Energy Load Forecasting

for the Global Energy Forecasting Competition 2014

Presentation of Semester Project
by Fabian Brix

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

January 27, 2015

Outline

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

GEFCom 2014

- Dates: 08/15/2014 to 12/15/2014
- Tracks: energy load forecasting, price forecasting, wind, solar
- Task: forecast the distribution, in quantiles, of the hourly district load of an energy utility.
- Includes: forecasting the temperature of different zones in the district.
- Forecasting horizon: 1 month
- Competition type: rolling forecast
  $\Rightarrow$ Target for preceding forecasting task published every week to prepare for next horizon

GEFCom 2014

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Dates: 08/15/2014 to 12/15/2014
- Tracks: energy load forecasting, price forecasting, wind, solar
- Task: forecast the distribution, in quantiles, of the hourly district load of an energy utility.
- Includes: forecasting the temperature of different zones in the district.
- Forecasting horizon: 1 month
- Competition type: rolling forecast
  $\Rightarrow$ Target for preceding forecasting task published every week to prepare for next horizon

GEFCom 2014

- Dates: 08/15/2014 to 12/15/2014
- Tracks: energy load forecasting, price forecasting, wind, solar
- Task: forecast the distribution, in quantiles, of the hourly district load of an energy utility.
- Includes: forecasting the temperature of different zones in the district.
- Forecasting horizon: 1 month
- Competition type: rolling forecast
  $\Rightarrow$ Target for preceding forecasting task published every week to prepare for next horizon

GEFCom 2014

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Dates: 08/15/2014 to 12/15/2014
- Tracks: energy load forecasting, price forecasting, wind, solar
- Task: forecast the distribution, in quantiles, of the hourly district load of an energy utility.
- Includes: forecasting the temperature of different zones in the district.
- Forecasting horizon: 1 month
- Competition type: rolling forecast
  $\Rightarrow$ Target for preceding forecasting task published every week to prepare for next horizon

GEFCom 2014

- Dates: 08/15/2014 to 12/15/2014

- Tracks: energy load forecasting, price forecasting, wind, solar
- Task: forecast the distribution, in quantiles, of the hourly district load of an energy utility.
- Includes: forecasting the temperature of different zones in the district.
- Forecasting horizon: 1 month
- Competition type: rolling forecast
  $\Rightarrow$ Target for preceding forecasting task published every week to prepare for next horizon

GEFCom 2014

- Dates: 08/15/2014 to 12/15/2014

- Tracks: energy load forecasting, price forecasting, wind, solar
- Task: forecast the distribution, in quantiles, of the hourly district load of an energy utility.
- Includes: forecasting the temperature of different zones in the district.
- Forecasting horizon: 1 month
- Competition type: rolling forecast
  $\Rightarrow$ Target for preceding forecasting task published every week to prepare for next horizon

GEFCom 2014

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

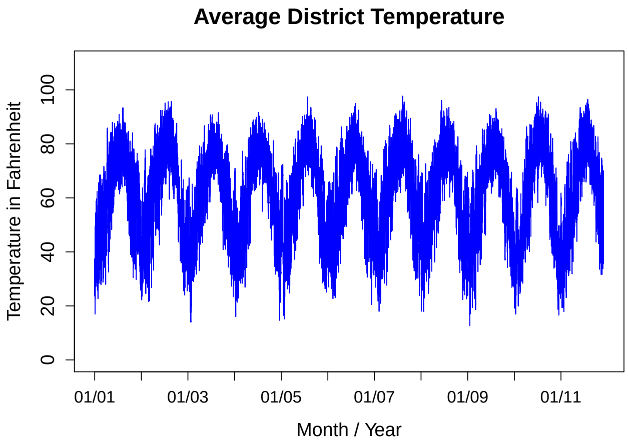
- Dates: 08/15/2014 to 12/15/2014
- Tracks: energy load forecasting, price forecasting, wind, solar
- Task: forecast the distribution, in quantiles, of the hourly district load of an energy utility.
- Includes: forecasting the temperature of different zones in the district.
- Forecasting horizon: 1 month
- Competition type: rolling forecast
  $\Rightarrow$ Target for preceding forecasting task published every week to prepare for next horizon

Outline

- Introduction
- Data
- Features
- Forecasting Methods
- Quantile Forecasting
- Error Measures
- Method Evaluation
- Results

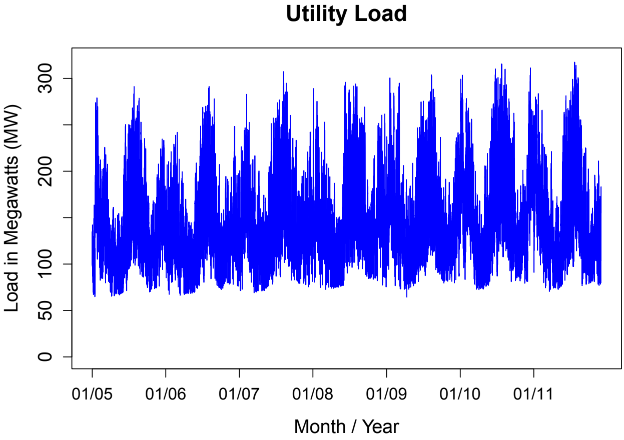
Temperature Data

- zonal: 25 series of hourly temperature (Fahrenheit)
- 01/01/2001 1am to 12/01/2011 midnight.



Load Data

- district level: hourly load (Megawatts)
- 01/01/2005 1am to 12/01/2011 midnight.



Outline

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Introduction
- Data
- Features
- Forecasting Methods
- Quantile Forecasting
- Error Measures
- Method Evaluation
- Results

Features

For temperature & load:

- YLAG: $y_{-365 \text{ days}}$
- DLAG: $y_{-35 \text{ days}}$
- TOY: time of year

For load only:

- CTEMP: current temperature
- DAYT: day type (holiday!)

Type	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Holiday!
SDAYT	2	2	2	2	2	1	1	-
WDAYT	2	3	4	5	6	7	1	-
DAYT	2	3	4	5	6	7	1	8

Features

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

For temperature & load:

- YLAG: $y_{-365 \text{ days}}$
- DLAG: $y_{-35 \text{ days}}$
- TOY: time of year

For load only:

- CTEMP: current temperature
- DAYT: day type (holiday!)

Type	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Holiday!
SDAYT	2	2	2	2	2	1	1	-
WDAYT	2	3	4	5	6	7	1	-
DAYT	2	3	4	5	6	7	1	8

Features

For temperature & load:

- YLAG: $y_{-365 \text{ days}}$
- DLAG: $y_{-35 \text{ days}}$
- TOY: time of year

For load only:

- CTEMP: current temperature
- DAYT: day type (holiday!)

Type	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Holiday!
SDAYT	2	2	2	2	2	1	1	-
WDAYT	2	3	4	5	6	7	1	-
DAYT	2	3	4	5	6	7	1	8

Outline

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Generalized Additive Model (GAM)

- Additive predictor with non-linear functions
- $y = s_1(x_1) + s_2(x_2) + \cdots + s_p(x_p) + \varepsilon$
- $s_i(x_i)$ are smooth functions estimated from the data x_i
- Noise ε assumed to be Gaussian
- Use of standard configuration of GAM in R package "mgcv"

Generalized Additive Model (GAM)

- Additive predictor with non-linear functions
- $y = s_1(x_1) + s_2(x_2) + \cdots + s_p(x_p) + \varepsilon$
- $s_i(x_i)$ are smooth functions estimated from the data x_i
- Noise ε assumed to be Gaussian
- Use of standard configuration of GAM in R package “mgcv”

Generalized Additive Model (GAM)

- Additive predictor with non-linear functions
- $y = s_1(x_1) + s_2(x_2) + \cdots + s_p(x_p) + \varepsilon$
- $s_i(x_i)$ are smooth functions estimated from the data x_i
- Noise ε assumed to be Gaussian
- Use of standard configuration of GAM in R package "mgcv"

Generalized Additive Model (GAM)

- Additive predictor with non-linear functions
- $y = s_1(x_1) + s_2(x_2) + \cdots + s_p(x_p) + \varepsilon$
- $s_i(x_i)$ are smooth functions estimated from the data x_i
- Noise ε assumed to be Gaussian
- Use of standard configuration of GAM in R package "mgcv"

Generalized Additive Model (GAM)

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Additive predictor with non-linear functions
- $y = s_1(x_1) + s_2(x_2) + \cdots + s_p(x_p) + \varepsilon$
- $s_i(x_i)$ are smooth functions estimated from the data x_i
- Noise ε assumed to be Gaussian
- Use of standard configuration of GAM in R package “mgcv”

Feedforward Neural Network (NN)

Configuration using R package nnet

- Hidden layer: 1
- Hidden units: 5, 10, 15, 20, 25, 30
- Linear output for regression

Feedforward Neural Network (NN)

Configuration using R package nnet

- Hidden layer: 1
- Hidden units: 5, 10, 15, 20, 25, 30
- Linear output for regression

Feedforward Neural Network (NN)

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Configuration using R package nnet

- Hidden layer: 1
- Hidden units: 5, 10, 15, 20, 25, 30
- Linear output for regression

Feedforward Neural Network (NN)

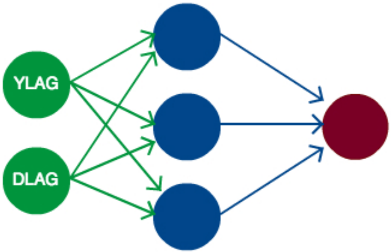
Configuration using R package nnet

- Hidden layer: 1
- Hidden units: 5, 10, 15, 20, 25, 30
- Linear output for regression

Feedforward Neural Network (NN)

Configuration using R package nnet

- Hidden layer: 1
- Hidden units: 5, 10, 15, 20, 25, 30
- Linear output for regression



Random Forest (RF)

Configuration using R package randomForest

- # trees: 50, 100, 150, 300

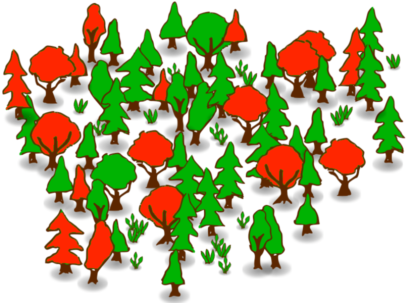
Configuration using R package randomForest

- # trees: 50, 100, 150, 300

Random Forest (RF)

Configuration using R package randomForest

- # trees: 50, 100, 150, 300



Forecasting Models

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Model: hyperparameter settings + formula
- Formula: collection of features used in model (R)
  $\Rightarrow$ Create point forecasts with different forecasting methods

Forecasting Models

- Model: hyperparameter settings + formula
- Formula: collection of features used in model (R)

$\Rightarrow$ Create point forecasts with different forecasting methods

- Model: hyperparameter settings + formula
- Formula: collection of features used in model (R)
  $\Rightarrow$ Create point forecasts with different forecasting methods

Outline

- Introduction
- Data
- Features
- Forecasting Methods
- Quantile Forecasting
- Error Measures
- Method Evaluation
- Results

Quantile Forecasting

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Future: random variable with "forecast distribution"
- "Quantile forecast": quantiles of this distribution (percentiles)
- Point forecasts correspond to mean of forecast distribution
  $\rightarrow$ Offset quantiles of training residuals by point forecast

Quantile Forecasting

- Future: random variable with "forecast distribution"
- "Quantile forecast": quantiles of this distribution (percentiles)
- Point forecasts correspond to mean of forecast distribution
  $\rightarrow$ Offset quantiles of training residuals by point forecast

Quantile Forecasting

- Future: random variable with "forecast distribution"
- "Quantile forecast": quantiles of this distribution (percentiles)
- Point forecasts correspond to mean of forecast distribution
  $\rightarrow$ Offset quantiles of training residuals by point forecast

Quantile Forecasting

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

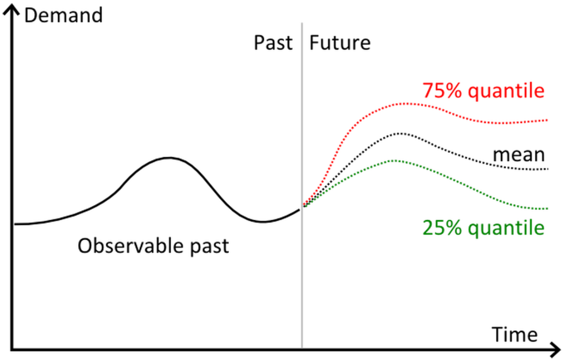
- Future: random variable with "forecast distribution"
- "Quantile forecast": quantiles of this distribution (percentiles)
- Point forecasts correspond to mean of forecast distribution
  $\rightarrow$ Offset quantiles of training residuals by point forecast

Quantile Forecasting

- Future: random variable with “forecast distribution”
- “Quantile forecast”: quantiles of this distribution (percentiles)
- Point forecasts correspond to mean of forecast distribution
  $\rightarrow$ Offset quantiles of training residuals by point forecast

Quantile Forecasting

- Future: random variable with "forecast distribution"
- "Quantile forecast": quantiles of this distribution (percentiles)
- Point forecasts correspond to mean of forecast distribution
  $\rightarrow$ Offset quantiles of training residuals by point forecast



Outline

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Introduction
- Data
- Features
- Forecasting Methods
- Quantile Forecasting
- Error Measures
- Method Evaluation
- Results

Mean Average Percentage Error

- Used to evaluate goodness of fit
- Scale-independent (%)

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{\xi_i}{y_i}\right|$$

- residual: $\xi_i = y_i - \hat{y}_i$
- y_i: time series value at time i
- $\hat{y}_i$: forecast at time i

Mean Average Percentage Error

- Used to evaluate goodness of fit
- Scale-independent (%)

$$\mathrm{MAPE} = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{\xi_i}{y_i} \right|$$

- residual: $\xi_i = y_i - \hat{y}_i$
- y_i: time series value at time i
- $\hat{y}_i$: forecast at time i

Mean Average Percentage Error

- Used to evaluate goodness of fit
- Scale-independent (%)

$$\text{MAPE} = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{\xi_i}{y_i} \right|$$

- residual: $\xi_i = y_i - \hat{y}_i$
- y_i: time series value at time i
- $\hat{y}_i$: forecast at time i

Mean Average Percentage Error

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Used to evaluate goodness of fit
- Scale-independent (%)

$$\text{MAPE} = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{\xi_i}{y_i} \right|$$

- residual: $\xi_i = y_i - \hat{y}_i$
- y_i: time series value at time i
- $\hat{y}_i$: forecast at time i

Mean Average Percentage Error

- Used to evaluate goodness of fit
- Scale-independent (%)

$$\mathrm{MAPE} = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{\xi_i}{y_i} \right|$$

- residual: $\xi_i = y_i - \hat{y}_i$
- y_i: time series value at time i
- $\hat{y}_i$: forecast at time i

Pinball Loss Function

- Used to evaluate quantile forecasts
- And model performance against competition leaderboard
- Strictly positive, the lower the more accurate the forecast
- Score at i given by mean$(L_{0.01}, L_{0.02}, \ldots, L_{0.99})$

$$L_\tau(\xi) = \begin{cases} \tau\xi & \text{if } \xi \geq 0 \\ (\tau - 1)\xi & \text{if } \xi < 0 \end{cases} \quad \text{where } \xi = (y_i - \hat{y}_{i,\tau})$$

Pinball Loss Function

- Used to evaluate quantile forecasts
- And model performance against competition leaderboard
- Strictly positive, the lower the more accurate the forecast
- Score at i given by mean$(L_{0.01}, L_{0.02}, \ldots, L_{0.99})$

$$L_\tau(\xi) = \begin{cases} \tau\xi & \text{if } \xi \geq 0 \\ (\tau - 1)\xi & \text{if } \xi < 0 \end{cases} \quad \text{where } \xi = (y_i - \hat{y}_{i,\tau})$$

Pinball Loss Function

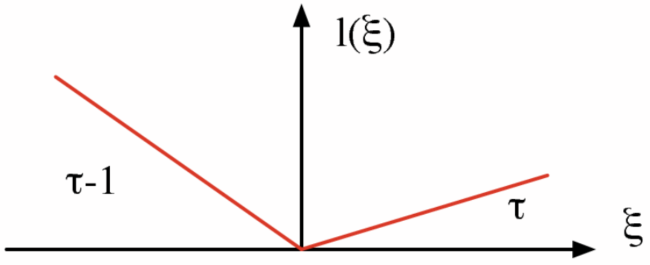
- Used to evaluate quantile forecasts
- And model performance against competition leaderboard
- Strictly positive, the lower the more accurate the forecast
- Score at i given by mean$(L_{0.01}, L_{0.02}, \ldots, L_{0.99})$

$$L_\tau(\xi) = \begin{cases} \tau\xi & \text{if } \xi \geq 0 \\ (\tau - 1)\xi & \text{if } \xi < 0 \end{cases} \quad \text{where } \xi = (y_i - \hat{y}_{i,\tau})$$

Pinball Loss Function

- Used to evaluate quantile forecasts
- And model performance against competition leaderboard
- Strictly positive, the lower the more accurate the forecast
- Score at i given by mean$(L_{0.01}, L_{0.02}, \ldots, L_{0.99})$

$$L_\tau(\xi) = \begin{cases} \tau\xi & \text{if } \xi \geq 0 \\ (\tau - 1)\xi & \text{if } \xi < 0 \end{cases} \quad \text{where } \xi = (y_i - \hat{y}_{i,\tau})$$



Outline

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Introduction
- Data
- Features
- Forecasting Methods
- Quantile Forecasting
- Error Measures
- Method Evaluation
- Results

Method Evaluation

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Use whole competition dataset
- Initial data + 14 months from subsequent tasks
- Use rolling forecast of 14 month period to evaluate forecasts

Method Evaluation

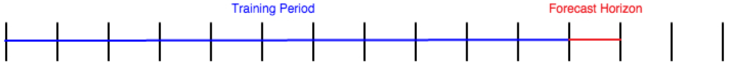
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Use whole competition dataset
- Initial data + 14 months from subsequent tasks
- Use rolling forecast of 14 month period to evaluate forecasts

Method Evaluation

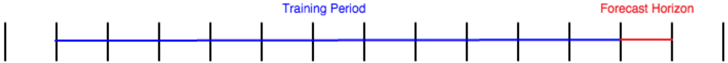
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Use whole competition dataset
- Initial data + 14 months from subsequent tasks
- Use rolling forecast of 14 month period to evaluate forecasts



Method Evaluation

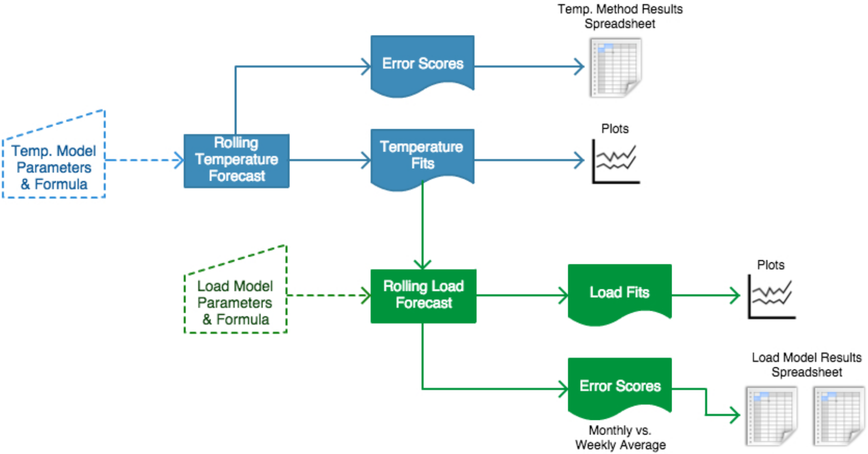
- Use whole competition dataset
- Initial data + 14 months from subsequent tasks
- Use rolling forecast of 14 month period to evaluate forecasts



Method Evaluation

- Use whole competition dataset
- Initial data + 14 months from subsequent tasks
- Use rolling forecast of 14 month period to evaluate forecasts

Workflow Diagram



Outline

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Introduction
- Data
- Features
- Forecasting Methods
- Quantile Forecasting
- Error Measures
- Method Evaluation
- Results

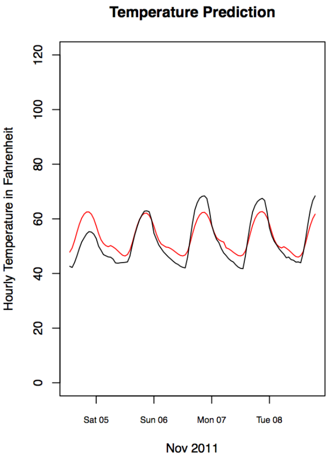
Results - Temperature

- Best score: 12.04% MAPE
- Model: neural network and 30 hidden units

Results - Temperature

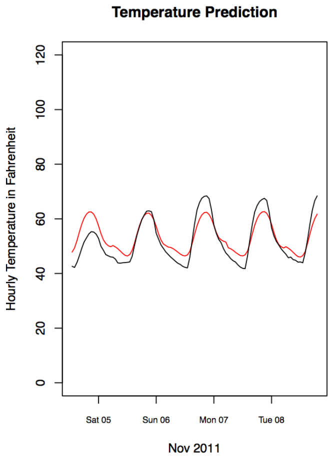
- Best score: 12.04% MAPE
- Model: neural network and 30 hidden units

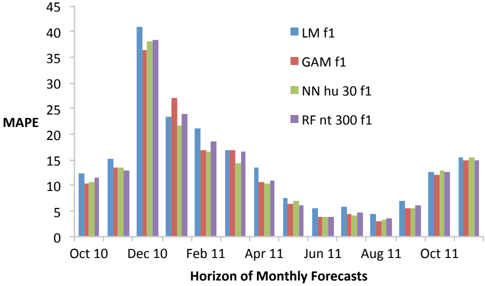
- Best score: 12.04% MAPE
- Model: neural network and 30 hidden units



Results - Temperature

- Best score: 12.04% MAPE
- Model: neural network and 30 hidden units





Results - Load

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Effect of temperature
  $\rightarrow$ Signifcant: -2.5% MAPE
- No effect of different temperature preprocessing (average, PCA, 1 series)
- Predicting temperature for training period of load forecast
  $\rightarrow$ Pinball error improves
- Effect of forecast horizon
  $\rightarrow$ results improved with 1 week horizon
- Effect of daytypes
  $\rightarrow$ better results without

Results - Load

- Effect of temperature
  $\rightarrow$ Signifcant: -2.5% MAPE
- No effect of different temperature preprocessing (average, PCA, 1 series)
- Predicting temperature for training period of load forecast
  $\rightarrow$ Pinball error improves
- Effect of forecast horizon
  $\rightarrow$ results improved with 1 week horizon
- Effect of daytypes
  $\rightarrow$ better results without

Results - Load

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Effect of temperature
  $\rightarrow$ Signifcant: -2.5% MAPE
- No effect of different temperature preprocessing (average, PCA, 1 series)
- Predicting temperature for training period of load forecast
  $\rightarrow$ Pinball error improves
- Effect of forecast horizon
  $\rightarrow$ results improved with 1 week horizon
- Effect of daytypes
  $\rightarrow$ better results without

Results - Load

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Effect of temperature
  $\rightarrow$ Signifcant: -2.5% MAPE
- No effect of different temperature preprocessing (average, PCA, 1 series)
- Predicting temperature for training period of load forecast
  $\rightarrow$ Pinball error improves
- Effect of forecast horizon
  $\rightarrow$ results improved with 1 week horizon
- Effect of daytypes
  $\rightarrow$ better results without

Results - Load

- Effect of temperature
  $\rightarrow$ Signifcant: -2.5% MAPE
- No effect of different temperature preprocessing (average, PCA, 1 series)
- Predicting temperature for training period of load forecast
  $\rightarrow$ Pinball error improves
- Effect of forecast horizon
  $\rightarrow$ results improved with 1 week horizon
- Effect of daytypes
  $\rightarrow$ better results without

Best Configuration

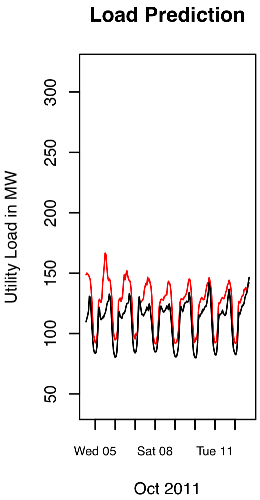
- Random forest, 100 trees, 1 week f. horizon
- MAPE: 16.1%, Pinball: 7.999

Best Configuration

- Random forest, 100 trees, 1 week f. horizon
- MAPE: 16.1%, Pinball: 7.999

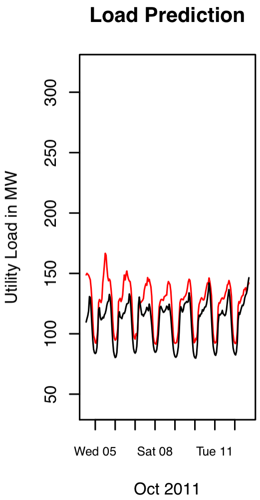
Best Configuration

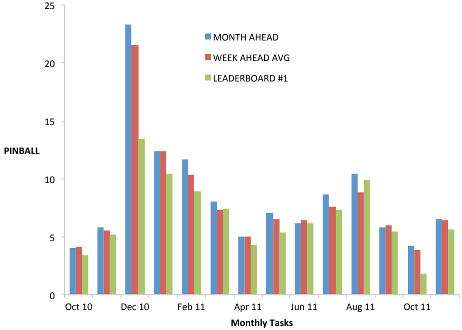
- Random forest, 100 trees, 1 week f. horizon
- MAPE: 16.1%, Pinball: 7.999



Best Configuration

- Random forest, 100 trees, 1 week f. horizon
- MAPE: 16.1%, Pinball: 7.999

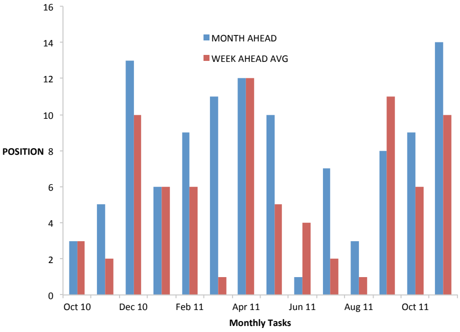




Best Configuration - Final Leaderboard

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Trimmed pinball mean: 7.16 $\Rightarrow$ #3
- Average position on task leaderboards: 5.64
- Not every contestant reaches equally high scores every week



Further Work

- Find a development environment which makes debugging easier (R in vim)
- Analyze training residuals $\rightarrow$ white noise
- Further research on random forests for regression
- Produce and evalute ensemble forecasts

Further Work

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Find a development environment which makes debugging easier (R in vim)
- Analyze training residuals → white noise
- Further research on random forests for regression
- Produce and evalute ensemble forecasts

Further Work

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

- Find a development environment which makes debugging easier (R in vim)
- Analyze training residuals $\rightarrow$ white noise
- Further research on random forests for regression
- Produce and evalute ensemble forecasts

Further Work

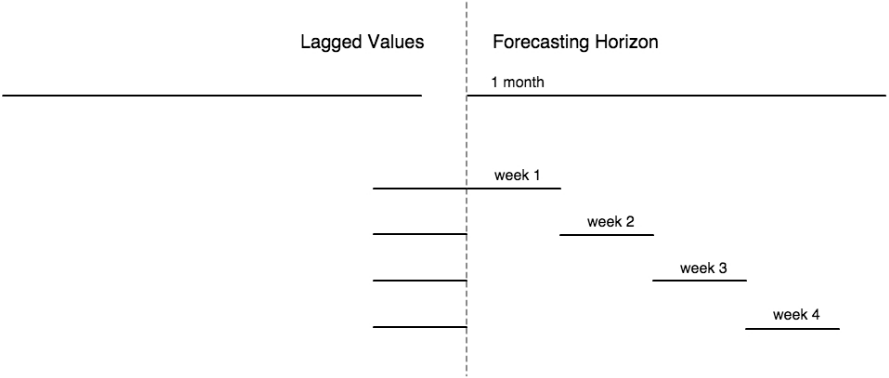
- Find a development environment which makes debugging easier (R in vim)
- Analyze training residuals $\rightarrow$ white noise
- Further research on random forests for regression
- Produce and evalute ensemble forecasts

Further Work

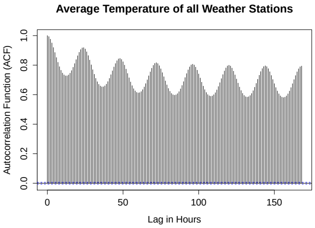
- Find a development environment which makes debugging easier (R in vim)
- Analyze training residuals $\rightarrow$ white noise
- Further research on random forests for regression
- Produce and evalute ensemble forecasts

Forecasting Horizon - Trick

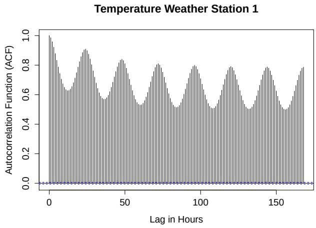
- Forecasts with two different time horizons: 1 month vs 1 week
- Forecasts of week 1 to 4 are combined + rest of monthly forecast



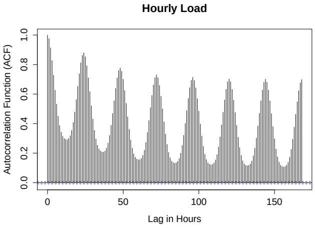
Features



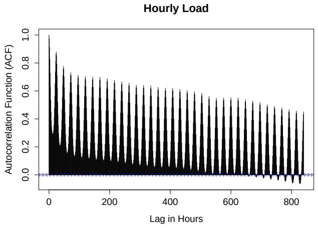
(a) time lag: 7 days, 168 hours



(b) time lag: 7 days, 168 hours



(c) lag: 7 days, 168 hours



(d) lag: 35 days, 840 hours

Residual Plot