2022

Electrical Signature Based Fault Detection and Classification Framework for Single-Speed, Unitary Heat Pumps Using an Adaptative Neural Network Approach

David Yuill

Andrea Mammoli

Thomas Caudell

Krish Gomatom

Yifeng Hu

See next page for additional authors

Electrical Signature Based Fault Classification and Detection Framework for Single-Speed, Unitary Heat Pumps using an Adaptative Neural Network Approach

David YUILL[1,*], Andrea MAMMOLI[2], Thomas CAUDELL[3], Phanikrishna GOMATOM[2], Yifeng HU[1], Olusegun SHOWUNMI[2]

[1]University of Nebraska - Lincoln, Durham School of Architectural Engineering and Construction, Omaha, NE, USA
(402-554-4176, dyuill@unl.edu)

[2]Electric Power Research Institute, Power Distribution and Utilization, Palo Alto, CA, USA

[3]Professor Emeritus of Electrical and Computer Engineering, Edgewood, NM, USA

* Corresponding Author

ABSTRACT

Automated fault detection and diagnosis (AFDD) for residential air-conditioning systems and heat pumps has significant potential to provide benefits, such as reducing environmental impact, operating cost, energy usage, and peak power required. This is accomplished by addressing faults that would otherwise go unnoticed, or by early detection of progressive faults, such as refrigerant leakage. The key challenge of AFDD development is to provide a method that is both accurate and cost-effective. Since there are very few accessible sensors included in most residential systems, this challenge is often focused on finding a balance between diagnostic accuracy and minimization of sensors. This paper describes a pilot study to test the hypothesis that high-resolution measurements of electrical power can be used to detect and diagnose faults. An experiment was conducted in which faults were simulated in an air-conditioner under controlled laboratory conditions. The system is a standard 10.5 kW nominal capacity single-speed split system with R-410A refrigerant and a scroll compressor. The faults that were imposed were liquid-line restrictions, improper refrigerant charge, and reduced evaporator airflow. The faults were imposed with a range of typical operating temperatures. Electrical data from startup and from steady operation were gathered in various resolutions, using multiple instrumentation sets capable of sampling frequencies in the 1-20 kHz range. The electrical measurements were made on the outdoor unit, which includes the condenser fan, compressor, and control board. The power signatures of the faulted and fault-free system were compared, using unsupervised machine-learning based methods. The LAPART (laterally primed adaptive resonance theory) approach explores detectable differences between fault and no-fault conditions and classifies voltage and current signatures into learned categories. The initial results show that this method can successfully discriminate between the faulted and fault-free condition at steady state operation, showing that there is potential to develop AFDD methods based upon single-location electrical measurements. Ongoing work will explore the limits of the diagnostic potential from this approach, explore possibilities for simplifying the measurements, and develop practical methods for cloud-based data processing.

1. INTRODUCTION

Heating and cooling services are still the largest individual energy consumer in residential buildings, although EV charging may eventually become the leader. The majority of heat pumps and air conditioners for residential space heating and cooling currently installed in the United States is of the single-speed type, in which the thermal load of the building is matched by cycling the compressor on and off. Specifically, most of these units are controlled by a thermostat that cycles the compressor and associated equipment on and off, to maintain space temperature within a

deadband around a temperature specified by the user. There is ample evidence that many of these machines operate in sub-optimal conditions, due to a variety of faults that result from improper installation, setup or maintenance. This often leads to excessive energy consumption, resulting in higher bills for the end user, and in higher electrical loads for distribution system operators and electricity providers in general. In a study to assess the impact of HVAC faults on electricity consumption, Domanski *et al.* (2014) report that a 30% increase in annual energy consumption is a plausible result of faults such as refrigerant overcharge or low indoor air flow. Pigg *et al.* (2016) report that average performance degradation due to installation or maintenance faults is on the order of 12%, with about one in six systems suffering a 25% degradation due to a combination of factors, primarily refrigerant charge and improper airflow. Moreover, while professional service calls for heating and cooling systems are common, there are few incentives promoting proper installation and maintenance, accompanied by little awareness or concern of associated problems. In a recent market survey, Butzbaugh *et al.* (2020) report that such problems are pervasive, with 70-90% of homes having some type of HVAC fault. So, in summary, while inexpensive maintenance interventions could produce large energy and cost savings for consumers and electric utilities alike, there is little awareness of the problem, and as a consequence little is done to resolve it.

It is possible to use thermomechanical data from various sensors on vapor compression HVAC equipment as features to detect and diagnose faults. For example, Rossi and Braun (1997) used seven temperatures – five refrigerant temperatures and two air temperatures) in a statistical rule-based method to detect five common faults. Yuill and Braun (2017) tested the effectiveness of several fault diagnostic methods that use thermomechanical features, and found that performance varied from very good to very bad, with a trend that the simpler methods, which used fewer measurement features, had weaker performance. Some manufacturers are beginning to offer fault diagnostics for their high-end systems as part of their service, with operational data originating from on-board control instrumentation. However, the penetration of these systems is currently low. One third-party fault diagnostics product can be retrofitted to most currently deployed systems, with the retrofit, in the form of sensors, data logging and communications equipment, implemented by service professionals. In most cases, diagnostics information is collected by the service provider, and service calls are sent to associated contractors (Butzbaugh *et al.* 2020). Unfortunately, high costs mean that few such diagnostics are currently present, or likely to be deployed on the existing fleet of heat pumps and air conditioners. Rather, with this paradigm, diagnostics capabilities are likely to grow organically, over one to two decades, following replacement of existing units, initially at the higher end of installation costs and gradually permeating all installations. To promote pervasive and fast deployment of fault detection and diagnostics, it is necessary to provide a solution that is easy to install, accurate, and has very low hardware cost. Non-intrusive load monitoring (NILM) is an approach that is intended to provide such a low-cost solution. NILM has been studied extensively, and an extensive review of the technology was recently provided by Rafati *et al.* (2022). A typical approach is to detect the on/off state of an HVAC unit, and determine whether degradation exists (e.g. short cycling), or whether the cycling is consistent with prevailing weather conditions. Another approach relies on analysis of electrical data at a high sampling rate, to detect faults (thermomechanical or electrical) that carry through to the electrical signal. Armstrong *et al*. (2006) tested NILM for fault diagnostics in rooftop units, using high frequency sampling electrical measurements, and concluded that several faults could be reliably detected, and in some cases diagnosed, using electrical measurements, particularly if they are combined with thermomechanical measurements. Analysis of electrical data at a high sampling rate is increasingly a possibility, due to the rapidly decreasing cost of computation, accompanied by the decreasing cost of sensing and communications equipment. This paper is based on the hypothesis that short-timescale thermomechanical processes in the compressor are reflected in the electrical signature of the power feed to the vapor compression unit of a residential type split system, and that such electrical signatures could reveal diagnostic information related to the operating condition of the system, including potential faults.

2. EXPERIMENTAL SETUP

The experiments were carried out in a university's environmental testing chambers. One chamber simulates the outdoor environment, allowing controlled temperatures ranging from -29 °C to 54 °C, at relative humidity between 5% and 95%. The outdoor unit, placed in the outdoor chamber, is outfitted with a refrigerant mass flow sensor, and is modified using a manually adjustable valve in the liquid line that simulates a liquid line restriction. This fault can result from several causes, including accumulation of debris or impurities in the filter-drier, or physical damage, such as a crimp in the line. The other chamber simulates indoor conditions, also with controlled temperature and humidity conditions. In the indoor chamber, an accurately controlled air flow can be imposed on the indoor unit heat exchanger, allowing the simulation of fault conditions such as undersized ducting or clogged filters that reduce evaporator airflow.

Two other common faults, undercharge or overcharge of refrigerant, are simulated by undercharging the system or overcharging the system by a controlled amount by weighing the amount removed or added.

The laboratory setup also includes temperature measurements at multiple points in the cycle, temperature and humidity measurements on the air side, and refrigerant pressure measurements on the high and low side. All of the thermomechanical measurements are recorded in a data acquisition system on one-second intervals. Future work will consider combinations of thermomechanical and electrical measurements as a potential path toward increasing accuracy and certainty of fault diagnostics.

The experiments conducted in this work are summarized in Table 1. The test matrix is meant to span the range of most commonly encountered faults, including multiple faults, in typical settings of outdoor temperature and indoor temperature. The fault intensity denotation in Table 1 uses the definitions in Hu *et al.* (2021). All tests were conducted for 20-30 minutes, allowing the system to reach steady-state conditions. Between tests, the system was turned off for approximately 10 minutes or a period of several hours, simulating typical cycling times for a range of loads. Electrical measurements were collected at startup and at steady-state.

Table 1: Experimental test matrix

Moderate Temp. (27 OD / 23 ID °C)	**High Temp. (40 OD / 23 ID °C)**
Baseline – no fault	Baseline – no fault
Evaporator airflow 60% (EA60)	Evaporator airflow 60% (EA60)
Liquid line restriction 30% (LL30)	Liquid line restriction 30% (LL30)
EA60 + LL30	EA60 + LL30
Overcharge 20%	Overcharge 20%
Undercharge 30%	Undercharge 30%

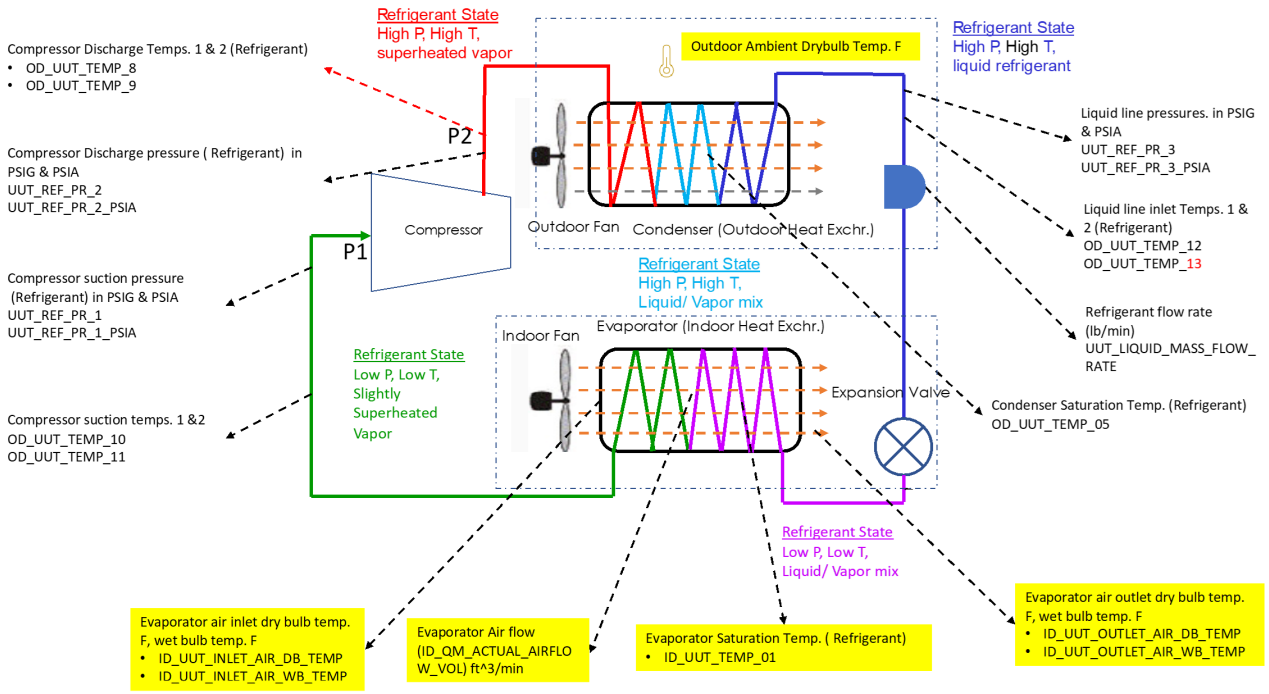

Figure 1: Mechanical system configuration and thermomechanical instrumentation.

2.1 HVAC system configuration

The mechanical configuration of the HVAC system, and the associated thermomechanical instrumentation, is shown in Figure 1. The system under test has a rated capacity of 34,700 Btu/h (10.2 kW), rated SEER of 13, and nominal (i.e. with no EA fault) indoor airflow of 1200 CFM (0.57 m³/s). The nominal refrigerant charge, including adjustment for the mass flow sensor and liquid line restriction valves, is 2.10 kg of R-410A. The system has a fixed orifice expansion device, and rated suction superheat of 6.67 °C at the AHRI (2017) rating condition A (35/26.7/19.4 °C outdoor drybulb/indoor drybulb/indoor wetbulb). The system has a single-speed scroll compressor, representative of the vast majority of modern compressors for systems of this class.

2.2 Electrical Measurements configuration

The electrical measurements were concentrated on the outdoor unit, as shown in Figure 2. Voltage common to all circuits was measured between two legs of the split phase, 240 V supply. Current transducers and voltage measurement probes were placed on the main power supply to the outdoor unit to measure electrical data. For specific tests current transducers were placed on the compressor and the fan, respectively. Further, currents through the run winding and the start windings of the compressor motor and fan motor were measured individually. The rationale for such a fine-grained measurement is that the source of the harmonic content in the waveform is not known a priori, and could result from the start circuitry or from the compressor / motor dynamics, or both simultaneously.

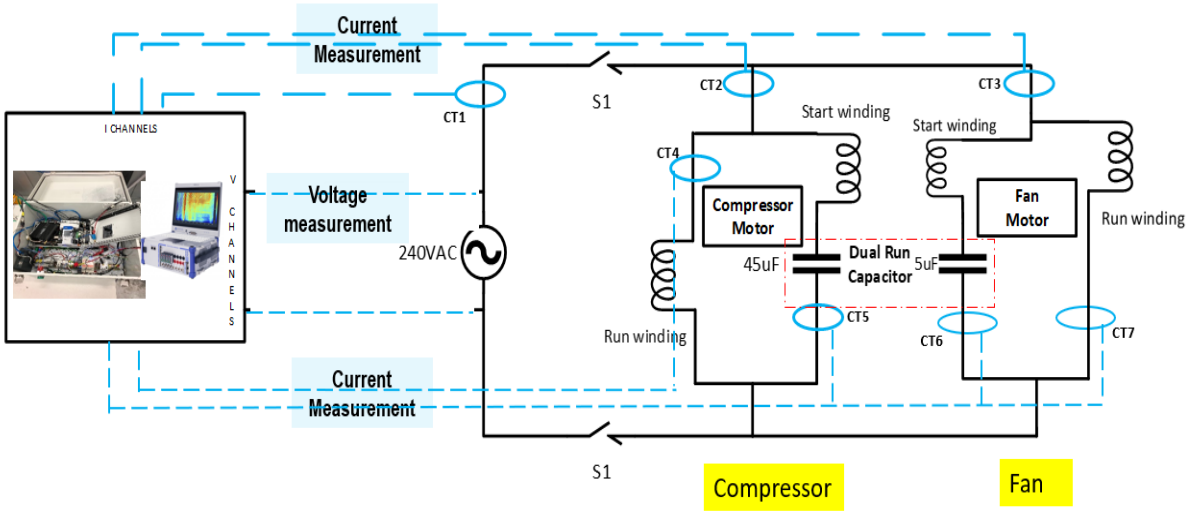

Figure 2: Configuration of the electrical measurements.

The electrical measurements were made, in parallel, with a series of instruments of varying capability and cost. The characteristics of the instruments are listed in Table 2.

Table 2: Waveform meter characteristics

Meter	Data resolution	Sampling	Reporting interval	Nominal cost	Application
Dewetron PQ Analyzer	Very high	20 kHz	Hundreds of measurements per second	$30,000	Power quality diagnostics
PQube3	High	10 kHz	10 measurement per second (Modbus), higher on trigger waveform	$5,000	Power quality disturbance diagnostics
Meazon	Low	1 kHz	1 measurement per 5 second interval	$500	Submetering and energy management
OpenZmeter	Medium	5 kHz	V,I measurements at 3 second intervals, reactive power every 15 minutes	$400	Standardized power quality diagnostics, smart metering (opensource)

The Dewetron PQ Analyzer can be considered as the "reference" instrument. The analysis conducted in the present paper is based on Dewetron data, as a first step in the development of AI solutions to detect and diagnose faults. If AI solutions work with the Dewetron data, then it will be possible to test them on data from lower cost instruments. In principle, there is good reason to believe that data from the OpenZmeter will be of sufficient quality, with the proviso that the AI developed in this project would be embedded in the onboard data processing. Ultimately, the goal is to deploy even lower-cost sensing and diagnostics, with a total cost of less than $100.

For the work presented in this paper, we utilized the voltage and current waveforms measured near the startup of the machine, but sufficiently distant from startup transients, which may be influenced by startup conditions such as scroll position. For the process of deciding how far away from actual startup would be ideal, several cycles were plotted on a current-voltage (I-V) diagram, as shown in Figure 3.

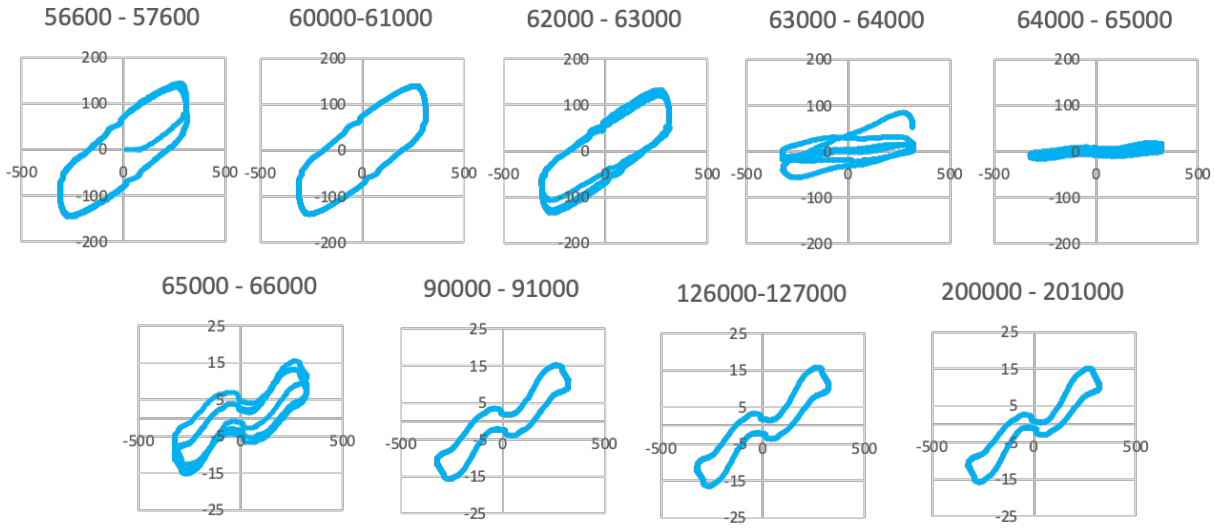

Figure 3: Scatter plots of I vs V over three 60 Hz voltage cycles at various locations in time series. Headers indicate the time step interval used in the plot. Each time step is 5.0e-5 seconds. Top and bottom rows of plots have different current scales.

The evolution of the I-V plots shows that there is a strong transient phase over the first 20 or so cycles (inrush currents on the order of 150 A), but that electrical steady-state is not reached until after about 10 s after startup, when the shape of the I-V curve completely stabilizes. Data prior to stable conditions were discarded for the purpose of this analysis.

3. DATA ANALYSIS

3.1 Overall strategy

Several I-V plots, generated for different operating conditions (i.e. combinations of operating temperatures and fault condition), indicate that there is "shape" of the waveform associated with each condition. The test data for each condition were subdivided into a set of test patterns, consisting of time series for a set of consecutive current cycles. The idea was to use a biologically-inspired neural network to see whether the waveform shapes collected for each operating condition could be associated with the operating condition itself in a reliable manner. The ultimate goal is to train a neural network that, based on a measured waveform, could then reliably indicate the operating condition. Among the wide variety of neural networks available, the Adaptive Resonance Theory (ART) class of networks was chosen, because these neural networks are known to be fast learners, and to also have the capacity to learn new patterns if needed.

3.2 The LAPART neural network

The *LAPART* (Healy and Caudell 1997) network architecture is based upon the lateral coupling of two *Fuzzy ART* subnetworks (Carpenter *et al.* 1991, Carpenter and Grossberg 1987), referred to as *A* and *B*, which perform unsupervised categorization of analog input patterns, as illustrated in Figure 4(a). *Fuzzy ART* class networks are generally composed of three interconnected layers of nodes: the *F0* input layer, the *F1* comparison layer, and the *F2* winner-takes-all unsupervised class labeling layer that recruits more class nodes as learning occurs. In addition, each *Fuzzy ART* network has an attentional subsystem manifest in its vigilance node, Vig. The vigilance parameters for these subnetworks are respectively, $\{0 < \rho_A, \rho_B \leq 1\}$. Note that ρ controls the level of generalization performed by the subnetwork. If its value is close to unity, the number of *F2* layer categories (nodes) rapidly increases during training, essentially memorizing each individual input pattern. Conversely, if ρ is near zero the number of *F2* categories increases very slowly during training, creating a system that greatly generalizes the input patterns, lumping many patterns into a small number of categories.

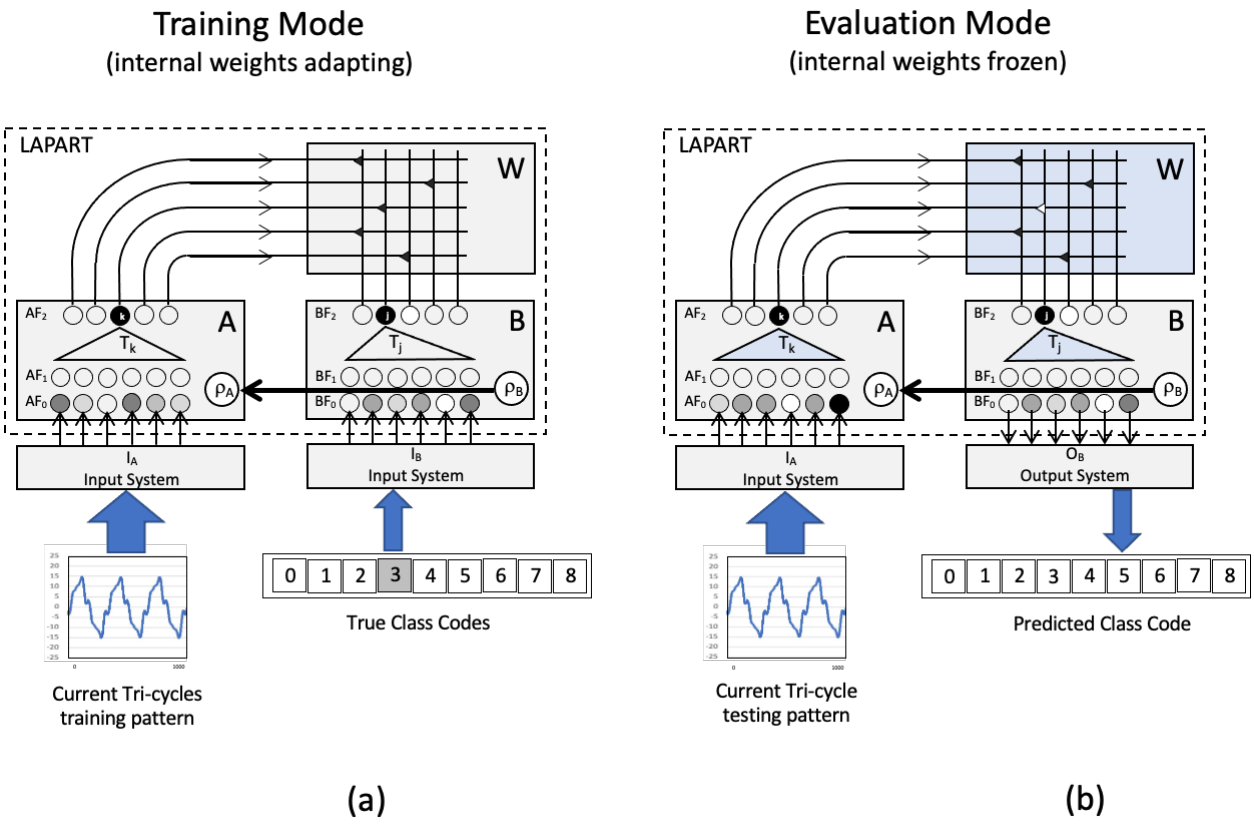

Figure 4: Details of the LAPART neural architecture. Note that the Input Systems perform the necessary normalizations required by Fuzzy ART. (a) training mode, (b) testing mode.

The interconnects between the two subnetworks force an interaction of the respective learning performed by the *Fuzzy ART* subnetworks on their inputs. This modifies their unsupervised learning properties to allow the learning of inference relationships or associations between the learned pattern categories representing their individual input domains. This can be thought of as supervised learning, or supervised classification. In actuality, however, it is much more general. The usual sense of classification is that of creating a partition of the inputs; that is, separating them into disjoint sets, with a label (the desired output specified by the "teacher") attached to each element of the partition. With the *LAPART* architecture, sets are labeled with *sets* - in other words, the network extracts rules with antecedent and consequent predicates. The sets are referred to as classes because they are labeled by *Fuzzy ART F2* nodes whose connections to the *F1* layer nodes have formed patterns of weights that are templates for the input patterns in their classes. The template weights for a class are the fuzzy AND of the input patterns adopted into the

class. The input patterns, *Fuzzy ART* layers, and templates will be labeled with an *A* or *B* referring to the *A* or *B* *Fuzzy ART* subnetworks.

During *LAPART*'s training mode, pairs of input patterns $I_A(k)$, and $I_B(k)$, where *k* is this pattern index, are passed through the Input Systems, which performs all appropriate scaling and normalizations, to the *F0* layers of subnetworks *A* and *B*, respectively. As *A* and *B* form class templates for their inputs, the *LAPART* network learns inference relations between their resonant *F2* nodes (category nodes) through a Hebbian (Hebb 1949) learning process, by forming strong $F2A \rightarrow F2B$ weighted interconnections between pairs of simultaneously-activated *F2A* and *F2B* nodes.

Convergence of a *LAPART* network in a finite number of passes through a training set requires that it reach the following operational state: Presentation of any input pair $(I_A, I_B)(k)$ from the training set results in pattern I_A being immediately assigned a class in *Fuzzy ART* subnetwork *A* through direct access to the class template. Through a strong, learned inferencing connection, the class *F2A* node signals a unique *F2B* node to which it is connected, forcing it to become activated. This results in the inferred *B* class template being read out over the *F1B* layer just as pattern I_B reaches the *F1B* layer. The ensuing vigilance test in subnetwork *B* confirms that the inferred class is an acceptable match for I_B, forcing the subnetwork *B* vigilance node to remain inactive. Further, the *B* class template is a subset template for I_B. In summary, a final pass through the data shall result in no resets and no synaptic strength changes (i.e., no learning). For most applications, this occurs in two presentation epochs of the training data.

During the evaluation (or testing) mode of operation, learning is disabled and no input is supplied to the *B* subnetwork, as illustrated in Figure 4. An input pattern $I_A(k)$, drawn from the testing data set, is passed through the Input System to the *F0* layer of subnetwork *A* where it functions independently of the *B* subnetwork, activating existing *AF2* nodes or attempting to recruit a new one. If an existing *AF2* node resonates, then its lateral inference connection primes a *BF2* node, causing the read out of its class encoding template across the *BF1* layer. The Output System decodes the class label from the *BF1* layer activation pattern. If the *A* subnetwork attempts to recruit a new *AF2* node, then the Output System is instructed to report an "unknown" classification of the current *A* input. Finally, at the end of a single testing epoch, performance statistics are computed for the trained LAPART. Three outcomes are possible for each testing input: 1) a correct prediction, 2) an incorrect prediction or error, and 3) unable to predict or unknown. These overall statistics and the actual predicted classes, summarized in the form of a confusion matrix, are results of the evaluation process.

3.3 The creation of the tri-cycle input representation.

For this current work, the *A* subnetwork input was chosen to be current waveforms. Ideally, each input pattern would be a single 60Hz cycle current wave. But given the Dewetron sampling period of 5.0e-5 sec, one 60Hz cycle does not contain an integral number of samples (333.333...). Alternatively, three 60Hz cycles do contain a whole number of samples (1000) and is therefore used for this analysis. These 1000-dimensional input patterns are referred to as "tri-cycle" input representations.

3.4 Preparation of the data set for neural network training and evaluation.

Each Dewetron time series data set from 31 trials and nine experimental conditions was divided into 1000 timestep tri-cycles. These were then concatenated into a single list of 146,114 condition (class) labeled tri-cycle input patterns. For clarity, the nine conditions selected from the experimental matrix in Table 1 are numbered as follows:

Table 3: Fault and operating conditions

Class	Condition	Class	Condition
0	Low temperature, baseline, no faults	1	High temperature, baseline, no faults
2	Low temperature, LL30 fault	3	High temperature, LL30 fault
4	Low temperature EA60 fault	5	High temperature EA60 fault
6	High temperature CH120 fault	7	Low temperature CH70 fault
8	High temperature CH70 fault		

3.5 Training *LAPART*

The order of the list of inputs pattern was shuffled multiple times using the Fisher-Yates (Fisher & Yates 1953) algorithm. The resultant list was divided into two sub-lists, by assigning the training set as even numbered shuffled list indices (73057) and the testing set as odd numbered (73057), producing a disjoint random sampling of class examples for the two modes. The training mode was performed over two presentation epochs during which *LAPART* convergence was verified. One important result of the training process is the resultant total number of *A* subnetwork templates. The ratio of this to the total number of training patterns is a measure of memorization/generalization. The closer to zero, the more generalization.

3.6 Evaluating *LAPART*

During evaluation mode, learning is disabled and the patterns not used in training are presented to the *A* subnetwork of *LAPART* for class prediction. Only a single testing presentation epoch is required. Multiple studies were performed using a range of vigilance parameters and numbers of shuffles. The overall statistics and the actual classifications, in the form of a confusion matrix (Stehman 1997), are the most important results of the evaluation process.

3.7 Examples of results

The following three plots are examples of the results of this study. The first, Figure 5(a), shows the number of *A* subnetwork categories versus its vigilance value, ρ_a. As can be seen, the number of categories grows exponentially towards the total number of testing patterns as ρ_a approaches unity. At its largest value in the plot, the ratio of number of *A* subnetwork categories to the total number of testing patterns is ~27%. This is far from memorization.

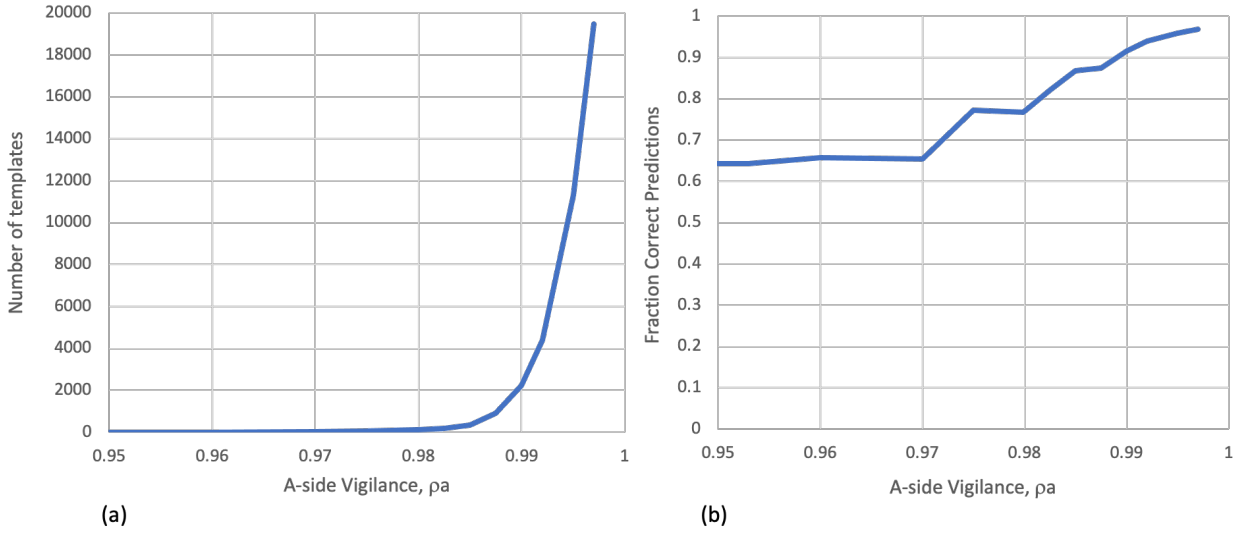

Figure 5: (a) Number of A subnetwork templates versus ρ_a and (b) fraction of correct predictions versus ρ_a

The second figure in the set, Figure 5(b), plots the fraction of correct predictions during testing as a function of *A* subnetwork vigilance value, ρ_a. Note that after a range of slow improvement, the fraction appears to be on an asymptote towards a value of unity.

The third figure in the set, Figure 6, shows the confusion matrix for the case where ρ_a = 0.995. The confusion matrix plots the relative frequencies of classifications and misclassifications. For each true class row, it plots in the columns a distribution of the predicted classes. In this figure, each row distribution has been normalized by the actual number of patterns in the class. For perfect classification performance, the normalized confusion matrix would be a diagonal of unity height. As can be seen in the figure, for this set of parameters, the matrix in nearly diagonal, with an average diagonal height of ~0.96.

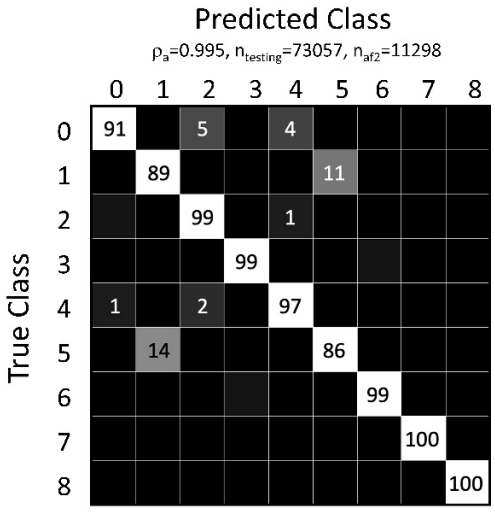

Figure 6: The confusion matrix for the results of testing data with ρ_a = 0.995. Values are percentages of the true class, rounded to the nearest whole number. The diagonal dominance indicates high performance of the classifier.

4. DISCUSSION AND CONCLUSIONS

While the voltage waveforms are sinusoidal as expected, the current waveforms are far from sinusoidal, and indicate an opportunity for waveform-based classification leading to the ability to detect and even diagnose faults. The mechanisms that lead to the deviation of the current waveform from a sinusoid are not well known, but are likely to result from a dynamic interplay between electrical and mechanical components, resulting from compressor pressure variations and the electromagnetic coupling of the induction motor rotor and stator. An indication that this could be the case comes from an experimental study on pressure variations in a scroll compressor by Picavet and Ginies (2014), where highly structured and variable pressure values are reported.

Waveforms associated with different operational conditions cluster very strongly, i.e. similar input patterns associate with the same condition in training. In testing, waveform input patterns fall into categories reliably, as demonstrated by the very strong diagonal dominance of the confusion matrix. Errors and uncategorized outcomes are rare and likely can be reduced by optimization of the training. The quality of the results presented here is somewhat surprising. The amount of information transmitted by the compressor shaft to the motor, and hence reflected in the electrical signal, is unexpectedly high. The results are very encouraging, supporting the hypothesis that it may indeed be possible to detect thermomechanical conditions reliably by examining the waveform.

Despite the encouraging results, this study only scratches the surface, and more testing is clearly needed. For example, it could be instructive to conduct a series of tests where the average torque experienced by the motor is the same, but is produced by different thermofluid conditions – e.g. high pressure lift at low mass flow rate (maybe a result of liquid line restriction), or low pressure lift and high mass flow rate. If a difference is detected in the current waveform, then it could be argued that small-scale torque variations do indeed communicate thermomechanical information. Another important consideration is to determine whether the methodology can be generalized from one air-conditioner or heat pump to another, especially if made by a different manufacturer. Test are ongoing to answer these and other questions.

If the outcome of this research continues to produce such promising results, there are two possible concepts for deployment. The first is that waveform-based diagnostic methods could be used in place of diagnostics that use thermomechanical measurements. Yuill and Braun (2013) showed that many such existing methods suffer from poor accuracy, with error rates often above 50% for a comparable set of fault conditions, so the performance of Figure 6 compares very favorably with those methods. A second concept is to deploy waveform-based diagnostics to electrical measurements made outside a home – potentially by electrical utilities – thereby providing an opportunity to sidestep the significant challenge of persuading homeowners to adopt diagnostics, which could lead to a greater number of monitored systems. Thus, the potential future impact of the findings in the current study could be substantial.

NOMENCLATURE

AFDD	Automated fault detection and diagnosis	HVAC	heating, ventilation, and air conditioning
ASHP	air source heat pump	LL	liquid line restriction
CH	refrigerant charge	NILM	Non-intrusive load monitoring
EA	evaporator airflow	SEER	seasonal energy efficiency ratio

REFERENCES

AHRI (2017), Performance Rating of Unitary Air-conditioning & Air-source Heat Pump Equipment. *AHRI Standard 210/240-2017*, Air-Conditioning, Heating & Refrigeration Institute, Arlington, VA, USA.

Armstrong, P.R., Laughman, C.R., Leeb, S.B. & Norford, L.K. (2006) Detection of Rooftop Cooling Unit Faults Based on Electrical Measurements, *HVAC&R Research*, 12(1), 151-175.

Butzbaugh, J. B., Tidwell, A. S., & Antonopoulos, C. A. (2020). *Automatic Fault Detection & Diagnostics: Residential Market Analysis* (No. PNNL-30077). Pacific Northwest National Lab. (PNNL), Richland, WA (United States).

Carpenter, G. A., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer vision, graphics, and image processing*, *37*(1), 54-115.

Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991). Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural networks*, *4*(6), 759-771.

Domanski, P. A., Henderson, H. I., & Payne, W. V. (2014). *Sensitivity analysis of installation faults on heat pump performance* (pp. 1-94). US Department of Commerce, National Institute of Standards and Technology.

Fisher, R. A., & Yates, F. (1953). *Statistical tables for biological, agricultural and medical research*. Hafner Publishing Company.

Healy, M. J., & Caudell, T. P. (1997). Acquiring rule sets as a product of learning in a logical neural architecture. *IEEE Transactions on Neural Networks*, *8*(3), 461-474.

Hebb, D. O. (1949). *The Organization of Behavior: A Psychological Theory*. New York: Wiley.

Hu, Y., Yuill, D.P., Ebrahimifakhar, A. & Rooholghodos, S.A. (2021). An experimental study of the behavior of a high efficiency residential heat pump in cooling mode with common installation faults imposed, *Applied Thermal Engineering*, 184, 116116,

Picavet, A., & Ginies, P. (2014). Experimental Pressure-Volume diagrams of scroll compressors. In *Proceedings of the International Compressor Engineering Conference*, Purdue University.

Pigg, S., Cautley, D., & Koski, K. (2016). Improving Installation and Maintenance Practices for Minnesota Residential Furnaces, Air Conditioners and Heat Pumps. *Conservation Applied Research & Development (CARD) FINAL REPORT (Sept. 2016)*.

Rafati, A., Shaker, H. R., & Ghahghahzadeh, S. (2022). Fault Detection and Efficiency Assessment for HVAC Systems Using Non-Intrusive Load Monitoring: A Review. *Energies*, *15*(1), 341.

Rossi, T.M. & Braun, J.E. (1997). A Statistical, Rule-Based Fault Detection and Diagnostic Method for Vapor Compression Air Conditioners. *HVAC&R Research*, 3(1), 19-37.

Stehman, S. V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote sensing of Environment*, *62*(1), 77-89.

Yuill, D.P & Braun, J.E. (2013). Evaluating the performance of FDD protocols applied to air-cooled unitary air-conditioning equipment, *HVAC&R Research* 19(7), 882–891.

Yuill, D. P. & Braun, J. E. (2017). A figure of merit for overall performance and value of AFDD tools. *International Journal of Refrigeration,* 74(2017), 649-659.