ETH zürich

Smart Meter Data-Driven Load Forecasting and Price Design in the Retail Market

Graduate Seminar @ KAUST

Yi Wang, Postdoc @ ETH
yiwang@eeh.ee.ethz.ch

Feb 21, 2021

Power Systems Laboratory

Appointment

2019.2- Postdoc, ETH Zurich (Prof. Gabriela Hug)

Education

2014.9-2019.1 Ph.D., Tsinghua University (Prof. Chongqing Kang)

2017.3-2018.4 Visiting Student, University of Washington (Prof. Daniel Kirschen)

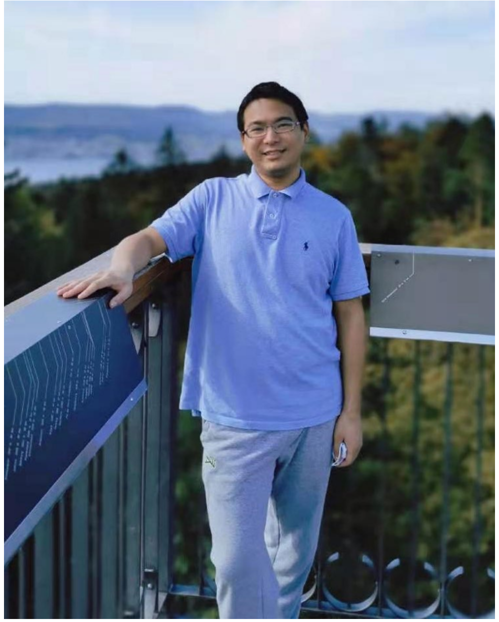
2010.9-2014.6 B.S., Huazhong University of Science and Technology

Research Interests

Data analytics for smart energy

Cyber-physical power and energy systems

Multi-energy systems integration



Acknowledgements

The works presented here are from the collaboration with the following professors and graduate students:

- **Prof. Gabriela Hug** and **Mr. Leandro Von Krannichfeldt** from ETH Zurich;
- **Prof. Chongqing Kang**, **Prof. Qixin Chen**, and **Mr. Cheng Feng** from Tsinghua University.

1. **Yi Wang**, Qixin Chen, Tao Hong, and Chongqing Kang, "Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges," *IEEE Transactions on Smart Grid*, 2019, 10(3):3125-3148.

2. **Yi Wang**, Qixin Chen, Mingyang Sun, Chongqing Kang and Qing Xia, "An Ensemble Forecasting Method for the Aggregated Load with Subprofiles," *IEEE Transactions on Smart Grid*, 2018, 9(4): 3906-3908.

3. Leandro Von Krannichfeldt, **Yi Wang**, and Gabriela Hug, "Online Ensemble Learning for Load Forecasting," *IEEE Transactions on Power Systems*, 2021, 36(1):545-548.

4. Cheng Feng, **Yi Wang**, Kedi Zheng, and Qixin Chen, "Smart Meter Data-Driven Customizing Price Design for Retailers," *IEEE Transactions on Smart Grid*, 2020, 11(3):2043-2054.

5. **Yi Wang**, Leandro Von Krannichfeldt, Gabriela Hug, "Probabilistic Aggregated Load Forecasting with Fine-grained Smart Meter Data," *IEEE PowerTech 2021*.

Slides available at http://www.eeyiwang.com/KAUST_Seminar_Yi.pdf

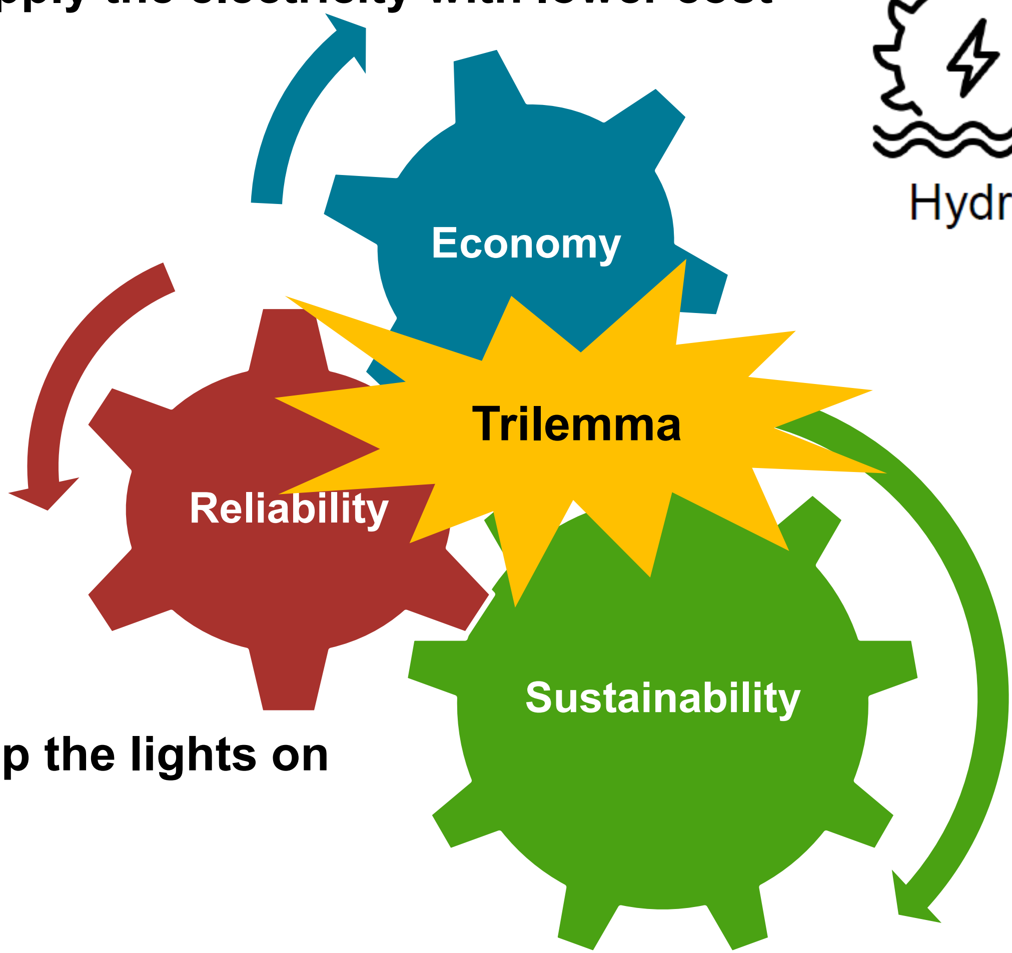
Outlines

- Backgrounds
- Aggregated Load Forecasting
- Personalized Retail Price Design
- Conclusions

ETH zürich

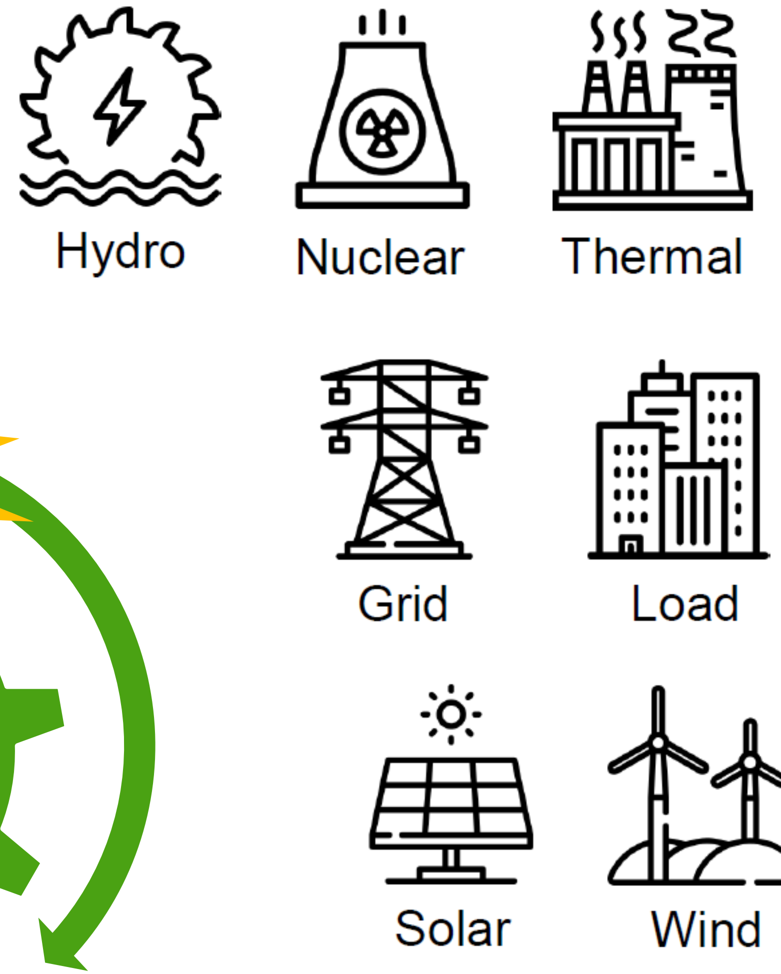
Backgrounds

Supply the electricity with lower cost



Keep the lights on

Accommodate more renewable energy

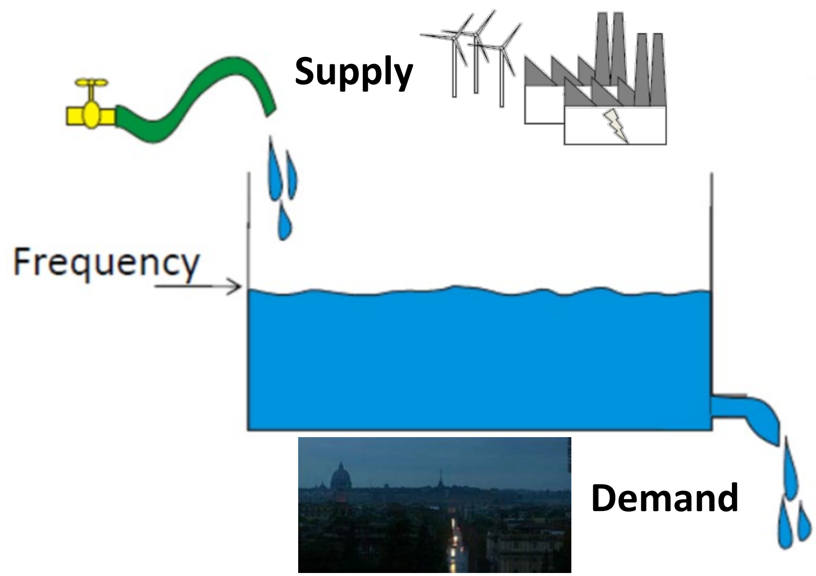


Backgrounds

The power generation and consumption should be balanced in real-time.

Traditional Generation + Renewable Energy = Energy Demand

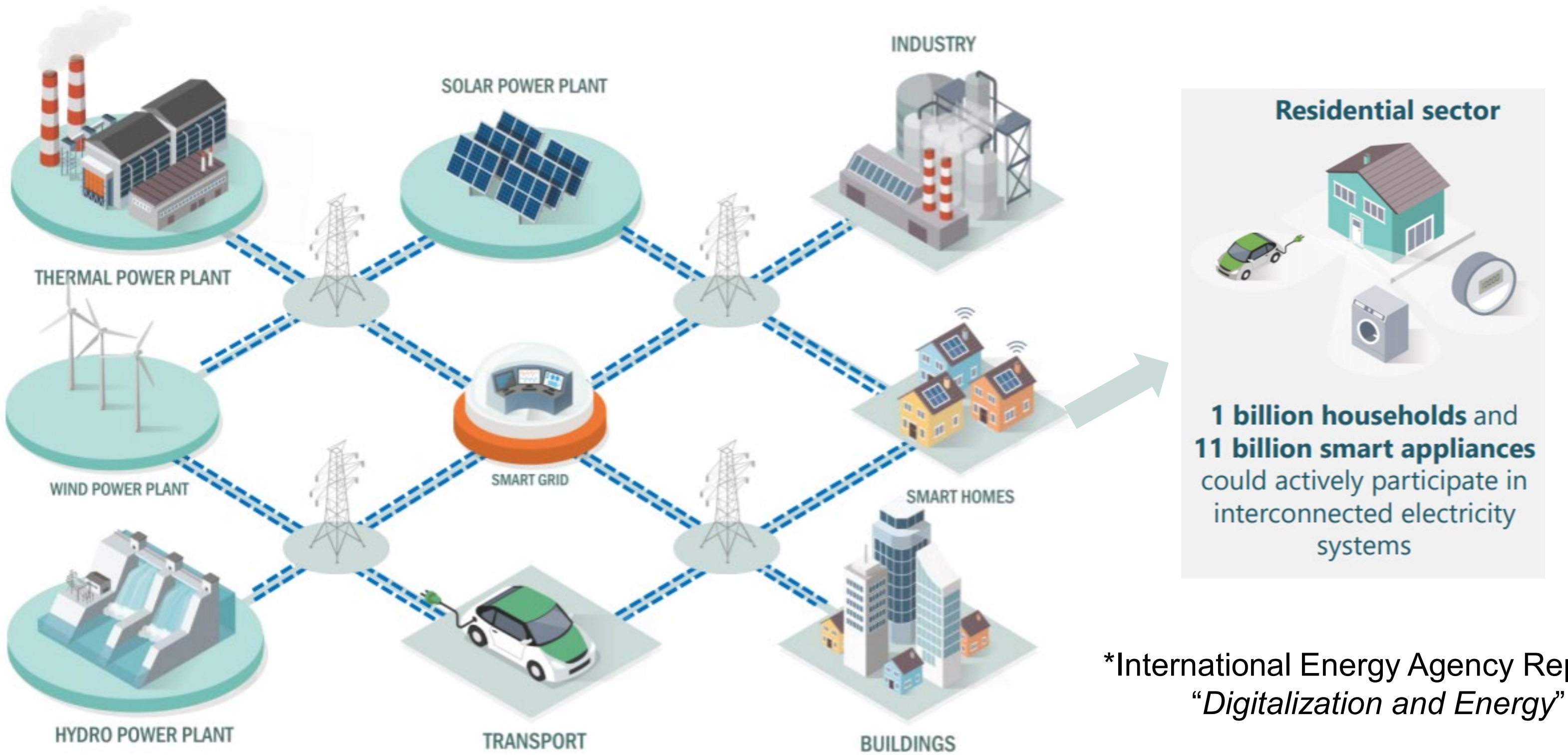
Real-time Balance in Power Systems



Flexibility: Better ways of **matching supply and demand** over multiple time and spatio-scales.

Backgrounds

Various sensors and controllers will be installed in the power and energy systems.



*International Energy Agency Report on "*Digitalization and Energy*"

Demand response in buildings, industry and transport could provide **185 GW** of flexibility, and avoid **USD 270 billion** of investment in new electricity infrastructure.

- https://www.iea.org/reports/digitalisation-and-energy

Backgrounds

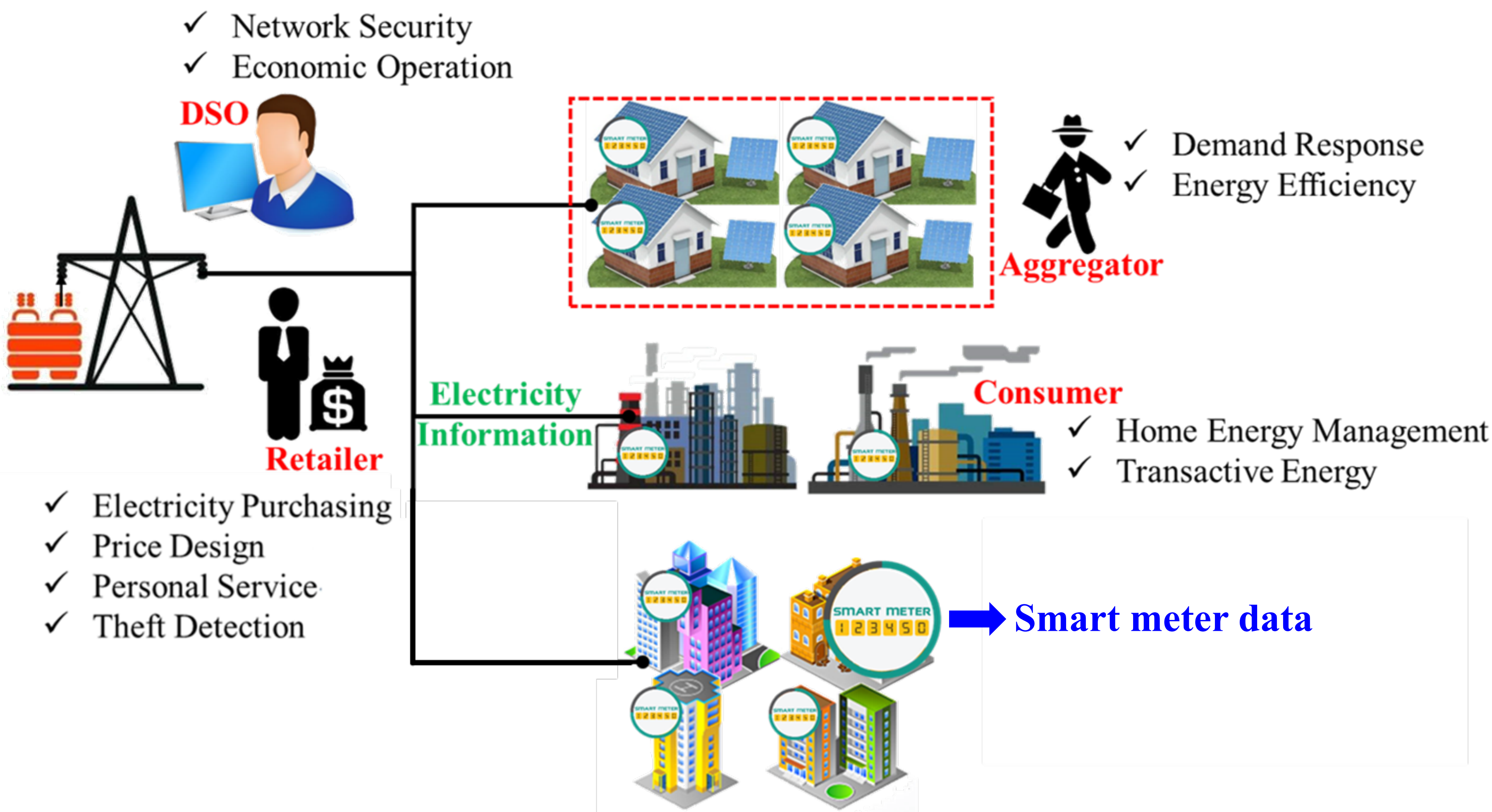
No.	System/ Data	Data Source	Data Type	Frequency	Data Structure
1	Economic Information	Statistic Bureau	GDP、CPI、PMI（Purchasing Managers Index）、Sales Value、Prosperity Index	Per Month	Non structural
2	Energy Consumption Data	Energy Efficiency Platform	Electrical Load、Output、Power Quality、Temperature	15Min	Non structural /Structural
3	Meteorological Data	Meteorological Bureau	Temperature、Humidity、Rainfall	Per Day	Structural
4	EV Charging Data	Charging-Pile RTU	Current、Voltage、Charging Rate、State of Charge	15Min	Structural
5	Customer Service Voice Data	Customer Service System	Customer Voice Data	Real Time	Non structural

Variety **Value???** **Velocity**

10 million Smart Meters, 15min ⇨ 60GB per day, 21TB per year. **Volume**

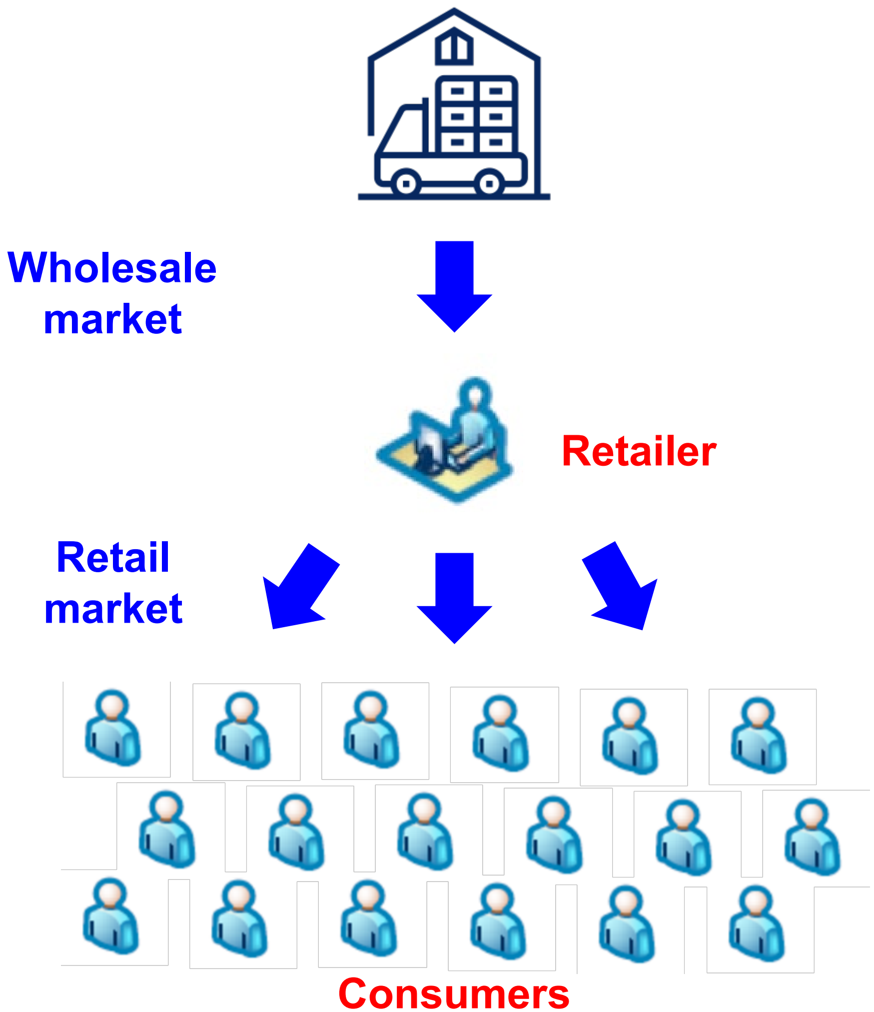
Backgrounds

Participators and their businesses on the demand side



○ Wang Y, Chen Q, Hong T, et al. Review of smart meter data analytics: Applications, methodologies, and challenges[J]. IEEE Transactions on Smart Grid, 2018, 10(3): 3125-3148.

Backgrounds



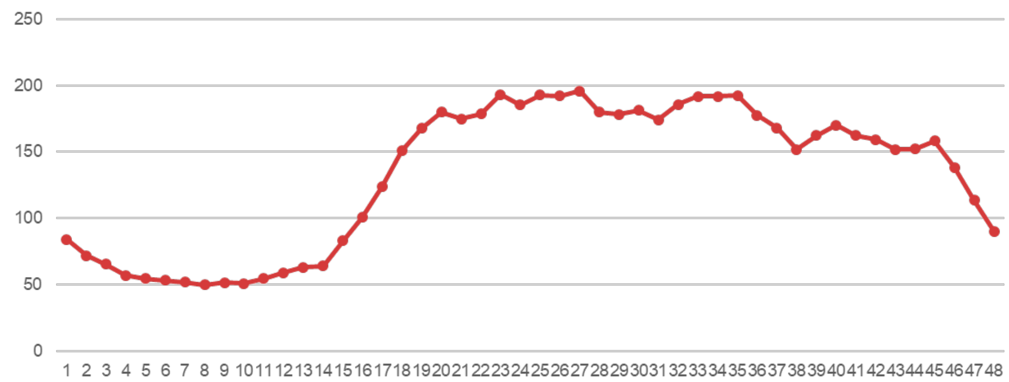
❖ **Aggregated Load Forecasting**

❖ **Personalized Retail Price Design**

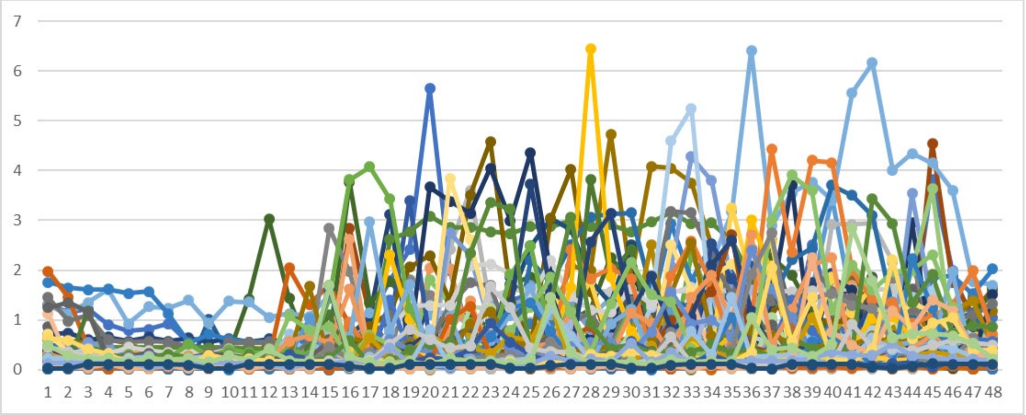
Aggregated Load Forecasting

➢ Introduction

Traditional load forecasting algorithms directly use historical data at the aggregation level.



With the prevalence of smart meters, fine-grained sub profiles reveal more information about the aggregated load and further help improve the forecasting accuracy.



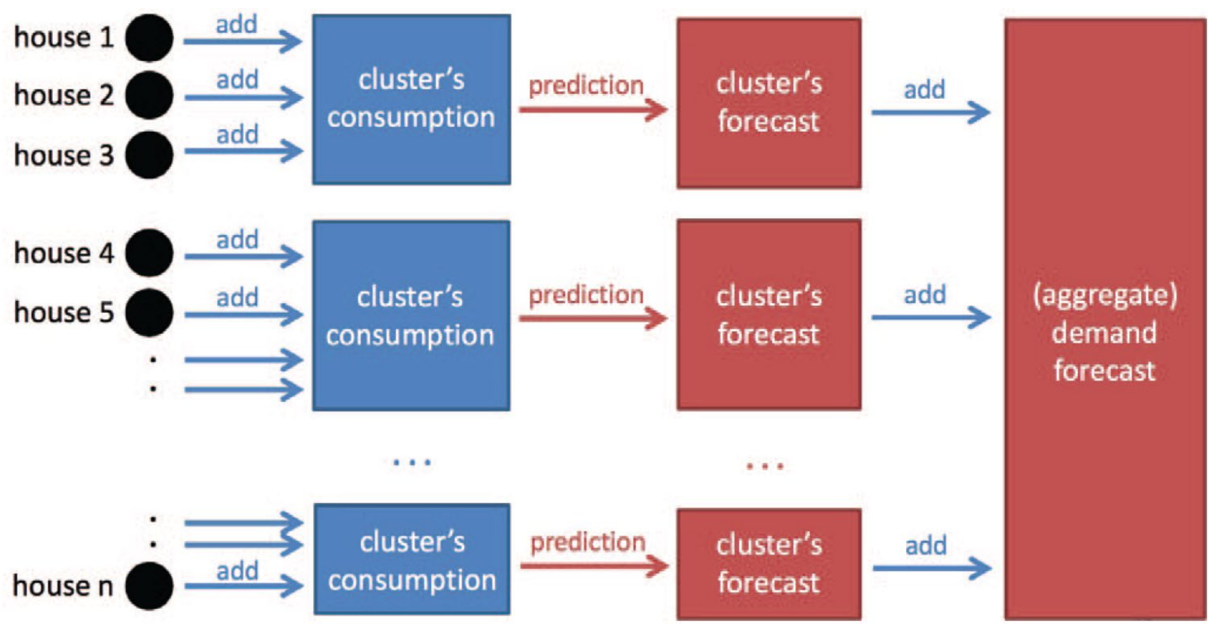
Aggregated Load Forecasting

➢ Introduction

Three strategies for aggregated load forecasting (ALF):

1) Top-down; 2) bottom-up; 3) clustering based.

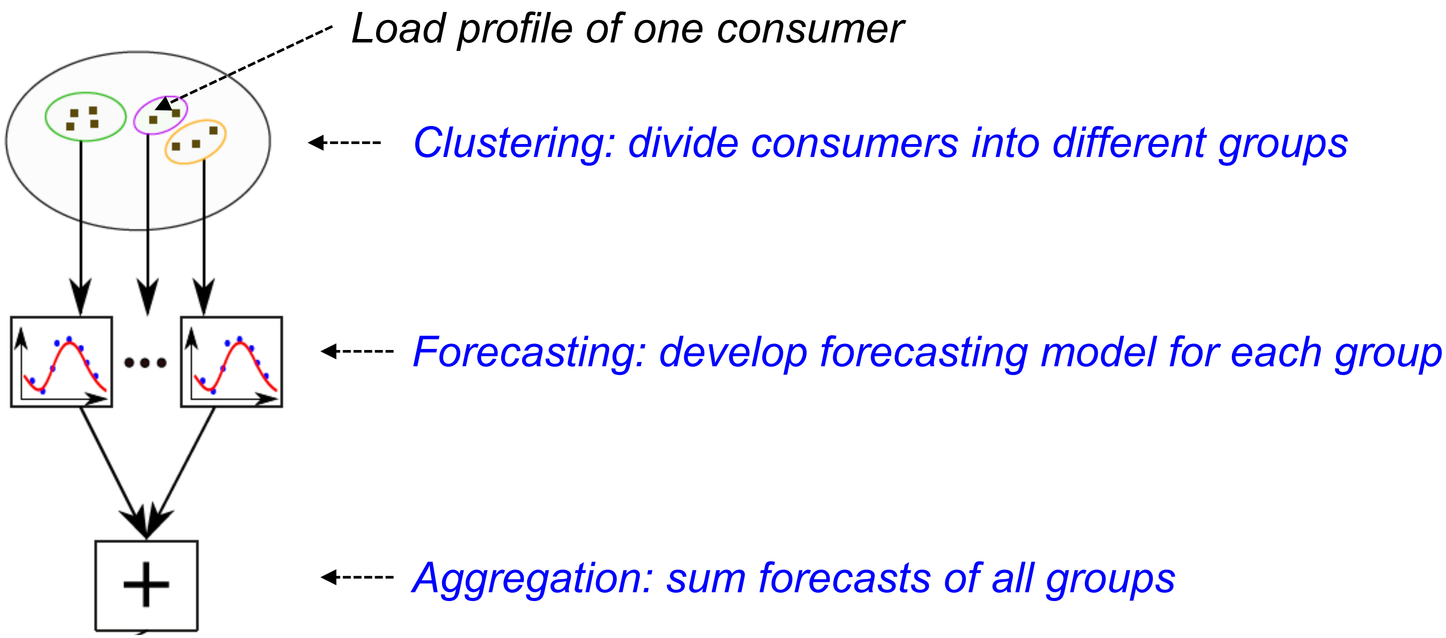
Is it possible to utilize both ensemble techniques and fine-grained subprofiles to further improve the aggregated load forecasting accuracy?



Aggregated Load Forecasting

➢ Introduction

Primary idea: instead of treating the aggregated load as a whole, partitioning consumers into several groups and making predictions might help improve load forecasting.
A three-stage approach for aggregated load forecasting with smart meter data:



Aggregated Load Forecasting

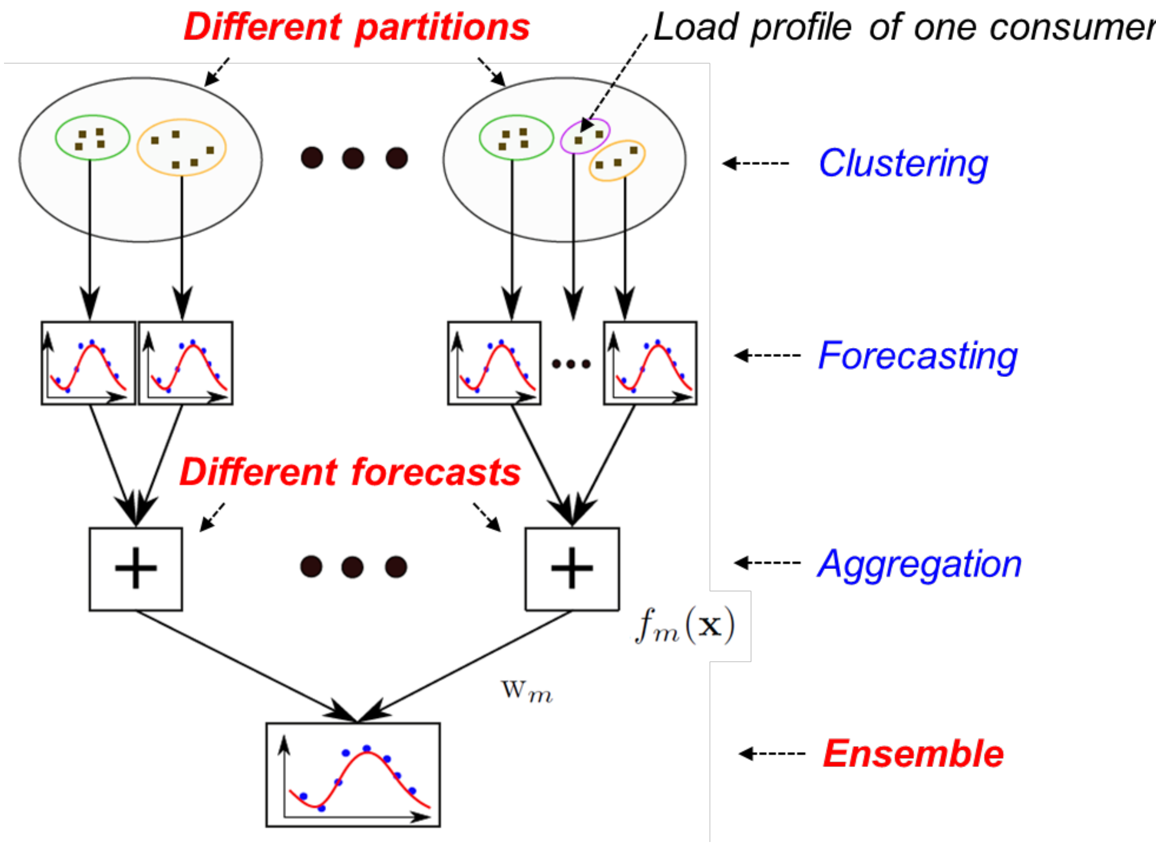
➢ Introduction

Go further steps by ensemble learning?

Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting

If there are different partitions of consumers, we can obtain different load forecasts.



$$G(\mathbf{x}) = \sum_{m=1}^{M} \mathrm{w}_m f_m(\mathbf{x})$$

Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting

How much weight should be given to each method for the optimal combination?

$$\min_{\mathbf{w}} \quad L(y - G(\mathbf{x}))$$
$$\text{s.t.} \quad \sum_{m=1}^{M} \mathrm{w}_m = 1, \quad \mathrm{w}_m \geq 0, \quad m = 1, ..., M$$

Real load **The *n*-th predicted load**

$$\hat{\omega} = \arg\min_{\boldsymbol{\omega}} \sum_{t=1}^{T} \frac{1}{T} \frac{|L_{en,t} - \hat{L}_{en,t}|}{L_{en,t}}$$

Minimize MAPE

$$s.t. \; \hat{L}_{en,t} = \sum_{n=1}^{N} \omega_n \hat{L}_{en,n,t}, \quad \sum_{n=1}^{N} \omega_n = 1, \quad \omega_n \geq 0.$$

It can be formulated as an LP problem.

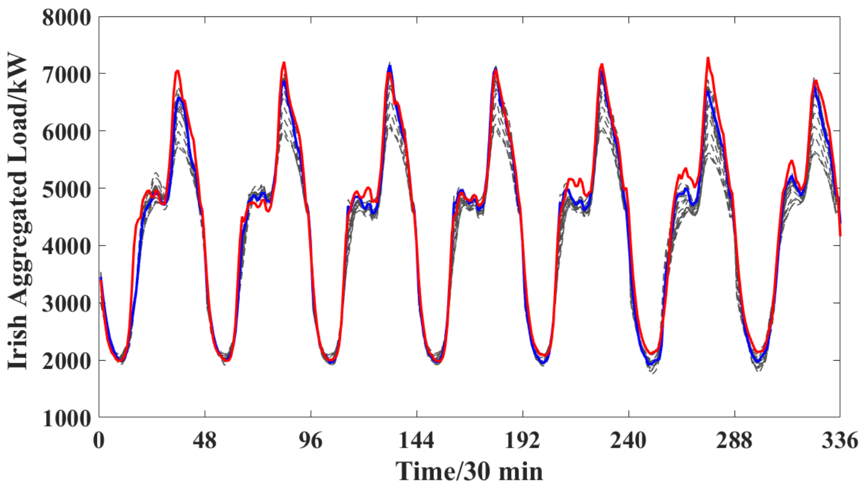
To determine the weights for the forecasts

Aggregated Load Forecasting

➤ Deterministic Aggregated Load Forecasting

Weights, MAPE, and RMSE of different forecasts with different groups

N	1	2	4	8	16	32	64	128	256	...	5237	Ensemble
ω	0.634	0	0	0.271	0	0	0.095	0	0	...	0	/
MAPE	***4.25%***	5.05%	5.29%	4.74%	5.55%	4.66%	4.79%	5.09%	5.59%	...	10.31%	**4.05%**
RMSE	***210.95***	229.73	228.01	217.68	244.9	217.64	227.36	232.61	250.27	...	441.33	**202.88**

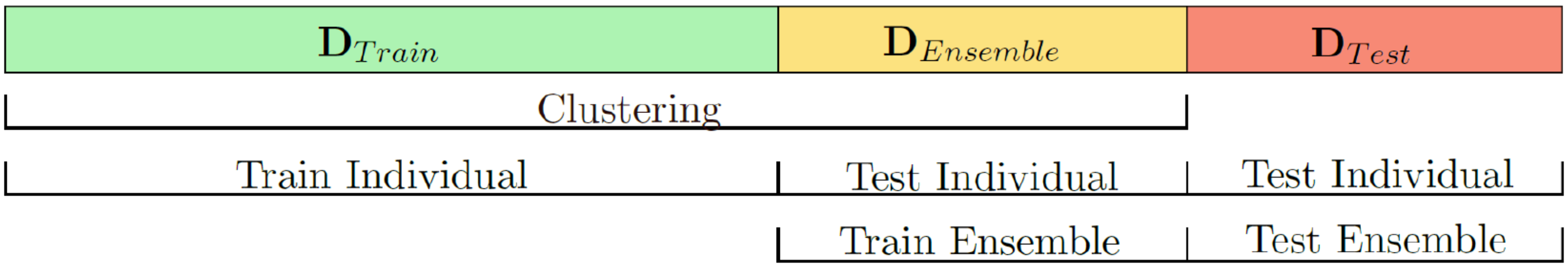


red line: actual load blue line: ensemble forecast
dashed lines: individual forecasts

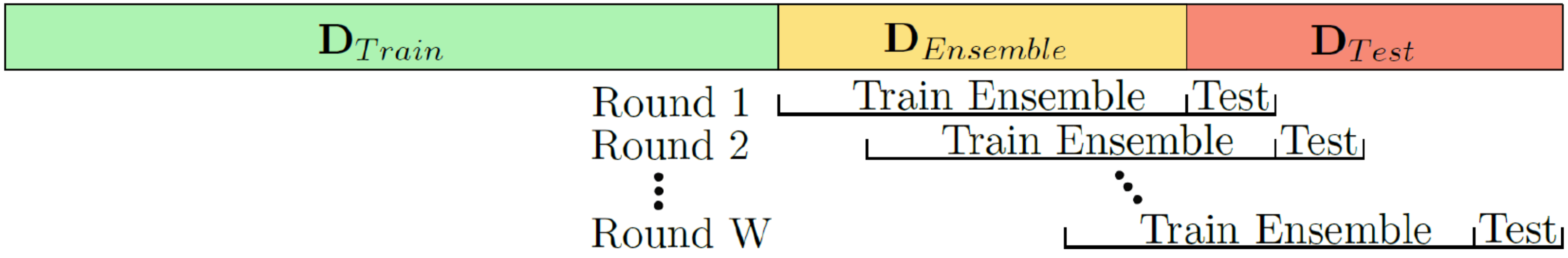
The MAPE and RMSE of the proposed ensemble method are 4.05% and 202.88 which gain 4.71% and 3.83% improvements, respectively compared with the best individual forecast.

Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting



Can we update the weights in a rolling window-based manner?

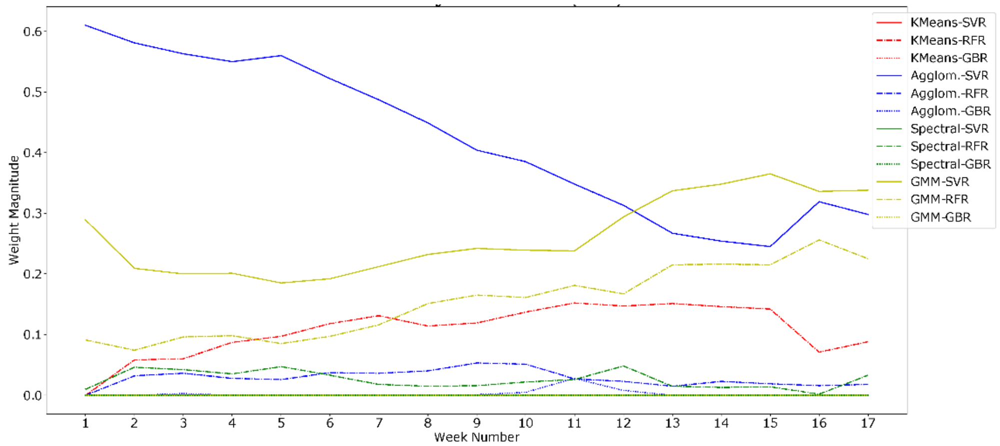


Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting

Benefits of window-based method

Ensemble Method	Error Metrics	Window	Benchmark
COP_{MAPE}	MAPE	2.85%	3.13%
	MAE	106.13	116.66
	RMSE	149.81	166.74
COP_{MSE}	MAPE	2.89%	3.15%
	MAE	107.3	116.8
	RMSE	151.26	166.92



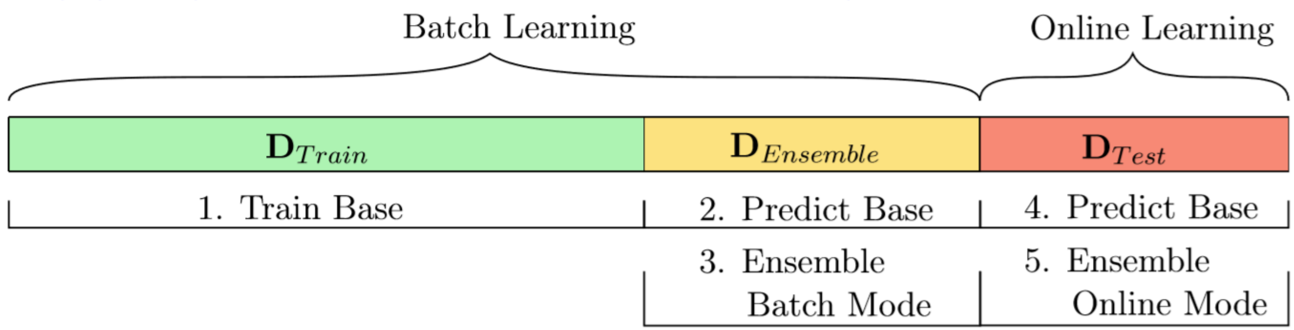
Ensemble weights over 17 weeks of the test set for all individual models.

Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting

Combined model:

$$G(\mathbf{x}) = \sum_{m=1}^{M} \mathrm{w}_m f_m(\mathbf{x})$$



Algorithm 1: Online Protocol

input: Initial model weights $\mathbf{w}_1 \in \mathbb{R}^M$,
convex loss function ℓ, weight update rule U

for $t = 1, 2, \ldots$
- Calculate individual predictions $\mathbf{f}_t \in \mathbb{R}^M$
- Predict $\hat{y}_t = \mathbf{w}_t \cdot \mathbf{f}_t$
- Reveal true value $y_t \in \mathbb{R}$
- Calculate loss $\ell(y_t, \hat{y}_t)$
- Update model $\mathbf{w}_{t+1} = U\big(\mathbf{w}_t; \ell(y_t, \hat{y}_t)\big)$

end

Online Convex Optimization (OCO) is a unifying framework for the analysis and design of online algorithms.

Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting

- **General formula**

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}} \left[\ d(\mathbf{w}, \mathbf{w}_t) + \eta_t \ell(y_t, \mathbf{w} \cdot \mathbf{x}_t)\ \right]$$

Distance d
Prevent information loss

Loss ℓ
Integrate new sample

Passive Aggressive Regression

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}} \left[\|\mathbf{w} - \mathbf{w}_t\|_1 + \ell_\varepsilon(y_t, \mathbf{w} \cdot \mathbf{f}_t) + \lambda \|\mathbf{w}\|_1 \right]$$

Aggressive: weights change if losses are big enough

$$\ell_\varepsilon(y_t, \mathbf{w} \cdot \mathbf{f}) = \begin{cases} 0 & \text{if } |y - \mathbf{w} \cdot \mathbf{f}| \leq \varepsilon \\ |y - \mathbf{w} \cdot \mathbf{f}| & \text{otherwise} \end{cases}$$

Passive: weights do not change every time slot

Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting

Update the weights online for a better performance

Errors on test set after online learning

Method	MAPE	SD	MAE	RMSE
SGDR	2.43%	0.025	86.05	122.71
FTRLP	2.23%	0.021	81.09	113.87
OSELM	2.80%	0.029	106.03	155.03
Online Bagging	2.07%	0.021	74.33	106.23
PAR	1.67%	0.015	61.83	86.68
Proposed	1.62%	0.014	59.59	83.21
Best SVR	3.18%	0.032	117.54	171.72
Best RF	2.89%	0.029	108.25	156.84
Best GBRT	3.53%	0.032	127.81	175.78
Batch OPT	2.89%	0.028	107.55	154.88
Window OPT	2.85%	0.028	106.13	149.81

SD: Standard deviation of the absolute percentage error

- ➢ All ensembles improve their forecasting performance through online learning.
- ➢ Nearly all ensembles outperform the benchmarks after online learning.
- ➢ The proposed method has the highest accuracy and stability among all examined ensembles.

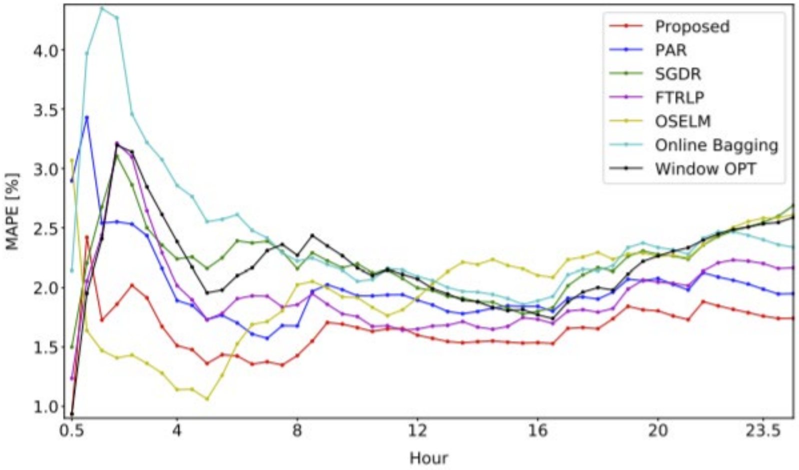
Aggregated Load Forecasting

➢ Deterministic Aggregated Load Forecasting

Update the weights online for a better performance

The hour of break-even for all ensembles

Method	Break-even [hour]			
	MAPE	SD	MAE	RMSE
SGDR	39.5	86.5	41.0	64.0
FTRLP	66.5	87.0	64.0	60.5
PAR	17.5	9.0	19.5	17.5
OSELM	112.0	2.0	2833.5	no
Online Bagging	22.5	4.5	23.0	35.5
Proposed	1.5	2.0	1.5	1.5

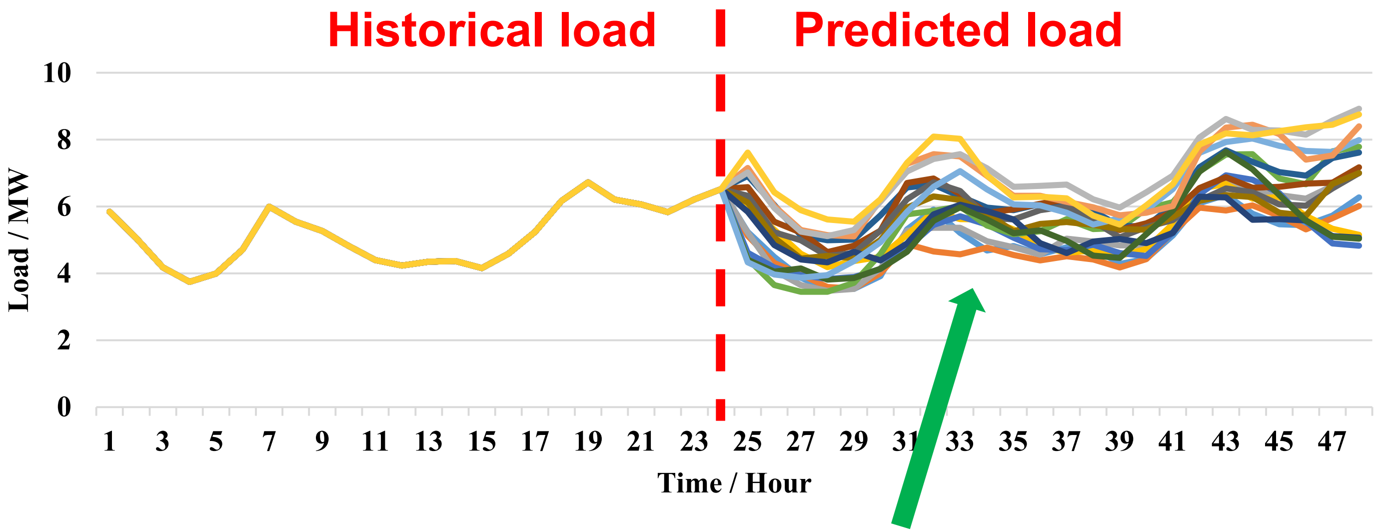


MAPE over the course of the first day of online learning

- ➢ The proposed method has the earliest break-even after 2 hours for all metrics.
- ➢ The other ensembles have the break-even approximately within one or two days.
- ➢ An ensemble employing online learning is able to pay off at a relatively early point in time.

Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting



Compared with deterministic forecasting, probabilistic load forecasts provide comprehensive information about future uncertainties.

Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

Pinball loss (PL) and **Winkler Score (WS)** assess the calibration and sharpness simultaneously.

$$\mathrm{PL}\left(\hat{y}_{t,q}, y_t\right) = \begin{cases} \left(y_t - \hat{y}_{t,q}\right) q & \hat{y}_{t,q} \le y_t \\ \left(\hat{y}_{t,q} - y_t\right)(1-q) & \hat{y}_{t,q} > y_t \end{cases}$$

Performance of overall quantiles

$$\mathrm{WS}(L_t, U_t, y_t) = \begin{cases} \delta_t + 2(L_t - y_t)/\alpha & y_t \le L_t \\ \delta_t & L_t < y_t < U_t \\ \delta_t + 2(y_t - U_t)/\alpha & U_t \le y_t \end{cases}$$

Performance of extreme quantiles

Average Coverage Error (ACE) evaluate the reliability of the forecasts.

$$ACE = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\{y_i \in [L_i, U_i]\}} - (1 - \alpha)$$

Performance of an certain interval

Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

$$\mathbf{x}_i = [1, \hat{y}_{1,i}, \ldots, \hat{y}_{K,i}] \qquad i \in [1, \ldots, n]$$

PCA

$$\hat{\mathbf{w}}_q = \underset{\mathbf{w}_q}{\arg\min} \sum_{i=1}^{n} \rho_q(y_i - \mathbf{z}_i \mathbf{w}_q)$$

Factor Quantile Regression Averaging

$$\hat{\mathbf{w}}_q = \underset{\mathbf{w}_q}{\arg\min} \sum_{i=1}^{n} \rho_q(y_i - \mathbf{x}_i \mathbf{w}_q)$$

Quantile regression averaging (QRA), a special form of quantile regression, is a kind of model averaging method.

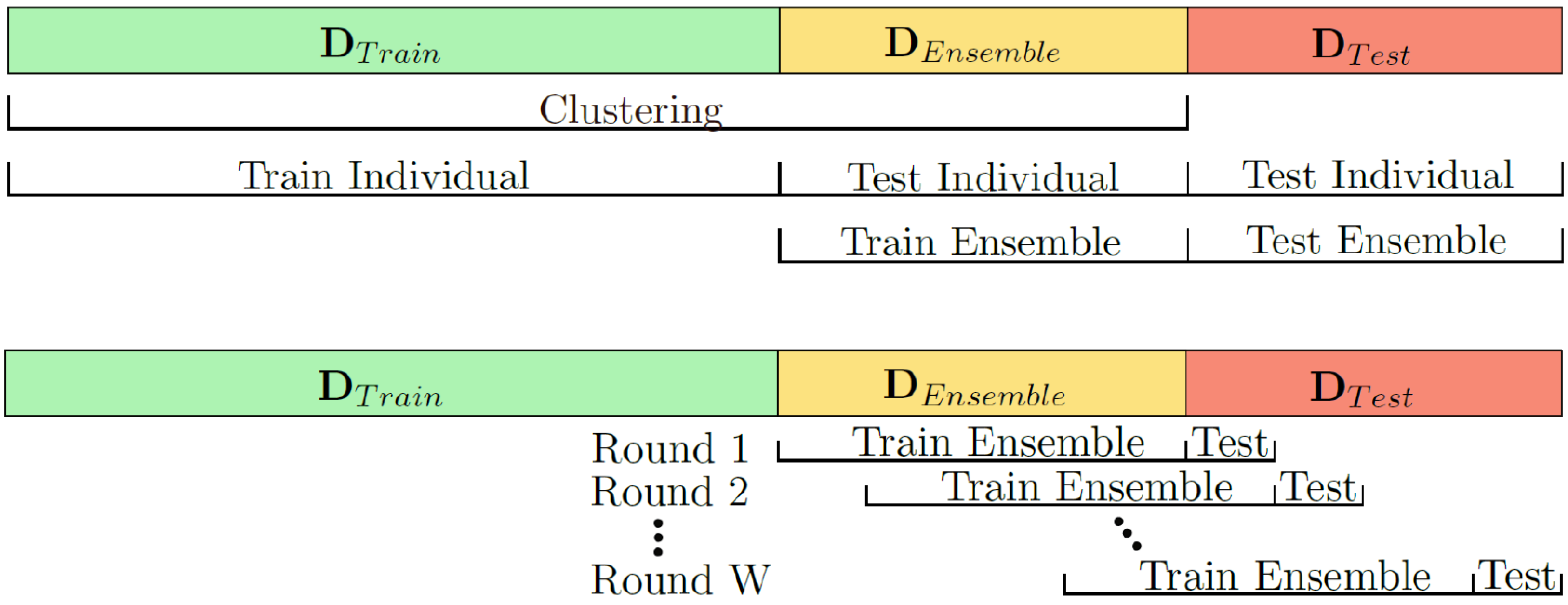
$$\hat{\mathbf{w}}_q = \underset{\mathbf{w}_q}{\arg\min} \sum_{i=1}^{n} \rho_q(y_i - \mathbf{x}_i \mathbf{w}_q) + \lambda \|\mathbf{w}_q\|_1$$

LASSO Quantile Regression Averaging

Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

Similar to deterministic forecasting……



Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

Error metric comparison for all ensemble methods with a Prediction Interval of 90%.

Ensemble Method	Error Metrics	Offline Ensemble	Benchmark 1	Rolling Window-based Ensemble	Benchmark 2
QRA	ACE	-1.73%	-1.85%	-0.56%	-0.92%
	PBL	45.82	50.19	42.28	46.52
	WKS	788.62	846.89	728.13	791.78
FQRA	ACE	-1.80%	-1.85%	**-0.45%**	-0.92%
	PBL	45.82	50.19	**42.26**	46.52
	WKS	787.26	846.89	**727.24**	791.77
LQRA	ACE	**-1.71%**	-1.83%	-0.63%	-0.98%
	PBL	**45.84**	50.2	42.26	46.53
	WKS	**785.77**	845.7	724.74	791.55

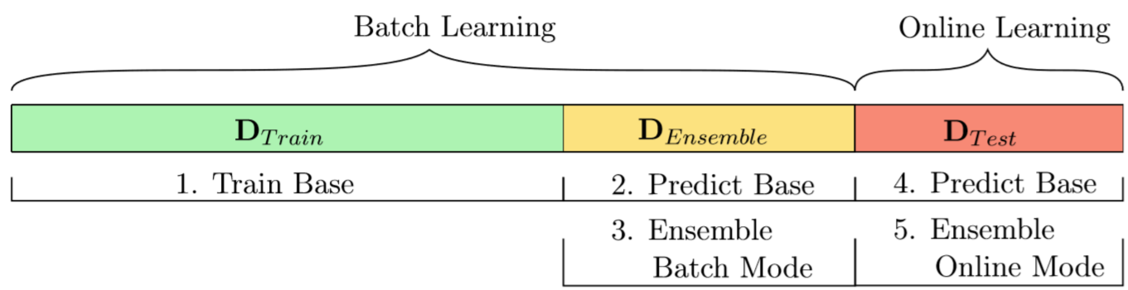
- ➢ The two naive benchmarks are obtained by directly forecasting the total loads without dimension reduction and clustering.
- ➢ Benchmark 2 updates the weights in a rolling window-based approach, while Benchmark 1 does not.

Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

Combined model:

$$G(\mathbf{x}) = \sum_{m=1}^{M} \mathrm{w}_m f_m(\mathbf{x})$$



- **General formula**

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}} \left[\ d(\mathbf{w}, \mathbf{w}_t) + \eta_t \ell(y_t, \mathbf{w} \cdot \mathbf{x}_t)\ \right]$$

Distance d
Prevent information loss

Loss ℓ
Integrate new sample

Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

- General Formula

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}} \left[\ d(\mathbf{w}, \mathbf{w}_t) + \eta_t \ell(y_t, \mathbf{w} \cdot \mathbf{x}_t)\ \right]$$

- L_2-distance :

$$d(\cdot) = \frac{1}{2} \|\cdot\|^2$$

- ε-insensitive quantile loss :

$$\ell_{\varepsilon,q}(\mathbf{w}_q; \mathbf{x}, y) = \begin{cases} q(y - \mathbf{w}_q \cdot \mathbf{x} + \varepsilon(q-1)) & \text{if } y - \mathbf{w}_q \cdot \mathbf{x} > \varepsilon(1-q) \\ 0 & \text{if } -\varepsilon q < y - \mathbf{w}_q \cdot \mathbf{x} < \varepsilon(1-q) \\ (q-1)(y - \mathbf{w}_q \cdot \mathbf{x} + \varepsilon q) & \text{if } y - \mathbf{w}_q \cdot \mathbf{x} < -\varepsilon q \end{cases}$$

- Solving KKT conditions:

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \eta_t sign(y_t - \mathbf{w}_t \cdot \mathbf{x}_t)\tau_t \mathbf{x}_t \qquad \tau_t = min\left\{C, \frac{\ell_{\varepsilon,q}(y_t, \mathbf{w}_t \cdot \mathbf{x}_t)}{q\,\|\mathbf{x}_t\|_2^2}\right\}$$

Aggregated Load Forecasting

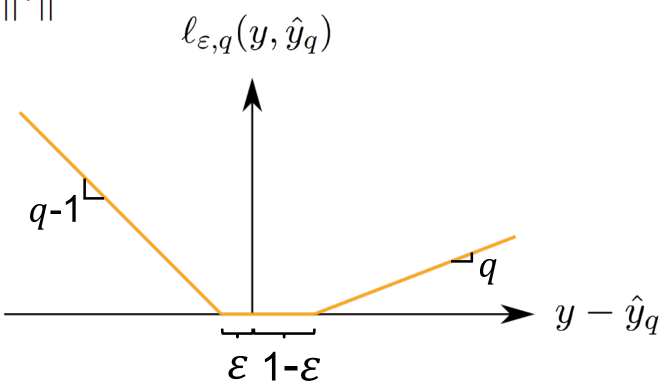
➢ Probabilistic Aggregated Load Forecasting

- General Formula

$$\mathbf{w}_{t+1} = \arg\min_{\mathbf{w}} \left[\, d(\mathbf{w}, \mathbf{w}_t) + \eta_t \ell(y_t, \mathbf{w} \cdot \mathbf{x}_t) \,\right]$$

- L_2-distance : $d(\cdot) = \frac{1}{2} \|\cdot\|^2$

- ε-insensitive quantile loss :



- Solving KKT conditions:

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \eta_t sign(y_t - \mathbf{w}_t \cdot \mathbf{x}_t) \tau_t \mathbf{x}_t \qquad \tau_t = min\left\{ C, \frac{\ell_{\varepsilon,q}(y_t, \mathbf{w}_t \cdot \mathbf{x}_t)}{q \|\mathbf{x}_t\|_2^2} \right\}$$

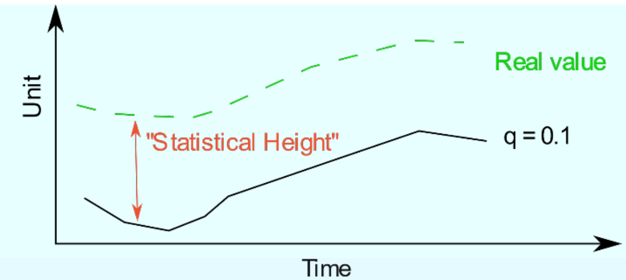
Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

Mechanism of Quantile Passive Aggressive Regression

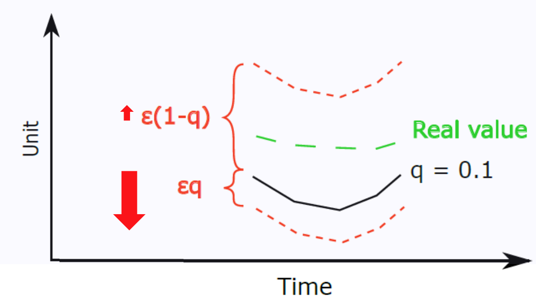
- Extension to probabilistic forecasting: **ε-insensitive loss -> ε-insensitive quantile loss**
- ε-insensitive region: Preserve «quantile height» between y_q and y

- Batch quantile regression
  - Access to whole data sequence
  - «Statistical height» implicitly given



$$\ell_q(y, \hat{y}_q) = \begin{cases} q(y - \hat{y}_q) & \text{if } y \geq \hat{y}_q \\ (q-1)(y - \hat{y}_q) & \text{if } y < \hat{y}_q \end{cases}$$

- Online quantile regression
  - Only access to one sample per round
  - «Statistical height» collapses → Real value
- ε-insensitive quantile: Preserve «statistical height»



Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

The performance on Irish load data

Errors on test set after batch learning

Method	ACE	PBL	WKS
QSGD	-0.92%	51.60	722.43
QPAR	2.23%	47.61	1075.02
QNN	-2.55%	54.94	776.86
Batch QRA	-5.25%	44.55	734.64
Window QRA	-1.90%	40.30	659.94

Errors on test set after online learning

Method	ACE	PBL	WKS
QSGD	-0.02%	30.04	527.94
QPAR	-1.69%	29.47	484.59
QNN	-0.64%	56.10	930.23
Batch QRA	-5.25%	44.55	734.64
Window QRA	-1.90%	40.30	659.94

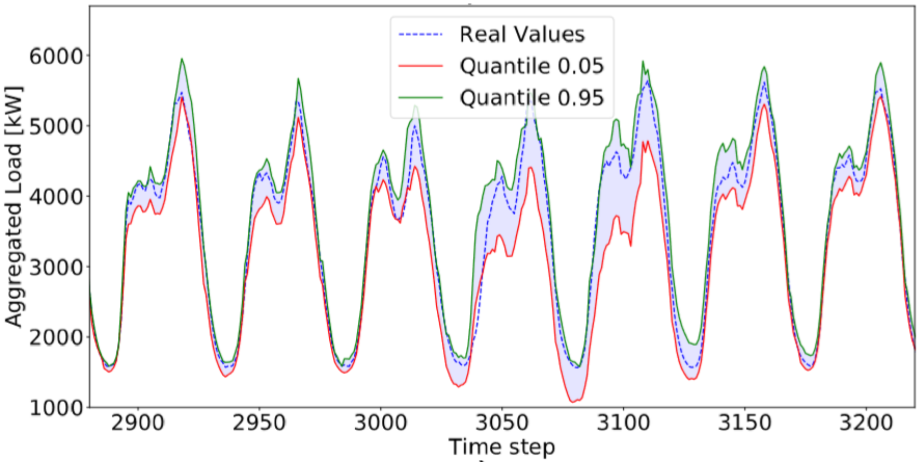
*QSGD: Quantile Stochastic Gradient Descent
*QPAR: Quantile Passive Aggressive Regression
*QNN: Quantile Neural Network

*Window OPT: window-based optimization

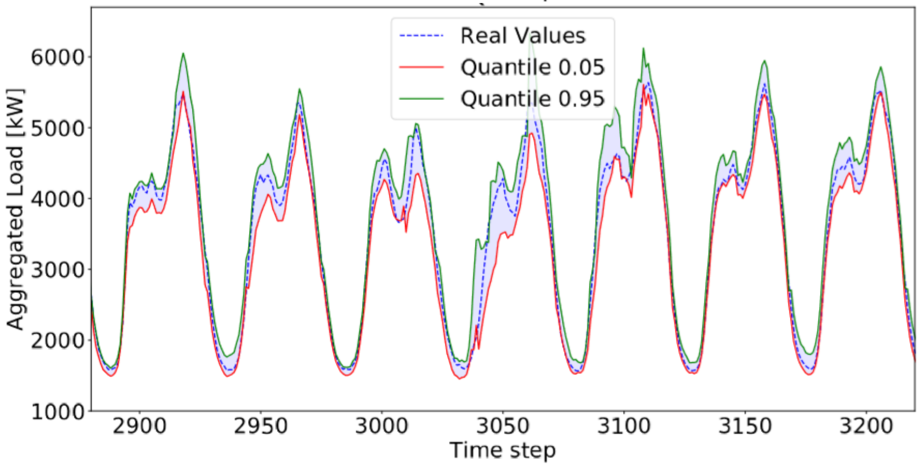
- All ensembles outperform the benchmarks after online learning except QNN
- The proposed method has the highest accuracy regarding pinball loss and winkler score
- A substantial performance improvement can be achieved by ensembles incorporating online learning.

Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting



QSGD online forecast over one week



QPAR online forecast over one week

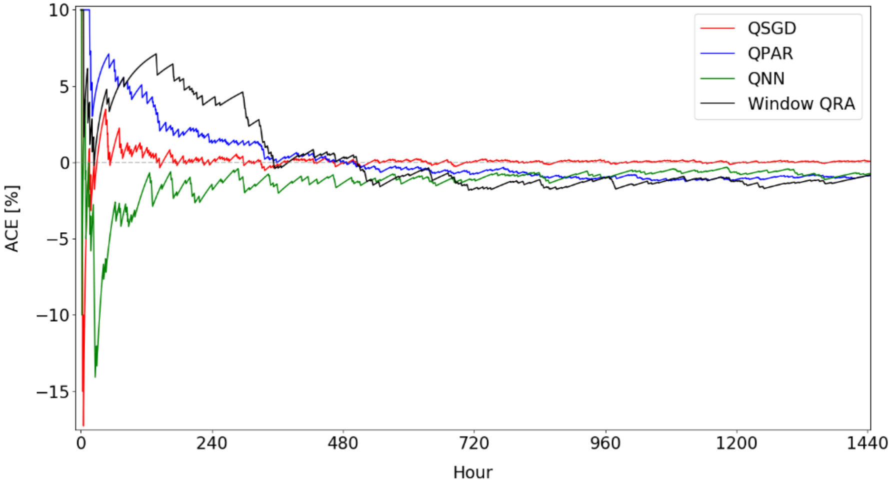
Aggregated Load Forecasting

➢ Probabilistic Aggregated Load Forecasting

The performance on Irish load data

The hour of break-even for all ensembles

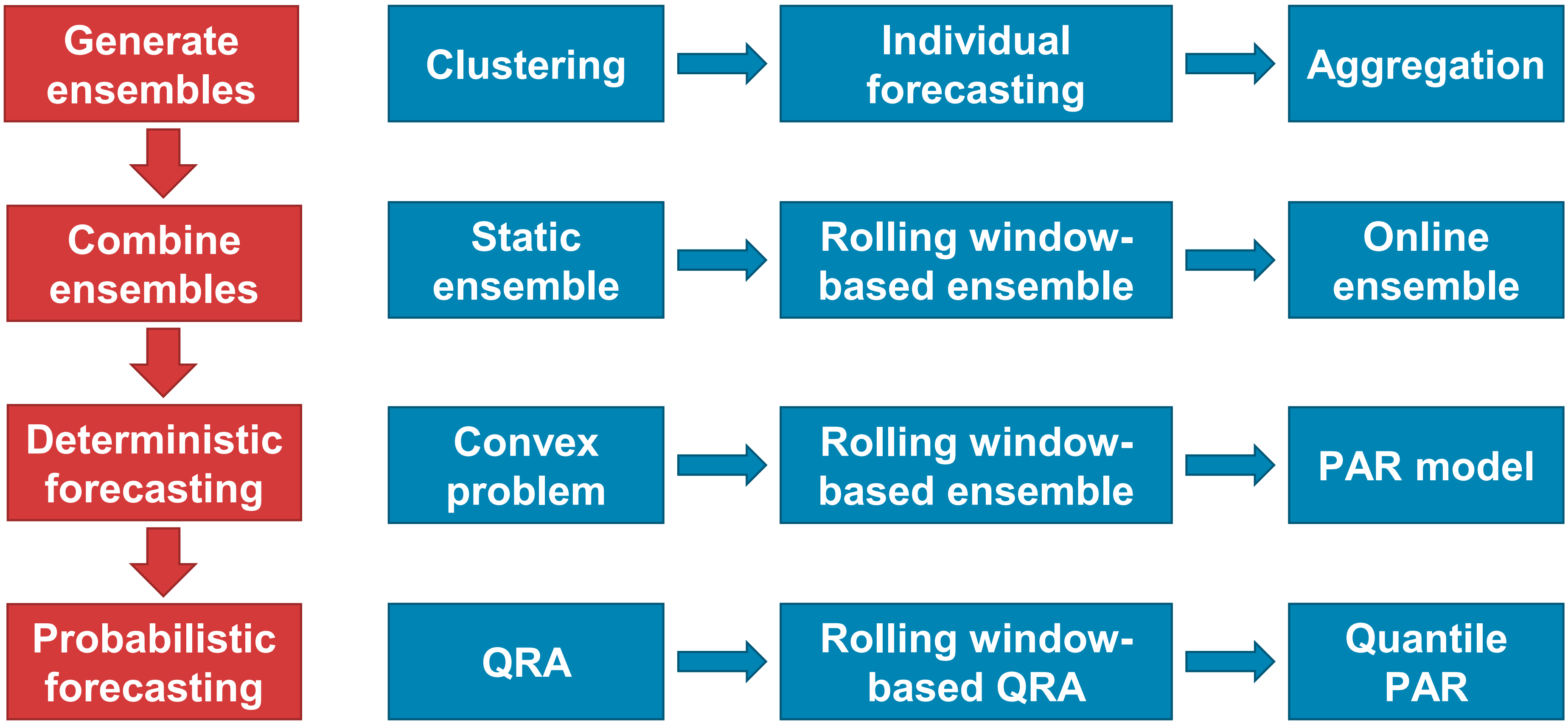
Method	Break-Even ACE	Break-Even PBL	Break-Even WKS
QSGD	508.0 h	35.0 h	307.0 h
QPAR	2810.0 h	138.5 h	253.5 h
QNN	687.0 h	no	no



- ➢ The proposed QPAR has earliest WKS break-even
- ➢ QSGD has earliest Break-even for ACE and PBL
- ➢ Online learning enables to outperform batch approach within a month.

Aggregated Load Forecasting

➢ Short Summary



Aggregated Load Forecasting

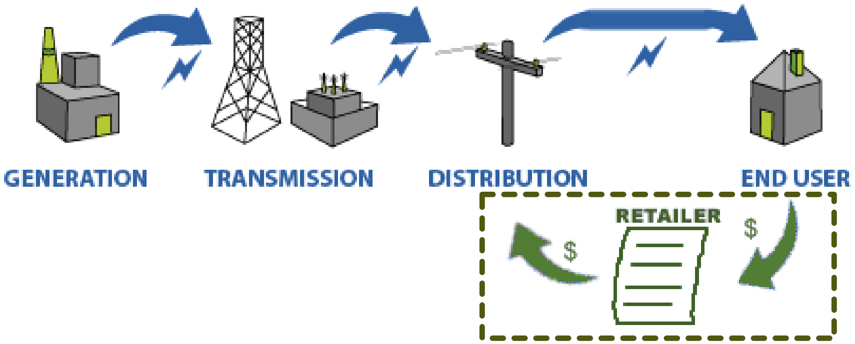
➢ Short Summary

- ❑ High quality point forecasting can be generated by making full use of the fine grained smart meter data;
- ❑ On this basis, we can utilize ensemble techniques to further improve the forecasting accuracy;
- ❑ Online learning can be a powerful tool in short-term load forecasting by integration new information and the proposed modified PAR model is very suitable in this context, especially as an online ensemble method;
- ❑ PAR model can be further extend to quantile PAR model using quantile regression averaging for probabilistic forecasting.

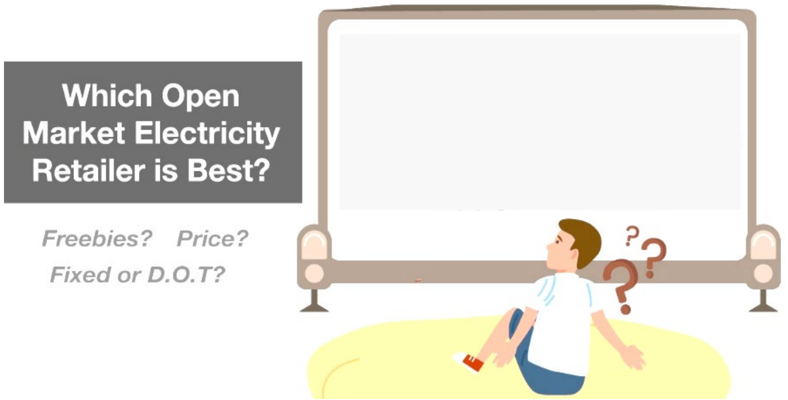
Personalized Retail Price Design

➢ Introduction

➢ The opening of electricity retailing market



➢ The need for diversified service



➢ Consumers choose freely in market

How to provide **diversified services** for different consumers to enhance the competitiveness of the retailers?

ETH zürich

Personalized Retail Price Design

Main Idea

- ***Data-driven price design.*** Smart meter data contains great value which may help retailing price design.
- ***Respect self-selection.*** Consumers' willingness and rights to choose must be respected.

Challenges

- Diversified service
- Mine consumers' inner need
- Satisfying consumers
- Self-selection in a real market
- Proper incentive

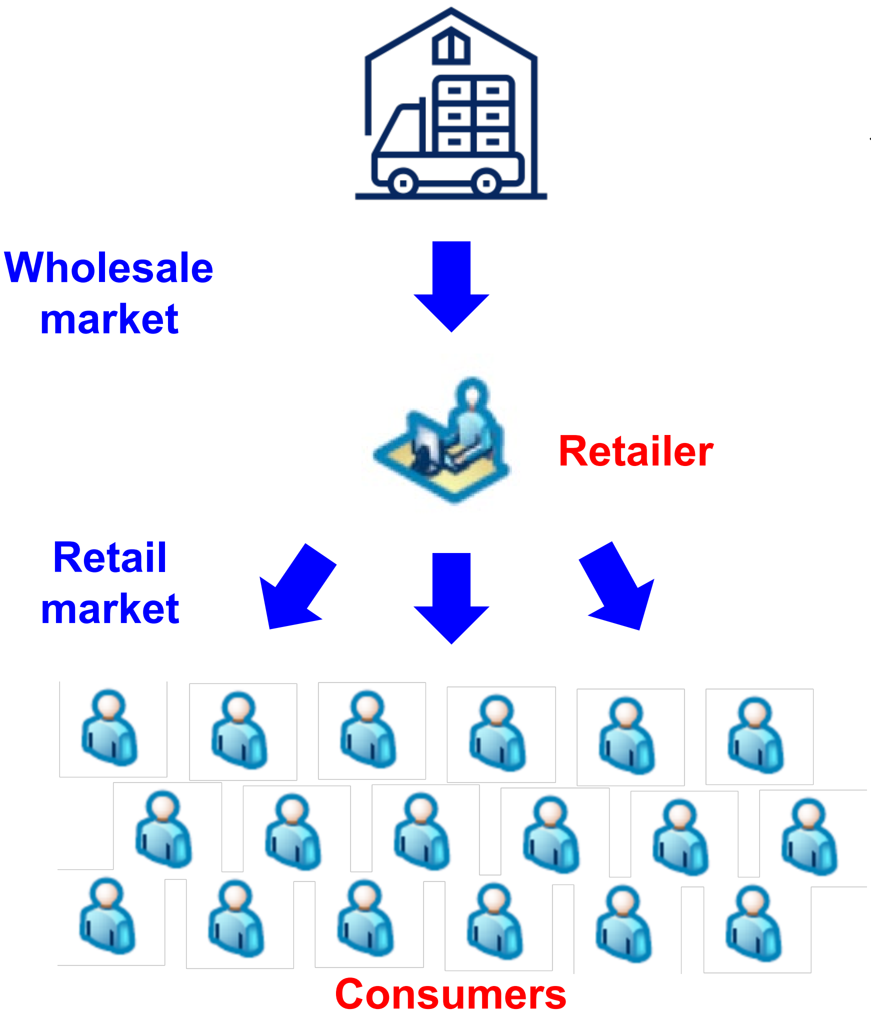
Data-driven price design

Compatible incentive design

Discover utility from data → Cluster load profiling data → Correlate preference with shape → Make centroids representatives → Form optimization problem

ETH zürich

Personalized Retail Price Design



❖ **A Stackelberg game**

Leader——Retailer

- Design pricing schemes
- Predict consumer behaviors

Followers——Consumers

- Choose one pricing scheme
- Adapt electricity consumption

Power
Systems
Laboratory

ETH zürich

Personalized Retail Price Design

Problem formulation - Consumer

Consumer Utility

- Measure satisfaction
- Comparison between different plans
- Diminishing marginal utility

$$F(\boldsymbol{p},\boldsymbol{q}) = u(\boldsymbol{q}) - \sum_{t=1}^{T} p_t q_t$$

Consumer Strategy

- Strategic and rational consumers:

Utility Maximization

$$\boldsymbol{q}^*(\boldsymbol{p}) = \arg\max_{\boldsymbol{q}} \{F(\boldsymbol{p},\boldsymbol{q})\}$$

$$U(\boldsymbol{p}) = \max_{\boldsymbol{q}} \{F(\boldsymbol{p},\boldsymbol{q})\} = F(\boldsymbol{p},\boldsymbol{q}^*(\boldsymbol{p}))$$

How can smart meter data be useful?

$$F(\boldsymbol{p}_{(0)},\boldsymbol{q}_{(0)}) = 0 \qquad \frac{\partial F(\boldsymbol{p}_{(0)},\boldsymbol{q}_{(0)})}{\partial q_t} = 0, \quad \forall t$$

Original electricity consumption is the realization of Utility Maximization!

ETH zürich

Personalized Retail Price Design

Problem formulation - incentive

Individual rationality

If the retailer wants consumer k to choose pricing scheme r, the retailer **must guarantee choosing r is consumer k's dominant strategy**

$$U_k(\boldsymbol{p}_r) \geq U_k(\boldsymbol{p}') \quad \forall k$$

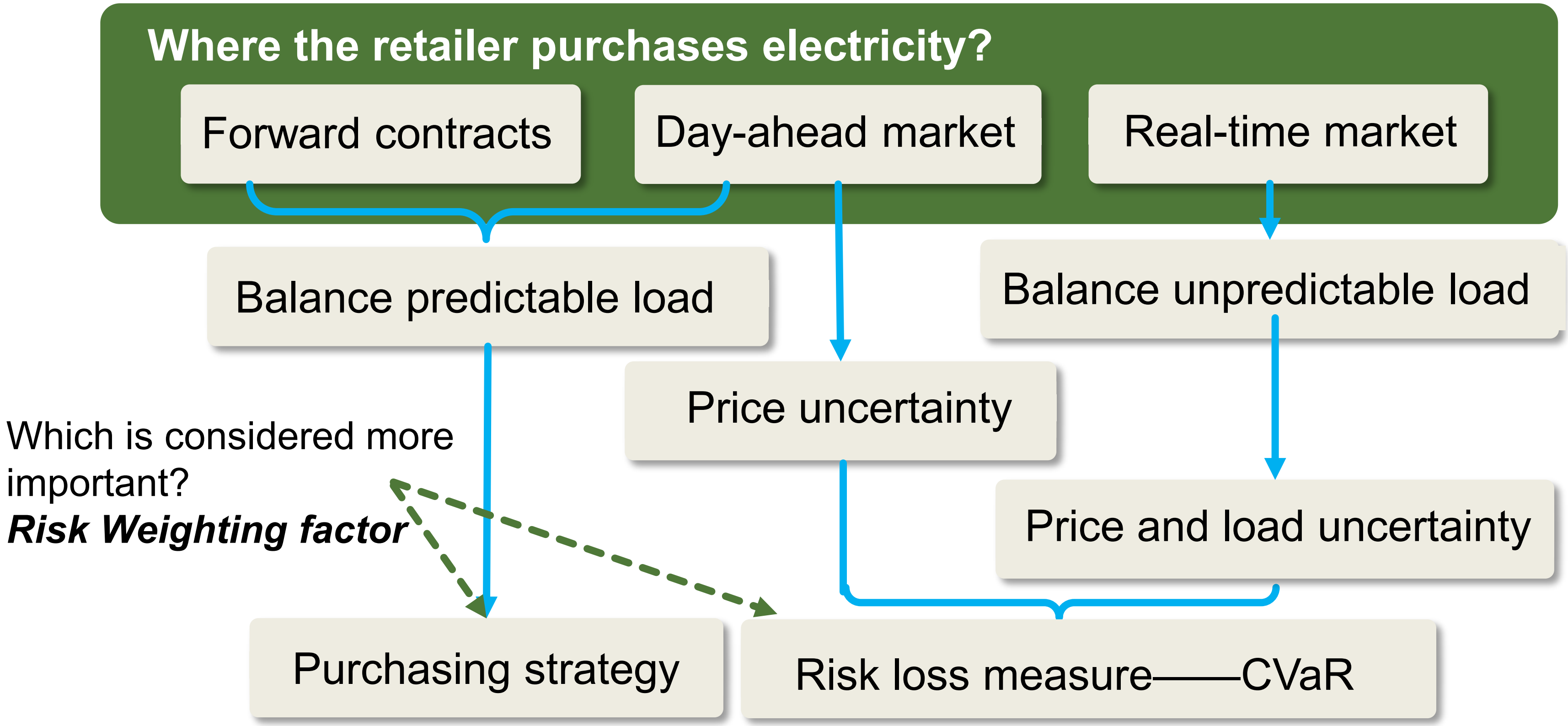
Compatible incentive

If the retailer wants consumer k to choose new pricing scheme r, the retailer **must guarantee choosing r is at least as good as previous situation**

$$U_k(\boldsymbol{p}_r) \geq U_k(\boldsymbol{p}_0) \quad \forall k$$

Personalized Retail Price Design

Problem formulation - Retailer



ETH zürich

Personalized Retail Price Design

- Electric Reliability Council of Texas (ERCOT)
- Extreme Cold Scenarios
- Rotating Outages
- Extreme-high Price !!!
- Price Uncertainty
- CVaR

Oper Day	Interval Ending	HB_BUSAVG	HB_HOUSTON	HB_HUBAVG	HB_NORTH	HB_PAN	HB_SOUTH
02/19/2021	0515	8964.32	8968.28	8967.64	8957.56	8967.52	8975.67
02/19/2021	0530	8963.28	8968.24	8966.67	8956.54	8966.46	8973.94
02/19/2021	0545	8962.84	8967.35	8966.32	8955.76	8965.93	8974.62
02/19/2021	0600	8964.70	8968.24	8967.99	8957.93	8967.91	8976.33
02/19/2021	0615	8963.11	8968.24	8966.52	8956.37	8966.29	8973.66
02/19/2021	0630	8963.13	8968.26	8966.53	8956.39	8966.31	8973.68
02/19/2021	0645	8964.98	8970.52	8967.60	8959.24	8964.72	8975.13
02/19/2021	0700	8966.01	8971.77	8968.21	8960.81	8963.90	8975.94
02/19/2021	0715	8968.81	8971.73	8970.81	8963.56	8966.76	8980.70
02/19/2021	0730	8965.21	8971.16	8967.20	8960.24	8961.66	8975.22
02/19/2021	0745	8965.18	8968.07	8966.64	8961.41	8963.00	8973.38
02/19/2021	0800	8960.38	8967.47	8962.38	8954.77	8953.21	8972.88
02/19/2021	0815	8989.40	8990.41	8988.92	8989.58	8984.42	8990.91
02/19/2021	0830	8977.41	8981.32	8978.40	8974.20	8970.94	8985.44
02/19/2021	0845	8987.87	8991.29	8986.79	8987.26	8975.92	8995.41
02/19/2021	0900	8987.12	8990.86	8985.60	8986.93	8971.74	8995.02
[illegible]	0915	3206.30	[illegible]	3205.68	3205.44	3198.12	3212.74
[illegible]	0930	35.61	37.85	35.44	34.39	31.05	41.33
[illegible]	0945	36.46	38.70	36.29	[illegible]	[illegible]	42.19
02/19/2021	[illegible]	27.05	27.84	[illegible]	26.62	25.45	29.06
02/19/2021	1015	[illegible]	[illegible]	27.54	23.59	19.61	44.55
02/19/2021	1030	25.78	27.65	26.02	22.41	18.40	38.85
02/19/2021	1045	23.35	24.92	23.52	20.44	16.86	34.86
02/19/2021	1100	26.18	28.53	26.01	24.83	22.90	32.52
02/19/2021	1115	25.17	28.25	25.02	23.29	20.80	33.49
02/19/2021	1130	22.49	24.31	22.16	21.38	20.29	28.65
02/19/2021	1145	21.58	23.06	21.12	21.28	21.05	24.99
02/19/2021	1200	18.64	19.78	18.16	18.62	18.77	21.02

100 times

https://www.dallasnews.com/opinion/commentary/2021/02/20/dont-just-blame-ercot-what-caused-outages-is-our-competitive-electricity-market/

PSL Power Systems Laboratory

ETH zürich

Personalized Retail Price Design

Problem formulation - Clustering

Different Clustering Methods

- Hierarchical Clustering
- *K*-means
- Fuzzy C-means
- Gaussian mixture

One method may not fit all data sets

Clustering evaluation

Davies Bouldin Index

With-cluster compactness

Between-cluster seperation

Centroid as representative

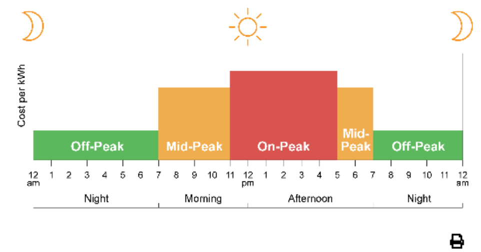
Personalized Retail Price Design

Problem formulation – Optimization framework

Optimization framework – an MINLP model

- Objective: Retailing profit maximization
- Constraints:
  - Load balance
  - Consumer reaction
  - Compatible incentive
  - Risk measure CVaR
  - Price structure: Various choices

- Price category: CPP RTP ToU



Lower Risk
Less changes

3. Personalized Retail Price Design

Optimization framework – an MINLP model

- **Objective**: Retailing profit maximization

$$\max\ R = \sum_{r=1}^{R}\sum_{t=1}^{T} \underbrace{K_r \times p_{k,t} \times q_{k,t}}_{\text{Consumer payment}} - \sum_{t=1}^{T}\sum_{n=1}^{N_F} \underbrace{p_n^F \times L_n^F \times o_{n,t}^F \times o_n}_{\text{Forward contracts}} - \sum_{t=1}^{T} \underbrace{p_t^{D,est} \times L_t^D}_{\text{DA}} - \underbrace{\xi \times CVaR}_{\text{Risk Loss in DA \& RT}}$$

- **Constraints**: Load balance

$$\sum_{r=1}^{R} \underbrace{K_r \times q_{r,t}}_{\text{Consumer load}} = \sum_{n=1}^{N_F} \underbrace{L_n^F \times o_{n,t}^F \times o_n}_{\text{Forward contracts}} + \underbrace{L_t^D}_{\text{DA}},\ \forall t$$

DA=Day-ahead market
RT= Real-time market

- **Constraints**: Compatible incentive

$U_r(\boldsymbol{p}_r) \geq U_r(\boldsymbol{p}')\ \ \forall k$ — Choosing $\boldsymbol{p}_r$ is consumer k's dominant strategy, k likes $\boldsymbol{p}_r$ than any other pricing schemes

$U_r(\boldsymbol{p}_r) \geq U_r(\boldsymbol{p}_0)\ \ \forall k$ — Choosing $\boldsymbol{p}_r$ is consumer k's rational choice, k likes $\boldsymbol{p}_r$ than the old pricing schemes

* nonlinear terms are marked in red

3. Personalized Retail Price Design

Optimization framework – an MINLP model

- **Constraints**： Utility and response

$$q_t = \left(\frac{p_t}{p_{t(0)}}\right)^{\frac{1}{\alpha-1}} \times q_{t(0)}$$ Reactions

$$U(\boldsymbol{p}) = \sum_{t=1}^{T}\left(\frac{1}{\alpha}-1\right)\left[\left(\frac{p_t}{p_{t(0)}}\right)^{\frac{\alpha}{\alpha-1}}-1\right]\times q_{t(0)}\, p_{t(0)}$$ Utility

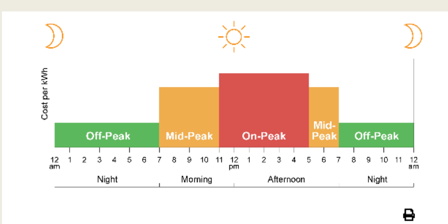
- **Constraints**： Risk measure CVaR

Loss in DA Loss in RT

$$CVaR = \inf_{a\in R}\left\{a + \frac{1}{(1-\alpha^{CVaR})\cdot N_S}\sum_{n_S=1}^{N_S}\left[(-\Delta R^{D} - \Delta R^{RT}) - a\right]^{+}\right\}$$

- **Constraints**： Price structure

Price category： CPP RTP $\boxed{\text{ToU}}$



Lower Risk
Less changes

$$\sum_{m=1}^{M} e_{r,t}^{m} = 1, \qquad \sum_{t=1}^{T} e_{r,t}^{m} \geq D_{\min}, \ \forall m, r$$

$$\left|e_{r,T}^{m} - e_{r,1}^{m}\right| + \sum_{t=2}^{T}\left|e_{r,t-1}^{m} - e_{r,t}^{m}\right| = 2, \ \forall m, r$$

$$p_{r,t} = \sum_{m=1}^{M} e_{r,t}^{m} \times p_{r}^{m}, \quad \forall t, r$$ m block ToU

* nonlinear terms are marked in red

MINLP model $\xrightarrow[\text{Piecewise linear approximation}]{\text{Big M method}}$ **MILP model**

Power Systems Laboratory

Personalized Retail Price Design

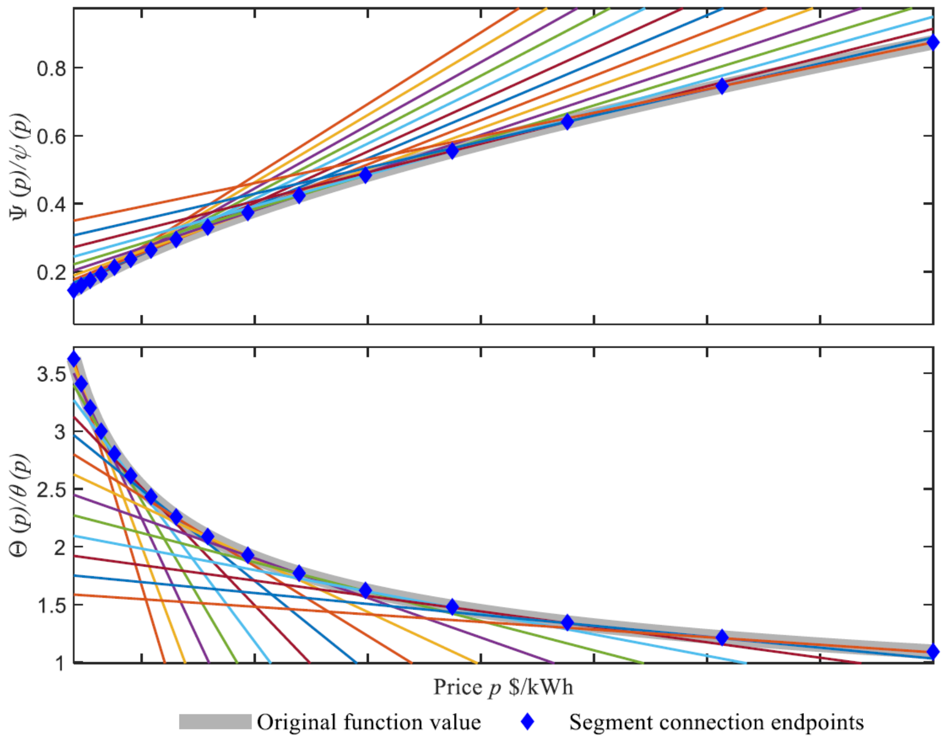
Solution method

Nonlinear model

- Power exponent $p_{r,t}^{\frac{\alpha}{\alpha-1}}$, $p_{r,t}^{\frac{1}{\alpha-1}}$
- Two variables' product $p_{r,t} \times q_{r,t}$

Linear model

- Linear segment approximation
  Take $p_{r,t} \times q_{r,t}$ as a whole



Personalized Retail Price Design

Solution method

Nonlinear model

- Binary variables times continuous variables
- Absolute value $|e^m_{r,t-1} - e^m_{r,t}|$
- CVaR

Linear model

- Add auxiliary variables

  Conversed to linear equations

$$\sigma_{r,t} \le M \times e^m_{t,r}$$
$$\sigma_{r,t} \le p^m_r$$
$$\sigma_{r,t} \ge p^m_r - M \times (1 - e^m_{t,r})$$
$$\sigma_{r,t} \ge 0$$

$$e_1 - e_2 \le A \le e_1 - e_2 + 2 \times B$$
$$e_2 - e_1 \le A \le e_2 - e_1 + 2 \times (1 - B)$$

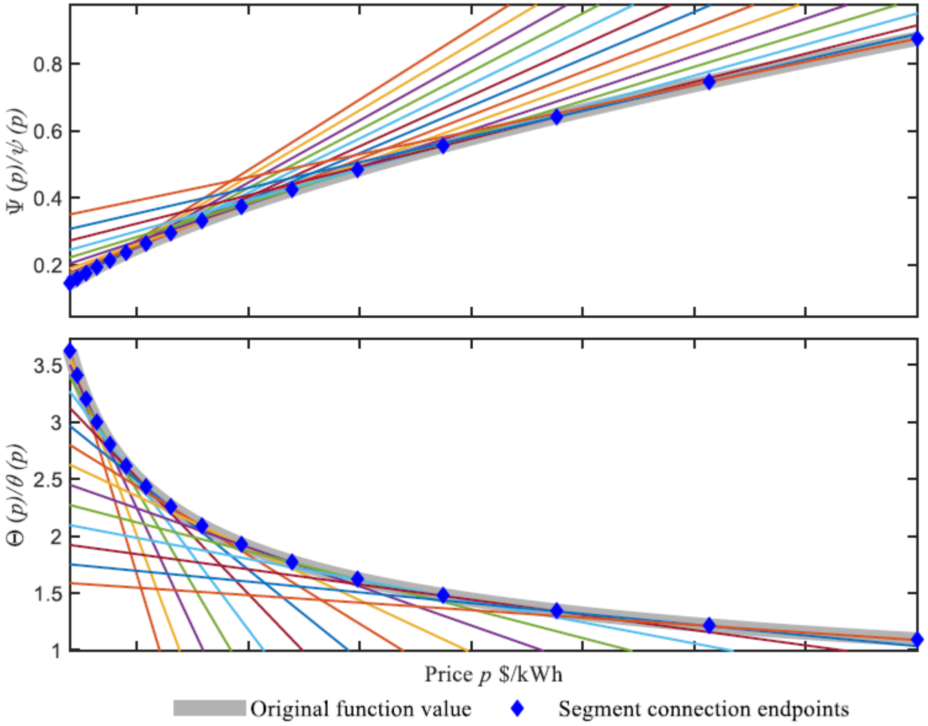
$$CVaR \ge a + \frac{1}{(1-\alpha^{CVaR}) \cdot N_S} \sum_{n_s=1}^{N_s} W_{n_s}$$
$$W_{n_s} \ge 0$$
$$W_{n_s} \ge [(-\Delta R^D_{n_s} - \Delta R^{RT}_{n_s}) - a]$$

* new variables are marked in blue

Personalized Retail Price Design

Case Study

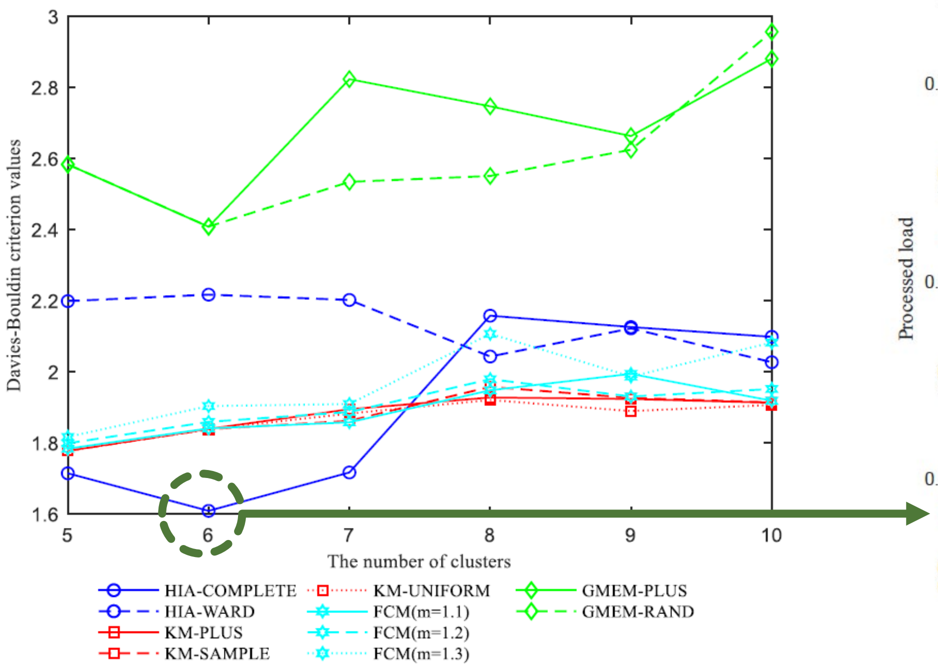
- 6435 consumers in Ireland.
- Data collected every 30 minutes.



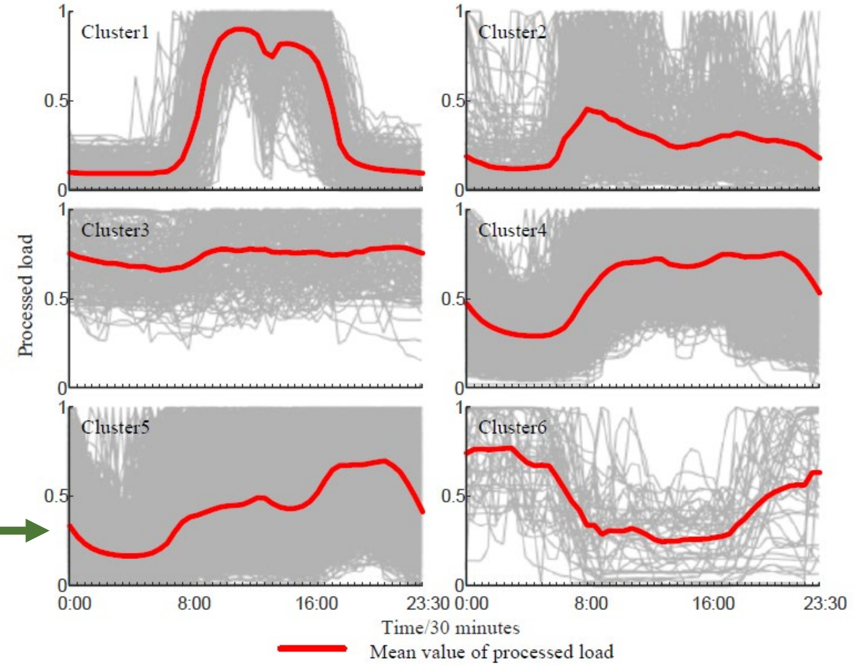
Linear segment approximation(12 segments)

Personalized Retail Price Design

Case Study - clustering



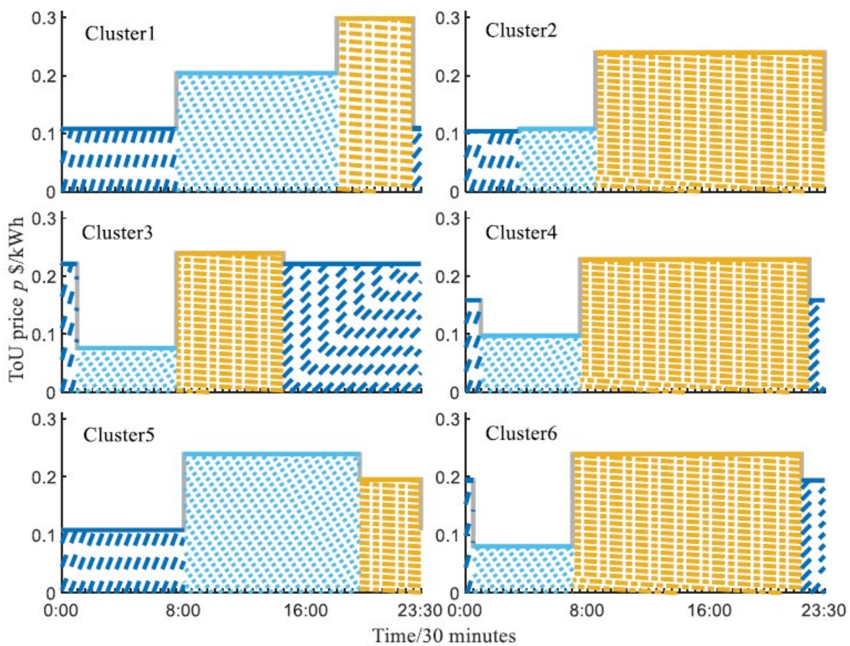
DB index result



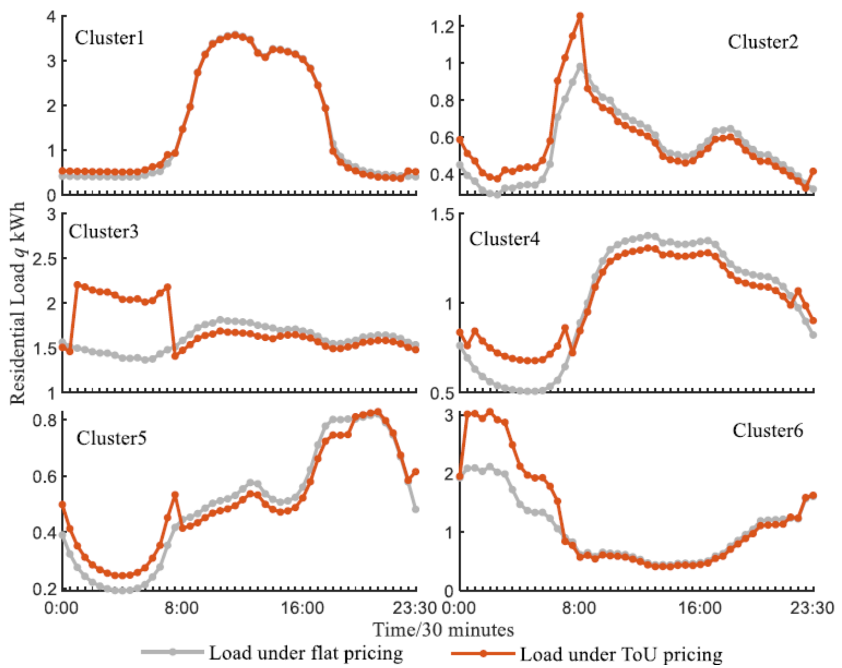
Clustering result

Personalized Retail Price Design

Case Study – prices and responses



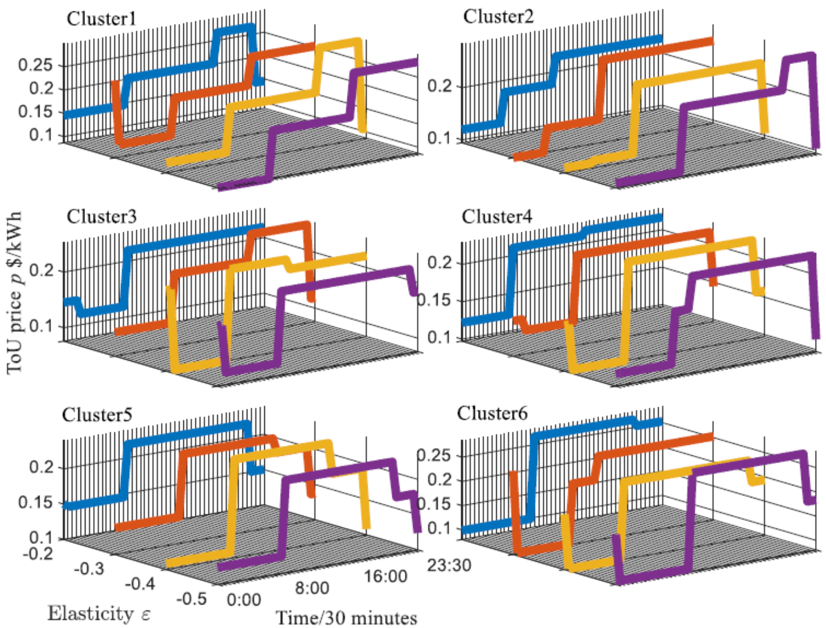
Personalized price



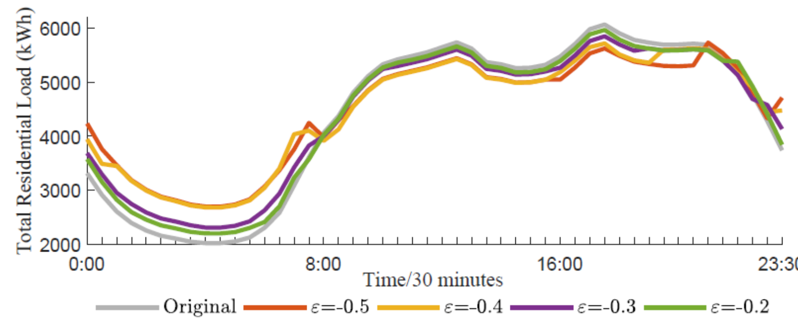
Consumer response

Personalized Retail Price Design

Case Study – sensitivity analysis on elasticity



ToU under different elasticity



Total load under different elasticity

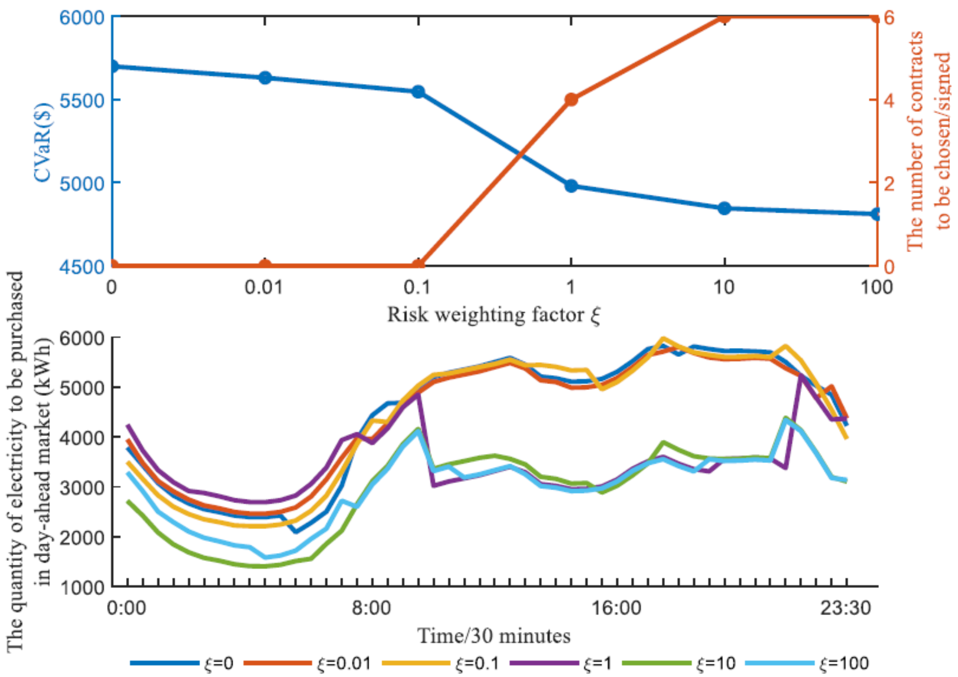
Retailing profit under different elasticity

Elasticity	Original	-0.2	-0.3	-0.4	-0.5
Retailing Profit($)	752	833	977	1186	1385

Elasticity ↓ 😢 Willingness to change ↓

Personalized Retail Price Design

Case Study – sensitivity analysis on risk weighting factor



How **CVaR, the quantity of power bought from day-ahead market and through forward contracts** changes with the change of risk weighting factor?

risk weighting factor rises ↑
attach more importance to risk ↑

minimize CVaR ↓

buy less from day-ahead market ↓
buy more through forward contracts ↑

Power Systems Laboratory

Personalized Retail Price Design

Case Study - sensitivity analysis on clustering methods

	RP	SW	AP	F/SC
Original	752.03	0	0.2	-/-
HIA-COMP	1186.01	339.72	0.1947	65%/89%
HIA-WARD	1188.70	10.01	0.1971	33%/59%
KM-PLUS	1145.68	7.01	0.1973	9%/20%
KM-SAMPLE	1137.61	4.50	0.1975	22%/48%
KM-UNIFORM	1142.61	15.76	0.1973	11%/31%
FCM(m=1.1)	1150.43	9.43	0.1970	30%/47%
FCM(m=1.2)	1176.08	18.64	0.1968	19%/35%
FCM(m=1.3)	1208.06	0.64	0.1970	8%/20%
GMEM-PLUS	1145.82	36.01	0.1965	13%/28%
GMEM-RAND	1144.85	46.60	0.1967	10%/24%

How much profit does the retailer get?

- RP=Retaling Profit($)

How much welfare do the consumers get?

- SW=Social Welfare
- AP=Average Price($/kWh)

How well does clustering perform?

- F/SC=First/Second Choice

> - The most accurate prediction
> - The most profitable for both retailer and consumers

Personalized Retail Price Design

➤ Short Summary

- The Stackelberg game between the retailer and the strategic consumers, an incentive-compatible market, and the retailer's costs, risks and purchasing strategy are considered in this model.
- The ToU tariff can achieve the effects of peak shaving and valley filling, thereby simultaneously increasing the retailer's profitability and ensuring consumers' willingness and preferences.
- How elasticity of consumers and risk weighting factor of retailer influence the designed price is studied.

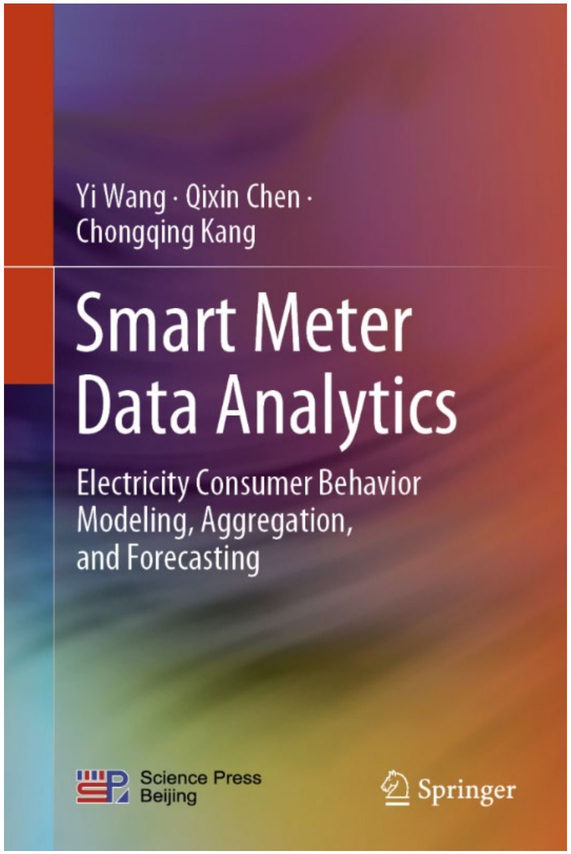
Conclusions

- How to make full use of fine-grained smart meter data?
- A better understanding of the consumer behavior helps to improve the accuracy/performance of aggregated load forecasting.
- A better understanding of the consumer behavior helps to make better decision for both retailer and consumers.
- Any other applications???

Any other applications???

No.	System/ Data	Data Source	Data Type	Frequency	Data Structure
1	Economic Information	Statistic Bureau	GDP、CPI、PMI（Purchasing Managers Index）、Sales Value、Prosperity Index	Per Month	Non structural
2	Energy Consumption Data	Energy Efficiency Platform	Electrical Load、Output、Power Quality、Temperature	15Min	Non structural /Structural
3	Meteorological Data	Meteorological Bureau	Temperature、Humidity、Rainfall	Per Day	Structural
4	EV Charging Data	Charging-Pile RTU	Current、Voltage、Charging Rate、State of Charge	15Min	Structural
5	Customer Service Voice Data	Customer Service System	Customer Voice Data	Real Time	Non structural

Advertisement……



Foreword

Smart grid is a cyber-physical-social system where the power flow, data flow, and business flow are deeply coupled. Enlightened consumers facilitated by smart meters form the foundation of a smart grid. Countries around the world are in the midst of massive smart meter installations for consumers on the pathway towards grid digitalization and modernization. It enables the collection of extensive fine-grained smart meter data, which could be processed by data analytical techniques, especially now widely available machine learning techniques. Big data and machine learning terms are widely used nowadays. People from different industries try to apply advanced machine learning techniques to solve their own practical issues. The power and energy industry is no exception. Smart meter data analytics can be conducted to fully explore the value behind these data to improve the understanding of consumer behavior and enhance electric services such as demand response and energy management.

This book explores and discusses the applications of data analytical techniques to smart meter data. The contents of the book are divided into three parts. The first part (Chaps. 1–2) provides a comprehensive review of recent developments of smart meter data analytics and proposes the concept of "electricity consumer behavior model". The second part (Chaps. 3–5) studies the data analytical techniques for smart meter data management, such as data compression, bad data detection, data generation, … The third part (Chaps. 6–12) conducts application-oriented research to depi… electricity consumer behavior model. This part includes electrical consumpt… tern recognition, personalized tariff design for retailers, socio-demograp… mation identification, consumer aggregation, electrical load forecasti… The prospects of future smart meter data analytics (Chap. 13) are also prov… the end of the book. The authors offer model formulations, novel algorith… -depth discussions, and detailed case studies in various chapters of this bo…

One author of this book, Prof. Chongqing Kang, is a prof…onal colleague. He is a distinguished scholar and pioneer in the power and e…rgy area. He has done extensive work in the field of data analytics and load forecasting. This is a book worth reading; one will see how much insight can be gained from smart meter data

viii Foreword

alone. There are definitely broader qualitative understanding that can be gained from massive data collected in the realm of generation, transmission, distribution, and end use of the smart grid.

September 2019

Prof. Saifur Rahman
Joseph Loring Professor and Founding Director
Advanced Research Institute at Virginia Tech
Arlington, VA, USA

President of the IEEE Power and Energy Society
New York, NY, USA

This is a book worth reading; one will see how much insight can be gained from smart meter data alone.

Prof. Saifur Rahman
IEEE Fellow
President of the IEEE Power and Energy Society

Yi Wang, Qixin Chen, Chongqing Kang, "Smart Meter Data Analytics," *Springer*, 2020.

Thank you for your attention

Yi Wang | yiwang@eeh.ee.ethz.ch | www.eeyiwang.com