
Customer classification based on heat load pattern recognition

Smart Energy Systems and 4th Generation District Heating,

Aalborg, 13-14 November 2018

Morten Karstoft Rasmussen, Data Scientist, Kamstrup



Morten Karstoft Rasmussen

Data Scientist

Central Region, Denmark

Add profile section ▼

More...



Kamstrup



Aarhus Universitet



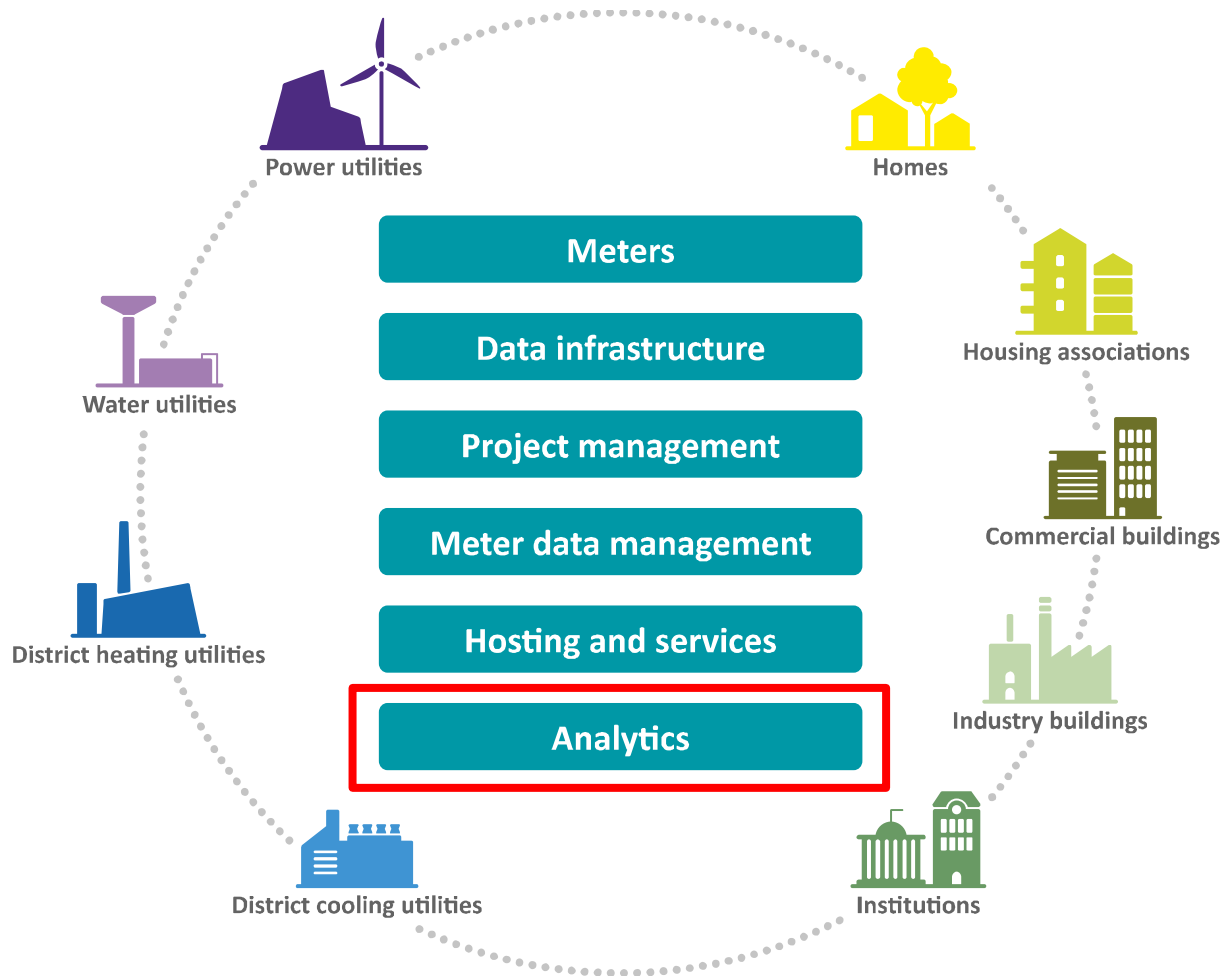
See contact info



See connections (500+)

Experiences include: - Project management - Consultancy - Metrology, test, measurement and calibration
- Numerical programming and applied statistics - Teaching

Our products and solutions



Imagine ...

kamstrup

What would it mean for a utility to get a **better overview** of the district heating network?

What if measured data could be used for **identifying distribution network efficiency**?

Imagine having a possibility to find **faulty substations**?

What if you could **visualize meter data** together with your distribution network?

Imagine having actual data and **no more hidden databases**?

What if you could **identify your heat loss** based on data you already have?



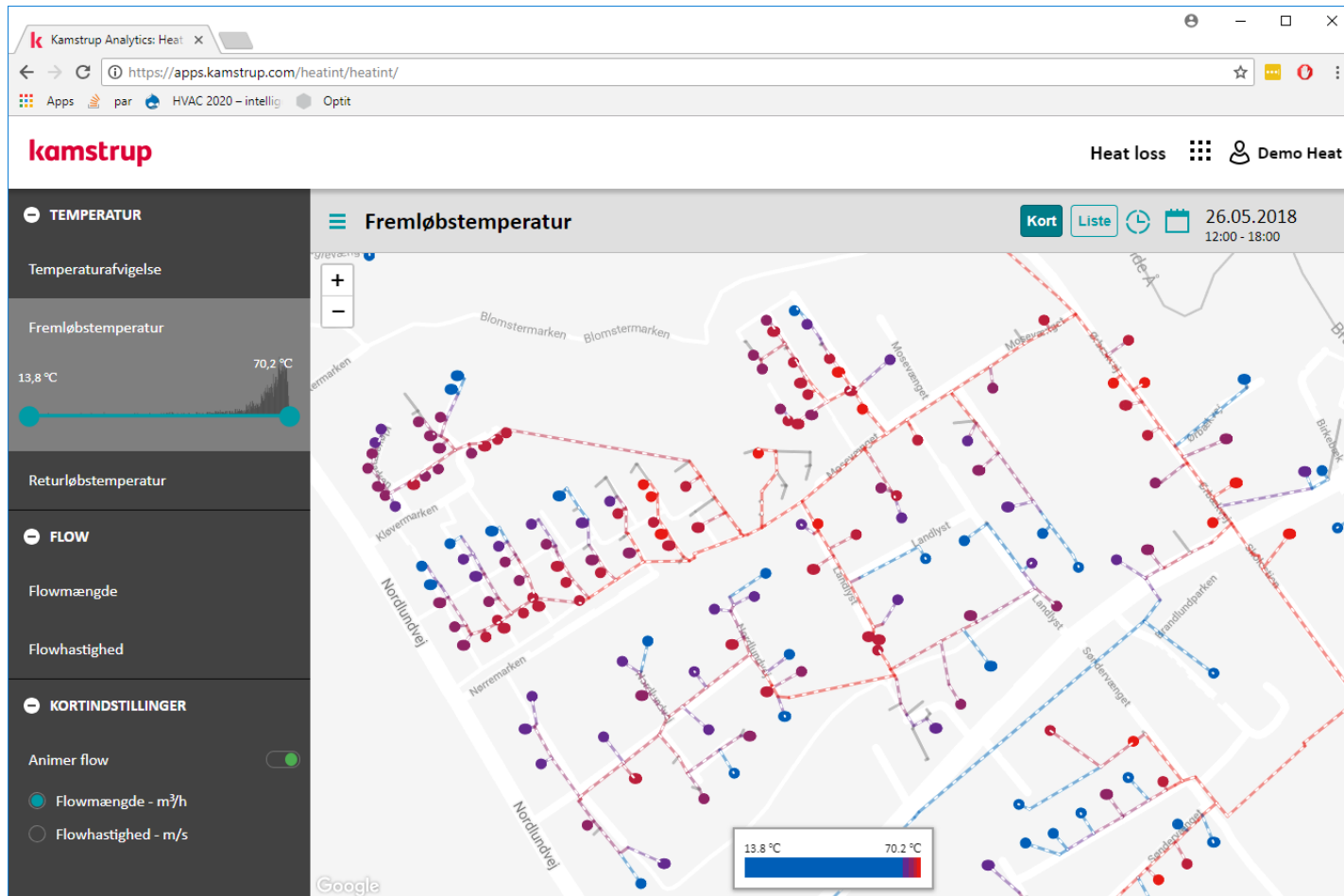
Digitalised District Heating

Digitalisation creates transparency
and reduces losses

Because you cannot optimise what
you do not measure



Heat Intelligence – monitor network and locate heat losses
(among other things)



use of smart meter data to look into the buildings and consumer behavior:

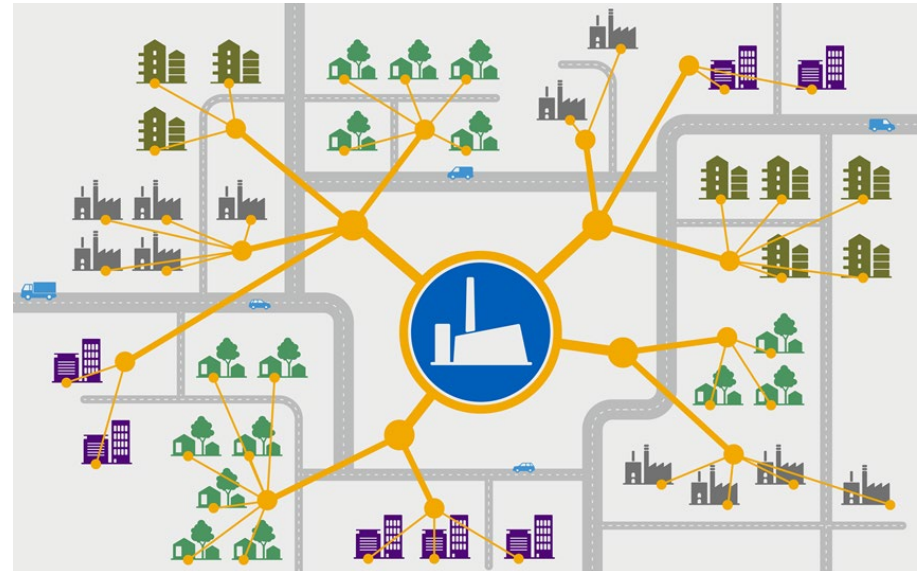
- many substations are faulty (up to 75 %) which stresses the network and cause high heat losses
- Frequent data and automated analytical tools can provide overview of this faulty or inappropriate behavior
- Reveal what are the causes and where to optimize
- identify building characteristics and end user behavior that stresses the network most
- show the “bad” buildings on a map and point at where to prioritise resources

• Benefits in relation to customers:

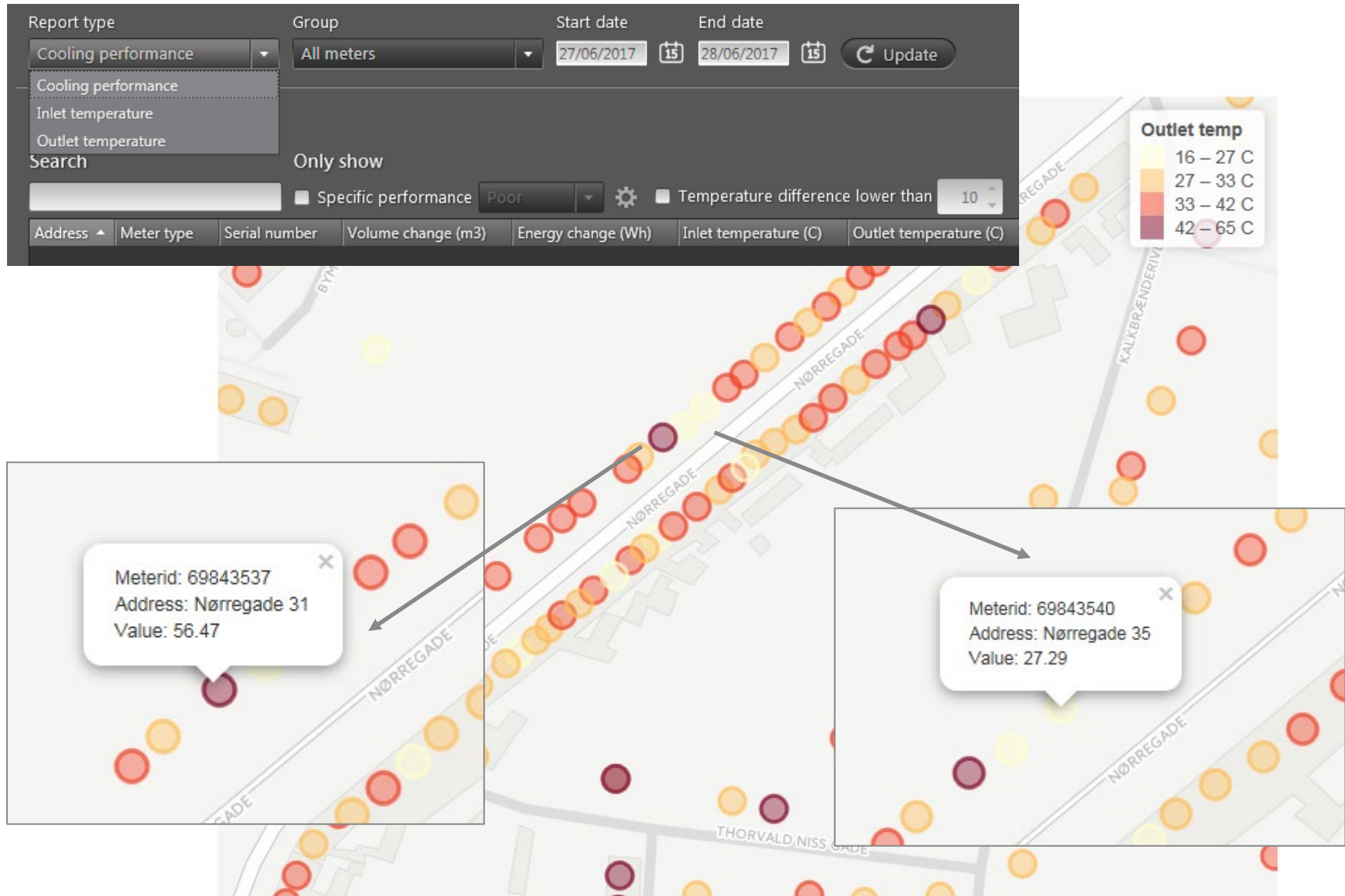
- get closer to customers
- fact based relevant dialog with customers
- targeted advice and guidance
- reduce customers heating bill
- optimize network performance



- Use of meter data to
 - classify consumers depending on consumption pattern
 - Input for forecasting, peak shaving, **building optimization, end user behavior**, calculating coincidence factors etc.
- For optimization of the distribution network



Find the buildings that stresses the network most

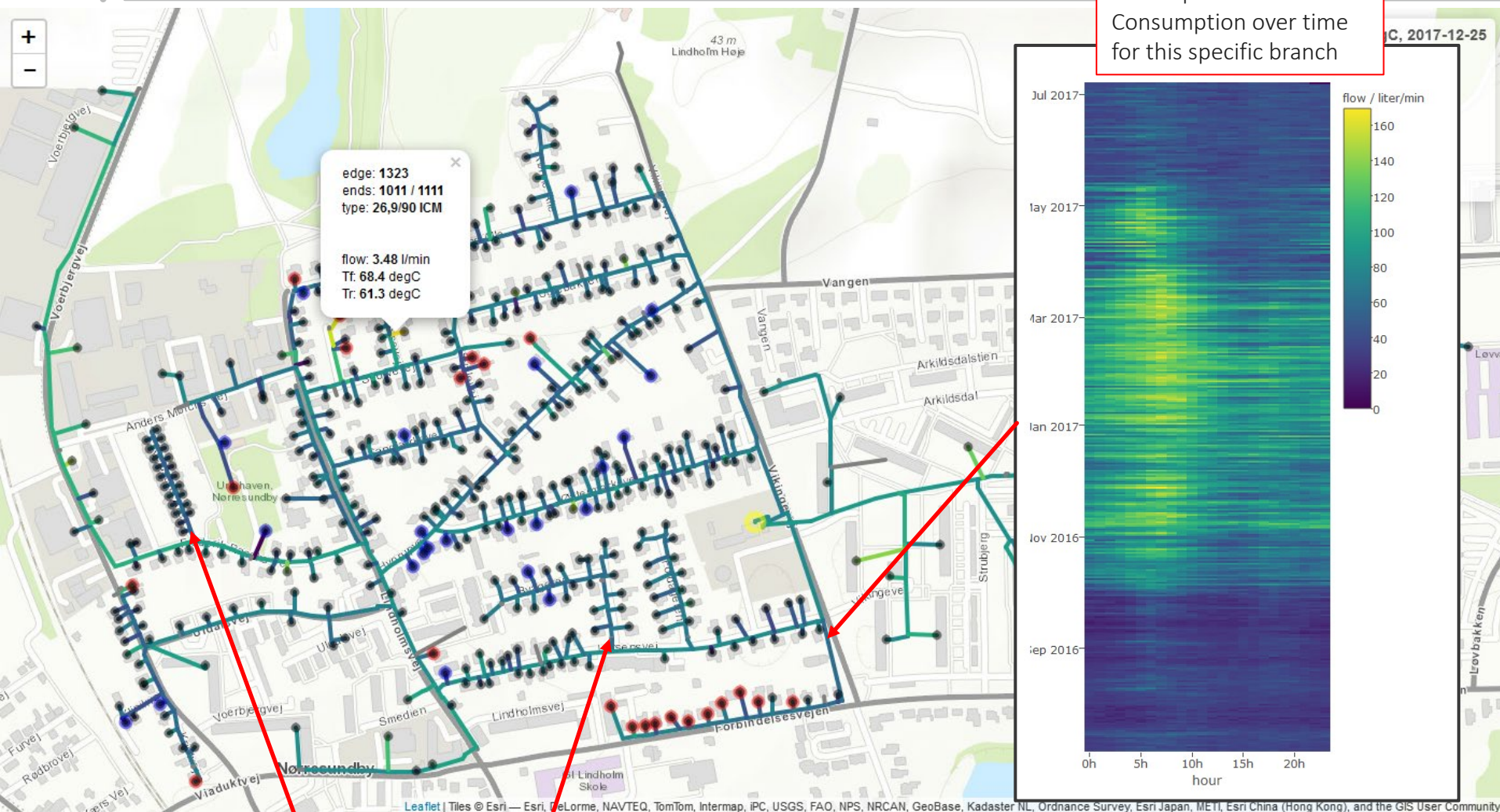


A building's true performance



Aggregated consumption in network subsection

kamstrup



cf = 0.91
22 consumers

cf = 0.65
22 consumers

Identify faulty or misadjusted substations

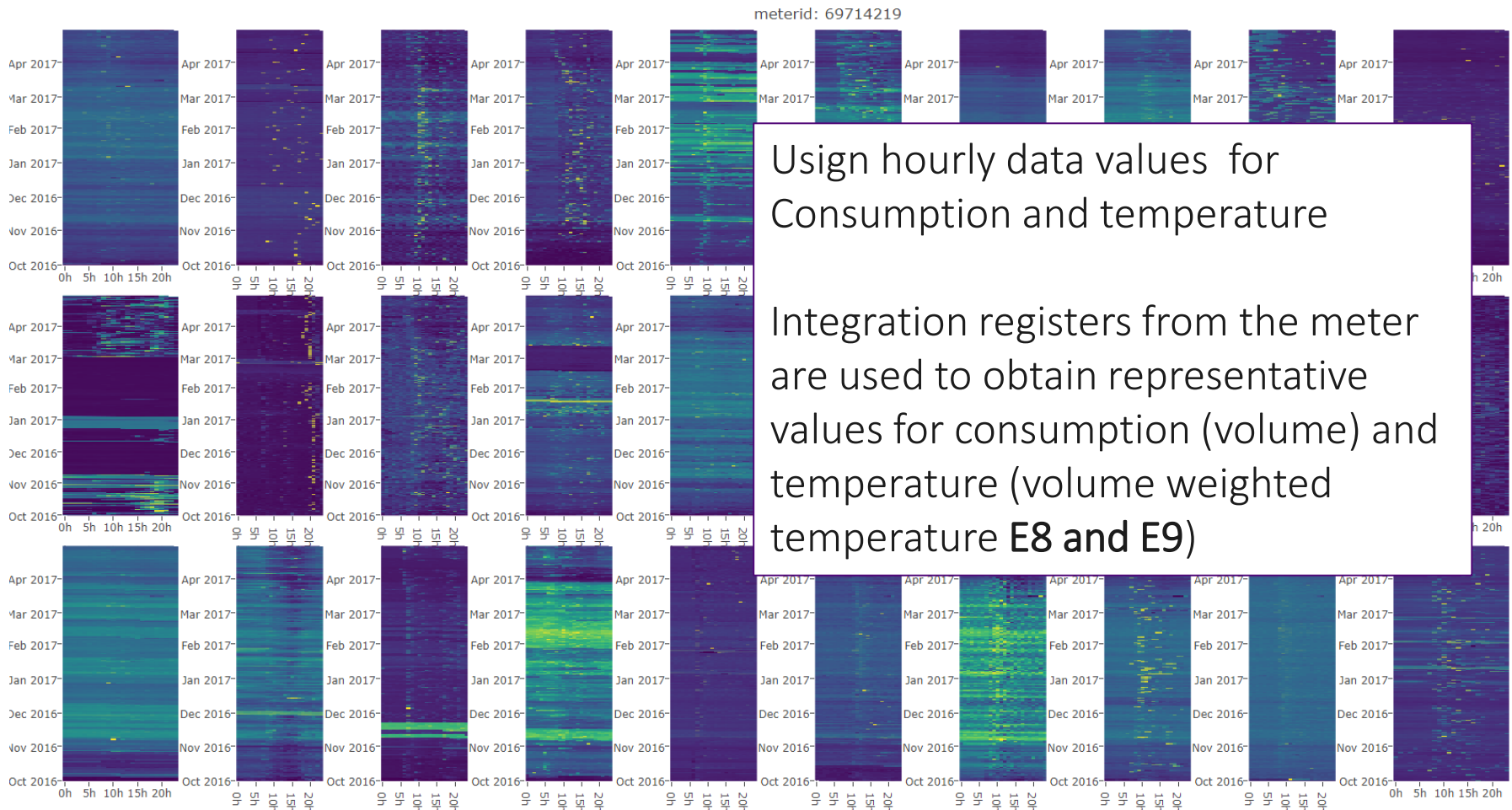


Heat consumption, pattern recognition, signal correlation (machine learning)

- **Heating system:**
 - Floor heating, radiator, both?
 - Control: Time operation, thermostat-controlled? Night setback? Well-regulated?
- **Hot water:**
 - Heat exchanger? Hot-water tank? Both? Well-regulated?



Heat load pattern / fingerprint for 30 customers during 1 heating season



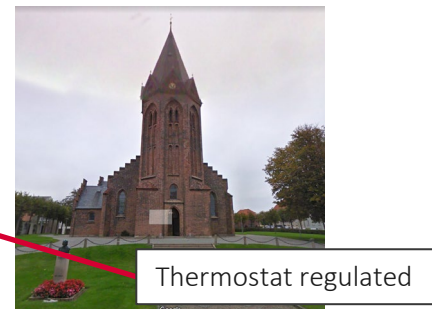
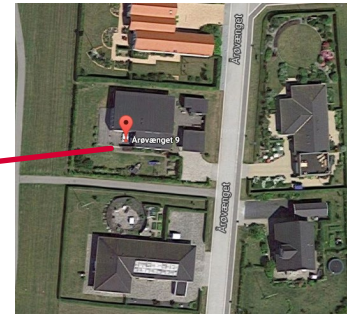
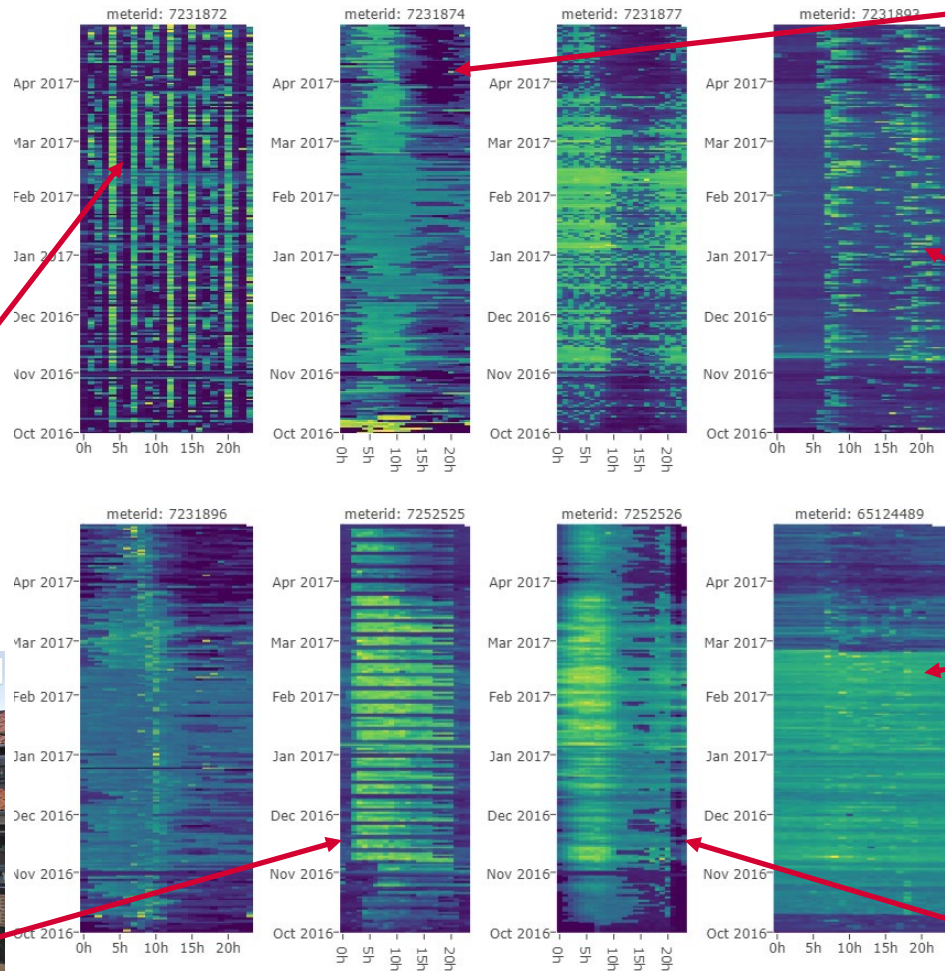
Customer classification

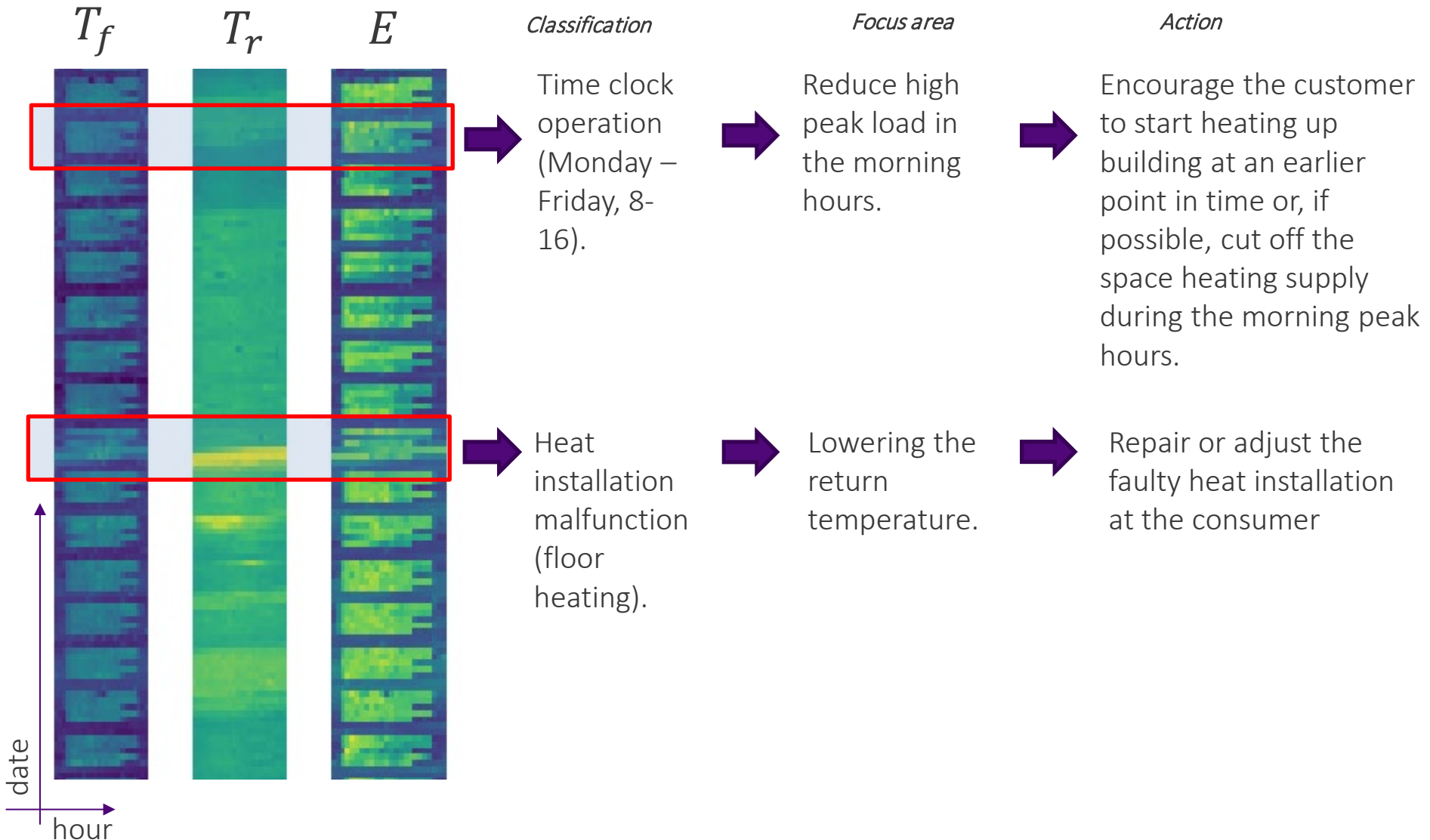
kamstrup

Understand the customers and their energy demand

Different types of customers and consumption is showed

Train a classifier for use in production planning, distribution network optimization etc.

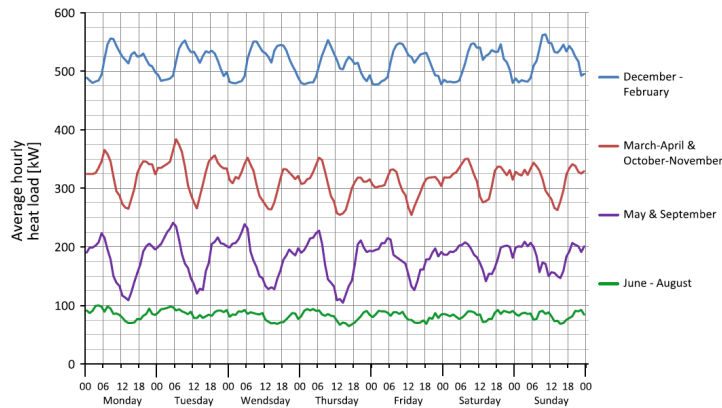




Heat load pattern recognition

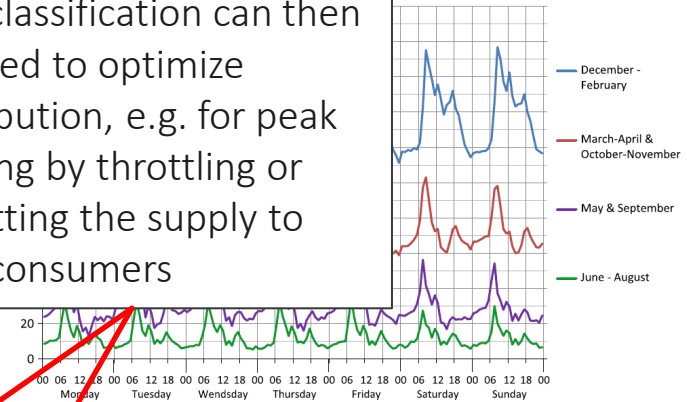
Continuous operation control

H. Gadd, S. Werner / *Applied Energy* 108 (2013) 176–183

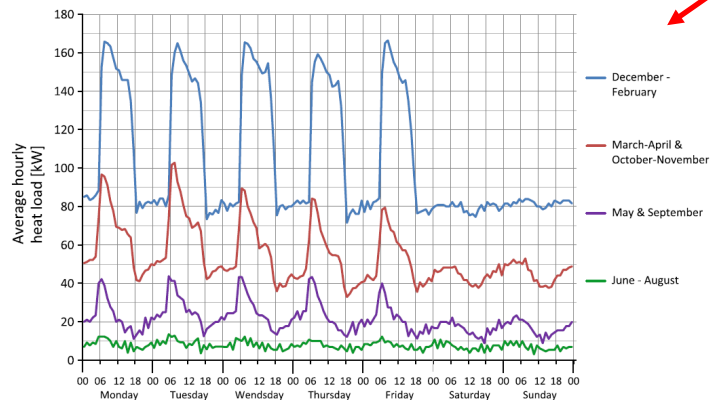


Automatically detect certain characteristics e.g. consumers with TCO

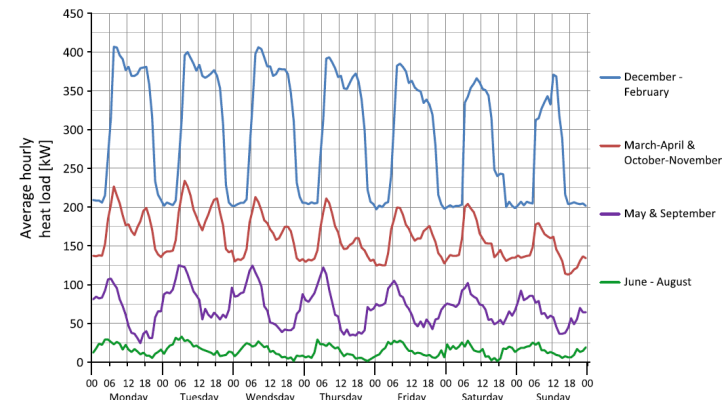
This classification can then be used to optimize distribution, e.g. for peak shaving by throttling or offsetting the supply to TCO consumers



Time clock operation, 5 days



Time clock operation, 7 days



Labelling of dataset (1 year, hourly data from ~2700 costumers), time clock operation, 5 days

DH label

With shorter load periods Wednesday and Friday

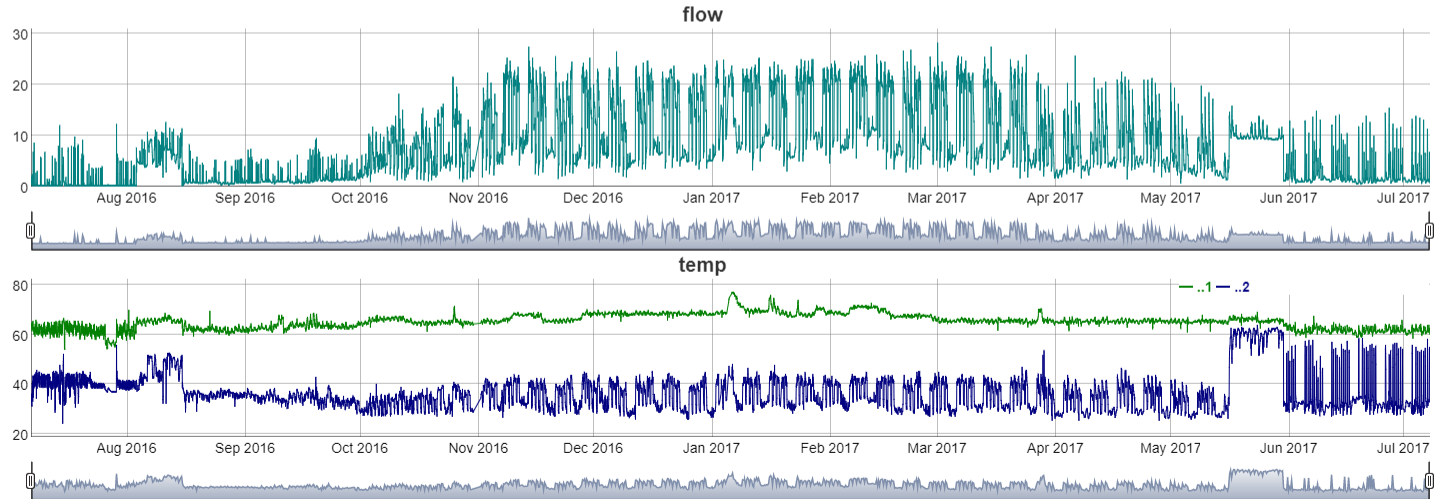
Meter ID (2662)

7252525

Previous Next

Labels:

- ☐ COC
- ☐ NSC
- ☐ TCO
- ☒ 5
- ☐ 7



Labelling of dataset (1 year, hourly data from ~2700 costumers), time clock operation, 5 days

kamstrup

DH label

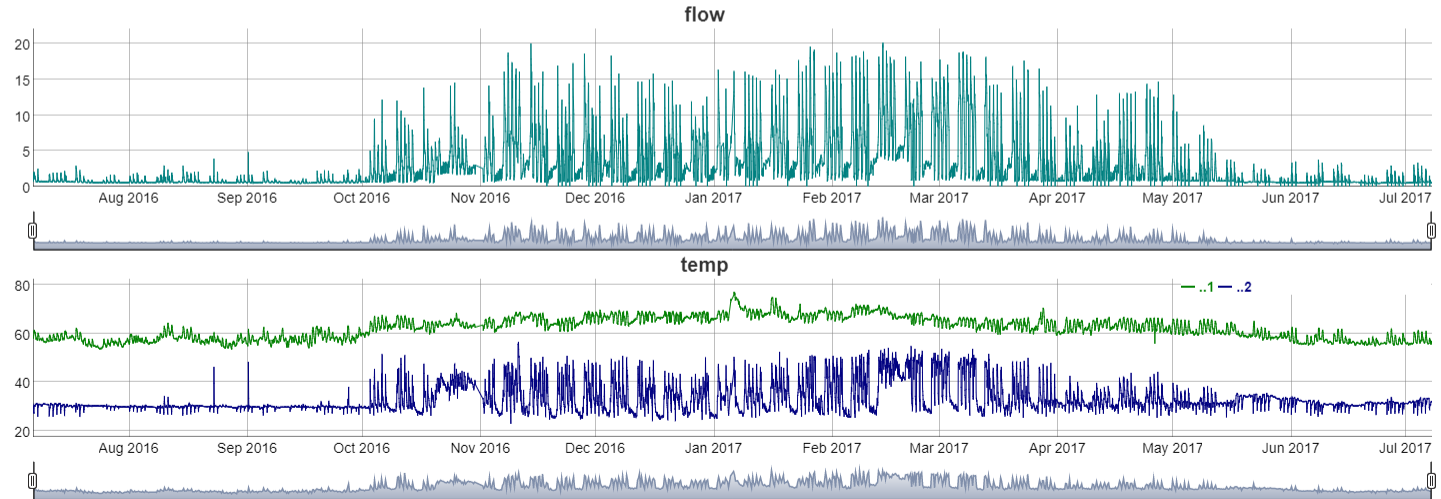
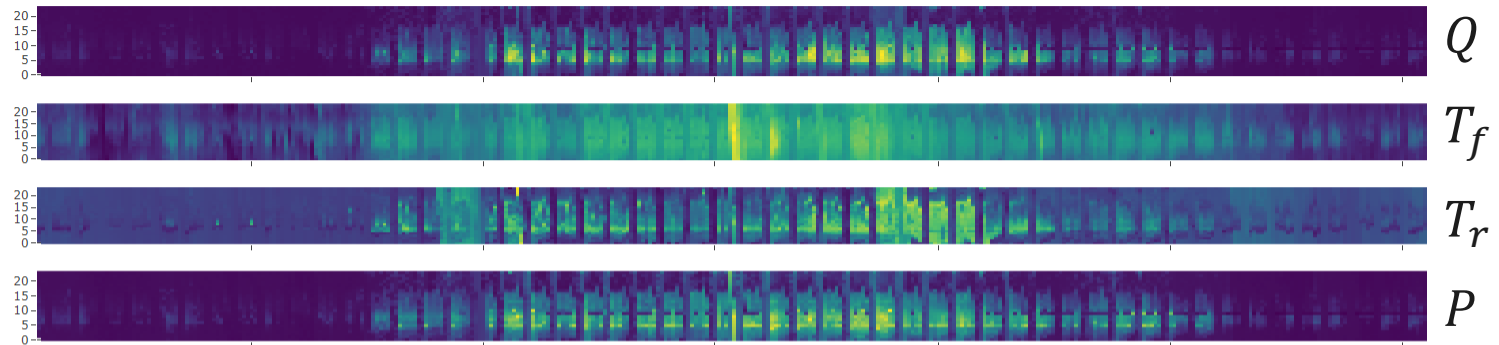
With high morning peak load

Meter ID (2662)
65124555

Previous Next

Labels:

- ☐ COC
- ☐ NSC
- ☐ TCO
- ☒ 5
- ☐ 7



Labelling of dataset (1 year, hourly data from ~2700 costumers), time clock operation, 5 days

kamstrup

DH label

With midday peak load

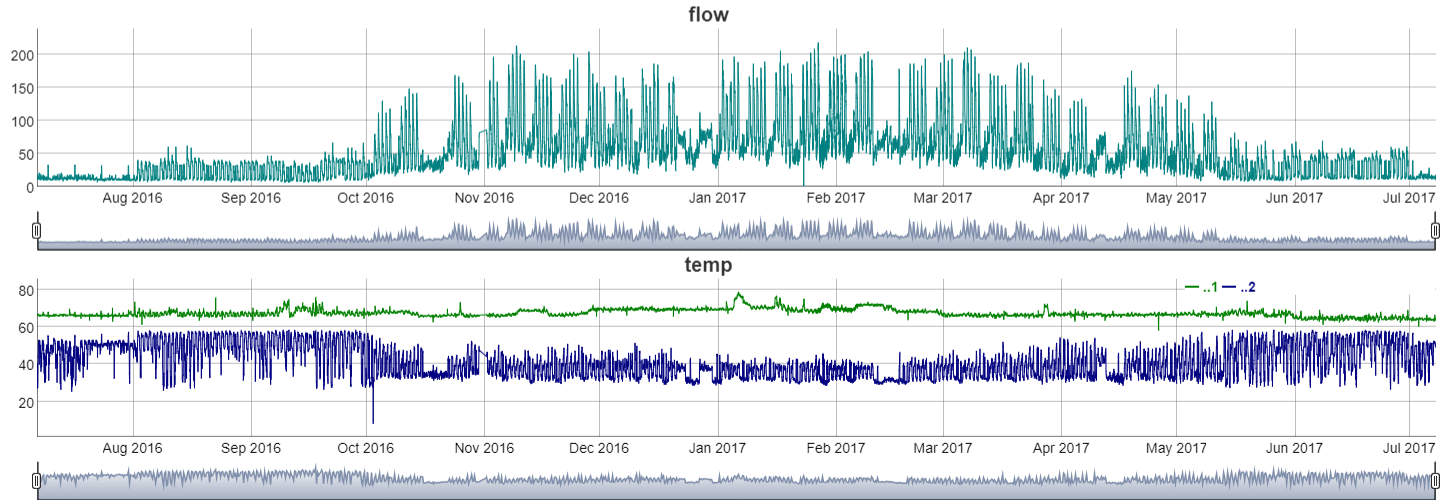
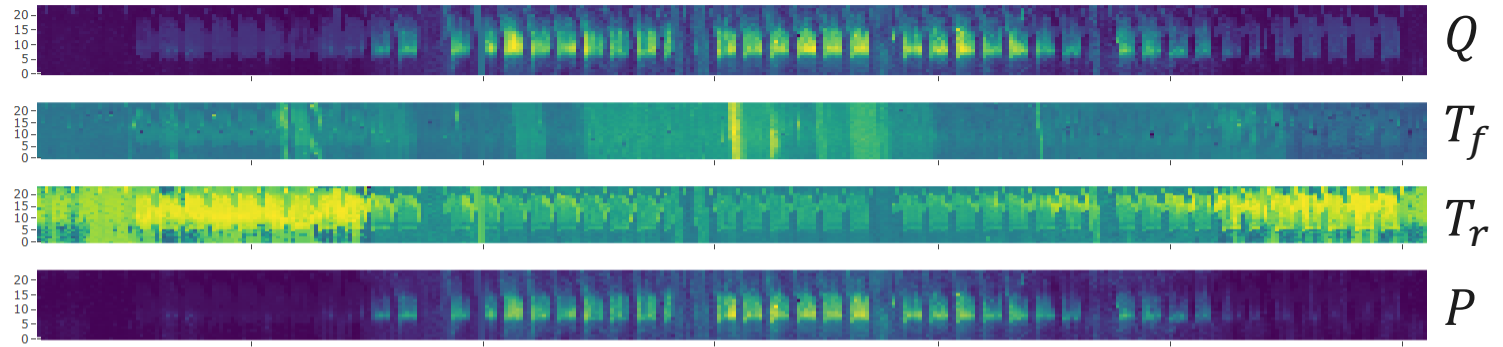
Meter ID (2662)

7248152

Previous Next

Labels:

- ☐ COC
- ☐ NSC
- ☐ TCO
- ☒ 5
- ☐ 7



Labelling of dataset (1 year, hourly data from ~2700 costumers), time clock operation, 5 days

kamstrup

DH label

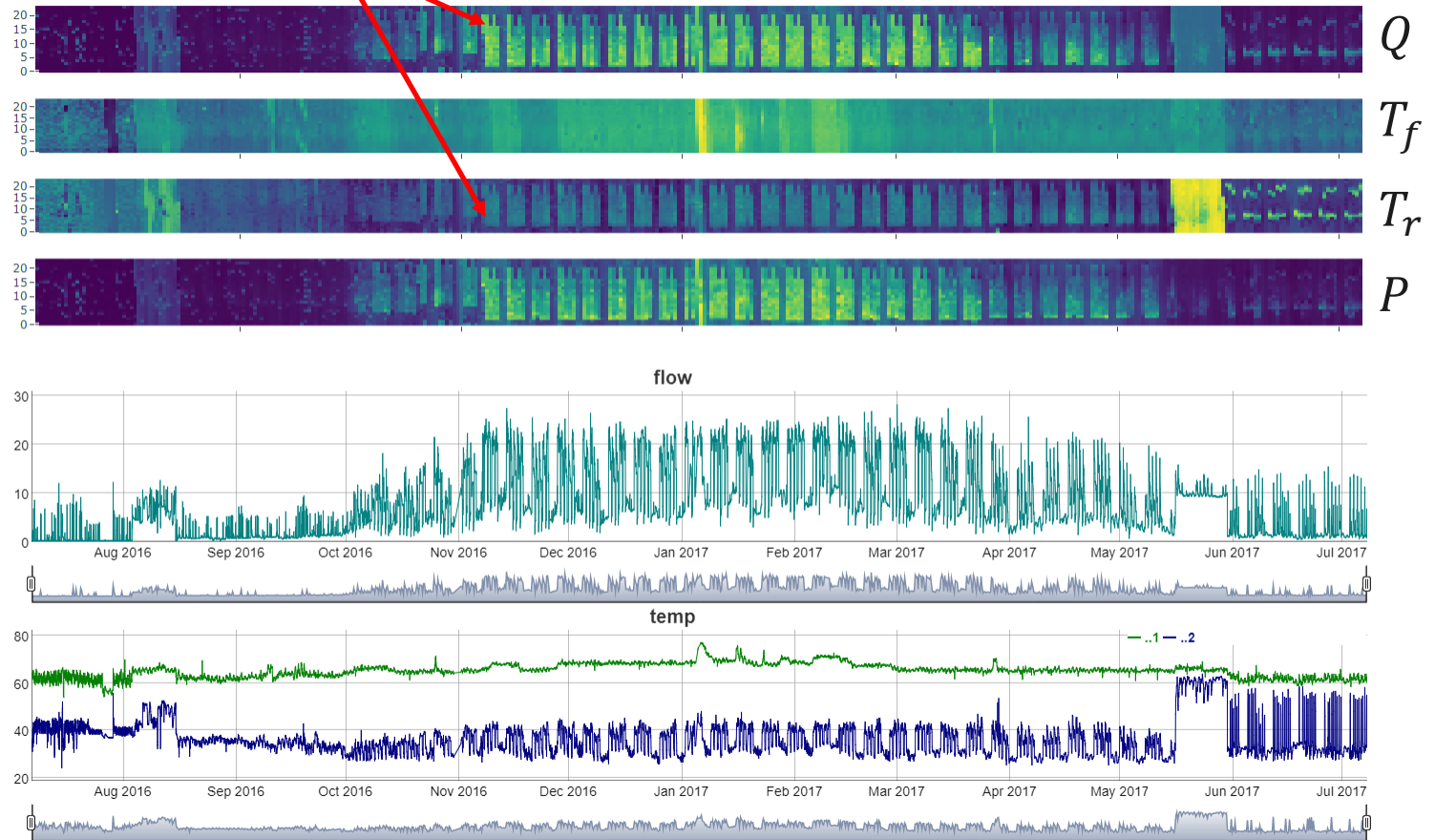
Positive correlation between heta load and return temperature

Meter ID (2662)
7252525

Previous Next

Labels:

- ☐ COC
- ☐ NSC
- ☐ TCO
- ☒ 5
- ☐ 7



Labelling of dataset (1 year, hourly data from ~2700 costumers), time clock operation, some days (warehouse)

kamstrup

~/HOME/MOK/shiny_apps/dhtrain - Shiny
http://127.0.0.1:4444 Open in Browser

— □ ×
Publish

DH train

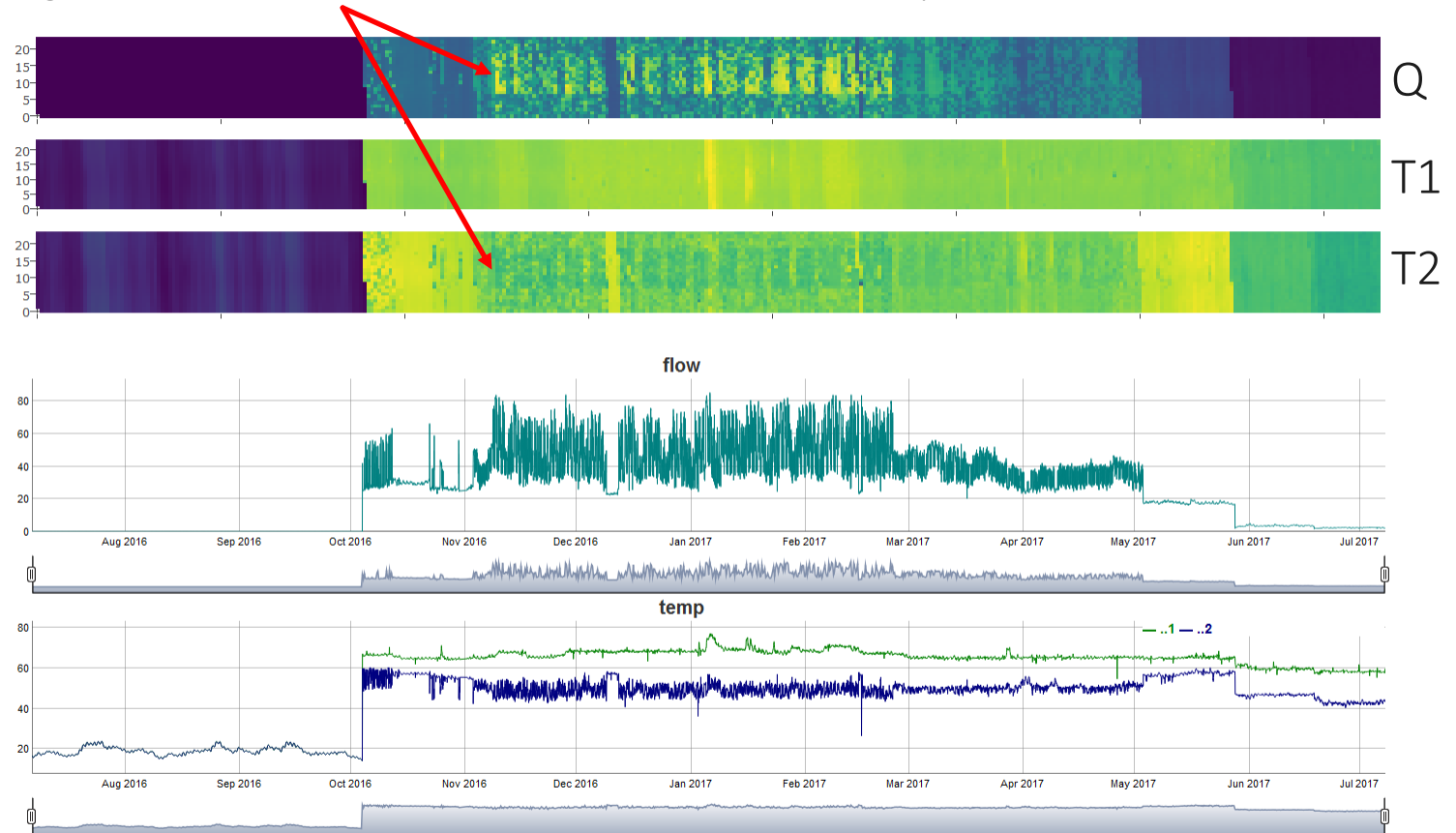
Meter ID (2662)

7248150

Previous

next

Negative correlation between heta load and return temperature



Labelling of dataset (1 year, hourly data from ~2700 costumers), continuous operation control

DH label

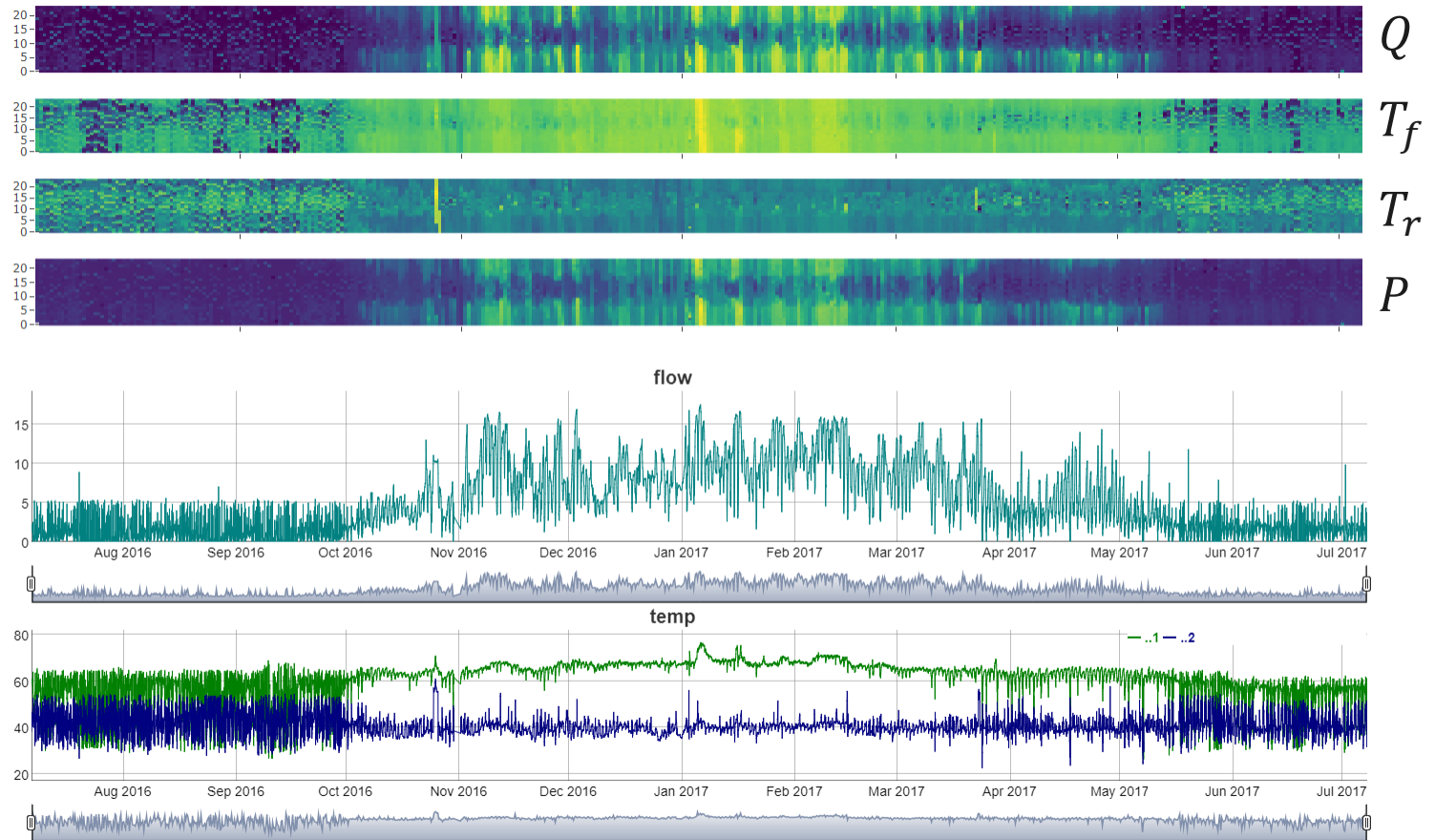
Highest load during morning (and at night)

Meter ID (2662)
69007804

Previous Next

Labels:

- ☒ COC
- ☐ NSC
- ☐ TCO
- ☐ 5
- ☐ 7

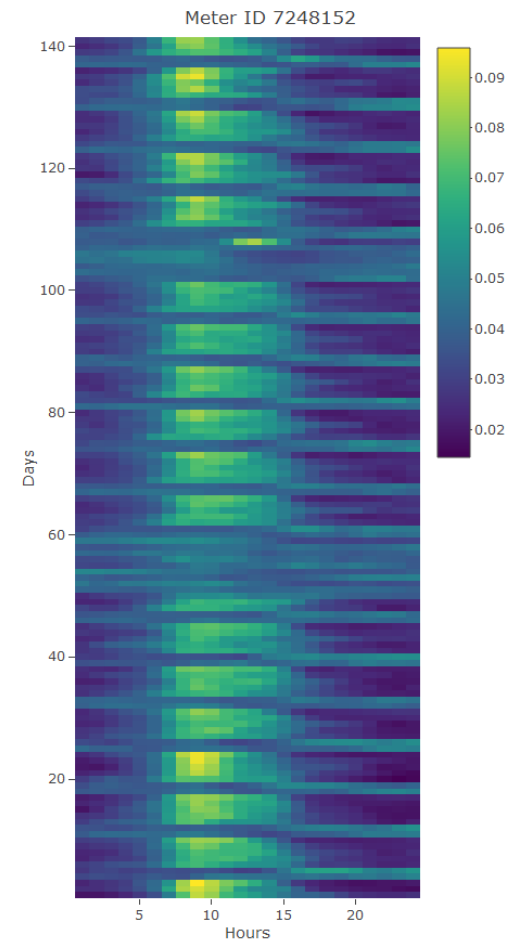


Time clock operation is classified with quite high accuracy

It is then possible to automatically identify all customers with time clock operation

Classifier automatically detects which days are showing time clock operation, e.g output for the meter shown here:

Day	TCO?
Monday	1
Tuesday	1
Wednesday	1
Thursday	1
Friday	1
Saturday	0
Sunday	0



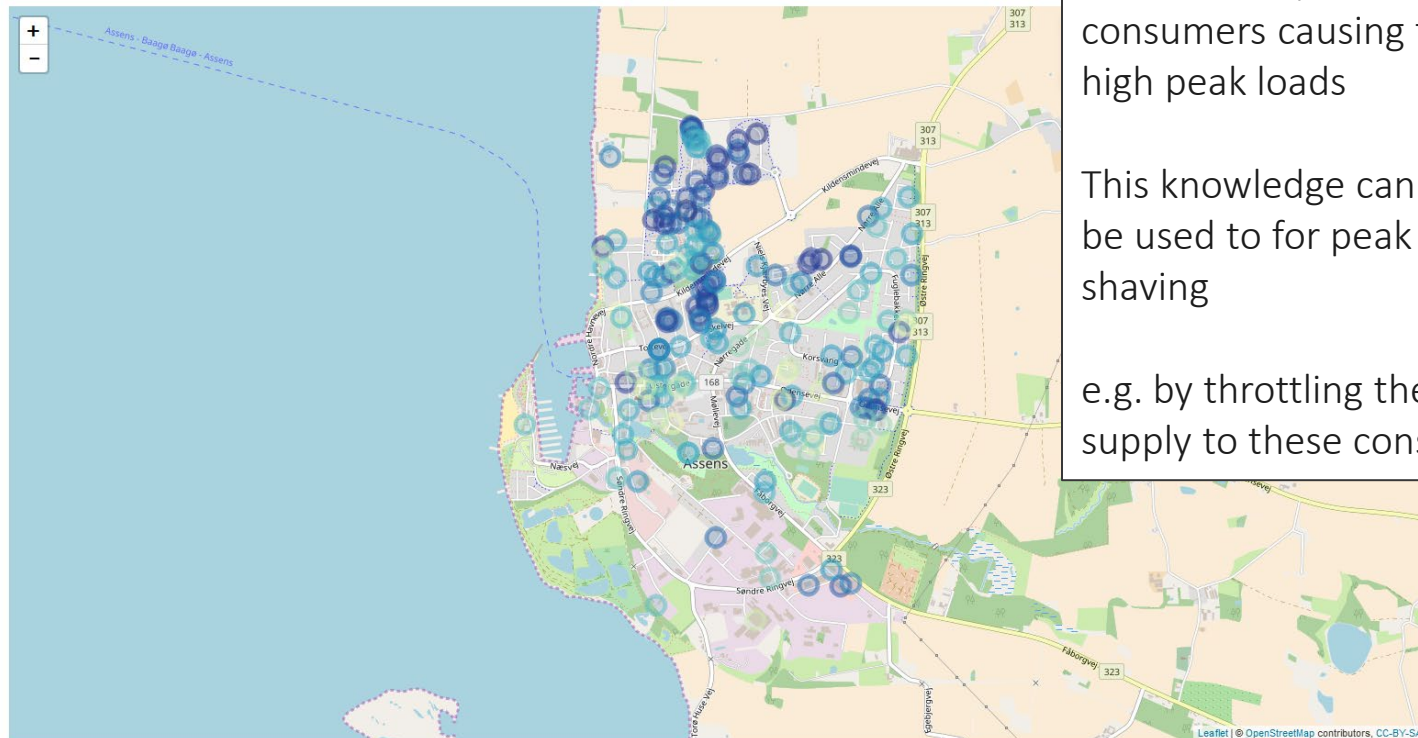
Filtering the TCO consumers on a map



Can e.g. be used to automatically detect the consumers causing the high peak loads

This knowledge can then be used to for peak shaving

e.g. by throttling the supply to these consumers



entries	
n	%
83	7.4
67	7.0
20	14.5
12	4.1
2	40.0
2	7.7
2	66.7
2	50.0
1	25.0
1	20.0

24 entries
1 2 3 Next

TCOC_days	
	%
0	1.91082802547771
4	
5	82.7586206896552
6	100
7	64.7058823529412

Showing 1 to 5 of 5 entries

Conclusion

