Purdue University

Purdue e-Pubs

International Refrigeration and Air Conditioning Conference

School of Mechanical Engineering

2014

Validation of a Fault-Modeling Equipped Vapor Compression System Model Using a Fault Detection and Diagnostics Evaluation Tool

David P. Yuill
Purdue University, United States of America, dyuill@purdue.edu

Howard Cheung
Purdue University, United States of America, cheung@purdue.edu

James E. Braun
Purdue University, United States of America, jbraun@purdue.edu

Follow this and additional works at: http://docs.lib.purdue.edu/iracc

Validation of a Fault-Modeling Equipped Vapor Compression System Model Using a Fault Detection and Diagnostics Evaluation Tool

David P. YUILL*, Howard CHEUNG and James E. BRAUN

Ray W. Herrick Laboratories, Purdue University, Mechanical Engineering
West Lafayette, IN, USA
(765) 494-2132, dyuill@purdue.edu

* Corresponding Author

ABSTRACT

A methodology for evaluating the performance of fault detection and diagnostics (FDD) tools for unitary air-conditioners has been developed (Yuill and Braun 2013). The methodology uses laboratory measurements of systems with and without faults to test FDD tools' effectiveness. A gray box modeling method capable of modeling systems with faults was developed by Cheung and Braun (2013a and 2013b) to provide input data, as an alternative to using laboratory data that had been collected. The simulation method was validated by direct comparison with experimental data, but a comparison of FDD evaluation results provides a more direct and useful validation of the model for its intended purpose. Eight different systems have been modeled using Cheung and Braun's method. Six FDD tools were evaluated using both experimental and modeled inputs under the same environmental and fault conditions. The fault conditions include non-standard charging, heat exchanger fouling, loss of compressor volumetric efficiency, liquid line restriction, and the presence of non-condensable gas in the refrigerant. The model's performance is characterized by comparing its outputs from the evaluation – false alarm rates, misdiagnosis rates, missed detection rates, and rates of undiagnosed faults – with the results based upon experimental data. The model is found to be highly suitable for its purpose.

1. INTRODUCTION

A methodology for evaluating the performance of fault detection and diagnostic (FDD) protocols applied to vapor compression unitary air conditioning equipment is presented by Yuill and Braun (2013). The paper describes the methodology, and a specific method in which data from laboratory measurements are fed to the candidate FDD protocol, and the responses are gathered and organized. It also provides the motivations for developing the method, and presents a case study of a widely used protocol. In a companion paper to the current paper, Yuill *et al.* (2014) summarize the original method, and present a significant modification to the FDD evaluation method. The modification is to use simulation data to provide inputs for evaluation, rather than using the laboratory measurement data.

The simulation data that are used come from gray-box models developed by Cheung and Braun (2013a, 2013b) expressly for this purpose. These multiple input, multiple output models are semi-empirical and represent steady-state operation. The models simulate each component of the system (condenser, evaporator, compressor, expansion device and piping) separately and pass outputs from one component to be inputs to the next. Wherever appropriate, the submodels are physics-based, but for the more complex components, such as heat exchangers, performance parameters are learned using data from the laboratory experiments described by Yuill and Braun (2013). Furthermore, physical constraints are imposed at several points in the system to further add to the realism of the models. One reason that this is necessary is that the measurements contain some level of error, both bias and random, that could combine with other parts of the model to generate non-physical results at the component level.

Some examples of constraints include energy balances, mass flow continuity, minimum superheat and subcooling values of zero, and maximum relative humidity of 100%. Inverse models do not contain such constraints.

For any model to be relied upon, it must be verified (checked to ensure that the model is correctly implemented) and validated (checked to ensure that the model represents the phenomenon it simulates with some acceptable level of accuracy for its intended application). Verification and validation are particularly important in the current case, for two reasons. First, the modeling methodology is novel, so that both the model and the modeling approach have the potential for errors. Second, the purpose for the model is to evaluate the performance of FDD protocols. The outcome of these evaluations may have large repercussions for the protocols' developers and for the overall market for FDD tools, so the cost of modeling error could be quite large. Furthermore, the FDD protocols are very sensitive to variances.

There is no single prescribed method for validating a model. Roach (2009) and Oberkampf and Roy (2010) provide comprehensive discussions on verification and validation in general. Many approaches can be used, and typically verification and validation are conducted iteratively along with model improvements. This iterative process and related terminology are described by Schlesinger (1979) and Thacker *et al.* (2004), as summarized in Figure 1.
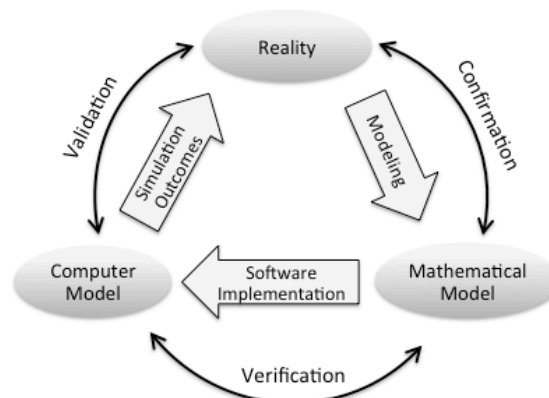

Figure 1: Modeling, verification and validation process

There are six methods that have been employed in the current project to test the simulation. These methods are listed below and each is associated with the corresponding parts of the process described in Figure 1. These six methods are described in section 2: Verification and Validation Process.

1. Third party description – confirmation and verification
2. Degeneracy testing – verification
3. Expert intuition – confirmation, verification and validation
4. Real system measurements – confirmation, verification and validation
5. Self consistency check – verification and validation
6. FDD evaluation comparison – confirmation, verification and validation

2. VERIFICATION AND VALIDATION PROCESS

The six methods used in the verification and validation of the gray-box models of Cheung and Braun (2013a, 2013b) are described below.

1. Third-party description: This process, sometimes referred to as "structured walk-through" or "step-by-step analysis", consists of the modeler describing the modeling approach to expert third parties for technical critique. It covers the choice of models (confirmation) and the mathematical and programming approach for implementing these models (verification). In the current case, the modeler described details both orally and in written form on a weekly basis to his academic advisor, biennially to his research sponsor, and as-needed to his colleague. Besides adding third-party input, this process caused the modeler to focus on his own approach from a different perspective.

2. Degeneracy testing: In this process, a model is run with inputs from the extremes of the intended input space, primarily to test the stability of the model. In the current case, it is allowable for the model to fail occasionally because it is intended to generate a data library in which a uniform distribution of modeled points is not necessary. There are, therefore, numerous cases in which the model fails to converge, or in which components (sub-models) fail to converge. These cases are simply abandoned. However, when degeneracy testing uncovered entire regions of model failure, then the model was modified.

3. Expert intuition: Since the process being modeled is physical – a vapor compression cycle – and the modeler and his colleagues are knowledgeable about the process, simulation outputs were studied to see whether they gave expected results. If they did not, then either the mathematical model or the software implementation had a problem. For example, it was found in many cases that an increase of inlet subcooling at the expansion valve led to a decrease of refrigerant mass flow rate. This is not realistic, so the parameter estimation process was modified so that the model always predicts an increase in refrigerant mass flow rate with an increase of inlet subcooling.

4. Real system measurements: Comparison of simulation outputs with experimental data is a powerful and most commonly used approach whenever experimental data are available. Discrepancies highlight the existence of a problem, but they don't necessarily indicate whether it's a problem with the selected model or the implementation of the model. Furthermore, the discrepancy may be caused by a problem with the measurement data. In the current case, since the model has inversely modeled components, the comparisons with measurement data were used in model verification, both at the component level and system level, and model validation. Direct comparison of several outputs, such as coefficient of performance (COP) and suction superheat, are shown in Cheung and Braun (2013a).

5. Self consistency check: Models of physical systems tend to contain continuous well-behaved functions; models can be expected to provide similar outputs for similar inputs. In the current model, except at cusp-points, the functions tend to be very well behaved – smooth and linear. (The cusp points are the points at which (a) superheat or subcooling decrease to zero; (b) the air-side of the evaporator begins to condense water. These points were given special consideration. They were found and specifically modeled, along with several points nearby. Cusp points are singular, but simpler to model physically, than non-cusps.) By plotting various combinations of output versus input, unexpected behaviors were identified, investigated, and addressed appropriately.

6. FDD evaluation comparison: The most important, but sometimes overlooked aspect of model validation is the last part of the definition given on page 2: the model must be suitable "for its intended application". In the current case, the model's purpose is to provide input data for use in evaluation of FDD protocols. Therefore, the ultimate validation is a comparison of the results of evaluating a sample of FDD protocols with: (a) measurement data, and (b) data generated by the model. This specific approach is novel, and is the primary focus of the current paper.

2.1 FDD Evaluation as a Basis for Model Validation
The FDD evaluation method developed by Yuill and Braun (2013) has been coded for implementation in a software package, *FDD Evaluator* (Yuill and Braun 2014). This software has been verified by replicating it in two different computing environments and comparing results. The software contains an input data library, which it feeds through the candidate FDD protocol then collects the protocol's outputs, comparing them to the reference condition in each case, and organizing the results as a function of fault impacts. To validate the gray-box model of Cheung and Braun (2013a, 2013b), *FDD Evaluator* has been modified to include special input data libraries, and six FDD protocols have been evaluated with each of these libraries.

To validate the model by comparing FDD evaluation results requires a direct comparison. To do this, the model is run using the measured laboratory test conditions as the inputs (independent variables). The five independent variables are: T_{amb}, T_{ra}, WB_{ra}, fault type (ϕ), and fault intensity (*FI*). For example, if a system was tested in the laboratory with T_{amb} = 35.1°C, T_{ra} = 26.8°C, WB_{ra} = 19.5°C, and 80% of nominal condenser airflow then the model was run with these same conditions. There are models for eight different systems, so this process was repeated for each. The measurement data library contained in *FDD Evaluator* contains 607 tests from nine systems, but one of these systems couldn't be modeled because it had an insufficient range of test conditions to train the model. Of the eight remaining systems, there were several cases for which the model didn't converge. To make an even

comparison, the *FDD Evaluator*'s measurement data library was reduced to contain only those cases for which a model output exists. There were 536 such cases.

3. GRAY BOX MODEL

The gray-box model described in Cheung and Braun (2013a, 2013b) is used to model the impact of faults on different vapor compression systems accurately and quickly. The model is built by first dividing a vapor compression system into multiple major components as shown in Figure 1.
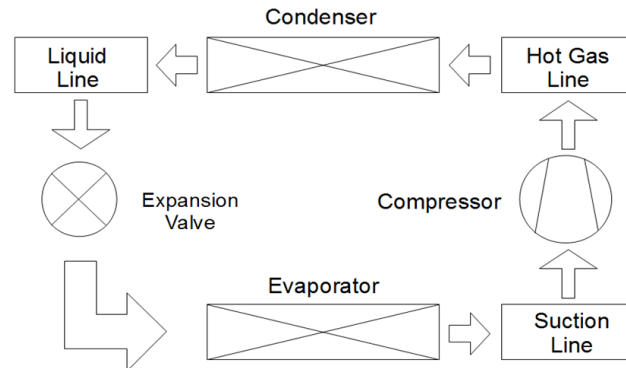

Figure 2: Schematic of component models in the gray-box system model (Cheung and Braun 2013a)

Gray-box component models are used to characterize component behavior, which follow the laws of physics and expert intuition and have their parameters estimated by minimizing the differences between the outputs of the models and valid experimental observations of normal and faulted vapor compression systems. The model outputs include pressures and temperatures of refrigerant between each of the components in Figure 2, air temperature and humidity leaving the evaporator, air temperature leaving the condenser, power consumption of the compressor and of the whole system, refrigerant mass flow rate, and several convenient secondary outputs such as capacity, COP, suction superheat, subcooling, etc.

After the parameter estimation process, the component models are combined together to form an implicit model of a vapor compression system. The system model is further tuned with experimental observations so that it estimates the amount of charge inside the system without bias, and the amount of charge can then be used as an input to the model. Simulating faults – non-standard charging, heat exchanger fouling, loss of compressor volumetric efficiency, liquid line restriction and presence of non-condensable gas in the refrigerant – is implemented by imposing fault models, as shown in Table 1, to the system model.

Table 1: Fault models

Fault type	Description
Non-standard charging	Change the charge level at input to the system model
Heat exchanger fouling	Reduce the airflow across the heat exchanger
Loss of compressor volumetric efficiency	Add a refrigerant flow bypass from the compressor discharge to the compressor suction
Liquid line restriction	Increase the pressure drop across the liquid line by adding a flow restriction at liquid line outlet
Presence of non-condensables	Use Dalton's law to estimate the refrigerant partial pressure and total pressure at the compressor discharge and condenser inlet

Integrating the fault models into the system model enables prediction of the impacts of faults to the system under a variety of environmental conditions and fault levels.

4. EVALUATION RESULTS USING BOTH DATA SETS

There are five types of aggregated result that are used to characterize FDD performance in a typical evaluation (as described in Yuill *et al.* 2014): 1. No Response rate; 2. False Alarm rates; 3. Misdiagnosis rates; 4. Missed Detection rates; 5. No Diagnosis rates. The latter four are organized on the basis of the impact that the fault imposes on the air-conditioners' capacity and efficiency. The impact for each scenario is quantified with a fault impact ratio (FIR), defined as:

$$FIR_{capacity} = \frac{capacity_{faulted}}{capacity_{unfaulted}} \qquad\qquad FIR_{COP} = \frac{COP_{faulted}}{COP_{unfaulted}} \tag{1}$$

The calculation of False Alarm rates requires use of FIR. This is because an arbitrarily small fault could be considered a fault, so detection of this fault would not constitute a False Alarm. If the magnitude of the fault is not considered, all systems could thereby be considered faulted. To address this in the False Alarm rate calculation, several fault thresholds – dividers between what is and is not considered faulted, based upon FIR – are imposed.

In a typical FDD evaluation, Misdiagnosis, Missed Detection, and No Diagnosis rates are grouped into bins based upon FIR. However, for the model validation presented in this paper, grouping these results by FIR does not provide additional understanding of the model's performance, so they are presented as overall rates.

Evaluation results have been generated using simulation and measurement data, and are presented in the following sections.

4.1 No Response Rates
When a protocol cannot be applied to a given operating condition or equipment type, the case is classified as No Response.
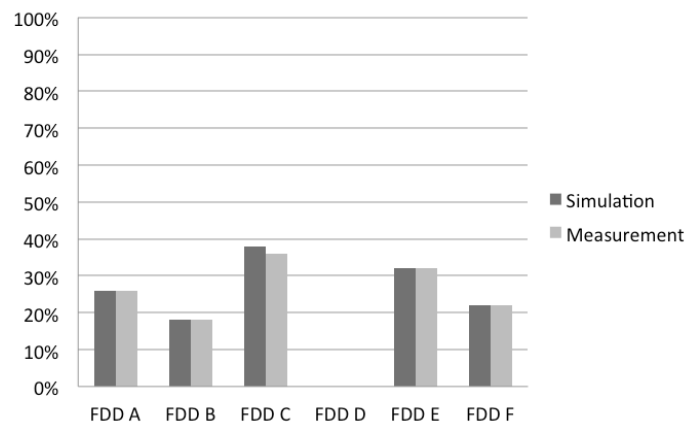

Figure 3: No Response rates for six FDD protocols using simulation data and measurement data

Since the operating conditions are identical for the two data input sets, the No Response rates shown in Figure 3 are mostly identical for the simulation and measurement data sets.

4.2 False Alarm Rates
In an FDD evaluation, the False Alarm rate is one of the most important results. A False Alarm is a case in which the FDD protocol detects a fault, but no *significant* fault is present. The rates are organized on the basis of the fault threshold, which distinguishes between a case that is considered *significantly* faulted, and a case that is not, as discussed above. Much of the difference in the measurement- and simulation-based evaluation results is caused by cases crossing from one category to the next because of minor differences in calculated FIR, and is most pronounced at FIR values closer to 100%, because there are fewer tests. The False Alarm rates for the six FDD protocols are divided between Figure 4 and Figure 5, to reduce clutter. In all cases, the measurement-based results are represented with a dashed line, and the simulation based results with a solid line.
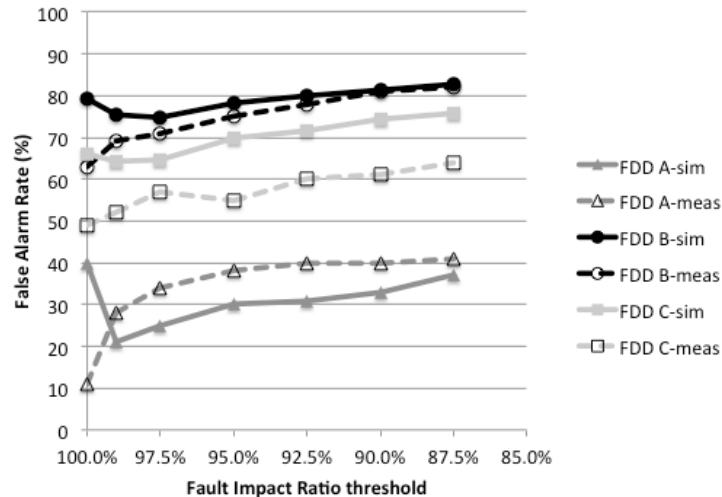
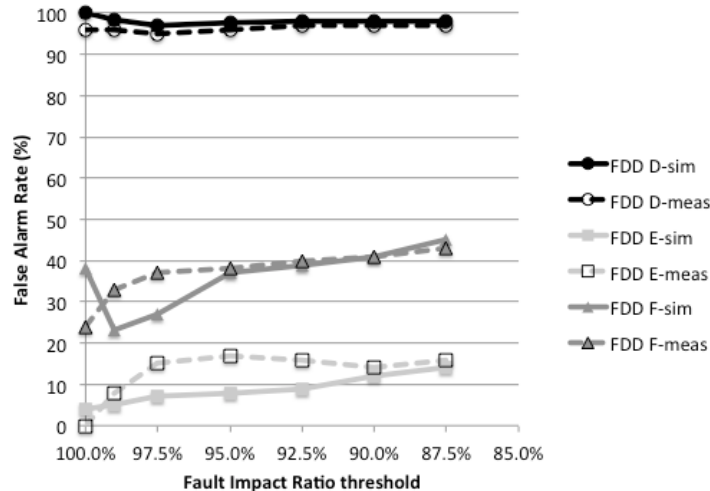**Figure 4:** False Alarm rates for protocols A to C using simulation and measurement data

Figure 5: False Alarm rates for protocols D to F using simulation and measurement data

In Figure 4, FDD A has reasonable agreement between the simulation and measurement data evaluations. Except at the 100% point, the difference ranges from 3-10% across the FIR range, with performance consistently better for simulation data. FDD B has excellent agreement, except at 100% FIR. FDD C doesn't agree quite as well, with a fairly consistent offset of 10%. However, the rank order for all six protocols is the same for the two data sets.

In Figure 5 we see excellent agreement in general. The best performing protocol, FDD E, also has the most complex algorithms. Its performance with simulation data is excellent, and it is better and more consistent than its performance with measurement data, exhibiting the expected upward slope with simulation data only. This suggests that the modeled data may be more reliable than the measurement data (which are subject to random error).

Half of the protocols have lower False Alarm rates with simulation data and half have lower False Alarm rates with measurement data. This suggests that the discrepancies are not caused by a bias, but by random differences.

4.3 Misdiagnosis Rates
A Misdiagnosis is a case in which a fault is present and the protocol detects a fault, but it gives an incorrect diagnosis. For example: an overcharged system being diagnosed as having non-condensable gas in the refrigerant. The Misdiagnosis rates presented below are the total rates, without regard to FIR.
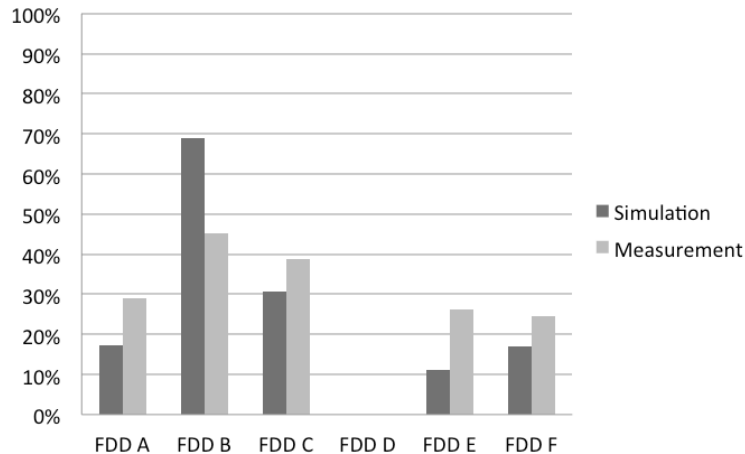
Figure 6: Misdiagnosis rates for six FDD protocols using simulation data and measurement data

Since FDD D is a detection-only protocol (it does not produce diagnoses), it always has a 0% Misdiagnosis rate. For the other protocols the differences in Misdiagnosis rate range from 7 to 24%. Although these discrepancies are not negligible, the rank order between the protocols is generally preserved. More importantly, the protocols tend to have lower rates for the simulation data, which may indicate that experimental error distorts the measurement data sufficiently that it doesn't match the expected effects of faults. The exception to this trend is FDD B. There were several cases in the measurement data for which FDD B couldn't make a diagnosis because the parameters didn't follow any of its diagnosis rules, but for the corresponding simulation case it was able to make a diagnosis. This is why FDD B's rate is higher for simulation (and vice versa for the No Diagnosis results in Figure 8), and bolsters the idea that experimental error causes the measurement data to be less reliable than the simulation data.

4.4 Missed Detection Rates
A Missed Detection is a case in which a fault is present, but the protocol indicates that there is no fault.
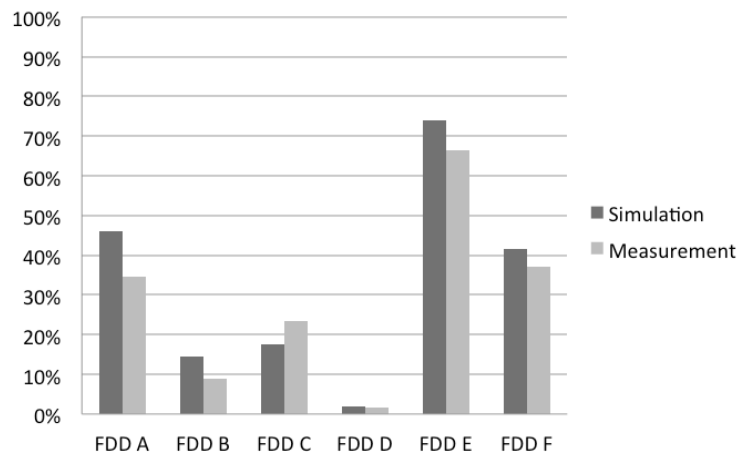

Figure 7: Missed Detection rates for six FDD protocols using simulation data and measurement data

In most cases the simulation and measurement rates agree quite well. The high rate for FDD E is related to that protocol's low False Alarm rate; it has looser tolerances to avoid spurious detections. Conversely, the low rate for FDD D is a result of faults being detected for nearly all input scenarios. Again, rank order is generally preserved.

4.5 No Diagnosis Rates
If a protocol detects a fault (and a fault is actually present) but does not provide a diagnosis of the type of fault, it is classified as a No Diagnosis case. Some protocols, such as FDD D, do not have diagnostic capabilities, so the No

Diagnosis rate is always 100%. Some protocols, such as FDD A and FDD F, do not provide detection without a diagnosis, so their No Diagnosis rate is always 0%.
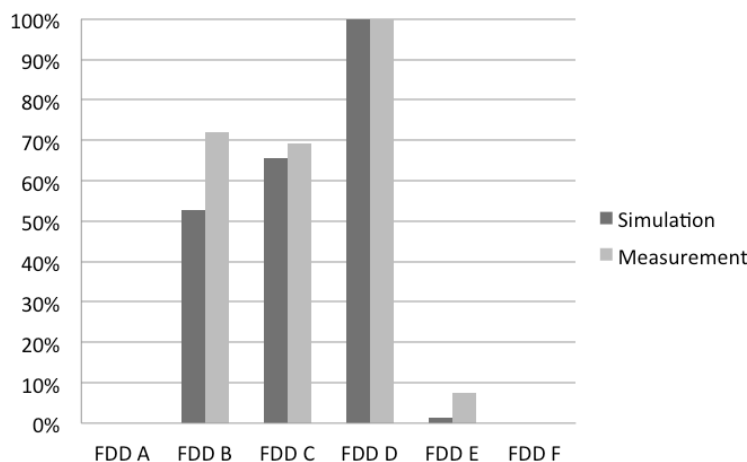

Figure 8: No Diagnosis rates for six FDD protocols using simulation data and measurement data

The No Diagnosis rates have very good agreement. This is a good indicator of model validity because No Diagnosis cases tend to occur for the more extreme scenarios – near the limits of range for operating conditions and fault levels – which also tend to be the most difficult to model. There are lower rates of No Diagnosis with simulation data, which may indicate that the simulation is providing fewer unrealistic scenarios than measurements with their attendant error.

4.6 Discussion of FDD Evaluation Results
The FDD evaluation results generated with simulation and with measurement data do not agree perfectly. However, they clearly show similar trends in all of the plots. More importantly, in almost all cases a transition between simulation and measurement data preserves the rank order of the performance for these six protocols. Since these protocols use very different approaches and different inputs for their diagnostic algorithms, it is reasonable to assume that these results will generalize to other protocols.

Returning to the question of the overall validity of this model for its intended application, the results above need to be considered within the context of the FDD evaluation method proposed by Yuill *et al.* (2014). That method is continuing to be developed, but in its current state it provides broad results of overall performance that are useful primarily for distinguishing the comparative performance of FDD protocols, rather than absolute performance levels for individual protocols. From that perspective, the simulation and measurement data give equivalent results. For example: FDD E is clearly the top performer overall, but it suffers a high Missed Detection rate; FDD A is second best, but also has a high Missed Detection rate; FDD D is not a useful protocol. Each of these results is the same whether measurement or simulation data are used.

What is not apparent from the results in this paper, but is made clear in Yuill *et al.* (2014) is that the choice of input conditions has a much larger influence on FDD performance than the shifts between simulation and measurement for identical conditions (these shifts being the focus of the current paper). This strengthens the validity of the model in two ways. First, the model-induced differences become less significant, or negligible, overall. Second, since the model allows control over the input conditions, its benefits outweigh any costs that may be imposed by the model.

A final consideration on the validity of the model is that it is not known whether the model or the measurements more accurately represent reality. The measurements are subject to both random and bias experimental error, whereas the model is free of random error, and in many of its components it is also free of bias error (for example, when energy balances are applied). Most likely there are some scenarios in which the model is better, and some in which the measurement is better. Therefore, the differences in Figures 3 to 8 should not be assumed to be model shortcomings.

5. CONCLUSIONS

A method for evaluating the performance of FDD protocols has been developed. It requires realistic input data. A gray-box model has been developed to provide these input data. A study of the validity of this model is presented in this paper. The cornerstone of this validation is the comparative performance of several FDD protocols when evaluated using an exactly equivalent set of input data from a) laboratory measurements, and b) gray-box simulations at identical conditions. Some key conclusions of this validation study are:

- In most cases the FDD evaluation results are very similar
- In most individual evaluation tests (False Alarms, Misdiagnosis, etc.) the rank order of performance for the protocols is identical for the two input data sets
- In all cases, the overall performance rank order is preserved
- The effects on performance when shifting from measurement to simulation are small compared with the effects of changing the set operating conditions that FDD protocols are tested with

Summarizing these conclusions, the gray-box modeling approach described by Cheung and Braun (2013a, 2013b) is considered a valid source for input data to be used in evaluating FDD protocols.

NOMENCLATURE

COP	coefficient of performance	(–)
FDD	fault detection and diagnosis	
FI	fault intensity	(–)
FIR	fault impact ratio	(–)
meas	measurement	
sim	simulated	
T	drybulb temperature	(°F or °C)
WB	wetbulb temperature	(°F or °C)
ϕ	fault type	

Subscript

amb	ambient
ra	return air

REFERENCES

Bayarri, M.J, Berger, J.O, Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J., Lin, C-H and Tu, J., 2007, A Framework for Validation of Computer Models, *Technometrics*, vol. 49, no. 2: p. 138-154.

Cheung, H. and Braun, J.E., 2013a, Simulation of Fault Impacts for Vapor Compression Systems by Inverse Modeling Part I: Component Modeling and Validation, *HVAC&R Research*, vol. 19, no. 7: p. 892-906.

Cheung, H. and Braun, J.E., 2013b, Simulation of Fault Impacts for Vapor Compression Systems by Inverse Modeling Part I: System Modeling and Validation, *HVAC&R Research*, vol. 19, no. 7: p. 907-921.

Oberkampf, W.L. and Roy, C. J., 2010, *Verification and Validation in Scientific Computing*, Cambridge University Press, Cambridge, UK, 790 p.

Roache, P., 2009, *Fundamentals of Verification and Validation*, Hermosa Publishers, Socorro, NM, 476 p.

Schlesinger, S., 1979, Terminology for Model Credibility, *Simulation*, vol. 32, no. 3: p. 103-104.

Thacker, B.H., Doebling, S.W., Hemez, F.M., Anderson, M.C., Pepin, J.E. and Rodriguez, E.A., 2004, Concepts of model verification and validation. *LA-14167-MS*. Los Alamos National Laboratory, Los Alamos, NM.

Yuill, D.P. and Braun, J.E., 2013, Evaluating the performance of FDD protocols applied to air-cooled unitary air-conditioning equipment. *HVAC&R Research*, vol. 19, no. 7: p. 882-891.

Yuill, D.P. and Braun, J.E., 2014, *FDD Evaluator 0.1.4*, Ray W. Herrick Laboratories, West Lafayette, IN, software available at *www.tinyurl.com/FDDeval014*

Yuill, D.P., Cheung, H. and Braun, J.E., 2014, Evaluation of Fault Detection and Diagnostics Tools by Simulation Results of Multiple Vapor Compression Systems. Paper 2605*, Proceedings of the 15th International Refrigeration and Air Conditioning Conference at Purdue, July 14-17*.

ACKNOWLEDGEMENT