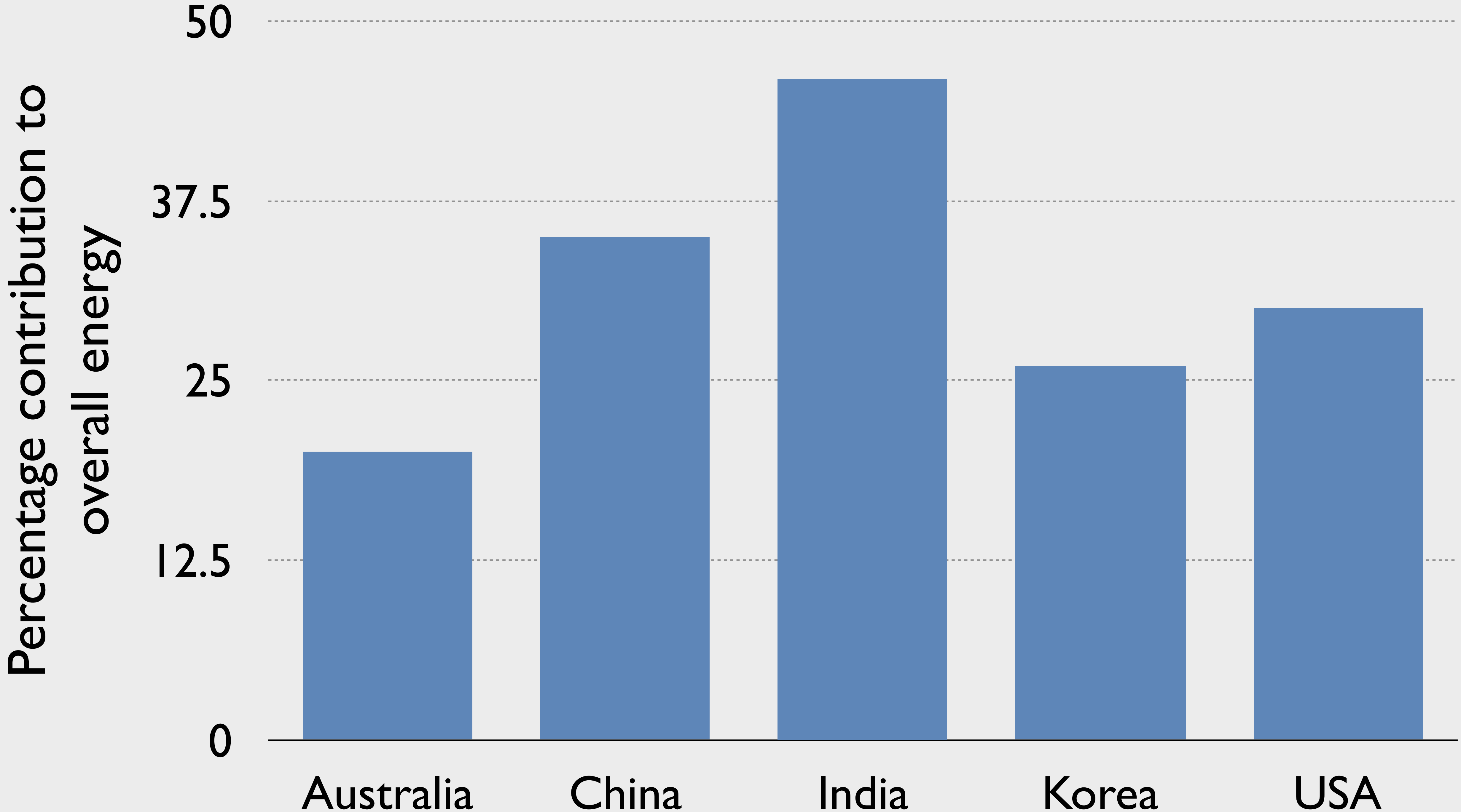
Making energy disaggregation **practical**

Nipun Batra
IIIT Delhi
November 1, 2015

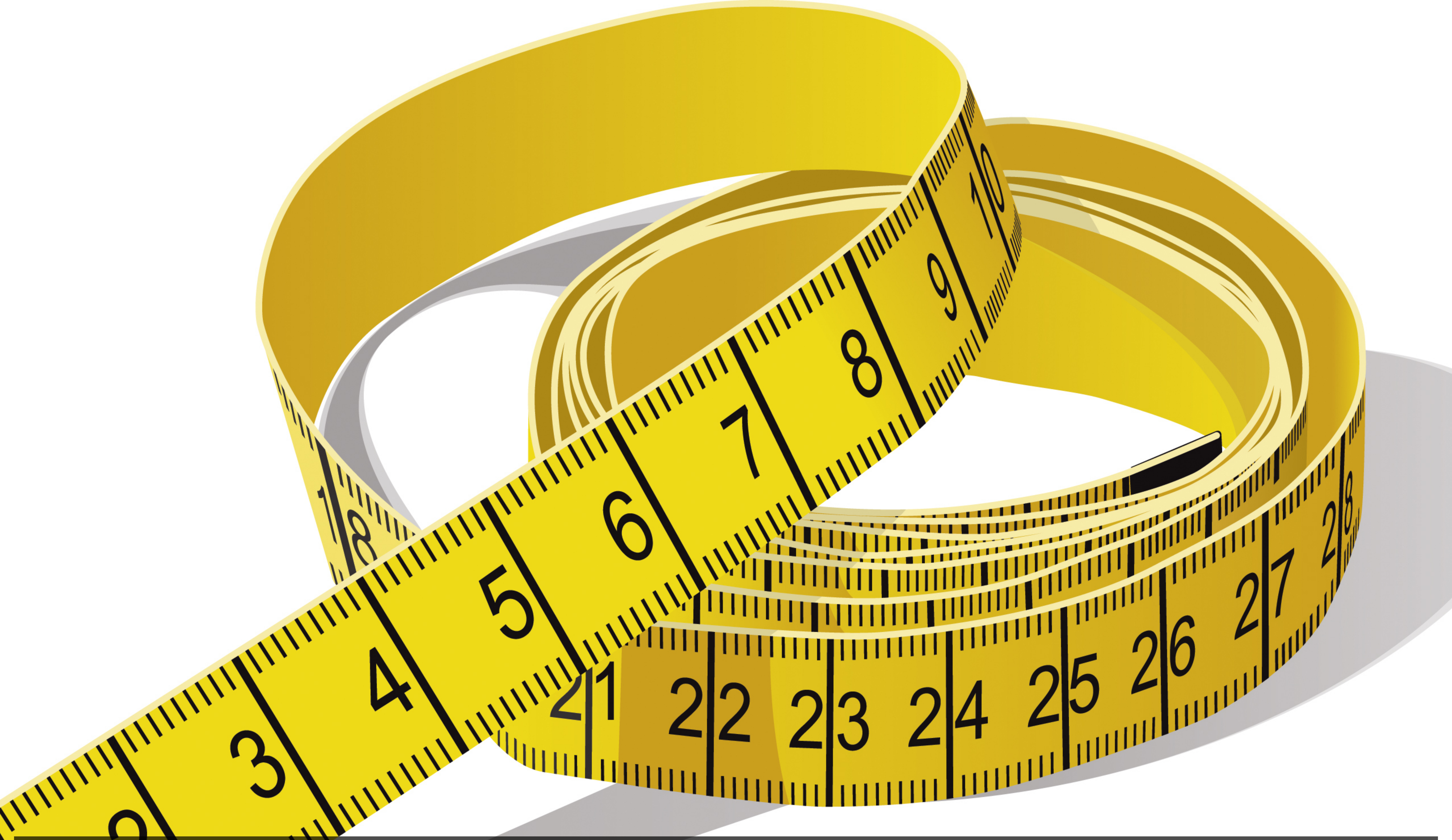
Buildings contribute significantly to overall energy consumption



Buildings getting constructed at rapid rate


Buildings are an attractive target
towards sustainability


Residential buildings can contribute upto
93% of building energy usage



"If you cannot measure it, you cannot improve it" - Kelvin

Sensor deployments have several challenges*

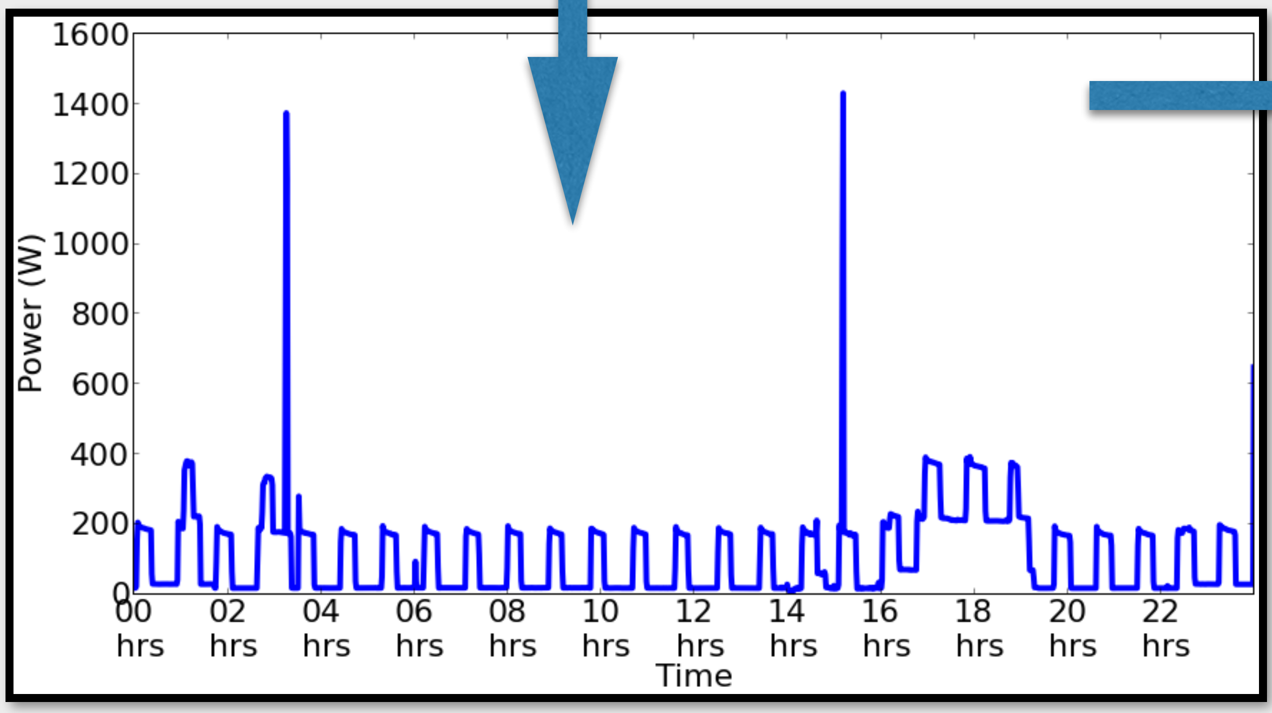
1. Homes are not a power panacea
2. Homes have poor connectivity
3. Homes are hazardous
4. Limited user interaction
5. Aesthetics matter

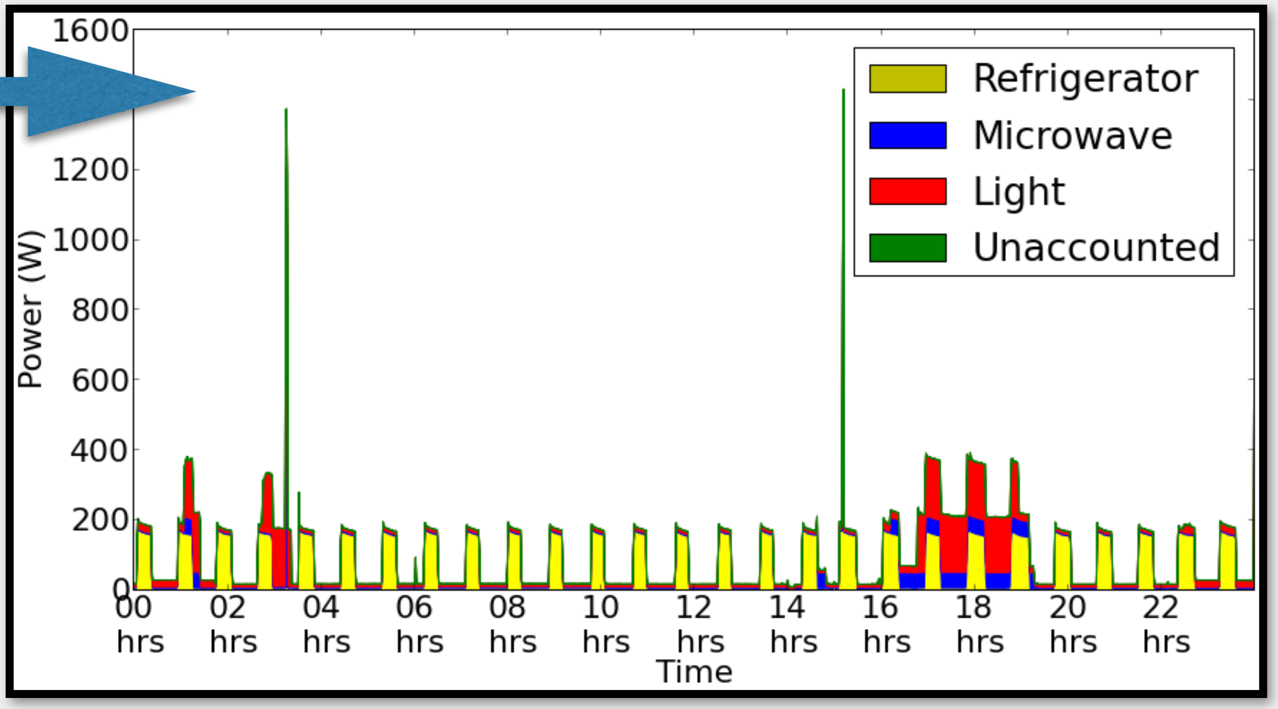
1. Hnat et al. "The hitchhiker's guide to successful residential sensing deployments". Sensys 2010
2. Batra et al. "It's different. Insights into home energy consumption in India". Buildsys 2013

Non **i**ntrusive **l**oad **m**onitoring (NILM) or Energy disaggregation

Smart meter

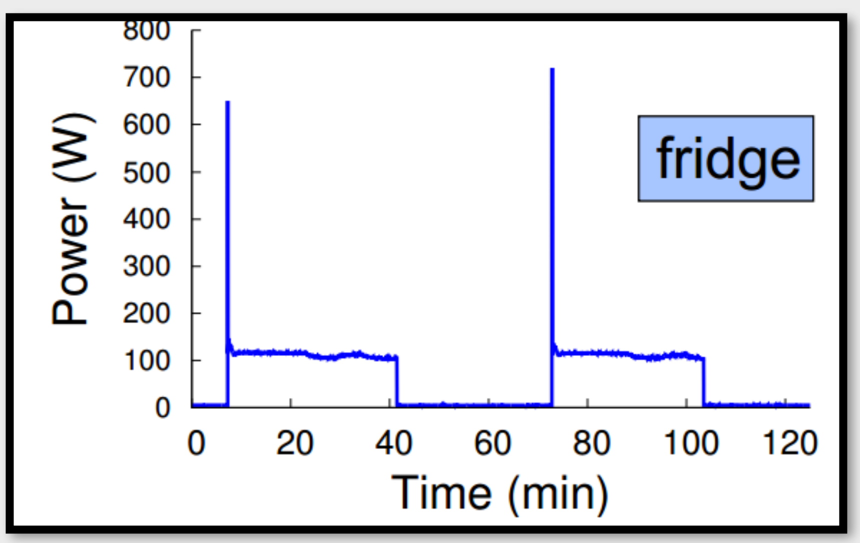
Machine learning

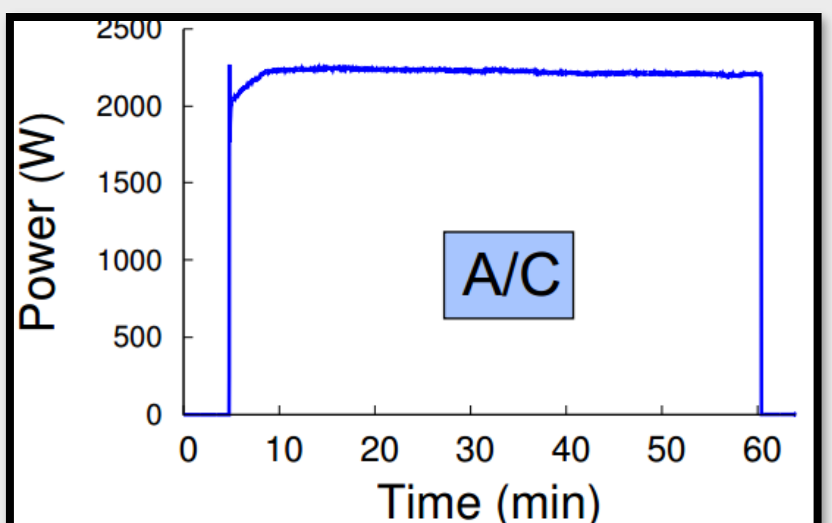


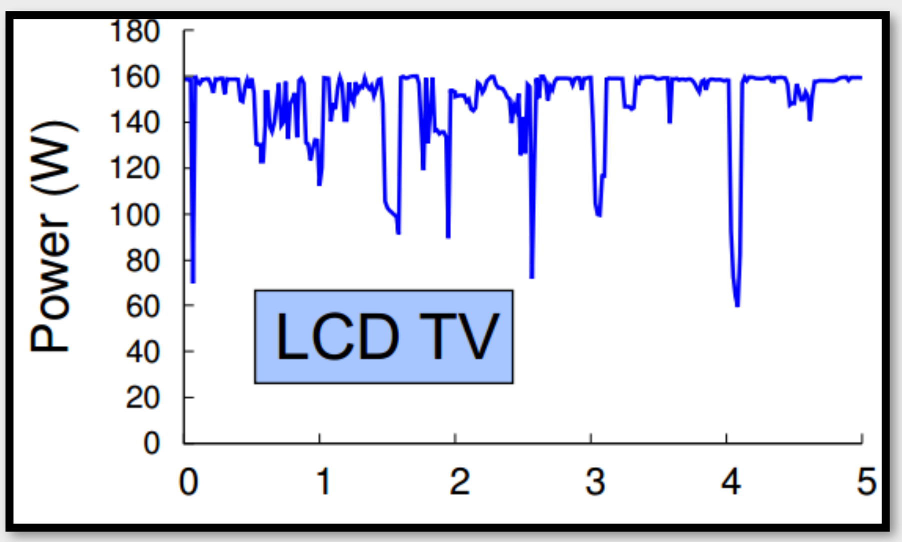


Why NILM can work

Different appliances can have unique "signatures"







Making NILM **practical**

1. Comparable
2. Utility-driven
3. Scalable

Making NILM **comparable**

eEnergy 2014 and Buildsys 2014

What is the best NILM approach?

Despite 30+ years of NILM research really hard question

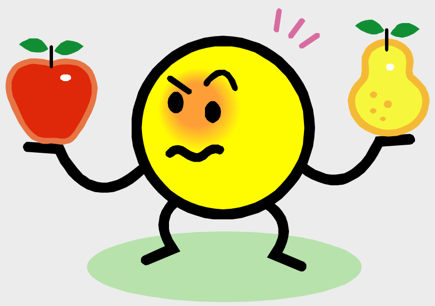
3 main problems

1: Hard to assess generality

- Previous contributions evaluated only on single dataset
- Non-trivial to set up similar experimental conditions for direct comparison

2: Lack of comparison against same benchmarks

- Newly proposed algorithms rarely compared against same benchmarks
- Lack of "open source" reference algorithms often lead to reimplementation



3: "Inconsistent" disaggregation performance metrics

- Different performance metrics proposed in the past
- Different formulae for same metric, eg. 4+ versions of "energy assigned"

$$\mathrm{Acc} = 1 - \frac{\sum_{t=1}^{T} \sum_{i=1}^{n} \left| \hat{y}_t^{(i)} - y_t^{(i)} \right|}{2 \sum_{t=1}^{T} \bar{y}_t}$$

$$\sqrt{\left(\sum_{t,i} \left\| y_t^{(i)} - \hat{y}_t^{(i)} \right\|_2^2 \right) / \left(\sum_{t,i} \left\| y_t^{(i)} \right\|_2^2 \right)}$$

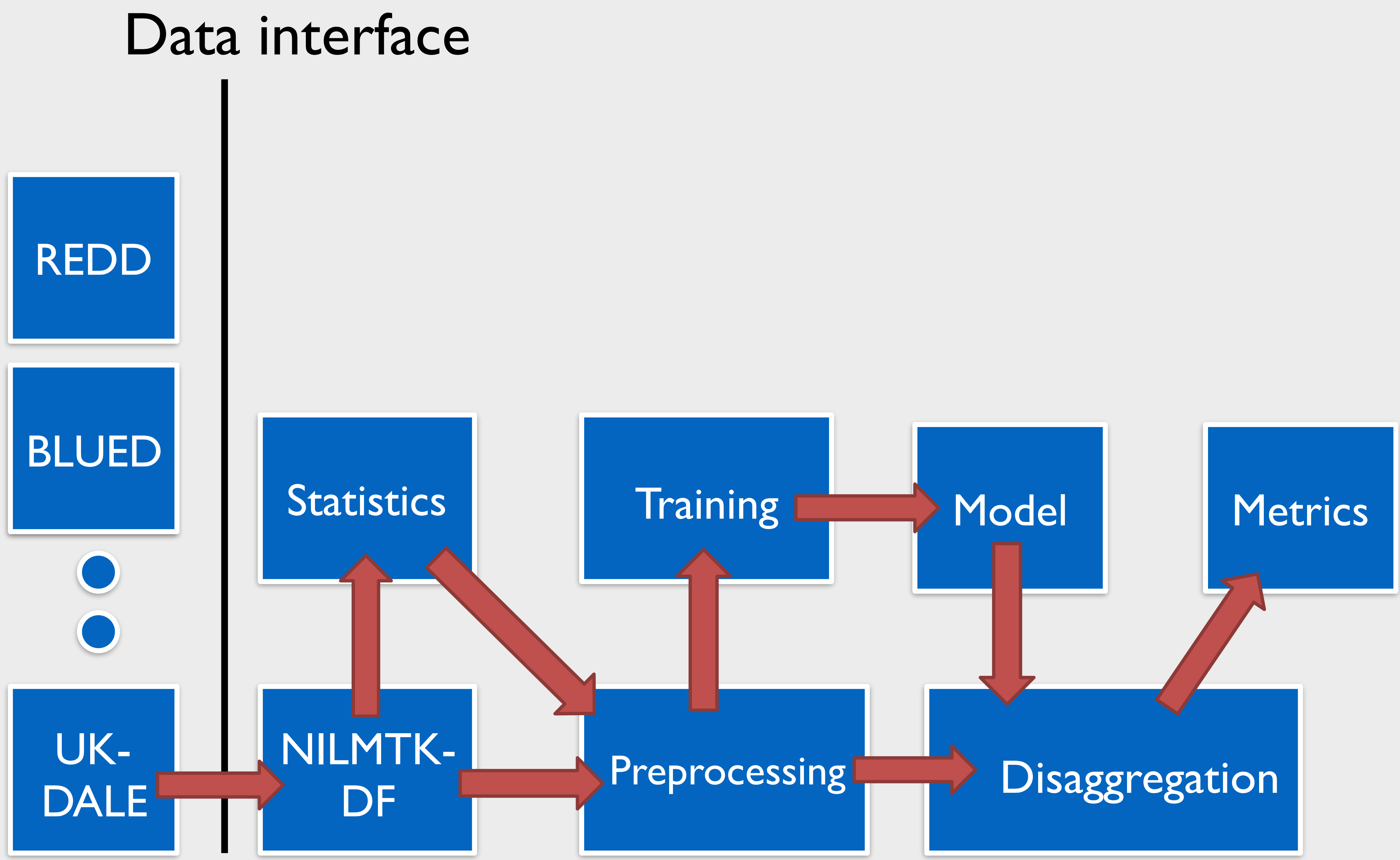
$$\left| \sum_t x_t^{(n)} - \sum_t \mu_{z_t^{(n)}}^{(n)} \right| / \sum_t x_t^{(n)}$$

$$MNE(n) = \frac{\sum_{t=1}^{T} |\theta_t^n - y_t^n|}{\sum_{t=1}^{T} \theta_t^n}$$

And NILMTK was born

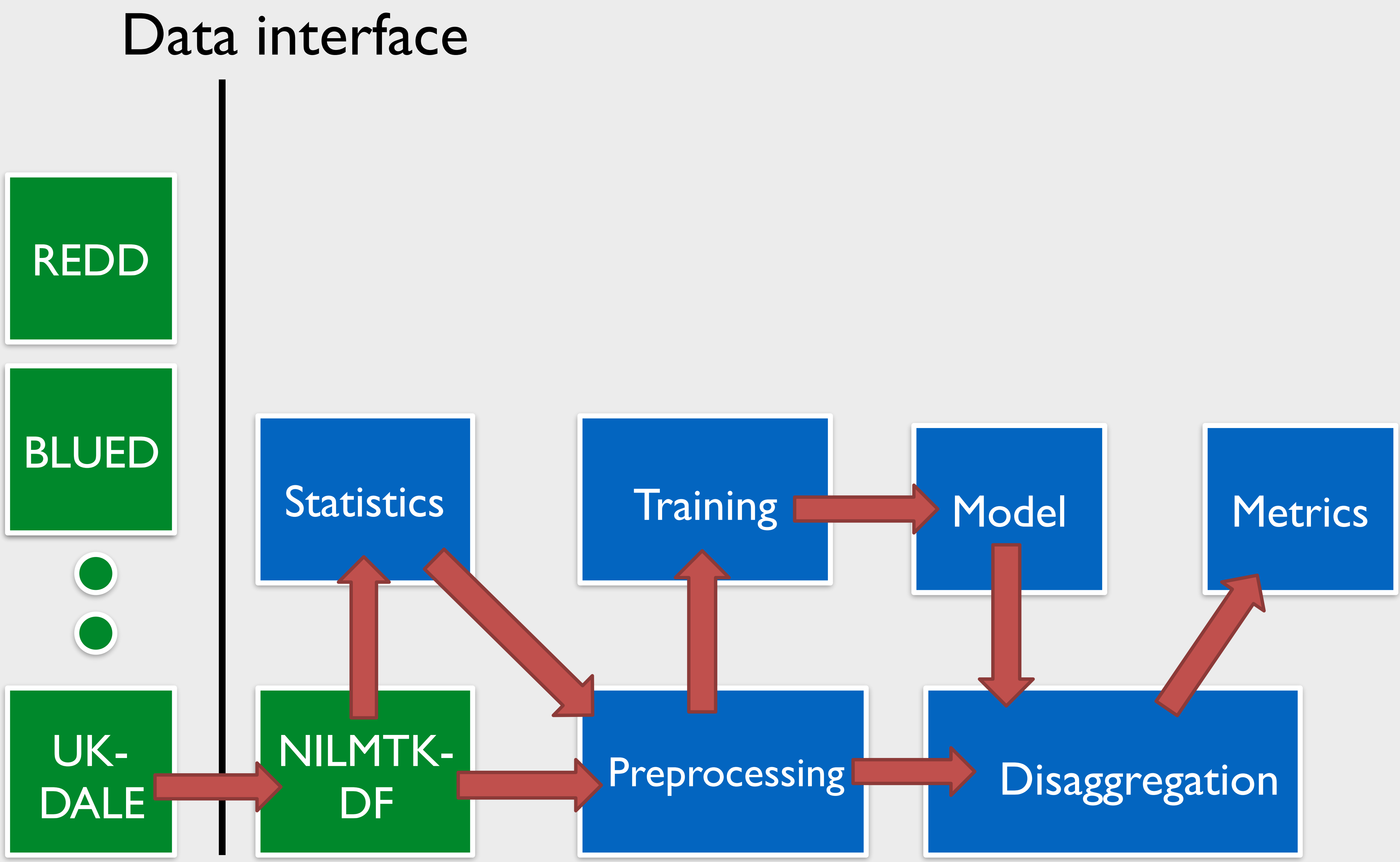
Open source **NILM toolkit** to enable **easy comparative** analysis of NILM algorithms **across data sets**

NILMTK pipeline



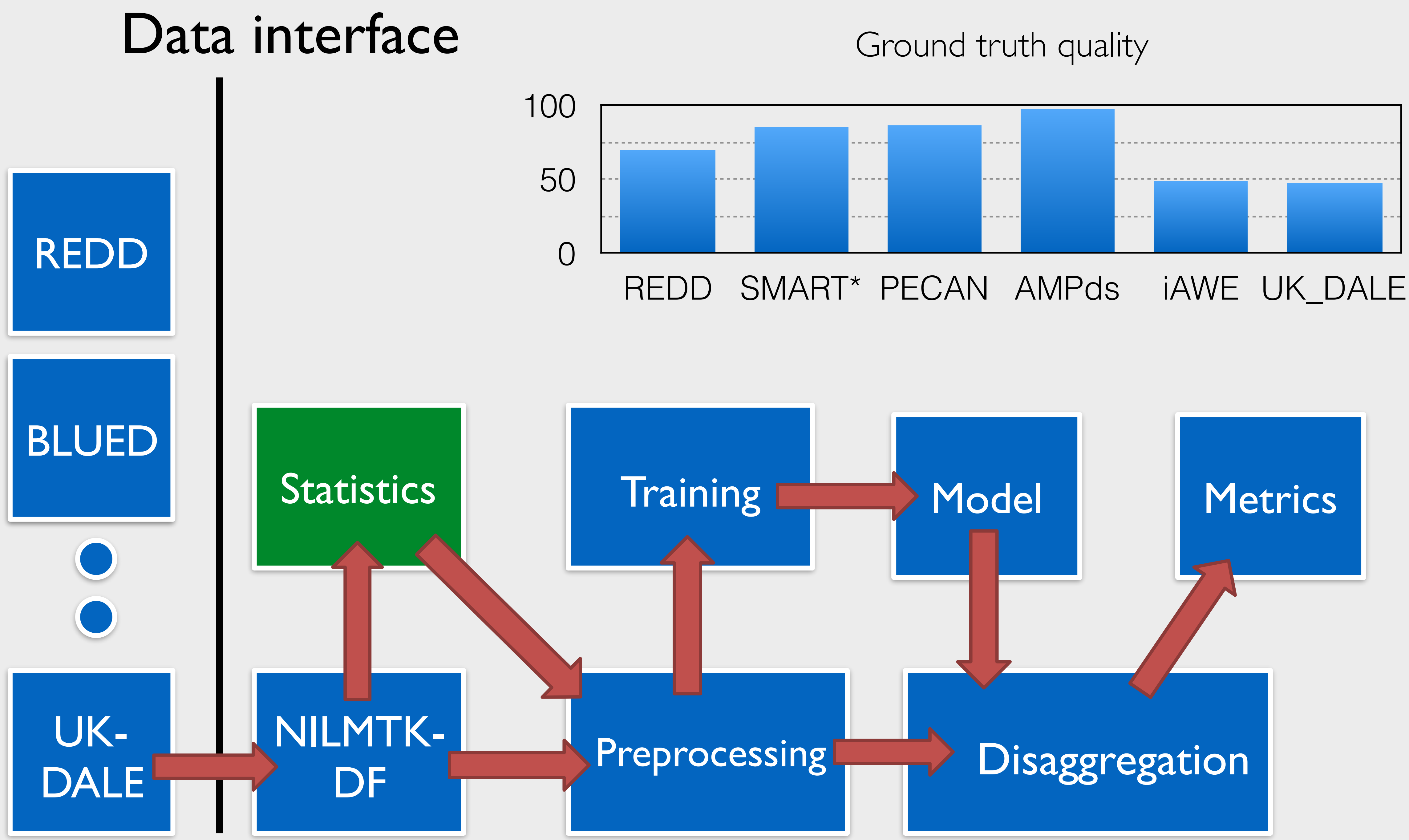
NILMTK-DF: Common data format

10 data sets released

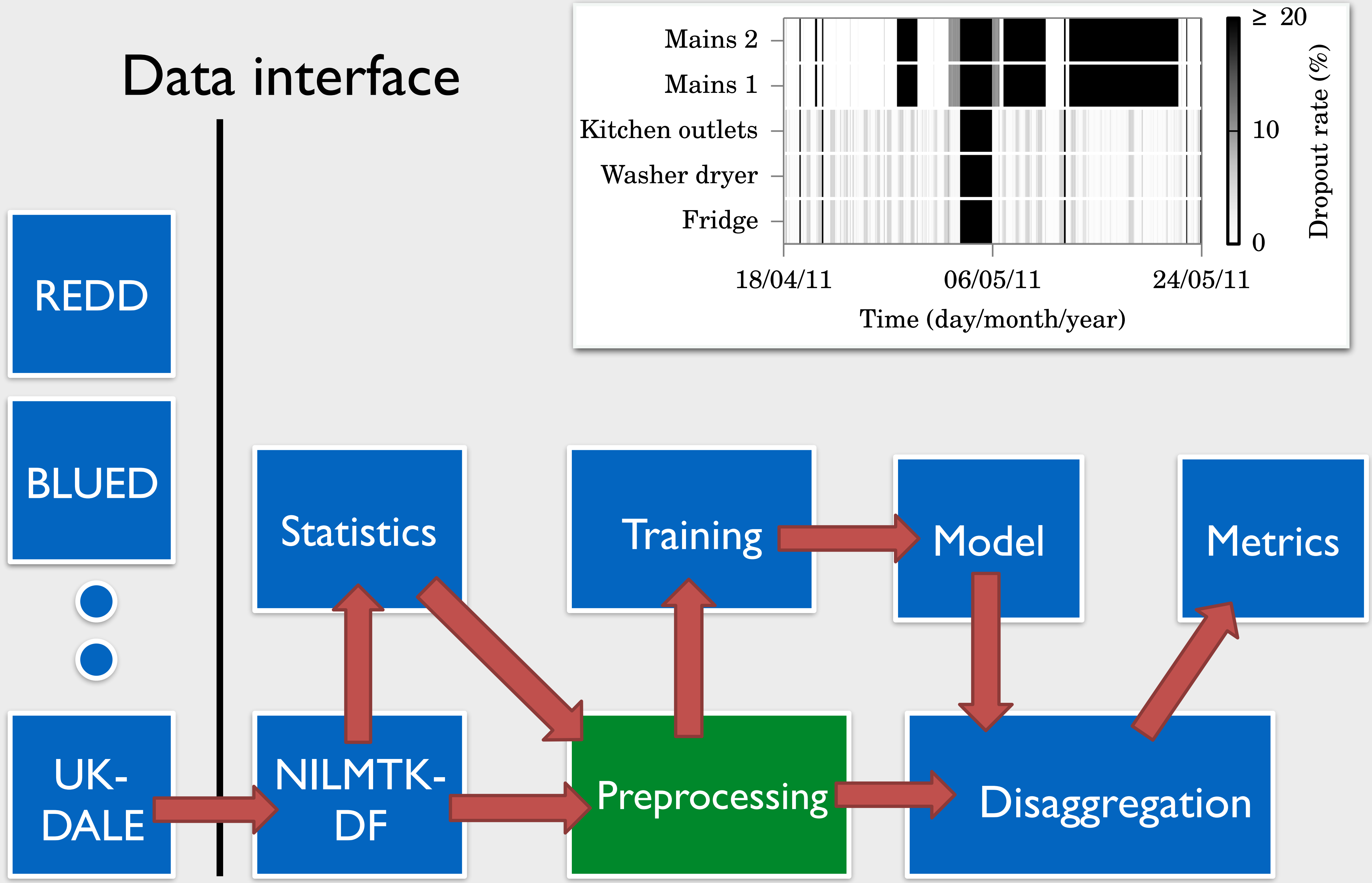


Statistical functions

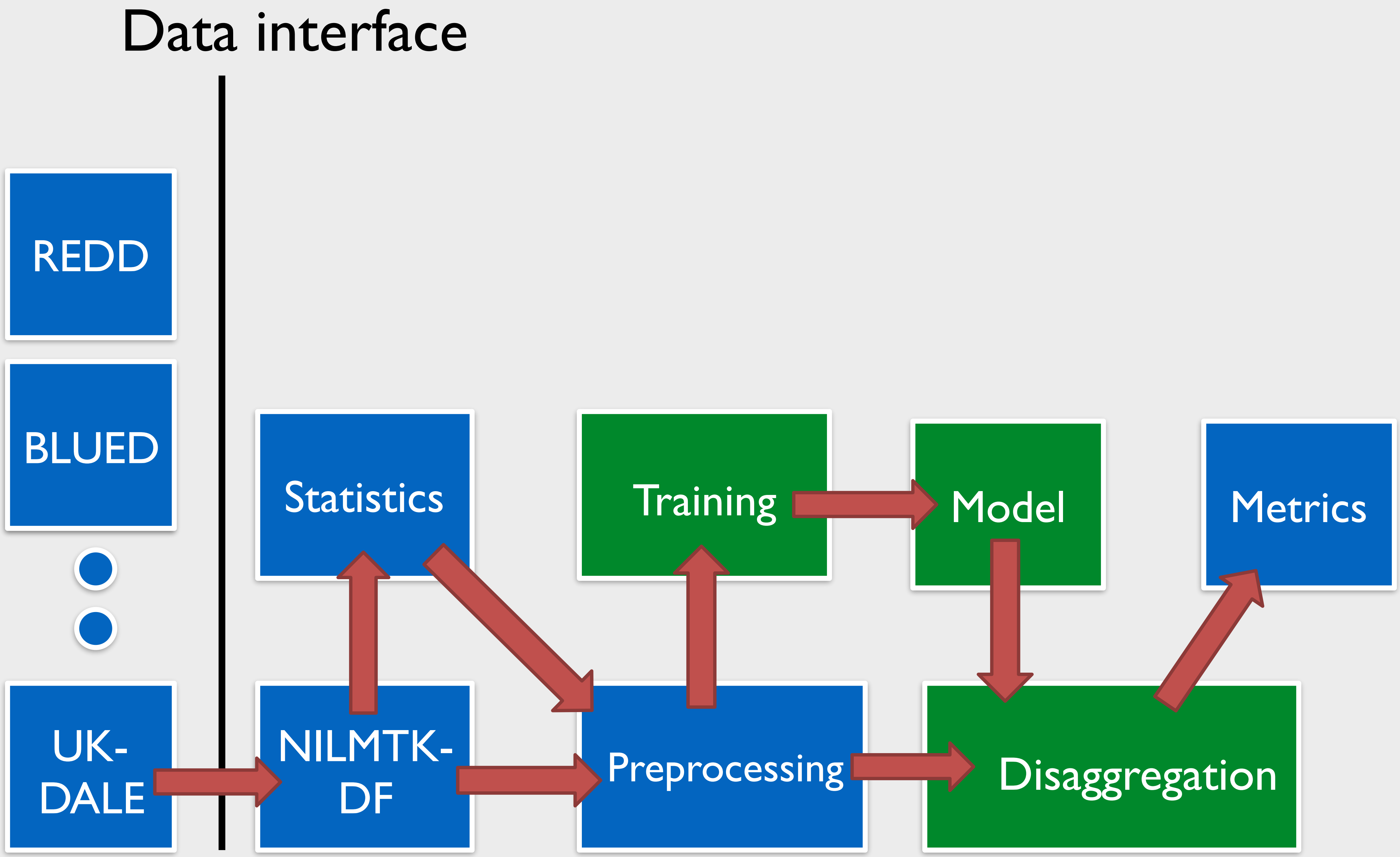
Suite of commonly used statistical functions



Preprocessing

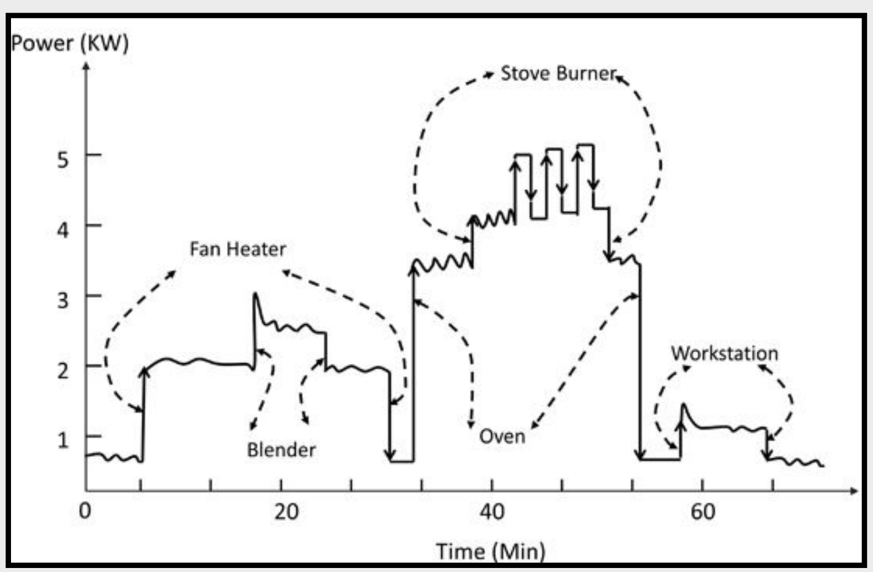


Train and Disaggregate

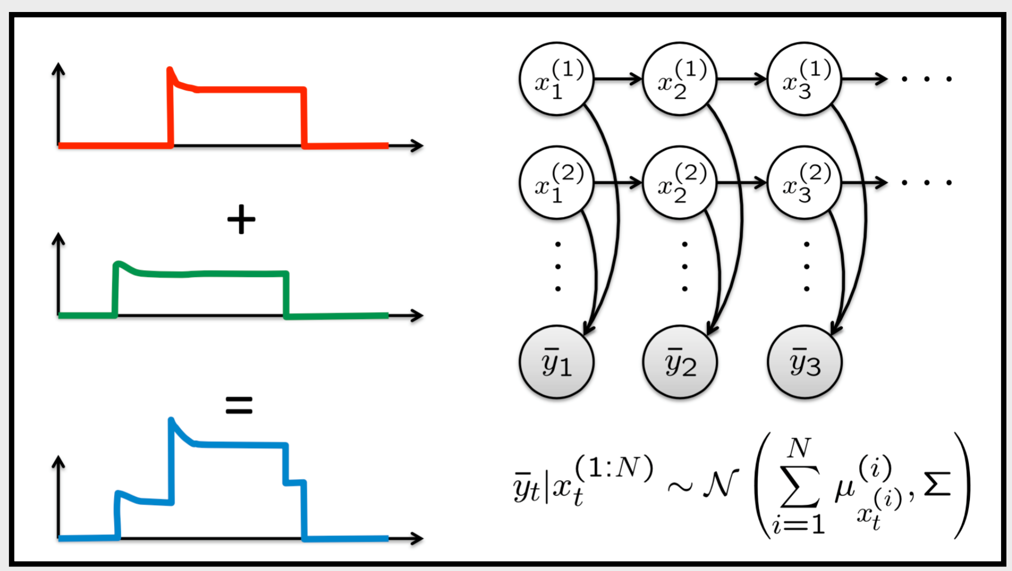


Train and Disaggregate

Hart's event detection algorithm



Factorial Hidden Markov Model (FHMM)



Combinatorial Optimisation

Appliance	Off power	On power
Light	0	200
Fridge	0	100

NILMTK impact

- 10+ papers using NILMTK (4 in Buildsys 2015)
- 2 user contributed NILM algorithms
- 3 user contributed NILM data sets
- Best demonstration award at Buildsys 2014

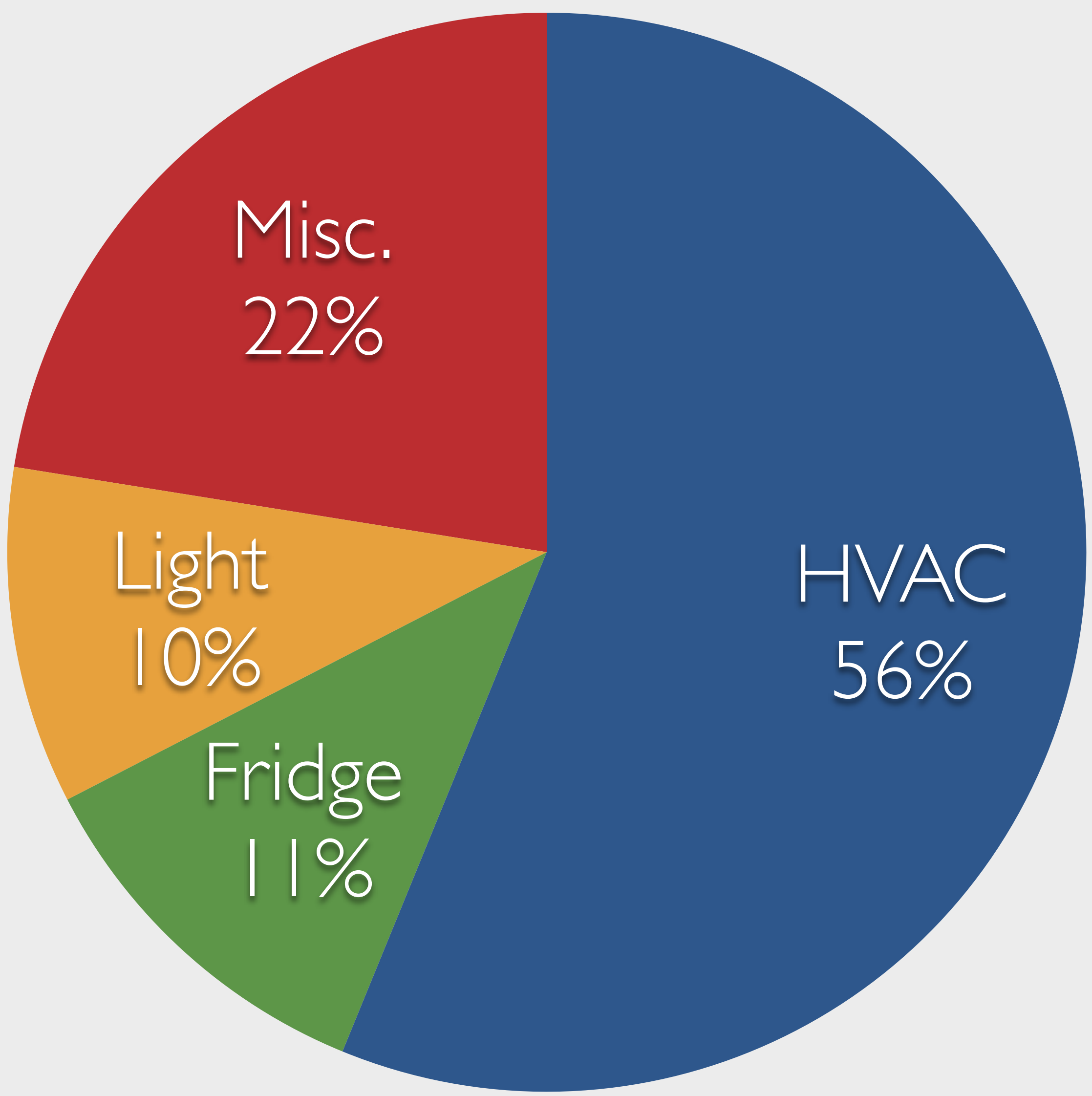
Making NILM **utility-driven**

Buildsys 2015 and Percom 2016 (under submission)

If you can measure, can you improve?

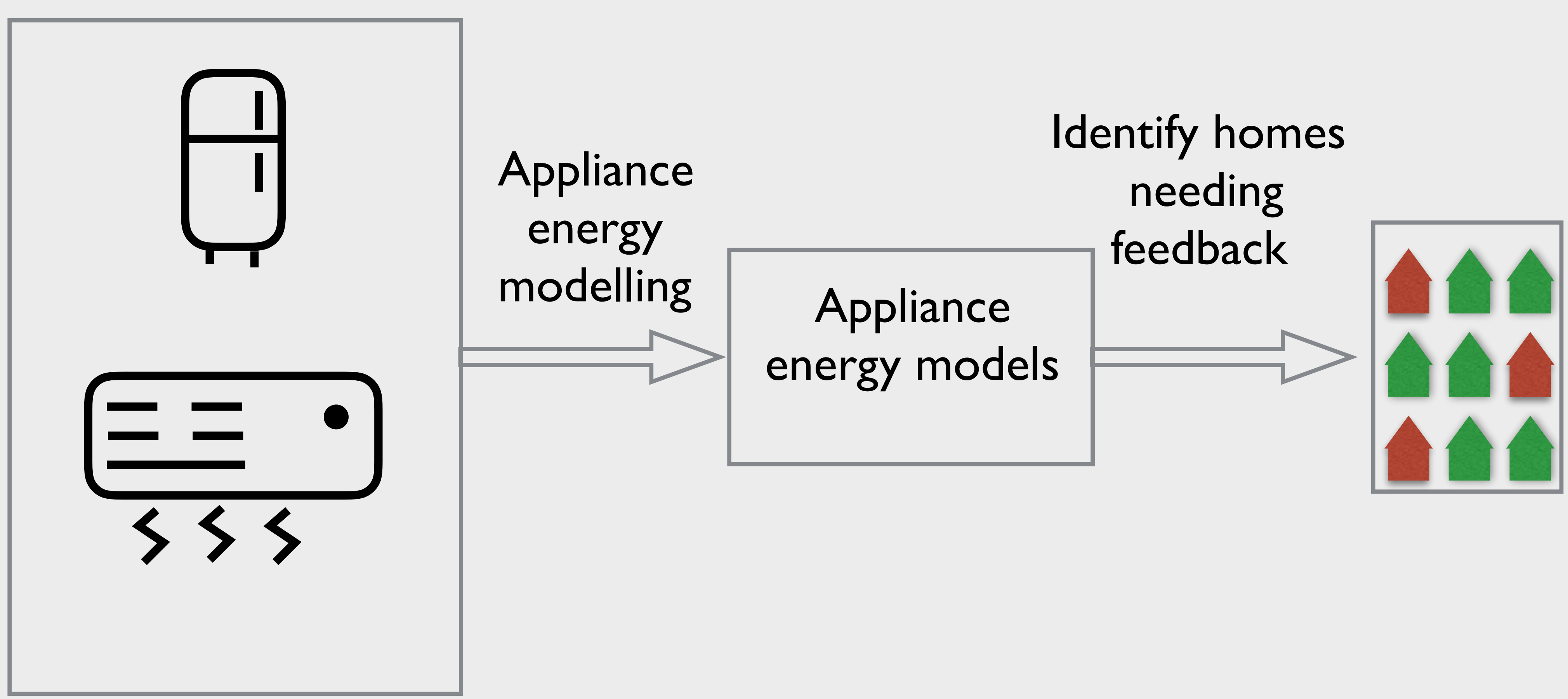
Does NILM **really** save energy?

Does telling you that HVAC takes 56% save you energy?
Lack of specific **actionable insights**



Exploring the value of Energy disaggregation

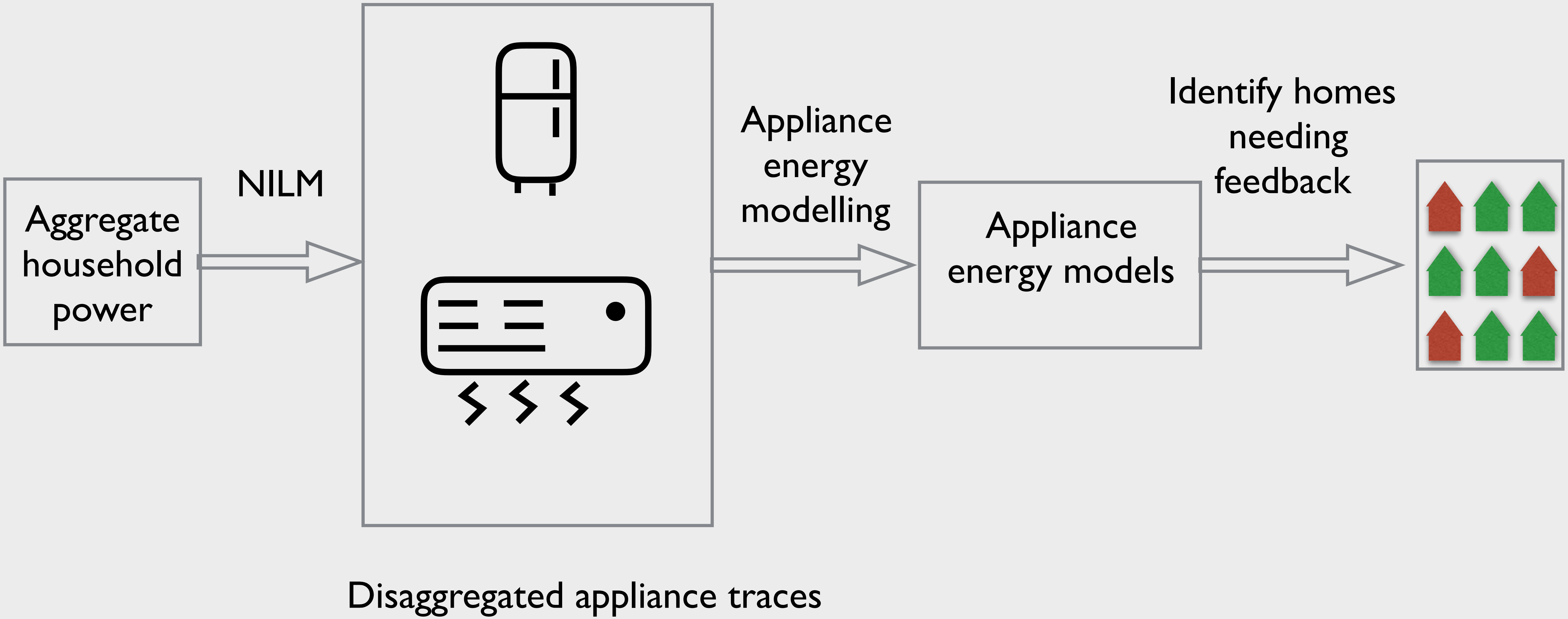
1. Can disaggregated traces provide actionable insights?



Submetered appliance traces

Exploring the value of Energy disaggregation

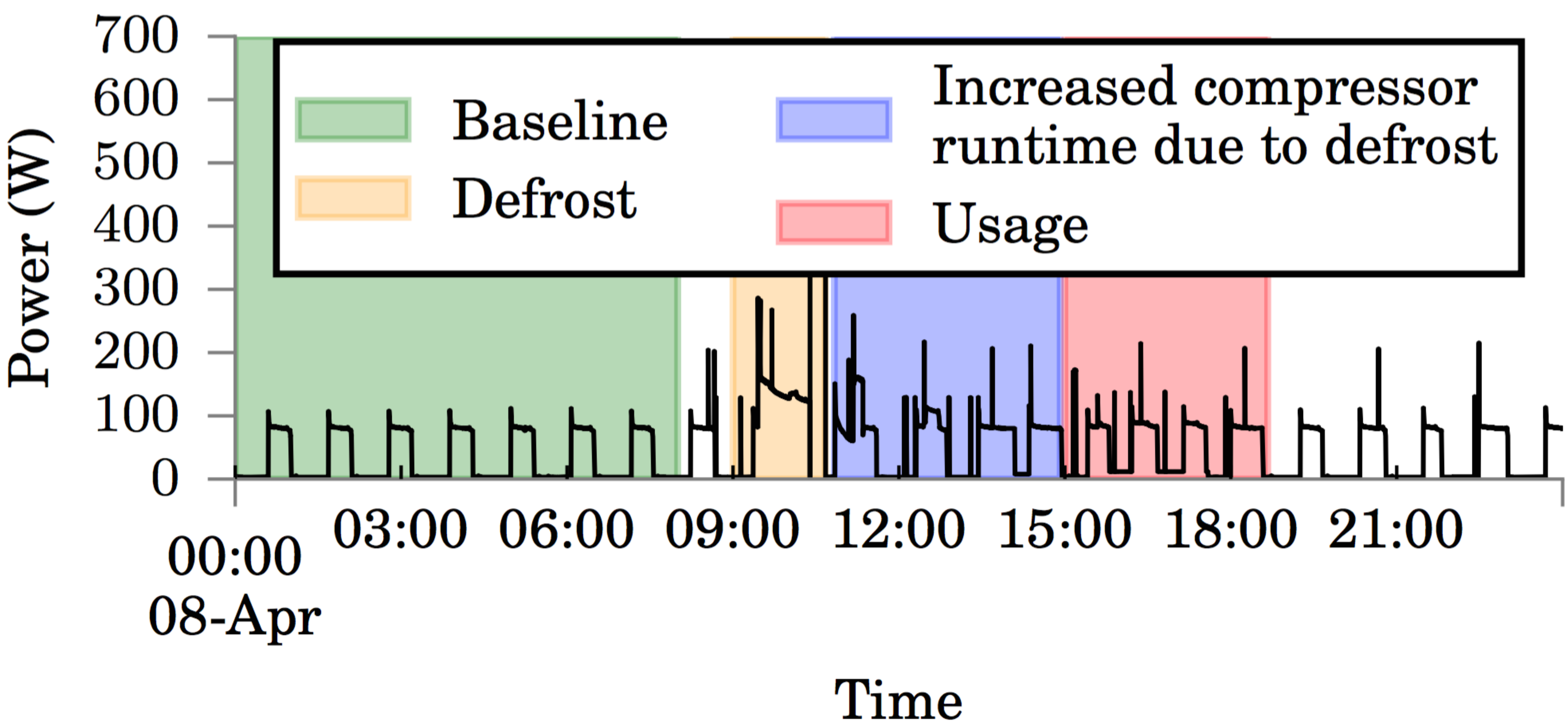
2. Do existing NILM techniques provide traces with sufficient fidelity to support feedback?



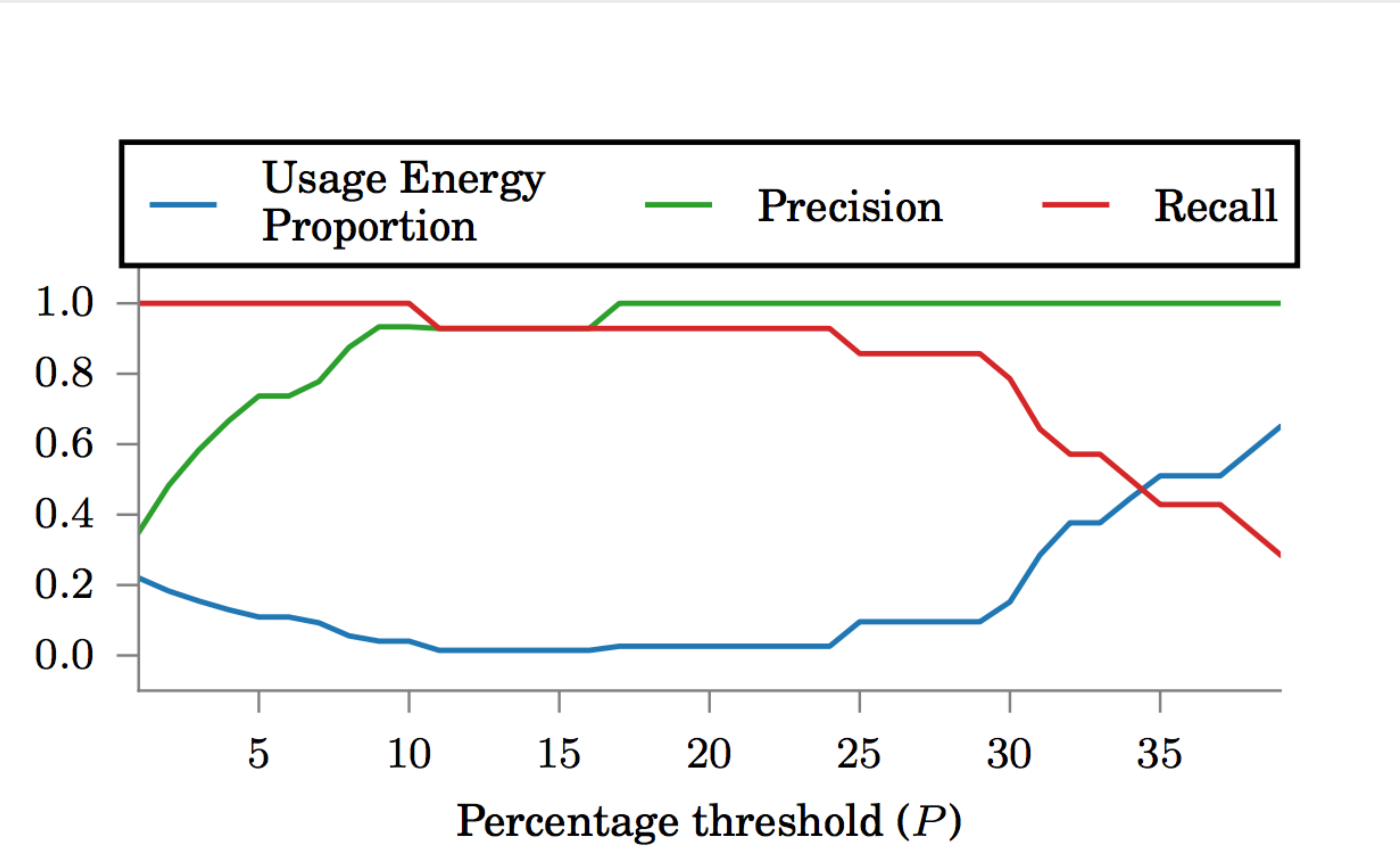
Feedback methods on Fridge and HVAC

- Both appliances common across homes
- Both appliances contribute heavily to overall energy consumption

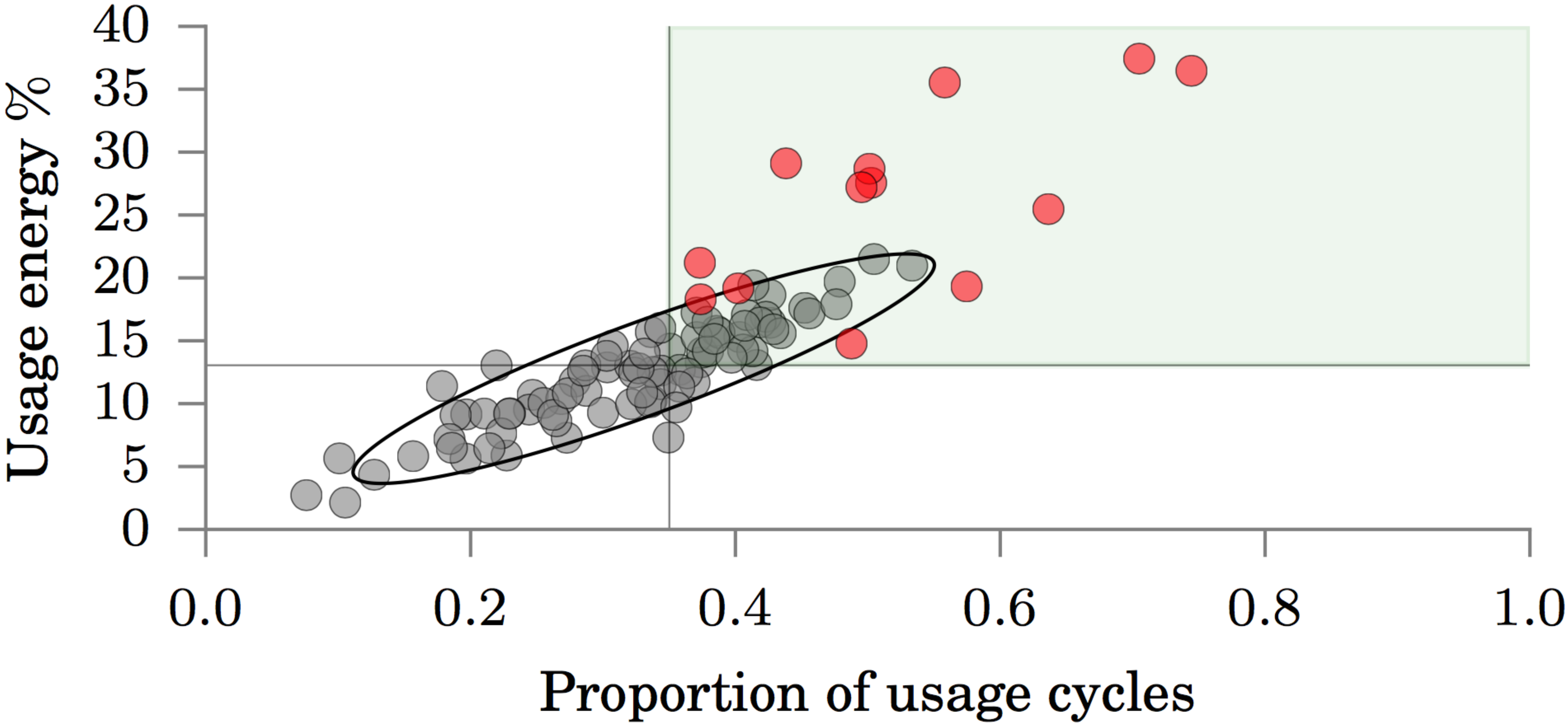
Fridge energy modelling



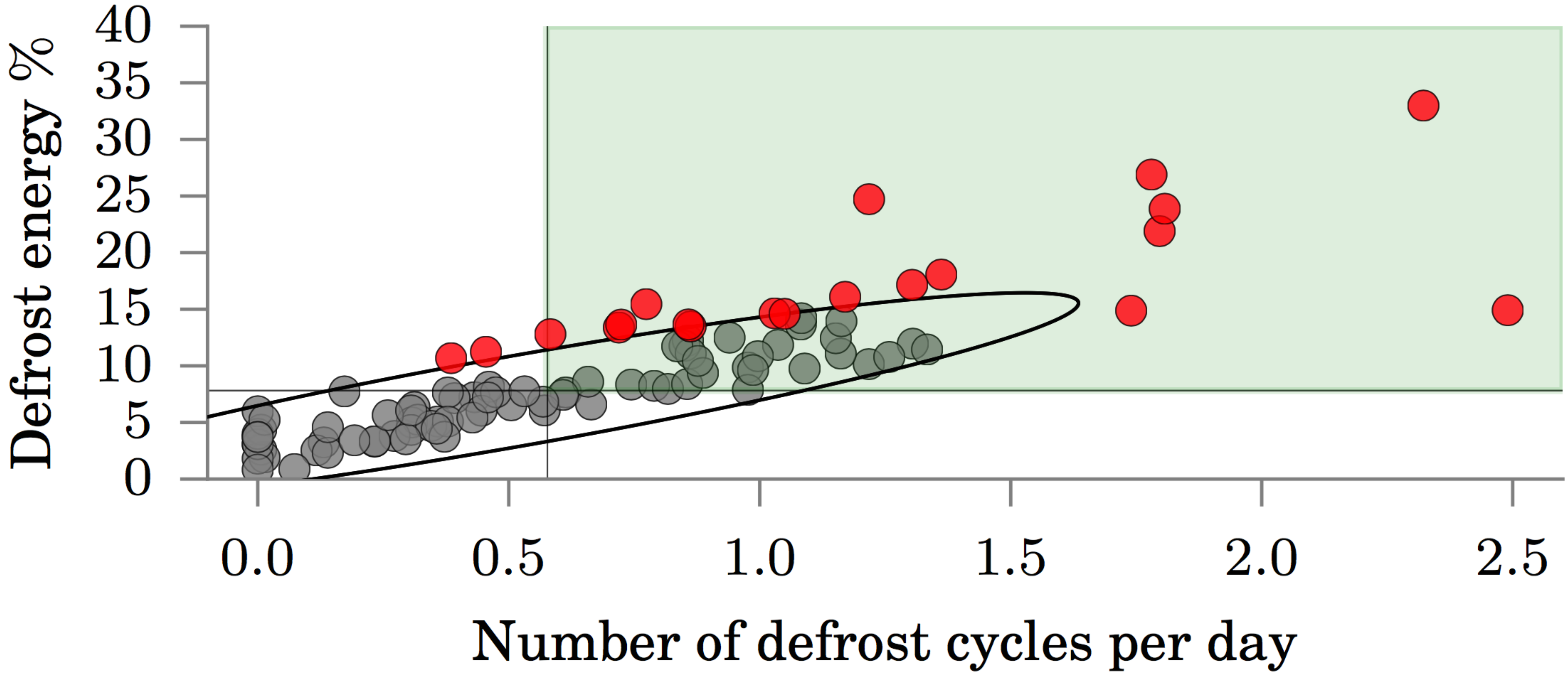
We can break down fridge energy with less than 4% error



13 out of 95 homes can be given feedback based on **usage energy** saving upto 23% fridge energy



17 out of 95 homes can be given feedback on **excess defrost** saving upto 25% fridge energy



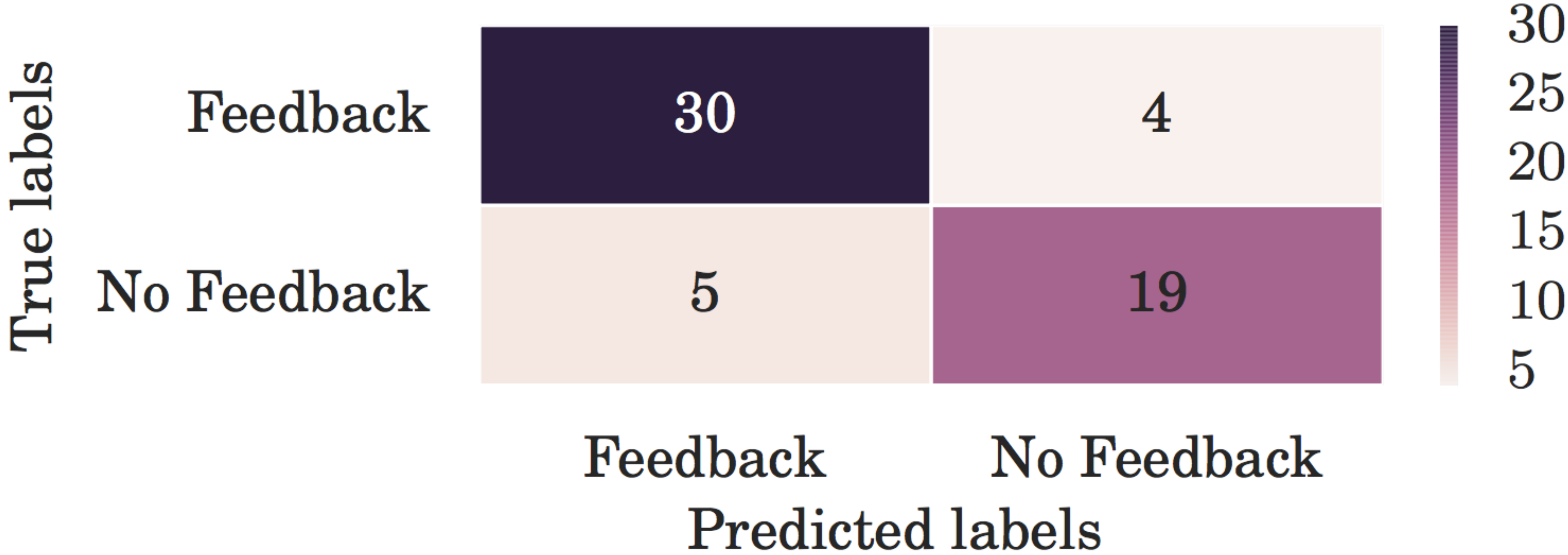
HVAC modelling

Objective

1. Learn set point from weather and energy data
2. Optimising setpoint can save upto 20-30% HVAC

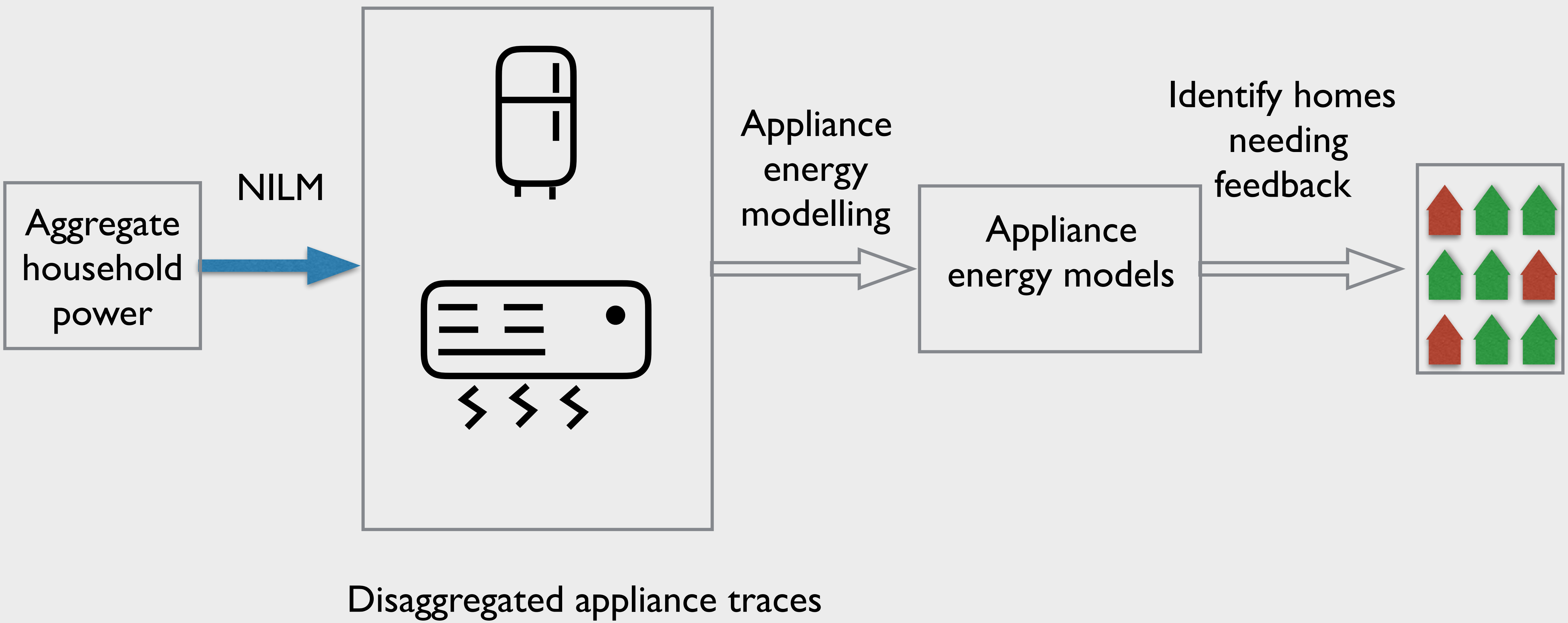
HVAC feedback

- 84% accuracy on giving feedback based on setpoint temperature



Exploring the value of Energy disaggregation

2. Do existing NILM techniques provide traces with sufficient fidelity to support feedback?

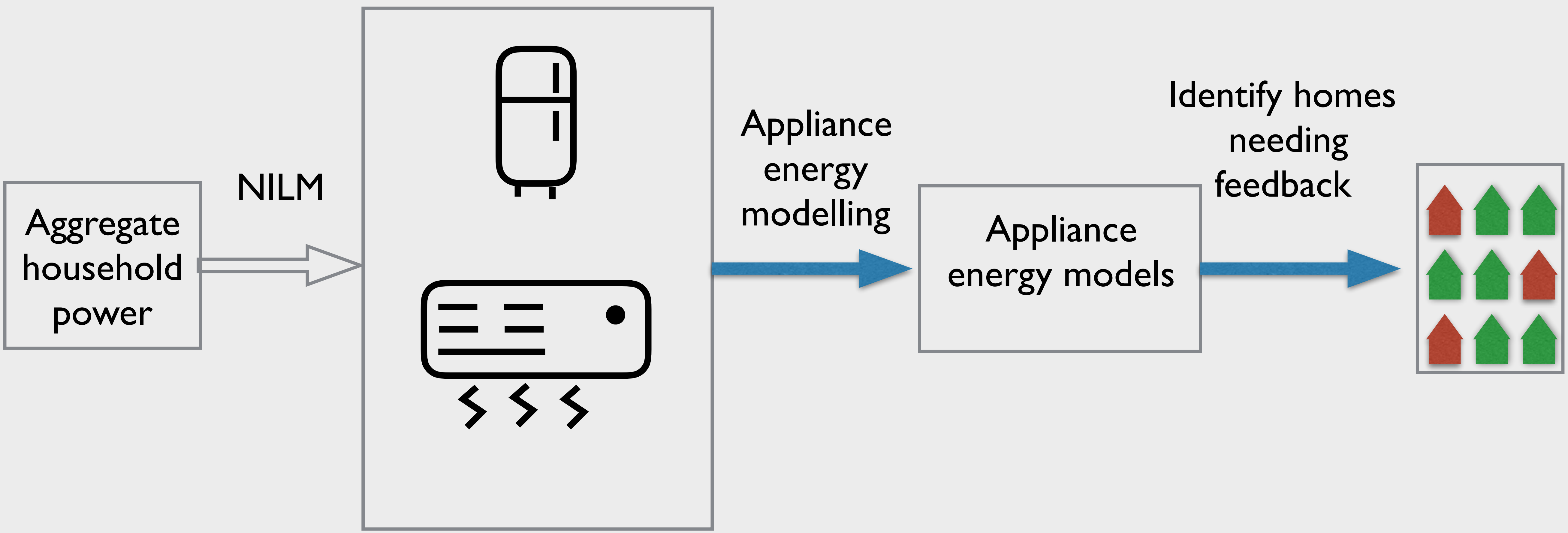


Benchmark NILM algorithms on our data set give accuracy comparable to state-of-the-art

Authors	Year	Dataset	#Homes	Algorithm	Fridge			HVAC		
					RMSE (W)	Error Energy %	F-score	RMSE (W)	Error Energy%	F-score
Kolter [15]	2012	REDD	6	Additive FHMM	-	62.5 △	-	-	-	-
Parson [18]	2012	REDD	6	Difference HMM	83	55	-	-	-	-
Parson [19]	2014	Colden$^{\Psi}$	117	Bayesian HMM		45				
Batra [5]	2014	iAWE	1	FHMM	-	50	**0.8**	-	30	**0.9**
Current work		Data port	240	CO*	85	**19**	0.65	**600**	**15**	0.87
Current work		Data port	240	FHMM*	95	20	0.63	650	18	0.89
Current work		Data port	240	Hart	**82**	21	0.72	890	23	0.76

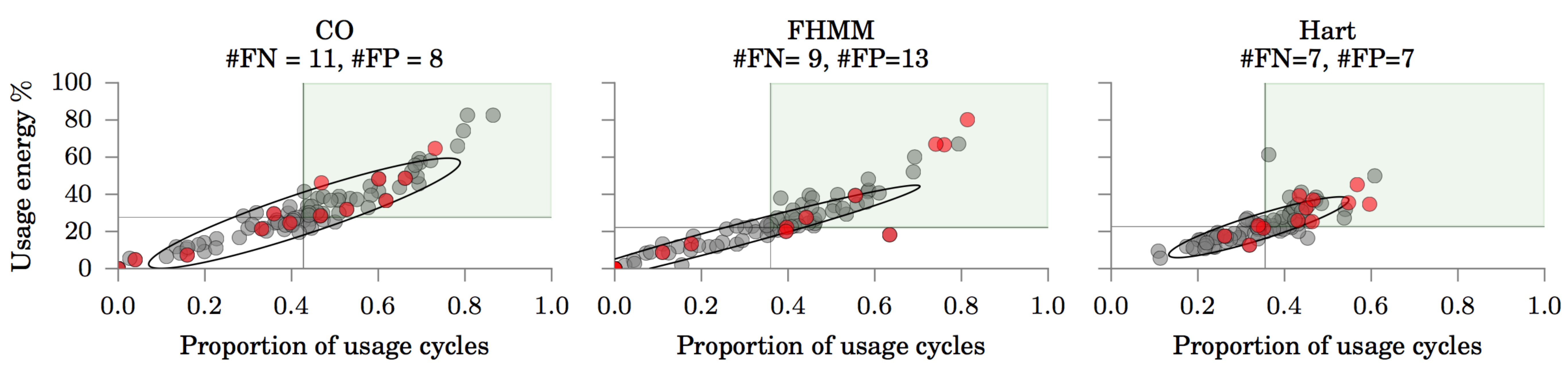
Exploring the value of Energy disaggregation

2. Do existing NILM techniques provide traces with sufficient fidelity to support feedback?



Disaggregated appliance traces

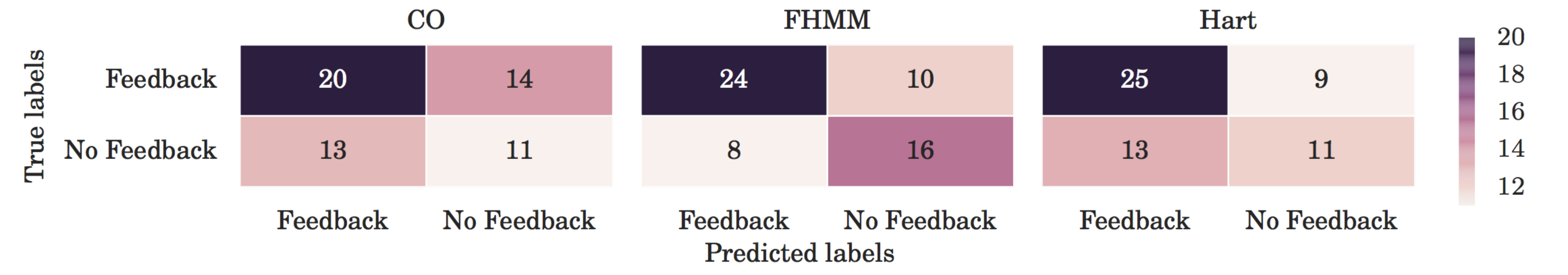
NILM algorithms show poor accuracy in identifying homes which can be given feedback based on **usage energy**



NILM algorithms don't identify the defrost state and thus prevent feedback based on defrost energy

Defrost state is hard to detect!

NILM algorithms show poor accuracy in identifying homes needing HVAC setpoint feedback



Take aways

1. Appliance level data **does** enable actionable energy saving feedback
2. **Feedback** accuracy can be **low** despite **good disaggregation** accuracy
3. We, the disaggregation community, need to **revisit the metrics** by which we measure progress

Making NILM **scalable**

IPSN 2016 (under submission)

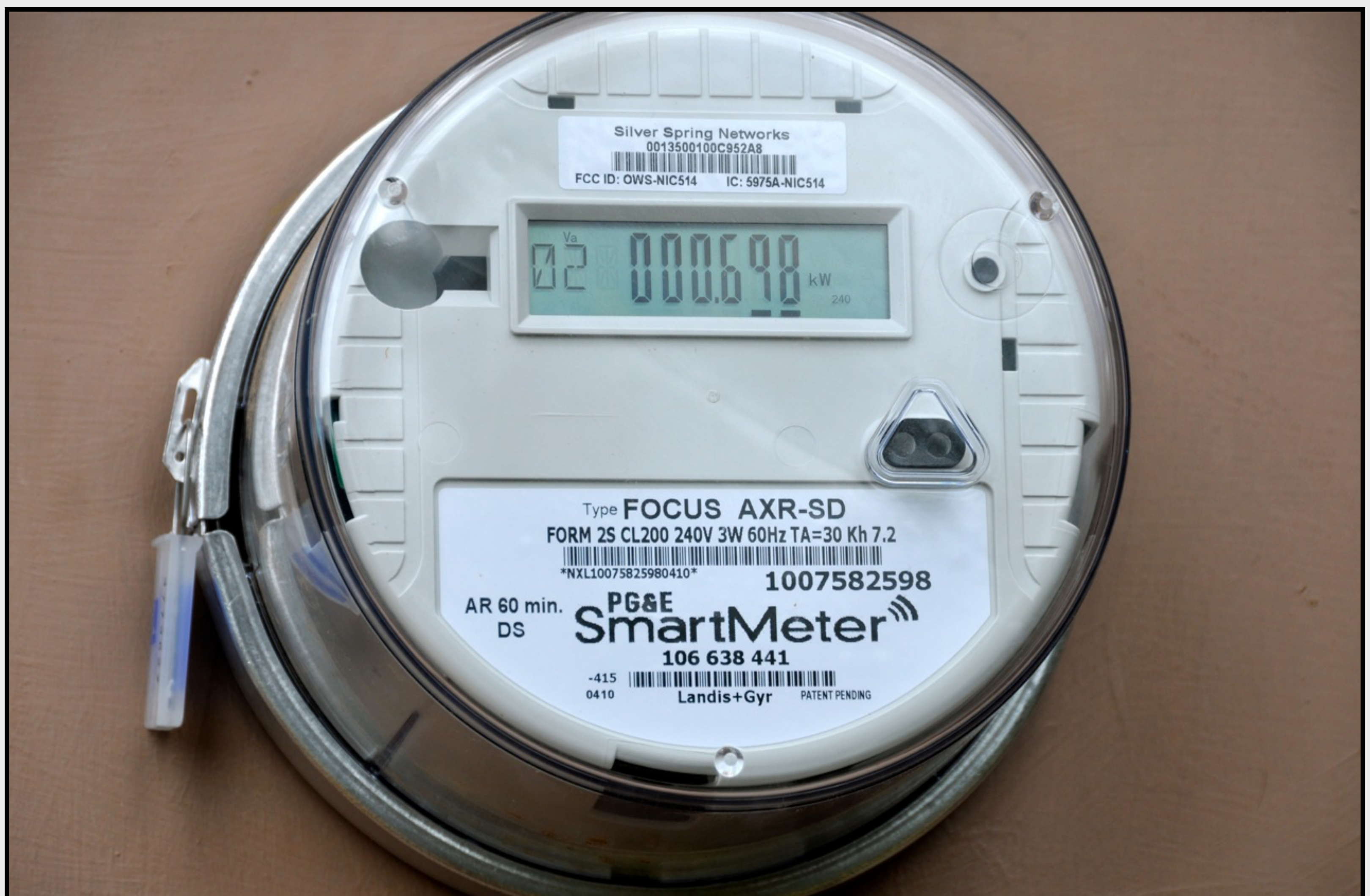
3 fundamental problems

1. Lights (and other low power appliances) show poor disaggregation accuracy. Light are third highest overall in terms of loads
2. Current NILM algorithms are often supervised and need careful tuning and model specification.
3. Most techniques assume 1 min. or less sampling interval. Existing smart meters sample once every 15 mins.

Can we leverage **big data** to make NILM more scalable?

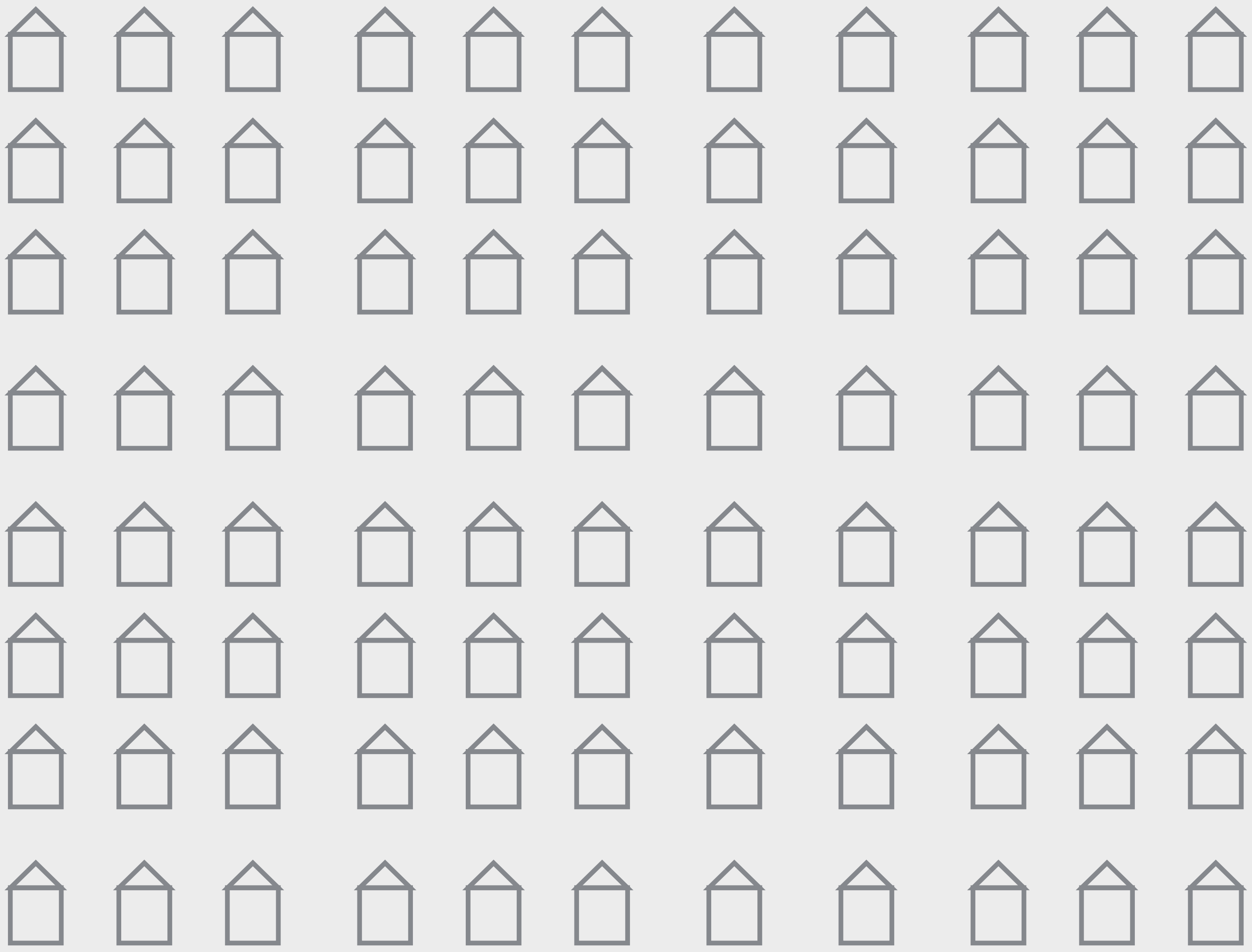
Is **big data** more valuable than **precise** data?

precise data



- Smart meter
- 1 min sampling
- Fine tune model per home

big data

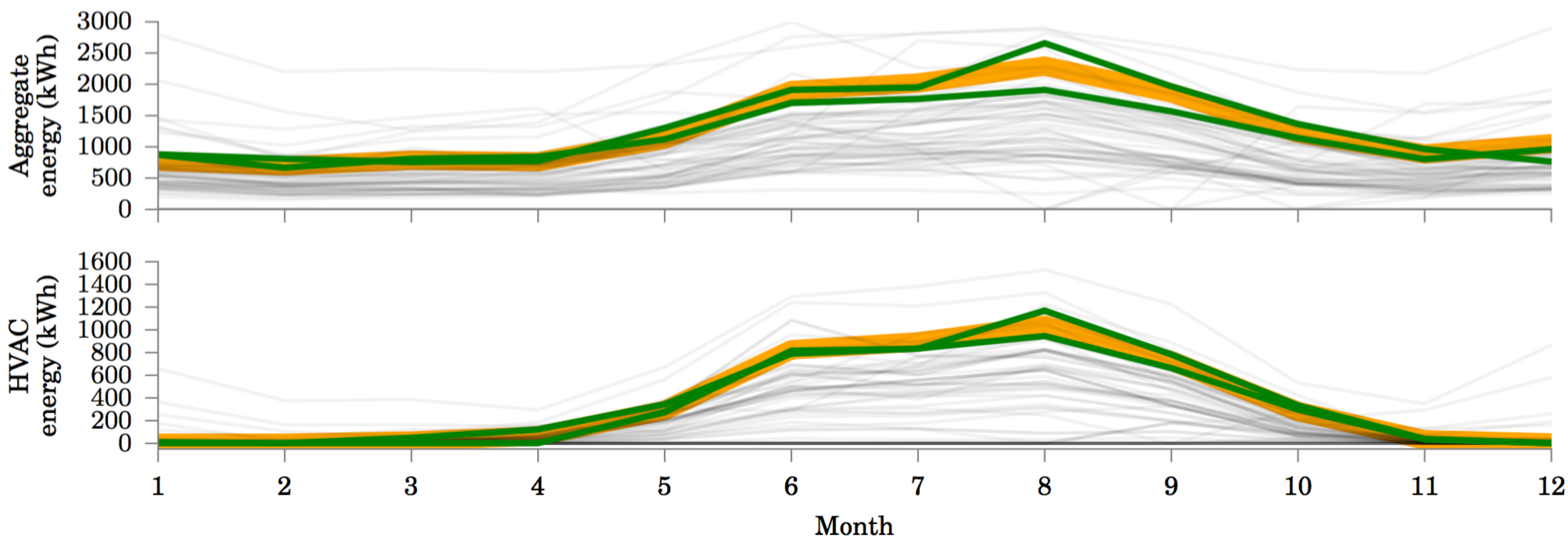


- Large number of homes

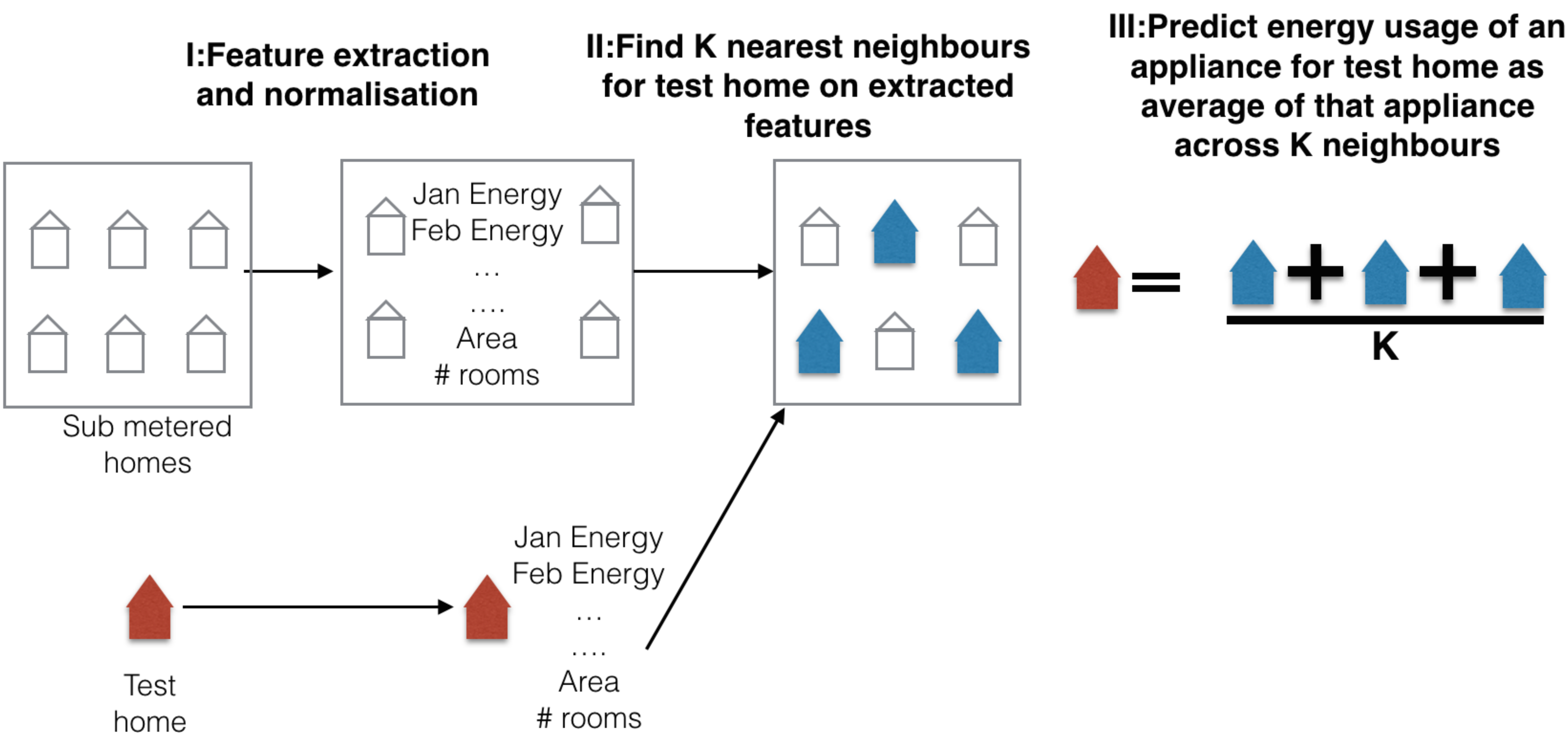
big data

- Large number of homes
- Submeter small subset of homes
- Use **single** reading per month

Similar homes have **similar** per-appliance energy consumption



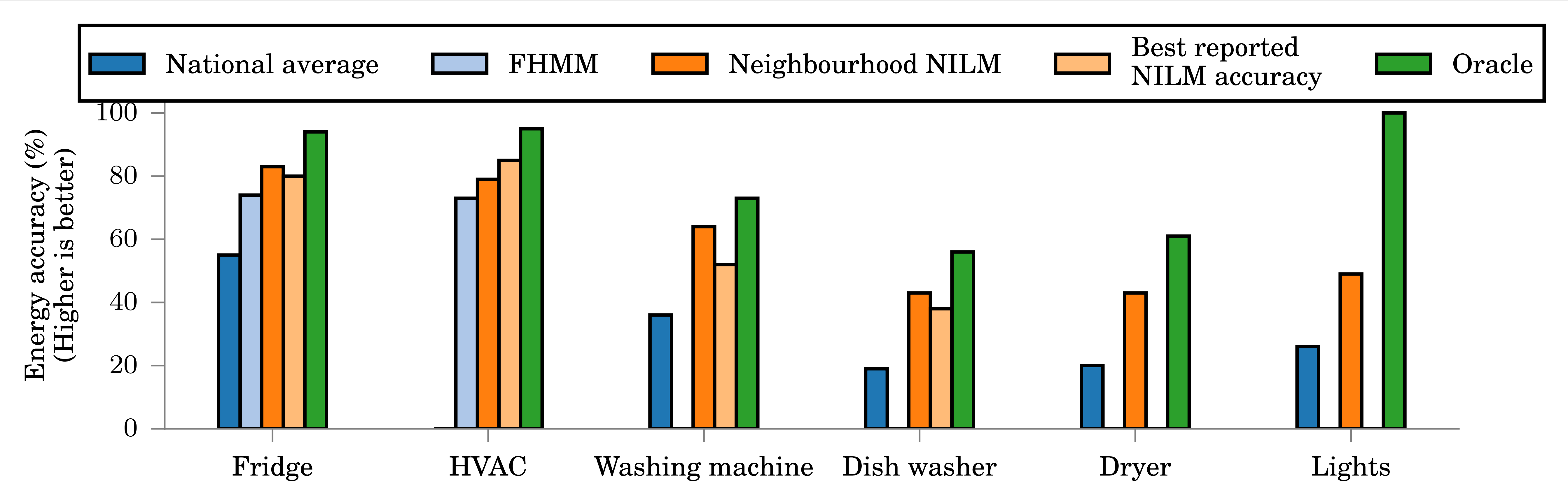
Approach: Neighbourhood NILM



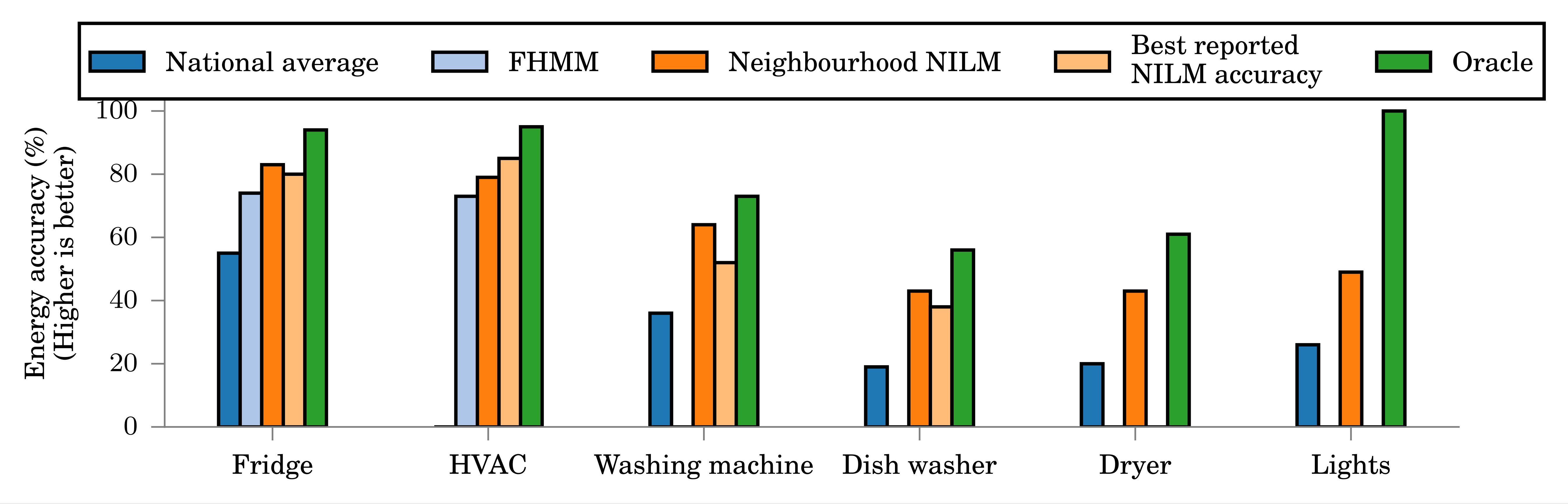
Features

- Energy consumption:
  - Past 12 months household aggregate
  - Ratios (Min. energy/Max. energy)
- Static household properties: #occupants, Area, #rooms

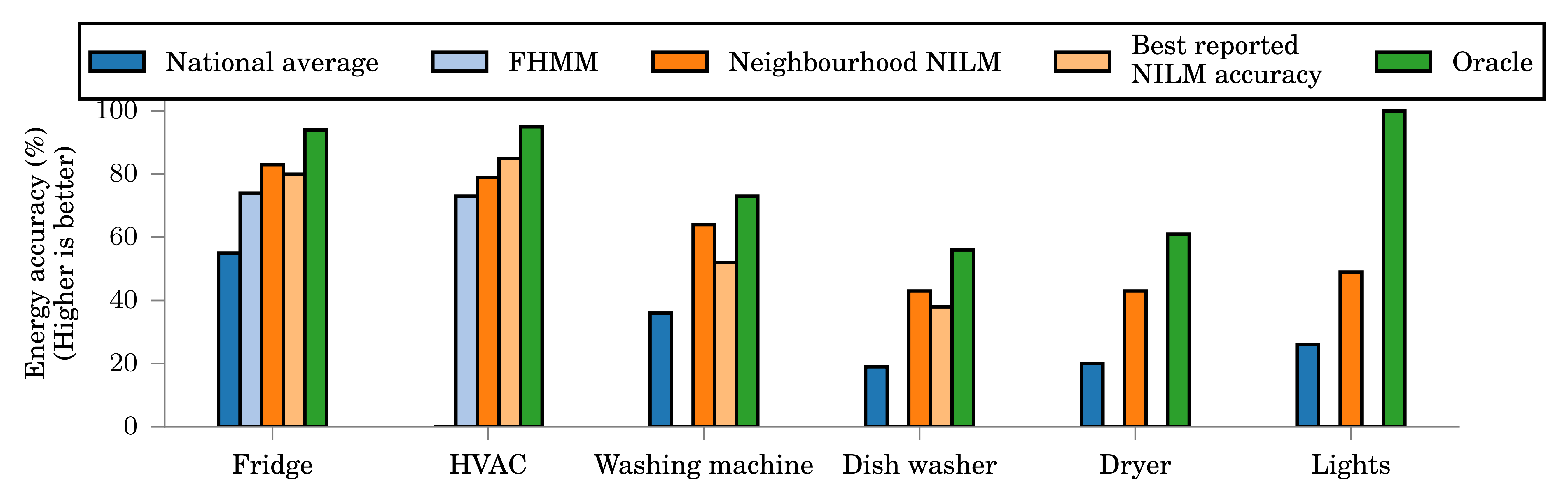
Neighbourhood NILM comparable or better than best reported NILM accuracy



Neighbourhood NILM significantly accurate in Washing machine, dish washer, dryer- all pain points for traditional NILM



High accuracy of "Oracle" suggests that our approach is promising



Take away

Big data more valuable than **precise data** for the problem of energy disaggregation

Conclusions

Making NILM practical in 3 ways:

1. **Comparable-** NILMTK
2. **Utility-driven-**Energy saving feedback
   Inferring household characteristics
3. **Scalable-** Neighbourhood NILM

Future work

Neighbourhood NILM with 15 minute meter data

1. Can we reduce the number of neighbours needed when we use 15 minute meter data
2. 15 minute data will present daily patterns, in addition to monthly patterns in current implementation
3. Metrics and utilities on 15 minute resolution

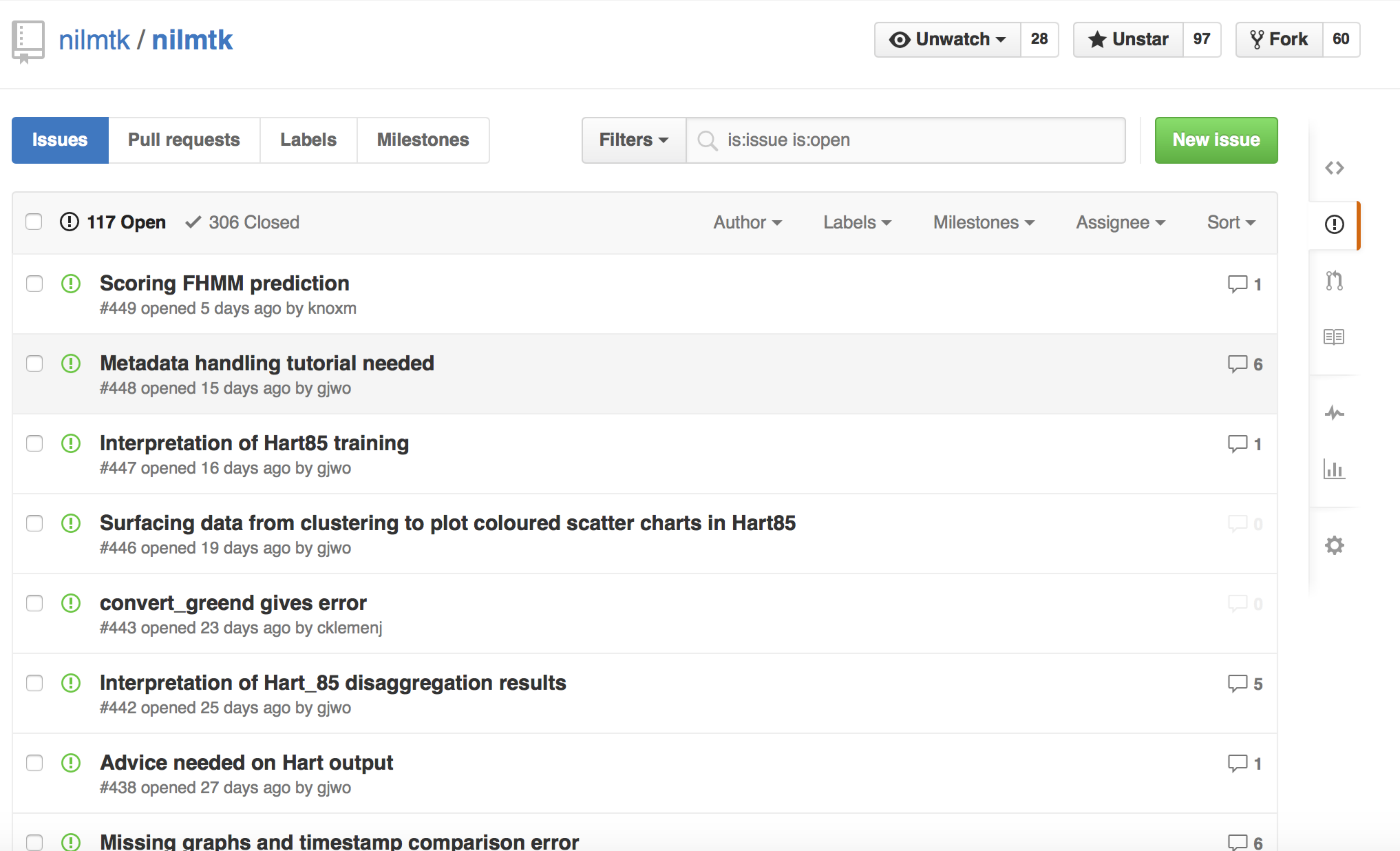
Homes "changing" behaviour pose an interesting challenge to Neighbourhood NILM

1. Balance between "historical" data and recent trends?
2. Continue having same neighbours?
    1. When to "change" the neighbours of a home

Scaling NILM to "similar" commercial buildings/different appliance types

1. Class of commercial buildings have exact same electrical infrastructure
2. Deployment across 10 dairy booths in New Delhi

NILMTK-"The cost of impact is a bug report/feature request a day on Github" :)



Conclusions

Making NILM practical in 3 ways:

1. **Comparable-** NILMTK
2. **Utility-driven-** Energy saving feedback
   Inferring household characteristics
3. **Scalable-** Neighbourhood NILM

Other work

1. Insights into home energy consumption in India [Buildsys 2013]
2. Inferring household characteristics from NILM [under submission Percom 2016]
3. Improving NILM performance using additional data [ICMLA 2013]