2018

Comparing Maintenance Strategies for Rooftop Units having Multiple Faults through Simulation

Andrew L. Hjortland
Purdue University, United States of America, andrew.hjortland@gmail.com

James E. Braun
Purdue University - Main Campus, jbraun@purdue.edu

Simulating and Comparing Maintenance Strategies for Rooftop Units with Multiple Faults

Andrew L. HJORTLAND*, James E. BRAUN

Purdue University
Department of Mechanical Engineering,
West Lafayette, IN, USA
Andrew.Hjortland@gmail.com
* Corresponding Author

ABSTRACT

Maintenance strategies currently used for commercial building rooftop units (RTU) can be classified into two categories: reactive strategies and proactive strategies. In reactive strategies, maintenance and service is performed only when needed, e.g. when a system is unable to maintain set point. In proactive strategies, maintenance is scheduled at routine intervals to avoid service interruptions regardless of whether the system actually needs it. While these strategies could not be more different, it is unclear which strategy is more optimal. Moreover, whether one strategy is more optimal than the other more than likely depends on the application – contributing to uncertainty. A third category of maintenance has been enabled by automated fault detection and diagnostics (AFDD) technologies that aims to provide building operators and service providers more detailed information about the actual state of equipment in the field. This third strategy, called condition-based maintenance, aims to optimize service and maintenance decisions throughout the life of equipment based on updated measurements of performance and service costs. In this work, these three types of maintenance strategies are compared using a commercial building simulation model utilizing a fault impact equipment model. Along with comparing different strategies under the same fault scenario, ambient conditions, and loads, optimal maintenance schedules are generated using dynamic programming. Benefits of a condition-based maintenance approach utilizing a suite of AFDD methodologies are highlighted with respect to reducing operating costs.

1. INTRODUCTION

Like almost any other mechanical system, direct-expansion air conditioners require routine or unscheduled measures to maintain reliable and efficient operation. If an air conditioner is ignored or regular maintenance goes unscheduled, performance of the system will deteriorate over time. Determining the frequency of when to inspect or perform maintenance can be difficult since impacts of deterioration or faults are relatively difficult to estimate. It is also difficult to quantify the benefits of performing different maintenance and service tasks. Furthermore, systems can develop multiple different faults at the same time which makes service decisions even more complex.

In some cases, the operation of an air conditioning system may become completely suspended by a fault. For example, an air conditioner may fail to start when a motor capacitor fails over years of operation. When this happens, no cooling will be available until the capacitor is replaced. It is relatively easy to detect faults that totally halt system operation – when cooling is not available, comfort in the condition space cannot be controlled. Additionally, for many applications in commercial buildings the decision about how to handle these faults is easy: the fault must be fixed or the system must be replaced as soon as possible. This is because preserving occupant comfort is usually a high priority in most commercial buildings.

Some other faults do not totally suspend the operation of a system, but rather degrade overall system performance. In other words, faults may decrease the amount of cooling capacity available or the efficiency of the system, but the system can still maintain comfort in the space. An example of a fault that does not completely disable operation of an air conditioner is condenser fouling. When a condenser becomes fouled, an air conditioner is still able to deliver cooling to a conditioned space; however, it does so less efficiently. These faults are more difficult to detect than faults that halt operation – from an occupants' perspective it may not be noticeable at all. It may also be difficult to determine if faults are present by comparing utility bills as well.

For these types of faults that degrade or deteriorate performance over time, maintenance decisions are less straightforward. This is because costs required to fix or repair a system may be comparable to the impact that the fault has on utility costs. For some faults, the cost to repair may be more than the benefit incurred. For others, economic benefits for repair may outweigh these service costs, though it still may be difficult to quantify this benefit leading to uncertainty.

Limited work on service decision support systems for direct-expansion (DX) equipment has been published. While there has been extensive work done related to the fields of industrial engineering, much of this work has been focused on infrastructure, manufacturing processes, and large engineered systems and fleets. Inspection and replacement decisions for devices subject to random failures, such as light bulbs, have also been well researched. Low-cost systems prone to long-term degradation, like DX equipment, have not been as widely studied. In this work, different types of maintenance strategies are compared using a commercial building simulation model utilizing a fault impact equipment model. Along with comparing different strategies under the same fault scenarios, ambient conditions, and loads - optimal maintenance schedules are generated using dynamic programming. Benefits of a condition-based maintenance approach that utilizes the outputs of automated fault detection and diagnostics (AFDD) systems are highlighted with respect to reducing operating costs.

2. SIMULATION AND COMPARISON METHODOLOGY

In order to compare the operational cost impacts of different faults and maintenance strategies, an hourly simulation program was implemented that models the sensible and latent cooling loads of a simple commercial building. The interaction between the building model and RTU cooling equipment was implemented to determine energy consumption at hourly time intervals. The equipment model implemented in this work is based on a grey-box model developed by Cheung and Braun that captures the effects of faults on system cooling capacity and energy consumption. The effects of faults are varied over time to simulate deterioration between successive service intervals. Finally, the benefits and costs of performing maintenance tasks were modeled. Details of the simulation framework used to generate the results in the work are discussed in a companion paper (Hjortland & Braun, 2018).

One of the variables that was selected at the start of each simulation was the fault evolution rate, e.g. the refrigerant leakage rate. Because choosing combinations of fault rates is somewhat arbitrary due to the lack of reliable fault prevalence data, rates for refrigerant charge leakage, condenser fouling, and evaporator fouling were sampled from random distributions (Yuill & Braun, 2017). To consider a relatively wide range of fault rate combinations, uniformly distributed random samples were chosen for each. For refrigerant leakage fault rates, a uniformly distributed random sample between 0% to 20% leakage per year was selected for each trial. Likewise, condenser fouling fault rates were sampled uniformly between 0% to 20% per 5000 hours of condenser fan runtime for each trial. For evaporator fouling fault rates, a uniformly distributed random sample between 0% to 20% per 5000 hours of evaporator fan runtime was selected. In all, a distribution of 200 random combinations of refrigerant charge leakage, evaporator fouling, and condenser fouling faults were simulated in this study.

For each scenario, consisting of a set of fault evolution rates, simulations were performed over a 15-year life of the equipment for buildings in Miami, Atlanta, and Chicago having different service policies. This included a benchmark optimal service decision policy that minimized lifetime operating costs as described in the following section. Simulations at each location and fault combination were also performed using different proactive, reactive, and condition-based service policies. The result of this process was a distribution of lifetime operating costs for each location and service policy studied. The lifetime operating costs of each policy were compared with the corresponding optimal operating cost. Sample and summary results from this study will be discussed in the remainder of this work.

3. MAINTENANCE AND SERVICE POLICY DESCRIPTIONS

2.1 Optimal Service Policies determined using Dynamic Programming

The goal of service and maintenance optimization is to determine an optimal sequence of service decisions that minimizes life-cycle operating costs for a system or group of systems while maintaining constraints on occupant comfort, safety, or environmental impact. For direct-expansion cooling equipment, significant life-cycle costs include utility costs (cost to consume energy, usually electricity), equipment replacement costs (due to premature failures or scheduled upgrades), and maintenance costs.

More explicitly, the goal of maintenance and service optimization is to determine an optimal sequence of service tasks (a) from the set of permissible service tasks (A) that minimizes the total operating costs (OC) of a system over a life-cycle $\tau_{life} = t_{life} - t_0$,

$$J_0 = \min_{a \in A} \left\{ \int_{t_0}^{t_l} OC(t) \mathrm{d}t \right\} \tag{1}$$

such that the temperature in the conditioned space (T_z) remains comfortable,

$$\tau_l \le T_z(t) - T_{z,sp}(t) \le \tau_u \quad \forall t \in [t_0, t_l] \tag{2}$$

where $T_{z,sp}$ is the space air temperature set point, and τ_l , τ_u are the minimum and maximum allowable deviations in space air temperature that maintain comfort. Operating costs (OC) at some time t in this work were estimated as the sum of utility costs (UC), equipment costs (EC), and service costs (SC),

$$OC(t) = UC(t) + EC(t) + SC(t). \tag{3}$$

The utility costs in this work were calculated using an effective flat rate electricity cost, C_{elec} ,

$$UC(t) = C_{elec} \cdot W_{elec} \tag{4}$$

where W_{elec} is the total energy consumption of the DX equipment over some interval. The equipment costs were estimated by assuming the installation cost of the system can be uniformly distributed over the expected finite life of the system (15 years)

$$EC(t) = C_{equip} \cdot \Delta t_{run} \tag{5}$$

where C_{equip} is the effective cost per unit time to operate the equipment assuming a finite life and Δt_{run} is the total amount of equipment run-time required over some simulation interval. Equation (5) has been included to capture the penalty of faults that result in longer run-time requirements which may lead to earlier replacements. Finally, the service costs were modeled as a function of service tasks

$$SC(t) = C_{service}(a(t)) \tag{6}$$

where the cost for different service tasks are described in Table 1. When multiple tasks were performed during the same interval, a 20% discount was assumed to account for potential savings in “trip costs.”

Table 1. Summary of fault evolution rate parameters used in second multiple fault simulation and service scheduling optimization.

Task	Cost
Add Refrigerant Charge	\$100 + \$50/lb refrigerant
Clean Condenser Coil	\$300
Clean Evaporator / Filter	\$80

Dynamic programming optimization (by backwards induction) was used to solve the service decision problem in Equation (1) and determine the optimal service decision policy for each simulation trial. The optimal solution is one way to measure the performance of any maintenance plan or policy. In other words, it is a useful benchmark that can be determined initially that provides some measure of how good or bad sub-optimal maintenance policies are.

To give an illustrative example, a system with only a refrigerant charge leakage fault was simulated. Figure 1 shows results for the optimal service policy for a building in Miami where the air conditioner leaks charge at a rate of 5% per year and has a maximum capacity that is 20% greater than the maximum load during the year. Also plotted on Figure 1 are the trended refrigerant charge levels for two systems with the same charge fault (5% leakage per year) but different initial charge levels (100% and 90%). When the amount of refrigerant in either system crosses the optimal service decision boundary, the optimal service policy would recommend adding refrigerant to the system.

Applying this decision rule at any instance in time, the optimal lifetime operating costs under the assumptions of the simulation will be obtained.

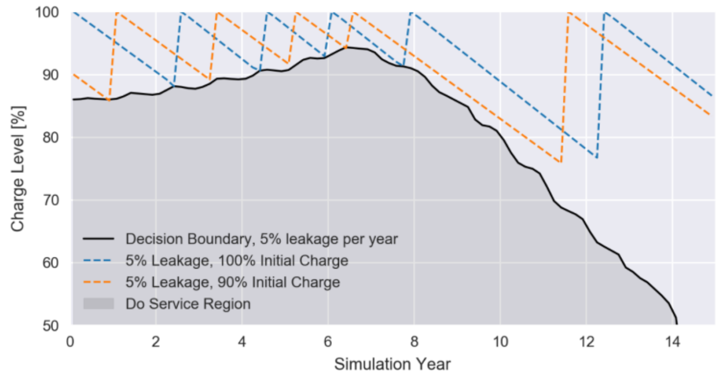

Figure 1. Example optimal service decision policy for refrigerant charge faults in Miami, FL for system that is 20% oversized to maximum annual load and refrigerant leakage rate equal to 5% per year.

Early in the simulation lifetime, the optimal boundary between adding refrigerant to the system and not performing service is largely a tradeoff between integrated energy and equipment cost impact and the cost to perform additional service tasks. Performing service tasks more often would save on energy and equipment costs, at the expense of much higher lifetime service costs. On the other hand, performing service less frequently would save on service costs by possibly performing less service tasks throughout the life, at the expense of much higher energy costs.

The decision boundary changes over time for two reasons. First, there is a small seasonality component to the decision which causes the small ripple in the decision boundary with a period of 12 months. Because refrigerant leaks throughout the year and there is less cooling load in Miami in the winter months, it is slightly better to wait until the warmer months to do service. The decision boundary for Miami reaches a peak around the 7th year. This is an effect of optimizing the lifetime operating costs over a finite interval. After the 7th year, the costs for performing service must be balanced by diminishing utility costs savings since there is no reward for finishing the simulation with an air conditioner with more charge (no salvage value). In other words, the optimal service policy tolerates more leakage since the possible future utility costs savings are less than earlier in the simulation.

Finally, it should be noted that each system shown in Figure 1 has charge added around the 11th and 12th years much before the systems' charge levels intersect with the decision boundary. These service decisions are the result of comfort violations – the significant reduction in refrigerant charge resulted in insufficient cooling capacity to maintain the space temperature of the building. Because there is a constraint to maintain comfort with the building, service must be performed.

The optimal service schedule for the building located in Miami, FL that has multiple faults evolving over time was also determined using dynamic programming for each simulation trial. The fault rates for an example trial are summarized in Table 2. The refrigerant charge in the system leaked 5% per year of simulation time. The condenser airflow rate was reduced by condenser coil fouling at a rate of 5% per 5000 hours of condenser fan runtime. The evaporator airflow rate was reduced by evaporator coil fouling at a rate of 5% per 5000 hours of evaporator fan runtime.

The optimal service schedule for the system in Miami, FL with the faults described in Table 2 is shown in Figure 2. Because the optimal decision boundaries for each service task also depend on the other fault levels, it is not possible to show the optimal policy decision threshold as was done previously. Instead, each trended fault level is plotted. It can be observed that the optimal service schedule tends to group multiple service tasks into each service interval. This occurs because of the 20% discount applied to the service costs when multiple faults are serviced at the same time. It can also be observed that the time between service intervals tends to increase as the simulation progresses due to diminishing returns from service.

Table 2. Summary of fault evolution rate parameters used in first multiple fault simulation and service scheduling optimization.

Task	Cost
Refrigerant Charge Leakage Rate	5% per year
Condenser Fouling Rate	5% per 5000 hours condenser fan runtime
Evaporator Fouling Rate	5% per 5000 hours evaporator fan runtime

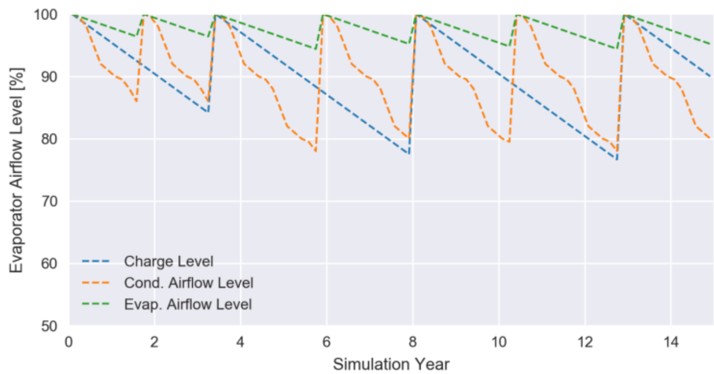

Figure 2. Example optimal service schedule for system located in Miami, FL with multiple faults. The optimal schedule tends to group multiple service tasks at each service interval.

2.2 Periodic Service and Maintenance at Regular Intervals

Implementing periodic service policies is relatively straightforward and can easily be compared with optimal service policies using the simulation framework. In these policies, a service technician is assumed to visit the air conditioning system at regular intervals and perform any preventative maintenance that is needed. In many commercial buildings, service contracts between the building owner and HVAC service providers are often implemented which approximate periodic service policies. In these contracts, the service provider generally agrees to visit the site a fixed number of times per year and perform a variety of preventative maintenance tasks in return for some fixed costs paid by the building owner. While service may not be performed at exact intervals (i.e. every six months), service time between service visits is approximately constant (i.e. annually, biannually, or quarterly).

In this work, an additional assumption about how periodic service is performed may not be exactly true in a real application. The periodic service contract that is implemented within the simulation requires each service task to be completed at every visit. In other words, evaporator cleaning, condenser cleaning, and refrigerant charge adjustment is performed whenever the service technician visits the site if the fault levels are not normal (i.e. service to repair a fault is not performed if it is not considered in the simulation). In a real scenario, the service technician may not perform all tasks during every visit. Rather, the technician may only inspect the system to determine if maintenance is needed based on experience. If these inspections are permitted within the service contract, rather than requiring that each task is performed per visit, it may decrease the service costs.

Inspection policies were not investigated or implemented in this work since it requires some assumptions about how the service technician perceives the equipment state and when service is needed. In real scenarios, performing on-site inspections is generally a good idea since it may provide valuable insights into how a system is performing. For instance, a service technician may clean the condenser coil if they notice it is covered in debris. On the other hand, minor condenser fouling would be ignored if the condenser looks mostly clean. In the simulation, these considerations are not modeled. Rather, the service technician will perform different service tasks regardless of the severity of the faults.

A comparison between the trended refrigerant charge levels for a system that is serviced annually and biennially and leaks 5% of its refrigerant charge annually for a system installed in Miami, FL is shown in Figure 3. In comparison to the optimal decision boundary, Figure 3 shows that the annual and biennial service schedules perform service too often for a 5% leakage fault. Using these schedules, annual utility costs are less than the utility costs obtained using the optimal schedule. However, service costs are much greater since service is performed more often over the life of the equipment. Figure 3 also shows that periodic service schedules do not consider the payback time required to break even towards the end of equipment life. This accounts for much higher lifetime service costs.

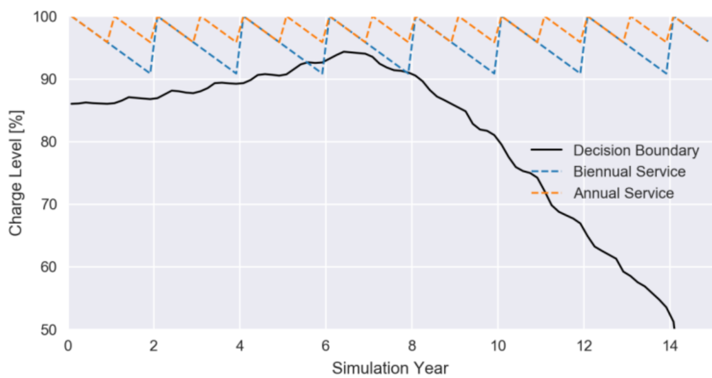

Figure 3. Example comparison of lifetime refrigerant charge levels for system located in Miami, FL with a leakage rate of 5% percent per year using biennial and annual periodic service schedules.

2.3 Emergency Service Policies

Whereas periodic service policies can be viewed as proactive, emergency service policies can be considered reactive. In this policy, service is performed only when a comfort violation occurs due to insufficient cooling capacity provided by the air conditioning equipment. In this work, comfort violations were defined when the temperature in the space exceeded the set point by 1.1 °C (2.0 °F) for a continuous interval of 4 hours or longer. When comfort was violated, service was performed immediately – evaporators were cleaned, condensers were cleaned, and refrigerant charge was adjusted. In real situations, a time lag between when comfort is violated and when service is performed may be significant. This is especially true depending on the time of the year service is needed: service technicians may be very busy during peak cooling months, while they may be more available during the shoulder seasons. The seasonal availability and costs are not considered in this work.

Emergency service policies may contribute to large operating costs for oversized systems since capacity violations may not occur until faults have degraded performance significantly. Additionally, emergency policies may incur higher operating costs than periodic service policies for faults that have limited impacts on cooling capacity, e.g. condenser fouling. When condenser fouling occurs, both head pressure and energy consumption increase. Comfort violations caused by condenser fouling may never occur and can lead to significant time between service intervals. One measure that has been implemented within the simulation that triggers comfort violations because of condenser fouling is a high-pressure limit. In normal systems, high pressure limit switches are typically installed to protect the compressor from operating outside the manufacturer's suggested operating envelope. When the head pressure exceeds the high-pressure limit, the air conditioner is disabled until service is performed.

2.4 Condition-based Service Policies using Virtual Sensors

Rather than performing service at fixed intervals or solely when comfort violations occur, an alternative policy based on the actual condition of the air-conditioner could be used instead. For example, service could be performed only when a significant fault is present if the air conditioner has an automated fault detection and diagnostics system installed. Furthermore, instead of calling for service when a fault is detected or diagnosed, virtual sensors could be used to call for service when performance has degraded past a certain point. For example, the virtual cooling capacity sensor could be used to monitor capacity degradation and call for service when system capacity decreases below 10% of the normal capacity.

Using real-time data to prioritize maintenance by observing the state of the system is known as condition-based maintenance. In comparison to emergency service policies, condition-based maintenance may reduce comfort violations since maintenance could be performed before capacity is degraded significantly. In comparison to periodic maintenance policies, condition-based maintenance may reduce operating costs by requiring service time only when significant faults exist. A problem with periodic service policies is that the underlying assumptions about the building loads and rate of performance degradation must be estimated a-priori or re-evaluated annually (or at some other interval). If building loads change or the rates of degradation change, the condition-based maintenance policy would be able to adapt.

Three condition-based service policies were implemented using the simulation framework to determine how operating costs are affected. The first policy performed service when the total cooling capacity degradation exceeded a threshold, δ_{cool},

$$FIR_{cool} \underset{\omega_0}{\overset{\omega_1}{\gtrless}} \delta_{cool} \tag{7}$$

where FIR_{cool} is the ratio of actual cooling capacity to normal cooling capacity at the current operating condition, ω_1 is the decision rule to perform service, and ω_0 is the decision rule to not perform service. The second policy performed service when the COP was degraded more than a threshold, δ_{COP},

$$FIR_{COP} \underset{\omega_0}{\overset{\omega_1}{\gtrless}} \delta_{COP} \tag{8}$$

where FIR_{COP} is the ratio of actual COP to normal COP at the current operating condition. The last policy considered the impact on energy consumption and performed service when the electrical energy consumed by the system exceeded a threshold, δ_{elec},

$$FIR_{elec} \underset{\omega_0}{\overset{\omega_1}{\gtrless}} \delta_{elec} \tag{9}$$

where FIR_{elec} is the ratio of actual energy consumed by the system to the normal energy consumption of the system at the same operating condition. In this work, the fault impact ratio thresholds were: $\delta_{cool} = 0.9$, $\delta_{COP} = 0.9$, and $\delta_{elec} = 1.1$. A review of previous work can be consulted regarding practical implementations of these fault impact ratios using virtual sensors (Li & Braun, 2007a, 2007b).

2.5 Historical Operating Cost Impact Service Policies using Fault Impact Estimates

The performance of the proactive, reactive, and condition-based maintenance strategies described thus far are far from optimal for most scenarios. Reactive strategies (emergency service policies) are particularly problematic when equipment is significantly oversized relative to the building load and require comfort violations before service is requested. Proactive maintenance strategies, like periodic service, may reduce comfort violations and decrease utility cost impacts. However, if periodic service is scheduled too often, additional service costs may outweigh any utility cost savings accrued by keeping equipment in tip-top shape. Periodic service intervals should be adjusted if building loads change or equipment starts to degrade at different rates overtime. To account for these changes, condition-based service strategies may be applied to equipment with automated fault detection and diagnostics systems. Comparing the actual performance of the equipment with a model of normal performance, service decisions can be requested when performance has been degraded significantly. Identifying the optimal degradation threshold is not trivial and depends on the equipment sizing and rate of degradation over time.

To overcome these limitations, automated fault detection and diagnostics systems can be extended to account for operating cost impacts of running equipment with faults and performing service over time. In other words, heuristics or simplifications to the underlying maintenance decision problem formulation can be applied to approximate the optimal solution in real-time. For example, Rossi and Braun developed an operating cost based service policy (Rossi & Braun, 1996). This policy is implemented using a service decision rule based on the cost of service (C_s) and electricity cost (C_u) given by

$$C_u \int_{t_0}^{t} \bar{h}_t\left(x_t, f_t\right) dt \underset{\omega_0}{\overset{\omega_1}{\gtrless}} C_s \tag{10}$$

where t_0 is the time service was last performed, and the net accumulated energy consumption benefit to perform service task a_i is given by

$$\bar{h}_t\left(x_t, f_t\right) = \bar{h}_{t-1}\left(x_{t-1}, f_{t-1}\right) + \frac{\Delta t_{run}}{\tau_h}\left[h_t\left(x_t, f_t\right) - \bar{h}_{t-1}\left(x_{t-1}, f_{t-1}\right)\right] \tag{11}$$

where x_t represents the external driving conditions that affect system performance, f_t is a vector containing each fault level, and τ_h is the time constant of a low pass filter used to reduce the diurnal fluctuations on system performance and ensure $\bar{h}$ is an increasing function. The estimated power consumption savings for performing service at any time are calculated using

$$h\left(x_t, f_t\right) = P\left(x_t, f_t\right) - P\left(x_t, f^0\right) \tag{12}$$

where $P(x_t, \cdot)$ is a function that estimates the power consumption of the system at given driving conditions and fault levels, and f^0 is the equipment state immediately after service (when all the fault levels have returned to normal). The rule states that a service task should be performed when the accumulated energy impact (the left-hand-side) since the last service is greater than the service cost (right-hand-side) to fix the fault.

A significant limitation to the methodology developed by Rossi and Braun is that decisions between different maintenance tasks cannot be directly handled (Rossi & Braun, 1996). This is because only the total utility cost impact is considered, and the relative importance of multiple faults is not estimated. Thus, it is impossible to calculate the net benefit of performing different service tasks and selecting the option that provides the maximum benefit at each decision stage. For example, when an air conditioner has relatively minor condenser fouling, yet significant evaporator fouling – the optimal service task is often to change the evaporator air filter only since it is relatively inexpensive and the evaporator fouling likely impacts the system more significantly than the condenser fouling. Using Rossi's method, this service task would be delayed until the utility cost impact became greater than the cost of evaporator fouling service and condenser fouling service. As a result, the system operates at a lower average net efficiency.

In order to improve the original simplified method developed by Rossi and Braun, previously developed virtual sensor approaches for automated fault detection and diagnostics and fault impact evaluation models (Kim & Braun, 2015, 2016; Li & Braun, 2007b, 2009) are used within a modified method to the estimate benefits of performing different service tasks. The inclusion of these measures provides two sources of information that can make deciding between service tasks possible: measured fault levels and isolated fault impacts. In addition to handling multiple fault service decisions, the operating cost function was modified to include equipment cost impacts that account for the effects of increased equipment run-time on replacement costs.

The classification rule of Rossi and Braun has been modified to inform the service action taken at any point in time, t,

$$\int_{t_{0,i}}^{t} \left[C_u \cdot \bar{h}_t\left(x_t, f_t, a_i\right) + C_e \cdot \bar{g}_t\left(x_t, f_t, a_i\right)\right] dt \underset{\omega_0}{\overset{\omega_i}{\gtrless}} C_s\left(a_i\right) \quad \forall a_i \in A \tag{13}$$

where $t_{0,i}$ is the time since the i^{th} component was previously serviced, C_u is the cost per unit time of electricity consumption, C_e is the time averaged equipment replacement costs assuming the system has a finite number of run-time hours, $C_s\left(a_i\right)$ is the service cost required to perform service task a_i to repair the i^{th} component, $\bar{h}_t$ is the net accumulated energy consumption benefit to perform service task a_i given by

$$\bar{h}_t\left(x_t, f_t, a_i\right) = \bar{h}_{t-1}\left(x_{t-1}, f_{t-1}, a_i\right) + \frac{\Delta t_{run}}{\tau_h}\left[h_t\left(x_t, f_t, a_i\right) - \bar{h}_{t-1}\left(x_{t-1}, f_{t-1}, a_i\right)\right] \tag{14}$$

and $\bar{g}_t$ is the net accumulated equipment runtime savings to perform service task a_i given by

$$\bar{g}_t\left(x_t, f_t, a_i\right) = \bar{g}_{t-1}\left(x_{t-1}, f_t, a_i\right) + \frac{\Delta t_{run}}{\tau_g}\left[g_t\left(x_t, f_t, a_i\right) - \bar{g}_{t-1}\left(x_{t-1}, f_t, a_i\right)\right] \tag{15}$$

where x_t represents the external driving conditions that affect system performance, f_t is a vector containing each fault level, and τ_h, τ_g are the time constants of low pass filters used to reduce the diurnal fluctuations on system performance and ensure $\bar{h}$ and $\bar{g}$ are increasing functions.

The estimated power consumption savings for performing service task a_i at any time are calculated using

$$h\left(x_t, f_t, a_i\right) = P\left(x_t, a_0\left(f_t\right)\right) - P\left(x_t, a_i\left(f_t\right)\right) \tag{16}$$

where $P(x_t,\cdot)$ is a function that estimates the power consumption of the system at given driving conditions and fault levels, a_0 is the "do nothing" service task which has a functional form given by

$$a_0\left(f_t\right) \rightarrow f_t \tag{17}$$

which states the "do nothing" service task has no impact on the fault levels of the system and a_i is the i^{th} service task which is expressed as

$$a_i\left(f_t\right) \rightarrow f_t^i. \tag{18}$$

Conceptually, this means when action a_i is applied to the system, the fault levels affected by the service task are returned to their normal values (as if the faults did not exist). The result of Equation (16) is the difference between power consumption for the current fault levels and the power consumption for the system if the service task was performed on the system.

In a similar manner, the runtime savings for performing service task a_i is given by

$$g\left(x_t, f_t, a_i\right) = \Delta t\left(x_t, a_0\left(f_t\right)\right) - \Delta t\left(x_t, a_i\left(f_t\right)\right) \tag{19}$$

where $\Delta t(x_t,\cdot)$ is a function that estimates the run-time requirement of the system at given driving conditions and fault levels. In order to evaluate Equations (16) and (19), simplified semi-empirical models developed by Hjortland can be used to estimate the relative impact of different faults that occur simultaneously (Hjortland, 2018).

Despite the rather complex mathematical formulation, the decision rule described by Equation (13) has a relatively straightforward interpretation. The rule states that for all service tasks that can be applied to the system at each decision interval, a service task should be performed when the accumulated energy and equipment cost impacts (the left-hand-side) is greater than the service cost required to perform the task (right-hand-side). In order to account for discounted service costs when performing multiple service tasks at the same time, unions between multiple service actions should be included in the set of available service actions, A.

The resulting distributions of additional lifetime operating costs relative to optimal costs using the original (unmodified) service scheduler developed by Rossi and Braun for the random sample of fault combinations are shown in Figure 4. In this policy, the total accumulated energy impact is estimated and used to calculate the additional utility cost due to faults. At each decision interval, the net utility costs are compared to the costs of performing the three service tasks: adding refrigerant, cleaning the condenser coil, and changing the evaporator coil filter. Because all three service tasks are considered, this requires greater accumulated utility cost impact before service is performed. The additional lifetime operating costs relative to the optimal costs in Miami, Atlanta, and Chicago were 6.9%, 7.4%, and 8.7% respectively. Because Miami has the longest cooling season, this location has higher normal electrical energy consumption for cooling than the other locations. This makes preventative maintenance more cost effective since larger energy consumption savings are possible. The additional costs in Atlanta and Chicago are greater than Miami due to lower cooling load requirements. These locations also have a winter season, where no cooling is required which makes the service policy less effective. This is because the policy is not able to quantify the impact of charge leakage during the winter months – which makes doing service late in cooling season possible and creates a lag in the spring before accumulated impacts become greater than the cost of service. These are two behaviors the optimal policy can avoid since the optimization horizon is over the entire equipment life.

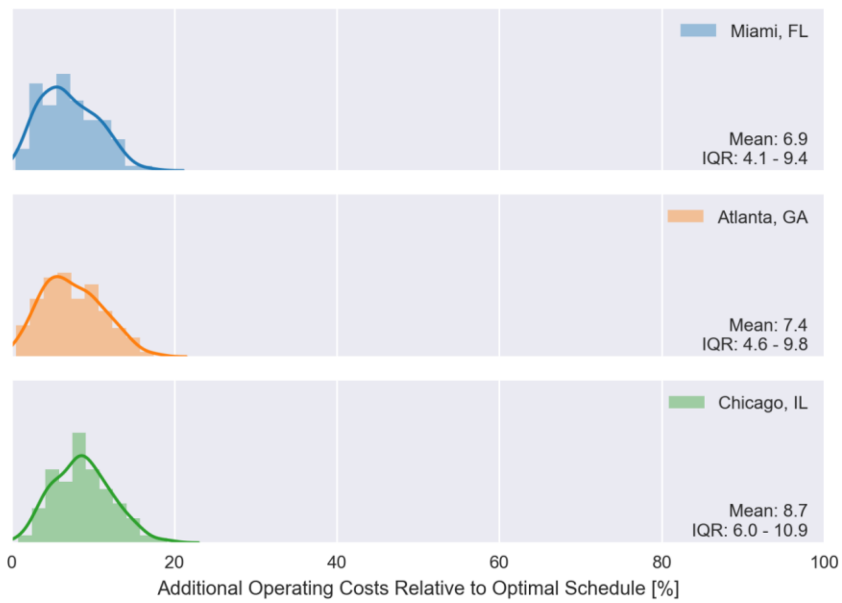

Figure 4. Additional lifetime operating costs relative to optimal lifetime costs when the simplified service decision methodology developed by Rossi and Braun is used to service multiple faults in three different locations.

The distributions of additional lifetime operating costs using the modified service scheduler that considers multiple service tasks for the random sample of fault combinations is shown in Figure 5. In this policy, the total accumulated energy impacts for different faults are estimated and used to calculate the additional utility cost consumed. At each decision interval, the net utility costs are compared to the costs of performing different combinations of service tasks. When the cost of one of the combinations of service tasks becomes less than the accumulated utility costs for the corresponding faults, service is performed. The additional lifetime operating costs relative to the optimal costs in Miami, Atlanta, and Chicago were 3.7%, 5.5%, and 5.7% respectively. Compared to the simplified service scheduler developed by Rossi and Braun, additional operating costs savings are possible using the methodology that isolates the impacts of different faults. The remaining costs above the optimal operating costs are caused by suboptimal scheduling around the winter season and not considering end of life payback intervals.

A summary comparing the inner-quartile range and mean additional lifetime operating costs for all the results comparing each service policy considered is included in Table 3. In general, policies that do not consider the condition of the system while determining when to do service (periodic service schedules) have the highest lifetime operating costs. Policies that consider the state of the equipment tend to have lower lifetime operating costs, though the metric used to determine when to do service has a significant impact on actual costs. Even in the extreme case, emergency service policies may have lower operating costs than periodic service policies since service costs are saved for systems that have minimal faults. Utilizing more information when determining when to do service generally reduces operating costs. These results were observed using the simplified service scheduler proposed by Rossi and Braun and with the modified service scheduler that considers different service tasks. These strategies tended to have the lowest lifetime operating costs.

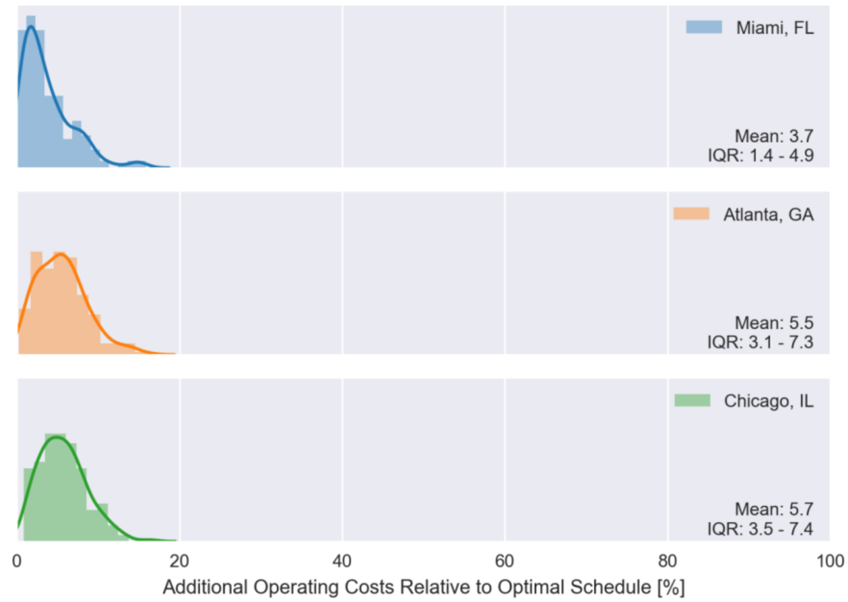

Figure 5. Additional lifetime operating costs relative to optimal lifetime costs when the modified service decision methodology that considers multiple service tasks is used to service multiple faults in three different locations.

Table 3. Summary of additional lifetime operating costs relative to optimal lifetime costs determined using dynamic programming for similar buildings in different locations. Inner-quartile ranges and means of 200 randomly sampled fault scenarios are compared.

	Miami, FL			Atlanta, GA			Chicago, IL		
Service Policy	25%	Mean	75%	25%	Mean	75%	25%	Mean	75%
Annual	30.0	31.5	33.0	41.7	44.6	47.2	61.6	65.2	68.6
Biennial	17.1	18.3	19.7	29.0	30.6	32.6	36.2	38.8	40.9
Emergency	28.8	34.2	41.2	20.0	23.4	28.2	18.4	22.4	27.1
CBM – Capacity 10% Threshold	11.5	16.8	21.8	17.2	20.8	24.2	17.7	21.0	25.1
CBM – COP 10% Threshold	6.2	10.4	13.8	12.8	15.3	18.1	15.7	20.2	25.8
CBM – Energy 10% Threshold	5.0	8.0	10.7	8.6	11.7	14.5	12.4	15.5	18.7
Accum. Total Cost (Rossi & Braun, 1996)	4.1	6.9	9.4	4.6	7.4	9.8	6.0	8.7	10.6
Accum. Indiv. Cost (Hjortland, 2018)	1.4	3.7	4.9	3.1	5.5	7.3	3.5	5.7	7.4

Even using the service decision strategies that consider accumulated impacts, the lifetime operating costs were still appreciably higher than the optimal service schedule costs. This occurs for two reasons: 1) the policies do not handle scheduling service around winter seasons optimally and 2) the policies do not optimize service well towards the end of equipment life. It is believed that these deficiencies could be corrected in future work. One possible solution to these problems would be to adapt the service schedulers to use a future optimization horizon. This would require a model for expected utility cost savings during the prediction horizon, as well as model for how the faults would evolve. Since the loads and operation throughout the year and life of HVAC equipment is largely cyclical, it may be possible to learn this model using trended data from past performance.

A second improvement to the models that may help to avoid thermal comfort violations would be to pre-schedule service using a model for the peak cooling loads and the capacity degradation measurements. Rather than allowing the system to cause thermal comfort violations due to insufficient capacity, an algorithm that estimates the peak cooling load over a prediction horizon could be used to calculate the minimum capacity needed. Using virtual sensor measurements, a service schedule could determine if the system will be able to meet all future loads and schedule service if it cannot.

4. CONCLUSIONS

Several different types of maintenance strategies have been implemented and compared using a simulation framework that models the interaction between building cooling loads and equipment performance while faults evolve over time. As a common benchmark, dynamic programming was used to find solutions to an optimal service scheduling problem that was formulated to minimize lifetime operating costs by performing service tasks when they are most cost effective during the equipment life. For each simulation scenario considered, the optimal service scheduling problem was solved, and an optimal service policy function was used to calculated optimal operating costs.

Simple, fixed interval service policies were compared with the optimal policies for different fault rates. These comparisons showed that periodic maintenance policies can often lead to significant increases in operating costs, especially if faults grow slowly over time. Reactive maintenance strategies that perform maintenance only when comfort is violated were also implemented and simulated using the framework. Policies that base their decisions on the condition of equipment tended to have less lifetime operating costs.

REFERENCES

Hjortland, A. L. (2018). *Automated Fault Detection, Diagnostics, Impact Evaluation, and Service Decision-making for Direct Expansion Air Conditioners.* (Ph.D.), Purdue University, West Lafayette, IN.

Hjortland, A. L., & Braun, J. E. (2018). *Artificial Neural Networks for Fast Rooftop Unit Fault Impact Modeling and Simulation*. Paper presented at the International Refrigeration and Air Conditioning Conference, West Lafayette, Indiana.

Kim, W., & Braun, J. E. (2015). Extension of a virtual refrigerant charge sensor. *International Journal of Refrigeration, 55*, 224-235. doi:10.1016/j.ijrefrig.2014.09.015

Kim, W., & Braun, J. E. (2016). Development and evaluation of virtual refrigerant mass flow sensors for fault detection and diagnostics. *International Journal of Refrigeration, 63*, 184-198. doi:10.1016/j.ijrefrig.2015.11.005

Li, H., & Braun, J. E. (2007a). Decoupling features and virtual sensors for diagnosis of faults in vapor compression air conditioners. *International Journal of Refrigeration, 30*(3), 546-564. doi:10.1016/j.ijrefrig.2006.07.024

Li, H., & Braun, J. E. (2007b). An Overall Performance Index for Characterizing the Economic Impact of Faults in Direct Expansion Cooling Equipment. *International Journal of Refrigeration, 30*(2), 299-310. doi:10.1016/j.ijrefrig.2006.07.026

Li, H., & Braun, J. E. (2009). Decoupling features for diagnosis of reversing and check valve faults in heat pumps. *International Journal of Refrigeration, 32*(2), 316-326. doi:10.1016/j.ijrefrig.2008.05.005

Rossi, T. M., & Braun, J. E. (1996). Minimizing Operating Costs of Vapor Compression Equipment with Optimal Service Scheduling. *HVAC&R Research, 2*(1), 3-25. doi:10.1080/10789669.1996.10391330

Yuill, D. P., & Braun, J. E. (2017). A Figure of Merit for Overall Performance and Value of AFDD Tools. *International Journal of Refrigeration, 74*, 651-661. doi:10.1016/j.ijrefrig.2016.11.015