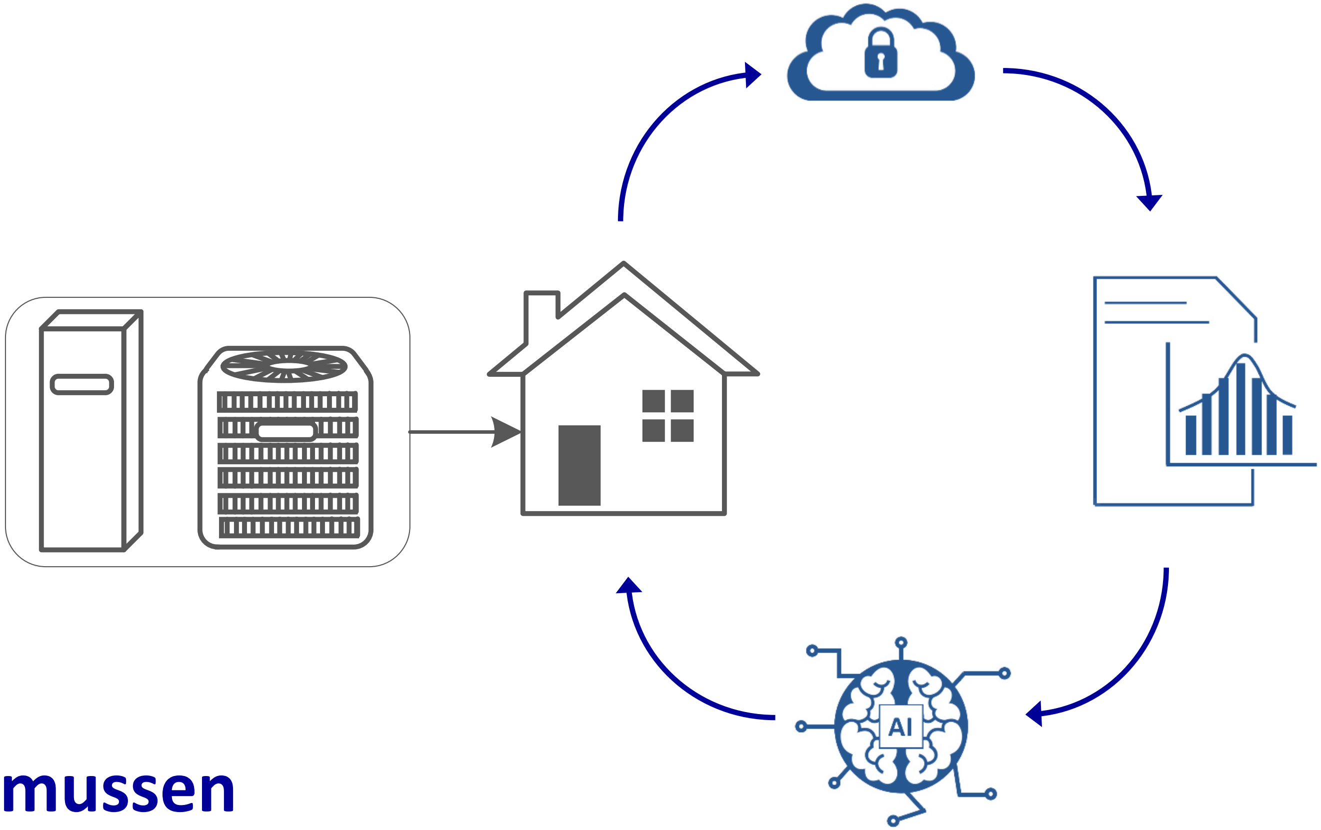
Cloud-based Fault Detection:

Leveraging Big Data to Reduce HVAC Energy Usage



Bryan Rasmussen

Professor
Mechanical Engineering
Texas A&M University

Aug. 20, 2020

TEXAS A&M ENGINEERING

Acknowledgements

Research Sponsor

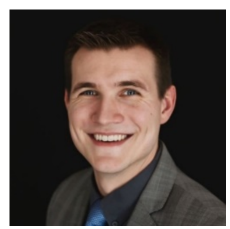
This material is based upon work supported by Trane Technologies. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Trane.

Specific Acknowledgment

Mechanical Engineering

Dr. Bryan Rasmussen

Austin Rogers
(2017-2019)



Fangzhou Guo
(2018-2021)

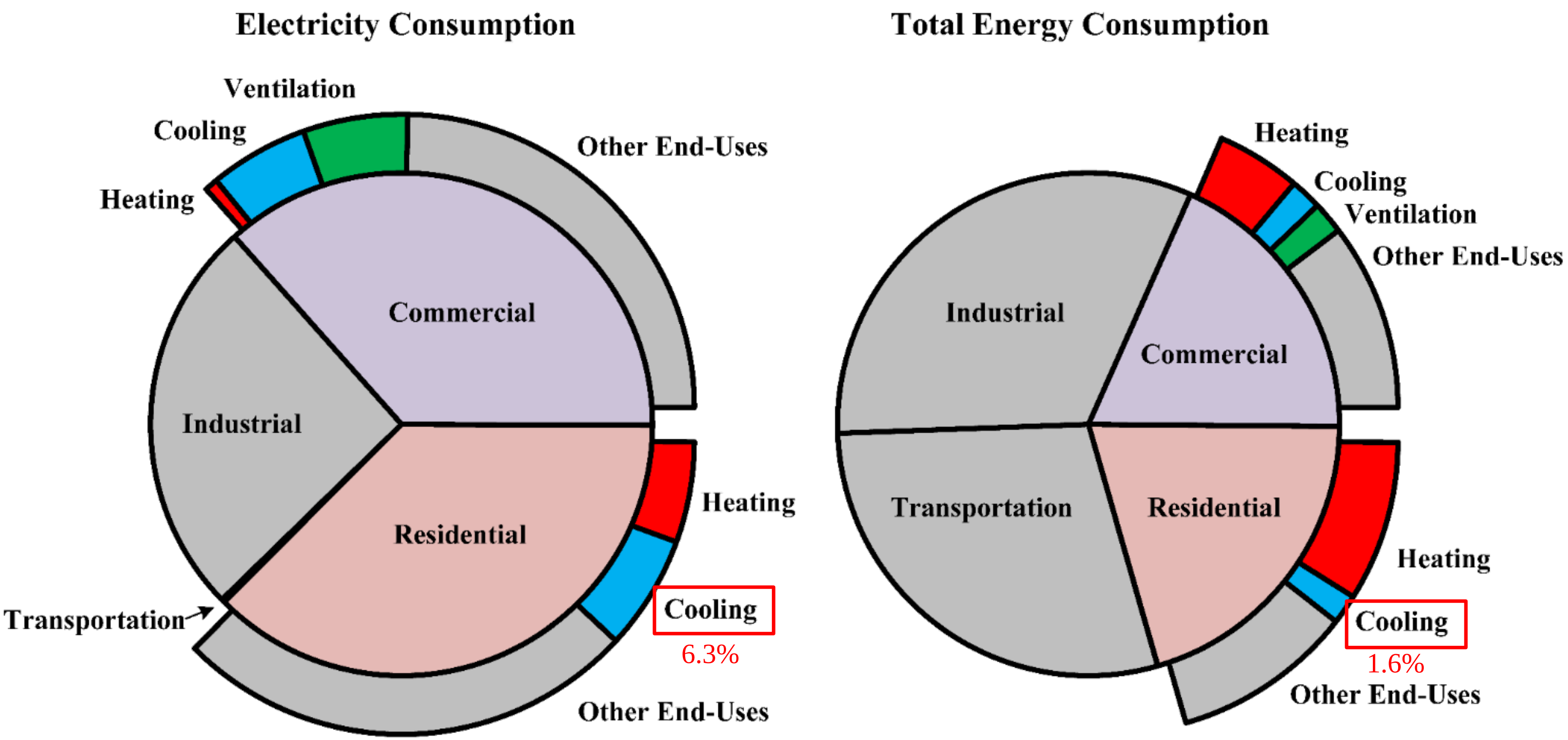
Arvind Phani Vadali
(2020-2021)

Computer Science and Engineering

Dr. Xia Ben Hu
Asst. Professor

Jonathan Martinez
(2017-2019)

Why focus on residential air conditioning?



Data obtained from U.S. EIA [1,2]

[1] EIA. 2015. Residential Energy Consumption Survey (RECS). U.S. Energy Information Administration.
[2] EIA. 2012. Commercial Buildings Energy Consumption Survey (CBECS). U.S. Energy Information Administration.

Residential A/C systems are designed to be robust

Air conditioner faults often result in decreased efficiency, while continuing to (partially) meet capacity requirements

Figure Credit: https://blog.ccacac.com/

Figure Credit: https://www.candcair.com/blog/

Field surveys indicate that faults are common and impact performance [3,4,5]

- Residential air conditioners operate at least 17% below their design performance.
- 68% of residential systems require repair
- 40% of systems had more than 17% undercharge
- A 16% increase in efficiency is possible by correcting only airflow and charge faults

[3] Li and Braun, 2007, A Methodology for Diagnosing Multiple Simultaneous Faults in Vapor-Compression Air Conditioners. *HVAC&R Research*.
[4] Katipamula et al., 2015, Rooftop Unit Embedded Diagnostics: Automated Fault Detection and Diagnostics (AFDD) Development, Field Testing and Validation, PNNL-23790, Richland, WA.
[5] Downey, T., and J. Proctor. 2002. What can 13,000 Air Conditioners Tess Us? In the Proceedings of the 2002 ACEEE Summery Study on Energy Efficiency in Buildings.

Breakdown of HVAC Fault Detection and Diagnosis (FDD) efforts

Research efforts have traditionally focused on **larger** systems and **simpler** systems.

The high cost of larger systems (chillers and RTUs) decreases the relative cost of FDD sensors.

Packaged units (chillers and RTUs) do not have the same variation across installations.

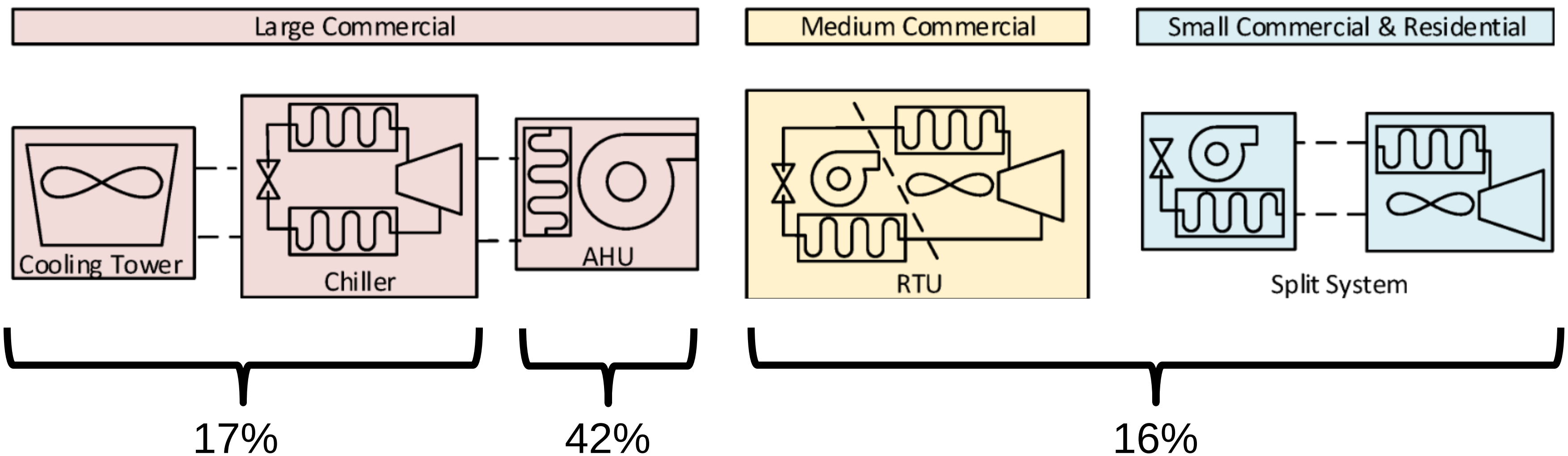
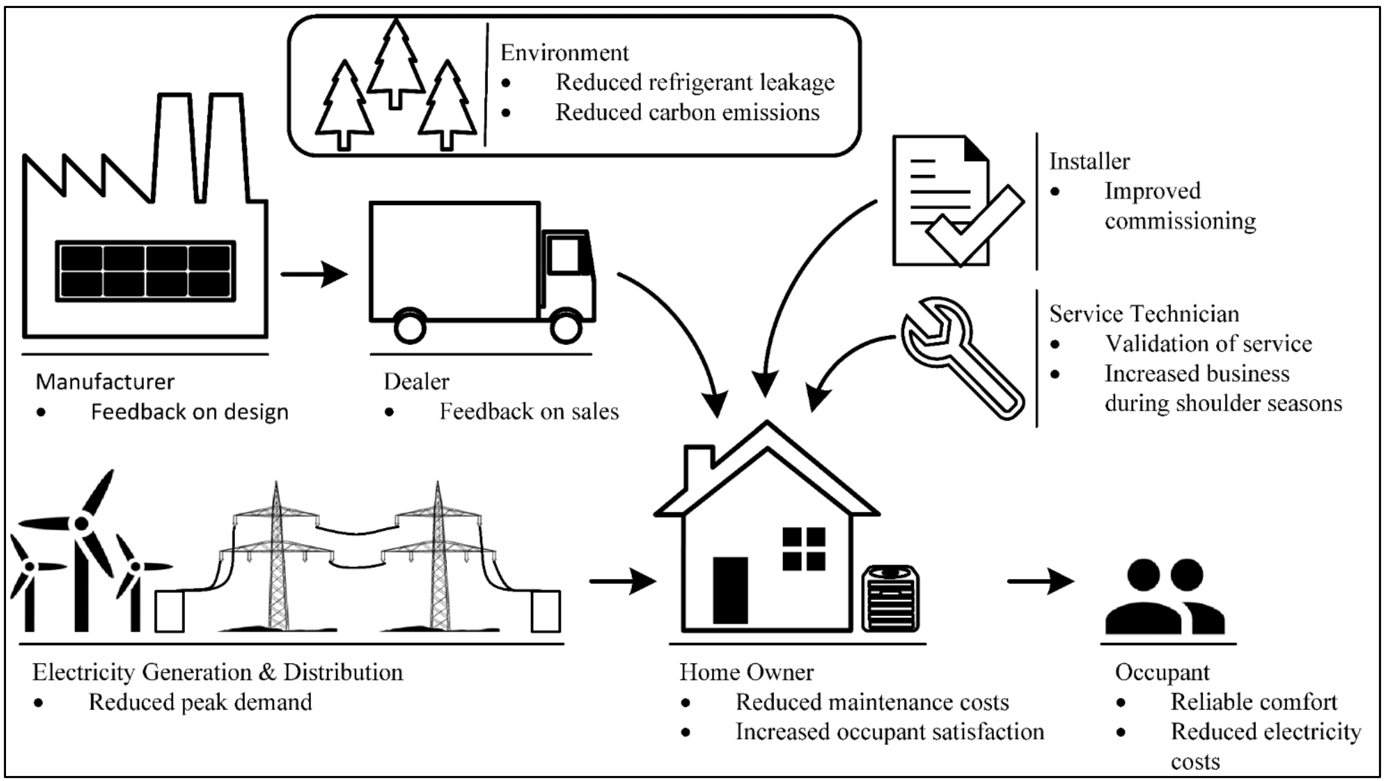


Figure: Portion of HVAC FDD literature focused on the respective systems [6]. Other categories (25%) include overall building, water heaters, refrigerators, lighting, etc.

[6] Kim and Katipamula, 2017, A review of fault detection and diagnostics methods for building systems. *Science and Technology for the Built Environment*.

Secondary benefits of fault detection and analytics

Fault detection and diagnosis (FDD) could provide other benefits throughout the value chain (besides energy savings).



Characterizing air conditioning faults

Many ways to characterize faults

- Hard vs. Soft [2]
- Control/Sensor/Electrical/ Mechanical [2]
- Operational vs. Service [3]
- Air- vs. Refrigerant-side [4]

Commonly Studied:

- Evaporator Filter
- Condenser Fouling
- Undercharge & Overcharge
- Liquid Line Restriction
- Compressor Valve Leakage

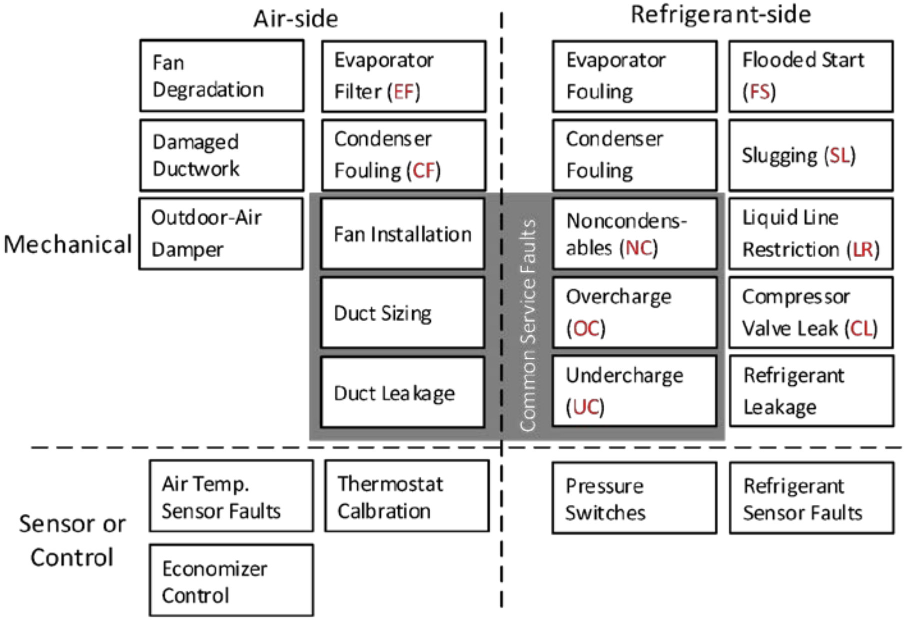


Figure: Several examples of soft faults in air conditioning systems.

[2] Breuker and Braun, 1998, Common Faults and Their Impacts for Rooftop Air Conditioners. *HVAC&R Research*.
[3] Li and Braun, 2007, A Methodology for Diagnosing Multiple Simultaneous Faults in Vapor-Compression Air Conditioners. *HVAC&R Research*.
[4] Katipamula et al., 2015, Rooftop Unit Embedded Diagnostics: Automated Fault Detection and Diagnostics (AFDD) Development, Field Testing and Validation, PNNL-23790, Richland, WA.

Common elements of HVAC FDD methods

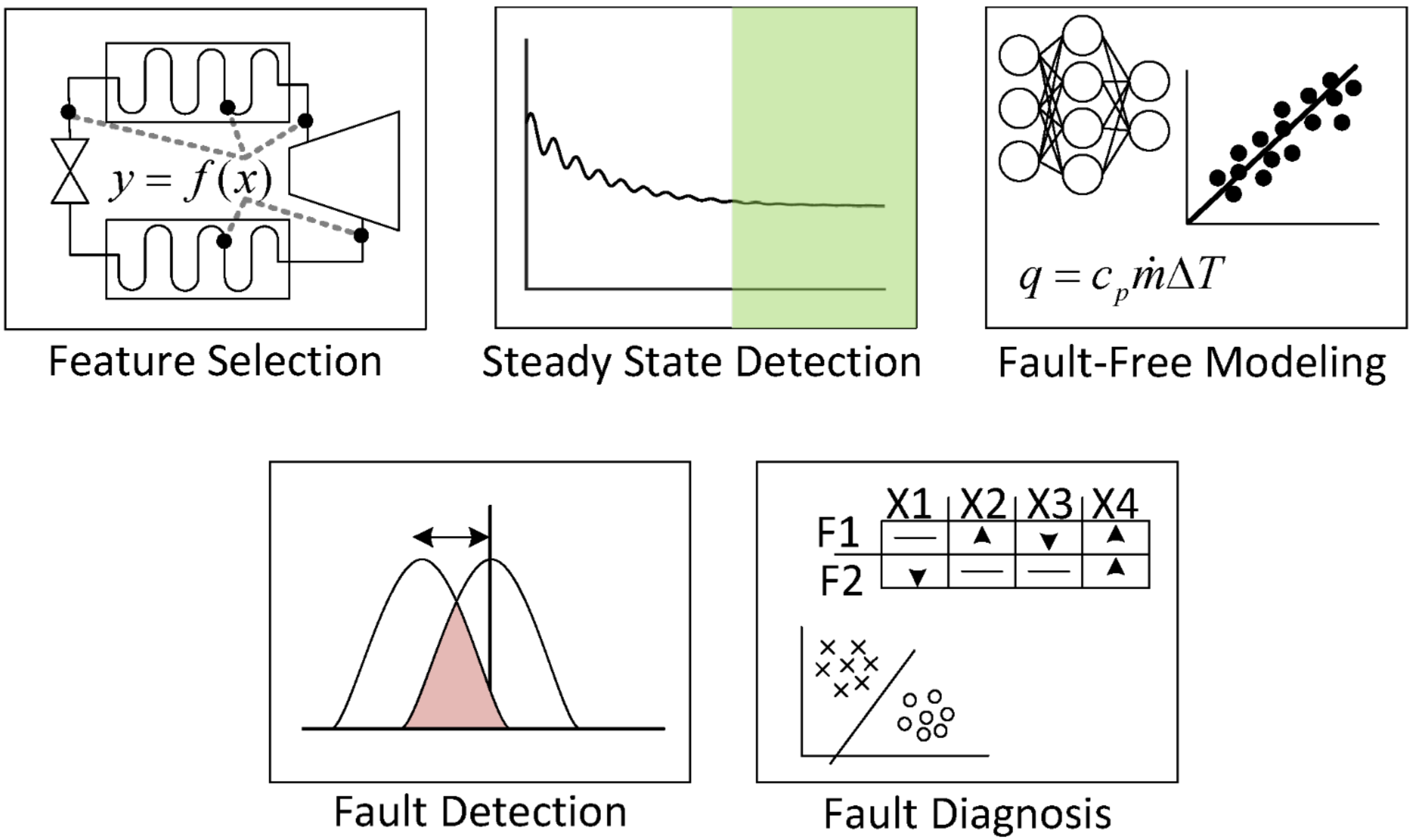
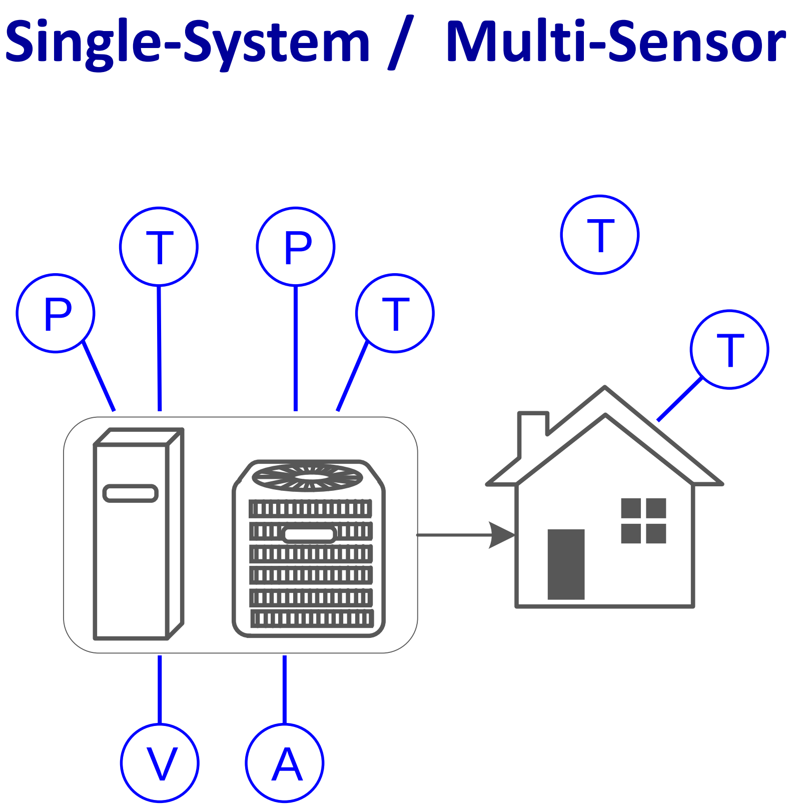
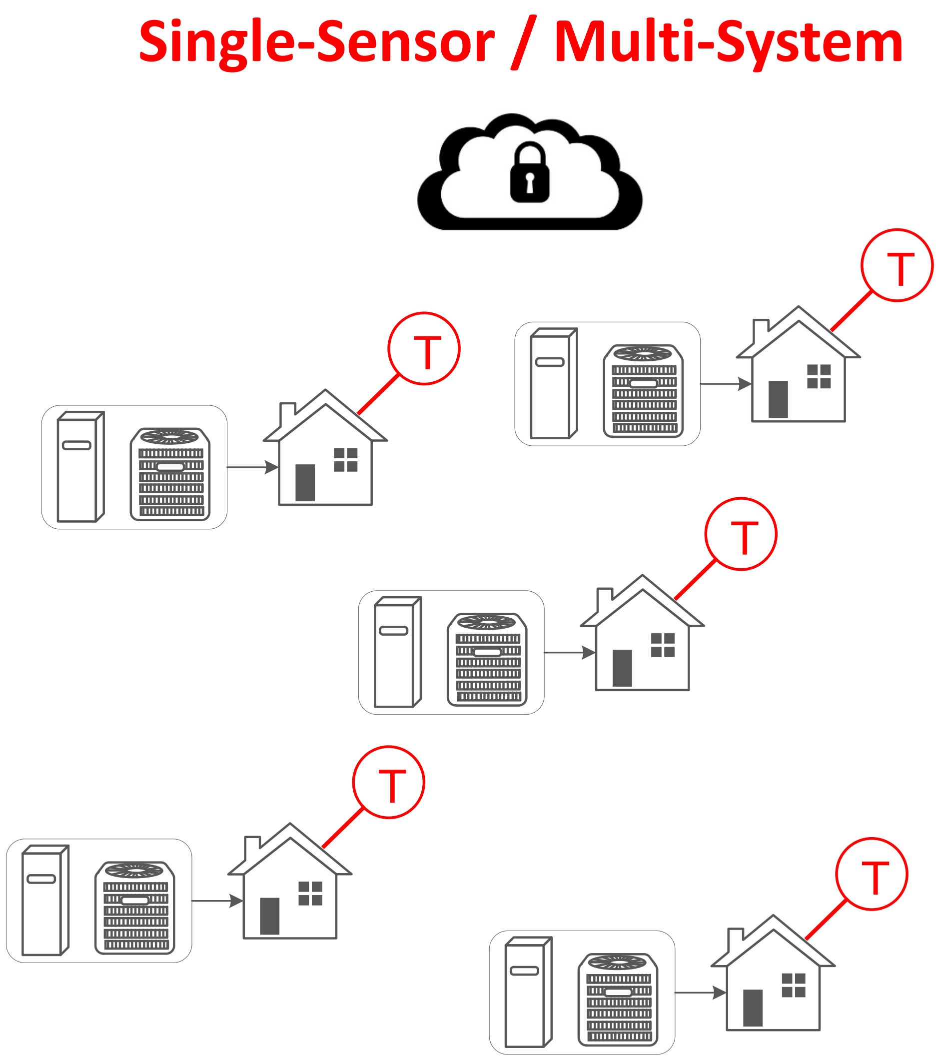


Figure: An overview of common FDD processes for air conditioning systems.

Paradigms for Residential Air Conditioning Fault Detection





Challenges and opportunities cloud-based FDD

Challenges:

- Limited sensor information
- Sensor/communication integrity problems
- Privacy and cybersecurity concerns
- Installation data inaccuracies
- Maintenance technician training

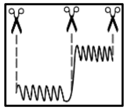
Opportunities:

- Potential for broader fault detection (e.g. recurring dealer installation faults)
- Baselining potential using field installed systems
- Leverage meta-data (e.g. home size, weather, equipment repair records)

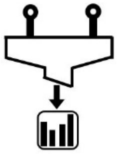
Proposed cloud-based FDD methodology

Data Management:

- Smart thermostat data retrieval and verification



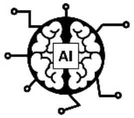
- Time based data parsing and mode labeling



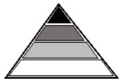
- Feature extraction

Analytics:

- Statistical analysis for outlier detection

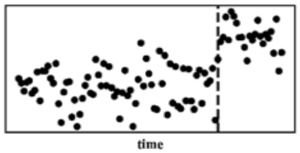


- Machine learning for fault diagnosis



- Fault impact quantification for repair prioritization

- Occupant behavior analysis
- Sudden change detection



- Degradation change detection

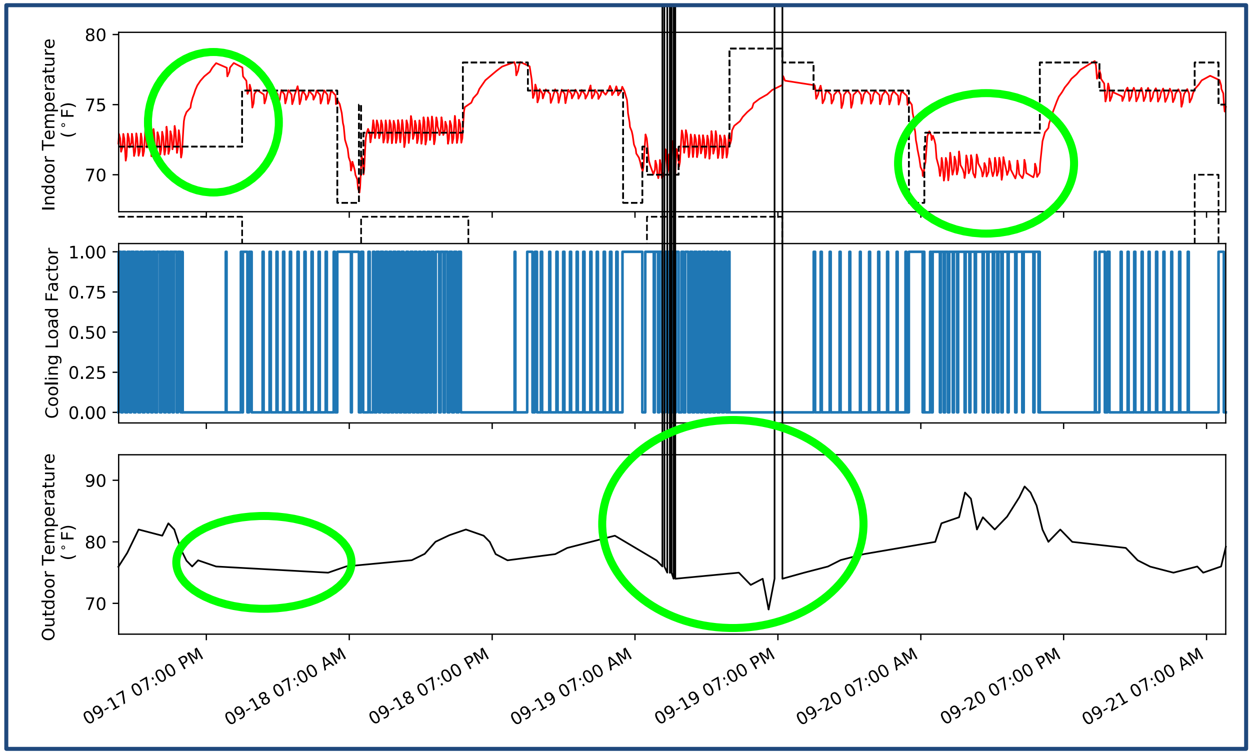
Data Management

Data retrieval and integrity verification

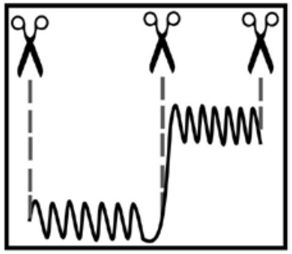
- Retrieve historical data from the cloud (10,000s to 100,000s of systems)
- Correct data issues caused by sensor or network faults.

Examples:

- Change in setpoint not reported
- Sensor errors



Data Management

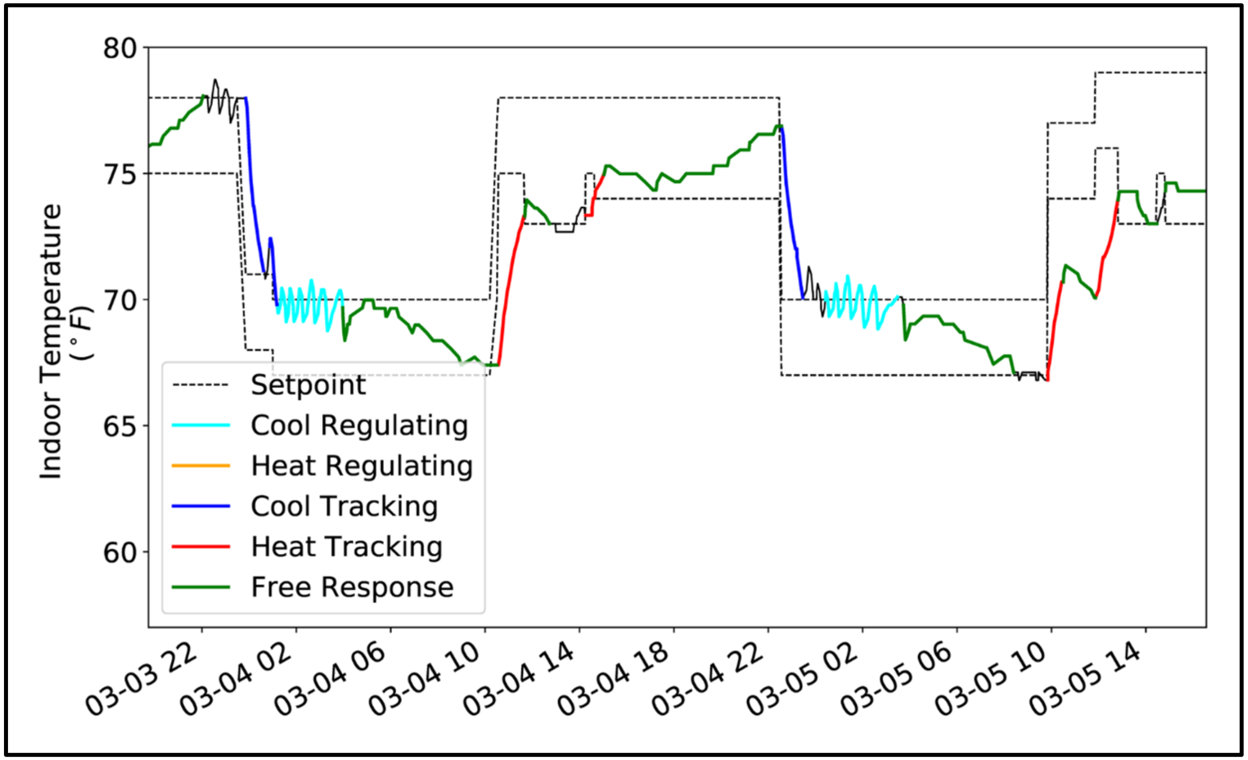


Data parsing and mode labeling

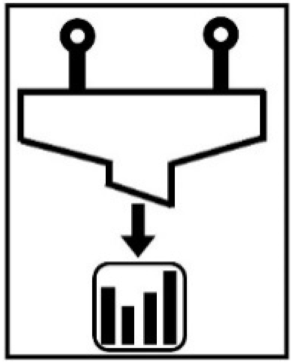
- Define and partition raw sensor data based on operational mode and transient behavior.

Examples:

- All modes experienced during a two day period



Data Management



Feature Extraction

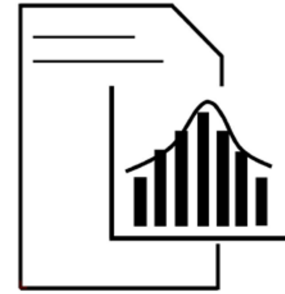
- Extract features from raw data that are indicative of faulty operating behavior.
- Different features of interest are extracted from both steady state and transient modes.

Examples:

- A few of the data features, and corresponding operating mode

Symbol	Feature Name	Regulating	Tracking	Free Response
ΔT_{oi}	Temperature difference [°F]	x	x	x
H_{id}	Indoor humidity ratio (kg H_2O/kg dry air)	x		
E_c, E_h	Cooling/heating effort [%]	x	x	
F	Cycle frequency [cycle/hr]	x		
T_{spc}, T_{sph}	Cooling/heating setpoint error [°F]	x		
$\Delta \dot{T}_{ii}$	Indoor temperature change rate [°F/hr]		x	x
$\Delta T_{i,inc}$	Indoor temperature increase [°F] (for cooling periods)		x	
Δt	Cycle duration [hr]		x	

Data Analytics

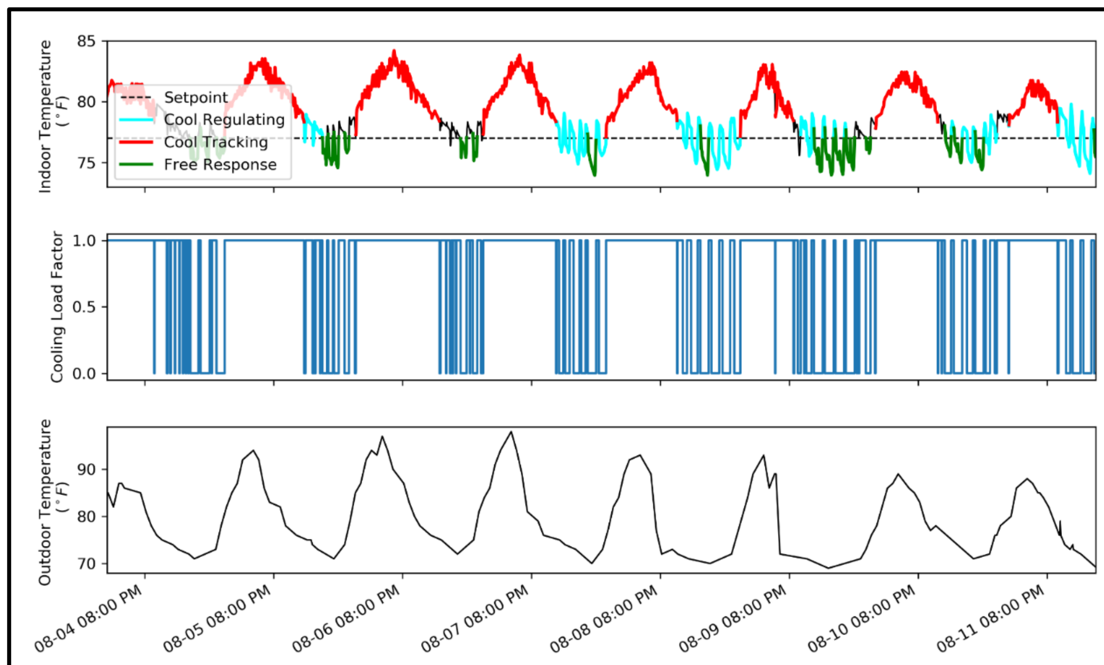


Multivariate statistical methods to identify:

- **abnormal behavior**
- degradation trends
- sudden changes

Examples:

- System fails to maintain comfortable temperatures despite working at 100% load



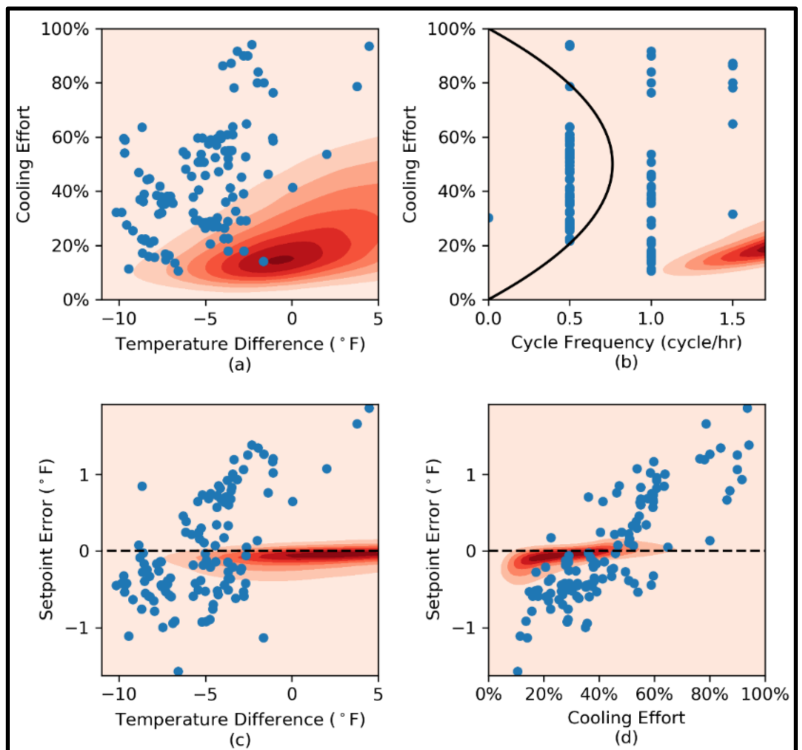
Data Analytics

Multivariate statistical methods to identify:

- **abnormal behavior**
- degradation trends
- sudden changes

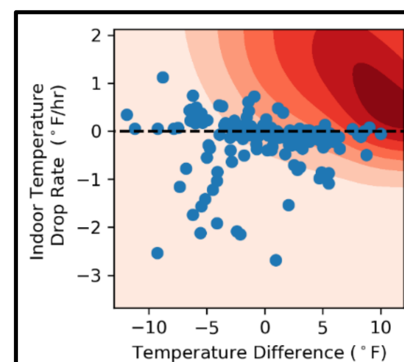
Examples:

- System fails to maintain comfortable temperatures despite working at 100% load



Poor pseudo steady-state behavior

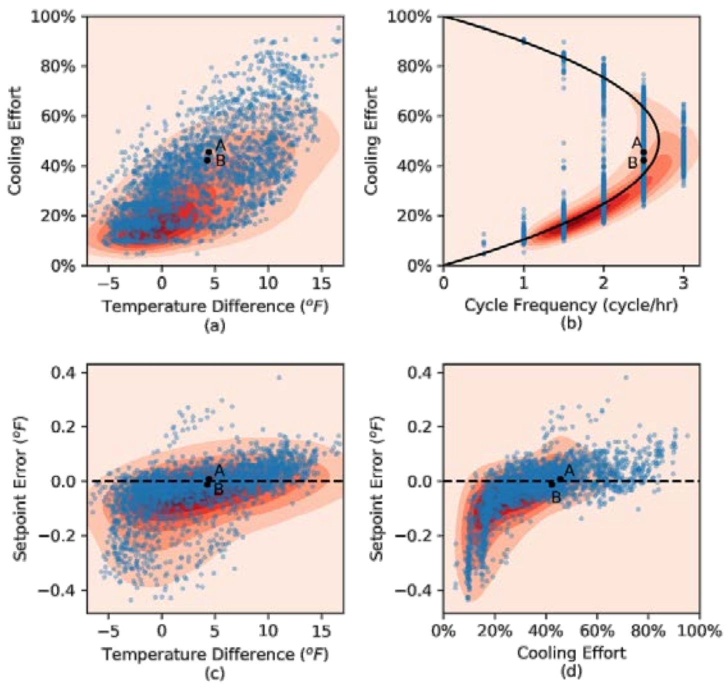
Poor transient behavior



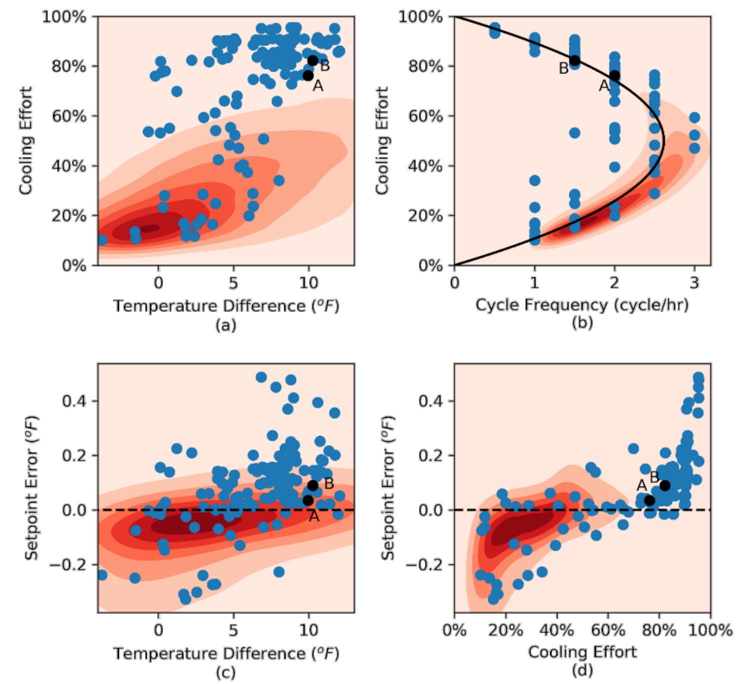
Data Analytics

Multivariate statistical methods

- Thermostat data is inherently sparse, but insight is possible using multi-system comparison
  - Features are chosen within the operating modes.
  - Population distribution is determined using Kernel Density Estimation (KDE).
  - Outlier detection performed using the Mahalanobis distance



A Normal System



Outlier System

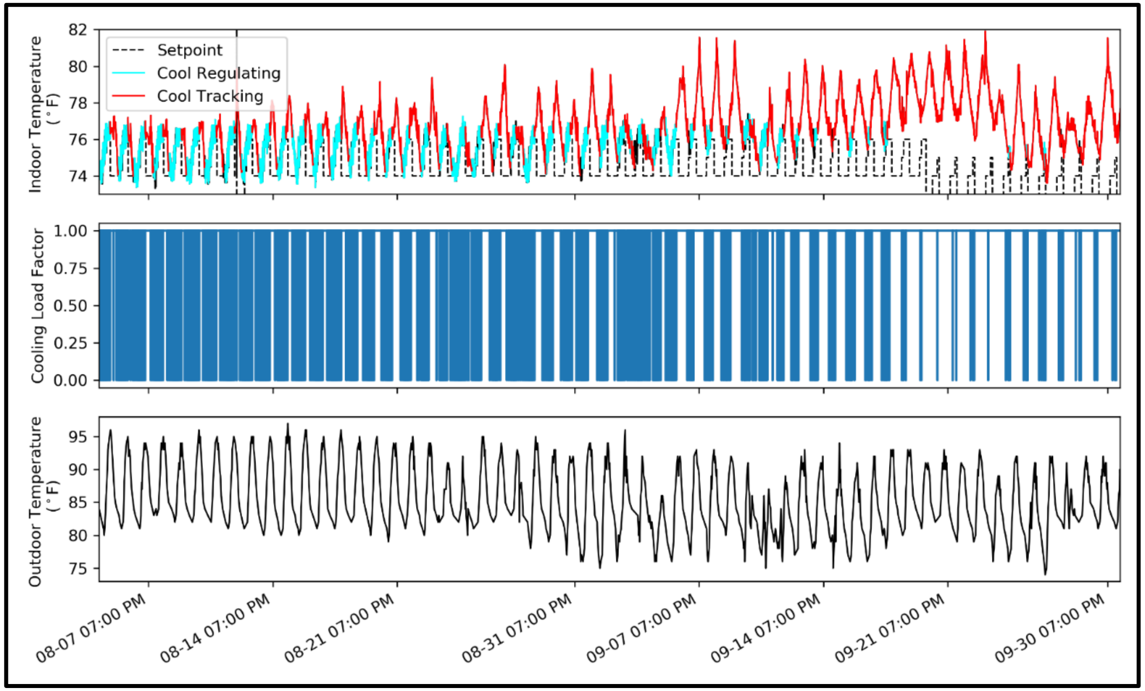
Data Analytics

Multivariate statistical methods to identify:

- abnormal behavior
- **degradation trends**
- sudden changes

Examples:

- Exhibits a slow degradation in performance over two months



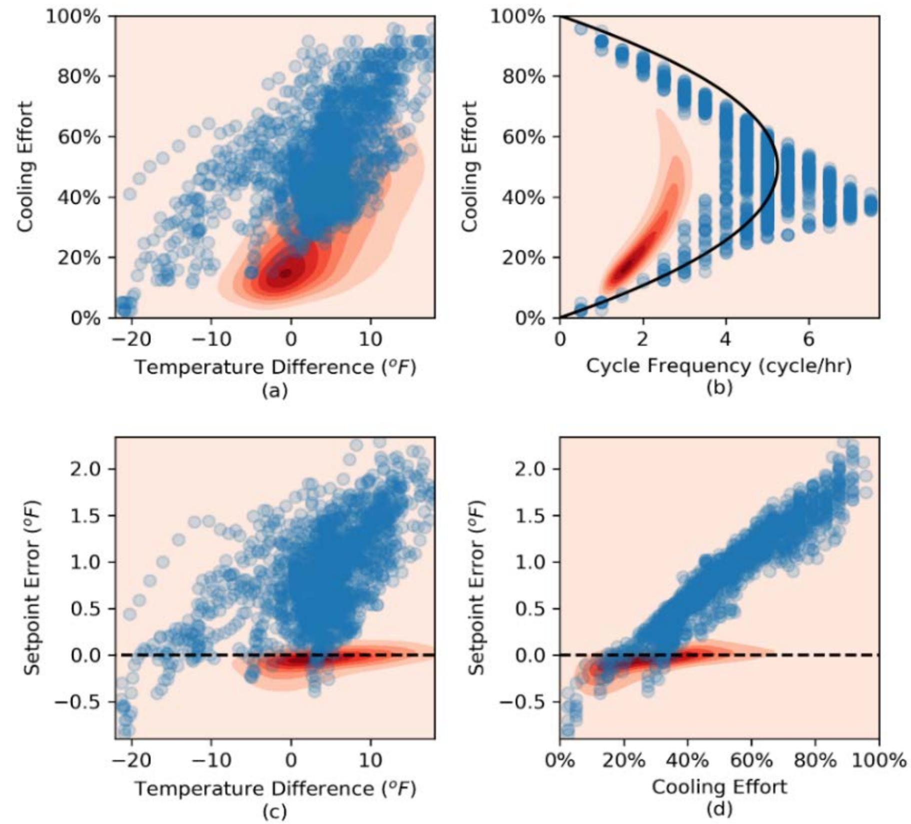
Data Analytics

Multivariate statistical methods to identify:

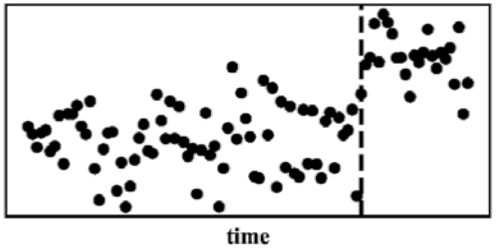
- abnormal behavior
- **degradation trends**
- sudden changes

Examples:

- Exhibits a slow degradation in performance over two months



Data Analytics



Multivariate statistical methods to identify:

- abnormal behavior
- degradation trends
- **sudden changes**

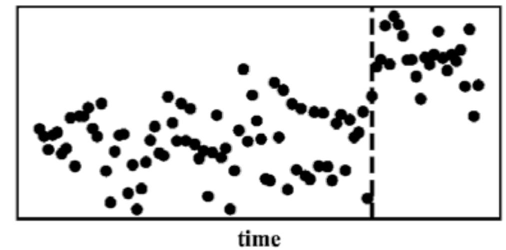
Examples:

- Computationally efficient code for detecting sudden changes in operation
- Does not require knowledge of underlying statistical distribution
- Implemented recursively

Example Python Code (10 lines)

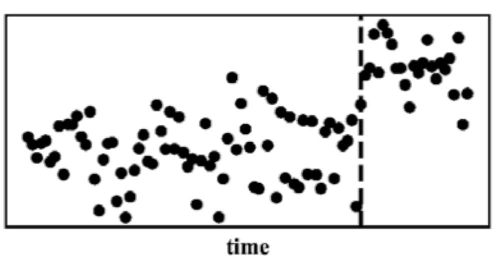
```
def f(x): return cumsum(x)
def g(x): return sum(x) – cumsum(x)
n_X = f(w)
mean_X = f(w*x) / f(w)
var_X = (f(w*x**2) + f(w)*mean_X**2 – 2*f(w*x)*mean_X) / f(w)
n_Y = g(w)
mean_Y = g(w*x) / g(w)
var_Y = (g(w*x**2) + g(w)*mean_Y**2 – 2*g(w*x)*mean_Y) / g(w)
sp = ((n_X-1)*var_X + (n_Y-1)*var_Y) / (n_X + n_Y – 2)
t = abs(mean_X – mean_Y) / sqrt(sp*(1/n_X + 1/n_Y))
```

Data Analytics

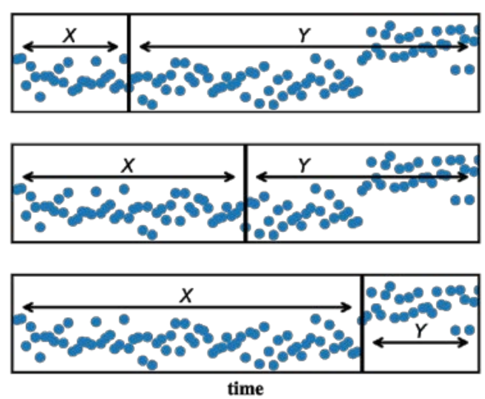


Multivariate statistical methods to identify:

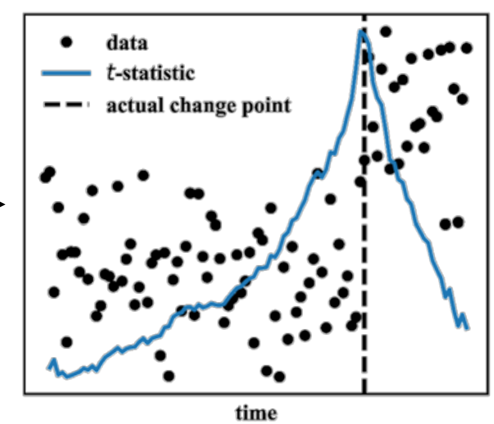
- abnormal behavior
- degradation trends
- **sudden changes**



(a) An example time series with a change point.

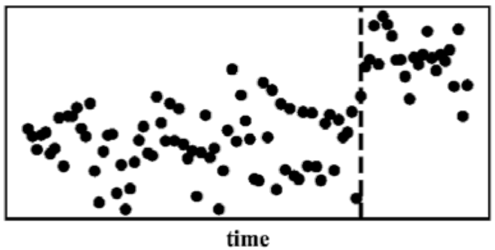


(b) The time series is bisected at <u>all</u> possible locations.



(c) The *t* statistic is calculated at all bisections. The maximum *t* statistic (MTS) occurs at the change point.

Data Analytics

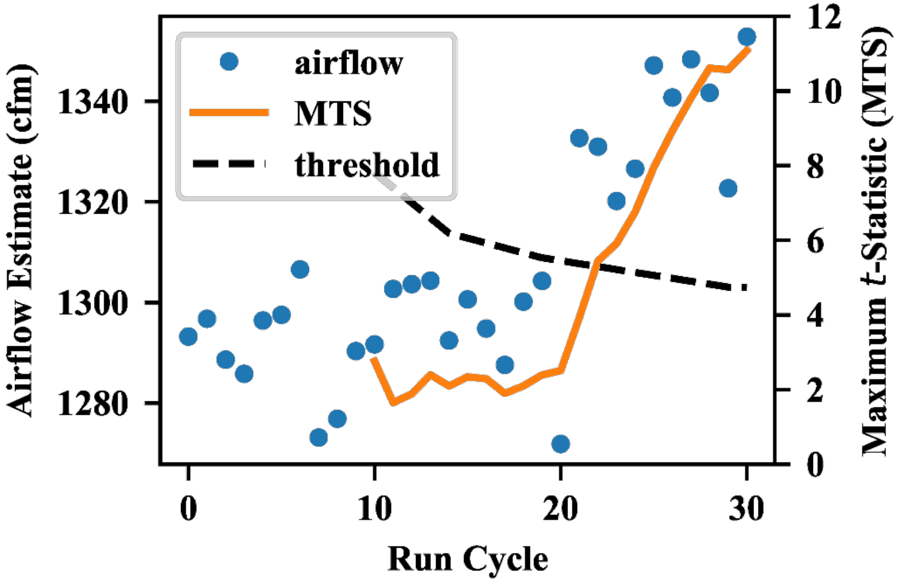


Multivariate statistical methods to identify:

- abnormal behavior
- degradation trends
- **sudden changes**

Examples:

- Used to detect replacement of air filter on residential system
- Detects a 3% change in estimated airflow after only two data points



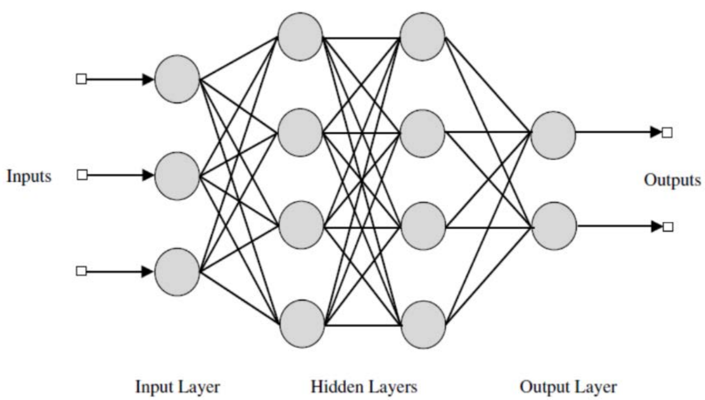
Data Analytics

Machine Learning methods can be used to leverage meta-data and identify previously unknown relationships

- Clustering algorithms
- Deep-learning networks

Challenges:

- Availability and quality of training data
- Computational load on server
    - **Millions** of systems
    - **Years** of recorded sensor measurements
    - **>> TB of data**



Data Analytics

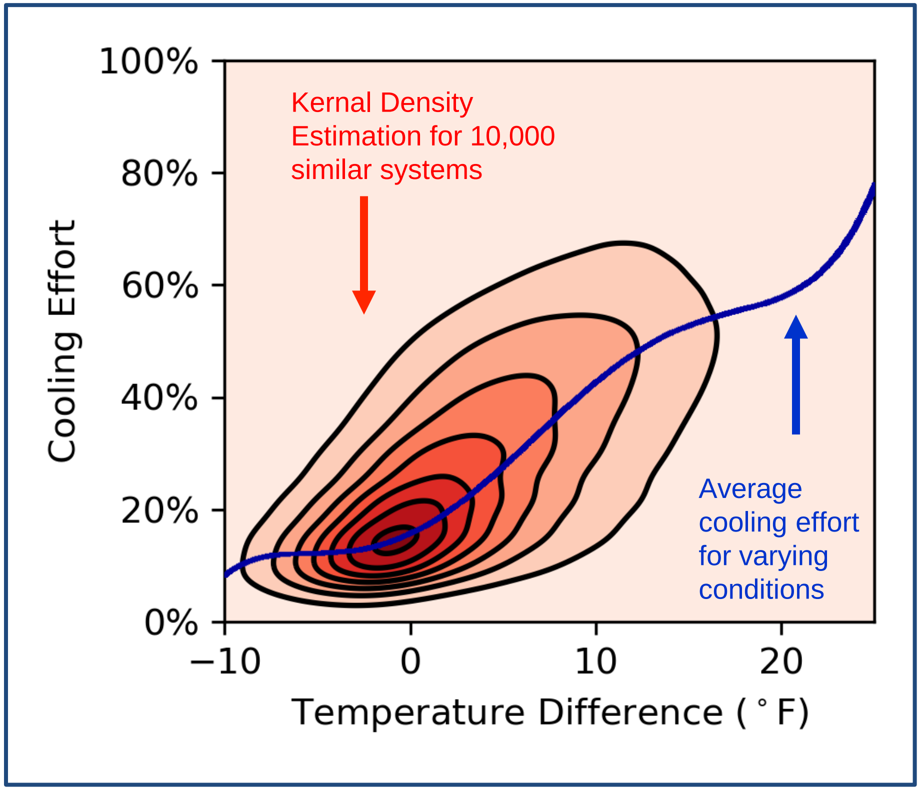
Fault impact quantification:

- Impact on energy use
- Impact on occupant comfort
- Allows for prioritization of repairs

Examples:

- Calculate the expected energy savings if the system operation were corrected and brought within $\pm 1\sigma$ of the average system performance

- Similar metrics are possible for occupant comfort



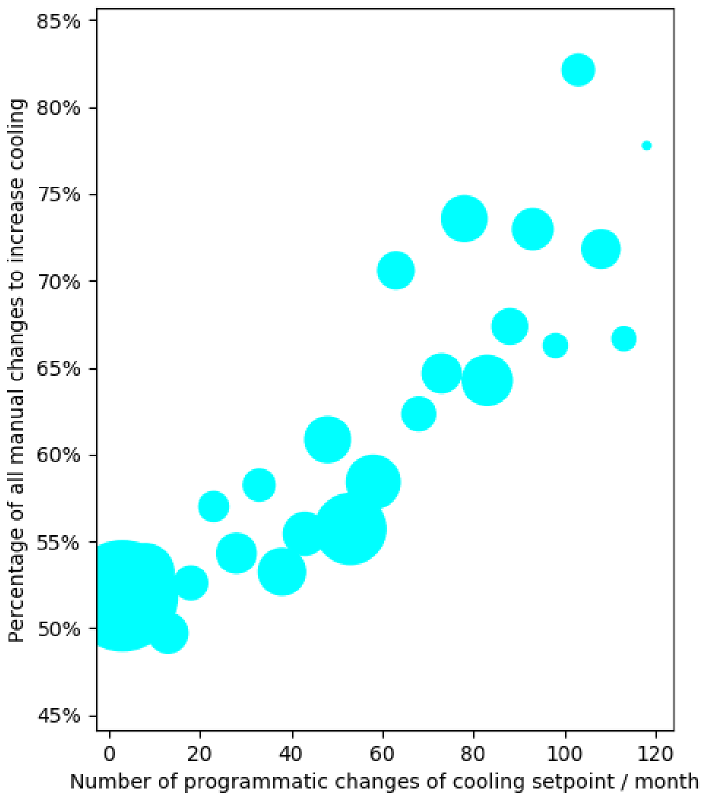
Data Analytics

Occupant Behavior:

- How do occupants interact with thermostats?
- What conditions prompt people to override thermostat programs?
- Are there demographic trends?

Examples:

- Manual vs. Programmatic Thermostat setpoint changes



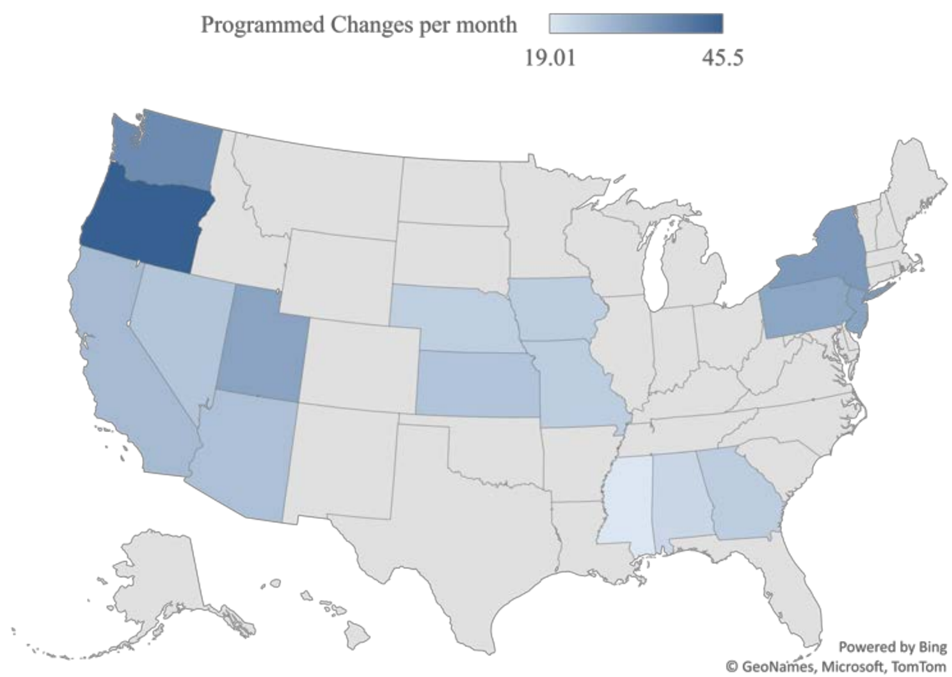
Data Analytics

Occupant Behavior:

- How do occupants interact with thermostats?
- What conditions prompt people to override thermostat programs?
- Are there demographic trends?

Examples:

- Number of thermostat programmatic temperature changes



Summary

Cloud-based data analytics can be leveraged for fault detection with limited sensor information.

Computationally practical methods for fault detection and analysis were demonstrated using database of smart thermostat data provided by Trane Technologies, Inc.

Smart thermostat data analytics has potential for enhancing occupant comfort and reducing energy usage, without additional equipment or sensors.

QUESTIONS?

Bryan Rasmussen

brasmussen@tamu.edu