Knowledge-based Short-Term Load Forecasting for Maritime Container Terminals

Evaluation of two approaches based on operation plans

Energie
Energy

12.06.2017 @International Data Science Conference

Norman Ihle
R&D-Division Energy
OFFIS – Institute for Information Technology

2 Outline

- Introduction & Motivation
- Available Data & Data Preprocessing
- CBR-based Short-Term Load Forecasting
- Forecasting based on Artificial Neural Networks
- Evaluation
- Summary & Outlook

3 Introduction & Motivation

Smart Grid & Demand Side Integration

- Smart Grid
  - Rising level of renewable energy sources
  - Energy feed-in in distribution grids
  - Volatile energy production

- One of the main challenges:
  - to keep production and consumption of electricity on the same level in all parts of the grid at all times

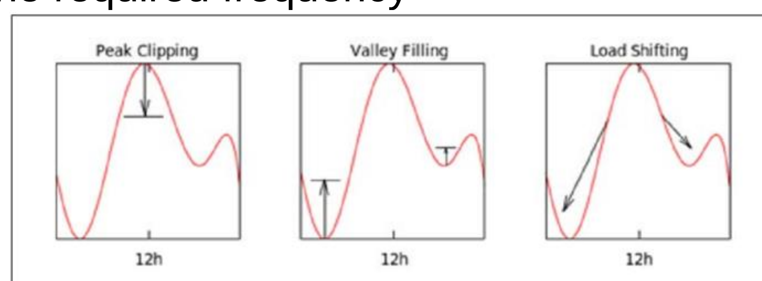
- One part of the solution:
  - **Integration of consumers** into the energy market
  - Demand Side Integration / Demand Response
    - Usage of flexible loads

4 Introduction & Motivation

Demand Side Integration for industrial sites

- Demand Side Integration

- Possible use-cases for industrial sites:
  - Optimize the energy procurement using variable pricing
    - Prices will be low if there is a large supply of renewable energy available
    - Energy suppliers might offer Real-Time Pricing or Time-Of-Use Tariffs or industrial sites can procure the energy themselves at the electricity exchange
  - Offering Balancing or Control Energy
    - Supporting the grid operator keeping the required frequency
  - Load shifting to avoid grid fees



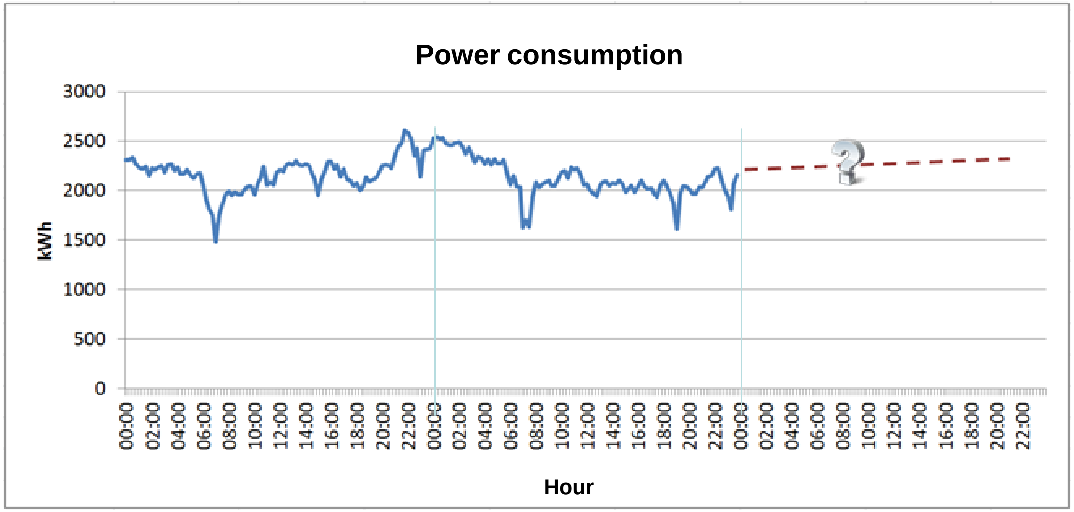
Source: M. Sonnenschein, B. Rapp, and J. Bremer, "Demand Side Management und Demand Response", in Handbuch Energiemanagement, vol. 3, Heidelberg, 2010.

- **All use-cases need accurate forecasts!**

5 Introduction & Motivation
Short-Term Load Forecasting

- A daily load curve consists of 96 values
  - each representing the power consumption of 15 minutes
- Short-Term Load Forecasting is a frequently discussed topic in scientific literature
  - Most methods are designed to forecast whole grids or parts of grids with a high number of consumers
  - None of these methods are discussed in the reference to a container terminal

- Established methods are:
  - Equivalent day approach
  - Time series models
  - Artificial neural networks
  - Simulation

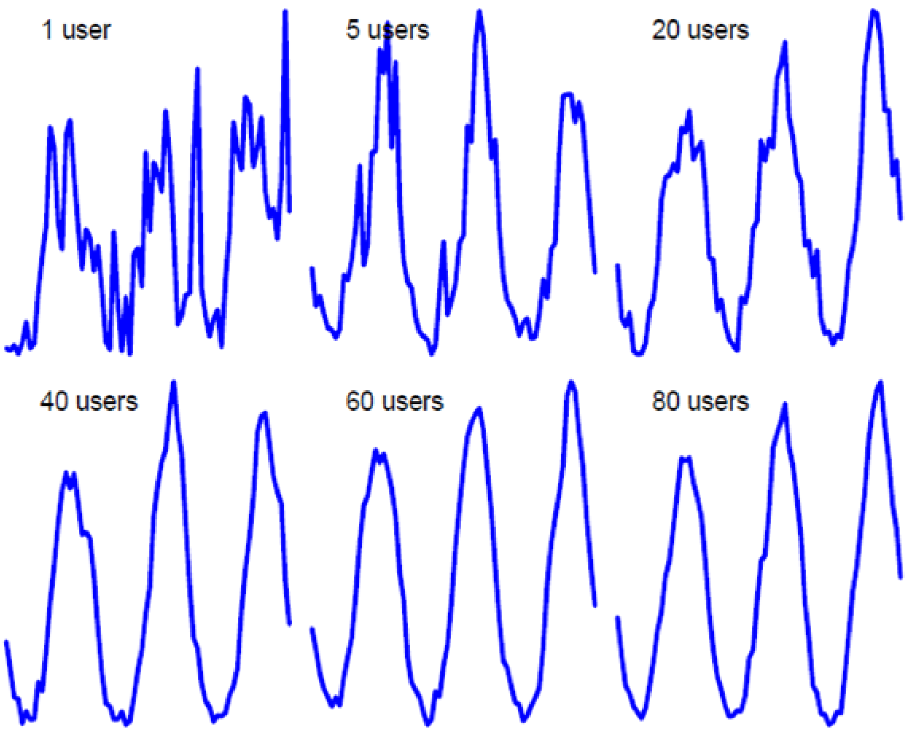


6 Introduction & Motivation

Aggregation levels

- The consumption pattern of a single customer generally has little structure to be exploited.
- Aggregating more and more customers „smoothens“ the signal so it can be more predictable

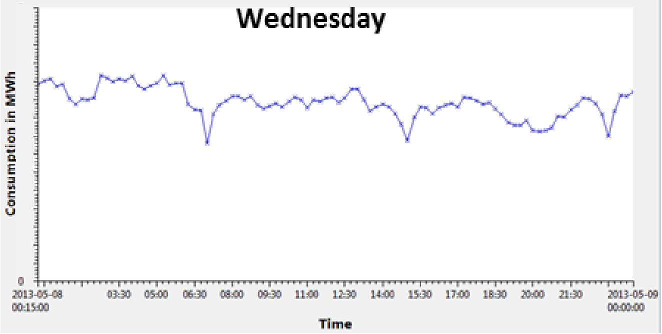
- Can we make use of additional data to improve the forecasting process?

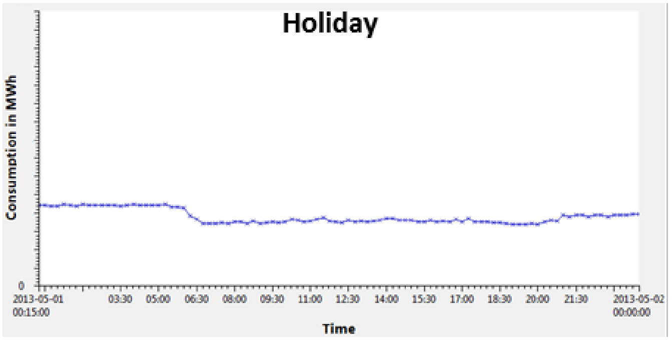


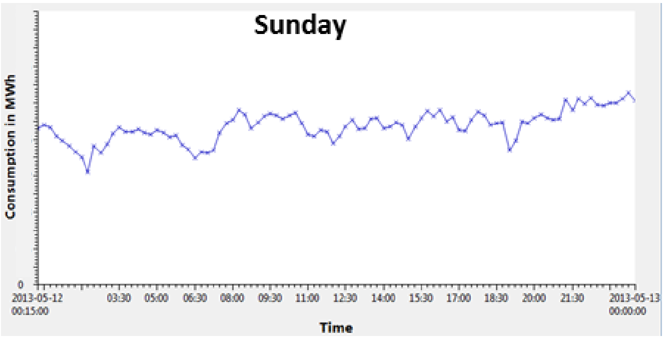
Source: Sevlina, R., Rajagopal, R. (2014): Short Term Electricity Load Forecasting on Varying Levels of Aggregation

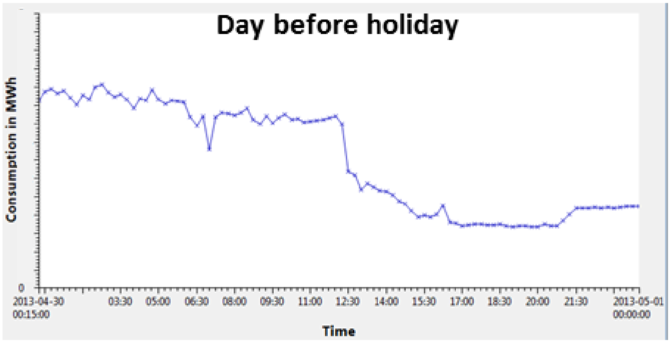
7 Introduction & Motivation
Load curves of a container terminal

- Example of different load curves of a container terminal





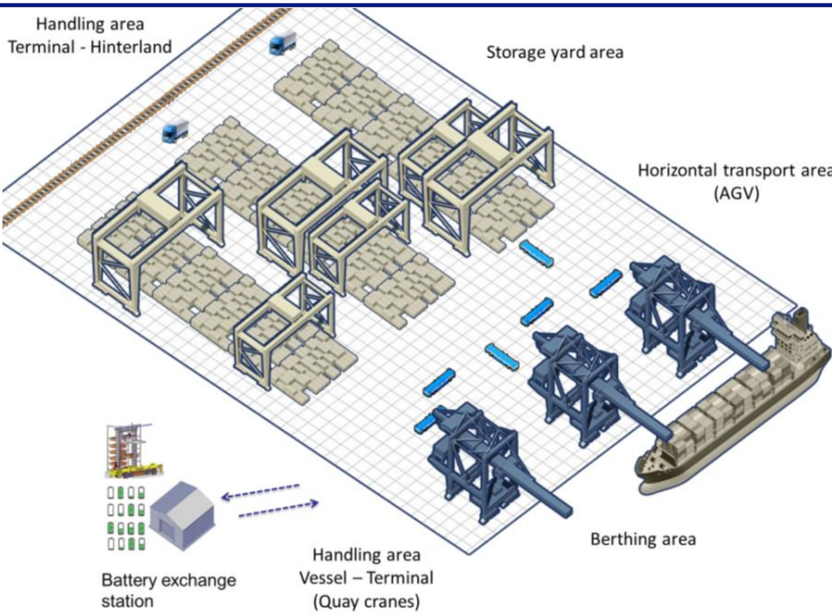


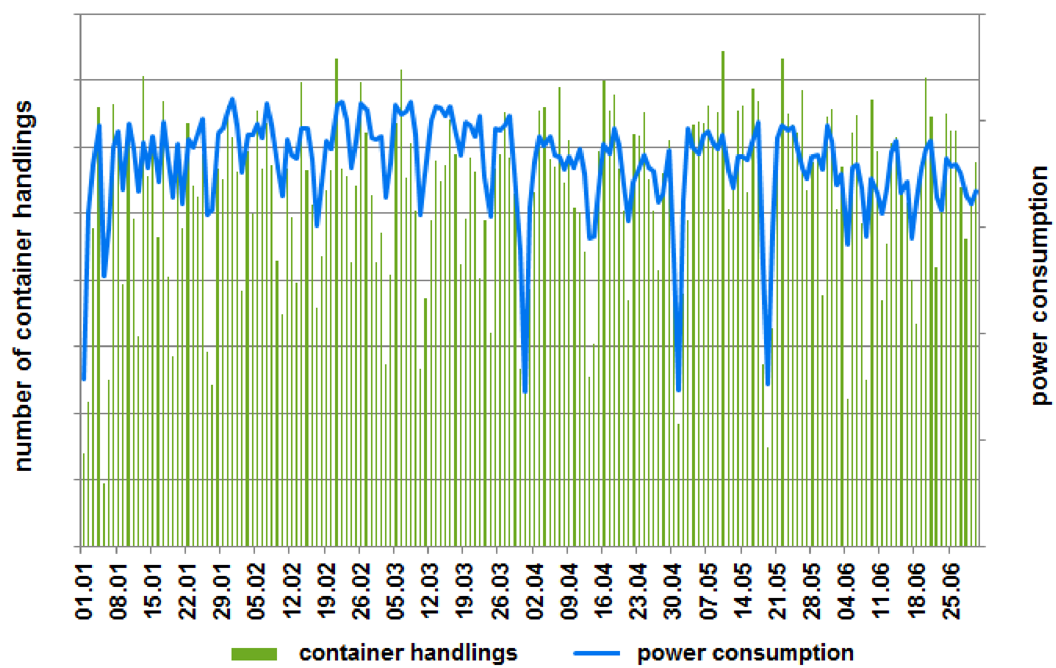


- Data is available for the years 2010 to 2013

8 Introduction & Motivation

Power consumption at a maritime container terminal





- A container terminal has a very high portion of electric handling equipment
  - Quay cranes, yard cranes, reefer storage area, lighting,…
  - The electricity consumption relates to the container handling numbers
- But: container handling numbers show only very little to none regularity

9 Available Data & Data Preprocessing

Can we make use of additional data to improve the forecasting process?

10 Available Data & Data Preprocessing

The sailing list

- The container terminal plans its operations based on the sailing list
- The sailing list contains information about planned ship arrivals and departures

JSNR	Ship Name	Ship Type	Expected Arrival	Expected Departure	Loading	Unloading
308505	AKACIA	Feeder	04.09.2013 15:45	05.09.2013 04:00	373	244
306757	OOCL KAOHSIUNG	ATX	05.09.2013 00:10	05.09.2013 15:35	1399	16
308442	A LA MARINE	Feeder	05.09.2013 07:00	06.09.2013 06:00	534	556
308538	OOCL RAUMA	Feeder	05.09.2013 18:05	06.09.2013 15:30	501	451
308579	KAHN DBR	Kahn	05.09.2013 07:15	05.09.2013 10:50	53	21
308632	KAHN LAUK	Kahn	05.09.2013 15:55	05.09.2013 16:30	0	5
306926	EMOTION	Feeder	06.09.2013 07:45	07.09.2013 00:55	458	333
307896	APL VANDA	LOOP_7	06.09.2013 17:55	09.09.2013 14:00	3024	3630
308543	LEONIE P	Feeder	06.09.2013 15:45	06.09.2013 20:15	22	73

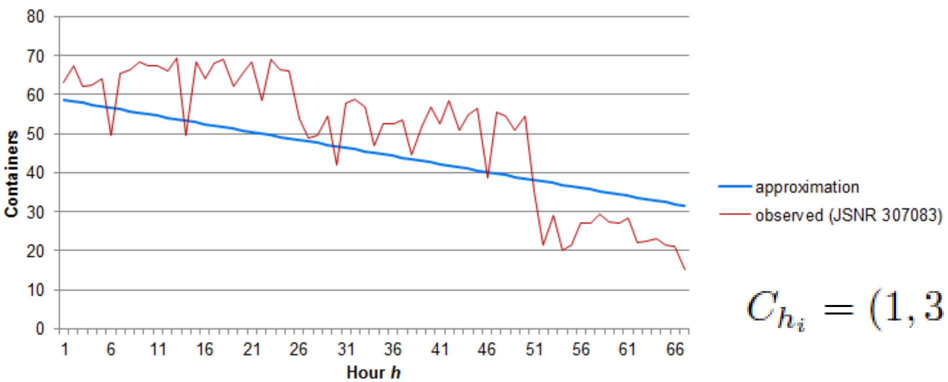
11 Available Data & Data Preprocessing

Concept – Case Base

- Build up of daily views on the sailing list by splitting up single berthing information

307896	APL VANDA	LOOP_7	06.09.2013 17:55	09.09.2013 14:00	3024	3630

307896a	APL VANDA	LOOP_7	06.09.2013 17:55	06.09.2013 23:59	346	415
307896b	APL VANDA	LOOP_7	07.09.2013 00:00	07.09.2013 23:59	1236	1484
307896c	APL VANDA	LOOP_7	08.09.2013 00:00	08.09.2013 23:59	1000	1200
307896d	APL VANDA	LOOP_7	09.09.2013 00:00	09.09.2013 14:00	442	531



$$C_{h_i} = (1,3 - \frac{0,6}{h_n - 1} * (h_i - 1)) * \frac{C_{total}}{h_n} \ with\ h_i = 1, ..., h_n$$

12 Available Data & Data Preprocessing

Additional available data

- Operational data
  - Well-defined sailing list entries with berthing information and container numbers for each vessel

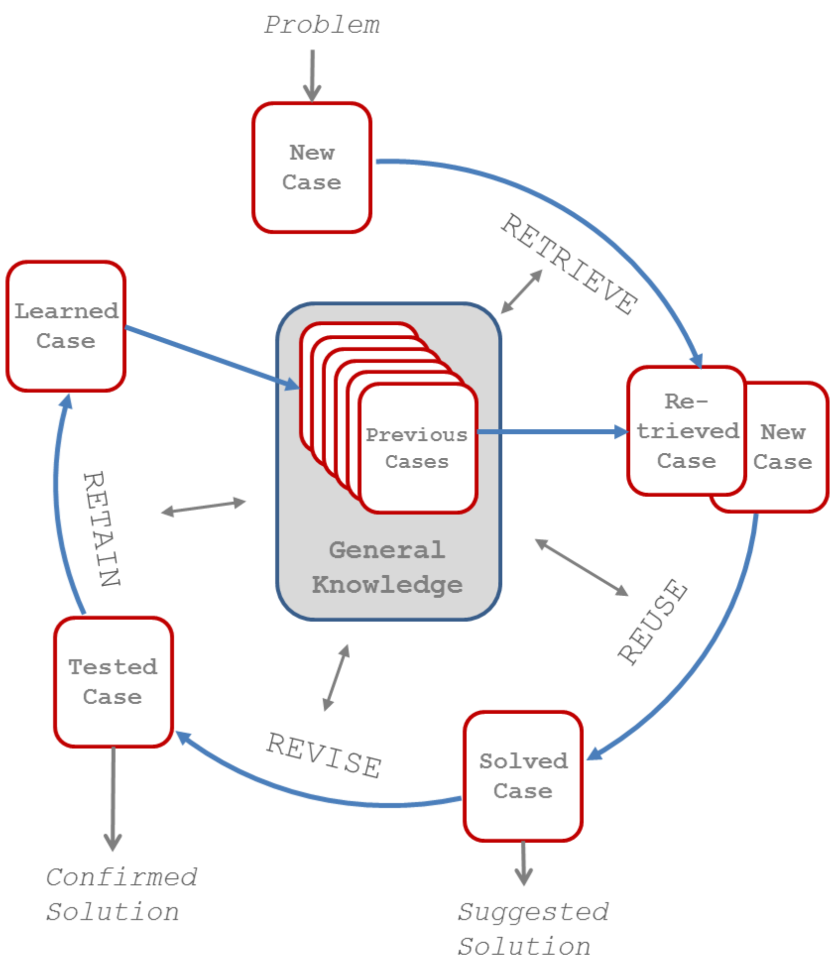
- Further (daily) data
  - Derived attributes
    - Number of overall ship arrivals and departures of one day
    - Overall number of container handled on one day

  - Weather data
    - Temperature
    - Wind Speed

  - Information according to the calendar
    - Weekday
    - Holiday

13 Case-Based Reasoning

14 CBR-based Short-Term Load Forecasting
Idea



(following Aamodt and Plaza, 1994)

Definition Case:

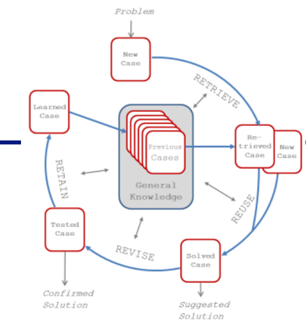
$$F_{case} = (probl, sol)$$

$$F_{query} = (probl)$$

- Basic idea:
  - The load curve of a similar day in the past is the foundation for a new forecast
  - The similarity is determined using data of the sailing list (= operation plan)
  - The load curve can be adapted using differences in ship arrival data
  - Additional common knowledge can be used for adaptation

15 CBR-based Short-Term Load Forecasting

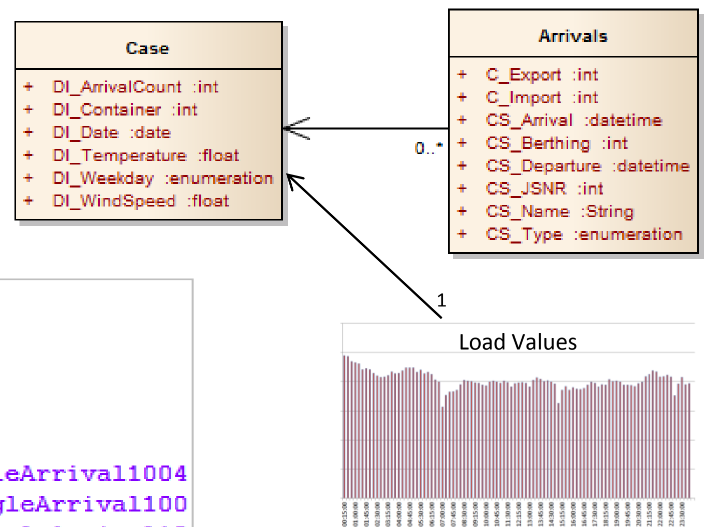
Concept: Cases and Case Base



- The available information is used as case
  - Description part of the case

```xml
<instance id="SingleArrival10041" >
  <att name="Unload" value="485" />
  <att name="BerthingTime" value="994" />
  <att name="ShipType" value="PAX" />
  <att name="ShipName" value="KIEL EXPRESS" />
  <att name="Load" value="838" />
  <att name="JSNR" value="109731" />
  <att name="Departure" value="05.11.11 23:59" />
  <att name="ArrivalTime" value="05.11.11 07:25" />
</instance>
```

```xml
<instance id="DailyArrivals20111105" >
  <att name="Weekday" value="Saturday" />
  <att name="Temperature" value="8.5" />
  <att name="NumberOfArrivals" value="15" />
  <att name="Date" value="05.11.11" />
  <att name="Arrivals" value=
  "SingleArrival10038;SingleArrival10039;SingleArrival10041;SingleArrival1004
  0;SingleArrival10043;SingleArrival10042;SingleArrival10046;SingleArrival100
  47;SingleArrival10044;SingleArrival10045;SingleArrival10048;SingleArrival10
  049;SingleArrival10050;SingleArrival10052;SingleArrival10051;" />
  <att name="NumberOfContainerHandles" value="4626" />
  <att name="WindSpeed" value="3.7" />
</instance>
```



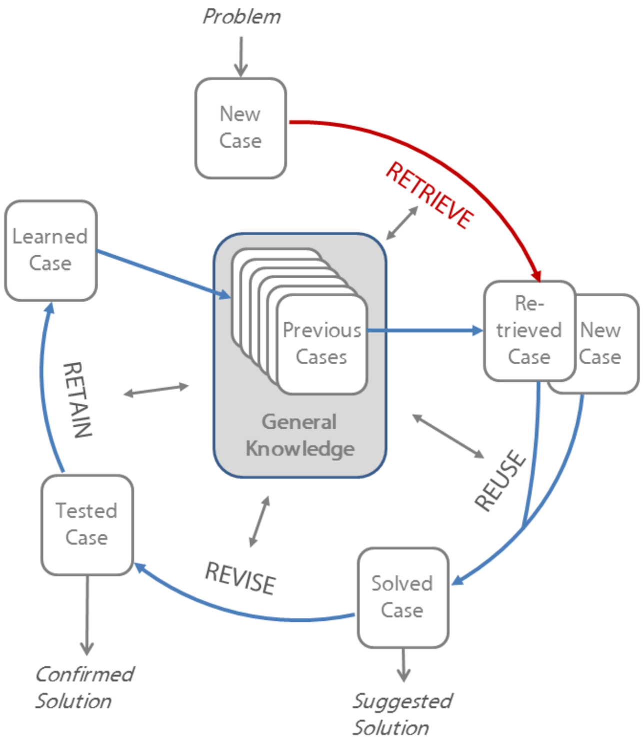
- Each case points to the according metered load curve for the day
  - Solution part of the case

16 CBR-based Short-Term Load Forecasting

Concept – Similarity Measures

“The purpose of this similarity assessment is to approximate the utility of a given solution for a new problem”

(Bergmann, 2003)



Local similarity (using distance):

$$sim_{a_i}(q_{a_i}, c_{a_i}) = \frac{d(q_{a_i}, c_{a_i})}{1 + d(q_{a_i}, c_{a_i})}$$

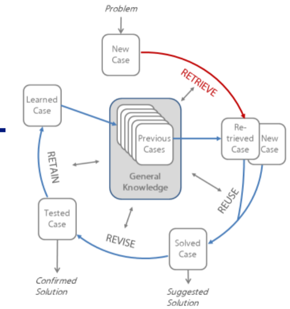
Global similarity:

$$sim(q,c) = \frac{\sum_{i=1}^{n} w_i * sim_{a_i}(q_{a_i}, c_{a_i})}{\sum_{i=1}^{n} w_i}$$

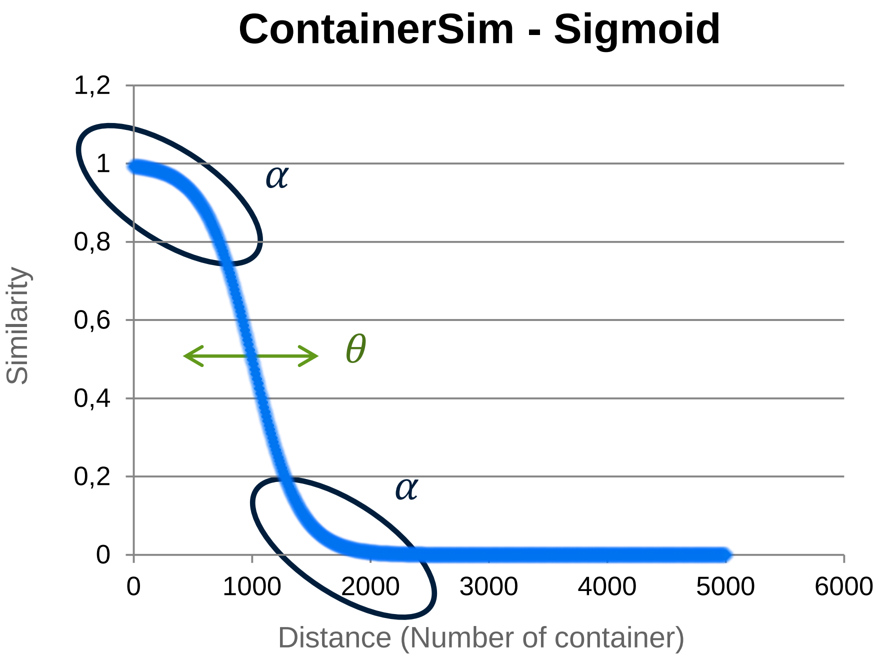
q = query; c = case; a_i = attribute i; w_i = weight of attribute i

17 CBR-based Short-Term Load Forecasting

Concept – Similarity Measures



- Choosing similarity measures for each attribute
  - Using similarity measures according to Bergmann (Bergmann 2002)
  - Example below: Similarity measure for container handling number



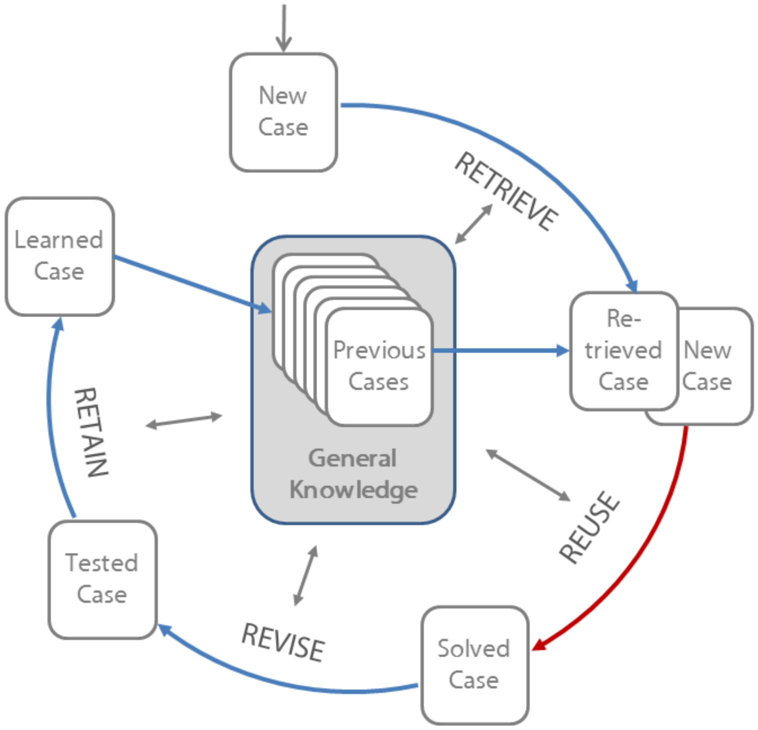
$$f(d) = \frac{1}{e^{\frac{d-\theta}{\alpha}}+1}$$

$$sim_{Container}(c,q) = \frac{1}{e^{\frac{d(c,q)-1000}{200}}+1}$$

q = query; c = case; d = distance

18 CBR-based Short-Term Load Forecasting

Concept - Adaptation

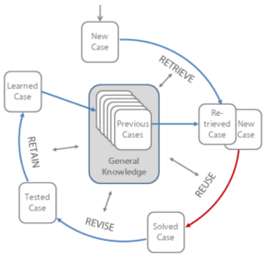


Adaptation describes the alignment of the found solution to the conditions of the query

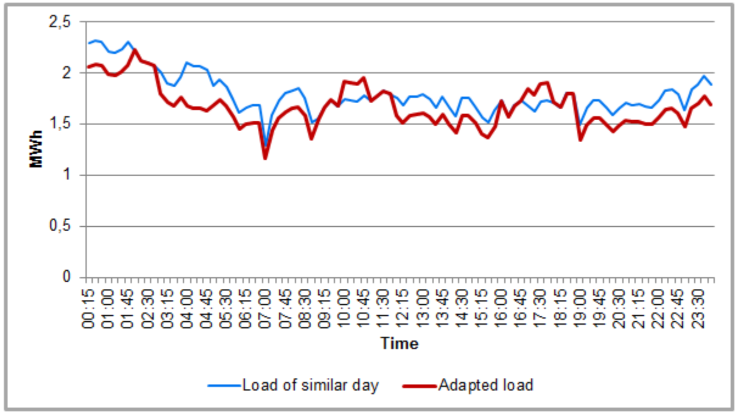
$$Adaption = ((case, Sol(case)), query) \rightarrow Sol(query)$$

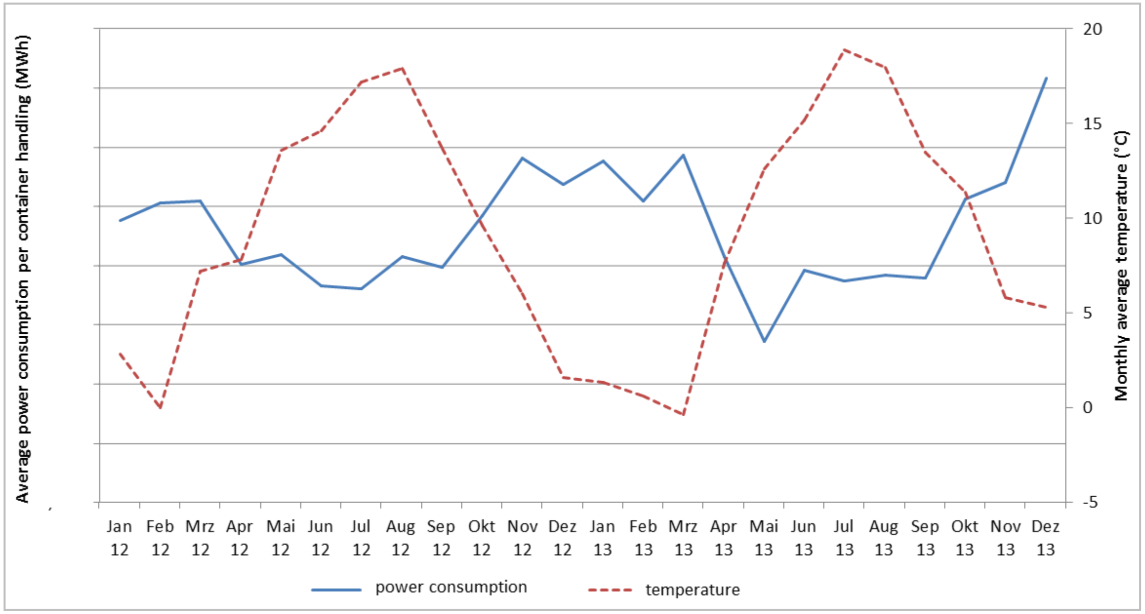
19 CBR-based Short-Term Load Forecasting

Adaptation based on general knowledge



- Contextual knowledge:
  - Difference in temperatures in winter and summer time have an impact on power consumption
  - Difference in lighting hours in summer and winter
  - Sunday truck driving prohibition

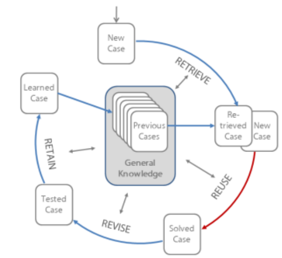




OFFIS Energy
INSTITUTE FOR INFORMATION TECHNOLOGY

20 CBR-based Short-Term Load Forecasting

Adaptation based on case knowledge



Case knowledge:
Difference in hourly container numbers

	Hour 0 00:00 – 01:00	Hour 1 01:00 – 02:00	Hour 2 02:00 – 03:00	Hour 3 03:00 – 04:00	...
query	207	216	295	188	...
case	304	287	304	312	...

hour i	d_i	d_{i-1}	d_{i-2}	d_{i+1}	*factor*
2	>200	>150	> 0	>150	0.4
2	>150	>100	> 0	>100	0.3
2	>100	> 50	> 0	> 50	0.2
2	> 50	> 30		> 30	0.1
2	< -50	< -30		< -30	-0.1
2	< -100	< -70	< 0	< -70	-0.2
2	< -150	< -120	< 0	< -120	-0.3
2	< -200	< -150	< 0	< -150	-0.4

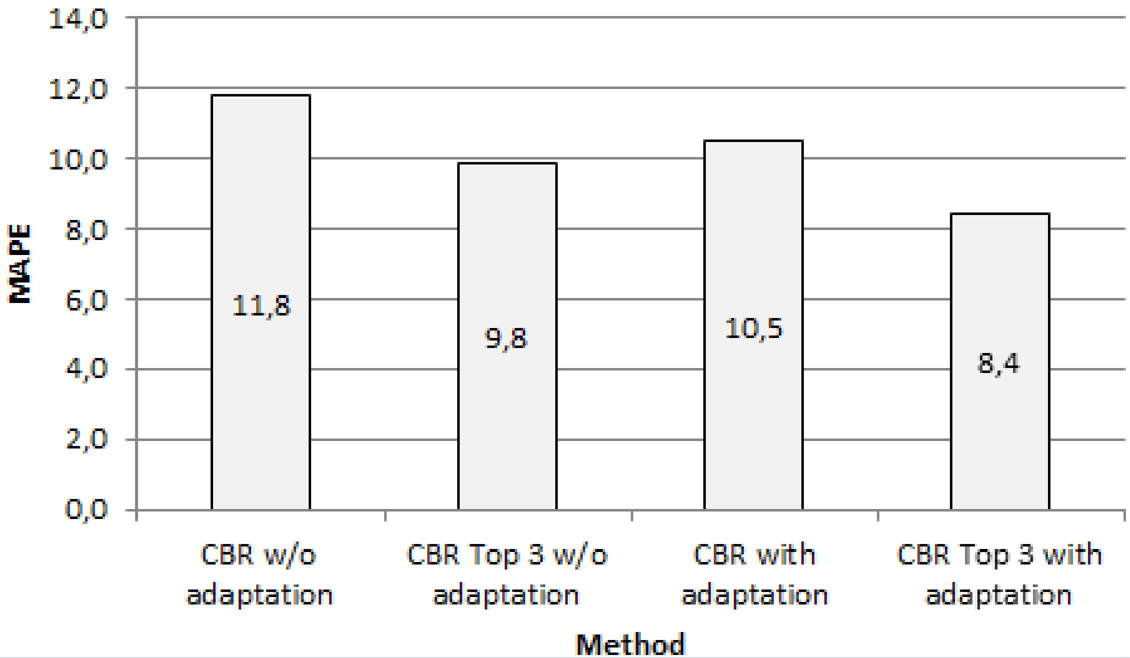
hour i	d_i	d_{i-1}	d_{i+1}	*factor*
8	>200	>150	>150	0.4
8	>150	>120	>120	0.3
8	>100	> 70	> 70	0.2
8	> 50	> 20	> 20	0.1
8	< -50	< -20	< -20	-0.1
8	< -100	< -70	< -70	-0.2
8	< -150	< -120	< -120	-0.3
8	< -200	< -150	< -150	-0.4

21 CBR-based Short-Term Load Forecasting

Evaluation - first results

- A first evaluation for one year showed that the CBR-forecasts underestimate the real consumption
  - Adding an additional factor for the yearly increase in electricity consumption
  - Using the mean of the 3 best CBR-results to smooth the load curve
  - Average results for one year:

$$MAPE = \frac{100}{N} \sum_{t=1}^{N} \frac{|F_t - A_t|}{A_t}$$

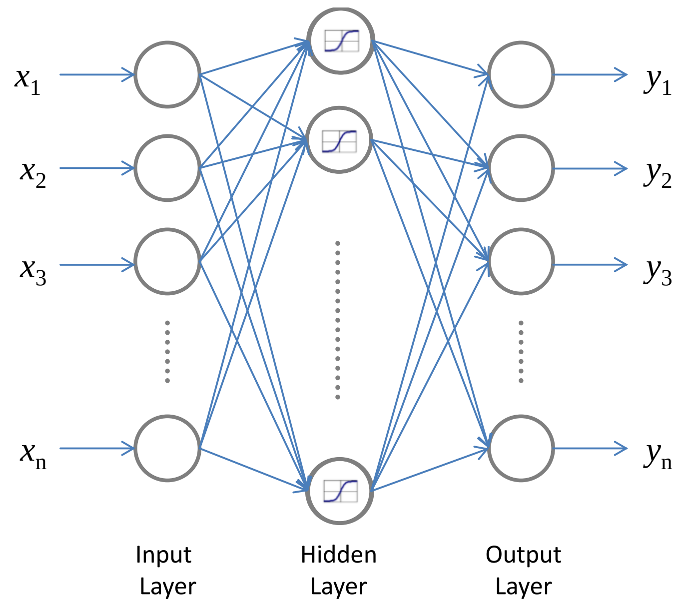


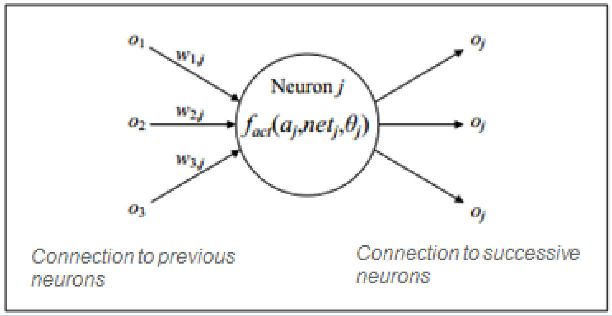
22 Artificial neural networks

23 Artificial neural networks

Multilayer perceptron

- Modeling the way the brain solves problems with large clusters of parallel working units (neurons)
- Challenges:
  - Selection of appropriate network structure, training methods and activation functions
- Idea: Training with operational data as input and load curve as output





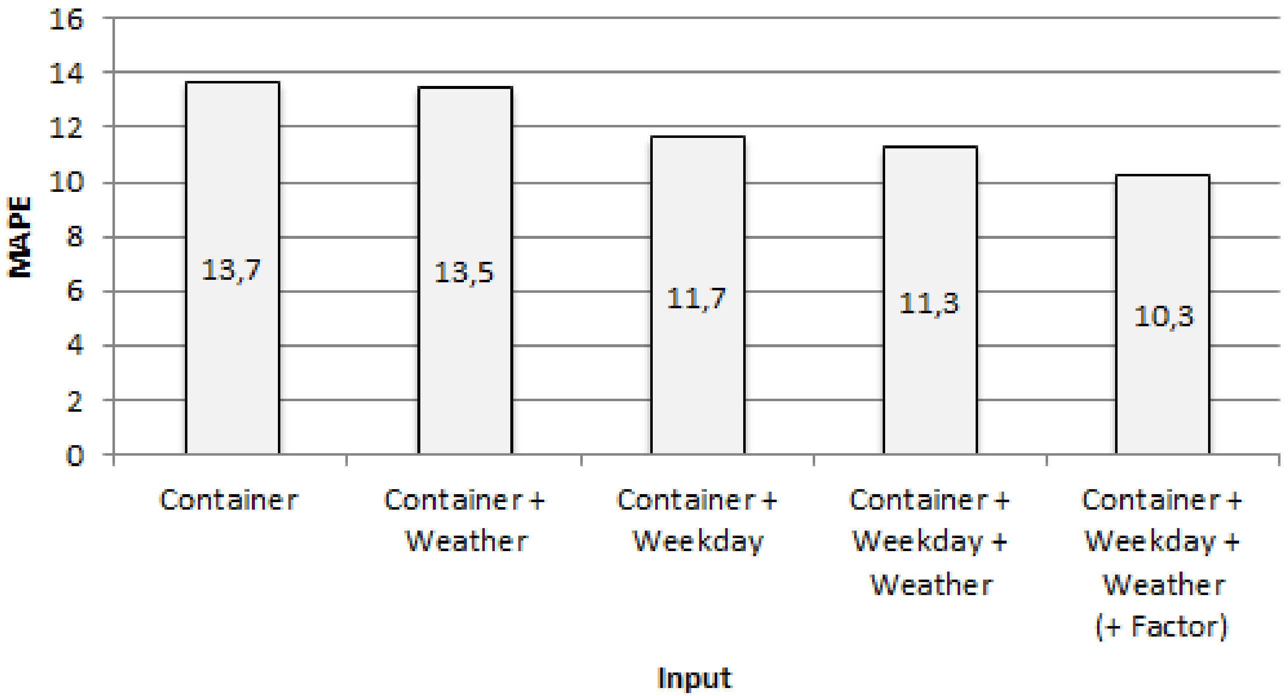
24 Artificial neural networks

Implementation

- Data normalization (MINMAX-normalization)
- Resilient Backpropagation as training method
- Starting with 96 input and 96 output values and 192 neurons in the hidden layer
  - Each input value represents the number of container handlings per quarter hour
  - Each output value represents the power consumption per quarter hour
  - Adding more input values with each iteration:
    - Weather
    - Calendar information
    - A factor representing the yearly increase rate in power consumption
- At the end: 99 input values, 198 neurons hidden and 96 output values

25 Artificial neural networks

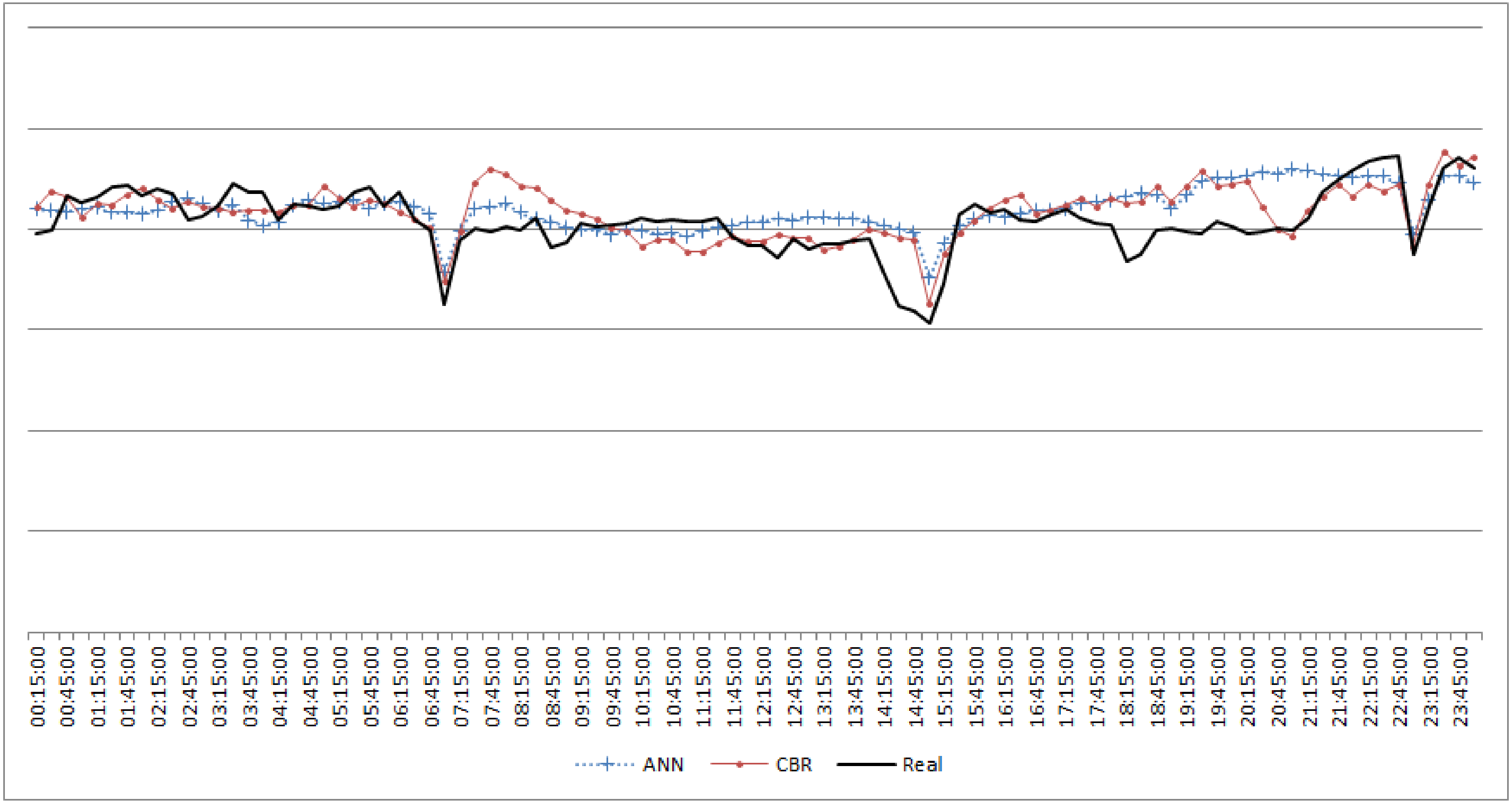
Evaluation - first results



$$MAPE = \frac{100}{N} \sum_{t=1}^{N} \frac{|F_t - A_t|}{A_t}$$

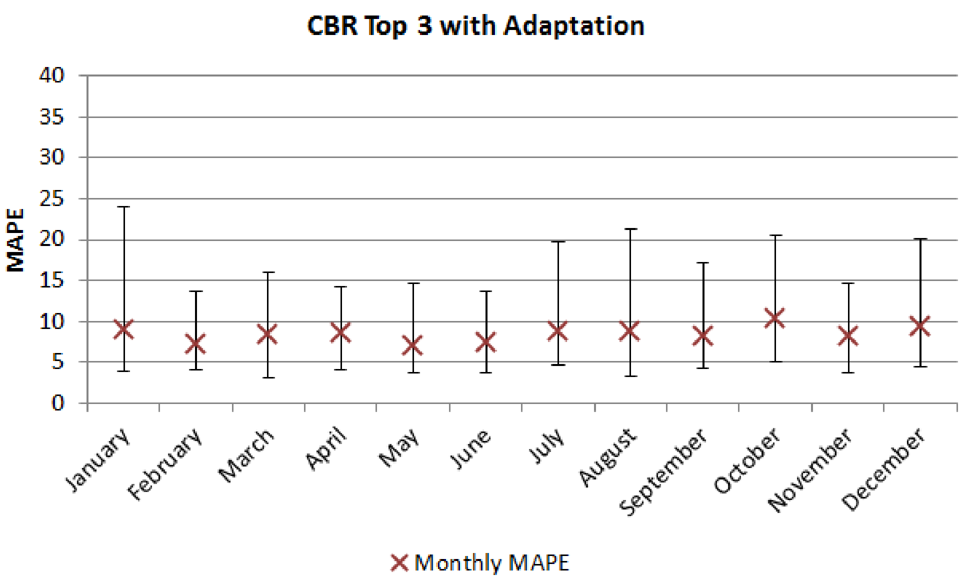
26 Evaluation

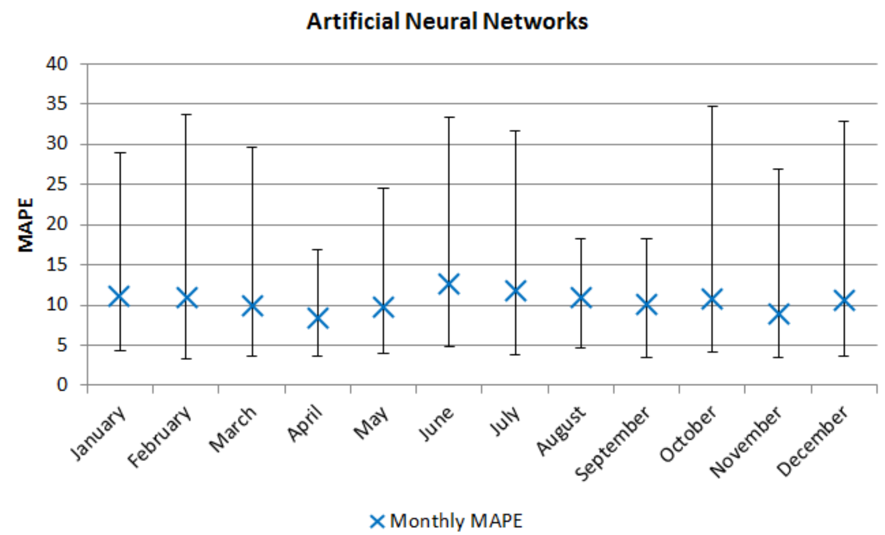
Comparing both approaches – example May 24th 2013



27 Evaluation

Comparing both approaches – monthly results





28 Evaluation & Outlook

Roundup

- Short-Term Load Forecasting for single industrial customers is essential in order to use the flexibility in consumption that they can offer to support the Smart Grid
- Case-Based Reasoning as well as Artificial Neural Networks can be used for Short-Term Load Forecasting based on operational data
- Both approaches offer reasonable results
- Case-Based Reasoning offers less fluctuation in the average results when the Top 3 approach is used
- Further work will include the research of other ANN topologies, e.g. recurrent neural networks and the integration of the previous days load figures

29 Questions and Discussion

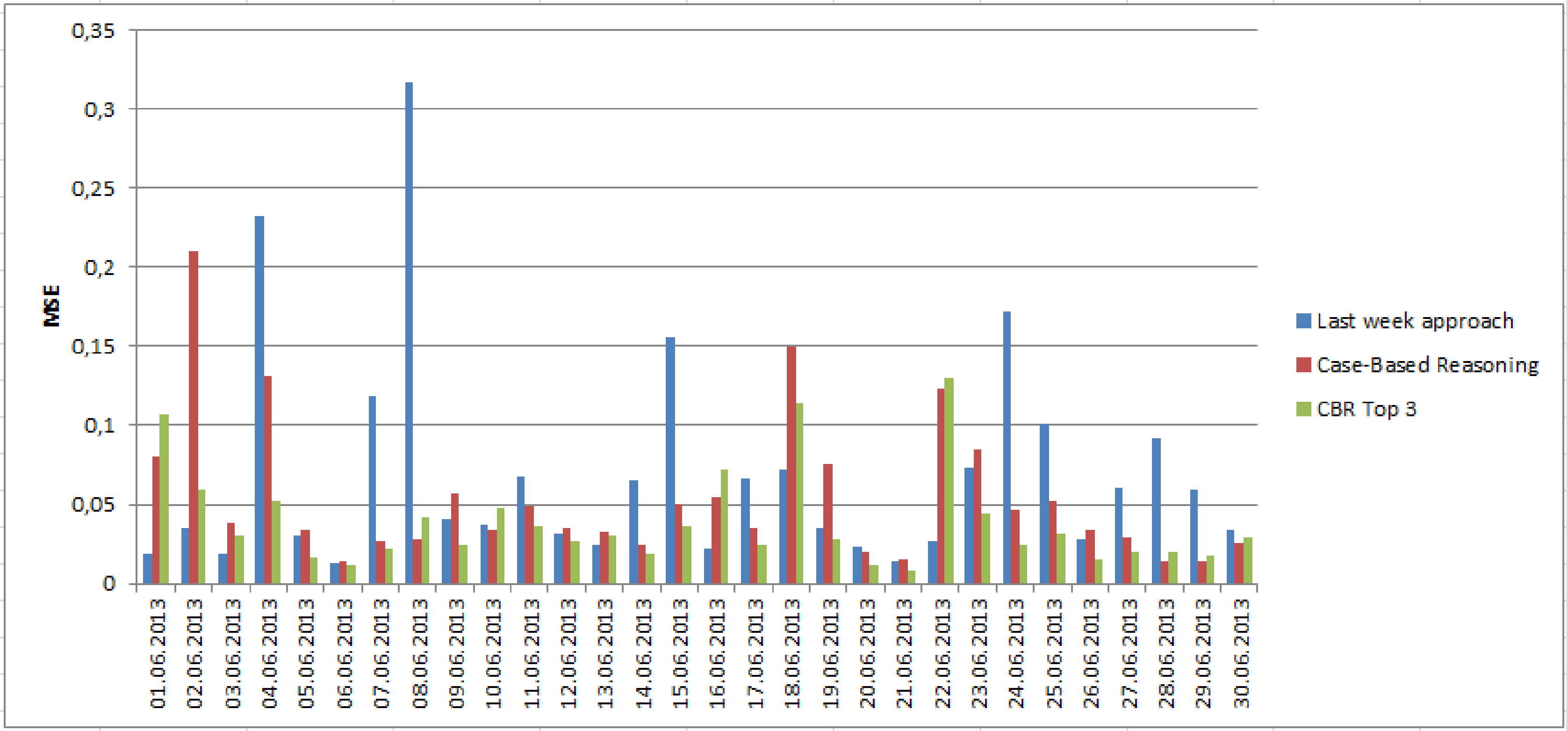
IKT FÜR ELEKTROMOBILITÄT
Gefördert durch:
Bundesministerium für Wirtschaft und Technologie
aufgrund eines Beschlusses des Deutschen Bundestages

30 Backup

31 Evaluation & Outlook

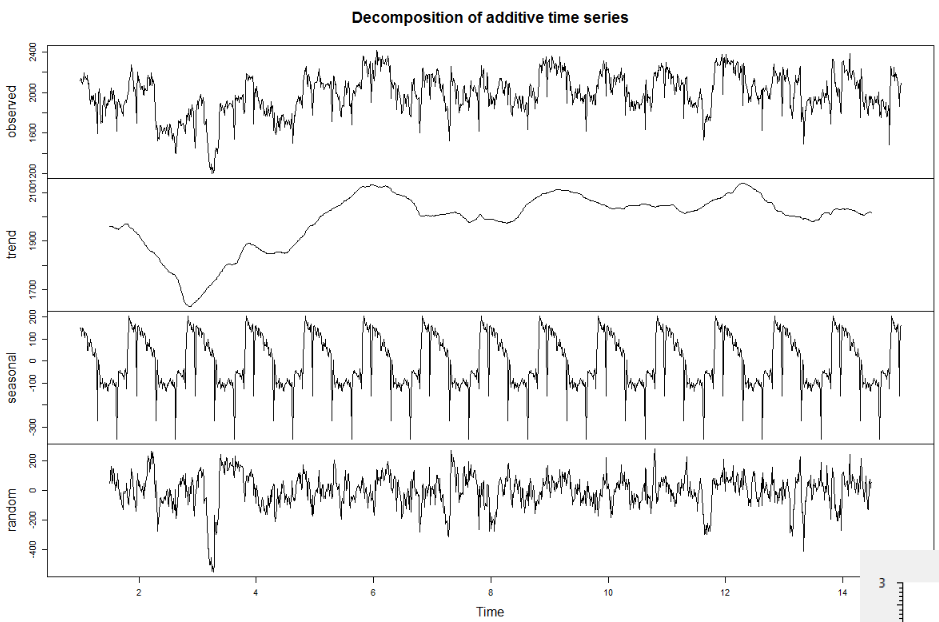
MSE June 2013

- Mean Squared Error -> emphasis large deviations in single values



32 Further Methods for Load-Forecasting

Time Series Analysis: ARIMA

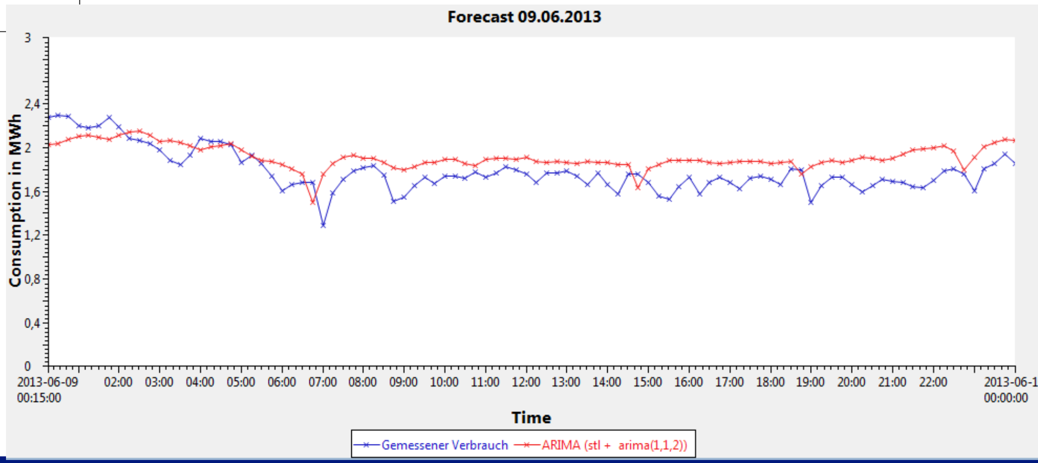


$$y_t = c + \sum_{i=1}^{p} a_i y_{t-i} + \sum_{j=0}^{q} b_j \epsilon_{t-j}$$

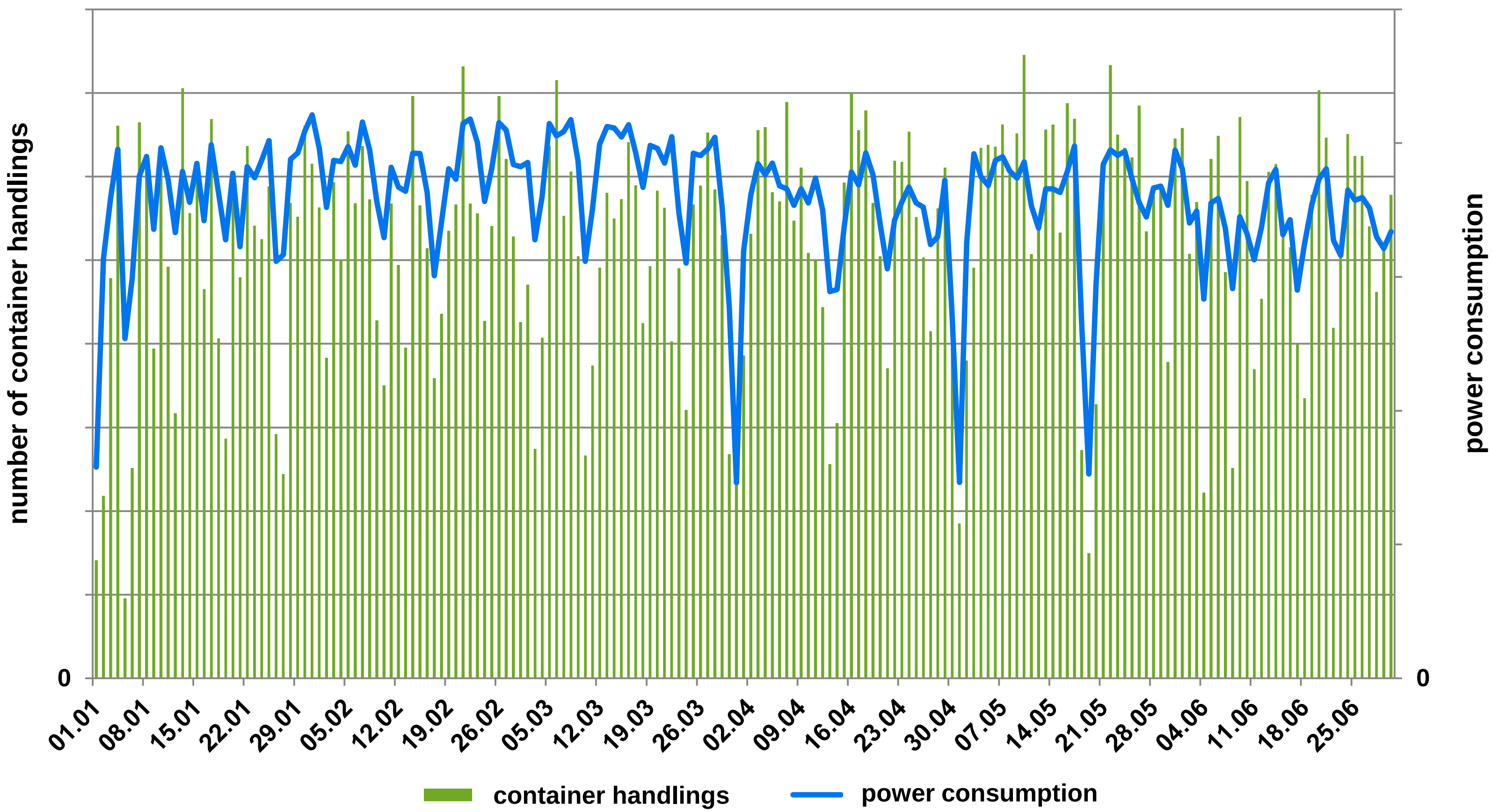
- Usage of the Hyndman-Khandakar algorithm for determination of the ARIMA-model for the random data

- External influences are not integrated yet (ARIMAX)

- Decomposition of time series data is the first step
  - Trend, Season, Random

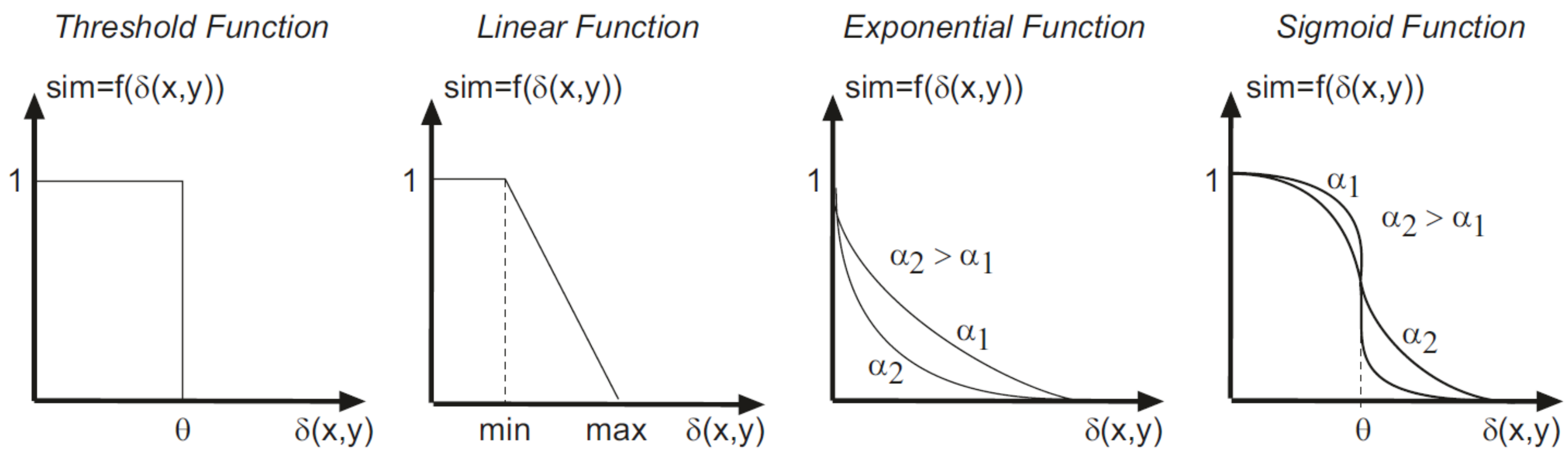


33 Power consumption at a container terminal



34 Case-Based Reasoning
Similarity Measures

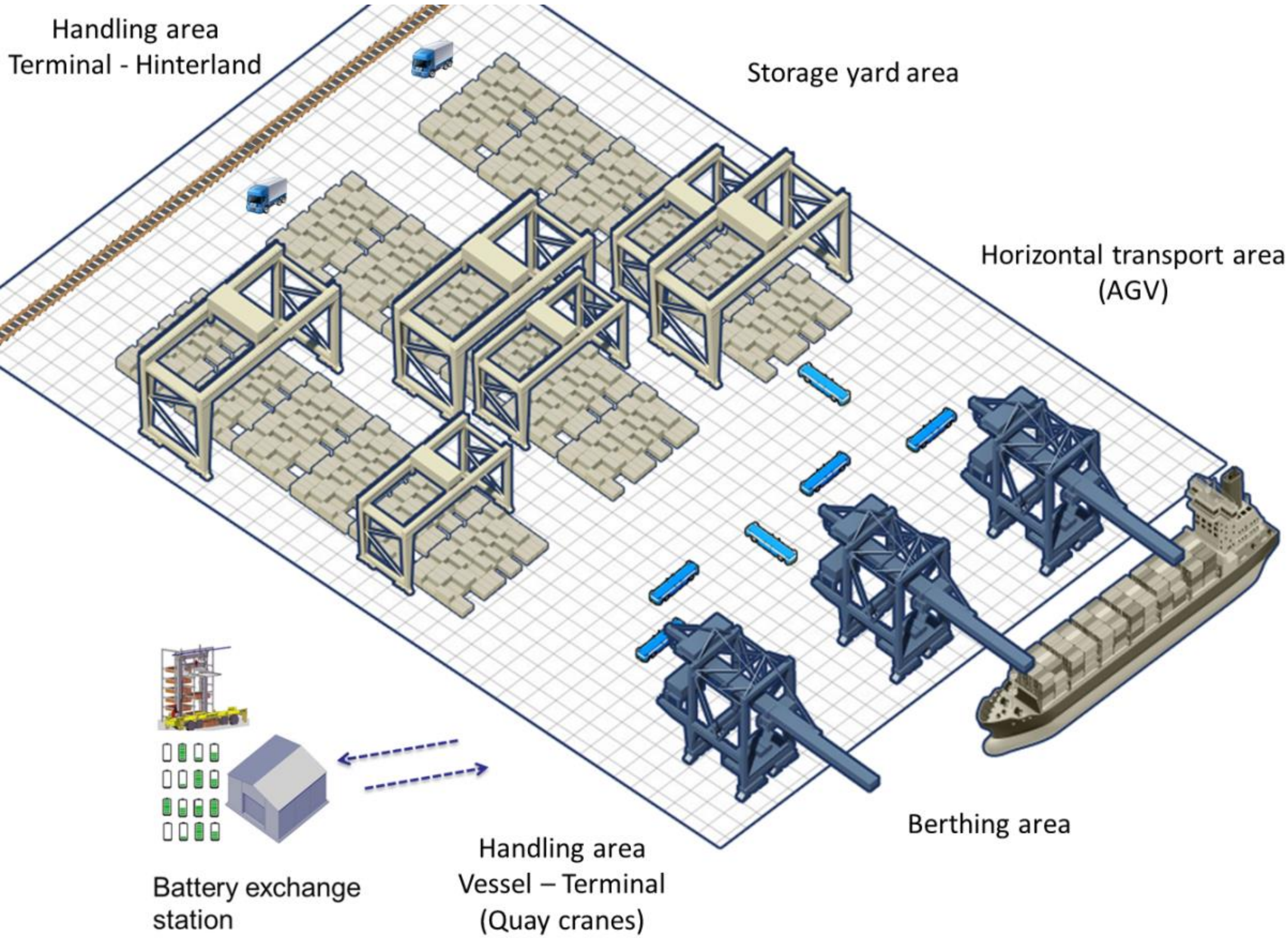
- Choosing similarity measures for each attribute
  - Using similarity measures according to Bergmann (Bergmann 2002)



Source: R. Bergmann: Experience management: foundations, development methodology, and Internet based applications; Springer 2002

35 Introduction & Motivation

An industrial site: maritime container terminal



36 CBR-based Short-Term Load Forecasting

Concept – Similarity Measure for arrival and departure times

- Idea: Use minutes from start of the day till the time as value for similarity assessment

Query:

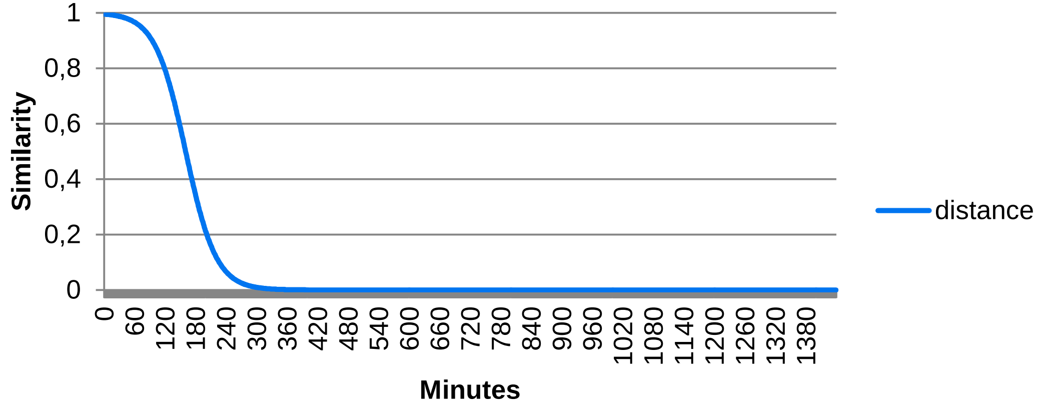
JSNR	Ship Name	Ship Type	Expected Arrival	Expected Departure	Loading	Unloading
308442	A LA MARINE	Feeder	60	1380	534	556

Case:

JSNR	Ship Name	Ship Type	Expected Arrival	Expected Departure	Loading	Unloading
306926	EMOTION	Feeder	405	1435	458	333

- Simple similarity measures for integer values can be used

$$sim_{time}(c,q) = \frac{1}{e^{\frac{d(c,q)-120}{30}}+1}$$



q = query; c = case; d = distance

37 CBR-based Short-Term Load Forecasting

Related Work

Case-based Reasoning in the Energy Domain

	Urosevic et al. 2010	Davidson et al. 2009 Taylor et al. 2010	Vilcahuaman et al. 2004	Wang 2006	Monfet et al. 2013
Object	Decision support for energy efficiency in production environments	Voltage control in distribution networks	Medium-Term Load Forecasting	Forecast of peak loads	STFL for office buildings
Similarity is based on	Metered values of the production site	Metered values of substations	Factors according to the calendar	Factors according to the calendar and weather data	Weather data, currently metered values
Load Forecasting	-	-	+	(+)	+
Usage of operation plans	-	-	-	-	-