Temperature correction in meteorological forecasts

● ● ●

Richard Ižip
Supervisor: doc. RNDr. Robert Lukoťka PhD.

Problem description

- Postprocessing of temperature forecasts created by Aladin
- Focus on general and stable weather model
- Combine models together

Dataset

- Data supplied by SHMI
- Approximately 3 years period
- 35 stations
- Observations for temperature, humidity, wind speed, wind direction, pressure and rainfall within last hour
- Temperature forecasts made by Aladin for 1 to 72 hours ahead
- Our models forecast 1 to 12 hours ahead

Related work

Numerical weather prediction postprocessing (Hajdin, 2016)

- Linear regression model
- Rolling window with 12 hour period
- General model (Aladin temperature, temperature, humidity, wind direction)
- $y_i^h = \theta_0 + \theta_1 P_i^h + \theta_2 T_i + \theta_3 H_i + \theta_4 W_i \qquad h \in 1,...,12$
- Autocorrection model
- MAE improvement from 13% up to 57%
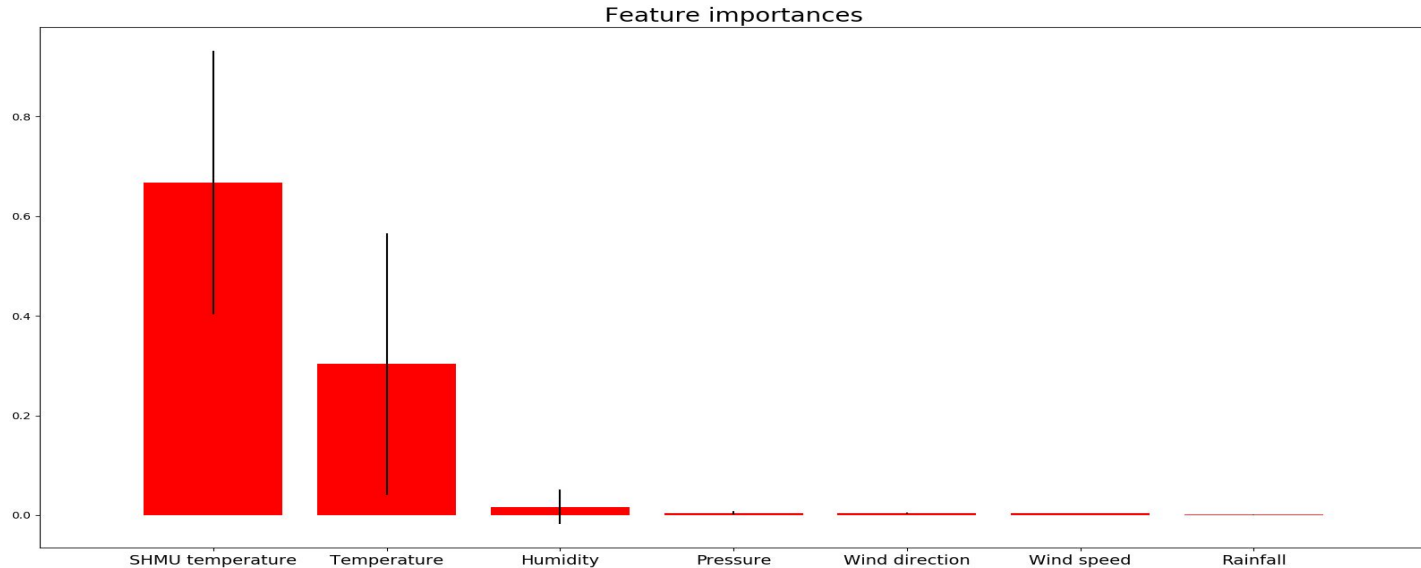
Samples

Rolling window

Adjust rolling window

- Change length and period
- Use station Bratislava-Airport for evaluation
- $y_i^h = \theta_0 + \theta_1 T_i + \theta_2 P_i^h$ $h \in 1,...,12$ (Reference model)
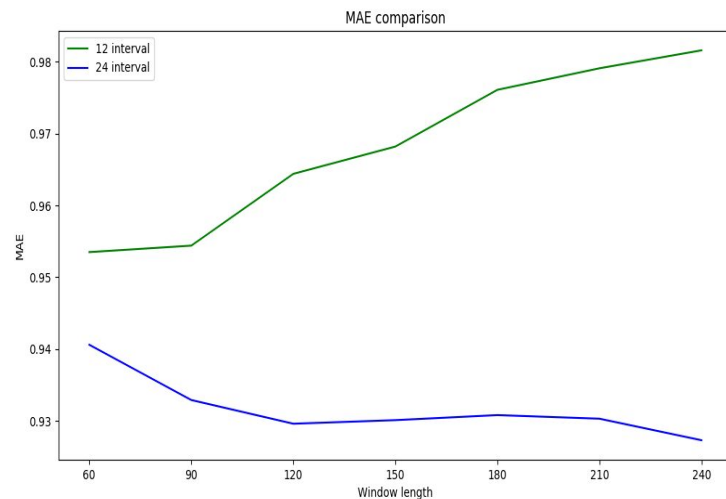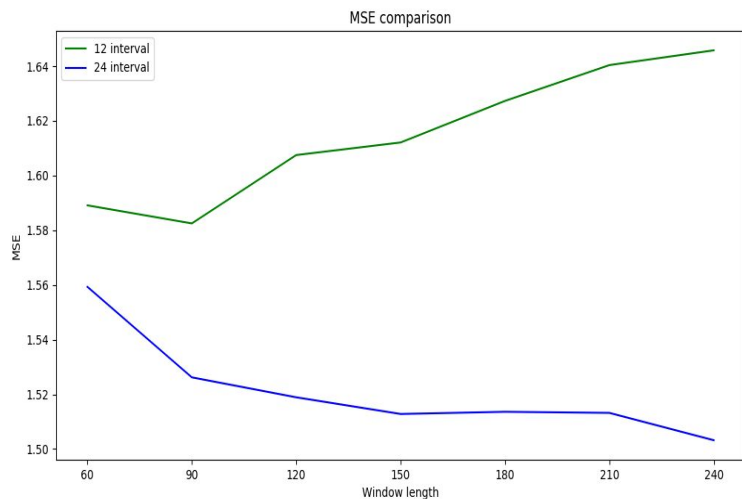
Selecting features

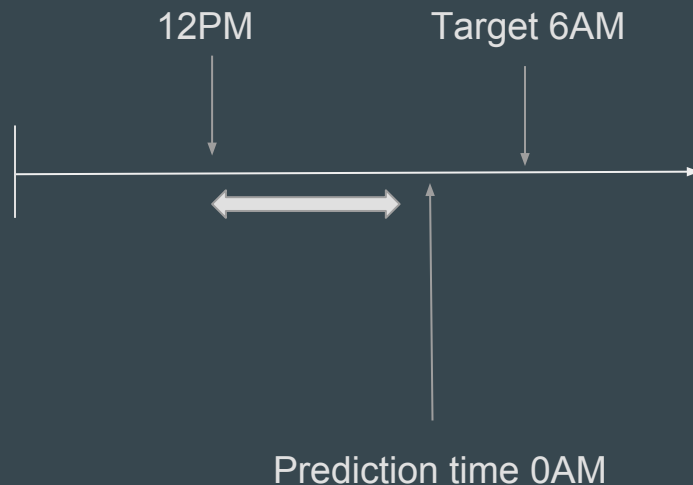- Feature importance (Extra trees regressor), missing values

Feature importances

Selecting features

- Mean
- Variance
- Min-Max

MAE, 120 window length, 14615 predictions

	From	To
Mean	0.9273	0.8999
Variance	0.9273	0.9076
Min-Max	0.9273	0.9147

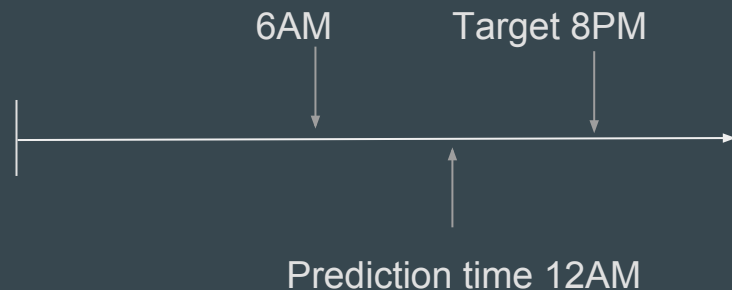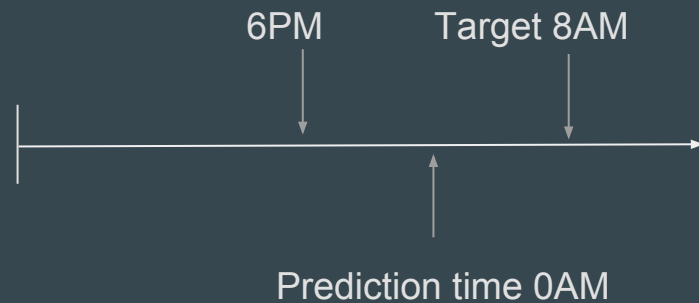12PM   Target 6AM

Prediction time 0AM

Selecting features

- Aladin error from prediction time
- Temperature from 6am or 6pm
- One hour lag for predicted temperature

120 window length, MAE, =~14615 predictions
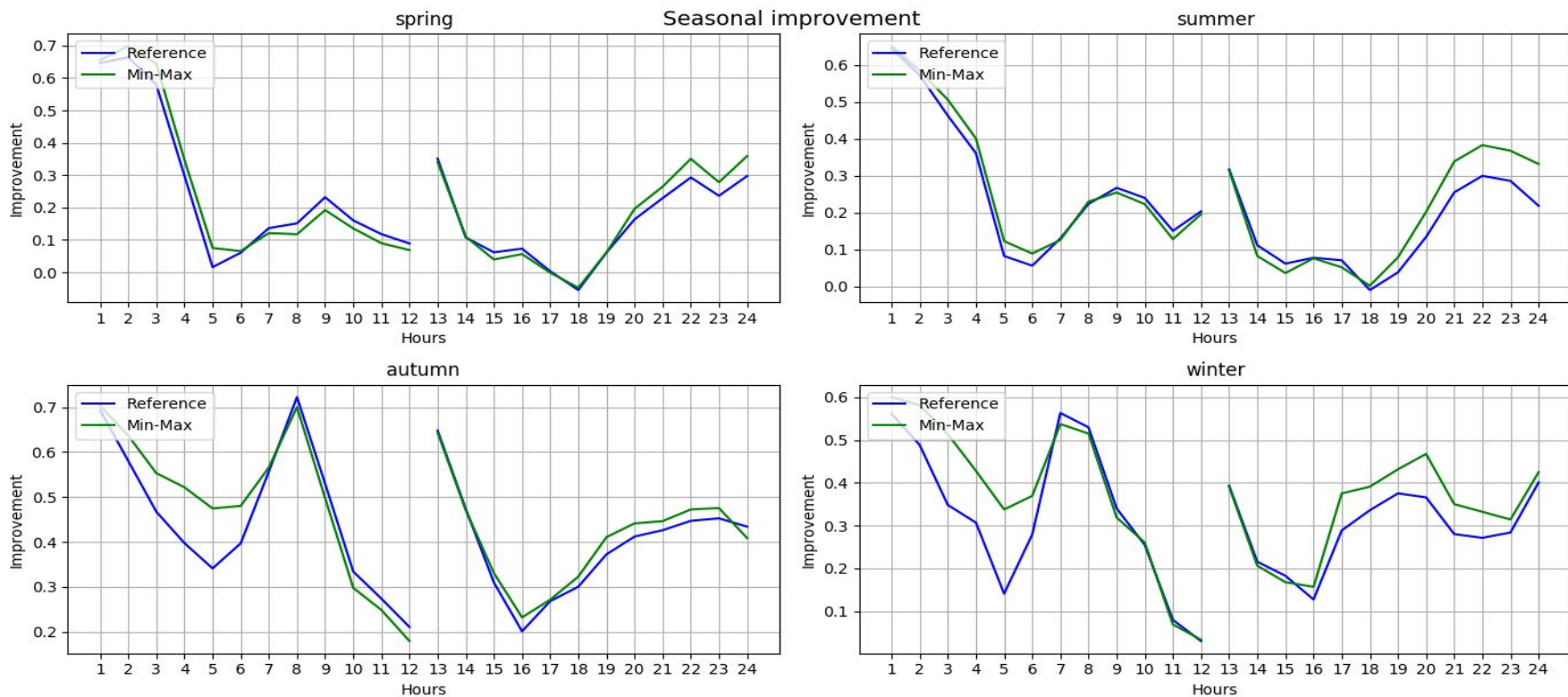
	From	To
Aladin err	0.9273	0.9151
6am-6pm	0.9273	0.9043
Temp. lagged	0.9270	0.9087

6PM   Target 8AM

Prediction time 0AM

6AM   Target 8PM

Prediction time 12AM

Improvement for min-max

Final ensemble model

$$y_i^h = \theta_0 + \theta_1 T_i + \theta_2 H_i + \theta_3 P_i^h + \theta_4 \min(T_{i-12},...,T_i) + \theta_5 \max(T_{i-12},...,T_i) + \theta_6 P_i^{h-1}$$

$$y_i^h = \theta_0 + \theta_1 T_i + \theta_2 H_i + \theta_3 P_i^h + \theta_4 \text{mean}(T_{i-12},...,T_i) + \theta_5 P_i^{h-1}$$

$$y_i^h = \theta_0 + \theta_1 T_i + \theta_2 H_i + \theta_3 P_i^h + \theta_4 T_{i-6} + \theta_5 P_i^{h-1}$$

$h \in 1,...,12$

H_i humidity at time i

T_i measured temperature at time i

P_i^h temperature predicted by Aladin for h hours ahead from time i

Rolling window length: 120

Weighting function: $w(i) = \sqrt{0.97}^i$

MAE improvement: from 0.9273 to 0.8811

Other learning algorithms

- Regularized regression (Ridge, Lasso, Elastic net)
- SVR
- MLP and recurrent networks (GRU, RNN + ReLU)
  - 2000 predictions, slow
 - 30 units, RMSprop optimizer, default Keras settings, reference model features

	OLS	GRU	Combined
MAE	0.8642	0.8733	0.8513
MSE	1.3578	1.3887	1.3361

Detecting stable weather

6 hours   6 hours

12PM   18PM   12PM   Target hour 18PM

- Difference between predicted temperatures must not be greater than 1, for any of the six temperature pairs

Detecting stable weather

Stable weather analysis from sample 16048
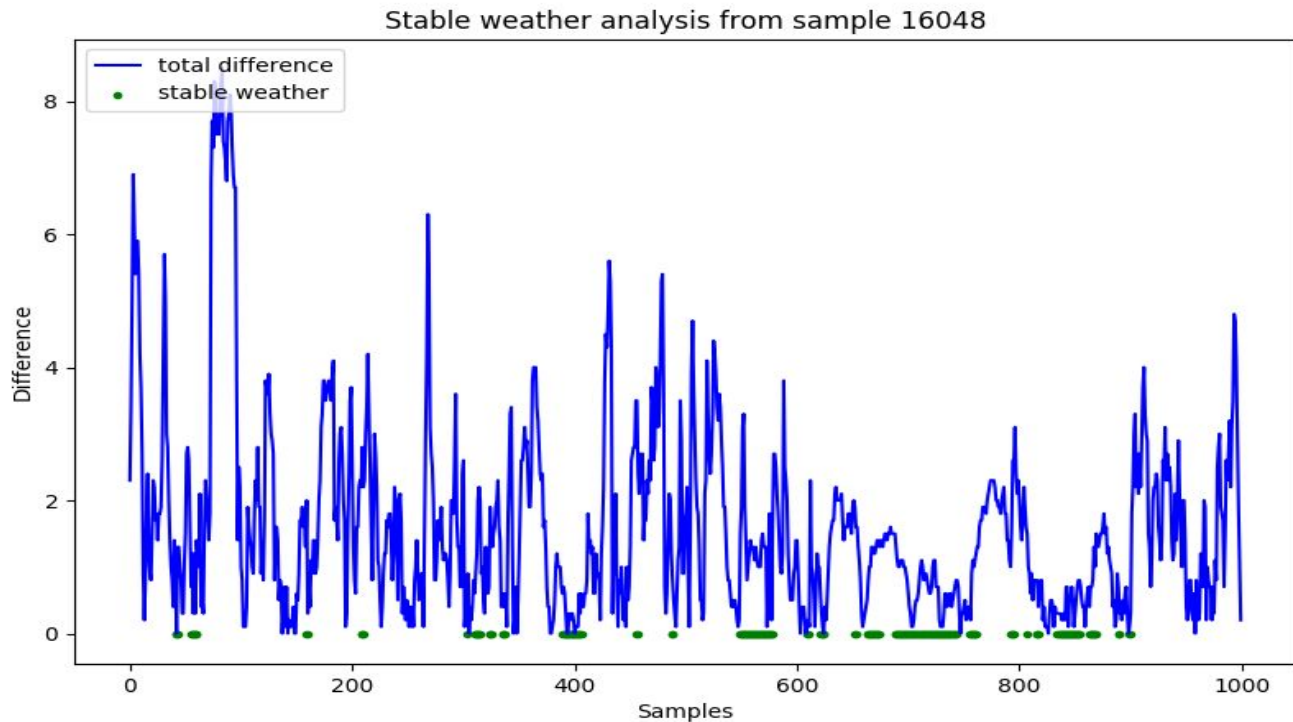
Stable weather models

- Include model error from 24/48 hours before
- Include mean of past model errors
- Improvement for simpler models
- Minor/non-significant improvement for models with more features

$$y_i^h = \theta_0 + \theta_1 T_i + \theta_2 P_i^h + \theta_3(\text{predicted}^h_{i-24} - y^h_{i-24}) + \theta_4 \text{mean}((\text{predicted}^h_{i-24} - y^h_{i-24}),...,(\text{predicted}^h_{i-12} - y^h_{i-12}))$$

Model	reference	Err 24	Err 24 + mean 24
MAE	0.8653	0.8451	0.8434
MSE	1.2766	1.2139	1.1956

Results for all stations (Final vs Aladin)

- MAE improvement from **18.95%** to **73.61%** (Hajdin 13% to 57%)

Percentage improvement for MAE
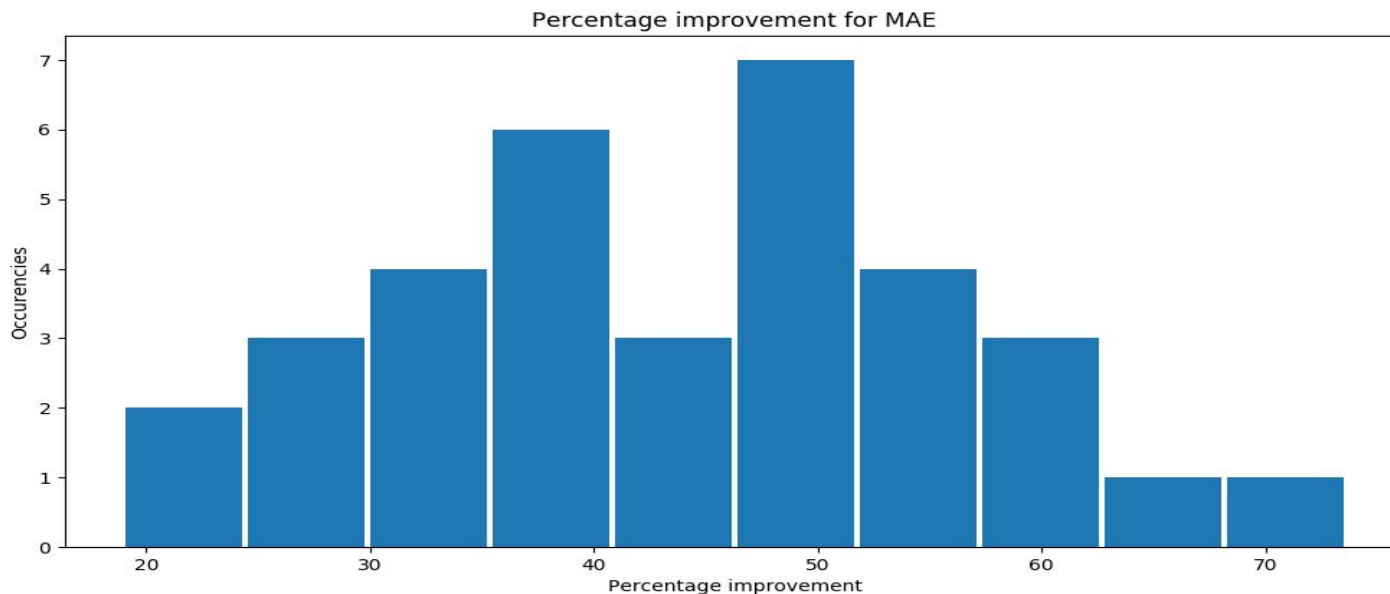
Results for all stations (Final vs Aladin)

- MSE improvement from 27.60% to 89.44%

Percentage improvement for MSE
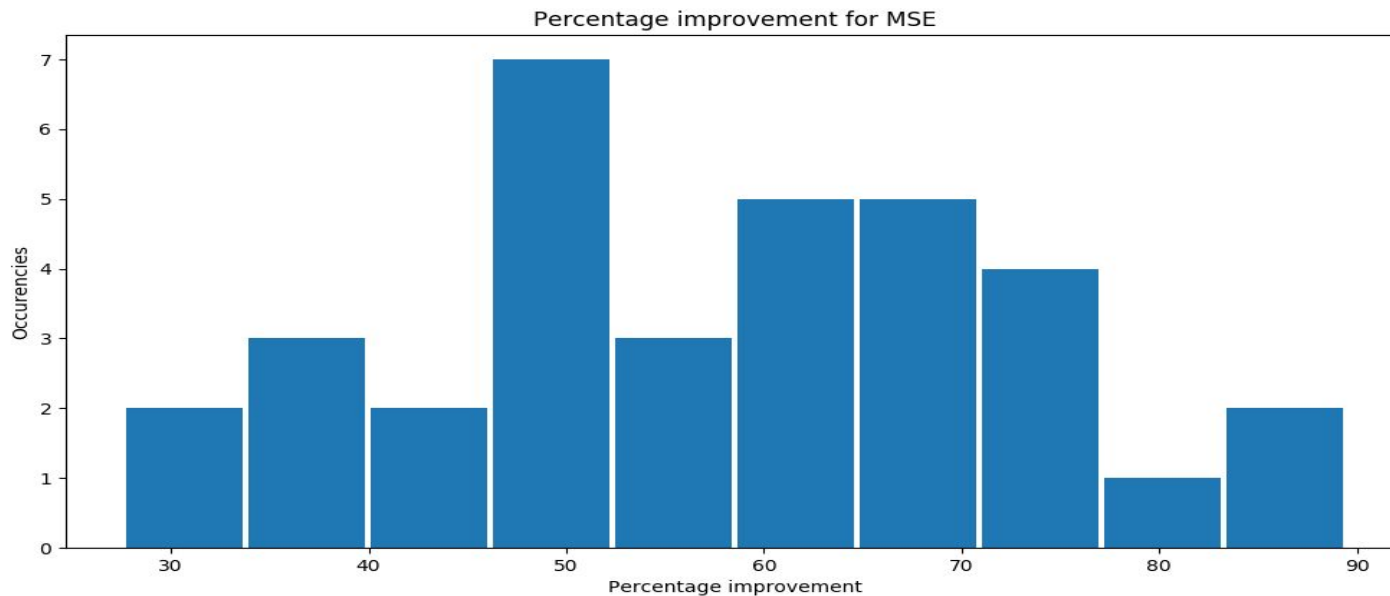
Results for all stations (Final vs Reference)

- MAE improvement from 2.15% to 5.65%

Percentage improvement for MAE
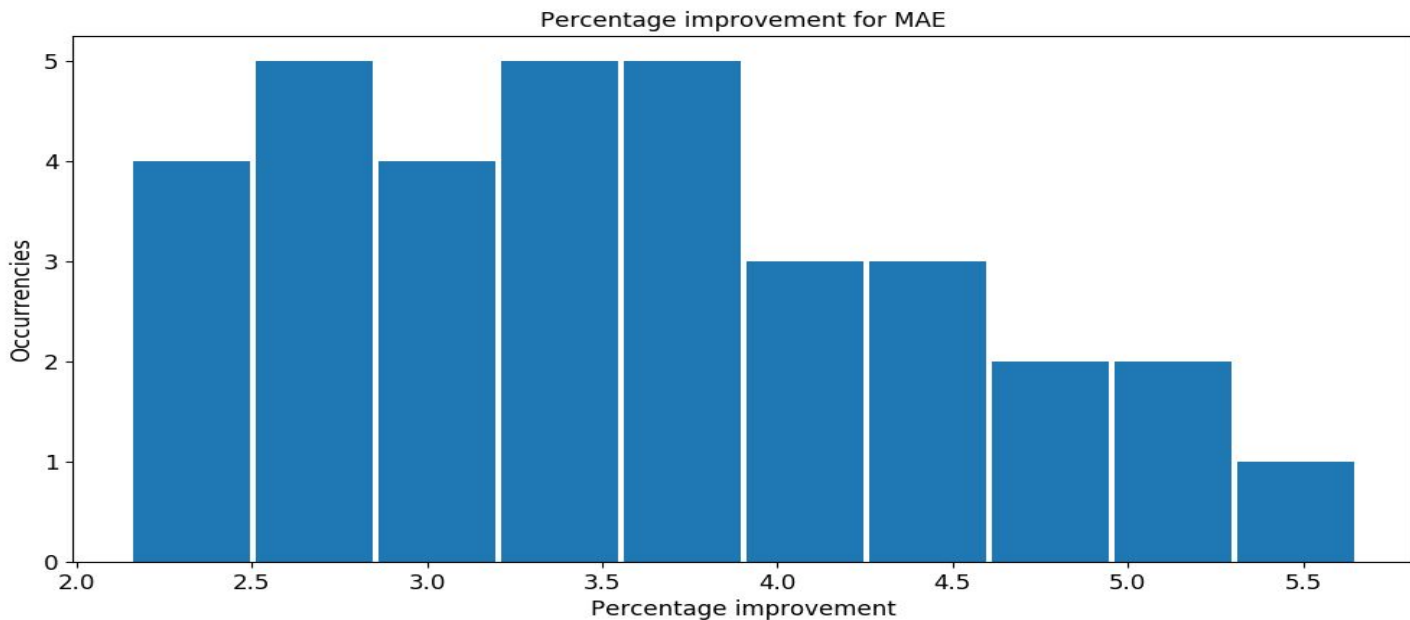
Results for all stations (Final vs Reference)

- MSE improvement from 1.67% to 9.68%

Percentage improvement for MSE
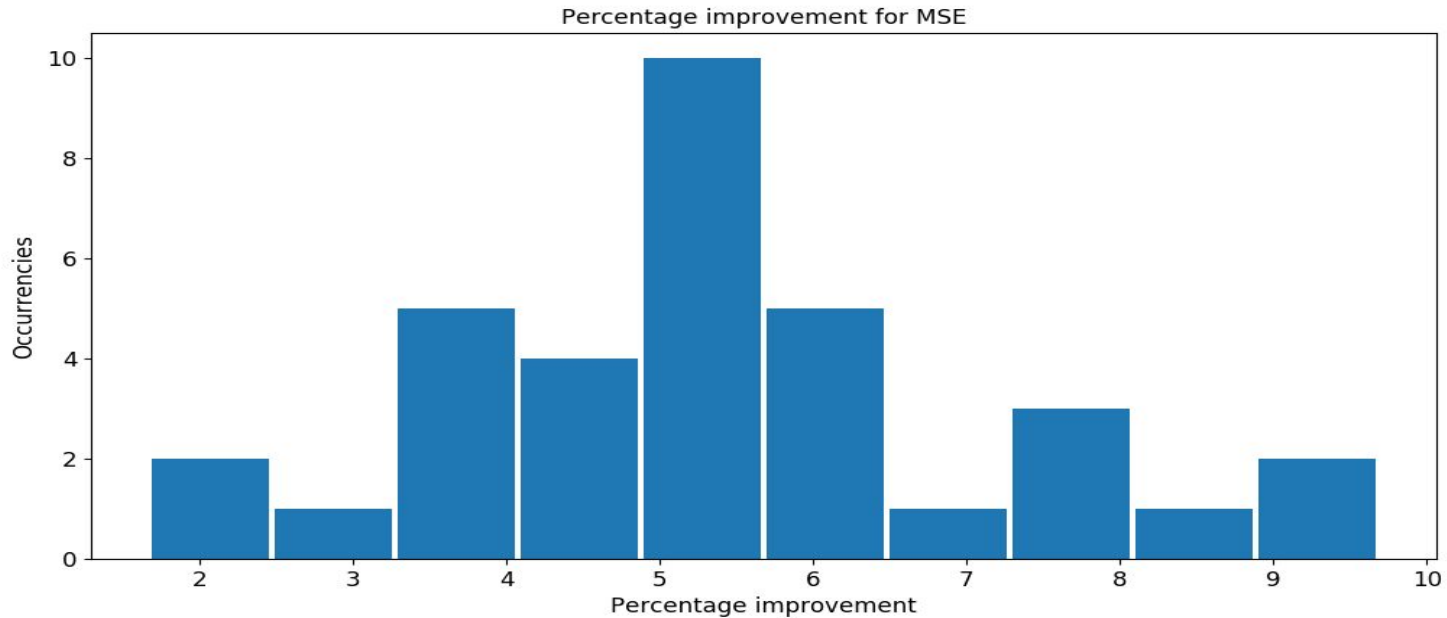
Results for all stations

- Model error from 24 hours before did not cause significant improvement
  - MAE: 19 better vs 10 worse
  - MSE: 15 better vs 13 worse
- Greatest MAE improvement
  - *Lomnický štít* from 4.1°C to 1.1°C
  - *Chopok* from 3.04°C to 0.97°C

Thank you for your attention!

- Meteorological interpretation of features
  - Different improvement across stations according to improvement graphs
  - Appropriate additional information
- Greatest improvement for *Chopok* and *Lomnický štít*
  - Aladin has altitude modeling problems