Applications and Verification

Report

Prepared by:	Area Leader Doina-Simona Tașcu	
Period:	2022	
Date:	March 2023	

Introduction and background

The main activities within the “Applications and Verification” area are to collect information and identify LACE members needs and create an action plan to systematically manage, organize and control the changes in the software codes. Furthermore, to maintain a documentary fund which should help the NWP colleagues.

Goals

The primary goal is to develop and adapt/use different specific applications into user-friendly mode. Many tools and software products were developed along the years for meteorological parameters. These days, it is imperative to have easy to use applications, maybe to find and to identify the operational activities, to make a common way for saving time and manpower resources. Nowadays, it is important to make the applications easy to implement without too much cost and to make a common way for saving time, computer costs and manpower. It is a big challenge to identify and to merge all the beneficial technical approaches and applications for all countries.

Main R&D activities

It is well known that the verification of numerical model is an important tool in order to ensure a better use of the meteorological products for operational and research usage. The main activities are the following:

- HARP implementation and verification for deterministic and probabilistic forecasts
- HARP linked to OPLACE database
- Multiple verification methods
- Post-processing of model output
- RC-LACE github platform
- Database of cases
- Trainings

1 Action: HARP implementation and verification for deterministic and probabilistic forecasts

In **Austria**, the work related HARP verification package was continued by extending the usage of HARP by colleagues from other departament. For this purpose, wrapper scripts for user friendly HARP and reading grib files were developed. Another important topic for the austrian team is the participation of the Florian Weidle at “Harp – harp standard verifications set”, 07-10.06.2022, Helsinki, 2022.

Also in Austria, the work related HARP verification package was continued by extending the spatial verification of simulated infrared channels from AROME-Aut for multiple thresholds. For this purpose, the observations from MSG2 - SEVIRI satellite channel Infrared 10.8 µm (brightness temperature) were used. The model is AROME ensuite and the used domain is covering Austria. The computed score is Fraction Skill Score (FSS) and the window sizes are: 1, 3, 5, 11, 21 for the thresholds:

- 273 K (0 -degree Celsius)
- 260 K (supercooled water droplets)
- 250 K (convective initiation)
- 240 / 230 K (overshooting tops)

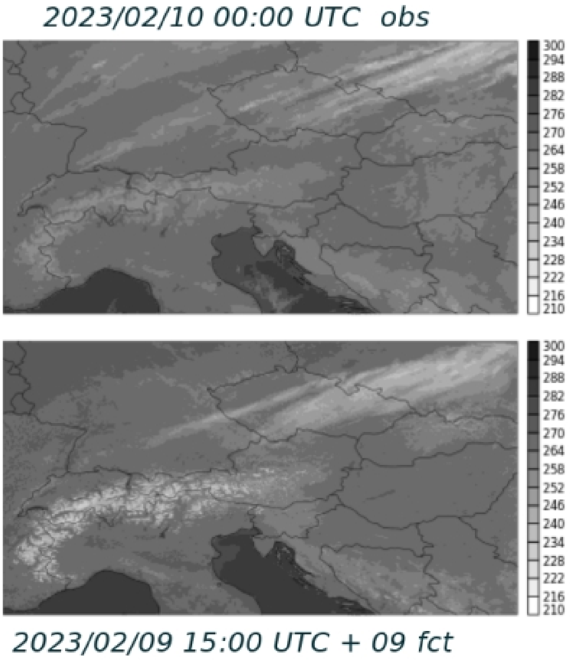
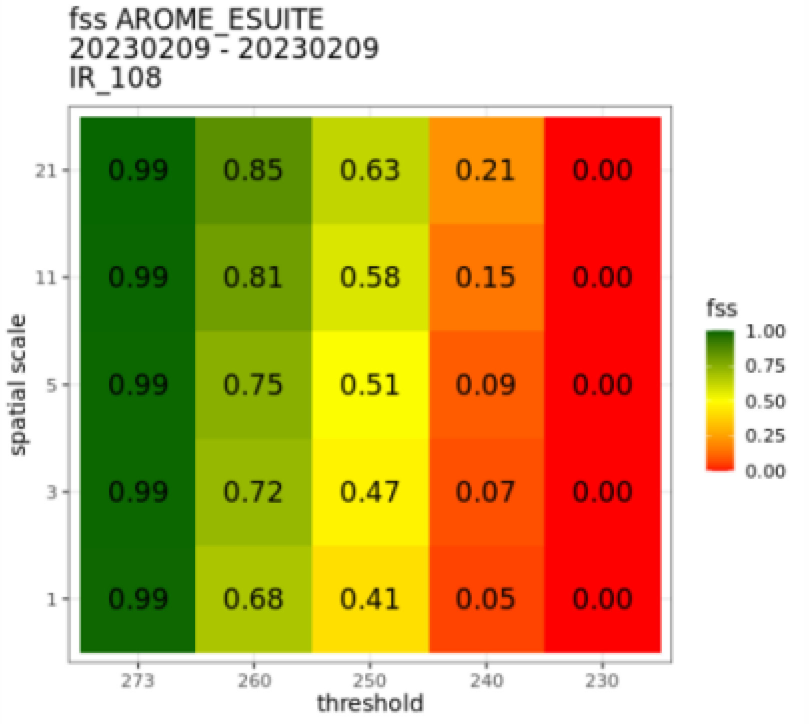


Figure 1: Example of verification of simulated IR channels in HARP, 20230210, 00 UTC. Observation *2023/02/10 00:00 UTC and forecast 2023/02/09 15:00 UTC + 09.*



spatial scale	273	260	250	240	230
21	0.99	0.85	0.63	0.21	0.00
11	0.99	0.81	0.58	0.15	0.00
5	0.99	0.75	0.51	0.09	0.00
3	0.99	0.72	0.47	0.07	0.00
1	0.99	0.68	0.41	0.05	0.00

Figure 2: Example of the FSS score computed for AROME-ESUITE, 20230209, IR_108

Contributors: Polly Schmederer, Christoph Zingerle (1pm - MQA 1.1)

Another topic is generating Fact Sheets based on model verification with HARP using Markdown for customers. To provide customers with easy accessible and easy to understand evaluations of model performance at GeoSphere Austria, fact sheets tailored to the needs of specific customers are created with HARP and Markdown. The fact sheets include short documentation about the NWP-models presented and the methods and interpretation of the demonstrated scores.

Informationsblatt Verifikation und Performance der Wettermodelle

Geosphere Austria

2023-01-18

Produktinformation

Für die Vorhersagen wird das von der ZAMG entwickelte Analyse- und Nowcastingsystem **INCA (Integrated Nowcasting through Comprehensive Analysis)** verwendet. Das Modell ermöglicht zeitlich und räumlich hochauflösende Analyse und Vorhersage für die nächsten Stunden unter besonderer Berücksichigtung regionaler und kleinräumiger topographischer Effekte. Es liefert auf einem 1-km Raster stündlich aktualisierte Prognosen von Temperatur, Luftfeuchte, Wind, Niederschlag. Das Ziel von INCA ist eine zeitlich und räumlich hochauflösende Analyse und Prognose des aktuellen atmosphärischen Zustandes im Nowcasting- und Kurzfristbereich. Nach 48 Vorhersagestunden geht die Prognose von INCA in die Prognose des Modells des **ECMWF (European Center for Medium-Range Weather Forecasts)** über. Das ECMWF ist sowohl ein Forschungsinstitut als auch Dienst, der u.a. globale numerische Wettervorhersagen erstellt und anbietet. Für die Berechnung der Prognosen wird das am 0.25 Grad Gitter vorliegende Modell auf die INCA Auflösung von 1 km verfeinert.

In diesem Bericht werden die Vorhersagen der beiden Modelle gegenübergestellt. Allgemein ist bei den berechneten Verifikationsmaßen zu beachten, dass es sich um eine Validierung an einem Punkt oder Ort handelt. Wird ein Ereignis für eine Statione vorhergesagt, aber an einer benachbarten Station beobachtet, wird die Vorhersage bestraft. Auch wenn die Distanz zwischen den beiden Punkten nur wenige Kilometer beträgt.

Trefferquote - relative Häufigkeit

Die Tabelle zeigt die relative Häufigkeit in Prozent (`sharpness`), mit der das Ereignis über einen bestimmten Zeitraum vorhergesagt wurde. Das Ereignis wird nicht direkt mit der Beobachtung verifziert, sondern die Differenz zwischen Vorhersage und Beobachtung betrachtet. Anschließend wird die relative Häufigkeit berechnet, mit der diese Differenz von einer vorgegebenen Klasse abweicht. Diese Klasse wird als "Abweichung" in der Tabelle bezeichnet. Je höher der prozentuelle Wert, desto häufiger liegt die Differenz in der Klasse. Eine gute Prognose ist, die höchsten Ergebnisse in der obersten Klasse zu erreichen.

ist die Verteilung der Niederschlagsintensität darzustellen. D.h. es zeigt, welche Niederschlagsbereiche ein Modell gut vorhersagt, oder ob es einen Bereich über- oder unterschätzt. Damit kann eine AUssage gemacht werden, ob das Modell eine nützliche und sinnvolle Vorhersage liefert.

Trefferquote - False Alarm und Hit Rate des Niederschlags

Figure 2: Hit Rate (h) und False-Alarm Rate (f) von Niederschlagsklassen (0-0.1, 0.1-1, 1-5, 5-20 und 20-60 mm/12h). Bild zum Vergrößern anklicken.

In Abbildung 2 ist die Hit Rate (h) und False Alarm Rate (f) verschiedener Niederschlagsklassen dargestellt. Weiters werden die Stationen anhand der Höhenniveaus unterteilt. Unterschieden wird zwischen Stationen im Flachland (flat), im Gebirge (mountain) und in Tälern (Valley). Die Werte geben an, wie häufig die Vorhersage richtig liegt (Hit Rate). Diese ist perfekt bei einem Wert von 1. Die False Alarm Rate gibt an, wie oft eine Vorhersage fälschlicherweise ein Ereignis prognostiziert (False Alarm Rate). Eine Prognose ist perfekt bei einem Wert von 0.

Figure 3: Example of generating Fact Sheets of HARP based verification for customers

Contributors: Christoph Zingerle, Judith Luftinger Svacina (1pm - MQA 1.1)

In **Slovenia**, were developed the procedure and the scripts necessary for automatic objective verification based on HARP. Currently, this still hasn't been put in operations, but it will be in next month. The purpose is the automatic production of weekly and monthly scores, initially for traditional parameters (Figure 4), also for scorecards (Figure 5).

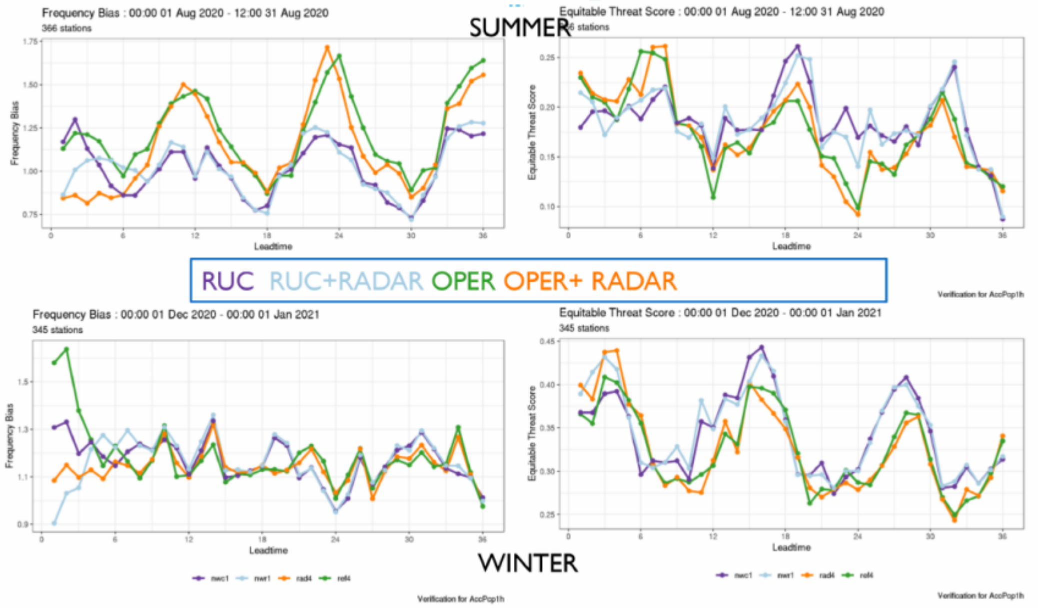


Figure 4. 2mm/h precipitation threshold, performance of RUC and impact of radar data: frequency bias for 01.08.2020 – 31.08.2020 (top left) and for 01.12.2020 – 01.01.2021 (bottom – left), equitable threat score for 01.08.2020 – 31.08.2020 (top right) and for 01.12.2020 – 01.01.2021 (bottom – right).

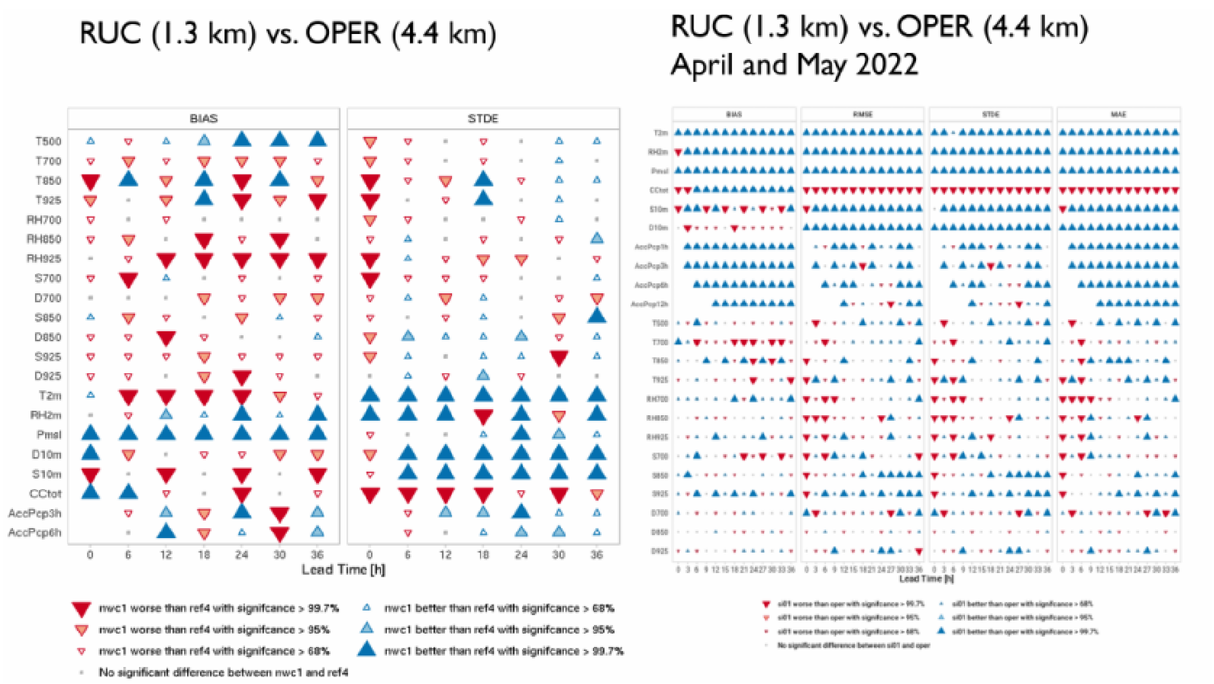


Figure 5. Scorecards for the evaluation of RUC in comparison of the operational version at 4.4 km horizontal resolution, for April and May 2022.

Contributers: Florian Weidle (0.25 pm – MQA1.4, 1 pm – WW), Jure Cedilnik (0.5 pm – MQA3), Neva Pristov (0.5 pm - MQA3), Benedikt Strajnar (0.5 pm - MQA3)

2 Action: HARP linked to OPLACE database

In **Slovakia**, starting from February 2022, Andrew Singleton added Martin Petras's development related to obsoul reading capability to the development version of harpIO:

- https://github.com/andrew-met/harpIO/tree/develop

Another further Martin's progress is related to the following:

- the capability to read from merged files, there is no need to define country anymore
- replace SID with numbers only if they contains string('AT', 'SI'......)
- dropping ships and working only with synop
- adding new parameters, for example RR1h, RR6h, etc

Contributor: Martin Petras (1 pm)

3 Action: Multiple verification methods

Validation of AROME SEKF (Simplified Extended Kalman filter)

In **Hungary**, AROME-TEST has been running in parallel to the operational AROME/HU forecasts since November 2021. Simplified extended Kalman filter (SEKF) was applied in the test run for surface data assimilation, while in the operational AROME/HU model used the optimum interpolation (OI-main) method. (More details about the experimental setup can be found in the data assimilation report.)

After the winter test period, an additional 1-month convective test period (2022. 05. 04 – 2022. 06. 01) was chosen to evaluate and compare the forecasts. Some objective and subjective evaluations were carried out:

- The forecasters were involved in the evaluation focusing on differences between the operational and test forecasts at 0, 6 and 12 UTC. They recorded daily their comments.
- At the same time, subjective verification was done by the model developers in interesting weather situations, ranking the 2-meter temperature, 10-meter wind and gust, precipitation and cloudiness forecasts from 1 to 5 (where 5 is the best).
- The objective pointwise verification was completed by the in-house developed Perl-based OVISYS verification system. RMSE and bias were computed for some surface parameters: mean sea level pressure, total cloudiness, temperature, dewpoint and relative humidity at 2 m, 10 m wind speed and wind gust, and for some upper-level parameters: temperature at 850 hPa, relative humidity at 700 and 925 hPa, geopotential height at 500 hPa. The verification contained frequency bias and some other scores calculated based on the contingency table of 12- and 24-hour precipitation amounts as well.
- In addition, SAL verification (Wernli et al., 2008) was also made to evaluate the location, amplitude, and structure of the precipitation patterns in comparison with radar data.

For several parameters, there was a consistent difference between the operational and the test version, which usually dominated at night (for minimum temperature RMSE and bias improvement of nearly 1 °C occurred). In the case of dewpoint, the operational run had a smaller error during the daytime hours, but the SEKF reduced the underestimation at night (Figure 6). A larger daily temperature cycle was observed in the AROME-TEST run, with more intense daytime warming, and nighttime cooling, which mostly affected the central part of the domain. This sometimes led to an improvement (mostly in clear, dry, anticyclonic weather situations, like on 19 May; Figure 7) and sometimes to deterioration. The larger differences that occur in these areas can be explained by moisture increments of SEKF.

Further differences appeared in intensity and structure of precipitation. During the test period, less than average precipitation fell, which appeared in the form of showers and thunderstorms. In this convective period AROME-TEST predicted precipitation fields with a much finer structure than the operational model and also reduced the underestimation of the number of objects (Figure 8). Regarding the wind forecasts, the models did not accurately reproduce the intensification associated with thunderstorms in time and space. Overall, the positive results in the case of convective precipitation and temperature extremes, furthermore the reduction of night-time inaccuracies, allowed the operational implementation of the SEKF on 29 June 2022.

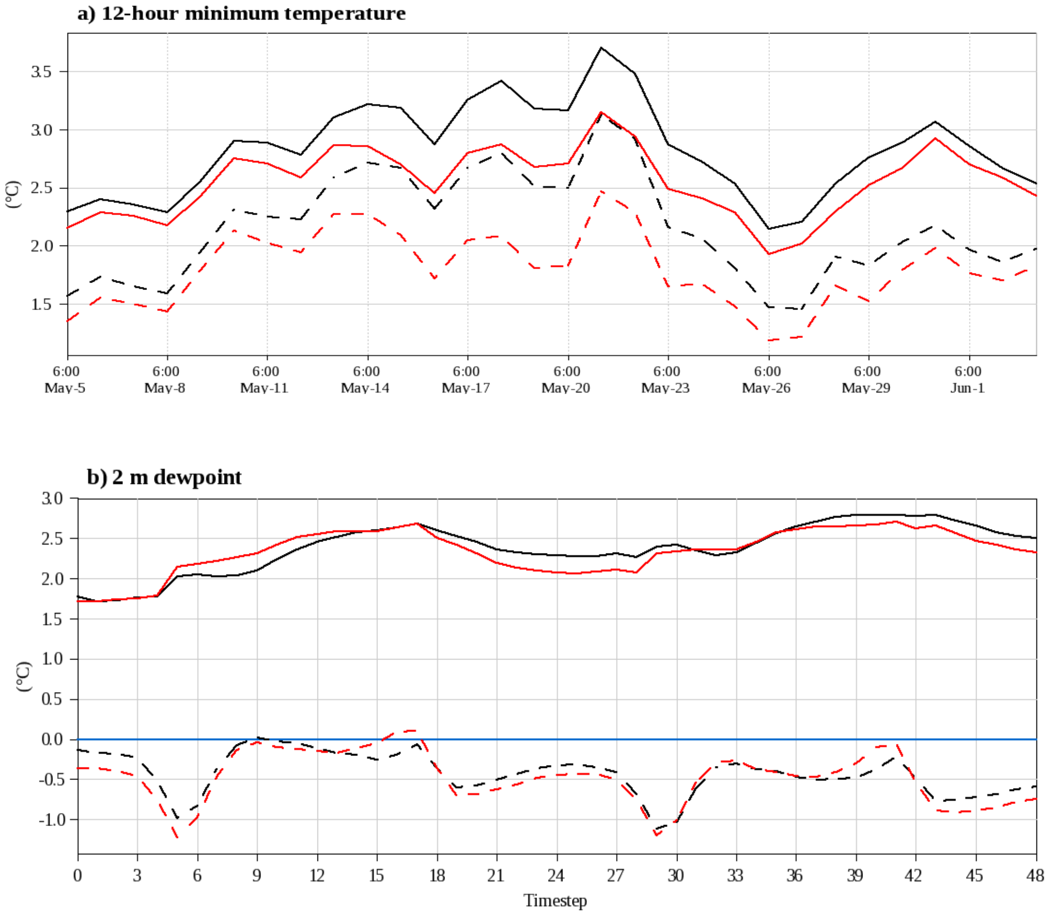


Figure 6: a) 5-day moving average of RMSE (solid lines) and bias (dashed lines) of the 12-hour minimum temperature (in °C) based on AROME (black) and AROME-TEST (red) forecasts for the 00 UTC +30h between 4 May and 1 June 2022. b) RMSE (solid lines) and bias (dashed lines) of 2 meter dewpoint (in °C) based on AROME (black) and AROME-TEST (red) forecasts as function of lead time.

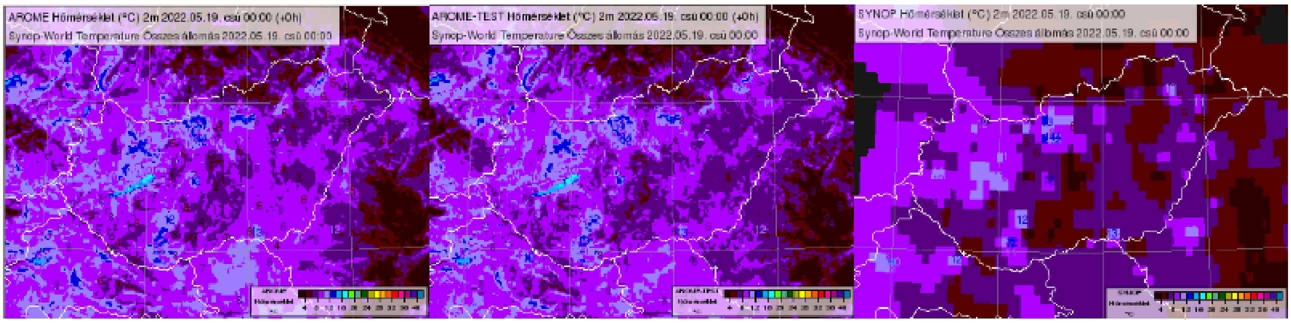


Figure 7: 2 meter temperature (in °C) based on analysis of AROME (left) and AROME-TEST (middle), SYNOP observations (right) at 00 UTC on 19 May 2022.

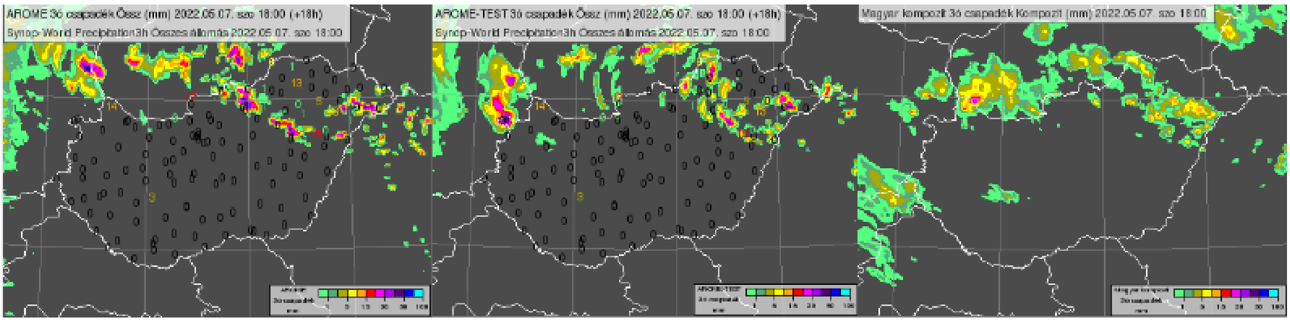


Figure 8: 3-hour precipitation sum (in mm) based on 18-hour forecasts of AROME (left) and AROME-TEST (middle), Hungarian radar measurements (right) and SYNOP observations (marked with numbers in the first two panels) at 18 UTC on 7 May 2022.

Contributor: Boglárka Tóth (1.5 pm)

References

Wernli, H., Paulat, M., Hagen, M. Frei, Ch., 2008: SAL – A novel quality measure for the verification of Quantitative Precipitation Forecast. Mon. Wea. Rev., 136, 4470–4487.

Adaptation of machine learning post-processing method on AROME forecasts

Also In **Hungary**, the main activity was to implement the post-processing methods developed in 2021 into the operational system. OMSZ provides forecasts from numerical weather prediction models to support partners producing renewable energy. With statistical post-processing, errors of global radiation and near-surface wind forecasts can be reduced. As part of a project, mathematician colleagues developed machine learning and EMOS (ensemble model output statistics) methods applicable to AROME and AROME-EPS ensemble forecasts in 2021. The work on applying these procedures in operational practice is ongoing.

The autoencoder method is chosen for improving both the global radiation and the 100m wind AROME forecasts. It is based on neural networks and every geographical point of interest uses its own neural network system. These systems consist of sub neural networks in every lead time between 0 and 36/48 hours, which improves the forecast value in the given timestep.

To re-train the neural networks a new 1-year long data series was used from August 2021 to September 2022. 80% of the period was randomly selected for the training and 20% was used for validation. 3 geographical points were chosen for 100m wind post-processing and 2 points for global radiation. In these locations there are power plants which provide regular measurements for the training. The validation period showed an improvement in RMSE by 17% for radiation forecasts and by 10-20% for 100m wind forecasts.

The operational implementation of the post-processing was tested on the 00 UTC AROME forecast runs. Beside the forecast data, the method uses observations from 12 hours preceding the forecast start. A script, which calls the python program for post processing, automatically starts when the forecast data becomes available. The results containing the improved value for every timestep is stored in a csv text file for each location.

For the next step the evaluation of the obtained results is required, and if it is necessary the modifications in the neural network training configurations will be made. If sufficient data are available it is possible to use a longer training period or expand the post-processing to more locations.

Contributor: Dávid Tajti (3 pm - E6)

Panelification

This topic was started in Austria also for precipitation verification and visualization. The idea was to calculate a number of scores and present them together with the visualized field, such that humans can quickly have a look at the fields and the scores in one single place. Panelification is being extended to accommodate additional variables other than precipitation, namely hail and lightning. These two were selected because there are data sets which can be used for gridded verification and because they are good indicators of strong convection. As such, they can be used to reliably located the position of thunderstorms.

In **Austria**, the code of Panelification was modified to read and process the corresponding model fields as well as the observation fields. In the case of lightning, MODIS-Data for the verification time nad location is read from a local database and converted into a grided field of lightning strikes. These are compared to the lightning density from AROME. For hail, a threshold approach will be used.

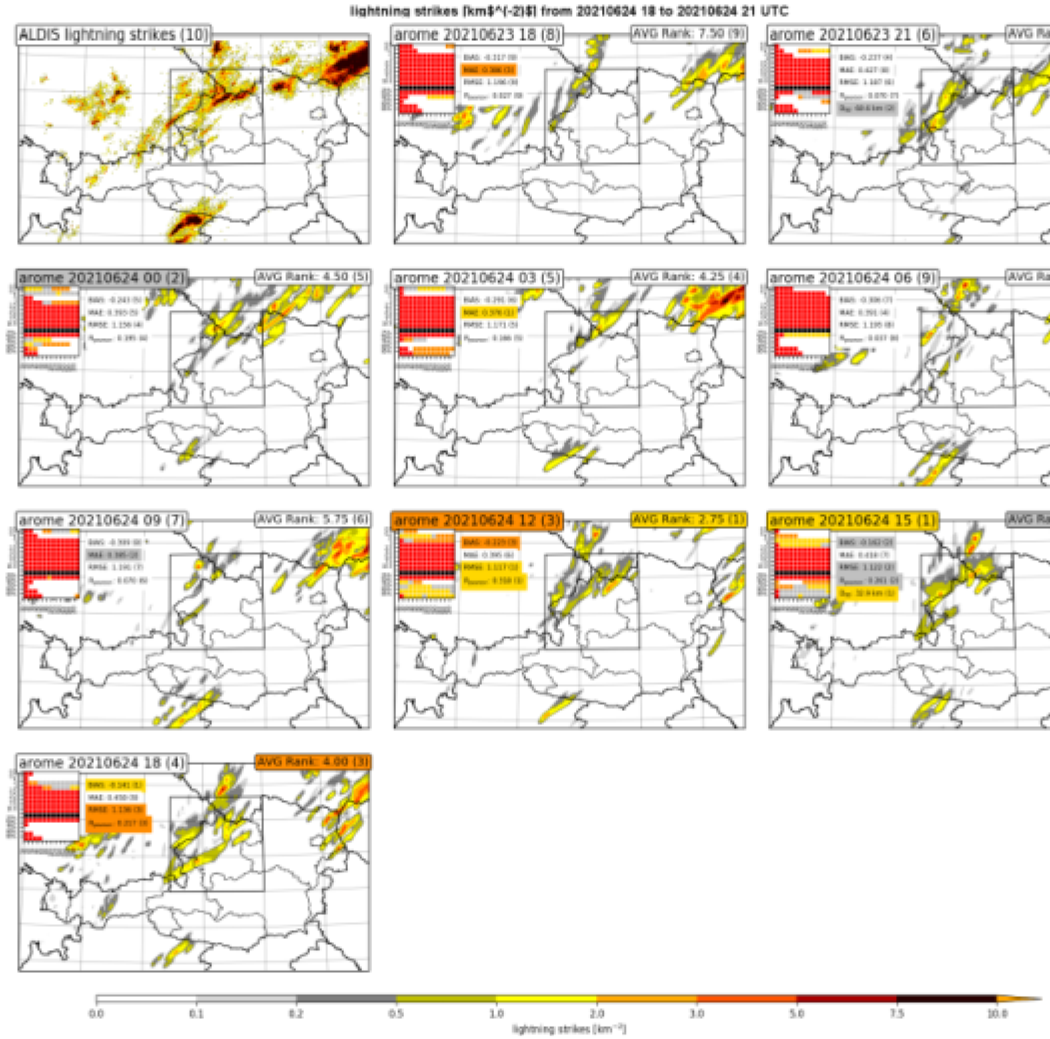


Figure 9: Lightning strikes from observations (top left) and 9 different AROME-Aut runs for 2021-10-24 between 18 and 21 UTC over Upper Austria, units are given in strikes per km^2.

Figure 9 shows an example of Panelification being used for a subdomain over Austria for a 3 hour period, comparing different runs of the same model. The tool shows the same scores that are also used for precipitation but with adapted thresholds where necessary.

The work on spatial lightning and hail verification using Panelification is still ongoing. The code of Panelification was adapted to use pygrib instead of epygram tool. As well, the code was adjusted to migrate from python2 to python3. The conda environment was built to run the code easily. The Panelification was installed on HPC ATOS (ECMWF).

Contributor: Phillip Scheffknecht: (1.5 pm - MQA2)

Visualization of 2D vertical cross-sections in Python

In **Croatia**, the work on the visualization of 2D vertical cross-sections in Python was finished and operational on the new HPC by the end of 2022. As well, a python-based verification dashboard was developed, which enables interactive work (selection of location(s), score(s)).

Finally, there is a plan to create an interactive system for real-time comparison of measured and modeled time series of near-surface parameters (python-based).

Contributors: Endi Keresturi (0.5 pm -MQA3), Iris Odak Plenkovic (1.5 pm)

Subjective validation of 1.3 km RUC model

In **Slovenia**, the purpose of this ongoing task is to evaluate the quality of the 1.3 km model. In particular, to run consistency and the ability to simulate the (severe) weather events while they are in progress. The main focus is on the development of the convection, particularly the onset of convection and the positioning of the convective systems. This work is also partly related to radar DA.

The procedure is to plot a time series of radar images in the top row and below that a large panel of many consecutive outputs in such a way that the cells in a column of this matrix all have the same time of validity but the lead time increases towards the bottom (Figure 10), in such view, an individual run lies on the diagonal. This enables to

1. study the performance of the assimilation to capture the current ongoing convection activity by comparing two top row,
2. to study the consistency of the model from run to run by comparing each column individually.

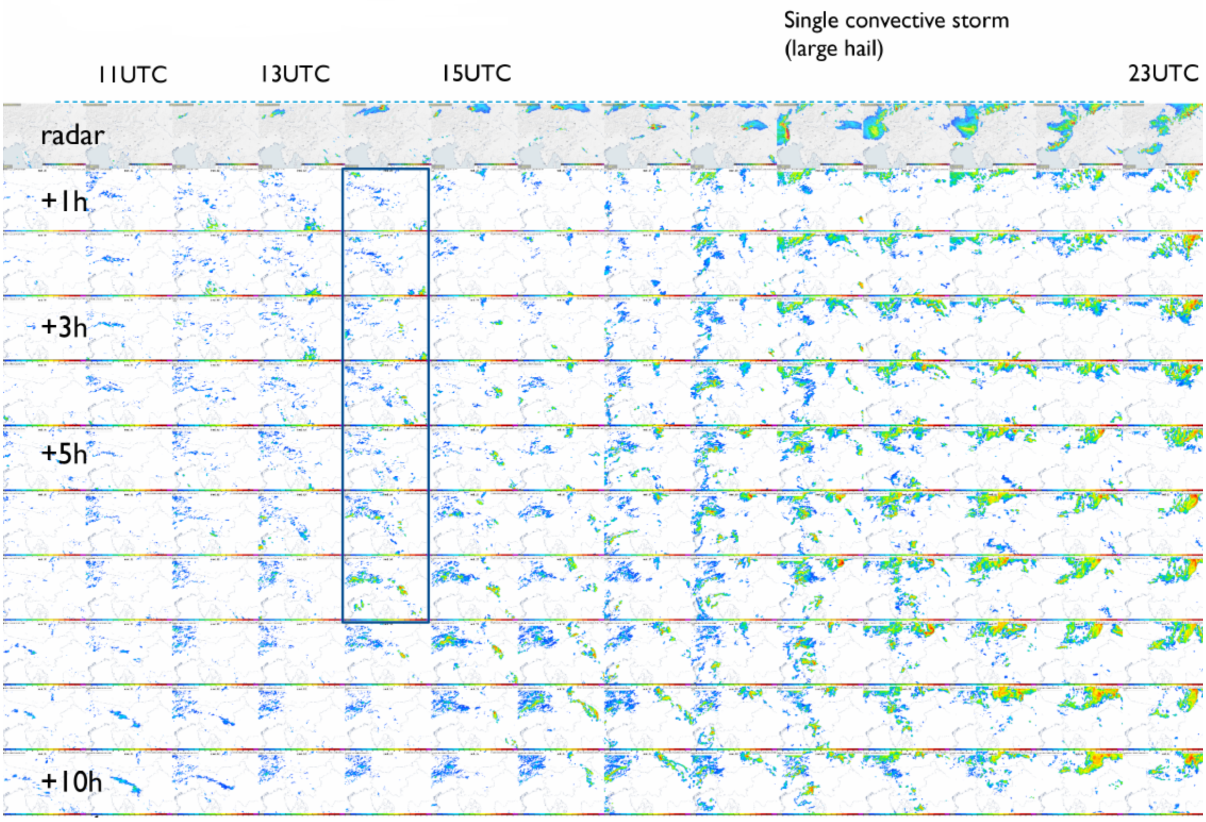
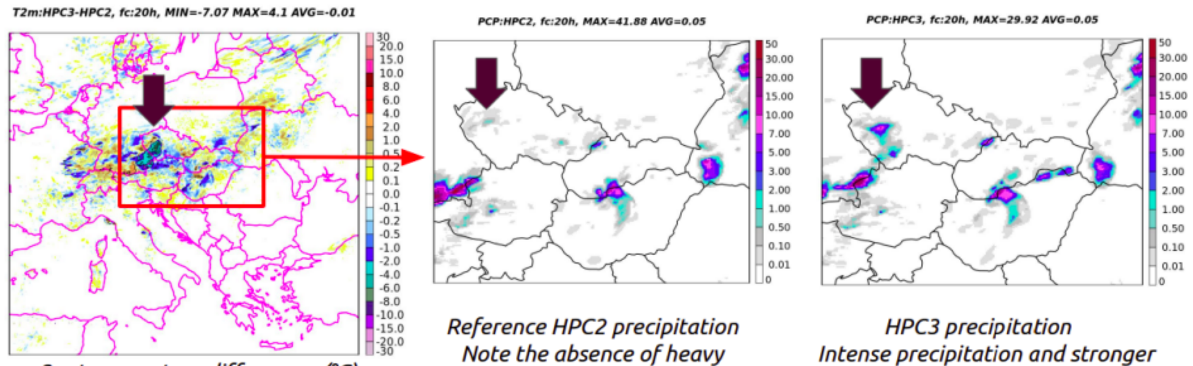


Figure 10: The vizualization of weather events while they are in progress.

Contributor: Jure Cedilnik (0.5 pm – MQA3), Neva Pristov (0.5 pm - MQA3), Benedikt Strajnar (0.5 pm - MQA3) – already added to HARP utilization

Validation of HPC3 forecasts vs HPC2

In **Slovakia**, the ALADIN/SHMU model was tested on HPC3 on a deep convection case of 15 August 2021, with the same initial and boundary conditions as on HPC2. Significant (exceeding 6°C) differences in 2m temperature were found after 20 h of integration, which seems to be related with convective precipitation (Figure 11). Similar differences can be found when comparing forecasts between other HPCs as well and even on the same HPC machine when changing the mode of optimalization (O2 vs O0).



2m temperature differences (°C)
HPC3-HPC2 forecast

Reference HPC2 precipitation
Note the absence of heavy
precipitation over Czech Republic

HPC3 precipitation
Intense precipitation and stronger
cooling over Czech Republic

Figure 11. 2m temperature difference HPC3-HPC2 forecast (left), HPC2 precipitation (middle), HPC3 precipitation (right).

Concerning long-term scores, the differences between the HPC2 and HPC3 operational productions were not that spectacular. The validation was done for the period of 20 January 00 UTC till 21 February 12 UTC over 95 Slovak AWS (Figure 12). The complete operational suite was tested here (including physics, DA, etc).

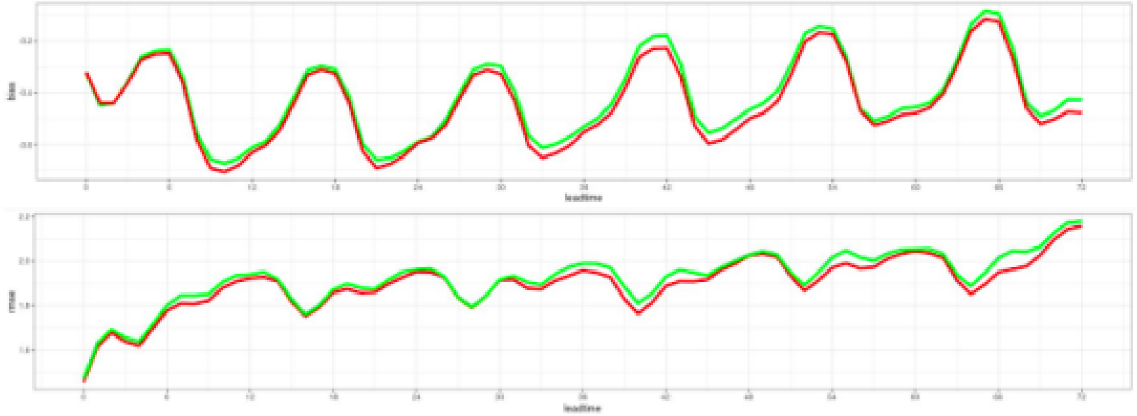


Figure 12. Scores of 2m temperature, bias (top), rmse (bottom) for the period of 20 January 00 UTC till 21 February 12 UTC over 95 Slovak AWS. HPC2 result is in green, HPC3 in red colour. Computed using **HARP**.

Regarding HPC3 mirror suite vs. HPC2 scores, RMSE and BIAS two different periods were computed. Figure 13 shows the results of RMSE and BIAS of relative humidity for winter period 23/01/-11/02/2022 and figure 14 shows the same scores for rainy period 12/05-27/05/2021 for temperature at 850 hPa. The validation was done for operational suite HPC2 and mirror suite denoted SHMU on HPC3. More results can be found in SMHU presentation at the ALARO-1 working days, from Prague this year. (https://www.rclace.eu/media/files/ALARO/alaro1_wd22/2022_ALARO1_WD_SHMU status.pdf)

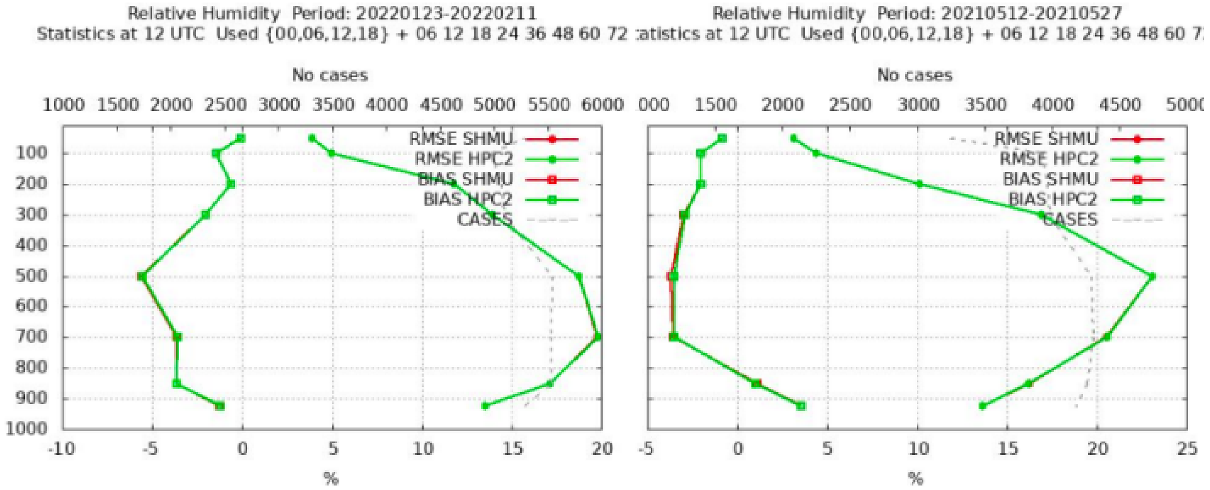


Figure 13. RMSE and BIAS of RH for winter period 23/01/-11/02/2022 for operational suite HPC2 (green) and mirror suite denoted SHMU on HPC3 (red).

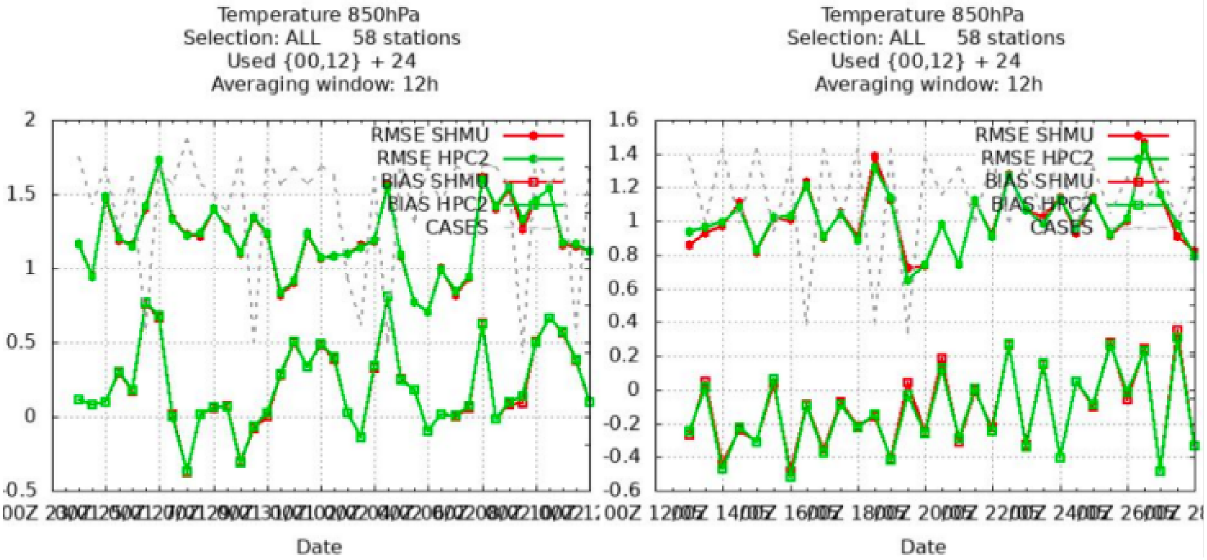


Figure 14. RMSE and BIAS of T850 hPa for rainy period 12/05-27/05/2021 (right) for operational suite HPC2 and mirror suite denoted SHMU on HPC3. Two right panels: the +24 h forecast scores evolution of temperature at 850 hPa over the same periods

Contributers : Slovakian team (1 pm - A. Simon, M. Petras, O. Spaniel, V. Tarjani, M. Derkova)

Total Contributors: Boglárka Tóth (1.5 pm), Phillip Scheffknecht: (1.5 pm – MQA2), Croatia team, Slovenian team, Slovakian team (1 pm - A. Simon, M. Petras, O. Spaniel, V. Tarjani, M. Derkova)

4 Action: Post-processing of model output

In 2021, in **Hungary**, work was started on post-processing of AROME and AROME-EPS outputs. The main objective is to improve the forecasts for global radiation and 100-meter wind speed. It will be decided if this topic can be kept for the future plans and extented to all the countries.

The aim was to analyse the quality of global radiation forecasts of local AROME-EPS (Jávorné et al., 2020), which are of great use for the renewable energy sector. OMSZ has been using a locally developed tool for ensemble (pointwise) verification. This script has been extended to process global radiation data. We use the global radiation measurements of OMSZ network (of 33 qualified stations) for the verification. Forecast data of the 11 members of AROME-EPS initialized at 0 UTC are taken at the grid points nearest to the stations. The three summer months in 2021 were chosen for the first tests. The evalution was made partly by Natália Gáspár, a BSc student in meteorology (Gáspár, 2022).

The weather for the summer of 2021 in Hungary was as follows: June was dry, with the least amount of cloud cover and the most sunshine hours of the three months, mostly due to persistent the anticyclone situation. Contrary to the climatological average, the maximum of global radiation was measured in June instead of July. August had the most cloudiness and the least surface solar radiation.

A negative bias (Figure 15) during daytime for EPS-mean was found, for all summer months. This is mainly caused by an average overestimation of cloudiness by AROME(-EPS) forecasts. However, after sunrise and before sunset we found slight positive bias in radiation forecasts. The maximum average CRPS values are around 90 W/m^2 (Figure 15). The EPS is clearly underdispersive shown both by spread-skill diagrams and rank histograms (Figure 15).

CRPS curves are similar for all summer months despite the differences in EPS mean RMSE, bias and spread. Mean RMSE was the highest in July, while lower values are seen in June due to the more stable anticyclonic weather situations. The J-shape of the rank histogram shown for 12-hour lead time (when maximal solar radiation reaches the surface) indicates the strong underestimation which is even higher at 9 hours. Histograms for the hours after sunrise, before sunset and the second day are somewhat more uniform.

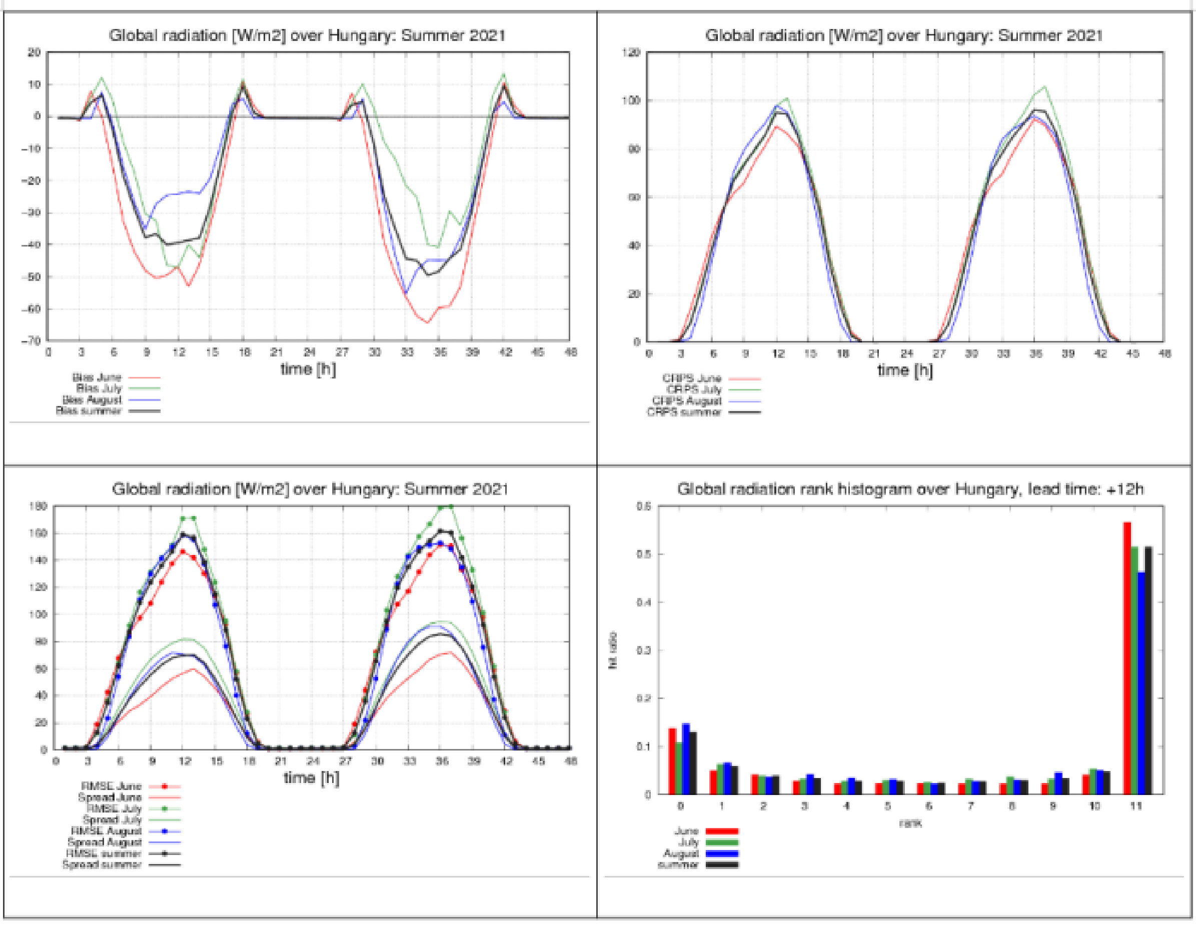


Figura 15: Global radiation (in W/m^2) bias of ensemble mean (top left), CRPS (top right), RMSE of ensemble mean and spread (lines with dots and solid lines, respectively, bottom left) in function of lead time, and rank histogram at lead time of 12h (bottom right) for AROME-EPS forecast for June (red), July (green), August (blue) and summer (black) of 2021 over Hungary. Values for a given lead time correspond to the average radiation for the previous hour.

Adaptation of EMOS and machine learning post-processing methods on AROME-EPS forecasts

In **Hungary**, the mathematician colleagues developed different methods for improving the global radiation and 100 m wind forecasts, taking into account the properties of these quantities. In both cases, a one-year data series was used by them for testing.

For radiation, the EMOS technique proved to be successful, in which the distribution of predictions was approximated with censored normal (CN) or censored logistic (CL) functions. Using a 31-day rolling training period, parameters of the distribution were estimated based on data of seven stations of OMSZ measuring network and 11 members of AROME-EPS (Jávorné et al., 2020) forecasts initialized at 0 UTC in the grid points to nearest the stations. As a result, the CRPS of the improved probabilistic forecast could be reduced by 16-18% with respect to the

CRPS of the raw AROME-EPS (Schulz et al., 2021). CN-EMOS method proved to be numerically somewhat more stable.

For wind, EMOS based on truncated normal (TN), log-normal (LN) and truncated generalized extreme value (TGEV) distributions were applied, and also, multilayer perceptron (MLP) machine learning method based on TN or LN distributions. Observations of three wind farms in the northwestern part of Hungary and corresponding AROME-EPS forecasts at hub height (100m) and nearest grid points were used. A 51-day long training period was found to be optimal. MLP was the most successful, with CRPS improved by 10-15% of the raw EPS (Baran and Baran, 2021). In the received R and python3 codes, sampling and verification routines are also included.

The work on integrating the received code into the operating system is ongoing. The improved forecasts are produced daily: the output gives the value of the improved forecast per station and every 15 minutes by determining 11 equal quantiles from the given distribution function. (Each method handles each lead time independently.) The running procedure consists of 3 steps:

3. The collection of forecast data for the given day and observations for the preceding day is done by the (pre-)operational suite on the HPC.
4. The EMOS fitting and the MLP training run on a separate computer and produce the outputs.
5. At the end, verification is done for the forecasts for the preceding day, some statistics are calculated for the previous period of about one week, and also few simple meteograms (Figure 16, left) and verification plots (Figure 16, right) are produced to compare the improved forecast and the raw AROME-EPS.

In autumn the test was done on the current operational AROME-EPS which is downscaling of ECMWF-ENS and we experienced the expected improvement as a result of the post-processing (Figure 17 and 18). In general, the operational AROME-EPS is under dispersive and biased, both features seem to be improved after the post-processing. During the recent winter days the test is done on the pre-operational AROME-EPS EDA suite (see EPS report for details) and the improvement is not so clear. The reason may be both the different AROME-EPS basis and different weather situations. We note that an improvement in CRPS does not mean an improvement in RMSE of ensemble mean to the same extent, i.e. end users should consider the probability information and not only EPS mean.

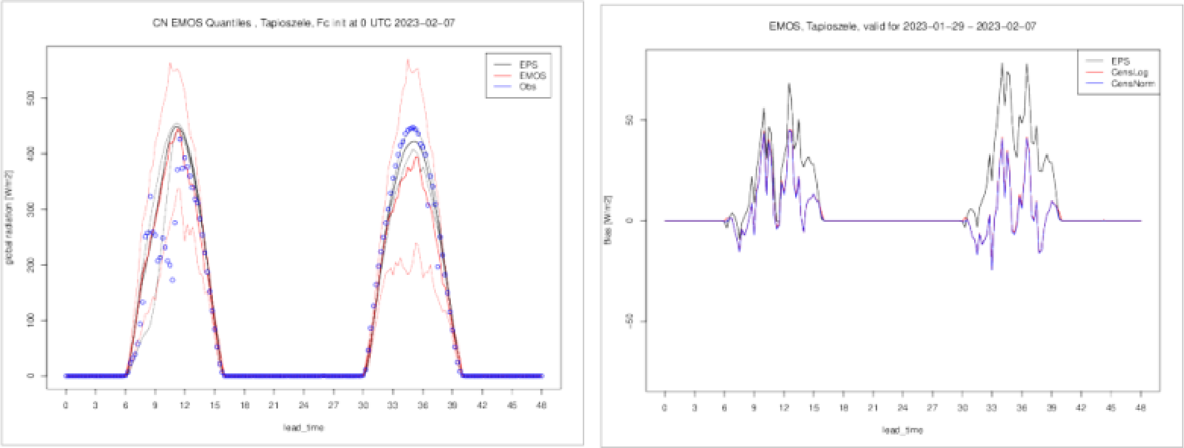


Figure 16: left - Ensemble meteogram for radiation on 7 February 2023 based on raw AROME-EPS forecasts (ensemble mean in black, minimum and maximum in grey), on CN-EMOS (ensemble mean estimated from the corrected distribution function with thick red line, upper and lower quantiles with thin red lines) and observations with blue circles. Right - bias of ensemble mean for radiation as function of lead time between 29 January and 7 February 2023 based on raw AROME-EPS forecasts (black), CN-EMOS (blue) and CL-EMOS (red). Both plots are valid for station of Tápiószele.

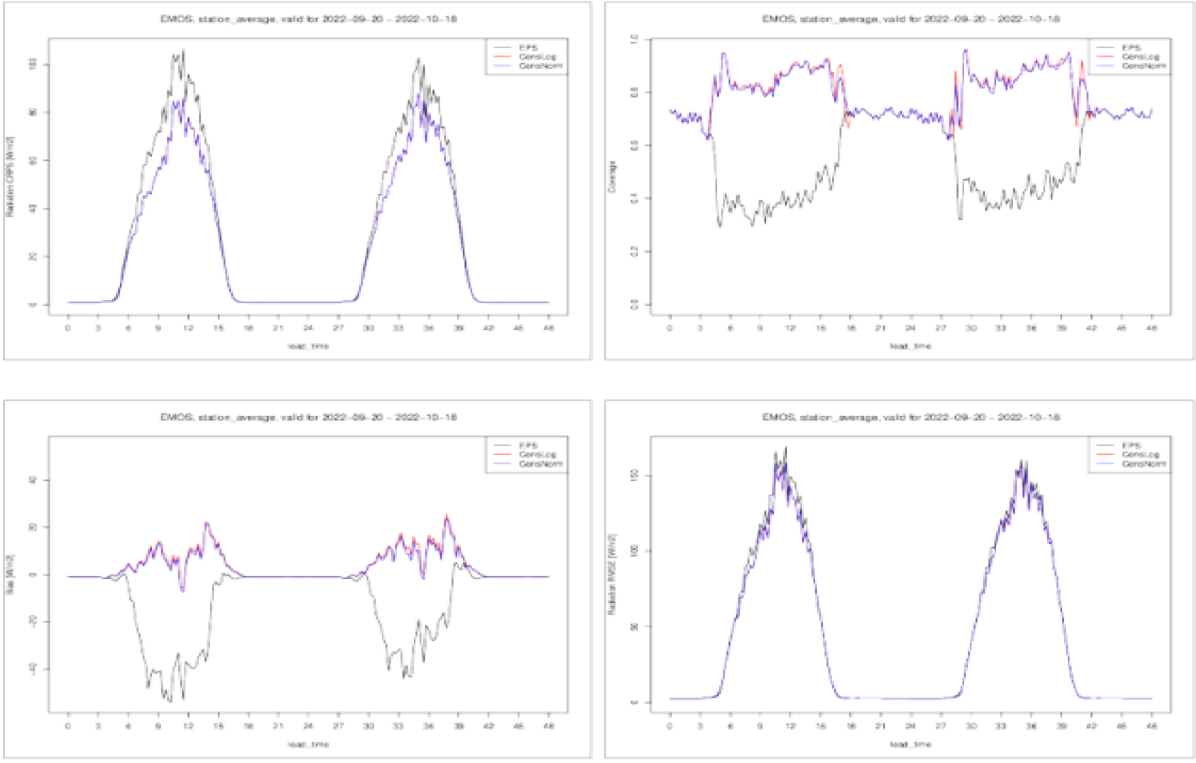


Figure 17: Verification scores for global radiation forecasts between 20 September and 18 October 2022 for 7 stations: CRPS (top left), coverage (top right), bias of ensemble mean (bottom left) and RMSE of ensemble mean (bottom right) based on raw AROME-EPS forecasts (black) and the predictive distribution functions provided by CN-EMOS (blue) and CL-EMOS (red).

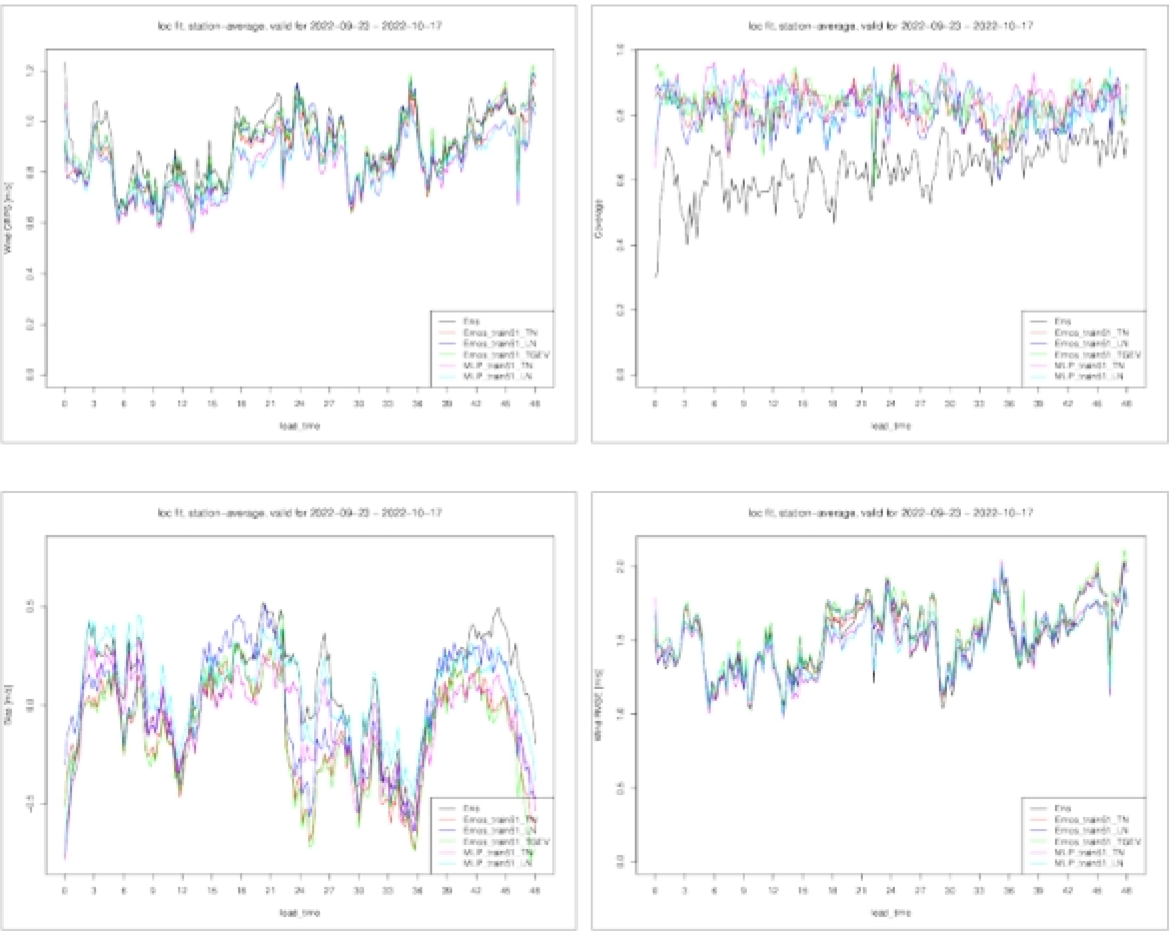


Figure 18: Verification scores for 100m wind forecasts between 23 September and 17 October 2022 for 3 stations: CRPS (top left), coverage (top right), bias of ensemble mean (bottom left) and RMSE of ensemble mean (bottom right) based on raw AROME-EPS forecasts (black) and the predictive distribution functions provided by TN-EMOS (red), TN-EMOS (blue), TGEV-EMOS (green), TN-MLP (magenta), LN-MLP (cyan).

Some smaller corrections and improvements on verification and plotting scripts are necessary. After that, the aim is to choose the best predictive distribution function and method, by running and verifying all of them over a longer period of time. In the future, these methods on additional stations and data needs to be tested, but also to develop the data pre-processing procedure.

References

Baran, S. and Baran, Á., 2021: Calibration of wind speed ensemble forecasts for power generation. Időjárás 125, 609-624.

Gáspár N., 2022: Evaluation of AROME-EPS global radiation ensemble forecasts. BSc Thesis, Eötvös Loránd University, Budapest. (in Hungarian)

Jávorné Radnóczi, K., Várkonyi, A., Szépszó, G., 2020: On the way towards the AROME nowcasting system in Hungary. ALADIN-HIRLAM Newsletter, 14, 65–69.

Schulz, B., El Ayari, M., Lerch, S., Baran, S., 2021: Post-processing numerical weather prediction ensembles for probabilistic solar irradiance forecasting. Sol. Energy 220, 1016–1031.

Contributor: Katalin Jávorné Radnóczi (1.5 pm + 3 pm - E6)

GAM (Generalized Additive Models) models to estimate wind speed

In **Romania**, A post processing method of the model output was tested for wind speed forecast of ALARO. This method consists of applying GAM (Generalized Additive Models) models to estimate wind speed. These are regression models that allow for more complex relations (than a simple regression model) between the response variable and the predictors. For this study, several GAM models were defined. They differ mostly in the predictors they use in the regression equation.

The first GAM model (m1) is practically a simple regression model (one single predictor which is the wind speed simulated by ALARO). Model m2 takes into consideration the coordinates (latitude and longitude) of the point where is applied; model m3 includes the altitude of the station; m4 includes the simulated wind direction; m5 adds the 24 hours lagged simulated wind speed; m6 takes into consideration two local characteristics: the distance to the Black Sea and the number of urban pixels within 3km radius for the point considered. These models were applied for a one year training period (2021), and then were used to estimate the wind speed for the period January – May 2022, at 157 meteorological stations in Romania.

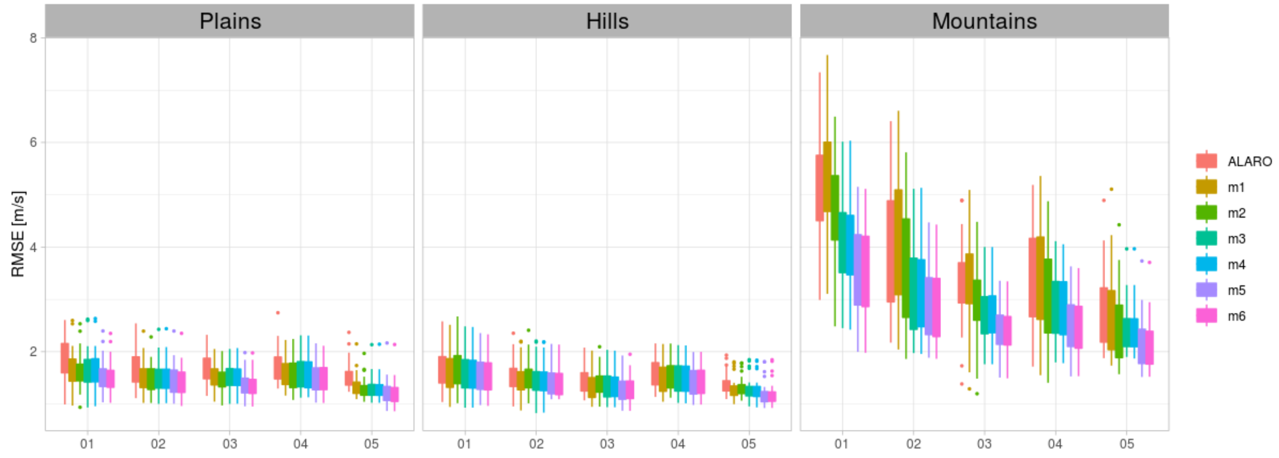


Figure 19. Boxplots of daily RMSE values for each month, model and altitude type (plains – stations located between 0 and 300 m, hills – between 300 and 800 m and mountains - over 800 m).

It was observed from the RMSE scores (Figure 19) indicate that the models show different results, depending on the month or altitude of the station.

Improvements are more visible for mountain stations, mostly for models m5 and m6. This result shows the significance of adding more predictors in the regression model. For example, figures 20 and 21 show the daily mean wind speed for station Vf. Omu (located at 2506 m altitude) from models m1 and m6 compared to ALARO and the observations. Model m1 shows very slight differences compared to ALARO in this case and both underestimate the real wind speed, while model m6 leads to wind speeds more closer to registered values. These results show that this type of post processing method (which is quite easy to implement using R tools) can be found useful for specific cases where the model forecast proves problematic.

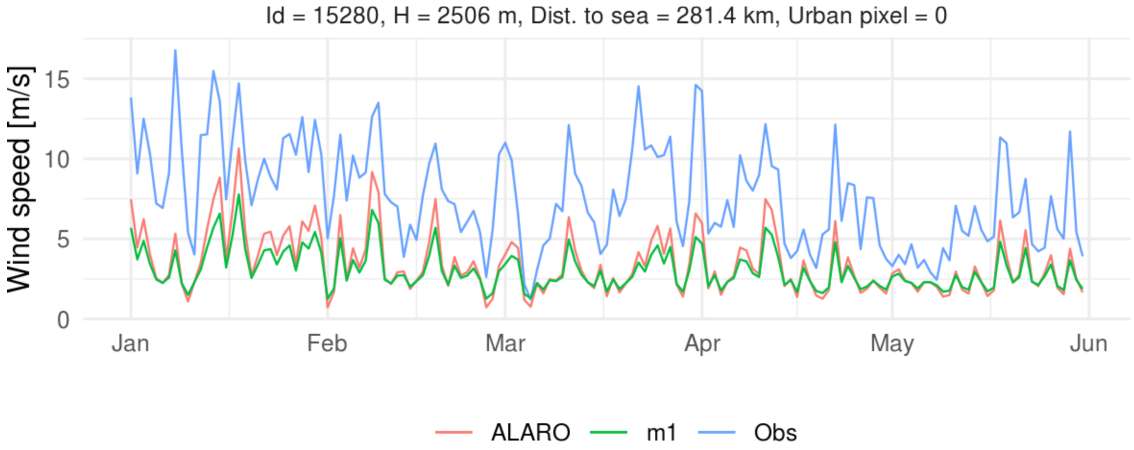


Figure 20. Daily mean wind speed from ALARO (red), model m1 (green) and observations (blue).

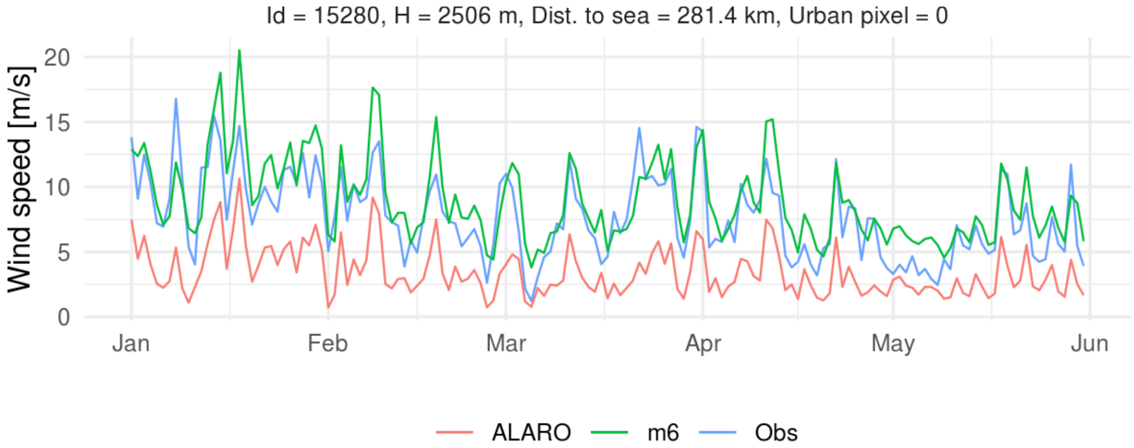


Figure 21. Daily mean wind speed from ALARO (red), model m6 (green) and observations (blue).

Contributor: Alexandra Craciun (1 pm – MQA3)

6 Action: RC-LACE Github platform

Regarding gitlub platform, it is not suitable for RC-LACE to host the platform there. A free platform, as github platform, is more convenable. A github platform dedicated for ACCORD consortium already exists. What application do you want to store on github platform?

Contributers: -

7 Action: Database of cases

Several case studies were provided by hungarian, austrian and slovakian teams. See previous report.

Contributors: Gabriella Tóth (0.25 pm), Christoph Wittmann (0.25 pm), Martin Bellus (0.25 pm)

8 Action: Trainings

Summary of resources [PM]

Subject/Action	Resource		LACE stays	ACCORD stays (MQA?)
HARP implementation and verification for deterministic and probabilistic forecasts	planned	realized		
	1.75	4.75		
HARP linked to OPLACE database	-	1		
Multiple verification methods	6.25	10.5		
Post-processing of model output	1.25	4.5		
RC-LACE github platform	2	-		
Database of cases	-	0.75		
Trainings	-	-		
Total	9.25	21.5		