TECHNISCHE UNIVERSITÄT DRESDEN

Dresden Database Systems Group

Defense of the Bachelor thesis:

Optimization options for the usage of forecast combinations

Florian Wittenbecher

26.08.2019

Structure

1. Forecast combination: motivation and application

2. Optimization approaches

3. Implementation

4. Results

5. Conclusion

Forecast combination: motivation and application

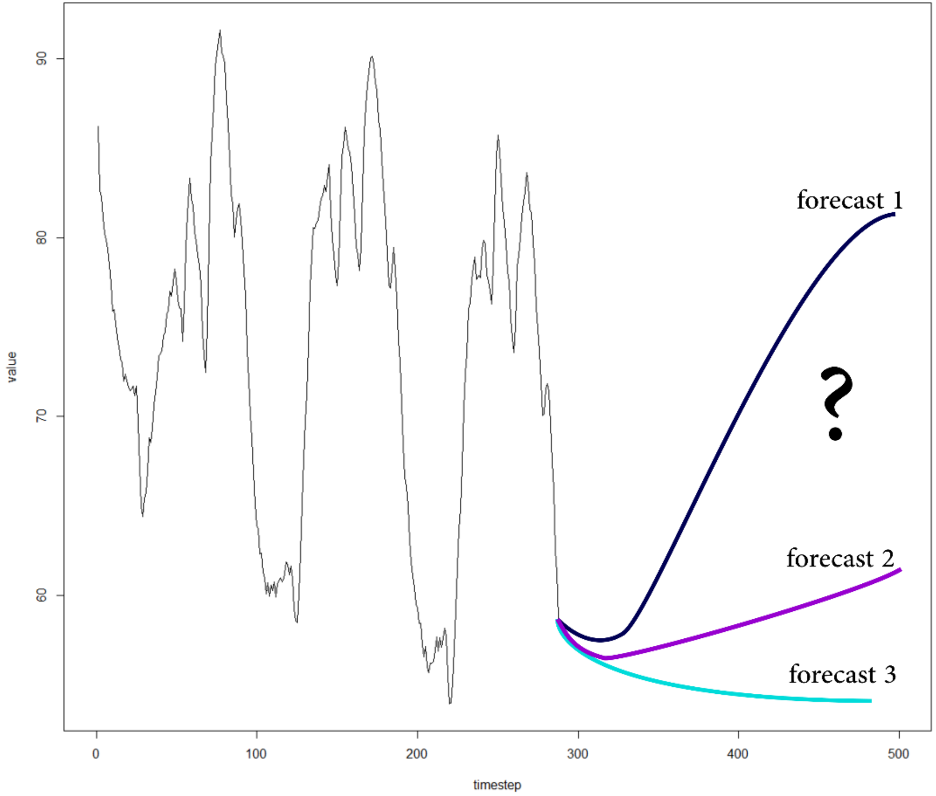
Use cases

Dresden Database Systems Group

Energy consumption

Generation of solar energy

- Idea: use characteristics of several forecasting models at once
- Alternative to model selection



General approach

Combination techniques

- Naive methods: average / median of all values for each time step
- More complex methods: weights are assigned to each forecast

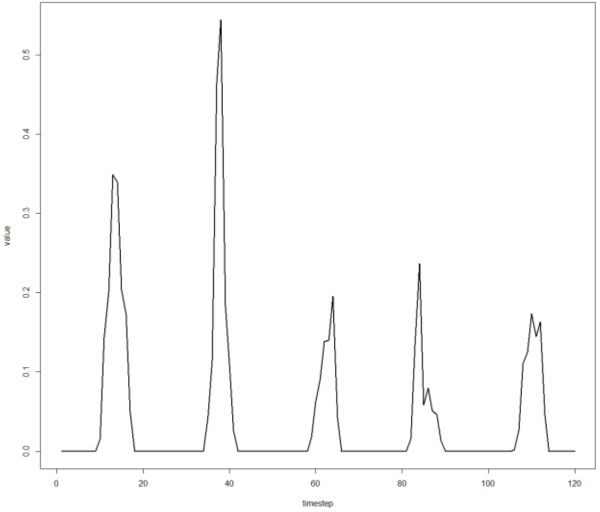
$$C_T = \sum_{i=0}^{N} k_i * f_{i,T}$$

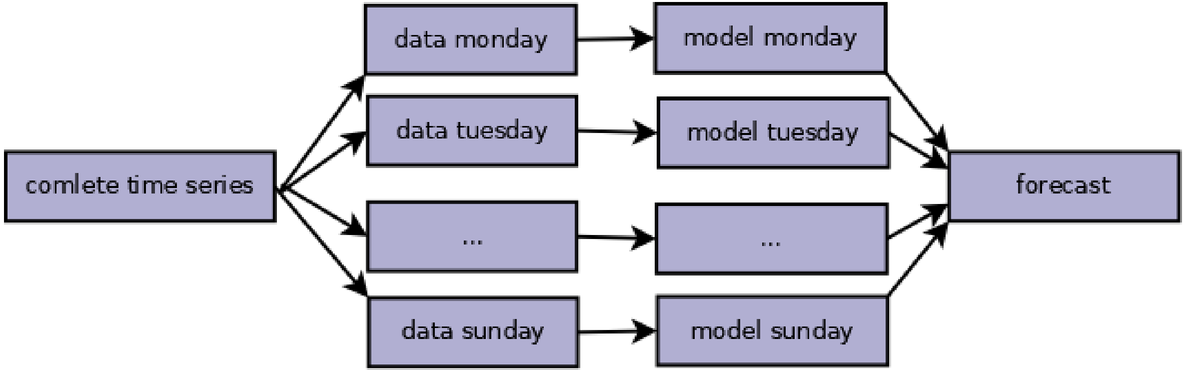
C_T - result of combination at time t
N - number of forecasts
k_i - weight of forecast i
f_{iT} - forecast i at time t

Optimization options

Seasonal filtering

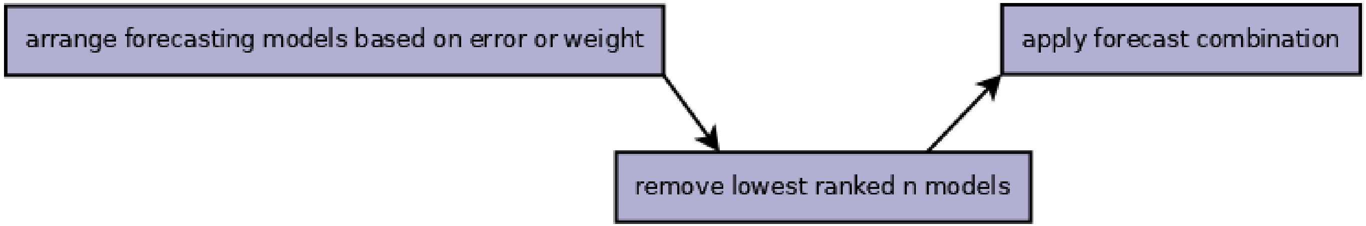
- *Idea: create several combination models for different seasonalities*
- *Assumption: forecast models show differences in their quality for different seasonalities*





Exit of experts

- *Problem: limit the impact of weak forecast models on the combination*
- *Use of less forecast models based on their error*
- *Only models with a high quality are used*



Implementation

Datasets and algorithm

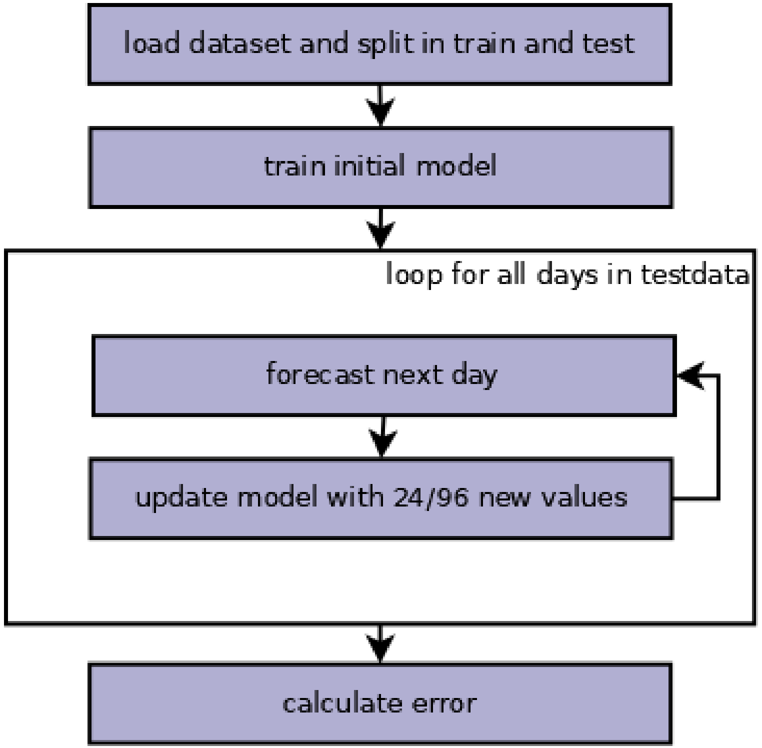
Dresden Database Systems Group

General approach

- Implementation in R
- Main focus on day-ahead forecast

Datasets

- Energy consumption: aggregated consumption of the Swiss canton of Valais every 15 minutes (2015 and 2016) - 70276 observations
  - 6 forecasts
- Production of solar energy: hourly generation in Germany, Denmark and Australia (2013/2015) - 8760 observations each
  - 15 forecasts



Combination techniques

Simple mean

- Average of all given values without weights

Linear Regression

- Using the lm()-package in R

L-BFGS-B

- Quasi-Newton-method using the optim()-package in R

Bayesian model averaging

- Probabilistic method

Opera-package for R

- Contains five methods
- Domain: energy consumption

Optimizations

Dresden Database Systems Group

Seasonal filtering:

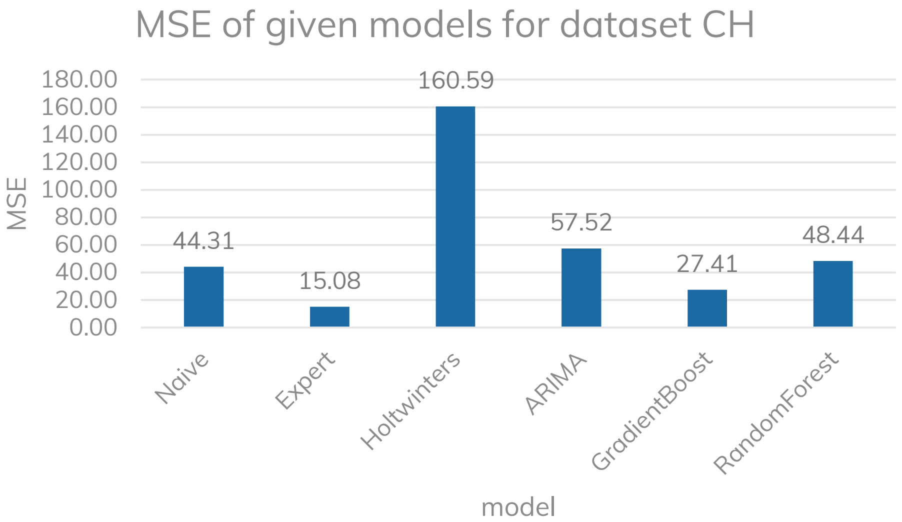
- Month of year (only for energy consumption)
- Day of week
- Hour of day

Exit of experts:

- Remove 50% of the models…
- …based on error (MSE)
- …based on assinged weight

Expectations

- *Models differ in their quality: complex combination techniques should be superior to the simple mean*
- *Exit of experts should improve the quality of the forecast combination*
- *Seasonal Filtering should work especially well on an hourly level*



Error measures

absolute:

$$MSE = \frac{1}{N} \sum_{t=1}^{N} (F_t - A_t)^2$$

percentage:

$$SMAPE = \frac{100\%}{N} \sum_{t=1}^{N} \frac{|F_t - A_t|}{(|A_t| + |F_t|)/2}$$

Results

Dresden Database Systems Group

TECHNISCHE UNIVERSITÄT DRESDEN

Combination methods

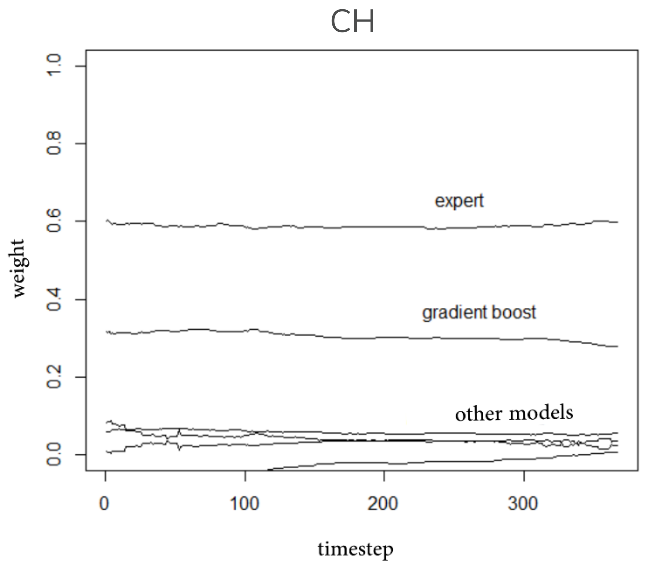
Dresden Database Systems Group

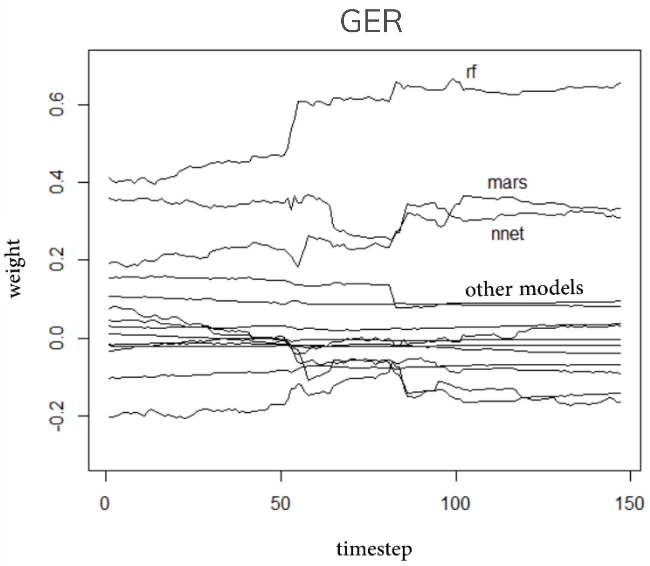
- *Linear regression and bayesian model averaging showed the best results (regression as basis for further optimizations)*
- *Methods of opera-package better for one-step-ahead forecast for energy consumption*

SMAPE	CH	GER	AUS	DEN
Best single model	6,66%	44,26%	16,05%	45,30%
Simple mean (improvement)	6,06%	44,28%	172,28%	37,38%
Linear Regression (improvement)	-18,65%	18,34%	-29,45%	-7,55%
Bayesian model averaging (improvement)	-18,65%	18,88%	-30,79%	-8,58%

Solar data

- *Unstable results for solar data*
- *Fluctuating quality of the given forecast models*





Exit of experts

Dresden Database Systems Group

- *Removing based on error often leads to worse results*
- *Removing based on weights can lead to a smaller error*
- *Mostly just slight changes*

SMAPE	Linear regression	Exit based on weights (improvement)	Exit based on error (improvement)
CH	5,42%	0,44%	2,42%
GER	52,32%	-0,31%	-0,36%
AUS	11,33%	-0,18%	0,93%
DEN	41,88%	0,26%	0,37%

Seasonal Filtering

Dresden Database Systems Group

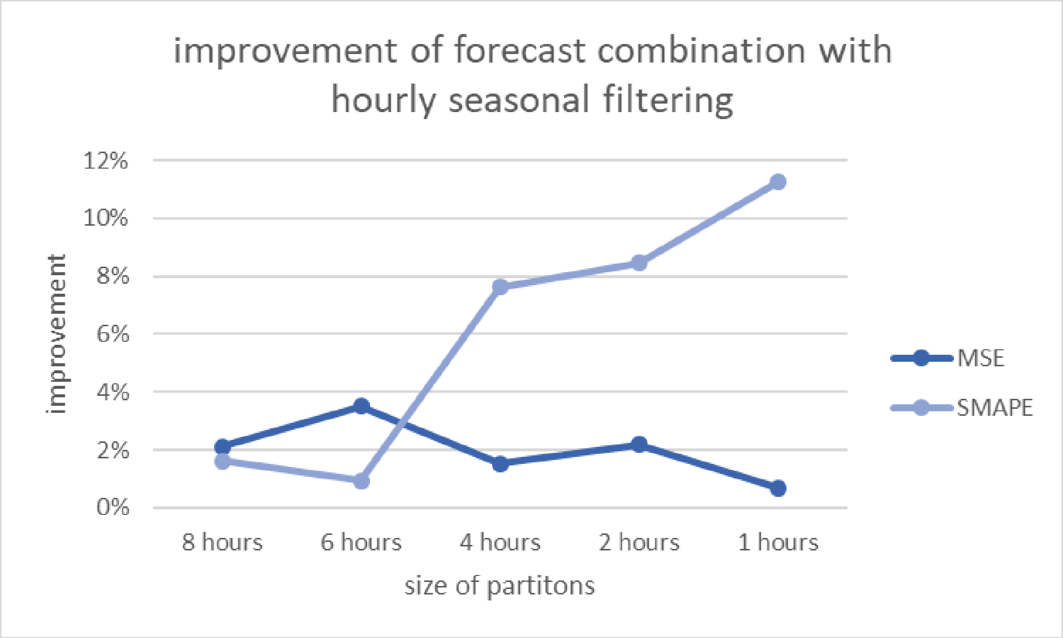
- *Filtering for month of year with a slightly higher error*
- *Filtering for day of week just suitable for energy consumption*
- *Filtering for hour of day improves the result of the combination in most cases*

SMAPE	CH	GER	AUS	DEN
Lin. Regression	5,42%	128,30%	95,85%	164,88%
Season (improvement)	1,17%	Not applicable		
Weekday/ Weekend (improvement)	-0,63%	0,01%	0,03%	0,10%
24 hours (improvement)	-7,24%	-0,67%	-32,57%	-4,56%

Seasonal filtering

Dresden Database Systems Group

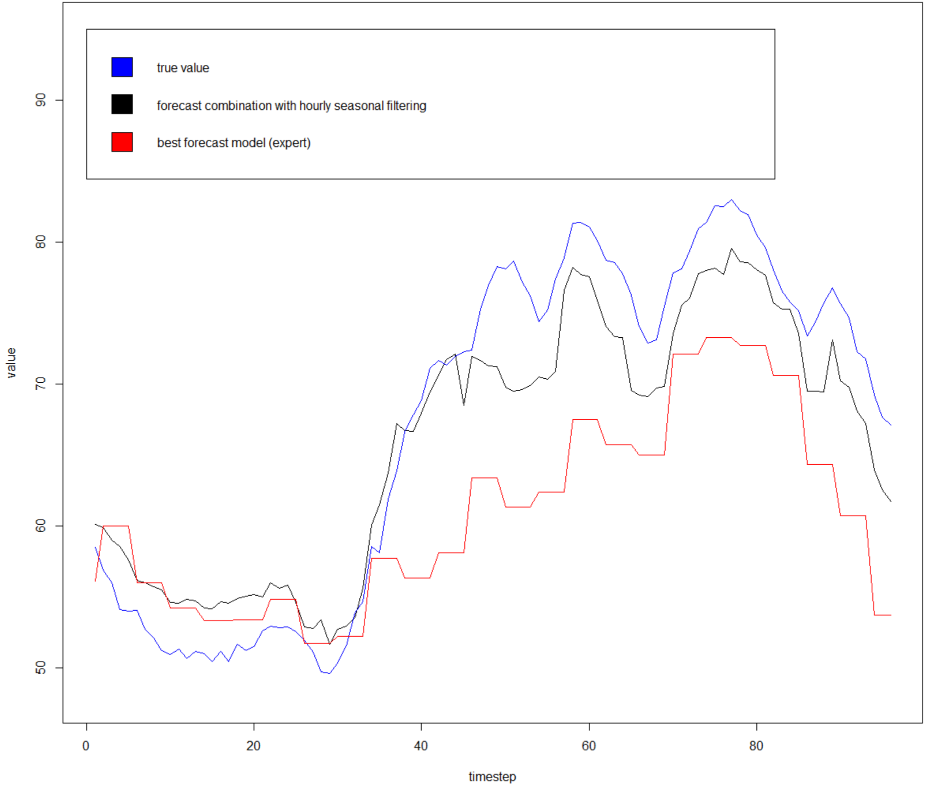
- *Principle: The smaller the partitions, the better the result*



Conclusion

Findings

- *Seasonal filtering has the potential to improve the results of forecast combinations*
- *Recommended especially on an hourly level*
- *Exit of experts can be successfull, but naive approach is not good enough*
- *BUT: combination for solar data partly with a higher error than best forecast model*



Future approaches

Dresden Database Systems Group

Forecast of optimal combination weights for the next day

- Overcome the naive method of using the optimal weights of the previous day

More complex method for the exit of experts

- Time-dependent entry and exit

External influences

- E.g. weather data

Evaluation with other error measures

- MSE and SMAPE showed major differences