ALADIN in Poland

Bogdan Bochenek, Piotr Sekuła, Marcin Kolonko
Małgorzata Szczęch-Gajewska, Gabriel Stachura

Institute of Meteorology and Water Management, Poland

Operational Forecasts

ALARO-v1B NH (CY43T2)

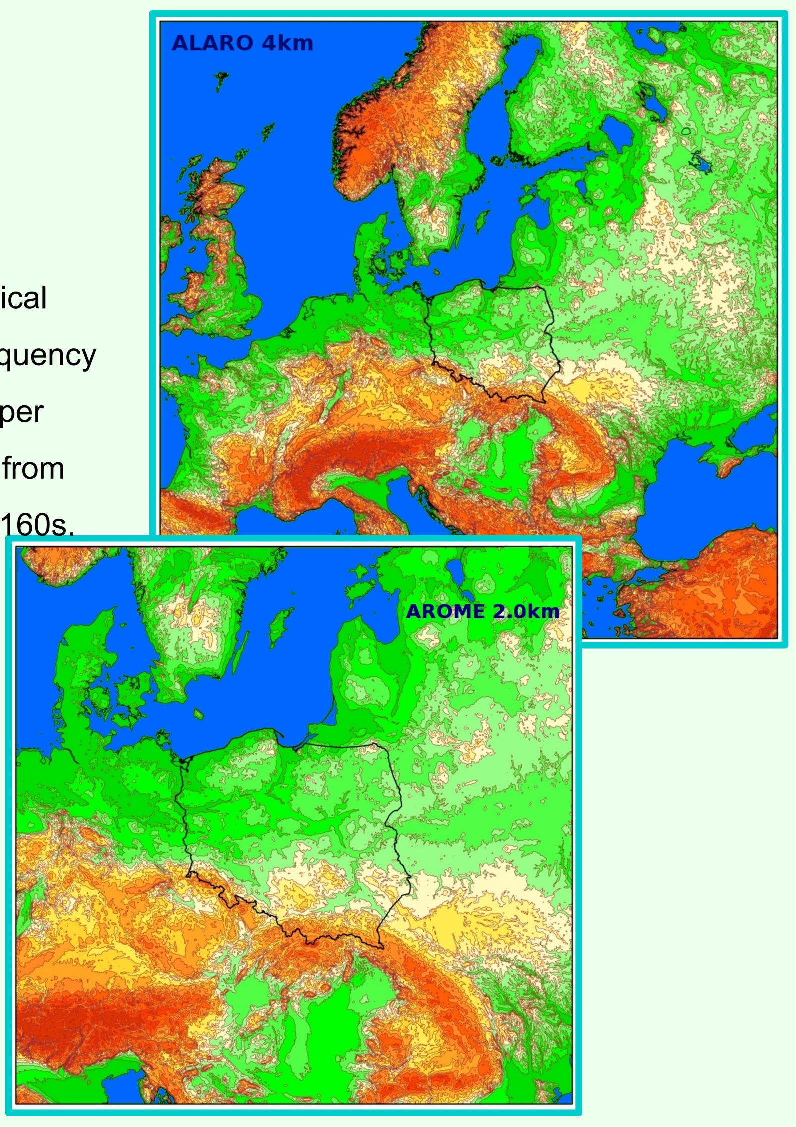
E040 domain:

4.0 km horizontal resolution, 789x789 grid points, 70 vertical model levels on a Lambert projection with 3h coupling frequency and 1h output, coupling zone with 8 points; Runs 4 times per day (00,06,12 and 18) with 72 hours forecast range; LBC from ARPEGE with 9.4 km horizontal resolution; Time step ~160s.

AROME (CY43T2)

P020 domain:

2.0km horizontal resolution, 799x799 grid points, 70 vertical model levels on a Lambert projection with 1h coupling frequency and 1 hour output. 4 runs per day (00,06,12 and 18) with 30 hours forecast range; LBC from ALARO 4.0 km; Time step ~50s.



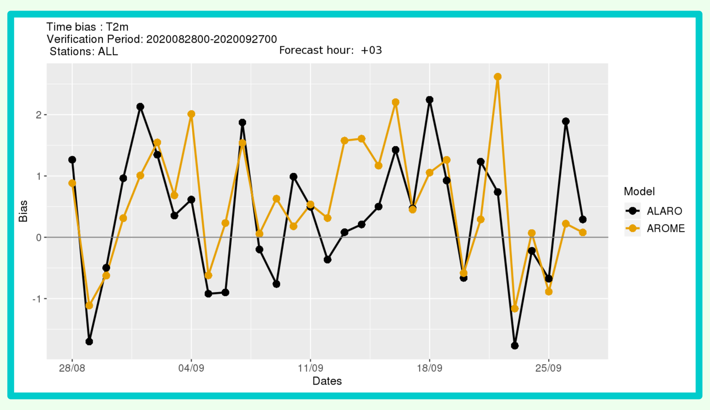
Operational machine characteristics

Cluster of HP BL460c_GEN8 servers connected with Infiniband network, OS Scientific Linux 6, Intel Xeon E5-2690 processors – with maximum 1552 cores (97 nodes with 16 cores each), each core RAM 128 GB, disc array – 64 TB.

Operational Verification

Point verification (SYNOP obs data) carried out for r00 and r12 runs of operational ALARO model. Continuous meteorological elements errors there are presented as mean error (bias) and root mean square error (RMSE) of model for last 30days and last 3 forecasts. Contingency tables are carried out for precipitation accumulated in 12 and 24 hours (7 categories in table) and cloud cover (5 categories in table),

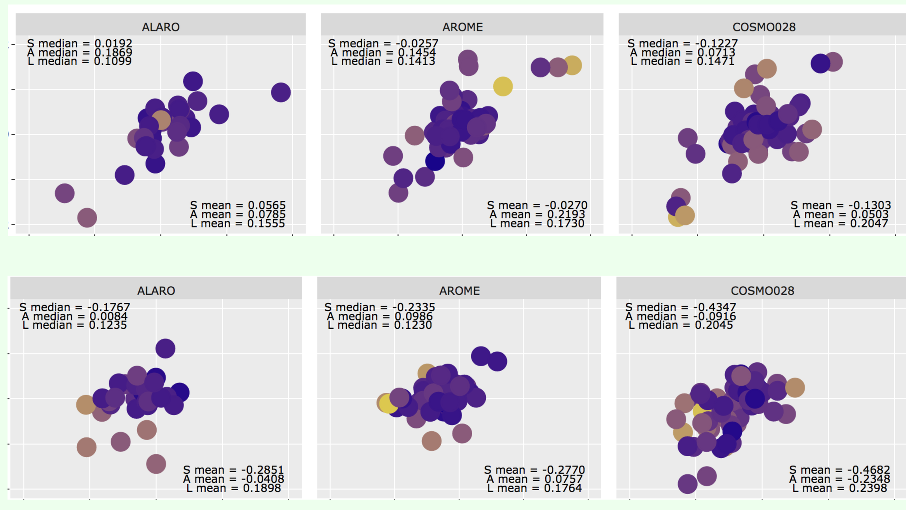
HARP system piont verification (SYNOP and automatic stations data) carried out for all runs of ALARO and AROME operational, for periods of 3, 7, 30 and 90 days.



HARP Spatial

SAL MEASURE

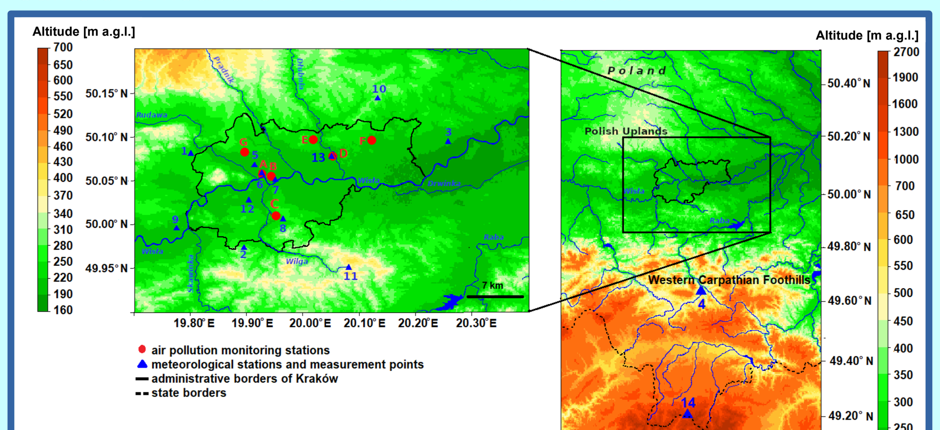
SAL verification is running operationally for ALARO, AROME and COSMO models used operationally and in research mode in IMGW. We investigated results for May and June 2020. In May there was more stratiform precipitation, while in June there was more convection events. For May (top row), scores are similar for ALARO (left column), AROME (centre column) and COSMO2.8 (right column), while for June (bottom row) ALADIN CMC's are slightly better.



Impact of Foehn Wind on PM10 Concentrations and Urban Boundary Layer Structure in the Complex Terrain

(Kraków, as an example)

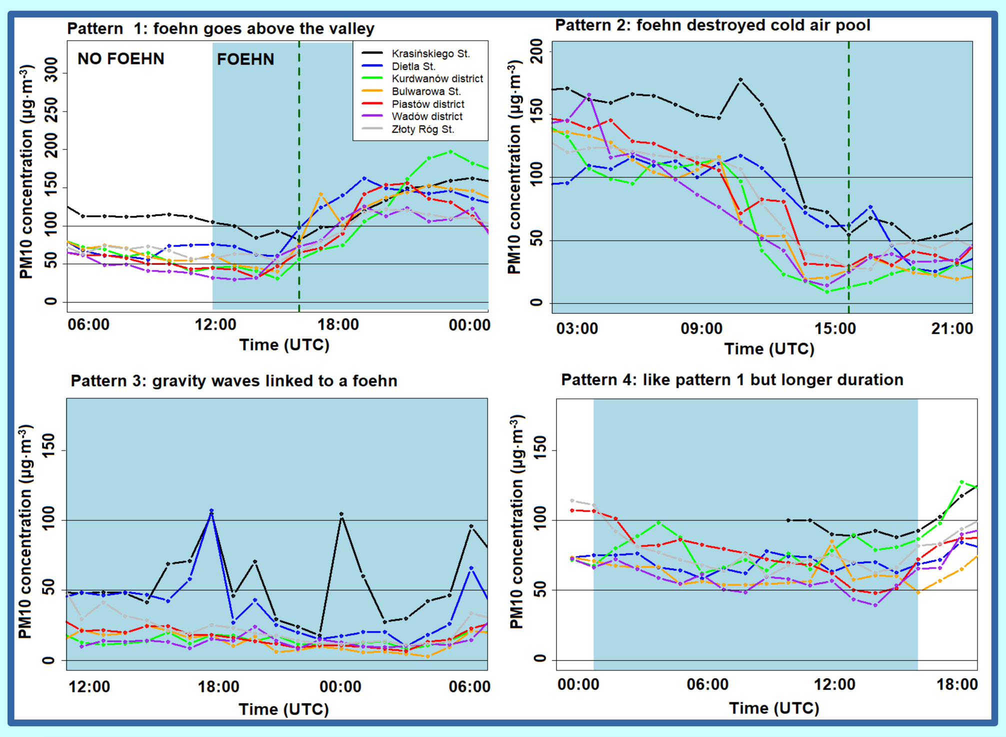
The present work aims at better understanding of foehn wind's impact on urban air quality, due to modification of air pollution dispersion conditions in cities located in large valleys. Kraków, southern Poland, is an example of a city with poor air quality, mainly due to abundant PM10 concentrations, located in a large valley of the Wisła River, within the area affected by a foehn wind 'halny' from the Carpathian Mts. There were 14 long (i.e. >24h) episodes of 'halny' analyzed (40 days, 591 hours in total), from the periods Sep. 2017-Apr. 2018 and Sep. 2018-Apr. 2019, for which sufficient measurement data were available.



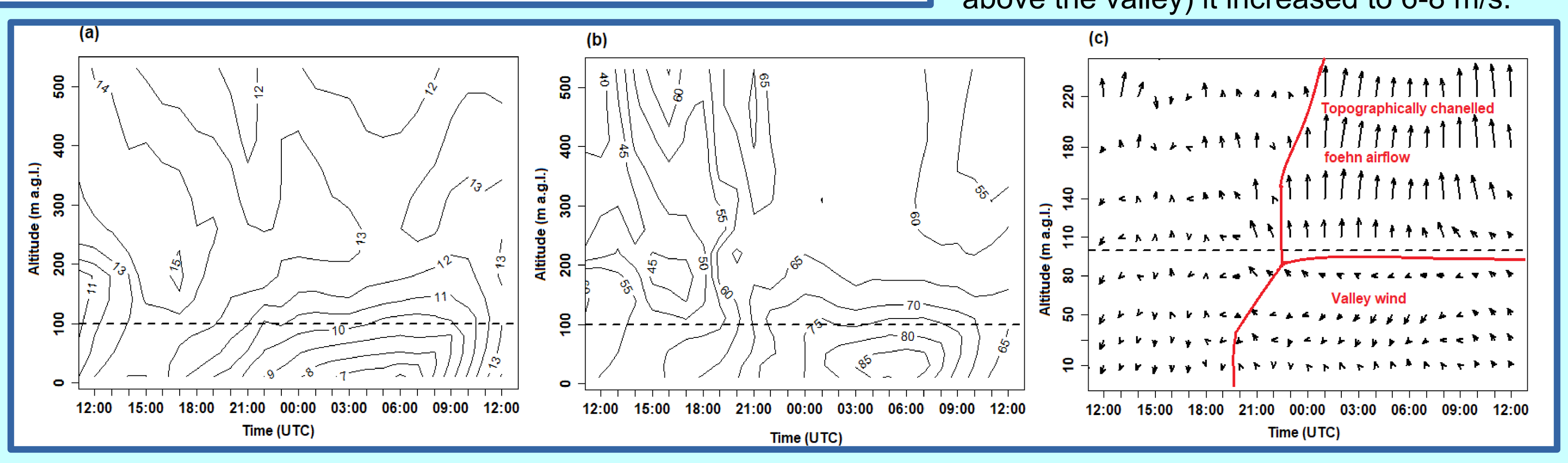
Topographic map of analyzed area and location of meteorological and air pollution stations.

Numerical forecasts of model AROME cy40t1r1 (87 vert. levels, res. 1km x 1km) were used to analyse meteorological conditions and dynamics of atmosphere at lowest part of troposphere.

Four spatial-temporal patterns of PM10 concentrations, were defined: (1) sudden, large increase of PM10 concentrations at all measurement points; (2) sudden, large decrease of PM10 concentrations at all measurement points; (3) short-term peaks or long-term periods of high PM10 concentrations in the western part of the city only; (4) long-term periods of high PM10 concentrations at all measurement points.



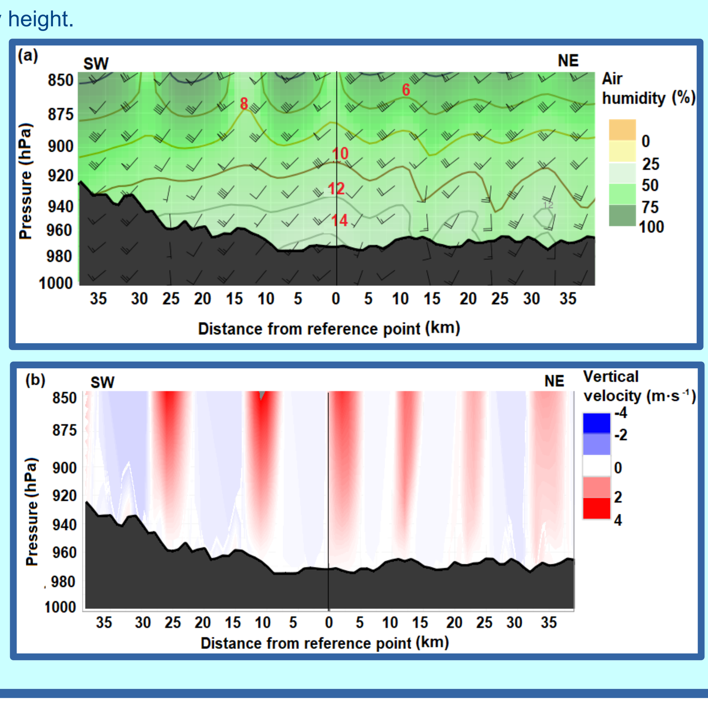
Pattern 1: Sudden, large increase of PM10 concentrations at all measurement points. PM10 increase was linked to the development of air temperature inversion in the Wisła River valley. At the same time, relative humidity reached 80-90% at the valley floor in rural areas which indicates fog occurrence, as the values were decreasing quickly with height. Wind speed inside the valley was below 3 m/s while at about 100 m higher (i.e. above the valley) it increased to 6-8 m/s.



Vertical profiles of: **a.** air temperature (°C); **b.** relative humidity (%); **c.** wind components, for the location of PM10 monitoring station in Krasińskiego St. (point B at topographic map), from AROME model, on 12-13.11.2018. Dashed line: Wisla River valley height.

Pattern 3: Short-term peaks or long-term periods of high PM10 concentrations in the western part of the city only

The differences in relief between the western and eastern part of the Wisła River valley in Kraków contribute to the occurrence of pattern 3. In all cases of pattern 3, much larger PM10 concentrations were observed in the western than in the eastern part of the valley. They were either long-lasting or occurred as concentration peaks. Such conditions were the effect the gravity waves impact. The gravity waves are one of foehn wind effects and can generate closed eddies in a valley which in turn leads to the accumulation of air pollution locally.



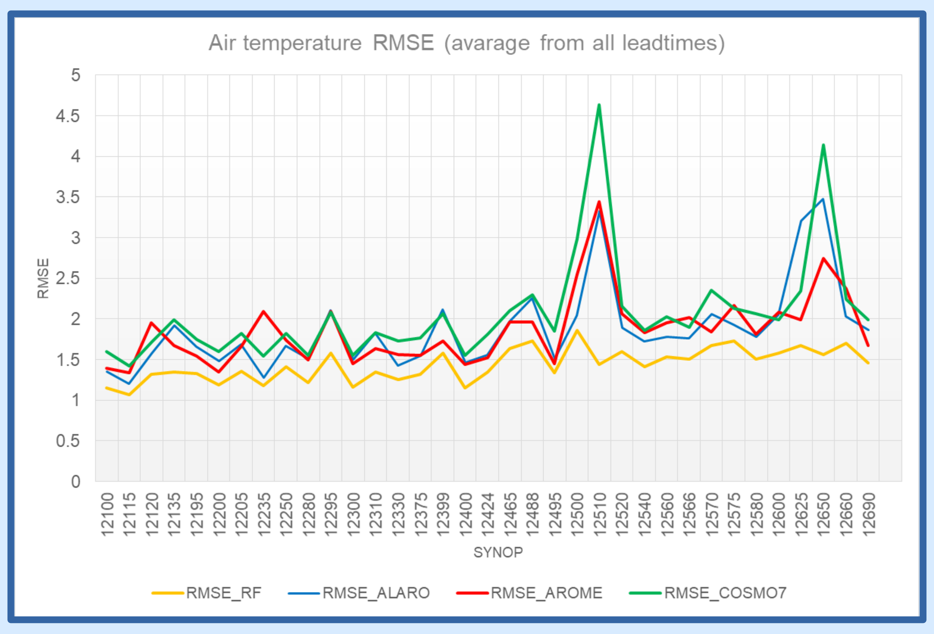
Spatial pattern of: **a.** air temperature (contour lines), air humidity (background) and wind speed (in knots) and direction (graphical symbols) in the SW-NE cross section through Kraków and its vicinities, and **b.** vertical velocity, at 7 UTC on 12.12.2017

Improving Air Temperature Forecast at Post-processing Stage - Machine Learning Methods

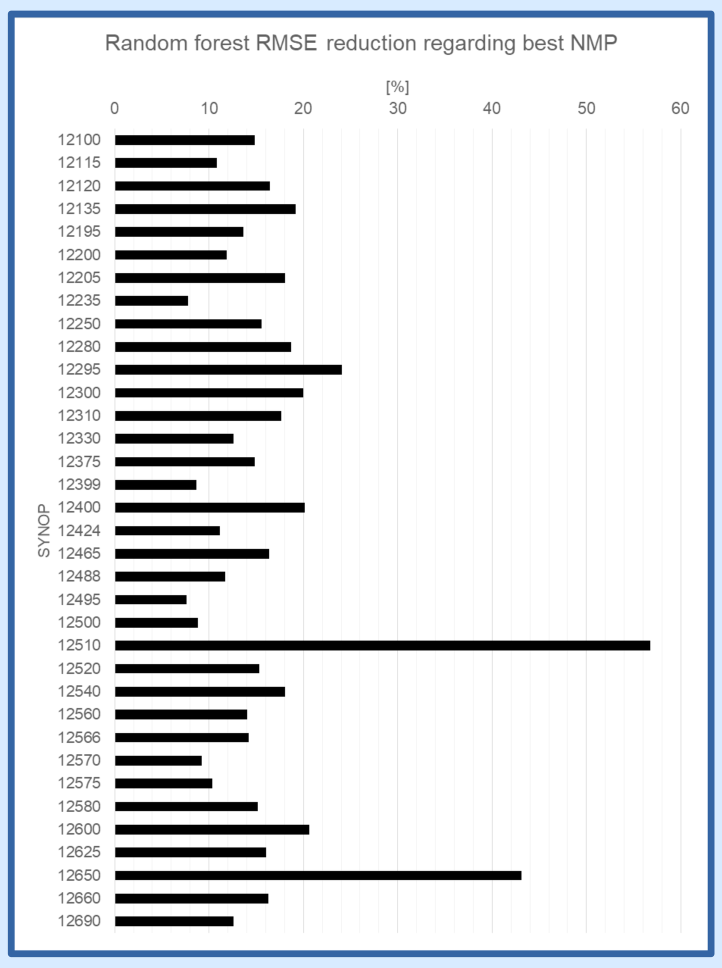
In the experiment forecasted 2m temperatures from three operational models (ALARO 4km, AROME 2km and COSMO 7km) were compared with observed 2m temperature SYNOP values. A machine learning algorithm *Random Forest Approach* was applied (package *randomForest* in R), which uses a collection of decision trees (random forest) with increased performance and can use both classification and regression techniques depending upon the user and target or categories needed. In our case each decision tree, based on the respective predictor variables, was trained on different model and finally Random Forest took the average of the results from all the decision trees.

For building a random forest all models datasets were used with the same specification:

- sampling frequency:3h with leadtime: 3-30h;
- predictors comprise also month, day of the year, hour and leadtime;
- training was conducted on the two years period (2018-2019), which gave around 6800 cases,



Averaged for each synop station (on all leadtimes and test days) 2m temperature RMSE of ALARO, AROME, COSMO forecasts and random forest (RF).

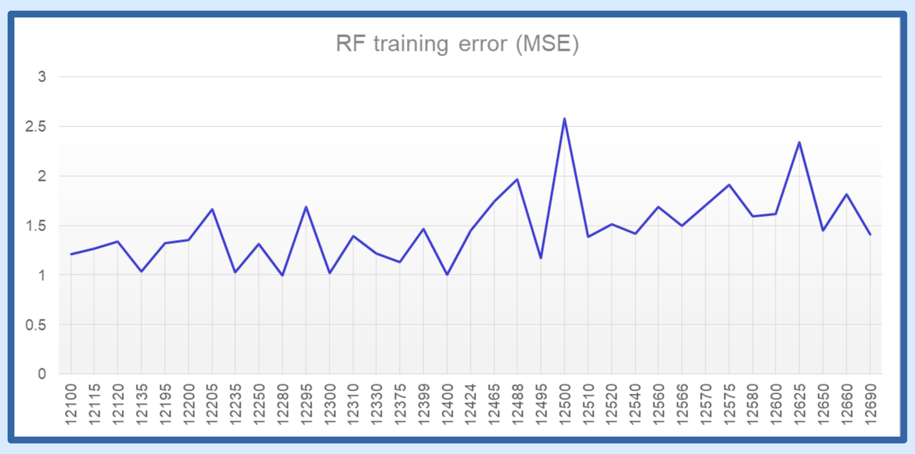


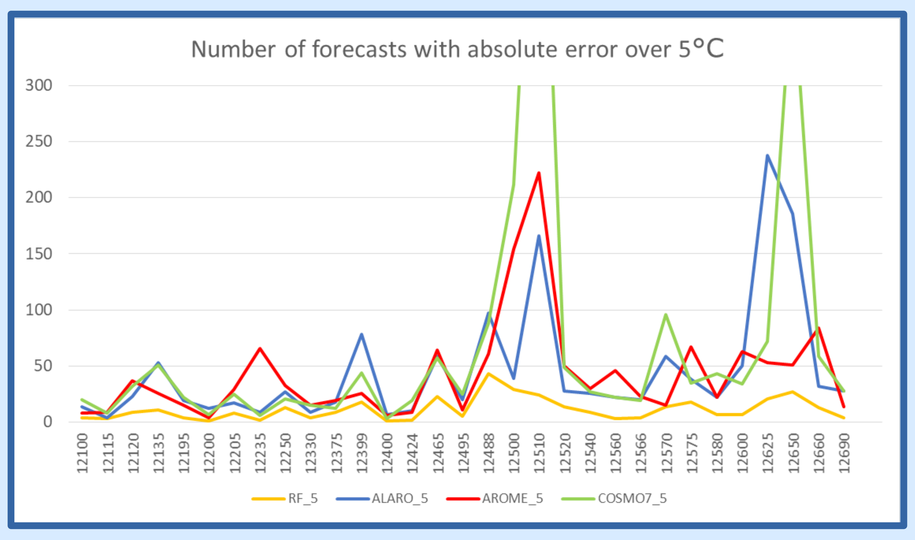
Comparison of random forest RMSE with respect to best performing model RMSE for certain station.

Built random forest was tested on the period from 2020 (until August 13th) and compared with forecasta of each of previously used models. Results were verified by calculating RMSE for 35 selected SYNOP stations from different regions in Poland (lowlands, uplands and mountains) . On figures they are named by its numbers.

Conclusions:

- RMSE reduction occurred at every station, with an avarage of 16%;
- The biggest improvement occured for mountain-top stations (12510,12650), probably because models were strongly biased there;
- Training error (on the right) is slighty higher in highlands (from 12500 on) than in lowlands;
- The biggest training errors are noticeable for stations 12500 and 12625, which lays in a bottoms of valleys at the mountains;
- Some predictors with information about topography and cloudiness//insolation should be added to training data;
- Although random forest performs better than the best of three models only in 30-40% of forecasts, its impact is the most visible when considering big errors (over 5°C) – e.g. for Cracow (12566) random forest made such an error only 4 times, while COSMO 20 times, ALARO – 21 and AROME – 23.





Number of forecasts of each model and random forest with absolute error value bigger than 5°C. 12510 and 12650 are mountain-top stations.