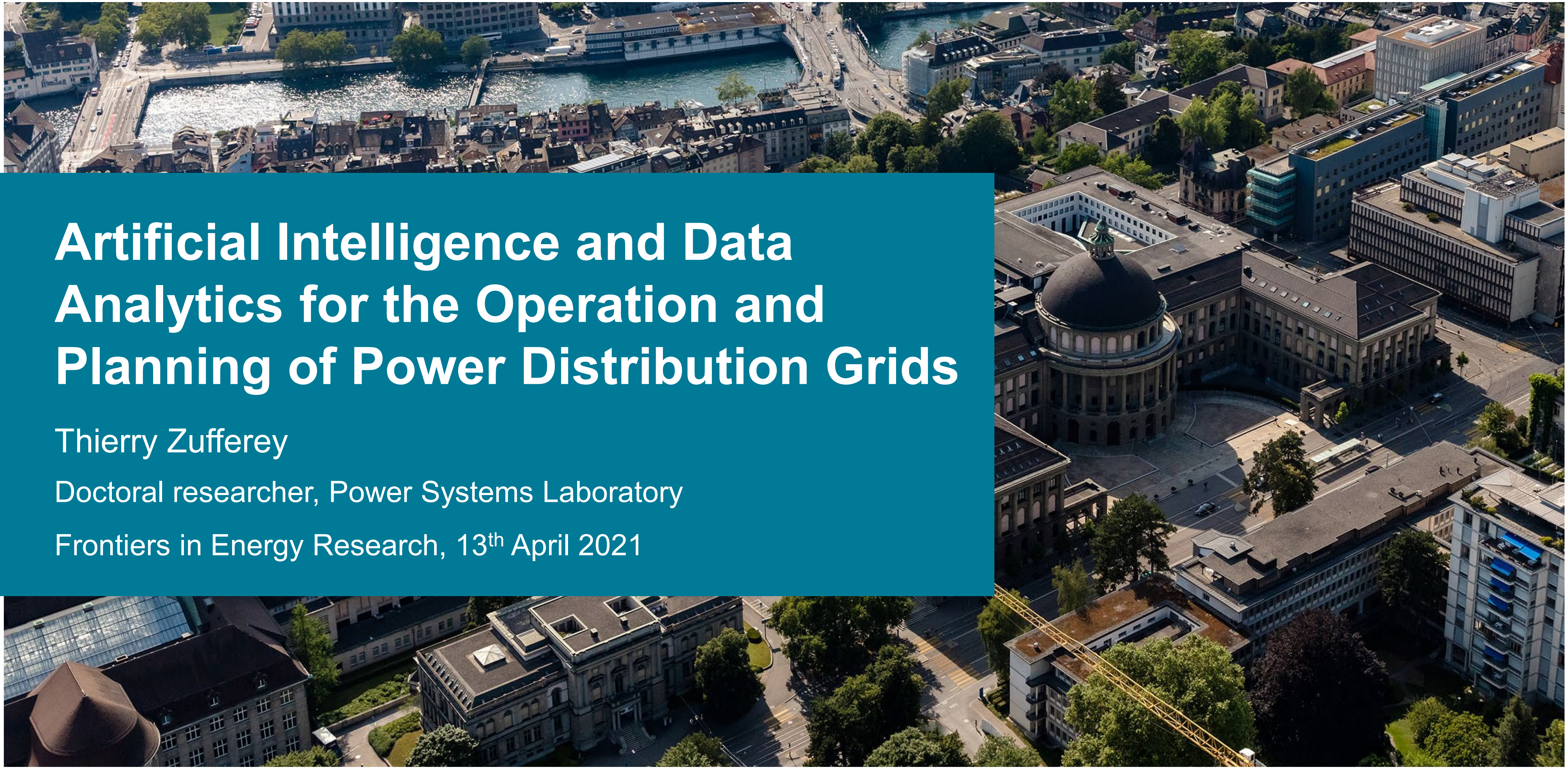
ETH zürich

Power
Systems
Laboratory



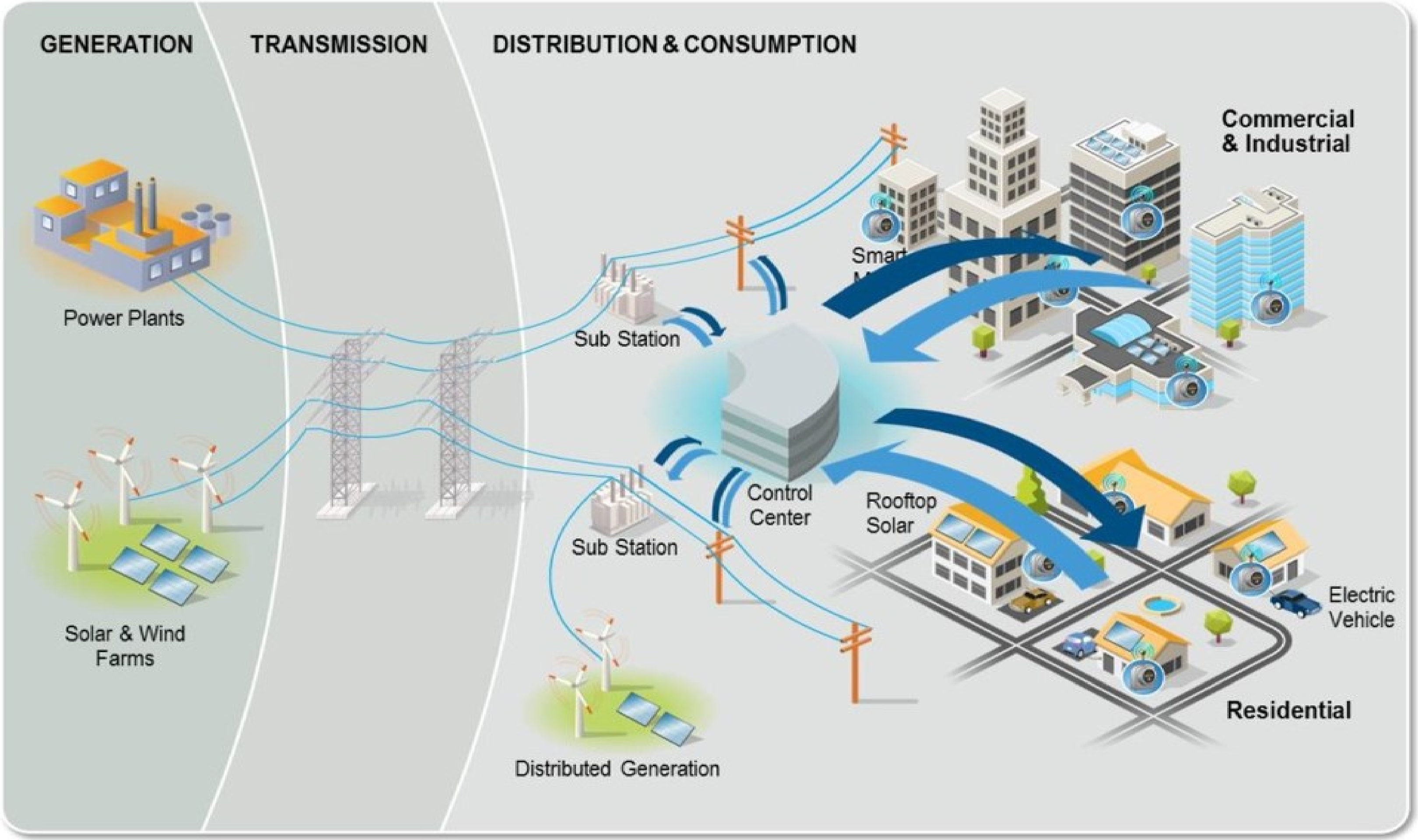
Artificial Intelligence and Data Analytics for the Operation and Planning of Power Distribution Grids

Thierry Zufferey

Doctoral researcher, Power Systems Laboratory

Frontiers in Energy Research, 13th April 2021

Digitalization of Power Distribution Grids



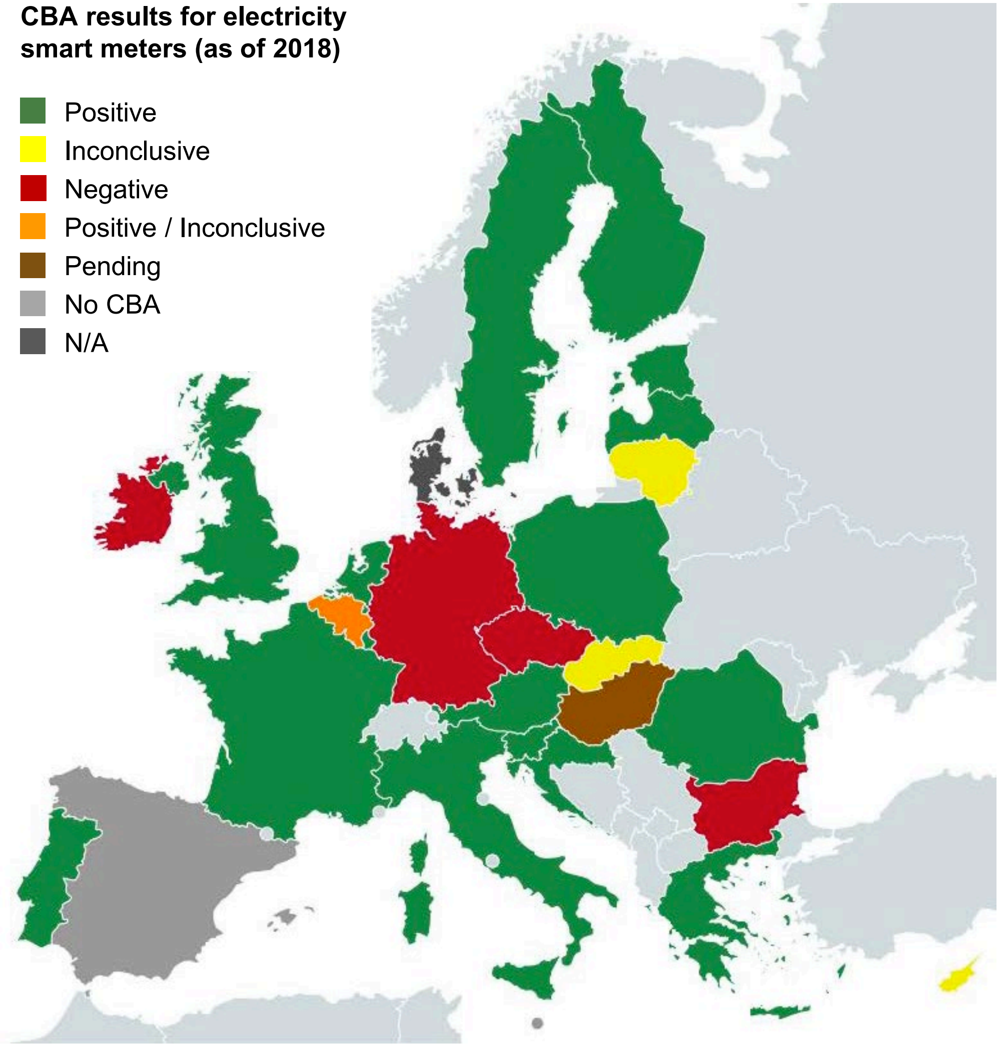
© Rethink Electric, *What is a Smart Grid and Why are We Building it?*

Advanced Metering Infrastructure (AMI)

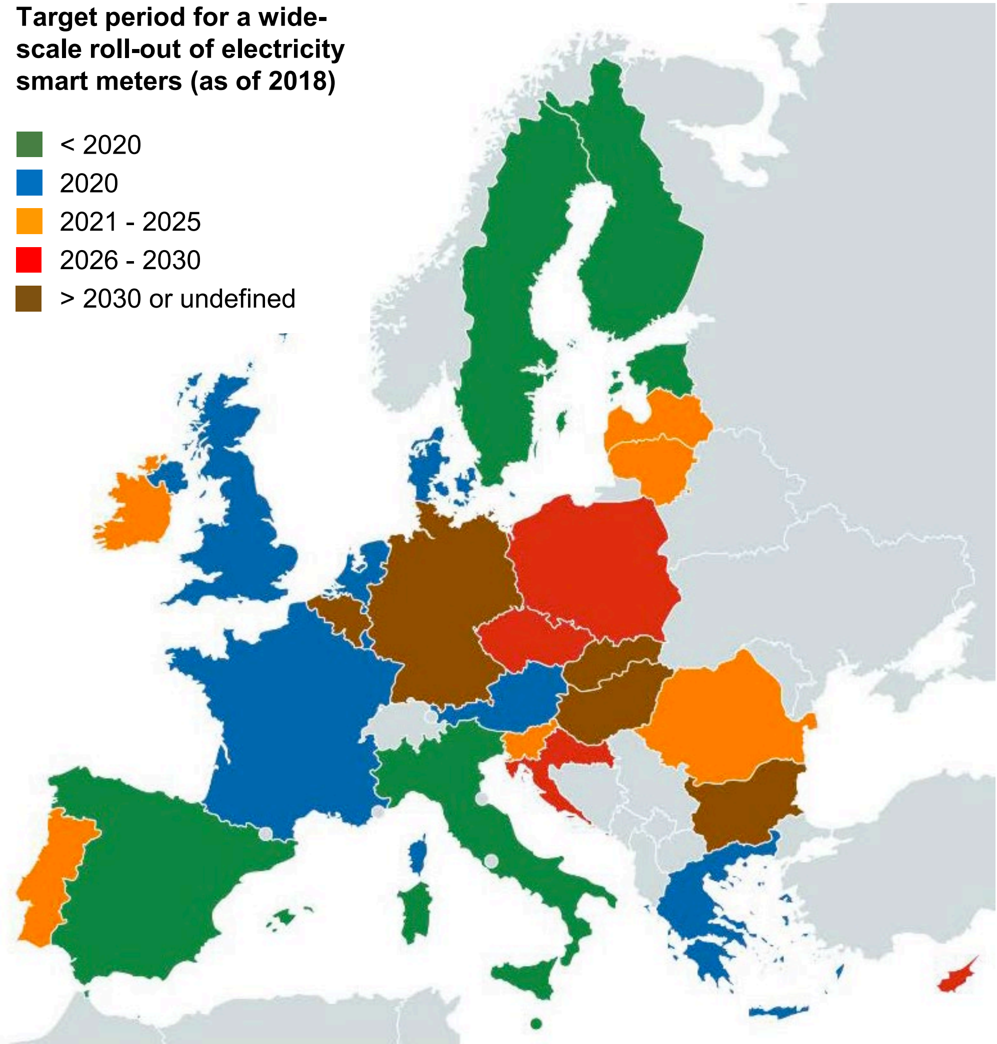
Smart meter = digital sensor + bidirectional communication

© Nuri Telecom, “Solutions – Advanced Metering Infrastructure”

Roll-Out of Electricity Smart Meters in the EU



CBA results for electricity smart meters (as of 2018)



Target period for a wide-scale roll-out of electricity smart meters (as of 2018)

© F. Tounquet and C. Alaton, “Benchmarking Smart Metering Deployment in the EU-28”, 2020.

Advanced Metering Data Application and Challenges

Data Preparation

Transactive Energy Systems

Data Privacy

Automated Meter Reading

Long-Term Forecasting

Customer Concerns

Detection of Non-Technical Losses

Short-term Forecasting

Big Data Visualization

Advanced Metering Infrastructure

Flexibility Estimation

Situational Awareness

Load Disaggregation

Grid Monitoring

Pseudo-Measurement Synthesis

Load Profiling

Grid Topology Estimation

Demand-Side Management

State Estimation

Demand Response

Outline

Characteristics of Data in Distribution Grids

- Data Preparation
- Big Data Visualization
- Impact of Temporal Resolution and Spatial Aggregation

Data-based Modeling of Distribution Grids

- Pseudo-Measurement Synthesis
- Optimal Load Allocation

Application of Data-based Approaches in LV Grids

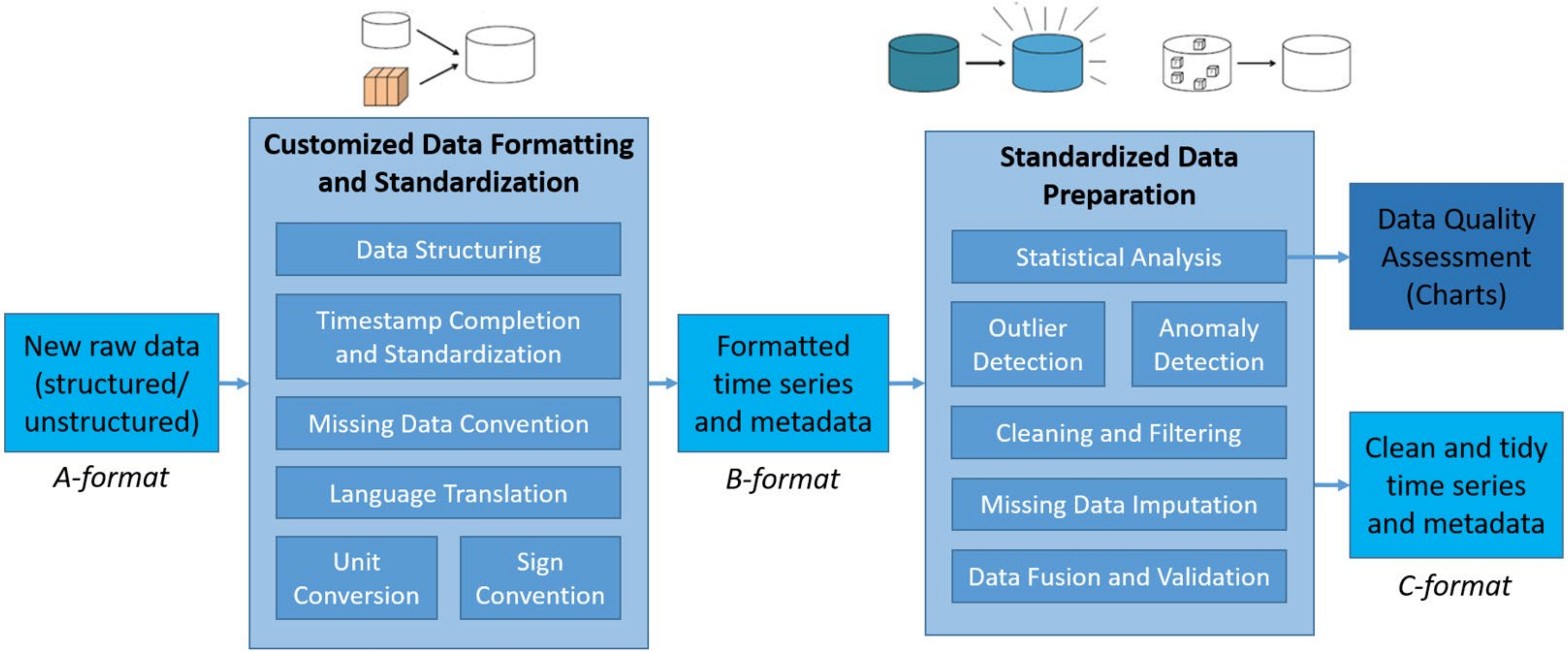
- Detection and Disaggregation of Flexible Loads
- Deterministic and Probabilistic Short-Term Forecasting
- Preventive Voltage Control

Conclusions and Outlook

Data Preparation

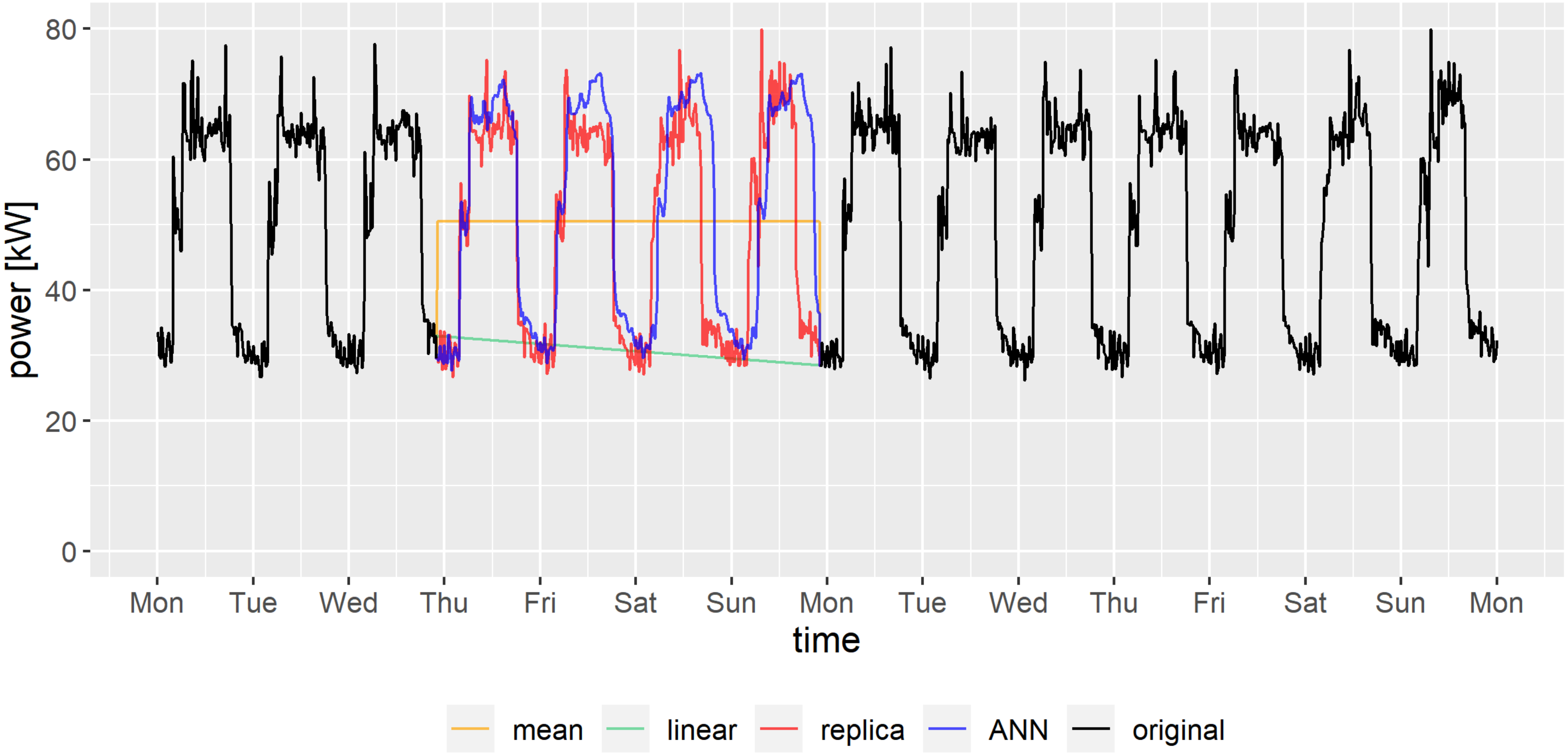
“If 80 percent of our work is data preparation, then ensuring data quality is the important work of a machine learning team.”

Andrew Ng



Data Preparation

Missing Data Imputation

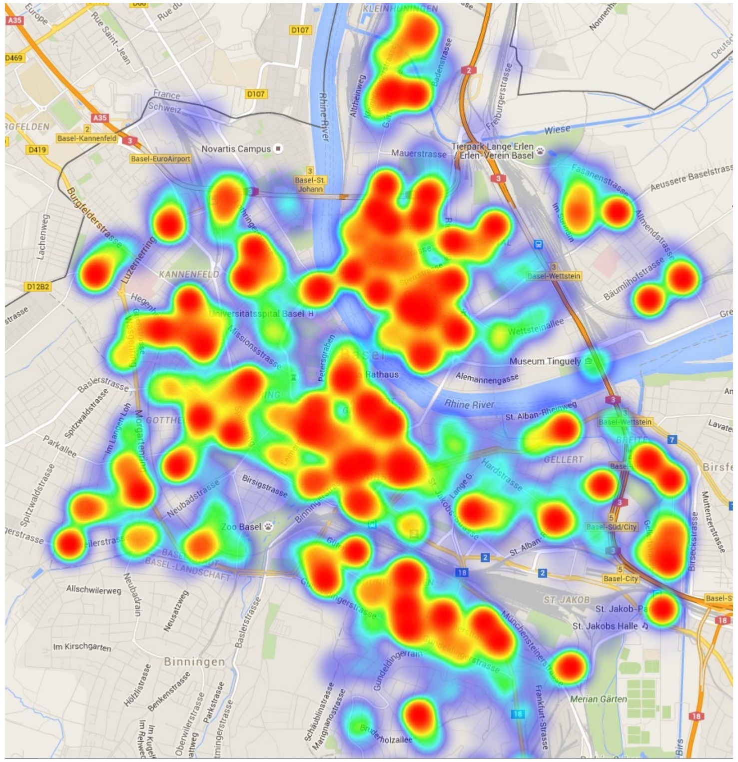


There is no one-size-fits-all solution in data preparation

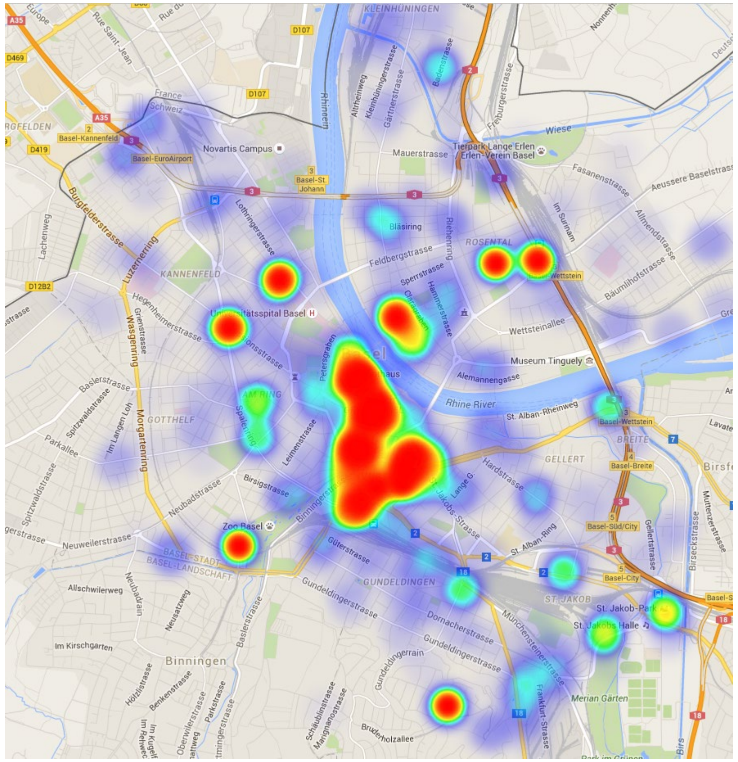
Big Data Visualization and Heatmaps

Spatial Representation

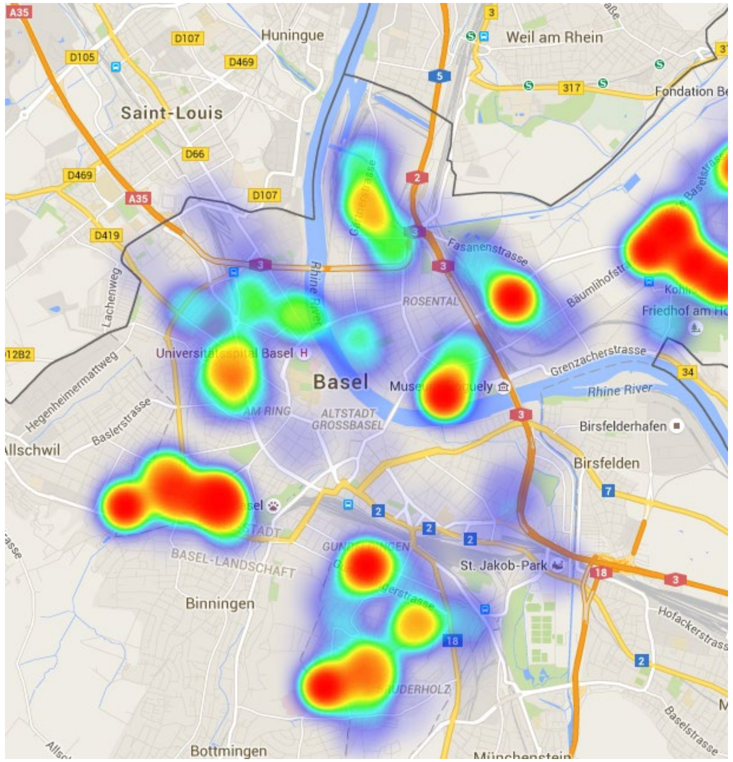
Consumption from Residential End-Users



Consumption from Industrial and Commercial End-Users



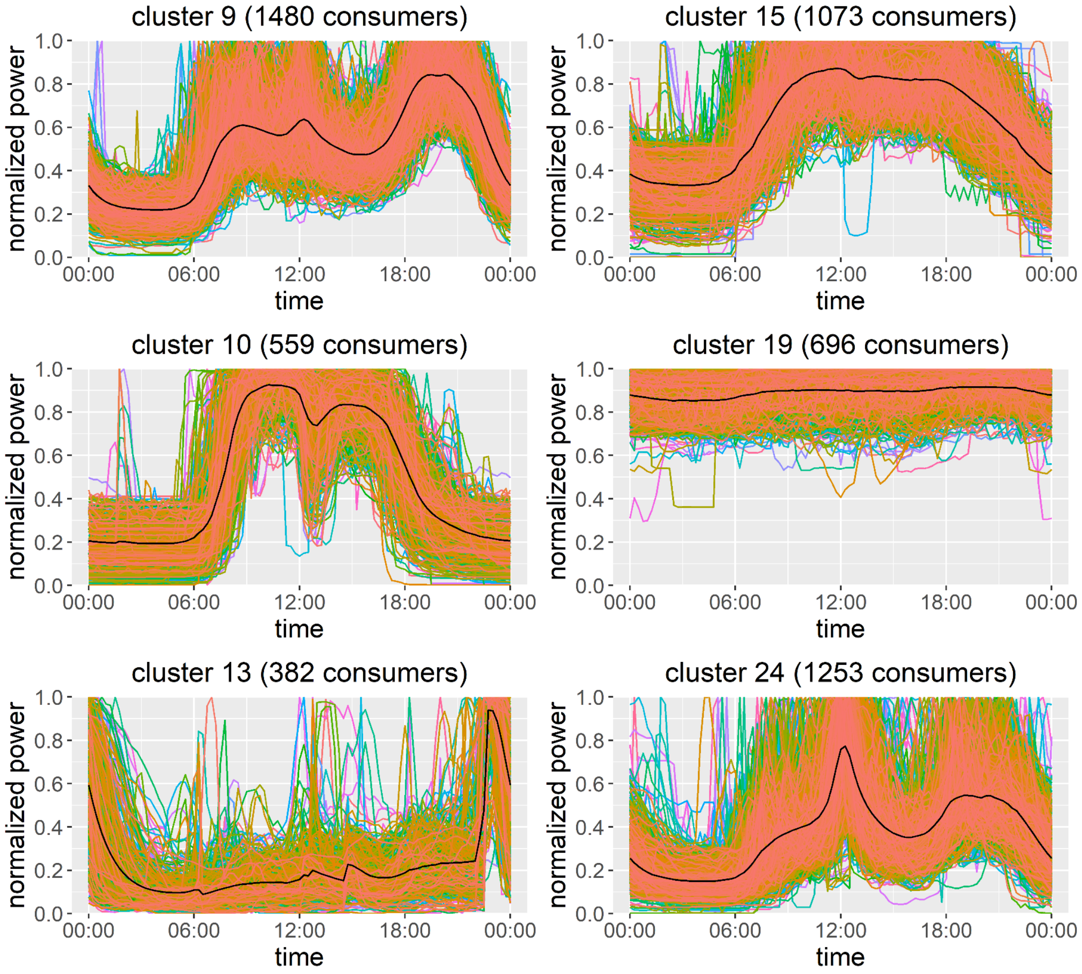
Production from PV Systems



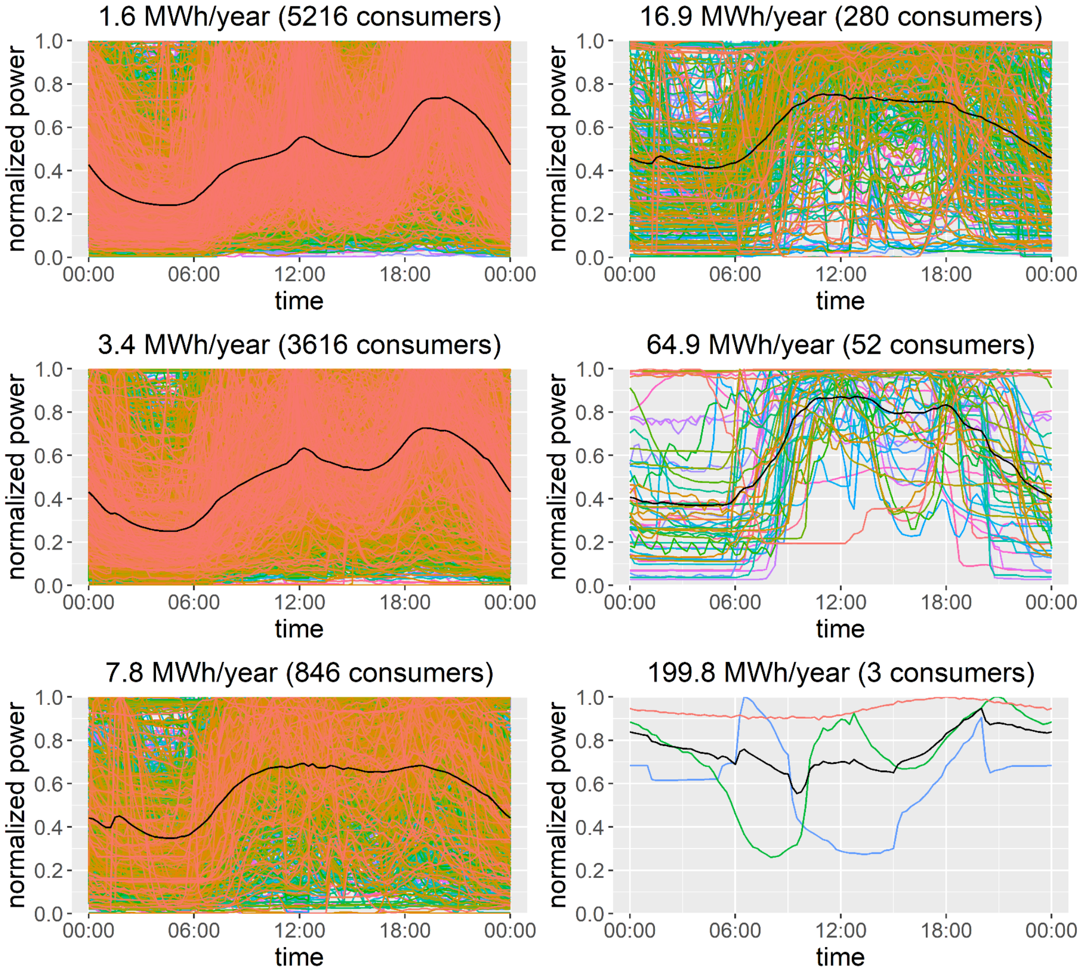
Big Data Visualization and Unsupervised Learning

Temporal Representation

Clustering based on typical daily load profile

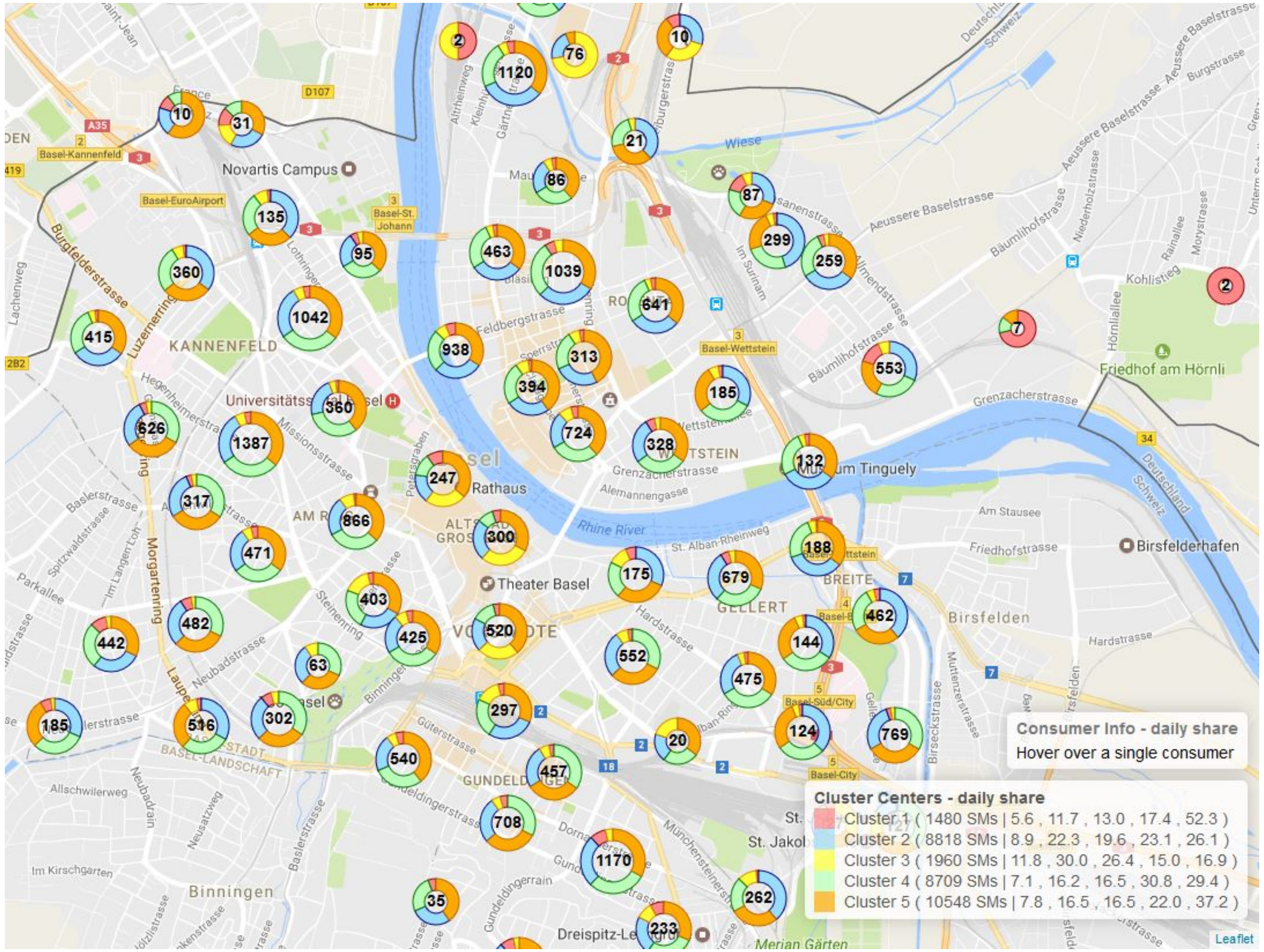


Clustering based on average energy consumption



Big Data Visualization and Unsupervised Learning

Spatial Representation



Consumer Info - daily share
Hover over a single consumer

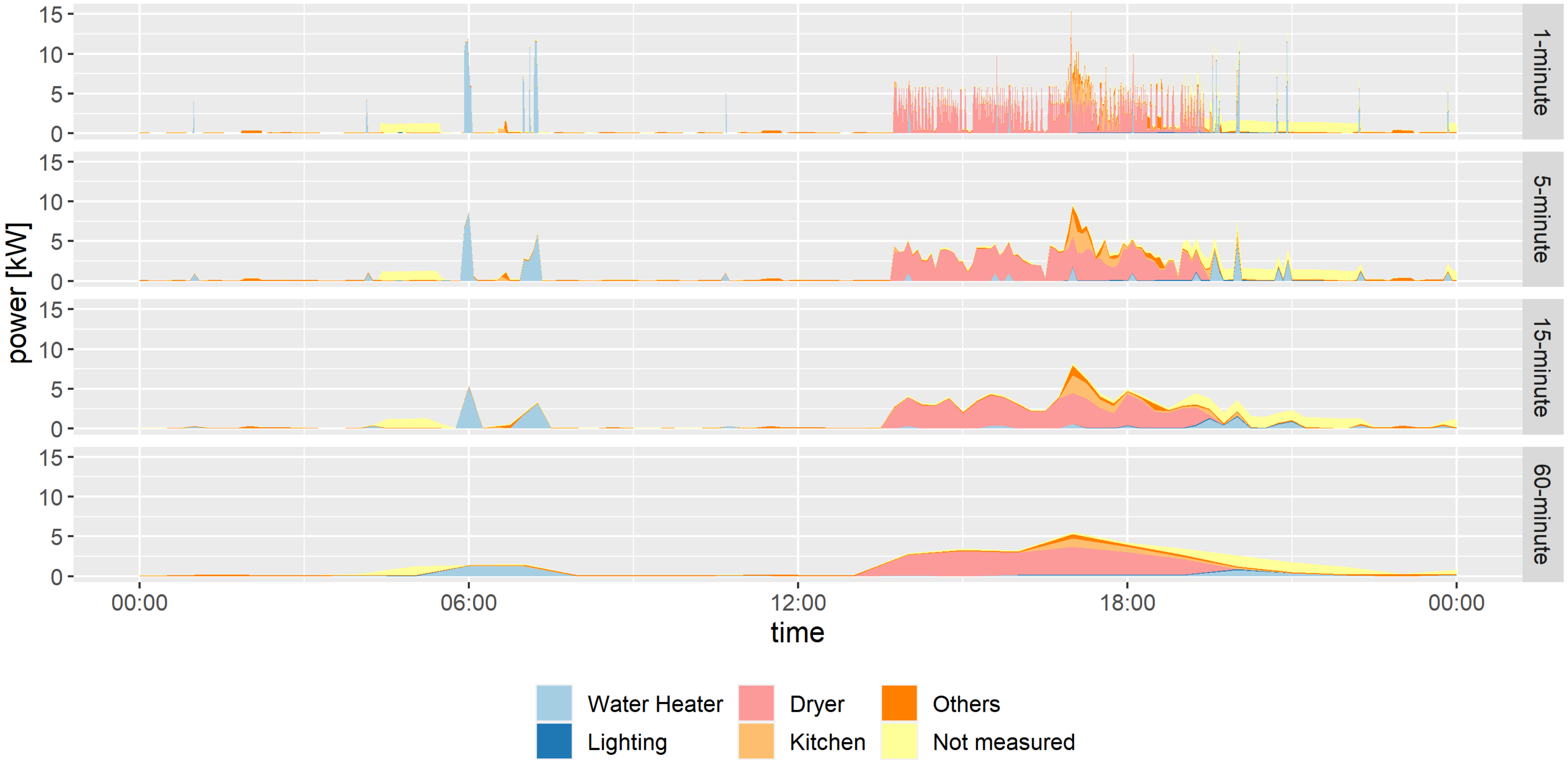
Cluster Centers - daily share
🟥 Cluster 1 (1480 SMs | 5.6 , 11.7 , 13.0 , 17.4 , 52.3)
🟦 Cluster 2 (8818 SMs | 8.9 , 22.3 , 19.6 , 23.1 , 26.1)
🟨 Cluster 3 (1960 SMs | 11.8 , 30.0 , 26.4 , 15.0 , 16.9)
🟩 Cluster 4 (8709 SMs | 7.1 , 16.2 , 16.5 , 30.8 , 29.4)
🟧 Cluster 5 (10548 SMs | 7.8 , 16.5 , 16.5 , 22.0 , 37.2)

Clustering based on Intra-Daily Consumption Share

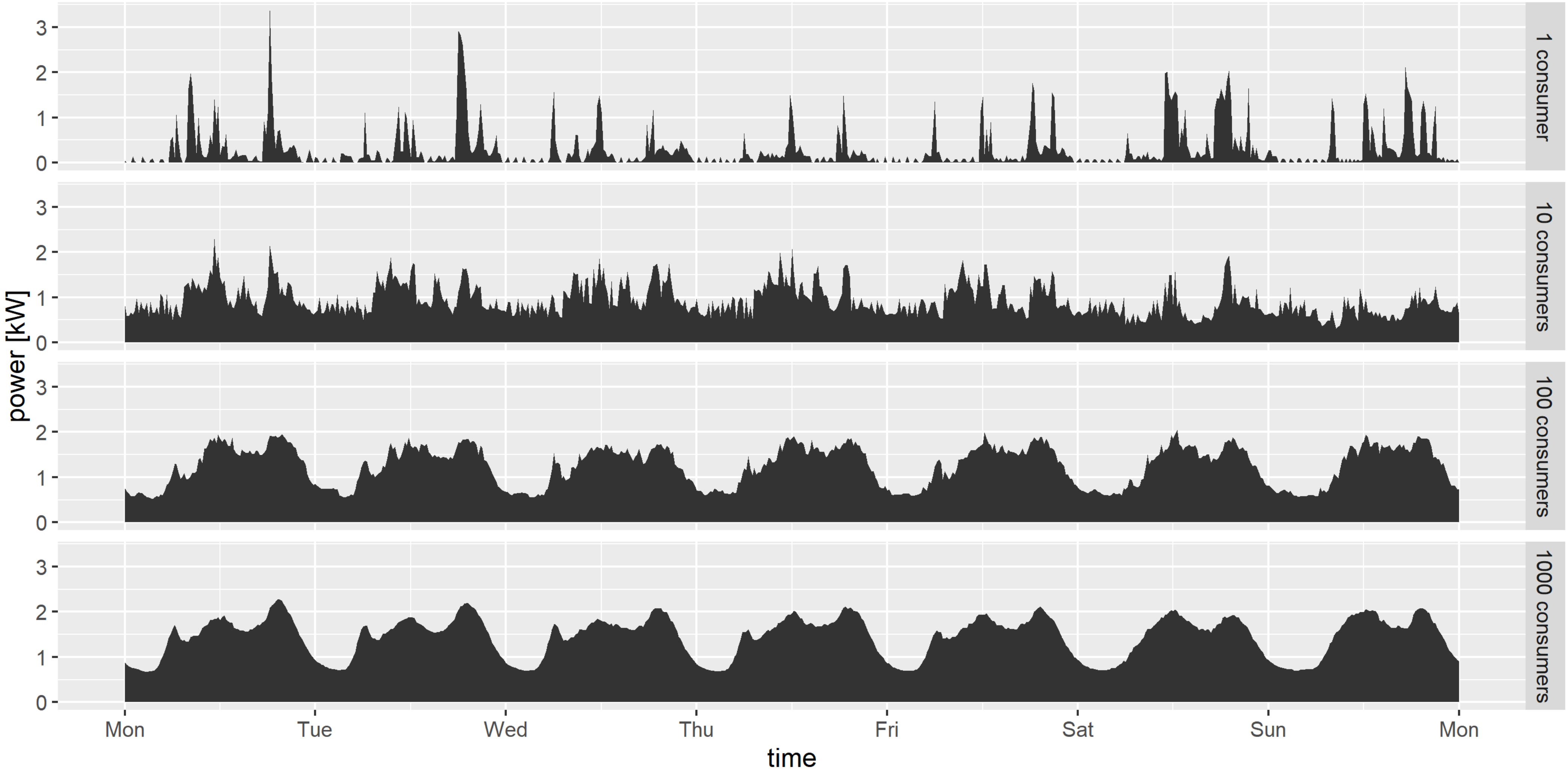
🟥	Active during the night (price-sensitive)
🟦	Active from morning to evening
🟨	Active during business hours
🟩	Active in the early evening
🟧	Active in the late evening

Simple unsupervised methods already provide valuable insights at a city level

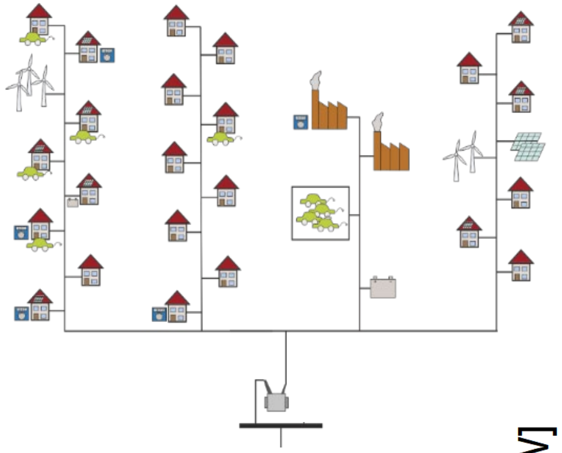
Impact of Temporal Resolution

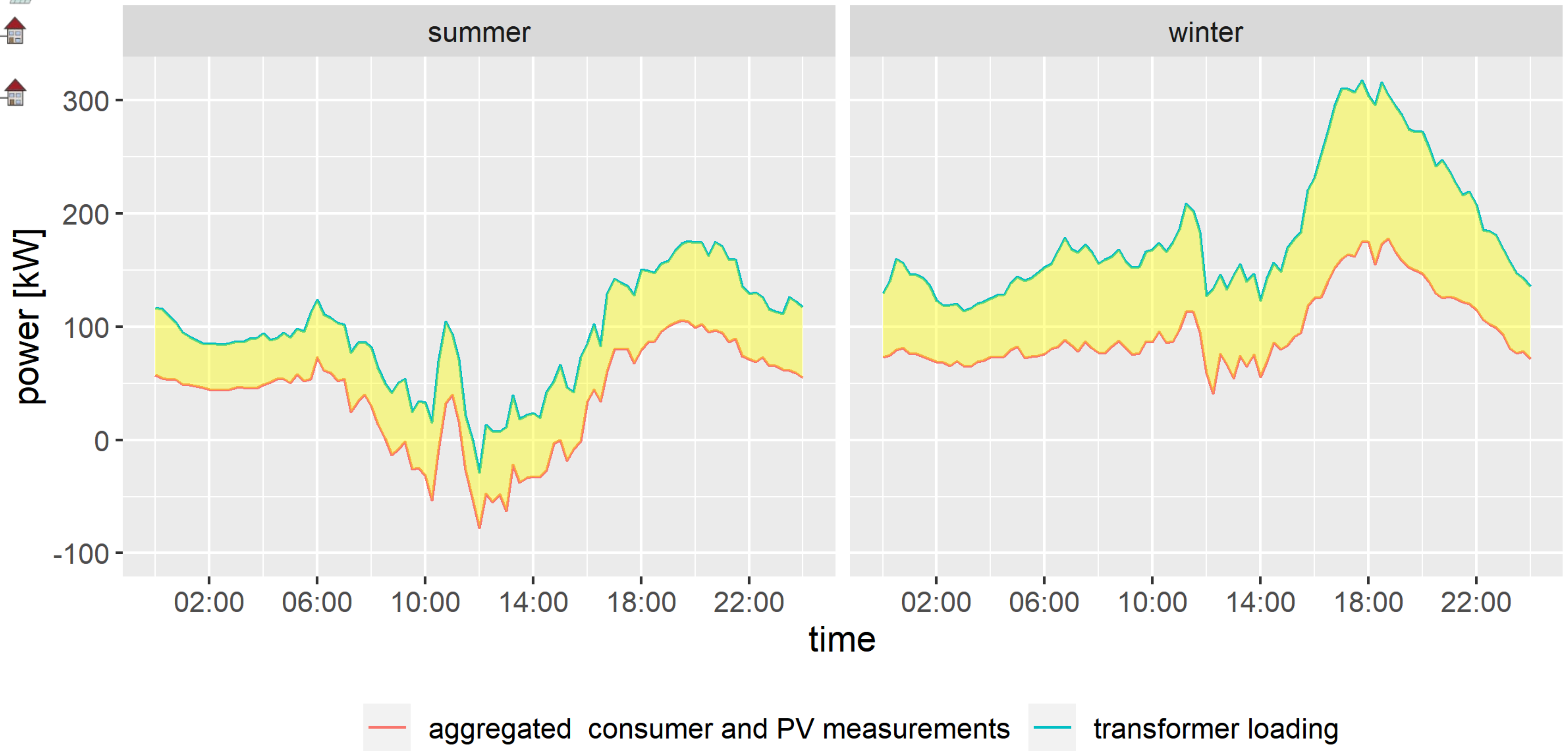


Impact of Spatial Aggregation



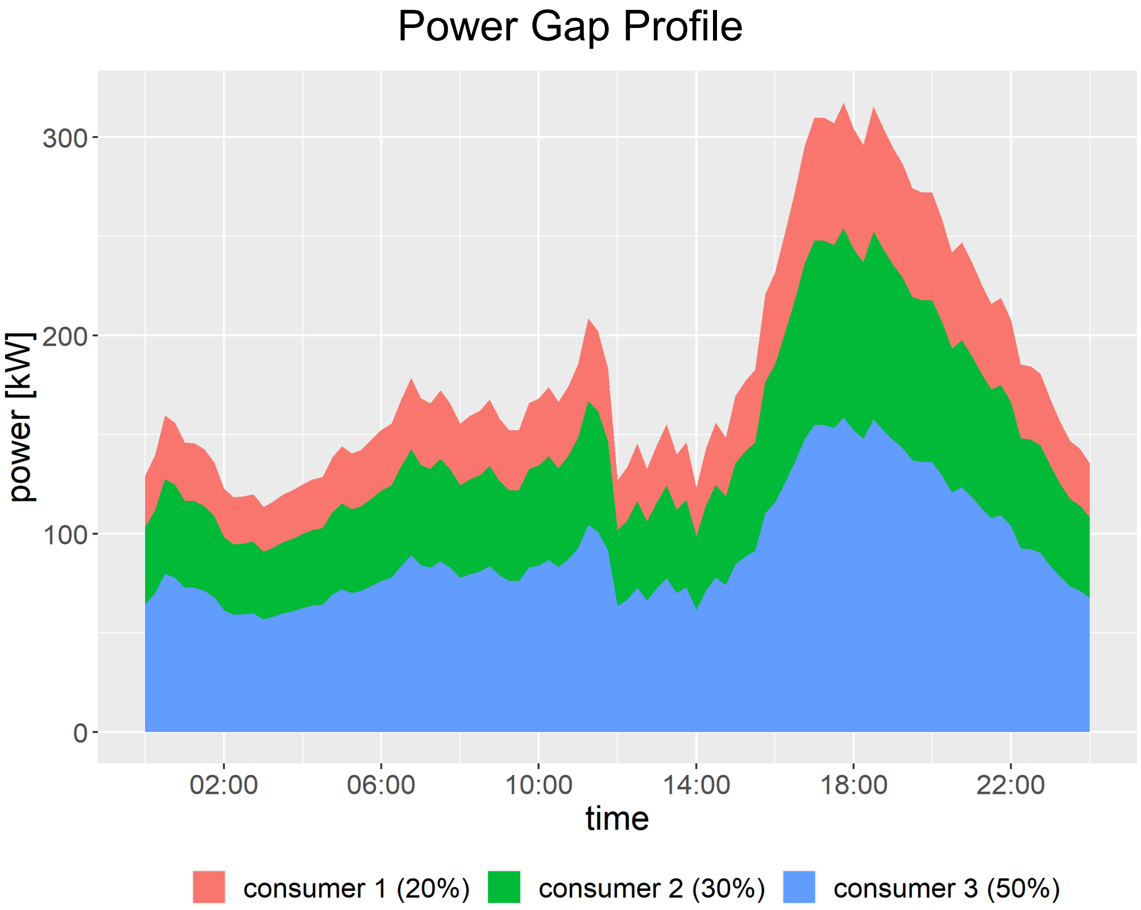
Synthesis of Active Power Pseudo-Measurements

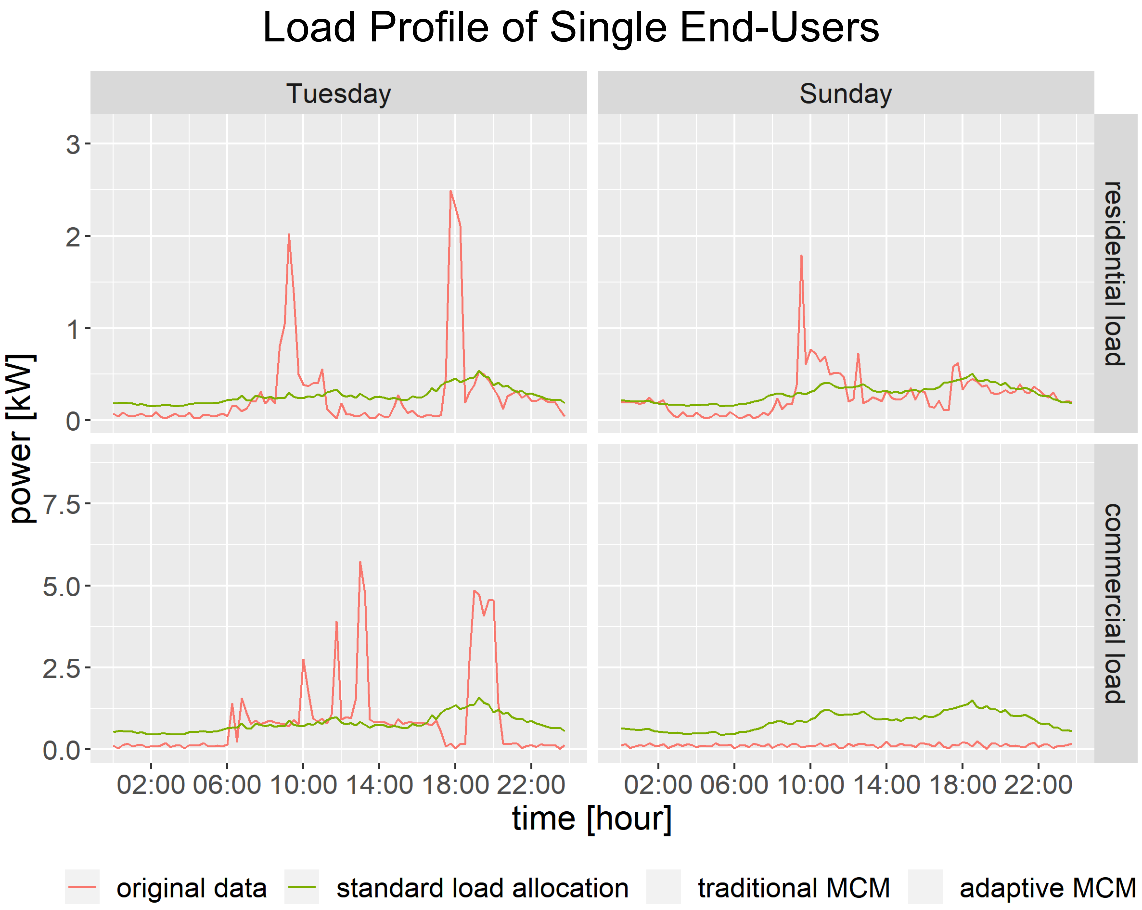




Synthesis of Active Power Pseudo-Measurements

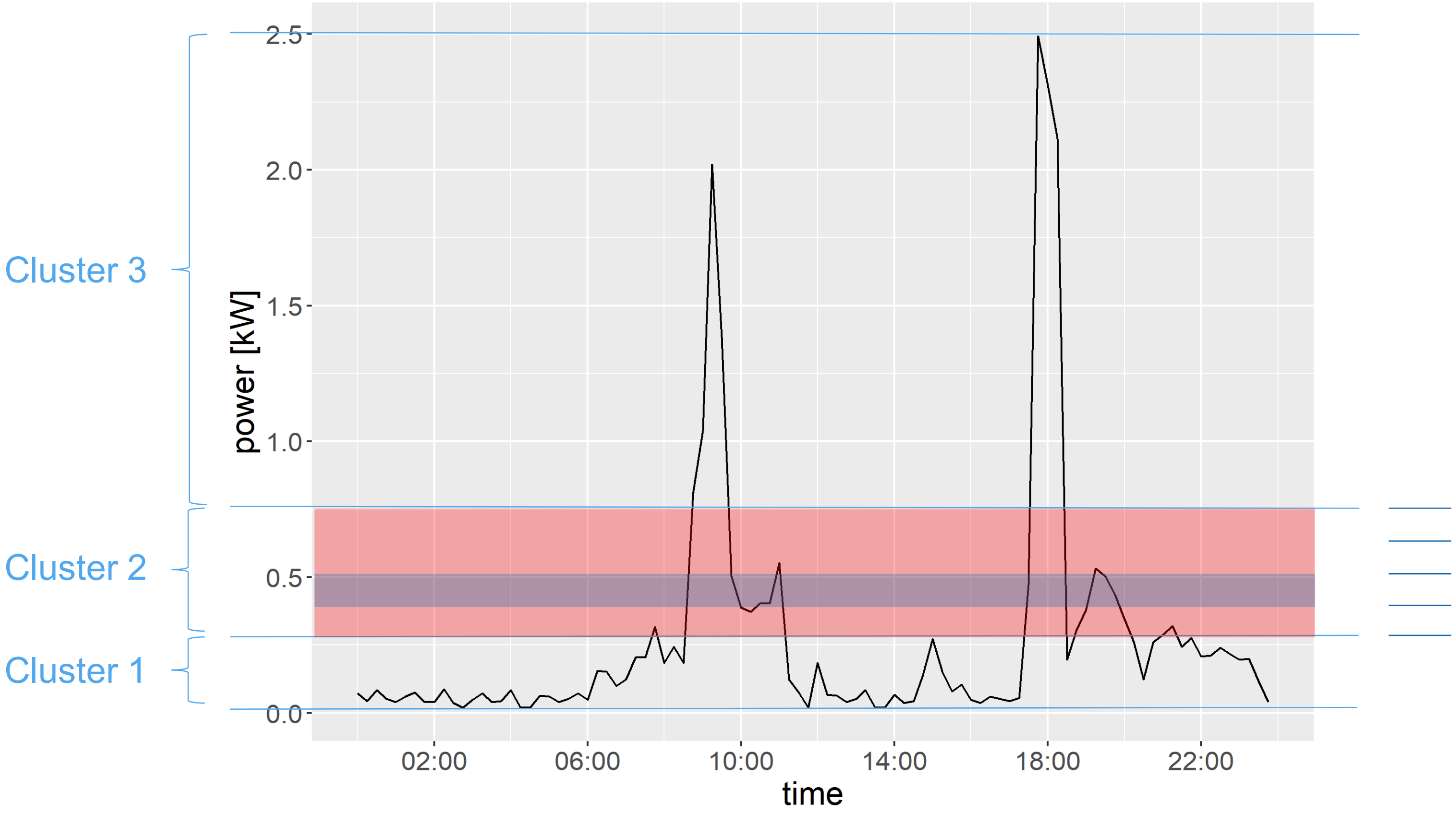
Standard Load Allocation





Synthesis of Active Power Pseudo-Measurements

Traditional Markov Chain Model (TMCM)



1. States → K-means clustering
2. Traditional transition matrix

state at time t

3	0.7	0.1	0.2
2	0.4	0.4	0.2
1	0.6	0.3	0.1
	1	2	3

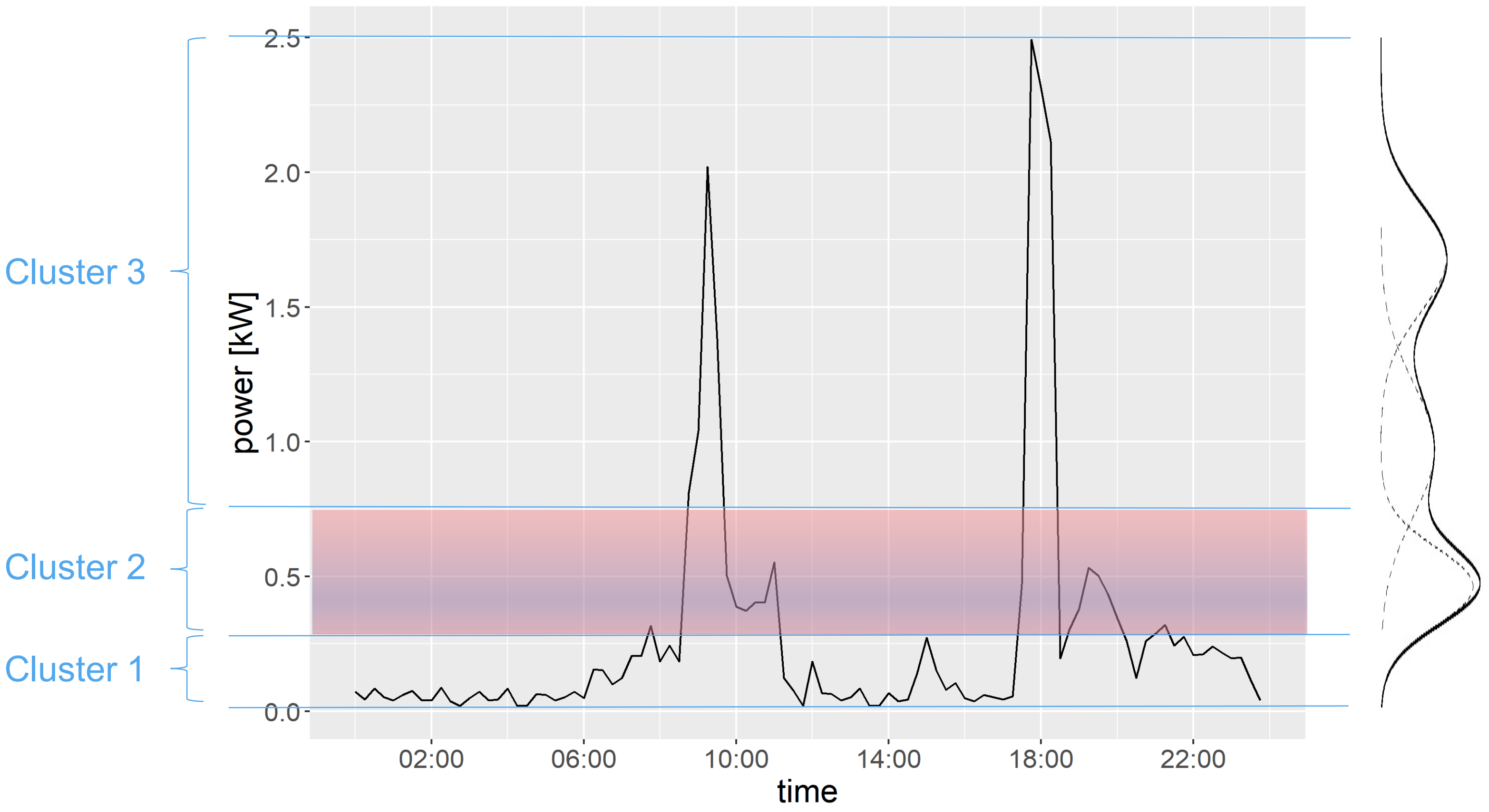
state at time t+1

3. Equidistant sublevels in each state

Load profile generation:
random walk + sublevel selection + noise

Synthesis of Active Power Pseudo-Measurements

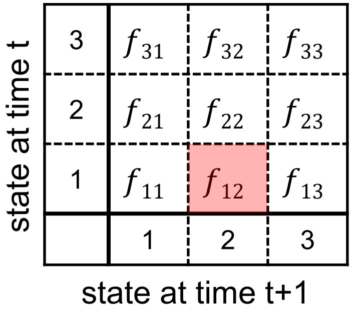
Adaptive Markov Chain Model (AMCM)



1. States → K-means clustering
2. Adaptive transition matrix → each transition element = logistic regression:

$$f(x) = \frac{1}{1 + e^{-\theta^{\mathsf{T}} x}}$$

$$x = (1, hour, weekday, month)$$
$$\theta = (\theta_0, \theta_1, \theta_2, \theta_3)$$



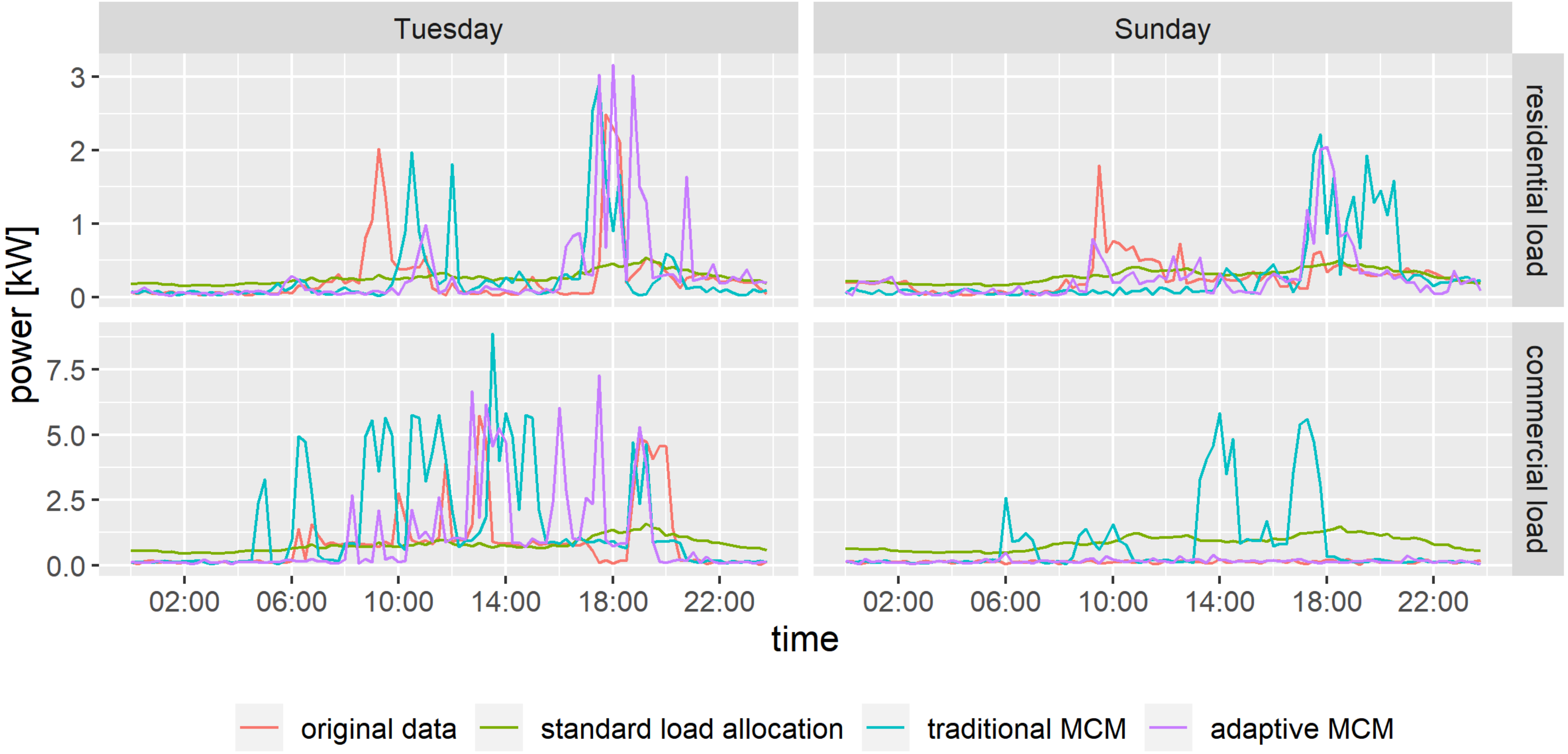
3	f_{31}	f_{32}	f_{33}
2	f_{21}	f_{22}	f_{23}
1	f_{11}	f_{12}	f_{13}
	1	2	3

3. Load distribution modelled as Gaussian Mixture Model (GMM)

Load profile generation: random walk + value selection in GMM

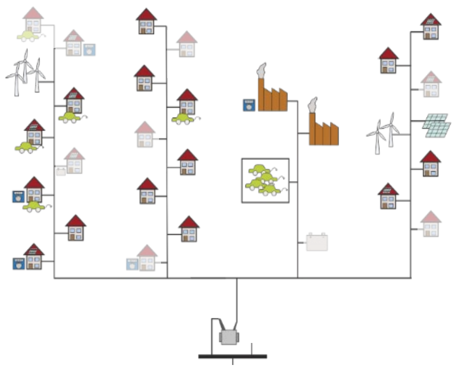
Synthesis of Active Power Pseudo-Measurements

Load Profile of Single End-Users



Optimal Load Allocation

Selection of Optimal Load Profiles – Binary Optimization Problem

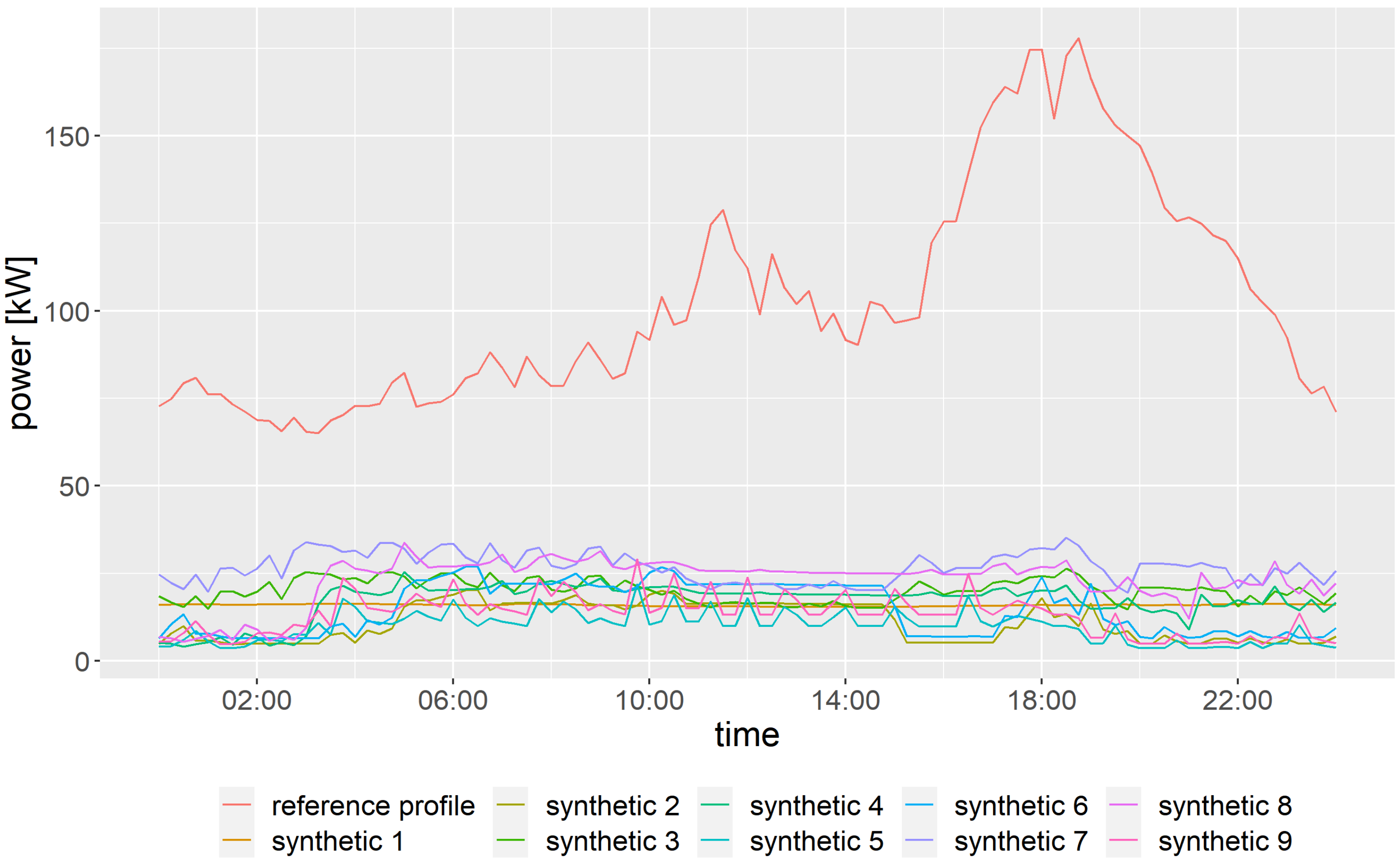


$$\min_{\beta} |y - X\beta|$$

$such\ that$: $\beta \in \{0,1\}$,

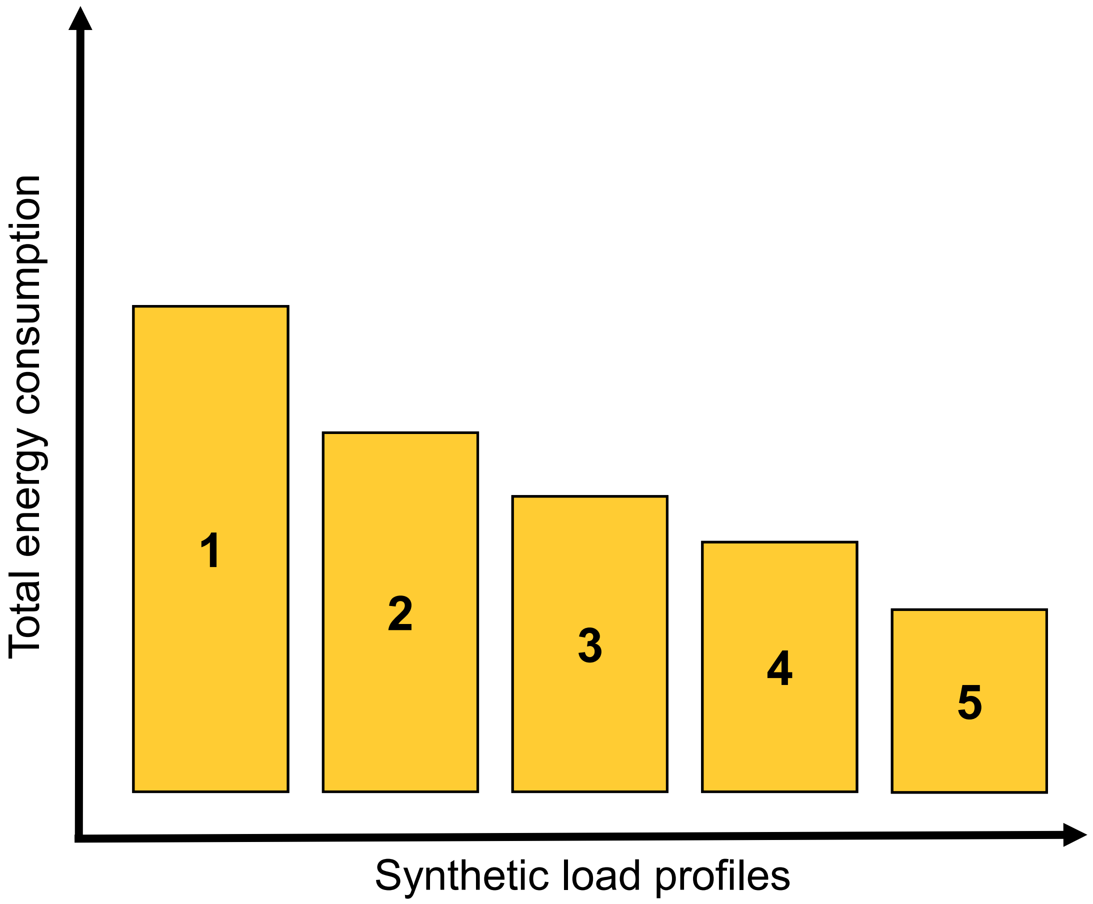
$sum(\beta) \geq M$

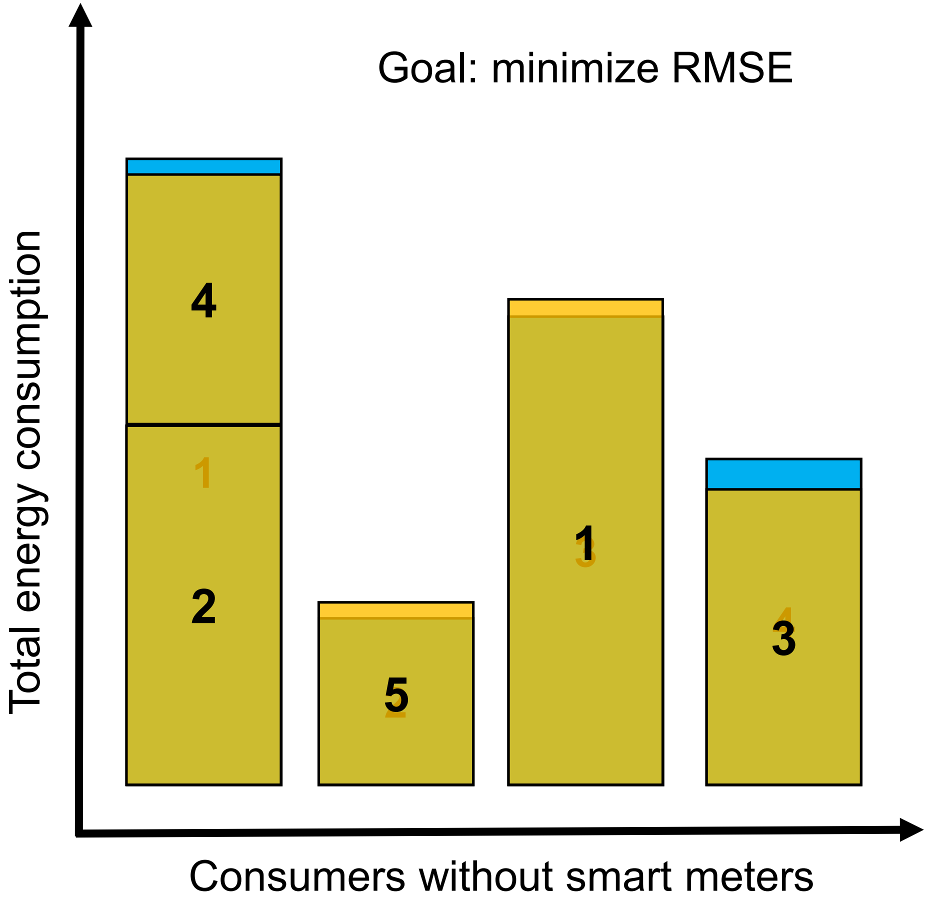
$y = (y_1, \dots, y_N)$	reference profile
$X = (x_1, \dots, x_m)$	design matrix
$x_j = (x_{1j}, \dots, x_{Nj})$	synthetic profile
$\beta = (\beta_1, \dots, \beta_m)^T$	binary factors
M	number of non-metered consumers
N	number of time steps
m	number of synthetic profiles



Optimal Load Allocation

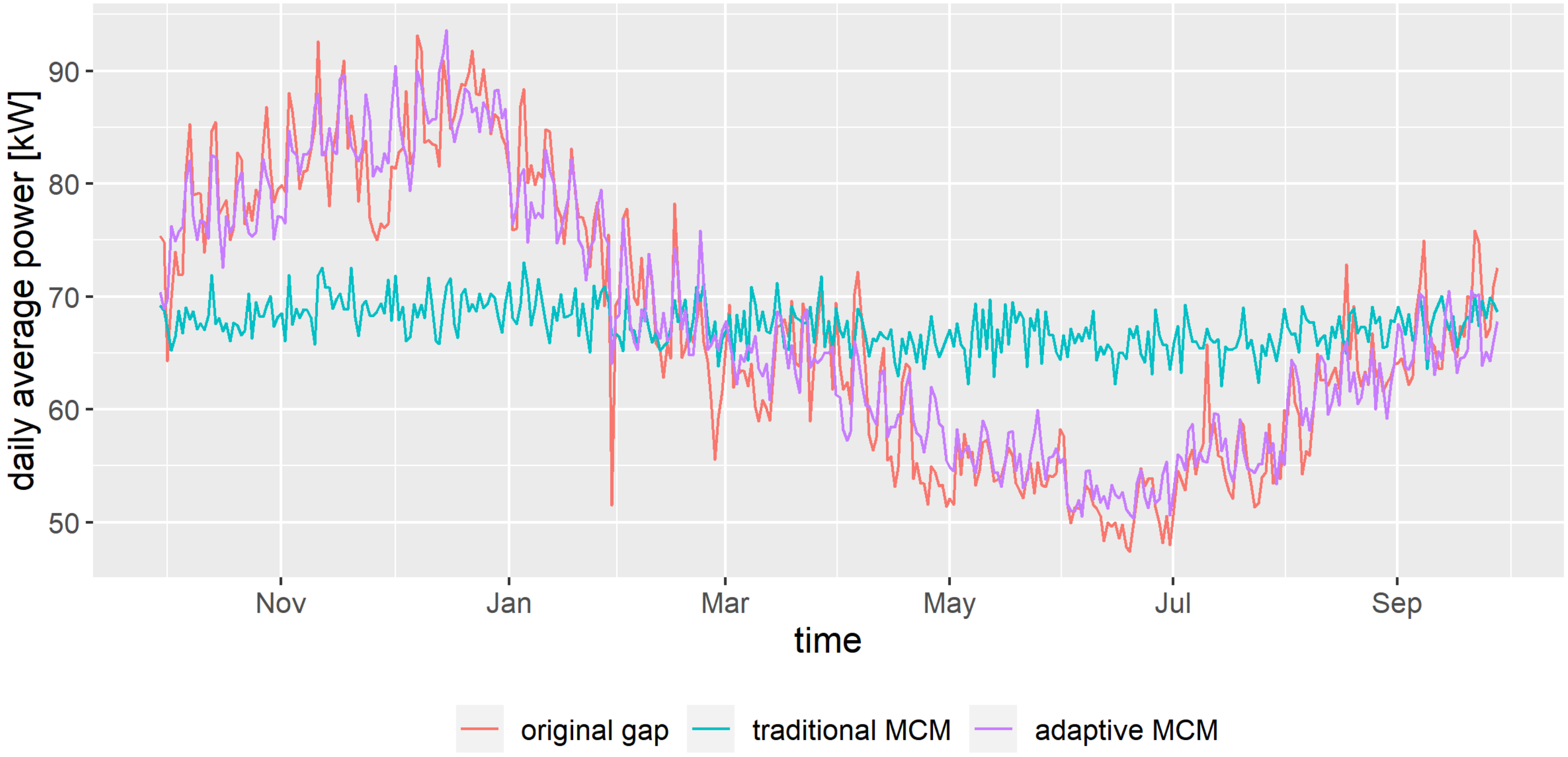
Bin Packing Problem





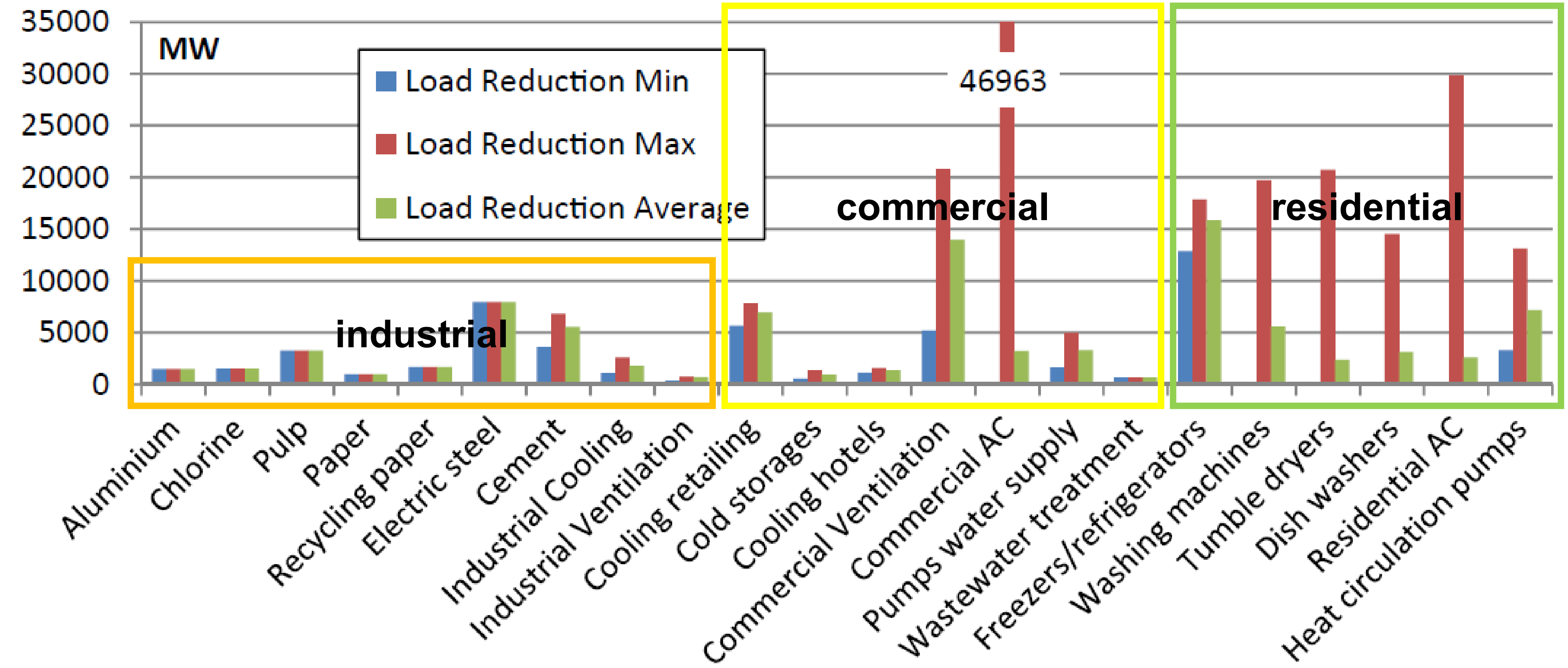
Optimal Load Allocation

Matching at Aggregate Level (Daily Average)



Detection and Disaggregation of Flexible Domestic Loads

Context and Motivation



© H.C. Gils, “Technology Shares in Potential Load Reduction in Europe”, *Energy*, vol. 67, pp.1–18, 2014.

Detection and Disaggregation of Flexible Domestic Loads

Context and Motivation

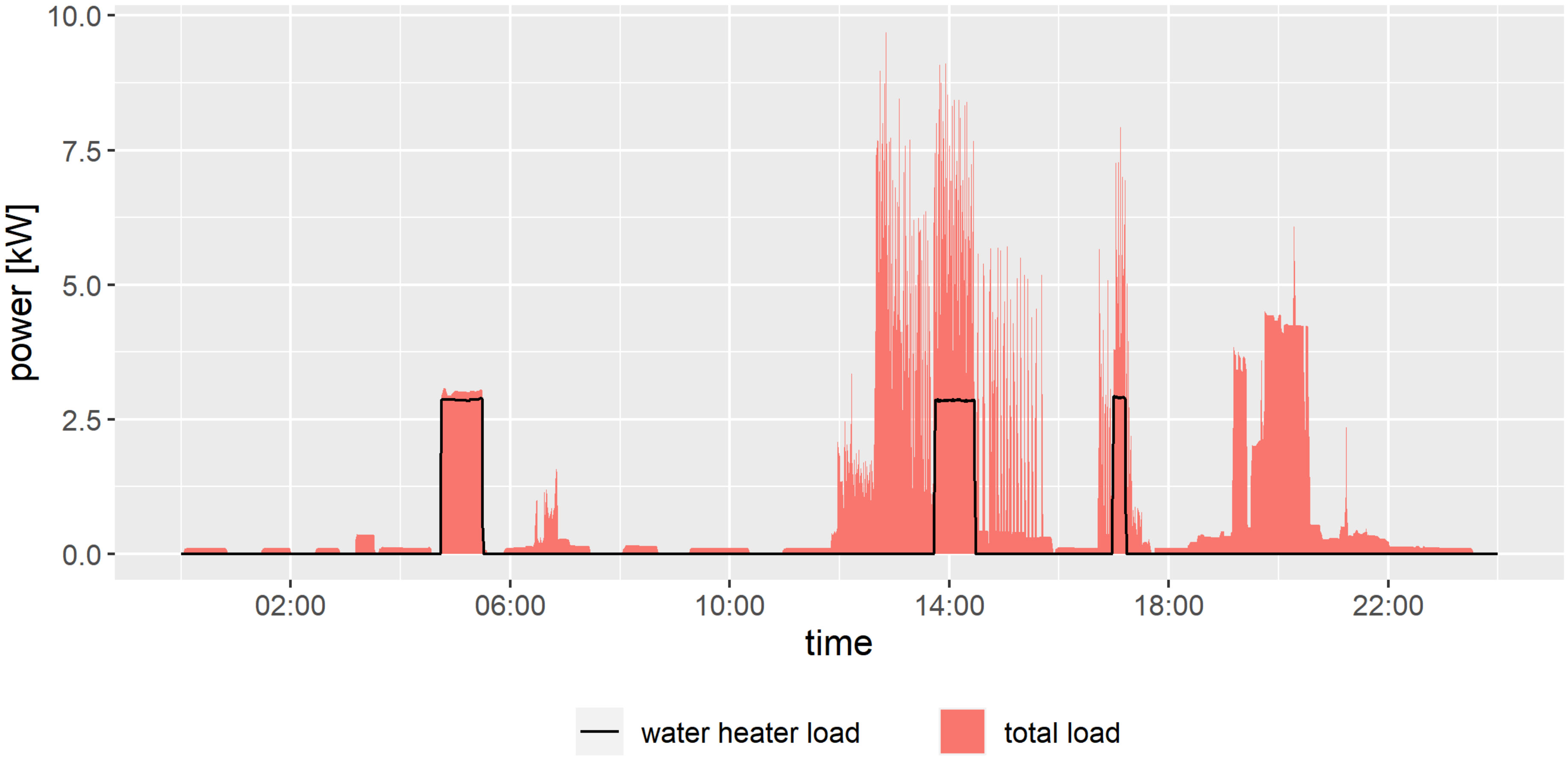
Domestic Appliance	**Existing Literature**	**Power consumption**	**Energy consumption**	**Availability for DR**	**User interaction required for DR**	**Detectable in SM data**
TCL (AC, heat pump)	abundant	medium	high	always	no	no
Water Heater (boiler)	moderate	high	high	always	no	yes
Wet appliance	substantial	medium	medium	sometimes	yes	no
Cold appliance	little	low	medium	always	no	yes
Electric Vehicle	abundant	high	high	sometimes	yes	-

- High flexibility potential of some domestic appliances for Demand Response (DR)
- BUT: Existing data-based approaches to estimate the load of individual appliances are not broadly applicable:
  - Intrusive Load Monitoring (sub-metering): too expensive, data privacy concerns
  - Non-Intrusive Load Monitoring: requires high-resolution data (at least 1 Hz)

➔ Unsupervised load disaggregation based on standard smart meter data (1- to 30-minute resolution)

Detection and Disaggregation of Water Heater Load

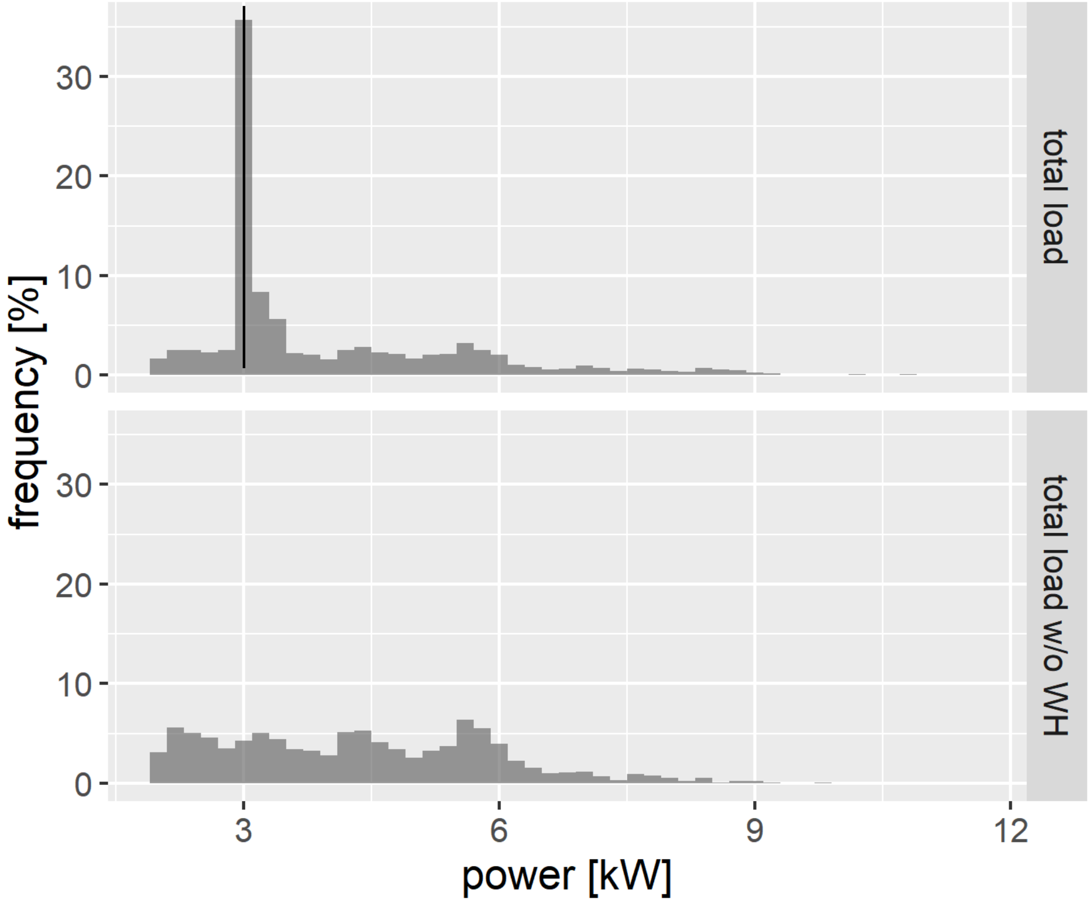
Starting Point



Detection and Disaggregation of Water Heater Load

Detection Process

3 kW = water heater's rated power

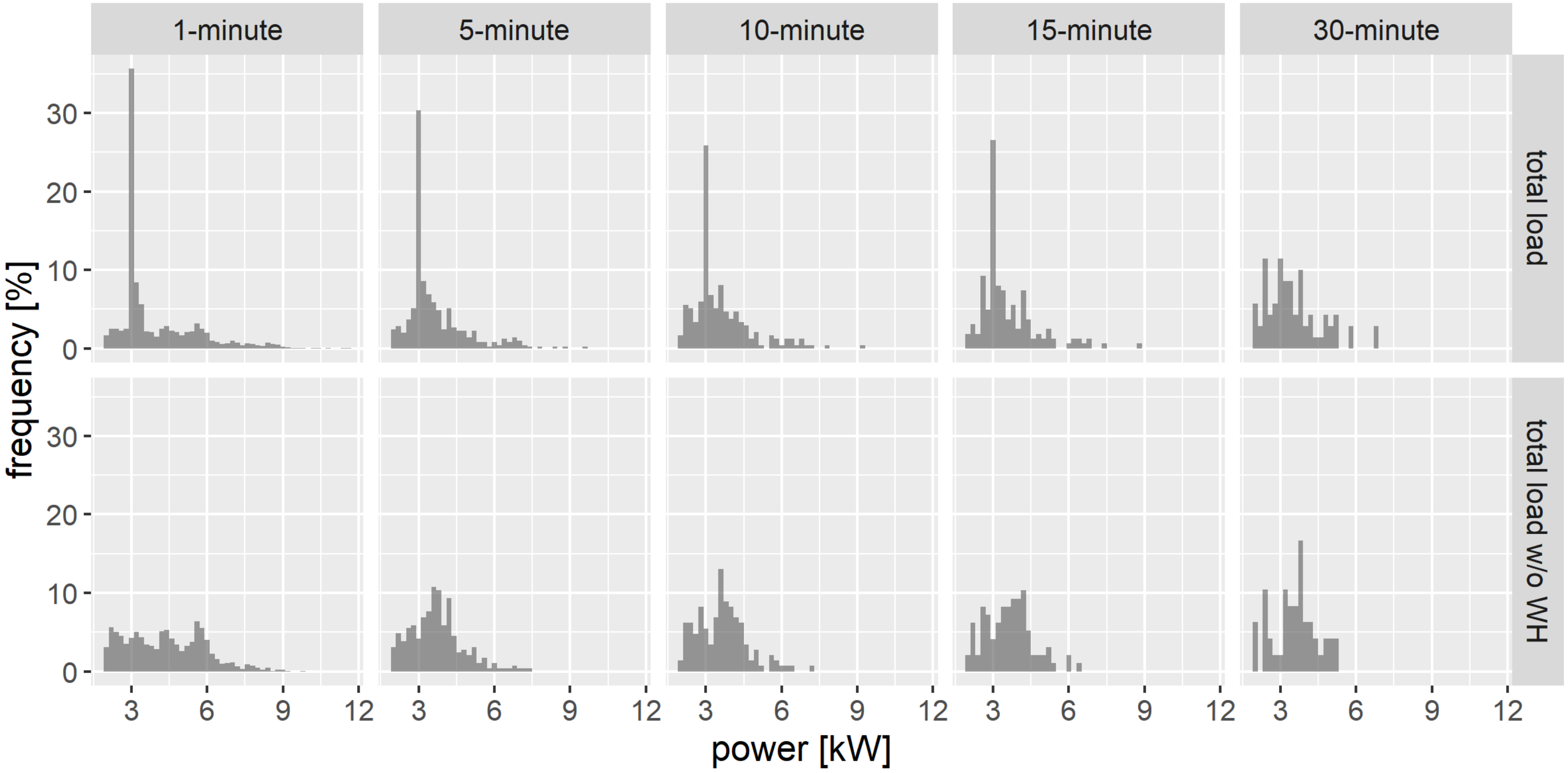


- Bin width of 200 W
- Power higher than 2 kW
- Detection of "outlier bin":

$$d_i^{\text{bin}} \geq \underset{\substack{\uparrow \\ \text{75th percentile in bin frequency values}}}{d_{Q3}^{\text{bin}}} + 1.5 \cdot \underbrace{(d_{Q3}^{\text{bin}} - d_{Q1}^{\text{bin}})}_{\text{interquartile range}}$$

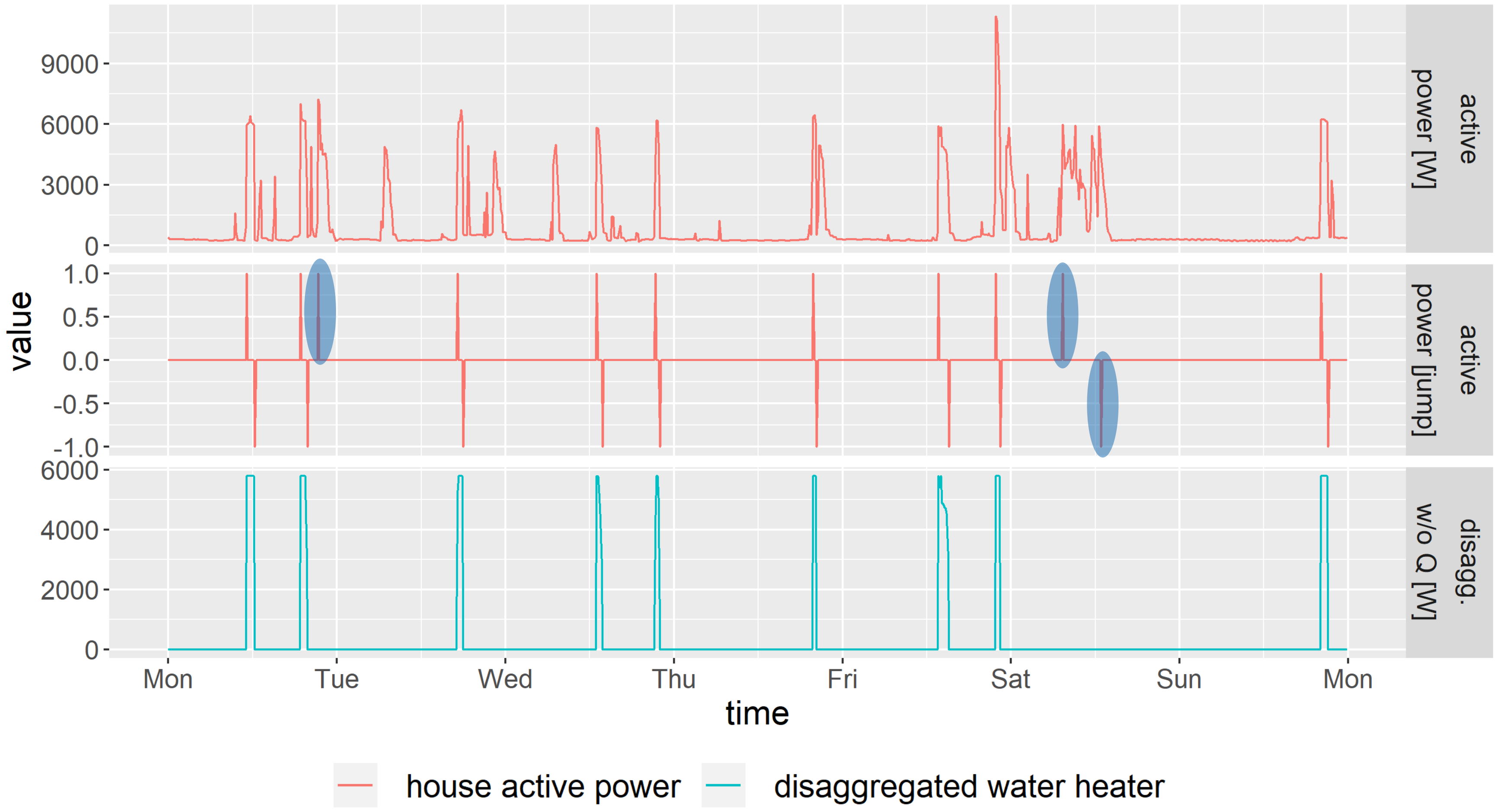
Detection and Disaggregation of Water Heater Load

Detection Process – Influence of Temporal Resolution



Detection and Disaggregation of Water Heater Load

Disaggregation Process



1. Detect jumps (up/down):

   $\Delta p \geq 90\% \cdot \hat{p}_r^{\mathrm{WH}}$ over 2 time steps

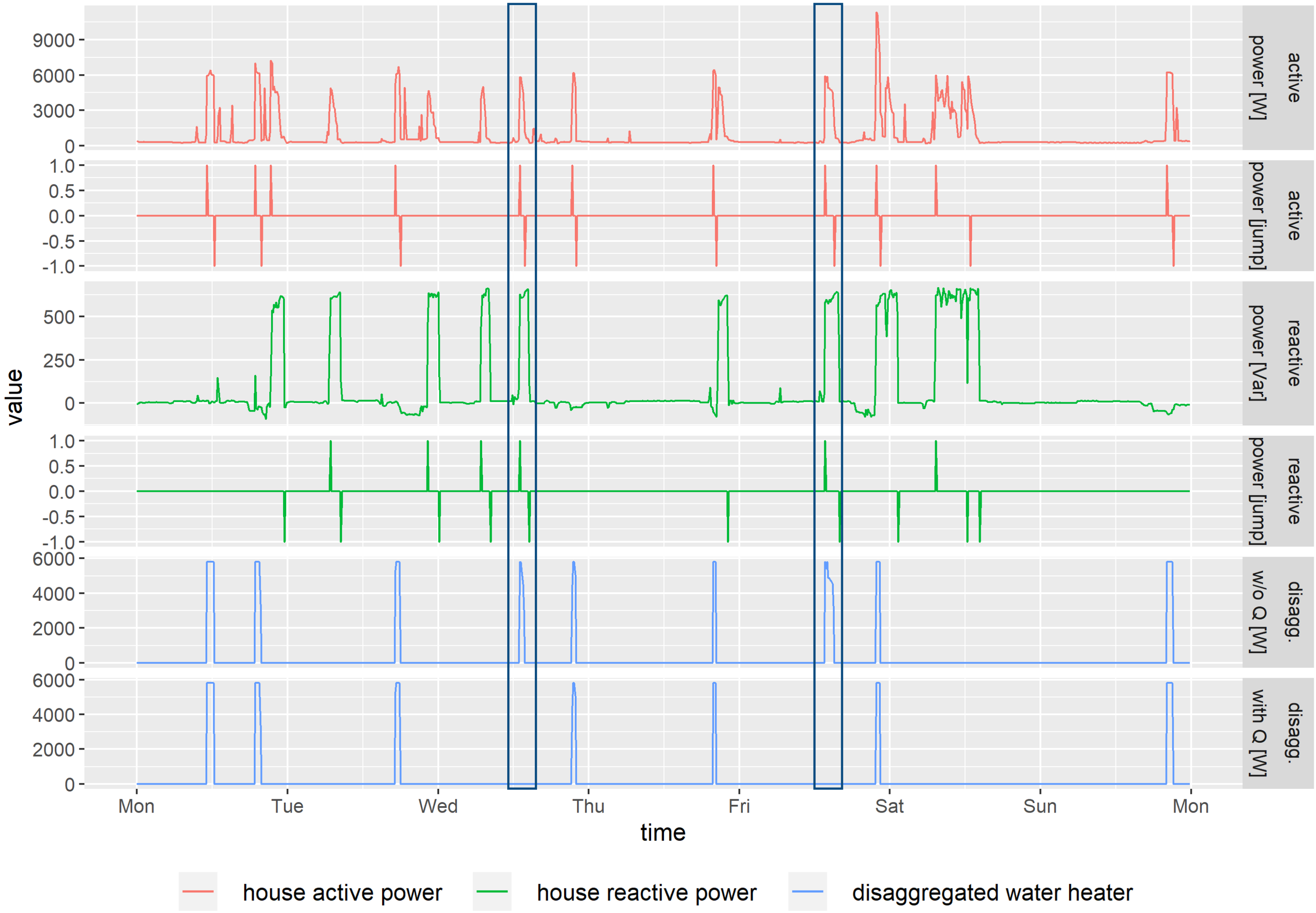
2. Clean jump profile:
   a) Alternation of up/down jumps
   b) max. 2 hours for ON periods

3. Build water heater load profile:

$$p_t^{\mathrm{WH}} = \begin{cases} 0\,, & \text{if } t \in OFF \\ \min\left(\hat{p}_r^{\mathrm{WH}}, p_t^{\mathrm{tot}}\right), & \text{if } t \in \mathrm{ON} \end{cases}$$

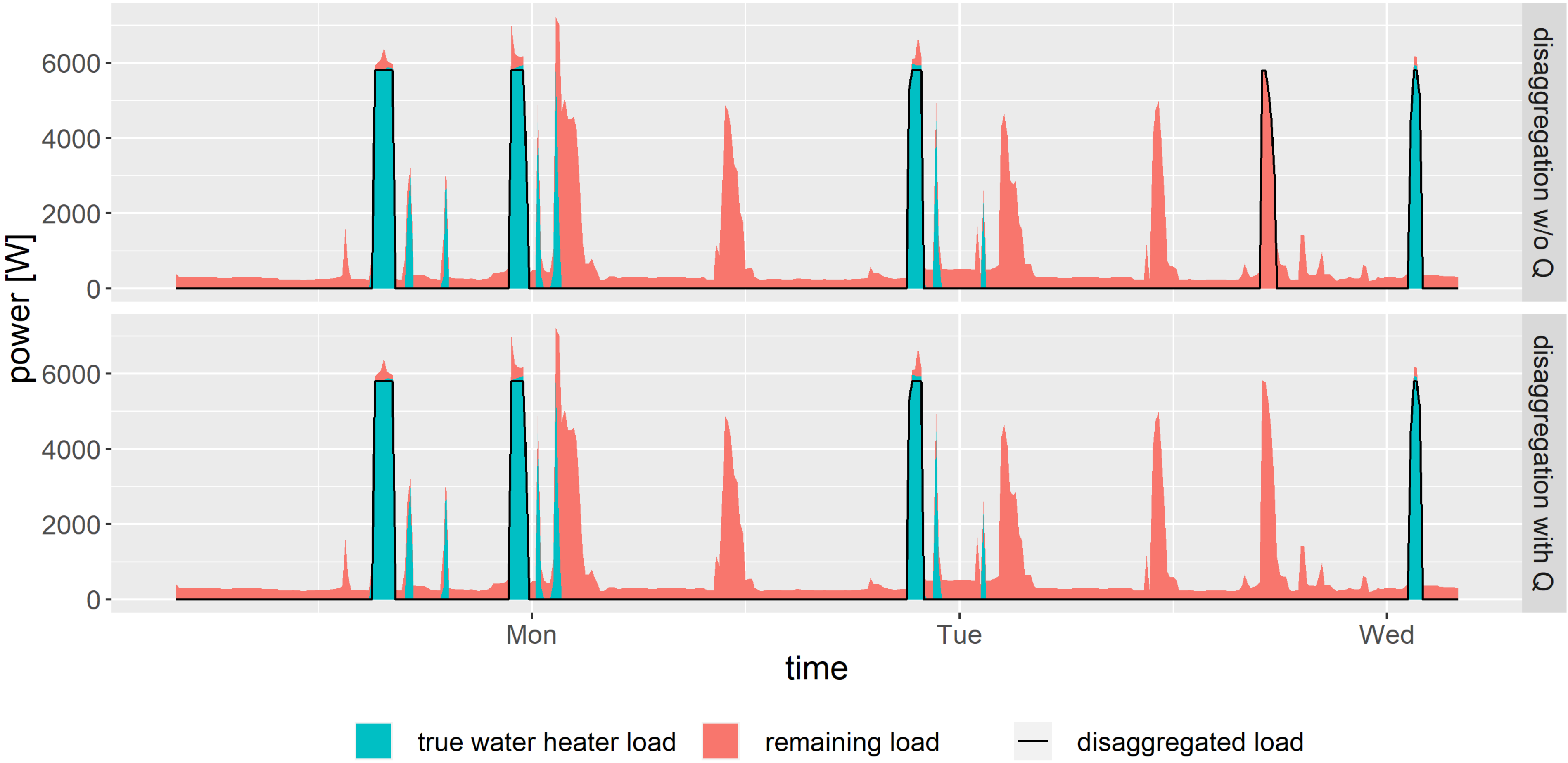
Detection and Disaggregation of Water Heater Load

Disaggregation Process



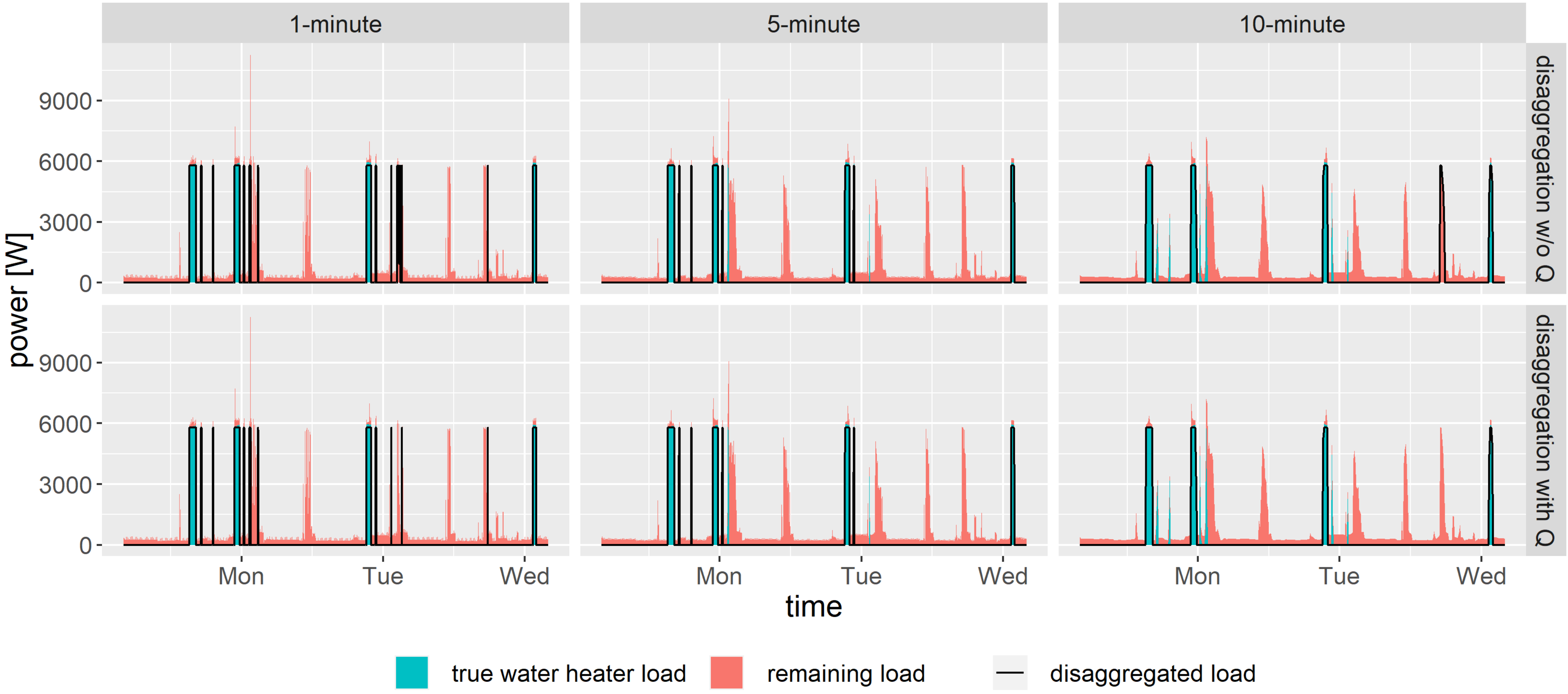
Detection and Disaggregation of Water Heater Load

Disaggregation Result



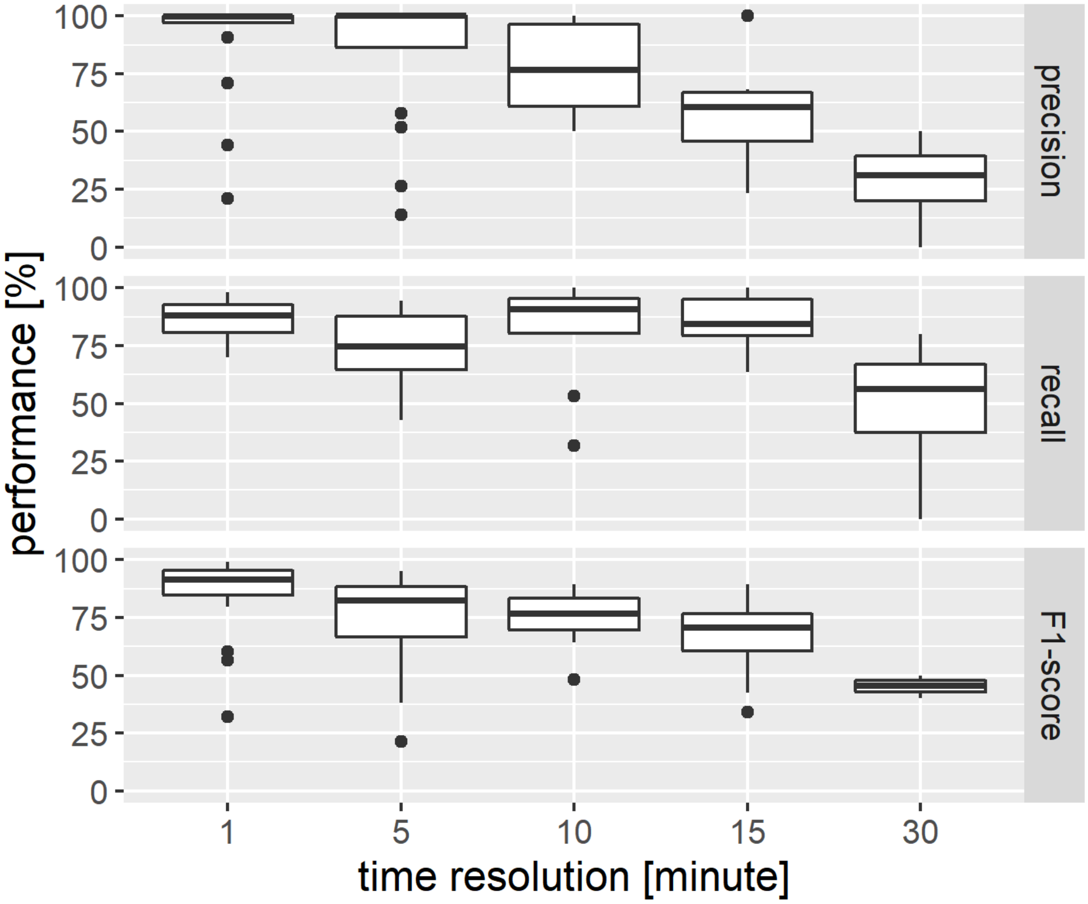
Detection and Disaggregation of Water Heater Load

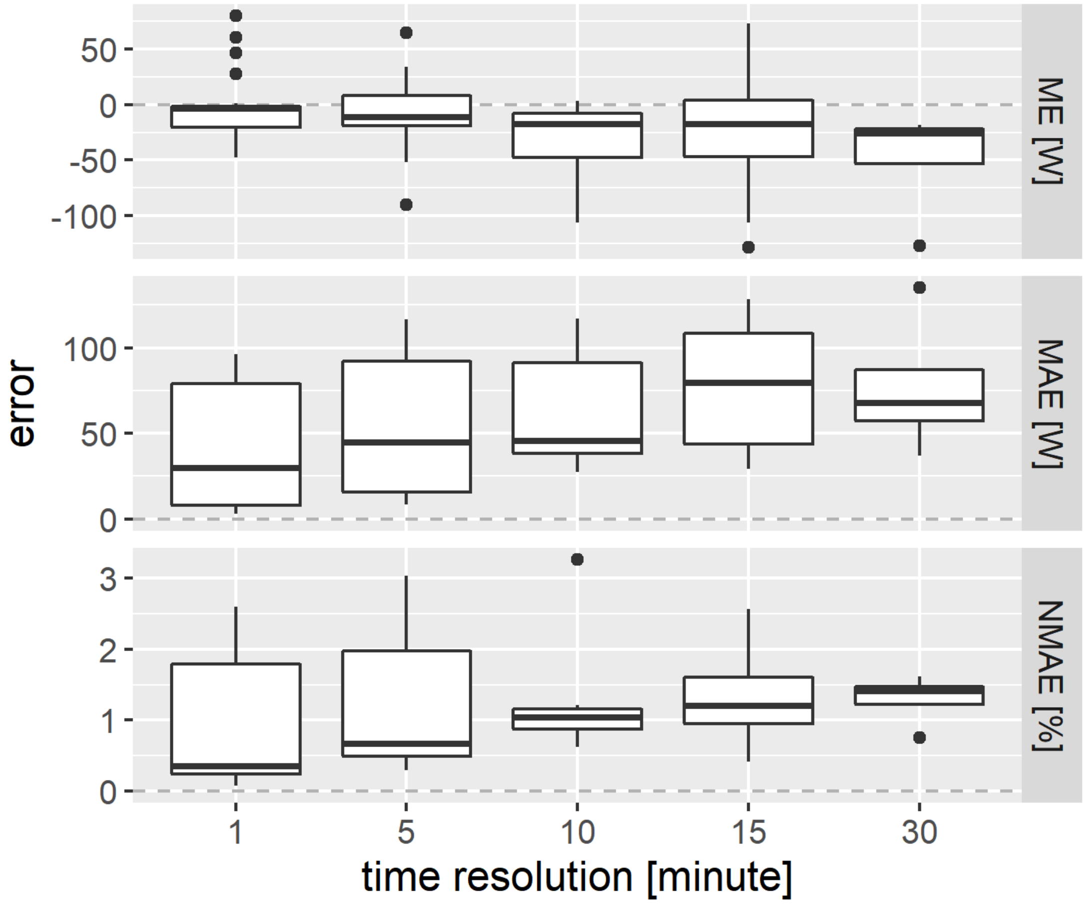
Disaggregation Result – Influence of Temporal Resolution



Detection and Disaggregation of Water Heater Load

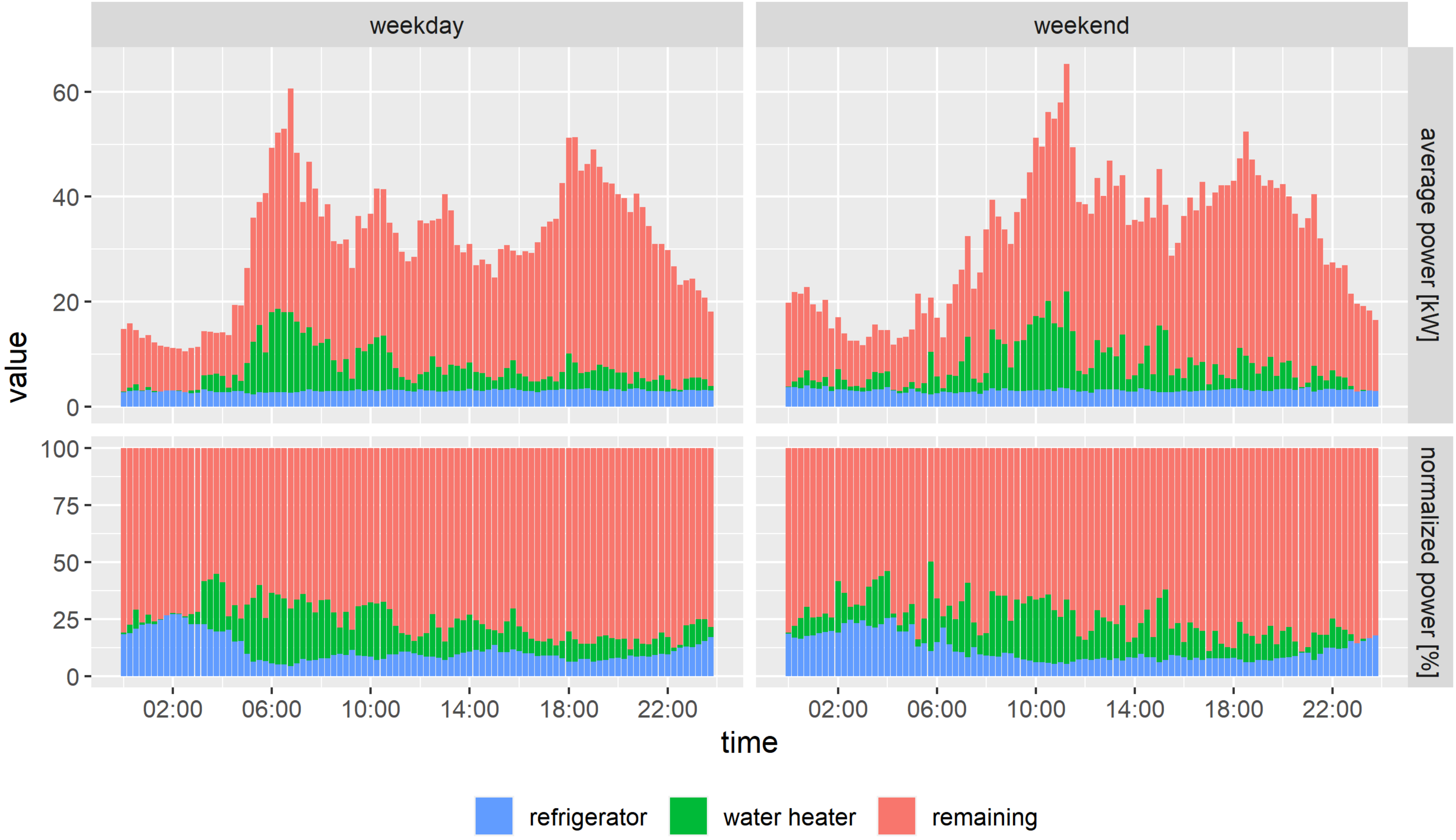
Disaggregation Performance





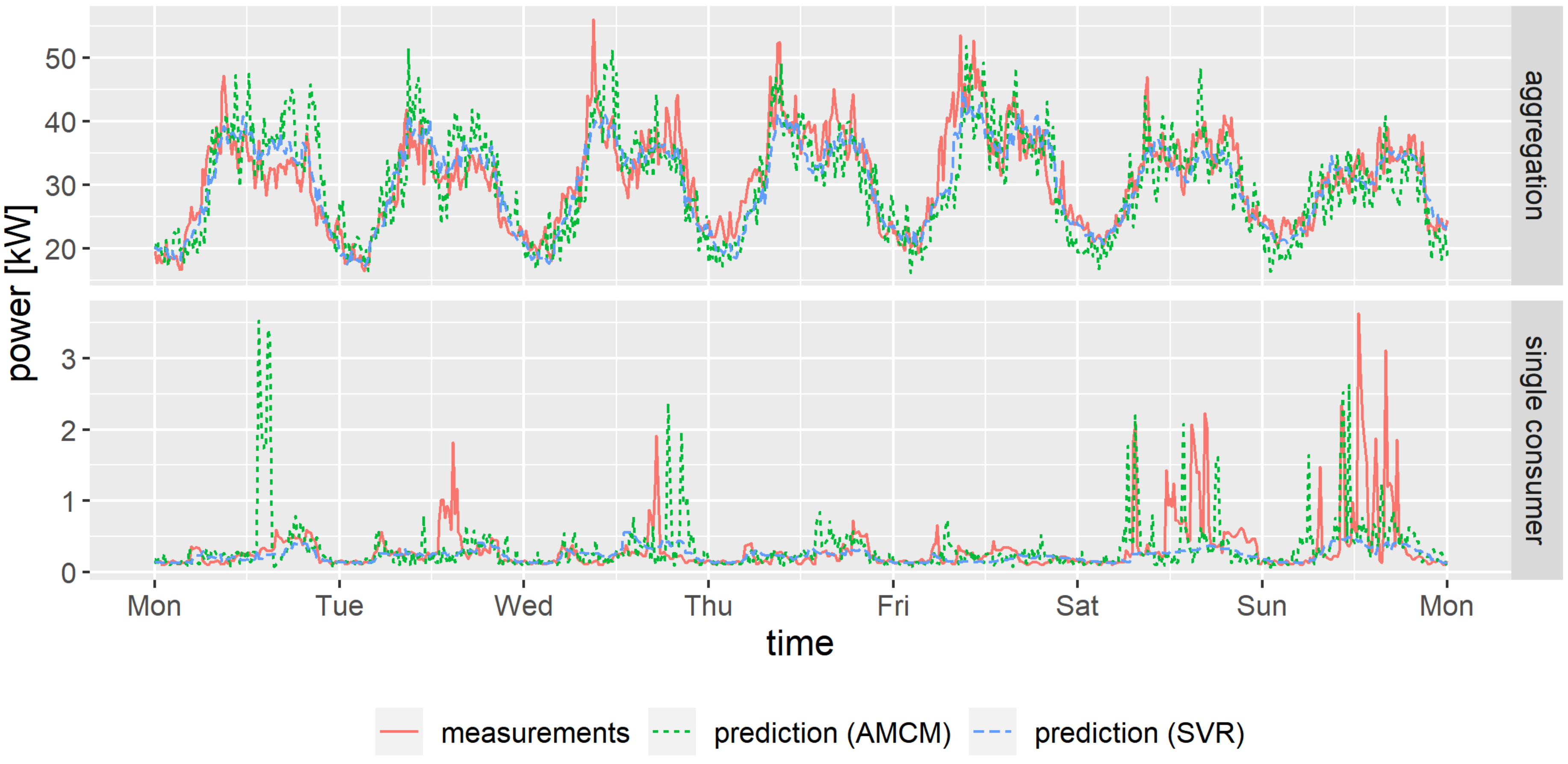
Detection and Disaggregation of Flexible Domestic Loads

Share of Disaggregated Flexible Loads at an Aggregate Level (70 Households)



Short-Term Deterministic Load Forecasting

What is a “good” prediction?



Short-Term Deterministic Load Forecasting

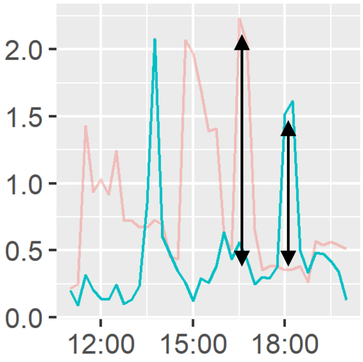
Standard Evaluation Metrics

- Mean Absolute Percentage Error (MAPE)

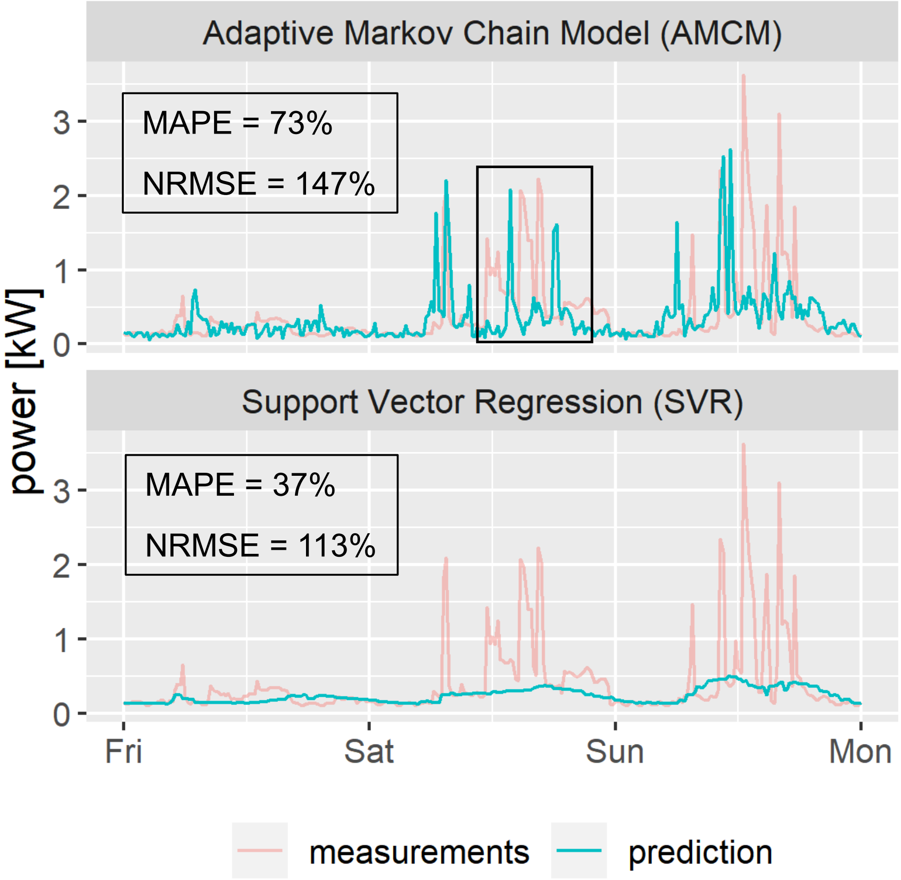
$$\mathrm{MAPE} := \frac{100\%}{T}\sum_{t=1}^{T}\left|\frac{y_t - \hat{y}_t}{y_t}\right|$$

- (Normalized) Root Mean Square Error (NRMSE)

$$\mathrm{NRMSE} := \frac{100\%}{\bar{y}}\sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}$$



double penalty effect



Short-Term Deterministic Load Forecasting

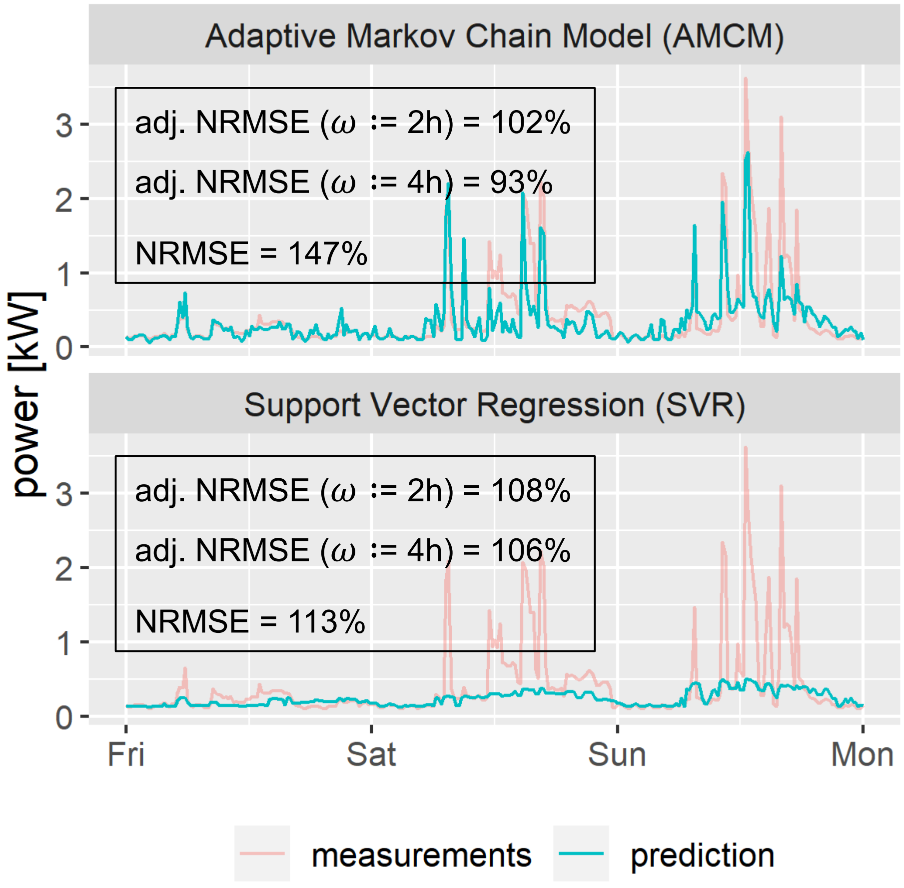
Adjusted Error Metric

- Absolute *p*-norm error: $E_p = \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|_p = \left(\sum_{t=1}^{T} |y_t - \hat{y}_t|^p\right)^{1/p}$
- Adjusted *p*-norm error[1]:

$$E_p^\omega = \min_{P \in \mathcal{P}} \|\boldsymbol{y} - P\hat{\boldsymbol{y}}\|_p$$

$$subject\ to\ P_{ij} = 0, \forall\ |i - j| > \omega$$

$\omega \geq 0 \coloneqq$ adjustment limit
$P \coloneqq$ permutation matrix
$\mathcal{P} \coloneqq$ complete set of restricted permutations

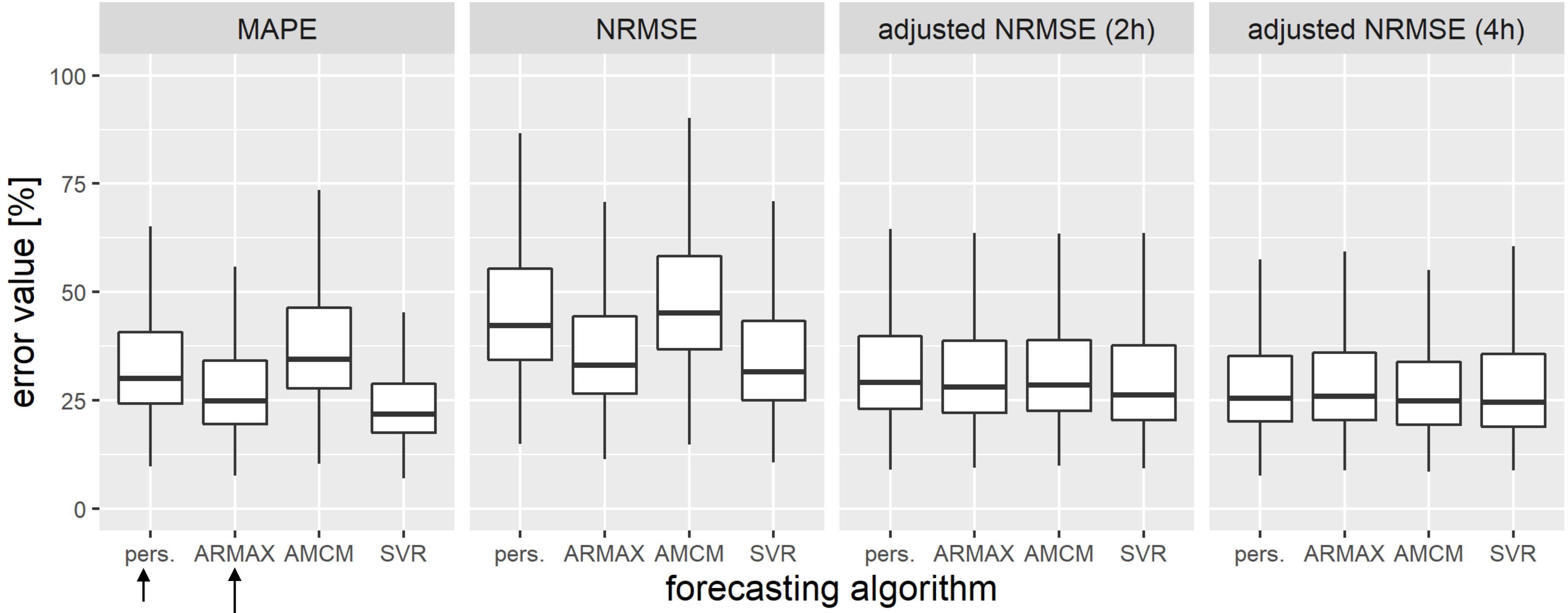


---

[1] S. Haben, J. Ward, D. V. Greetham, C. Singleton, and P. Grindrod, "A new error measure for forecasts of household-level, high resolution electrical energy consumption", *International Journal of Forecasting*, vol. 30, no. 2, pp. 246–256, 2014.

Short-Term Deterministic Load Forecasting

Evaluation of 24h-ahead forecasting on 1000 load profiles



persistence model

autoregressive moving average model with exogenous inputs

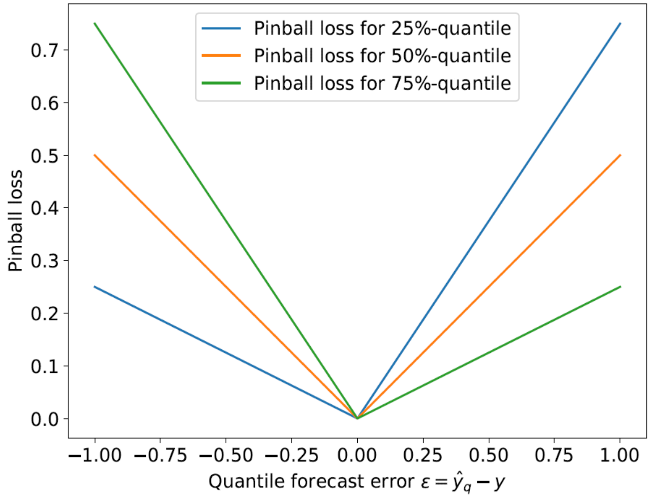
Short-Term Probabilistic State Forecasting

Setup, Pinball Loss, and Quantile Neural Network

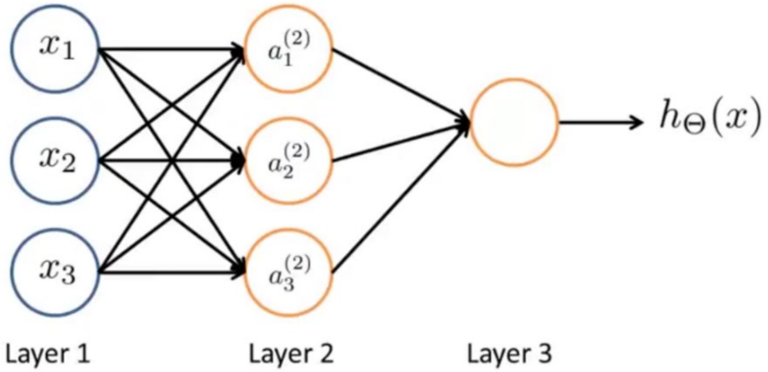
- Hour-ahead prediction of $P_{\text{cons}}, Q_{\text{cons}}, V, P_{\text{flow}}, Q_{\text{flow}}$ in a low-voltage grid
- One ML model per grid component (i.e., line and bus), per quantity and per quantile

Pinball Loss

$$J_q = \frac{1}{N}\sum_{n=1}^{N}\begin{cases}(y_n - \hat{y}_{q,n})\,q & \text{if } y_n \geq \hat{y}_{q,n}\\ (\hat{y}_{q,n} - y_n)\,(1-q) & \text{if } y_n < \hat{y}_{q,n}\end{cases}$$



Quantile Neural Network

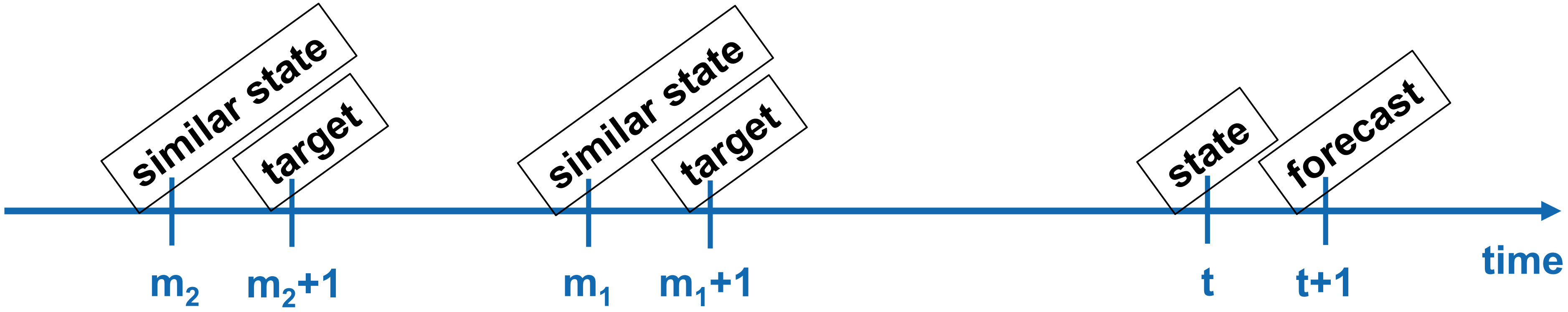


Forecast = Combination of input features

Training process = minimizing pinball loss

Short-Term Probabilistic State Forecasting

Proposed Quantile K-Nearest Neighbor



Forecast = linear combination of target values related to similar states

Deterministic Forecast

$$J_{\text{knn}} = \min_{\boldsymbol{w}^{\text{knn}}} \frac{1}{N} \sum_{n=1}^{N} \overbrace{\left|(\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} - y_n\right|}^{\text{MAE}}$$

$$\text{subject to} \quad w_i^{\text{knn}} \geq 0,\ \forall i \in \{0, \dots, k\},$$

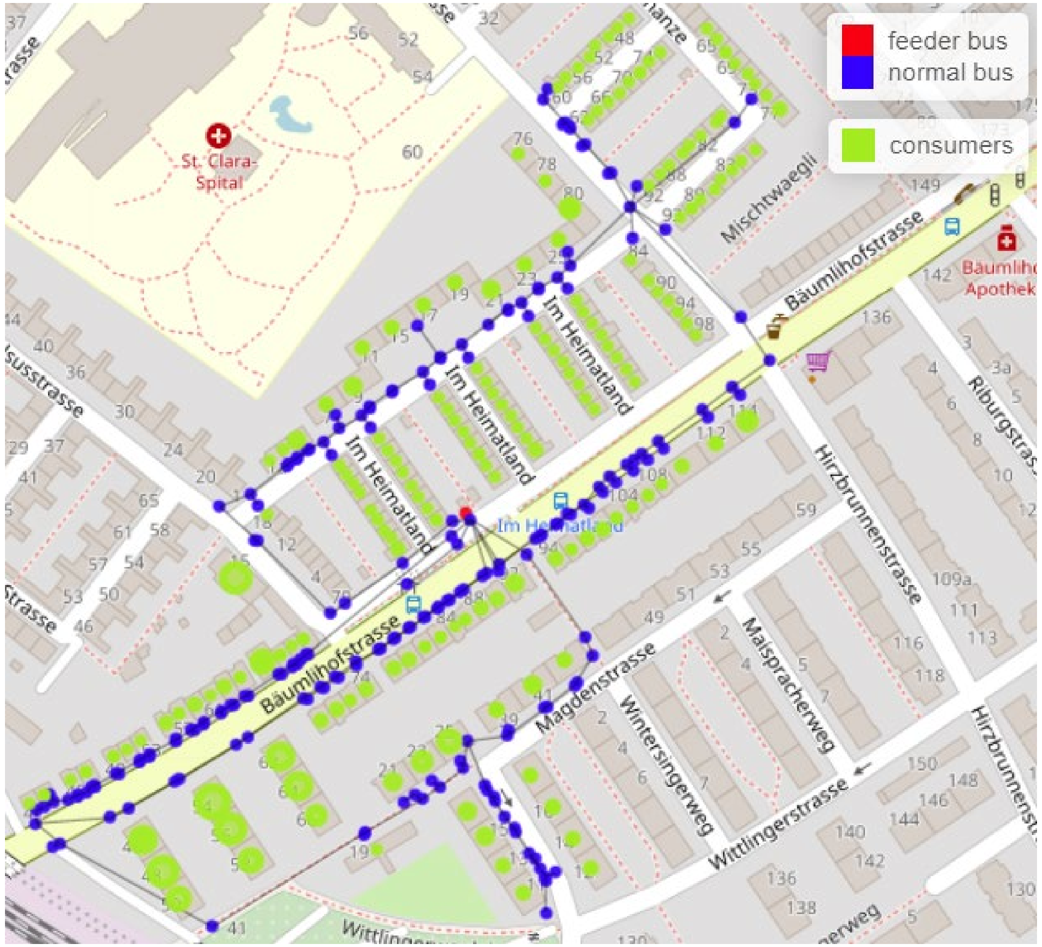
Quantile Forecast

MAE replaced by pinball loss

$$J_{\text{knn}} = \min_{\boldsymbol{w}^{\text{knn}}} \frac{1}{N} \sum_{n=1}^{N} \begin{cases} ((\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} - y_n)\, q & \text{if } y_n \geq (\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} \\ ((\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} - y_n)\,(1-q) & \text{if } y_n < (\boldsymbol{w}^{\text{knn}})^T \boldsymbol{y}_n^{\text{knn}} \end{cases}$$

Short-Term Probabilistic State Forecasting

Case Study



Grid Topology

- Weakly meshed LV-grid in the City of Basel
- 198 lines, 196 buses
- 583 smart metered residential consumers (15-min. resolution)
- 60% of houses equipped with PV system
- 30% of households equipped with EV

Assumptions

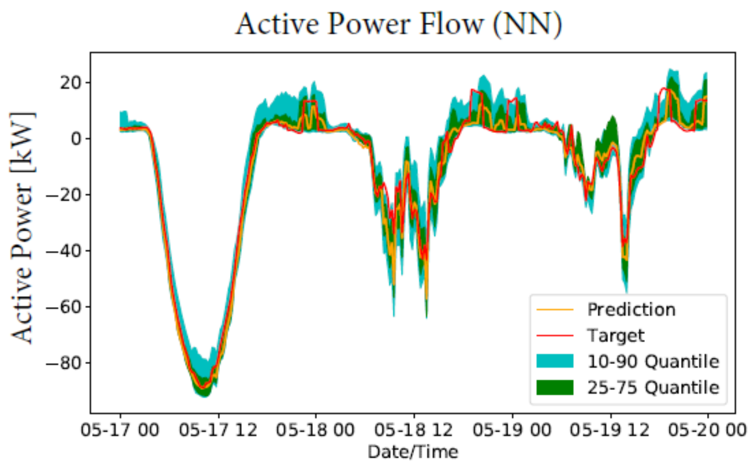
- Perfect knowledge of grid topology and parameters
- Voltage-independent loads and PV systems
- Feeder voltage maintained after optimization

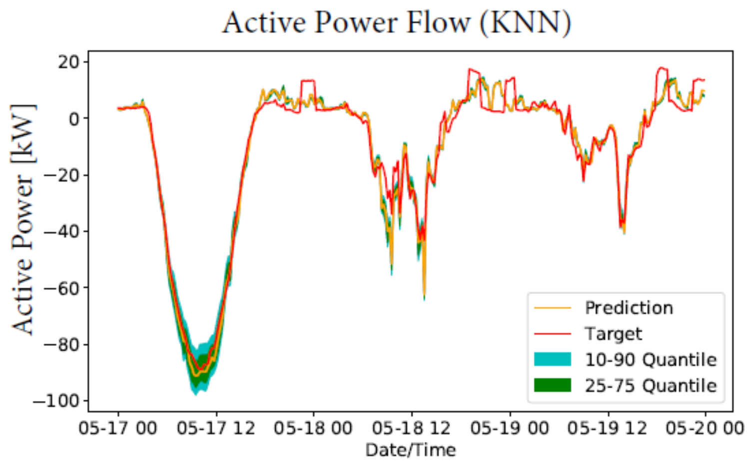
Training = 42 weeks

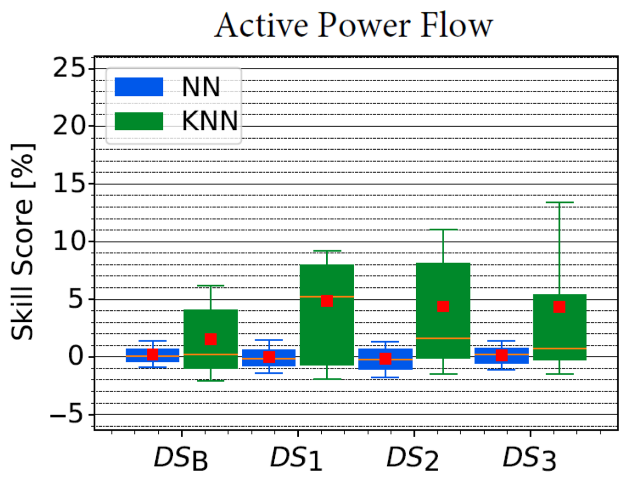
Testing = 10 weeks

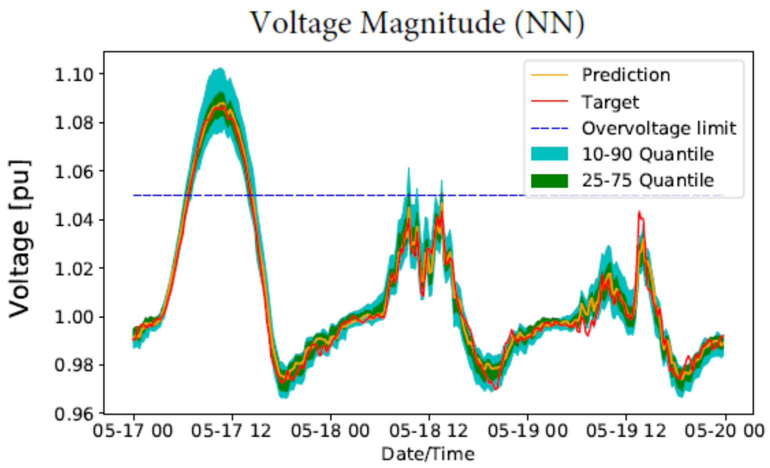
Short-Term Probabilistic State Forecasting

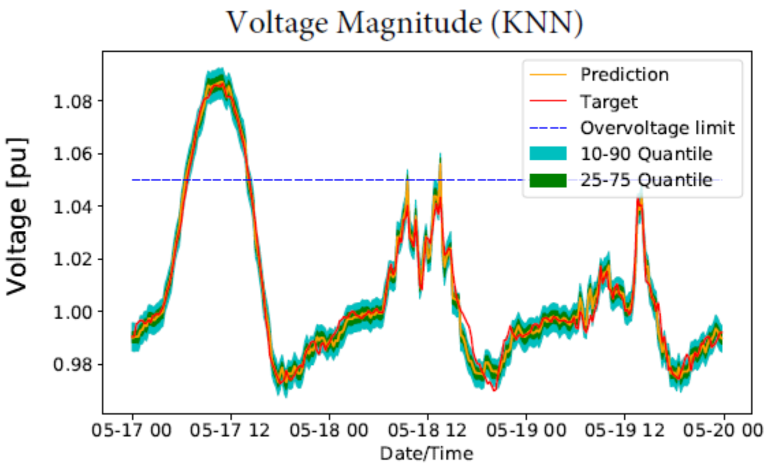
Performance Evaluation

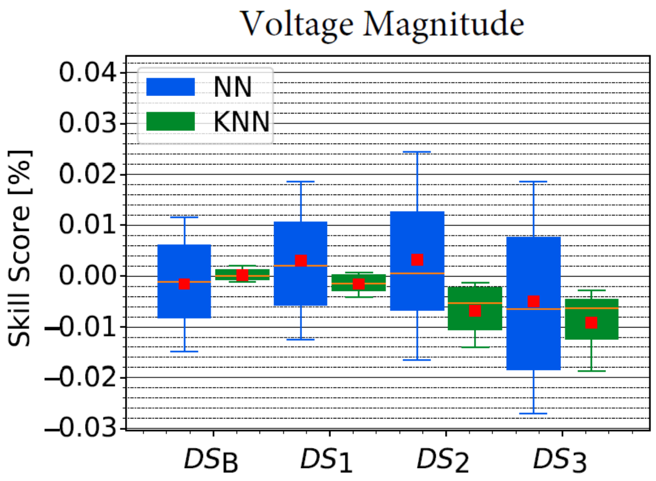












Preventive Voltage Control under Uncertainty

Real-Time Optimization Strategy (Benchmark)

Real-time: If overvoltage detected ➔ real-time optimization

$$\min_{P^{\text{curt}}_{n_{\text{PV}}}} \sum_{\Psi_{\text{PV}}} \frac{1}{4} C_{\text{ib}} P^{\text{curt}}_{n_{\text{PV}}},$$

Minimize cost of active power PV curtailment in real-time

subject to

$$0 \leq P^{\text{curt}}_{n_{\text{PV}}} \leq P^{\text{prod}}_{n_{\text{PV}}}, \ \forall n_{\text{PV}} \in \Psi_{\text{PV}},$$

Limit on active power PV curtailment

$$P_k + \sum_{\Psi_{\text{PV},k}} P^{\text{curt}}_{n_{\text{PV}}} + \sum_{m \in \Omega_k} P_{km} = 0, \ \forall k \in \Psi_{\text{B}},$$

Active power node balance

$$P_{km} = (V_k)^2 g_{km} - V_k V_m (g_{km} \cos \theta_{km} + b_{km} \sin \theta_{km}), \ \forall k, m \in \Psi_{\text{B}}$$

AC active power flow equation

$$Q_k + \sum_{m \in \Omega_k} Q_{km} = 0, \ \forall k \in \Psi_{\text{B}},$$

Reactive power node balance

$$Q_{km} = -(V_k)^2 b_{km} + V_k V_m (b_{km} \cos \theta_{km} - g_{km} \sin \theta_{km}), \ \forall k, m \in \Psi_{\text{B}}$$

AC reactive power flow equation

$$0.95 \leq V_k \leq 1.05, \ \forall k \in \Psi_{\text{B}},$$

Limit on voltage magnitude

$$V_1 = V_1^{\text{meas}},$$
$$\theta_1 = 0,$$

Voltage magnitude and angle at slack bus

Preventive Voltage Control under Uncertainty

Optimization Strategy using Point Forecasts (50%-Quantile)

Hour-ahead: If overvoltage predicted ➔ optimization strategy using point forecasts

$$\boxed{\min_{P^{\text{curtHA}}_{n\text{PV}}} \sum_{\Psi_{\text{PV}}} \frac{1}{4} C_{\text{m}} P^{\text{curtHA}}_{n\text{PV}},}$$

Minimize cost of active power PV curtailment hour-ahead

subject to

$$0 \leq P^{\text{curtHA}}_{n\text{PV}} \leq \hat{P}^{\text{prod}}_{n\text{PV}}, \forall n_{\text{PV}} \in \Psi_{\text{PV}},$$

Limit on active power PV curtailment

$$\hat{P}_{50,k} + \sum_{\Psi_{\text{PV},k}} P^{\text{curtHA}}_{n\text{PV}} + \sum_{m \in \Omega_k} P_{km} = 0, \forall k \in \Psi_{\text{B}},$$

Active power node balance

$$P_{km} = (V_k)^2 g_{km} - V_k V_m (g_{km} \cos \theta_{km} + b_{km} \sin \theta_{km}) \ \forall k, m \in \Psi_{\text{B}},$$

AC active power flow equation

$$\hat{Q}_{50,k} + \sum_{m \in \Omega_k} Q_{km} = 0, \ \forall k \in \Psi_{\text{B}},$$

Reactive power node balance

$$Q_{km} = -(V_k)^2 b_{km} + V_k V_m (b_{km} \cos \theta_{km} - g_{km} \sin \theta_{km}) \ \forall k, m \in \Psi_{\text{B}},$$

AC reactive power flow equation

$$0.95 \leq V_k \leq 1.05, \ \forall k \in \Psi_{\text{B}},$$

Limit on voltage magnitude

$$V_1 = \hat{V}_{50,1}, \quad \theta_1 = 0,$$

Voltage magnitude and angle at slack bus

Real-time:

If overvoltage still detected
➔ real-time optimization

Preventive Voltage Control under Uncertainty

Optimization Strategy using Point Forecasts (50%-Quantile)

Hour-ahead: If overvoltage predicted in quantile q ➔ optimization strategy using quantile forecasts

Voltage forecast uncertainty: $d\hat{V}_{q,k} = \hat{V}_{q,k} - \hat{V}_{50,k}, \ \forall q \in \ (0, 100), \ \forall k \in \Psi_{\mathrm{B}}$

Minimize cost of PV active power curtailment hour-ahead

subject to

Limit on PV active power curtailment

Active/reactive power node balance

AC active/reactive power flow equations

Limit on voltage magnitude: $0.95 \leq V_k + \boxed{d\hat{V}_{q,k}} \leq 1.05, \ \forall k \in \Psi_{\mathrm{B}}$

Voltage magnitude at slack bus: $V_1 = \hat{V}_{50,1} - \boxed{d\hat{V}_{q,1}}$

Angle at slack bus

Real-time:

If overvoltage still detected
➔ real-time optimization

Preventive Voltage Control under Uncertainty

Cost of Optimal Voltage Control Strategies

Voltage control strategies

- S_{B} = RT power curtailment
- S_{50} = HA power curtailment using point forecasts + RT power curtailment
- S_{44} = HA power curtailment using 44%-quantile forecasts + RT power curtailment
- $S_{62.5}$ = HA power curtailment using 62.5%-quantile forecasts + RT power curtailment

Cost for active power curtailment

- Situation 1: $C_{\mathrm{m}} = 40\frac{€}{\mathrm{MWh}}$; $C_{\mathrm{ib}} = 63.8\frac{€}{\mathrm{MWh}}$
- Situation 2: $C_{\mathrm{m}} = 40\frac{€}{\mathrm{MWh}}$; $C_{\mathrm{ib}} = 127.6\frac{€}{\mathrm{MWh}}$

	S_{B}	S_{50}	S_{44}	$S_{62.5}$
HA power curtailment [MWh]	0.0	22.3	21.6	23.7
RT power curtailment [MWh]	22.3	2.1	2.4	1.5
Total power curtailment [MWh]	22.3	24.4	24	25.2
Total cost in situation 1 [€]	1422	1022	1017	1044
Total cost in situation 2 [€]	2843	1154	1172	1142

Conclusions and Outlook

- Although the Advanced Metering Infrastructure (AMI) is still under development in most distribution grids, the generated data already offers excellent opportunities for power utilities and their customers.

- BUT: There is a large gap between state-of-the-art practices in the power industry and the various data-driven applications proposed in the literature:
  - The power industry is conservative (principally trusts well-known and time-tested solutions) and lacks expertise to properly make use of the gathered data.
  - Literature often relies on unrealistic assumptions (full system observability, perfect data quality, same approaches as in the transmission level, case studies based on synthetic data, etc.).
  - Data protection and privacy concerns prevent the implementation of intrusive applications.

➔ Closer collaboration between the energy sector, start-up companies, and the scientific community.

➔ Set-up of real-world demonstrators (e.g., NEST on Empa site)

THANK YOU!

Thierry Zufferey
Doctoral Researcher
Power Systems Laboratory

thierryz@eeh.ee.ethz.ch