factors such as increased electric demand, and other non-energy-related issues (such as loss of cooling capacity, increased maintenance, reduced component life) are neglected in this simplified formulation

$c_e$ = cost of electric energy

$C_{FP}$ = cost of a technician's time to verify the complete system in response to a false alarm

$F_{N,f}$ = false negative rate for fault $f$ (or missed opportunity rate)

$F_P$ = false positive rate (or false alarm rate)

$f$ = index for fault type

$N_F$ = total number of possible faults in system

$P_f$ = probability of occurrence of fault type $f$

$P_0$ = probability of occurrence of no-fault (i.e., fractional time of chiller operation during a given period, say, a year, during which the chiller is fault free)

$\Delta E_f$ = extra electric power required to provide necessary cooling due to performance degradation as a result of fault type f

$\Delta t_f$ = time period (hours) in which the fault f has gone undetected or un-rectified

Alternatively, assuming $\Delta t_f$ to be the same for all faults, we could re-express Equation 1 as

$$\text{or } J' = \min\left\{P_0 \cdot F_P \cdot \frac{C_{FP}}{c_e \cdot \Delta t} + \sum_{f=1}^{N_F} (P_f \cdot \Delta E_f \cdot F_{N,f})\right\} . \qquad (2)$$

It is difficult to assign realistic values for many of the quantities appearing in the first term of the above objective function. For example, the time for the fault to go unrectified is dependent on the particular installation. Further, the optimal threshold values may be specific to the FDD method. A practical manner to avoid dealing with such issues is to compare the various FDD methods with **detection thresholds adjusted, so that they have the same false positive or false alarm rate** (say, $F_P$ = 0.05, or 5%). We can re-express the objective function as an FD evaluation criterion of a specific FDD method as

$$\Phi'_{Detect,s} = \sum_{f=1}^{N_F} [P_f \cdot \Delta E_f \cdot (1 - F_{N,f})] . \qquad (3)$$

Note that the second term on the right-hand side of Equation 2, which expressed the cost penalty in terms of the false negative rate, is now to be interpreted in terms of its complement, namely, the correct fault detection rate $(1 - F_{N,f})$.

A better way is to normalize the **fault detection criterion** so as to obtain a **normalized score,** or rank between 0 and 1 (where the basis of evaluation is with respect to an ideal detector with score of unity and with no false alarms whatsoever):

$$\Phi_{Detect,s} = \sum_{f=1}^{N_F} [P_f \cdot \Delta E_f \cdot (1 - F_{N,f})] / \sum_{f=1}^{N_F} (P_f \cdot \Delta E_f) \qquad (4)$$

### Evaluation of Different Methods Based on Overall FDD Capability

Fault diagnosis involves four quantities or fractional rates as described above. Each of these has a different implication in terms of the time taken (and, hence, the cost) for the technician or the serviceman to diagnose the fault and then evaluate its implication and chose an appropriate course of action. The **normalized overall FDD score criterion** is similar to Equation 4 with, however, weighting given to the four different outcomes at the fault diagnosis stage:

$$\Phi_{FDD} = \sum_{f=1}^{N_F} [P_f \cdot \Delta E_f \cdot (w_{cu} \cdot r_{cu,f} + w_{cn} \cdot r_{cn,f} + w_{ic} \cdot r_{ic,f} + w_{ud} \cdot r_{ud,f})] \Big/ \sum_{f=1}^{N_F} (P_f \cdot \Delta E_f) \qquad (5)$$

where

$r_{cu}$ = correct and unique diagnosis rate expressed as a fraction of the signaled faulty data,

$r_{cn}$ = correct but non-unique diagnosis rate,

$r_{ic}$ = incorrect diagnosis rate,

$r_{ud}$ = unable to diagnose rate,

$w_{cu}$ = weighting factor for correct and unique diagnosis rate (same for each fault type),

$w_{cn}$ = weighting factor for correct but non-unique diagnosis rate,

$w_{ic}$ = weighting factor for incorrect diagnosis,

$w_{ud}$ = weighting factor for inability to diagnose

and other terms as defined in Equation 1.

## PRACTICAL FEEDBACK FROM A LARGE CHILLER MANUFACTURER AND SERVICE COMPANY

This section summarizes discussions with the engineering and service managers of large-capacity chiller units of a major company on issues such as probability of occurrence of different faults, energy implications of various faults, and what other issues need to be considered during evaluation of different FDD methods.

a. **It is difficult to generalize the effect of different faults on centrifugal chillers.** The specific chiller design, selection of specific components, type of control—all have an important effect on the onset of faults in a chiller. Consider ASHRAE research project RP-1043 where fault-free and faulty operation of a chiller under different types and fault severities was experimentally determined in laboratory tests (Comstock and Braun 1999). The types of faults studied and their levels of severity and the numerous tests performed are summarized in Table 1. The laboratory chiller was a 90-ton centrifugal chiller with thermostatic expansion valve (TXV) control. Such a control mechanism regulates the refrigerant supply to the evaporator, which under part-load operation would result in refrigerant backing up in the evaporator and some subcooling capability being lost. However, large chillers, in general, are more often regulated with an orifice with inlet guide vane control. In such chillers, the orifice control is based upon condenser refrigerant level, and the very top of the evaporator bundle becomes uncovered during part-load condition. Under faulty operation, the exact design and component selection is important. For large chillers under orifice control, loss of refrigerant will cause loss of capacity only if the amount lost is large enough to uncover the evaporator tubes. If this occurs

**Table 1. Summary of RP-1043 Lab Chiller Data sets (Comstock and Braun 1999)**

| | Description of Fault | Normal Operation | SL1 | SL2 | SL3 | SL4 |
|---|---|---|---|---|---|---|
| 1 | Reduced condenser-water flow | 270 gpm (0.98–1.0) | 0.87–0.93 | 0.77–0.81 | 0.69–0.70 | 0.59–0.61 |
| 2 | Reduced evaporator-water flow | 216 gpm (0.99–1.0) | 0.90–0.91 | 0.81–0.82 | 0.72–0.72 | 0.63–0.65 |
| 3 | Refrigerant leak | 300 lb | 0.1 | 0.2 | 0.3 | 0.4 |
| 4 | Refrigerant overcharge | 300 lb | 0.1 | 0.2 | 0.3 | 0.4 |
| 5 | Excess oil | 22 lb or 2.75 gal. | 1.14 | 1.32 | 1.5 | 1.68 |
| 6 | Condenser fouling | 164 tubes in total | 0.06 | 0.12 | 0.20 | 0.30 |
| 7 | Noncondensables in system (by volume) | No nitrogen | 0.01 | 0.017 | 0.024 | 0.057 |

**Note:** Fractional values indicate the level of fault severity. For example, the range 0.59–0.61 under SL4 for reduced condenser water flow indicates that the flow was reduced to about 60% of the normal value.

during part-load condition and if the evaporator bundle were slightly oversized, the effect on power may be minimal. In the RP-1043 lab chiller, on the other hand, flow regulation is being achieved by the TXV, and during a refrigerant loss condition, the subcooler bundle is very likely to become uncovered, thereby resulting in a loss of subcooling (effect on power), or, under an extreme condition, condenser gas could by-pass through the orifice, resulting in higher power and a loss of capacity. Another example has to do with fault implication due to cooling water flow rate reduction. One can select different sizes of heat exchangers (HX) for the condenser and evaporator for the same centrifugal compressor. These HX are designed for internal water flow velocities between 3 and 12 ft/sec. If one picked an HX with the higher flow rate, reduction in condenser water flow (or evaporator water flow) is likely to have much less adverse effect since it is still likely to be in turbulent regime as compared to the case when an HX with a lower design water flow were initially selected.

b. **Fault severity levels are manufacturer-specific.** For example, certain chiller designs cannot be overcharged by 20% excess refrigerant (the RP-1043 chiller data include 40% refrigerant overcharge tests). The same is the case with refrigerant loss—surging and other undesirable phenomena would occur even when 20% refrigerant loss occurs in such machines. The importance of the noncondensables in system fault would depend on whether it is a low- or high-pressure machine (i.e., on the working refrigerant used).

c. **Thermal load under which equipment is operated is a very important factor during faulty operation**. The energy penalties are likely to be very much different, with maximum penalty due to faults occurring at maximum cooling load.

d. **Loss of cooling capacity is too intangible and plant-specific.** Though this is an important issue, it is difficult to quantify this effect and include it in the FDD evaluation methodology since it depends too much on the annual cooling load profiles of the buildings being served, the type of building (how critical it is to meet the desired cooling load), and specific cooling plant design (whether multiple chillers are used and what mix). Given so many site-specific factors, it is better not to include this effect in the FDD methodology evaluation.

e. **RP-1043 fault survey results** (Comstock et al. 1999) were done mainly from the point of view of maintenance issues (probability of occurrence, cost implications, etc.) and energy expense was not included. Hence, these results could not be directly used in our research. Most of these types of maintenance faults are difficult to identify by monitoring; for example, motor burnout is difficult to detect by thermal/pressure measurements. Hence, it was concluded that we limit the current evaluation to include energy savings benefit only.

f. **Heuristic ranking for frequency and importance/impact of different faults.** Table 2 assembles numerical values to rank the frequency of occurrence of different faults (related to probability of occurrence $P_f$) in large chillers and their importance/impact in terms of associated effort for remedial action and occurrence frequency. The higher the numerical value of the frequency rank, the lower the critical nature of that fault. Thus, condenser fouling (with a ranking of 1) is to be interpreted as the most critical in terms of cost of repair, with refrigerant overcharge (with a rank of 6) as the least critical, since this is a fault that ought to be detected at start-up commissioning and not during routine operation.

g. **Diagnosis occurrences**. The FDD evaluation methodology was discussed, and it was deemed acceptable. The

**Table 2. Ranking of Different Faults Following Discussions with Personnel from a Large Chiller Company**

| Fault # | Description | Impact Rank* | Frequency Rank* |
|---|---|---|---|
| F1 | Reduced condenser-water flow | 3 | 4 (1/yr) |
| F2 | Reduced evaporator-water flow | 4 | 3 (1–3/yr) |
| F3 | Refrigerant leak | 2 | 1 (1–2/yr) |
| F4 | Refrigerant overcharge | 6 | 6 (never) |
| F5 | Condenser fouling | 1 | 2 (1/yr) |
| F6 | Noncondensables in system | 5 | 5 (1/yr or never) |

* 1—highest rank order.

distinction between fault detection and diagnosis was an important and relevant one. Further, it was felt that the four possible outcomes of diagnoses were also relevant and would have practical associations. The weights for the four possible diagnosis outcomes would depend on how quickly the operator is able to diagnose the fault. Since unique and correct diagnosis is the best outcome, we shall assume a value of unity and reduce the scores for the other three. The weight numbers that were suggested are assembled in Table 3.

h. **Energy penalty of different faults.** Service personnel from the large chiller manufacturing company kindly agreed to use an internal software program to simulate energy use under various faults at different fault severities and at the 27 operating points similar to those adopted during RP-1043 lab chiller tests. This allowed energy penalties under different faults to be determined. Table 4 assembles the results of energy penalties at the three cooling load conditions assumed for the RP-1043 lab chiller, corresponding to the highest fault severity level SL4. This table is a convenient way of comparing the results from simulations from the large chiller company and those of the RP-1043 lab chiller. We note that agreement is fairly good, except for the low charge case, where the lab chiller indicates an increase in COP (since the values are less than 100%).

Table 5 assembles our suggested values for use in the FDD evaluation for excessive electric energy use (in percentage) of different faults at different severity levels (as described in Table 1) along with their frequency of weights, which are numerical values inversely related with the frequency rank shown in Table 2. However, it is important to note that these are preliminary suggestions and are bound to be different for different generic chiller types and sizes. In any case, whatever FDD method is used, a certain amount of customization to the specific chiller system would be required in the light of the issues discussed above. An important issue impacting experimental design can be gleaned from Table 5. We note that the severity levels selected by RP-1043 for faults F2–F5 are too small in terms of the associated energy penalties. An experimental design where the fault severity levels were selected so as to result in a preselected energy penalty would have been the better procedure instead of the current one, which was based on physical considerations (such as a predetermined decrease in fluid flow rate or a certain percentage of condenser tube blockage to mimic fouling).

**Table 3. Suggested Values of the Weights for the Four Diagnostic Outcomes After Discussions with Service Personnel of a Large Chiller Company**

| Diagnosis Outcome | Symbol | Weights |
|---|---|---|
| Correct and unique diagnosis | $w_{cu}$ | 1.0 |
| Correct but non-unique diagnosis | $w_{cn}$ | 0.75 |
| Unable to diagnose | $w_{ud}$ | 0.5 |
| Incorrect diagnosis | $W_{ic}$ | 0.4 |

## SUMMARY

This paper began by providing a thorough literature review of published work on evaluation of FDD tools, suggested a structure to the various relevant issues in terms of assigning them to one of four categories (element-defining site-specific criteria; element-defining performance criteria; element-defining implementation costs, training, and maintenance; and elements related to sequence and cost benefit analysis). It was also pointed out that much of the needed information to use the above criteria is lacking at this time. An analytical method whereby the FDD evaluation is cast into an objective function of two competing considerations (cost associated with false alarms and penalties associated with the onset of faults) is also proposed. We point out that, from a practical viewpoint, FDD evaluation should be based on two criteria: (1) the normalized fault detection criterion, which allows different methods to be evaluated based on their sensitivity, i.e., their ability to correctly detect faults without exceeding a pre-stipulated false alarm rate, and (2) the combined fault detection and fault diagnosis criterion, where in addition we consider four different diagnostic outcomes, all of which have different practical consequences. The general

**Table 4. Comparison of Excess Energy Consumed as Predicted by the Simulation Program of a Large Chiller Company and the RP-1043 Lab Chiller at the Highest Fault Severity Conditions**

| Chiller Load | Rated Load % | Reduced Condenser Flow | Reduced Evaporator Flow | Low Charge | Condenser Fouling |
|---|---|---|---|---|---|
| **Chiller Company Simulations** | | **SL4** | **SL4** | **SL4** | **SL4** |
| 500 T | 100 | 108.93 | 100.27 | 101.22 | 101.24 |
| 333 | 66.6 | 104.96 | 100.13 | 100.67 | 100.67 |
| 167 | 33.4 | 102.24 | 100.33 | 100.43 | 100.54 |
| **RP-1043 Chiller** | | | | | |
| 90 ton | 100 | 109.53 | 100.41 | 97.30 | 102.73 |
| | 66 | 106.52 | 101.92 | 98.81 | 102.6 |
| | 33 | 98.90 | 100.18 | 97.66 | 99.68 |

**Note:** A value of 102.27 implies that the electric use increased by 2.27% w.r.t fault-free use.

**Table 5. Values Suggested for Use in Chiller FDD Evaluation for Different Faults and Fault Severities (as Described in Table 1)***

| | Fault | Energy Penalty ($\Delta E_f$) SL1% | SL2% | SL3% | SL4% | Frequency Weights ($P_f$)‡ |
|---|---|---|---|---|---|---|
| F1 | Reduced condenser-water flow | 0.70 | 1.9 | 3.0 | 5.3† | 1.0 |
| F2 | Reduced evaporator-water low | 0.0 | 0.0 | 0.40 | 0.90 | 2.0 |
| F3 | Refrigerant leak | 0.14 | 0.31 | 0.47 | 0.71 | 1.5 |
| F4 | Refrigerant overcharge | 0.80 | 0.94 | 3.8 | 7.6 | 0.25 |
| F5 | Condenser fouling | 0.50 | 0.50 | 0.50 | 1.8 | 1.0 |
| F6 | Noncondensables in refrigerant | 4.5 | 6.2 | 7.4 | 15.6 | 0.5 |

* These values were selected based on feedback from a large chiller manufacturer and from the RP-1043 lab chiller data.
‡ These weights are heuristic and have been selected after discussions with service personnel of a large chiller company.
† A value of 5.3 implies that the electric use increased by 5.3% w.r.t fault-free use.

formulation has been simplified to apply specifically to chillers and to the case where the penalties of fault occurrence are limited to excess electric energy consumption to meet the stipulated cooling load. It is suggested that a reasonable approach is to first tune the detection thresholds of the relevant measurement features for each of the various FDD methods being evaluated using fault-free data so that they have the same false alarm rate and then evaluate them separately on their detection capability and their diagnostic capability for different faults and fault severity levels. Another paper (Reddy 2007) will describe its application and usefulness to assess four chiller FDD methods customized to the same performance data sets.

A special effort was made to determine the types of penalties associated with various faults in chiller installations (energy increase, loss of cooling capacity, reduced life, etc.). After discussion with service personnel of a large chiller company, it was decided that an initial practical choice would be to limit the FDD evaluation to the energy penalty alone. Data from the RP-1043 lab chiller, as well as simulations from an in-house computer program of a large chiller manufacturer, have been analyzed, which in conjunction with personal discussions with the service managers of a large chiller company, yielded preliminary, but realistic, values of energy penalties that could be used in the evaluation of chiller FDD tools. The ranks or weights associated with the occurrence frequency of various faults, as well as the weights for four possible outcomes of a fault diagnosis, were also determined based on personal discussions. Such an evaluation methodology can be used by a particular FDD tool manufacturer or an independent evaluation agency to frame the sequencing of its test results and to subsequently analyze the data in a manner that will provide a single quantitative index of how well the FDD method or tool fares with respect to an ideal one.

## ACKNOWLEDGMENTS

This research was sponsored by ASHRAE TC 7.5, Smart Building Systems. Constructive comments and suggestions by the PMSC members are greatly appreciated. We acknowledge advice and helpful feedback from Ashish Singhal, John Seem, Bill McQuade, Justin Kauffman, Dennis Dietz, and John House.

## NOMENCLATURE

| | | |
|---|---|---|
| $C_{FP}$ | = | cost of a technician time in verifying the complete system in response to a false alarm |
| COP | = | coefficient of performance of chiller |
| $c_e$ | = | cost of electric energy |
| $E$ | = | electric power input to compressor motor |
| $F_{N,f}$ | = | false negative rate for fault $f$ |
| $F_P$ | = | false positive rate |
| $f$ | = | index for fault type |
| $J_1$ | = | cost penalty due to false negatives (technician's time to verify the system) |
| $J_2$ | = | energy cost due to increased energy use as a result of missed detection of faults (due to false negatives) |
| $N_F$ | = | total number of possible faults in system |
| $P_f$ | = | probability of occurrence of fault type $f$ |
| $P_0$ | = | probability of occurrence of no-fault |
| $Q$ | = | thermal heat load or capacity |
| $r_{cu}$ | = | correct and unique diagnosis rate expressed as a fraction of the signaled faulty data |
| $r_{cn}$ | = | correct but non-unique diagnosis rate |
| $r_{ic}$ | = | incorrect diagnosis rate |
| $r_{ud}$ | = | unable to diagnose rate |
| $s$ | = | index for fault severity level |
| $w_{cu}$ | = | weighting factor for correct and unique diagnosis rate (same for each fault type) |
| $w_{cn}$ | = | weighting factor for correct but non-unique diagnosis rate |
| $w_{ic}$ | = | weighting factor for incorrect diagnosis |
| $w_{ud}$ | = | weighting factor for inability to diagnose |
| $x$ | = | regressor variable |
| $y$ | = | response variable |
| $\Delta E_f$ | = | extra electric power required to provide necessary cooling due to performance degradation as a result of fault $f$ |
| $\Delta t$ | = | time period (hours) for which the fault has gone undetected or un-rectified |

### Subscripts

| | | |
|---|---|---|
| $cd$ | = | condenser |
| $ch$ | = | chiller, evaporator |
| $ev$ | = | evaporator |
| $r$ | = | refrigerant |

## REFERENCES

ASHRAE. 2001. *ANSI/ASHRAE Standard 140-2001, Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs*. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

Comstock, M.C., B. Chen, and J.E. Braun. 1999. Literature review for applications of fault detection and diagnostic methods to vapor compression cooling equipment. HL 99-19: Report #4036-2, Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN.

Comstock, M.C., and J.E. Braun. 1999. Development of Analysis tools for the evaluation of fault detection and diagnostics in chillers. ASHRAE Research Project RP-1043, also HL 99-20, Report #4036-3, Ray W. Herrick Laboratories, Purdue University, West Lafayette, IN.

Corcoran, J.P., and T.A. Reddy. 2003. Improving the process of certified and witnessed factory testing for chiller procurement. *ASHRAE Transactions* 109(1):193–201.

Dexter, A., and J. Pakanen, eds. 2001. Demonstrating automated fault detection and diagnosis methods in real buildings. International Energy Agency Annex 34 Final Report, Technical Research Center of Finland, Espoo, Finland.

Gruber, P. 2001. Section E2, Feedback from industrial partners in national projects including issues affecting commercialization. in A. Dexter and J. Parkanen (eds.), Demonstrating automated fault detection and diagnosis methods in real buildings, International Energy Agency Annex 34 Final Report, Technical Research Center of Finland, Espoo, Finland.

Himmelblau, D.M. 1978. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*. Amsterdam: Elsevier Scientific Publishing Co.

House, J.M., J.E. Braun, T.M. Rossi, and G.E. Kelly. 2001. Section D, Evaluation of FDD Tools. In A. Dexter and J. Parkanen, eds., Demonstrating automated fault detection and diagnosis methods in real buildings. International Energy Agency Annex 34 Final Report, Technical Research Center of Finland, Espoo, Finland.

Katipamula, S., R.G. Pratt, and J.E. Braun. 2001. Building systems diagnostics and predictive maintenance. In *CRC Handbook for HVAC&R Engineers*, pp. 7-25–7-85J, J. Kreider, ed. Boca Raton, FL: CRC Press.

Norford, L.K., J.A. Wright, R.A. Buswell, and D. Luo. 2000. Demonstration of fault detection and diagnosis methods in a real building. Final report of ASHRAE RP-1020, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Atlanta, GA.

Patel, S.A., and A.K. Kamrani. 1996. Intelligent decision support system for diagnosis and maintenance of automated systems. *Computers and Industrial Eng*. 30(2):297–319.

Reddy, T.A. 2006. Evaluation and assessment of fault detection and diagnostic methods for centrifugal chillers—Phase II. Final project report of ASHRAE RP-1275, June, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Atlanta, GA.

Reddy, T.A. 2007. Application of a generic methodology for assessing four chiller FDD methods (RP-1275). *HVAC&R Research* 13(4).