Energy disaggregation on hourly whole-building electricity data

Qing Gao
School of Data Science, Fudan University
2017/9/11

Outline

- Background and motivation
- Disaggregation Model
- Data
- Disaggregation Results
- Summary

Background

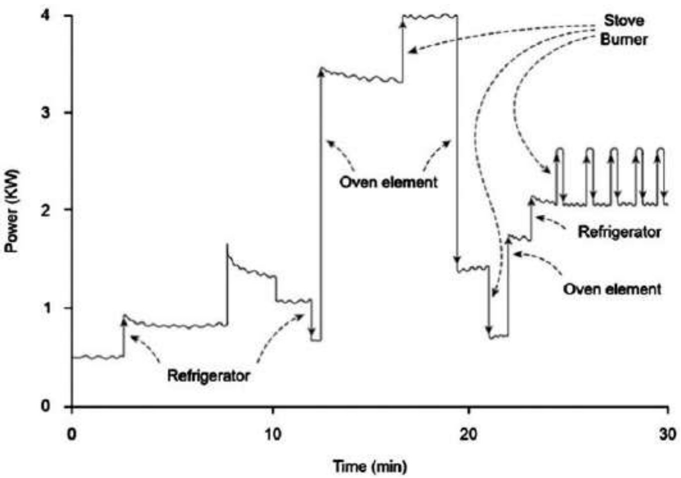
- Why disaggregation
  - Efficient energy arrangement, redesign better appliance
  - Improve building operational efficiency
  - Energy saving, reducing cost of energy supply
- Research status[1]

Data freq	1h-15min	1min - 1s (Hz)	1-60Hz	60Hz-2kHz	10-40kHz	>1MHz
Data features	Duration and time of appliance use	Steady state steps/ power transitions		Current, voltage, low order harmonics	Current, voltage, medium order	Current, voltage, high order
Algorithm	Machine learning, sparse coding	Steps signatures database matching		Event identification	Machine learning, FHMM	Machine learning, neural network
Appliances identified	loads related with temp, continuous, time	Top<10 types: refrigerator, ACs, heaters, dryers, etc.	10-20 types	Not known	20-40 types: toasters, computers, etc.	40-100 specific appliances: light 1, light 2

[1] K. Carrie Armel, et. al., Energy policy 52 (2013) 213-234.

Background

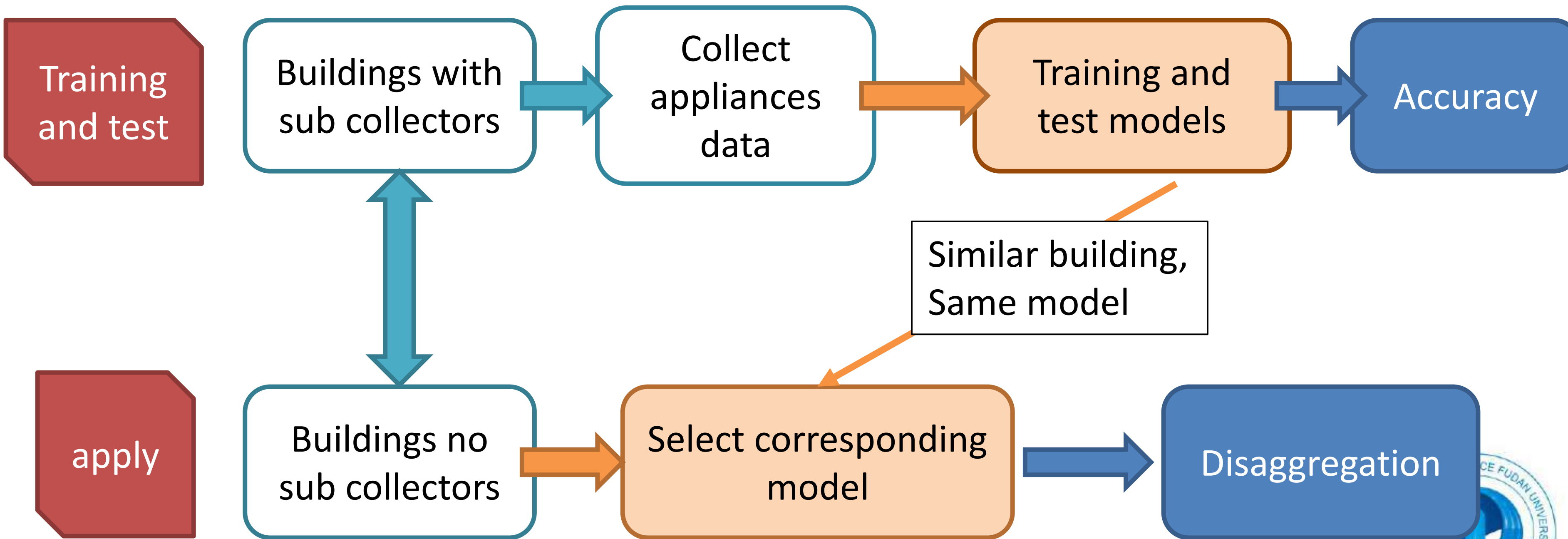
- Traditional methods
  - Event based disaggregation, electricity data alone
    - ✓ Unsupervised models
  - High frequency electricity data
    - ✓ Appliance power curves
    - ✓ lab experiments
- For low frequency data
  - supervised
  - Sparse coding
  - FHMM model
    - ✓ more information input (adopted)



Disaggregation to appliances.
(Hart, 1992, IEEE 80 (12), 1870–1891.)

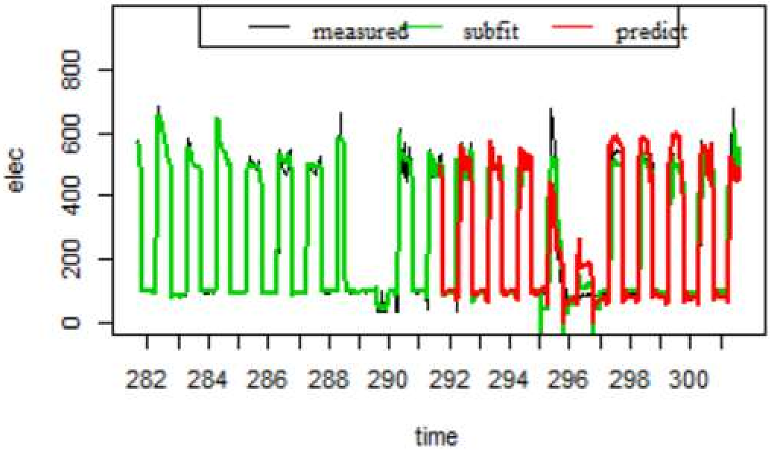
Motivation

- Supervised disaggregation for commercial buildings
  - Calibrate the meters for appliances
  - Training model, improve accuracy
  - Disaggregation for buildings without sub meters



Disaggregation

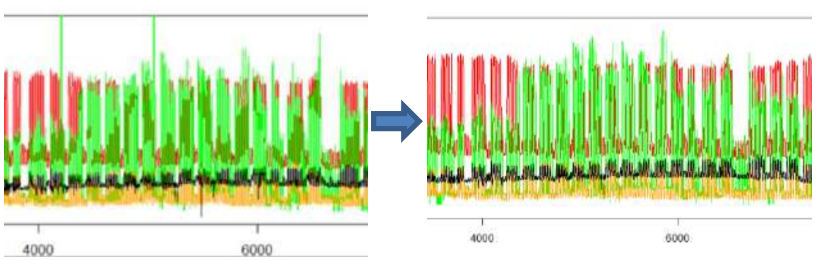
- **Model set**
  - Training + test sets
    - ✓ First part, training model with all observation
    - ✓ Second part, test the model with only total data
- **Steps**
  1. Data loading and cleaning
  2. Model training for appliances
  3. Predict for each appliance
  4. Repair by total data
  5. Calculate the accuracy

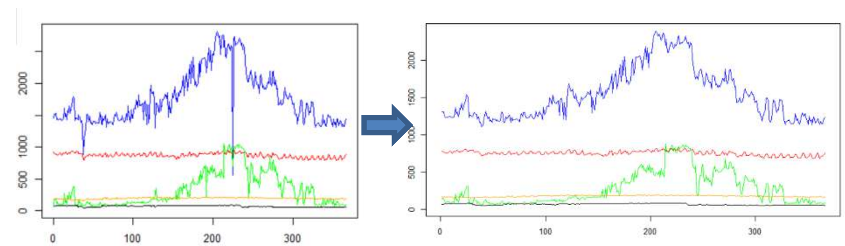


Model training and disaggregation

Disaggregation: data cleaning

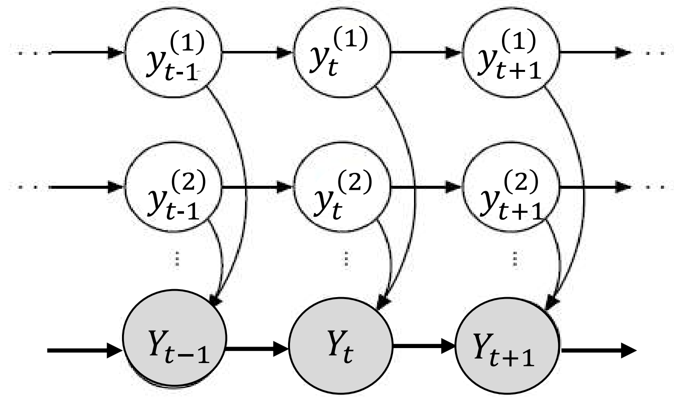
- Clean methods
  - Outliers: 3 standard deviation, 4 weeks for reference
  - Missing: Linear interpolation
- For real data
  - clean respectively, appliances, weekday, weekend, hours
  - Aggregation data, compare to the sum of separated appliances
    - ✓ Larger parts, smaller parts





Disaggregation: FHMM model training

- FHMM Model



$$p\left(Y_t\left(S_t^{(1:N)}\right)\right) \sim N\left(\sum_{i=1}^{N} \bar{y}_t^{(i)}\left(S_t^{(i)}\right), \Sigma^2\left(S_t^{(1:N)}\right)\right), \qquad 1 \leq S_t^{(i)} \leq m^{(i)}$$

where i for separated appliance, $m^{(i)}$ is the number of states for i-th appliance, and

$$\bar{y}_t^{(i)}\left(S_t^{(i)}\right) = c_i\left(S_t^{(i)}\right) + \vec{\beta}_i\left(S_t^{(i)}\right) \times \overrightarrow{\mathrm{Out}_t} = f^{(i)}\left(S_t^{(i)}\right)$$

$$p\left(y_t^{(i)}\left(S_t^{(i)}\right)\right) \sim N\left(\bar{y}_t^{(i)}\left(S_t^{(i)}\right), \sigma_i^2\left(S_t^{(i)}\right)\right)$$

- State probability, transition matrix are combination of appliances
- Number of states $\prod_{i=1}^{N} m^{(i)}$

[1] Kolter, et.al.(2012) International Conference on Artificial Intelligence and Statistics Pp. 1472–1482.

Disaggregation: FHMM model training

- HMM model for each appliance
  - Initial state probability

$$\sum_{j=1}^{m^{(i)}} \delta_j^{(i)} = 1, \qquad 1 \leq j,k \leq m^{(i)}$$

  - Transition matrix

$$\Gamma^{(i)} = \left(\gamma_{jk}^{(i)}\right), \qquad \gamma_{jk}^{(i)} = p\left(S_t^{(i)} = j \middle| S_{t+1}^{(i)} = k\right)$$

$$\sum_{k=1}^{m^{(i)}} \gamma_{jk}^{(i)} = 1, \qquad 1 \leq j,k \leq m^{(i)}$$

  - Number of degree of freedom

$$\mathrm{NDF} = \left(m^{(i)} - 1\right) + \left({m^{(i)}}^2 - m^{(i)}\right) + m^{(i)} \times (n_{out} + 1) + m^{(i)}$$

Initial state | Transition matrix | Outer effects | Normal dis

Disaggregation: FHMM model training

- Estimation with EM algorithm by definite $m^{(i)}$

$$\log \mathbf{L_T}^{(i)} = \underbrace{\log \boldsymbol{\delta}^{(i)}}_{\text{Initial state}} + \underbrace{\sum_{t=2}^{T} \log \boldsymbol{\gamma}^{(i)}_{S_{t-1},S_t}}_{\text{Transition matrix}} + \underbrace{\sum_{t=1}^{T} \log \boldsymbol{p}\left(y_t^{(i)}\left(S_t^{(i)}\right)\right)}_{\text{Conditional probability density}}$$

$$= \sum_{j=1}^{m^{(i)}} u_j^{(i)}(t) \log \delta_j^{(i)} + \sum_{j,k=1}^{m^{(i)}} \left(\sum_{t=2}^{T} v_{jk}^{(i)}(t)\right) \log \gamma_{jk}^{(i)} + \sum_{j=1}^{m^{(i)}} \sum_{t=1}^{T} u_j^{(i)}(t) \log p\left(y_t^{(i)}\left(S_t^{(i)}\right)\right)$$

- The third part

$$u_j(t)\big(y_t - \hat{y}_t(j)\big) \sim N(0, \sigma_j^2)$$

$$\sigma_i^2 = \frac{\sum_{t=1}^{T} u_i(t)\big(y_t - \hat{y}_t(i)\big)^2}{T}$$

the parameters equals to that from linear regression

$$u_j(t) y_t = c_j u_j(t) + u_j(t) \vec{\beta}_i\left(S_t^{(i)}\right) \times \overrightarrow{\mathrm{Out}_t} + \varepsilon_t(j)$$

Disaggregation: FHMM model training

- Decide the number of states $m^{(i)}$
  - Loop from 2 to 25

$$\text{residual} = y_t^{(i)} - \sum_{j=1}^{m^{(i)}} \hat{y}_t(S=j) \times p(S_t^{(i)}=j)$$ weak stable

$$\text{BIC} = \log \frac{\text{SSR}_\text{p}}{T} + \log(T)\frac{ndf}{T}$$ least

$$\frac{\left(y_t^{(i)} - mean\left(y^{(i)}\right)\right)}{sd\left(y^{(i)}\right)} < 5$$ no outliers

  - Repeat fitting, take best fit result

Disaggregation: predict

- Appliance, multi stages (h), get the expected value

1. State prob now $\vec{u}(t) = \vec{\boldsymbol{p}}_{\boldsymbol{t}}(\boldsymbol{S}|\boldsymbol{y}_{\boldsymbol{t}}) = \frac{\vec{\boldsymbol{p}}_{\boldsymbol{t}}(\boldsymbol{S})^{\boldsymbol{T}}\vec{\boldsymbol{p}}_{\boldsymbol{S}}(\boldsymbol{y}_{\boldsymbol{t}})}{\sum \vec{\boldsymbol{p}}_{\boldsymbol{S}}(\boldsymbol{y}_{\boldsymbol{t}})}$
2. State prob future $\vec{\boldsymbol{p}}_{\boldsymbol{t}+\boldsymbol{h}}(\boldsymbol{S}) = \vec{u}(t)^T \boldsymbol{\Gamma}^h$
3. State value future $\vec{\boldsymbol{y}}_{\boldsymbol{t}+\boldsymbol{1}}(\boldsymbol{S}) = c_j + \sum_{k=1}^{n} \beta_{jk} \times \mathrm{Out}_{k,t+h}\left(S_t^{(i)} = j\right)$
4. Expect value $\hat{y}_{t+1} = \vec{\boldsymbol{p}}_{\boldsymbol{t}+\boldsymbol{h}}(\boldsymbol{S})^{\boldsymbol{T}}\vec{\boldsymbol{y}}_{\boldsymbol{t}+\boldsymbol{h}}(\boldsymbol{S})$

- Aggregation data, determine state probability

1. State probs $U_s(t) = \prod_{i=1}^{N} u_{j_i}^{(i)}(t)\,, s = \sum_{i=1}^{N} j_i \times \prod_{k=1}^{i-1} m^{(k)}$
2. State prob next $\vec{\boldsymbol{p}}_{\boldsymbol{t}+\boldsymbol{1}}^{\boldsymbol{tot}}(\boldsymbol{S}) = \vec{U}_s(t)^T \boldsymbol{\Gamma}^{\boldsymbol{tot}}$
3. Measured value next Y_{t+1}
4. Conditional prob $\vec{U}_S(t+1) = \vec{\boldsymbol{p}}_{\boldsymbol{t}+\boldsymbol{1}}^{\boldsymbol{tot}}(\boldsymbol{S}|\boldsymbol{Y}_{\boldsymbol{t}}) = \frac{\vec{\boldsymbol{p}}_{\boldsymbol{t}+\boldsymbol{1}}^{\boldsymbol{tot}}(\boldsymbol{S})^{\boldsymbol{T}}\vec{\boldsymbol{p}}\left(Y_{t+1}\left(S_{t+1}^{(1:N)}\right)\right)}{\sum \vec{\boldsymbol{p}}\left(Y_{t+1}\left(S_{t+1}^{(1:N)}\right)\right)}$

Disaggregation: result

- Repair appliances prediction by deviation

$$y_{t+1day}^{(i)} = \hat{y}_{t+1day}^{(i)} + per_1 \times per_2 \times \left(Y_{t+1day} - \sum_{i=1}^{N} \hat{y}_{t+1day}^{(i)} \right)$$

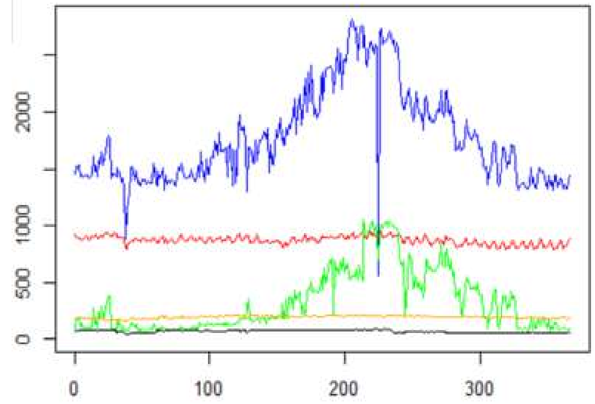
where $\mathrm{per}_1(\Delta) = \frac{\hat{y}_{t+1day}^{(i)} - y_t^{(i)}}{\sum_{i=1}^{N} \hat{y}_{t+1day}^{(i)} - y_t^{(i)}}$, $\quad \mathrm{per}_2(\Delta) = \frac{\sum_{j=1}^{m^{(i)}} prob_j^{(i)} \times {\sigma^{(i)}}^2}{\sum_{i=1,j=1}^{n,m^{(i)}} prob_j^{(i)} \times {\sigma^{(i)}}^2}$

- Accuracy (relative uncertainty)

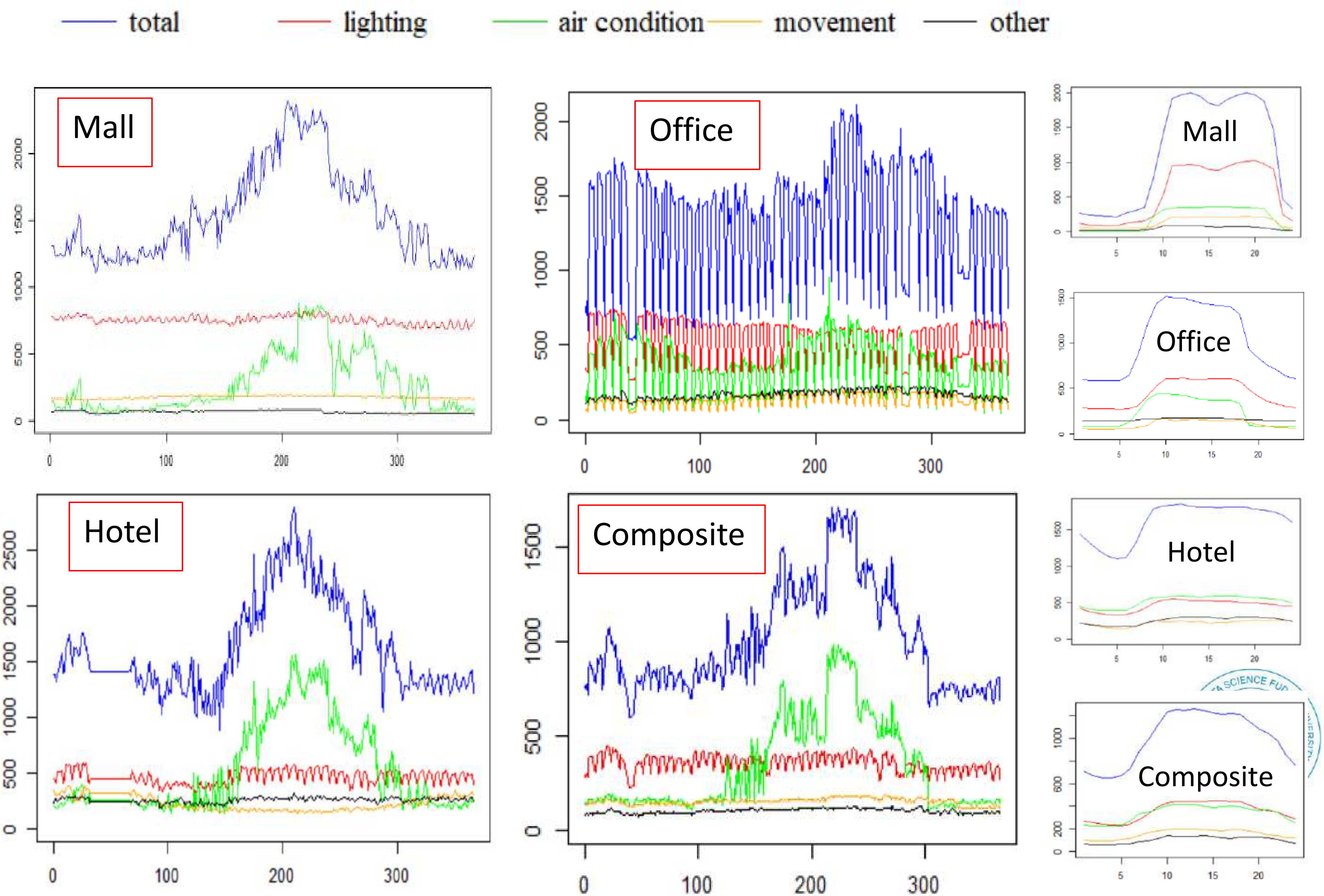
$$\mathrm{rela} = \sqrt{\frac{\sum_t (\mu_t - y_t)^2}{\sum_t y_t^2}}$$

Data

- Electricity
  - Mall, office, hotel, composite
  - Time: 2016-1-1 0:00 to 2016-12-31 23:00, hourly
  - Measured items: total, lighting, air condition, movement, others
- More
  - Temperature[2], raining, wind velocity, pressure, humidity
  - Holiday: 10 legal holiday, 11 weekend, 00 workday;
  - Day-night: dummy variable; hour, 0~23
- Cleaning
  - Outliers
  - Missing values
  - Unknown = total-sum



Data status


total
lighting
air condition
movement
other
Mall
Office
Hotel
Composite
Mall
Office
Hotel
Composite

Data status

- Electricity (day night obvious, air condition season sensitive)

mall	hour	Total (kWh)	Light (kWh)	AC (kWh)	Mv (kWh)	Other (kWh)
Day	9-22	1200~2300 (season)	800	0-800 (season)	200	100
night	23-8	~200	100	0	20	0

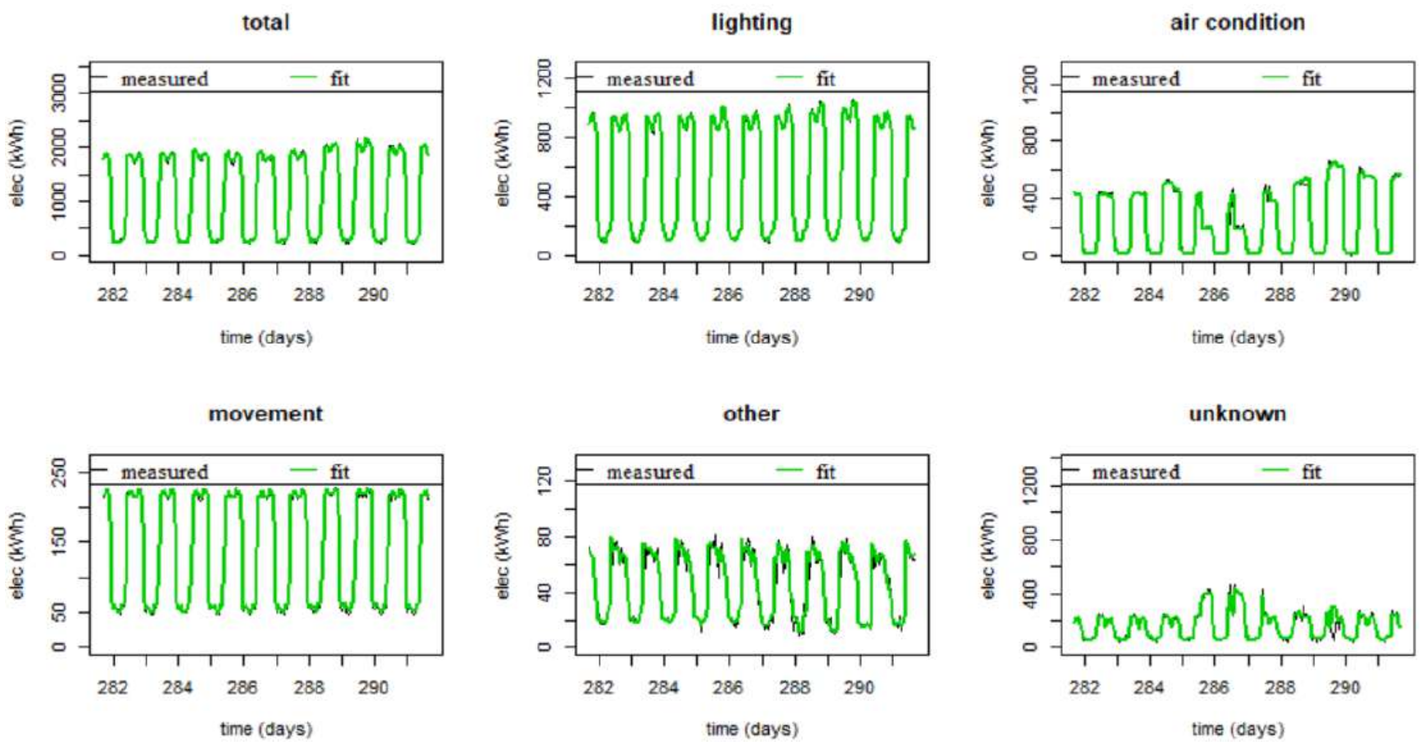
- Comparison all

	Mall	Office	Hotel	Composite
Max (kWh)	2500	2000	3500	2500
Day	9-22	6-18	8-23	8-22
Dominant	Lighting	lighting	Acs summer, lighting other	Acs summer, lighting other
Week cycle	No	Lighting, Acs, movement	Lighting	Lighting
Common	Air conditions sensitive to season, large fluctuation; spring festival effect obvious, difference between appliances small			

Disaggregate result: mall

- Training

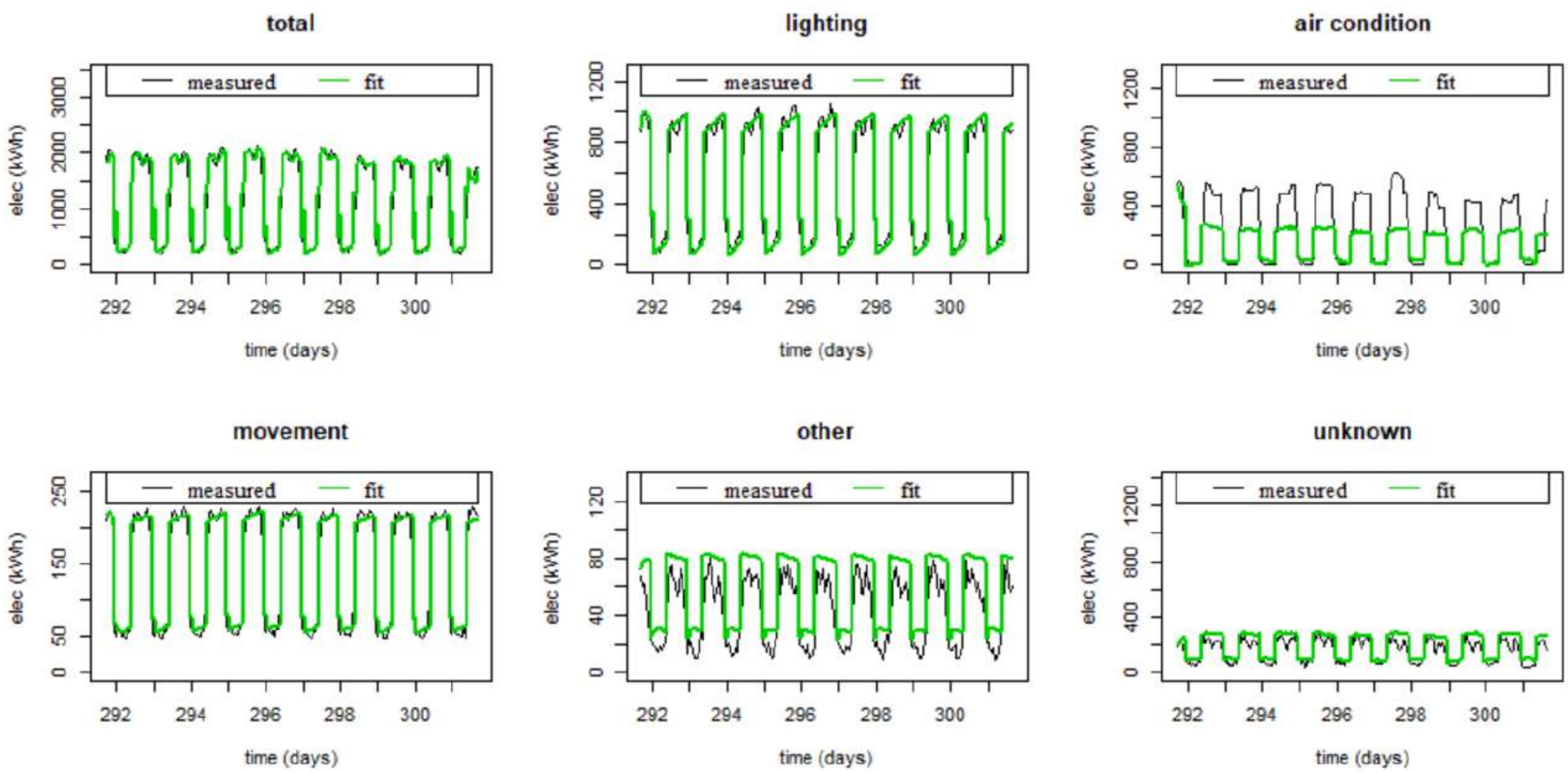
Item	Lighting	Air condition	Movement	Others	Unknown	Tot
N states	16	13	12	16	11	12
Sigma	15.02	20.9	4.98	3.41	31.38	67.04
Rela err (%)	1.98	5.33	2.91	15.25	8.61	3.90



Disaggregate result: mall

- Predicting in testing set

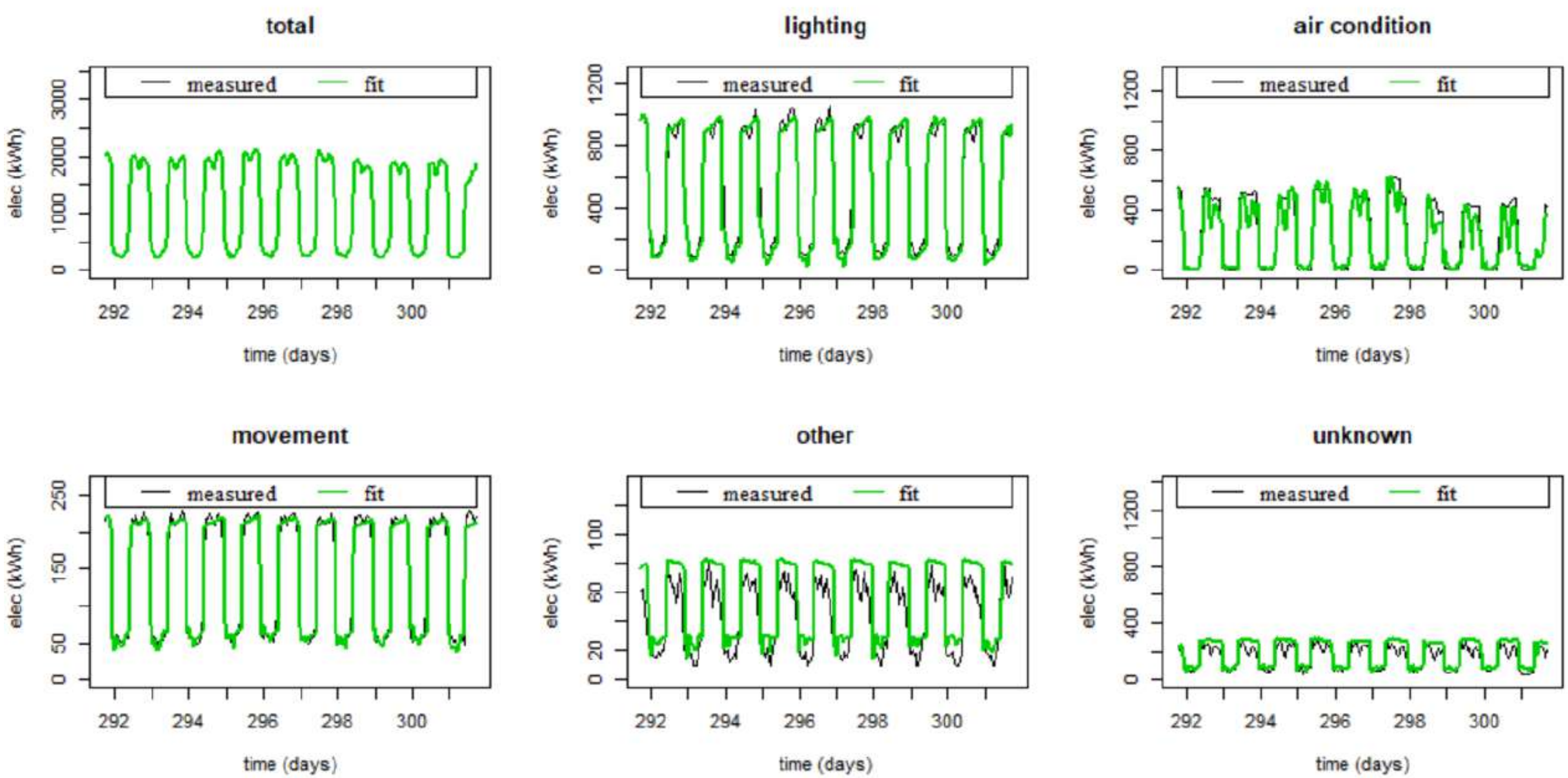
Item	Lighting	Air condition	Movement	Others	Unknown	Tot
Stages	Multi	Multi	Multi	Multi	Multi	One
Rela err (%)	13.99	55.39	12.52	38.32	34.10	16.33



Disaggregate result: mall

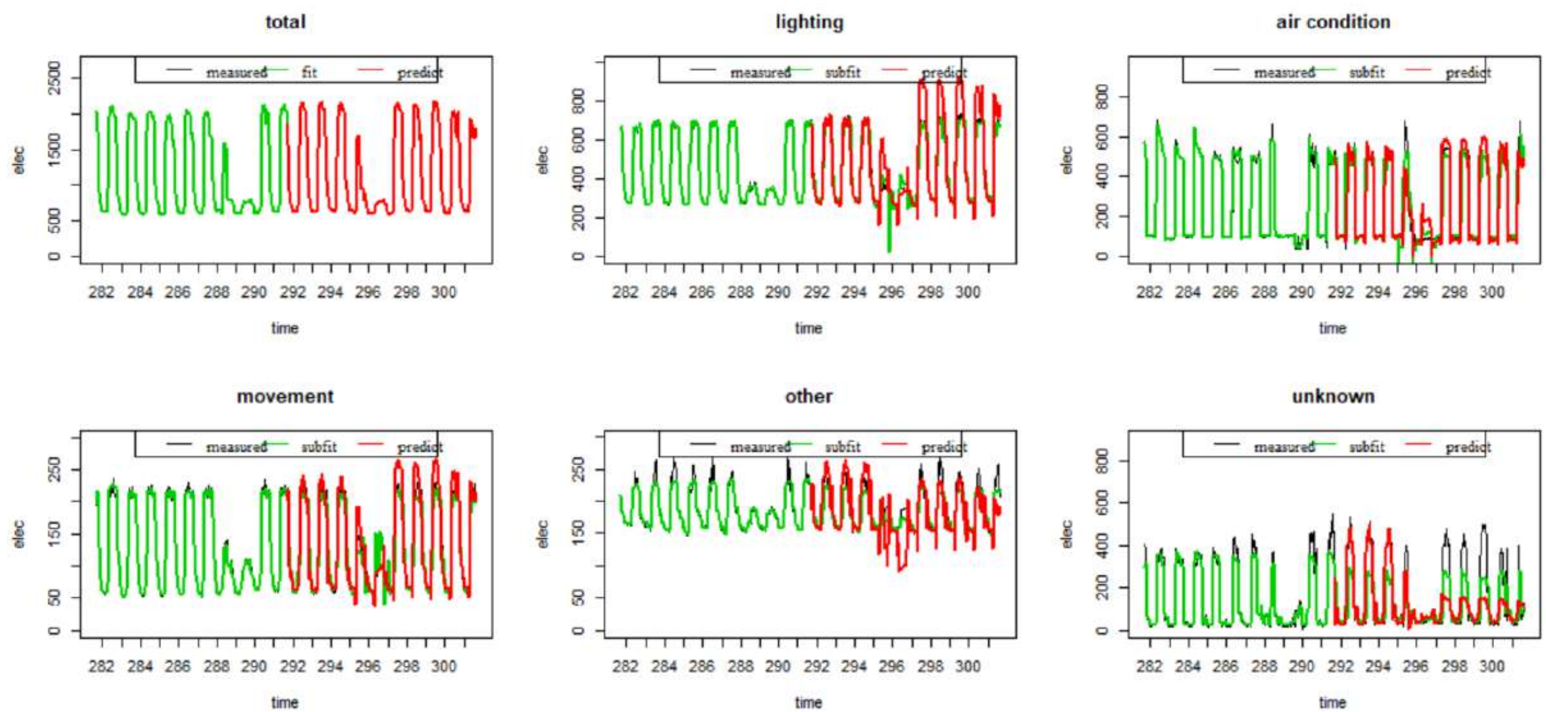
- Repair by aggregation data

Item	Lighting	Air condition	Movement	Others	Unknown
relative err(%)	12.46	37.53	10.52	32.02	32.25



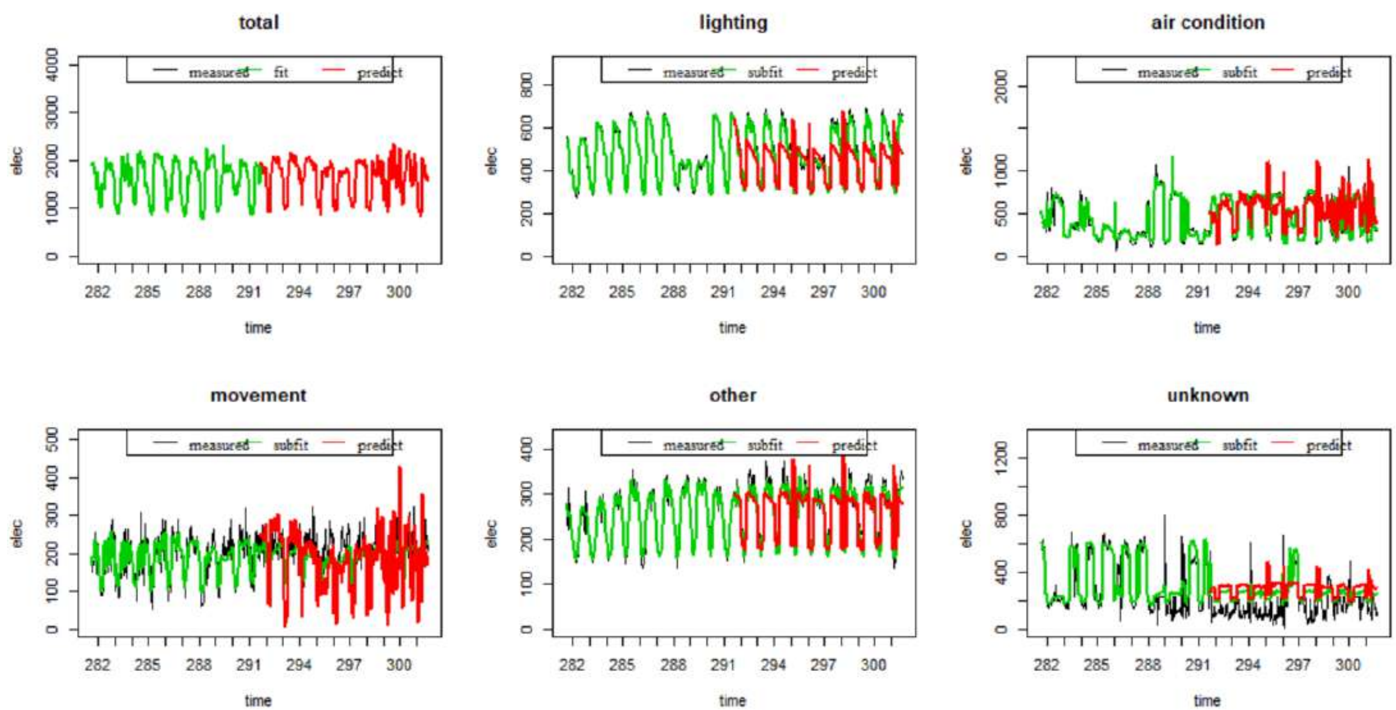
Disaggregate result: office

Item	Lighting	Air condition	Movement	Others	Unknown	Tot
N states	13	12	13	10	4	18
Training rela err (%)	2.13	6.32	4.29	3.76	27.13	2.87
sigma	11.61	23.19	6.31	7.12	43.81	40.4
Predict rela err (%)	19.59	43.47	25.70	15.93	72.43	19.11
Last rela err (%)	14.01	29.50	20.63	19.35	61.26	—



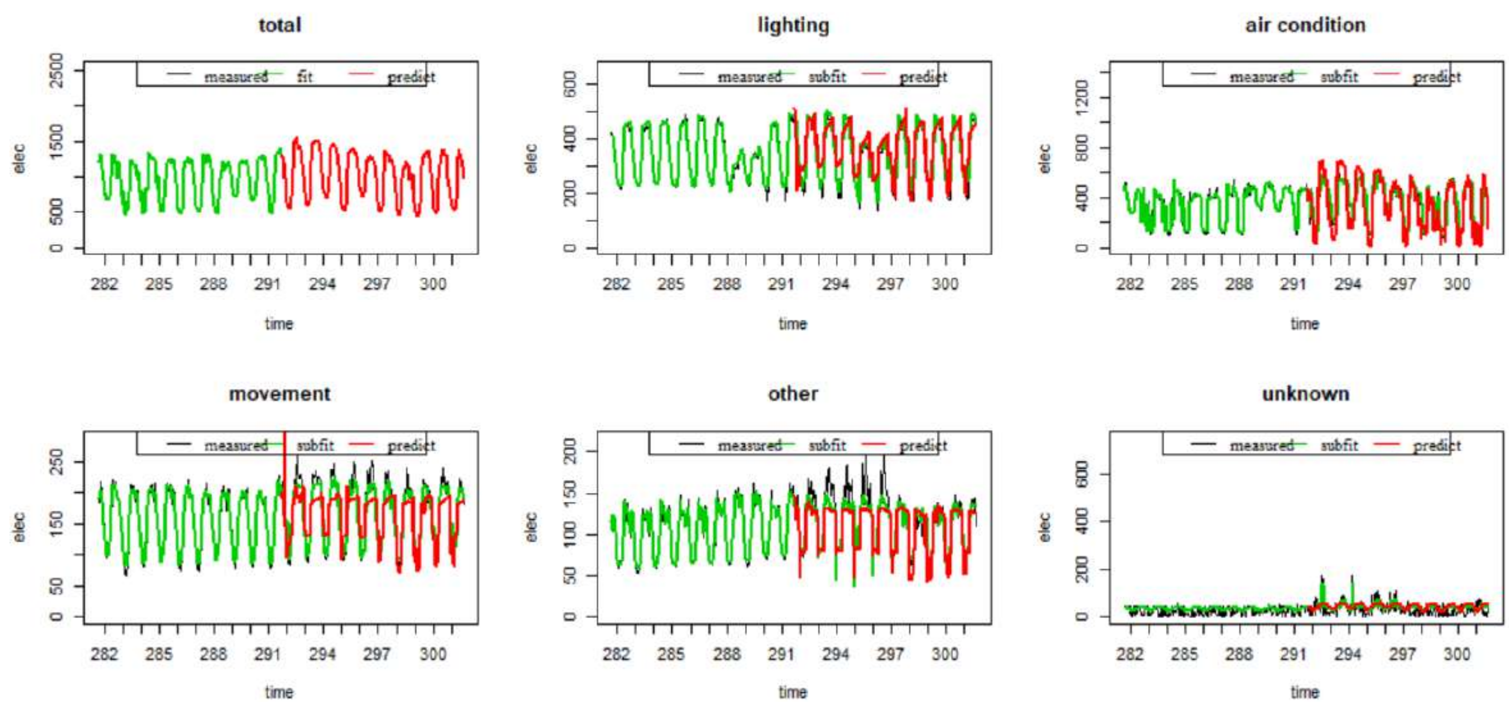
Disaggregate result: hotel

Item	Lighting	Air condition	Movement	Others	Unknown	Tot
N states	10	7	3	8	3	7
Training rela err (%)	3.06	6.67	14.16	5.64	25.17	4.04
sigma	17.09	62.44	30.03	16.98	92.08	98.82
Predict rela err (%)	16.14	50.67	22.85	13.79	95.20	11.21
Last rela err (%)	16.92	36.21	33.91	14.93	96.57	—



Disaggregate result: composite

Item	Lighting	Air condition	Movement	Others	Unknown	Tot
N states	18	9	10	9	3	10
Training rela err (%)	2.45	5.38	4.04	4.63	35.93	3.15
sigma	10.93	29.48	8.05	6.13	25.94	44.56
Predict rela err (%)	22.14	46.68	23.92	18.13	64.91	9.09
Last rela err (%)	13.97	35.47	18.72	15.82	62.97	—



Disaggregation comparison

Item		Lighting	Air condition	Movement	Others	Unknown	Tot
Training relative error (%)	Mall	1.98	5.33	2.91	15.25	8.61	3.90
	Office	2.13	6.32	4.29	3.76	27.13	2.87
	Hotel	3.06	6.67	14.16	5.64	25.17	4.04
	Composite	2.45	5.38	4.04	4.63	35.93	3.15
Result relative error (%)	Mall	12.46	37.53	10.52	32.02	32.25	—
	Office	14.01	29.50	20.63	19.35	61.26	—
	Hotel	16.92	36.21	33.91	14.93	96.57	—
	Composite	13.97	35.47	18.72	15.82	62.97	—

- Testing relative uncertainty larger than training
- The larger of relative uncertainty for training, the larger disaggregation
- Performance similar for buildings
- Air condition, unknown largest both training and disaggregation, for large fluctuation

Summary

- Extend FHMM model with bonus data to disaggregate hourly whole-building electricity consumption into appliances
- Apply the method to several commercial buildings
  - Successfully disaggregate and get rules of appliances
  - Performance for different buildings are similar
  - Model training perfect, relative uncertainty lower than 7%
  - Model testing, air condition not good for large fluctuation
- Extend to similar buildings without collectors
  - Input the characters of buildings into the model
  - Training different models for different type buildings
  - Important for energy monitoring, need response, accurate prediction

Thank you!

SCHOOL OF DATA SCIENCE FUDAN UNIVERSITY 复旦大学大数据学院