Newsletter n°4

19 June 2023

ACCORD
A Consortium for COnvection-scale modelling
Research and Development

CONTACT US

Claude Fischer, ACCORD Programme Manager
@ pm@accord-nwp.org
Patricia Pottier, Consortium Scientific Secretariat
@ css@accord-nwp.org
http://www.accord-nwp.org/

Contents

Edito

Claude Fischer, ACCORD Programme Manager

The 4th ACCORD Newsletter contains a fair limited number of contributions (seven) however it reaches slightly more than 50 pages in length. The recently held ASW, with its large number of sometimes very high-quality contributions, might eventually have been a reason preventing more contributions from being received (rather than the opposite). In this volume, you will be able to read an update about the progress on “harp” and a recent working week, complemented by two contributions that illustrate efforts on meteorological verification by two Member Institutes. The other scientific articles provide insight in research work on data assimilation and EPS. In link with probabilistic modelling, we welcome a new Doctor from Belgium !

Talking about welcoming … and saying goodbye. The last ASW in Tallinn gave the opportunity to ACCORD staff to meet a last time with Patricia before she leaves the consortium activity, and her role as CSS, for retirement. I use the opportunity of this editorial to thank Patricia again for her instrumental role in our consortium efforts, during the ALADIN times, in the context of the ALADIN-LACE-HIRLAM convergence process and in the very first years of ACCORD. Her steady involvement and ability to propose imaginative organizational solutions definitely have been a tremendous advantage for ACCORD. The conclusion to this Newsletter introduction however shall remain a welcoming, since from 1 September 2023 our new CSS Anne-Lise Dhomps will officially join the ACCORD activities. Anne-Lise has been working under several project contracts at MF in the field of observations for data assimilation, and the CSS position is a new motivating challenge for her. We are looking forward to a fruitful collaboration with our new CSS and wish her all the best in this important central position in the consortium.

Claude Fischer

Patricia Pottier and Anne-Lise Dhomps, the CSSs doing the liaison

ACCORD events in 2023

Patricia Pottier, ACCORD Consortium Scientific Secretary

1 Introduction

The outcomes of the ACCORD events such as governance bodies meetings (Assembly, PAC, STAC, LMT), Management Group (MG) meetings, Working Days, Working Weeks, thematic WG meetings can be found on the ACCORD "Events" webpages: slides, minutes, summary, photos, videos, … (when available). The material and conclusions of the thematic meetings (WD) organised by the MG with the team or WG in their area are available on the relevant part of the ACCORD wiki.

2 ACCORD governance meetings in 2023

Committees (PAC, STAC), Assembly and Bureau

- Selection committee meetings for DA and Physics ALs, 27 February
- ACCORD Bureau (DA and Physics ALs), 28 February
- STAC spring meeting: 25 May morning (video)
- Bureau meeting: 7 June (video)
- 6th Assembly video meeting, 26 June afternoon
- autumn STAC, two half-day, in Brussels (RMI) and hybrid, date to be defined
- PAC meeting, if convened by the Assembly
- Regular Bureau meetings in the autumn, in addition to the ACCORD Bureau to prepare the Assembly (3 weeks before the Assembly)
- Assembly meeting, 4 December, Reading (hybrid)

Management Group

- The Management Group resumed their every other Friday morning meetings at the beginning of January and on 25 August after the summer break.
- MG visit to INM and in-person meeting: 4-5 May 2023, Tunis (hybrid)
- possible in-person meeting and visit to Estonian team: 28 or 29 September 2023, Reykjavík, after EWGLAM/SRNWP meetings.
- Additional meetings are dedicated to the preparation of the DAP2023, the preparation of the RWP2024, the reporting of RWP2023, …

LTM meetings

- 6th LTM meeting, Tallinn, Estonia, 27 March 2023 (besides the All Staff Workshop)
- 7th LTM meeting, Reykjavik, Iceland, 26 September 18:00-19:00 CEST (ie 4-5pm UTC, besides EWGLAM)

3 Scientific and technical meetings

Thematic regular video-meetings

- Transversal activities on future software infrastructure, more information on ACCORD wiki SPTR dedicated page
- WG on Very High Resolution Modeling (VHR-MOD): meetings information
- WG on Machine Learning (ML): meetings information
- DA Research Teams and Support Teams meetings, more information on the ACCORD wiki dedicated pages
- Surface monthly meetings, more details on the ACCORD wiki pages
- O2R WG: WG on Operation to Research, on-line meetings from October 2022 through March 2023
- harp community on-line meetings (1st meeting on 17 April 2023, more details on the ACCORD wiki pages)

All Staff Workshop

The big ACCORD event in the first semester of 2023 was the **3rd ACCORD All Staff Workshop, 27-31 March 2023, Tallinn.** It was organised as a hybrid meeting, with participants in Tallinn Nordic Hotel and remote participants. For this second edition in an hybrid format, it did run smoothly: WARM CONGRATULATIONS to the Estonian hosts. Future ASWs will be hybrid as well: 4th ASW on 15-19 April 2024 (SMHI, Norrköping). A few figures of the ASW2023:

- 198 registered participants of which 90 on site,
- in addition 32 remote participants just popped up without registering,
- 65 to 95 remote participants during the sessions,
- 8 sessions plus the opening & closing sessions, and 3 side-meetings,
- 78 speakers including 19 poster presenters,
- video-recording: 30 hours. |

A big **THANK YOU** to the participants, in-situ or remotely, who made this All Staff Workshop a real success.
And a big **PERSONAL THANK YOU** from the 1st CSS ACCORD for all the touching attention on the occasion of the announcement of her upcoming departure (see photos).

All the material is available on the ASW2023 page on the ACCORD website.

EWGLAM

On the kind invitation of the Icelandic Meteorological Office the 45th European Working Group on Limited-Area Modelling (EWGLAM) and 30th Short Range NWP (SRNWP) EUMETNET meetings will take place in Reykjavík, Iceland between 25-28 September 2023, in a hybrid form. The special topic for the meeting will be in connection with ensemble prediction over the working title "uncertainty of modelling components and their impact". Although not an ACCORD event, but with the participation of many ACCORD colleagues, the meeting may offer the opportunity to the LTMs and to staff to meet in-person during the week. The MG intends to meet the Estonian team on 29 September.

Working Days and Working Weeks

- ACCORD DA WW on Algorithms (including OOPS) and New obs. Obsconvert, SMHI, Norrköping, Sweden, 6-10 Nov 2023 see Wiki dedicated page
- Snow, SBL, fog and MUSC WW, Sodankylä, Finland,18-22 Sep 2023 see Wiki dedicated page
- LACE WD and Nowcasting/Mode-S, CHMI, Czech Republic, 11-15 Sep 2023 see Wiki dedicated page
- Surface in-person WW, SMHI, Norrköping, Sweden, 22-26 May 2023 see Wiki dedicated page
- ACCORD DA WW DA diagnostics/tuning and satellite observations, Madeira, Portugal, 22-26 May 2023 see Wiki dedicated page
- EPS WW, Met Norway, Oslo, 24-28 April 2023 see Wiki dedicated page
- ACCORD DA WW on Algorithm (incl. OOPS), ML, GNSS, radar, KNMI, De Bilt, Netherlands, 6-10 Mar 2023 see Wiki dedicated page
- WW on harpSpatial, DMI, Copenhagen, Denmark, 20-24 Feb 2023 see Wiki dedicated page
- Very-High Resolution Modelling Workshop, SMHI, Norrköping, Sweden, 14-16 Feb 2023 see Wiki dedicated page
- WW about the diurnal cycle of T2m and Q2m WW, France, 9-13 October.
- Davaï developers WW, place and date, t.b.c.
- Horizontal length scales across CSCs / 3D-turbulence related topic, MF, Toulouse, France, date t.b.c.
- TEB WW, MF, Toulouse, France, Autumn, date t.b.c.
- ALARO/SURFEX WW, place and date, t.b.c.
- PH6/CAR WW, place and date, t.b.c.

MQA working week on spatial methods and data in harp

Carl Fortelius

1. Introduction

A working week (WW) on spatially distributed data (i.e. data not tied to fixed stations) and spatial methods in harp was arranged as a hybrid event in February and hosted by DMI at the institute headquarters in Copenhagen. The 4-day week attracted 12 on site participants and 14 registered remote participants. The programme contained a review of methods and data in use today, working on further development, sharing of experiences, and making plans for the future.

Presentations and other material related to the ww, including a survey of available data supporting spatial methods, are available at the ACCORD MQA wiki:
https://opensource.umr-cnrm.fr/projects/accord/wiki/MQAWW202302

2. Data and methods in use

Spatially-oriented verification methods in harp are being used for research or operationally at several institutes, typically involving software developed locally for interfacing to observations and forecasts. Presentations at the WW contained examples of using traditional spatial metrics such as the Fractions Skill Score (FFS, Roberts and Lean, 2008) or Structure, Amplitude Location score (SAL, Wernli et al., 2008), but also novel approaches such as the SLX (Sass, 2021), error-fractions skill score and dispersion fraction skill score (eFSS, dFSS, Singleton, 2023) or verification by cyclone tracking and pattern recognition aided by machine learning. Data used during the WW are listed in Table 1 (Section 5).

3. Developments in harpSpatial and related software

Recent developments in harp and harpSpatial were presented in talks by Alex Deckmyn and Andrew Singleton. The SAL, FSS, and neighbourhood adapted contingency tables are all fully implemented in a development version of harp which is available in github:
https://github.com/harpSpatialTeam, or https://github.com/harphub/harpSpatial-team. Other scores such as the neighbourhood-aware CRPS (Stein and Stoop, 2022) or the error-fractions skill score (eFSS) and dispersion fractions skill score (dFSS) (Singleton, 2023) are being implemented.

At the working week a tutorial example of applying utilities of harpSpatial was developed by Andrew Singleton, and posted in the repository: https://github.com/harphub/harp-ww-feb2023. The repository contains R-code and configuration files as well as a small sample of data supporting verification of precipitation using FSS and SAL. Two examples of plots produced by the software are given in Figs 1 and 2 (Section 5). The repository contains detailed instructions for installing and and applying the code. The software is well suited to be used as a template for other applications besides the examples given, and thus provides an accessible point of entry to harpSpatial as a whole.

Other contributed software developments include a bash script for accessing IMERG precipitation data using wget by Juan Jesús Gonzáles Alemán, and utility functions for easily defining and extracting sub domains of harp geofields by Martin Petras. Using these functions, sub domains can be defined in terms of geographical coordinates, thus eliminating any explicit reference to matrix indexes.

4. Next steps

New developments from the harpSpatial-team branch will be included in the next harp version. Scientific and technical documentation as well as user's guides need to be reviewed and updated as appropriate. Developers are strongly encouraged to consider sharing properly documented software for acquisition and ingestion of data in the harp users repository for harpUserScripts on github.com/harphub.

An interface to the ODB database is essential for utilising the data screening methods applied in data assimilation, and would facilitate the use of all assimilated data for verification, thus extending enormously the observational basis (see Randriamampianina et al, 2023). Common interfaces scatterometer winds, IMERG precipitation analyses, and MSG SEVIRI brightness temperatures should likewise be items of high priority in the rolling work plan. It is recommended to start with data that are available via MARS.

In the medium term, application of spatial verification methods should be incorporated into the accord-verification scripts along with other metrics.

It was decided at the ww to continue with regular meetings open to everybody interested in the development of a harp and especially harpSpatial. Material related to the meetings is posted on the ACCORD MQA wiki:
https://opensource.umr-cnrm.fr/projects/accord/wiki/Harp_community_meetings

5. Figures and table

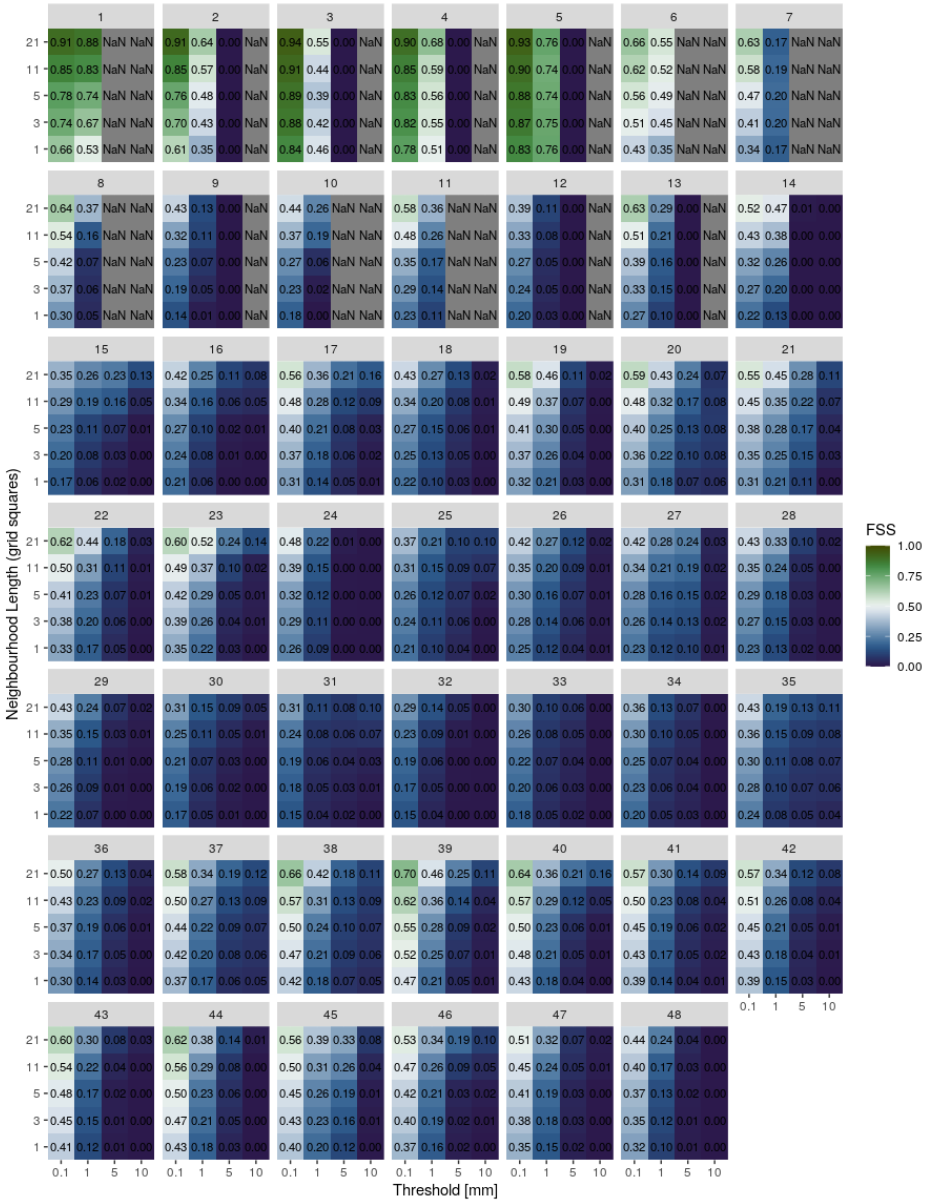

Figure 1: FSS for hourly precipitation amount for each hourly lead time in a single 48-hour forecast by the MetCoOp ensemble (MEPS) initiated at 00Z on the 15th of August 2022. Observations are Norwegian quantitative precipitation estimates based on a combination of radar, rain gauges, and NWP. The results for different lead times reflect the evolution of the observed precipitation pattern, as well as the evolution of forecast errors.

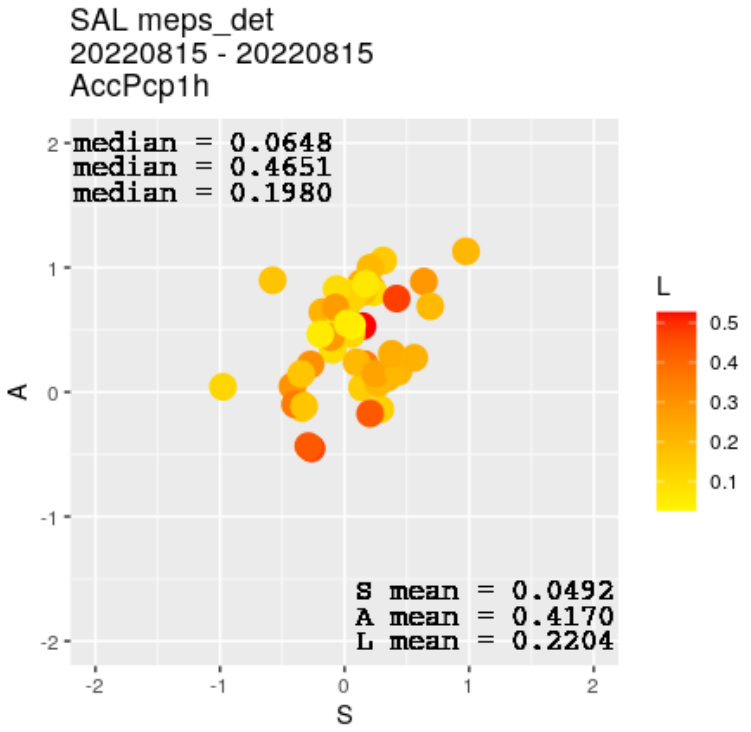

Figure 2: Errors in Structure, Amplitude, and Location (SAL) for hourly precipitation amount for 48 hourly lead times in a single 48-hour forecast by the MetCoOp ensemble (MEPS) initiated at 00Z on the 15th of August 2022. Observations are Norwegian quantitative precipitation estimates based on a combination of radar, rain gauges, and NWP.

TABLE 1. Data used in work presented at the ww

Parameter	**Data source**	**Institute/*presentation(s)***
QPE	IMERG (Integrated Multi-satellitE Retrievals for GPM) precip. analysis National radar+gauges	AEMET/*Gonzáles Alemán* Met. Norway/*Singleton* DMI/*Feddersen,* Sass
Reflectivity	National radar	AEMET/*Gonzáles Alemán*
Brightness temp. (10.8 µm)	MSG 2 SEVIRI	AEMET/*Gonzáles Alemán* Geosphere/*Schmederer*
Near surface wind speed and direction	OSI SAF product (HY-2b, HY-2c, METP-B/C)	UWC-W/*Fannon*
PTU-profiles	Correctly-located soundings	UWC-W/*Fannon*

6. References

Roger Randriamampianina, Roger, Eoin Whelan, and Per Dahlgren, 2023: Verification against all (assimilated) observations, ACCORD Working Week DAWW202305, https://opensource.umr-cnrm.fr/attachments/download/5212/VerAllObs_2023.pdf

Roberts, N., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of con-vective events. *Mon. Wea. Rev.*, 136, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Sass, B. H. (2021). A scheme for verifying the spatial structure of extremes in numerical weather prediction: Exemplified for precipitation. *Meteorological Applications*, 28(4), e2015. https://doi.org/10.1002/met.2015

Singleton, Andrew, 2023: Fractions Skill Score for ensembles with harp, 3rd ACCORD All Staff Workshop, 27-31 March 2023, Tallinn & hybrid. http://www.umr-cnrm.fr/accord/?3rd-ACCORD-All-Staff-Workshop-27-31-March-2023-Tallinn-hybrid

Stein, Joël and Fabien Stoop, 2022: Neighborhood-Based Ensemble Evaluation Using the CRPS. Monthly Weather Review, 150 (8), pp.1901-1914. https://doi.org/10.1175/MWR-D-21-0224.1

Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL—A Novel Quality Measure for the Verification of Quantitative Precipitation Forecasts. *Mon. Wea. Rev.*, **136**, 4470–4487, https://doi.org/10.1175/2008MWR2415.1.

Verification of precipitation forecasts simulated by ALADIN and AROME over Algeria

Islam BOUSRI[1], Mohamed MOKHTARI[1]

(1) Numerical Weather Prediction Department, Office Nationale de la Météorologie, Algiers, ALGERIA.

Abstract

Precipitation forecasts have always been of great interest to forecasters because they influence daily life. In this paper we present the first results of verification of precipitation simulated by ALADIN and AROME over Algeria using a contingency table approach. The main synoptic messages over the whole domain of ALADIN-Algérie has been used as reference.

A fuzzy theory is applied to compare the forecasts with a specific perimeter, thus allowing to take into account the spatial variability of precipitation in a more accurate way. This allows the performance of precipitation forecasting models to be evaluated in terms of detection probability, false alarm rate, precision and accuracy.

Examples of results obtained during September to December of 2022 period are presented, highlighting the evolution of precipitation forecast model performance over time. These results provide important informations on the ability of ALADIN and AROME operational models to predict precipitation in Algeria and give the identification of the strengths and weaknesses of these two models.

1 Introduction

Precipitation is a discrete variable with stochastic behaviour that on small scales presents fractal properties. For these reasons, it is difficult to simulate and verify. Several precipitation verification methods have been developed. The contingency tables was one of these method and they were entered the field of forecast verification more then one hundred years ago and they continue to be an area of intensive research. The reason is the ability of contingency table to condense and clearly display the properties of some set of forecasts and corresponding observations. Utilization of contingency tables is particularly useful for the verification of quantitative precipitation forecasts, which is for many reasons difficult to verify.

For example, Gold et al. (2019) used a contingency table to verify the performance of probabilistic rain and lightning forecasts and determine necessary adjustments using an established verification process. Baig et al (2023) made a study to examine the consistency and applicability of four satellite precipitation products, namely PERSIANN, PERSIANN-CCS, PERSIANN-CDR and PDIR-Now, over the UAE using the contingency table. Giarno et al. (2018) evaluated the accuracy of satellite precipitation estimates in regions with large variations in rainfall patterns. Montesarchio et al. (2015) compared different precipitation threshold estimates using empirical data, hydrological simulations, and probabilistic methods, and evaluated system performance using contingency tables.

The purpose of this paper is to explore the use of the contingency table as a statistical tool to verify the precipitations simulated by the operational numerical weather prediction models ALADIN and AROME over Algeria. We will present the basic principles of the contingency table and explain how it can be used to evaluate the performance of precipitation forecasting models. We will also discuss the benefits of this approach, emphasizing its relevance for forecast validation and continuous improvement of numerical models.

In addition, we will emphasize the importance of taking into account the different precipitation thresholds in the analysis of model performance. The contingency table is used to quantify forecast errors, identify systematic patterns of under- or over-forecasting, and evaluate the ability of models to capture extreme precipitation events.

We also show an example of verification results obtained during September to December, 2022 period. These results were compared against ARPEGE.

2 Data and methods

2.1 Data

The observations used in our study were extracted from the main SYNOP messages, specifically from section 1 of group 6RRRtR. Since this group is coded differently according to each National Meteorological and Hydrological Services as shown in the Table 1, we have perform arithmetic operations to obtain a cumulative precipitation over a 6-hour period, which is the frequency used in this study to evaluate our models.

The verification domains for ALADIN and AROME and the observation sites used are shown in figure 1.

Table 1: Coding of group 6 in SYNOP .

	Europe	**Algeria & Morocco**	**Tunisia**
0600 Z	**tR=2** Rainfall totals accumulated 12h	**tR=4** Rainfall totals accumulated 24h	**tR=4** Rainfall totals accumulated 24h
1200 Z	**tR=1** Rainfall totals accumulated 6h	**tR=1** Rainfall totals accumulated 6h	**tR=1** Rainfall totals accumulated 6h
1800 Z	**tR=2** Rainfall totals accumulated 12h	**tR=2** Rainfall totals accumulated 12h	**tR=2** Rainfall totals accumulated 12h
0000 Z	**tR=1** Rainfall totals accumulated 6h	**tR=1** Rainfall totals accumulated 6h	**tR=3** Rainfall totals accumulated 18h

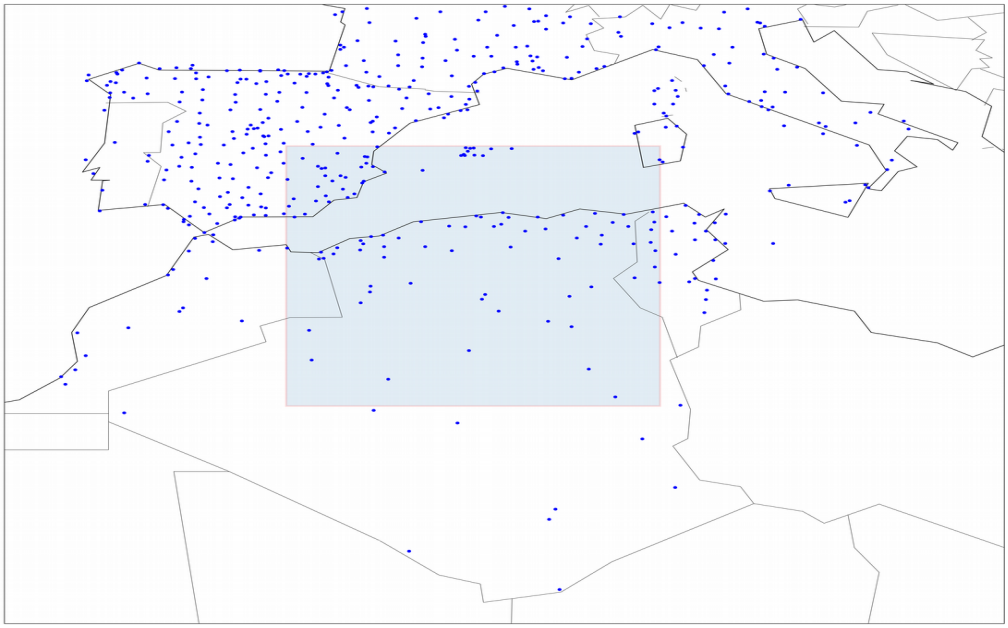

Figure 1 : Observations used for model checking.

2.2 Methods:

A contingency table is a table that summarizes the results of a classification based on two variables.

In meteorology, a contingency table can be used to compare weather forecasts to actual observations. It is often used to evaluate the quality of weather forecasts and to calculate performance measures such as accuracy, false positive rate and false negative rate.

Table 2 : Example of a contingency table for comparing rain forecasts with observations.

	Positive Observations	**Negative** Observations
Positive Forecast	True positive (TP)	False positive (FP)
Negative Forecast	False negative (FN)	True negative (TN)

In this table, positive forecasts indicate that rain was forecast, while negative forecasts indicate that rain was not forecast. Positive observations indicate that rain was observed, while negative observations indicate that no rain was observed. The TP, FP, FN, and TN values represent the numbers of true positive, false positive, false negative, and true negative results, respectively.

From this contingency table, performance measures can be calculated, such as:

- False alarm rate (FAR) = $\frac{FP}{(FP+TP)}$ (1)

- Probability of detection (POD) = $\frac{TP}{(TP+FN)}$ (2)

- Precision = $\frac{TP}{(TP+FP)}$ (3)

- Accuracy = $\frac{(TP+VN)}{(TP+FP+TP+TN)}$ (4)

These measurements allow us to evaluate the quality of rain forecasts in comparison with real observations in several thresholds of rainfall (0.2mm , 0.5mm , 1mm , 5mm , 10mm , 20mm).

When high-resolution forecasts from numerical models are evaluated using traditional measurements, they often show poor results due to difficulties in obtaining an exact match with high-resolution observations. In this study, we propose to use a "fuzzy" verification approach (Ebert, 2008) that relies on the use of a spatial window or neighborhood surrounding the forecast points over a radius of 50km . This fuzzy approach better accounts for spatial variability and overcomes the limitations of traditional measurements in the evaluation of high-resolution numerical model forecasts.

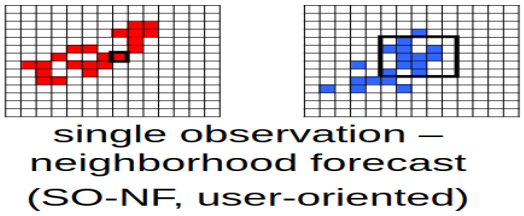

Figure 2 : Single observation neighborhood forecast.

3 Results

In this section, we will present some examples of the results obtained, highlighting the contingency tables and the evolution of several metrics such as detection probability, false alarm rate, precision and accuracy. Figure 3 shows verification scores of six-hour accumulated precipitations simulated by ALADIN and AROME for 0.2mm, 5mm and 20mm threshold values. These scores are compared with ARPEGE.

While the metrics are almost similar between ALADIN and ARPEGE, AROME stands out with better accuracy and precision for the 0.2 mm threshold and then joins the other models for the 5 mm threshold.

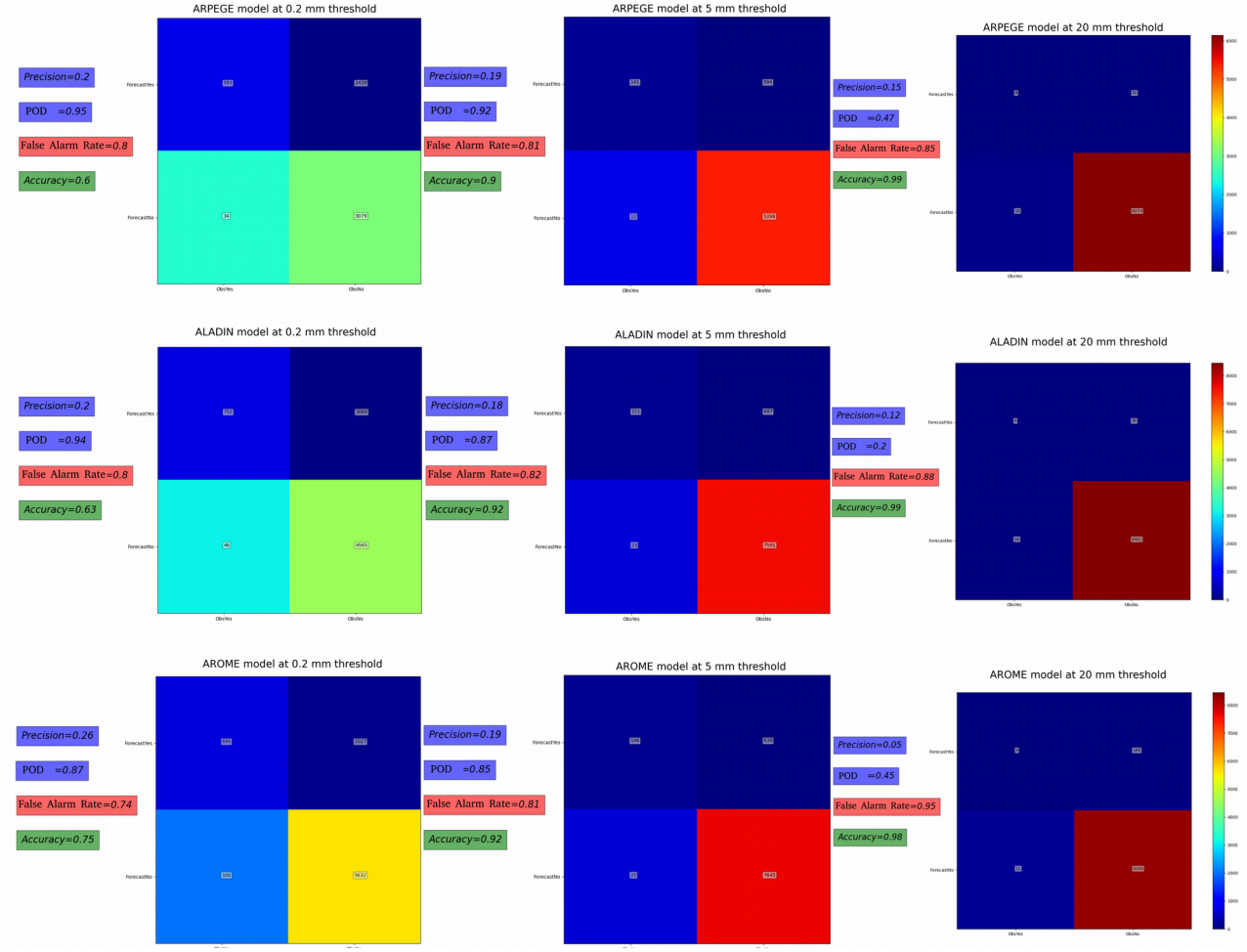

Figure 3 : Contingency table of six-hour accumulated precipitations simulated by ALADIN and AROME for 0.2mm, 5mm and 20mm threshold values.

Figure 4 shows the evolution by threshold of the probability of detection, false alarm rate, precision and accuracy for six-hour accumulated precipitations.
This graph provides the same information as the contingency tables in the previous figure, but with more details. It confirms the similar behaviour between ARPEGE and ALADIN, while highlighting the favourable performance scores for AROME for thresholds below 5 mm.

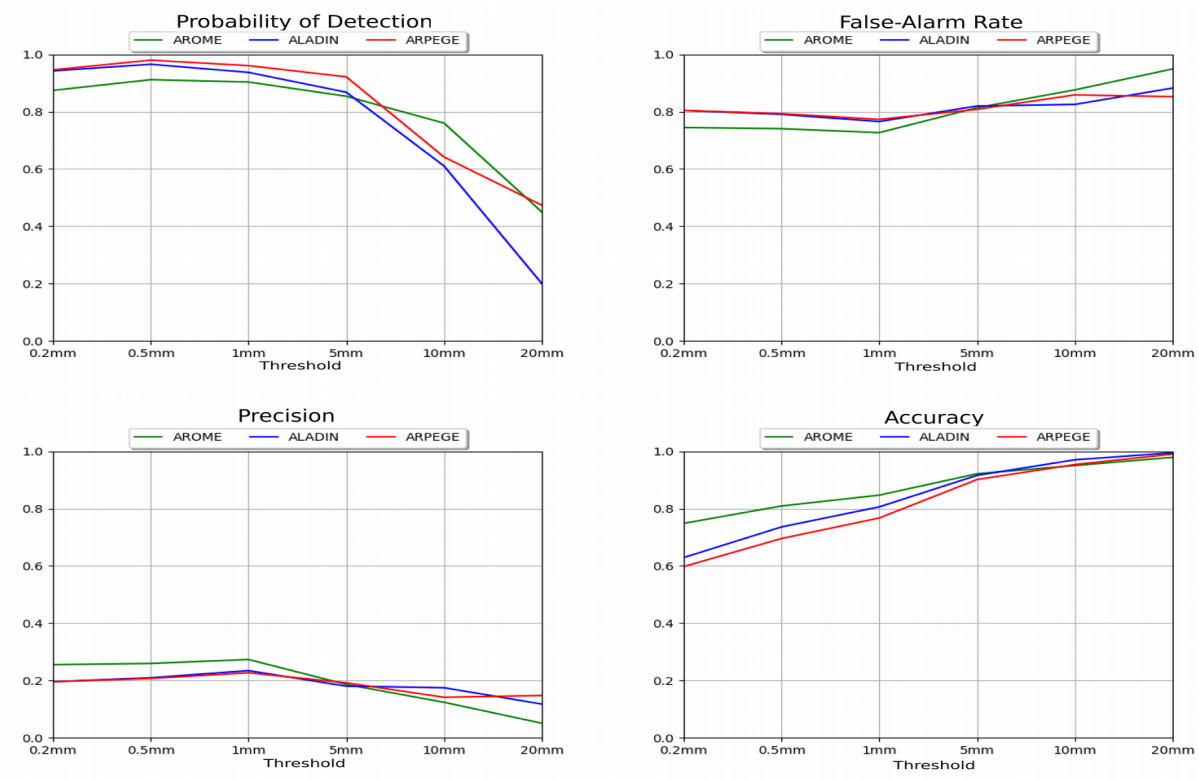

Figure 4 : Evolution by threshold of the probability of detection, false alarm rate, precision and accuracy for six-hour accumulated precipitations.

In figure 5 we will see the evolution by forecast time of the probability of detection, the false alarm rate, the precision and the accuracy .

No significant trend is observed in the evolution of the scores, indicating a stable behaviour where the models produce almost similar results for the different metrics.

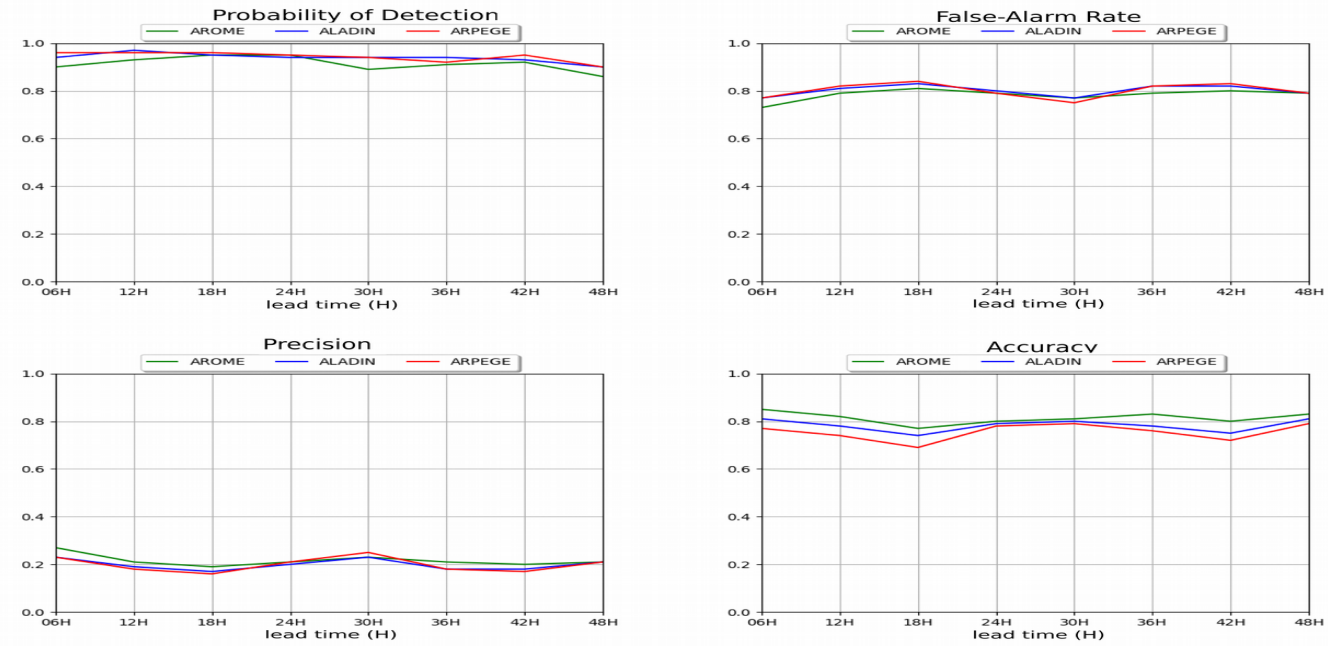

Figure 5 : Evolution of the probability of detection, the false alarm rate, the precision and the accuracy.

Figure 6 shows the monthly evolution of the probability of detection, false-alarm, precision and accuracy for six-hour accumulated precipitations.
What is particularly notable in this figure is that all of the scores (POD, FAR, Precision and accuracy) show a very pronounced peak in the month of November. This result makes sense, as that month was characterized by a period of heavy rainfall.

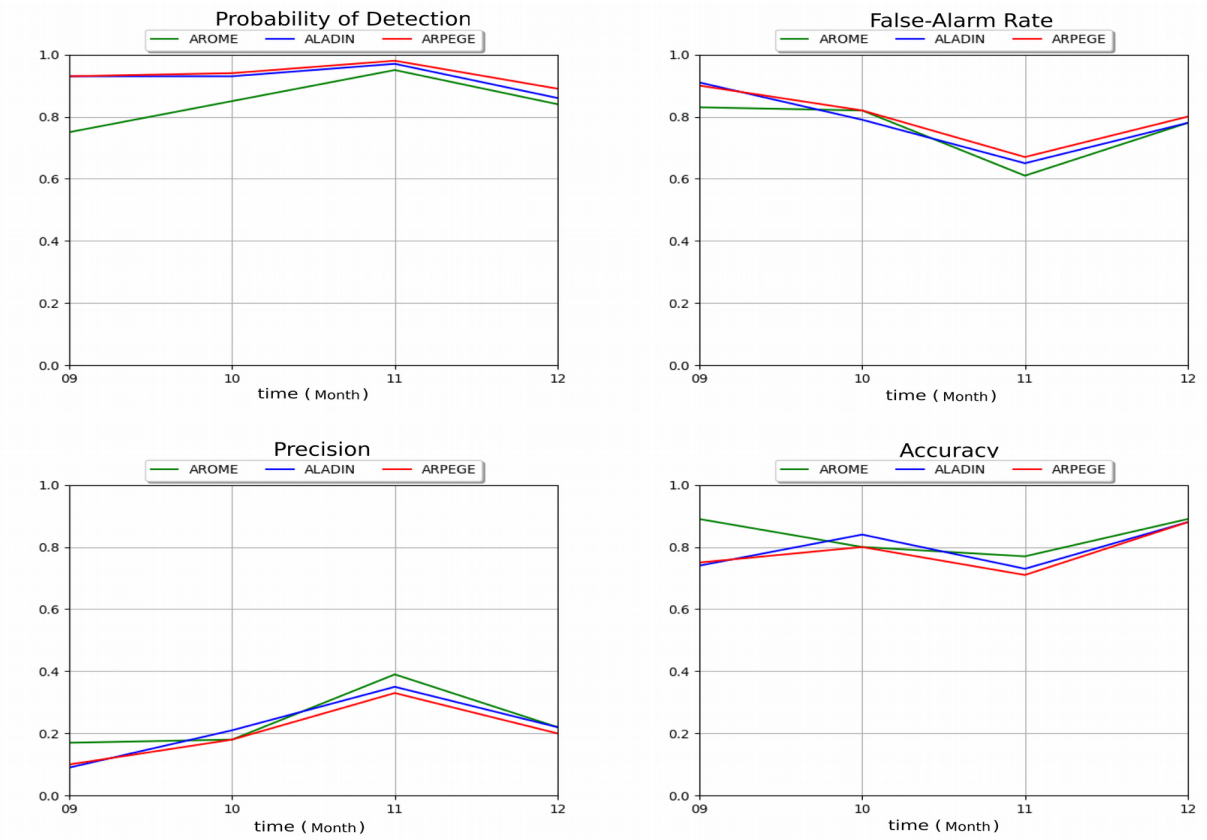

Figure 6 : Monthly evolution of the probability of detection, false-alarm, precision and accuracy for six-hour accumulated precipitations.

4 Conclusion

This paper presents the first results of verification of six-hour accumulated precipitation simulated by ALADIN and AROME over Algeria against observation using a contingency table approach. The results obtained over September to December of 2022 period showed interesting trends, highlighting the improvement of model accuracy over time.
Programs developed in this work will be deployed in operational to progress our verification chain already in place for surface parameters (temperature, wind, humidity and pressure).
However, improvement perspectives remain to be explored. Among them, it would be relevant to introduce other extreme parameters such as wind gusts, in order to better understand the behaviour of our models such AROME.

5 References

Baig, F., Abrar, M., Chen, H., & Sherif, M. (2023), Evaluation of Precipitation Estimates from Remote Sensing and Artificial Neural Network Based Products (PERSIANN) Family in an Arid Region. Remote Sensing, 15(4), 1078. https://doi.org/10.3390/rs15041078.

Ebert, E. E. (2008), Fuzzy verification of high-resolution gridded forecasts : a review and proposed framework. Meteorological Applications, 15(1), 51-64. https://doi.org/10.1002/met.25.

Giarno, G., Hadi, M. N. S., Suprayogi, S., & Murti, S. H. (2018), Distribution of Accuracy of TRMM Daily Rainfall in Makassar Strait. Forum Geografi, 32(1), 38-52. https://doi.org/10.23917/forgeo.v32i1.5774.

Gold, S., White, E. B., Roeder, W. P., McAleenan, M., Schubert, C. M., & Ahner, D. K. (2019), Probabilistic Contingency Tables : An Improvement to Verify Probability Forecasts. Weather and Forecasting, 35(2), 609-621. https://doi.org/10.1175/waf-d-19-0116.1.

Montesarchio, V., Napolitano, F., Rianna, M., Ridolfi, E., Russo, F., & Sebastianelli, S. (2015), Comparison of methodologies for flood rainfall thresholds estimation. Natural Hazards. https://doi.org/10.1007/s11069-014-1357-3.

Recent NWP activities in Romania

Simona Taşcu, Alina Dumitru, Alexandra Crăciun, Mirela Pietrişi, Mihaela Neacşu, Meda Andrei

1. Introduction

This is a summary of some activities and results obtained recently within the ACCORD group in Romania. The national operational setup is based on two parallel configurations, which differ in horizontal resolution and physical packages: ALARO 0 baseline with 6.5 km resolution and ALARO 1 with 4 km resolution. They both use the model version cy43t2 and use lateral boundary conditions files from the global model ARPEGE. In the last year, longer period integrations were performed also for experiments running in higher resolution (2.5 and 2 km). Here, we show verification results for operational and experimental configurations.

2. Testing ALARO with different resolutions

1. Subjective verification

Results of a subjective evaluation of the precipitation forecast for two cases are presented. The cases were selected to show the impact of the change in horizontal resolution for the most challenging parameter. Figures 1 and 2 present the 24 hours cumulated precipitation for June 14th and April 28th 2022. The top panels show the results from the highest resolution experiments (2 and 2.5 km) and the observations. The bottom panels present the results from the 4 km, the big domain which is covering the whole Black Sea, the 4 km small domain and 6.5 km simulations.

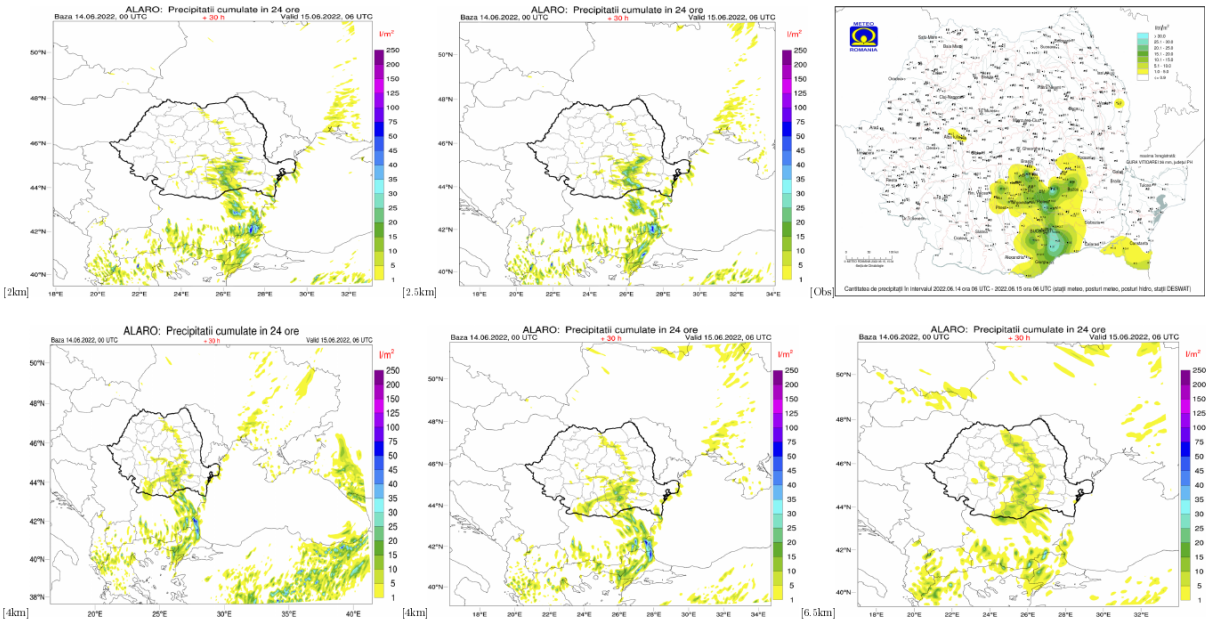

Figure 1: 24 hours cumulated precipitation for 14th June 2022. Top panels show the precipitation forecasts from the highest numerical simulations (2 and 2.5 km) and the observations, while the bottom panels show the 4 km (big and small domain) and 6.5 km results.

For the convective event from 14th June, the benefit of high resolutions models is noticeable. The numerical simulations at 2 and 2.5 km reduced the amounts in the north-eastern and west-southern parts of the country, leading to more realistic precipitation representation. For the 28th April case, it is worth noting that all configurations tend to overestimate the precipitation amounts, more visible in the 4 km numerical simulations.

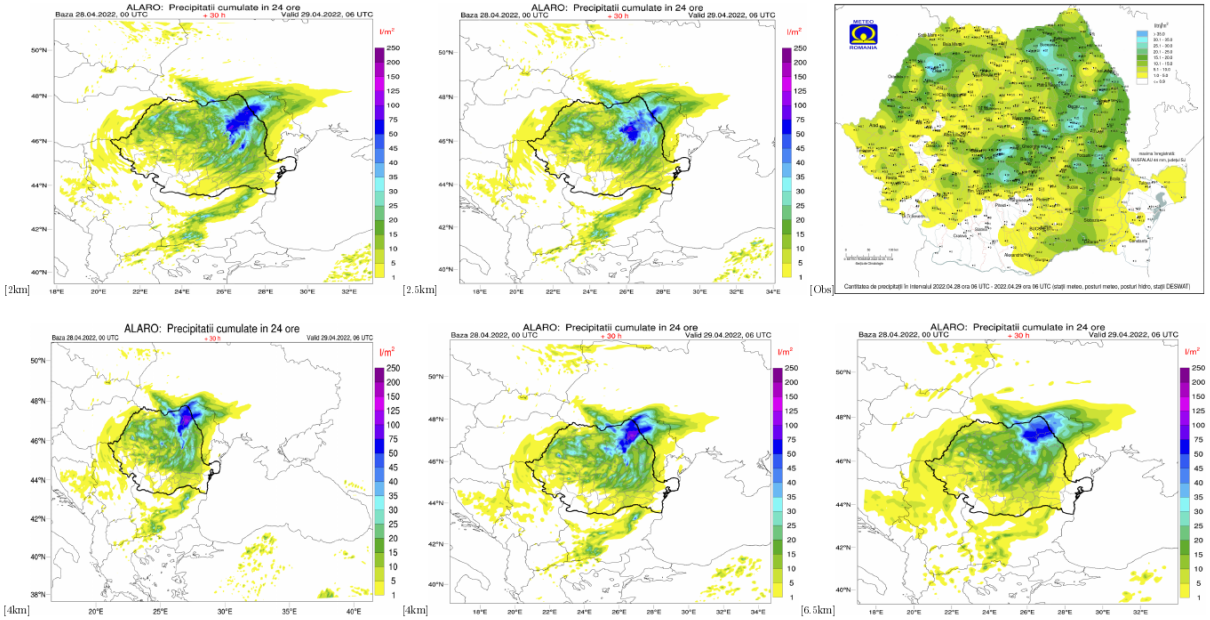

Figure 2: 24 hours cumulated precipitation for 28th April 2022. Top panels show the precipitation forecasts from the highest numerical simulations (2 and 2.5 km) and the observations, while the bottom panels show the 4 km (big and small domain) and 6.5 km results.

2. Objective verification

In order to analyse the ALARO model performance, an objective verification of three different configurations with 6.5 km, 4 km (big domain) and 2.5 km horizontal resolutions was done. The model was run for a period of 5 months starting from 01.04.2022 to 31.08.2022, for 00 UTC runs. The forecast length is up to 78 hours for 6.5 km and 4 km configurations and up to 30 hours for 2.5 km configuration. The traditional scores (as BIAS and RMSE) were computed by using 157 synop stations over Romania (Figure 3) for three meteorological fields: 2 m temperature (T2m), mean sea level pressure (MSLP) and 10 m wind speed (W10m).

In Figure 4, the BIAS and RMSE scores for temperature at 2 m are shown. It can be noticed in general better performance of ALARO at 6.5 km configuration in terms of both scores. For ALARO at 4 km resolution it can be seen there are larger values of the RMSE score. Regarding ALARO at 2.5 km, starting from 12 - 24 hours, the RMSE score is better than the other configurations. For MSLP, the results are presented in Figure 5. A slight improvement is visible for the highest resolution numerical simulation (2.5 km) for BIAS scores in the first interval. Similar results are obtained for RMSE score for all configurations. Figure 6 shows the BIAS and RMSE scores for W10m. It is worth mentioning the negative values of BIAS for model configurations at 6.5 km and 4 km and the positive values for ALARO at 2.5 km horizontal resolution for the interval between 00 to 06 hours. In general, ALARO at 6.5 km has a positive BIAS, while ALARO at 4 km has predominant negative values. RMSE shows the similar performance for all three configurations.

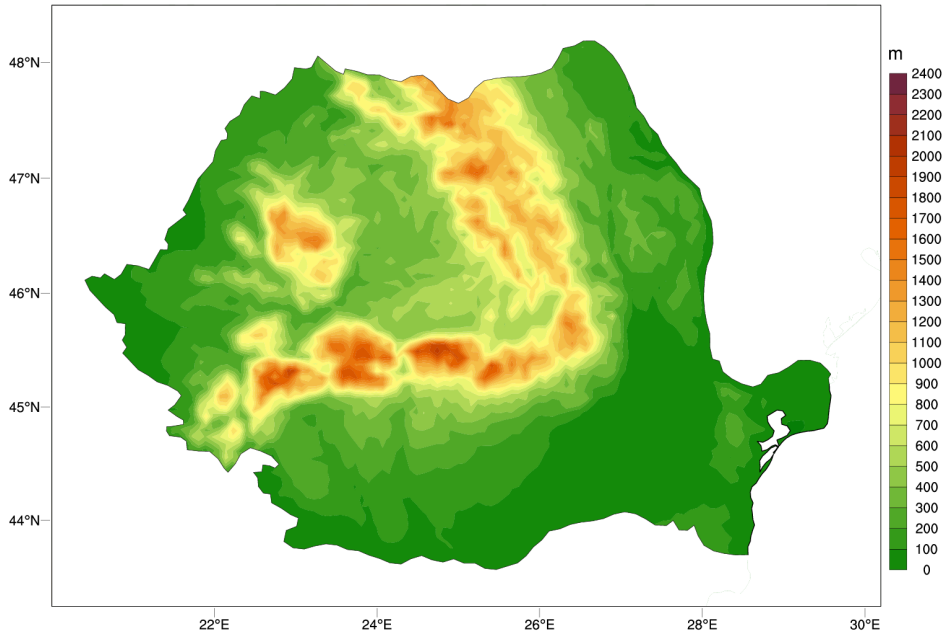

Figure 3: The operational ALARO orography at 6.5 km horizontal resolution (only Romanian territory).

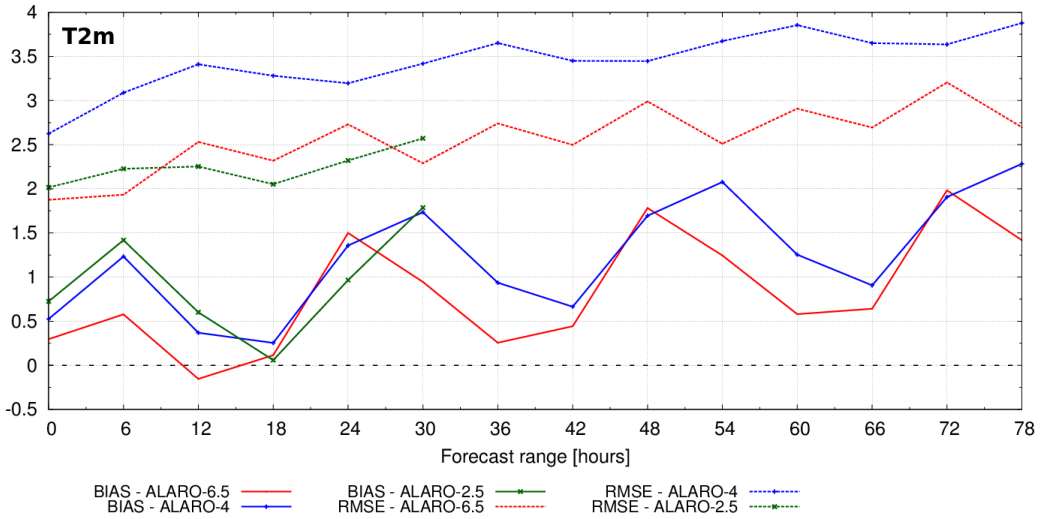

Figure 4: 2 m temperature: BIAS and RMSE for ALARO model at different horizontal resolutions (6.5 km, 4 km and 2.5 km) for 5 months verification period from 01.04.2022 - 31.08.2022, 00 UTC runs.

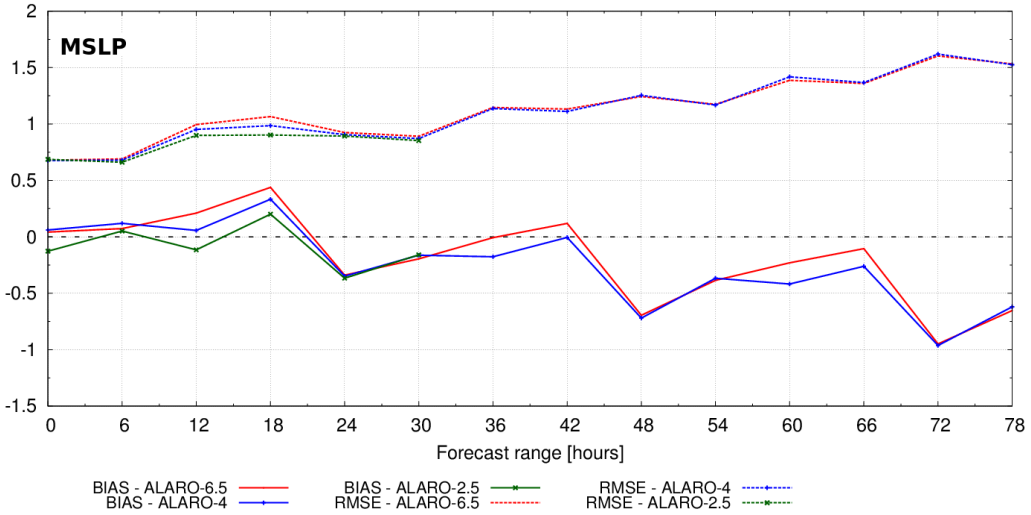

Figure 5: MSLP: BIAS and RMSE for ALARO model at different horizontal resolutions (6.5 km, 4 km and 2.5 km) for 5 months verification period from 01.04.2022 - 31.08.2022, 00 UTC runs.

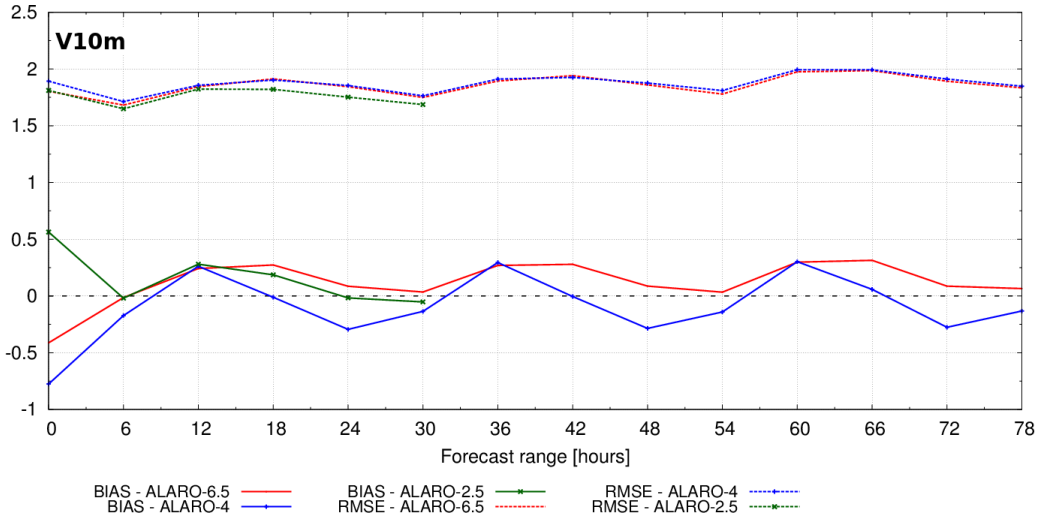

Figure 6: 10 m wind speed: BIAS and RMSE for ALARO model at different horizontal resolutions (6.5 km, 4 km and 2.5 km) for 5 months verification period from 01.04.2022 - 31.08.2022, 00 UTC runs.

3. Feedback on ALARO performance, 6.5 km configuration

Since an important user group of the ALARO products is represented by the local forecasters, it is useful to participate in regular meetings where such type of feedback is gathered and discussed. Forecasters reported several features that they encountered related to their experience with the ALARO at 6.5 km forecast, such as:

- in most cases, 2 m temperature is overestimated in the southern and eastern parts of the country, while in the inter-Carpathian areas it is usually underestimated. Figure 7 shows the BIAS and RMSE scores for 2 m temperature for the month of April 2023 for two regions in Romania: Dobrogea (all stations in the south - eastern part of the country) and Transilvania

(all stations in the central part of the country, situated between the eastern, southern and western parts of the Carpathians).

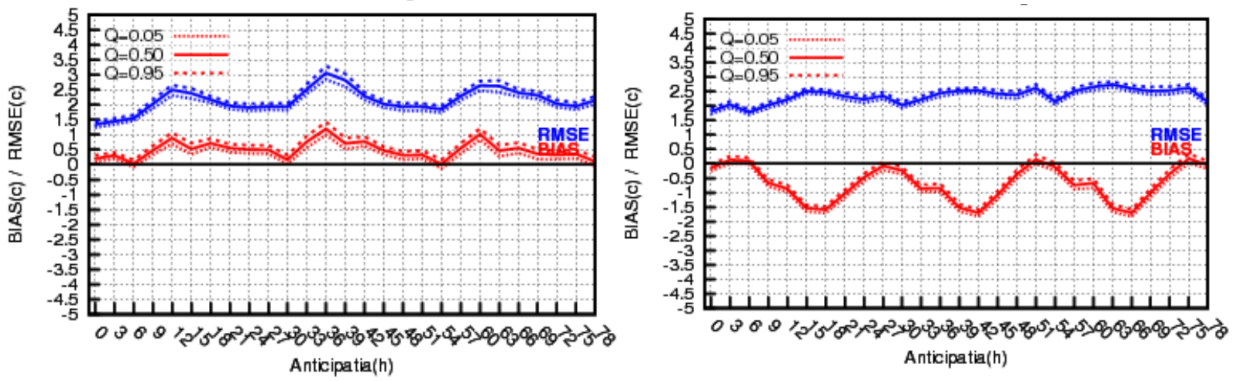

Figure 7: Bias and RMSE scores for 2 m temperature for 2 regions: Dobrogea (left) and Transilvania (right), for April 2022.

- a similar pattern was observed for 10 m wind speed; biases in the southern and eastern parts are usually larger than those in the inter-Carpathian regions. Such behaviour can be observed in figures 8 and 9 that show boxplots for simulated and observed wind speed for the month of July, 2022 for two stations that are located in these regions.

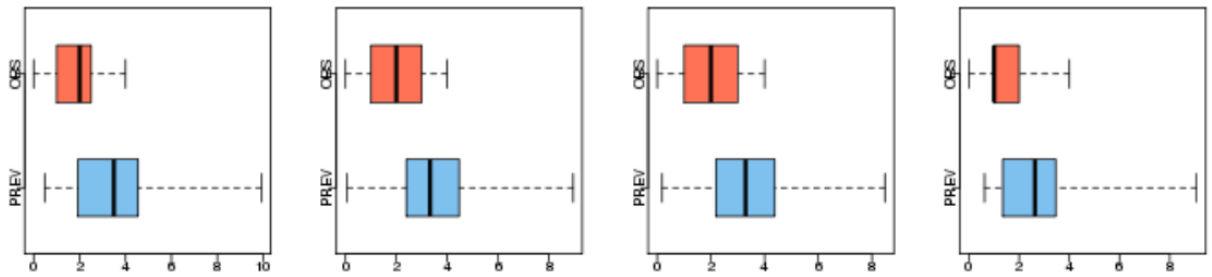

Figure 8: Wind speed (m/s) boxplots: observed (red) and forecast (blue) for station Constanța, for July 2022.

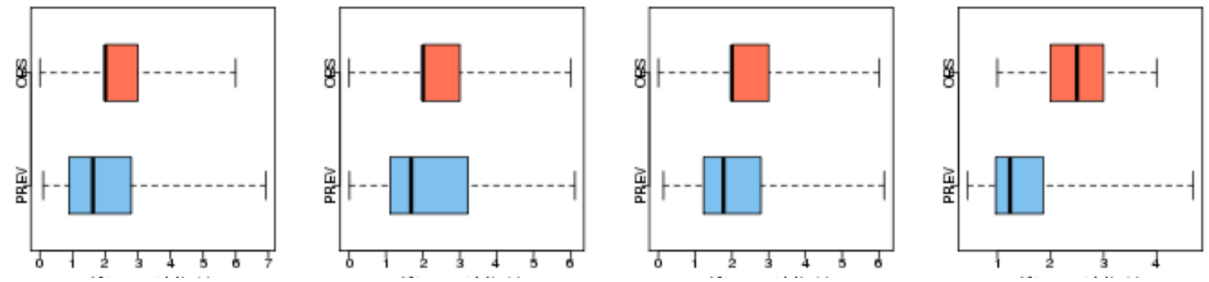

Figure 9: Wind speed (m/s) boxplots: observed (red) and forecast (blue) for station Cluj, for July 2022.

- for such verification purposes, it is important to take into consideration when performing the statistical validation that some part of the information can be lost when scores are averaged on the whole region of interest compared to results obtained when the verification is done on smaller regions. An example of this can be observed in Figure 10 where the BIAS and RMSE for 10 m wind speed are shown.

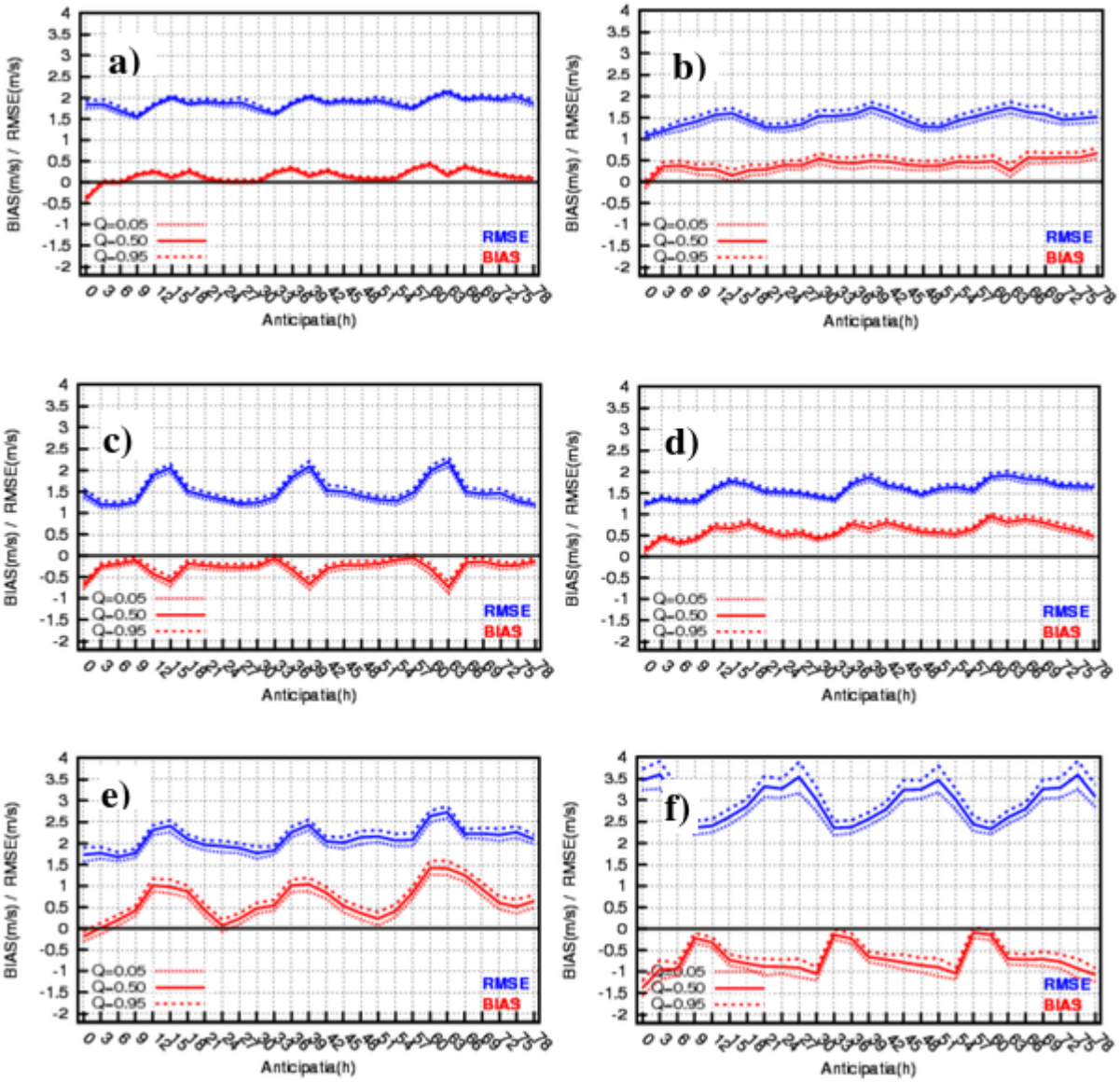

Figure 10: Bias and RMSE for 10 m wind speed for meteorological stations in the whole country (a) and several regions in Romania: Banat (b), Transilvania (c), Muntenia (d), Dobrogea (e), mountain stations (f).

We can see the overall pattern is different depending on the considered area, although general values of the scores computed for the whole country indicated only very small BIAS for 10 m wind speed. In this case, the forecaster's input is significant in order to have a more detailed overview regarding the model performance.

4. MOS for ALARO at 4 km

Statistical adaptation using the Model Output Statistics (MOS) method was prepared and implemented operationally for the new configuration of ALARO 4 km, which has been running in parallel since last year. This technique is applied for 166 meteorological stations; the training period used in the development is 2018 – 2021. The new forecast is obtained for several meteorological parameters: 2 m/maximum/minimum temperature, 10 m wind speed and direction, cloudiness, 6h cumulated

precipitation and is applied daily for 00 and 12 UTC runs. The results are disseminated on the *intranet* webpage.

Statistical verification of the forecast obtained with the MOS technique was performed. BIAS and RMSE are shown in Figure 11 in comparison with the persistence method for 2 m temperature, extreme temperature, 10 m wind speed and direction for the month of January 2023. While for the bias we can see some daily variations compared to the persistence method, it can be noticed that a significant decrease in the RMSE scores is obtained with the MOS forecast.

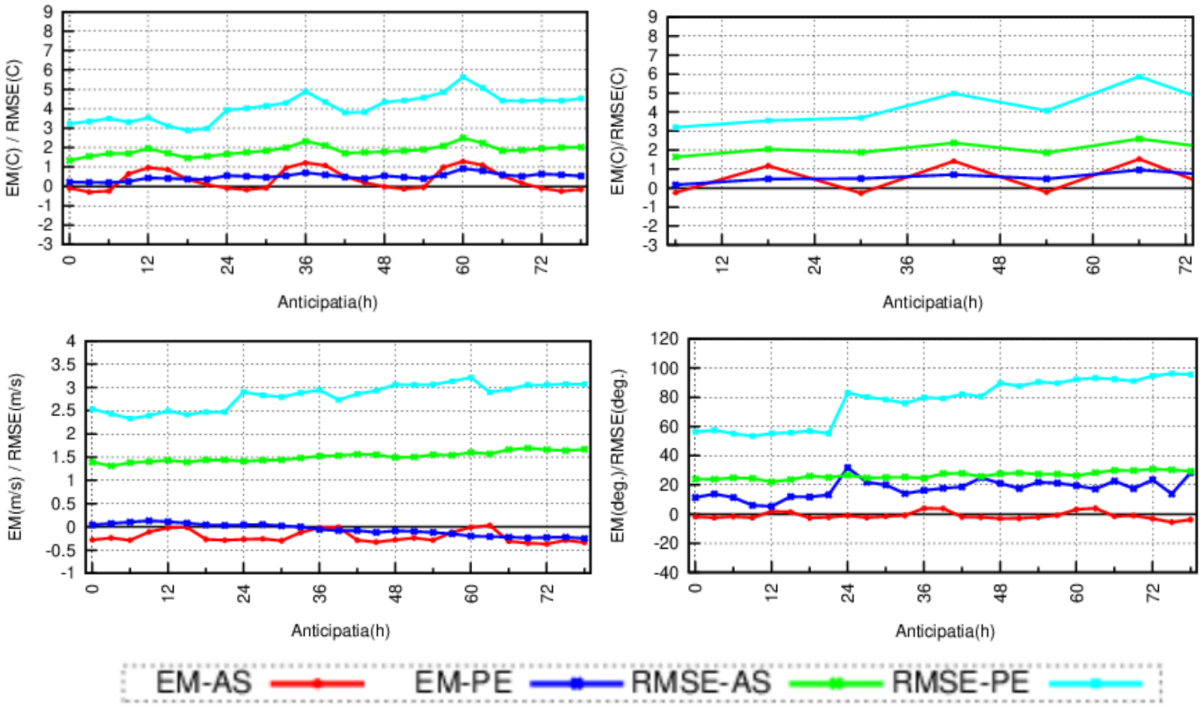

Figure 11: Bias and RMSE for 2 m temperature (upper panel - left), extreme temperature (upper panel - right), 10 m wind speed (lower panel - left) and direction (lower panel - right), for the MOS forecast (AS) and the persistence method (PE), January 2023.

Assimilation of Microwave Link Signal Attenuation as Precipitation Rates using Machine Learning, 1D-VAR and 3D-VAR

Phillip Scheffknecht and Christoph Wittmann

1. Introduction

In the framework of the LINK Project, GeoSphere, together with its project partners Drei Hutchison Austria GmbH and FH St. Pölten, has been exploring the potential of microwave links between cell phone towers for the generation of precipitation analyses and assimilation into AROME. These observations are widely available over Austria, with several thousand connections between cell phone towers reporting data every 15 minutes. This makes the observations interesting for state of the art forecasting models as cycling intervals keep getting shorter and resolutions keep increasing.

2. Methodology

1. Conversion of raw data into precipitation rates

The microwave links are mostly bi-directional radio links, that is, they consist of two emitters and receivers, one of each at both ends of the link, and the hardware records the emitted and received signal strength. These data can be obtained with relatively short delay (less than 5 minutes, if needed). Over Austria, there are about 5000 such links in operation. They have frequencies varying between 13 and 80 GHz, with the majority of about 75% being 80 GHz link. This fraction is expected to increase in the future, as high frequency links and their higher bandwidth will be needed to satisfy the increasing demand for rapid data transfers.

There have been multiple attempts to use such microwave links in the generation of precipitation analyses (Overeem et al, 2021; Leijnse et al., 2010, 2007). They all only used links with frequencies up to 40 GHz, with some just using a single frequency. The novel approach of our project partner FH St. Pölten uses machine learning to convert the signal attenuation data into precipitation rates.

A first attempt was to extract only qualitative information, i.e. whether rain was observed or not, and this was used to generate 100% relative humidity (RH) observations at the midpoint of each link. The second method converted the signal attenuation data into precipitation link. The machine learning algorithms used the geographic location, frequency, altitude, and attenuation data from previous time steps and were trained to produces precipitation rates as seen in the INCA analysis at the same time and location.

This approach allowed to make use of the 80 GHz links, which are also impacted by fog due to their short wave lengths. Machine learning was able to differentiate between rain and fog with sufficient accuracy to make use of these observations in the model. However, the test period was a summer month, when fog is rare, but low clouds might sometimes be an issue instead. A test during a winter period has not been performed.

2. 1D-VAR + 3D-VAR

Combining 1D-Var and 3D or 4D-Var isn't a new approach in and of itself (see, e.g., Sahlaoui et al., 2019; Caumont et al., 2010; Lopez and Bauer, 2007). However, using microwave link data as a source of precipitation rates as input for the assimilation is a novel approach. The 1D-Var + 3D-Var approach for these experiments consists of several steps that handle the conversion from raw hourly precipitation

rates generated via neural network into OBSOUL files usable by AROME. The 1D-Var binary (Lopez et al., 2007) also needs information on the B-matrix of the model, namely correlations of *T* and *q* between the different model levels, and standard deviations for those variables. These are obtained from the current operational B-matrix for AROME-Aut 2.5 km. It also uses the A and B coefficients for the definition of the hybrid model levels. The following steps are performed:

1. Simple conversion from ASCII to ASCII into a simple three column `lon lat rr` file, readable by the 1D-Var binary.
2. Extraction of profiles from the first guess at all observation locations for temperature, specific humidity and the tendency for those two variables obtained from the model physics for the respective time step (`ZTENT` and `ZTENQ`), for this purpose, the AROME code was modified to write these variables into the history files.
3. Extraction and conversion of the profiles from the first guess into an ASCII file readable by the 1D-Var binary.
4. Run the 1D-Var binary, this uses the correlations from the B-matrix, the observations, and the first guess profiles and returns an ASCII file containing *T* and *q* increments for all model levels and all profiles, as well as quality flag to indicate whether the 1D-Var converged successfully.
5. Read the ASCII file generated by the 1D-Var and the first guess profiles and add the increments to the first guess values, then convert the profiles of *T* and *q* into profiles of *RH*.
6. Convert the *RH*-profiles into OBSOUL format using observation type 5 (TEMP) and code 135 (dropsonde). Dropsondes are an extremely rare observation code and are generally absent, which allows to easily keep track of the new observations for later analysis. However, this also comes with the downside of no thinning.

Analogous to other steps of the operational AROME-Aut setup, these steps are combined into a single job file which is submitted and executed before the 3D-Var assimilation is started. All steps take about 2-5 minutes, depending on the number of observations for the respective initialization time.

3. Simulation Period and Cycling Setup

For the final test of the microwave link data, AROME was run for one month from 1 to 30 June 2022 with two experiments (runs) being performed. Below, the two simulations will be referred to as *reference* and *link* for the runs without and with microwave link data, respectively.
The model was initialized on 31 May 2022 and run for 24 hours with 3-hourly cycling until 21 UTC. The 3-hour forecast from this run was used as first guess for both experiments and starting at 1 Jun 2022 00 UTC the model was cycled 3-hourly until 30 Jun 2022 21 UTC. Only two of the eight runs, 00 and 12 UTC, provide forecasts up to +24 hours, the others were 3 hour runs used only to maintain the cycling between the 00 and 12 UTC runs.

4. Verification Methodology

The two runs were verified against surface stations within the model domain for temperature, humidity, wind, and precipitation. In addition, the runs were verified against available radiosondes within the domain, which are, unfortunately, sparse. There is a 03 UTC sounding available over Austria in Innsbruck, western Austria, as well as 00 and 12 UTC soundings at Wien Hohe Warte, in the east of Austria. Lastly, precipitation was also verified against the INCA precipitation analysis, which provides precipitation on a 1 km grid over Austria and around 0.5° around the country. The verification of T, q, and 10 m wind speed is limited to the first 20 days, as all data was not yet available at the time of writing.

Table 1: Verification parameters and corresponding sources of observational data

Variable	**Observations**
Temperature (1 to 19 June)	Surface stations, radiosondes
Humidity (1 to 19 June)	Surface stations, radiosondes
Wind (1 to 19 June)	Surface stations, radiosondes
Precipitation	Surface stations, INCA analysis

3. Results

To illustrate the distribution of the data over Austria, Fig. 1 shows a slice of western Austria with microwave links data as points on a map. For this initialization time, all observations were located over the west of Austria, because no dry points are assimilated, and the rain was limited over this region. Figure 1 shows that, while most points have a positive analysis increment, some points show a drying of the profile. The figure shows data from an earlier test using the same data and mainly serves illustrative purposes.

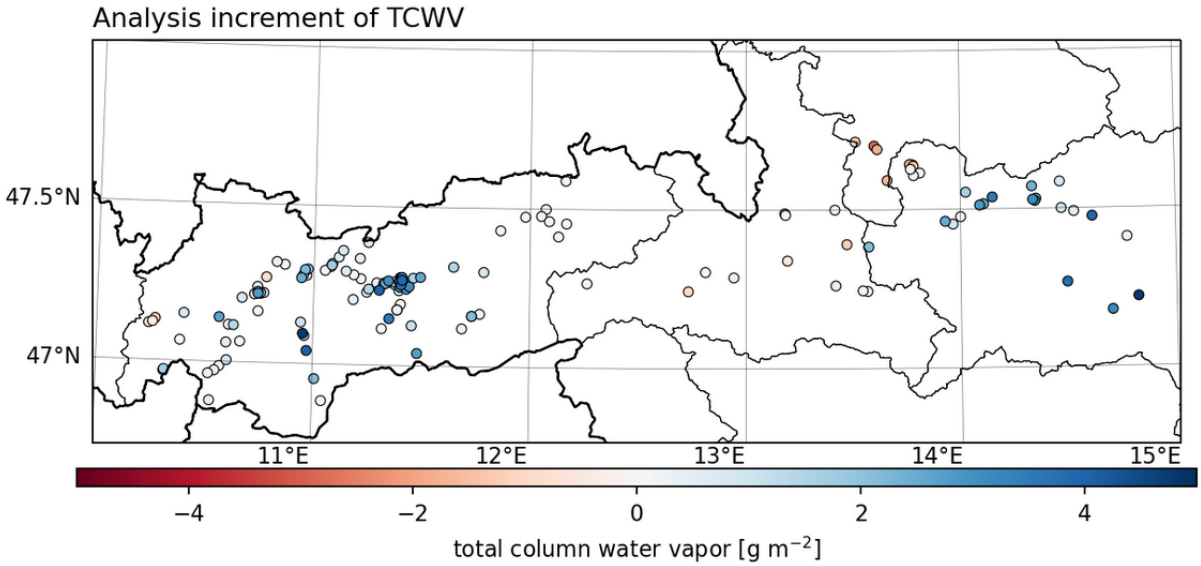

Figure 1: Analysis increments for 2021-07-17 04 UTC showing the positions of the successfully converged 1D-Var profiles. The dots are coloured by the increment of total column water vapor (TCWV).

While the overall impact of link observations is largely mixed and neutral, there are some systematic differences between the two runs, which will be shown below. Firstly, the bias of 2 m temperature and relative humidity is visibly reduced over the first 6 to 7 hours of run time for the 12 UTC simulations while the wind speed bias is about 0.2 m s^{-2} worse during the first 3 hours (Fig. 2). For the 00 UTC simulations, the link observations have little impact during the first 4 – 6 hours but then result in a larger bias for *T* and *q* later during the simulations. However, wind speed improves in the 00 UTC simulations between 5 and 12 hours after the initialization. The overall differences are small at around 0.1 °C, 1% RH, and less than 0.1 m s^{-2} for wind speed.

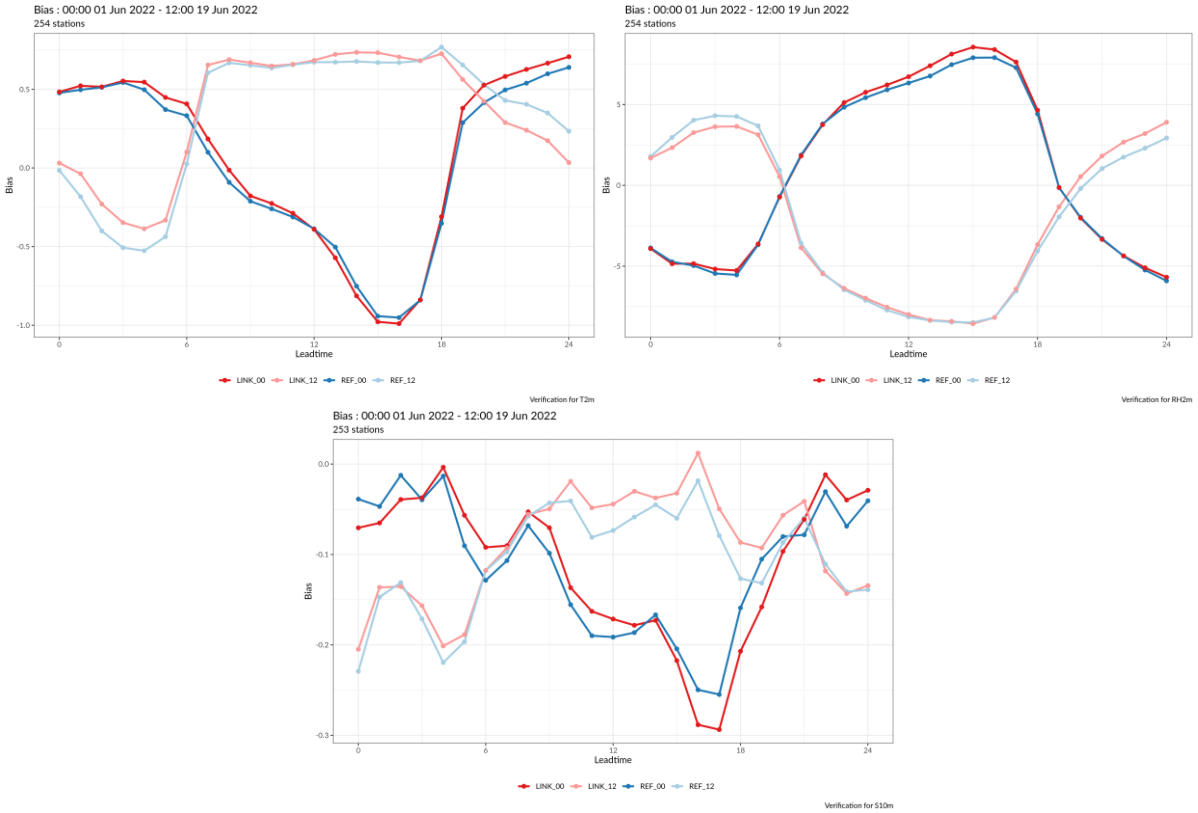

Figure 2: Bias by model lead time for the 00 (dark) and 12 UTC (light) reference (blue) and link (red) runs for T, RH, and 10 m wind speed.

For the mean absolute error, the differences are even smaller than for the bias, and the improvement or deterioration shows a less clear picture. Where the bias improved for the 12 UTC runs during the first 6 – 7 hours of the model runs, the MAE is slightly increased. There is a small reduction of the MAE for temperature and wind speed in the 00 UTC runs starting around 7 hours into the simulation, but the values are small and such a delayed impact is hard to attribute solely to the quality of the link observations.

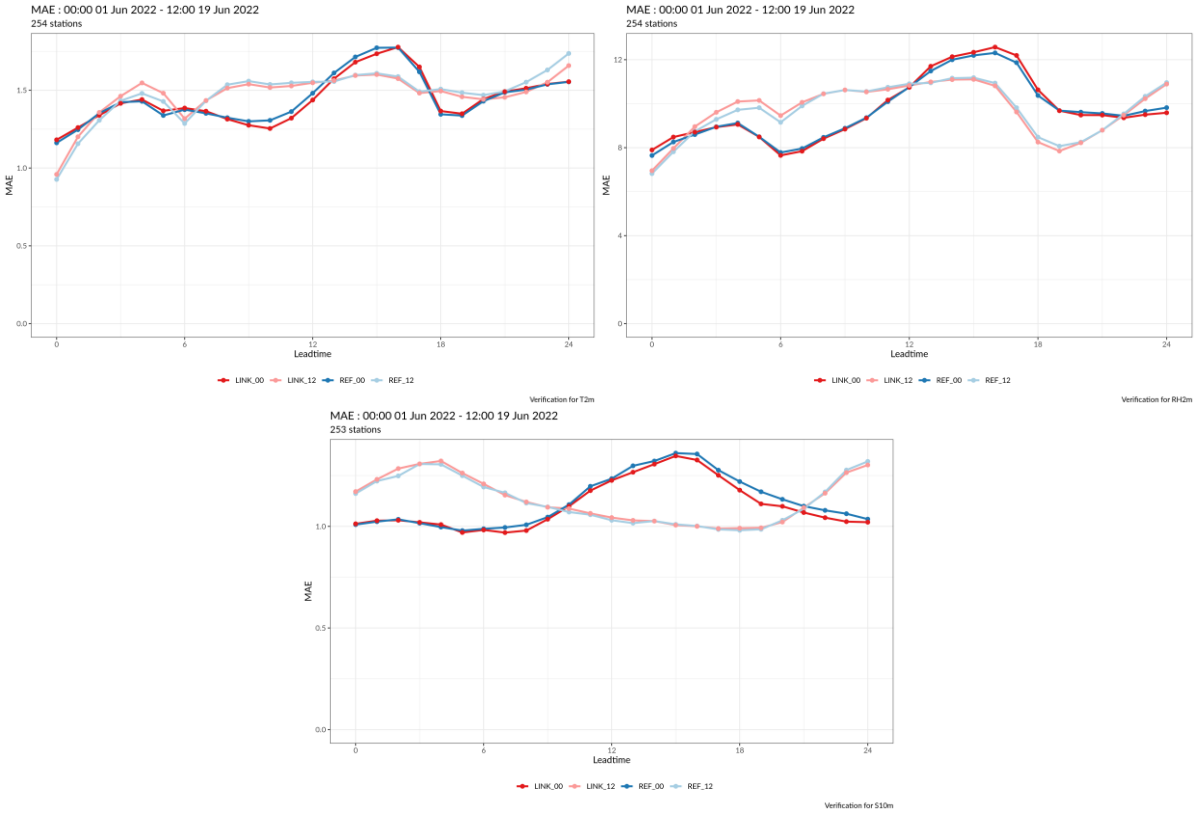

Figure 3: Like Fig. 2, for mean absolute error by model lead time for the 00 (dark) and 12 UTC (light) reference (blue) and link (red) runs for T, RH, and wind speed.

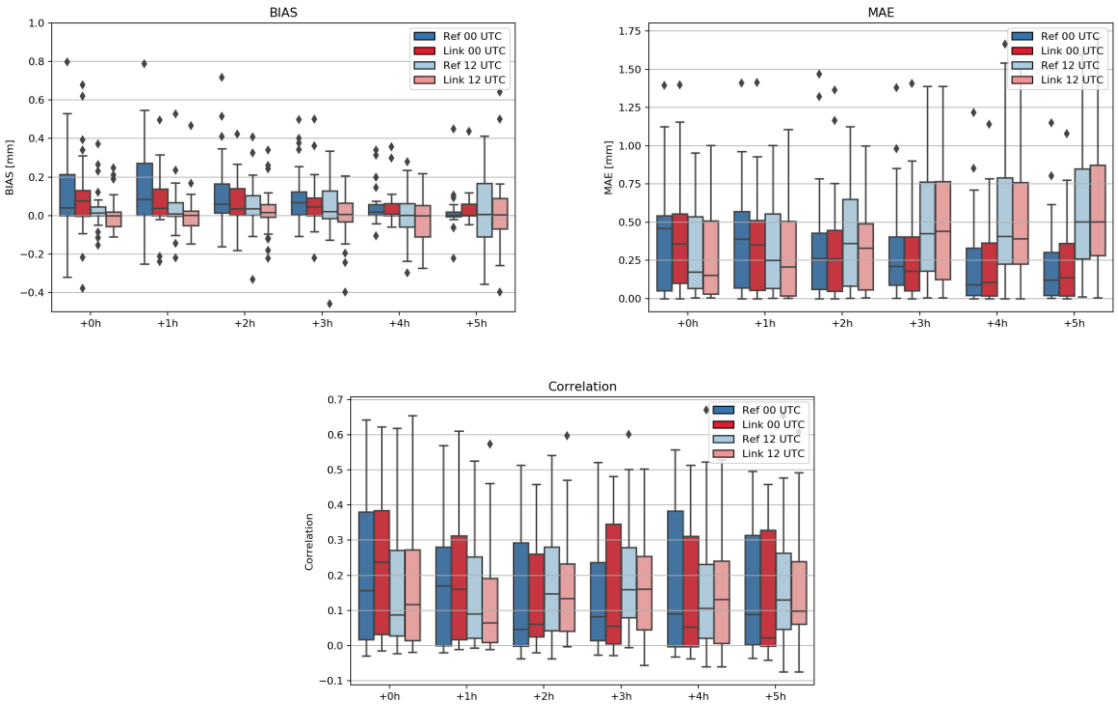

Figure 4: Verification metrics for hourly precipitation verified against the gridded INCA analysis, showing bias, mean absolute error, and correlation coefficient by model lead time for the 00 (dark) and 12 UTC (light) reference (blue) and link (red).

Precipitation was verified in more detail, as the project was focused on flooding events. Figure 4 shows box and whisker charts for the bias, MAE, and correlation coefficient for precipitation, verified for hourly precipitation against the gridded INCA analysis. A reduction in bias during the first hours of the simulations is also visible for precipitation for both, the 00 and 12 UTC runs. This effect is visible until 4 hours after initialization. However, this does not translate into an improvement of the MAE or correlation, as can be seen in Fig 4. Such a reduction in bias can sometimes be linked to an overall reduction in predicted precipitation. Examination of the predicted values (not shown) showed that this might be a contributing factor in this experiment as well.

To illustrate the impact of the link observations on the rain field on an example, Fig. 5 shows panel-based verification from GeoSphere's Panelification tool. This tool is used to verify case studies where multiple model configurations need to be compared. In this case, it shows the reference and link runs for two initialization times for a comparison of four simulations for the three-hour accumulated precipitation from 15 to 18 UTC.
The precipitation fields are similar, which is seen throughout the entire testing period. For this three hour window, the link observations improved the forecast over multiple classical metrics for the 12 UTC runs but worsened the results for the 12 UTC runs. However, the fraction skill score shows that in both cases the link runs provide the better forecast. No clear direction is visible in the change in the other metrics. This is exemplary for the entirety of results, where small improvements and small deteriorations occur to lead to an almost neutral impact overall.

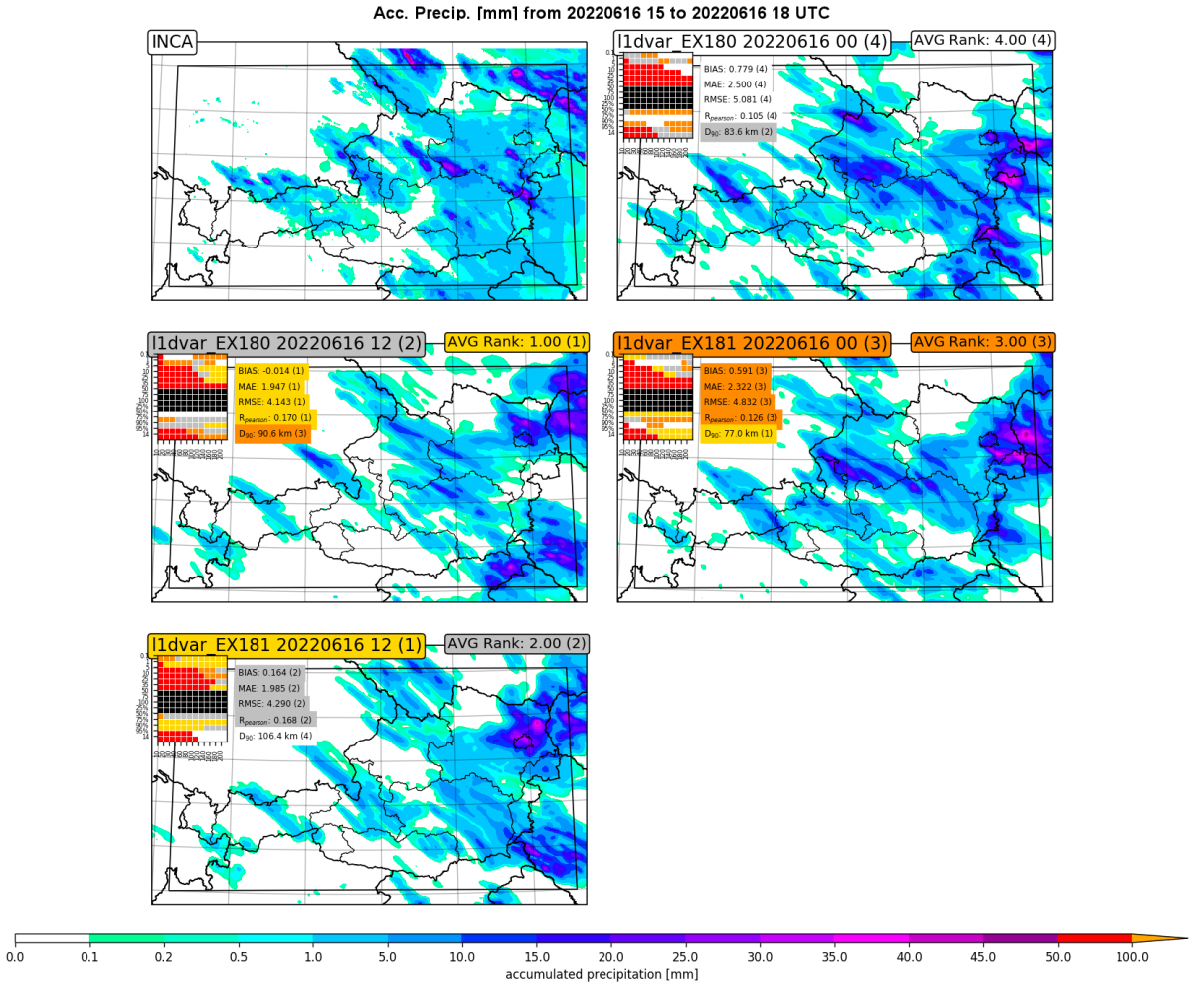

Figure 5: Example of detailed verification panel used to examine precipitation forecasts for individual time periods.

4. Discussion and Conclusion

The experiment showed that microwave link signal attenuation data can be converted into rain rates, which can then be assimilated into AROME using a 1D-VAR + 3D-Var approach. There are some simplifying assumptions involved, which need to be discussed. Firstly, the 2D-Var code was developed for the use with ECMWF's integrated forecasting system (IFS), and thus the physics schemes used within the program are the physics schemes used in IFS at the time of the code development. AROME uses more sophisticated schemes, such that the tendencies obtained from the AROME schemes are not perfectly equivalent to what the 1D-Var would obtain given the same data. However, the schemes are qualitatively similar, in that they are responsible for the same physical processes which result in similar energy conversions. Sahlaoui et al. (2019) demonstrated that this can still lead to an overall positive impact on AROME forecasts.

Secondly, the observations are dense in some regions, and the model does not apply any kind of thinning to dropsondes. While the relative rarity of such observations is an advantage when tracking them, the lack of thinning causes a violation of the implicit assumptions of uncorrelated errors. In practice, this has little effect the convergence of the 3D-Var analysis, and the observations do not cause any problems in the model, despite being numerous with sometimes more than 1000 profiles per assimilation cycle.

Lastly, the link observations were added into a well configured forecasting system, which already receives a wide variety of different observations, ranging from surface stations over radio soundings to aircraft data. For such a system, the impact of a new observation is often small and mostly neutral. It would be interesting to apply the same method over an area where observations are less dense. In such a scenario, the impact could be larger and potentially more beneficial.

In summary, the impact of the link observations over Austria is mostly neutral with a slight improvement in temperature and relative humidity bias, mainly for the 12 UTC runs, which does not translate into an equivalent improvement in any other metric or variable. The assimilation process consists of several steps, which include extraction of data from a first guess, but the overall process is sufficiently fast to be implemented operationally, if desired. The assimilation works well, without causing any technical problems or model instabilities.

5. References

Caumont, O., Ducrocq, V., Wattrelot, É., Jaubert, G. and Pradier-Vabre, S. (2010), 1D+3DVar assimilation of radar reflectivity data: a proof of concept. Tellus A, 62: 173-187. https://doi.org/10.1111/j.1600-0870.2009.00430.x

Jayaram Pudashine, Adrien Guyot, Aart Overeem, Valentijn R.N. Pauwels, Alan Seed, Remko Uijlenhoet, Mahesh Prakash, Jeffrey P. Walker, Rainfall retrieval using commercial microwave links: Effect of sampling strategy on retrieval accuracy. Journal of Hydrology, Volume 603, Part B, 2021, 126909, https://doi.org/10.1016/j.jhydrol.2021.126909.

Leijnse, H., Uijlenhoet, R., & Stricker, J. N. M. (2007c). Rainfall measurement using radio links from cellular communication networks. *Water Resources Research*, 43, W03201. https://doi.org/10.1029/2006WR005631

Leijnse, H., Uijlenhoet, R., & Berne, A. (2010). Errors and uncertainties in microwave link rainfall estimation explored using drop size measurements and high-resolution radar data. *Journal of Hydrometeorology*, 11, 1330– 1344. https://doi.org/10.1175/2010JHM1243.1

Lopez, P. and Bauer, P. 2007. "1D+4DVAR" assimilation of NCEP Stage-IV radar and gauge hourly precipitation data at ECMWF. *Mon. Wea. Rev.* 135(7), 2506– 2524. doi:10.1175/MWR3409.1.

Overeem, A., *et al* 2021 *Environ. Res. Lett.* **16** 074058

Zahra Sahlaoui, Soumia Mordane, Eric Wattrelot, Jean-François Mahfouf, Improving heavy rainfall forecasts by assimilating surface precipitation in the convective scale model AROME: A case study of the Mediterranean event of November 4, 2017, Meteorological Applications, 10.1002/met.1860, 27, 1, (2019).

Exploring Sub-Hour DA in HARMONIE-AROME

Carlos Geijo, AEMET (Spanish Met Agency), cgeijog@aemet.es

1 Introduction

This letter summarizes the main aspects and results of a set of first experiments carried out with the HARMONIE-AROME LAM-NWP system in the topic of upper-air DA at sub-hourly cycling frequency. The motivation for exploring the performance of this system in this range of DA cycling rates hardly needs to be justified in a context of developing LAM-NWP capacity to manage hazardous and extreme weather sometimes associated to atmospheric processes of intrinsically short predictability timescales. To the knowledge of the author of this article, there is a lack of previous studies with the current HARMONIE-AROME system in this DA sub-hourly range. This situation forces us to make a number of system configuration choices which, in the end, may not result optimal, not even correct, to the intended goal of implementing a practical NWP-NWC sub-hour DA method.

2 Experiments Description

2.1 Generalities

The experiments were carried out with version cy43h2 of the HARMONIE-AROME system on the domain IBERIAxxm_2.5 (2.5Km 800x648 L65). During the evening of the 23rd and early morning of the 24th of March 2022, heavy precipitation swept the area covered by the meteorological radar of Málaga, in the South of Spain. At 23rd midday, an "orange" grade warning was issued by AEMET due to forecasted precipitation up to 80 l/m^2 in 12 hours, which turned out to be fairly accurate. Social impact was modest, mainly restricted to traffic disturbances due to several roads cut-off by flash-floods. This episode is an example of the kind of meteorological situations which NWP-NWC systems can be useful for. The meteo centre responsible for monitoring and management support of this potentially hazardous scenario might have wished to have the capacity to spawn short-range forecasts initialized at a pace synchronized with observations acquired from the meteorological radar(s) located in the area. The information so obtained would have been certainly not contained in the latest forecast available from the EDS (Early Delivery Service) by ECMWF, as its effective DAW, even with the Co-DA extension, still leaves a temporary gap of about 8 hours without observational coverage between deliveries, a meaningful time length when compared with typical life spans of this sort of phenomena.

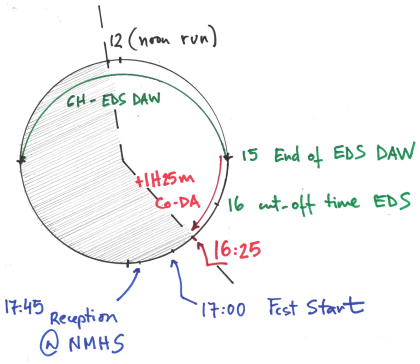

Figure 1: Diagram depicting production time schedule by EDS-ECMWF. The grey zone represents the time interval blind to observations in the time between successive deliveries.

2.2 Twin-Experiments with Synthetic Doppler Wind Radar Observations (DOW)

Given the exploratory nature of these experiments and to facilitate the analysis of the very different issues involved, a "twins-experiment" with synthetic observations was considered the best experiment design option. In this ideal concept framework, observation quality and availability are of no concern as the data are produced at will by means of the observation operators used to generate the "model equivalents" in more realistic DA experiments. In this "twins" configuration, two runs are generated for the time period of interest, one of the runs is used as a proxy of "nature" (the truth), and the other run acts as "model" (the reference or baseline). Differences between the two and impact on these differences caused by whatever changes applied on the baseline that are regarded as worth studying can be easily computed and analysed.

Truth: Cy43 HARMONIE-AROME 18H forecast starting at 12:00 UTC March 23rd. Initial state and LBCs are taken at 6H intervals from the ECMWF operational analyses ("analysis_only" strategy, 4 ELSCF*LBC files, ~2.6 GB). No UA or Surface analyses are done, SURFXINI.fa surfex initialization ("cold start"). Climate, Physics and Dynamics options by default. Output upper-air FA files are archived every 10 minutes from +240 till +1080 minutes (85 ICMSH*+*:000 files, ~75GB). The first 6 hours are disregarded to let the model spin-off the noise introduced by the interpolations necessary to handle the change in resolution between the IFS host and the nested LAM. In order to activate the extended FA header format for time at minute resolution NTIMEFMT=1 in namelist NAMOPH, and this is the same for all the runs in these experiments. Also, time-step is 60 secs in all cases.

Reference: Another 18H forecast starting again at 12:00 UTC March 23rd , but now the initial state and LBCs are taken at 1H intervals from the ECMWF 06UTC March 23rd operational forecast ("same_forecast" strategy, 19 ELSCF*LBC files ~10.2 GB). SURFXINI.fa surfex initialization. No UA or Surface analyses. Climate, Physics and Dynamics options by default. Output upper-air *and surfex* FA files are archived every 10 minutes from +240 till +1080 minutes (85 ICMSH*+*:000 files, ~75GB and 85 ICMSH*+*:00.sfx files, ~187GB!). The costly outputting of surfex FA files is necessary to have complete specification of the initial conditions for intermediate sub-hourly cycles (see below 2.2.1 section).

Synthetic DOW data: DOW radar pseudo-images can be generated from the 10-minute ICMSH*+*:000 FA files of the "Truth" twin. This is achieved by running in mode BOG the Field-Alignment (FA) software on these files. The DOW obs-operator is incorporated in this FA software. It also contains the utility to code these false Doppler wind observations in OPERA-BUFR format. At the end of the process, we have a set of BUFR files totally alike to those provided in this same BUFR format by the AEMET weather radars operational network. One important advantage of working with this synthetic data is that the scanning strategy can be selected at will. For these experiments radar volumes of 4 low elevations (1.8, 1.6, 1.4 and 1.2) and 120 Km range were produced (see table 1). Operational exploration schedule at AEMET contains just 2 low elevations (below 5.0) in DOW mode.

2.2.1 Exploring the impact of sub-hourly DOW DA on the reference twin

These 10-minute DOW data can be used to correct the trajectory of the reference forecast. Assimilation of these data can be done by different methods and with different frequencies. The former issue (methodology) requires more elaboration than the second. The 3D-VAR algorithm is widely used across the ACCORD consortium for UA DA. It is also a fundamental element in the EnVAR systems operated by those ACCORD members with enough computing power to exploit Ensemble DA. It is therefore a natural first choice. Key issue to the 3D-VAR algorithm is the B-matrix specification. In these experiments the default B-matrix for the IBERIAxxm_2.5 domain is used. Clearly, this decision is objectionable and casts doubts on the reach of the results of these experiments. Estimations of model error variances, correlation lengths and multivariate correlations

for forecasts of lead times of hours surely are different from those characteristic of the error-growth during the first time steps of model integration. Trying to estimate these, involves the same sort of problems that are encountered in the estimation of ensemble-based B-matrix: noise filtering, localization in multivariate analysis, etc. The choice of using the default B-matrix is based on economy reasons and also on the reflection that, after all, a statistical interpolation approach to sub-hourly cycling is just only one of several possibilities. It may turn out not to be the best option. It is left for future experiments to address more carefully this issue.

The exploration of the impact due to different cycling frequencies on the other hand requires only technical changes. To this end, the cy43 HARMONIE-AROME system has been adapted to run with cycle periods of minutes. The detailed set-up is best described with the following figure

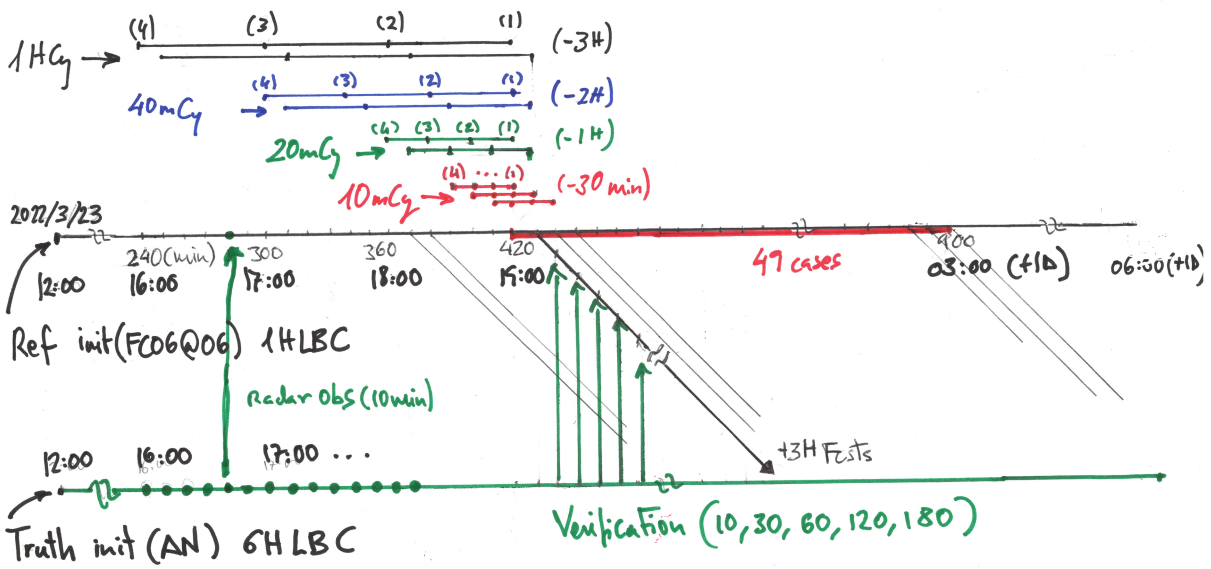

Figure 2: Diagram depicting the twin-experiments scheme at sub-hourly cycling frequency. Read the text for explanation.

In figure 2, the bottom green horizontal line running from left to right represents the "truth" forecast. The green dots on this line indicate the 10-minute output files which are used for the generation of synthetic DOW observations that will be used to initialize the sub-hourly experiments (thick vertical green line) and also for the computation of differences used to measure the impact ("Verification" thin vertical green lines). The long horizontal line running from left to right in the middle of the figure represents the "reference" or "baseline" forecast. On the top of the diagram the four sets of coloured shorter horizontal lines represent the four cycling frequencies explored in this work: hourly frequency (black lines), 40 minutes (blue), 20 minutes (green) and 10 minutes (red). One very important parameter in these experiments is the number of consecutive initializations used to "wind-up" the model. In this study a fixed number of 4 cycles for the four different tested cycling frequencies is selected, this means that the 1H experiment gathers information over the previous 3 hours (at 1H sampling rate), while the 10minutes experiment does it over the last half-hour (at 10 minutes sampling rate). It is perhaps a limitation in the design of these experiments that this wind-up number is the same in all cases and future experiments can explore variations in this parameter. After these four consecutive DA cycles are completed a short forecast (up to +3H) is launched. The process is repeated by sliding the colour lines 10 minutes to the right. The 49 cases between 19:00 UTC (23rd) and 03 UTC (24th) make a homogeneous set suitable for comparisons. This time period also covers the phase of highest intensity of precipitation.

The only aspect in which all these 3H forecasts are intended to be different is in the specification of the initial state of the UA fields. This strict control of the experiments configuration is achieved by

taking for all the exps for the surface initial state the corresponding ICMSH*+*:00.sfx FA 10minute files outputted by "reference". As for the LBCs, they are interpolated to the minute of the initial start times from the LBCs used in the "reference". This interpolation is done however in a new way that does not use the "gl" utility and that guarantees that the values of the fields on the boundaries correspond exactly with those used by "reference" for the given minute (or time-step, as both coincide in these experiments). There are also practical reasons that make this new preparation of LBCs attractive. NWP-NWC systems for closing-up on rapidly evolving weather must streamline ancillary processes, like the generation of LBCs, required by nested LAM-NWP systems. In the current HARMONIE-AROME system this generation starts anew for each cycle even when the initialization time has shifted ahead by only few minutes, as is the case in the experiments presented in this letter. It is not much trouble however to spare all this overhead, and at the same time be certain that LBCs are fully consistent between exps and reference, by introducing a pointer that keeps track of the time-step shift between the initial time of the forecast and the nearest previous LBC file. This new method has been implemented and is activated with a new switch LSBHOURLY in NAMCT0. A detailed description of this implementation is out of the scope of this letter.

2.2.2 2D-wind retrievals from DOW data with the Field-Alignment (FA) algorithm

The synthetic DOW observations are processed by the Field-Alignment (FA) algorithm to retrieve 2D-wind fields at different model levels. The procedure has been successfully tested already several times in other occasions, see for instance (1). The radar scanning parameters for these experiments are the following

Table 1: Radar Volume Parameters for Synthetic DOW Data

scan#	Elev (º)	BeamW (º)	Bins (#)	Rays (#)	BinSize (m)	StartRange (Km)	Nyquist (m/s)
1	1.8	0.9	240	480	500	0.13	48,2
2	1.6	0.9	240	480	500	0.13	48,2
3	1.4	0.9	240	480	500	0.13	48,2
4	1.2	0.9	240	480	500	0.13	48,2

These wind pseudo-observations are ingested in an ODB database created by the own FA package. These data use ODB tables lay-out and code types as those of AMDAR data. The FA software also performs quality control and thinning and determination of observations and FG errors ("σ_o" and "σ_b"). At the end of the process we have an ODB database ready to be presented to the 3D-Var algorithm.

2.2.3 Initialization. The Variational Constraints (VC) algorithm

Spin-up is certainly another relevant issue in these experiments. Previous studies have shown that spin-up effects in the ALADIN-NH dynamics are small after one hour of integration. One of the goals of these experiments is to study these spin-up effects and the performance of the Variational Constraints initialization algorithm (VC) in the context of sub-hour DA. This algorithm has been tailored made to ALADIN-NH dynamics because it uses the set of equations in its SI system to define a weak-constrained variational problem on the increments to four model prognostic variables: vertical velocity, horizontal divergence, temperature and surface pressure. This algorithm, which has also been tested several times in other experiments (2), employs a cubic-spline vertical discretization scheme and Green's functions kernels to obtain the solution by numerical quadrature. In spite of being a method developed from and for the ALADIN-NH dynamics, the set of prognostic variables it considers, and the numeric methods it employs to find a solution, differ from the ones actually used by ALADIN-NH during the forecast, and this introduces incongruences in the interface DA-dynamics that surely detract effectiveness to the method and ideally should be removed. The VC cost to minimize is

$$2J(\Delta x) = \int_0^{\xi_{tot}} d\xi \; (w(\Delta x - d)^2 + (M\Delta x)^2) = \int_0^{\xi_{tot}} d\xi \; (w(\Delta x - d)^2 + C_1^2 + C_2^2 + C_3^2 + C_4^2)$$

Here "w" is a free tuneable parameter for the relative weight between departures and constraints, "M" stands for the SI system , "ξ" is the top-down vertical coordinate, "Δx" represent the increments to find for the four different above mentioned variables, and "d" the corresponding departures between observations and background for these variables (if available ! Unfortunately no vertical wind obs are available. For no available departures d=0). "Δx" and "d" are both suitably made non-dimensional.

$$\Delta x^T = \left(\Delta gw \; , \Delta D \; , \Delta T \; , \Delta \pi_s\right) \quad ; d^T = \left(gw_o - gw_b \; , D_o - D_b \; , T_o - T_b \; , \pi_{s,o} - \pi_{s,b}\right) = \left(d_{gw}, d_D, d_T, d_{\pi_s}\right)$$

The "M" operating on "Δx" represents the SI system of equations in its VC form. It and its adjoint (with respect to the inner product defined by ∫ dξ) can be written as follows

$$M = \begin{bmatrix} L & 0 & 0 & 0 \\ -K^2\left(1+\gamma\,\partial\right) & \left(1+K^2\,\gamma\right) & 0 & 0 \\ -\dfrac{R}{c_v}\partial & \dfrac{R}{c_v} & 1 & 0 \\ 0 & N[\;] & 0 & 1 \end{bmatrix} \quad M^+ = \begin{bmatrix} L^+ & -K^2\left(1-\gamma\,\partial\right) & \dfrac{R}{c_v}\partial & 0 \\ 0 & \left(1+K^2\,\gamma\right) & \dfrac{R}{c_v} & N^+[\;] \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The interested reader is referred to (3) for the precise definition of all these symbols and also for a detailed description of the method. By expanding "M Δx" one can see that the C_i are given by

$$C_1 = \left(-\lambda + \partial\left(\partial + 1\right)\right)\Delta gw \equiv L\left[\Delta gw\right] \quad ;$$

$$C_2 = -K^2\left(1+\gamma\,\partial\right)\Delta gw + \left(1+\gamma\,K^2\right)\Delta D \quad ;$$

$$C_3 = \Delta T + \frac{R}{c_v}\left(\Delta D - \partial\Delta gw\right) \quad ;$$

$$C_4 = \Delta\pi_s + N\left[\Delta D\right] \quad ;$$

They correspond to the vertical momentum, horizontal momentum (div part), T full compressible equation and surface pressure tendency equation in the linearized SI system. The VC cost is made stationary for the "Δx" that satisfy the following fourth order 1D elliptical boundary value problem:

$$M^{+}M\,\Delta x + \lambda\,\Delta x = w\,\Delta x \;;\; \Delta gw(0) = \Delta gw(\xi_{tot}) = \partial\Delta gw(0) = \partial\Delta gw(\xi_{tot}) = 0$$

Finally, the value for ΔPD is obtained from the first order ordinary differential equation

$$\Delta gw - N^2(\partial + 1)\Delta PD = 0 \qquad \Delta PD(0) = 0 \quad N^2 = {}^{g^2}/_{RT}$$

As mentioned above the B-matrix employed in this work incorporates the so-called "statistical balances". Because of economy of efforts for first tests, these regressions computed from differences between forecasts with lead times of a few hours are utilized here. Again, this choice may well be revisited in future investigations. The VC algorithm is applied on the analysis increments, that is, the processing flow is as the following scheme shows:

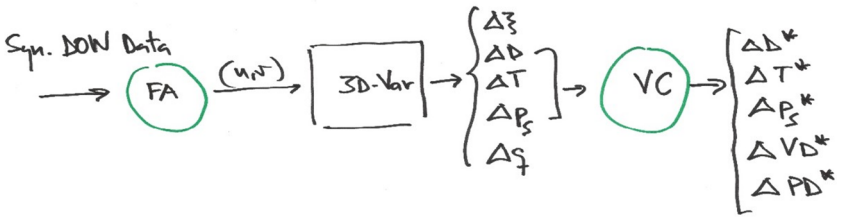

Figure 3: Data Flow and Initialization Strategy for the sub-hour UA-DA experiments

The FA algorithm aligns DOW radar images and first-guess Line-Of-Sight wind fields. The result of this process are 2D-wind pseudo-observations which are presented to the 3D-Var analysis algorithm. After the ordinary minimization the analysis increments are filtered by the VC algorithm. The initialized fields now also include the NH variables vertical divergence and pressure departure.

3 Results of the Experiments

3.1 Rainfall

Figure 4 compares the timeseries of **area averaged rain (AAR) accumulated in 1H (left) and in 3H (right)**. The averaging area is 50 x 50 grid points (125Km x 125Km) centered at the Málaga radar location. The purple bars correspond the "Truth" (TRU) and the thick green line corresponds to the "Reference" (REF). The x-axis signals each of the 49 short forecasts in the sample, comprising the time period from 19 UTC on the 23rd until 03 UTC on the 24th. The thinner colour lines display the results of the 10-minutes cycling (10mCy, cyan), 20-minutes cycling (20mCy brown) and 1-hour cycling (1hCy, deep blue). These lines have jagged aspect as compared with the REF line which is smoother. For the 1H accumulated AAR, the best score corresponds to the 20mCy and the 10mCy is second best. Both sub-hour experiments have corrected the phase or "timing" with respect to REF quite well. The 1hCy experiment performs clearly worse. For the 3H accumulated rain (on the right), there seems to be still some improvement in the 10mCy and 20mCy over the 1hCy, but statistical dispersion is higher rendering more uncertain any conclusions.

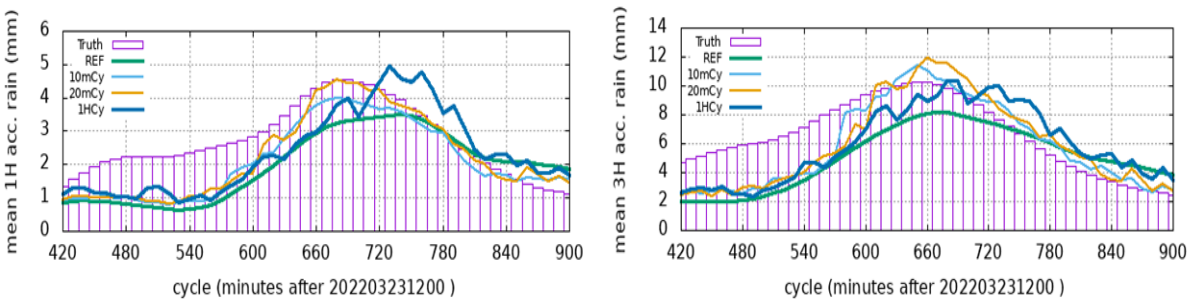

Figure 4: Time series of area averaged rain (AAR) accumulated in 1H (left) and in 3H (right)

Figure 5 shows the **ETS (Equitable Threat Score) for accumulated rain in 30 minutes, 1H, 2H and 3H**. These ETS are computed from contigency tables with the rain values in the 50 x 50 grid points in the area around the Málaga radar location, for the 49 cases in the experiment sample. The sub-hour experiments (10mCy and 20mCy) give for the 2H and 3H accumulated rainfall better scores than the 1hCy experiment, but the results are mixed in the 1/2H and 1H accumulated values. All of them, of course, beat the REF values (no DOW DA).

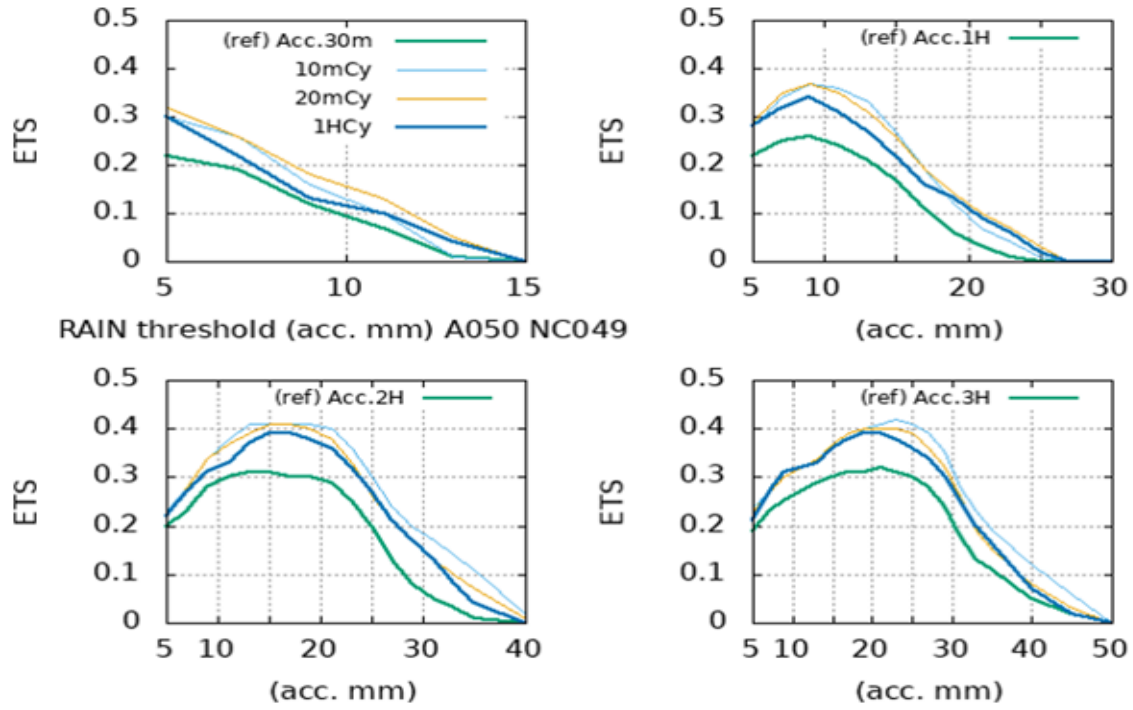

Figure 5 : ETS (Equitable Threat Score) for accumulated rain in 1/2H (top-left), 1H (top-right), 2H (left-bottom), and 3H (right-bottom)

3.2 Rainfall Extremes

Having in mind the implementation and deployment of NWP-NWC systems to manage hazardous weather, validation of extremes forecast is clearly important.

Figure 6 displays the time series for the **maximum of 1H (left) and 3H (right) accumulated rain over the 50 x 50 grid points area**. As above, TRU is represented by the purple boxes, REF by the green thick line and the 10mCy, 20mCy and 1hCy by the cyan, brown and deep-blue lines respectively. The over-forecasted values by REF are not corrected by the sub-hour DOW DA, although there seems to be an slight improvement in timing. Actually, the sub-hour experiments still

increase more the forecast values. The 1hCy performs clearly better now, in particular for the max 3H accumulated rain. This result can be seen more clearly in figure 7 which displays the histograms (in accumulated way) for the distribution of the 49 max values in the sample under analysis. The 1hCy adjust better than the sub-hour exps to the high-end of the histogram of TRU in all cases, and for the max 3H accumulated rain, in fact very well.

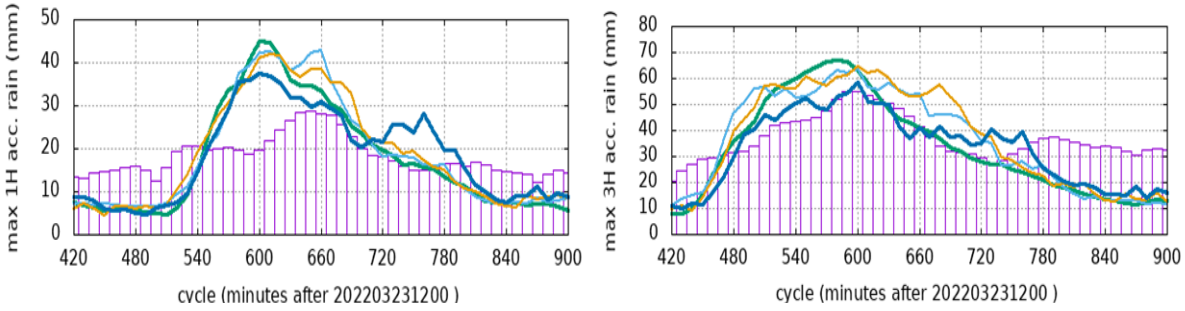

Figure 6: Time series of maximum of 1H (left) and 3H (right) accumulated rain over the Málaga radar area, for the different experiments.

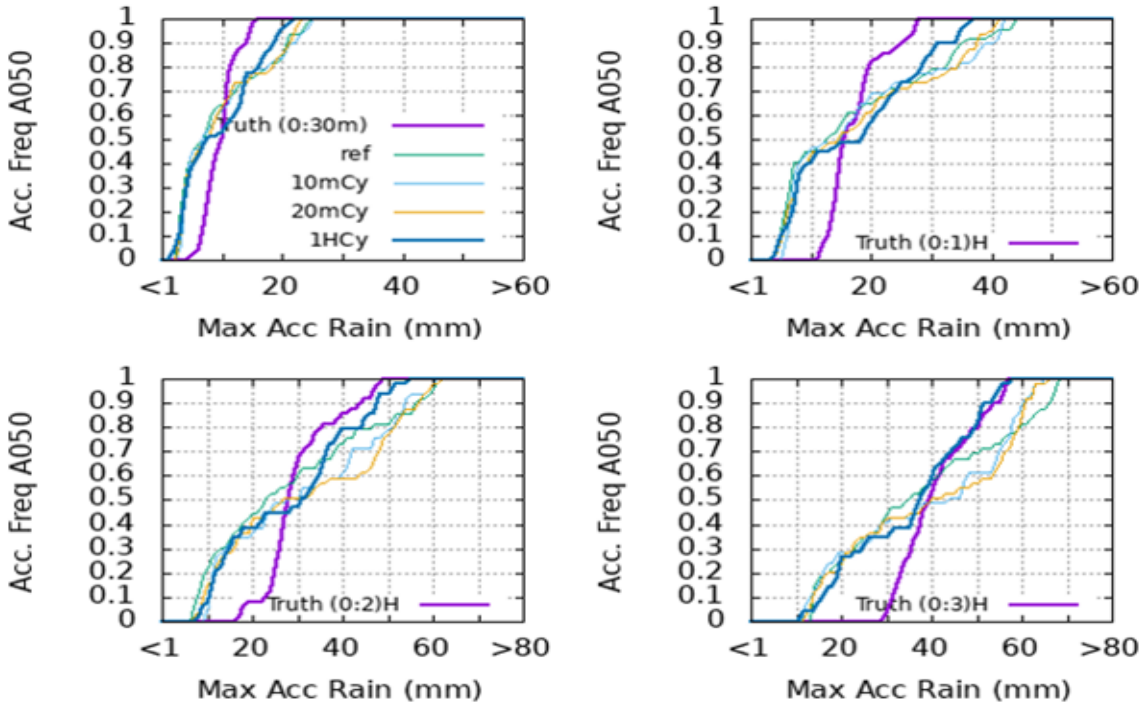

Figure 7: Histograms for the maximum of 1H (left) and 3H (right) accumulated rain over the Málaga radar area, for the different experiments.

4 Conclusions and Outlook

Comprehensive tests of UA DA of DOW data at cycling frequencies below the hour (10, 20 and 40 minutes) have been carried out with the Cy43 of the HARMONIE-AROME system. The validation results for the precipitation forecasts during an event of heavy rainfall, in particular for the extreme rainfall forecasted values in the very short range (up to 3 hours), may indicate some problems with the initialization of the model. While the 1 hour cycle DA correctly moderates over-forecasted values in the reference experiment, the forecasts generated by cycling at sub-hour rates failed to achive this. The result is at contrast with validation results for mean precipitation over the high-impact weather area. In

this case the sub-hour experiments perform better, they give better timing and intensity. More puzzling is the fact that the validation in terms of ETS does show improvement in the high precipitation thresholds (≥ 30mm) for 2H and 3H accumulated rain in the sub-hourly experiments, which is somewhat contradictory with the results obtained for the absolute maximum values. One possible way to explain this lack of congruence is that although sub-hour DA happens to improve the location of the maxima, and hence the ETS scores as they are computed by comparing values co-located on the model grid, it fails to improve the intensity of these maxima.

Another important issue addressed in these experiments is the initialization of the NH-ALADIN dynamics. The validations done indicate that the VC algorithm has a small positive impact on the results. The plots are not displayed for brevity but can be consulted in (4). Although positive, the magnitude is perhaps disappointingly small. This points to a weak link between DA and dynamics that deserves further investigation.

5 References

(1) Carlos Geijo (AEMET) hosted by Magnus Lindskog (SMHI), 27 June - 1 July 2022 : Meteorological Radar Doppler Wind Data Assimilation in HARMONIE-AROME using the Field-Alignment Algorithm
http://www.umr.cnrm.fr/accord/IMG/pdf/geijo_accord_visit_to_smhi_2022.pdf

(2) C.Geijo and P.Escribà (2018): "Variational Constraints for Data Assimilation in ALADIN-NH Dynamics"
https://www.researchgate.net/publication/326479446_Variational_Constraints_for_Data_Assimilation_in_ALADIN-NH_Dynamics

(3) C.Geijo (2018): "ANNEX to Variational Constraints for DA in ALADIN-NH Dynamics"
https://www.researchgate.net/publication/327117950_ANNEX_to_Variational_Constraints_for_DA_in_ALADIN-NH_Dynamics

(4) C. Geijo (2023): Oral presentation at the 3rd ACCORD ASW meeting.
http://www.umr-cnrm.fr/accord/IMG/pdf/exploringsubhourda.pdf

EPS research and development in LACE in 2022

Clemens Wastl, Martin Belluš, Gabriella Szépszó

1 Introduction

Within the LACE consortium three ensemble systems are operated: The common A-LAEF system (4.8 km horizontal resolution, ALARO based) and the two AROME based systems of Austria (C-LAEF) and Hungary (AROME-EPS) on a 2.5 km grid.
Major operational upgrades in the past year mainly comprised the migration of the two systems A-LAEF and C-LAEF to the new HPC facilities of ECMWF in Bologna and the introduction of EDA into the Hungarian system.

Furthermore several research and development activities are currently ongoing in the EPS area in LACE. A strong focus has been put on an improved representation of the model error in our ensemble systems. SPP (stochastically perturbed parameters) is becoming more popular in the EPS community and hence also LACE has put a lot of effort into this topic. For C-LAEF SPP is almost ready for operationalisation and also Hungary started to implement SPP into their AROME-EPS. In this context a new approach has been developed which considers a kind of flow dependency into the parameter perturbations.
Also in the post-processing area a lot of research is ongoing within LACE (SAMOS, EMOS, analog-based post-processing, etc.).

2 Operational systems upgrades

A-LAEF and C-LAEF migration:
The biggest change of the LACE operational EPSs in 2022 was the migration of A-LAEF and C-LAEF ecFlow suites to the new ECMWF HPC in Bologna (Atos). Originally, Atos was planned to be fully available at the end of 2020, but first access to the system was not possible before Q2 2022. After migrating the scripts/environment to Atos and the code compilation, the first test runs on the new machine have been performed in May and June. At the beginning, a lot of stability issues appeared (jobs suddenly crashed, etc.), which were eventually solved together with the ECMWF staff. A first, stable version of C-LAEF E-suite on Atos was launched at the beginning of July, the one of A-LAEF started in September. Some time has also been spent on the optimization of the task arrangement on the new facilities (number of cores, CPUs, etc.).

Verification of the C-LAEF E-suite during the summer months (July, August) showed comparable results for all the relevant parameters. However, when ECMWF provided new coupling files based on the new cy47r3 release candidate on Atos in September, the results were not reproducible. During the validation procedure of the A-LAEF E-suite, significant differences were observed in the lateral boundary conditions (LBCs) depending on the used global ECMWF inputs. These differences were identified by comparing the coupling files produced from ECMWF suites running on Cray and Atos HPC systems. Figure 1 shows that the production of LBCs (based on ECMWF-ENS) with c903 is reproducible on Atos (first row),

but when using ECMWF cy47r3 (release candidate) inputs of Atos in Bologna instead, considerable differences pop up e.g. for temperature fields. This also had an essential impact on the A-LAEF forecasts when comparing the two suites on Cray and Atos HPCs in Figure 2. For the unperturbed control run (upper six panels of Figure 2) the forecast differences are rather small but they increase with lead time. This corresponds to the observed discrepancy in the LBCs generated out of ECMWF's operational vs. release candidate outputs (both cy47r3, but running on different platforms, Figure 1). For the perturbed ensemble members (bottom six panels of Figure 2), the differences are about twice as big, and they tend to increase faster with lead time. In addition to the forcing by coupling, stochastic physics has a strong impact on the errors' growth too. Due to the stochastic nature of perturbation patterns, the same ensemble members are mutually incomparable between the two platforms (Reading vs. Bologna), even though the physics setting is equal for a given A-LAEF member (the perturbations arise from a combination of multi-physics and stochastic physics).

Nevertheless, we had to cope with these discrepancies. The overall ensemble mean characteristics were preserved, and the differences noticeable in the post-processed ensemble products were only marginal (not shown). As a result, we migrated the A-LAEF and C-LAEF TC2 suites to Bologna and became operational on October 18th, coinciding with the final release of Atos HPC for operational use. This was shortly followed by the termination of Cray HPC services in Reading in November. After some stability problems in the first weeks, the suites are now running reliably.

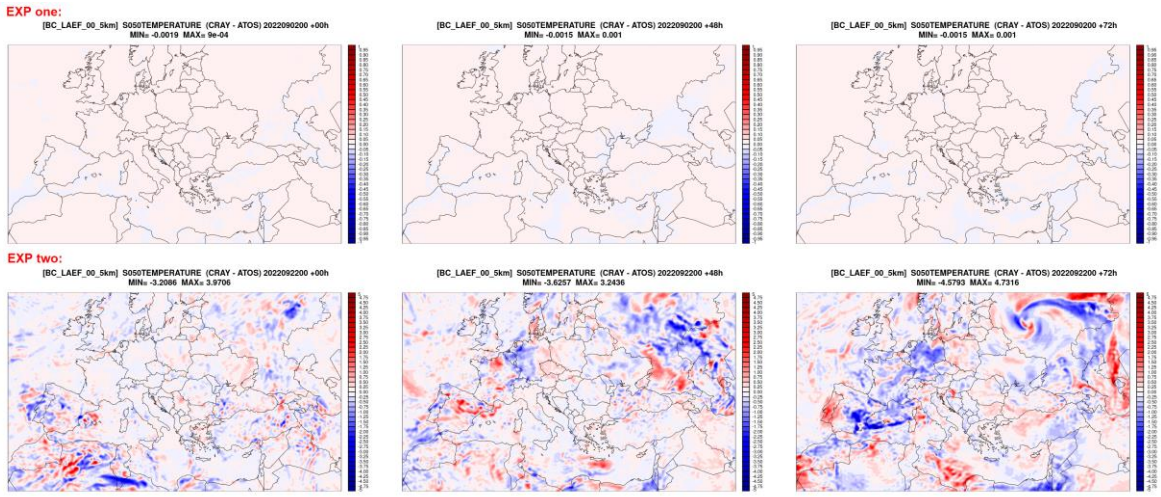

Figure 1: Differences in temperature at 50th model level for LBCs created on Atos HPC for the two experiments and different lead times +00, +48 and +72 hours (shown in columns). Experiment 1 used ECMWF cy47r3 inputs from Reading, while experiment 2 used ECMWF cy47r3 inputs from Bologna (release candidate at that time). In both cases, the differences were made against the operational LBCs created on Cray HPC (Reading).

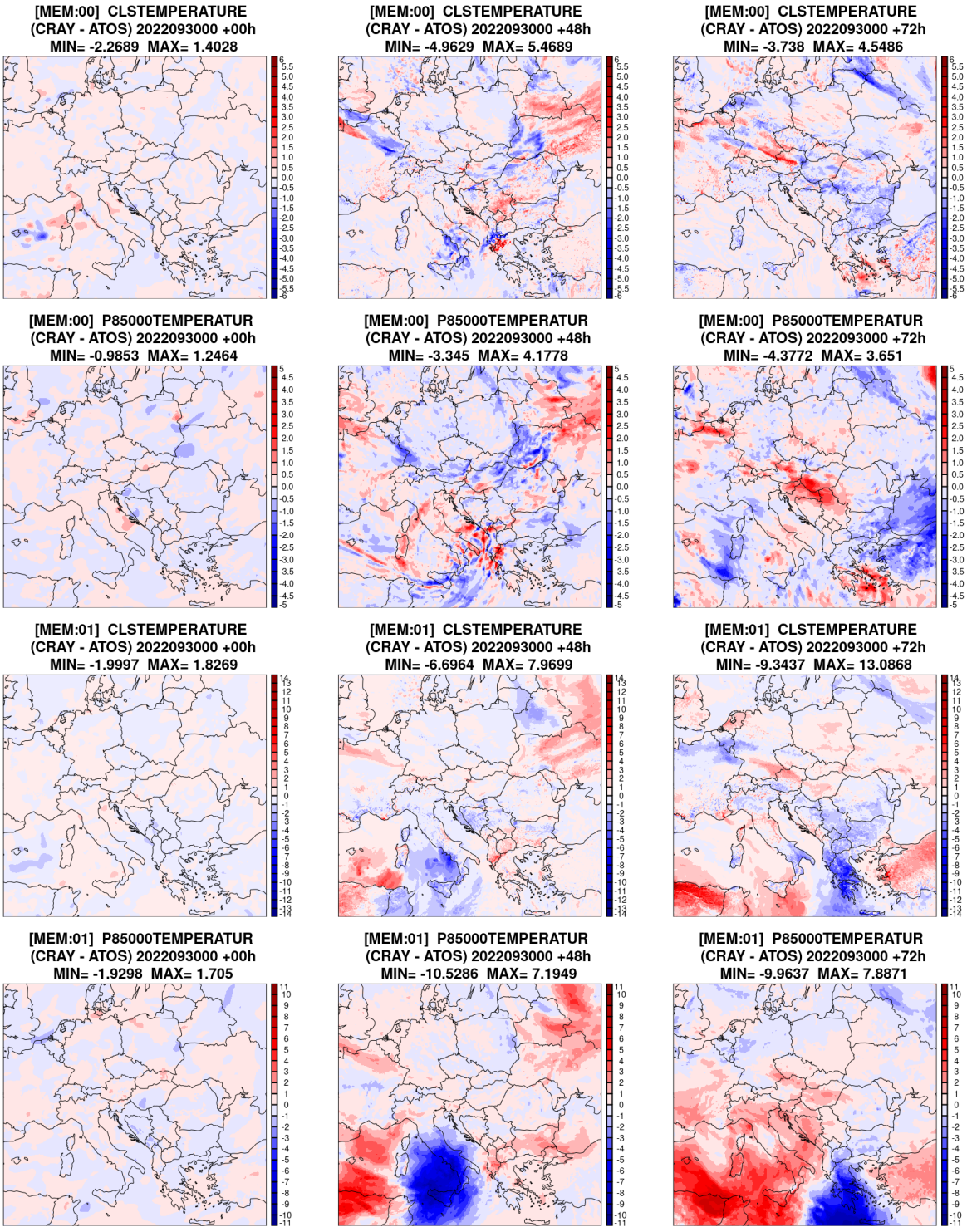

Figure 2: Differences in temperature between A-LAEF runs on Cray and Atos HPCs for unperturbed control member (upper six panels) and ensemble member 01 (bottom six panels). The differences are shown for 2m and 850 hPa (in rows), with three lead times +00, +48 and +72 hours (in columns).

EDA in AROME-EPS:

In March 2023 ensemble data assimilation (EDA) has been finally integrated into the operational AROME-EPS in Hungary. Before it was a pure dynamical downscaling of the first 11 members of ECMWF-ENS at 2.5 km horizontal resolution and 60 vertical levels (Jávorné et al., 2020). First experiments to introduce local initial perturbations with EDA have already started with CY40T1 in 2020. After an upgrade to CY43T2 an AROME-EPS version with EDA has been running for a 1-month test period in summer (July 2021) and winter (January 2022). After a cold start, the experiment was started with a 10-day spin-up period. The observations were perturbed offline before the surface assimilation and after screening. The results were compared to the 00 UTC operational AROME-EPS run.

During the summer period, applying EDA in AROME-EPS caused noticeable improvements in the surface parameters in general. For the 10m wind speed, 10m wind gust, 2m temperature and 2m relative humidity, the forecast was usually better during day time: the CRPS values and bias of the ensemble mean and control member were higher during the night. EDA rather decreased the error in the first 6 hours and during day time, and it had slight impact during the evening hours. The spread of these parameters increased for the whole forecast time. Precipitation and cloudiness results for the summer period were variable: both RMSE and spread were increasing during the first few hours, while the impact of EDA was almost neutral later. Wind speed forecasts performed best with EDA on most pressure levels in comparison with the operational AROME-EPS. There was also noticeable improvement by temperature and relative humidity on different heights. Compared to the results of the summer experiment, EDA caused less improvements in winter. In most cases, the impact was rather neutral. The best results were obtained for the 10m wind speed and wind gusts. Both RMSE of the ensemble mean and the CRPS decreased during the first 9 hours. For the 2m temperature, slight increase was seen in both scores, except the initial time. For relative humidity, the error increased during daytime and the impact of EDA was mostly neutral. The ensemble spread increased in the first 9-18 hours, however, the ensemble is still underdispersive.

An AROME-EPS-EDA E-suite has been started in August 2022, and it was intensively tested in comparison with the operational AROME-EPS by forecasters and model developers from 15 November to 15 December 2022. Compared to the EDA experiments before, this E-suite contained two runs a day (00 and 12 UTC) and used a Simplified Ensemble Kalman Filter (SEKF) in surface assimilation (identical to the operational AROME-HU deterministic model). The most positive effects of EDA were seen for the wind speed and wind gust (Figure 3). For relative humidity, the impact was rather neutral, except the initial time. EDA slightly increased the error of the 2m temperature forecast and it generally predicted lower temperature than the operational AROME-EPS. It had rather small effect on the precipitation forecast (Figure 4) and it was mainly limited to daytime. In forecasting of snow and freezing precipitation, EDA sometimes gave lower probabilities than the operational AROME-EPS.

As final conclusion, applying ensemble data assimilation in the Hungarian AROME-EPS had less positive effect in winter than in summer. The highest impact was seen at the beginning of the forecast time, both in growing spread and reduced error. The best results were obtained for the 10m wind gust and wind speed. Considering the results of all experiments, EDA has been finally put into operations in March 2023.

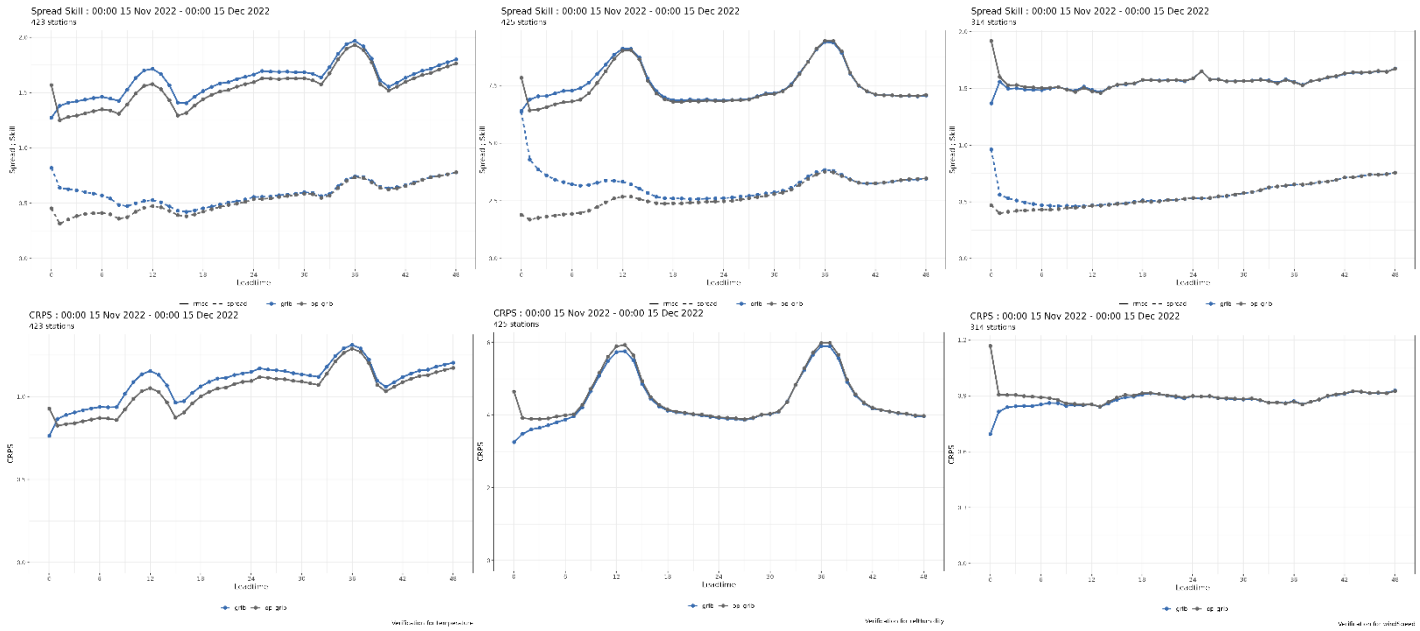

Figure 3: Spread and RMSE (top), CRPS (bottom) of 2m temp. (left), 2m rel. hum. (middle) and 10m wind speed (right) based on the operational AROME-EPS (grey) and AROME-EPS-EDA (blue) forecasts for the November/December period 2022.

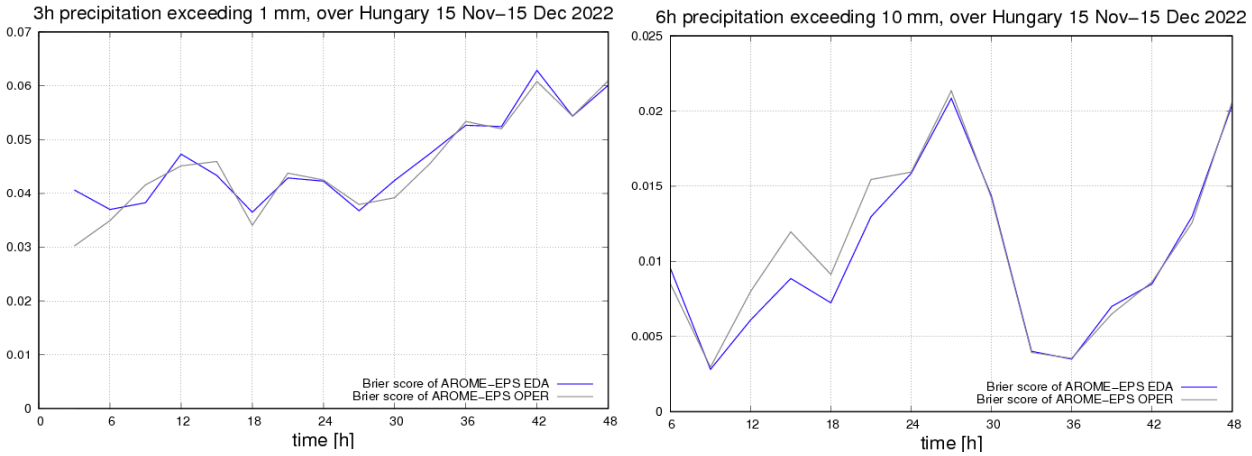

Figure 4: Brier-score for 3 hourly sum of precipitation exceeding 1mm (left) and 6 hourly sum of precipitation exceeding 10mm (right), based on the operational AROME-EPS (grey) and AROME-EPS-EDA (blue) forecasts for the November/December period 2022.

3 Case Study

20 August is a national holiday in Hungary celebrated with a big firework event in the evening. Weather situation over Central Europe on 20 August 2022 was governed by such complex processes, due to which the 6-18 hour range forecasts were characterized by large uncertainty and became accurate only approaching the event. The corresponding operational (ECMWF-ENS, A-LAEF, AROME-EPS-EDA) forecasts and some research experiments were carefully evaluated to understand the background (Szépszó et al., 2022).

The ECMWF-ENS (as well as other global model) forecasts predicted (false) high and intense convective precipitation patterns over the region for evening on 20 August already for several days. Thorough analysis of the clusters formed by the A-LAEF ensemble members (2 out of the 4 clusters are highlighted in Figure 5) showed that the situation during the day and the evening was just partly induced by the local circumstances (PBL stability, convective lifting). The evolution of thunderstorm systems was largely determined by cascade of interacting events (evolution of cells, outflows etc.) developing along several hundreds of km track in advance the evening. All these made the forecasts very sensitive to any small perturbations in the analysis, the lateral boundary conditions (LBCs) and the parametrization setups. Although the extra radiosondes erected in that morning already indicated some departure from the forecasts, these data had major impact only on the 6 UTC global forecast runs (when they were first assimilated) and on the 12 UTC LAM runs (when the LBCs were also updated). The case study demonstrated the importance of the ensemble predictions as well as their proper interpretation.

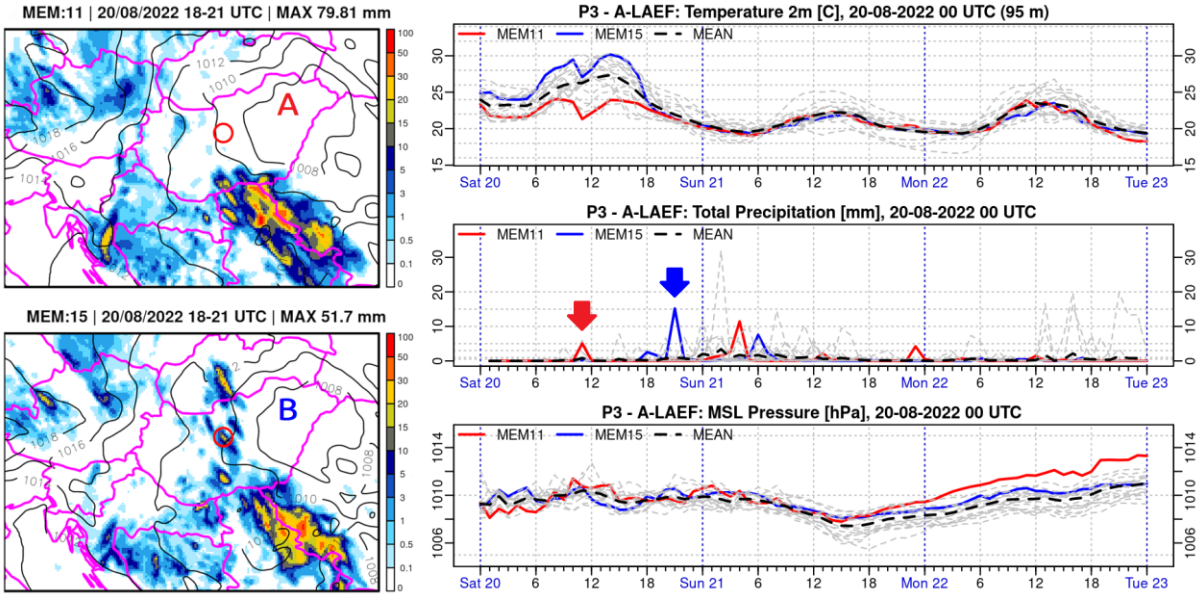

Figure 5: Forecasts of A-LAEF at 00 UTC on 20 August 2022. Top left: 3-hour sum of precipitation between 18 and 21 UTC 20 August according to member 11 of A-LAEF representing cluster A (with correct rejection); bottom left: the same for member 15 representing cluster B (with false alarm); right: plume diagrams for 2m temperature, hourly precipitation and mean sea level pressure according to the A-LAEF ensemble members (grey, while members 11 and 15 are highlighted in red and blue, respectively) and the ensemble mean (black).

4 Research and Development

Research in the EPS area within LACE is currently mainly focussing on an improved representation of the model error, but also in the post-processing area some work is ongoing.

Stochastically perturbed parametrisations (SPP):

The model error represenation in the operational C-LAEF system of Austria is based on a hybrid stochastic perturbation scheme, where perturbations of tendencies in shallow convection, radiation and microphysics are combined with parameter perturbations in the turbulence scheme (Wastl et al., 2020). Following the general trend in the EPS community towards more pyhsically consistent stochastic perturbation schemes, the idea came up to replace this hybrid system with a pure parameter perturbation scheme (SPP - stochastically perturbed parametrisations; Ollinaho et al., 2017). In SPP uncertain parameters are directly perturbed in the physics parametrisations with some random noise generated by a pattern generator (SPG, Tsyrulnikov and Gayfulin, 2017). A first version of the SPP scheme has already been implemented in a C-LAEF E-suite in 2021. This first version included a set of 13 stochastically perturbed parameters - 11 of those parameters are listed in the following Table 3. Additionally, 2 microphysics parameters were perturbed which are controlling the sublimation of graupel and snow hydrometeors (ZRDEPGRED, ZRDEPSRED). They have been added because of too strong orographic precipitation influence on the precipitation field in the operational C-LAEF (too much precipitation on the mountains and in the luv, too less in the valleys and in the lee). By stochastically perturbing these parameters, the precipitation field in the Alps could be improved significantly.

Table 1: Parameters which are perturbed stochastically in the SPP scheme currently implemented in a C-LAEF E-suite (yellow boxes).

Scheme	Parameter	Physical meaning	Default	Range
Radiation	RSWINHF	Shortwave inhomogeneity factor	1	0.6 - 1
	RLWINHF	Longwave inhomogeneity factor	1	0.6 - 1
Microphysic	RCRIAUTI	Snow Autoconversion threshold	0.2e-3	0.2e-4 - 0.25e-3
	RCRIAUTC	Rain Autoconversion threshold	1e-3	0.4e-3 - 1e-3
	VSIGQSAT	Constant for subgrid condensation	0.02	0 - 0.1
Turbulence	XLINI	Minimum mixing length	0	0 - 0.2
	XCTD	Constant for dissipation	1.2	0.98 - 1.2
	XCTP	Constant for T-P correlations	4.65	1.035 - 22.22
	XCEP	Constant for V-P correlations	2.11	0.225 - 4.0
	XCED	Constant for dissipation of TKE	0.85	0.4 - 2
	XPHI_LIM	Threshold value for Sc^{-1} and Pr^{-1}	3	1 - 4.5
	XCET	Constant for transport of TKE	0.4	0.072 - 1.512
Diffusion	SLHDEPSH	Strength of SLHD	0.060	0.01 - 0.09
	SLHDKMIN	Diffusion function minimum	0	-1 - 1
	SLHDKMAX	Diffusion function maximum	6	4 - 12
Surface	XRIMAX	Critical Richardson Number	0.2	0 - 0.2
	XFRACZ0	Coefficient of orographic drag	5	2 - 10
Convection	XCMF	Closure coefficient at bottom level	0.065	0 - 0.1
	XABUO	Coefficient of the buoyancy	1	0.7 - 1.5
	XBDETR	Coefficient of the detrainment	1e-6	0 - 1
	XENTR_DRY	Coefficient for dry entrainment	0.55	0.1 - 0.699

In 2022 a lot of tuning considering the perturbation scales and ranges has been made based on verification results from previous test periods and some case studies of severe weather events. In spring 2022 a full C-LAEF E-suite with SPP has been set up on the old ECMWF HPC Cray and it was running in May and June. The performance of the new SPP scheme was for most variables comparable to the previous hybrid scheme. For some parameters (e.g. relative humidity and temperature at 2m) the scores were even slightly better.

In summer 2022 this SPP scheme has been migrated to the C-LAEF suite on the new Atos HPC in Bologna. Furthermore the way how the perturbations are applied has been adapted according to a new perturbation code of the colleagues from HIRLAM. A full C-LAEF E-suite on the Atos HPC has been set-up at the beginning of November 2022 and was running until end of December. The verification results were very similar to the summer period on the Cray HPC (Figure 6). RMSE could be reduced for most parameters together with a slight decrease in ensemble spread. With these (more or less neutral) verification results it is planned to operationalize the new SPP scheme with the next upgrade of C-LAEF in May 2023.

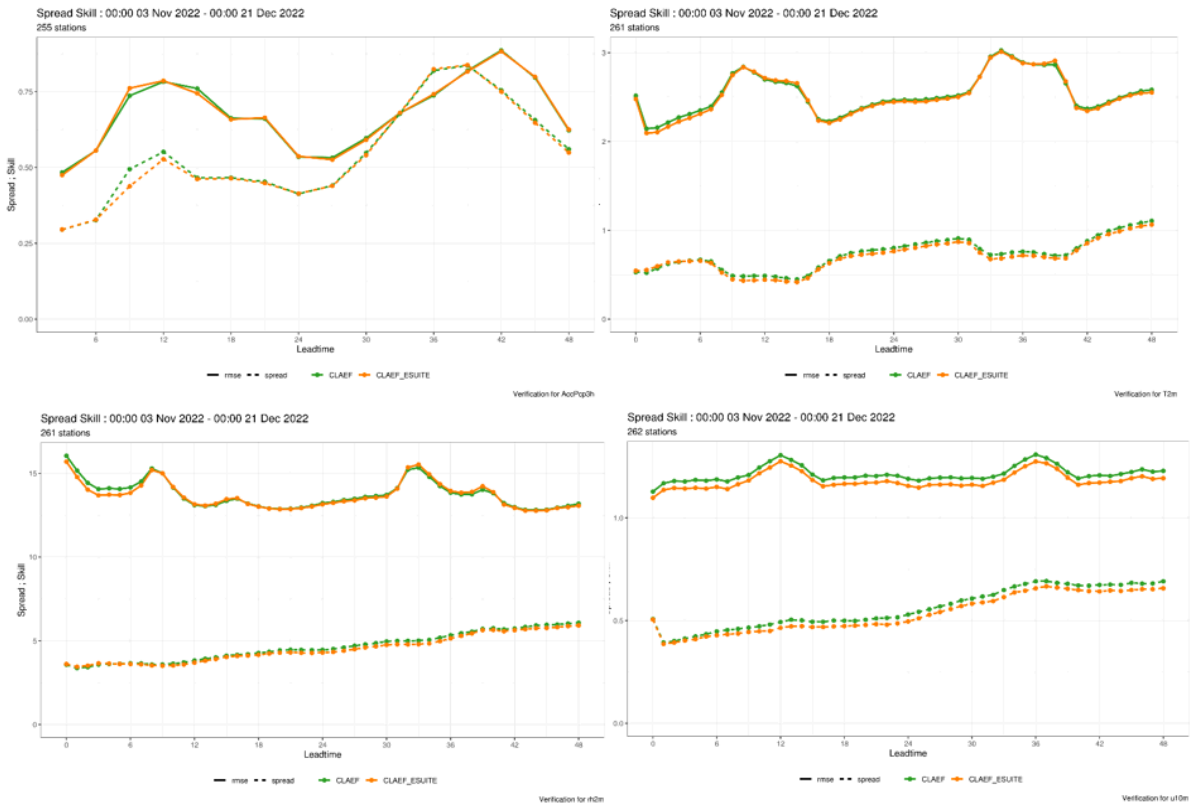

Figure 6: Spread (dashed) and RMSE (full) of operational C-LAEF with hybrid stochastic perturbation scheme (green) and C-LAEF E-suite with new SPP scheme (orange) for 3h accumulated precipitation (upper left), 2m temperature (upper right), 2m relative humidity (bottom left) and v-component of 10m wind (bottom right) for the period 03 November – 21 December 2022.

Flow dependent SPP:

SPP scheme is a widely used perturbation scheme to represent model uncertainties but it is purely stochastic - the perturbations are applied completely randomly without any consideration of the weather/flow situation. Therefore the idea came up to develop a kind of intelligent perturbation scheme which applies perturbations especially in areas where most impact can be expected (frontal zones, convective areas, etc.). First preparatory work (literature research, code study, etc.) in this area has been made in summer 2022, but main work started with the stay of Endy Keresturi (DHMZ) at Geosphere Austria in Vienna in October.

In the first version of this flow dependent perturbation scheme, the focus was set on the microphysics parametrisation (parameters ZRDEPSRED, ZRDEPGRED, RCRIAUTI and RCRIAUTC, see Table 1). The pattern generator and the perturbation strategy of SPP has not been modified, instead the approach is based on modifying the existing pattern by some weights. The idea is to diagnose which areas in the model are the most unstable for each

parameter and then to modify the pattern so that it perturbs more in those areas. The question is how to find which areas of the domain to target, i.e., how to determine the magnitude and the spatial distribution of weights. For microphysics, several model fields have been tested (precipitation, vertical moisture profile, cloudiness). The problem is how to convert these fields to weights (for example, to between 0 and 1) because they cannot be normalized easily as we do not have the access to the whole fields (parallelization). For this reason, cloud fraction has been chosen because it is already in fractions – between 0 and 1 so it is almost ready to be used as weight. First, as this field is given for all model levels separately, it needs to be summed up over all model levels. The normalized cloud fraction (Equ. 1) can now be applied to the SPP scheme (Equ. 2). w is the weighting factor, w' the vertically integrated cloud fraction, N_l the number of vertical levels, N an arbitrary factor to weight the influence, P the unperturbed parameter value, c a constant, φ the stochastic pattern and P^ the final flow dependent perturbation field.

(1) $$w = \left(\frac{w\prime}{N_l} \times N\right) + 1$$ (2) $$\hat{P} = P\, e^{c+w\varphi}$$

Figure 7 shows the impact of the cloudiness on the stochastic pattern for a selected microphysics parameter.

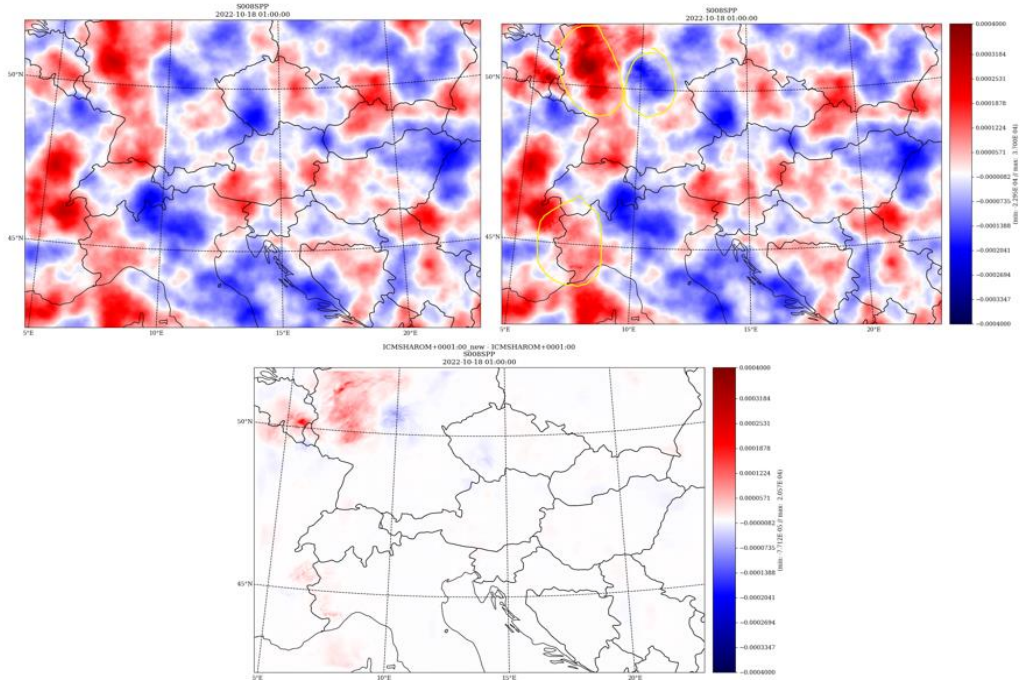

Figure 7: Impact of the cloudiness to the stochastic perturbation field of microphysics parameters in SPP. Upper left: SPP without flow dependency, upper right: SPP with flow dependency, lower panel: difference.

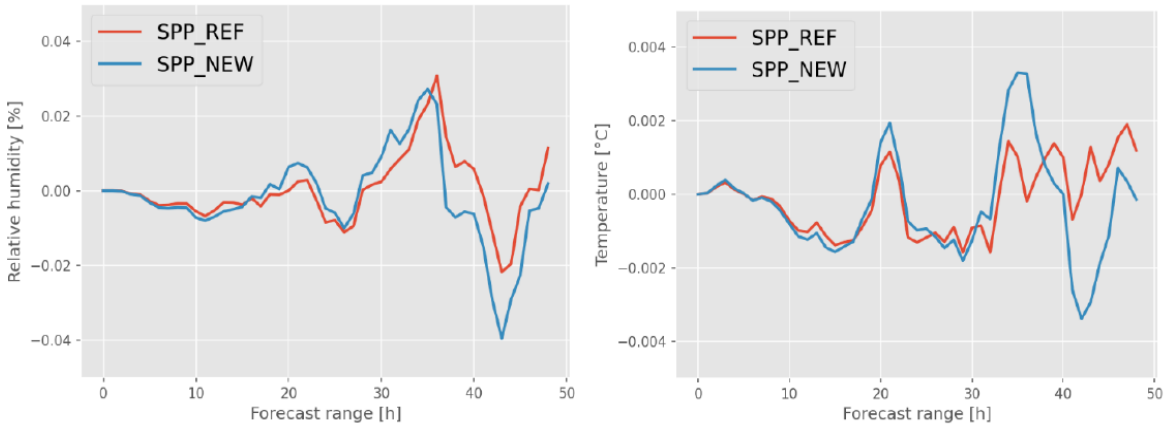

Figure 8: Domain average of relative ensemble spread of two SPP experiments for a test case on November 3rd 2022. SPP_REF is standard SPP, SPP_NEW is flow dependent SPP. Spread is relative to an experiment without any model error representation.

The new flow dependent parameter perturbation scheme has been tested with a full C-LAEF ensemble for a case study on 3 November 2022 (cold front passage). Three different experiments were made: NO_SPP – no model perturbations are used, SPP_REF – default SPP is used, SPP_NEW – new flow dependent SPP is used. In the two SPP experiments, only microphysics parameters were perturbed (4 parameters) and the forecasting range was 48 hours.

Figure 8 shows the impact on the ensemble spread in comparison to SPP without flow dependency. The impact is rather small, but it is nice to see that the method is principally working, since a difference is not visible before the cold front enters the domain (e.g. lead time +20h). No systematic differences could be observed which indicates that no bias is introduced by the new method. More details can be found in Endy's stay report on the LACE webpage.

This is just a first step into the direction of a full flow dependency in the model error representation. Much more impact can be expected when perturbing more parameters and for other cases (e.g. convective situation in summer). In a next step (planned research stay of Endy in 2023) it is planned to expand this methodology to other parametrisations (turbulence, shallow convection, radiation), to optimize it and test it over a longer period.

Statistical Post-processing:

There is some ongoing work in this topic in Austria, Hungary and Croatia. SAMOS (standardized anomaly model output statistics) has been developed and implemented at Geosphere Austria to improve direct model output from ensembles (EMCW-ENS, C-LAEF) especially for costumers. It has been put into operations in 2022. At the moment it has been implemented for 2m temperature and relative humidity, precipitation and 10m wind speed, gusts will follow soon. The training of the data is done every three hours with a rolling 45 days period in the past.

Verification shows that SAMOS is able to improve the BIAS of the EPSs significantly (Figures 9 and 10) and is also able to correct the under-dispersion. SAMOS is providing spatial forecasts and offers a seamless forecast from analysis over short-range to middle-range forecasts.

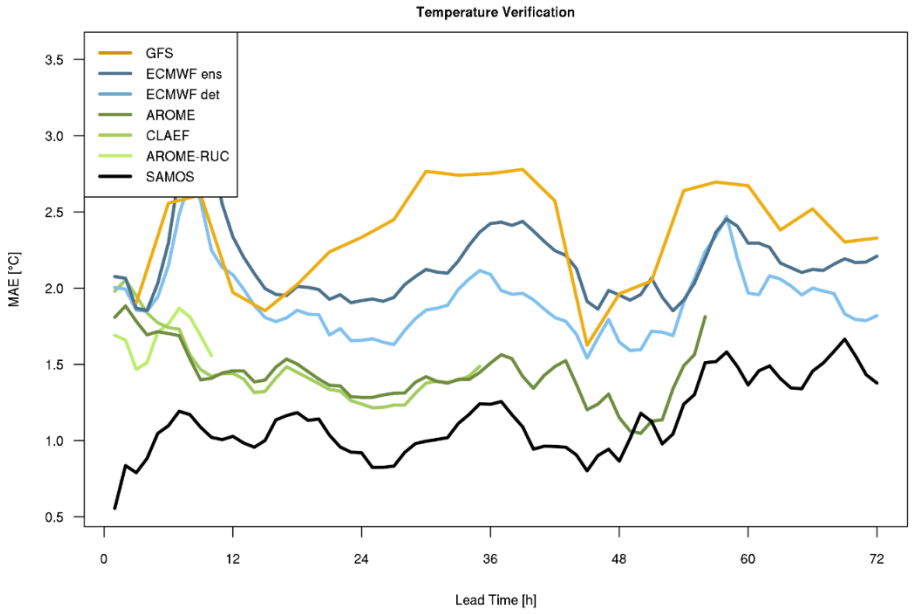

Figure 9: MAE of 2m temperature for different models and SAMOS for a case study in Austria.

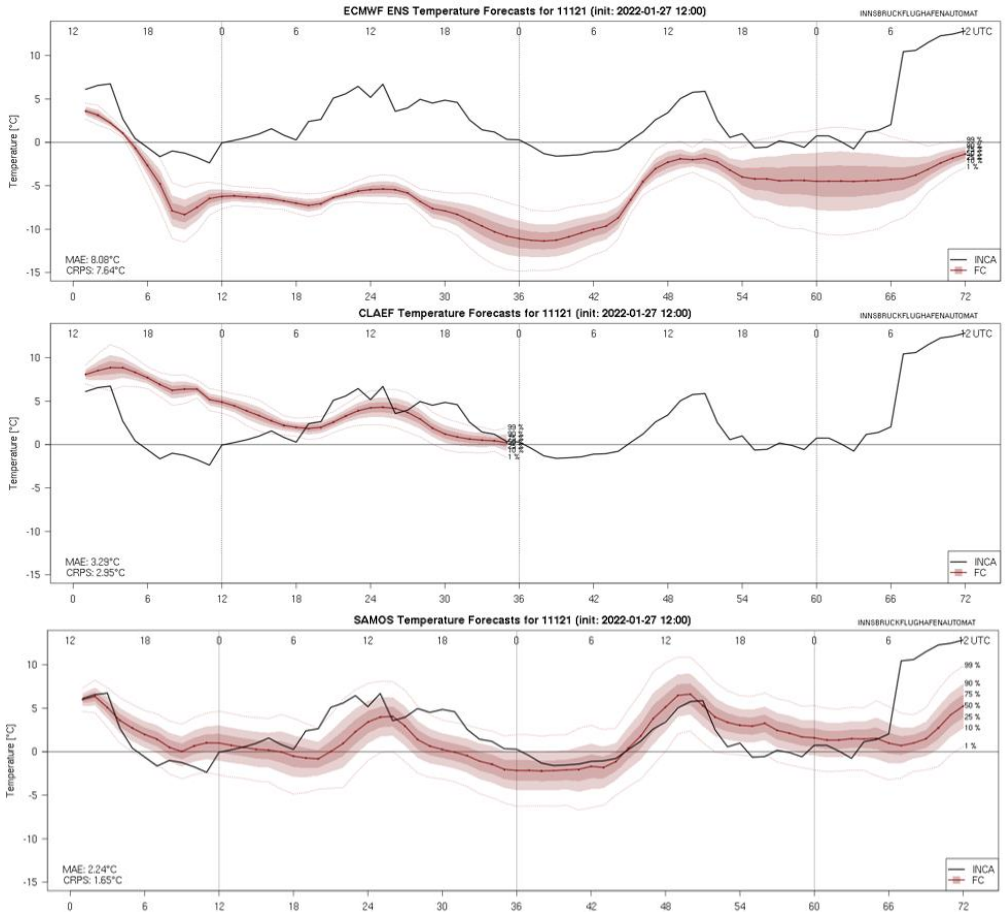

Figure 10: Temperature forecast for the station of Innsbruck on 27 January 2022 based on ECMWF-ENS (upper panel), C-LAEF (center) and SAMOS (lower).

OMSZ in Hungary 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. OMSZ worked on applying these procedures in operational practice including pre-processing.
To improve global radiation and 100 m wind forecasts, different methods were developed, taking into account the properties of these quantities. One-year data series were used for testing. For radiation, the EMOS technique proved to be successful, in which the distribution of radiation 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 forecasts initialized at 00 UTC in the grid points to nearest the stations (Schulz et al., 2021). 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 (Figure 11). 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 (Baran and Baran, 2021). MLP was the most successful (CRPS improved by 10-15 %; Figure 12). 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.
OMSZ integrated the code into the operational system and applied daily for the 00 UTC run of AROME-EPS. The procedure consists of 3 steps:

1. Collection of forecast data for the given day and observations for the preceding day, checking data quality and availability.
2. Using EMOS fitting and MLP training on the 15-minute outputs and producing new forecast values per station and timestep by determining 11 quantiles from the given distribution function.
3. Verification of forecasts: statistics for the preceding week, meteograms and verification plots to compare the post-processed and the raw forecasts.

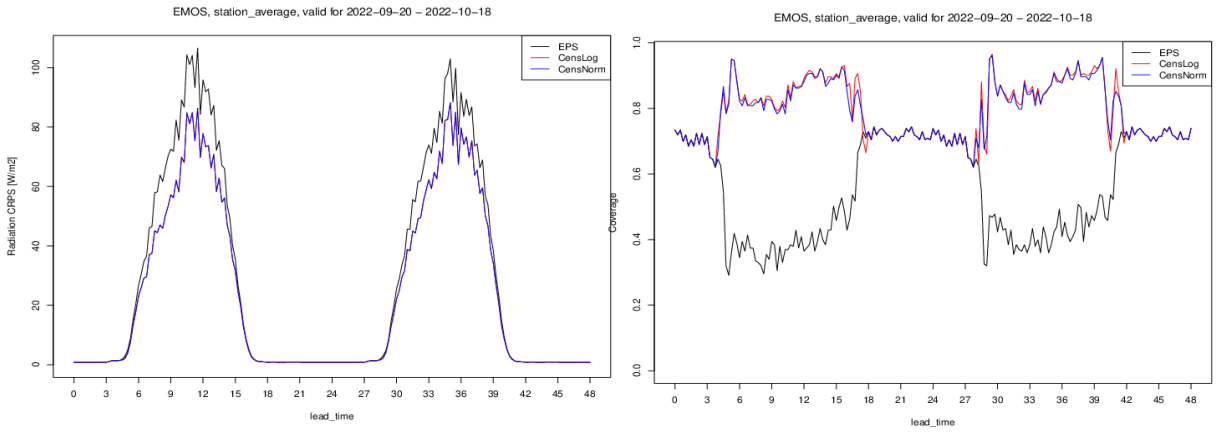

Figure 11: Verification scores for global radiation forecasts between 20 September and 18 October 2022 for 7 stations: CRPS (left) and coverage (right) based on the raw AROME-EPS forecasts (black) and the predictive distribution functions provided by CN-EMOS (blue) and CL-EMOS (red).

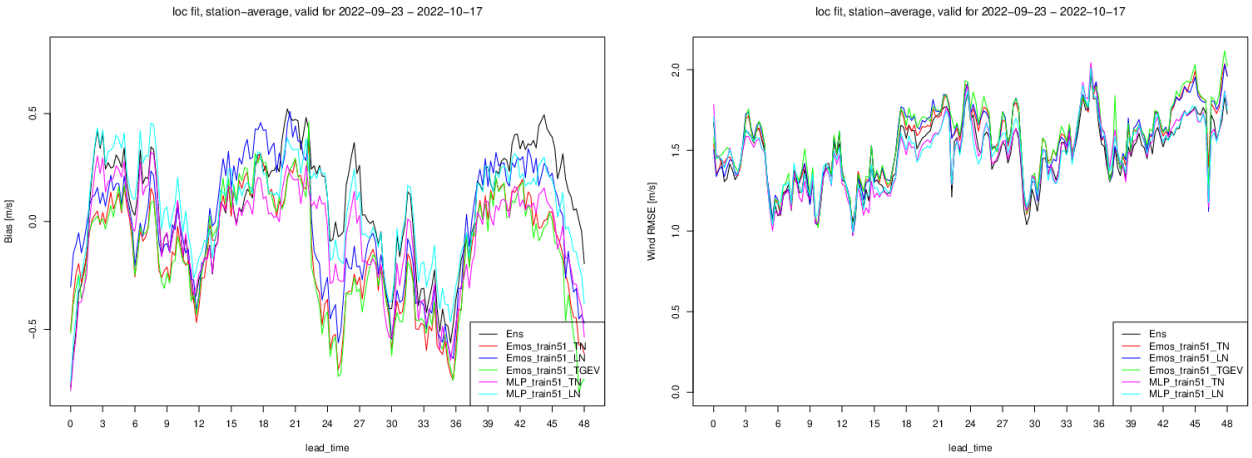

Figure 12: Verification scores for 100m wind forecasts between 23 September and 17 October 2022 for 3 stations: CRPS (left) and coverage (right) based on the raw AROME-EPS forecasts (black) and the predictive distribution functions by TN-EMOS (red), LN-EMOS (blue), TGEV-EMOS (green), TN-MLP (magenta), LN-MLP (cyan).

5 References

Baran, S. and Á. Baran, 2021: Calibration of wind speed ensemble forecasts for power generation. Időjárás 125, 609-624, https://doi.org/10.28974/idojaras.2021.4.4

Belluš M., M. Tudor, X. Abellan, 2022: "The mesoscale ensemble prediction system A-LAEF", ECMWF Newsletter, 172 - Summer 2022, 27-34, https://doi.org/10.21957/xa927ug5k0

Ollinaho, P., Lock, S. J., Leutbecher, M., Bechtold, P., Beljaars, A., Bozza, A., Forbes, R. M., Haiden, T., Hogan, R. J. and Sandu, I. (2017). Towards process-level representation of model uncertainties: Stochastically perturbed parametrisations in the ECMWF ensemble. Quart. J. Roy. Meteor. Soc. 143, 408–422, https://doi.org/10.1002/qj.2931

Radnóczi, J., 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.

Szépszó G., Csirmaz K., Kardos-Várkonyi A., Lancz D., Simon A., F. Prates, M. Belluš, M. Neštiak, 2022: Meteorological background of the forecasts on 20 August 2022 (in Hungarian). Légkör 67, 4, 182–188, https://doi.org/10.56474/legkor.2022.4.1

Tsyrulnikov, M. and D. Gayfulin, 2017: A limited-area spatio-temporal stochastic pattern generator for simulation of uncertainties in ensemble applications, Meteorologische Zeitschrift 26(N 5), 549-566, https://doi.org/10.1127/metz/2017/0815.

Wastl C., Y. Wang, A. Atencia, F. Weidle, C. Wittmann, C. Zingerle, E. Keresturi, 2021: CLAEF: Convection-permitting Limited-Area Ensemble Forecasting system. Quarterly Journal of the Royal Meteorological Society, 147, 1431-1451, https://doi.org/10.1002/qj.3986

Previous editions of the ACCORD Newsletter

1st ACCORD Newsletter,
published on 5 October 2021

2nd ACCORD Newsletter,
published on 28 February 2022

3rd ACCORD Newsletter,
published on 24 October 2022

The joint ALADIN-HIRLAM Newsletter (2013-2021, 16 editions) can be found on the ALADIN dedicated webpage.

Elaboration of the ACCORD Newsletters

For the concrete writing of an article, please refer to the editorial guidelines, accessible at:
http://www.accord-nwp.org/?Recommendations-templates

The newsletter content is based on voluntary contributions by the scientists and the teams in the consortium. We want it to be a useful tool for sharing both “practical” information and experience (code engineering, quality assurance, system aspects) and “more fundamental” results (advances in research work, outcome of specific meetings or working days etc.).

NL5 is expected for late 2024, however contributors can upload their material at any time during the year on this shared folder. Do also not hesitate to encourage scientific contributions by young scientists (PhDs, post-docs etc.) or technical contributions (codes, porting, optimization etc.).