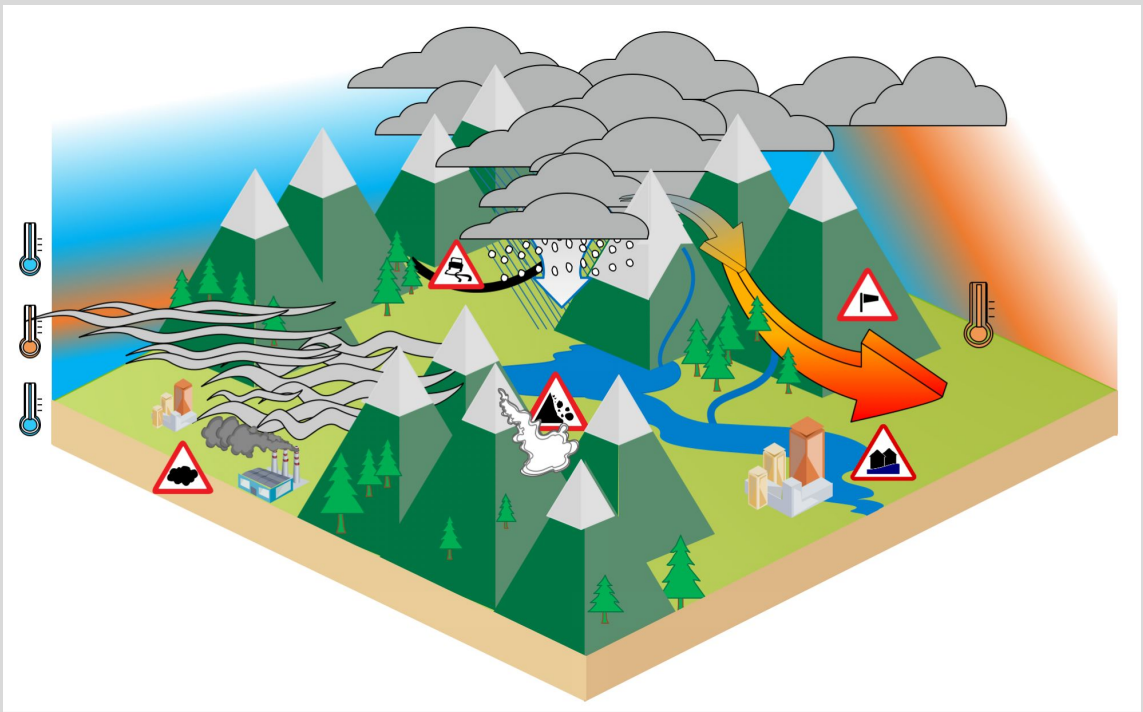
Newsletter n°1

5 October 2021



Schematic representation of mountain hazards related to small-scale atmospheric conditions from Gabriel Arnould and Danaé Préaux

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

We welcome herewith the very first ACCORD Newsletter, which gives an opportunity to look back at the first steps of our new consortium. The Assemblies of November 2020 and March 2021 have nominated the full compositions of the Management Group (MG) and the Support Team, as well as those of PAC and STAC. All of our Governance bodies have met at least one time. The MG started to work in March right in time to help prepare for the 1st All Staff Workshop. Several LTM meetings have been held as well.

The Rolling Work Plan (RWP) and the Common Manpower Register (CMR) are important tools for monitoring the consortium activity and resources. The MG has drafted the RWP-2022, the first version for which the MG and the associated co-leads will be fully responsible. The LTMs are now registering the quarterly manpower for their teams on the RWP-2021, using the new rules and the updated tools prepared by our Consortium Scientific Secretary (CSS, aka “Patricia”). The budgetary mechanism for the consortium, based on bilateral agreements of the members with Météo-France, is now in place and the financial exchanges have started. Our CSS is proposing an overview article about the many facets of governance, management, bodies and tools that now form the structural backbone of ACCORD.

Beyond these blunt “formal” facts, much of the management and scientific work in ACCORD was about our continued efforts to enhance bridges and practical collaboration across the teams. The MG members have devoted much of their energy to bring the scientists even closer together and encourage cross-family planning and working. Attention was paid to well discuss and agree on the goals, tasks and work practices per Area, and we know this whole process is work in progress for the next months and probably years.

The newsletter is one tool to make visible the concrete activity of the teams. In this first edition, model evaluation and validation appear in many of the contributions. Proposals of novel evaluation methods (like for extreme precipitation events), implementation of a new routine verification environment by one ACCORD member, specific evaluation of a new code cycle version installed on a home HPC, all provide different points of view about the quality assurance efforts in the consortium. Fog or visibility are being addressed in four different papers of this newsletter, from various perspectives (climatology, verification, research with very high resolution configurations), illustrating the scientific challenges in this area and certainly addressing (if only indirectly) the expectations by our users. Two articles address the challenge of forecasting temperature in mountain areas with deep valleys. One paper explains how a high resolution LAM can be used in comparison with a coarser resolution global model for assessing the impact of sub-grid orographic drag in the global model. Physiography data are addressed in several contributions and are likely to be a recurrent topic in the newsletter, given the needs for both accurate and new information for very high resolution modeling. A few contributions describe progress done within the CSC physics, EPS and data assimilation systems. The continuation of the DasKIT program within ACCORD is illustrated both by an overview article and by the realization of one member institute.

Techniques from Artificial Intelligence, like neural networks, are now more and more used in post-processing and downstream applications and there are several examples provided in this newsletter. Their use within the NWP codes remains an open area of research. On the side of technology, making the NWP codes more easily portable and ready to face the challenges of novel HPC architectures, is discussed in two articles. One, regarding cloud services, illustrates furthermore the potential possibilities offered by cloud-based technology.

A few words about the elaboration of our ACCORD newsletters. For the concrete writing of an article, I kindly refer to the editorial guidelines, that are 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.).

I wish to thank the contributors to newsletter number 1, and hope all ACCORD staff members find a good time reading its content.

Claude Fischer.

Presentation of ACCORD

Patricia Pottier, ACCORD Consortium Scientific Secretary

1 ACCORD Consortium

On the 27th of November 2020, 26 Euro-Mediterranean National Met Services (fig. 1) have decided to enter into a larger partnership and create a single consortium: ACCORD (**A** **C**onsortium for **CO**nvection-scale modelling **R**esearch and **D**evelopment), built on the ALADIN, LACE and HIRLAM consortia[1]. The ACCORD Memorandum of Understanding is valid from the 1st of January 2021 until the end of 2025.
More details about the Members.

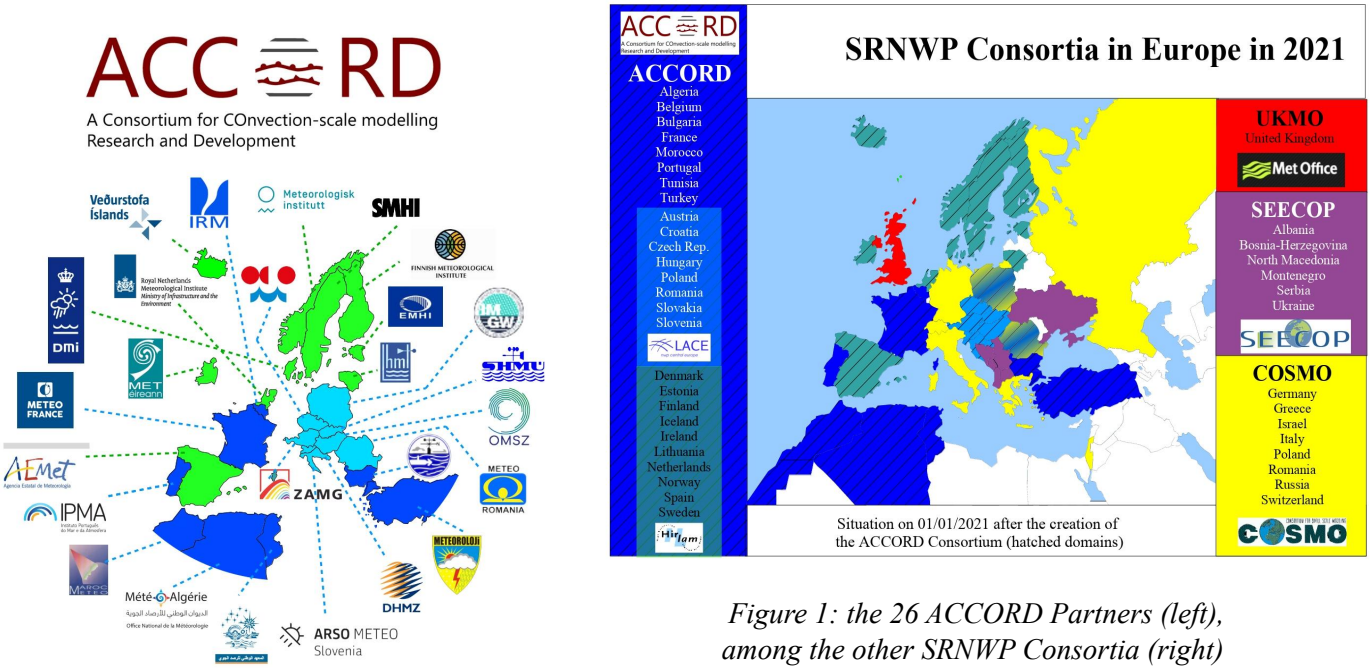


Figure 1: the 26 ACCORD Partners (left), among the other SRNWP Consortia (right)

Who does what ?

The **Assembly of the Directors** of the 26 NMSs Members of ACCORD is the upper governance body of ACCORD.
The **Programme Manager (PM)** is the highest executive officer of the management structure of the Consortium.: Claude Fischer from Météo-France.
The **Policy Advisory Committee (PAC)** advises the Assembly on strategic and policy matters of the Consortium. It is chaired by Daniel Gellens (Belgium) and Florinela Georgescu (Romania).

[1] ALADIN (Algeria, Austria, Belgium, Bulgaria, Croatia, Czech Republic, France, Hungary, Morocco, Poland, Portugal, Romania, Slovakia, Slovenia, Tunisia, Turkey),
LACE (Austria, Croatia, Czech Republic, Hungary, Poland, Romania, Slovakia, Slovenia),
HIRLAM (Denmark, Estonia, Finland, Iceland, Ireland, Lithuania, Netherlands, Norway, Spain, Sweden)

The **Scientific and Technical Advisory Committee (STAC)** advises the PM and the Assembly on scientific and technical issues. It is chaired by Saji Varghese (Ireland) and Christine Lac (France).
The **Project team** is composed of staff provided by Members to prepare and execute the Rolling Work Plan (RWP). It is led by a **Management Group (MG)**, chaired by the PM. The Management Group is composed of the three **CSC leaders** (Arome, Alaro and Harmonie-Arome), the **Integration leader**, the **Area Leaders for Dynamics, Surface, Meteorological QA, EPS, Transversal activities, Physics, System and Data Assimilation**.
The Project team also includes the position of a **Coordinator for Network Activities (CNA)**.
The **Support team** is composed of staff provided by Members to support the work of the Project team, including a **Consortium Scientific Secretary (CSS)** to assist the PM and the MG in their duties, a **Coordinator for the Data Assimilation starters KIT applications (DAsKIT coord.)** and **Local Team Managers (LTM)** to assure the good liaison between the Member and the project team.
The ACCORD organisational chart summarises the bodies and their compositions (fig. 2).

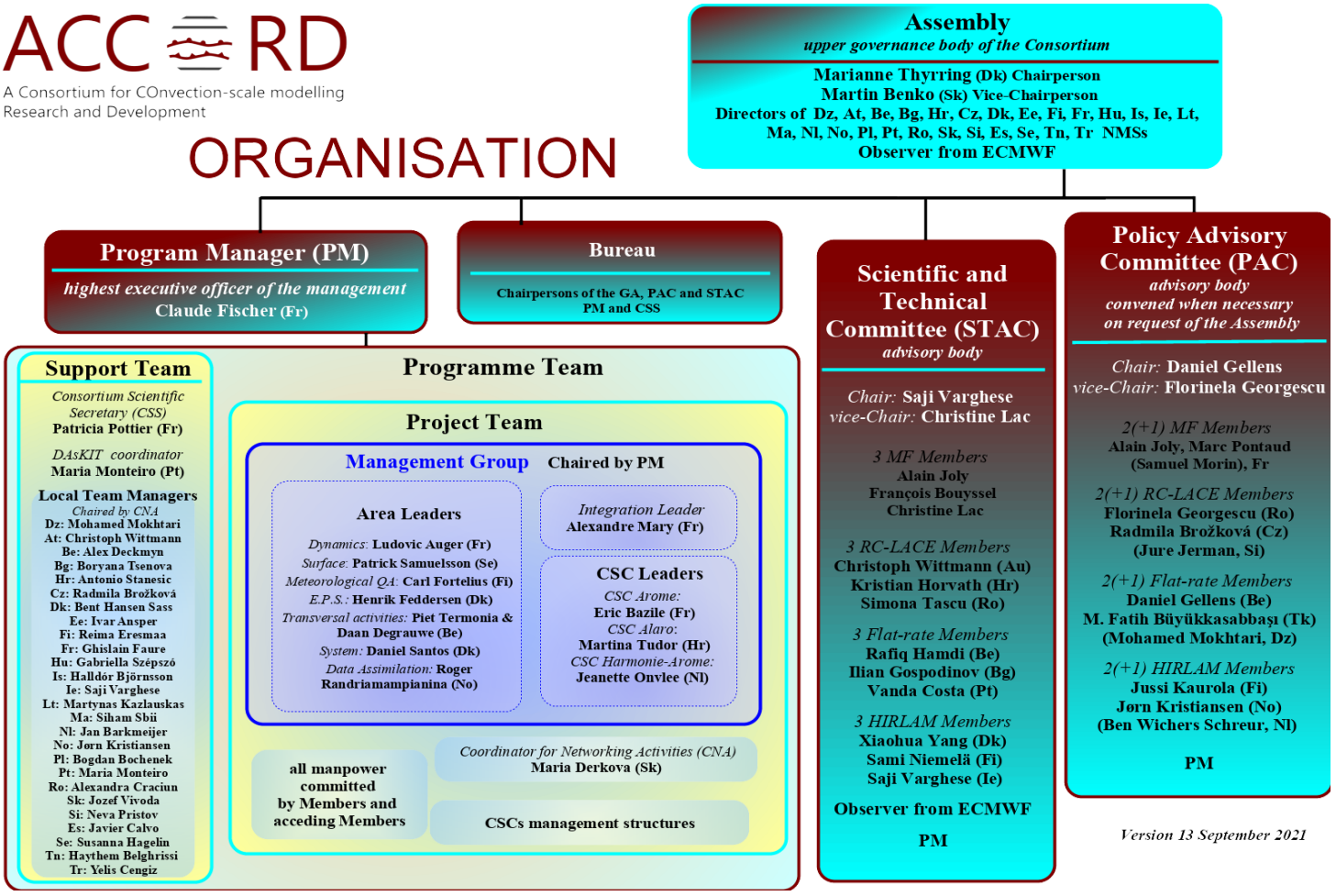


Figure 2: ACCORD organisation chart

All the bodies have met during the first semester of 2021, although always virtually. We took the opportunity of these meetings to prepare some "trombinoscopes" (a French word for photo gallery). Below are presented these trombinoscopes (fig. 3).

Figure 3: trombinoscopes

2 ACCORD strategy

The document Strategy 2021-2025 defines the strategic objectives for 2021-2025, and identifies the activities and resources needed to achieve them. Common key objectives were agreed on:

- for the usual scientific areas as described in the Rolling Work Plan: data assimilation, dynamics, model parametrisations, surface, EPS, quality assurance and system
- the key actions needed to achieve greater interoperability between CSC's and a more common working environment
- the possibilities and challenges arising from new technologies: computational science aspects, optimization of code/algorithms, performance and scalability on a variety of HPC architectures, possibilities from machine learning, the handling of increasingly large data volumes, etc.

3 Canonical System Configurations and operations

The NWP codes shared and developed in ACCORD use, to some extent, the IFS/ARPEGE global codes as backbone. Additional codes, not directly related to IFS/ARPEGE, are needed to form a full ACCORD model executable file (SURFEX surfaces cheme, Méso-NH physics package, specific surface assimilation codes, etc.).
The full ACCORD NWP system is currently being developed along 3 main model configurations, the so-called Canonical System Configurations ("CSC"):

- AROME
- HARMONIE-AROME
- ALARO

Within ACCORD Consortium, the 3 CSCs (ALARO, AROME, HARMONIE-AROME) and the ALADIN configuration run operationally on a wide range of computers, from single workstations to vectorial or scalar computers, in shared or distributed memory, and on a cluster of workstations under Linux.
The figure 4 illustrates the operational configurations in April 2021: details can be found on the ACCORD website (see "Operational configurations").

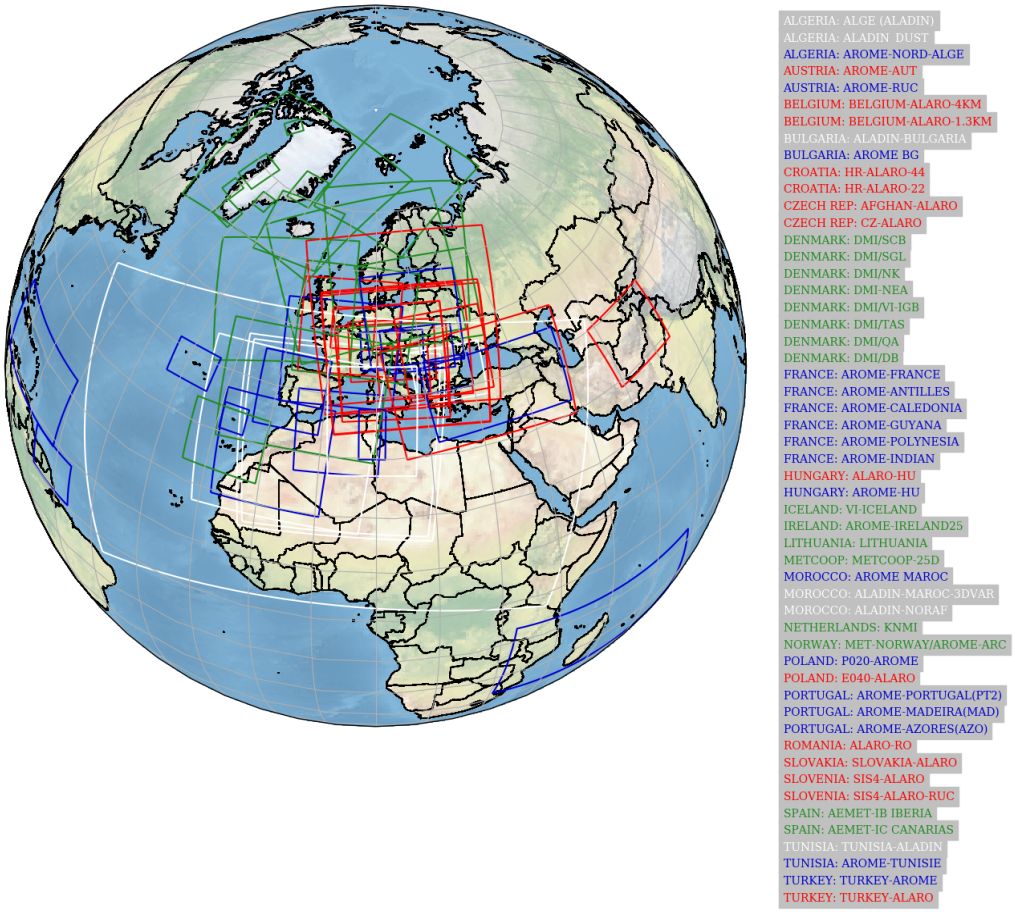


Fig. 4: operational configurations:
AROME (blue), ALARO (red), HARMONIE-AROME (green), ALADIN (white)

The aim in ACCORD is to progressively make a transition towards fully transversal, CSC-agnostic system configurations. This so-called “convergence” should be achieved by reaching a high degree of interoperability across the CSCs on various aspects such as physics, components of DA and EPS, scripting, and quite notably a common working environment and common work practice.
ACCORD codes are regularly updated with new R&D developments, technical code overhauls and phasing with the IFS/ARPEGE backbone codes. This evolution leads to the definition of new releases over time, aka as T-cycles. T-cycles are regularly defined by the successful integration and validation of code developments from the different partners (in ACCORD, including Météo-France contributions, and impact of the IFS/ARPEGE code changes on the LAM components when relevant).

The code evolution process is currently undergoing a modernization in terms of work environment, process and tools.The aim in ACCORD is to move to an incremental code integration with systematic evaluation of non-regression (of existing testing configurations), reproducibility of numerical results (unless otherwise stated by a contributor), new tests when appropriate. The intention is furthermore to make use of modern code configurations (e.g. like available with OOPS) and new testing tools (e.g. “DAVAÏ”[2]). All model configurations are addressed in the test procedure (ie IFS, ARPEGE and the LAMs).

4 R&D activities and Work Plans

1. Rolling Work Plans: a two-year process

The Management Group plans the work of the Consortium staff in order to utilise the available competences in the most efficient possible way. A specific tool used for planning is the yearly Rolling Work Plan (figure 5).
From the start of its redaction to the full reporting of its realisation, each RWP has a two-year life time (figure 6). Put in different words, the MG deals each year with 2 RPWs simultaneously, for the realisation of the current RWP and for the preparation of the next year RWP.
The redaction of the Rolling Work Plan for year YYYY generally begins at the very beginning of the 2nd quarter of year YYYY-1. Based on the RWP YYYY-1, the ACCORD MoU and Strategy 2021-2025, the guidance and remarks from the last Assembly, STAC (possibly PAC), the MG reviews the Work Packages definition and their redactors. Then, each member of the MG works on the content of the Work Packages relevant in their area. This step requires technical discussions and sometimes efforts of prioritization or negotiation, in close coordination with co-editors (lead scientists) and more generally the teams working in the area. At the end of June, the MG reviews the draft RWP and it is proposed to the LTMs for them to check that nothing misses with respect to their local scientific working plans. The redaction still goes on during summer. The MG finalises the text of the RWP early September. Then, the LTMs communicate their commitments to the redactors of the Work Packages where their team will contribute.
The draft RWP including the commitments is proposed to STAC, amended by the MG according to the STAC remarks, and proposed to the end-of-the-year Assembly for approval. The approved RWP is then made available for the teams.
During year YYYY, the teams work on the RWP under the coordination of the MG. The LTMs quarterly report in the Common Manpower Register the work done on the RWP during the previous

[2] The DAVAÏ testing system enables to test a code version (new development or merge result, for instance).
The steps encompass: fetching the codes to be tested, building executables, running sets of integrated or elementary test-cases (representative of canonical configurations including IFS, ARPEGE, AROME, ALARO, HARMONIE-AROME), automatic comparison of outputs to reference outputs,user-friendly display of these results. It is a crucial step in the process of integration and validation of code changes.

quarter. The MG follows the execution of the RWP all along the year and regularly checks the manpower register.
From September, the MG prepares a scientific report, to be scrutinized by STAC, then proposed to the Assembly. At the very beginning of the second quarter of YYYY+1, a summary report of the execution of the RWP YYYY is published, including the manpower realisation.

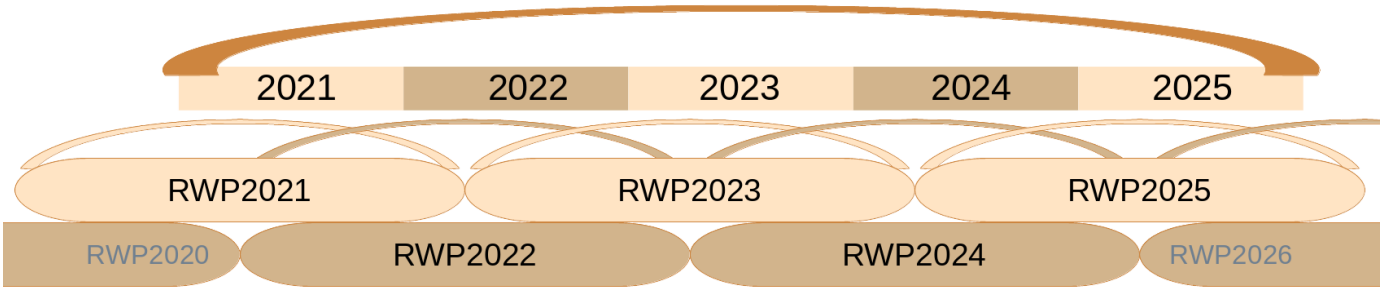


Fig. 5: schedule of the Rolling Work Plans during MoU1 period (2021-2025)

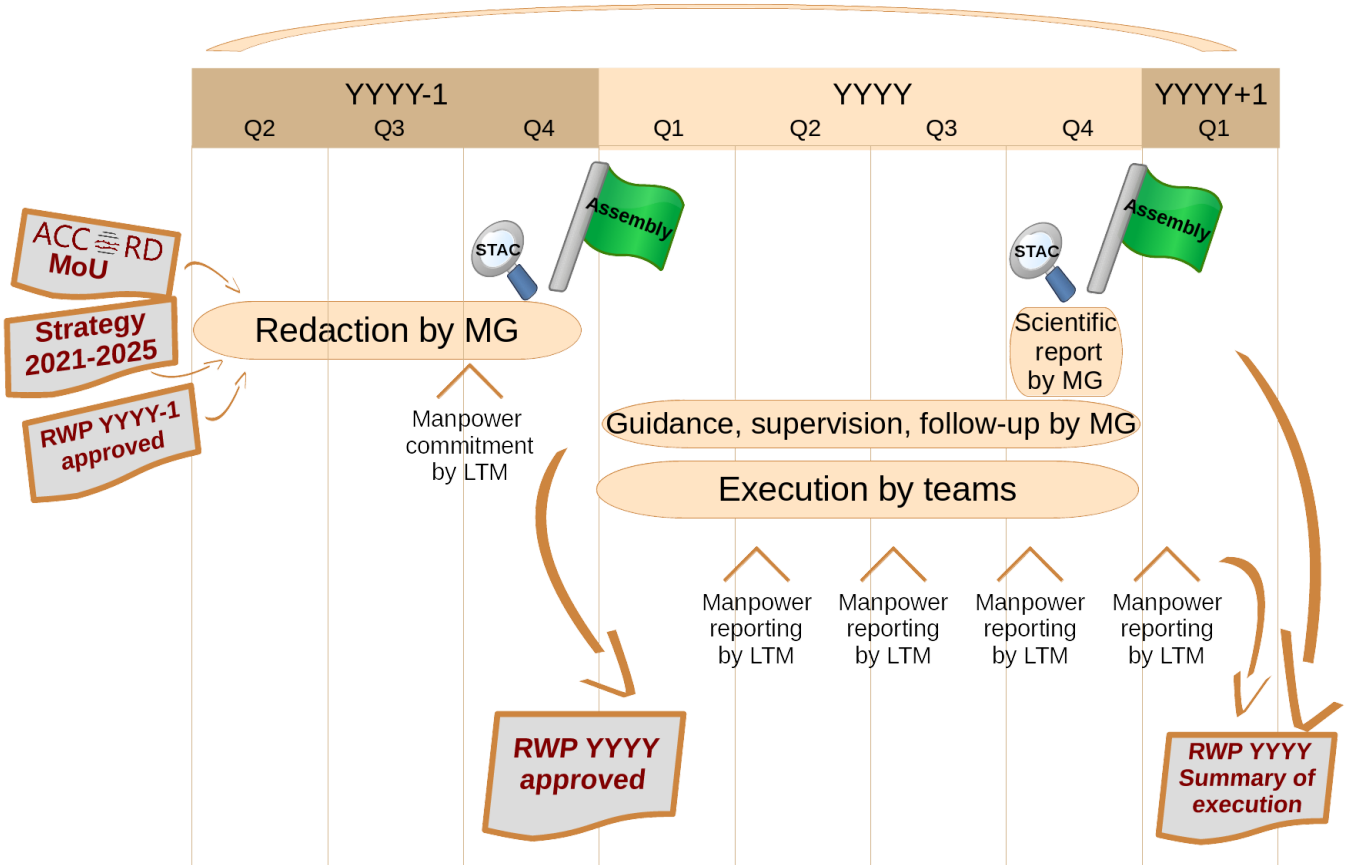


Fig. 6: Rolling Work Plan YYYY: redaction, approval, execution and reporting

2. Rolling Work Plan 2021

With respect to the previous ALADIN-HIRLAM common work plans, the content of the first ACCORD Rolling Work Plan (RWP2021) had been significantly adapted (while keeping the general structure in Work Packages) to strengthen the work on mid-term specific objectives while not sacrificing steady progress in the long term goals described in the Strategy 2021-2025 (see chapter 2).
Some headlines of the RWP2021:

- adapting the codes to new HPC architectures, new options for/in dynamical core (while continuing to improve the existing one),
- developments on upper-air DA (OOPS, EnVar, support for observations),
- surface (SURFEX in all CSCs, assimilation, coupling),
- physics (including increased interoperability across physics, 3D effects, use of Machine Learning approaches),
- EPS (new transversal WPs),
- extended Meteorological Quality Assurance (general verification of model performances and alleviation of model weaknesses),
- new WP for System (exploration of solutions and tools for converging to a more common working environment), towards sub-km modelling.

The Management and common activities WPs have also been renewed with a new WP for the ACCORD management, the common WPs have been updated (cycles, collaboration on porting the codes to Members, training) and new WPs added (design for modernized working environment, All Staff Workshop and EWGLAM meetings).

The Assembly approved the RWP2021 at their kick-off meeting in November 2020. Since January 2021, the teams are working along this work plan and their effort is quarterly reported by the LTMs and monitored by the MG. The MG is preparing the RWP2022 while organising the R&D activities within the thematic areas or through thematic working groups. Dedicated email lists are available by Area.
Each thematic area or working group has also dedicated pages on the ACCORD wiki where the teams exchange about the relevant Work Packages in the Rolling Work Plan, the Working Weeks, the video-meetings, etc ...:

- Thematic Areas:
  - System, AL :Daniel Santos Muñoz (Dk)
  - Data Assimilation, AL: Roger Randriamampianina (No), team organised in DA Research Teams (RT) and Support Teams (ST)
  - Surface, AL: Patrick Samuelsson, team organised around the RWP work packages, complemented with 2 thematic groups: Snow Analysis lead by Ekaterina Kurzeneva (Fi) and Urban group lead by Rafiq Hamdi (Be)
  - Dynamics, AL: Ludovic Auger (Fr)
  - Meteorological Quality Assurance, AL: Carl Fortelius
  - Physics: Physics area is lead by the Physics Working Group composed of the PM (Claude Fischer), AROME CSC Leader (Eric Bazile), HARMONIE-AROME CSC Leader (Jeanette Onvlee), ALARO CSC Leader (Martina Tudor), Transversal activities Area Leaders (Piet Termonia and Daan Degrauwe), Sylvie Malardel (Fr)
  - EPS, AL: Henrik Feddersen
  - Transversal Activities, AL: Piet Termonia and Daan Degrauwe
- Thematic Working Groups:
  - Machine Learning (ML)
  - Very High Resolution Modeling (VHR-MOD)
  - Seamless Data Assimilation, VHR-DA for NWP and Nowcasting (DA-SEAM)

The PM will report on the execution of the RWP2021 to STAC and to the Assembly and propose the RWP2022, prepared with the MG (the LTMs will commit the manpower of their team on the different Work Packages and tasks of the RWP2022).

5 Communication

As mentioned in the previous chapters, the communication inside ACCORD Members and teams and outside the consortium is already well organised:

- many Partners issued press releases to celebrate ACCORD birth and a video-animation was published
- ACCORD has its domain name (accord-nwp.org) and its logo (selected by the Assembly after many proposals by the teams)

ACCORD
A Consortium for COnvection-scale modelling Research and Development

This logo was designed by Peep Jürmann and Piret Pärnpuu from Estonian Environment Agency. The circle with 5 musical lines represents unity, commitment and community. Circle has no beginning or end, it represents life and the cycle of life. Two of the lines were converted to cold and warm fronts because every co-operation has its ups and downs and new possibilities that come of it. The 26 ACCORD NMSs are like musicians in a symphony orchestra who practice and rehearse together and the outcome to the ACCORD users is like music to their ears.

- website: http://www.accord-nwp.org/
- wiki: https://opensource.umr-cnrm.fr/projects/accord/
- email lists for the different governance bodies, for the teams, by topic or area, by CSC …
- many events (all of them virtual, so far) have been held and will be held inside ACCORD (see the next article “Events in 2021) and with colleagues from other consortia (ACCORD will be presented at the next EWGLAM meeting).

ACCORD events in 2021

Patricia Pottier, ACCORD Consortium Scientific Secretary

1 Introduction

With the effective start of ACCORD, many events were held, all of them on-line, from the MoU signature and the kick-off Assembly last November to the 2nd Assembly meeting in July. The next events are mostly also planned as video-meetings, at least for the second semester of 2021. We do hope to resume in-situ meetings in 2022, especially for the All Staff Meeting.

2 January - August 2021

When details (slides, minutes, summary, photos, …) about the events below are available on the ACCORD website, the link to the relevant pages are indicated:

- 14 January 2021: Kick-off LTM meeting
- 20 January, 12-19 February: interviews and meeting of the selection panel for the Area Leaders and CNA positions
- 2 March 2021: ACCORD Bureau meeting
- 8 March 2021: 1st ACCORD Assembly
- 25 March - 27 August: 13 Management Group meetings, generally on Friday mornings twice per month: summary of the meetings
- 30 March: 1st PAC meeting
- 12-16 April: 1st All Staff Workshop fully virtual BIG conference:
    - 250 colleagues joined us for the 1st ACCORD All Staff Workshop, many ACCORD colleagues of course, but also colleagues from ECWMF and the COSMO scientific PM.
    - The All Staff Workshop does befit its name: many of the 290 ACCORDers (persons who worked on the Rolling Work Plans of the consortium over the last 3 years, representing 125 annual Full Time Equivalent) attended the sessions they were interested in and participated to the exchanges with the speakers and the chairs, including in the chat tool.
    - The biggest session gathered 180 simultaneous attendees (opening), the others between 130 (Data Assimilation, dynamics, EPS and Meteorological Quality Assurance), 150 (physics, posters introduction), 160 (surface), 170 (code and system development).
    - During the closing session, the new Area Leaders presented a summary of the session in their area and their plans for the future, with still almost 130 persons connected on this Friday afternoon session.
    - The speakers, posters' presenters and chairs have prepared more than 100 documents for 70 talks, 20 posters, 9 sessions and 2 side-meetings.Automatic word wrap
    - The sessions have been recorded and the 31 hour video recordings are available. The content of the discussions in the chat have also been kept.
    - All this material is available on the ASW2021 webpage.
- 12 April: 1st LTM meeting

- 4 May: 1st STAC meeting
- 16 June: Bureau meeting
- 17 June: LTM information meeting on DestinE
- 2 July: 2nd ACCORD Assembly meeting

Many on-line thematic meetings (WD) were organised by the MG with the team or WG in their area. They are announced on the relevant part of the wiki, that also contains the material and the conclusion of these meetings.

3 Future events

The Management Group resumed their every other Friday morning meetings at the end of August. Additional meetings may be added (reporting on the RWP2021, finalization of the RWP2022, ..).

Although not an ACCORD event, but with the participation of many ACCORD colleagues, the EWGLAM meeting, initially planned in Brussels will take place as a fully online event, between 27 September and 1 October.

The 2nd LTM meeting will be organised on the 4th of October, just after the EWGLAM/SRNWP virtual meetings as it was difficult to find a comfortable time slot during the EWGLAM.

The STAC will meet on the 15th of November to review the RWP2021 status and advice on RWP2022

The end of the year Assembly will be held on the 8th of December (2021 reporting and preparation for 2022), and will be prepared by a Bureau meeting on the 19th of November 2021.

The big event in 2022 will be the 2nd All Staff Workshop on 4-8 April in Ljubljana.

Precipitation extremes in NWP

Bent H Sass, Danish Meteorological Institute

1 Introduction

There is a significant focus on the prediction of extremes today, both in climate prediction and in numerical Weather prediction (NWP), e.g. due to potential large damage from flooding due to extreme precipitation events and related loss in life and properties. As an example, huge amounts of precipitation resulted in severe floodings 14 July 2021, affecting Germany, Belgium, Luxembourg and the Netherlands.

Traditional verification schemes used in NWP do not verify extreme precipitation explicitly as part of routine verification. As a consequence, there appears to be a need for developing and testing verification schemes suited for verifying the model's ability to forecast large precipitation amounts with sufficient skill. Adequate assessment of skill will also make it possible to diagnose if new model developments lead to improved forecasting of extreme precipitation.

It is a prerequisite that precipitation can be analysed and forecasted as full fields in order to compare the details of the fields spatially. A precipitation analysis based on combining information from calibrated radars in Denmark with measurements from ground based rain gauges is available and may be compared with fields from HARMONIE-AROME run at DMI, see documentation of the method for verification in Sass (2021).

There are several dilemmas when considering extreme precipitation, e.g. `true´ extremes are rare phenomena which complicates verification from a statistical point of view, and one may ask how large values are needed to qualify as extremes. Since the model prediction has a finite resolution it is necessary to consider spatial (and temporal) effects, e.g. the effect of neighborhood width. Specific user needs may be taken into account when forecasting extreme precipitation. A basic consideration involves defining and identifying extreme points in a NWP domain, and to measure how close is the agreement between analysed extremes and forecasted extremes.

In section 2 a scheme is presented that verifies the forecasting of extremes based on the accumulation of information from daily runs of HARMONIE-AROME in DMI. An early version of the present idea, based on verifying the model's ability to forecast high precipitation amounts in observation points, goes back to Sass and Yang (2012). A documentation of a more general scheme named **SLX** (**S**tructure of **L**ocal E**X**tremes), based on spatial verification, has recently become available. (Sass 2021, Sass 2020). A setup in DMI has been developed which enables verification of the model's ability to forecast analysed precipitation as a function of the level of the extremes. The setup for this is described in section 2. An example of a cloudburst occurring in Denmark on 28 May 2021 is described in section 3. The example highlights current challenges related with predicting extreme precipitation. Section 4 provides examples of statistics from using SLX in 2021. Finally, some conclusions and outlook is presented in section 5.

2 A setup for verifying the prediction of precipitation extremes

The present verification scheme computes for every forecast its ability to forecast the highest value(s) analysed, becoming the `extreme´ value of the day. The advantage of this approach is that an assessment in terms of a score can be computed for all forecasts. A verification of `true extremes´ can

be done as a `post processing´ if a computed score is stored together with the actual extremes occurring on that day. When considering accumulated information over many forecasts the statistics of the score may be computed as a function of the `extreme´ values involved. Minima is included in the assessment since minima in some cases, that are dominated by large precipitation amounts, appear as `extremes´ in the context of the actual forecast. The nomenclature of 5 involved scores is:

SLX_{ob_max} : How well does forecast maximum match observed maximum in its neighbourhood ?

SLX_{ob_min} : How well does forecast minimum match observed minimum in its neighbourhood ?

SLX_{fc_max} : How well does observed maximum match forecasted maximum in its neighbourhood ?

SLX_{fc_min} : How well does observed minimum match forecasted minimum in its neighbourhood ?

SLX_{ave} : Combined average score = ¼ ($SLX_{ob\_max} + SLX_{ob\_min} + SLX_{fc\_max} + SLX_{fc\_min}$) .

The scores are computed using all points in a neighborhood of width L. If L = 0 then comparisons are done in single points. Multiple extreme points may occur, e.g. in the case of minima (zero precipitation as an example). Maxima may be defined with a certain tolerance, e.g. points qualifying as maxima may deviate from the highest value with a small absolute value of tolerance, alternatively with a small percentage of the absolute maximum.

The final value of one of the scores is defined as the weighted average between scores of individual points qualifying as extremes. A score function defines the assessment of the score between 0 and 1, with 0 meaning the poorest agreement between forecast and analysis, and 1 meaning perfect match. The current choice gives S=0 for forecast exceeding observation by a factor of 5. A score of 0 is achieved if the forecast is less that 10 % of the observed value. For details of the computational procedure, see the documentation of the scheme: (Sass 2021, http://dx.doi.org/10.1002/met.2015)

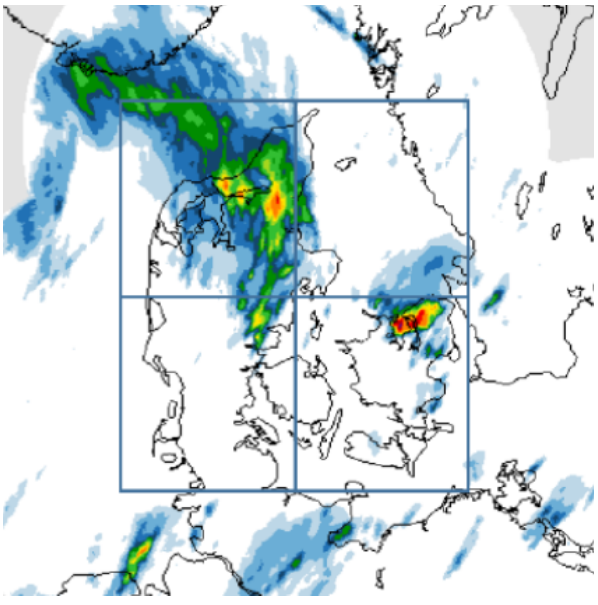


Figure 1: Model setup used in DMI covers Denmark. Analyses and forecasts involve 4 subdomains, and scores are computed for each subdomain. An example of 3 hour analysed accumulated precipitation is shown in the figure.

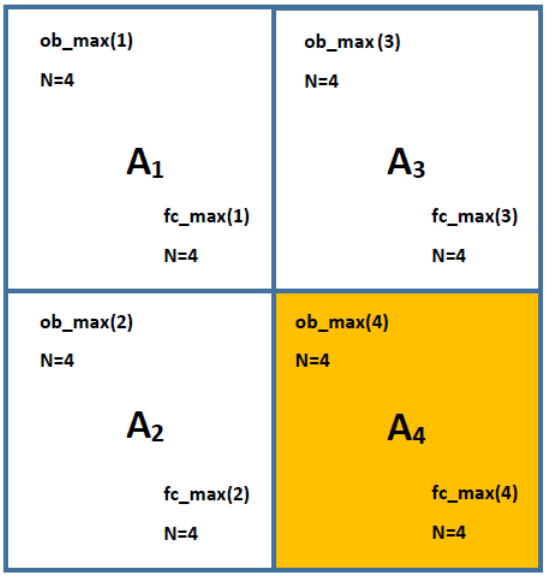


Figure 2: Schematic representation of 4 subareas where forecasted and analysed accumulated precipitation are compared. Values of the subarea with largest extreme (analysed or forecasted) are stored for use in conditional verification (see text).

For the reasons of flexibility the scheme is defined in a number of subareas (figure 1 and figure 2) The different score components are computed separately for each subarea. Currently the verification domain with forecast and analysis available covers Denmark, using 4 subareas (figure 1).

The output of the scores of any forecast is accumulated in a SQLite file. In order to assess forecast skill as a function of the forecasted- and observed maxima, in a statistical sense, the maximum values of the subarea with the largest extreme (analysed or forecasted) are stored for use in conditional verification. In this way various statistics may be computed as a postprocessing of the output file. By comparing the stored forecast maximum with the associated score (SLX_{fc_max}) of the given forecast , and the observed maximum with the associated score (SLX_{ob_max}) , we may compute the scores statistically using all forecasts in a period. Figure 3 shows these scores as a function of maxima in certain intervals. We see that the values of the two scores (in percent on the y-axis) tend to decrease as the maximum precipitation increases. This is an important and expected result since one will expect that it is increasingly difficult to get a high score as the precipitation threshold increases towards `true´ extremes. Figure 3 confirms this. It is based on all forecasts at DMI of the main HARMONIE-AROME model run over Denmark during the 3-month period from May 2021 – July 2021.

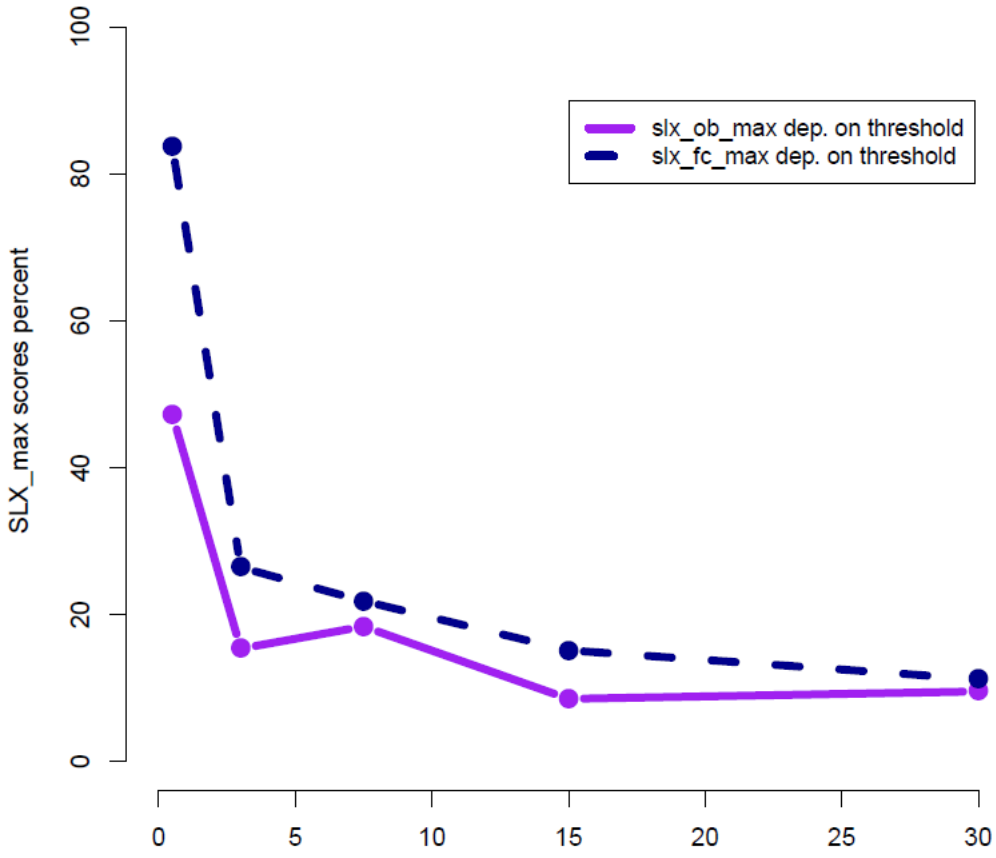


Figure 3: The scores SLX_ob_max (purple curve) and SLX_fc_max (dark blue) as a function of maximum precipitation thresholds, computed as averages over all forecasts during May-2021 –July 2021. Forecast lengths up to +48 hours and 3 hour accumulated precipitation are considered. The neighbourhood width =0 (grid scale)

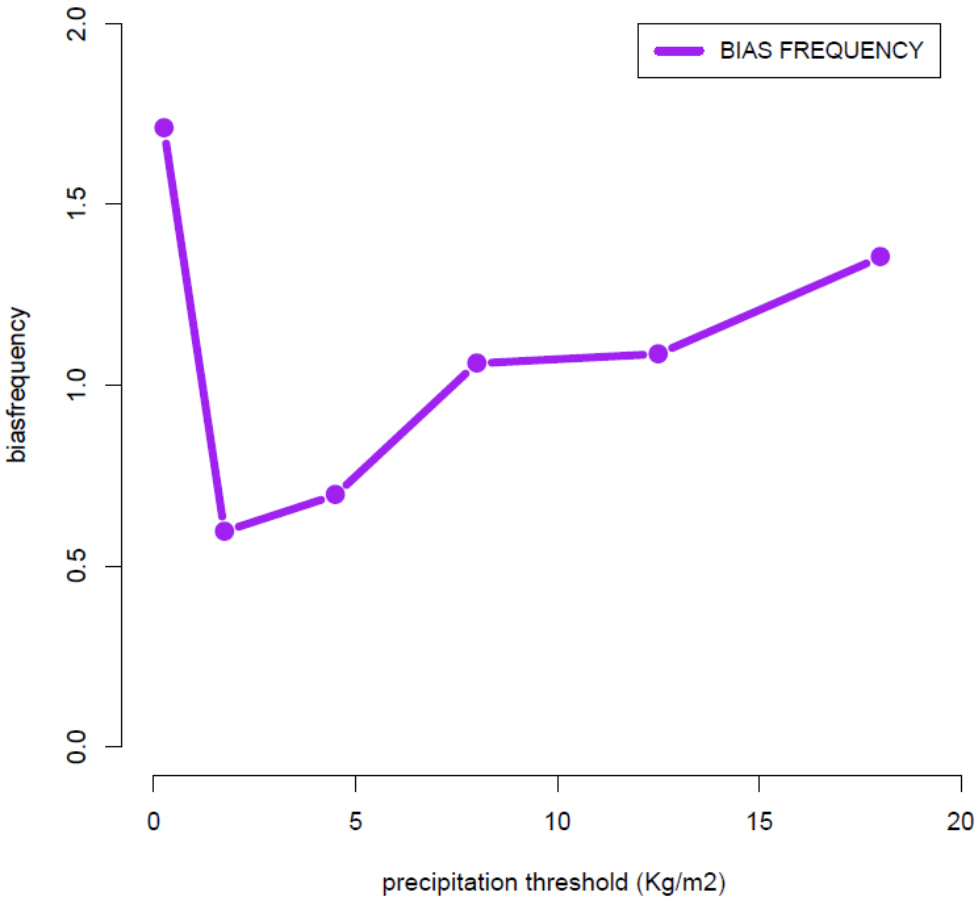


Figure 4: Bias frequency as a function of precipitation threshold, considering stored forecast maxima and observed maxima in relevant subareas. All forecasts during May 2021–July 2021 are used.

Similarly a special frequency bias may be computed from the accumulated statistics, based on the stored values of forecast maxima versus observed maxima in a subarea. This is shown in Figure 4 this for May 2021 –July 2021. It is seen that the frequency bias is positive for small amounts (< 1 Kg/m^2) and is less than 1 for accumulation between 1 and 7 Kg/m^2 . This seems to be in agreement with subjective estimates among forecasters that showers with relatively modest precipitation amounts are often missed by the forecast, e.g. in cold air advection. Frequency bias for accumulation thresholds above 7 Kg/m^2 has been somewhat higher than 1 in this type of assessment.

3 Forecasting cloudbursts : An example

An extreme cloudburst episode happened on 28 May 2021 in southern part of Denmark (location near the city of Svendborg on the Island of Funen).
The accumulated precipitation in an area of size less than 10*10 km was up to ~ 80 kg/m^2 in 3 hours, between 15 UTC and 18 UTC, see Figure 5. The convection evolved in a convectively unstable atmosphere, with weak surface winds and some warm air advection in the lower troposphere. Figure 6 shows the 3 hour accumulated precipitation at 18 hours forecasted with HARMONIE-AROME (left), run at initial time 00 UTC. To the right is shown is the corresponding result of a 12 hour forecast from 6 UTC.

At first sight both forecasts appear reasonable since they represent the observed cloudburst in the correct region of the country. However the accumulation is about 15-20 % of the observed one. The forecast from 00 UTC has some localization error. The SLX scores related with maxima are therefore small using the current SLX function considering extreme maxima. With neighborhood size = 0 the score is zero at some forecast origin times. Moreover the forecast maxima in the eastern parts of Jutland occur at locations with no observed precipitation and thus represent “false alarms”. Hence small scale cloudburst events can be very challenging ! The SLX scheme provides a low score - as it should - in the case of inaccurate precipitation amount and location.

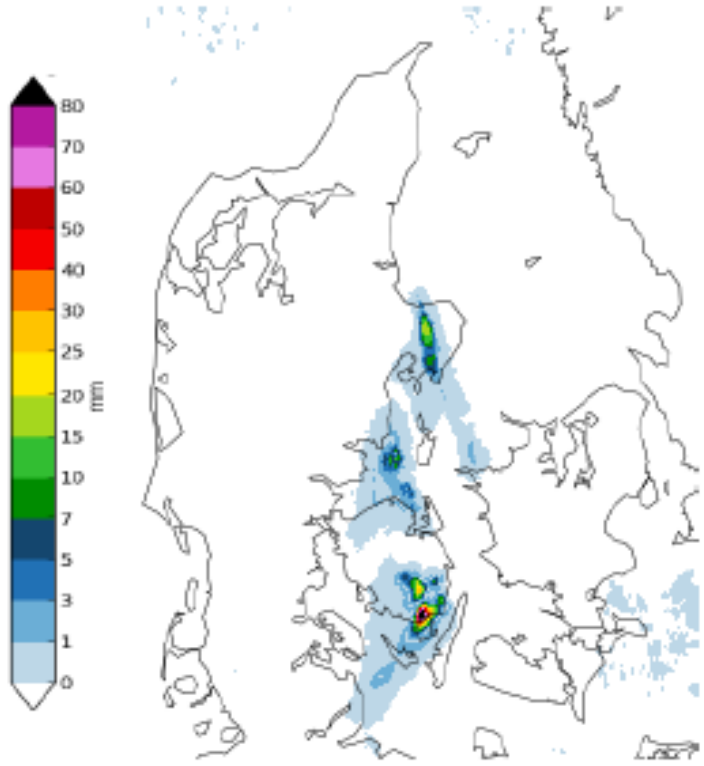


Figure 5: Observed 3 hour precipitation across Denmark between 15 UTC and 18 UTC on 28 May 2021. The small scale cloud burst is visible in the southern part (Island of Funen).

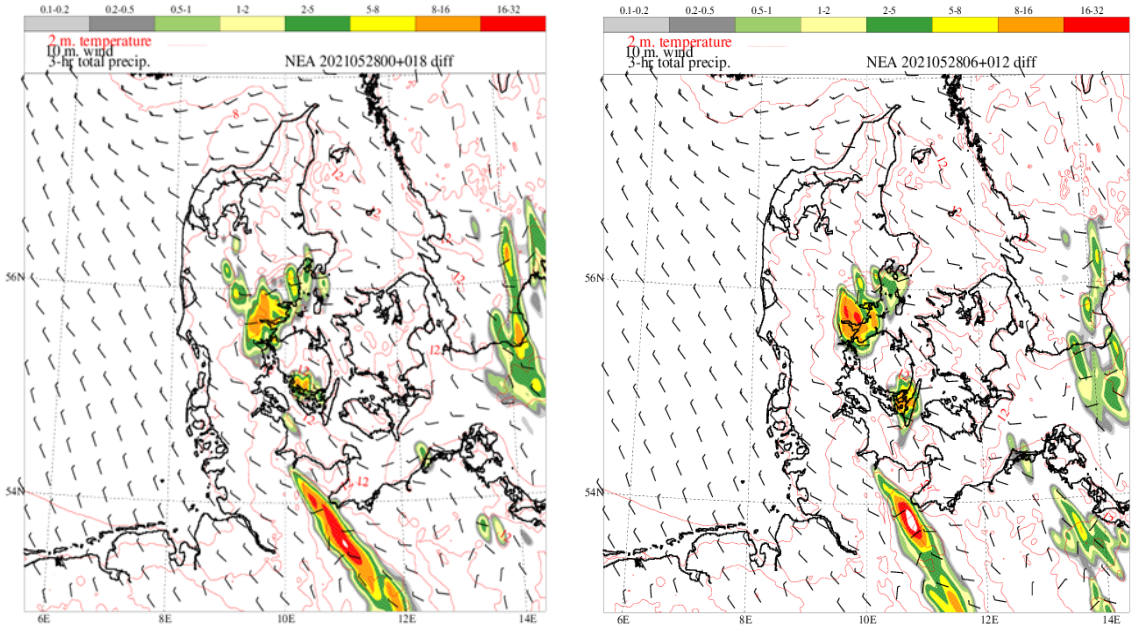


Figure 6 : 18 hour hour forecast (left) with HARMONIE –AROME at DMI valid a the same 3 hour period as figure 5. The 12 hour forecast valid at the same period is shown to the right.

4 Statistics from SLX scheme in daily runs

The results from running the SLX scheme on a daily basis are currently available on *shiny.hirlam.org*. The average of all forecasts in a month are verified separately for each forecast length up to +48 hours. Examples of verifying all scores are shown in figures 7- 10. The results have been interpolated to a high resolution grid of 0.75 km *0.75 km.

In figures 7 - 8 the results of March 2021 are compared with the corresponding ones of May 2021. It is seen that scores of maxima are smaller than those of minima (when verified on a monthly basis). The comparison also reveals that the scores of May are substantially lower than for March.

In Figures 9 – 10 the impact of neigborhood size is illustrated, comparing results of March 2021 at neighborhood widths of 0 and 15. The scheme tends to produce higher scores when increasing the neighnorhood width. This is understandable for extreme cases such as the cloudburst mentioned in section 3 where the localization error of forecast maximum is an important issue. - In single cases higher scores are not always obtained when neighbourhood width is increased, but it is normally the case when many forecasts are averaged.

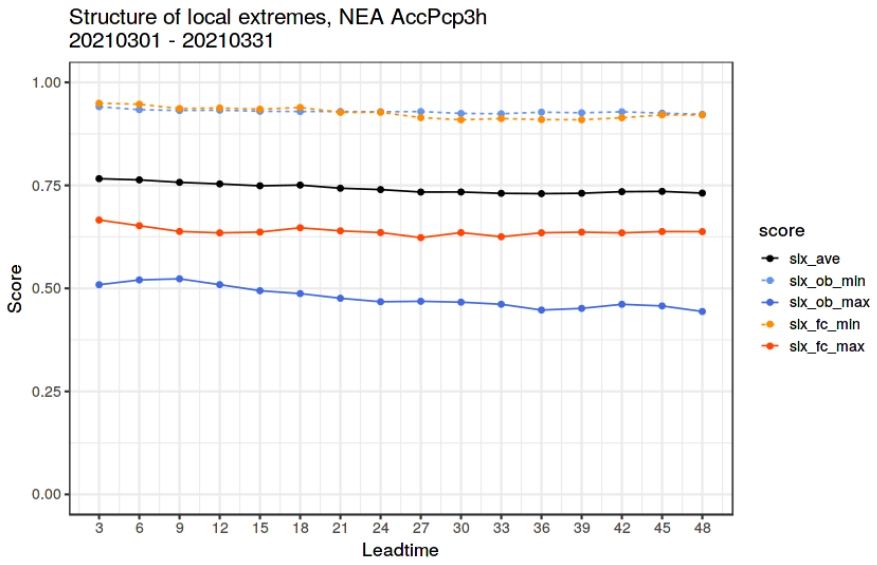


*Figure 7: Verification results of all SLX-score components during **March 2021**.*
All components and the average (slx_ave) are computed according to forecast length.
Average scores are computed using all forecasts available during the month.
[Neighborhood size = 0, points used in comparison =1 point]

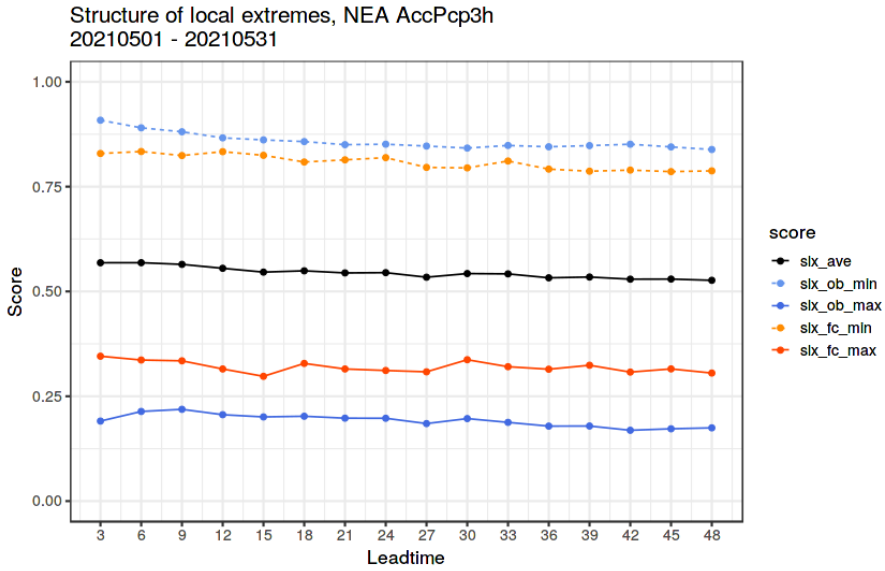


*Figure 8: Verification results of all SLX-score components during **May 2021**. All components and the average (slx_ave) are computed according to forecast length.*
Average scores are computed using all forecasts available during the month.
[Neighborhood size = 0 , points used in comparison =1 point]

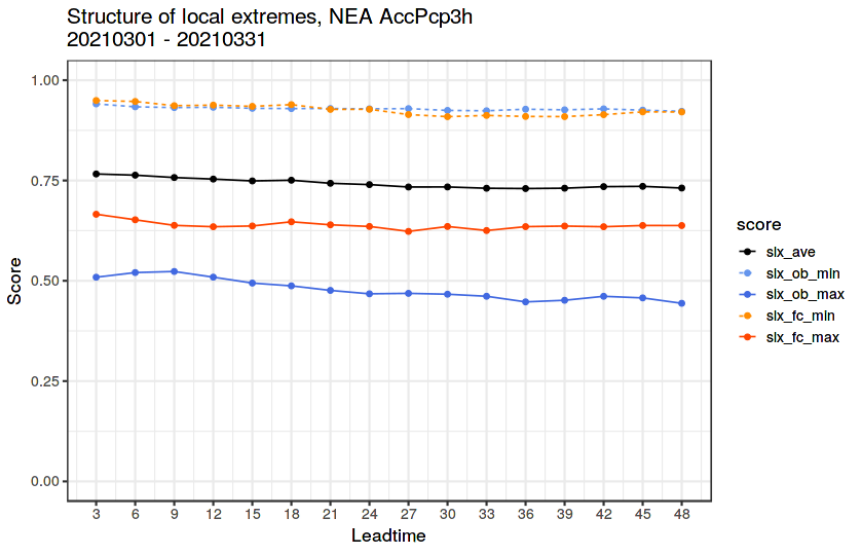


*Figure 9: Verification results of all SLX-score components during **March 2021**.*
All components and the average (slx_ave) are computed according to forecast length.
Average scores are computed using all forecasts available during the month.
[Neighborhood width = 0 points, - points used in comparison =1 point]

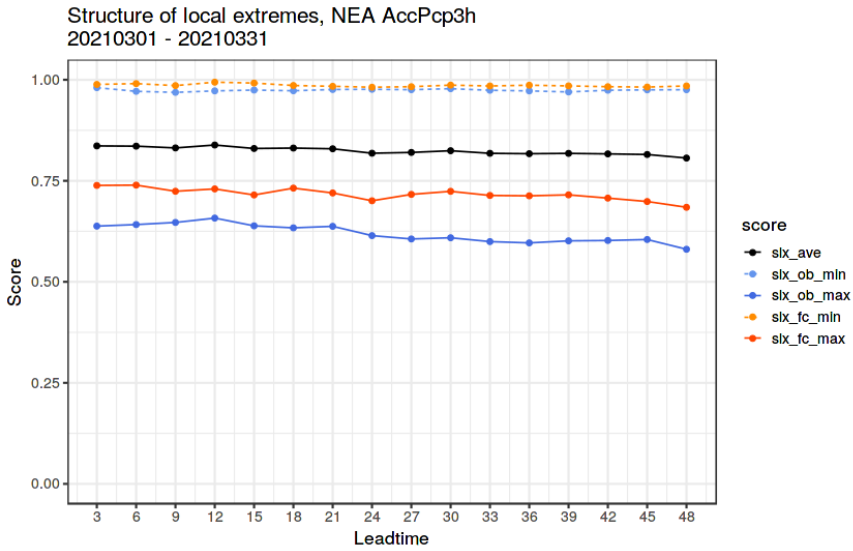


*Figure 10: Verification results of all SLX-score components during **March 2021**.*
All components and the average (slx_ave) are computed according to forecast length.
Average scores are computed using all forecasts available during the month.
(Neighborhood width = 15 points) .
*[points used in comparison = (2*15+1) * (2*15+1) points].*

5 Conclusions and outlook

Documentation of a scheme named **SLX** (**S**tructure of **L**ocal E**X**tremes), verifying the ability of a NWP model to forecast `extreme´ highest and lowest precipitation occurring on a given day, has recently become available [http://dx.doi.org/10.1002/met.2015]. It is based on spatial verification of a forecast- and analysis field.

Operational use of the SLX scheme in combination with storage of maxima in subareas enables conditional verification of scores depending on the level of maxima. When studying the impact of a new model version it is possible to compare if large values of extreme precipitation are predicted with higher scores compared with a current model version.

A further level of detail may be obtained by studying histograms of SLX components or alternatively absolute errors of forecasts. This provides insight to the fraction of the largest errors occurring in a given model version. Generalization of SLX to ensembles is rather straight forward. e.g. the histograms will change as a result.

6 References

Sass, Bent H. (2021) : A scheme for verifying the spatial structure of extremes in numerical weather prediction: exemplified for precipitation, *Meteorological Applications* , **http://dx.doi.org/10.1002/met.2015**

Sass, B.H. 2020: Forecasting spatial structure of local precipitation extremes : International Verification Methods Workshop Online (2020-IVMW-O), Nov.2020, https://jwgfvr.univie.ac.at

Sass, B.H. and Yang, X. 2012: A verification score for high resolution NWP: Idealized and preoperational tests. HIRLAM Tech. Rep. No 69, 28 pp, 2012. [Available from www.hirlam.org]

Operational Numerical Weather Prediction Models Verification at Météo Algérie

Sara Chikhi[(1,2,*)], Mohamed Mokhtari[(2)], Abdenour Ambar[(1,3)],Islam Bousri[(2)], Mohamed Arab Benamara[(2)]

(1) Faculty of Physics, Sciences and Technologies University USTHB, Algiers, ALGERIA.
(2) Numerical Weather Prediction Department, Office Nationale de La Météorologie, Algiers, ALGERIA.
(3) Faculty of Earth Sciences and Terittory Planning, Sciences and Technologies University USTHB, Algiers, ALGERIA.
(*) Contact person : ***scusthb@gmail.com***

1 Introduction

The weather forecast models verification is a vital step in the numerical weather prediction (NWP) system. This step gives us a general view of the uncertainties and the behaviour of NWP models against the reference. This information is crucial for forecasters who use these models in operations.

To improve the existing model verification suite at Météo Algérie, an action was included in the support program for the implementation of the Association Agreement between Algeria and European Union (P3A). This program having started on March 2020 and it has supported by Météo France and the Finish meteorological institute (FMI).
This action has two main objectives:

- Rewrite the code of the verification chain of our operational models, ALADIN and AROME, according to international standards and the criteria defined in the WMO note 485 (OMM, 2019).
- Define and compute some monitoring indicators for our NWP system in order to verify that the forecasts improve with each modification in our system.

Ultimately, this approach will facilitate the interpretation of the verification scores of our models and the comparison with other scores, particularly within the ACCORD community.

2 NWP models verification at Météo Algerie

The operational NWP system at Météo Algeria is based on the ALADIN, AROME and ALADIN-DUST models. This system runs twice a day for 00 and 12 networks. Table 1 summarizes the main characteristics of the operational configurations.

A new automatic chain of NWP models verification which meets international standards has been set up at Météo Algérie. In this chain we used three statistical indictors: Bias, root mean square error (Rmse) and standard deviation (Sd).

Bias: The bias measures the systematic error of the model. It is positive when the forecast overestimates the value of the parameter, negative when it underestimates it. It is calculated as follows:

$$Biais = \frac{1}{n_j x n_p} \sum_{j=1}^{n_j} \sum_{p=1}^{n_p} (Forecast - Reference) \quad \text{(E.1)}$$

Root mean square error (Rmse): Used to measure the distance between the expected field and that of the reference. It is calculated as follows:

$$RMSE = \frac{1}{n_j x n_p} \sum_{j=1}^{n_j} \sum_{p=1}^{n_p} (Forecast - Reference)^2 \quad \text{(E.2)}$$

Where : (n_j) is the number of day, (n_p) is the number of grid points and (Reference) is the verifying observation or analysis.

Standard Deviation (Sd): The standard deviation represents the variability of the error. When the value is low, the majority of errors are explained by a systematic bias of the forecast model. It is given by equation 3.

$$ECT = \sqrt{RMSE^2 - Biais^2} \quad \text{(E.3)}$$

The model verification is carried out against observation and ARPEGE analysis.

Table1: Operational models configurations at Météo Algérie

Models		ALADIN		AROME		ALADIN-DUST	
		Forecast	Fullpos	Forecast	Fullpos	Forecast	Fullpos
Cycle		CY43T2.bf.10	CY43T2.bf.10	CY43T2bf.10	CY43T2.bf.10	CY43T2.bf.10	CY43T2.bf.10
LBC Model		ARPEGE	-	ALADIN	-	ARPEGE	-
LBC Frequency		3h	-	1h	-	3h	-
Output Frequency		1h	1h	1h	1h	1h	1h
Horizontal Resolution		6km	0.08° X 0.08°	3km	0.03° X 0.03°	14km	0.11° X 0.11°
Vertical resolution		70	-	41	-	70	-
Time step		514s	-	60s	-	514s	-
Domain	**Lat**	-	18.5°N-46.5°N	-	28°N-40°N	-	18.5°N-46.5°N
	Lon	-	11°W-17°E	-	03°W-09°E	-	11°W-17°E
Number of points		600X600	350 X 350	500X500	400 X 400	250X250	240 X 240
Network		00, 12 UTC	00, 12 UTC	00, 12 UTC	00, 12 UTC	00, 12 UTC	00, 12 UTC
Time range		72h	72h	48h	48h	72h	72h

2.1 Verification against ARPEGE Analysis

The first step of the verification against ARPEGE analysis is the preparation of the ARPEGE analysis files for 00h, 06h, 12h and 18h networks. This task is realised using directly Fullpos configuration by projecting the ARPEGE analysis fields over the ALADIN and AROME Fullpos grids. The verification is done for the following parameters :

- **Surface level:** Temperature at 2 meters, Wind speed at 10 meters, relative humidity (Rh) and Mean sea level pressure (MSLP)

- **Altitude levels (1000, 850, 700, 500, 400, 300, 200 and 100 hPa):** Temperature, Wind, Humidity and Geopotential.

It should be noted that the temperature variability above the sea is low. This will impact the averages of the scores on the whole domain. To avoid this inconvenience, a land-sea filter is applied for temperatures at 2 meters.

Figure 1 shows the correspondence between the ARPEGE analysis and the forecast output.

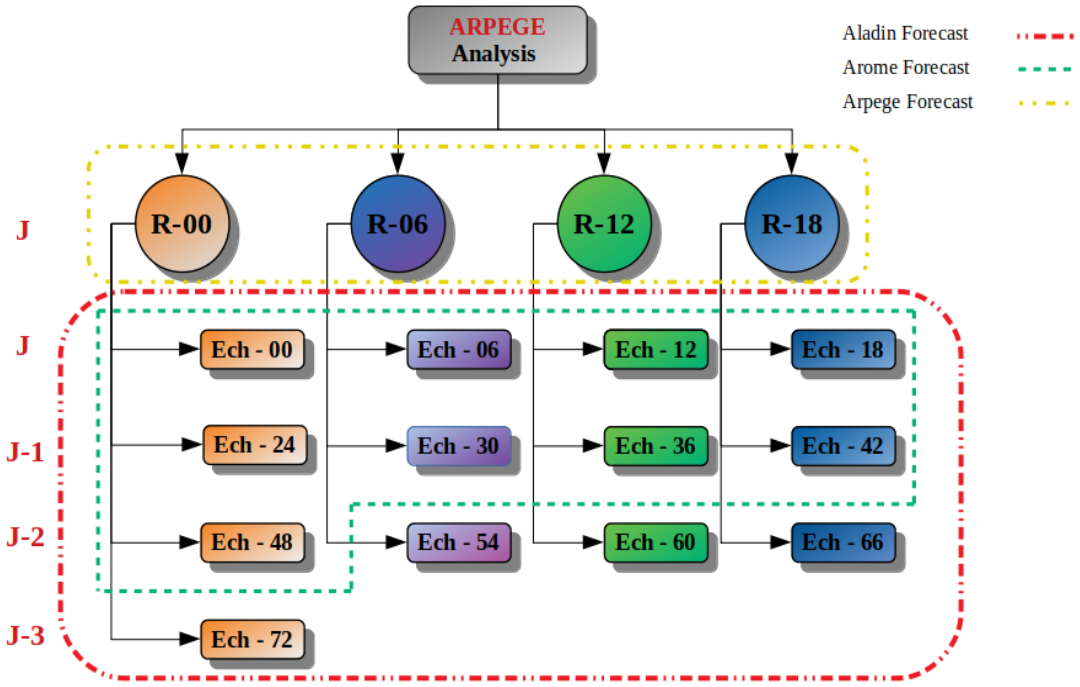


Figure 1: Correspondence between ARPEGE analysis and Forecast output.

2.2 Verification against observation

Currently, the verification against observations is performed only based on data from SYNOP messages, so the verification time step in three hourly. This data is retrieved from our GTS network. Sometimes we use the OGIMET platform (https://www.ogimet.com/synopsc.phtml.en) to complete the missing data. Figure 3 shows the plotting of observation stations included in the ALADIN domain for a typical day. The model data is brought back to the observation point by taking the nearest grid point.

Usually, the altitude of the observation point and the one of the corresponding grid point are not the same, which causes systematic differences that impact the scores. A correction of the temperatures due to the vertical gradient is carried out by taking into account this difference in altitude.

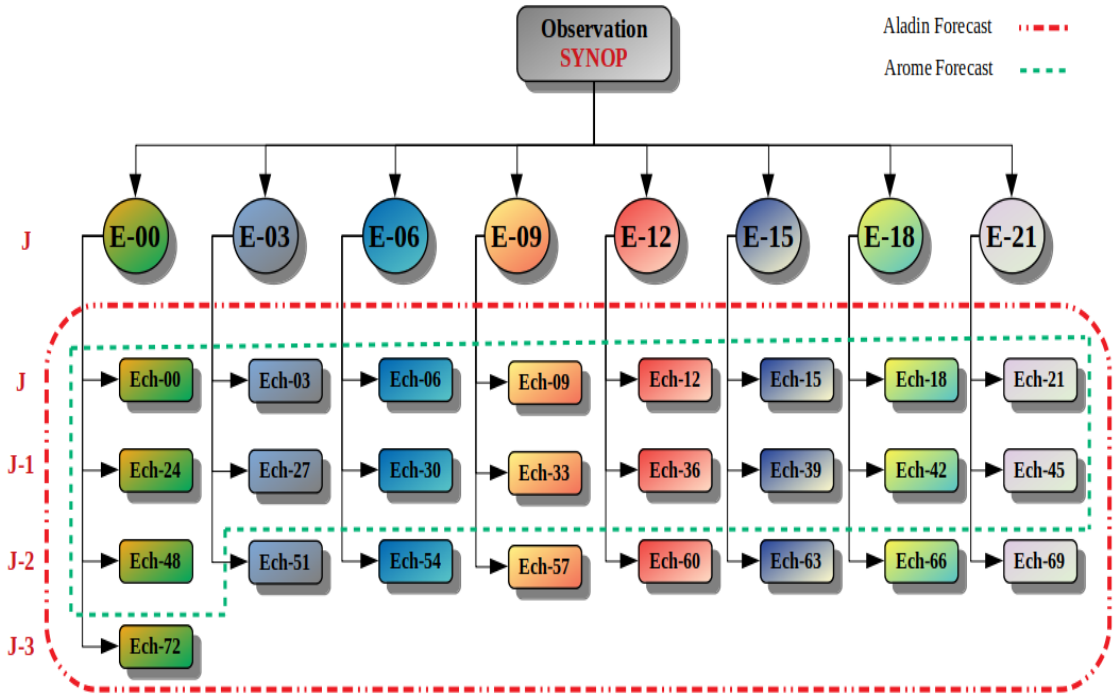


Figure 2: Correspondence between observation data and Forecast output.

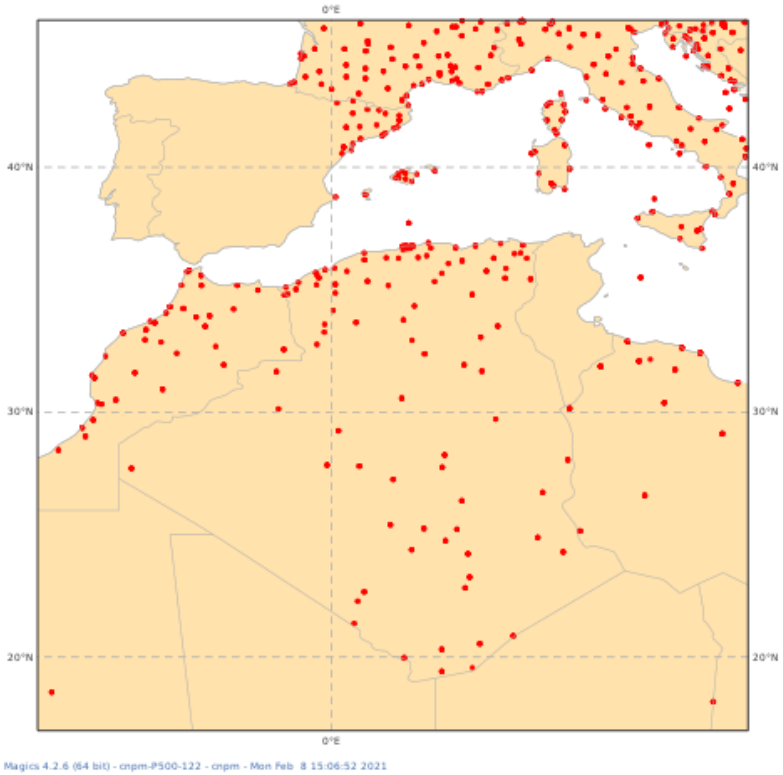


Figure 3: Spatial distribution of the SYNOP observation stations for a typical day received on GTS at Météo Algérie

3 Results

In this section, we present the monthly scores for ALADIN and AROME models compared to the ARPEGE analysis for March 2021. Figures (4.a) and (4.b) shows the scores of the ALADIN and AROME models for the surface parameters. The scores evolution shows the degradation of the forecasts over the hours with a reasonable values. This means at least that the new verification system is working correctly. We also notice strong temporal variability of AROME compared to ALADIN. For example, AROME temperatures showed a strong positive biases (3.5°C) during the night and a negative biases at noon (-0.5°C). This variability is less important for the other AROME parameters. However, ALADIN showed a weaker temporal variability compared to AROME, which could be clearly seen for the temperature parameter.

These strong positive biases of temperature calculated for AROME overnight can be explained by the damping of the cooling process of the Earth-Atmosphere system by the mechanism of energy loss by infrared radiation emitted by the surface. The comparison between the scores obtained for the two models showed clearly that the ALADIN scores are much better to those of AROME, especially for the 2 meters temperature parameter. This is probably due to the AROME vertical resolution of the operational configuration which is very low (41 levels).

It is also important to notice the significant bias for wind parameter (figure 4.b) simulated by ALADIN during the first time step, which is not the case for AROME.

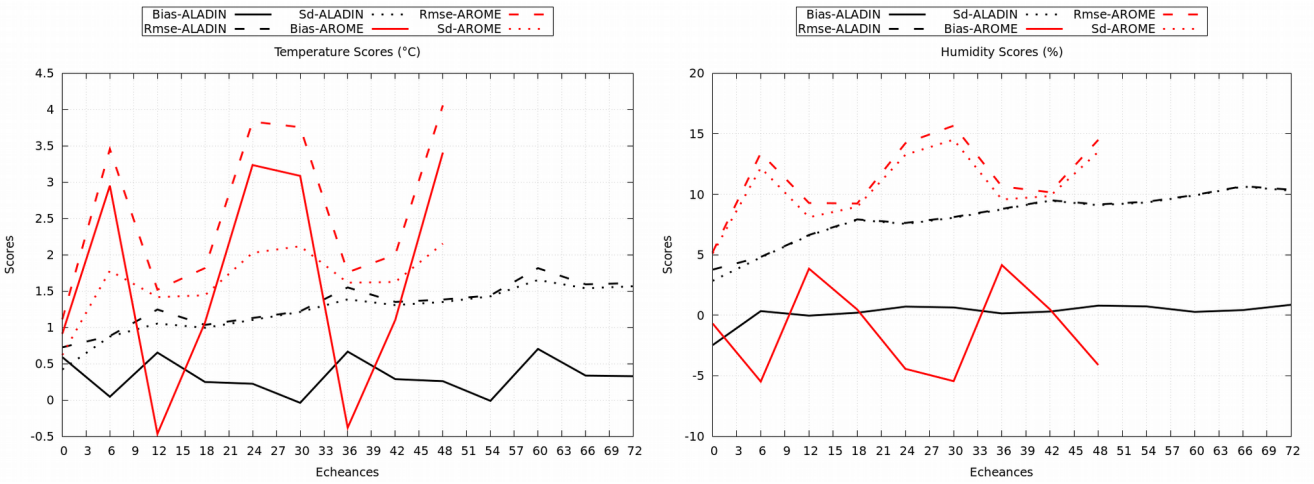


Figure 4.a: ALADIN scores (in black) and AROME scores (in red) for the surface parameters: Temperature and relative Humidity.

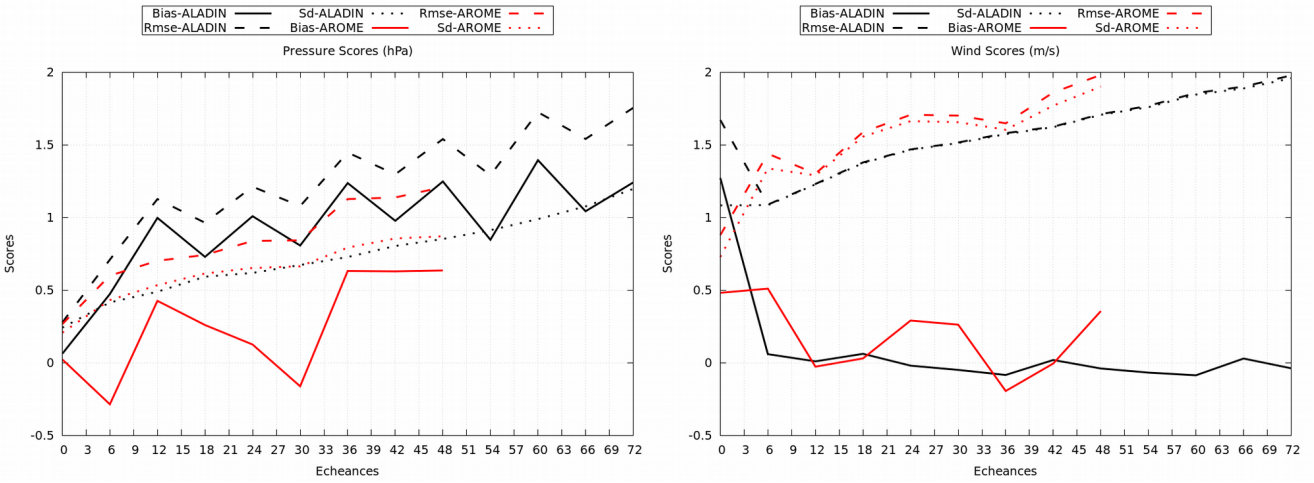


Figure 4.b: ALADIN scores (in black) and AROME scores (in red) for the surface parameters: Wind speed and MSL Pressure.

Figures (5.a) and (5.b) shows the scores of ALADIN and AROME models for the upper air parameters at 850hpa level. Unlike the surface level, temperatures bias at 850hpa showed a quite steady temporal variation for both AROME and ALADIN models. This variation is also noticed for humidity and wind speed, but not for the geopotential where it is a bit more important. Forecast degradation over time is detected only for geoptential and wind speed.

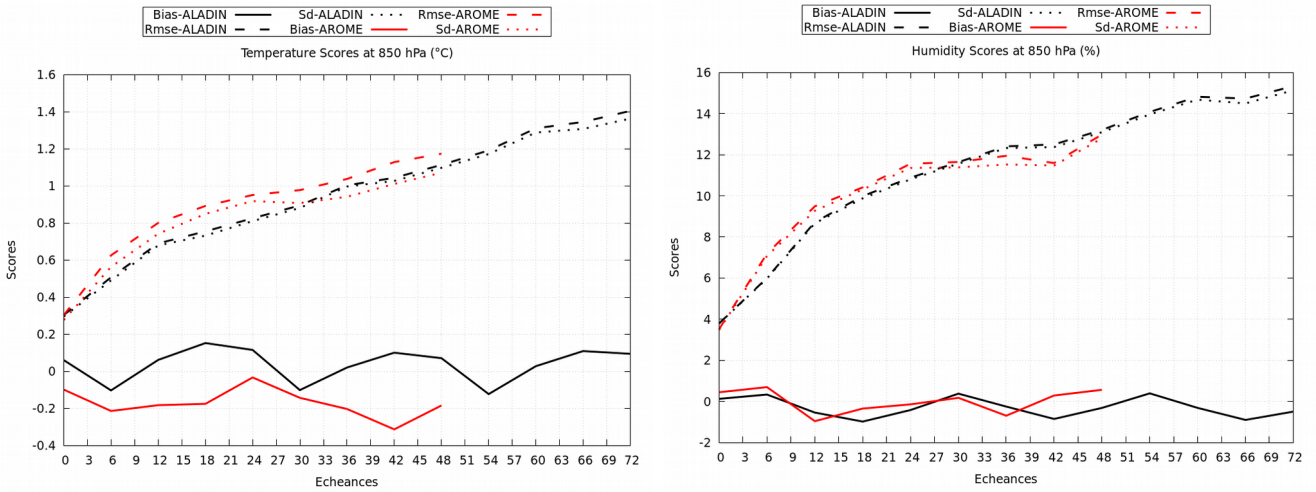


Figure 5.a: ALADIN scores (in black) and AROME scores (in red) for the upper air parameters (at 850hpa): temperature and relative humidity.

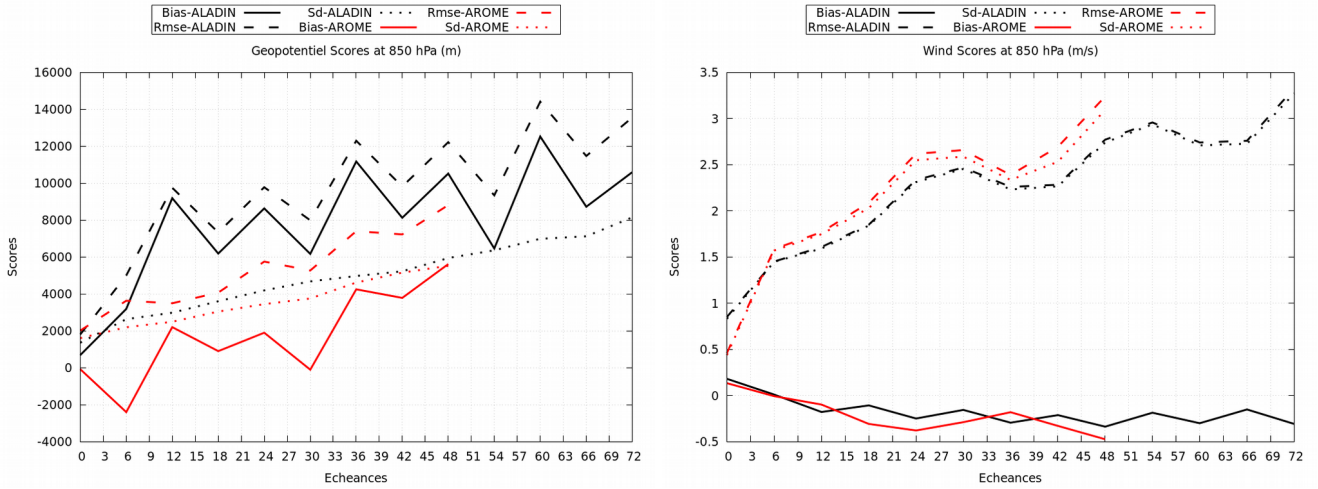


Figure 5.b: ALADIN scores (in black) and AROME scores (in red) for the upper air parameters (at 850hpa): geopotential and wind speed.

Figures (6.a) and (6.b) represents the scores of ALADIN and AROME models for the upper air parameters at 500hpa level. The results are almost similar to those obtained at 850at level. Temporal variation are steady for temperature, humidity and wind speed, while it is a bit perturbed for geopotential. Degradation of temperature forecast over time is much more clear at this level then the previous one (850hpa). However, this degradation is not actually detected for the other parameters at this level.

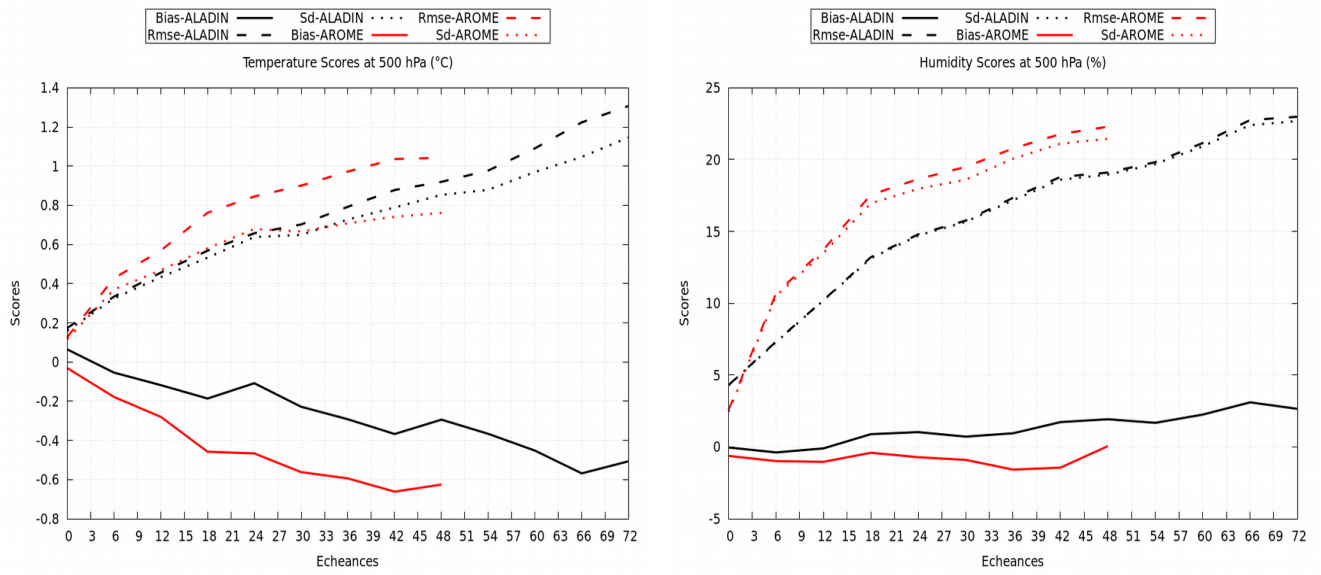


Figure 6.a: ALADIN scores (in black) and AROME scores (in red) for the upper air parameters (at 500hpa): temperature and humidity.

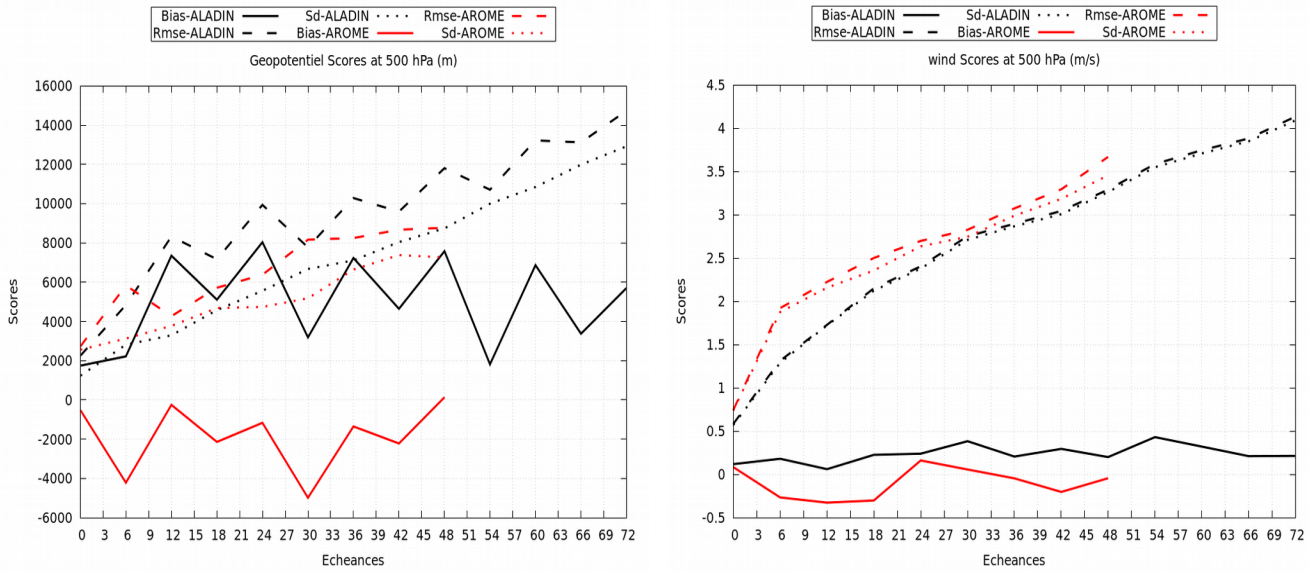


Figure 6.b: ALADIN scores (in black) and AROME scores (in red) for the upper air parameters (at 500hpa): geopotential and wind speed.

Figures 7, 8, 9 and 10 represents the root mean square error (rmse) vertical profile variation of temperature, humidity, geopotential and wind speed, respectively.

The temperature RMSE (Figure 7) showed the same shape along the vertical profile in both cases (ALADIN and AROME): we found high values at low levels (under 850hpa), it become less important between 850hpa and 300hpa, then it increased considerably to reach its maximum at around 200hpa. The difference between ALADIN and AROME is particularly seen in terms of temporal

evolution, where we noticed that RMSE values starts to increase after 45h to 48h for ALADIN, and after 21h-24h for AROME.

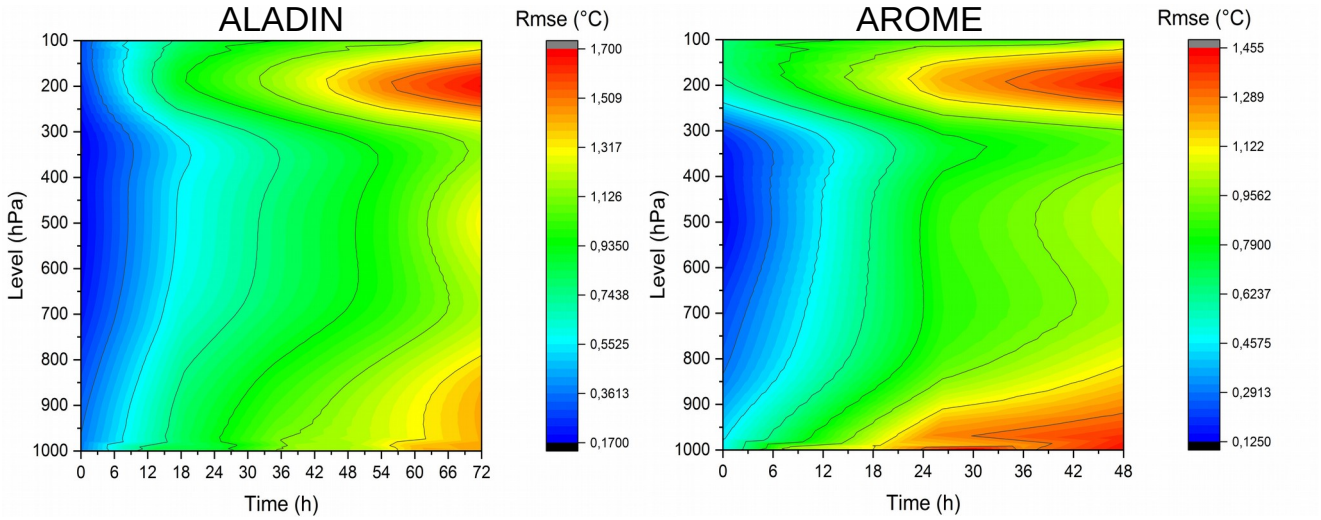


Figure 7: Vertical profile variation of the root mean square error (rmse) for ALADIN and AROME temperature forecasts.

The humidity RMSE (Figure 8) showed also the same shape along the vertical profile in both cases (ALADIN and AROME): the highest RMSE values are detected between 400hpa and 300hpa levels. The RMSE is getting worse moving forward in time for both configurations, but this degradation appears quicker in AROME (+18h) than in ALADIN (+36).

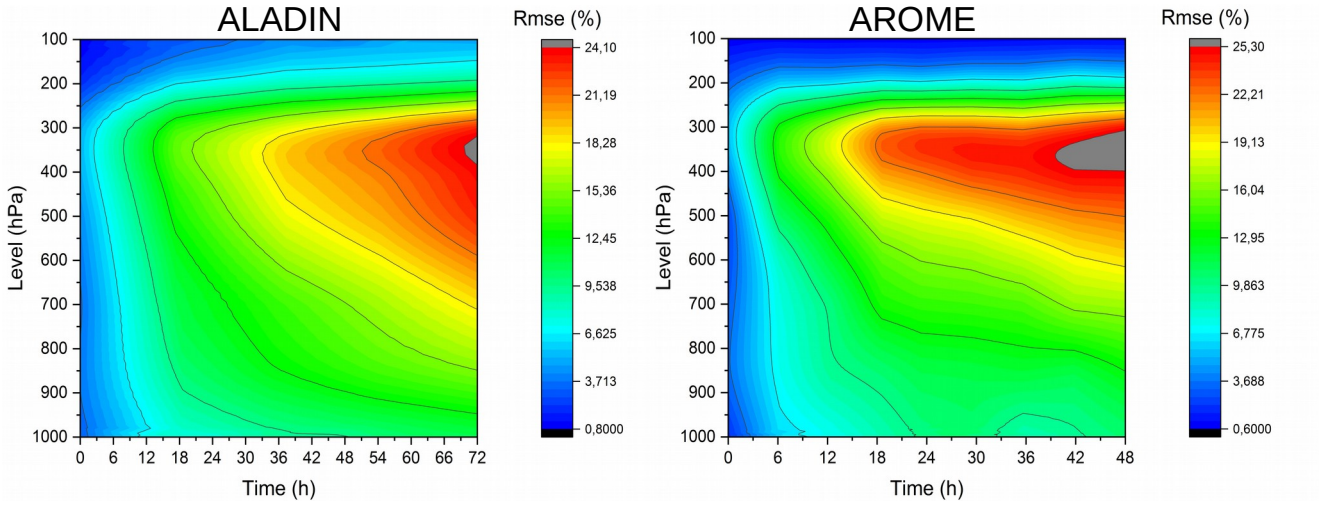


Figure 8: Vertical profile variation of the root mean square error (rmse) for ALADIN and AROME humidity forecasts.

Compared to the temperature and humidity RMSE, the geopotential RMSE (Figure 9) shows a bit different shape between ALADIN and AROME. It's true that, in both cases, the highest values are detected between 400hpa and 150hpa, but at lowest levels, the degradation started earlier in ALADIN (+06h) than in AROME (+21), particularly between 925hpa and 500hpa.

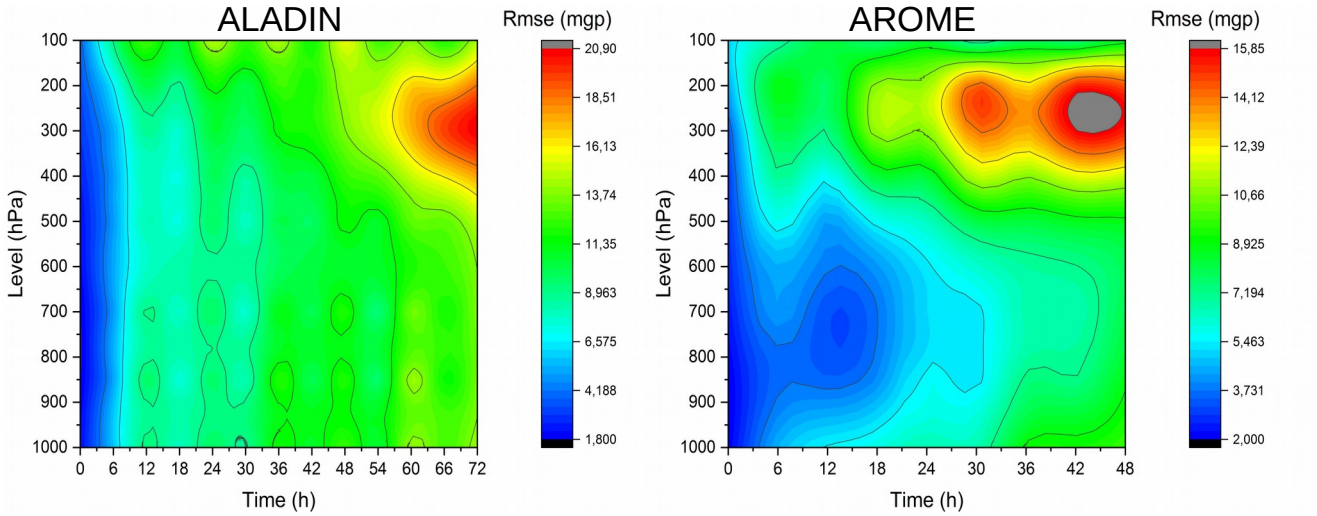


Figure 9: Vertical profile variation of the root mean square error (rmse) for ALADIN and AROME geopotential forecasts.

For the wind speed RMSE (Figure 10), we actually observe a similar results as those obtained for the humidity : the highest values are spotted between 350hpa and 250hpa. As noticed previously for the other parameters, the degradation starts much earlier in AROME (+21h) than in ALADIN (+45h).

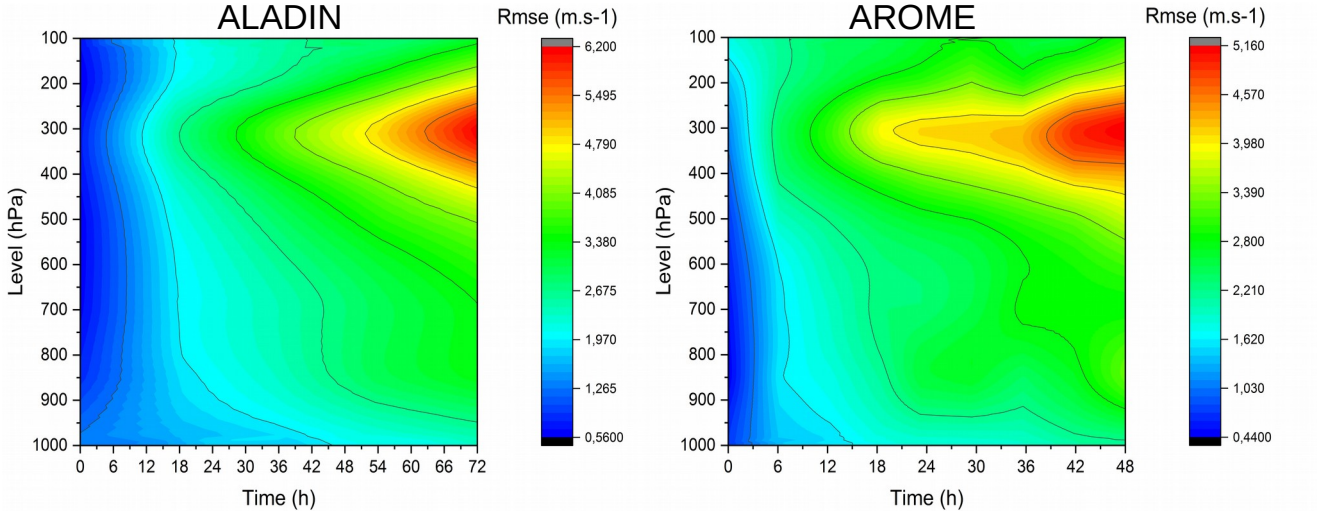


Figure 10: Vertical profile variation of the root mean square error (rmse) for ALADIN and AROME wind speed forecasts.

4 Conclusion

The main objective of this article is to present the new operational model verification system at Météo Algérie. The major task was to rewrite all the codes of the verification system (processing and post-processing) and to make sure that the verification methods used at ONM meet the international standards.

As a practical example, we presented in this article the first results obtained by using the new model verification system. It worth to mention that generally ALADIN forecasts showed a better scores than AROME. This could be explained the optimized configuration of AROME used at ONM (3km horizontal and 41 level), which is due to computational costs considerations. However, AROME is widely used by our forecasters in case of deep convection situations particularly during the transition seasons, spring and autumn.

The next step will be the computation of the verification scores over one year in order to define the monitoring indicators for our verification system. Also, we're working to extend this system to cover other parameters such as precipitation, extreme temperatures (maximum and minimum) and cloud cover. Another important work, in progress, is the use of the radiosonde observations.

Acknowledgement

We thank Mr Bruno Lacroix, expert at Météo France, for supporting this work in the framework of the P3A twinning project. His guidance are very useful.

5 References

[1] Joël Stein ; Jean Pailleux ; Fabien Stoop ; Marielle Amodei ; Olivier Dupont ; Mireille Mayoka ; Francis Pouponneau ; Isabelle Sanchez. La vérification des prévisions météorologiques à Météo-France. La Météorologie, 2015, N° 90 ; p. 40-49. DOI : 10.4267/2042/56837

[2] Daniel S. Wilks. Statistical Methods in the Atmospheric Sciences. 2006; Vol. 100; 2nd Edition., 2006. Academic Press; London.

[3] Organisation Météorologique Mondiale, Manuel du système mondial de traitement de données et de prévision, Annexe IV du Règlement technique de l'OMM, Edition 2019.

[4] Kallee Erola, Twenty-One Years of Verification from the HIRLAM NWP System, American Meteorological Society, DOI: 10.1175/WAF-D-12-00068.1,

Publications of ACCORD Poland team

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

1 Introduction

There were several peer-reviewed papers written in cooperation of the members of ACCORD Poland team last year. The range of topics cover multi-model (random forest method and its usage for the improvement of the quality of model forecasting products), PM10 particles' concentration in the bottom part of boundary layer in Kraków (and its investigation by UAV) as well as its classification of PM10 abundance by the vertical profile. Another point was the use of machine learning methods in forecasting the wind speed with the application of wind farms effectivity in Poland. The last one is the analysis of the derecho occurrence in Poland (in 2017) and possibilities of forecasting it by the ALARO and AROME models configurations nowadays.

2 PM10 concentration above Kraków

Publication I - summary

The paper shows wind shear impact on PM_{10} vertical profiles, in Kraków, southern Poland. The data used consist of background data for two cold seasons (Sep. 2018 to Apr. 2019, and Sep. 2019 to Apr. 2020), and data for several case studies from November 2019 to March 2020. The data is composed of PM_{10} measurements, model data, and wind speed and direction data. The background model data come from operational forecast results of AROME model. PM_{10} concentration in the vertical profile was measured with a sightseeing balloon. Significant spatial variability of wind field was found. The case studies represent the conditions with much lower wind speed and a much higher PM_{10} levels than the seasonal average. The inversions were much more frequent than on average, too. Wind shear turned out to be the important factor in terms of PM_{10} vertical profile modification. It is generated due to the relief impact, i.e., the presence of a large valley, blocked on one side with the hills. The analysis of PM_{10} profiles from all flights allows to distinguish three vertical zones of potential air pollution hazard within the valley (about 100 m deep) and the city of Kraków: 1. up to about 60 m a.g.l. – the zone where during periods of low wind speed, air pollution is potentially the highest and the duration of such high levels is the longest, i.e. the zone with the worst aerosanitary conditions; 2. about 60-100 m a.g.l. – transitional zone where the large decrease of PM_{10} levels with height is observed; 3. above 100-120 m a.g.l. – the zone where air quality is significantly better than in the zone 1, either due to the increase of the wind speed, or due to the wind direction change and advection of different, clean air masses.

Publication II - Summary

Kraków, southern Poland, is a city with poor air quality (abundant PM_{10} concentrations). It is located in the large Wisła (Vistula) valley, and affected by a foehn wind, the '*halny'*, from the Tatra Mountains in the Carpathians. There were 14 long episodes of the foehn analyzed from the periods Sep 2017 - Apr 2018 and Sep 2018 - Apr 2019. Data used included measurements of PM_{10} concentrations, air temperature and relative humidity, wind speed and direction from ground stations and tower measurements up to 100m a.g.l., along with model analysis results. Non-operational

configuration of the AROME CMC 1km x 1km (AROME CMC 1km) was applied. A conceptual model concerning the impact of a foehn on urban air pollution was developed. The occurrence of a particular effect of a foehn on the PM_{10} spatial-temporal pattern depends on its mode of transfer through the city and nearby: a. a foehn flows above the valley where a strong cold air pool and a return flow can be found; b. a foehn enters the valley from the eastern, wider part or from the valley top and destroys the cold air pool; c. gravity waves generated by a foehn are strong enough to enter the western narrower part of the valley and cause large spatial differences in turbulence parameters within the city. The first transfer mode worsens air pollution dispersion conditions throughout the city and leads to large increases in PM_{10} levels, the second mode improves dispersion and leads to large decreases in PM_{10} levels throughout the city, and the third generates large spatial differences in PM_{10} levels within the city. There is no single effect of a foehn on air pollution dispersion conditions.

3 Multi-model machine learning

A machine learning-based tool for post-processing of air-temperature numerical forecast has been tested. It is based on a concept of a multi-model since it uses forecasts from three limited-area models operating at the Institute of Meteorology and Water Management: ALARO, AROME and COSMO. A direct output of the models is often biased and needs to be adjusted to observed values. Forecasters often have to reconcile forecasts from several NWP models during their operational work bas on their knowledge and experience. The tested method could be supportive for them during the decision-making process. Predictors include forecasts of elementary weather elements produced by the NWP models for 35 selected synoptic stations in Poland as well as station-embedded data on ambient orography. They were then compared to real values of 2-m air temperature according to observations at the stations. A training set contains data from two years: 2018 and 2019. A Random Forest algorithm (RF) has been used to produce bias-corrected forecasts on a test set spanning a whole one year (2020). Results were verified against forecast from single NWP models as well as the mean of the models. The improvement of forecast accuracy was noted at every station and for every lead time, with the average reduction of RMSE equal to 20.4% with respect to the mean (26.6% with respect to the best NWP model). The biggest training error occurred for stations located in mountain basins (12500 and 12625), while for mountain-top stations (12510 and 12650) it was significantly smaller.

4 Day-Ahead Wind Power Forecasting in Poland Based on ALARO NWP

Introduction

The role of renewable energy sources in the Polish power system is growing. The highest share of installed capacity goes to wind and solar energy. Both sources are characterized by high variability of their power output and very low dispatchability. Taking into account the nature of the power system, it is, therefore, imperative to predict their future energy generation to economically schedule the use of conventional generators. Considering the above, this paper examines the possibility to predict day-ahead wind power based on different machine learning methods not for a specific wind farm but at national level. A numerical weather prediction model used operationally in the Institute of Meteorology and Water Management–National Research Institute in Poland and hourly data of recorded wind power generation in Poland were used for forecasting models creation and testing. With the best method, the Extreme Gradient Boosting, and two years of training (2018–2019), the day-ahead, hourly wind power generation in Poland in 2020 was predicted with 26.7% mean absolute percentage error and 4.5% root mean square error accuracy. Seasonal and daily differences in predicted error were found, showing high mean absolute percentage error in summer and during daytime.

Data

Hourly data time series of power generated on country level from wind turbines was obtained from Polskie Sieci Elektroenergetyczne (Polish Transmission System Operator). Together with information from Urząd Regulacji Energetyki (Energy Regulatory Office) about the localization and power of installed wind turbines in Poland, we constructed a database for every hour from 2018 to 2020 with power generated values and mean wind speed forecasts for County regions of Poland with installed wind turbines. From 2015 onwards, a stagnation in installed capacity in wind sources was observed, which provides a sizable learning sample for machine learning based methods. Additionally, dates from 2018 to 2020 were characterized with the most similar ALARO model characteristics (model version, physics configuration, boundary conditions), which is why this period was selected for this study. The majority of wind turbines in Poland are installed in the north and central parts of Poland.

The background model data come from operational forecast results of nonhydrostatic ALARO CMC model. Operational model ALARO CMC has a horizontal resolution of 4 km × 4 km and 70 vertical levels, and the forecast length is 72 h. The size of ALARO CMC domain is 799 × 799 points, centred on the geographical point 17.5° E 52.5° N. The location of the lowest model level is at 9 m above ground level, and the model top is located at 65 km above ground level. Lateral and boundary conditions for the ALARO CMC model were obtained from the forecast of the global model ARPEGE.

Results

Hourly and monthly statistics of MAPE are presented in Figure 1. For all models, the highest values (yellow colour) are present in June, while the lowest occurred in January and February (dark blue colour). Between 0 and 8 UTC, all models produced more accurate wind forecasts, which were represented by smaller MAPE. The ANN model had the lower variability of MAPE during days, as well as the whole year, while the XGB model has the highest differences between low error in January during nights and high error in July during days.

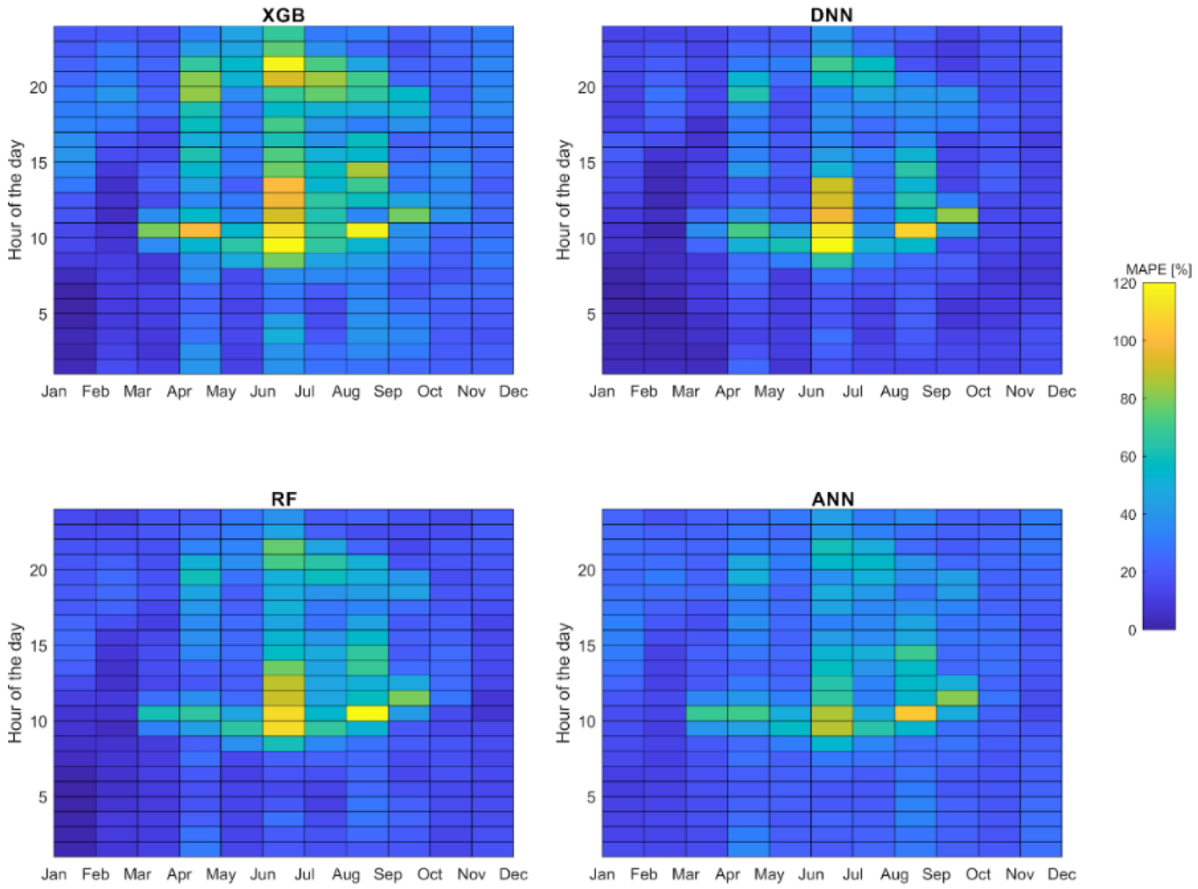


Figure 1. *Hourly and monthly distribution of mean absolute percentage error (MAPE) for all models.*

Discussion

We proposed a day-ahead wind power forecasting system based on multiple machine learning methods, an accurate limited area NWP model and hourly data time series of power generated from wind turbines that can be applied on the country level in Poland and possibly in other countries with similar climatological conditions. We have shown that all models (Random Forest - RF, Extreme Gradient Boosting - XGB, Artificial Neural Network - ANN and Deep Neural Network - DNN) produced forecasts with similar, high accuracy. The XGB model, with MAPE equal to 26.7% was the most accurate for hourly predictions, while for daily sums of produced energy, the ANN method, with MAPE equal to 13.6%, was the best. Our method, with forecasting wind energy production at national level and not for specific wind farms, has not been very often examined in the literature so far, so it is hard to compare our results **with** others' research.

Although the comparison of machine learning methods was not the main topic of this publication, we have shown differences in terms of performance in different seasons, hours of the day and whether there is a high or low real energy production in the system. Two methods based on decision trees (XGB and RF) were found to perform better than ANN and DNN in situations with very high hourly energy production. On the other hand, results of two methods based on neural networks (ANN ad DNN) are characterized with lower daily and monthly variability of MAPE. For all methods, June was the month with the highest and January with the lowest MAPE, while the lowest variances of results are found in winter months and the highest in summer.

5 Derecho case in Poland, 2017, seen by ALARO NH CY43T2

On average, the derecho occurs once a year in Poland while a bow echo happens few times per year. On 11 August 2017, severe meteorological phenomena were observed in Poland, including extremely strong wind gusts. One may particularly outline the storm of a derecho type which occurred on that date in northern and north-western Poland.

A synoptic situation which led to that storm is described in the paper of Mańczak et al., 2021. Developing storms in a warm air mass before a waving front transformed from single cells and chaotic multi-cell systems to the supercells and gathered into the large MCS (Mesoscale Convective System). In that developed a squall line which turned into a strong bow echo (Figure 2) and MCV (Mesoscale Convective Vortex). The evolving convective system, supported by the jet stream, fulfilled criterion of derecho occurrence and the observed maximal values of wind gusts exceeded 42 $m \cdot s^{-1}$ (150$km \cdot h^{-1}$; Mańczak, 2021; Taszarek, 2019). The post-factum weather forecast by the means of the presently available model ALARO CY43T2 a non-hydrostatic one with 4km horizontal resolution (unavailable in 2017), mapped reasonably well the possible severe wind gusts. The highest precision in position of the MCV were achieved by the ALARO NH 00UTC model forecasts while its strength was best predicted by 12UTC forecast (Figure 3), and can be seen on relative vorticity maps of the middle and lower troposphere. Other severe weather phenomena, such as cold pool and rear inflow jet (RIJ) (Figure 4, Figure 5), were clearly seen on forecasts on the maps of the wind velocity at 850 and 925hPa pressure levels and on temperature at 2m. The pressure signatures of MCS at Fig. 10 is low pressure ahead of the front, high pressure near the convergence line, and again low pressure on the tail of the MCS. Position of the MCV at 22UTC is clearly visible at Figure 5 a) and b) at the left (north-western) part of the MCS.

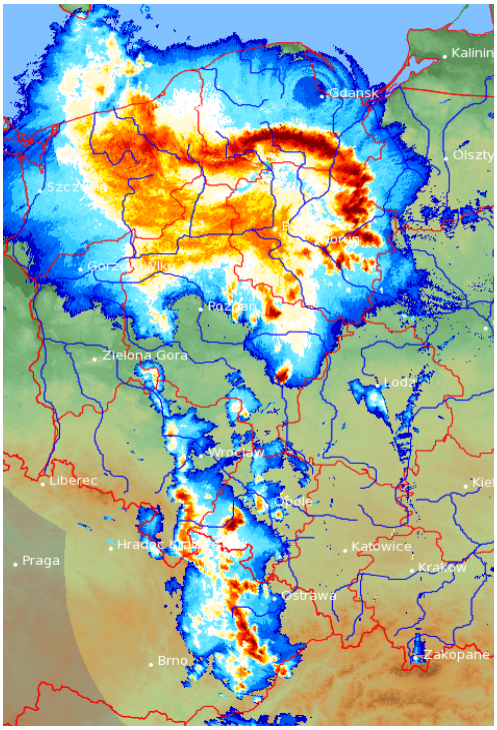


Figure 2. *Radar reflectivity, CMAX, observed for a couple of radar stations. 21UTC of 11.08.2017.*

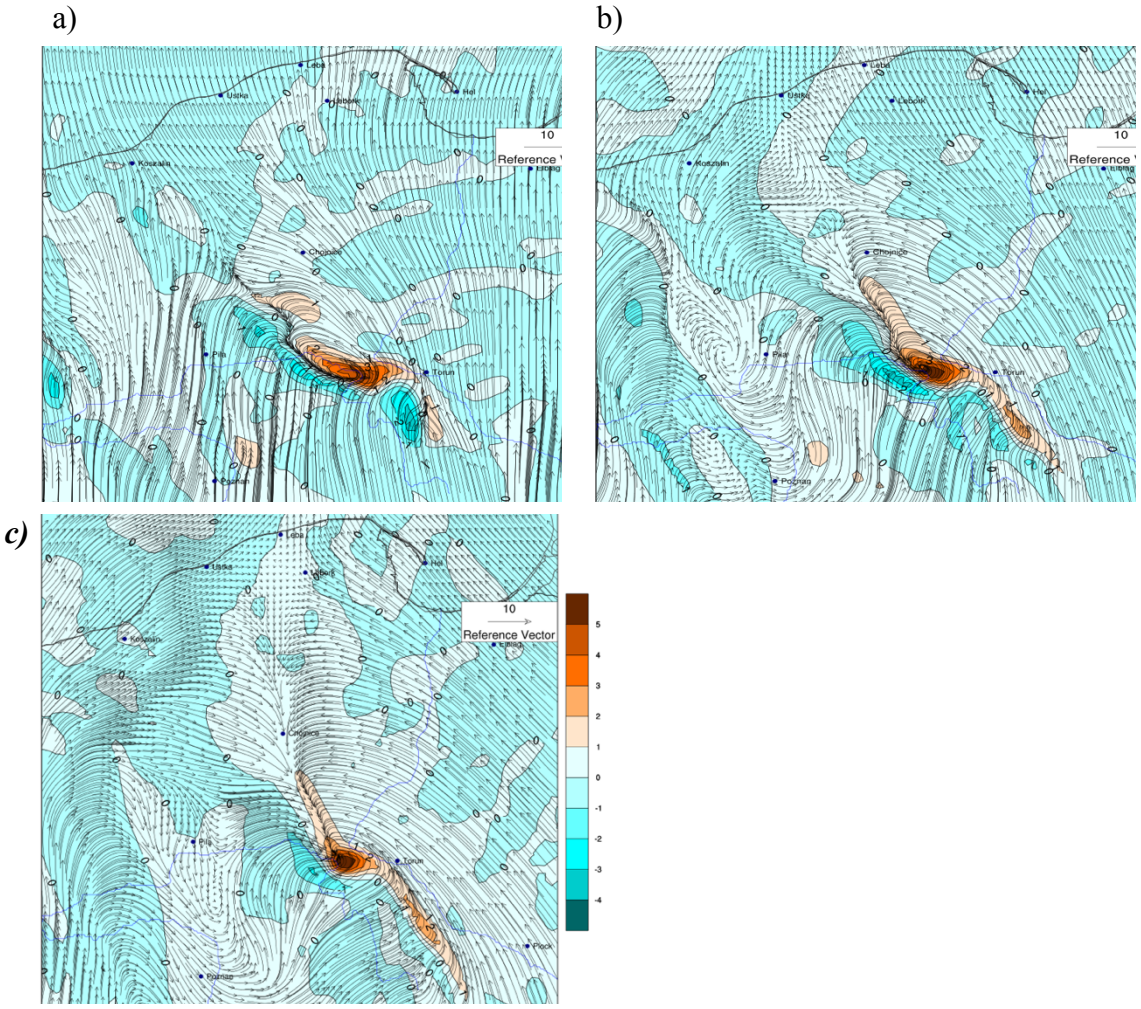


Figure 3. *The forecast of MCV. Wind ($m \cdot s^{-1}$, vectors) and relative vorticity (s^{-1}, colour) at the levels of a) 700, b) 850 and c) 925hPa. ALARO r12 on 21UTC 11 August 2017.*

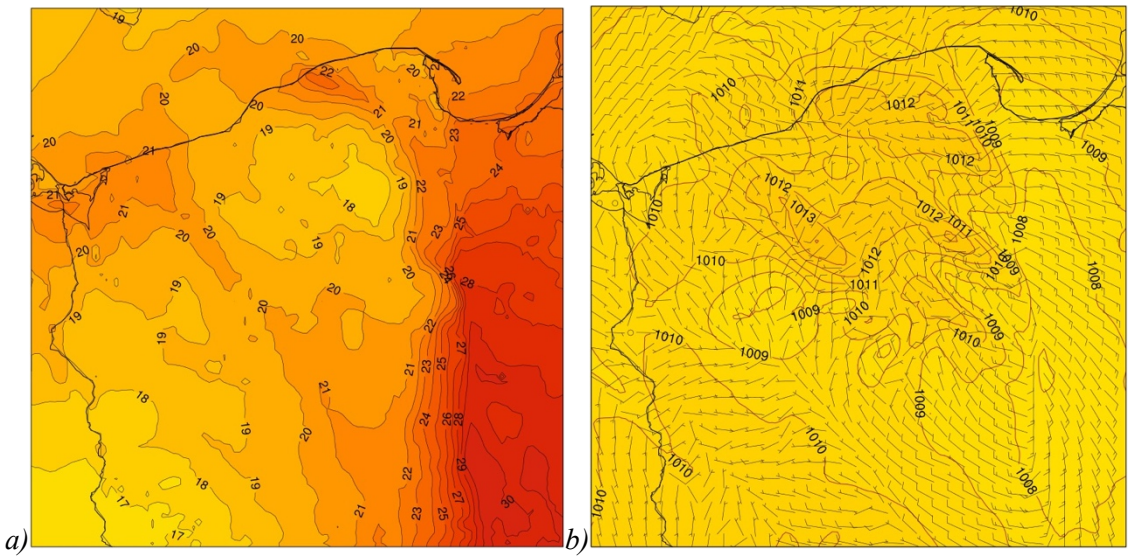


Figure 4. *Cold pool from ALARO a) temperature at 2m AGL b) atmospheric pressure and surface wind. The forecast for 21UTC 11 August 2017, r00.*

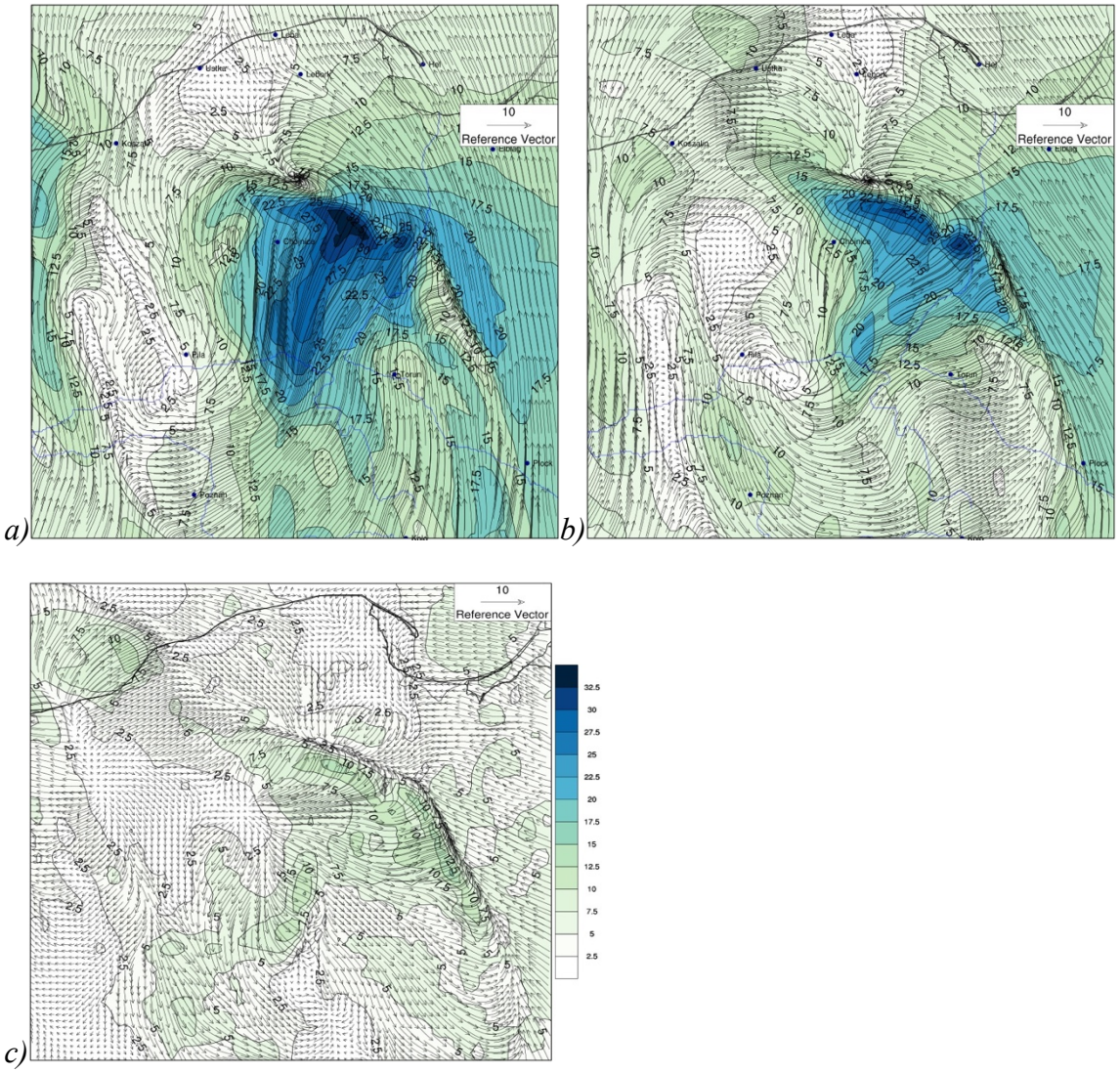


Figure 5 *a) and b) wind and the module of velocity (colour) at pressure levels of 850 and 925hPa; c) wind and the module of velocity (colour) at 10m AGL. ALARO r12 forecast for 22UTC, 11 August 2017.*

6 References

Bochenek, B., Jurasz, J., Jaczewski, A., Stachura, G., Sekuła, P., Strzyżewski, T., Wdowikowski, M., Figurski, M. Day-Ahead Wind Power Forecasting in Poland Based on Numerical Weather Prediction. Energies. 2021; 14(8):2164. https://doi.org/10.3390/en14082164

Kolonko, M., Szczęch-Gajewska, M., Stachura, G., Bochenek, B., Sekuła, P.: An analysis of NWP model forecast ALARO and AROME for the derecho case on 11 August 2017 – submitted to Meteorology, Hydrology and Water Management.

Mańczak, P., Wrona, B., Woźniak, A., Ogrodnik, M., Folwarski, M., 2021, „Synoptyczne uwarunkowania nawałnicy typu derecho, która przeszła przez północno-zachodnią część Polski w dniu 11.08.2017", submitted to Meteorology, Hydrology and Water Management.

Sekuła, P., Bokwa, A., Bartyzel, J., Bochenek, B., Chmura, Ł., Gałkowski, M., and Zimnoch, M.: Measurement report: Effect of wind shear on PM10 concentration vertical structure in the urban boundary layer in a complex terrain, Atmos. Chem. Phys., 21, 12113–12139, https://doi.org/10.5194/acp-21-12113-2021, 2021

Sekula, P., Bokwa, A., Ustrnul, Z., Zimnoch, M., and Bochenek, B.: The impact of a foehn wind on PM10 concentrations and the urban boundary layer in complex terrain: a case study from Krakow, Poland, Tellus Series B-Chemical and Physical Meteorology, 73, 10.1080/16000889.2021.1933780, 2021.

Stachura, G., Bochenek, B., Sekuła, P., Szczęch-Gajewska, M., Kolonko, M.: Machine-learning-based post-processing of model-derived near-surface air temperature – a multi-model approach - submitted to Monthly Weather Review.

Taszarek, M., Pilguj, N., Orlikowski, J., Surowiecki, A., Walczakiewicz, S., Pilorz, W., Piasecki, K., Pajurek, Ł., Półrolniczak, M., 2019. Derecho Evolving from a Mesocyclone—A Study of 11 August 2017 Severe Weather Outbreak in Poland: Event Analysis and High-Resolution Simulation. Monthly Weather Review 147, 2283–2306. https://doi.org/10.1175/MWR-D-18-0330.1

Implementation and validation of AROME CY43T2 at the Hungarian Meteorological Service

Balázs Szintai, Katalin Radnóczi, Viktória Homonnai, Helga Tóth, Kristóf Szanyi, Gabriella Tóth, Boglárka Tóth, László Kullmann, Gabriella Szépszó
Hungarian Meteorological Service

1 Implementation

The HPC of the Hungarian Meteorological Service (OMSZ) used for running the operational AROME suite is a HPE Apollo 6000 server. The operational part of the machine consists of 12 nodes running Intel XeonE5-2698 processors at 2.2 GHz, with each node having 40 cores and 128 GB RAM. The AROME code was compiled with Intel compiler version 18.0 using gmkpack 6.7.3.
Implementation work of AROME CY43T2 at OMSZ began in May 2020. The main pack is cy43t2_bf10 and cy43t2_bf11 is installed on top of it. Additionally, some local modifications are also implemented, which are the following:

- change of lake temperatures in SURFEX according to water temperature observations at Lake Balaton
- copying snow properties (albedo, snow-water equivalent, snow density) from the SURFEX first guess file into the SURFEX initial file
- writing snow properties in the upper air FA file
- introducing the upper limit of 1.0 for saturated water vapor mixing ratio over liquid water and ice in condensation.F90 to avoid unrealistic cloud cover at the uppermost level (with 60 level configuration)

For the CY43T2 implentation new climate files were generated using the CLIMAKE software at the HPC of Meteo-France. For the new climate files, the GMTED-2010 orography was used (for previous cycles at OMSZ we have used the GTOPO30 orography).

2 Validation and implementation of the deterministic AROME model

After installation of the model, some first tests were carried out to check the model forecasts and the assimilation. After these first encouraging runs three longer periods were computed with the new model cycle. For these longer runs the same model configuration was used as in the operational AROME model, using CY40T1, as described in Szűcs et al., 2019.

Summer period

The summer period was computed for the period between 9th and 31st July 2021, with forecast runs at 00 and 12 UTC for +36 hours. Both the reference (CY40T1) and the test (CY43T2) runs were initialized from the upper air analysis of ECMWF/IFS and the surface was taken from the operational AROME cycle. For this the initial SURFEX file had to be converted in the test run due to the different SURFEX version (SURFEX_v7.3 in CY40T1 and SURFEX_v8.0 in CY43T2). The assimilation cycle in the reference and test runs involved the same archived observations, the only difference from the operational configuration was that GNSS data were not used in these runs (reference and test) for technical reasons.

Pointwise verification was performed applying the in-house verification system of OMSZ, called OVISYS. For the verification of near-surface parameters all surface stations from Hungary were used (218 stations), for upper air parameters three radiosounding stations (Budapest, Szeged, Zagreb) were investigated. Largest difference between reference and test was observed in the case of 2 metre temperature in the nighttime (Fig. 1), for other parameters the impact was neutral.

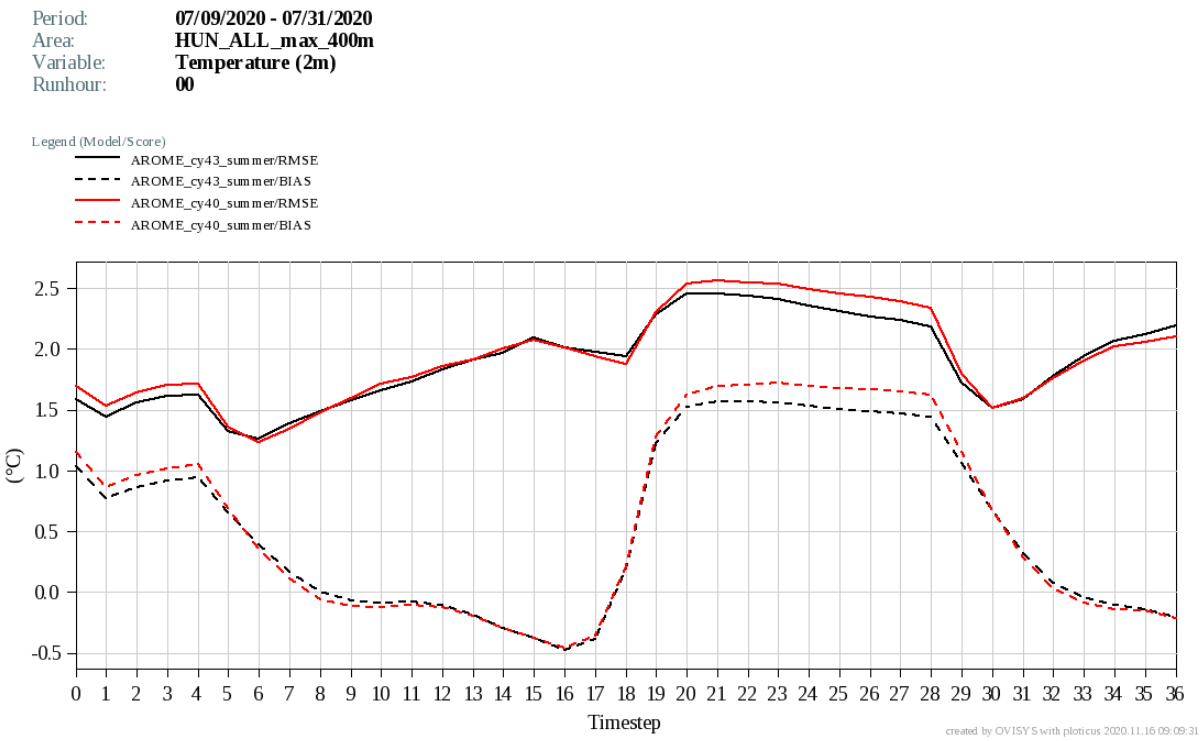


Figure 1: 2 metre temperature verification scores (bias: dashed line; RMSE: solid line) for the summer period for the reference (CY40T1, red line) and the test (CY43T2, black line) runs. Verification domain is Hungary with around 220 automated weather stations.

For precipitation forecasts the SAL verification method (Wernli et al., 2008) was also applied using precipitation data calculated from the Hungarian radar measurements. In the case of classical SAL scores only small differences could be observed between the reference and test runs (not shown). Slightly larger differences can be detected if new characteristics are computed based on the SAL object identification method (Řezáčová at al., 2015). By investigating the three strongest precipitation objects it can be noted that in the test run objects are slightly weaker than in the reference run, which is closer to the measurements (Fig. 2).

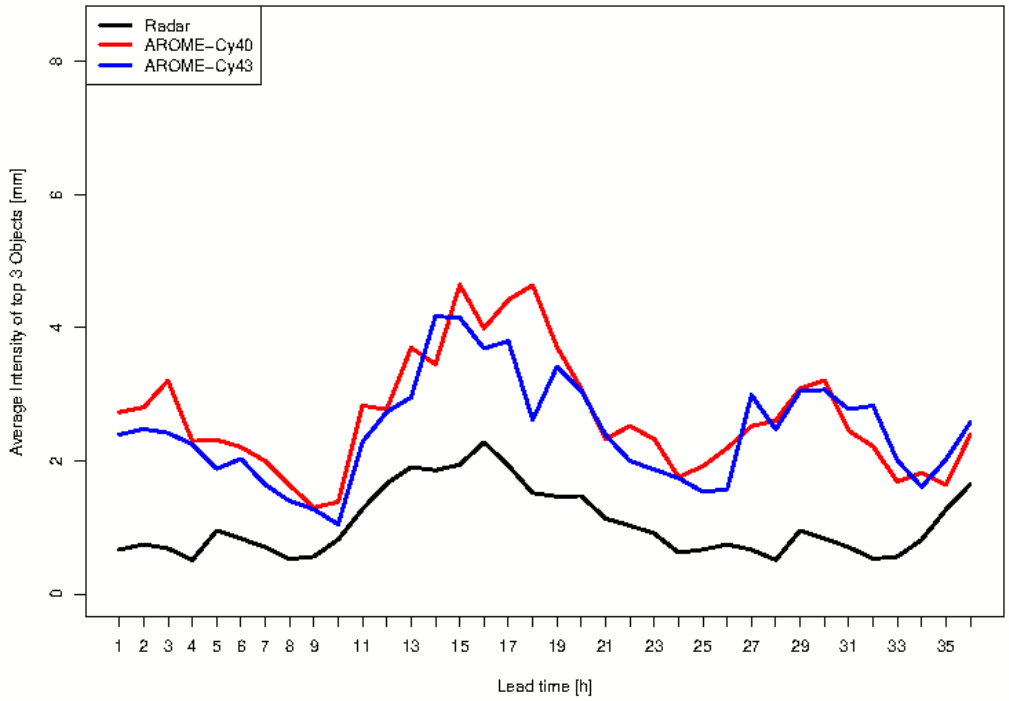


Figure 2: Average intensity of the three strongest precipitation objects for the summer period for the reference (CY40T1, red line) and the test (CY43T2, blue line) runs, as compared to radar measurements (black line). For this plot only the 00 UTC runs were used.

Next to the objective verification methods some case studies were also investigated, focusing mainly on 2 metre temperature, wind gusts and convective precipitation. On the days investigated, only small differences were observed between the reference and test runs. In the case of precipitation, mainly the structure of precipitation objects was different (Fig. 3).

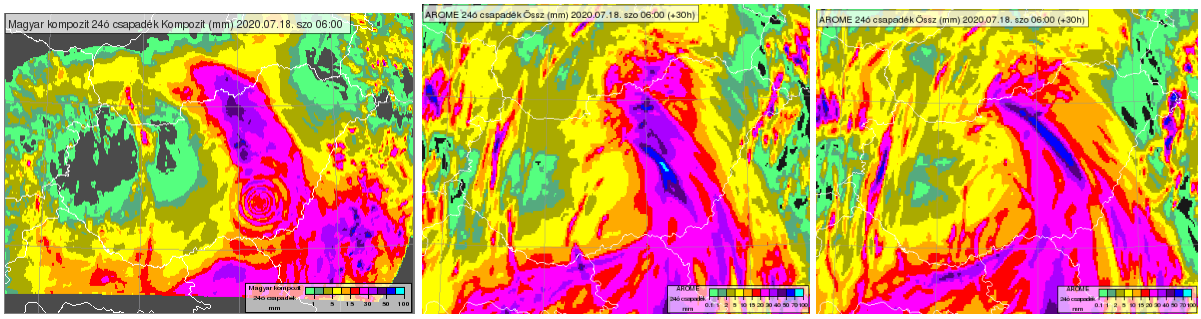


Figure 3: 24 hour accumulated precipitation at 06 UTC on 18 July 2020. Left: radar measurement; middle: CY43T2 run (+30h forecast); right: CY40T1 run (+30h forecasts).

Winter period

For the winter period the time interval between 25th November and 17th December 2019 was chosen. The experimental settings were similar to the summer period. The only difference was that in these winter experiments GNSS data were also used in the assimilation.

Pointwise verification showed mostly neutral impact for near surface and upper air parameters. Largest difference could be detected for 10 metre wind speed where the CY43T2 run had lower bias and RMSE than the CY40T1 run (Fig. 4).

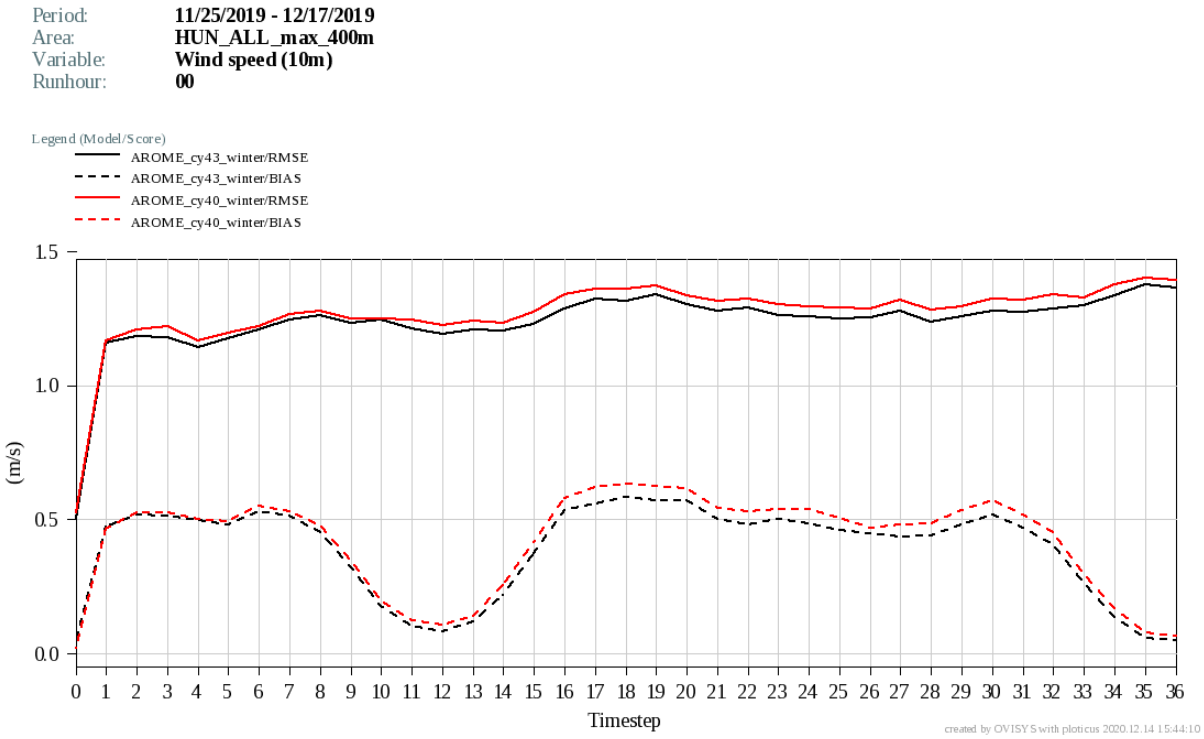


Figure 4: 10 metre wind speed verification scores (bias: dashed line; RMSE: solid line) for the summer period for the reference (CY40T1, red line) and the test (CY43T2, black line) runs. Verification domain is Hungary with around 100 automated weather stations (there are less stations measuring wind than temperature).

Case studies were also investigated from the winter period, mainly focusing on 10 metre winds and precipitation. For most case studies the impact of CY43T2 was neutral. Figure 5 depicts a case (17 December 2019) where the 10 metre wind gust forecast was slightly improved by CY43T2; the overestimation of CY40T1 over the western part of Hungary was corrected to some extent.

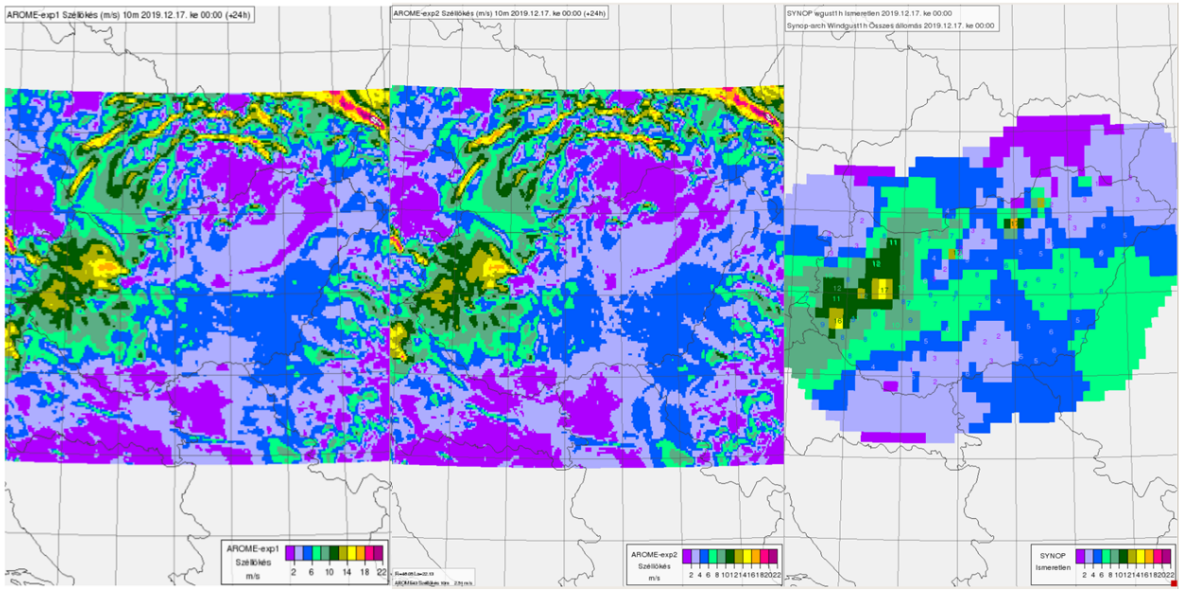


Figure 5:Hourly maximum wind gust at 00 UTC on 17 December 2019. Left: CY40T1 run (+24h forecasts); middle: CY43T2 run (+24h forecast); right: interpolated wind gust field computed from SYNOP measurements.

Real time test by forecasters

A pre-operational parallel run was also conducted between 8th February and 8th March 2021. During this period four forecast runs (at 00, 06, 12 and 18 UTC) were computed with CY43T2 at the same time as the operational AROME (CY40T1) runs and products were generated from CY43T2 as well for forecasters. Forecasters were systematically observing these parallel runs during their day-to-day work and gave written feedback to model developers. The general conclusion of these feedbacks was in line with the experience gathered from the previous validation: CY43T2 gives similar forecasts to CY40T1, some small improvements can be observed in near surface temperature and wind.

Forecasters also highlighted the improved wind forecasts of CY43T2 over Lake Balaton, which is a consequence of the different computation of the Charnock formula over water surfaces in CY43T2. In CY40T1 a systematic underestimation of 10 metre wind speed could be observed over Lake Balaton which is corrected in CY43T2 (Fig. 6).

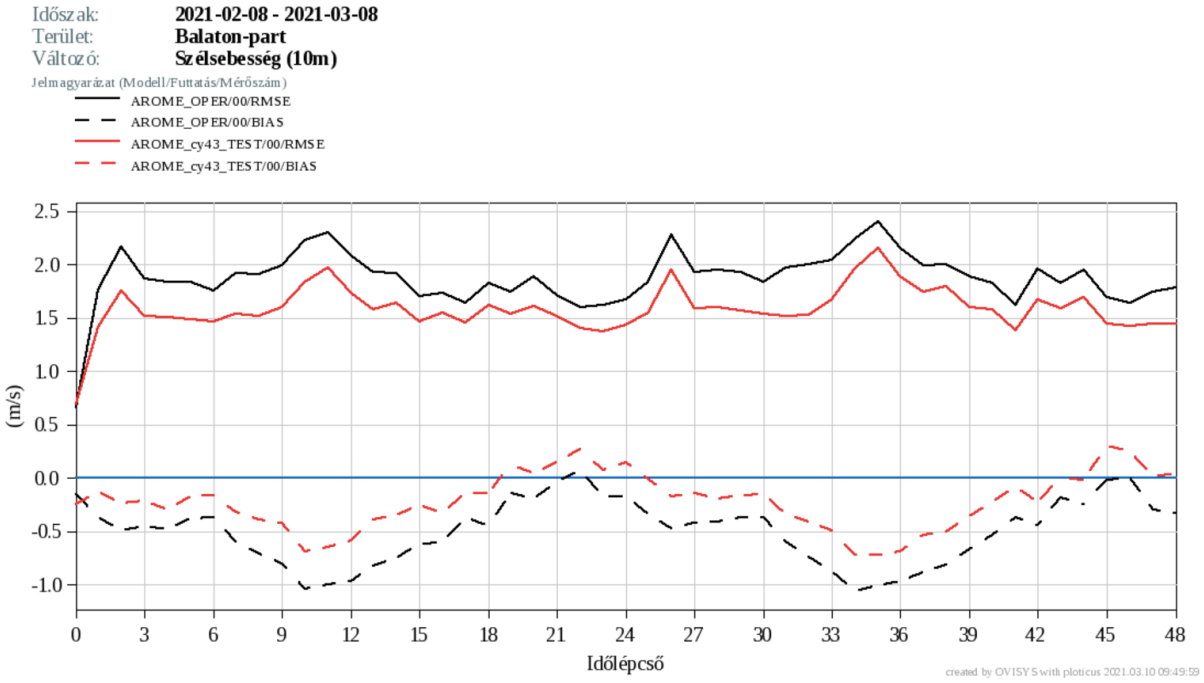


Figure 6: 10 metre wind speed verification scores (bias: dashed line; RMSE: solid line) for the real time test period for the operational AROME (CY40T1, black line) and the parallel experiment (CY43T2, red line) runs. Verification domain is covering Lake Balaton with 6 automated weather stations positioned on the shore of the lake.

Operational implementation

After the successful parallel experiment, CY43T2 was introduced in the operational AROME deterministic suite on 17th March 2021.

3 Validation and implementation of AROME-EPS

After the implementation to deterministic model run, CY43T2 was also tested in AROME-EPS, which uses the same model configuration, but runs without data assimilation, as a downscaling of the first 11 members of ECMWF ENS, and using surface initial state of AROME deterministic run (*Jávorné-Radnóczi et al.*, 2020). Differences between the CY40T1 and CY43T2 were examined in a case study for 18 and 19 March 2021, when cold advection in the common flow system of an anticyclone in the Northwest and a cyclone in the Southeast of Hungary caused small amount of rain and snow in the country. When comparing the two model versions, similar results were found in single EPS members as in the evaluation of deterministic AROME. Precipitation and cloudiness were overestimated by both configurations. 10 metre wind speed was slightly underestimated in the northwestern part, and overestimated in the southeastern plains of the country in the afternoon hours, CY43T2 provided slightly better predictions; in ensemble spread, we found no significant difference (Fig. 7). The new model version runs operationally since 30th March 2021.

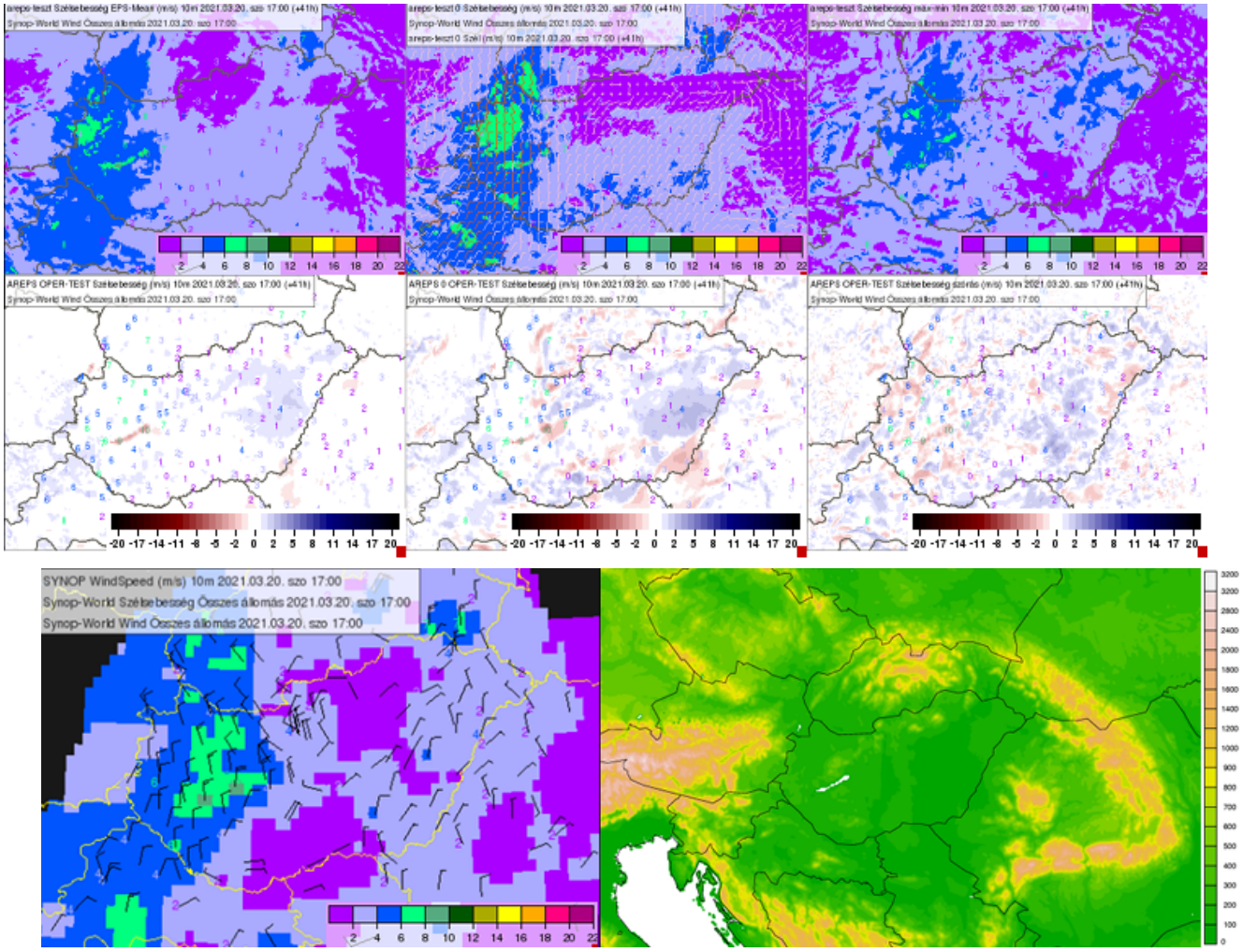


Figure 7: 1 hour average of 10 metre wind speed forecast (CY43T2, upper row) and difference between CY40T1 and CY43T2 forecasts (middle row) at 17 UTC on 20 March 2021 (+41 hours forecasts). In the first column EPS mean, in the second column control member, in the third column EPS maximal spread (calculated as difference of EPS minimum and maximum) are shown. In the lower left picture SYNOP observations are displayed for the same date, on the lower right model orography (CY43T2) is shown.

4 Acknowledgements

The authors would like to thank the help of Meteo-France colleagues: Claude Fischer, Florian Suzat, Ryad El Khatib and Yann Seity.

5 References

Jávorné-Radnóczi, K., Várkonyi. A., Szépszó, G., 2020: On the way towards the AROME nowcasting system in Hungary. ALADIN-HIRLAM Newsletter No. 14, 65–69.

Řezáčová D, Szintai B, Jakubiak B, -I Yano J, and Turner S. 2015: Verification of High-Resolution Precipitation Forecast with Radar-Based Data. In *Moist Atmospheric Convection: An Introduction and Overview*, 173–214. doi:10.1142/9781783266913_0024.

Szűcs, M., Homonnai, V., Lancz, D., Mester, M., Radnóczi, K., Suga, R., Szintai, B., Várkonyi, A., 2018: Modeling Activities at the Hungarian Meteorological Service, ALADIN-HIRLAM Newsletter No. 12, 62-66

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

AROME-Morocco on the new HPC : Recent developments and validation

Najla Marass, Zahra Sahlaoui, Fatima Zahra Hdidou, Kamal El Karouni

1 Introduction

The operational suite of the non-hydrostatic model AROME (Seity et al, 2011) started at Morocco in 2012 using dynamical adaptation mode. Thanks to the installation of the new HPC system, many progress were achieved to enhance the AROME forecast quality. Indeed, the implementation of cycle 43t2 and the preliminary test of surface data assimilation were accomplished, followed by an assessment of the impact of these changes.

2 Recent development on AROME-Morocco

2.1 The Moroccan new HPC system

Since the end of 2019, Maroc-Meteo has been renewing its intensive computing system . The chosen solution is based on Lenovo servers powered by Intel Xeon processors, with a peak power of 1069 TFLOPS. The deployment of the new platform made it possible to multiply by more than 100 times the power of the computing system compared to the previous one. The new HPC, named AMTAR, is used for both operational suites and research studies.
AMTAR has a set of 128 nodes, grouped into 8 racks and distributed as follows:

- 120 compute nodes with 96 cores and 384 GB of memory on which applications are run.
- 2 login nodes with 48 cores and 384 GB of memory for compilation, debugging and user access.
- 2 I/O nodes with 20 cores and 256 GB of memory for storage management and GPFS.
- 2 management nodes with 20 cores and 256 GB of memory for cluster monitoring.

The acquisition of this new supercomputer is a major issue for Maroc-Meteo. In this regard, since its implementation, collective efforts have mainly been made to improve the quality and accuracy of atmospheric forecasts.

2.2 Implementation of cycle cy43t2

Retrieved from the Météo-France oper packs, the export version of cycle cy43t2_op4 (based on cy43t2_bf8) was initially installed. Then a phasing of the incremental corrections bf09, bf.10 and bf.11 was carried out in the local version.

2.3 Implementation of surface analysis OI_MAIN

The objective of the surface data assimilation (SurfDA) is to refresh the model surface from the analysis of 2m temperature and relative humidity by the OI_MAIN method (Giard and Bazile, 2000). To achieve that, the following steps are required:

1. adding missing files in first guess.
2. refreshing SST on the first guess using the ARPEGE coupling file.
3. producing the SURFEX parameters (TG1,TG2,WG1,WG2) from 2 meter temperature and humidity increments produced by CANARI analysis.

The participation of the Moroccan NWP team to the effort of the DAsKIT group allowed the indoor installation of the surface data assimilation on the new HPC system for cycle cy43t2.

3 Verification method and results

3.1 Verification experiments and methods

The assessment of the new implementations was performed as a comparison of the 24 hours forecast against surface observations. The truth is 2 meter temperature and humidity provided by an observing network made up of 40 hourly synoptic stations and around 100 automatic stations over the AROME domain. The observations network is represented in figure 1.

The configurations of AROME-Morocco used during the impact assessment study are described in table 1.

Table 1: The studied configurations of AROME-Morocco

	AROME REF	AROME ADYN	AROME OIMAIN
Cycle	cy41	cy43t2	cy43t2
Resolution.	2.5 Km	2.5 Km	2.5 Km
Grid point	800 X 800	800 X 800	800 X 800
Number of levels.	90	90	90
Time step integration.	60 s	60 s	60 s
Coupling model.	ARPEGE	ARPEGE	ARPEGE
Coupling frequency.	1 hour	1 hour	1 hour
Forecast range	72 hours	72 hours	48 hours
Type of initialisation	ARPEGE coupling file	ARPEGE coupling file	Surface analysis

To evaluate these different configurations, The bias B and the root mean squared error RMSE were calculated at the nearest grid point using the following formulation:

$$B = \frac{1}{N}\sum^{N}(F_i - O_i)$$

B represents the arithmetic mean of the deviation from the forecast (F) to the observation (O). It indicates the direction of the deviation from the observation and reflects the systematic error of the model for a given meteorological parameter. This score indicates whether the model tends to underestimate (B <0) or overestimate (B> 0) the forecast of the studied parameter.

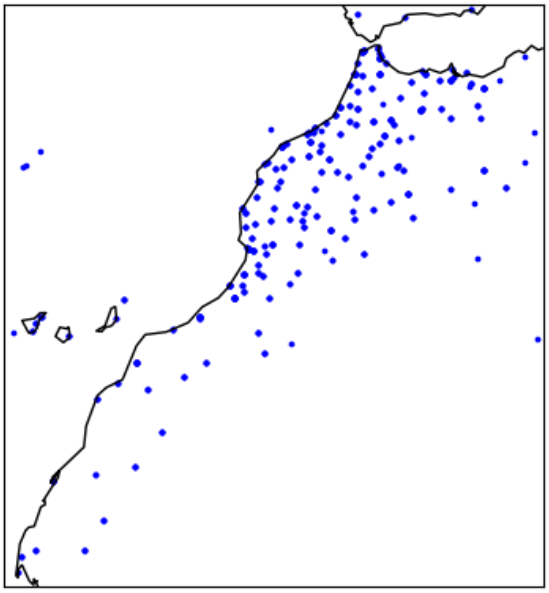


Figure 1: Synoptic and automatic weather stations over AROME-Morocco domain

$$RMSE = (\frac{1}{N}\sum^{N}(F_i - B - O_i)^2)^{\frac{1}{2}}$$

RMSE measures the dispersion of the values compared to the observation, and thus reflects the variability within the data sample.

3.2 Validation of cycle cy43 against cycle cy41

The evaluation of the cycle cy43 was performed in the summer of 2019 as a comparison with the cycle cy41 AROME forecasts. REF has a stronger bias in 2 meter temperature (left panel in figure 2) with a pronounced diurnal cycle. ADYN experience improves the RMSE in terms of daytime temperature but the impact is negative over the night. The 2 meter relative humidity forecast (right panel in figure 2) shows an overestimation in both experiences. However the run with cycle cy43t2 had allowed an important reduction of the RMSE.

3.3 Validation of surface analysis against dynamical adaptation mode

The usefulness of running a surface data assimilation is assessed by comparing the scores of forecasts that start from OI_MAIN analyses to those starting from a dynamical adaptation of ARPEGE analyses. The study period covers January and February 2021. The ADYN experiment shows a cold and wet bias. The AROME SurfDA provides an increase of the forecast biases for both 2 meter temperature and relative humidity during daytime. In terms of RMSE scores, the two forecasts sets of 2 meter temperature are very close to each other during the first 12 h, then the RMSE of OIMAIN experiment becomes more accentuated. Concerning 2 meter relative

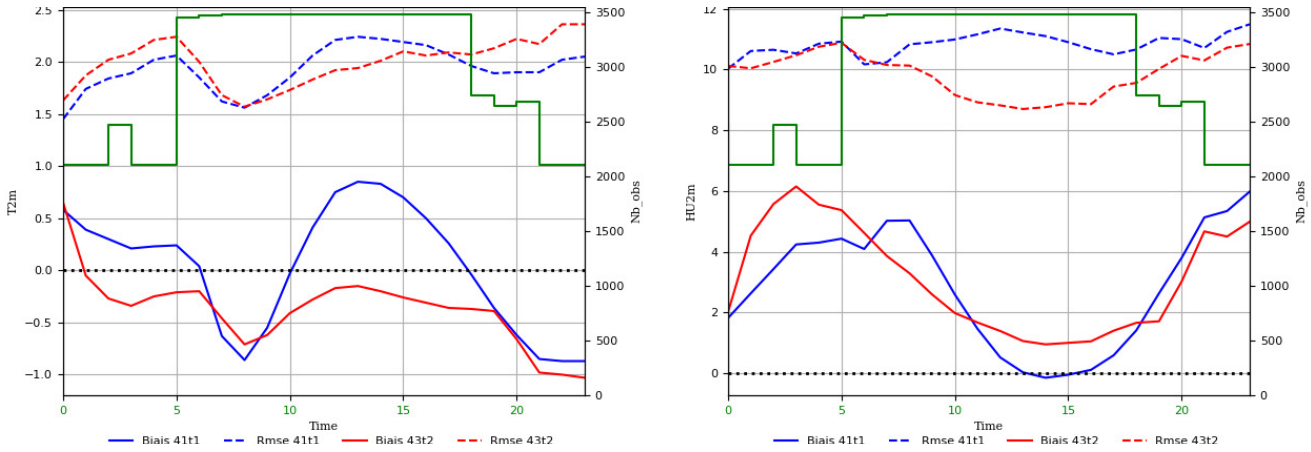


Figure 2: Comparative scores of bias (solid lines)) and RMSE (dashed lines) from AROME REF (cy41 blue lines) and AROME ADYN (cy43 in red lines) forecasts over July, August and September 2019. Left is 2 meter temperature and right is 2 meter relative humidity.

humidity, the RMSE of OIMAIN experiment is higher than ADYN experiment during all time of the forecast range.

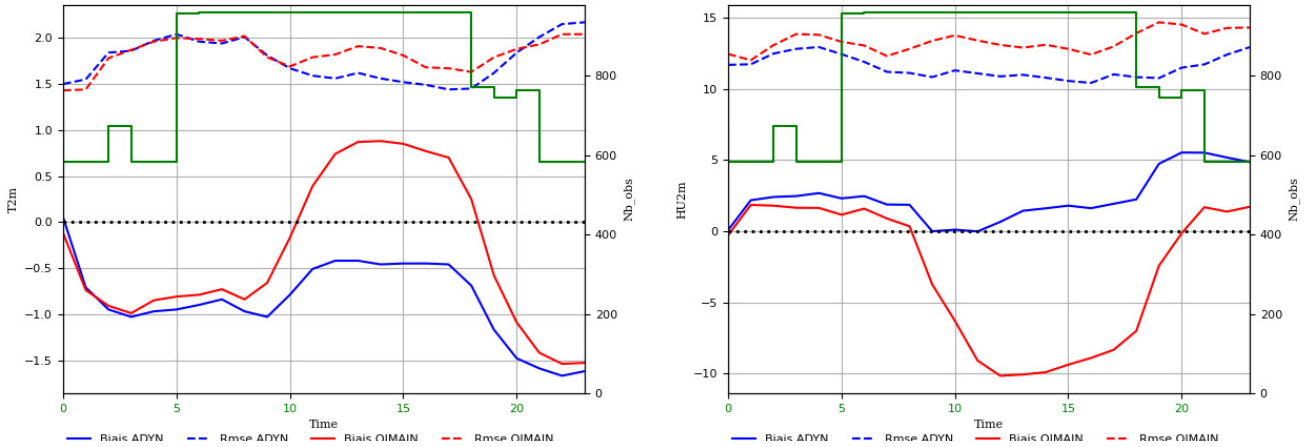


Figure 3: Comparative scores of bias (solid lines)) and RMSE (dashed lines) from AROME ADYN (blue lines) and AROME OIMAIN (red lines) forecasts over January and February 2021. Left is 2 meter temperature and right is 2 meter relative humidity.

4 Future plans

Thanks to the research efforts and the new version of the HPC system, the following evolutions are planned: (i) the implementation of a combined upper air data assimilation system (3D-Var) with surface assimilation; (ii) the increase of the horizontal resolution up to 1.3 km for AROME in assimilation mode and (iii) the supply of additional observation systems for the upper air analysis.

5 References

Giard, D., and E. Bazile, Implementation of a new assimilation scheme for soil and surface variables in a global NWP model. Mon. Wea. Rev., 128 , 997–1015, 2000.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., Masson, V., The AROME-France Convective-Scale Operational Model, Monthly Weather Review, 139(3), 976-991, 2011.

Installation of cy43t2_bf10 and Comparison of ALARO cy43t2_bf10 and ALARO cy40t1_bf7

Yelis Cengiz, Duygu Üstüner

1 Introduction

In this article, the installation of cycle 43t2_bf10 on HPC at TSMS and the comparison of ALARO cycle 43t2_bf10 and cy40t1_bf7 were described.

2 Installation of cycle 43t2_bf10

HPC at TSMS

The features of HPC SGI ICE XA at TSMS are as follows:

1. 4032 cores – Intel Xeon E5 2690 -2.6 GHz
2. 27 TB memory
3. ~167Tflops
4. Linux (SLES 12) operating system
5. PBS Professional workload management system

Installation

The following of the source code directories and scripts were removed before the installation as advised:

* rm -rf src/local/cope src/local/oops src/local/mitraille/* src/local/scripts/*
* rm -rf src/local/surf/offline/* src/local/odb/scripts/* src/local/odb/lib/Magics_dummy.F90
* rm -rf src/local/odb/extras/mpi_serial/* src/local/odb/extras/emos/* src/local/odb/extras/ec/
* rm src/local/odb/lib/Dummies_netcdf.c src/local/odb/lib/Ctxprint.F90 src/local/odb/lib/Bexit.F90
* rm src/local/odb/ddl/PRESCREEN.ddl src/local/odb/ddl/BUFRBASE.ddl
* rm src/local/odb/ddl.ECMA/matchup_update_$i.sql (i=4,5,6,7,8,9,10)
* rm src/local/odb/ddl.CCMA/matchup_update_$i.sql (i=4,5,6,7,8,9,10)

The pre-processing issue in the codes sfxfagrok.F90 and modd_io_surf_aro.F90 were solved by replacing FA with FAFA in both of the codes.
The compilers used are : mpif90, icc, icpc.
The verisons of some of the libraries are namely: ecCodes version 2.10.0, netcdf version 4.4.4 , hdf5 version 1.10.1, openjpeg 1.5. The source code was compiled by gmkpack version 6.6.8.

3 ALARO Model at TSMS and Verification Scores

For the verification ALARO model cycle 40t1_bf7 and cycle 43t2_bf10 were used. The model is coupled with ARPEGE and it produced forecasts up to 72 hours. The size of the domain was also changed while updating the cycle version. The new model domain is shown in Figure 1. The domain information of the model is given in Table 1.

Table 1: Domain information of Alaro-cy40 and Alaro-cy43

	Alaro 40	**Alaro 43**
cycle	cy40t1bf7	cy43t2bf10
center lon	32	29.67
center lat	40	40.67
resolution	4.5 km	4.5 km
total points x(ndlon)	720	1000
total points y(ndgl)	450	600
number of vertical levels	60	72
tstep	180	180

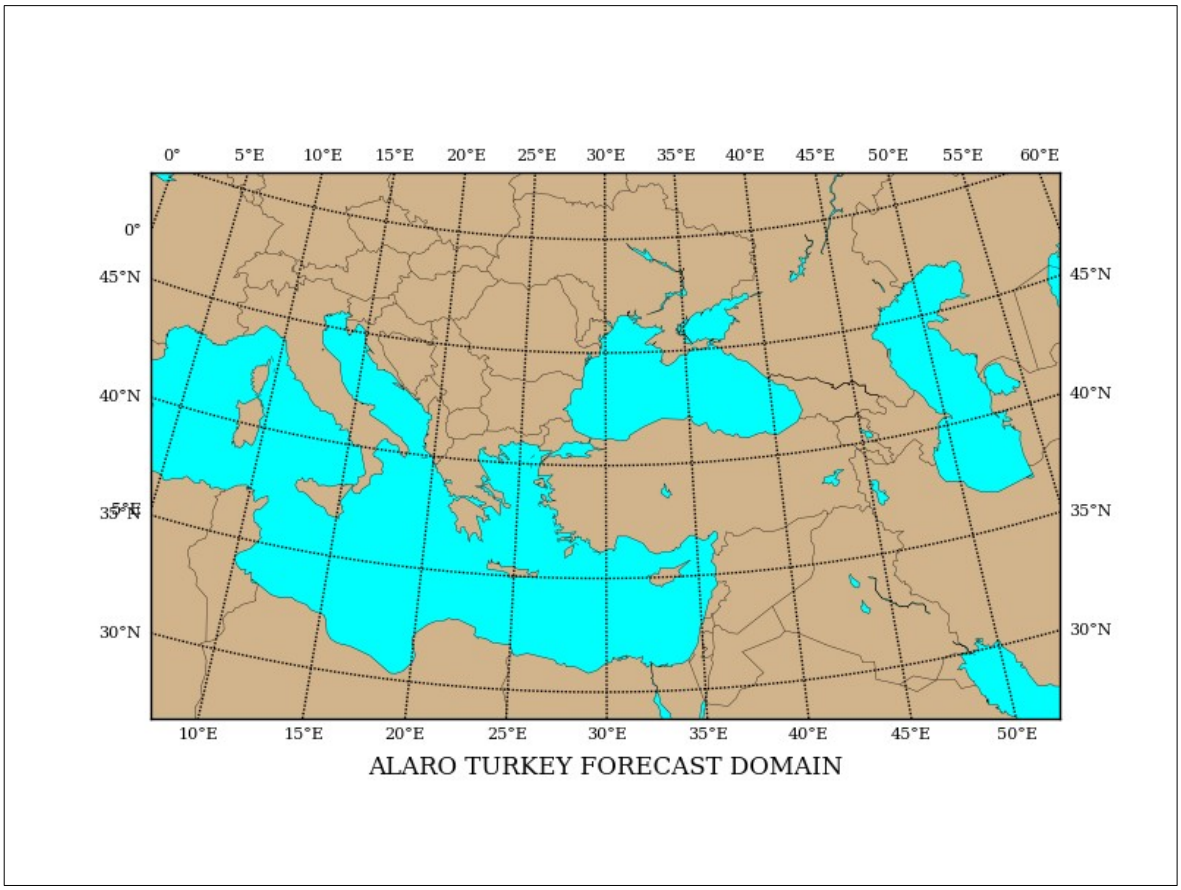


Figure 1: New Forecast Domain of ALARO Turkey

For the verification scores only 00 UTC runs and 6 parameters (2m temperature, 2m relative humidity, cloud cover, 10m wind, geopotential) and 3 months (April, May, June) are chosen. Harmonie verification package (Yang, 2008) was used to compare these two cycles . The figures of the verification scores are shown belown:

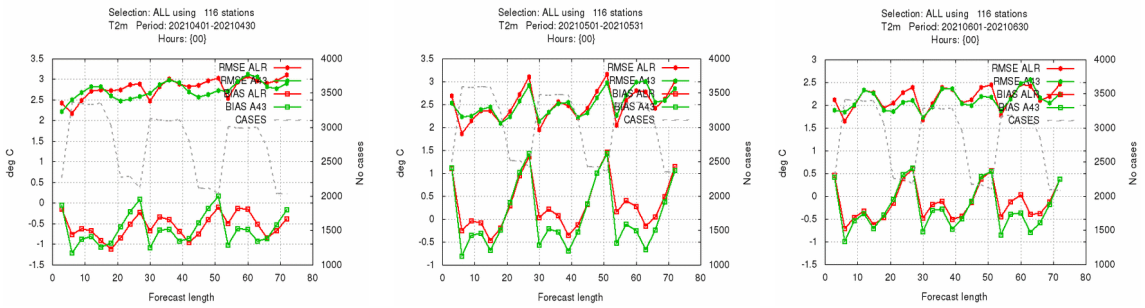


Figure 2: 2m Temperature rmse and bias in 2021 April, May and June for ALR (Alaro-cycle 40) and A43 (Alaro-cycle 43).

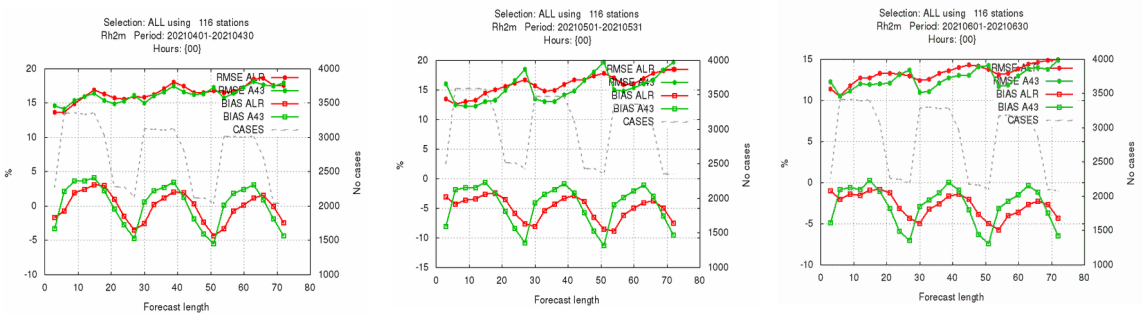


Figure 3: Relative Humidity rmse and bias in 2021 April, May and June for ALR and A43.

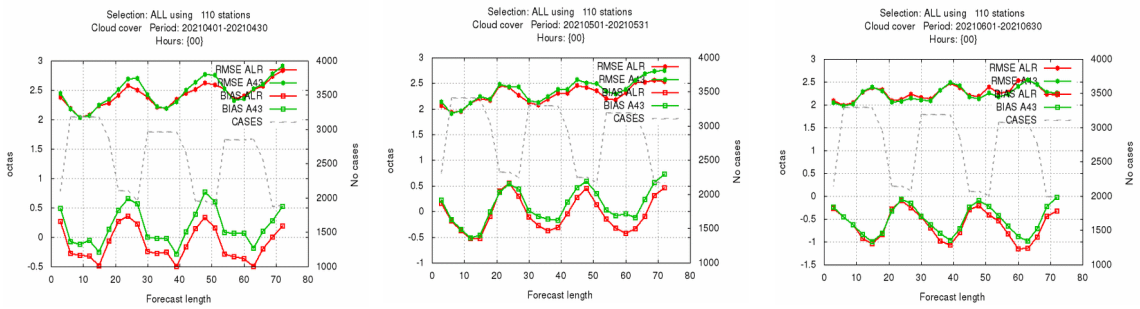


Figure 4: Cloud cover rmse and bias in 2021 April, May and June for ALR and A43.

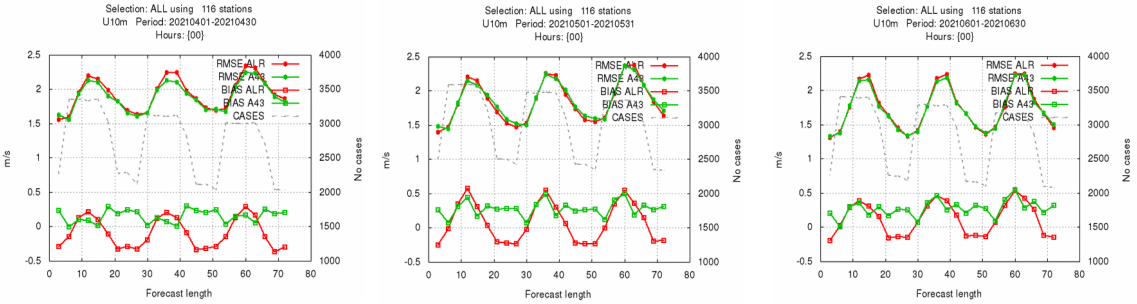


Figure 5: U10m rmse and bias in 2021 April, May and June for ALR and A43.

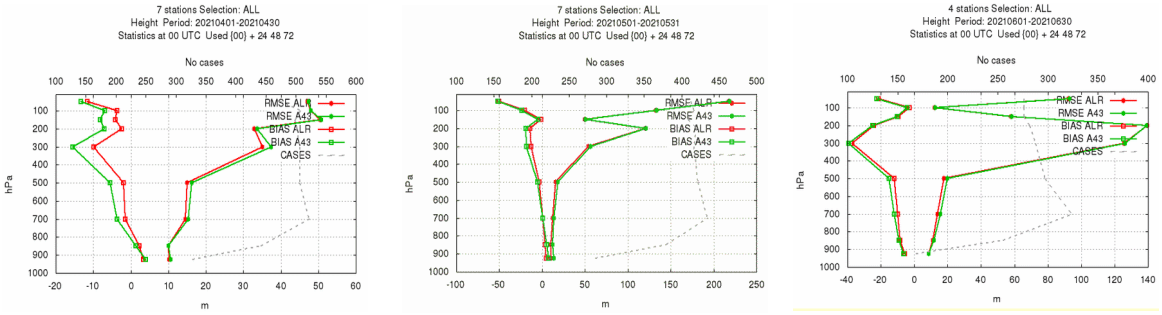


Figure 6: Geopotential rmse and bias in 2021 April, May and June for ALR and A43.

The significant points of verification results;
- cy43 has less rmse than cy40 over the night period for temperature,
- cy43 has positive bias during the three months for 10m wind,
- for cloud cover rmse of cy43 is higher than cy40 during the night period in April when both cycle rmse are almost same in June.
-No significant change was observered for the upper air so that cycle43t2_bf10 will be the operational cycle for ALARO.

4 References

Yang, X. 2008, Development of Hirlam/Harmonie monitoring system. HIRLAM Newsletter no 54, 62-65.

NWP activities in Romania

Alina Dumitru, Simona Taşcu, Alexandra Crăciun, Mihaela Neacşu

1 Introduction

In the present work, two different cycles of ALADIN/ALARO model were used in order to assess the performance of the actual operational version (cy43) against the previous model version (cy40). Other work was focused on finding an optimal setup in order to increase the model horizontal resolution. Some preliminary results are presented in this paper.

2 Forecast validation between CY43t2 and CY40t1

The verification of numerical models is a highly important issue, demanding a long series of data for a more complete validation. Due to some hardware problems, the migration on the new machine of the actual model version was done very quickly. Firstly, a validation for a 2-month period was done which was presented in the previous newsletter [1]. As a second step, in order to have a long series of data (leading to a more complex verification), the cycle 40 model version was also installed on the new machine. The same settings were used for both cycles: 6.5 km horizontal resolution, 60 vertical levels, 240 s time step, ALARO-0 baseline physical package. For both experiments, in line fullpos is used, with the following settings in *NAMFPD*: *NLAT=201, NLON=201, RLONC(1)=25.5, RLATC(1)=45.5, RDELX(1)=0.085, RDELY(1)=0.06.*

After obtaining the parallel datasets (integrations of cy40 and cy43), 78 hour forecasts provided by 00 UTC runs, for a period of six-month, from 1st of March till 31st of August 2020, were used for comparison. The performance evaluation was done for surface parameters (2 m temperature, 10 m wind speed, mean sea level pressure, total cloudiness and 6-hour accumulated precipitation) against 157 synop stations from Romania. The verification domain covers the operational area, from 17. to 34.º E and 39.5 to 51.5º N, using a resolution of 0.085º x 0.06º. The scores used in this comparison were the classical statistical scores BIAS and RMSE.

The results of BIAS and RMSE for monthly mean over the 6-month period indicates similar performance for all evaluated parameters. So, only two parameters were chosen for a further evaluation. Figure 1 contains monthly mean BIAS and RMSE of 2 m temperature forecast, over the whole period and for several forecast ranges: 12, 24, 36, 48, 60 and 72 hours. Figure 2 shows the same scores, but for 10 m wind speed. In terms of RMSE, the value of the score is almost the same for both cycles, for all parameters and all months which are under investigation. For 2 m temperature forecast BIAS, there are no significant differences. Even if similar results are obtained for 10 m wind speed, in this situation we can notice that in all forecast ranges, the values of the BIAS score for cycle 43 are constantly positive and generally negative for cycle 40. As a general conclusion, in the case of wind speed at 10 m, we can learn that cycle 43 might lead to some degree of overestimation of the wind speed, while the opposite was noticed for some forecast ranges and months in the forecast of cy40.

Another important aspect of the comparison between these two cycles is related to the running time: for cycle 40 it takes around 51 minutes for 78 forecast ranges, while for cycle 43 only 15 minute are spent for the same forecast length.

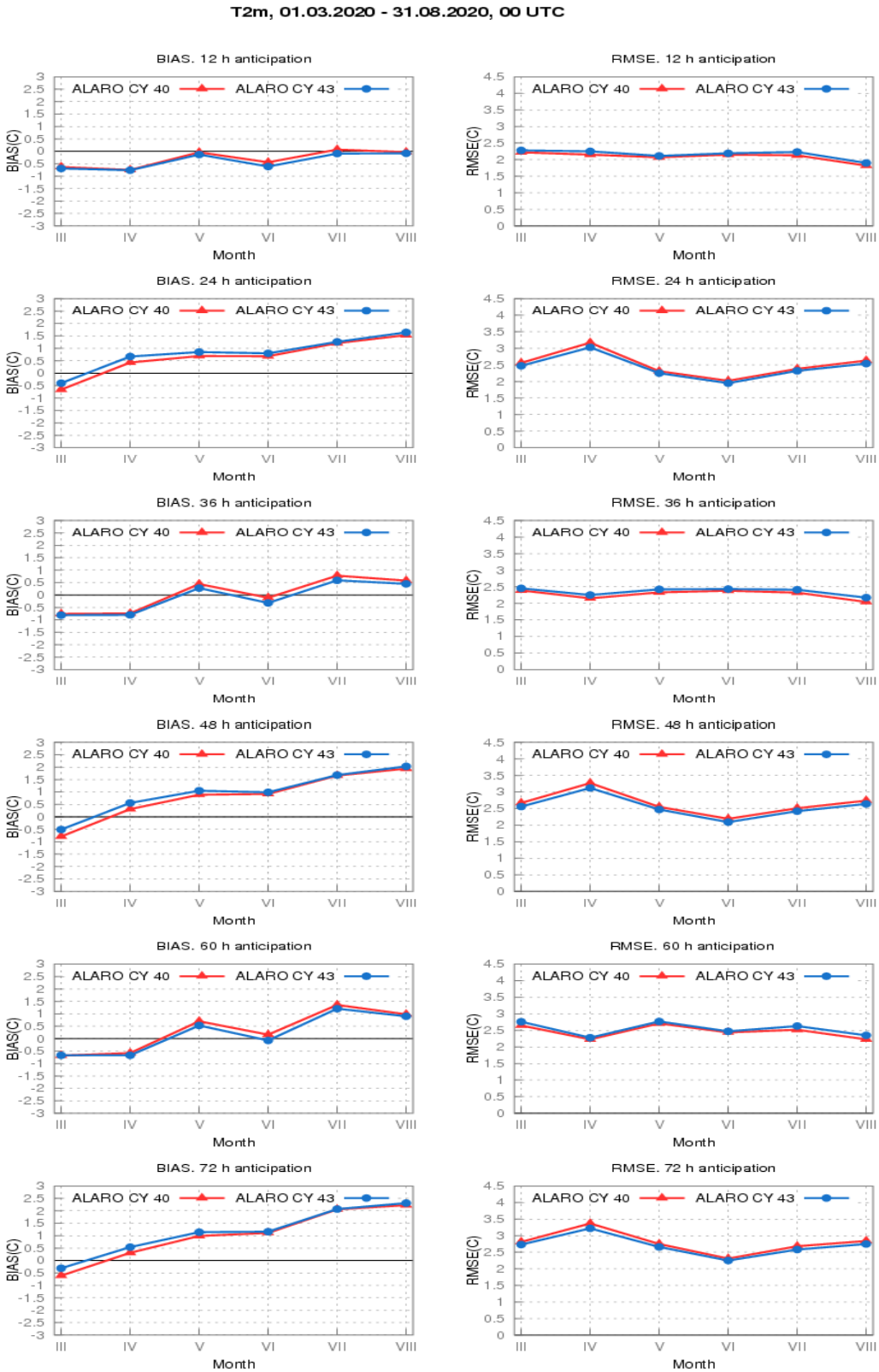


Figure 1: Monthly mean BIAS and RMSE of 2 m temperature, over the period 1st of March till 31st of August 2020, based on 00 UTC runs, for 12, 24, 36, 48, 60 and 72 forecast ranges.

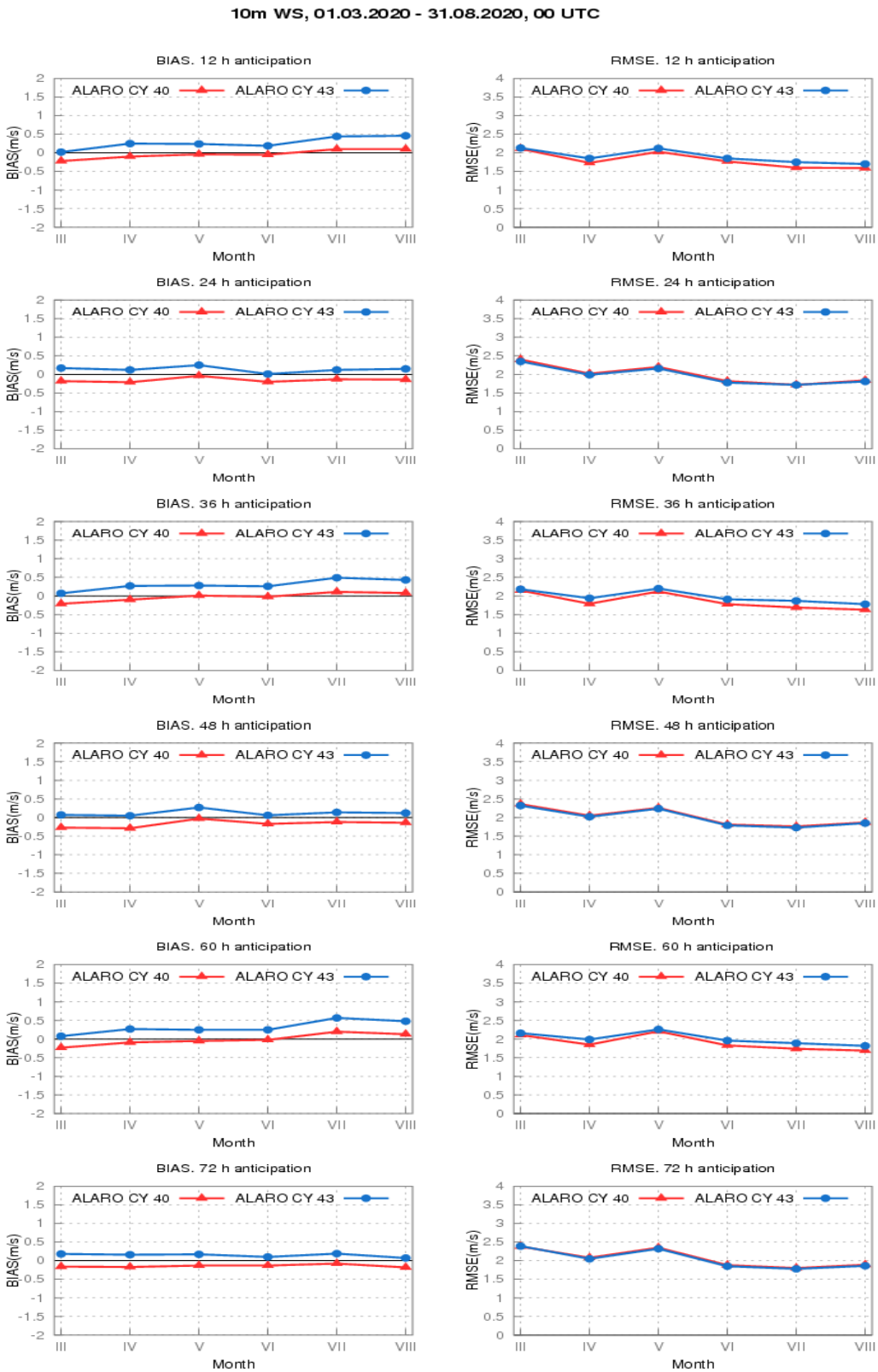


Figure 2: Monthly mean BIAS and RMSE of 10 m wind speed, over the period 1st of March till 31st of August 2020, based on 00 UTC runs, for 12, 24, 36, 48, 60 and 72 forecast ranges.

3 Preparation of the new operational version

One important point in our future objectives is to build an optimal setup in order to increase the horizontal resolution from 6.5 km to 4 km. Several integration domains were prepared for this purpose and in Figure 3 three domain sizes can be observed: 6.5 km horizontal resolution (blue) and two domains at 4 km horizontal resolution (dark green and red). Table 1 shows the difference in terms of grid points for these 3 domains. At the moment, all necessary configurations (ee927, e001) are running for both domains at 4 km horizontal resolution, with 60 vertical levels in dynamic adaptation mode and they are prepared for future implementation in operational mode.

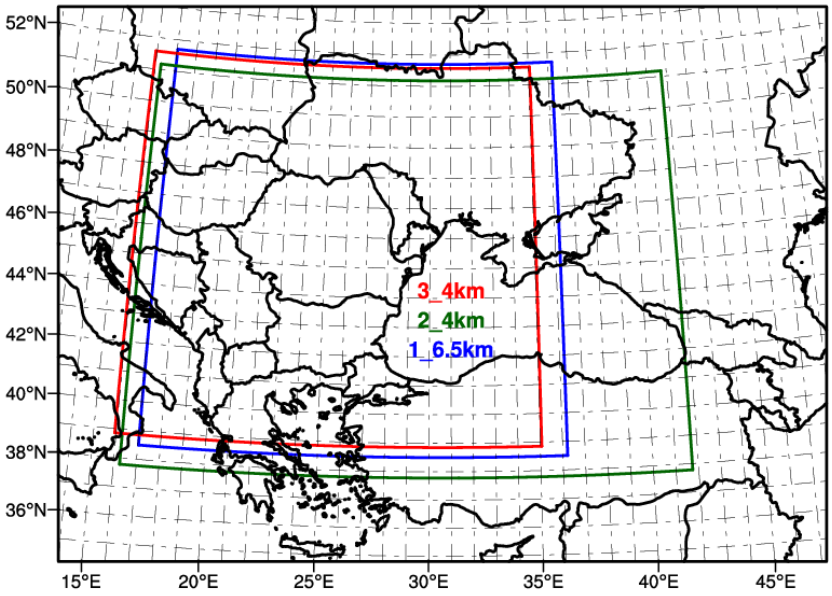


Figure 3: Representation of the tested domains: 6.5 km horizontal resolution – blue (exp 1), 4 km horizontal resolution – dark green (exp 2) and red (exp 3).

Table 1: Number of grid points for the tested domains

Resolution	6.5 km (exp 1)	4 km (exp 2)	4 km (exp 3)
Number of points on OX	240	600	450
Number of points on OY	240	432	400

We have chosen two domains for 4 km horizontal resolution in order to perform more experiments. In the future we plan to implement in the new operational chain a data assimilation system which is under development at this moment. However, depending on the computing power availability, we will identify the most suitable domain for our operational constraints.

4 Preliminary results at 4 km horizontal resolution

In order to start assessing the impact of switching to higher resolution as opposed to the operational one, a few experiments realized at 4 km resolution based on ALARO-1vb were done. Some preliminary results can be seen in Figure 4. It can be observed that the amount of precipitation from the south-eastern part of Romania is better captured in the forecast obtained from the bigger domain at 4 km horizontal resolution. It is expected to have better results when a bigger domain is used, especially when the domain covers the Black Sea. Past studies indicated that when different

convective systems occurred above or near this region, because of the right side of the domain boundary conditions, meteorological phenomena happening there were not properly simulated.

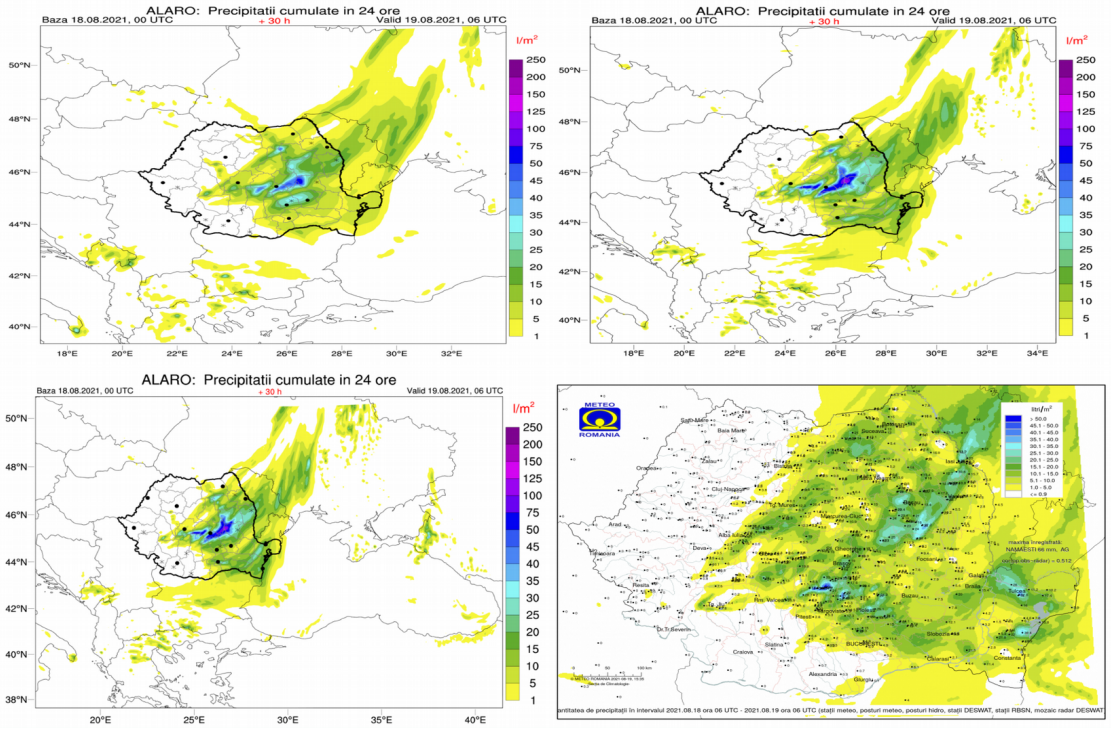


Figure 4: 24-hour accumulated precipitation: operational domain at 6.5 km horizontal resolution (top-left), smaller domain at 4 km horizontal resolution (top right), bigger domain at 4 km horizontal resolution (bottom left), observed precipitation (bottom – right).

For both experiments at 4 km horizontal resolution, in line fullpos is used, with the following settings in *NAMFPD*:

Table 2: NAMFPD details

NAMFPD	**4 km (exp 2)**	**4 km (exp 3)**
NLAT	373	358
NLON	501	371
RLONC(1)	29.	25.5
RLATC(1)	44.5	45.5
RDELX(1)	0.05	0.05
RDELY(1)	0.035	0.035

In order to evaluate the computing resources required for the integration of the model at different setups, some scalability tests were performed. The tests were done by variation of the number of nodes used in the e001 configuration and the results are expressed in terms of time necessary for the job to complete (Figure 5).

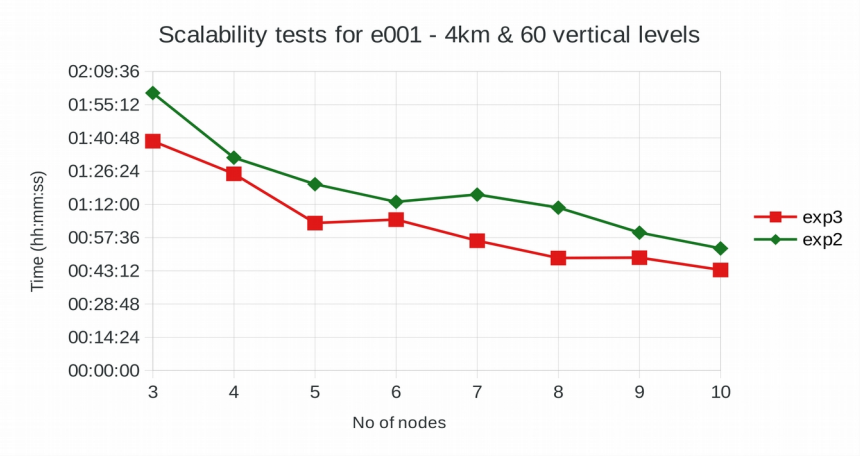


Figure 5: Scalability tests for configuration e001 for both domains at 4 km horizontal resolution.

5 Preliminary tests with higher resolution

Several tests were done to investigate the technical possibilities to go to even finer resolution, since nowadays high (or very high) resolution started to be more and more required for different means and applications. Some preliminary experiments were done. Figure 6 shows the following domain sizes: 2.5 km horizontal resolution (green) and 2 km horizontal resolution (magenta). In Table 3, the difference in terms of grid points for the high resolution domains used in these experiments can be seen. At the moment, only configuration ee927 works for 2 and 2.5 km horizontal resolution while for configuration e001 we are still facing some problems that are under investigation.

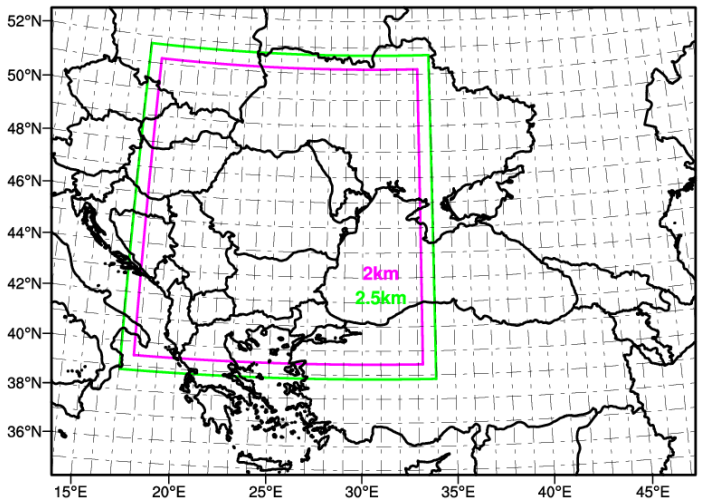


Figure 6: Representation of the tested domains: 2.5 km horizontal resolution – green, 2 km horizontal resolution – magenta.

Table 3: Number of points for the finer resolution tested domains

Resolution	2.5 km	2 km
Number of points on OX	640	720
Number of points on OY	640	720

The *clim files* necessary for these tests were obtained on *belenos* machine at Météo-France using two tools: *epygram* and *climake*. Firstly, the domain was created by using *domain_maker.py* from *epygram*.

The results from this step were used as input for the *climake* tool in order to generate *clim* files for all the tested domains.

6 Conclusion

The basics of implementing a new operational setup at 4 km horizontal resolution were done. Considering the results obtained and taking into account the computing constraints and characteristics of the future data assimilation system, only one setup will be selected for operational mode, while the others may be used for different purposes. A part of the results are obtained within a project related to the improving of air quality assessment and monitoring system at national level (SMIS 139703).

7 References

[1] Alina Dumitru, Răzvan Dobre, Simona Taşcu, Alexandra Crăciun (2021). ALADIN activities in Romania. ALADIN-HIRLAM Newsletter 16:100-105,
http://www.umr-cnrm.fr/aladin/IMG/pdf/nl16.pdf

Validation of AROME CY43T2 configurations at IPMA

João Rio, Manuel João Lopes, Maria Monteiro, Nuno Lopes

1. Introduction

The aim of this article is to report some aspects and conclusions on the porting and validation of the Portuguese version of the AROME dynamical adaptation configuration to cycle CY43T2_bf10 (from now on designated just by CY43). Three geographical domains of the local numerical system are considered.

At the present time, Portugal runs in operations a local version of the AROME model to CY40 [1] by dynamical adaptation of the ARPEGE model. The three domains where this configuration is run cover the Iberian Peninsula and the Atlantic Archipelagos of Madeira and Azores (Figure 1). The configurations are known, respectively, by AROME-PT2, AROME-MAD and AROME-AZO and their main characteristics are described in Table 1. Besides, it runs a surface Data Assimilation cycling by the so-called CANARI-OI-MAIN (see, for instance, [2]), which is used to provide the background to a CANARI hourly analysis, based on AROME. The validation of the Data Assimilation configurations at CY43T2 will be delivered afterwards.

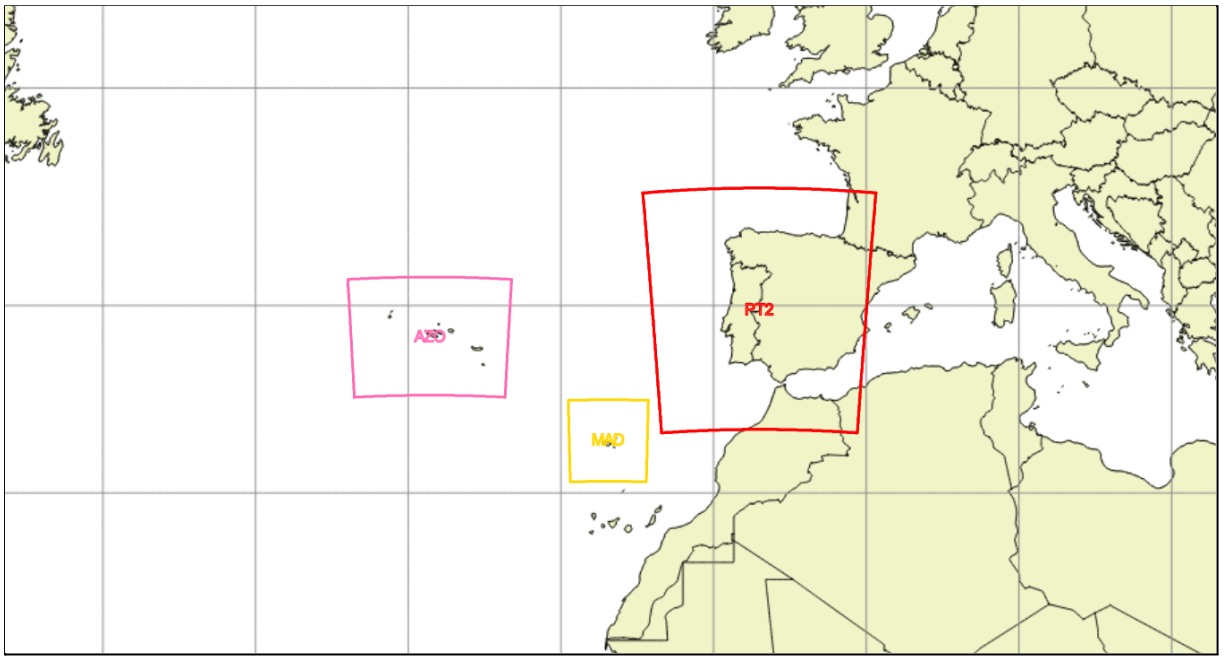


Figure 1: Portuguese geographical location of the actual Portuguese AROME configurations used in operations: PT2, covering the Iberia and adjacent Atlantic region (in red); MAD, covering the Madeira Islands (in yellow); and AZO, overing the Azores Islands (in magenta).

Recently, the local High-Performance Computing (HPC) infrastructure, an IBM P7+ with 9 cores, was considered obsolete to cope with new versions of the ALADIN system [1], hence the porting of the numerical configurations has taken place on ECMWF platforms. In this way, it was possible to

proceed with a systematic validation of the actual operational configuration (and the creation of a remote backup system).

Table 1: Portuguese AROME dynamical adaptation configurations at CY43.

Geographical domain	Iberia	Madeira	Azores
Designation	AROME-PT2	AROME-MAD	AROME-AZO
Resolution	2.5 km	2.5 km	2.5 km
Levels	60	60	60
Initial conditions	ARPEGE (H+00)	ARPEGE (H+00)	ARPEGE (H+00)
Boundaries	ARPEGE	ARPEGE	ARPEGE
Starting times	00 UTC	00 UTC	00 UTC
CANOPY scheme	NO	YES	YES
N2M (2-metre diagnostic)	2	1	1
LNOTS_T flag	false (default)	true	true

The structure of this article is as follows: the introduction is given in this Section; Section 2 delivers the general conclusions of the validation study as well as the illustration of selected results in each of the domains; Section 3 reports a long-term issue related to the Azores domain; and, Section 4 registers some short conclusions and a possible outlook.

2. Validation of AROME CY43

In order to get a general view of the new cycle's performance, several experiments have been made over the Portuguese domains, keeping the actual operational namelists, tuned for the CY40, as shown in Table 1. For each domain, the 00 UTC AROME dynamical adaptation was integrated up to a 48-hour lead time, for two seasonal periods: a 55-day Summer period, from 16 July 2021 to 08 September 2021; and a 90-day Winter period, from 01 December 2020 to 28 February 2021.

In this study, only surface and screen-level parameters have been examined using the local processing and verification tools. The validation of the new cycle was done using the operational forecasts at CY40 as the reference. The observations used are from the surface network available at IPMA, hence the maximum number of weather stations used is 153 in Mainland, 20 in Madeira and 10 in Azores. The validation comprises time series, scatterplots, classical statistical parameters and examination of spatial forecast differences, for the 2-metre temperature (T2M) and relative humidity (H2M), 10-metre

wind speed (W10M), mean sea-level pressure (MSLP), 24-hour precipitation (24PP), maximum (TMAX) and minimum (TMIN) daily temperatures.

The scores shown in this report are the Root-Mean-Square Error (RMSE), forecast accuracy (ACCU) - probability of the forecast to be within a given range of the observations - and the Heidke Skill Score (HSS). Frequency (PDF) and cumulative (CDF) distribution functions are also used.

Table 2 summarises the overall forecast skill of CY43, when compared to the operational (CY40). The view takes into account several verification results and is valid for the whole forecast range.

Table 2: Qualitative validation of CY43 vs, CY40(*).

Geographical domain	Iberia		Madeira		Azores	
MAX number of stations	153		20		10	
Season	Summer	Winter	Summer	Winter	Summer	Winter
T2M	↑↑	↓↑	↓↑	↓↑	↓↑	↓↓
H2M	↑↑	↓↑	↑↑	↑	↑↑	↓
V10M	↑	↓	↑	↓↑	↓	↓
MSLP	↑	↓↑	↓↑	↓↑	↑↑	↓↑
24PP	↑	↓↑	↑	↓↓	↑	↓
TMAX	↓↑	↓	↓↑	↓↑	↑	↑
TMIN	↑↑	↓	↓	↓	↓	↓↓

(*) 'upward green' arrows shows the new cycle improves the forecast; 'downward red' arrows shows the new cycle does not improve the forecast; 'double direction yellow' arrows suggest the change is neutral.

Overall, AROME CY43 improves the forecasts over Iberia and Madeira in the Summer period and is neutral in winter. In the Azores, the cycle change is neutral during Summer, but shows a degradation in Winter.

In the next subsections some of the available results are shown more in detail, followed by some comments.

2.1 Scores over Iberia

Figure 2 shows the RMSE for the 2-metre temperature and relative humidity, the 10-metre wind speed (for values above or equal to 4 m/s) and the HSS for 24-hour precipitation. Figure 3 shows the CDF and the PDF of the 2-metre temperature and 10-metre wind speed. These results are valid for the Summer 2020. Figure 4 is similar to Figure 2, but valid in the Winter 2020/21. Given the much larger availability of precipitation data, in the Winter period the HSS is shown for 3-hour periods.

As it can be seen from Figures 2 and 4, in Mainland Portugal CY43 is clearly beneficial during Summer, although neutral in the Winter period. In particular, as illustrated in Figure 3, the good results in Summer are in the whole range of values.

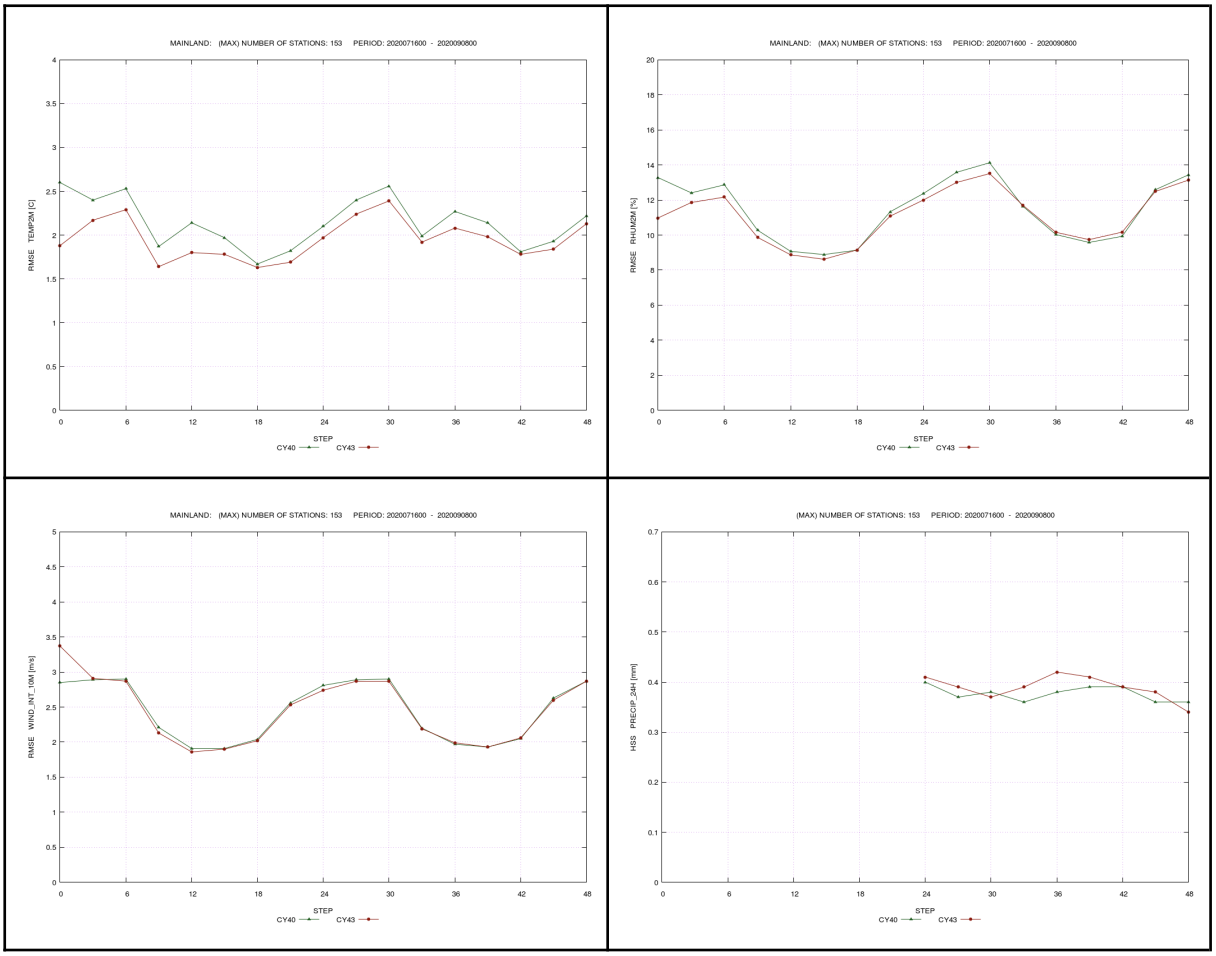


Figure 2: RMSE for AROME-PT2 Summer forecasts, of: 2-metre temperature (upper left panel); 2-metre relative humidity (upper right panel); and 10-metre wind speed (for values above or equal 4 m/s, bottom left panel); and HSS for AROME-PT2 Summer 24-hour precipitation (bottom right panel).

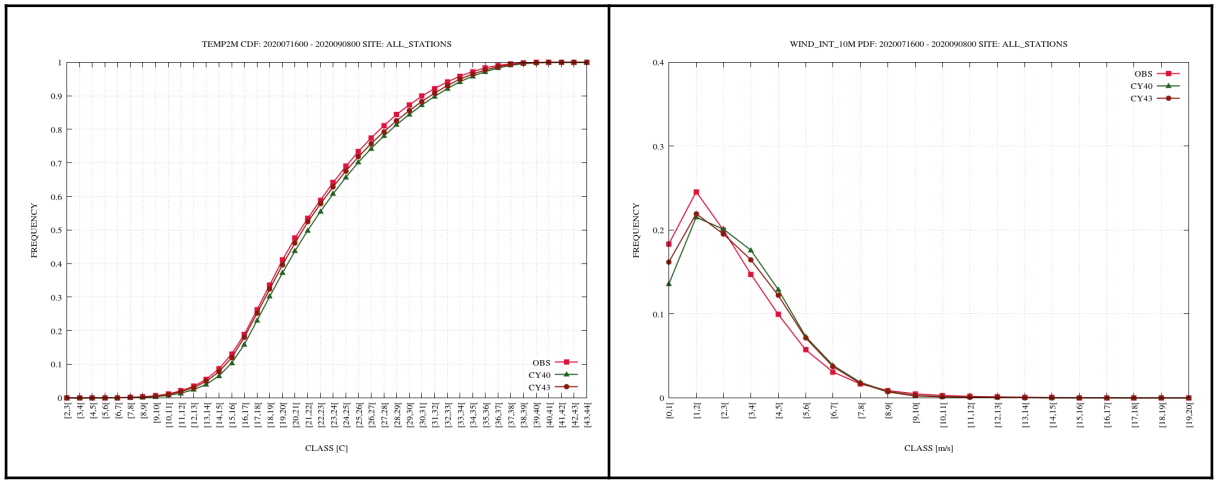


Figure 3: CDF for AROME-PT2 Summer forecasts, of 2-metre temperature (left panel); and PDF for AROME-PT2 Summer forecasts, of 10-metre wind speed (for values above or equal 4 m/s, right panel).

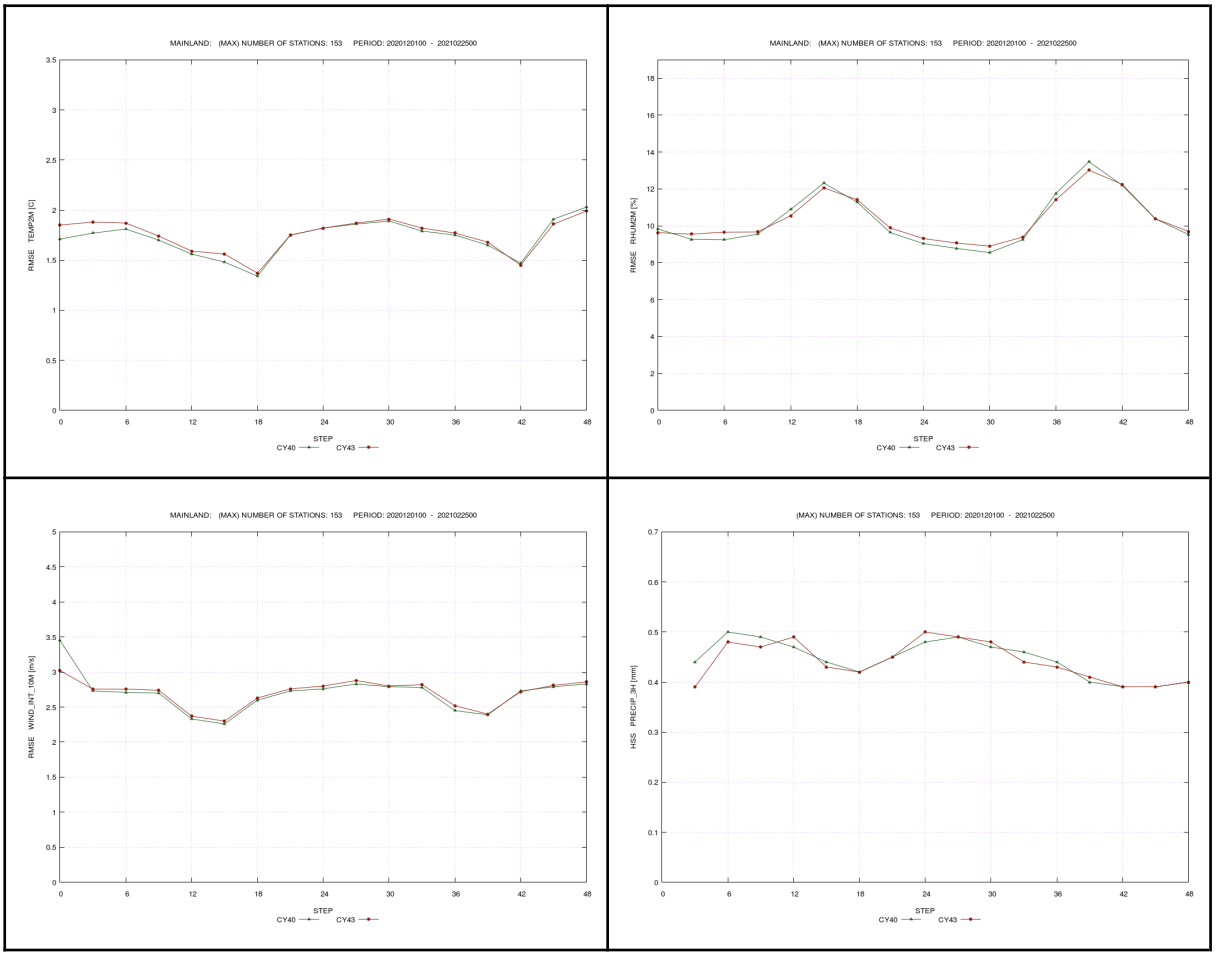


Figure 4: RMSE for AROME-PT2 Winter forecasts, of: 2-metre temperature (upper left panel); 2-metre relative humidity (upper right panel); and 10-metre wind speed (for values above or equal 4 m/s, bottom left panel); and HSS for AROME-PT2 Summer 24-hour precipitation (bottom right panel).

2.2 Scores over Madeira

Figure 5 shows the RMSE for the 2-metre temperature and relative humidity, the 10-metre wind speed (for values above or equal to 4 m/s) and the CDF of the 2-metre relative humidity. Figure 6 shows the spatial distribution of the RMSE for the daily 2-metre maximum temperature. These results are valid for the Summer 2020. Figure 7 is similar to Figure 5, but the CDF is for the 10-metre wind speed during the Winter 2020/21. Likewise, Figure 8 is similar to Figure 6 and illustrates the same results for the Winter period.

Overall, over Madeira, CY43 is neutral or beneficial, with improvements seen mainly during Summer. The Summer CDF for the 2-metre relative humidity confirms the added value of CY43, which is clearly achieved for all ranges of values.

The 10-metre wind speed CDF shows that CY43 is a better match to the observations only for values below 3 m/s. Above this threshold, CY40 has a better agreement with the observational CDF, both in Summer (not shown) and Winter.

Besides, Figures 6 and 8 show there is no major local impact on the maximum temperature when moving to CY43 (during both Summer and Winter).

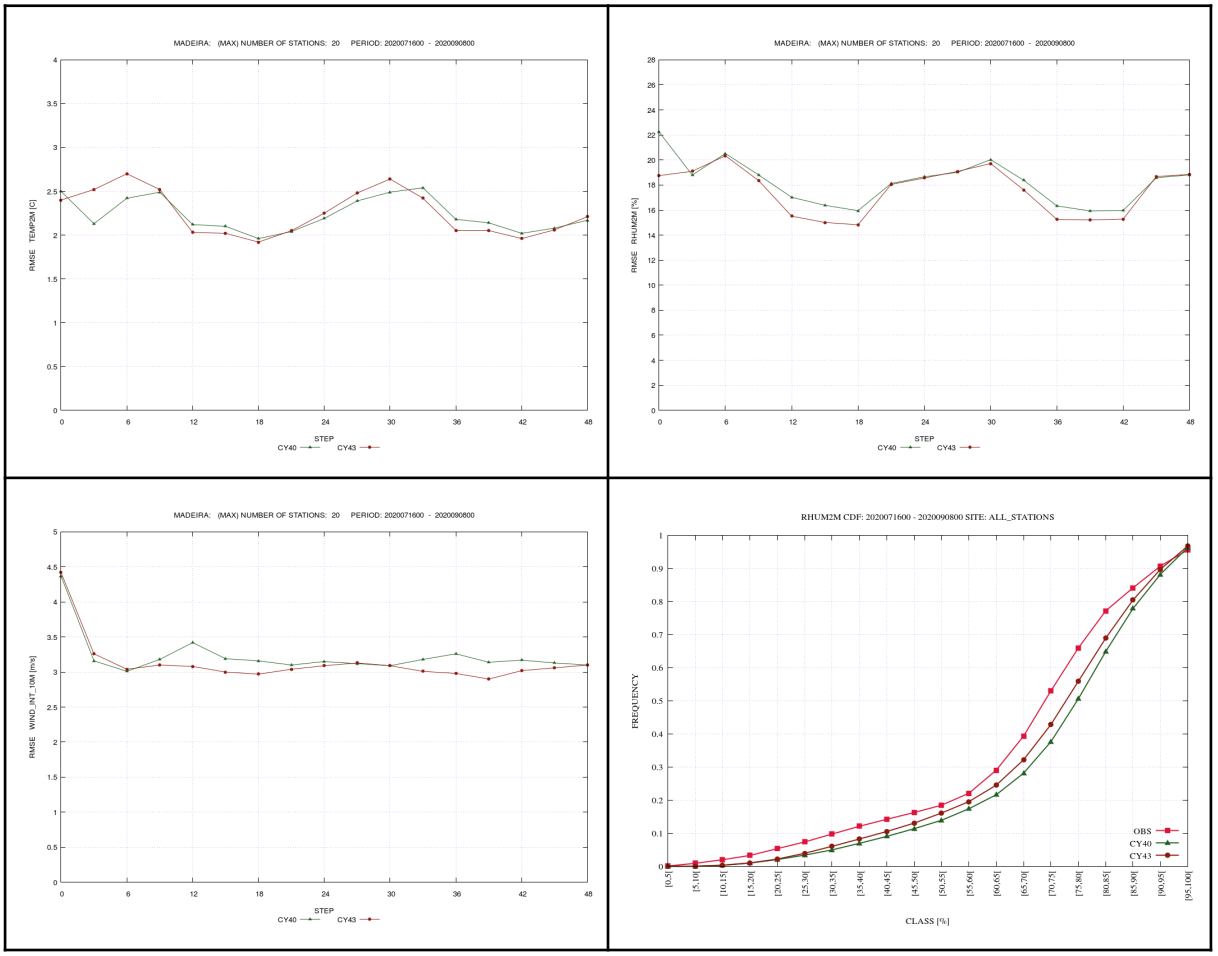


Figure 5: RMSE for AROME-MAD Summer forecasts, of: 2-metre temperature (upper left panel); 2-metre relative humidity (upper right panel); and 10-metre wind speed (for values above or equal 4 m/s, bottom left panel); and CDF for AROME-PT2 Summer 2-metre relative humidity (bottom right panel).

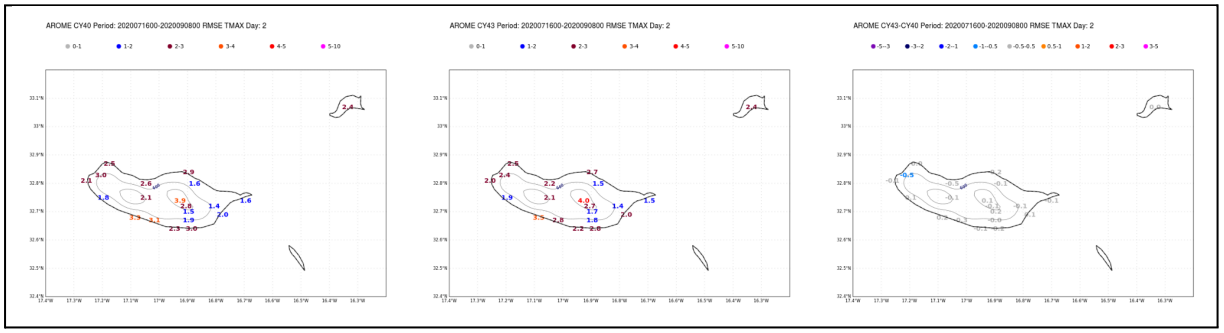


Figure 6: Spatial distribution of RMSE, and difference, for AROME-MAD Summer forecasts, of 2-metre maximum temperature (24-48 hour period), obtained with different cycles: CY40 (left panel); CY43 (central panel); CY40-CY43 (right panel).

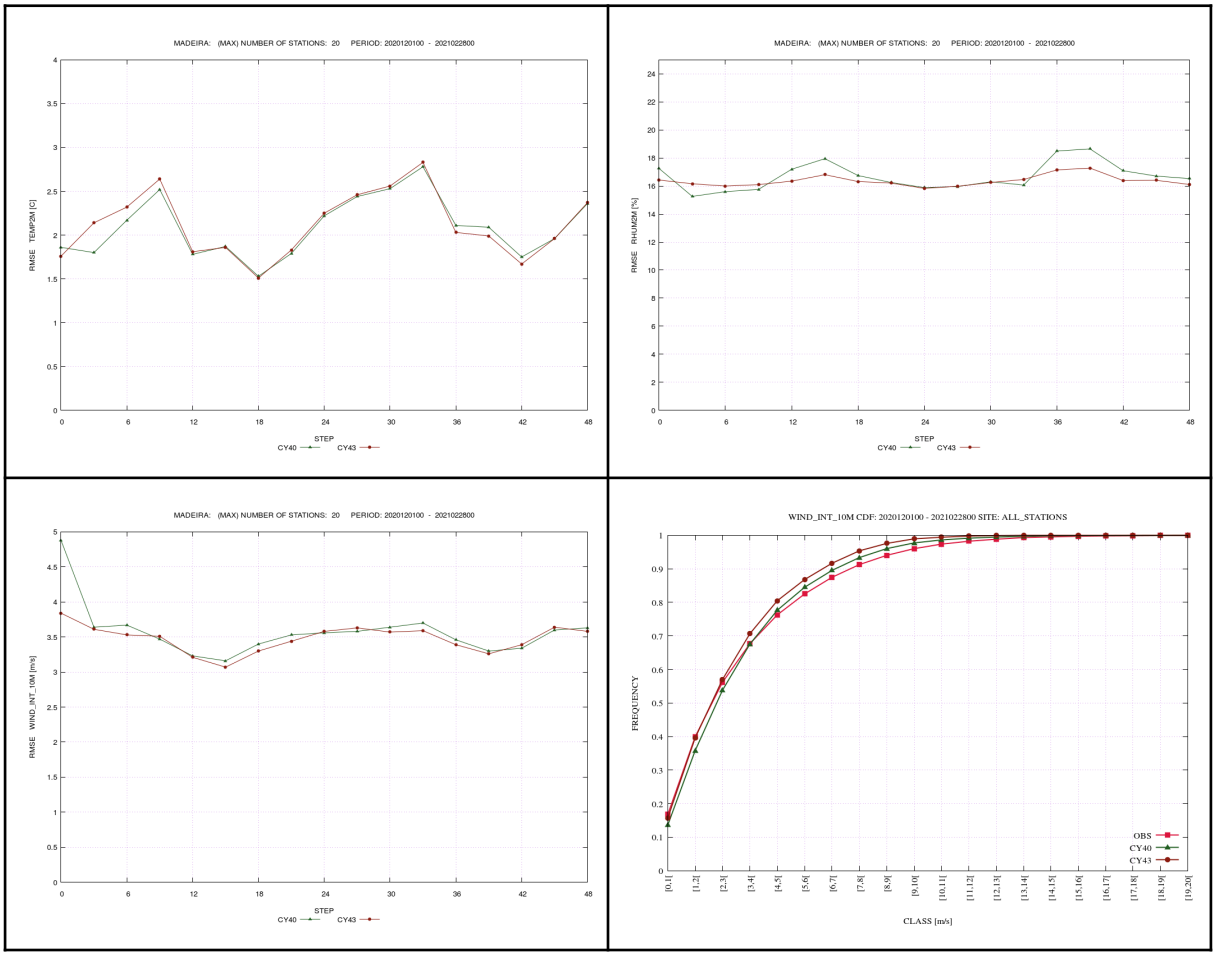


Figure 7: RMSE for AROME-MAD Winter forecasts, of: 2-metre temperature (upper left panel); 2-metre relative humidity (upper right panel); and 10-metre wind speed (for values above or equal 4 m/s, bottom left panel); and CDF for AROME-PT2 Summer 10-metre wind speed (bottom right panel).

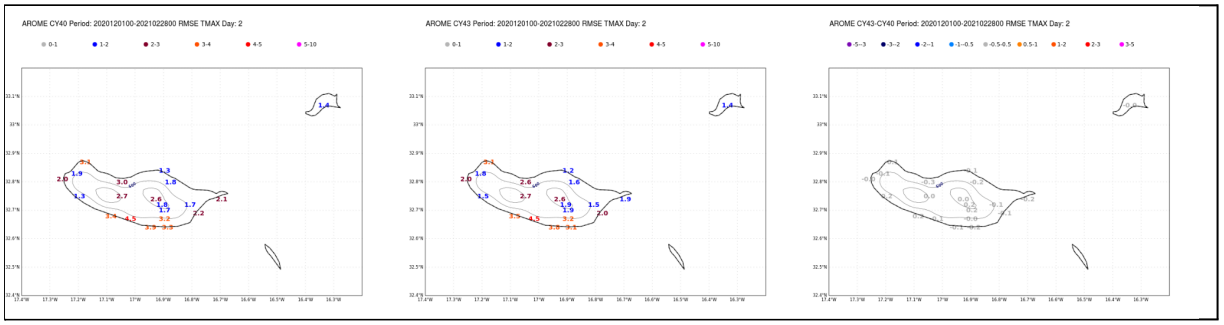


Figure 8: Spatial distribution of RMSE, and difference, for AROME-MAD Winter forecasts, of 2-metre maximum temperature (24-48 hour period), obtained with different cycles: CY40 (left panel); CY43 (central panel); CY40-CY43 (right panel).

2.3 Scores over Azores

Figure 9 shows the RMSE for the 2-metre temperature and relative humidity, the 10-metre wind speed (for values above or equal to 4 m/s), as well as the accuracy of the latter. These results are valid during

Summer 2020. Figure 10 is similar, but is valid in Winter. Figure 11 shows the CDF of the 2-metre temperature, in both periods.

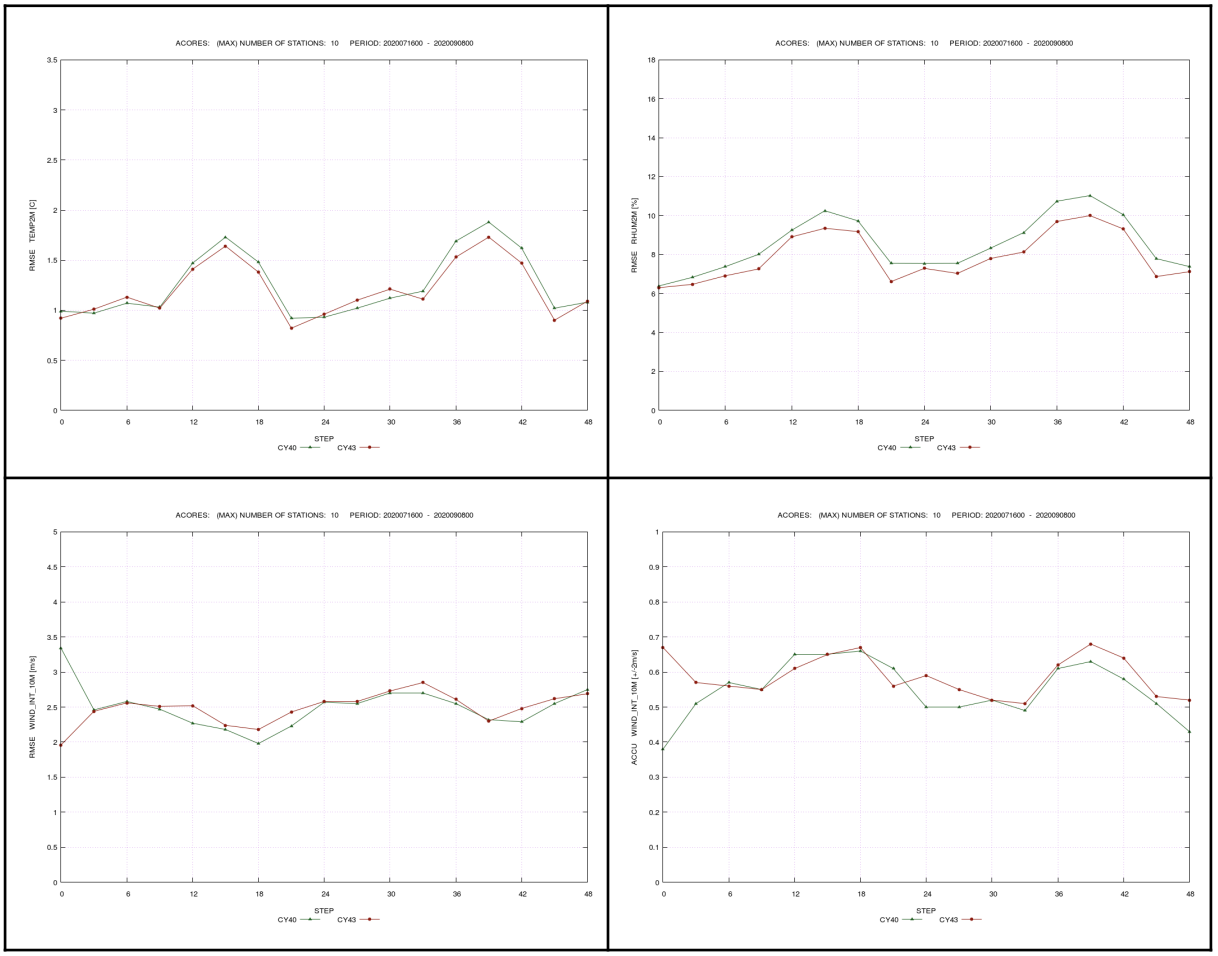


Figure 9: RMSE for AROME-AZO Summer forecasts, of: 2-metre temperature (upper left panel); 2-metre relative humidity (upper right panel); and 10-metre wind speed (for values above or equal 4 m/s, bottom left panel); accuracy for AROME-PT2 Summer 10-metre wind speed (bottom right panel).

In the Azores, CY43 shows an improvement in the Summer 2020. However, in the Winter 2020/21 the scores suggest that, at best, CY43 is neutral.

In the case of the 10m wind speed, the RMSE of CY43 is higher than CY40, in the Summer period. However, as the same happens with the accuracy, this suggests that the increase in RMSE may be due to a higher variability in CY43. The 24-hour precipitation scores (not shown) suggest CY43 is beneficial in Summer, with a slight degradation in Winter.

Furthermore, the CDF for the 2-metre temperature during the Summer period of CY43 is closer to the observations in all ranges of values. This is not seen in the Winter period.

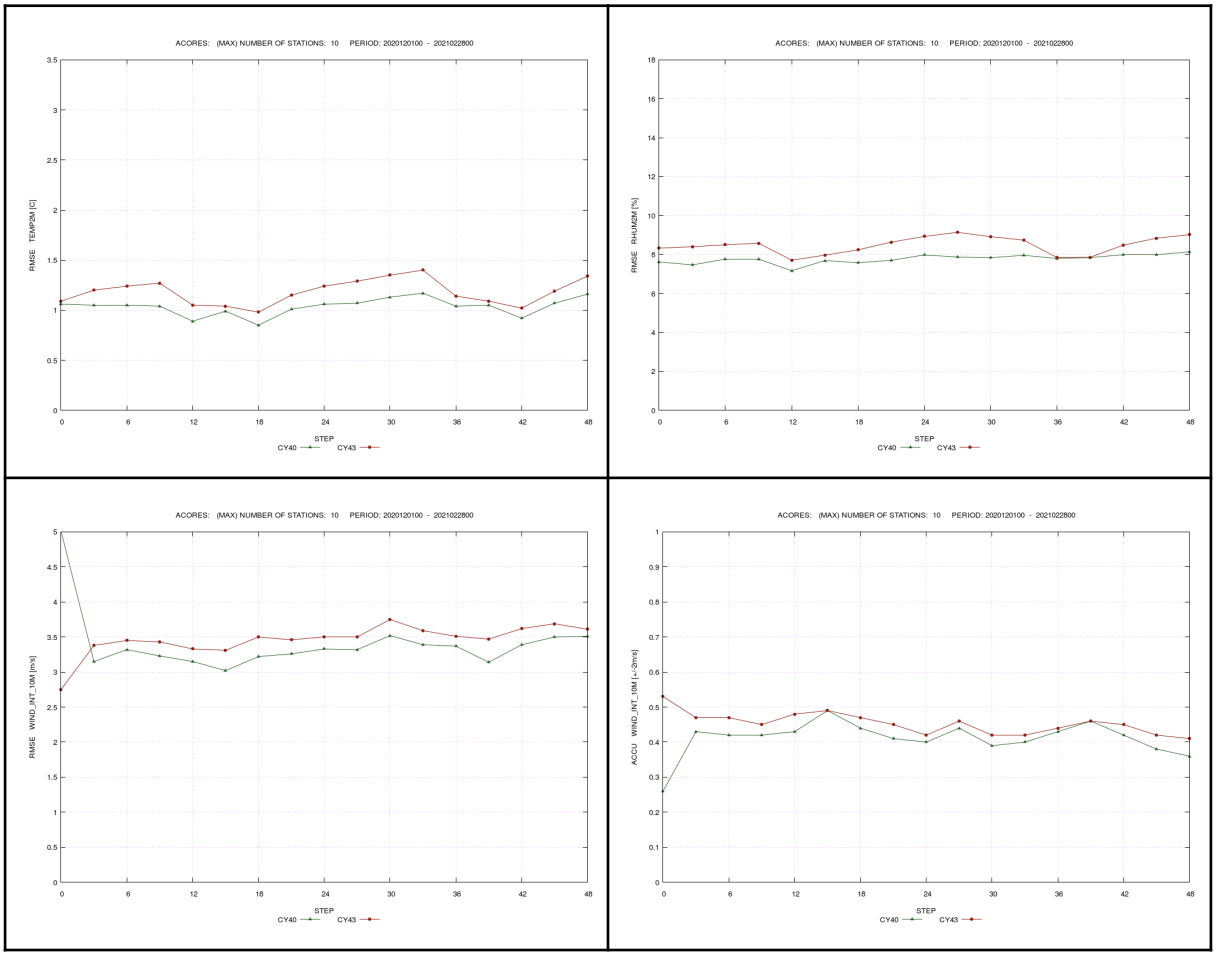


Figure 10: RMSE for AROME-AZO Winter forecasts, of: 2-metre temperature (upper left panel); 2-metre relative humidity (upper right panel); and 10-metre wind speed (for values above or equal 4 m/s, bottom left panel); accuracy for AROME-PT2 Winter 10-metre wind speed (bottom right panel).

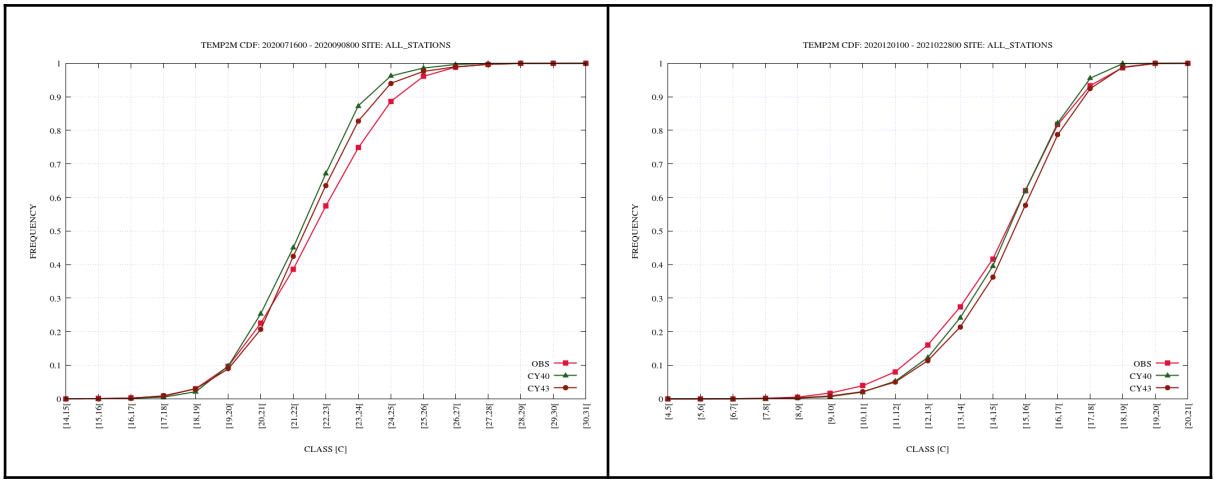


Figure 11: CDF for AROME-AZO forecasts, of: 2-metre temperature: during Summer (left panel); and Winter (right panel).

3. Particularities on AROME climatologies (CY43T2) over Azores

The porting of the Portuguese configurations to CY43 (which supports both the current operations and developments at IPMA) required the update of their climatologies to CY43. In fact, in this new cycle version, new ECOCLIMAP files were at play (due to the upgrade of SURFEX), on both "e923" and "001"' configurations, requiring the re-computation of AROME climatologies for the current Portuguese domains and, in particular, for the Azores domain in the Atlantic region. For this domain, the following subsections report a long-term issue, present in this computation since CY38T1 (subsection 3.1), and share the conclusions obtained with local sensitivity tests on the usage of some surface physiography fields (subsection 3.2).

3.1 Issues related to SAND and CLAY

During the update of the climatological files for the Portuguese operational AROME domains – Azores, Madeira and Iberia – to CY43T2 (GENV cycle) it was detected that the physiographic fields SAND and CLAY (soil texture) were "null" over the Azores, causing the abortion of the AROME integration (configuration "001"). These fields were "not null" over the other two domains.

This issue is seen since the local porting of AROME to CY38T1, forcing the use of the old CY35T2 climatologies in operations for the Azores domain up to now.

To overcome this problem during the porting to CY43, both fields were replaced by specific SAND and CLAY physiographic fields provided by Météo-France (namely: sand.HWSD_MOY_v1.10_ECO_I.tgz and clay.HWSD_MOY_v1.10_ECO_I.tgz), on the climatologies update (configuration "e923", from CLIMAKE).

To understand how the changes in SAND and CLAY could affect the forecasts of AROME-Azores (version CY43), the model was executed on a specific date (26 March 2021) with different values for SAND and CLAY:

- REF (reference value), using SAND=CLAY=0.33 (the PDG fields as provided by Météo-France);
- LOW (low value), using SAND=CLAY=0.03;
- HIGH (high value), using SAND=CLAY=0.9.

The impact evaluation of the changed SAND and CLAY was done by examining the forecast differences obtained from these executions. In order to maximise the differences it was chosen a stable weather situation. The variables focus were surface variables rather than the soil prognostic variables (water and temperature), whose impact should be larger. In particular, the variables considered were surface temperature (TSURF), 2-metre temperature (T2M) and 2-metre relative humidity (RH2M).

The evaluation of the forecasts was done with the experiment differences "HIGH" minus "REF" (HIGH-REF) and "LOW" minus "REF" (LOW-REF, using forecasts up to a 24-hour lead time, for the 00 UTC run on 26 March 2021. Figures 12 to 14 illustrate some of the results.

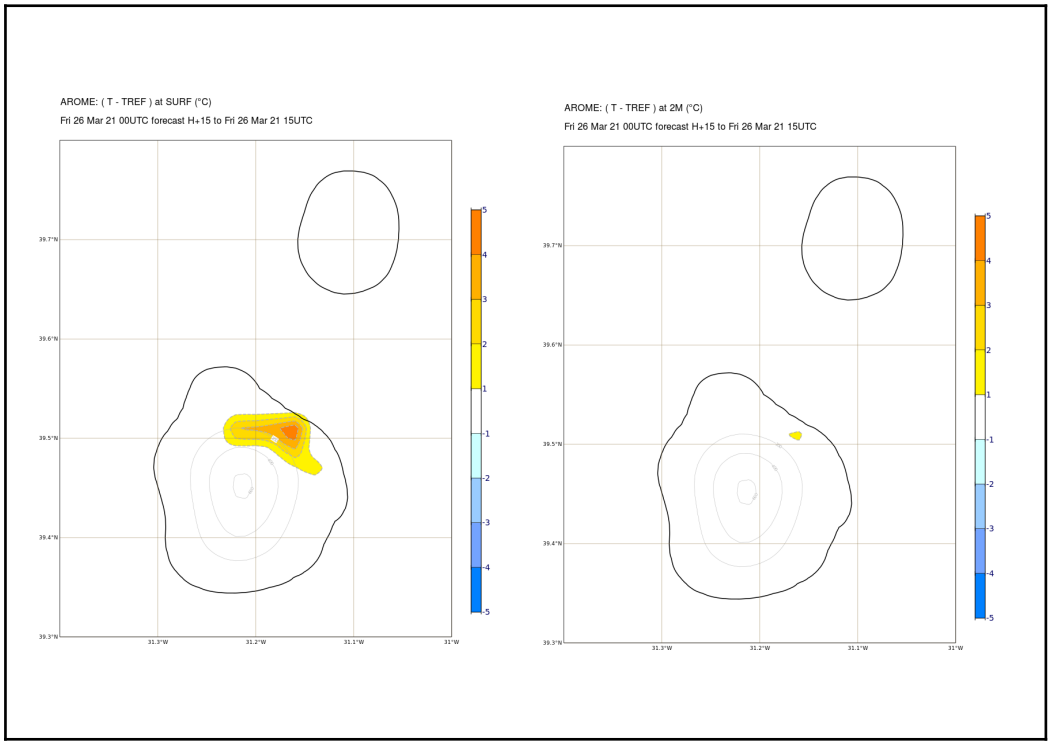


Figure 12: Spatial differences (LOW-REF) in AROME-AZO forecasts (step H+15), at the Western Islands of Azores (Flores), valid at 15 UTC on 26 March 2021, of: surface temperature (deg C, left panel); and 2-metre temperature (deg C, right panel). Orography is represented by grey isolines (contour interval of 200 metres).

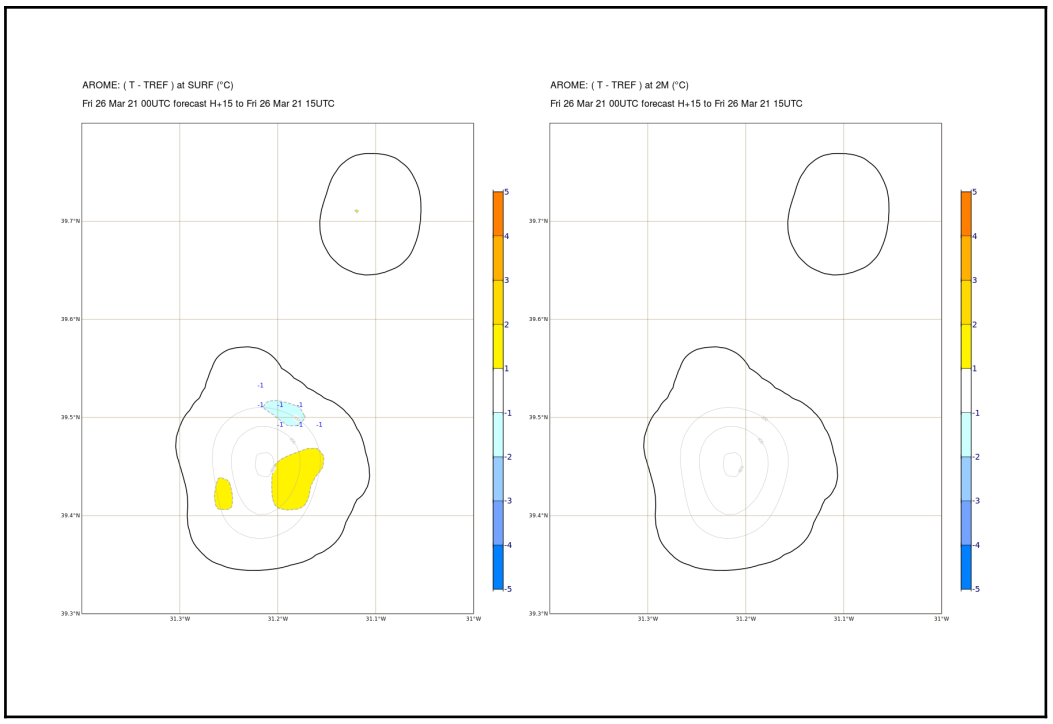


Figure 13: Spatial differences (HIGH-REF) in AROME-AZO forecasts (step H+15), at the Western Islands of Azores (Flores), valid at 15 UTC on 26 March 2021, of: surface temperature (deg C, left panel); and 2-metre temperature (deg C, right panel). Orography is represented by grey isolines (contour interval of 200 metres).

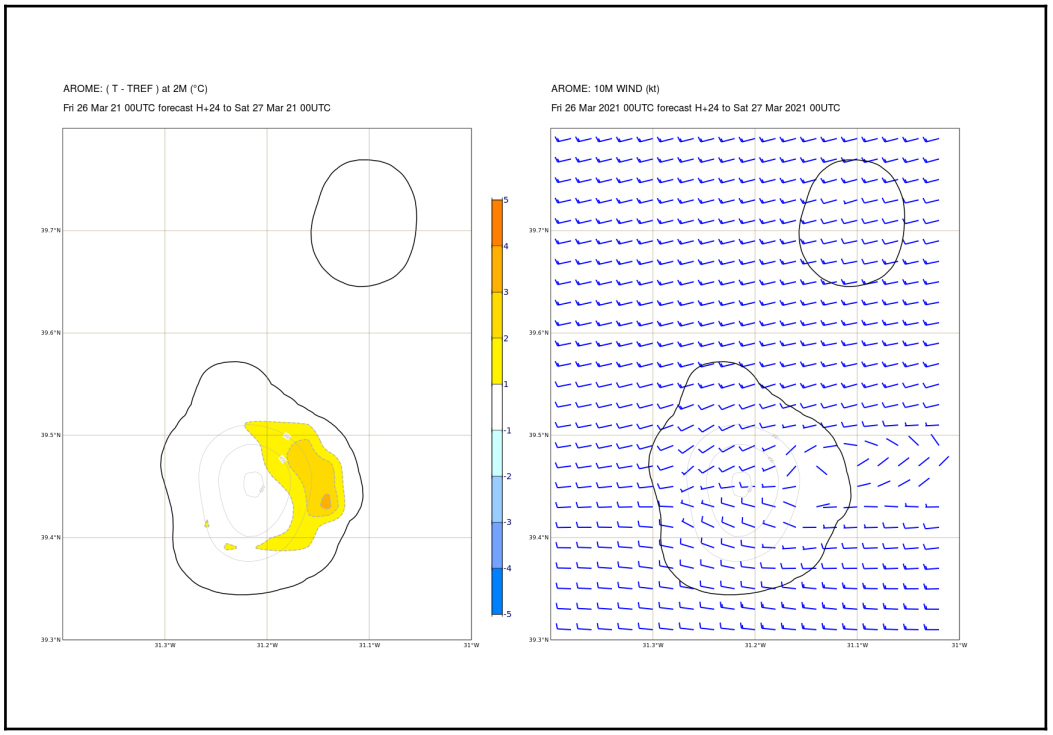


Figure 14: Spatial differences (HIGH-REF) in AROME-AZO forecasts (step H+24), at the Western Islands of Azores, valid at 14 UTC on 26 March 2021, of 2-metre temperature (deg C, left panel); and AROME-AZO forecast (step H+24) of REF 10-metre wind (kt, right panel). Orography is represented by grey isolines (contour interval of 200 metres).

In all experiments, the model runs were successful and the overall results show that:

- a few small forecast differences have been found in the comparisons (LOW-REF and HIGH-REF), mainly over land points;
- the maximum differences (considering all time steps) occured in the most Western island (Flores), where the model orography reaches 617 metres;
- the surface temperature forecast shows the largest difference in step H+15 - around 5ºC - corresponding to the comparison LOW-REF (left panel on Figure 12), when the difference of 2-metre temperature is around 1ºC (right panel on Figure 12). The forecasts H+15 with respect to HIGH-REF are shown in Figure 13;
- the 2-metre temperature shows the largest difference in step H+24 - around 3ºC - correspondent to HIGH-REF (left panel on Figure 14);
- the larger differences of 2-metre temperature seems to occur during the night in small areas, at the leeside of the mountains, which are associated to lower intensities of the 10-metre wind (right panel on Figure 14) and less mixing;
- as the difference in 2-metre temperature forecasts is somewhat small, the impact in the 2-metre relative humidity is also limited (not shown).

Taking into account the above results, the forecasts made with the SAND=CLAY=0.33 seem to be a good compromise while the issues that exist with the T-codes databases for the Azores domain are solved.

3.2 Sensitivity tests to SURFPROP.VEGETAT and SURFPROP.VEG.MAX

In addition to the identification of the above issue, it was seen during the climatologies update that the fields SURFPROP.VEGETAT and SURFPROP.VEG.MAX were equally “null” over Azores (though “not null” over the other two domains). Despite this lack of realistic climatological data over the Azores, the execution of AROME is not affected.

In order to verify if SURFPROP.VEGETAT and SURFPROP.VEG.MAX are really used by the model, sensitivity experiments, discussed with Météo-France, have been performed. The methodology consisted in changing the mentioned fields inside the climatological files and check if the outputs of configuration “ee927_surf_prep” (which prepares the SURFEX initial file) are modified when compared to a reference run (i.e., a run of AROME-Azores using the climatological files with the “null” fields). In fact, it could be possible that both fields, which are taken from old databases, could have become obsolete since SURFEX is at play in the local version of AROME. In addition to a reference run (REF), three experiments were performed for 00 UTC of 22 March 2021, as described below:

- AZO1, with both fields removed from the model climatological files;
- AZO2, with SURFPROP.VEGETAT=1 (the highest value);
- AZO3, with both fields identical to 1.

The main aspects of these experiments are summarized in Table 3. The output files of “ee927_surf_prep” were compared using *epygram*.

Table 3: Sensitivity experiments description to SURFPROF.VEGET and SURFPROF.VEG.MAX.

EXP ID	**SURFPROF.VEGETAT**	**SURFPROF.VEG.MAX**	**ee927_surf_prep**	**001**	**Remarks**
REF	0	0	OK	OK	--
AZO1	removed	removed	ABORT: “ *FACIL1 - KREP= -89, KNUMER=54, CDPREF='SURF', KNIVAU= 1, CDSUFF='PROP.VEGETAT', LDCOSP= F* *FACIL1 - ARTICLE-CHAMP "SURFPROP.VEGETAT" INEXISTANT, UNITE 54* “	-	SURFPROF.VEGETAT needs to be present to run config. ee927_surf_prep
AZO2	1	0	OK	OK	no differences found in the fields of init_surfex.fa
AZO3	1	1	OK	OK	no differences found in the fields of init_surfex.fa

As a conclusion, it can said that regardless of the value of SURFPROP.VEGETAT and SURFPROP.VEG.MAX (0 or 1) the fields of the output file of ee927_surf_prep (init_surfex.fa) are

unchanged, which suggests the information of both climatological fields is probably no longer used in configuration ee97_surf_prep. However, considering the abort message (corresponding to removing both fields inside the climatological file), SURFPROP.VEGETAT (at least) needs to be present to run the configuration.

Acknowledgements

To Alena Trojakova for the support in the installation of CY43T2 at ECMWF. To Florian Suzat for providing additional SAND and CLAY data and the useful discussions on these experiments.

4. Conclusions and outlook

At the end of this work, it is possible to say, that:

1. CY43T2 ported to all the Portuguese dynamical adaptation configurations at ECMWF; this was only possible due to
2. the update of the operational climatologies to CY43T2. Besides,
3. the long-term issue with the Azores domain climatologies was identified. A temporary solution was validated but a long-term solution for the ECOCLIMAP problem over the Azores region should be solved and will be report to the project.

As future work, it is foreseen:

1. the creation of a backup operational system at ECMWF, while the acquisition of a new local HPC acquisition will take place;
2. the systematic validation of the data assimilation solutions (surface by CANAI-OI_MAIN) and combined (by CANARI-OI_MAIN plus 3D-Var) to Iberia and Madeira, using as reference the local dynamical adaptation configuration of AROME..

5. References

[1] Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derková, M., El Khatib, R., Hamdi, R., Mašek, J., Pottier, P., Pristov, N., Seity, Y., Smolíková, P., Španiel, O., Tudor, M., Wang, Y., Wittmann, C., and Joly, A.: The ALADIN System and its canonical model configurations AROME CY41T1 and ALARO CY40T1, Geosci. Model Dev., 11, 257–281, https://doi.org/10.5194/gmd-11-257-2018, 2018.

[2] Monteiro, M., Aitmeziane, O., Belghrissi, H., Bogatchev, A., Cengiz, Y., Deckmyn, A., Dehmous, I., Hdidou, F., Khalfaoui, W., Kolonko, M., Rio, J., Sahlaoui, Z., Sezer, M., Szczęch-Gajewska, M., Tsenova, B.: The DAsKIT programme: status and plans, ALADIN-HIRLAM NL 15, June 2020.

Met Éireann's Physics Testing in HARMONIE-AROME CY43

Colm Clancy, Emily Gleeson, Ewa McAufield

1 Introduction

The operational NWP suite at Met Éireann is called the Irish Regional Ensemble Prediction System (IREPS). On the 23rd of March 2021, IREPS was upgraded to use cycle 43h2.1 of HARMONIE-AROME. Bessardon et al. (2021) provides more detail on this upgrade and the testing carried out ahead of this. In particular, it was found that a number of changes from the default code configuration were required to improve the forecast performance for Ireland. These are listed below in Table 1.

Table 1: Changes from default cycle 43h2.1 in the upgraded operational suite

Component	**Description**
Observations	CONV, MODES (winds only), AMSU-A, MHS, IASI, ATMS, MWHS2, ASCAT
Surface DA	XSODELX(0)=1.0, XSODELX(1)=2.0
Dynamics	Quadratic grid
Dynamics	LGWADV=T, LRDBBC=F (to control MSLP noise)
Surface	LFAKETREE=T
Surface	XRIMAX=0.0
Microphysics	Cloud droplet number concentration (CDCN)=50E6/m^3
EPS	EDA, multi-physics, scale_pert=yes, WG perturbations off Lagged 1+15 EPS: 1+10 at 0000/0600/1200/1800; 1+5 at 0300/0900/1500/2100 UTC

The Surface DA settings and choice of XRIMAX=0.0 were deemed necessary to improve the cold bias in 2 m temperature. However, it would be preferable to take a more physically-realistic approach. In addition, the new suite did not fully solve the ongoing problems with fog and low visibility forecasts particularly over sea areas.

This article summarises the results of over 100 physics-related experiments carried out to investigate more favourable settings, specifically aimed at temperature biases and fog forecasts. Cycle 43h2.1.1, which was officially tagged on the 4th of February 2021, was used for this work. However, the settings investigated were some of those that were chosen as defaults in cycle 43h2.2.

2 Seasonal testing

2.1 Test periods

The general testing of the physics settings was carried out for four two-week periods, roughly representing each season: May 8th–22nd 2020, June 1st–14th 2018, November 1st–14th 2019 and February 3rd–17th 2020, with the periods chosen as they included weather types of interest for case study analysis (e.g. strong convection, fog,

strong winds, etc). A five-day "warm-up period" was used in each case to spin up surface parameters, and tests were run on the IRELAND25 domain (540×500×L65, slightly smaller than the operational 1000×900).

Verification was carried out using observations from selections of various stations in the domain. For most variables, other than MSLP, we only consider Irish stations: not only because this is the region of interest, but also because of the strong influence of local factors. In addition to the approximate 20 synoptic stations in Ireland, the CAMP climate network of stations was used extensively. This network adds a further 60 or so observations of temperature, relative humidity and hourly rainfall, but data are not available before 2019. Finally, the network of approximately 400 volunteer rainfall stations were also used to verify forecasts of 24-hour rainfall accumulations.

2.2 Experiment details

Cycle 43h2.1.1 of HARMONIE-AROME was used for the testing. The default set-up, EXP0 in each case, matches the settings in the new operational suite; that is, includes the changes listed in Table 1. The following physics options were considered. Note that a few of these settings were only tested in relation to fog case studies (RSMIN and VSIGQSAT)

- Choice of XRIMAX, the maximum Richardson number
- ZTINER option: this refers to the Surface DA settings (Table 1) which had been found to be necessary in both cycles 40 and 43 to improve the 2 m temperature bias
- ECUME6: new sea-flux scheme
- LMIXUV: affects mixing of winds in the shallow convection scheme, default is TRUE
- LICERAD: change to cloud ice cover in the radiation scheme - making it consistent with microphysics, default is FALSE
- Cloud inhomogeneity factor: used to remove a fraction of the cloud condensate in the radiation calculations; default is 1.0 - no cloud inhomogeneity
- LTOTPREC: the evaporation/melting in the convection scheme is removed if LTOTPREC is true ; default is FALSE
- The Kristian Pagh Nielsen LW scheme (hereafter the KPN scheme): this refers to the KPN LW cloud liquid optical property scheme and consistent use of CDCNs in radiation as well as microphysics
- RFRMIN(10): control of precipitation release from the Kogan condensation scheme; default is 10
- RSMIN: minimum stomatal resistance for C3 crops and grassland - default scaling factor is 1.5
- VSIGQSAT: Saturation at lower levels; default is 0.02

As discussed in the introduction, the main aim of these tests was to find a configuration with a more physically-plausible approach to improving the model cold bias; specifically to find a non-zero XRIMAX and to remove the ZTINER settings. To this end the set of experiments detailed in Table 2 were carried out across the four test periods mentioned above. The exceptions to this were the cloud inhomogeneity and LTOTPREC experiments which were restricted to one season - the results of these are not included in this report.

Table 2: Experiment listing

Name	XRIMAX	ZTINER	Other settings
EXP0	0.0	Yes	default settings in Table 1
EXP1	0.2	No	
EXP2	0.2	No	ECUME6
EXP3	0.2	No	LMIXUV=FALSE
EXP4	0.2	No	LICERAD=TRUE
EXP5	0.2	No	LICERAD=TRUE, LMIXUV=FALSE
EXP6	0.2	No	LICERAD=TRUE, LMIXUV=FALSE, ECUME6
EXP7	0.0	No	LICERAD=TRUE, LMIXUV=FALSE, ECUME6
EXP0_NINHOM1	0.0	No	as EXP0 but with cloud inhomogeneity on
EXP5_LTOTPREC	0.2	No	as EXP5 but with LTOTPREC on
EXP6_LTOTPREC	0.2	No	as EXP6 but with LTOTPREC on
EXP8	0.1	No	LICERAD=TRUE
EXP9	0.1	No	LICERAD=TRUE, LMIXUV=FALSE
EXP10	0.2	No	LICERAD=TRUE, LMIXUV=FALSE, KPN scheme
EXP11	0.2	No	LICERAD=TRUE, LMIXUV=FALSE, KPN scheme, RFRMIN(10)=20

2.3 Initial testing of temperatures, winds and humidity

The first round of tests consisted of experiments 0-7 from Table 2. With so many experiments, it was necessary to identify promising candidates for further detailed analysis. Verification results showing the daily variation of 2 m temperature are displayed in Fig. 1. Looking at the night-time cold bias, particularly during the winter case (second from left), suggested excluding the configurations in experiments 1-3.

Of the remaining experiments 4-7, EXP7 appears to give the best improvement in night-time temperatures. Unfortunately, this experiment retains the zero value for the maximum Richardson number. The remaining experiments show a benefit for the day-time biases. This is mainly due to the use of LICERAD. The impact of LICERAD is that it increases the ice cloud cover while keeping the original ice cloud condensate. This results in a lower cloud water path and a higher cloud emissivity, which acts to increase the surface temperatures.

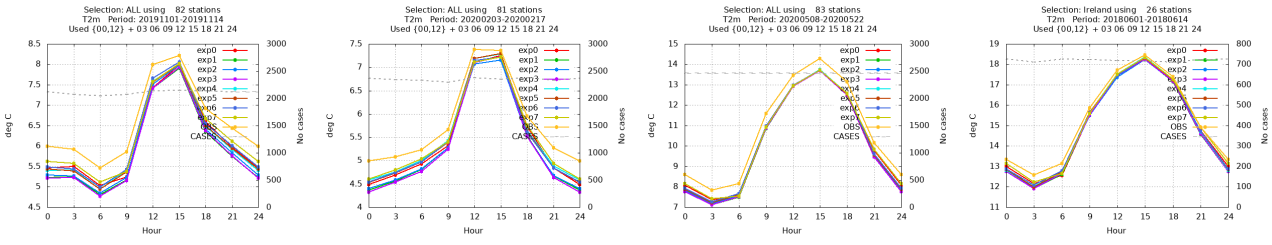


Figure 1: Point verification of daily variation of 2 m temperature using Irish synoptic and CAMP stations for the four test periods. Note that CAMP data are unavailable for the June 2018 period.

Verification of near-surface wind-speeds is presented in Fig. 2 for experiments 4-7. It can be seen immediately that EXP4 results in an increased positive bias, which is undesirable. The deviation of EXP7 from the default behaviour is most visible in the spring and summer cases, with a negative bias appearing. Finally, experiments 5 and 6 increase the positive bias slightly in Autumn and Winter. Note that there are positive biases in near-surface wind-speeds in Autumn and Winter in all of the experiments carried out; of the order of 0.5 to 0.7 m/s on average in Autumn, lower when XRIMAX=0.0, and exceeding 1.0 m/s in Winter. The biases for Spring and Summer are much lower - +/- 0.2 m/s.

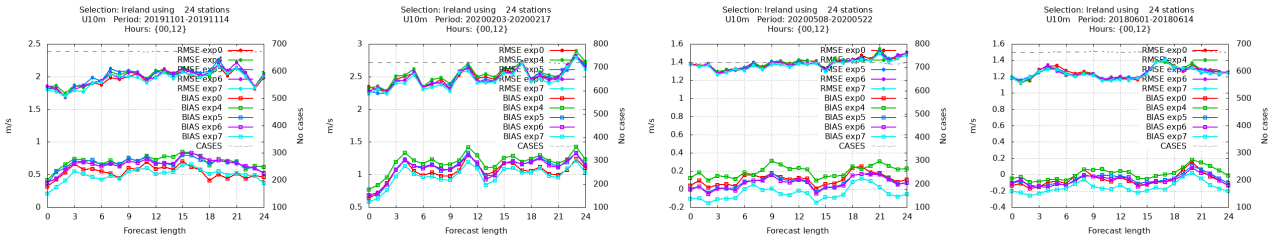


Figure 2: Point verification of 10 m wind-speed using Irish synoptic for the four test periods (left to right).

Although EXP7 showed decent results in these seasonal tests, the specific fog experiments (next section) showed a benefit to having a more realistic non-zero value of XRIMAX. Values of 0.1 were tested in experiments 8 and 9 (c.f Table 2). Experiments 8 and 9 otherwise correspond to 4 and 5, respectively. Given that 5 was superior to 4, 9 was considered more promising. It was hoped to find a compromise between the improved fog results with XRIMAX>0, and the degraded wind-speeds from XRIMAX=0.2. Unfortunately this wasn't to be: the wind verification (not shown) for 5 and 9 were almost identical. Nevertheless, this impact on the winds from XRIMAX>0 was deemed to be acceptable given the improvement in fog forecasts.

The effect of the ECUME6 sea-flux scheme can be seen in Fig. 3 below, where the verification of 2 m relative humidity is shown comparing EXP5 with EXP6, the only difference being that the latter (green curves) uses ECUME6. We see a general increase in humidity bias, albeit with some reduction in RMSE. Note that in general the humidity biases are higher in Spring and Summer - something that needs further attention over the coming months. This also ties in with the positive biases seen in day time dew point temperatures in Spring and Summer, consistent with too much evaporation, compared to negative biases in Autumn/Winter (Figure 4).

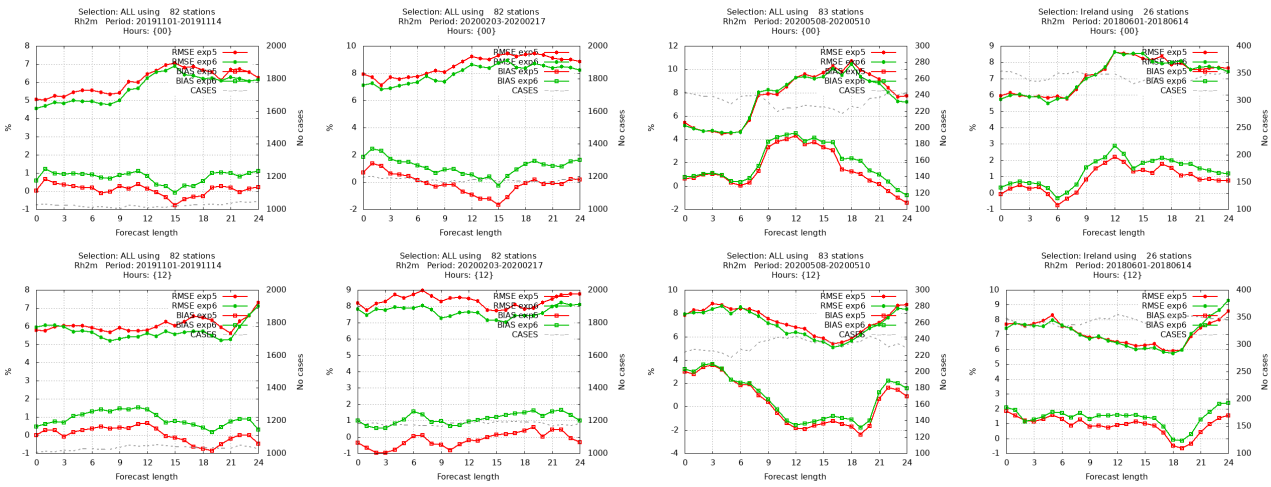


Figure 3: Point verification of 2 m relative humidity forecasts starting at 0000 UTC (above) and 1200 UTC (below) using Irish synoptic and CAMP stations for the four test periods. Note that CAMP data is unavailable for the June 2018.

Figure 5 shows the impact of ECUME6 on precipitation, with distributions of 24-hour rainfall for all four periods. Differences are most noticeable in the autumn and winter cases. This is emphasised by the spatial plots of bias in Fig. 6, with the bias becoming more positive when the ECUME6 scheme is used. Further work is required regarding the analysis of ECUME6.

Figure 7 highlights some of the outstanding issues that we have with HARMONIE-AROME cycle 43 in general. Figure 7 left indicates the degradation in bias in nighttime winds when positive values of XRIMAX are used

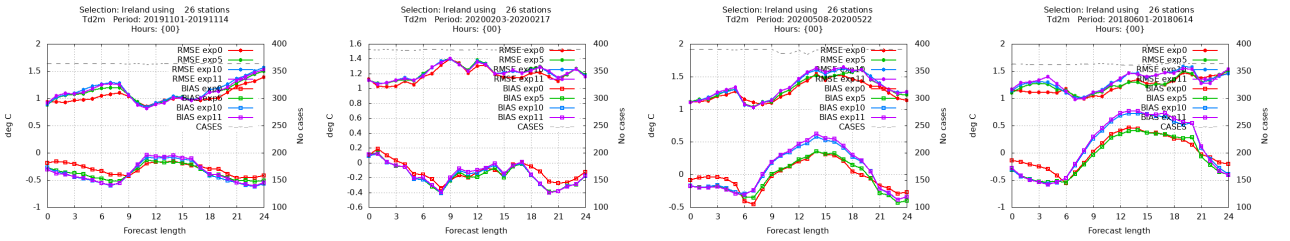


Figure 4: Point Verification of 2 m dewpoint temperature for the four test periods.

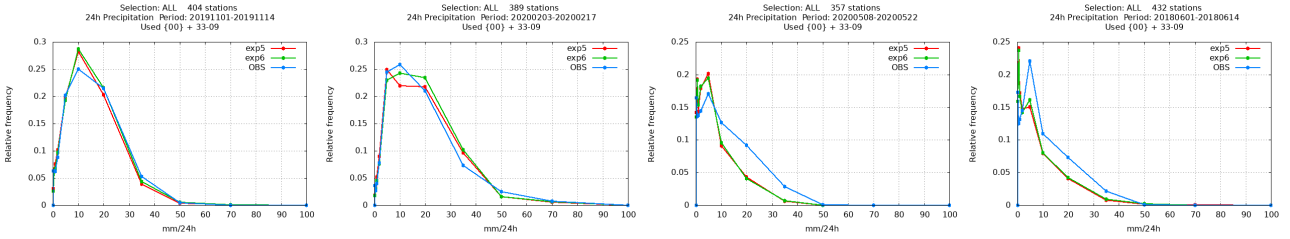


Figure 5: Frequency distributions for 24-hour rainfall accumulation for the four test periods.

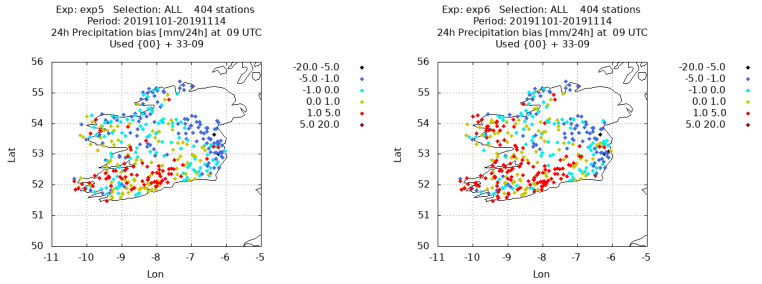


Figure 6: Verification of 24 hour precipitation for the Autumn period against our rainfall observation network of approximately 400 stations for (left) EXP5 and (right) EXP6. The only difference between the experiments is that EXP6 was run with the ECUME6 sea flux scheme.

(all bar red curve use positive XRIMAX). Without positive XRIMAX, the modelling of fog is vastly degraded. Since fog is a high impact weather event, and costly for the aviation sector in particular, XRIMAX=0.2 will be used in our upcoming upgrade of cycle 43 at the small expense of nighttime winds. Regarding T2m, the second and third panels of Fig. 7 show the slight degradation in bias in nighttime near surface temperatures when non-zero XRIMAX is used. In addition, the KPN changeset causes a positive bias in daytime temperatures in Spring and Summer (not shown) but improves the bias for Autumn and Winter (not shown). The positive bias is due to changes in the optical properties of the clouds - a negative bias in cloud cover is also seen (right-hand panel of Fig. 7) and is consistent with the results shown in the clear sky index plots in Section 2.4.

2.4 Clear Sky Index

The clear sky index (CSI) is the ratio between global shortwave irradiance and clear sky global shortwave irradiance. It is a very useful metric for not only evaluating global radiation but also clouds and in particular cloud condensate. Low values of CSI are associated with thick frontal cloud whereas high values are representative of clearer weather or cases with thin ice clouds. Gleeson et al. (2015) gives further information and references

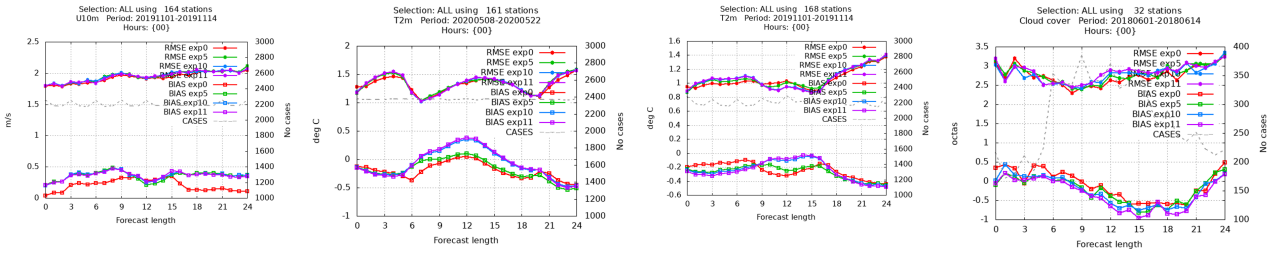


Figure 7: Verification of (from left) U10 Autumn, T2m Spring, T2m Autumn and cloud cover Summer. All synoptic stations in the domain are used.

on the CSI used here. Over the years, HARMONIE-AROME has always had a positive bias in CSI at low CSI values, linked to the fact that there has been too much cloud condensate in the thickest clouds. This was also shown when frontal cases were compared to the cloud water path (CWP) in the KNMI MSG SEVIRI cloud product. The reduction in cloud droplet number concentration (CDCN) from default values to a uniform value of 50E6 over land and sea, resulted in improved cloud condensate in frontal clouds.

Here we show sample results for each season for experiments 0, 4, 5 and 10 in Fig. 8. With the exception of EXP0, all experiments were run with LICERAD=TRUE. The impact of this setting on high CSIs can be seen, with better agreement between the EXPs and observations at high CSI. Note that these plots were created using modelled and observed hourly mean global shortwave irradiance at 20 sites around Ireland. Another interesting feature is the fact that the use of the KPN scheme (LW and SW changes) improves the histograms at very low CSI compared to observations. There are newer versions of the KPN scheme and changes to the profile of CDCNs have not yet been tested for the Irish domain (see Section of fog for further information). Thirdly, for the summer period, the use of the KPN scheme (pink curve) highlights a negative bias in cloud in the CSI 0.2 to 0.4 range, consistent with the cloud and temperature biases shown in Figure 7.

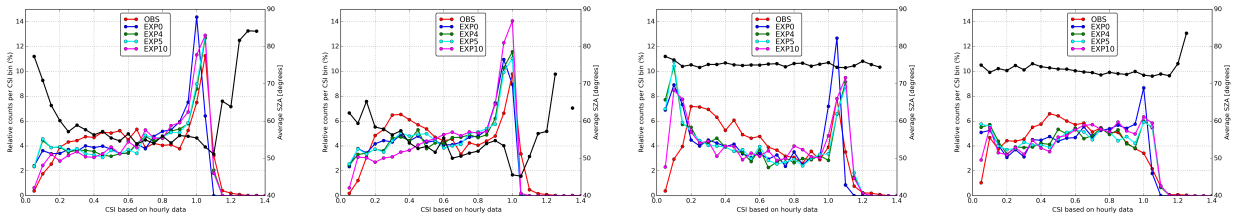


Figure 8: Histograms of CSI for (left to right) Spring, Summer, Autumn, Winter for experiments 0,4,5 and 10. Black curves show the average solar zenith angle in degrees.

2.5 Final proposed configuration

Following the analysis described above, the potential "optimal" configurations with non-zero XRIMAX seemed to be given by EXP5 or EXP9, with very little difference seen in verification between the two (not shown). The higher value of XRIMAX is preferable from a fog-forecasting point of view, and so we now focus on this as our proposed configuration for cycle 43h2.1.1, and compare with the default set-up matching our current operational cycle 43h2.1 IREPS suite (Table 1). The changes from this default EXP0 to the proposed EXP5 are highlighted again in Table 3 below. A detailed comparison is given for surface variables in Figs. 9 and 10.

As stated at the outset, one of the main aims of this work were to remove the non-physical ZTINER option and zero maximum Richardson number. The 2 m temperature verification at the top of Fig. 9 suggests that we can do this with our proposed configuration, without any significant degradation in terms of bias. Results in Fig. 9

Table 3: Differences between default operational and proposed configuration.

Name	XRIMAX	ZTINER	LICERAD	LMIXUV
Default EXP0	0.0	Yes	FALSE	TRUE
Proposed EXP5	0.2	No	TRUE	FALSE

for 2 m relative humidity show a general improvement, while those for dewpoint temperature are mixed. The bias in MSLP (bottom Fig. 9) is slightly reduced in three of the four test periods.

Near-surface winds are verified in Fig. 10. As before, the synoptic stations used in the analysis are also categorised into inland and coastal, and the marine buoy network is also used. The proposed EXP5 shows a larger positive bias in winter and autumn, which was discussed already and deemed a necessary trade-off. This bias seems to affect the inland stations rather than the coastal ones. Furthermore, the results from the marine buoys (bottom row of Fig.10) suggest a reduced bias. These should be interpreted with caution, however, due to the small number of sites, and the scaling required to adjust their values to 10 m.

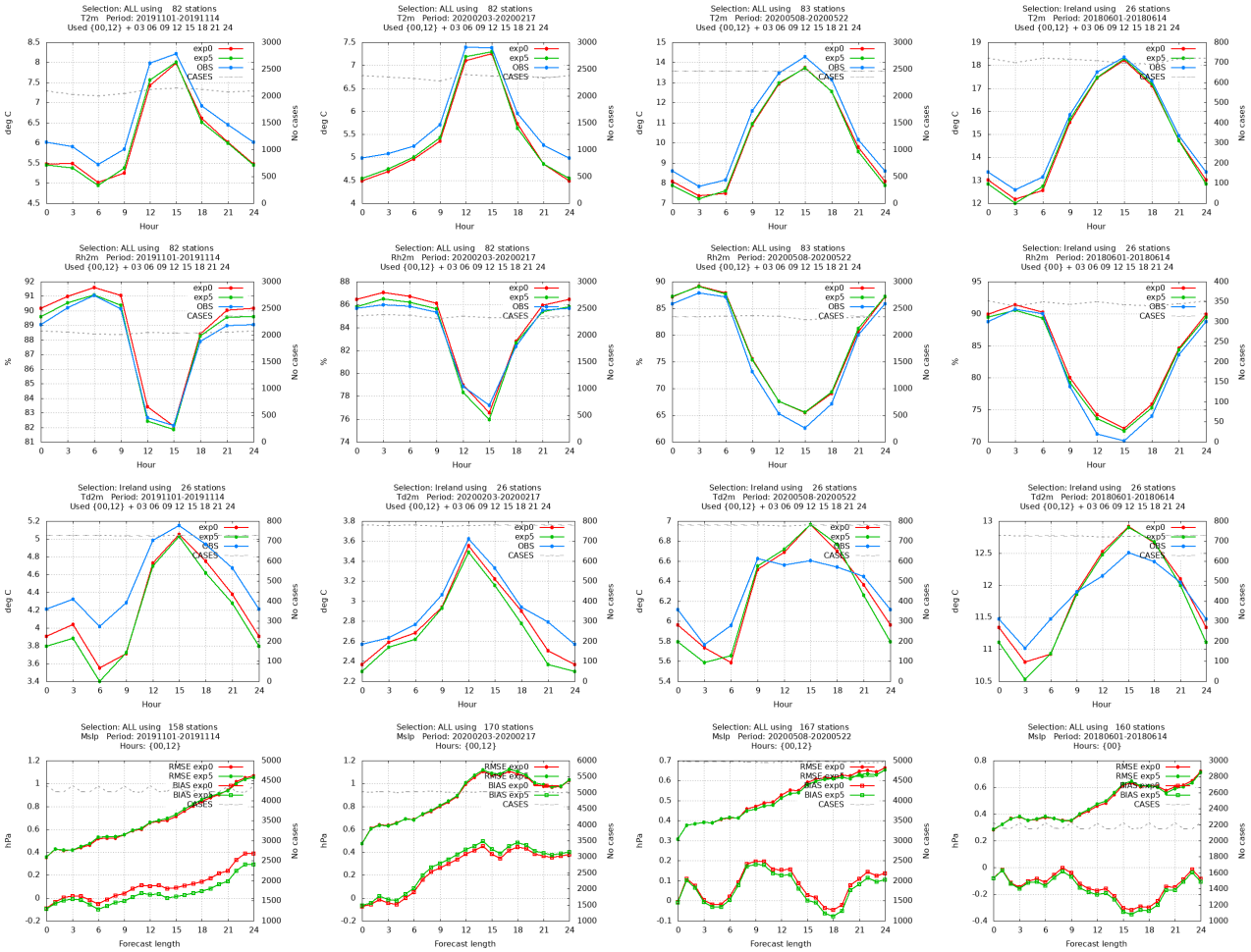


Figure 9: Point verification of surface parameters using Irish synoptic and CAMP stations for the four test periods (left to right). Note that CAMP data is unavailable for the June 2018. Parameters, from top to bottom: 2 m temperature daily variation, 2 m relative humidity daily variation, 2 m dewpoint temperature, MSLP.

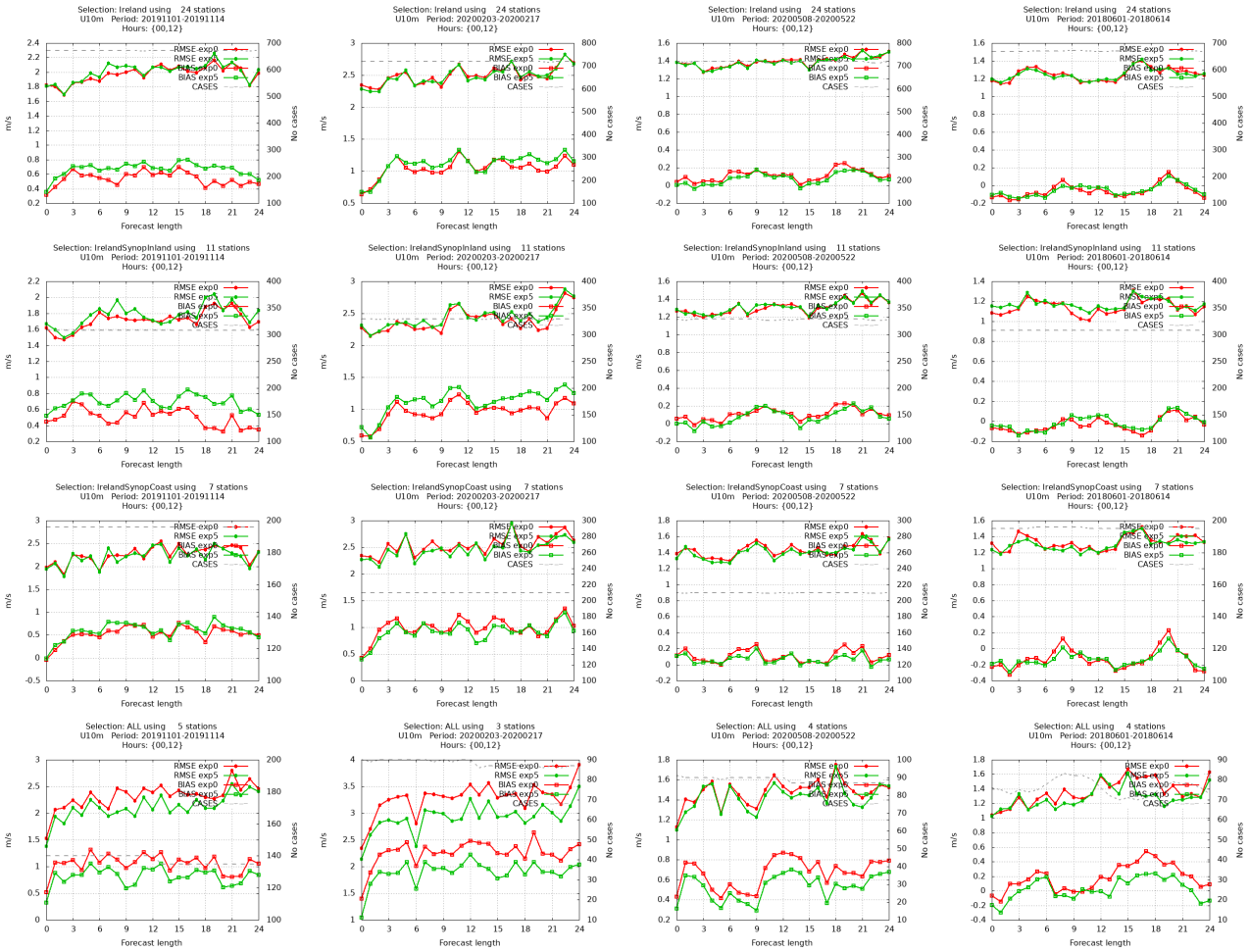


Figure 10: Point verification of 10 m wind-speed using different sets of stations for the four test periods. From top to bottom: all Irish stations, inland synoptic stations, coastal stations, marine buoys.

3 Fog experiments

3.1 Experiment Details

While there were some fog cases within the 4 x 2-week seasonal tests that we ran (details in Table 2), we also ran a selection of shorter runs covering fog cases of interest. Five days of spin-up preceded each fog period. The cases included were: 29/11/2020-07/12/2020, 28/02/2021-04/03/2021, 03/07/2020-11/07/2020, 26/02/2019, 17/03/2021 and 12/04/2021-14/04/2021. Details of the experiments are given below in Table 4.

3.2 Sample Fog Results

This section includes results from the March 17th 2021 fog case over land and the April 14th 2021 sea fog which was very much over-predicted by the operational HARMONIE-AROME cy43 at Met Éireann. Results from the other cases have not been included here but overall XRIMAX, the KPN changeset and RFRMIN(10) have an impact in many of the cases. Although the KPN changeset now need editing due to temperature/cloud

Table 4: Experiment listing for fog cases.

Name	XRIMAX	ZTINER	Other settings
EXPF0	0.0	No	default settings in Table 1
EXPF1	0.1	No	default settings in Table 1
EXPF2	0.2	No	default settings in Table 1
EXPF3	0.3	No	default settings in Table 1
EXPF4	0.0	No	RSMIN=2.0 (for C3 crops and grass)
EXPF5	0.0	No	RFRMIN(10)=20
EXPF6	0.0	No	VSIGQSAT=0.03
EXPF7	0.0	No	KPN LW and SW changes
EXPF8	0.2	No	KPN LW and SW changes
EXPF9	0.2	No	KPN LW and SW changes and RFRMIN(10)=20

biases, the results shown here are for the original KPN changeset (new LW cloud liquid optical properties and consistent use of 50E6 for CDCN at all levels in microphysics and radiation).

3.2.1 March 17th 2021 Fog Case

Figure 11 shows the visibility forecast for the control member of IREPS (the Irish Regional Ensemble Prediction System) at 06 Z on March 17th 2021 from the 00 Z run for (left) CY40 and (right) CY43. The main difference that has led to the improvement in cycle 43 (though fog is still very much over-predicted compared to the observations in Figure 11 bottom) is the change in CDCNs from 300/100 per cm^3 over land/sea in cycle 40 to 50 per cm^3 everywhere in cycle 43. Figure 12 shows the results of experiments done using configurations 0, 1, 5, 7, 8 and 9, also from the 00 Z run and valid at 06 Z on March 17th 2021. The impact of using a positive XRIMAX can clearly be seen (compare 0 to 1) as can the impact of using the KPN changeset (compare 0 to 7). In this case RFRMIN(10) had little effect. These effects can also be seen in the histograms in Figure 13. Note that in the upper figure EXPF0 and EXPF4 overlap as RSMIN showed no impact but the effect of increasing XRIMAX is clear. Similarly, in the lower figure the impact of the KPN changes is evident with an additional reduction in the lowest visibilities when a positive XRIMAX is used with the KPN changes. Note that only land points were used in the calculation of these histograms.

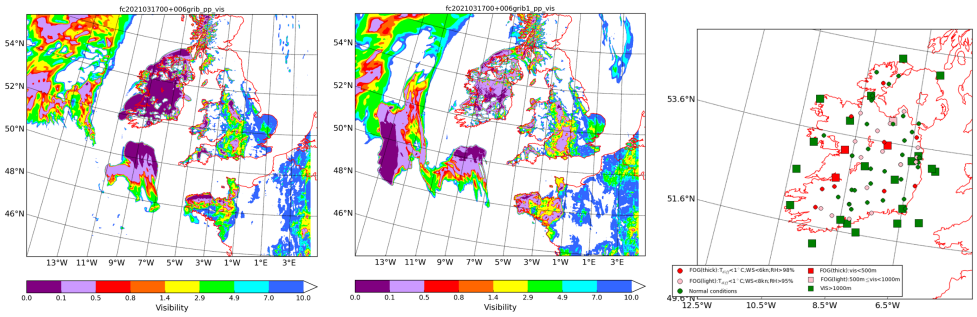


Figure 11: 06 Z visibility forecast for CY40 (left), CY43 (middle), observations (right).

3.2.2 April 14th 2021 Sea Fog Case

The second fog case included is from April of this year when HARMONIE-AROME erroneously forecast widespread fog in the seas around Ireland and also dense fog over land. Figure 14 shows the 06 Z forecast visibilities from the 00 Z run. Note that the ensembles cycle through values of 0, 0.1 and 0.2 for XRIMAX - in

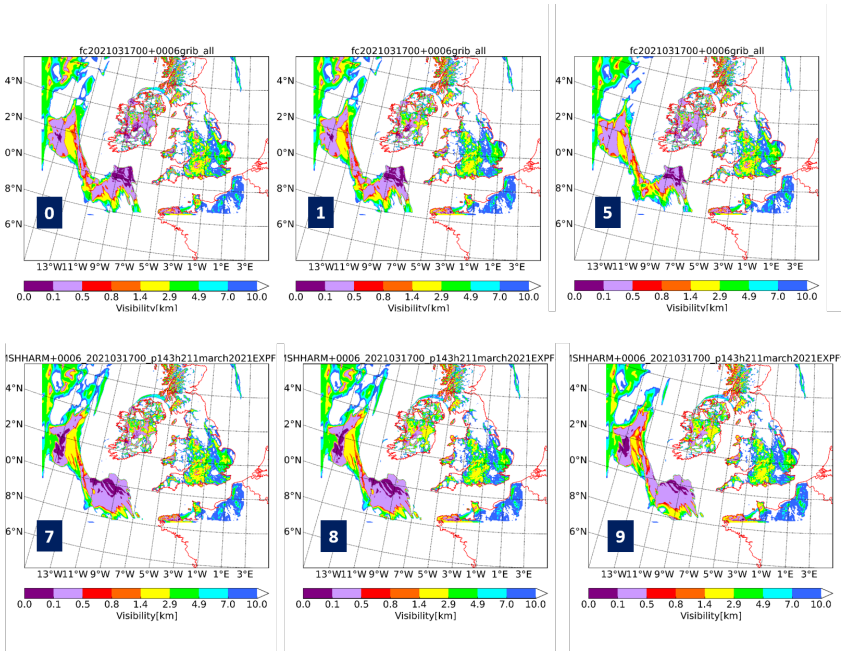


Figure 12: 00 Z + 6 hour forecast of visibility for: EXPF0, EXPF1, EXPF5, EXPF7, EXPF8, EXPF9.

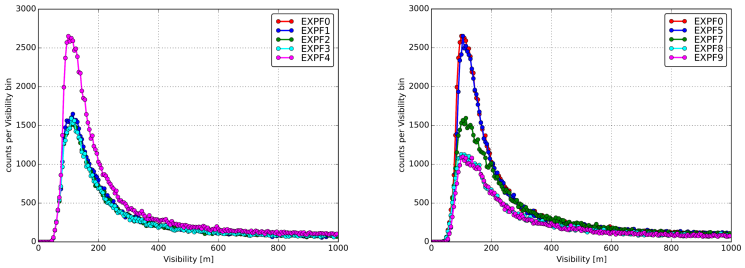


Figure 13: Histograms of visibility from the 00 Z run on March 17th 2021 using 1-24 hour forecasts over the day and focussing on on the lowest visibility range of the spectrum.

some cases we find that the members with positive XRIMAX perform better in terms of fog. You can see that several of the members include fog forming over the sea - much of this was initiated over the land however. For comparison the 06 Z IR MSG satellite image is shown in Fig. 15 (left) and the corresponding observations in Fig. 15 (right). While there are patches of fog around, it is not as widespread as in operational HARMONIE-AROME. 06 Z visibility forecasts from 00 Z runs of various configurations of CY43h2.1.1. are shown in Fig. 16 where again the impact of XRIMAX and the KPN changeset are evident. In this case RFRMIN(10) degrades visibility over the sea whereas in other cases (not included here) this setting improves the fog situation.

Similar plots but valid at 12 Z on April 14th are shown in Figures 17 and 18. The overestimation of fog at sea can clearly be seen in IREPS - there is some fog to the northeast of Ireland but most sea areas are fog free - backed up by buoy observations. A positive XRIMAX and the KPN changeset really help to improve the erroneous sea fog. As for the March 17th case, histograms of visibility forecasts over April 14th 2021 show the impact of each of the physics changes very clearly (Fig. 19).

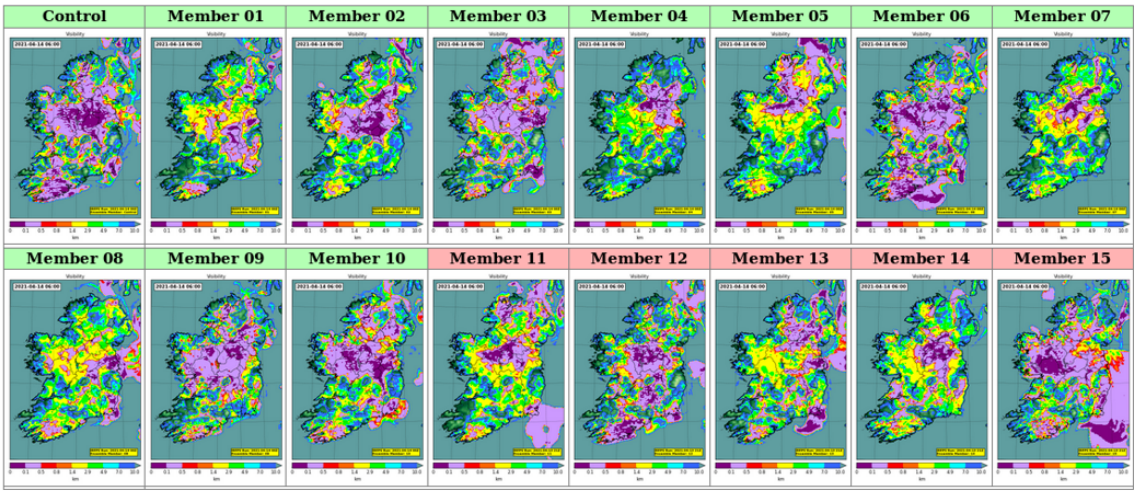


Figure 14: IREPS 06Z forecast of visibility from the 00 Z run on April 14th 2021.

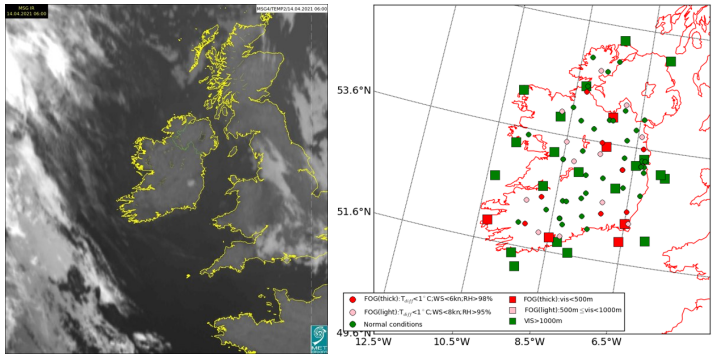


Figure 15: MSG IR satellite image (left) for 06 Z on April 14th 2021 and 06 Z observations (right).

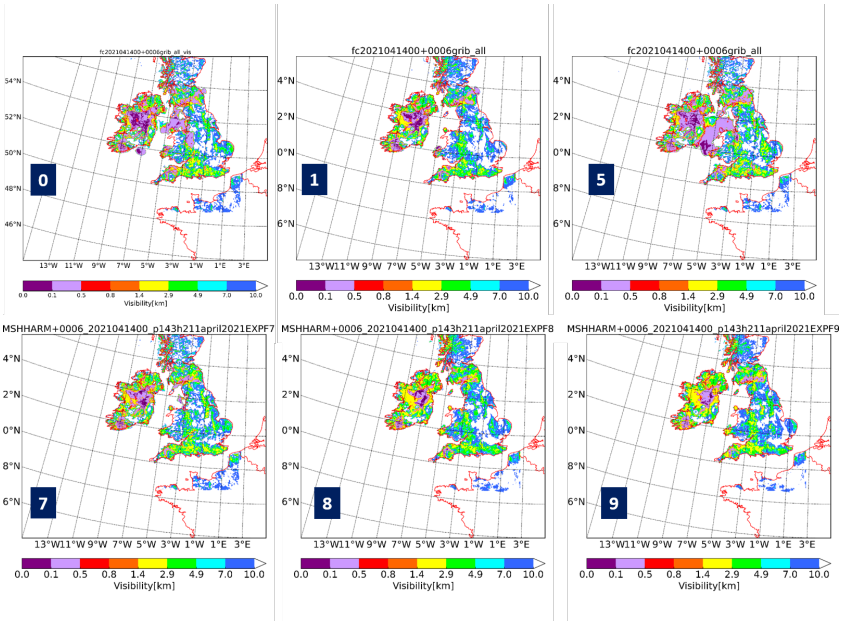


Figure 16: 00 Z + 6 hour forecast of visibility for: EXPF0, EXPF1, EXPF5, EXPF7, EXPF8, EXPF9.

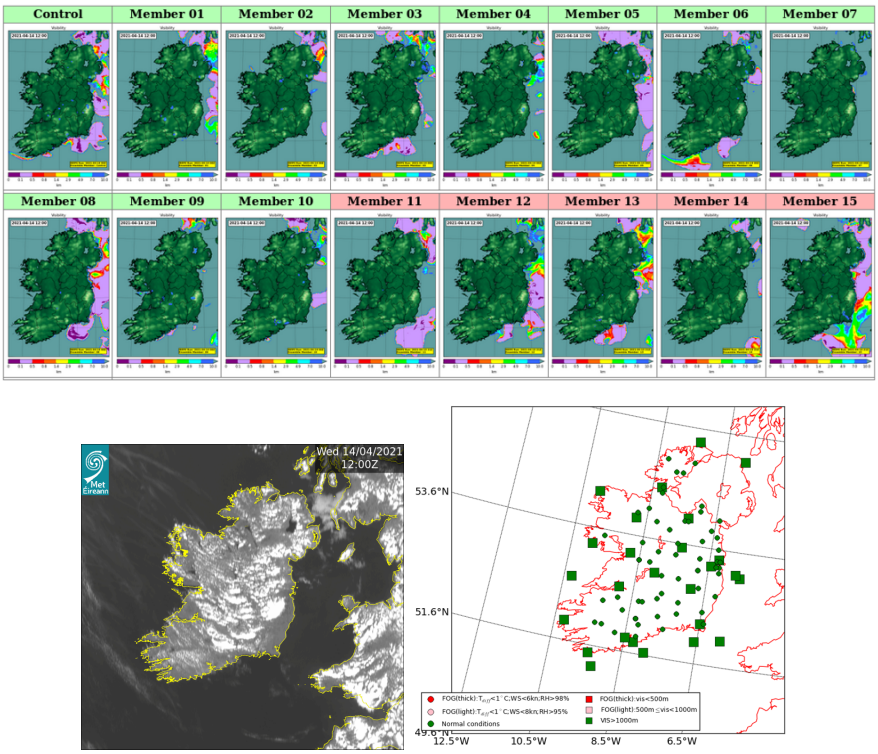


Figure 17: Valid at 12 Z on April 14th 2021. (top) IREPS ensemble members visibility forecasts. (bottom left) MSG VIS satellite image (bottom right) visibility observations.

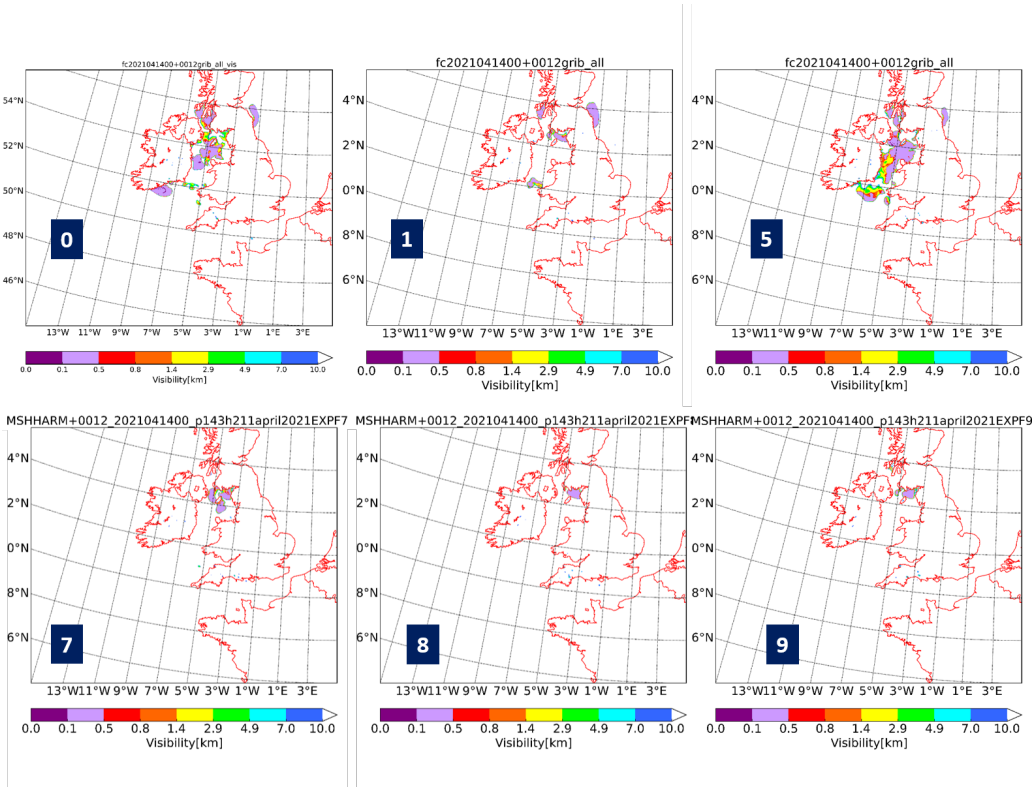


Figure 18: 00 Z + 12 hour forecast of visibility for: EXPF0, EXPF1, EXPF5, EXPF7, EXPF8, EXPF9.

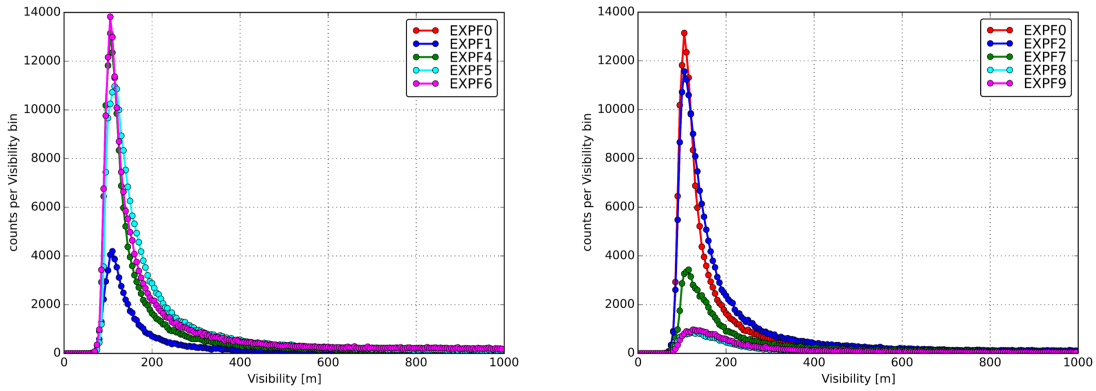


Figure 19: Histograms of visibility from the 00 Z run on April 14th 2021 using 1-24 hour forecasts over the day and focussing on the lowest visibility range of the spectrum.

4 Summary

The last few tests that we ran highlighted an issue with the KPN changeset, also seen in MetCoOp results. The consistent use of CDCNs in both microphysics and radiation seems to have led to a positive bias in T2m and a negative bias in clouds. This was not seen when CDCNs of 50E6 were only used in microphysics. In addition, the use of the KPN LW cloud liquid optical properties was not the reason behind the biases. The use of lower CDCNs in the radiation scheme (the same value as in microphysics) impacts on the cloud droplet effective radii and impacts on both the LW and SW cloud liquid optical properties. Further testing is currently underway.

The Irish domain has positive biases in dew point temperatures and humidity in Spring and Summer which exacerbate our issues with fog. The biases in dew point temperatures indicates an issue with too much evaporation. Further tuning of the minimal stomatal resistance will be tested in the coming weeks. While the issues with evaporation and humidity remain, we will not be in a position to use the ECUME6 sea flux scheme, which is a default in cycle 43h2.2.

We are currently testing all seasons with cycle 43h2.2 target 2, where CDCNs increase with height. In addition a range of fog cases are also being tested with this latest version to ensure that the new CDCN profile has no unexpected adverse effects.

For the past few releases of HARMONIE-AROME we have tested the same time periods for each season so that we can look back and gauge the progress over the past few years and have our own benchmarks to compare against. In this phase of testing, we have tested a range of physics parameters one by one in order to see the influence of each separately before combining a group of changes. As fog as been a large issue for the past few years, we also have a suite of fog cases that we will now test for future releases also.

5 Acknowledgements

Computing resources for this work was provided by Special Projects at ECMWF.

6 References

Bessardon, G., C. Clancy, C. Daly, R. Darcy, J. Fannon, E. Gleeson, A. Hally, E. Harney, E. McAufield and E. Whelan, Met Éireann Updates, ALADIN-HIRLAM Newsletter 16, 118–130, February 2021

Gleeson, E., K. P. Nielsen, V. Toll, L. Rontu and E. Whelan, Shortwave Radiation Experiments in HARMONIE. Tests of the cloud inhomogeneity factor and a new cloud liquid optical property scheme compared to observations, ALADIN-HIRLAM Newsletter 5, 92–106, August 2015

AROME-Based Climatology of Fog and Low Clouds over Morocco

Driss BARI, Nabila LASRI, Rania SOURI and Redouane LGUENSAT

1 Introduction

Poor visibility associated with fog and/or low clouds affects many socioeconomic sectors such as aviation, marine, and road transportation. This is why the fog phenomenon has become a scientific topic of interest [9, 10, 11, 12, 13]. In recent years, with the rapid development of socio-economic sectors, fog detection and forecasting have become more and more important; thus, the possibility of obtaining accurate hazardous weather warnings at least one or few hours in advance could help in the planning and mobilization of the responsible agencies and may contribute to reducing losses, mitigating damage and saving lives.

Different techniques have been proposed for short-term weather forecasting [1, 2, 3, 4, 5, 6]. Traditional techniques used for nowcasting are highly parametric, hence complex, and these methods require long computation times when applied to large areas, which, added to the data reception time, often makes the first forecasts useless. Recently, there has been a shift towards the use of artificial intelligence techniques for weather nowcasting [1, 7, 8]. These include the use of standard machine learning techniques such as artificial neural networks.

Deep Learning techniques, while they were only developed in the recent years, show impressive results in the field of image processing such as image classification, object detection or image segmentation, and they require very short computation times for inference tasks.

Hence, following the recent successes of deep learning techniques, the Moroccan General Directorate of Meteorology has undergone a project, to leverage the power of state-of-the-art image to image translation techniques in the field of remote sensing. This project is dedicated to explore the applicability of Convolutional Long Short-Term Memory architecture to obtain timely fog/ low clouds nowcasting and accurate position's estimation using satellite imageries from MSG satellites.

A preliminary step consists of an exploratory data analysis of the outputs of the operational numerical weather prediction AROME model as well as the observed data from the synoptic network of the General Directorate of Meteorology in order to delimit the areas most affected by the studied phenomena as well as its temporal distribution over the region during the day. In this step, an investigation of the possible physical mechanisms leading to the occurrence of this phenomenon is performed.

2 Data and Methods

First of all, the final database has to be established. For our purposes we have used two data-sets, covering 5-year period (2016-2020) over Morocco (the study domain : from 20.5 N to 37.0 N by 0.025 degrees for latitudes and from -17.5 W to -0.25 W by 0.025 degrees for longitudes with a total of 661 x 691 = 456751 grid points.) :

1. hourly forecasts, from 0000UTC run, of the operational NWP model AROME of Liquid Water Content (LWC) at the lowest level of the model (about 5m) as indicator of fog occurrence at each grid point. The low cloudiness forecasts are used to detect the low clouds over the study domain. These forecasts are issued from AROME model cycle 41t1, with the configuration described in Table 1.

2. hourly observations from 41 synoptic meteorological stations (CPM = Provincial Meteorological Center) of visibility and low cloudiness are used to compare qualitatively the spatial and temporal performance of the operational NWP model.

Table 1: Configuration of the operational NWP AROME model used in this study

Lateral boundary conditions coupling	ALADIN 36t1 (10km) since 01/01/2016 and ARPEGE (7.5km) since 15/11/2019
Coupling frequency	1 hour
time step	60 seconds
Vertical levels	90
Surface coupling	Surfex

3 Results

For each phenomenon, we first analyze the forecasts provided by the model to derive the climatology over the study area, and then we compare the main findings with the climatology issued from the actual observations.

3.1 Occurrence of fog only

On Fig. 1, we have plotted The probability of occurrence based on AROME outputs at 0100UTC (midnight), 0600UTC (early morning), 1500UTC (daytime) and 2100UTC (early night). From this figure, it appears that a **high probability of the occurrence of fog is noticed in the coastal cities from the side of the Atlantic Ocean in the north part**. A low-frequency area of fog occurrence is observed on the inland cities faraway from the sea and off-shore over the ocean.

[21] classified fogs formed over coastal areas as either coastal fogs or sea fogs, on the basis of the extent of the area affected. Coastal fogs are formed directly over the coastal land areas and generally extend only to the coastal region, while sea fogs develop over the sea and might extend or move into the coastal area. It should be noted that our understanding of the physical processes involved in fog formation is incomplete due to the great sensitivity of fog to environmental factors, and to the diversity of significant topographical features: local topography, land–sea heterogeneity, environmental conditions (urban, suburban, and rural), and coast shape [14].

Regarding **the temporal evolution**, we noticed that **the probability of fog occurrence decreases from north to south, especially during the daytime**. In addition, Regarding the diurnal and seasonal distribution of the fog events, **a maximum of their occurrence is classically observed during nighttime, particularly in winter**. This result is inline with the findings of [22] over the Grand Casablanca region. This behavior could be attributed to the radiative fog that dissipates under the action of solar radiation during the day or to the advection–radiation fog which results from the radiative cooling of humid air advected on land from the sea.

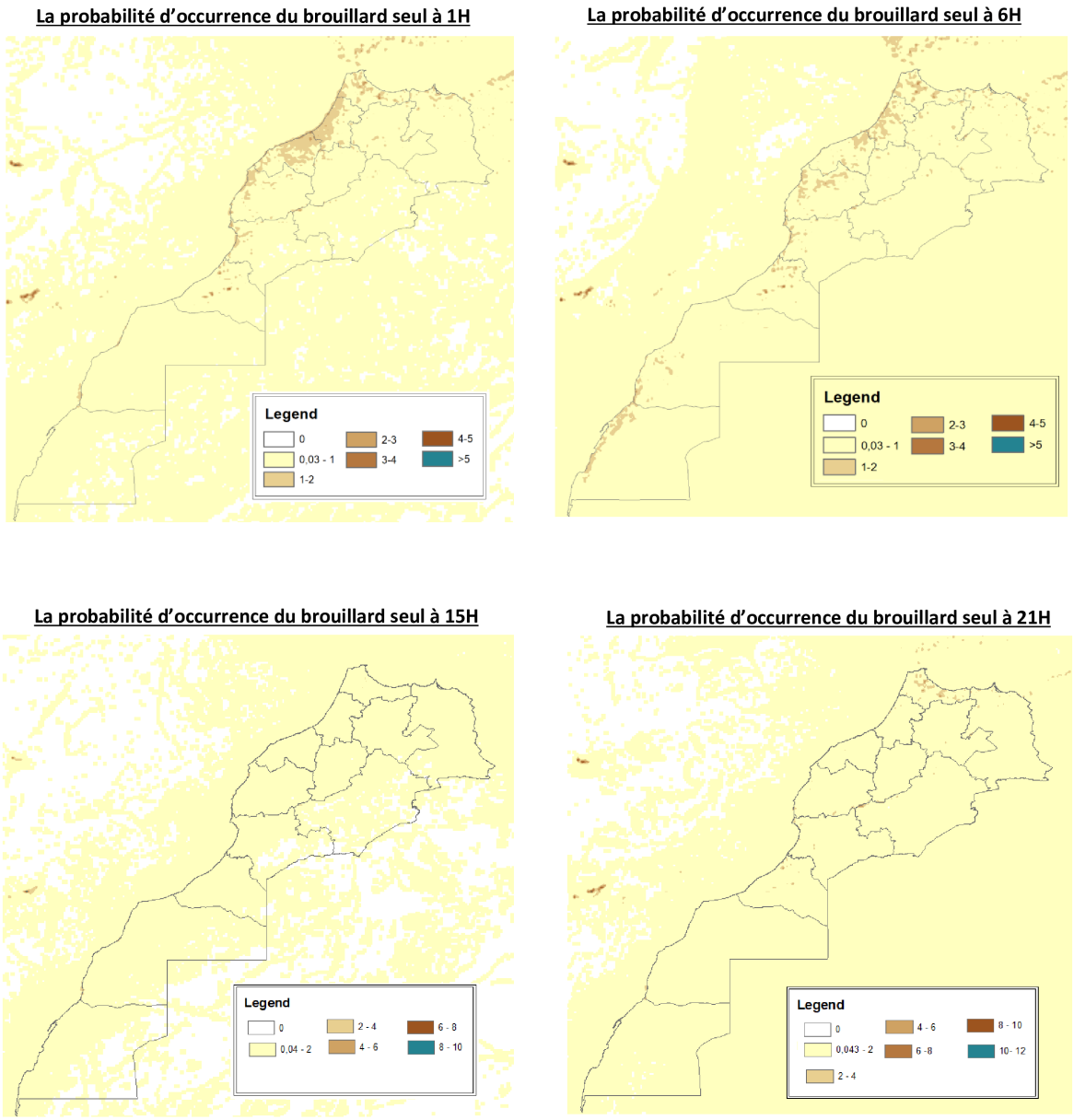


Figure 1: The probability of occurrence of fog only based on AROME forecasts at: a) 0100 UTC, b) 0600 UTC, c) 1500 UTC, and d) 2100 UTC.

In Fig. 2, we have plotted the observed frequency of fog only at the 41 synoptic meteorological stations based on the available data at each station over the studied period (2016-2020). It is seen clearly that the fog is frequent in the coastal cities either on the side of the Atlantic Ocean or the Mediterranean Ocean.

From midnight to 0600UTC, we notice an increase of the probability of fog occurrence in the coastal cities on the northern side of the Atlantic Ocean with peaks that are recorded by the stations of Rabat, Ifrane, Casablanca, El Jadida and Agadir. For the inland cities, they registered a very low probability of occurrence.

During the day and precisely until 1700UTC, the fog dissipates in most coastal cities and still rare at : Casablanca-El Jadida-Safi, Agadir-Essaouira and Sidi-Ifni.The station of Ifrane still records a high probability of occurrence of fog along the whole day with high values during the night and lower during the day. In fact, this is a mountainous station and the most frequent type of fog in this location is the up-slope fog related to the low clouds

over the region.

As a consequence, there is an agreement between the observed frequencies of the fog occurrence and the probabilities of fog occurrence in AROME forecasts based on the liquid water content at the lowest level of the model (about 5m). However, some discrepancies are observed in the Mediterranean and mountainous sites.

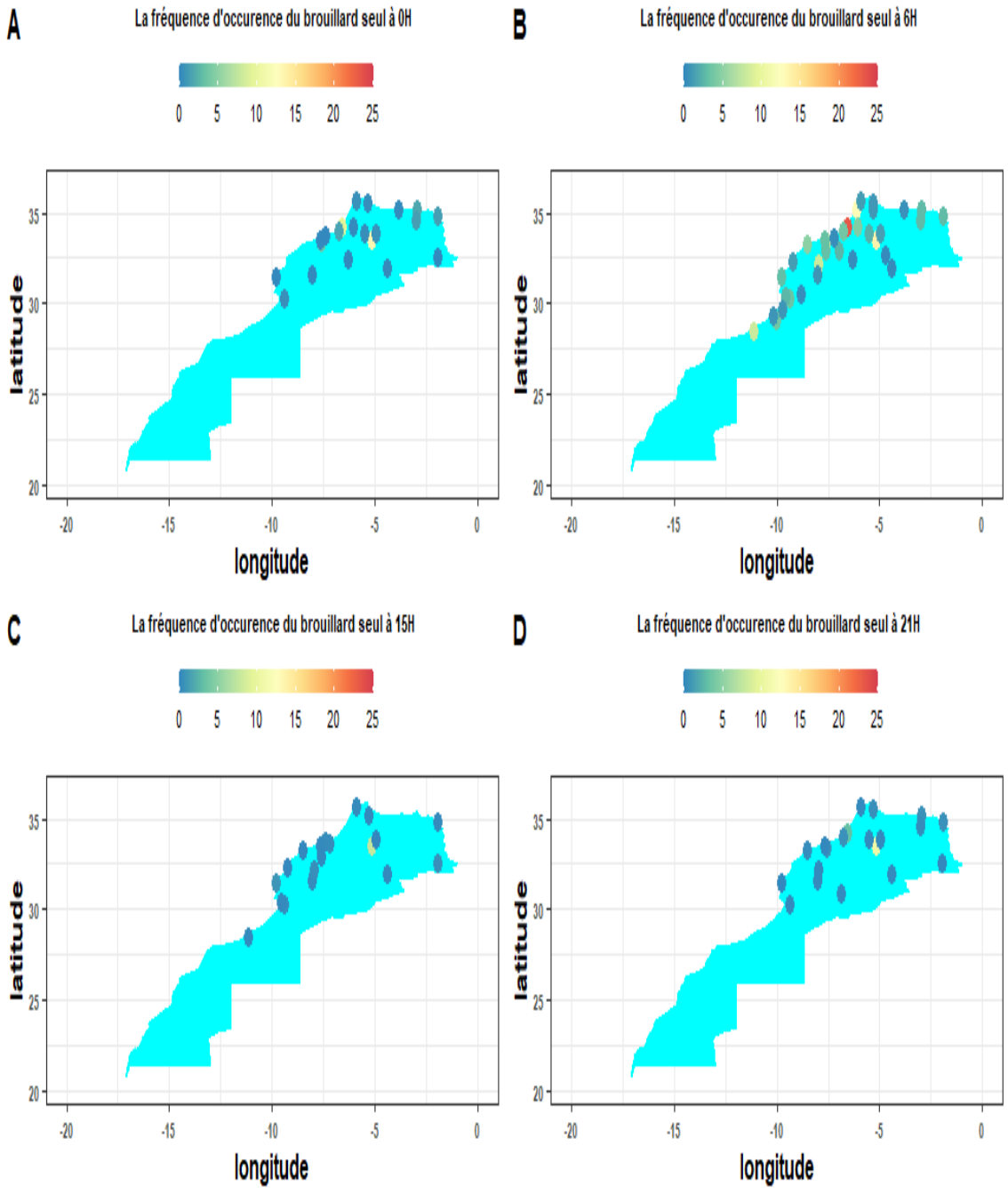


Figure 2: The observed frequency of fog occurrence based on observations at 41 synoptic meteorological stations over the study domain at a) 0100 UTC, b) 0600 UTC, c) 1500 UTC, and d) 2100 UTC.

3.2 Occurrence of low clouds only

Plotting low clouds probabilities derived from AROME model over our the study area (Fig. 3) reveals many features with respect to low clouds formation and their probability of occurrence based on the assumption of a perfect model.

The first finding concerns low clouds probability occurrence over ocean compared to surface. Obviously, **this probability is higher over the ocean which was mainly between 25% and 40% during the first half of the night (until 0300 UTC) and up to 80% during the second half of the night**. This maritime-continental contrasting characteristics is also confirmed in the literature e.g [15]; it is driven by a combined effect of topography, continental features and aerosols characteristics. In addition, Morocco is generally characterized by a cold and moist northerly to northwesterly wind associated to the atmospheric perturbations coming from the American continent and also from the north of the Atlantic Ocean, resulting in the formation of a stratiform low-level clouds over the ocean that extend inland over Morocco.

The second finding, is that **low clouds probability occurrence decreases from west to east**. In fact, the processes of stratus cloud (low-level liquid cloud) formation over the ocean are complex and driven by coupled parameters including heat flux (sensible and latent) at the sea surface, cloud radiative cooling/heating in the top and bottom layers, sea surface temperature (SST), and aerosol characteristics. These parameters are closely correlated [15]. In addition, [16] have studied factors controlling stratocumulus cloud lifetime over coastal land. The authors demonstrated that in contrast with stratocumulus over ocean, strong sensible heat flux over land prevents the cloud layer from decoupling during day. As the cloud thins during day, turbulence generated by surface flux becomes larger than turbulence generated by longwave cooling across the cloud layer.

The third finding, is **an important gradient in the probability of low clouds occurrence observed near the coast where this probability ranges from 10% to 30 %**. This could be explained by a differential heating between land and sea; the formation of stratus denotes the equilibrium point between two opposing processes; net condensation caused by cooling over the sea and net evaporation caused by heating (and the presence of drier air) over the land. This phenomenon is known as **the Sea-Breeze**, which is a circulation that develops due to differential heating of air over land and sea. As the sun heats the boundary layer over land, the resulting pressure gradient causes the movement of low-level air from the sea to land (Sea-Breeze) with a return flow aloft (return current). It should be noted that the existence and intensity of the Sea-Breeze depends strongly on seasonal and latitudinal factors as well as on the time of day. In fact, on many tropical and subtropical coastal regions the Sea-Breeze is a regular phenomenon throughout the year, while in cooler regions the Sea-Breeze is a common feature during spring and summer, when the temperature difference between land and sea is at its maximum.

On the other hand, warm moist air from the seaward side of **the upwelling** zone could be transported across the latter. There, low-level clouds form, which are then transported onshore [17]. From the geographical standpoint, Cropper et al. 2014 have determined two zones that are interested by the upwelling occurrence: (a) 21–261N with strong permanent annual upwelling zone; and (b) 26–351N with weak permanent annual upwelling zone. Upwelling is typically a year-round occurrence, but its magnitude is weaker than in the permanent upwelling zone. Stronger upwelling is present in summer, associated with the trade-wind migration. According to [19], for coastal upwelling to occur across NW Africa, three factors are required: (1) persistent winds, (2) a solid boundary, and (3) the Coriolis effect.

Regarding the temporal evolution, the cloud moves inland in the course of the night, as land temperatures, and hence evaporation rates, fall. In fact, Marine layer clouds often reach their maximum extent around sunrise as this is when the air near the surface usually reaches a minimum temperature. The colder

surface temperatures enhance the inversion layer and also make it more likely the air beneath the inversion base is saturated with a relative humidity of 100%. As the day progresses, sunlight that penetrates through the clouds will warm the surface and the air above. The warming is greatest over land areas as land heats up much faster than water. As the air warms it is mixed upwards and will begin to mix into the clouds. This warming of the cloudy air decreases the relative humidity below the 100% level and the cloud begins to evaporate. Strong winds above the clouds can mix in drier air also leading to more evaporation. A thicker marine layer will dissipate slower as it will take more time for enough warm and/or drier air to mix in and evaporate all the cloud.

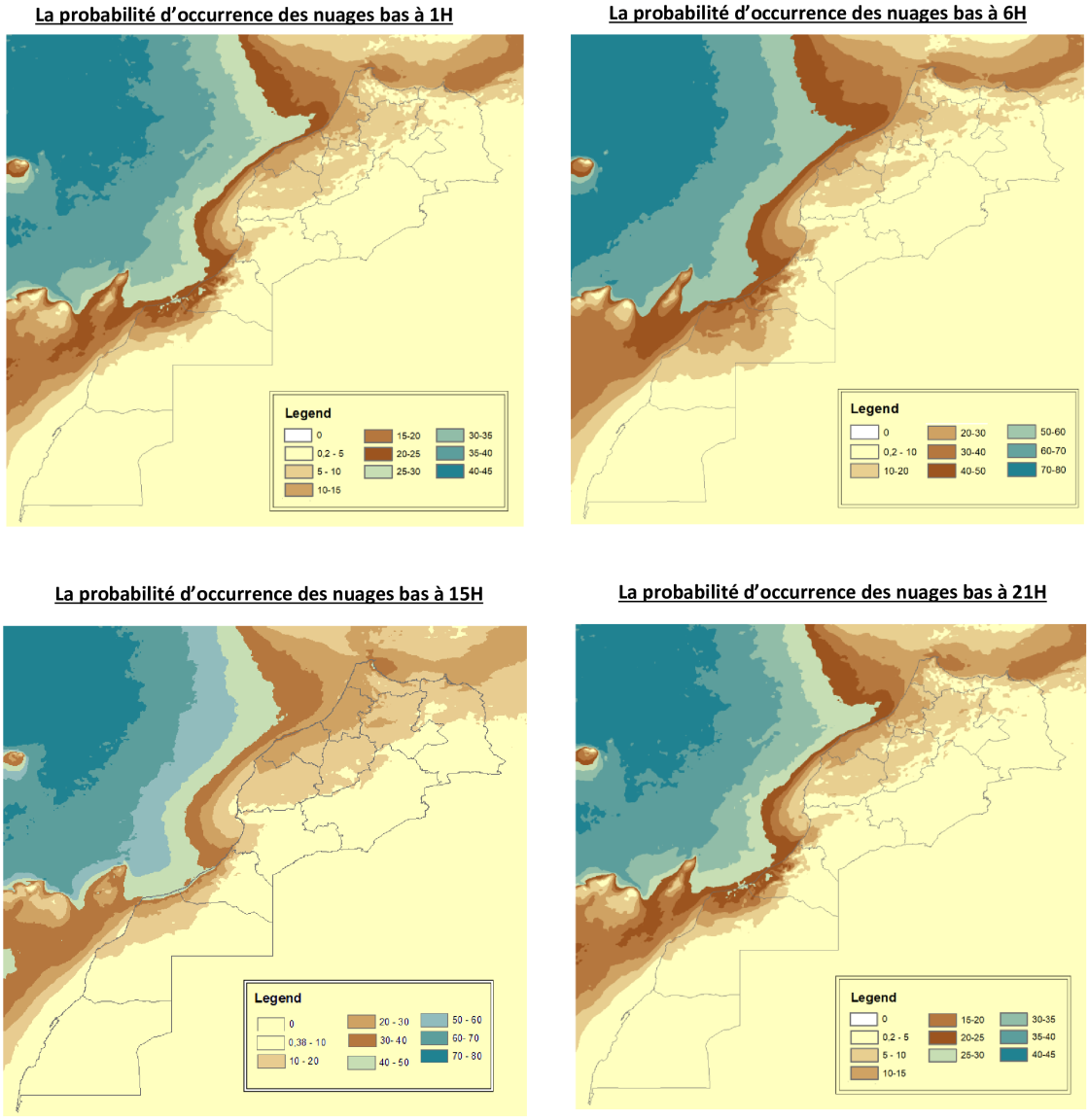


Figure 3: The probability of occurrence of low clouds only based on AROME forecasts at: a) 0100 UTC, b) 0600 UTC, c) 1500 UTC, and d) 2100 UTC.

Regarding the distribution of low clouds in reality, the observed frequencies show a strong presence of low clouds for all CPMs along the whole day as found in the probability of occurrence in AROME forecasts. It should be noted that the observed frequencies are affected by the availabbility of data and the working hours at each CPM. But, generally speaking, there is a great qualitatively agreement between the spatial distribution issued from AROME and observations.

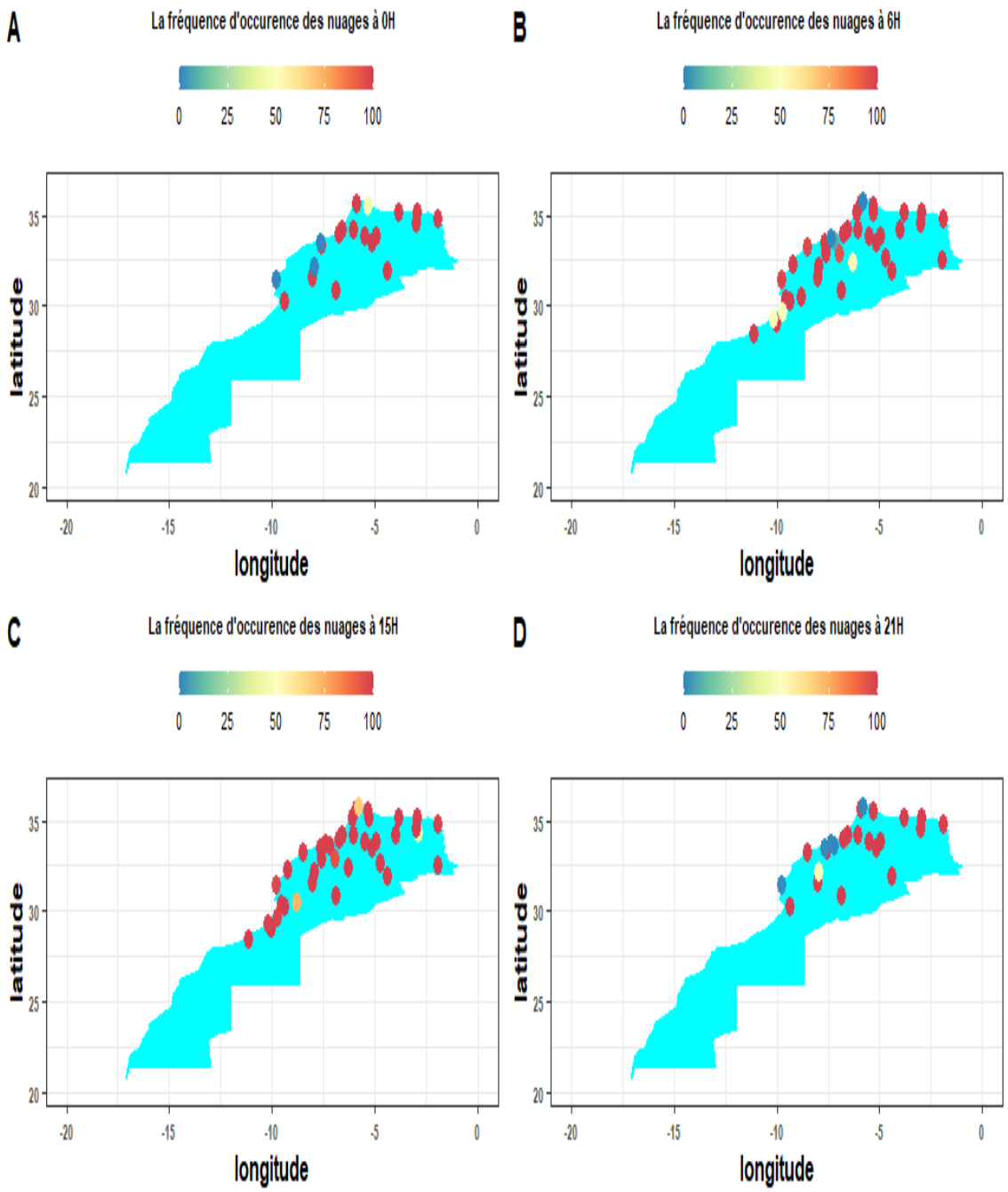


Figure 4: The observed frequency of low clouds occurrence based on observations at 41 synoptic meteorological stations over the study domain at a) 0100 UTC, b) 0600 UTC, c) 1500 UTC, and d) 2100 UTC.

3.3 Occurrence of fog and low clouds

It is seen clearly from Figure 5 that the maximum of the probability of fog and low clouds occurrence take place along the coastal areas, espacially the Atlantic ones. Indeed, low clouds transported onshore frequently lead to fog conditions as their bases touch the terrain. Advection of low clouds along the coast mostly occurs as a result of low pressure systems. Two noticeble areas with probabilty peaks : Casa-Safi, and Essaouira-Guelmim. The probability of occurrence is very high in these cities during the night, but in the morning they record lower values.

In fact,[22] have studied the fog climatology at the Grand Casablanca region by investiagting the local meteorological and spatial characteristics of coastal fog occurrence over this region. Fog-type analysis shows that advection–radiation fog events are the most common fog type in the area, followed by fog resulting from cloud-base lowering and radiation fog. Regarding the diurnal and seasonal distribution of the fog events, a maximum of their occurrence is classically observed during nighttime in winter. In addition, the synoptic analysis demonstrates that the advective processes associated with the sea-breeze circulation during daytime, followed by nocturnal radiational cooling early in the night, can often lead to fog formation over the GCB region.

Regarding the southern area, from Agadir to Gulemim, many studies have focused on harvesting fog water in the Ifni region to help a small community that lives in an arid region and that needs this water. [18]. The authors stated that even if the climate in this part of the Moroccan Atlantic coast is that of a desert, cloudiness and relative humidity are high, especially during the summer ($\geq$ 87 % of humidity) when trade winds bring stratocumulus clouds that predominate below 1,200 m. Stratiform clouds are possibly the most frequent and characteristic feature of this Atlantic coast, with a high number of fog days, above all in the summer [19]. At this time of the year the wet oceanic air cools as it flows over the cold waters of the Canary sea current; and it is compressed by an inversion layer, thereby increasing its stability and facilitating the formation of low-lying cloudiness that, on reaching the African continent, hits the coastal relief more easily. [20]. In fact, the relationship between these two phenomena is that low clouds transported over the coast often lead to fog conditions when their bases touch the ground.

And on the other hand, many radiation fog events are often dissipated or lifted from the surface after sunrise when daytime convection begins. This process can lead to clear sky after a few minutes/hours, but in some cases the fog turns into low stratus that can persist in the area even throughout the day.[23]

Regarding the temporal evolution of the probability of occurrence of both phenomena, it is found that the peak of occurrence has a south-north decreasing direction from 0100UTC to 0600UTC. In fact the maximum occurred first in the southern area at 0100 UTC, then the probabilities have similar values at the two areas at 0200UTC and 0300UTC. Then, the peak of probability decreases in the southern area while it increases in the northern one with the highest values at 0600UTC. By that, we notice that the peaks of the values do not appear during the day. And at night, we can note that the probability of occurrence of the two phenomena in the coastal cities starts to have again higher values.

Based on observations, the stations recorded a high probability of occurrence of fog and low clouds in the coastal cities especially those on the Atlantic Ocean side. Besides, some peaks are observed during the night in the stations of Rabat, Casablanca and EL Jadida. However, during the day, the values of the probability of occurrence showed a decrease, with a dissipation towards the southern region.

References

[1] Ivana Bartoková, Andreas Bott, Juraj Bartok, and Martin Gera. Fog prediction for road traffic safety in a coastal desert region: Improvement of nowcasting skills by the machine-learning approach. *Boundary-layer meteorology*, 157(3):501–516, 2015.

[2] Driss Bari and Mohamed El Khlifi. Lvp conditions at mohamed v airport, morocco: local characteristics and prediction using neural networks. *International Journal of Basic and Applied Sciences*, 4(4):354, 2015.

[3] Philipp Kneringer, Sebastian J Dietz, Georg J Mayr, and Achim Zeileis. Probabilistic nowcasting of low-visibility procedure states at vienna international airport during cold season. *Pure and Applied Geo-*

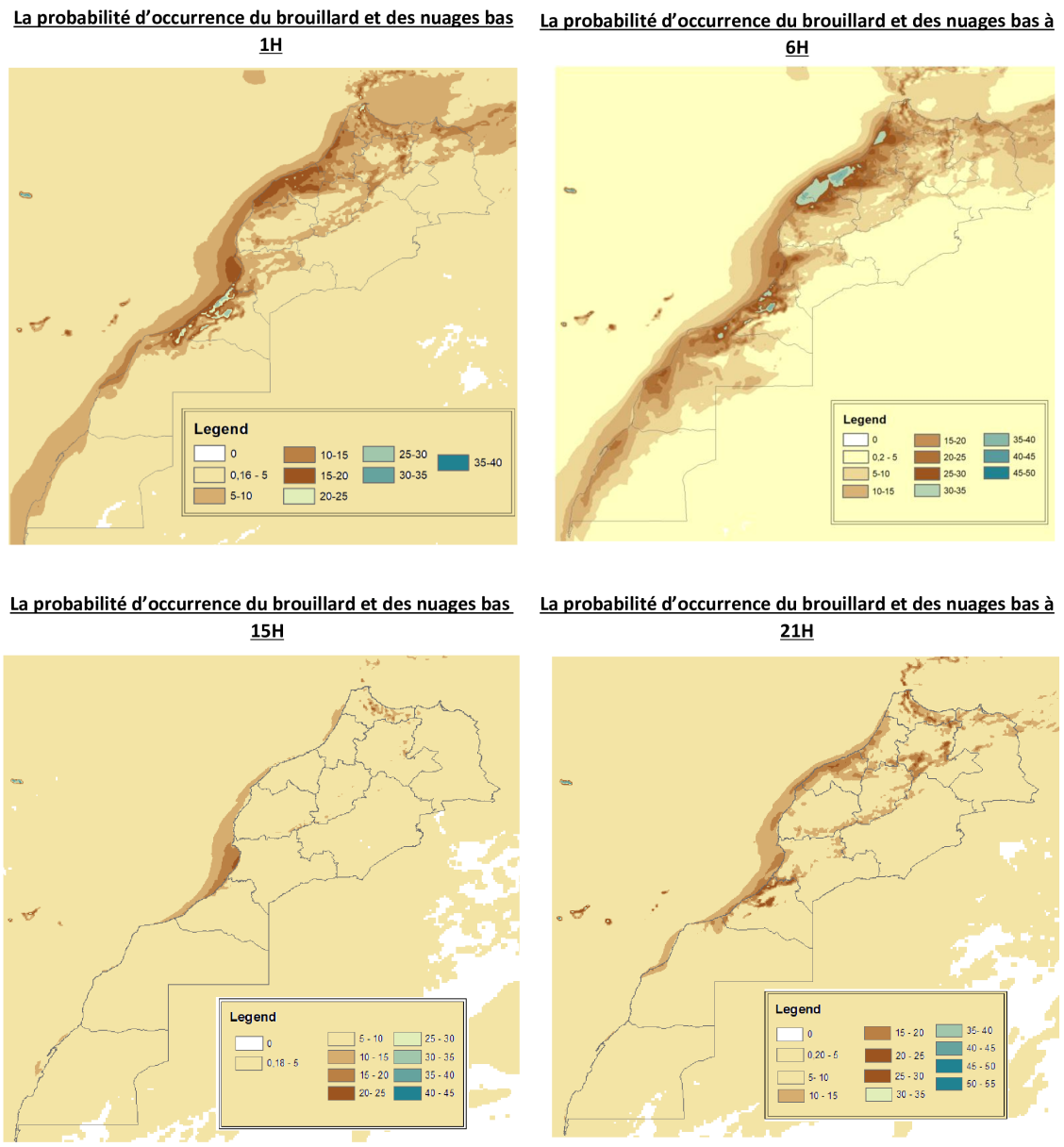


Figure 5: The probability of occurrence of fog and low clouds only based on AROME forecasts at: a) 0100 UTC, b) 0600 UTC, c) 1500 UTC, and d) 2100 UTC.

physics, 176(5):2165–2177, 2019.

[4] Lei Zhu, Guodong Zhu, Lei Han, Nan Wang, et al. The application of deep learning in airport visibility forecast. *Atmospheric and Climate Sciences*, 7(03):314, 2017.

[5] B Zhou and J Du. Fog prediction from a multimodel mesoscale ensemble prediction system. *Weather and Forecasting*, 25(1):303–322, 2010.

[6] Peter D Dueben and Peter Bauer. Challenges and design choices for global weather and climate models based on machine learning. *Geoscientific Model Development*, 11(10):3999–4009, 2018.

[7] Driss Bari. Visibility prediction based on kilometric nwp model outputs using machine-learning regression. In *2018 IEEE 14th international conference on e-Science (e-Science)*, pages 278–278. IEEE, 2018.

[8] Driss Bari and Abdelali Ouagabi. Machine-learning regression applied to diagnose horizontal visibility from mesoscale nwp model forecasts. *SN Applied Sciences*, 2(4):1–13, 2020.

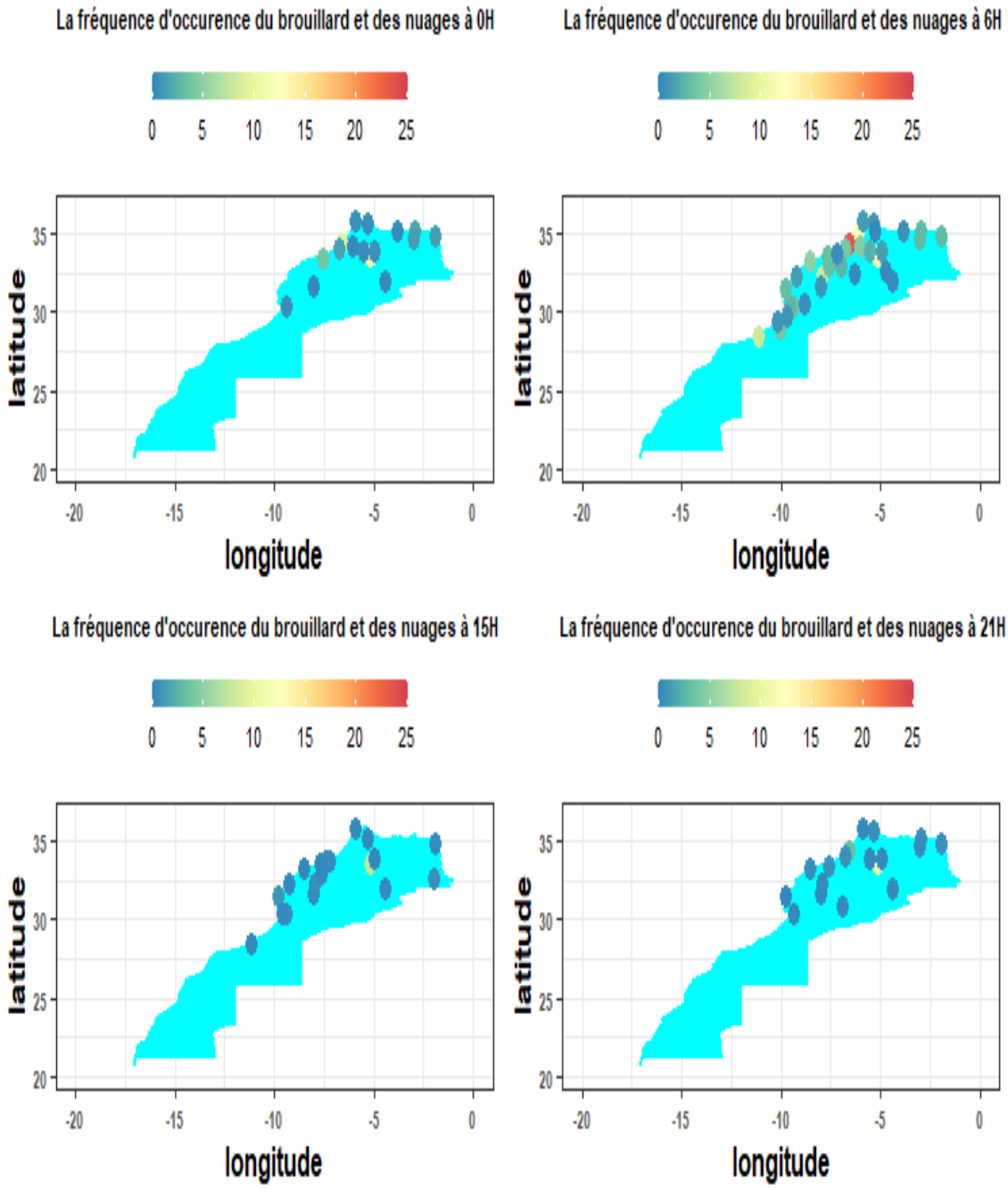


Figure 6: The observed frequency of fog and low clouds occurrence based on observations at 41 synoptic meteorological stations over the study domain at a) 0100 UTC, b) 0600 UTC, c) 1500 UTC, and d) 2100 UTC.

[9] PG Duynkerke. Radiation fog: A comparison of model simulation with detailed observations. *Monthly Weather Review*, 119:324–341, 1991.

[10] T Bergot and D Guedalia. Numerical forecasting of radiation fog. part i: Numerical model and sensitivity tests. *Monthly Weather Review*, 122(6):1218–1230, 1994.

[11] T Bergot, D Carrer, J Noilhan, and P Bougeault. Improved site-specific numerical prediction of fog and low clouds: A feasibility study. *Weather and Forecasting*, 20:627–646, 2005.

[12] D Bari, T Bergot, and M El Khlifi. Local meteorological and large-scale weather characteristics of fog

over the grand casablanca region, morocco. *Journal of Applied Meteorology and Climatology*, 55(8): 1731–1745, 2016.

[13] D Bari, T Bergot, and M El Khlifi. Numerical study of a coastal fog event over casablanca, morocco. *Quarterly Journal of the Royal Meteorological Society*, 141:1894–1905, 2015. DOI:10.1002/qj.2494.

[14] I Gultepe, R Tardif, SC Michaelides, J Cermak, A Bott, J Bendix ans MD Müller, M Pagowski, B Hansen, G Ellrod, W Jacobs, G Toth, and SG Cober. Fog research: A review of past achievements and future perspectives. *Pure and Applied Geophysics*, 164:1121–1159, 2007.

[15] Nawo Eguchi, Tadahiro Hayasaka, and Masahiro Sawada. Maritime-continental contrasts in the properties of low-level clouds: A case study of the summer of the 2003 yamase, japan, cloud event. *Advances in Meteorology*, 2014, 2014.

[16] Mohamed S Ghonima, Joel R Norris, Thijs Heus, and Jan Kleissl. Reconciling and validating the cloud thickness and liquid water path tendencies proposed by r. wood and jj van der dussen et al. *Journal of the Atmospheric Sciences*, 72(5):2033–2040, 2015.

[17] J Olivier and PL Stockton. The influence of upwelling extent upon fog incidence at lüderitz, southern africa. *International Journal of Climatology*, 9(1):69–75, 1989.

[18] Maria Victoria Marzol and J Sánchez. Fog water harvesting in ifni, morocco. an assessment of potential and demand. *Die Erde*, 139(1-2):97–119, 2008.

[19] Inocencio Font Tullot. El clima del África occidental española. *Servicio Meteorológico Nacional*, 1949.

[20] Ma V Marzol, A Alaeddine, J Sanchez, and A Derhem. Evaluation of fog collection in ifni, morocco. *Biggs, A*, pages 387–390, 2007.

[21] Chang Ki Kim and Seong Soo Yum. Local meteorological and synoptic characteristics of fogs formed over incheon international airport in the west coast of korea. *Advances in Atmospheric Sciences*, 27(4): 761–776, 2010.

[22] Driss Bari. *Etude du brouillard en zone côtière par modélisation des processus physiques de la couche limite atmosphérique: cas du Grand Casablanca (Maroc)*. PhD thesis, Université Paul Sabatier-Toulouse III, 2015.

[23] Carlos Román-Cascón, Carlos Yagüe, Gert-Jan Steeneveld, Gema Morales, Jon A Arrillaga, Mariano Sastre, and Gregorio Maqueda. Radiation and cloud-base lowering fog events: Observational analysis and evaluation of wrf and harmonie. *Atmospheric research*, 229:190–207, 2019.

Improvement of fog forecast at hectometric scales in AROME

Salomé Antoine, Rachel Honnert and Yann Seity (Météo-France, CNRM/Toulouse, France)

1 Introduction

Fog is a meteorological phenomenon defined by a low visibility (less than 1000 m) due to airborne hydrometeors. The economical and safety impact of fog is very important, in particular for aeronautics. Fog is usually a local phenomenon as depending from the local environment. Thus, improving fog predictability at very fine scales is heavily requested by the economic world.

For these reasons, the international field campaign SoFog3D took place in the south west of France between October 2019 and March 2020. The French numerical weather prediction (NWP) model AROME (Seity et al. 2011, Brousseau et al 2016) has been used during the campaign in order to plan the intensive observation periods (IOPs), which gave a first glimpse of the potential and the weakness of the operational model AROME in the forecasts of fog.

However, the ultimate goal is to forecast fog for aeronautics. It seems that the model horizontal resolution may have an impact on the representation of fog. Indeed, MetOffice runs a 333 m resolution model (the London Model, Boutle et al., 2016) on Heathrow. Although, the impact of the vertical resolution seems to be predominant (Philip et al., 2016). It is also well known that microphysics plays a major role in fog forecasts (Bergot, 2013). We also know that the current operational AROME 1-moment microphysics scheme has some limitations and suffers from a problem of over-estimation of cloud water contents in fogs (Zhou and Ferrier, 2008 ; Stolaki et al., 2015). Thus, we intend to evaluate the 2-moment microphysics scheme LIMA (Liquid Ice Multiple Aerosols, Vié et al., 2016) with the help of the SoFog3D campaign observations.

This particular work aims to quantify the benefit of both a fine resolution and the new 2-moment microphysical scheme LIMA in the representation of fog in the AROME model.

In this article, the SoFog3D field campaign is presented in the first section. Then numerical experiments are presented in section 2. Finally, AROME is compared to SoFog3D observations. At first, a statistical study compares AROME using the different configurations with observations, in order to put forward the systematic behavior of the model. Then, a case study allows to analyze in details some of the issues and ways in which the forecast could be improved.

2 SoFog3D field Experiment

To improve the comprehension of fog characteristics in the three dimensions, the Centre National de Recherches Météorologiques (CNRM) organized a field campaign: SoFog3D (South west Fogs 3 Dimensions).

In this study, we are particularly interested by some data from the SoFog3D campaign. Visibility has been routinely measured by stations from the french operational RADOME network complemented by stations installed for the campaign. Between 11 and 18 stations were available. This variability is due to the timing of the installation of SoFog3D specific stations and also some technical issues on them. Additional information about water content in atmosphere are given by 6 radiometers and 3 BASTA radar. During Intensive Observation Periods (IOPs), one tethered balloon measured microphysical and turbulence parameters at the super site and 4 radiosondes (at the super site, in Agen, Toulouse and Bordeaux) were launched. The launch time depends

from the site, for example in the super site, it was at 18 UTC (initial condition of night), at the beginning of fog development, when fog is mature and at the dissipation of fog. The campaign numbered 15 IOPs with 20 nights of observation.

3 AROME and sensitivity tests

The french mesoscale operational model AROME has been compared to observations of SoFog3D. Different setups would be evaluated with two grids :

- 1250-m horizontal resolution with 90 vertical levels (referred as L90) and a first level at 5 m : close to the operational setups,
- 500-m horizontal resolution with 156 vertical levels (referred as L156) and a first level at 1 m : the hectometric version.

Moreover, two microphysical schemes are tested : the operational one-moment scheme ICE3 (Pinty and Jabouille, 1998) and the new two-moments scheme LIMA (Vié et al. 2016). In addition, 2 processes will be switched on or off : a sink of liquid water by deposition (Mazoyer et al. 2017, Zhang et al. 2014) and subgrid condensation of the microphysic schemes.

4 Results

4.1 Statistical study

At first, the ability of AROME to forecast fog has been tested over the 4-month period of the campaign by using visibility observations.

The hit rate (HR) is the percentage of accurate forecast over the percentage of occurrence. The false alarm rate (FAR) is the percentage of false alarm over the percentage of (good or bad) forecast events. When the Frequency Bias Index (FBI) is larger than 1 the fog is over-forecast. The Critical Success Index (CSI) is the relation of accurate forecast over the percentage of fog events (forecast or observed).

Table 1: Impact of the grid on fog forecast. Comparison between 1250-m L90 grid and 500m L156 grid with ICE3 microphysic scheme, with subgrid condensation and without deposition, 00 UTC run from lead time +25 to +30. Statistically significant results (Bootstrap test with 95% confidence level) are in color or in bold. Positive impacts are in green and negative are in red.

	CSI	HR	FAR	FBI
1250m L90	0.41	56	38	0.91
500m L156	**0.44**	**67**	**44**	**1.2**

Preliminary results in Table 1 show that the detection rate (TD) is better in the 156 levels AROME-500m configuration. Even if the model over-forecast fog events, which leads to false alarms, the CSI is significantly better. This is consistent with a previous study by Philip et al. (2016), which showed that an improved vertical resolution and the height of the first level in particular is crucial to forecast fog properly in AROME. As expected, removing the subgrid condensation scheme has a significant negative impact on the detection rate and on the bias (not shown).

Tab. 2 shows a comparison between LIMA and ICE3 in the hectometric version with the deposition scheme and the subgrid condensation. LIMA is comparable to ICE3.

Table 2: Impact of the microphysic scheme on fog forecast. Comparison with 500-m L156 grid between ICE3 and LIMA schemes, with deposition and subgrid condensation, 00 UTC run from lead time +25 to +30. Statistically significant results (Bootstrap test with 95% confidence level) are in color or in bold. Positive impacts are in green and negative are in red.

	CSI	HR	FAR	FBI
500m L156 ICE3	0.43	64	43	1.13
500m L156 LIMA	0.4	64	49	**1.26**

4.2 Case study : SoFog3D IOP-11

The IOP-11 is one of the most interesting case of the campaign in terms of vertical fog development. During the night from February 8 to 9, 2020, fog appears at about 22 UTC. Fog thickens until it changes into status at about 2h.

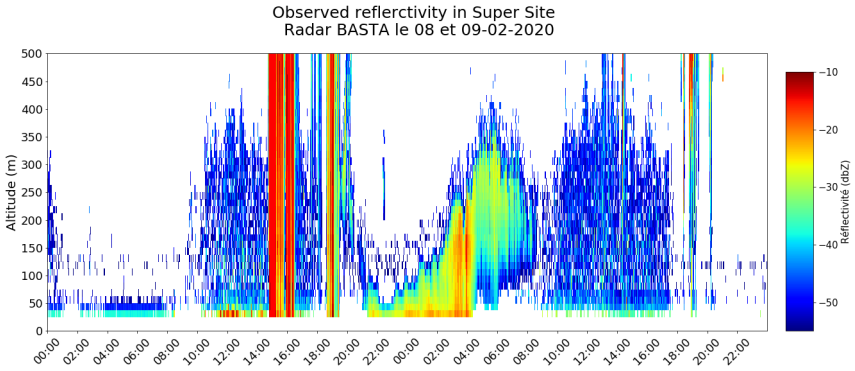


Figure 1: BASTA-radar reflectivity during the IOP-11 of SoFog3D.

Unfortunately, in the 1250m L90 configurations, a small piece of clear sky persists just over the Super Site. The 500-m configurations however provide a better representation of the spatial expanse of the fog during the night. Generally, fog is finer in the simulations than in the observations. Fig. 1 presents Radar reflectivities of the IOP-11 and Fig. 2 shows the associated simulated reflectivities in various AROME configurations. In the simulation 1250 m L90 ICE3, a delay is observed in both formation and dissipation. The delay in dissipation exists in all configurations, because fog lifting into stratus is ill-represented by the model, even if fog lifts at the very end of the case. Simulated reflectivity values are too strong in ICE3 and may be a little better in LIMA.

5 Conclusion

In this preliminary work, we took advantage of the SoFog3D field campaign (dedicated to understanding fog) to test the AROME model ability to forecast this phenomenon.

Sensitivity tests were carried out : the resolution (hectometric and close to operational), the microphysical scheme (1-moment ICE3 and 2-moment LIMA), as well as a new deposit term and the impact of subgrid condensation. At this stage the hectometric configuration is working well, in particular due to a lower first level. Even if delays are observed during both the phases of formation and of dissipation, LIMA simulates a more realistic reflectivity. We are optimistic about LIMA's ability to predict fog in the future.

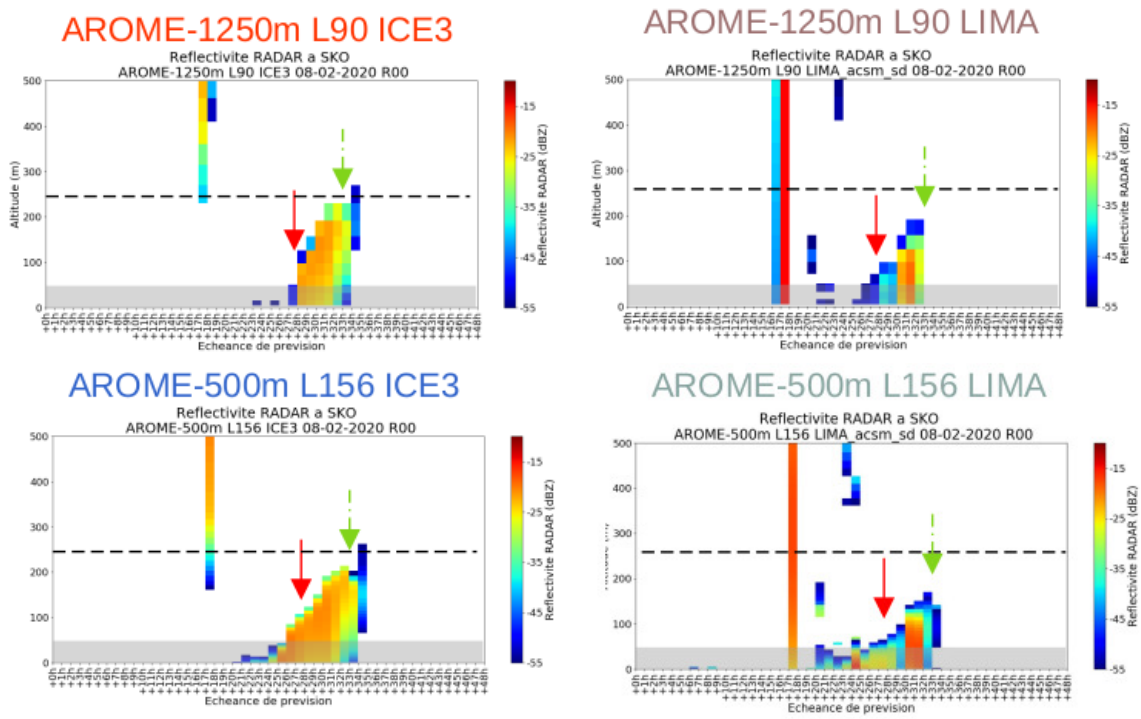


Figure 2: Simulated reflectivity during the IOP-11 of SoFog3D by 8th February 2020 00 UTC run with AROME 1250m L90 ICE3 (Top-left), AROME 500m L156 ICE3 (Bottom-left), AROME 1250m L90 LIMA (Top-right) and AROME 1250m L90 ICE3 (Bottom-right). The fog to stratus transition (respectively the stratus dissipation) of the fog event (from Basta data in Fig. 1) are pointed out by a red (respectively green) arrow.

6 References

Bergot T.. Small-scale structure of radiation fog : a large-eddy simulation study. Quarterly Journal of the Royal Meteorological Society, 139(673) :1099–1112, 2013. doi : 10.1002/qj.2051.

Boutle I. A., A. Finnenkoetter, A. P. Lock, and H. Wells. The London model : forecasting fog at 333m resolution.Quarterly Journal of the Royal Meteorological Society, 142(694) :360–371, 2016. doi : 10.1002/qj.2656.

Brousseau P., Y. Seity, D. Ricard, and J. Léger. Improvement of the forecast of convective activity from the AROME-france system. Quarterly Journal of the Royal Meteorological Society, 142(699) : 2231–2243, 2016. doi : 10.1002/qj.2822.

Mazoyer M., C. Lac, O. Thouron, T. Bergot, V. Masson, and L. Musson-Genon. Large eddy simulation of radiation fog : impact of dynamics on the fog life cycle. Atmospheric Chemistry and Physics, 17(21) :13017, 2017.

Philip A., T. Bergot , Y. Bouteloup, and F. Bouyssel. The Impact of Vertical Resolution357on Fog Forecasting in the Kilometric-Scale Model AROME: A Case Study and Statistics.358Weather and Forecasting,31 (5), 1655–1671, 2016. doi : 10.1175/WAF-D-16-0074.1

Pinty J. and P. Jabouille. A mixed-phase cloud parameterization for use in a mesoscale non-hydrostatic model : simulations of a squall line and of orographic precipitations. Proceedings of the AMS conference on cloud physics, 1998.

Seity Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. The AROME-France Convective-Scale Operational Model. Monthly Weather Review, 139(3) :976–991, 2011. doi : 10.1175/2010MWR34

Stolaki S., M. Haeffelin, C. Lac, J.-C. Dupont, T. Elias, and V. Masson. Influence of aerosols on the life cycle of a radiation fog event. A numerical and observational study. Atmospheric Research, 151 : 146–161, 2015. doi : 10.1016/j.atmosres.2014.04.013.

Vié B., J.-P. Pinty, S. Berthet, and M. Leriche. LIMA (v1.0) : A quasi two-moment microphysical schemedriven by a multimodal population of cloud condensation and ice freezing nuclei. Geoscientific ModelDevelopment, 9(2) :567–586, 2016. doi : 10.5194/gmd-9-567-2016.

Zhang X., L. Musson-Genon, E. Dupont, M. Milliez, and B. Carissimo. On the influence of a simple microphysics parametrization on radiation fog modelling : A case study during parisfog. Boundary-layer meteorology, 151(2) :293–315, 2014.

Zhou B. and Brad S. Ferrier. Asymptotic analysis of equilibrium in radiation fog. Journal of Applied Meteorology and Climatology, 47(6) :1704–1722, 2008. doi : 10.1175/2007JAMC1685.1.

Testing visibility diagnostics in AROME at high resolution

Walid CHIKHI[(1,*)], Abdenour AMBAR[(1,2)], Mohamed MOKHTARI[(1)]

(1) Numerical Weather Prediction Department, Office Nationale de La Météorologie, Algiers, ALGERIA.
(2) Faculty of Earth Sciences and Terittory Planning, Sciences and Technologies University USTHB, Algiers, ALGERIA.
(*) Contact person : ***waliidchikhi@gmail.com***

1 Introduction

The main objetive of this work has been to test and evaluate new visibility diagnotiscs fields coded in AROME at Météo France by Ingrid Etchevers (Dombrowski-Etchevers et al., 2018). In order to take full advantage of the physical parameterization of AROME, two configurations at very high resolution have been prepared : 1.3km and 0.5km resolution.

2 Visibility diagnostics in AROME CY46t1

Basically, visibility is defined as the distance up to which a meteorologist observer near the ground or the sea can see and identify an object in a given direction, at a specific time and place. This observation can be quantified and represented by a distance measure unit (meter, kilometers, feets, etc..). This parameter is reported regularly at the METAR observation messages (METeorological Aerodrome Report), and when it is significally reduced, the obscuring weather phenomena is also reported (fog, haze, rain, duststorm, etc..).
The current study is based on visibility diagnostics coded at Météo France by Ingrid Etchevers (Dombrowski-Etchevers et al., 2018). This new diagnostic is calculated separately due to hydrometeors (precipitation) and clouds (fog) at the height of the lowest model level, which is 5 meters. However, this height could be defined at namelists by the key ‘HVISI’ :

- If HVISI < 5 meters, visibility will be calculated automatically at the lowest model level.
- If HVISI > 5 meters, visibility will be calculated at the predefined height at namelist.

The computation of the visibility is based on Koschmieder's formula which relates the horizontal visibility to the extinction coefficients of the different hydrometeors (liquid water, ice, snow, rain and graupel). The basic equations are the following :

$$\mathrm{PVISICLD}=\frac{-\ln(0{,}05)}{0{,}013+\beta\mathrm{cld}+\beta\mathrm{ice}} \qquad (1)$$

PVISICLD : Visibility due to clouds (liquid water and ice water) in meter;
βcld : Extinction coefficient due to cloud liquid water.
βice : Extinction coefficient due to cloud ice water.

$$\mathrm{PVISIHYD}=\frac{-\ln(0{,}05)}{0{,}013+\beta\mathrm{rain}+\beta\mathrm{snow}+\beta\mathrm{graupel}} \qquad (2)$$

Where :
PVISIHYD : Visibility due to hydrometeors in meter;

βcld : Extinction coefficient due to rain.
βice : Extinction coefficient due to snow.
βgraupel : Extinction coefficient due to graupel.

Extinction coefficients β are calculated using the following relation : $\beta = a \cdot C^b$

Where : - **C** **:** Hydrometeors content (water - ice - rain - snow - graupel).
- **a , b :** Coefficients determined empirically.

Those formulations are coded in the new routine "acvisih.F90" which is directly called from "apl_arome.F90", and are activated by switching-on the key LXVISI in the forecast namelist (LXVISI=.TRUE.).

3 Experiment configuration

Study area

This study have been established on a small domain in Northern Algeria, covering the red area in figure 1. This region have been particularly picked because of the international airport of Algiers, where low visibility is frequently observed and thus it present a serious challenge for aeronautical forecasters due to its direct impact on airlines economics. The domain used for AROME 0.5km is smaller than the one used for AROME 1.3km just to reduce the computational costs of AROME 0.5km.

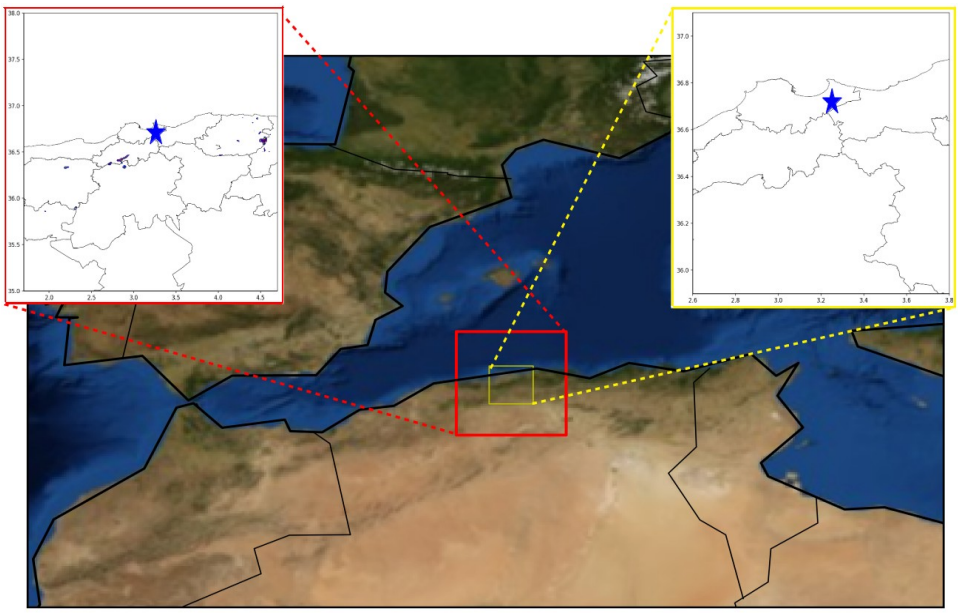


Figure 1: Study area location (red box : AROME 1.3km domain, yellow box : AROME 0.5km domain, blue star : international airport of Algiers).

Experiment setup

The occurrence of parameters reducing visibility (haze 1km to 5km, fog < 1 km) remain very sensitive to spatial resolution. So two configurations have been setup to evaluate the visibility diagnostics at different high resolution (1.3 and 0.5 km). We prepared two AROME configurations (table 1), both had 90 vertical levels based on CY46T1 and performed on the same study domain (figure 1).

Table 1: AROME configurations characteristics

Characteristics		AROME 1.3	AROME 0.5
Cycle		CY46T1.bf.06	
Initial conditions		ARPEGE	
Coupling frequency		1H	
Horizontal resolution		1.3 km	0.5 km
levels		90 level	
First level from ground		5 meters	
Time-step		40 s	10 s
Grid		232 x 232	241 x 241
Domain	Center	3.2°E – 36.5°N	
	Latitude	35°N – 38°N	35.9°N – 37.1°N
	Longitude	1.7°E – 4.7°E	2.6°E – 3.8°E

Simulations and evaluation

Simulations have been performed for two different months : December 2018 (28 day) and May 2018 (13 day). Observations data were obtained from METAR messages of the International airport station DAAG (36.69°N ; 3.21°E). Visibility data were organized into 3 categories :

< 1000 m : very bad visibility conditions,
[1000 m – 5000m[: bad visibility conditions,
> 5000 m : good visibility conditions.

The evaluation will be based on **contingency tables** (table 2) in order to determine the model efficiency in predicting visibility conditions categories. This method is described as a best way of testing the skill especially in weather forecasts. The forecasting skill is evaluated by calculating various skill scores. From the contingency table, the following skill score will be calculated :

Table 2: Contingency table.

		Observation			
		Very bad	Bad	Good	**Total**
Forecast	Very bad	a	b	c	**P**
	Bad	d	e	f	**O**
	Good	g	h	i	**N**
	Total	**J**	**K**	**L**	**M**

- Percentage Correct (PC) : this score shows a fraction of the forecasts that is correct. It ranges from 0-1, with the perfect score being 1. This score can be misleading since it is heavily influenced by common category, usually no event in the case of rare events. It is determined by the formula :

$$PC = \frac{a+e+i}{M} * 100$$

- Probability of detection/ hit rate (POD) : POD shows usually the fraction of observed yes events that were correctly forecasted. The score ranges from 0-1, with score of 1being the perfect score. It is sensitive to hits but ignores false alarm making it good for rare events. But for our case, operationally speaking, we'll consider that if the observed visibility is "Bad" and the forecast is "Very Bad", then the event is well predicted. The formula is :

$$PC = \frac{a+b+e}{J+K} * 100$$

- False Alarm Ratio (FAR) : It shows the percentage of the predicted yes events that actually did not occur. The FAR skill score ranges from 0-1 with the perfect score being o. The skill score is sensitive to false alarms, but ignores misses. It is also very sensitive to the climatological frequency of the event. FAR skill score can be determined by applying the formula below :

$$PC = \frac{c+f}{P+O} * 100$$

Simulations have been performed for a total of 41 days : 28 days for December 2019 and 13 days for May 2020. The situations have been selected based on the METAR observations. For each situation, we performed 24h run starting from the day before the event (j-1).

Figures 2 and 3 shows the simulations of visibility due to clouds simulated by AROME 1.3km and AROME 0.5km at two different dates :

- On May 13th 2018 at 09h (figure 2), we notice in both configurations that the same area is concerned by low visibility values, few differences are detected inside the area where more details are provided in AROME 0.5km.
- On December 29th 2018 at 06h (figure 3), the differences are more important and the area concerned by the low visibility values is significantly more important in AROME 0.5km (red circle in figure 3).

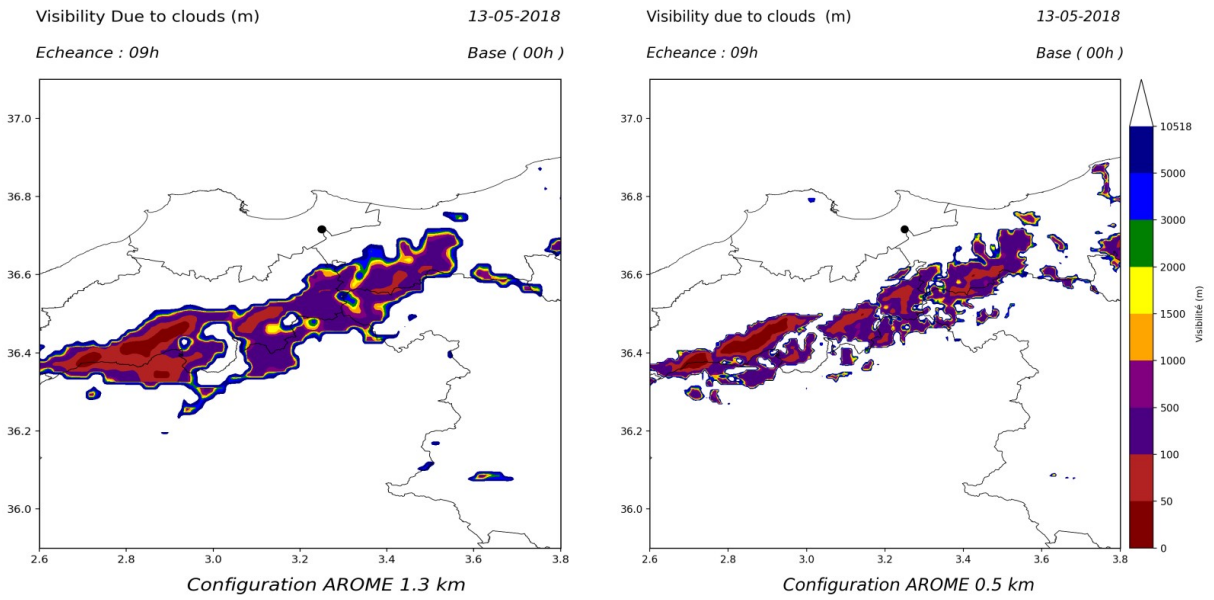


Figure 2: Comparison of visibility due to clouds simulated by AROME 1.3km and AROME 0.5km on 13 May 2018 at 09utc.

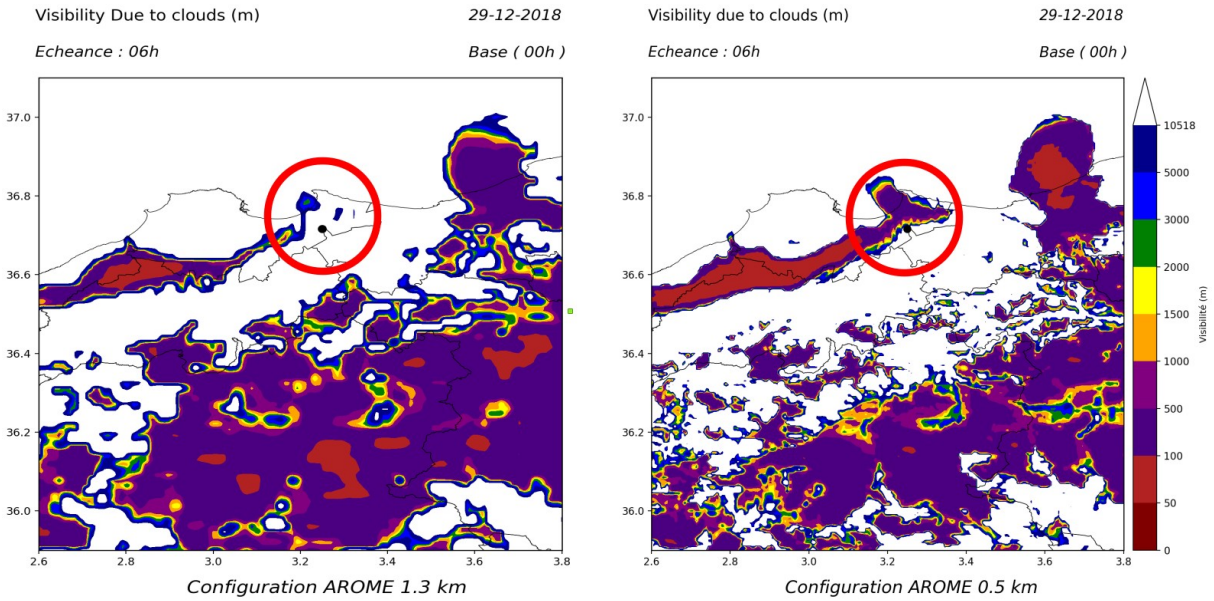


Figure 2: Comparison of visibility due to clouds simulated by AROME 1.3km and AROME 0.5km on 29 December 2018 at 06utc.

Table 3 and table 4 represents the contingency tables calculated based on hourly AROME 1.3km and AROME 0.5km simulations, respectively. Each one is decomposed into three parts : the whole period, December and May. Tab.05 represents the skill scores calculated for each configuration and for each period.

Table 3: Contingency table for AROME 1.3 km

		Observation (METAR)			
		All period			
Forecast (ARO 1.3)		VERY BAD	BAD	GOOD	Sum
	VERY BAD	17	40	46	103
	BAD	1	14	39	54
	GOOD	21	59	745	825
	Sum	39	113	830	982
		December			
		VERY BAD	BAD	GOOD	Sum
	VERY BAD	4	19	25	48
	BAD	1	1	22	24
	GOOD	16	43	539	598
	Sum	21	63	586	670
		May			
		VERY BAD	BAD	GOOD	Sum
	VERY BAD	13	21	21	55
	BAD	0	13	17	30
	GOOD	5	16	206	227
	Sum	18	50	244	312

Table 4: Contingency table for AROME 0.5 km

		All period			
		VERY BAD	BAD	GOOD	Sum
	VERY BAD	22	48	75	145
	BAD	2	10	28	40
	GOOD	15	55	727	797
	Sum	39	113	830	982
		December			
		VERY BAD	BAD	GOOD	Sum
	VERY BAD	8	24	45	77
Forecast (ARO 0.5)	BAD	1	0	16	17
	GOOD	12	39	525	576
	Sum	21	63	586	670
		May			
		VERY BAD	BAD	GOOD	Sum
	VERY BAD	14	24	30	68
	BAD	1	10	12	23
	GOOD	3	16	202	221
	Sum	18	50	244	312

Table 5: Skill scores calculated for AROME 1.3 km and AROME 0.5 km.

	AROME 0.5	AROME 1.3
	All period	
PC	77,29	79,02
POD	0,53	0,47
FAR	0,55	0,54
	December	
PC	79,55	81,19
POD	0,38	0,29
FAR	0,64	0,65
	May	
PC	72,43	74,35
POD	0,71	0,69
FAR	0,46	0,44

The results are very promising for both configurations, the percentage of correct forecasts is quite high: 79% with AROME 1.3km and 77% with AROME 0.5km. In most cases, simulated visibility due to hydrometeors and clouds were in the same categories of the observed visibility from METAR.

The probability of detection of “very bad” and “bad” visibility classes is acceptable but not good enough for both resolutions : 0,47 for AROME 1.3km and 0,53 for AROME 0.5km. It shows that improving the horizontal resolution of AROME from 1.3 km to 0.5km, improved slightly the probability of low visibility conditions forecasts. In terms of seasons, we noticed that this skill score is significantly more important during May than during December. This should be proved by performing simulations for a longer period (a year at least).

The false alarm ratio score is just acceptable and quite similar for both configurations : 0,54 for AROME 1.3km and 0,55 for AROME 0.5km. It turns out that false alarms are more frequent during December than during May.

4 Conclusion

Two sets of simulations at different high horizontal resolution have been performed in this study in order to test new visibility diagnostics in AROME (Cy46T1). The evaluation of visibility due to clouds and hydrometeors forecasts showed some interesting results over the international airport of Algiers, with acceptable results regarding “very bad” and “bad” visibility conditions detection. It seems that the scores are slightly better in December than in May, but this still need more investigation by evaluating the forecasts at other airports and over a longer period.

5 References

Dombrowski-Etchevers, I., Sanchez, I., Seity, Y., 2018: La visibilité : nouveau parametre issu des modeles de prévision météorologique. Presentation. Ateliers de Modélisation de l'Atmosphere, AMA, February 2018.

Kizito Amua Oundo : Visibility forecast verification at JOMO KENYATTA international airport, I50/80664/2012, July 2013.

Kunkel, B. A., 1984: Parameterization of droplet terminal velocity and extinction coefficient in fog models. J. Climate Appl. Meteor., 23, 34–41.

Philip, Alexandre, 2016: Apport d'une résolution verticale plus fine dans le calcul des tendances physiques pour la modélisation du brouillard dans le modèle AROME. Thesis. Université de Toulouse, 122 pp.

Sekula, P., 2018: Testing of new diagnostic fields: convective pack in cy43t2 and visibility in ALARO.ALADIN flat-rate stay report. Ljubljana 12.08.2018 – 8.09.2018

Simon, A., Dian, D., : Report on visibility tests with ALARO on cycle 43 t2 bf10 at SHMÚ. SHMÚ

Adapting the screening level diagnostics to improve AROME temperature forecasts in Alpine areas

Florian Meier, Clemens Wastl, Florian Weidle, Christoph Wittmann

1 Introduction

When it comes to the diagnostics of screening level parameters (e.g. 2m temperature and humidity) several options are available in the AROME/SURFEX code: CANOPY scheme (V. Masson et al, 2009, Geleyn 1988 and Paulsson 1970). At ZAMG the operational 2.5km models (AROME-Aut and its probabilistic counterpart C-LAEF) and the 1.2km nowcasting version AROME-RUC currently rely on the CANOPY scheme. Evaluations show that the methods have different strengths and weaknesses in different weather situations. And to make it even more complicated, it seems they do not work equally well for different types of locations, i.e. points located in flat areas vs. Alpine valleys vs. mountain tops. Figure 1 shows the BIAS averaged over 1month for all stations in Austria (blue) and for stations located in flatlands (green), mountains (red) and valleys (violet).

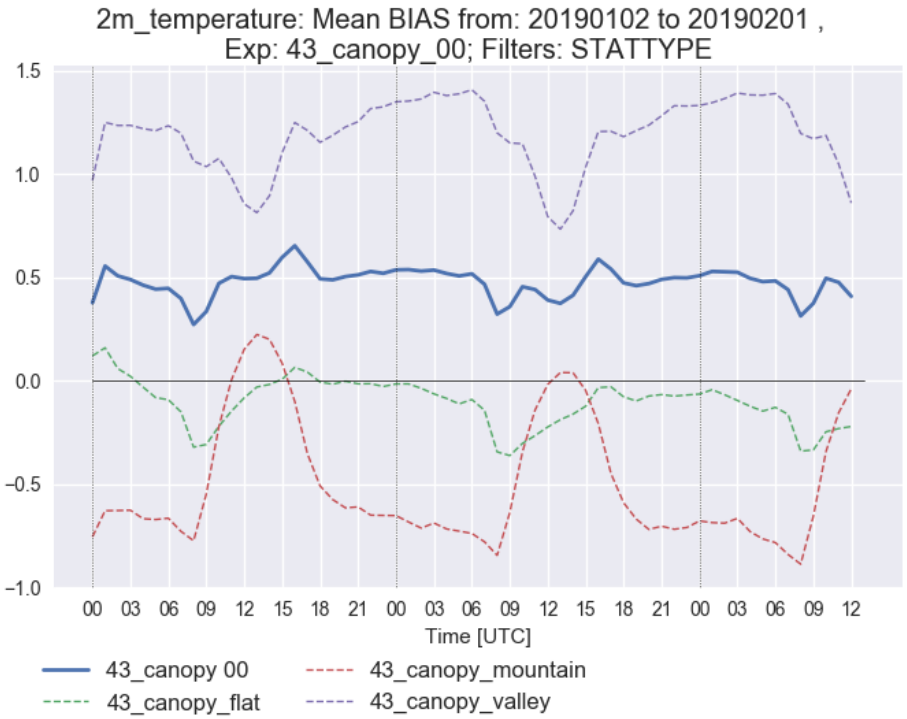


Figure 1: BIAS of 2m temperature (height corrected) for the period 20190102 – 20190201 of AROME-Aut using the standard CANOPY scheme for all Austrian stations (solid), and station selections according to station location (dashed).

The largest forecast errors for 2m temperature regularly occur during wintertime in Alpine valleys when shallow but pronounced stable boundary layers develop during nights with significant radiative cooling.
It is clear that any adaption of the diagnostics can just be used to somehow mitigate model deficiencies and in general it would be better to eliminate or at least improve the error source in the model physics and/or setup itself. Within this short article we summarize the work carried out to improve the current 2m temperature performance in the operational AROME models at ZAMG by 1) introducing some orographic dependency into the current diagnostics scheme (see section 2) and 2) introduce another method (Dian et al, 2016) as N2M=3 within SURFEX/AROME (see section 3). A representative test case (November 23rd 2020) was selected to illustrate the behaviour of the various options in a situation with (shallow) inversion layers developing in Alpine valleys during the night.

2 Introducing a "surface layer index" into 2m diagnostics

The different behaviour or systematic model error for different locations (valley vs. flatland vs. mountain top) led to the idea to slightly adapt the diagnostic scheme according to the topographic conditions and thus try to mitigate existing model deficiencies without major investments. As already mentioned, all operational AROME versions used at ZAMG currently apply the Canopy scheme (LCANOPY= .T.). By doing so, five extra SBL levels are introduced between the model surface and the lowest model level with surface fluxes from SURFEX as input. The model output 2m temperature and humidity is always based on the values at the second level from the ground. However, in certain situations and locations it turned out that it would lead to significantly better scores or reduced systematic errors when information from higher (e.g. on mountain tops) or lower (e.g. Alpine valleys) canopy levels is included instead of just using level CAN_T02 values. One may of course argue at this point that systematic errors could be left to post-processing (e.g. bias correction) afterwards, but the effectivity of this approach shown later justifies its implementation.
To include some dependency to orography into the 2m diagnostics scheme, we need at first a parameter which helps us to distinguish between grid points located in open flatland, in valleys or on mountain tops. We decided to use and adapt the "surface layer index" (aka "inversion factor" IFAC) which was already implemented within the INCA system at ZAMG (Haiden et al., 2011). The index helps to characterize to which extent a local terrain or grid point and its environment supports the formation of a distinct surface layer, i.e. an inversion. Such inversions are very likely e.g. in Alpine valleys and basins, because cold air is concentrated at the bottom of such areas, while on slopes and mountain peaks much warmer air dominates. The IFAC is calculated on the basis of the model orography, by scanning the surrounding orography at each grid point with a radius of 20km (tuneable via namelist). In equation 1 and 2 $\overline{z_{Ha}}$ and $\overline{z_{Hb}}$ represent the average altitude of the grid points situated above and below the current centre point z_H. If the majority of the surrounding grid points are located below the centre point Eq. 1 is applied, otherwise Eq. 2 is considered.

$$I_{IFAC} = max\left(0, min\left(\frac{\overline{z_{Hb}} - z_H - z_s}{z_s}, 1\right)\right) \qquad \text{(eq. 1)}$$

$$I_{IFAC} = max\left(-1, min\left(0, 0 - \frac{\overline{z_{Ha}} - z_H - z_s}{z_s}\right)\right) \qquad \text{(eq. 2)}$$

The predefined parameter z_s (tunable via namelist) definines the minimum altitude difference between the centre and surrounding grid points. In our case it is set to z_s=¿175m, which means that only differences with more than 175m are considered in the calculation. We introduced z_s to reduce noise

in the IFAC field, especially in relative flat areas. By using this IFAC parameter we can now distinguish whether a grid point is located in an Alpine valley (IFAC negative with minimum of -1), the flatlands (IFAC around 0) or on the mountain ridges/tops (IFAC positive with maximum of +1). Figure 2 shows the IFAC field for the operational AROME/C-LAEF domain of Austria.

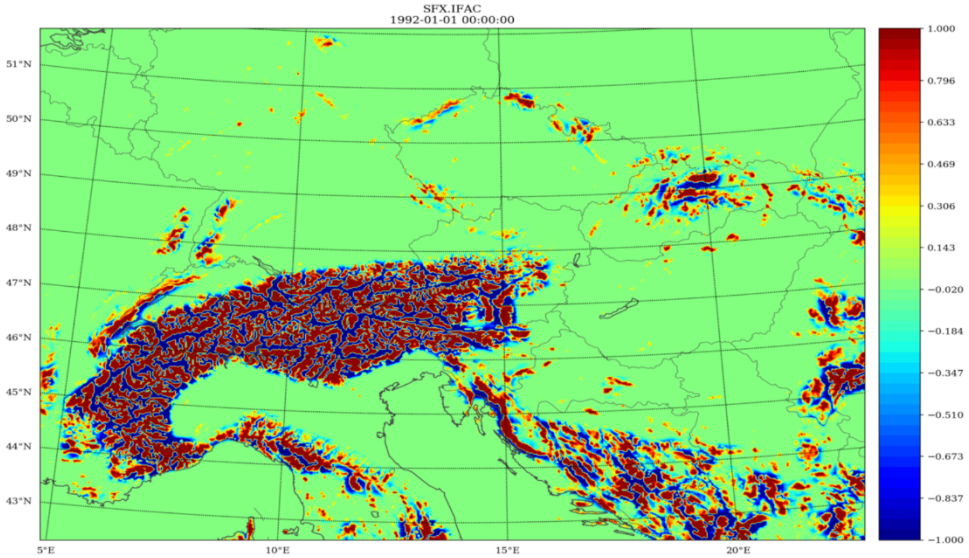


Figure 2: IFAC (inversion factor) for the AROME-Aut/C-LAEF domain with values varying between -1 (valley) to 1 (mountain tops) and values around 0 for flatland areas.

In the next step the IFAC is used to define the weight for each canopy level for the final composition of the extracted value for temperature and humidity. The principle behind is rather simple: Give more weight to the higher canopy levels or free atmosphere for exposed grid points on mountain tops and more weight for lower canopy levels for grid points located in narrow Alpine valleys. In between, in flatland areas the weighting is close to the default canopy diagnostics. To allow a smooth transition of weights quadratic polynomial functions are used (see Fig. 3). Technically, the adaption within the CANOPY scheme was implemented as a new namelist parameter (N2MTG) within SURFEX (coupling_isba_canopyn.F90), where different values (N2MTG = [1, 2, 3, 4, 5]) define different sets of weighting functions.

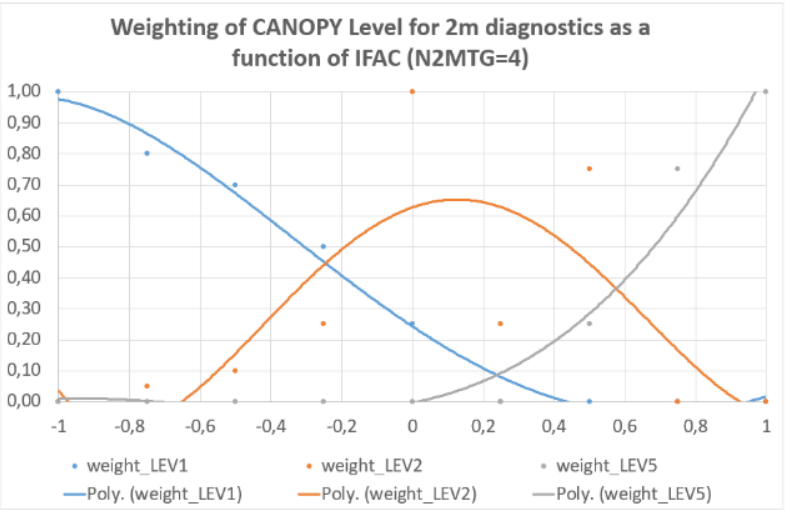
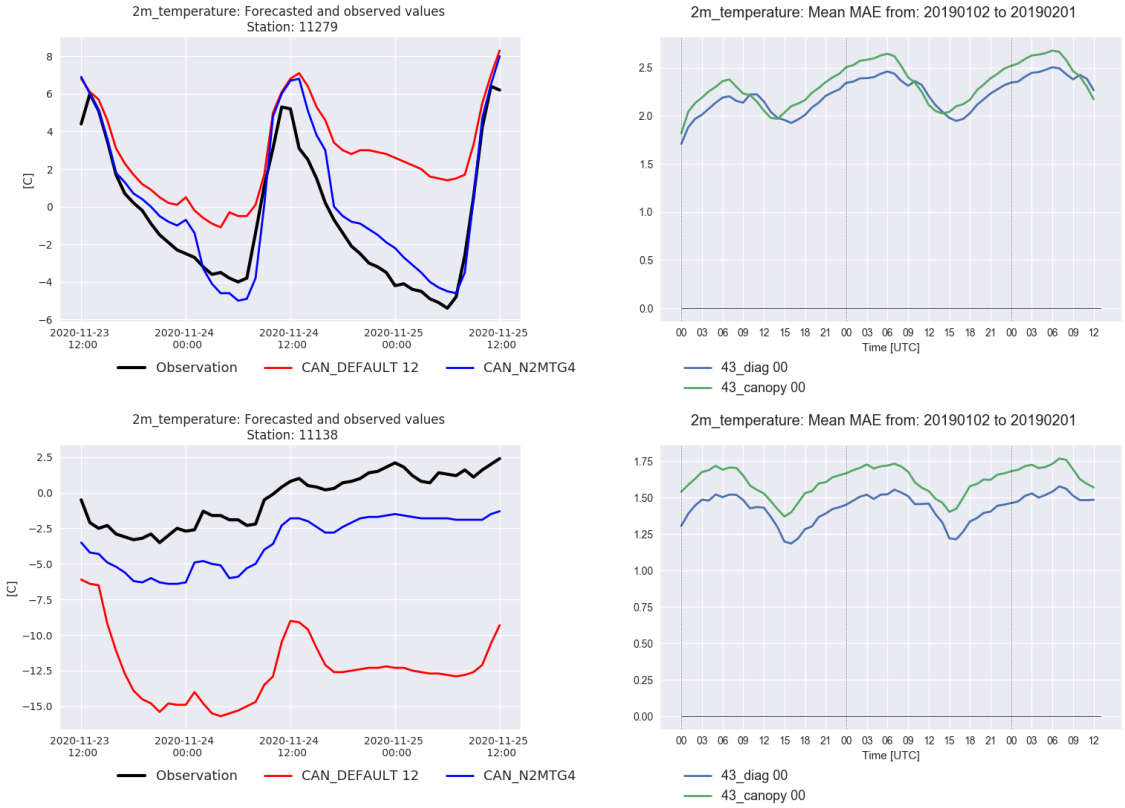


Figure 3: Polynomial weighting functions used in the modified screening level diagnostics within CANOPY. Curves describe the weighting for LEV1 (blue), LEV2 (orange) and LEV5 (grey) as a function of IFAC [-1, 1].

In order to demonstrate the effect of this modified diagnostics, the different versions are shown for a representative case (23rd November 2020, 12 UTC run). Figure 5 shows the impact of the modified diagnostics (labelled CAN_N2MTG4) in comparison to the default method used in CANOPY (CAN_DEFAULT). The figures in the left column include the forecasts and the observed 2m temperature for a station located in a valley (top), on a mountain top (middle) and in the flatlands (bottom). The right column finally shows the overall mean absolute error for the two versions averaged for all stations located in valleys (top), mountains (middle) and flatlands (bottom). The results of a longer verification period in January 2019 are shown in the left column Fig. 5 and Fig 6. It can be seen that the effect of the new diagnostic is as expected: It improves the performance for this case study in valleys and on mountain tops while it performs equally in flatland areas.
As already mentioned, this method can be also seen as some kind of bias correction placed within the model diagnostics to mitigate model deficiencies to resolve either shallow inversions in Alpine valleys and to ignore/damp misplaced inversion layers on exposed grid points.



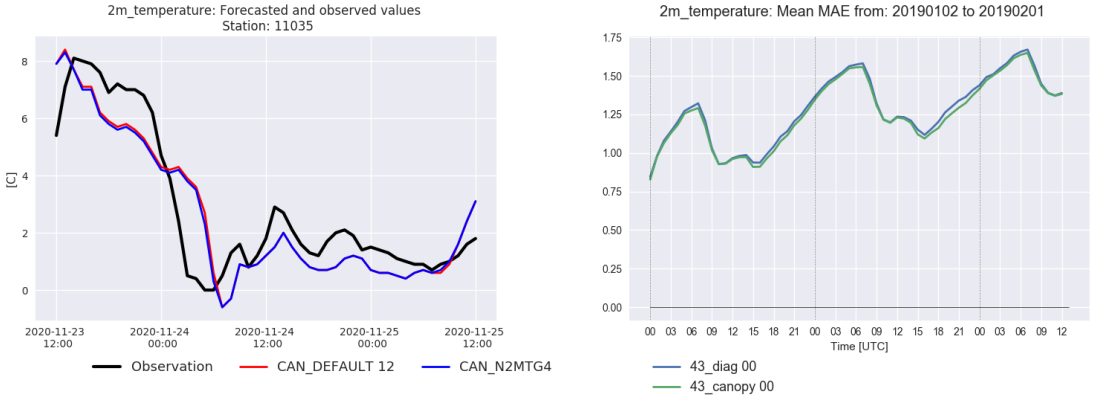


Figure 4: left column: 2m temperature forecast (blue = modified diagnostics, red = default canopy) for a station in a valley (top), on a mountain (middle) and in the flatlands for the case 20201123. Right column: 2m temperature MAE (light blue = modified diagnostics, green = canopy default) averaged over all station in valleys (top), on mountains (middle) and in flatland (bottom) for the period 20190102 - 20190201

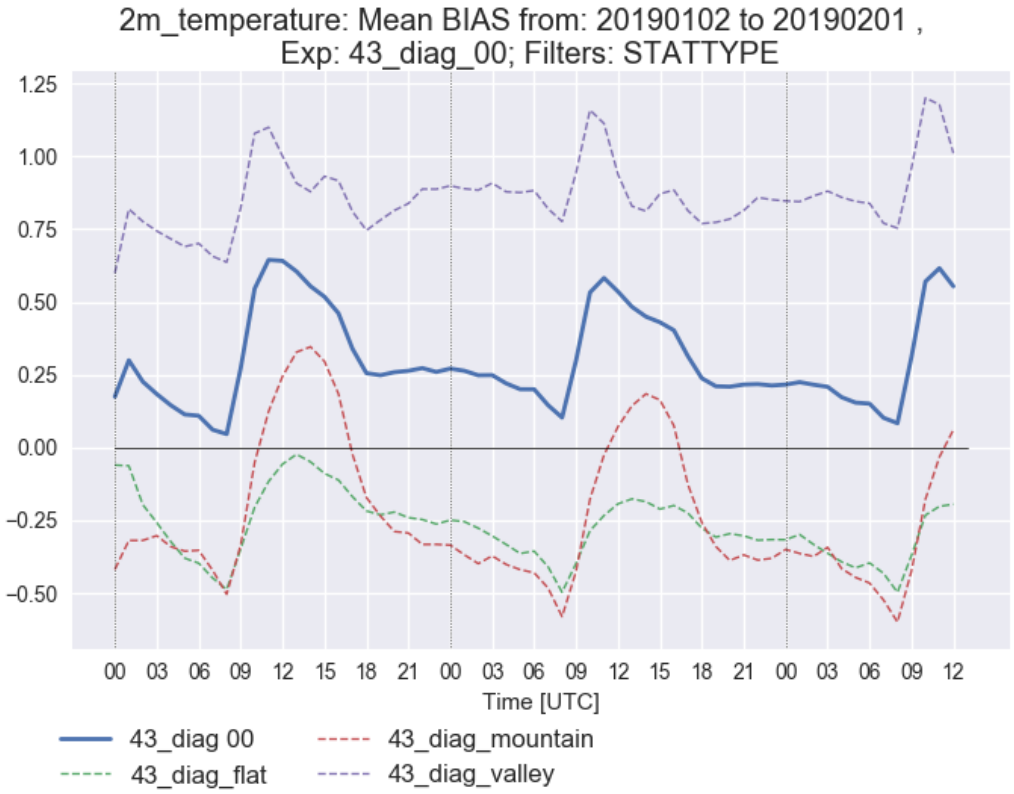
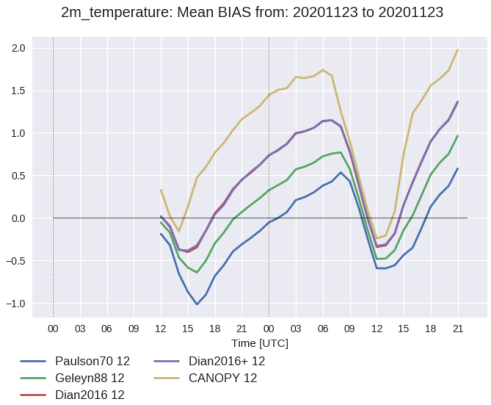
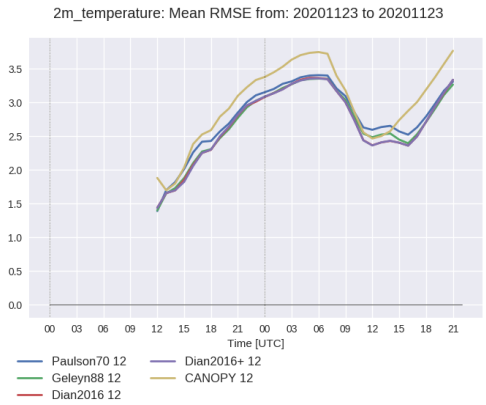


Figure 5: BIAS of 2m temperature (height corrected) for the period 20190102 – 20190201 of AROME-Aut using the proposed 2m diagnostic scheme for all Austrian stations (solid), and station selections according to station location (dashed).

3 Diagnostics without CANOPY scheme

As the diagnostics results with CANOPY scheme in the horizontally higher resolved AROME-RUC (1.2km) are also not satisfying, other options not using CANOPY were tested. CANOPY in current form is not ideal for steep slope conditions as it assumes gravity perpendicular to the CANOPY levels. AROME provides so far two more possibilities to calculate the screen level fields T2m, RH2m and 10m wind, if the CANOPY SBL scheme is switched off. The Paulson (N2M=1) option (Paulson, 1970) and Geleyn88 (N2M=2, Geleyn, 1988), where the former extrapolates from lowest model level to screen level, while the latter interpolates between surface and lowest model level using a vertical gradient assumption based on Monin-Obukhov theory and exchange coefficients. It was found that Geleyn88 is differently implemented in SURFEX routine surfex/SURFEX/cls_tq.F90 compared to the original formulation, which is completely used in the old inline 2L-ISBA scheme used in ALADIN/ALARO for long time: Instead of static energy, the thermal diagnostic equation is simplified taking into account a weighting between T at lowest model level and T at Surface only (assuming no heat capacity gradient). The more complex original formulation for static energy was implemented into SURFEX and compared to the standard SURFEX N2M=2 solution. However, for the cases considered no significant impact on the results was found by the simplified version. Dian and Masek propose another modification of Geleyn88 using a different assumption for the structure function ψ under stable conditions (Dian, 2016). This formulation was put to SURFEX as Option N2M=3 by adding a new routine cls_tq_dian.F90 and modification of the calls in diag_inline*F90 routines for sea, nature and lake. The diagnostics at TEB points does not use the Geleyn88 approach for 2m fields anyway, but takes the canyon values as 2m values. The free tuning parameters in Dian (2016) is set to a=1 hereafter. Finally, an additional modification of ψ to the Dian scheme was implemented for the unstable case (N2M=4, Dian2016+), where the tuning parameter is set to a=0.8 in unstable und a=1 in stable cases. This version reduces slightly the cold 2m bias in unstable cases. We consider the stable case of 23rd November 2020. AROME_RUC forecasts initated at 12 UTC were run with the different versions of diagnostics (Fig.6). The canopy scheme suffers of a strong positive bias especially during night and for stations in valleys. Paulson70 on the other hand shows strong negative bias during daytime especially for mountain stations. Geleyn88, Dian2016 and Dian2016+ are close where Dian2016 is warmer during nighttime as expected and has therefore lower bias at the beginning of the night, but higher bias (only due to valley stations, others have still negative bias) later than Geleyn88. Paulson70 performs best at valley stations, but worst at mountain stations in this case. Although all schemes have its strengths and weaknessess, the standard version of canopy diagnostics is outperformed by the other approaches considered.





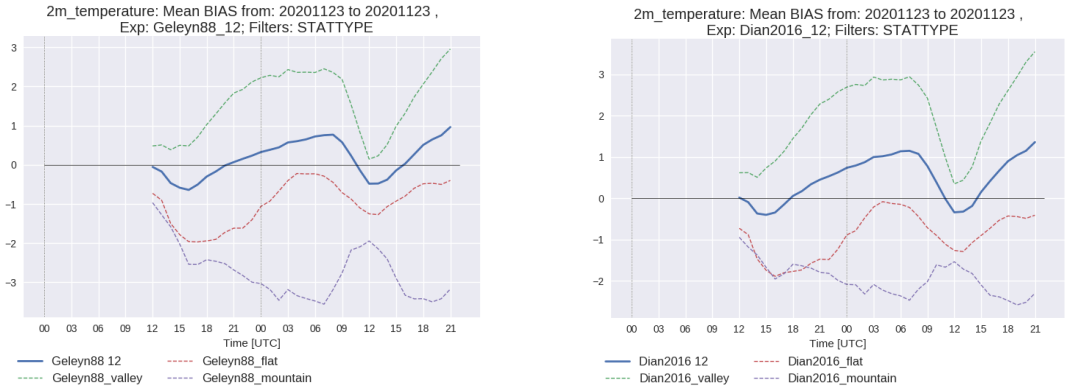


Figure 6: Bias (top left) and RMSE (top right) of T2m AROME-RUC 1.2 km forecasts for the 23rd November 2020 12UTC using different 2m diagnostics as described against Austrian surface stations. Bottom: Bias for different station types and Geleyn 1988 (left) and Dian 2016 (right)

4 Summary and Outlook

Evaluations of the forecast quality for screening level parameters, in particular for 2m temperature, revealed deficiences in particular for locations situated in Alpine valleys and on mountain tops. The error characteristics differ between stations/grid points located in valleys, mountains or flatlands. Major forecast errors regularly occur, in particular in winter time during situations when inversion layers form during nights with significant radiative cooling. These problems have a long history and so far the improvements when moving to higher model resolution were just marginal. So, beside looking for the source of the problem inside the model physics in connection to vertical and horizontal resolution, it is in a first step easier to modify the diagnostic scheme and try to achieve some improvements. Thus a modification of the CANOPY scheme, i.e. for the extraction of 2m temperature and 2m humidty was proposed an described in section 2. By implemeting a surface layer index (or inversion factor) based on the model orography, we have the possibilty to let the diagnostic act slightly differently for grid points in different topographical conditions. This is realized by applying different weights to the canopy levels which are finally used to dermine the screening level values. Results show that the impact of the modified CANOPY diagnostic is positive. The proposed method is currently running in cy43 based esuites for AROME-Aut and C-LAEF 2.5km at ZAMG and will probably enter the next operational suite in autmn. However, as this method is just a mitigation of model deficiencies the issue will be further investigated in the future during the preparation of the future C-LAEF / AROME-Aut 1km systems.In higher resolved AROME-RUC similar problems could be observed. The options without CANOPY scheme including also Dian (2016) approach were considered. Dian (2016) behaves as expected also within SURFEX and can reduce the negative bias of Geleyn (1988) and is assumed to be used in next operational version. However, this does not help for valley stations in cases as described in section 2, where all schemes show positive bias.

5 References

Dian, M. (2016): Improving the computation of screen level fields (temperature, moisture). LACE report. Available from https://www.rclace.eu/File/Physics/2016/dian_screeninter_pragueApr2016.pdf

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., & Gruber, C. (2011). The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region, *Weather and Forecasting*, *26*(2), 166-183

Geleyn J.F. (1988) Interpolation of wind, temperature, humidity values from model levels to the height of measurement. Tellus 40A, 347-351.

V. Masson and Seity, Y., 2009: Including Atmospheric Layers in Vegetation and Urban Offline Surface Schemes, J. Appl. Meteor.,48,1377-1397

Paulson, C. A. ,1970: The Mathematical Representation of Wind Speed and Temperature Profiles in the Unstable Atmospheric Surface Layer. J. Appl. Meteor. 9,857-861.

GMTED2010 global digital elevation model and subgrid-scale orography parametrization in ARPEGE: evaluation with the AROME fine-scale model

Sophie Marimbordes (CNRM/GMAP/PROC), Eric Bazile (CNRM/GMAP/PROC),
Yves Bouteloup (CNRM/GMAP/PROC), Didier Ricard (CNRM/GMME/PHY-NH)

August 25, 2021

1 Introduction

Orographic gravity waves result from the effect of the terrain on the atmospheric circulation and have a significant impact on the vertical transport of momentum. The ARPEGE global forecasting model is unable to characterize precisely these orographic effects. Its resolution, too low, induces a loss of information for scales smaller than the dimensions of the horizontal grid. Proper consideration of these sub-grid effects has a significant impact in the winter over the northern hemisphere, due to the presence of a disturbed flow. Their parametrization in global numerical forecasting models is therefore of major importance, in order to properly represent the general atmospheric circulation, as well as more local effects related to low-level turbulence.

Section 1 describes the general characteristics of orographic gravity waves and their parametrization in ARPEGE as well as the topographic data used in the models. Then, section 2 introduces the context and the method followed to perform the study of subgrid orographic effects. The results are presented in section 3. Finally, the conclusions and perspectives are detailed in section 4.

2 General information

2.1 Orographic gravity waves

Orographic gravity waves are waves which are considered to be stationary, induced by the action of the wind over terrain (figure 1). They are present on horizontal scales of between 5 and 100 km and are characterized by the size and shape of the terrain as well as by the properties of the flow. When an air mass undergoes orographic uplift in a stable stratified environment, it oscillates downstream of the mountain around an equilibrium height under the action of a restoring force, called buoyancy force. The generated orographic gravity waves have an oscillation frequency that corresponds to the Brunt-Väisälä frequency, expressed as a function of the gradient of the potential temperature of the flow. The presence of these waves is sometimes revealed by particular cloud structures, such as lenticular clouds over peaks or downstream of mountains, as well as parallel cloud bands in the form of rollers.

Orographic waves develop vertically and propagate a momentum flow along this axis. Thus, due to drag effect generated by the relief, these waves disturb the mean flow and modify the atmospheric momentum flow: this is called the orographic effect, also named orographic drag. Moreover, the amplitude of these waves tends to increase exponentially with altitude. At a given level, the associated disturbances are so important that the flow eventually becomes unstable. The waves are then dissipated, break up and give up momentum to the mean flow.

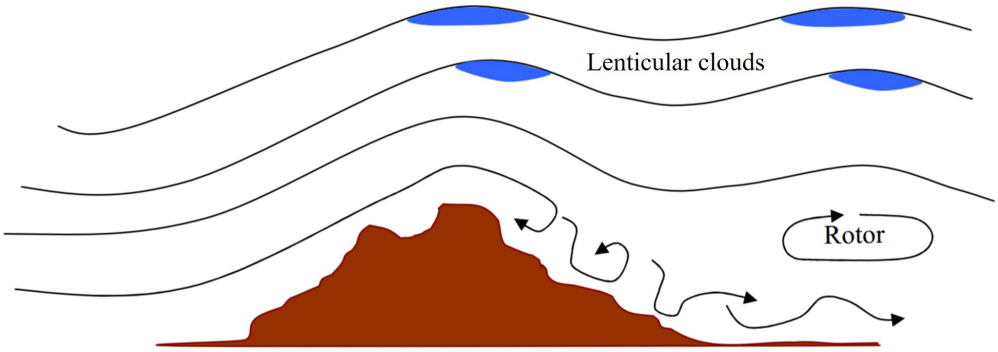


Figure 1: *Diagram illustrating the effect of landforms on flow.*

In the context of the linear theory, considering $\overline{u}$ the mean value of the zonal wind u over a sufficiently large domain, u' the deviation from this mean, w' the vertical component of the wind disturbance, and ρ the density, we can express the mean evolution equation under the effect of the wave disturbance as follows:

$$\frac{\partial \rho \overline{u}}{\partial t} = -\frac{\partial \rho \overline{u'w'}}{\partial z} \tag{1}$$

The equation 1, constructed using Reynolds axioms, states that the wind tendency is proportional to the vertical divergence of flux momentum.

2.2 Orographic drag parametrization

The parametrization, which is based on the equation 1, evaluates the momentum flux on the vertical axis generated by the unresolved orographic waves.

Through several equations (Lott 1994), the parametrization takes into account the dissipation (K_{GW}) as well as the resonance and reflection effects of gravity waves on the one hand, and the blocking (C_D) and lift effects of the flow induced by the relief on the other hand. The different coefficients K_{GW}, C_D and the effect of the lift acting on the parametrization are adjustable.

Another parametrization technique involves using an envelope relief in the model. It is supposed to represent the blocking effect of the highest peaks of a mountain range. This increases the orographic uplift and surface pressure drag. The method consists in increasing the mean relief by adding the standard deviation of the unresolved relief. Among global models, only ARPEGE still has this specificity, which we seek to remove and replace with a more adequate parametrization.

2.3 Global digital elevation model

The topographic data used in the models represent the surface elevation without considering vegetation or buildings. They are essentially constructed from satellite data (radar and lidar) and data acquired on the ground.

GTOPO30 (*Global Digital Elevation Model 30 arc seconds*) has a spatial resolution of 30 arc seconds of latitude and longitude, corresponding to a resolution of 1 km at the Equator. The data sources come from different continents. However, their quality is variable depending on the region of the globe. Major topographic features such as ridges and valleys are often poorly represented. Thanks to new satellite data sources, a new and more accurate terrain model, named GMTED2010 (*Global Multi-resolution Terrain Elevation Data 2010*) has been developed. Different spatial resolutions are available: 30 arc seconds (about 1 km), 15 arc seconds (about 500 m) and 7.5 arc seconds (about 250 m).

The AROME model uses topographic data from GMTED2010 with 250 m resolution while ARPEGE uses data from GTOPO30.

3 Methodology

3.1 Context

The study consists in analyzing the subgrid orographic effects over the Rocky Mountains and the Great Plains region of the United States. The effects of this mountain range on atmospheric circulation have an impact on the quality of forecasts over Europe. This region also benefits from numerous radiosonde data, allowing a better evaluation of the results. The analysis is carried out for a period spanning from December 1 to December 31 2020, which was characterized by a disturbed flow over the domain (figure 2).

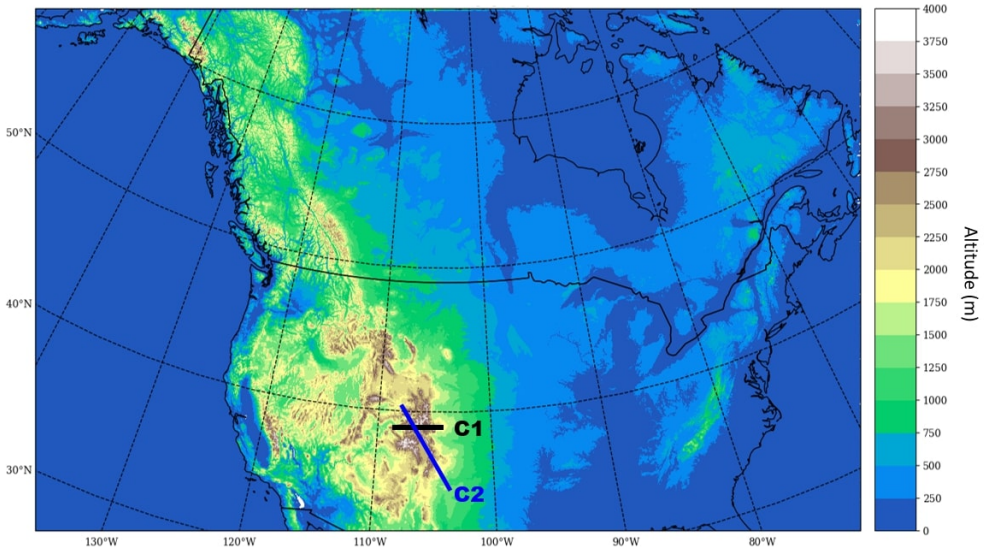


Figure 2: *North America domain topography (m).*
C1 and C2 refer to vertical sections used in the results.

3.2 Tools

The fine-scale AROME model (Seity et al. 2008) is used for the study. The horizontal resolution of the AROME model is taken at 2.5 km to reduce the computational cost, the domain defined over North America being quite large (about 6000 km by 3000 km). Additionally, 105 vertical levels, the same as ARPEGE, are taken into account. The ARPEGE model has a spatial resolution of about 10 km over the region. The parametrization tests are carried out using the ALADIN model which has the same dynamic core and physical schemes as ARPEGE. For clarity, ALADIN simulations are named ALADIN/ARPEGE in the following sections. As AROME and ALADIN/ARPEGE are limited area models, they are coupled every three hours to the version of the ARPEGE model that will be operational in 2022. Moreover, data assimilation is not performed for our tests. The analysis is thus the coupling model. Finally, the orographic effects are studied on forecasts up to +48 h.

3.3 Method

First, the study aims to characterize the impact of GMTED2010 topographic data compared to GTOPO30 in ARPEGE as well as the effect of envelope relief.

Next, orographic effects are studied on a particular day in the Colorado mountains. The large-scale ARPEGE model, which has a spatial resolution of about 10 km over the domain, is compared to the AROME model with a 2.5 km resolution in both hydrostatic (H) and non-hydrostatic (NH) modes.

Finally, the study consists in carrying out 2.5 km AROME simulations with their own relief, then with the 10 km ARPEGE relief, in H and NH mode, including the microphysics and shallow convection schemes of AROME and ARPEGE. These characterize the effect of subgrid-scale orography in both modes and for both physics by computing the difference of the simulation with 2.5 km relief with the equivalent simulation using the smoothed 10 km relief . The field considered for this study is the mean kinetic energy tendency because this physical quantity gives us information on the wind speed. These reference AROME simulations allow us to evaluate the subgrid-scale orography parametrization in ALADIN/ARPEGE. The impact of topographic data, envelope relief, K_{GW} and C_D coefficients and lift effect in the orographic drag parametrization on the mean kinetic energy tendency (December 2020, North America domain) is precisely studied. In order to evaluate the new parametrization settings with respect to the operational ALADIN/ARPEGE configuration, wind force field scores are calculated with respect to radiosonde data (up to +48 h).

3.4 AROME over the North America domain

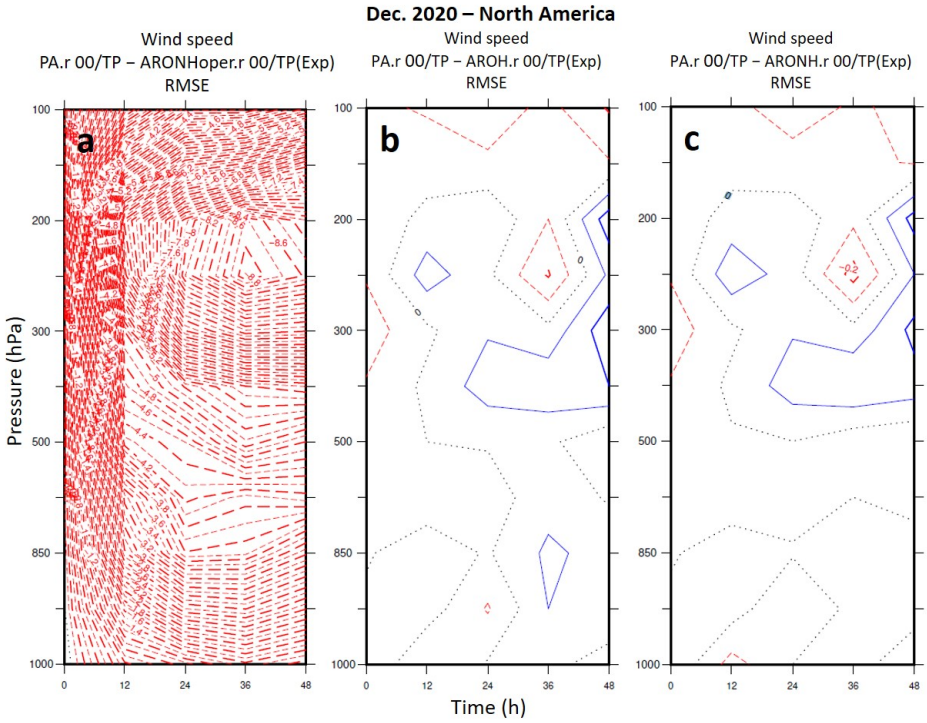


Figure 3: *Comparison of root mean square deviations computed from radiosonde data (TP), between operational ARPEGE (PA) and an AROME experiment. (a): AROME NH operational version, (b): AROME H and (c): AROME NH after dynamics modification.*

In the implementation of the AROME NH configuration over the North America domain, poor forecasts were found with operational settings and 2.5 km resolution (AROME-NH-oper-2.5km). Computations of root mean square deviation from radiosonde data for the wind force field over the month of December and the whole domain show a very clear deterioration of the AROME-NH-oper-2.5km model compared to the ARPEGE model (figure 3 (a)). This deterioration is largely attenuated for a 2.5km AROME configuration in hydrostatic mode (figure 3 (b)). Three modifications in the model dynamics have been retained by the CNRM/GMAP ALGO team to configure a new AROME NH at 2.5 km in order to improve the numerical stability of the model

and to have better predictions over the domain (figure 3 (c)):

- decrease of the reference cold temperature from 100 K to 70 K;
- deactivation of the predictor-corrector scheme and time extrapolation necessary again;
- decrease of the time step from 50 s to 30 s.

4 Results

4.1 Topographic data and envelope relief

This section is dedicated to the evaluation of GMTED2010 and GTOPO30 topographic data and envelope relief.

The vertical sections (figure 4) allow visualization of the topography across the Colorado mountains (black section C1 figure 2). The black curve represents the reference relief (2.5 km resolution, GMTED2010 topographic data). The red and green curves describe the relief with 10 km resolution, using topographic data from GTOPO30 and GMTED2010. These curves seem to be almost identical: they illustrate the fact that the GTOPO30 and GMTED2010 data are similar over the study area. The blue curve represents the effect of the envelope on the relief, allowing to simulate the highest peaks that we visualize thanks to the black curve. It is comparable to the red curve with a constant elevation that is associated with the addition of standard deviation of subgrid-scale orography on the mean orography. Computations of scores comparing a simulation with GTOPO30 terrain data (without envelope relief) with a simulation using GMTED2010 data show a slight but significant improvement in near-surface wind forecasts (strength and direction) up to +48 h when the GMTED2010 topographic data is used (reduction of bias, standard deviation and root mean square deviation). We also note that the effect is globally neutral at altitude. For the other atmospheric parameters such as temperature, humidity or pressure, the results are similar to those observed for the wind.

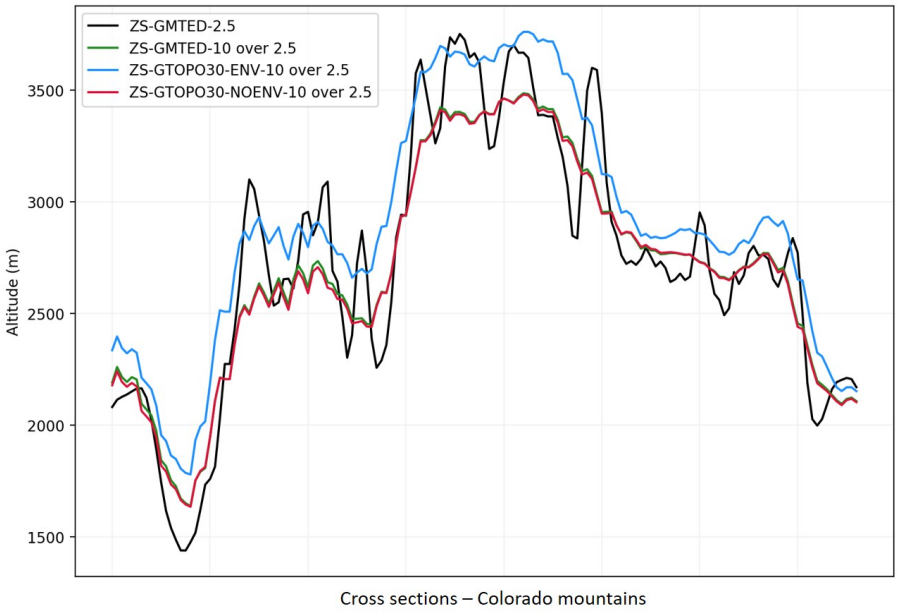


Figure 4: *Cross sections (C1 figure 2) of topography (m) over the Colorado region.*

4.2 Non-hydrostatic effects

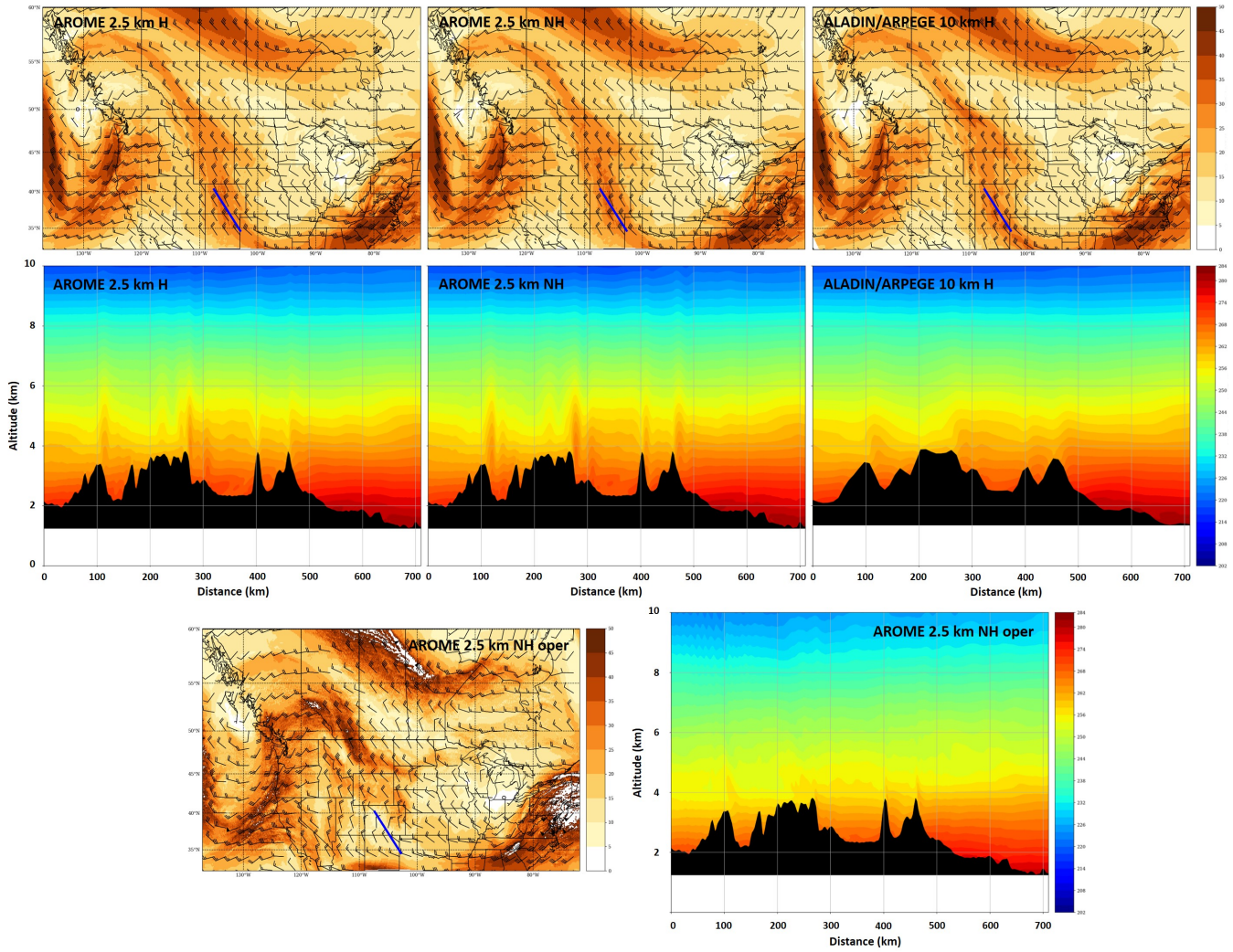


Figure 5: *First row: wind direction and wind speed (m/s) over the North America domain. In blue, C2 cross section (figure 2). Left: AROME H; center: AROME NH; right: ALADIN/ARPEGE. Day of 12/17/2020 at 00:00 (run 15/12/2020 00:00, +48 h).*
Second row: vertical cross-sections (C2 figure 2) of temperature (K) across the Colorado mountains. Left: AROME H; center: AROME NH; right: ALADIN/ARPEGE. Day of 12/17/2020 at 00:00 (run 15/12/2020 00:00, +48 h).
Third row: wind direction and wind speed (m/s) over the North America domain (left); vertical cross-sections (C2 figure 2) of temperature (K) across the Colorado mountains (right). Day of 12/17/2020 at 00:00 (run 15/12/2020 00:00, +48 h).

This section focuses on the non-hydrostatic effects on the representation of orographic waves for the day of 17/12/2020 at 00 h (forecast +48 h) over the Colorado region (C2 blue section figure 2). The flow over the area on that day is strong (30 m/s at 500 hPa, figure 5) and is perpendicular to the axis formed by the high peaks of the mountain range. Vertical temperature sections across the Colorado mountains (figure 5) identify the presence of waves downstream of highest peaks. The NH mode provides a better representation of orographic waves (hollows and humps) compared to the H mode. The ALADIN/ARPEGE simulation indicates that the horizontal resolution plays a major role in the representation of the relief: peaks and valleys are perceived inaccurately by the large-scale model. The orographic waves are then less well detected on the temperature field, which confirms the need to parametrize the unresolved effects in ALADIN/ARPEGE.

The wind forecasts for the AROME-NH-oper-2.5km model (third row, figure 5) are different from those obtained with the AROME H and NH and ALADIN/ARPEGE models (first row, figure 5). In the Colorado region, the wind speed reaches only 15 m/s with AROME-NH-oper-2.5km compared to about 30 m/s for the other models. Additionally, the vertical temperature section (third row, figure 5) shows numerical anomalies at high altitude. These results confirm the poor forecasts found with operational settings and 2.5 km resolution (figure 3).

4.3 Subgrid-scale orographic effect with AROME

This section aims to characterize the orographic effect under 10 km resolution identified thanks to AROME fine-scale model.

The subgrid-scale orographic effect is obtained by differencing an AROME simulation containing the 2.5 km relief with the equivalent simulation using the 10 km smoothed relief. The green kinetic energy tendency curve in Figure 6 characterizes the average subgrid-scale orographic effect for the North America domain over the month of December 2020 detected by AROME NH. The high resolution relief (2.5 km) leads to a higher energy loss than a smoothed relief (10 km), which implies a negative kinetic energy tendency.

Conducting the same experiment with AROME H yields a curve similar to the green curve. Similarly, when simulations are run with ARPEGE physics in H and NH mode, the results are similar (not shown). At 2.5 km, only the microphysics and shallow convection schemes differ in the physics of AROME and ARPEGE models.

The orographic effect below the 10 km grid detected by AROME is therefore not influenced by the mode (H/NH) and the physics of the model (AROME/ARPEGE).

4.4 Parametrization evaluation in ALADIN/ARPEGE

This last section presents a new configuration of the orographic drag parametrization in ALADIN/ARPEGE.

Figure 6 shows vertical profiles of mean kinetic energy tendency (North America domain, December 2020, D+2 forecast). The green curve represents the orographic effect under the 10 km grid detected with AROME; the black curve, the effect of the current orographic drag parametrization in ARPEGE; the blue and red curves, the effect of the parametrization taking into account new settings for the GTOPO30 (blue) and GMTED2010 (red) terrain data. ALD-NOSSO represents an ALADIN/ARPEGE configuration for which the orographic drag parametrization is disabled.

We first notice that the black and green curves have a similar evolution. Although they have the same dynamics, the black curve moves significantly away from the green curve between 700 and 250 hPa. The current parametrization, for which the envelope relief plays a crucial part (not shown), does not fully represent all of the subgrid-scale orographic effects, considering AROME simulation as our reference.

After performing various sensitivity tests, the ALADIN/ARPEGE configuration with no envelope, activation of the lift, K_{GW} multiplied by 6 and C_D multiplied by 2 is the one that best reproduces the effect of the current parametrization, with the blue and red curves close to the black curve.

Root mean square deviation calculations relative to the wind force field, computed with respect to radiosonde data up to +48 h, suggest that these new settings are encouraging for no longer using the envelope in ARPEGE, although they are not completely optimal (figure 6). The scores also show that these new settings are better suited to the GTOPO30 topographic data at altitude. Nevertheless, close to the surface, the deterioration of wind forecasts is less marked for the simulation that uses GMTED2010 data. If GMTED2010 data are eventually used in ARPEGE, it would be relevant to calculate validation scores over a longer period (up to +96 h), and then to consider adjusting the different coefficients of the parametrization taking into account this new digital elevation model.

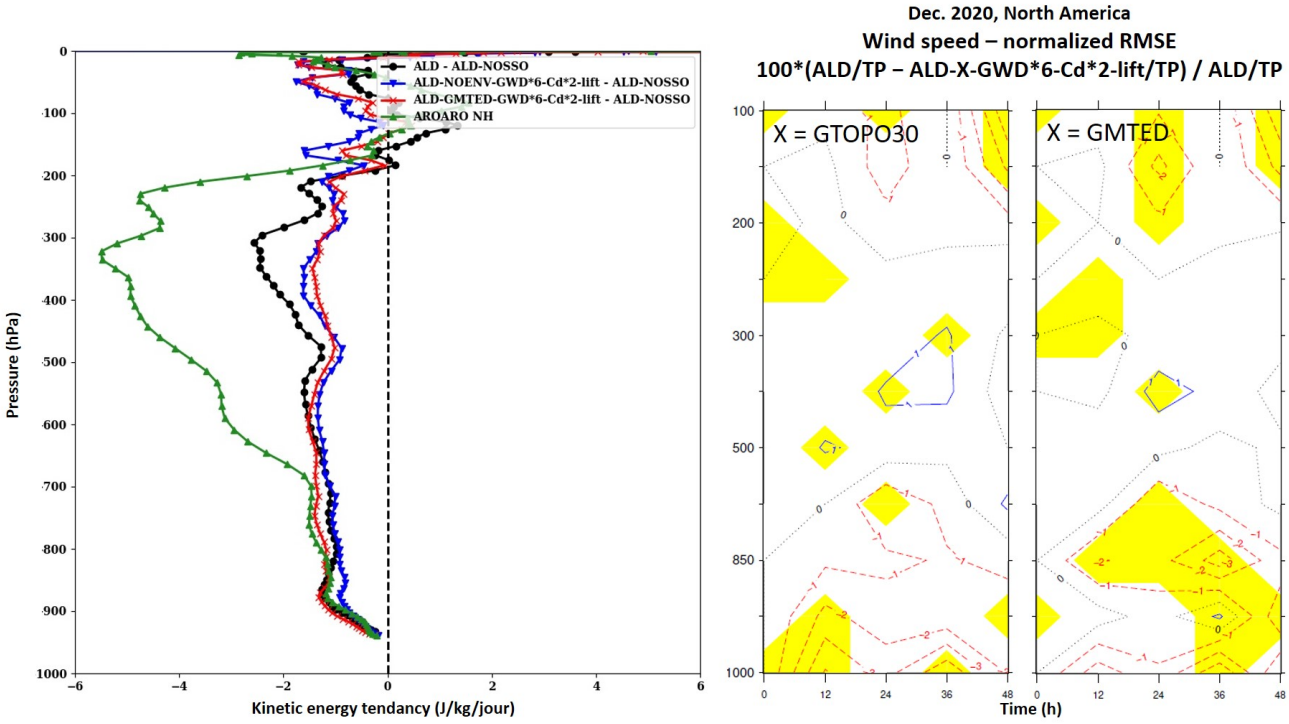


Figure 6: *Left: vertical profiles of mean kinetic energy tendency (J/kg/day), December 2020, North America, D+2. Sensitivity tests on the C_D coefficient with K_{GW} *6, for simulations with no envelope relief.Right: normalized root mean square deviations (RMSE), computed with respect to radiosonde data (TP) up to +48 h, to evaluate the ALD-NOENV-GWD*6-Cd*X simulations (with X = 2 or 3) with respect to ALD (ALADIN/ARPEGE). The field considered is the wind strength, December 2020, North America domain. The color red corresponds to a degradation of the wind force score of ALD-NOENV-GWD*6-Cd*X compared to ALD and the color blue, an improvement. The yellow areas indicate that the trend (improvement or degradation) is significant at 95%.*

5 Conclusion

The resolution of the ARPEGE large-scale model, which is too low (about 10 km over the North America domain), does not allow all orographic processes to be described explicitly. This is why it is necessary to parametrize them. The fine-resolution, non-hydrostatic AROME model has provided a better understanding of the orographic effects below the 10 km grid. New settings for the orographic drag parametrization in ARPEGE using GMTED2010 and GTOPO30 data are suggested in this study. It still seems delicate to think about removing the envelope relief in the large-scale model without degrading forecasts. The IFS model parametrization uses a form drag to represent turbulent processes for scales smaller than 5 km. One line of research would be to implement this parametrization in ARPEGE. It would also be interesting to integrate GMTED2010 terrain data into ARPEGE, as it contributes slightly to the improvement of forecasts.

References

F. Lott. The significance of subgrid-scale orography and problems in their representation in gcms. *Seminar on parametrization of sub-grid scale physical processes, ECMWF*, pages 277–304, 1994.

Y. Seity, P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. The arome-france convective-scale operational model. *Journal Of American Meteorological Society*, 139:976–991, 2008. doi: 10.1175/2010MWR3425.1.

Study of AROME Temperature in mountains regions

Gabriel Arnould and Danaé Préaux

These studies were done during our master thesis. We were supervised by I. Etchevers, I. Gouttevin and Y. Seity. We acknowledge L. Auger, E. Bazile, R. El Khatib, P. Marquet, J.-M. Piriou for their advices. D. Préaux will continue this work with a thesis in May 2022. It is entitled "Towards a sub-kilometer Arome for mountains meteorology and hazards modelling" and supervised by I. Gouttevin (CNRM/Snow Research Center) and I. Etchevers (CNRM/GMAP)

1 Introduction

Weather forecasting is a challenging task in mountainous areas because of large topographic variability that leads to e.g. very high temperature and precipitation gradients. Surface heating is not uniform due to slopes orientations, narrow valleys can be responsible for cold and moist air accumulation or local winds, giving raise to small-scale atmospheric conditions. These local processes deeply influence the synoptic circulations as well as energy and mass balance over mountains, so that they can play a key role in the understanding and forecast of hazards like low-altitude snowfalls, avalanches and floods (figure 1).

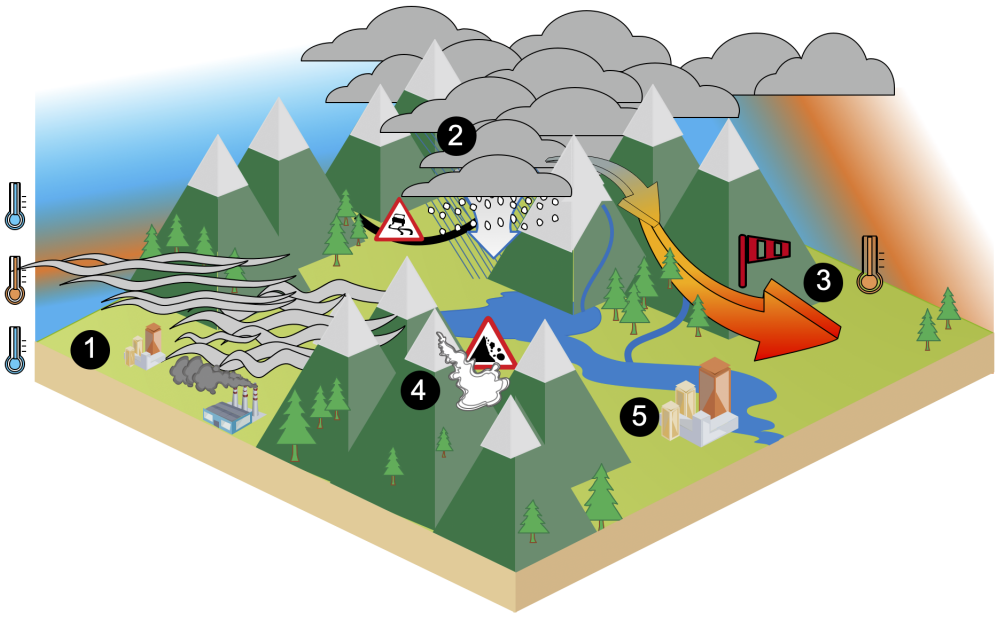


Figure 1: *Schematic representation of mountain hazards related to small-scale atmospheric conditions: (1) air pollution below temperature inversion, (2) low-altitude snowfall with isothermal conditions, (3) foehn effect, (4) avalanche, (5) flood due to snow melting.*

With its low horizontal grid spacing of 1.3 km, the Arome numerical model is an essential tool in mountain meteorology. However, its skills at temperature forecasting are altered by several temperature biases. These biases are explained in the first part of this article.

In a second part, we investigate a first hypothesis for the temperature biases in AROME, incriminating the basic grid interpolation of the post-processing system of Arome fields, Fullpos. In Fullpos, the transition from the original model grid to a regular grid proceeds via a a simple height correction on temperature fields. It is based on a constant lapse rate that is not compatible with atmospheric properties in complex terrain. Here, we investigate a more sophisticated approach using downscaling, and detail its impact on the Arome biases. The conclusions show that the error is poorly correlated with Fullpos treatments. Finally, in a third part, sensitivity tests in the model it-self are carried out in a challenging situation, which helps understanding the model's intrinsic limitations and perspectives to overcome them.

This article ends with conclusions and perspectives on the temperature biases in AROME.

2 Arome temperature biases in the mountains

Three features mainly characterize the Arome temperature error in mountain regions: (1) a cold bias at high altitude, (2) a low-altitude warm bias occurring in stably stratified layers and (3) a warm bias during snowfall situations. All of them are described in the following three subsections.

2.1 Cold bias in high mountain

First of all, a cold bias increasing with altitude has been detected by Vionnet et al. (2016). Temperature data collected over 4 years (2010-2014) show a significant difference between the Arome 2.5 km model and the observations. This difference of less than 1°C below 1500 m reaches **3°C** at night between 1500 and 2500 m altitude. Above this, the bias can **exceed 4°C** in winter, as shown in figure 2. Moreover, the bias is seasonal, being more important in winter, when the snow cover dominates at high altitude, than in summer (Vionnet et al. 2016). It thus impacts the evolution of the snowpack which is particularly sensitive to thermal gradients, with consequences for the avalanche hazard forecast (Gouttevin et al. in prep). In addition, this cold bias delays the snow-melt and therefore limits the potential of the Arome model in support of flood forecasting.

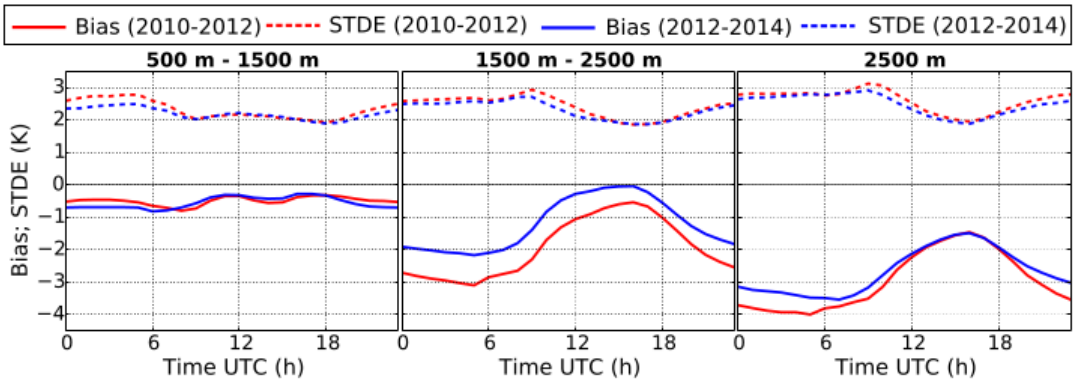


Figure 2: *Bias and standard deviation of the error (in K) for three altitude bands for Arome forecasts of 2 m temperature in the Alps. Two periods are considered: 2010-2012 (red) and 2012-2014 (blue). According to Vionnet et al. (2016)*

Finally, this cold bias induces an over-accumulation of snow at high altitude, which calls into question the capacity of the regional climate model based on Arome, Arome-Climat, to adequately represent climate change in the Alps (Monteiro 2020).

2.2 Warm bias in stable conditions

A second bias, a warm bias in valleys, appears during long anticyclonic episodes. It was highlighted during the 2015 observational campaign held in Passy, in the Arve Valley (Haute-Savoie). The study of this valley is essential since it is particularly polluted by fine suspended particles with a diameter of less than 10 µm (PM10). This campaign was the first to be carried out in a French Alpine valley in stable and persistent winter conditions, that are particularly prone to intense pollution events. This campaign revealed that the warm biases of the model during such events, hinders the forecasting and representation of these pollution events. Besides, temperature is a meteorological parameter of interest in the construction of a winter pollution risk indicator (Paci et al. 2016). The quality of its estimation is therefore a relevant issue for the forecasting of such events.

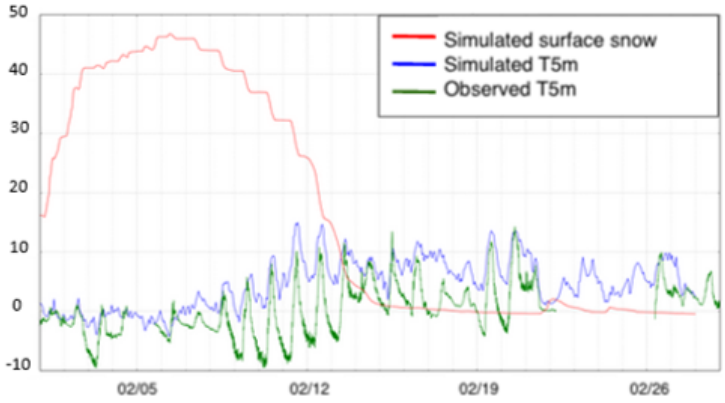


Figure 3: *5 m temperature observed and modelled by Arome 2.5 km during the campaign in Passy from 5 to 26 February 2015, and Arome snow cover (cm). According to Paci et al. (2016)*

2.3 Warm bias during snowfall situations

A third bias lies in the Arome temperature forecasts. This is, again, a warm bias. It occurs in valleys when the warm front of a disturbance meets the Alpine relief, and induces a too fast rise of **snowfall line** in the model. This can lead to a wrong qualification of the type of precipitation and to an underestimation of the quantity of snow precipitation in valleys, where the major roads are. As a matter of fact, over the whole period from December 2019 to January 2021, the snowfall line modelling proved to be **erroneous for at least 11 snowy episodes**, according to the operational forecasting services of the French Alpine region. The problems described are a snowfall line that is too high or rises too quickly, especially during isothermal episodes that were not captured by Arome. **In 6** of these situations, the Cosmo model is described as **much better** by the forecasters. This warm bias is not new. Beauvais (2018) describes three snow events with a snowfall line problem during the winter of 2017-2018 and even mentions a similar situation dating back to 2009. This study concluded that the resistance of cold air and isothermal conditions in valleys under warm fronts is underestimated by Arome

3 Air Temperature : downscaling and altitude correction

A first hypotheses was that the biases previously described originate from the post-processing of temperature fields. Coarse corrections are encoded in Fullpos for the transition from the original grid to the regular one in latitude/longitude format. A bilinear interpolation is applied to the temperature with a simple height correction based on standard atmosphere temperature profile. In order to test this hypothesis, a study was conducted in 2020 with two purposes: (1) use more sophisticated methods to project fields in regular grids within Fullpos and (2) make the projection on a 500-m resolution grid instead of 1.3 km to assess the impact of mesh size on AROME biases.

3.1 Methods

3.1.1 Downscaling experiments

The study only focused on 2 meters temperature (T2m) for which two kinds of downscaling methods have been developed. The first one is called "corrective" since the method directly acts on the 1.3-km original grid, applying corrections that take into account physical constraints on T2m such as small scale topographic variability. Three variants have been tested:

- a bilinear interpolation associated with a height correction based on a constant lapse rate of -6.5 °C/km corresponding to the standard atmosphere temperature profile. This is the approach currently implemented in Fullpos.
- a nearest-neighborhood interpolation associated with a height correction based on a constant lapse rate of -6.5 °C/km.
- a nearest-neighborhood interpolation associated with an "environmental lapse rate" as defined by Sheridan et al. (2010). A value of $\frac{dT}{dz}$ is computed in each point of the finer grid thanks to a linear regression over $T(z)$ data associated with the surrounded points in the coarse grid. In this study, the 16 nearest grid points were selected for the regression.

The second downscaling approach is dynamic and allowed by the high resolution model Arome 500 m currently in development and mainly used for research applications.

3.1.2 Domain and scores

The four downscaling approaches were evaluated during the 2019-2020 winter period (November to March included). As inputs, the corrective methods take the first twenty-four hours related to the run of 00TU of operational Arome 1.3 km (cycle 46T1). The observational dataset used as reference comes from stations located in the study domain. This one encompasses a large part of French and Italian Alps (figure 4). The comparison between observations and predictions is quantified by an hourly mean bias defined as follow:

$$Bias = \frac{1}{N}\sum_{n=1}^{N}(X_n - X_{obs})$$

where N is the total number of stations and days during the studied period. To avoid any competing effects on corrective downscaling outputs, no further altitudinal correction is applied on T2m to account for the possible altitude difference between the observation and the 500-m grid cell altitude. To compensate, only stations whose altitude difference is less than 150 meters with the nearest point on the 500-m grid are conserved. This option has already been tested in previous works (Vionnet et al. 2016).

3.2 Results and discussions

Figure 5 displays the diurnal cycle of mean bias calculated for the four downscaling simulations over the studied period. The stations used to compute the bias are split into three altitude bands. Corrective methods have similar performances, except a small difference in the lower band (altitude < 1000 m) where the two approaches based on nearest neighborhood interpolation are closer to the observations. Among them, the method of Sheridan et al. (2010) performs better, especially during night time when the air is colder and temperature inversions more prevalent. Computing an environmental lapse rate in each grid point, this technique indeed allows to deal with complex profiles of temperature which are very common in mountainous areas. The dynamical downscaling

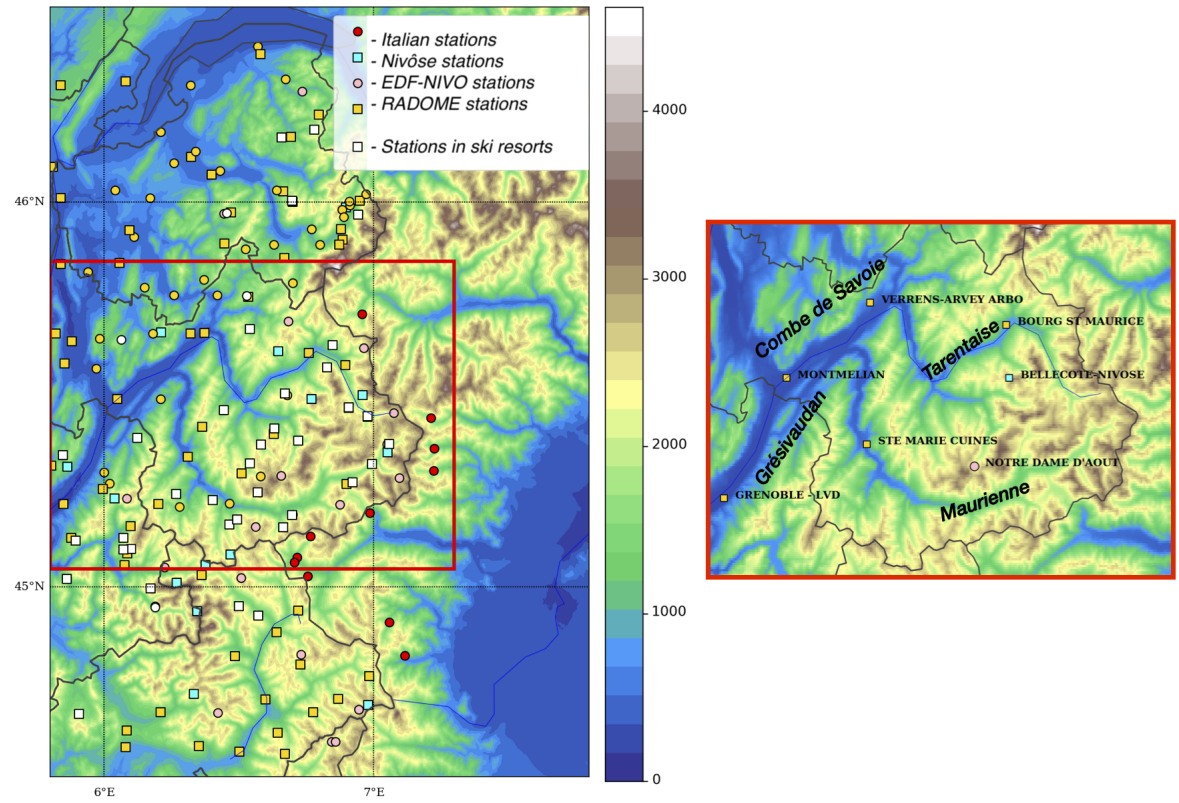


Figure 4: *Relief of the studied domain (in meters) corresponding to the 500 m resolution grid. Dots indicate the localisation of the stations. Their origin is specified into the legend box. The sub-domain introduced in the next subsection is framed in red and displayed on the right with the main valleys and a well-selected set of stations.*

has the smallest error in lower altitudes. However, its negative bias is increased in higher altitudes, being nearly 0.5°C colder than corrective methods beyond 1800 m.

On the other hand, the biases described in part 1 are persistent. Despite the initially assumed benefits of downscaling, the four simulations still overestimate the temperature in lower altitudes, especially at night when the bias exceed 0.5°C. The prediction is even more warmer when compared with observations recorded at the bottom of the valleys (not shown). In contrast, the cold bias remains in higher altitudes, reaching 1°C during the afternoon to 3°C and more at night and early morning. The underestimation is even worse with Arome 500 m. These conclusions are clearly illustrated by the situation of the 31th of December 2019 when a large temperature inversion was observed in Savoie. On figure 6 at 6TU, the inversion is pronounced within the valleys of Isere (*Grésivaudan* and *Combe de Savoie*). The inversion simulated by Arome 500 m is more realistic as it better matches the orography. The overestimated prediction in Ste-Marie-Cuines is probably due to a coarser model orography in this narrower portion of the Maurienne valley. Nevertheless, the Arome anticyclonic warm bias remains in valleys of all sizes (nearly +1°C in Montmelian and Grenoble-LVD). The corrective downscalings keep the initial error of Arome 1.3 km, as suggested by figure 7 concerning the more sophisticated tested method inspired by Sheridan et al. (2010). The graphics indicate the local lapse rate computed in the environment of three stations from the 16 nearest points (red or blue dots) in the Arome 1.3 km grid. Graphic **A** exhibits the drawbacks of the technique as it computes a local lapse rate from a linear regression which is not relevant in a hook temperature profile. In graphics **B** and **C**, the regression is more consistant (high r-squared) but the gap between model data and the observation (black dot), resulting from the warm bias, deeply limits the downscaling potential. On this day, the cold bias in higher altitude roughly ranges from 5 (in Notre Dame d'Août) to 8°C (in Bellecôte-Nivôse). Arome 500 m worsens the error found in the operational model.

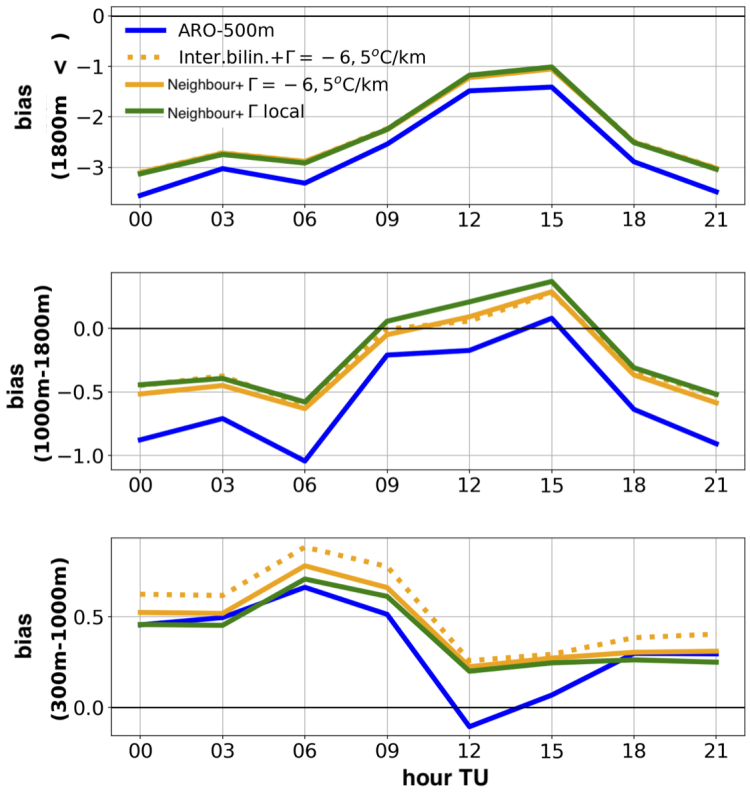


Figure 5: *Diurnal cycle of T2m bias (in °C) related to the four downscaling simulations. The stations used for the comparison are split in three altitude bands.*

3.3 Conclusion

Even doubling the resolution, the modifications tested in Fullpos do not correct the Arome surface temperature biases. Moreover, the cold bias in high altitude is increased with Arome 500 m. The continuing counter-performances confirm another source of error. If unbiased, corrective downscaling methods have a real potential in enhancing the model forecast while limiting time-consuming simulations. A deeper understanding of Arome biases is hence necessary to address fine-scale prediction in mountainous area. A study in this direction is developed in the next section.

4 Air Temperature : sensitivity tests

4.1 Methods

In this section, sensitivity tests were done to investigate the origin (or origins) of the warm bias in valley during snowfall situations. In other words, some physical, assimilation and dynamic modifications were successively activated in Arome 1.3 km and their performance was compared to the operational version. The situation of the 12th of January 2021 in Haute-Savoie is a typical case of this bias. On the night of 11 to 12 January, the air is cold and stable in the lower layers of Haute-Savoie. In the morning, a pseudo-warm front reaches the Arve valley, with snowflakes seen first at 10 a.m. The snow lies on the ground in all the area, whatever the altitude. Concerning road trafic assistance in winter, the temperature at 2 m is close to 0°C in the valley with -2.2°C at 2 p.m. However Arome is too warm, leading to a wrong forecast for the snow type and altitude of the snowline.

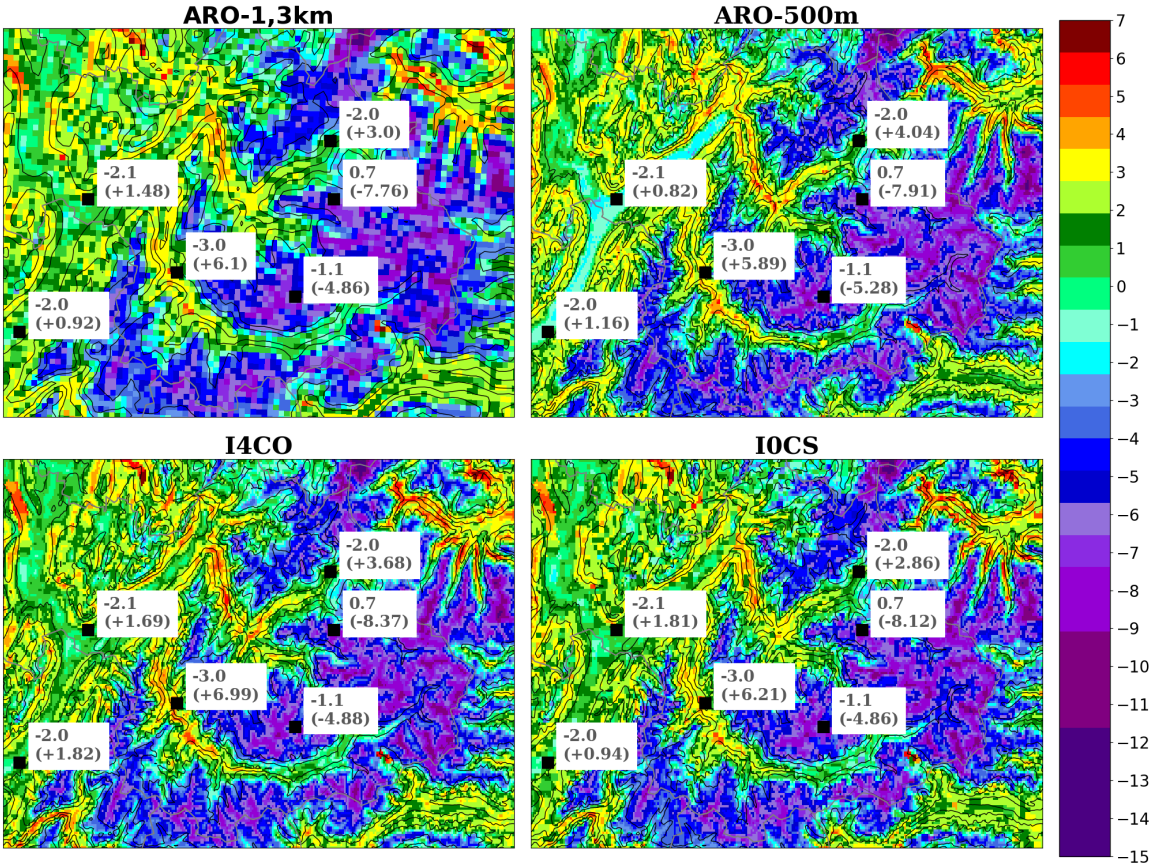


Figure 6: *T2m (in °C) predicted in Savoie (red sub-domain on figure 4) by Arome 1.3 (run 00TU), Arome 500 m, the corrective downscaling based on bilinear interpolation and constant lapse rate (I4C0) and the corrective downscaling based on nearest neighborhood interpolation and local lapse rate (I0CS), the 31th of December 2019 at 6TU.*

Around 4-5 p.m., 2000 vehicles get stuck on the roads until late in the night. The snowfall continues until the next morning and in Chamonix, 33 cm of fresh snow are recorded over the day (figure 8).

4.1.1 Sensitivity experiments

- **REF.** A similar experience to the operational version of Arome 1.3 km is launched for the 12th of January at midnight and for a 24 hours forecast.
- **SURF.** A more complex land **surf**ace model is used : Interaction Soil-Biosphere-Atmosphere-Diffusion (ISBA-DIF). This multi-layer model works with the Explicit 12 layers Snow (ES) scheme to model the snowpack instead of the single-layer D95 model (Douville et al. 1995) used in operational.
- **A.** Compared to REF, only boundary conditions are modified. Instead of forecasts from the operational version, they come from hourly analyses.
- **156_a.** The vertical resolution of Arome is increased from 90 to 156 levels. This change has already shown its efficiency in improving the modeling of local phenomenons as fog.

Once obtained, results are evaluated and compared with the reference and observations.

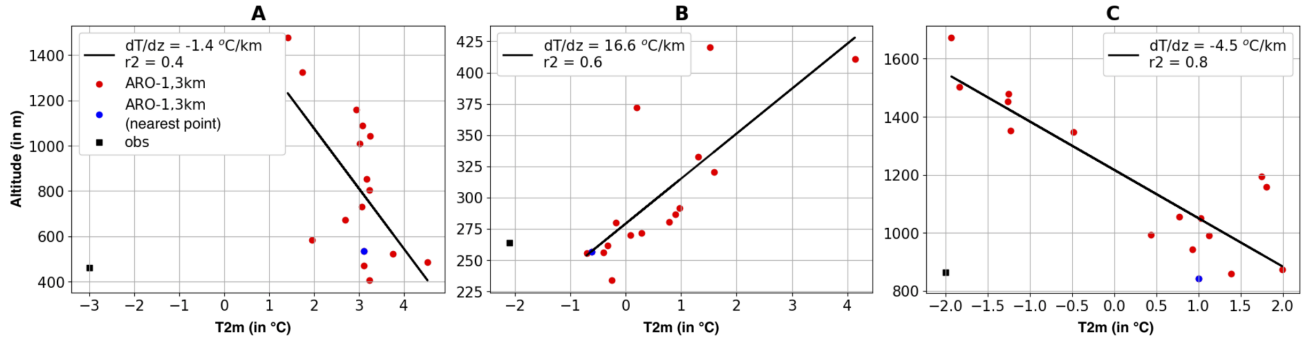


Figure 7: *Temperature profiles plotted from the $T_{2m}(z)$ data which correspond to the 4×4 grid points of Arome 1.3 km (run 00TU) located around the stations Sainte-Marie-de-Cuines (**A**), Montmélian (**B**) and Bourg-Saint-Maurice. (**C**), the 31th of December 2019 at 6TU. The black square indicates the observation, the red dots indicate the 16 model points used to compute the linear regression (black line), with the blue one representing the closest point to the observation, considering horizontal distances. The legend contains the lapse rate value and the r-squared of the regression.*

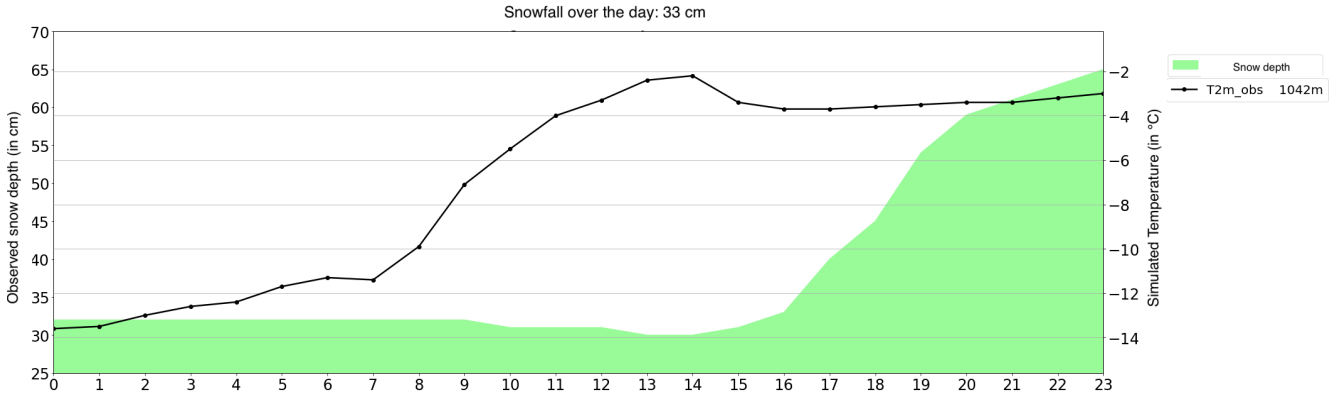


Figure 8: *Evolution of the observed temperature at 2 m (in black) and snow-pack height (in green) for the automatic weather reporting station of Chamonix the 12th of January 2021.*

4.1.2 Domain and scores

Simulations are run over the whole French Alpine region, but the study will focus on the area shown by figure 10.

This domain is centered on the Arve valley were the event of January 12 took place. This valley is characterized by **huge variations in altitude** (from 542 m for the bottom of the valley to 4809 m at the top of the Mont Blanc). Observations used to calculate the following scores come from stations of the RADOME network and the Nivôse station of Aiguilles Rouges :

- The **bias** score is used as in part 2.
- The mean absolute error or **MAE** calculates a bias without any error compensation. However, it doesn't show if the simulation is too warm or too cold compared to reality. The MAE will so be studied in conjunction with the bias. This calculation will also be done for each hour of the day.

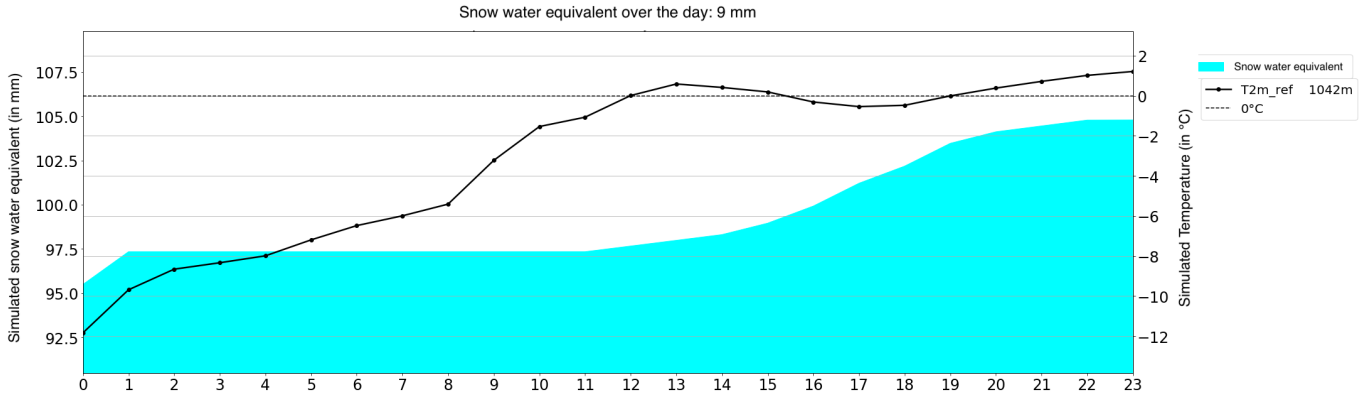


Figure 9: *Evolution of the modeled temperature at 2 m (in black) and snow water equivalent (in blue) for the automatic weather reporting station of Chamonix the 12th of January 2021.*

$$MAE = \frac{1}{N}\sum_{n=1}^{N} X_n - X_{obs}$$

These scores were calculated for each hour from 00h to 23h on 12 January. Only the stations with an altitude difference of less than 150 m with the nearest model point are presented and used here, as explained in part 2. In addition, the scores were also calculated by altitude bands: those **below 1100 m excluded**, those **between 1100 m and 2000 m excluded** and those **above 2000 m**.

4.2 Results

The simulations are all shown in figures 11 and 12, but only the most relevant will be considered and described below.

4.2.1 Investigating the warm bias

From the beginning of the episode, the model shows a warm bias in the valley with a value of 5.5°C for the station of Les Houches, almost 4°C in Chamonix and 2.8°C in Samoëns. The reference has a MAE of 2.3°C in the valley for the whole duration of the episode. The model shows a mostly positive temperature in the valley (figure 9) while the observations show temperatures below -2°C (figure 8).

SURF. Over the duration of the episode, this simulation is better than the reference whatever the altitude. Moreover, the warming linked to the arrival of the warm front is better represented in the valley (figure 11). Indeed, from 10h to 14h, its MAE is the lowest among all the simulations. The current snow scheme (D95) of Arome therefore seems too simplistic in this situation to correctly model the soil-atmosphere interactions. Moreover, the effectiveness of this correction suggests that it is one of the origins of the warm bias.

A. The use of the analysis does not significantly improve the bias in the valley. This score is due to a significant difficulty of **A** to model well the onset of precipitation. Indeed, before the snowy episode and after the beginning of isothermal conditions at about 3 pm, the bias and MAE are the lowest in the valley (figure 11). However, the experiment degrades the scores in the middle and high mountains (figure 12).

156_a. One of the first avenues considered for the dynamics was to increase the vertical resolution. For the duration of the episode, the result shows no difference with **A**. The experiment therefore does not seem to

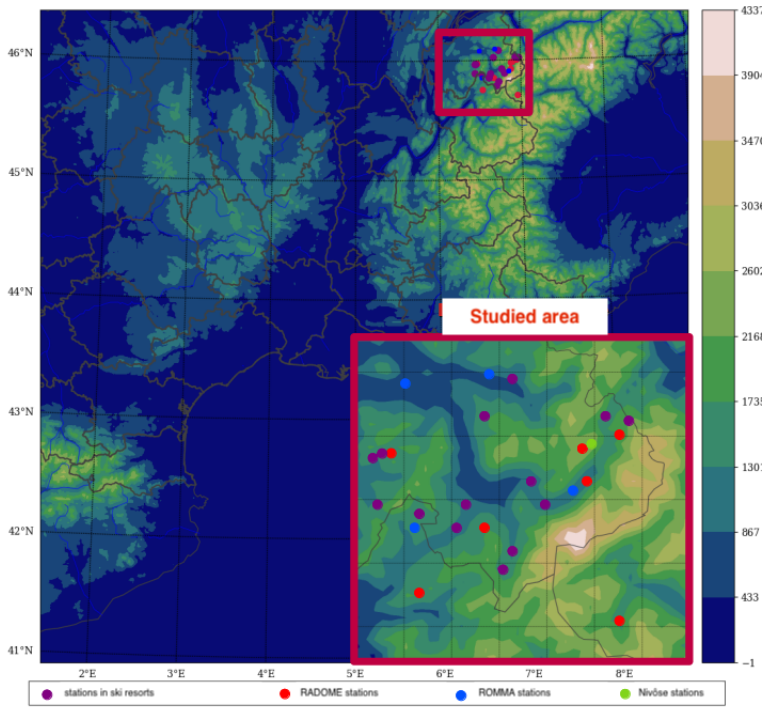


Figure 10: *Relief of the model for the Arve valley and its measuring stations. Automatic weather stations (RADOME) and stations in ski resorts are respectively in red and violet. Those of the ROMMA network are in blue and a Nivôse station (Aiguilles Rouges) is in green. Numbered stations correspond by ascending order to Chamonix (1042 m), Plan de l'Aiguille (2250 m), Aiguilles Rouges-Nivôse (2365), Les Houches (1005 m) and Samoëns (749 m).*

improve the performance of the model during the snowy episode in the valley. However, the MAE is reduced from 1.4 for **A** to 1.1°C in the mid-mountain and from 1.5 to 1.2°C above 2000 m. It should nevertheless be reminded that the use of the analysis (external constraint) to feed the boundary conditions may have increased the scores in the mid and high mountains compared to a simulation of 156 levels fed by the forecast. The scores could thus be even better.

4.2.2 A situation prone to biases

During the night and early morning of 12 January 2021, the model shows two errors corresponding to the warm bias in the valley under anticyclonic conditions, and the cold bias at altitude before the start of the snowy episode. Thus, the reference shows an error of +4.8°C at Chamonix, +5.9°C at Samoëns, +6.6°C at Les Houches and -5.3°C at Plan de l'Aiguille.

SURF. In the valley, the simulation increases the bias. Over the entire period before the snowfall, the mean absolute error is significantly higher than the reference. Conversely, over the same period in the mid and high mountains, the surface change seems to be very effective in decreasing the cold bias. Indeed, the experiment presents a MAE of 4.9°C against 6°C above 2000 m. However, the underestimation of the snow cover at Chamonix and its overestimation at higher altitudes raises questions, because the snowpack has an impact on the temperature. Further study would be required to analyze these results in more depth.

A. The scores of this experiment are clearly the best in the valley among all our sensitivity tests, with a MAE of 3.3°C over the period. The greater use of observations enabled by the analysis is therefore beneficial in the valley for the bias during snowy episodes and especially for the bias under anticyclone conditions.

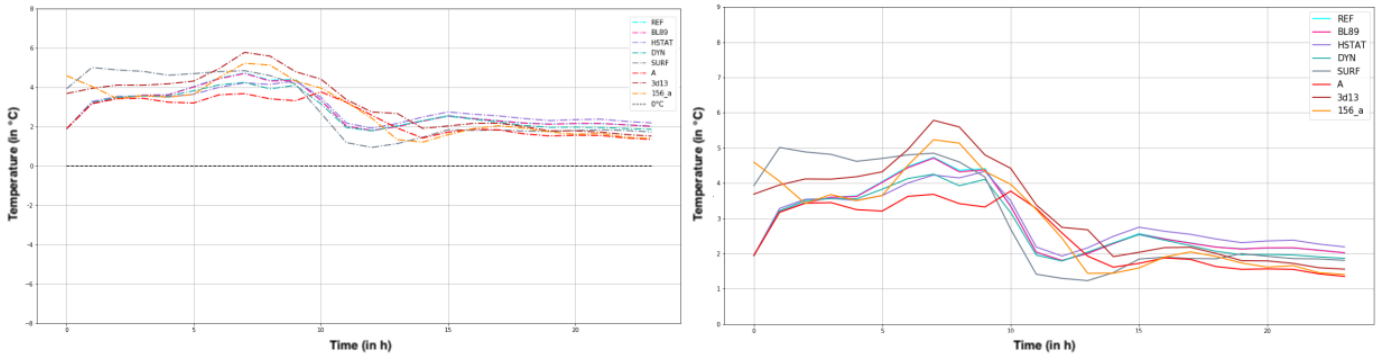


Figure 11: *The left figure shows the 0°C isotherm in dotted lines and the temperature bias over the day below 1100 m in dashed lines. The right graph shows the MAE for the same altitude as a solid line. The reference is in cyan, "A" in red, "SURF" in grey and "156_ a" in orange.*

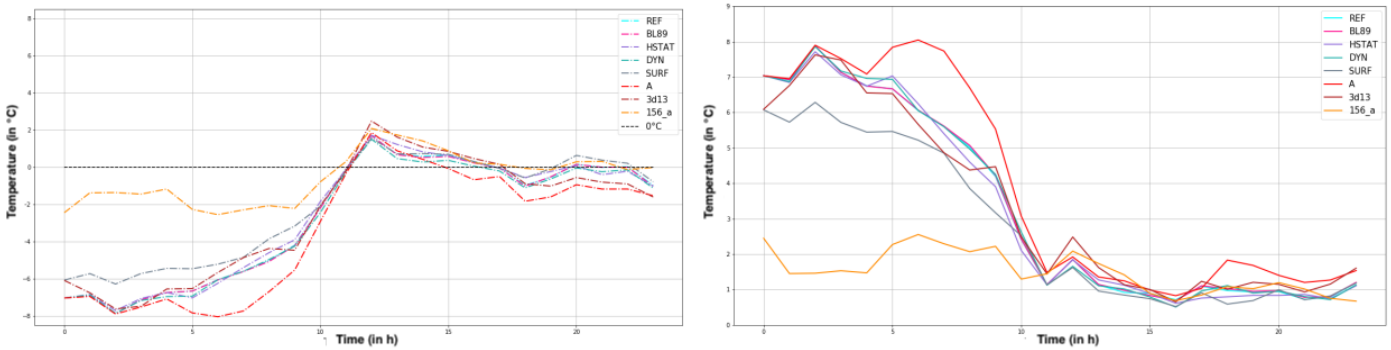


Figure 12: *The upper figure shows the 0°C isotherm in dotted lines and the temperature bias over the day above 2000 m inclusive in dashed lines. The lower graph shows the MAE for the same altitude as a solid line. The reference is in cyan, "A" in red, "SURF" in grey and "156_ a" in orange.*

156_a. Below 1100m, the simulation degrades the scores compared to **A** (figure 11). However, this experience stands out in the mid-mountain region, since it is the only one to show a warm bias. Moreover, its MAE is the worst there over the period and even exceeds 3°C of MAE during its temporal evolution. However, the simulation is by far the best in the high mountains with a MAE of 1.9°C. Consequently, this option amplifies the anticyclonic bias, but appears clearly necessary to correct the cold bias at altitude.

4.2.3 Inadequate data assimilation in mountains regions

Improving the data assimilation may be another way of correcting bias. Indeed, its use degrades the MAE scores, mainly in the high mountains, whatever the period (figure 12). Taking valley observations into account therefore seems to be detrimental to the temperature forecast at altitude. Indeed, as the model generally has a warm bias in the valley, the calculated analysis increment tends to cool the air, and is propagated to higher altitudes where fewer to no observations are assimilated. This cooling further increases the cold bias of the model, as we can see at Plan de l'Aiguille for the situation of January 12, 2021 (figure 13). A more appropriate data assimilation from the Nivôse stations could therefore improve the quality of the analysis in altitude. An other way of improvement is to use temperature increments dependent from the meteorological situation, as it

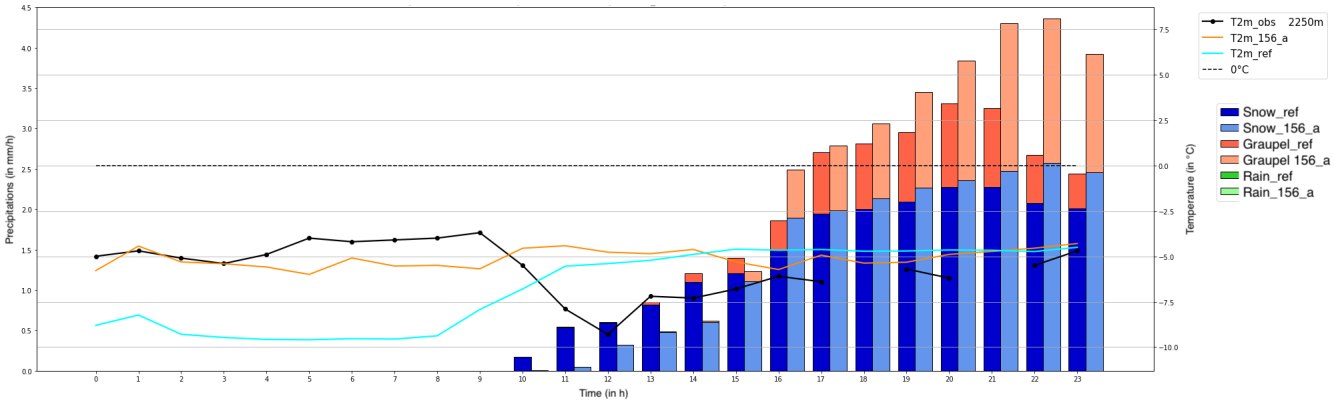


Figure 13: *Precipitation and temperature at 2 m at Plan de l'Aiguille. Snow is represented by a blue bar, graupel by a coral and rain by a green one. The darkest colours correspond to the reference and the lightest to "A". The temperature is illustrated by lines with the observations in black, the reference in cyan and "A" in red.*

is the case in 3DEnVar assimilation schemes for instance.

5 Conclusions and outlook

Several detrimental biases affect the Arome three-dimensional temperature fields in mountainous areas: (1) the model underestimates temperatures at high altitudes, (2) the inversion layer is too warm in stable conditions and (3) a warm bias occurs in valleys during snowfall situations. The correction of such biases is a necessary step before exploiting Arome 500 m for operational purposes in mountain regions.

Corrective downscaling from Arome 1.3 km outputs has been first attempted on 2 meters temperature. However, the biases remain and limit the downscaling potential, even in a dynamical approach allowed by Arome 500 m. These conclusions led to a focus on the model it-self, covering its assimilation, dynamical and physical parts.

Thus the work was continued by taking an interest in the snowfall situation of January 12, 2021. It is a typical situation of stable conditions followed by the arrival of a warm front through the Alps.

The sensitivity tests show that the warm bias has multiple origins. Indeed, the change of snowpack model improves the representation of the temperature in the valley. In addition, the situation has a poor starting condition since biases are present even before the beginning of the snow event. Thus, a data assimilation, more adapted to mountain areas, could have a positive impact on the forecasted temperature field. These results are valid for the warm bias in the valley during this episode of January 12, 2021.

Among the simulations, in the valley, the use of the analysis improves the warm anticyclonic bias, while the ISBA-ES surface scheme worsens it. Moreover, the increase in vertical resolution propagates the warm bias into the mid-mountain region. Therefore the warm anticyclonic bias does not seem to have exactly the same origins than the other two temperature biases. Indeed, as for the bias during snowy episodes, the snowpack model and, above all, the vertical resolution decrease the cold bias at altitude while the analysis worsens it. Thus, we anticipate that the assimilation of Nivôse stations data and the increase in vertical resolution would significantly reduce the cold bias. An interesting approach would be to use CANOPY (Masson and Seity 2009), a simplified turbulence scheme with additional levels for the surface, in order to better take into account the surface-atmosphere exchanges without drastically increasing its numerical cost.

Futures works need to deeply investigate the origins of temperature biases in Arome. In particular, the scores presented in this article are computed over a small number of stations and concern only a few situations, requiring a more general and exhaustive study to confirm and maybe detail these preliminary results. Moreover,

the case of the January 12n 2021 mainly concerns the warm bias during snowfall situations and does not fully focus on the two other biases for which a special work is desirable.

Correcting the Arome temperature biases paves the way to a performing deterministic and ensemble forecast in mountain areas. Projected on a high-resolution grid thanks to effective corrective downscaling, such products could supply the new forecast chain related to avalanche hazard.

References

Beauvais, L., 2018: Fronts chauds sur les alpes : Hiver 2017-2018. comportement des modèles. Ateliers de la prévision du Centre-Est.

Douville, H., J. Royer, and J.-F. Mahfouf, 1995: A new snow parameterization for the m??t??o-france climate model. part i: Validation in stand-alone experiments. *Climate Dynamics*, **12**, 21–35, doi: 10.1007/BF00208760.

Gouttevin, I., V. Vionnet, Y. Seity, A. Boone, M. Lafaysse, Y. Deliot, and H. Merzisen, in prep: To the origin of a wintertime screen-level temperature bias at altitude in a kilometric nwp model. *in prep for Journal of Hydrometeorology*.

Masson, V. and Y. Seity, 2009: Including atmospheric layers in vegetation and urban offline surface schemes. *Journal of Applied Meteorology and Climatology - J APPL METEOROL CLIMATOL*, **48**, 1377–1397, doi: 10.1175/2009JAMC1866.1.

Monteiro, D., 2020: Évaluation du modèle arome pour appréhender l'évolution passée et future du climat dans les alpes françaises. *Rapport Stage IENM3*.

Paci, A., et al., 2016: La campagne passy-2015 : dynamique atmosphérique et qualité de l'air dans la vallée de l'arve. Dernière connexion : 21 juillet 2020, dernière connexion : 21 juillet 2020.

Sheridan, P., S. Smith, A. Brown, and S. Vosper, 2010: A simple height-based correction for temperature downscaling in complex terrain. *Meteorological Applications*, **17 (3)**, 329–339, doi:10.1002/met.177.

Vionnet, V., I. Dombrowski-Etchevers, M. Lafaysse, L. Quéno, Y. Seity, and E. Bazile, 2016: Numerical weather forecasts at kilometer scale in the french alps: Evaluation and application for snowpack modeling. *Journal of Hydrometeorology*, **17**, doi:10.1175/JHM-D-15-0241.1.

An overview of the latest developments and tests regarding physics in MetCoOp

Karl-Ivar Ivarsson

1 Introduction

There have been new interesting physics developments during the last two years within the MetCoOp cooperation and also tests of other developments. A brief summary of this is presented here. The outline is the following: First some news about testing ECUME6 are presented, followed by presentations of two new tuning options, LHGT_QS and LICERAD. Then, there is a presentation of a simple way of improving the forecasts of supersaturation with respect to ice for especially contrail forecasts. It is followed by a new method of deriving the cloud cover, which is based on the optical thickness of solid water species. At last there is a short summary.

2 Tests and developments

Further tests of ECUME6

ECUME6 is the successor of the currently used ECUME scheme of fluxes over sea in SURFEX, developed by Meteo-France. Earlier studies have shown improvements of the forecasts of temperature and humidity in lower troposphere, probably as an effect of increased fluxes of heat and moisture. Some improvements of 2m temperature, clouds and precipitation forecasts where also found. But the drawback was an increase of fog, especially over sea. Since the fog was over-predicted also without ECUME6, the introduction of ECUME6 was postponed.
However, some resent developments of physics, e.g. LHGT_QS (see below) which reduces the fog, made it interesting to renew the tests together with those new updates. The new tests confirm earlier results, including an increase of fog, but now the occurrence of fog is acceptable also with ECUME6.

LHGT_QS

A way of tuning VSIGQSAT, which depends on the thickness of the model level has been developed. With VSIGQSAT, some extra variance of the distance between total humidity and saturation-humidity is added. This distance is an important variable in the statistical cloud scheme, since it is closely related to the cloud cover fraction. The variance is expected to increase with with increasing size of the grid-box, but in a complex and less well known way, so for the moment a simple tuning for adjusting VSIGQSAT depending on the model level thickness is considered.
Here is an example of how it works:
Assume VSIGQSAT is 0.02. Then VSIGQSAT remains 0.02 if the model thickness (Δz) is 30m, but will be set to VSIGQSAT x Δz / 30 otherwise. Δz / 30 is limited to the range [0.5:1.5], so in this case VSIGQSAT is limited to the range [0.01:0.03],
The parameters 0.5,1,5 and 30 may be changed by using the RFRMIN array in the namelist:
RFRMIN(25): The model level thickness for which VSIGQSAT is unchanged. Default is 30m.
RFRMIN(23): Lower limit for Δz / RFRMIN(25), default is 0.5. Lower value may give less fog.
RFRMIN(24): Upper limit for Δz / RFRMIN(25), default is 1.5. Higher value gives more low- and middle level clouds.

The verification results show somewhat less fog and some increase of low- and middle level clouds. Both are mostly beneficial, but it seems better to increase RFRMIN(24) to 2 or 2.5 in order to get enough low- and middle level clouds. The optimal choice is dependent of other updates/settings.

LICERAD

The option LICERAD makes ice clouds and mixed phase clouds more radiative thick. The reason for introducing this option is that currently ice- and mixed phase cloud fraction is reduced in order to account for that the visual effect of thin ice clouds may be very small, and this reduction is necessary for avoiding both too much cirrus and too dense ice fog.

But the radiation scheme needs the sub grid-scale fraction of grid box containing the solid water species, so although the reduction improves the cloud cover, the radiative effect of solid water species becomes too small. With LICERAD, this is solved by not using the cloud cover directly for radiation. Instead a larger value (C_m) is used in the following way:
$C_m = F_i + (1- F_i) C_t$, where C_t is the original cloud cover and $F_i = r_{stot} / (r_{stot} + r_c + tiny)$, which is basically the fraction of the non vapour water species that is ice. (rain and hail are currently excluded.) r_{stot} is the weighted sum of mixing ratios for ice, snow and graupel, r_c is the the mixing ratio of liquid water and 'tiny' is a small value used for numerical reasons (1.E-15).
When there is no solid water, F_i is zero, so the cloud cover enters radiation scheme untouched, but if it is one, the value entering the radiation scheme is also one, which actually is consistent with the ice microphysics, since the solid water species are assumed to cover the entire grid box.

LICERAD improves forecasts of mean sea level pressure, temperature (t2m , soundings) for all seasons, except in summer, and has been introduced in MetCoOp runs during spring 2021. However, combined with ECUME6, the forecasts become too warm, especially during summer, so ECUME6 is currently (summer 2021) not introduced in MetCoOp runs.

There have been other improvements of physics that make it possible to use both LICERAD and ECUME6 without a notable deterioration in the warm season. Applying recent developments of EDMF scheme, radiation scheme, new handling of liquid number concentration of droplets and adjustment of the tuning parameters RFRMIN(24), RADSN and RADGR gives the results labelled 'ALL-INCLUSIVE' in the following verification plots. The other three is 'REF' which is the c43 version used pre operationally in autumn 2020. The two others are REF with ECUME6 and REF with both ECUME6 and LICERAD.

The results for mean sea level pressure for summer (left) and winter (right) is seen in figure 1. The same for T2m in figure 2, low clouds in figure 3, temperature against soundings in figure 4 and specific humidity against soundings in figure 5.

Figure 1: Mean absolute error and bias for mean sea level pressure in July 2020 (left) and January 2021 (right) over the MetCoOp D-domain. The different forecasts are explained in the text.

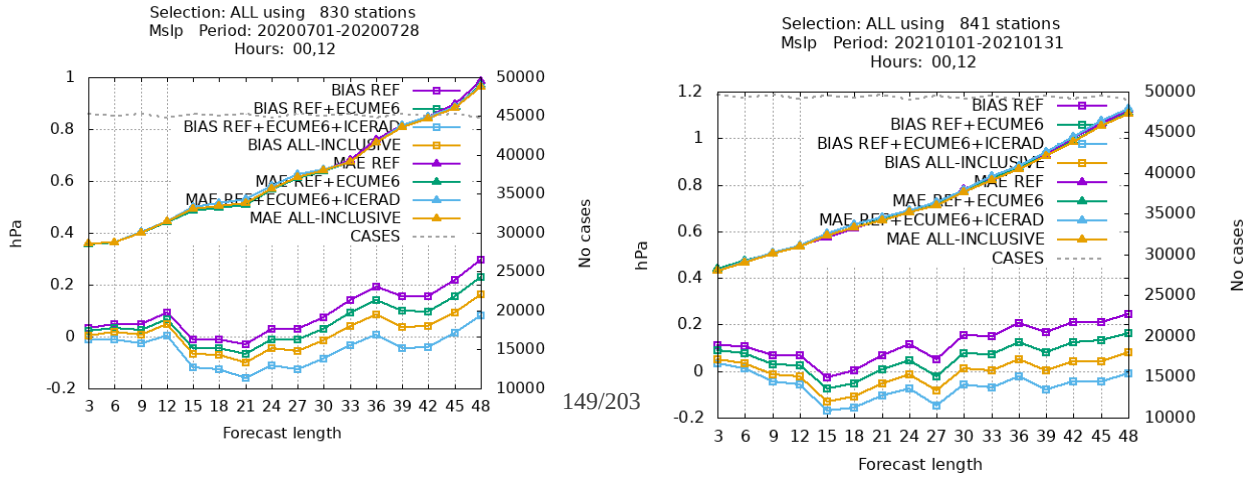


Figure 2: Mean absolute error and bias for 2m - temperature in July 2020 (left) and January 2021 (right) over the MetCoOp D-domain.

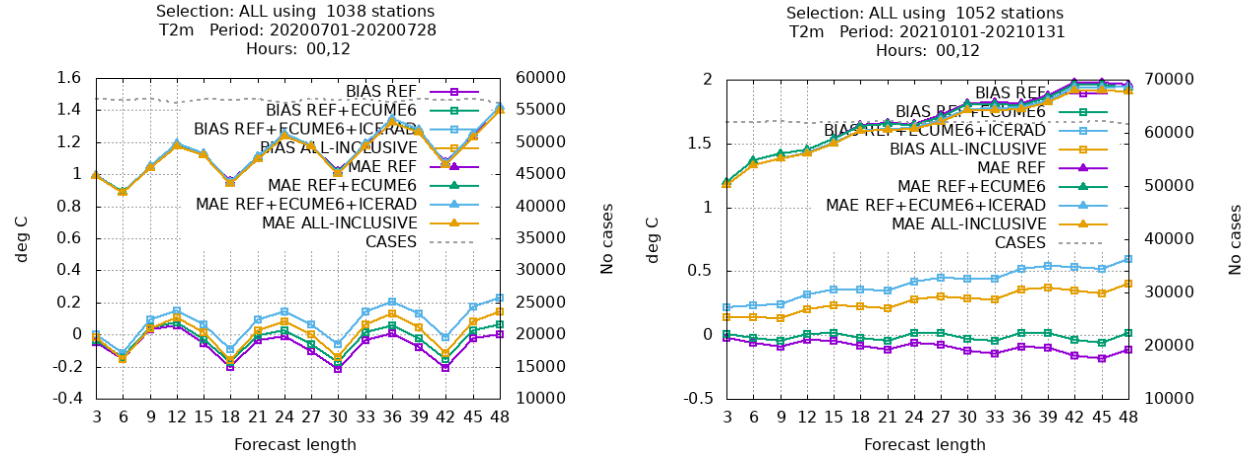


Figure 3: Mean absolute error and bias for low clouds in July 2020 (left) and January 2021 (right) over the MetCoOp D-domain.

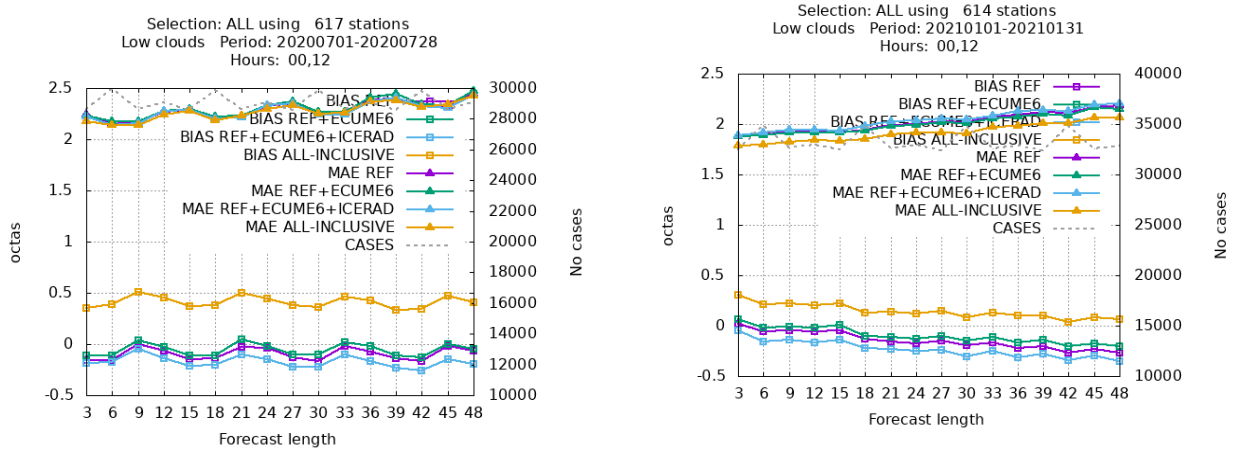


Figure 4: Mean absolute error and bias for temperature against soundings in July 2020 (left) and January 2021 (right) over the MetCoOp D-domain.

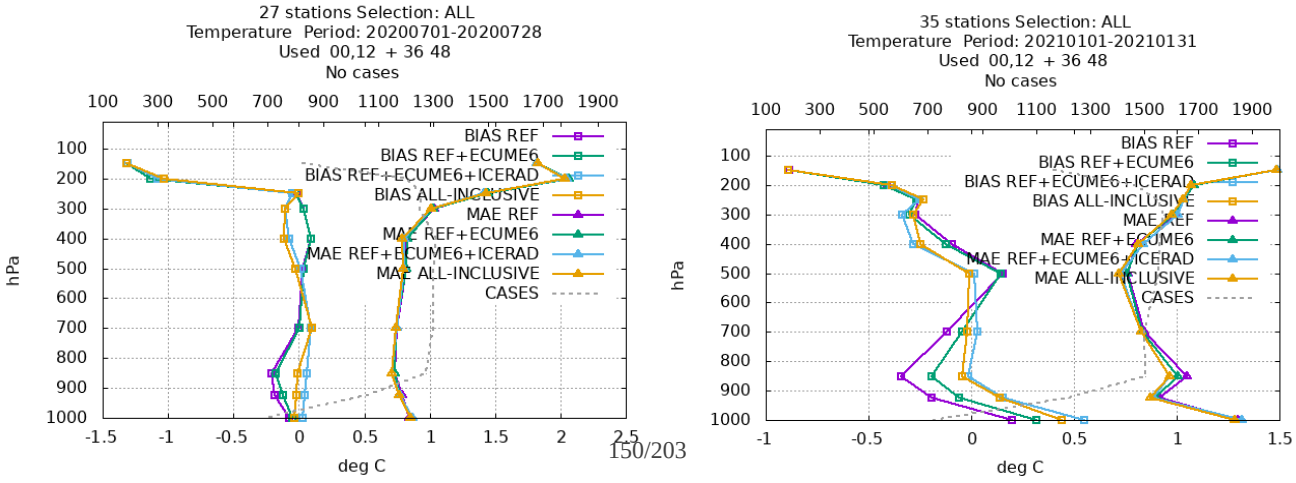
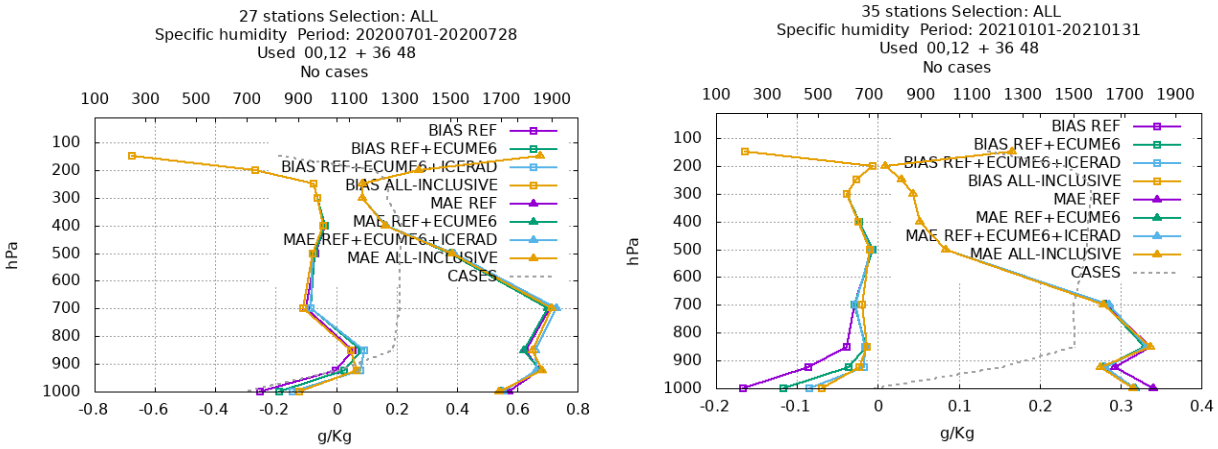


Figure 5: Mean absolute error and bias for specific humidity against soundings in July 2020 (left) and January 2021 (right) over the MetCoOp D-domain.



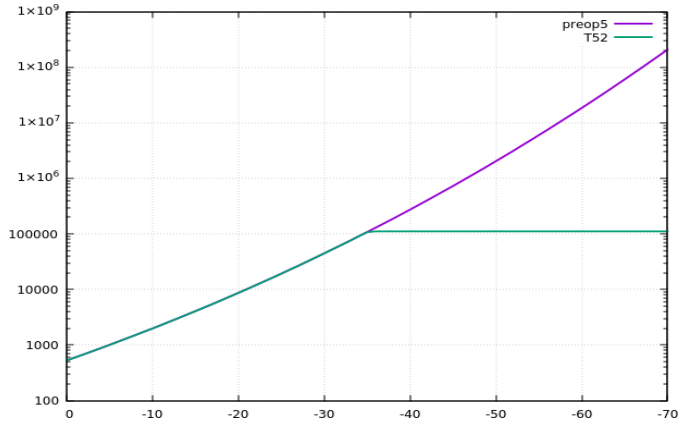
The absolute error of mean sea level pressure is similar for all tests and for both summer and winter. The bias is closest to zero in the 'ALL-INCLUSIVE' run. It is evident that the t2m-temparure becomes too high both summer and winter when ECUME6 and LICERAD are combined without other tunings. The bias is reduced in the 'ALL-INCLUSIVE' experiment. In order to increase low clouds, RFRMIN(24) is set to 2.5 (reference value 1.5) in the 'ALL-INCLUSIVE' test. It gives lower absolute error in winter, but the positive bias for both July and January is a sign that it may be too much. To be sure, more tests are needed. The forecast of temperature against soundings in July is quite similar between the runs. In winter, the biases are generally reduced with ECUME6 and LICERAD, as well as with the 'ALL-INCLUSIVE' test. Lowest absolute error with the 'ALL-INCLUSIVE' test. The forecasts of specific humidity are not improved in summer with the 'ALL-INCLUSIVE' test, but in winter both bias and absolute error are comparable with that of LICERAD+ECUME6 in lower stratosphere.

Improving forecasts of supersaturation with respect to ice

Aviation contributes to the global warming, not only by CO2 emissions, but also by the formation of contrails which reduce outgoing long-wave radiation. By avoiding areas with supersaturation with respect to ice, aircraft may reduce the effect on climate by reducing the contrails. This demands good forecasts of supersaturation with respect to ice. Verification results reveals that the occasions of supersaturation with respect to ice are under-forecast near the tropopause, where most of the aviation takes place. A simple way of improving the forecasts is presented here.

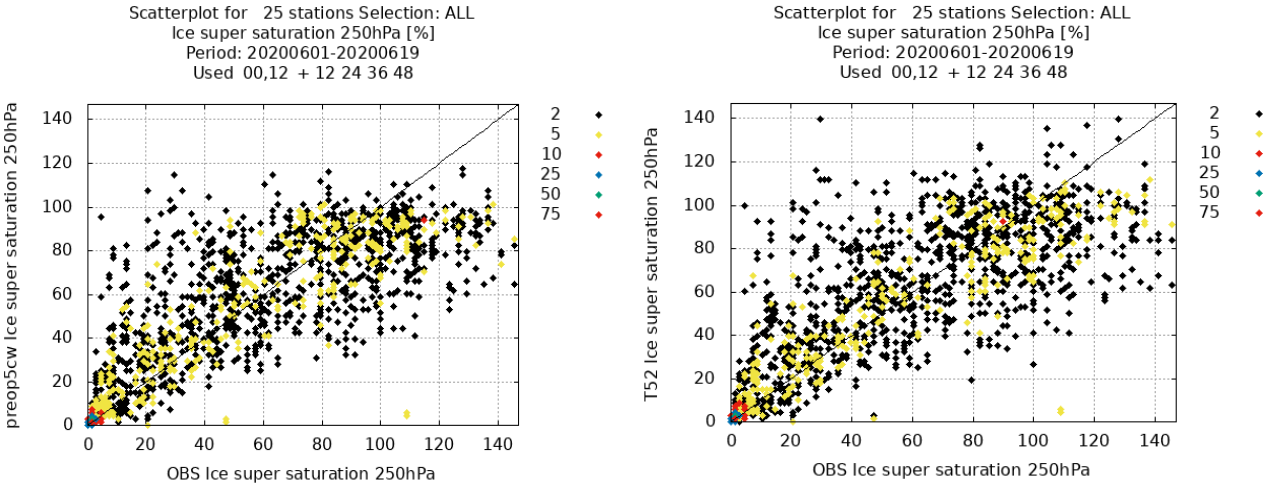
Reducing the number concentration of cloud ice water crystals is the simplest way to slow down the deposition of ice and thus prolonging a supersaturated environment. The present number concentration is based on the work by Meyers et al (1992). It originates from measurements between -10C and -30C, but is used also for the colder temperatures commonly present near the tropopause, thus less valid in this region. The prescribed number concentration is approximately tenfold for every drop of the temperature of 10 degrees, see figure 6. The test presented here, is just to limit the number concentration to what is used at 238 K (about -35 C).

Figure 6: Number concentration of ice at water saturation (vertical axis) for different temperatures in Celsius (horizontal axis). Original formulation in purple and the test in green. (for warmer than -35C they are equal)



The result is some improvement of the forecasts supersaturation with respect to ice, see figure 7, but there is still some under prediction. The remaining under prediction may partly be caused by poor vertical resolution.

Figure 7: Forecast relative humidity with respect to ice (vertical axis) compared to observations (horizontal axis) for soundings at 250 hPa. The degree of supersaturation is obtained by subtracting 100. Reference experiment to the left and the test to the right.



The verification result in other respect is mainly neutral. A winter period has also been tested with a similar result (not shown). An alternative to an upper limit of the ice number concentration, is to assume a less steep increase with decreasing temperature below e.g. -30 C. This has not been tested yet.

Deriving the cloud cover from the optical thickness of solid water species

The current cloud cover parametrization for ice clouds is partly based on the relative humidity over ice. If humidity exceeds a threshold, often called 'the critical relative humidity', the cloud cover becomes non-zero. However, in order to avoid too much cirrus and too much ice fog, a reduction of this cloud cover fraction is needed. The method of doing this reduction (B_{ci}) is inspired on the work by Xu and Randall (1996). It can be written as

$$B_{ci} = \min(1, A_{ci}\, r_{stot}\, (\Delta z / r_{si})^{0.5})$$

Δz is the thickness of the model layer, A_{ci} is a tuning coefficient, r_{stot} is a weighted mixing ratio of solid water and r_{si} is saturation mixing ratio of water vapour over ice. This pragmatic solution works quite well, but does not account for neither the number concentration of the solid water species nor the size distribution of them.

An alternative cloud cover parametrization is presented here. It is based on a new method, refereed to as the 'slice method', relying on the size distribution used in the microphysics (ICE3). The dependence of relative humidity over ice is dropped.

Consider a volume of air with a tiny thickness (or depth) of the same order as the particles of water species, B (m). Assume that the area *per volume* hidden by solid water is A (m^{-1}). Choosing the thickness of the layer to be of the same order as the diameters of the particles, one may neglect the overlapping effect and the fraction hidden by the particles becomes AB. The other part is then 1 – AB.

The next step is to derive the radiative blocking effect of several such thin layers. Starting with two layers only and random overlapping between the layers one obtains four alternatives, AB(1-AB) ,(1-AB)AB, ABAB and (1-AB)(1-AB). Only the part (1-AB)(1-AB) has no blocking so the part with blocking becomes $1 – (1\text{-}AB)^2$.

With K layers the result becomes $1 – (1\text{-}AB)^K$, and for a model layer with the thickness Δz one needs $\Delta z/B$ thin layers and the fraction blocked (B_2) becomes

$$B_2 = 1 – (1\text{-}AB)^{\Delta z/B}$$

A questionable assumption used here is to go from no overlapping within the layer to pure random overlapping between the nearest layers. It is more natural the assume a gradual transition. In order to test the weakness of this assumption, the sensitivity with respect to the choice of thickness B is tested. Consider a simple example: Spheres, all with the radius r_x and
a number concentration of N m^{-3}. B is also set to r_x. This leads to $A = N \pi r_x^2$ and

$$B_2 = 1 - (1 - B\, N\, \pi\, r_x^2)^{\Delta z/B}$$

With N = 50.E6, m^{-3}, Δz =100m and r_x = 1.E-5m one obtains B_2 = 0.792111. Here, B is set to r_x. If B is instead set to $100r_x$, the result becomes 0.792113. Changing it to a 10000 times lower value,
$B = 0.01r_x$ leads to the the result 0.792111. This indicates that
the value of B is not that critical, which is important considering that r_x may vary a lot within a given size distribution. It also indicates that going from no overlapping within the layer to random overlapping between the layers is an acceptable approximation in this context.

Using the standard microphysics formulas in ICE3, A can be derived for each solid water species which is valid for spheres and plates. (See the table below, table 1, for the different constants and the derivation of the slope parameter λ to the left. Here, r is mixing ratio of water specie and ρ is density of the air.)

$$A = \frac{\pi}{4} C \lambda^x \int_0^\infty D^2 g(D) d(D) = \frac{\pi}{4} C \lambda^{x-2} \frac{\Gamma(\nu + 2/\alpha)}{\Gamma(\nu)} \qquad \lambda = \left[\frac{r \rho\, \Gamma(\nu)}{aC\, \Gamma(\nu + b/\alpha)} \right]^{\frac{1}{x-b}}$$

Table 1: Values of different constants in the ICE3 scheme

Water specie	a	b	α	ν	C	x
Cloud liquid	524	3	3 on sea, 1 on land	1 on sea, 3 on land	variable	0
rain	524	3	1	1	$8\ 10^6$	-1
Cloud ice	0.82	2.5	3	3	variable	0
snow	0.02	1.9	1	1	5	1
graupel	19.6	2,8	1	1	$5\ 10^5$	-0.5

The cloud cover from water clouds based on the statistical cloud scheme is left untouched here, so the new parametrization deals with the ice clouds (C_i) only. Therefore, the cloud cover (C) is computed as

$$C = C_w + (1 - C_w)\, C_i ,$$

where

$$C_i = \min(1, B_{2S} R).$$

Here, $\min(1,B_{2S}R)$ is the contribution from solid water species and C_w is the water cloud fraction (based on the statistical cloud scheme). R is a tuning parameter representing how transparent a cloud should be to be regarded as 'overcast'. B_{2S} is the sum of B_2's for the three solid water species considered here (ice,snow,graupel). Some reduction of B_2 is needed near the surface for snow and graupel, thus avoiding confusion between cloudiness and horizontal and vertical reduction of visibility caused by precipitation. According to observation-practice, this visibility reduction is not regarded as cloudiness.

Presented here, is basically only a new post-processing of cloudiness. However, since low clouds also affects the surface analysis scheme Canari, the analysis becomes a little different, so there are some small other differences for other parameters as well.

The result for the total cloud cover is seen in figure 8 and for low clouds in figure 9.

Figure 8: Bias and mean absolute error for total cloudiness, July 1 to 28 2020 to the left and January and February 2021 to the right. The present cloud calculation (REF) is in purple and the experiment in green (TEST). The domain is MetCoOp D.

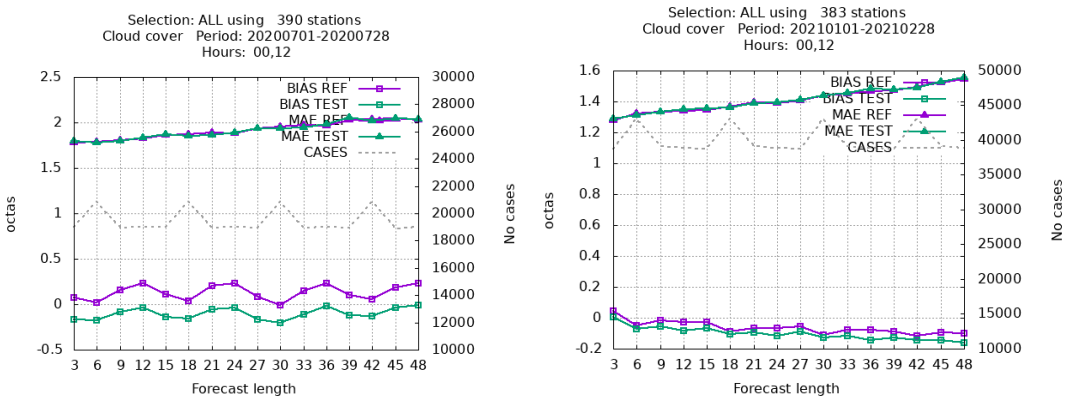
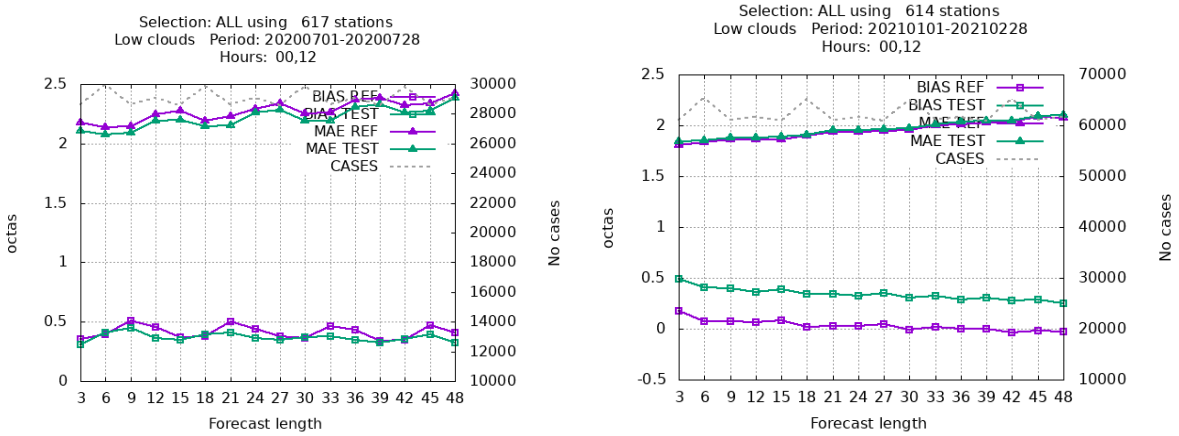


Figure 9: Bias and mean absolute error for low clouds, July 1 to 28 2020 to the left and January and February 2021 to the right. The present cloud calculation (REF) is in purple and the experiment in green (TEST).



The cloudiness is surprisingly similar, despite that the calculations are quite different. However, in July there is a little less low clouds as well as total clouds in the test run. The main reason is that in the test run, the cloud-fraction of pure water clouds comes entirely from the statistical cloud scheme, but in the reference run a contribution also comes from a calculation based on the cloud liquid water. This part is often larger than that from the statistical cloud scheme. Despite this, the amount of low clouds becomes larger for the test in winter. This is caused by an increased contribution from falling solid precipitation, especially snow. The largest differences between the runs are seen for cirrus clouds (pure ice clouds), which in general are less thick in the test run. This can be explained by less effect from cloud ice water. That leads to somewhat less cloudiness as a whole, also in winter. Some tuning may be needed to increase the amount of ice clouds. One may also consider including rain in the new version.
It should be noted that currently, radiation is implicitly treated as being particles, and not also as waves, as it is in real nature. The transparency of water species particles is also not accounted for. This has to be considered for updates of this parametrization.

3 A short summary

Further tests with ECUME6 show promising results and the new option LICERAD reduces some systematic errors, but no improvement in summer. Some further tunings are needed for combining them.
Tuning with LHGT_QS, together with RFRMIN(23), RFRMIN(24) and RFRMIN(25) gives some more low-, and middle level clouds without increasing fog.

More climatological realistic occurrence of supersaturation with respect to ice is achieved with a simple method by just set an upper limit of the ice number concentration. This can be activated by setting RFRMIN(27) to e.g. 238 K. There are also other methods, that need to be tested.

Deriving cloud cover from the optical depth of solid water species gives reasonable results, but longer tests and some adjustments are needed.

4 Acknowledgements

The author wishes to thank Ulf Andrae for technical help, especially with verification of ice super saturation. And also Emily Gleeson, Sander Tijm, Kristian Pagh Nielsen, Bent Hansen Sass and Wim de Rooij for valuable input and suggestions for the testing and development.

5 References

Xu and Randall, 1996 :
A semiempirical cloudiness parameterization for use in climate models.
J.A.S Vol 53 3084-3101

Meyers et al 1992 : New primary ice-nucleation parameterization in an explicit
cloud model J. Appl. Meteor. 31, 708-721

Physics parametrizations developments in RC LACE

Bogdan Bochenek, Martina Tudor, Mario Hrastinski,
Ján Mašek, Radmila Brožková, Andre Simon

1 Introduction

Developments on the turbulence scheme TOUCANS are progressing well, included shallow convection scheme is completed and already used in the operational application at CHMI. Study of the stability of the solver inside TOMs part and research on mixing length computation have continued.
Work on coupling ALARO-1 with the SURFEX scheme has continued. In order to ensure the ascending operational compatibility with using the ISBA scheme, many problems were identified (surface roughness, drag coefficients treatment) and solved.

2 Turbulence schema

The starting point is revised $L=L_{TKE}$ formulation, with smooth transition from the surface "κz" layer to the upper layer where pure TKE-based solution (L_{TKE}) prevails. On top of that there is so-called "crossing parcels" treatment, which significantly improves the model performance during the summer convection.
Starting formulation is further upgraded by setting the free atmosphere lower bound on LTKE and consequently lm. At first it was set as a constant value (for all conditions; LTKE=180 m), which is also done for the referent Geleyn-Cedilnik formulation. This modification resulted in necessary increase of mixing in middle and upper PBL and to smaller extent in the lower PBL. It also resulted in drying the surface layer and moistening of upper PBL, as well as reducing the difference of the magnitude of individual terms (DDH budgets) compared to the reference.
The impact on temperature is smaller and significant only near the surface. The concept of free atmosphere lower bound is further upgraded by allowing its variation depending on the maximum of LTKE, which shows strong daily and seasonal variability. This resulted in necessary additional warming and drying of the surface layer, as well as further moistening of middle and upper PBL.
The verification scores are significantly improved and TKE-based formulation is now comparable to the reference, both for summer and winter.

3 Prognostic graupel

Validation of e-suite with prognostic graupel in Poland is running. Scores of point-to-point (Fig. 1) and SAL (Fig. 2) verifications shows similar results compare to reference run with diagnostic graupel. The only difference is present at mountain stations, where prognostic graupel scheme performs better.

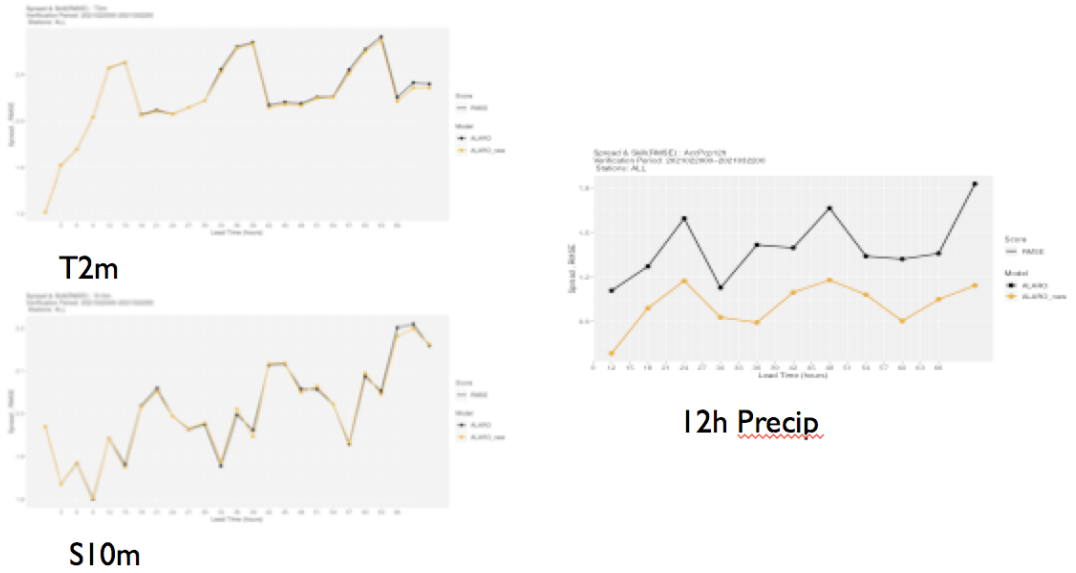


Figure 1: Prognostic graupel scores for 22.02.2021 – 22.03.2021. Black operational ALARO, yellow ALARO with prognostic graupel.

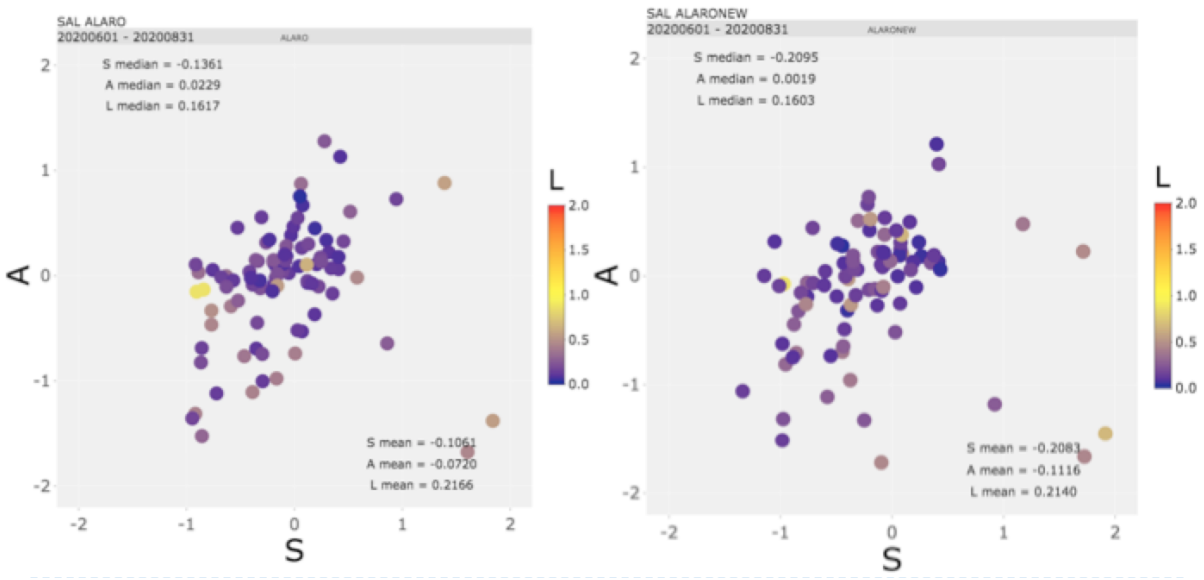


Figure 2: Prognostic graupel SAL scores for 22.02.2021 – 22.03.2021. Left: operational ALARO, right: ALARO with prognostic graupel.

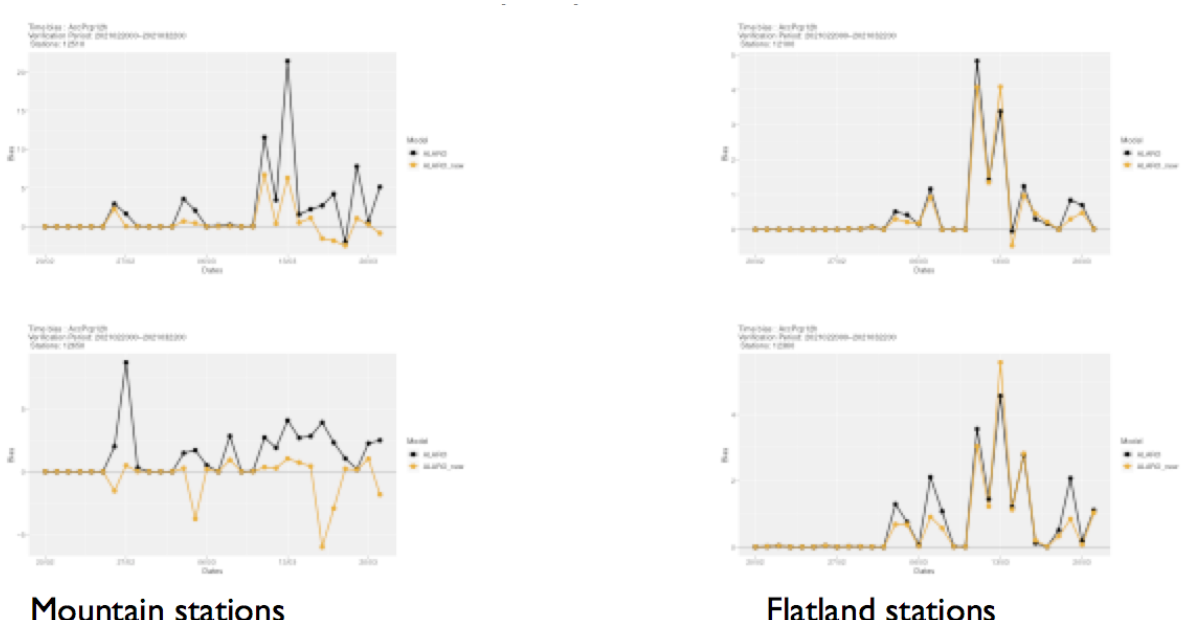


Figure 3: Prognostic graupel scores for 22.02.2021 – 22.03.2021. . Black operational ALARO, yellow ALARO with prognostic graupel.

4 VHR tests

Odd flows appear in the model forecast. The phenomenon was already recognized in 2019 in operational AROME model at OMSZ as an outflow and wave spreading very fast from the Balaton Lake shores. The model was exceptionally run as pure dynamic adaptation due to sudden technical problems, in "normal" runs with assimilation it did not occur

Very similar traits could be recognized in outflows in high-resolution (325 m) experimental dynamic adaptation run for the domain of Slovakia (NH dynamics, cy43). The concentrically spreading 10m wind velocity maxima indicated a problem with convection but appearing in very stable stratification on 26 January 2020 (12 UTC run).

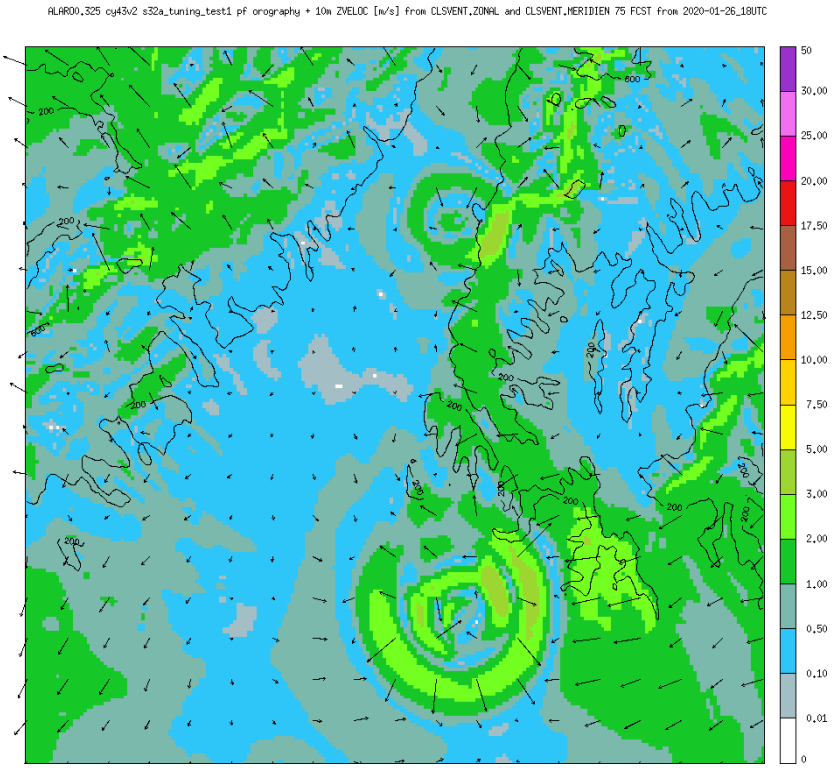


Figure 4: Example of outflow

The vertical cross-sections reveal an interesting fact that the temperature anomaly over the lakes was created in the e927 configuration due to vertical interpolation. Not at surface but about 1400 m high. Switching LESCALE_T=.F. In e927 "cleaned" the anomaly, which otherwise induces deep convection immediately.

Although the presented problem was artificial, lake-induced convection does exist. This can cause snowstorms (e.g. at large lakes as Great Lakes in the USA) and local thunderstorms, especially by warm-water shallow lakes as Balaton or in the tropics (Lake Victoria). At high resolution, there are already many local heat sources, not only lakes and water surfaces but also urban areas, soil/vegetation patterns, etc. However, in stable stratification, shallow heat sources should not induce deep convection as shown in the cases above. It is probable that one would not observe such violent effects in a hydrostatic model. This problem was visualised thanks to non-hydrostatic dynamics and it is possible that it could even cause numerical instability under some circumstances (as by LESCALE_U=.T.) The next task would be to find how to keep positive features of the vertical interpolation, while not inserting new anomalies from too warm surfaces

5 Surface physics

Orographic roughness length

Technical work (preparation of climatological files) towards using GMTED2010 instead of GTOPO30, and benefit from ECOCLIMAP I, II, SG. Wind speed verification over Central Europe (600 stations) with ALARO 2.3 km at CHMI for orographic roughness length calculated from GMTED2010 with/without scaling and smoothing. Similar experiment with vegetation roughness length with ECOCLIMAP I and ECOCLIMAP II with the same smoothing.

6 References

Mario Hrastinski: TKE-based mixing length in TOUCANS, RC LACE stay report, Prague, a very preliminary version, final version after the last stay
Ján Mašek, Radmila Brožková, 2020:Determination of the orographic and vegetation roughness in the ALADIN model at CHMI , CHMI technical document, Prague 2020

Code adaptation to hardware accelerators: activities and plans

Daan Degrauwe, Piet Termonia, Philippe Marguinaud, Florian Suzat, Thomas Burgot, Claude Fischer

1. Introduction

It is an understatement that the computer hardware market is undergoing fast changes, and that the ability to respond to these changes is of crucial importance to our NWP community. This is acknowledged in the strategy of the ACCORD consortium "[...] our codes should be highly adaptable to guarantee that the quality of forecasts will not suffer from unforeseen evolution of the computing market".
This document gives an overview of activities that have taken place with the goal of preparing our code for a changing hardware landscape. While the focus of these activities so far has been mainly on GPU's (more specifically, Nvidia GPU's), we should keep in mind that also other computing hardware is emerging, such as vector accelerators, ARM processors, FPGA's. That being said, we are convinced that the work that is carried out to improve the flexibility of the code and the tools that are being developed for code adaptation, will also benefit the porting to such platforms.
Besides documenting the past work, this document will also briefly outline the upcoming activities.

2. IDRIS GPU Hackathon

In May 2021, the French HPC centre IDRIS, together with Nvidia, organized a hackathon in which several teams were invited to port and/or improve their codes on GPU's, under the guidance of Nvidia experts, and with access to a hybrid CPU/GPU cluster from IDRIS. A Météo-France team took this opportunity to study the feasibility of porting various parts of the ACCORD code to GPU, and to familiarize themselves with the development tools provided by Nvidia, including compiling, debugging and profiling.

A first part of the ACCORD model that was worked on are the Fast Fourier transforms that lie at the basis of our spectral dynamical core. These Fourier transforms were ported to GPU, using the CUDA library, and a significant effort was made to reduce the impact of host (CPU) - device (GPU) data transfers.
While this work was successful in itself, with the Fourier transforms running considerably faster on GPU than on CPU, the current implementation does not scale beyond 4 GPU's. The challenge is that while the calculations on GPU are quite fast, the communications between GPU and CPU (and from there, to the interconnect) become the bottleneck that determines the overall speed. It will be investigated whether such limitations will disappear on other hardwares. From this perspective, the further monitoring of new hardware developments continues to be crucial.

Besides the FFT's, two physics parameterizations were addressed during the IDRIS hackathon: the deep convection scheme acpcmt and the radiation scheme acraneb2. The focus of the work on these parameterizations has been to compare different strategies of porting to GPU by means of OpenACC directives. One such strategy is to make the piece of code (kernel) to be offloaded to the GPU in a single go as large as possible; in this case that would come down to the entire parameterization. The

advantage of this approach is that the amount of OpenACC directives is quite limited. On the other extreme end, one could take an approach where many small kernels are defined (basically around each horizontal loop in the code). While less attractive in the sense that this approach will require many more OpenACC directives, it turns out to be significantly better in terms of performance. A detailed comparison of the porting strategies for the acraneb2 parameterization can be found on https://opensource.umr-cnrm.fr/projects/accord/wiki/NVidia_Idris_Hackathon. It should also be noted that annotating the code with OpenACC directives can be largely automatized if the original source code is in a proper shape (e.g. avoiding array operations, no implicit dimensioning of arguments).

In summary, several pieces of the ACCORD code were ported to GPU during this hackathon, and, probably more importantly, know-how and tools were built to also perform this task for other parts of the code.

3. Upcoming activities

While work on individual pieces of code, as was done during the IDRIS hackathon, is useful in itself, there's also a challenge in putting these pieces back together. For example, when running on a CPU, it is advantageous to split the work into smaller "NPROMA" blocks, in order to reduce the memory footprint, and improve cache efficiency. However, on a GPU it's important to expose as much parallelism as possible, and treating the different NPROMA blocks one-by-one will not lead to good performance. On top of that, the user may want to run some parameterizations on an accelerator (not necessarily GPU), while other parameterizations remain on the CPU. To deal with such a situation, a very flexible code is required where the choice for the parallelization model (OpenMP/OpenACC/vector/…) is made at the level where the parameterizations are called (i.e. routines aplpar/apl_arome in the ACCORD code), instead of at the higher level where it is made now (routine cpg_drv). The dataflow between the parameterizations, which currently relies on raw Fortran arrays, should be organized with smart data structures that are able to keep track of the context (e.g. allocation status on CPU host and accelerator device).
Parts of the ACCORD code are shared with other partners. The overall code structure is common with IFS. ECMWF has already made significant progress in restructuring their physics parameterizations calls and using smart data structures for the dataflow. In the future, these data structures may be integrated with Atlas classes. A similar effort is needed for the structure and dataflow of the ACCORD physics parameterizations.
Finally, a lesson learned from the hackathon is that while it is not feasible to maintain a separate code version for every possible hardware platform, it is possible to achieve reasonable performance with tools that translate/annotate the code for a specific platform. At ECMWF such a tool (LOKI) is being developed. Whether LOKI will also suit the needs of the ACCORD partners, or a custom tool should be developed by ourselves, remains open for discussion.

Porting Harmonie-arome to a public cloud service - AWS

Jacob Poulsen, Xiaohua Yang, Eoin Whelan

1 Introduction

This report provides a high-level description of our experience of porting the Harmonie-arome weather forecast system to Amazon Web Services (AWS), a public cloud service.

Working with Numerical Weather Prediction (NWP) models involves the need to adapt code and systems to a variety of High Performance Computation (HPC) platforms with different types of hardware and software stacks. Furthermore, in view of the usual needs in collaborative development involving multiple partners, both the system portability and transparency are of vital importance. In recent years, cloud services such as AWS, with a wide range of computation architectures, have become increasingly relevant for exploitation within the weather forecasting community. In connection with the work on a forecast benchmark for United Weather Centre-West (UWC-West) project and the upgrade of the operational forecast system at the Danish Meteorological Institute (DMI), both of which are based on the Harmonie-arome 43h2.1 released by the Hirlam-C consortium, porting and testing of the Harmonie forecast system onto AWS has been conducted. In these explorations, we tested both a stand-alone benchmark configuration and a full system configuration. The latter includes boundary preparation and observation pre-processing, data assimilation and forecast cycling, and the subsequent postprocessing; i.e. execution of a full NWP cycling dataflow with the required typical suite of applications using the Harmonie scripting system. Moreover, we have tested different compilers (GCC and Intel), OS versions, networking interfaces as well as filesystems. AWS offers an overwhelming set of compute instance types that we can use as node types. Merely as a proof of concept, we ported the Harmonie system to both x86-64-based nodes as well as Arm-based nodes and deployed them using AWS ParallelCluster.

Background

Project goals

Needless to mention, the public cloud vendors keep increasing their offerings, and nowadays cloud configurations that implement a full HPC-like cluster are available too, cf. interesting posts like AWS Top40 blog post and AWS blog post on WRF. Thus, it seems reasonable, for any NWP centre to consider, if these offerings are interesting to them and if so to what extent. Our initial goal was merely to learn what AWS had to offer HPCwise. We wanted to get a first-hand impression of how an AWS HPC system would compare with our current private or shared-within-the-community HPC systems. We decided that the best way forward was to try to port the Harmonie system to AWS and deploy using AWS ParallelCluster. Secondly, we aimed at getting a better understanding of what to expect from a performance and cost perspective, in order to compare with the traditional on-site systems that we are so used to. Getting these hands-on experiences would allow us to take in-depth discussions on which use-cases would be relevant, and eventually form a more solid strategy on how to potentially use HPC offerings in the cloud. At this early point in time, we certainly do not have all the answers yet, but the journey has begun, and we have been encouraged to share our initial findings. By doing so, we hope that we can engage with community members having similar interests.

Our exploration with Harmonie-arome system on AWS consists two flavours of the system,

- UWC-W benchmark 2021, which consists of three components: a 48 hour forecast on a large (1920x1620x90) domain [**MASTERODB** in single precision], a 4D-Var screening of observations with two hour window [**MASTERODB** in double precision], and a lateral boundary condition pre-processing task [**GL** in double precision].
- A complete assimilation-forecast cycle with the pre-operational DMI adaptation of the Harmonie-arome 43h2.2 including build, pre-processing, assimilation-cycling and postprocessing, on the (1200x1080x65) DMI-NEA domain.

Before we started, it was unclear to us how much effort would be required to port Harmonie to AWS, nor was it clear if AWS had an interconnect and IO service that would allow us to run a tightly coupled application with a lot of MPI communication and I/O operations like a Harmonie forecast. We were also unsure whether or not we would be able to meet runtime targets similar to those on our on-site HPC systems. Another open question was if we actually had use cases that would be economically viable when comparing with use of an on-site installation.

HPC on-site versus HPC on AWS

In this section we will try to describe how HPC on-site systems differ from those of AWS and we will also identify the parts that are truly similar. Comparing with either remote or local traditional HPC systems that we have used for the past decades, there were a couple of things that were quite different from the usual experiences and hence seems worth mentioning here. It should be noted that the NWP community is used to switching between different underlying architectures, e.g. vector systems versus MPP systems, different queuing systems, different UNIX variants, … and we have realized that these system changes always pose non-trivial porting and performance challenges initially. However, the fundamentals remain the same: on top of the HPC hardware carefully selected by an HPC vendor to meet tender requirements, there is also a well-established HPC software stack targeting the requirements of NWP. We rely upon that base when we start the porting work. That is, the HPC systems that we have been exposed to throughout decades all come with the following:

1. Multiple compilers with multiple versions that can easily be switched between by the user.
2. Complete Software stacks with all the scientific libraries used by the earth-science community (including compiler specific fortran modules) pertaining to the different compilers.
3. Parallel debuggers and profilers.

On AWS on the other hand, one will find all the hardware infrastructure that one could dream of. However, picking the proper components, picking a proper base OS and related kernel and maybe even tweaking these for sole HPC purposes, building the software stack of compilers, scientific libraries is the sole responsibility of the user of the HPC system. Same is also true for buying software licenses for compilers and HPC products like parallel debuggers and profilers. This is a stark contrast to the usual on-site HPC system experience, but indeed it follows the trends that we have seen building up in IT in general for a very long time. We will refrain from judging whether or not this approach is good or bad but confine ourselves to share the observation and stress that this difference is something one should keep in mind when cross-comparing these systems.

The fact that the user has to establish and maintain the software stack does impose a hidden cost, not to mention an additional risk in the challenge of porting these applications and the stacks they rely upon. Dealing with bugs down the software stack becomes an additional risk, which today is handled by the vendors or administrators of the on-site HPC system. This is something one should keep in mind when cross-comparing with typical on-site HPC systems, especially when it comes to evaluating the total cost of ownership. On the other hand, once a port is complete, the port including the full stack

can be shared with the whole community or any external collaborator for that matter and they will be able to mimic any results obtained on the cluster. In this context, a difficulty we often face working on a variety of HPC systems is that sometimes findings on one system cannot be reproduced on another and that is a true challenge for a closer collaboration on the source code. We do not always have the option to run on each other's systems and this impacts our ability to help each other pursue these issues. Reproducibility of results is a very important issue for NWP and by using AWS this is something that we can now easily achieve. In principle there is nothing that prevents either our own community, AWS or someone else to establish a complete NWP stack in the future that would mimic the on-site HPC experiences that we have been used to but currently this is not the case.

Once the stack is established, it's easy for others to use it and the experiences of working with AWS ParallelCluster become very close to what we are used to from on-site HPC systems, i.e. one will ssh into a head node and launch jobs via a queuing system in the exact way as we know from on-site systems. As an end-user, the differences between running on AWS versus on-site HPC systems seem negligible. The more technical details of what it takes to build the stack and Harmonie itself is shared at **HarmonieAWS on github.com**. Moreover, some information on initial performance results of the UWC-W benchmark test case can also be found in this AWS blog post. We were able to share the generated benchmark AMI containing the binary MASTERODB file with another specific AWS account. This was achieved without sharing any of the common source code or any scripts from the Harmonie system. With this AMI and a static input set, Karthik Raman from AWS was able to establish a Harmonie cluster on AWS and confirm that all reported findings with respect to running the binary MASTERODB on AWS were indeed reproducible from another AWS account. It should be stressed that the focus so far has been on getting a new cycle and corresponding new domain to work properly, completing the port to AWS of the Harmonie source as well as the scripting system and not so much on performance, neither on the on-site system nor on AWS. Having said that, we have not been able to find indicators to suggest that we will not be able to get comparable performance on the same kind of node-type at the scales that are relevant to our interests. Before this proof of concept work, it was unclear to us if the interconnect and the Lustre file-system offering at AWS would be able to keep up with those that we find on the on-site systems at scale. Our initial performance studies show that the performance is on-par when we stick to the same node-types.

We have faced several porting challenges in the process but as mentioned above this is expected when moving to a new system (and in dealing with a new untested cycle of the code and a brand new established testcase) but this is not our focus in this paper.

2 The AWS cluster

As described in the previous section, the full cluster is specified by two files, namely the binary Amazon Machine Image (AMI) with the complete Harmonie software stack including all Harmonie binaries and the config file for AWS ParallelCluster which contains a reference to the AMI that will be installed on the HEAD node as well as the compute nodes. So while the AMI specifies the full software stack on the Amazon EC2 Instances, the config file describes the infrastructure of the cluster. The details of this are all available on github.com/Hirlam/HarmonieAWS. Here we will confine ourselves to stress that the two files are all that we need to run a Harmonie HPC cluster on AWS. Now, the AMI files shall obviously be updated whenever there is an update to the source code. In order to handle this maintenance burden, we did a proof-of-concept where we added the building of the AMI into our experimental CI/CD pipeline. The artifacts of this build job is the logfile for establishing the AMI, the newly generated AMI-tag and finally a ready-made config file referring to the newly established AMI. After the job has completed successfully, one can take the config file from the job artifacts and create and access the cluster as shown below:

```
pcluster create mycluster -c artifact/config.c5n -r us-east-1
pcluster ssh mycluster
```

Figure1 illustrates the components that are specified by our sample config file and how they interplay whereas figure2 shows the end of a gitlab CI/CD job log pertaining to the build of the AMI. For those interested in the technical details please visit the github.com project.

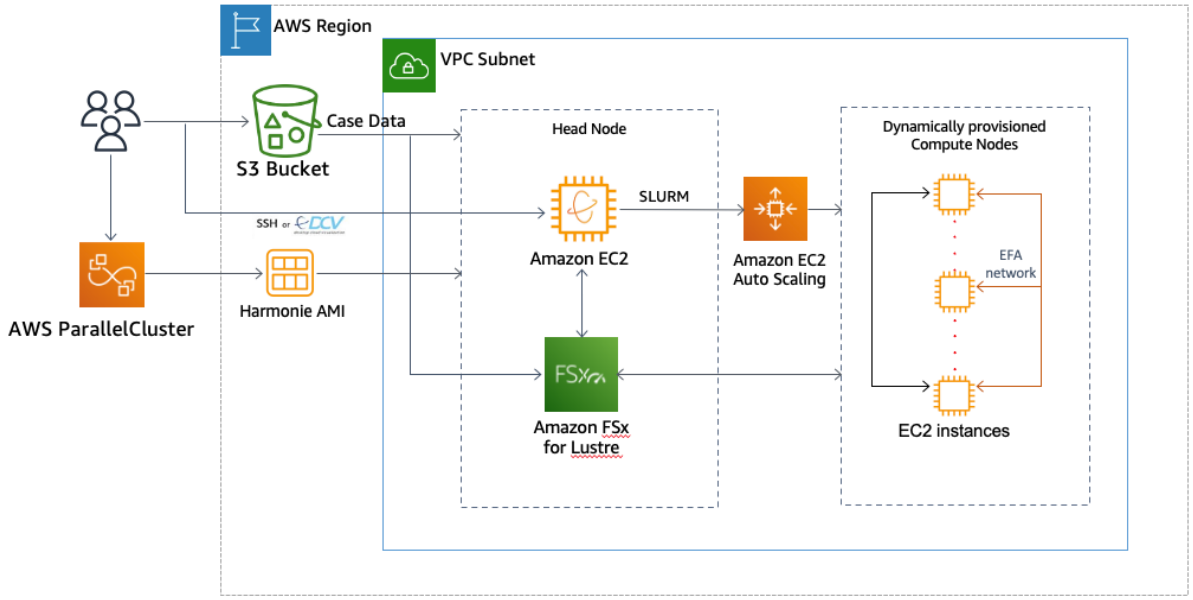


Figure 1: High level description of the Cluster as specified by the config file for AWS ParallelCluster software. The Amazon S3 bucket (object storage) is used to store the test cases whereas Amazon FSx for Lustre is the Lustre parallel file system used while running the model. The EC2 instances are c5n.18xlarge with Intel Skylake-SP CPUs and supporting the Elastic Fabric Adapter (EFA) interconnect (100Gbps throughput using libfabrics via the Intel MPI). The user interacts with the cluster through SSH into the head node and from there via the Slurm queuing system.

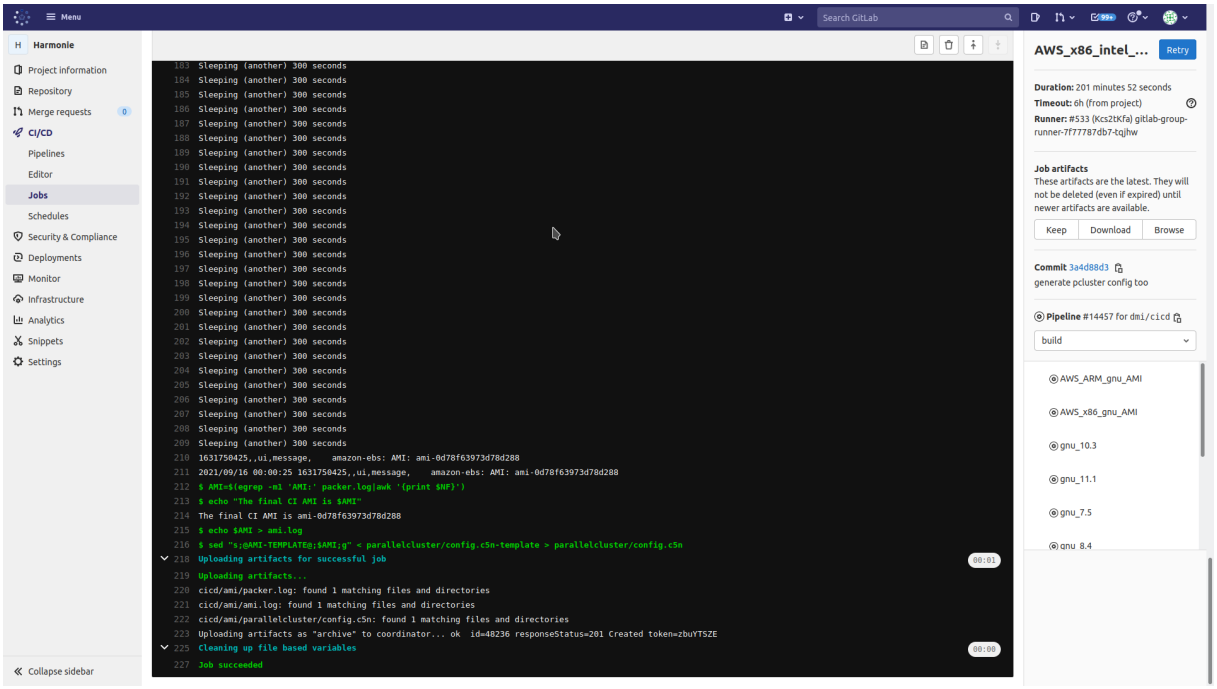


Figure2: The final job output from a gitlab CI/CD stage where the AMI is being built and eventually stored in the AWS EBS for later usage. The config file is adapted to the auto-generated tag for this AMI and kept as an artifact of the job so that one does not have to manually adjust a config file to point to the new AMI before creating the new cluster.

3 Use-cases for HPC in the cloud

While the experience of running the Harmonie-arome model system on AWS as reported in this note is merely at an initial stage, looking into the use of public cloud computing in the coming future, there seems to be numerous use-cases in the fields of weather forecasting and climate research.

Extreme HPC: It could be relevant for some of the extreme HPC research projects, e.g. where one needs to run at scales that is not possible on any on-site installation or with processor types that are not featured in the present on-site systems today. One such example is the development of the digital replica of the Earth system for different scenarios in climate change and extreme hazards. As one of the hypotheses is to examine the forecast capability for weather extremes with hyper-resolution, it necessarily needs experimentation of forecast software on newer generations of the HPC hardware and software environment to illustrate relevant scenarios. The ability to establish systems that can run across multiple architectures, say x86, Arm, GPUs and accelerators could also open new opportunities that we currently do not have. Such types of proof of concept experiments in a fast-forwarding manner may best benefit from the large public cloud services which could provide a future-like environment with exascale capacity.

Long simulations: Another obvious use case (assuming cost is not the limiting factor) would be reanalysis or reforecasting. As these types of work often involve duplication of the production suites for a long period, it may be challenging to secure HPC resources on-site. Amazon Spot Instances Pricing may become a good alternative for such. Similar could be suggested for climate simulations and in general research projects, where funding and hence planning rarely fit into the on-site investments in terms of capacity and schedule.

Backup solution: For applications in routine operational production, the public cloud, given its abundance in capacity and easy access with no need of advanced scheduling, also appears to be a

candidate for targeting a disaster backup service**.** Should the on-site HPC become unavailable for a larger period of time, one could consider switching to public cloud until the on-site service is restored. Naturally, such types of production system switchover requires preparation in the model system infrastructure. Essentially, this calls for an externalisation of input and output data flow in an operational NWP production chain, so that the public cloud services like AWS can be used realistically as a backup platform for a limited time period, but the process could be rewarding in the coming future.

Boosting capability: There are of course limits to how many ensemble members that makes sense to run on a regular basis but the limiting factor here is most often *not* set by this limit by merely dictated by the capacity of the on-site HPC. In extreme weather cases where say a 100-years event appears to build up (assuming that is apparent well ahead of the occurrence of the event) and this is foreseen to have a devastating impact, a naive wish could be to temporary increase the quality of the forecasts, e.g. by running additional ensembles in the cloud (or another data centre). Naturally, as stated above, this requires that cloud capability is already built into the production chain and that the ensemble scripting system is capable of handling a varying number of members or combining results from sub-ensembles.

Cloud computing in the context of operational NWP seems to become increasingly relevant, given the trends towards an increasingly larger scale joint operational production such as those in the UWC (with participation of 10 European national weather services), and the needs for exascale capacity and energy efficient computing in hyper-resolution modelling such as those in the EU DestinE initiative. Many inspirations as well challenges arise in the process of porting the Harmonie-arome system onto AWS. Essentially, we need a robust and versatile system that is transparent across hardware and software environments such as HPC platforms, processors, compilers, etc., ensuring an easily portable system across HPC platforms with reproducible forecasts. To secure such a platform equivalence, there is in addition a need for efficient schemes to perform systematic and comprehensive checks of technical and meteorological qualities. In this connection, establishment of a model system around an easily available repository with infrastructure for continuous integration and continuous deployment pipeline is necessary.

4 Acknowledgement

The research reported here have been conducted in association with the preparation of operational weather forecasting suites at the Danish Meteorological Institute (DMI), and the benchmark development at the United Weather Center - West (UWC-W) project, a collaboration for joint operational weather forecast production between national weather services in Denmark (DMI), Iceland (IMO), Ireland (Met Éireann) and the Netherlands (KNMI). The model systems in both of these works are based on the Harmonie-arome forecast system, a joint development in the HIRLAM-C NWP research consortia through collaboration with partners in the ACCORD programme involving 26 European national weather services and ECMWF. The numerical experiments have been carried out on the HPC platforms on the DMI Cray XC50, ECMWF Cray XC40, and Amazon Web Services. We acknowledge contributions by the AWS HPC team (Karthik Raman in particular), colleagues in the UWC-W NWP working group and DMI weather research group, without which the achievements reported here would not have been possible.

Forecasting surface currents using neural networks

Hrvoje Kalinić, Jadranka Šepić, Ivica Vilibić, Vlado Dadić, Hrvoje Mihanović, Damir Ivanković, Jasenko Blaž, Nedjeljka Žagar, Simone Cosoli, Ivica Janeković, Martina Tudor

1 Introduction

This paper describes a downstream application of forecasting surface currents using neural networks. The work was done during a two year project, NEURAL, from 2013 to 2015. The main goal was to find the link between the ocean currents and the wind in an advanced statistics kind of a way and get a finite number of sea surface current spatial patterns that are related to specific wind spatial patterns. This link would allow forecasting the surface currents using wind field as input.

2 Neural networks

Input data

Here we will use 2km resolution wind forecast from the operational nonhydrostatic run using ALARO CSC and surface current measurements from the area obtained with ocean radars. Two sets of ocean current radar measurements were used, one from the Northern Adriatic for the period from 2007 to 2010 (part of the data was used for training and another part was used to verify the forecast performance). The other dataset comes from two ocean radars installed in the central Adriatic. Both wind and ocean data are available with hourly interval.

A lot of input data is needed to train the artificial intelligence (AI) / machine learning (ML) algorithms well and more data should lead to better results. In this work, the hourly input data were quality checked, de-tided and interpolated (with some spatial gap filling). Often, radar returns measurements on only part of the area covered by the radar and this area varies in time. Therefore the ocean current data coverage varies over the domain and is lowest close to the edges.

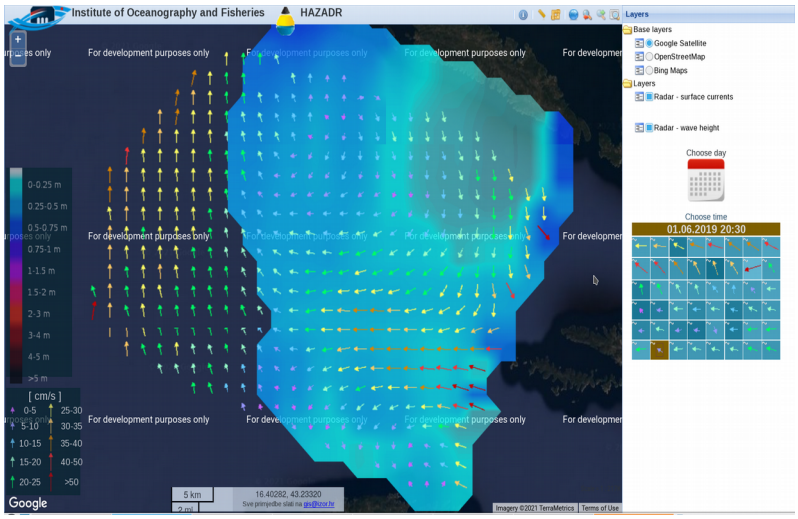


Figure 1. Example of the measurements from the ocean radar system, ocean currents are shown in arrows and the wave height is shown as shaded. The data coverage varies.

Self organizing maps (SOM)

In this work, we use Kohonen neural networks or self organizing maps (SOM) method applied on the wind and ocean current fields (not just one gridpoint at a time). The basic assumption is that local surface current field depends on the wind field more than anything else. Therefore, only the wind field would be used as a driver for the SOM ocean surface current forecast.

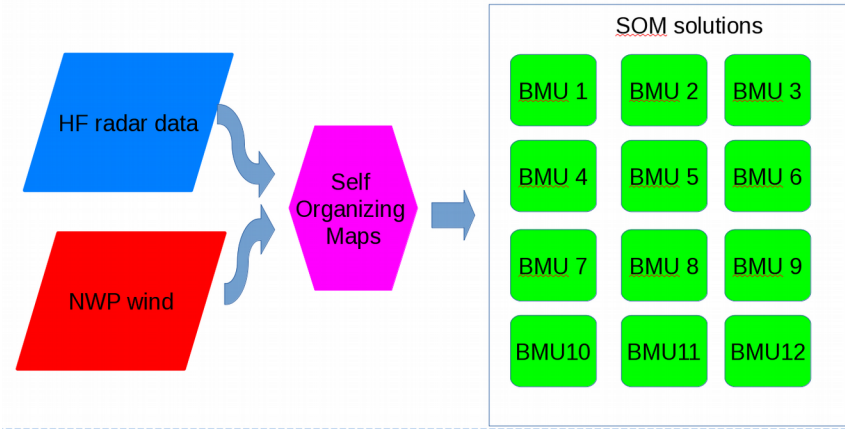


Figure 2. Illustration of the SOM training. Measured ocean current data and wind forecast fields are fed into SOM algorithm to yield 12 possible pairs of solutions.

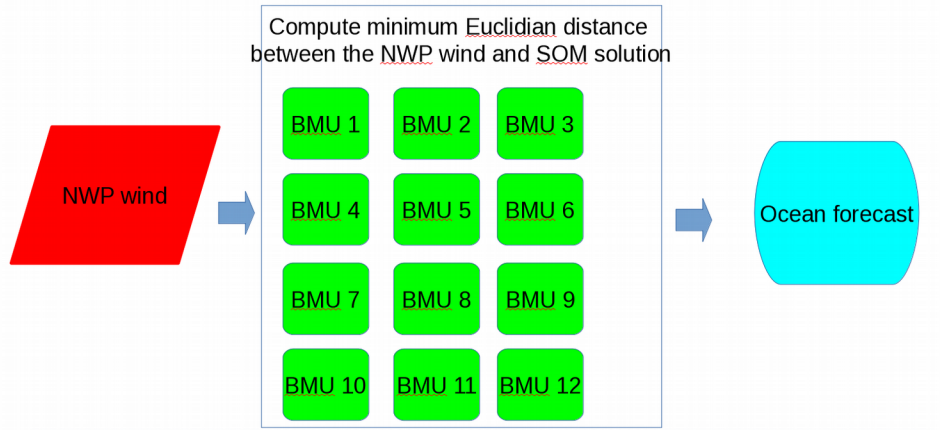


Figure 3. The SOM forecast takes the forecast wind field as input and finds the closest BMU, the ocean surface current field associated with the BMU is the ocean forecast.

The more nods (BMUs in Figures 2 and 3) we use, the larger data set we need for training. On the other hand, the more nods we use, the forecast error is lower as the system is able to represent larger set of the forecast fields.

3 Results

The ocean surface currents field forecast using SOM can only take one of the shapes that was determined by SOM, one of the nodes. The larger number of nodes, the larger number of shapes that the field can assume, therefore, the forecast error reduces substantially with increasing number of

SOM nodes, but at the same time we need more data to train the SOM. Forecast error is also large close to the edges of the ‘domain’.

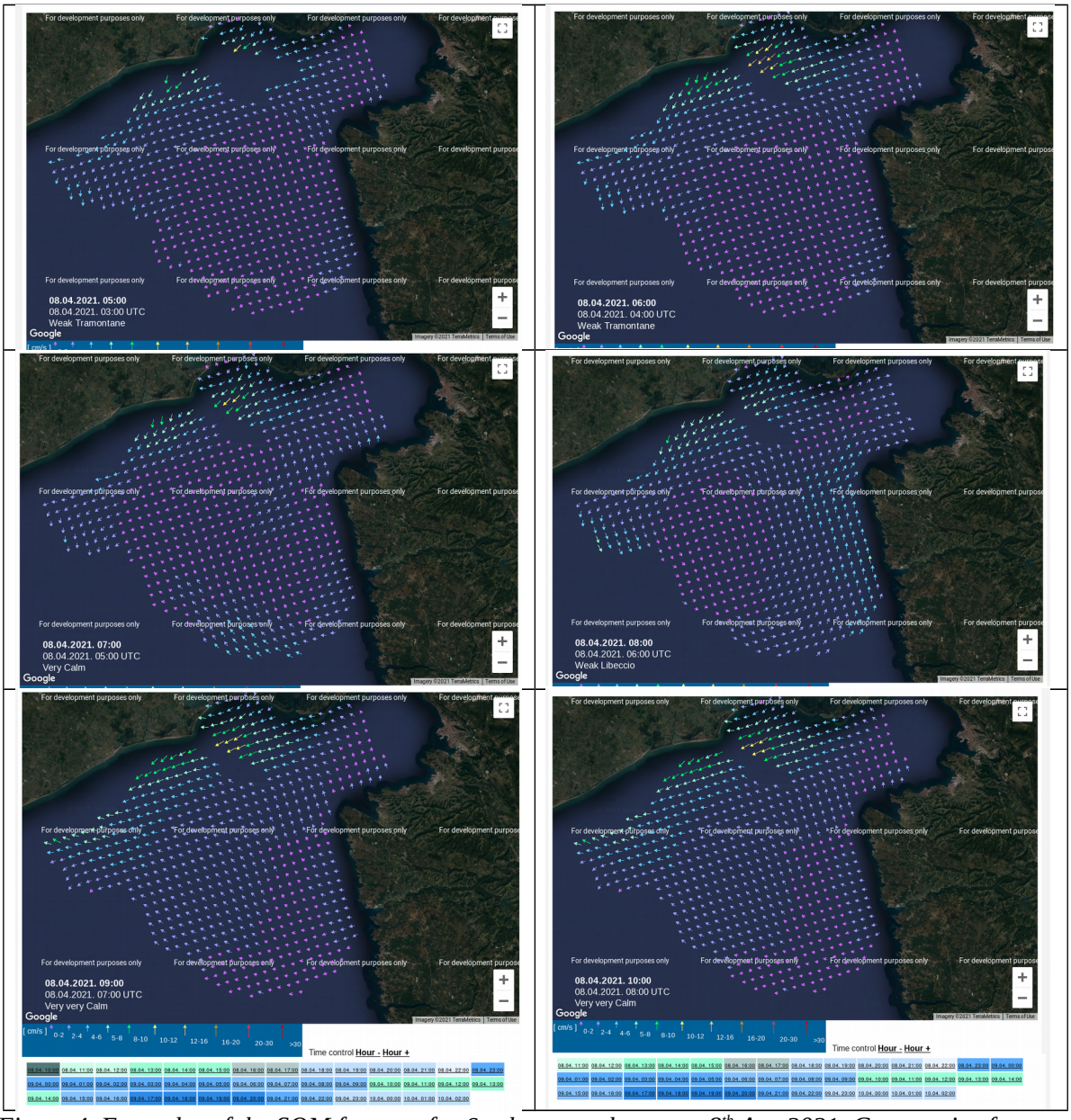


Figure 4. Examples of the SOM forecast for 6 subsequent hours on 8th Apr 2021. Consecutive forecast fields can be identical if the input wind field belong to the same node (shape of field).

Conclusions

AI/ML methods are only as good as the input data. In the work presented here, SOM give an organized spatial pattern for the ocean currents forecast and provide reliable forecast of the surface current field. The same method can be applied to other ocean fields provided they are driven primarily by the meteorological conditions.

There are also limitations to the method. The variability of the ocean current forecast field is substantially reduced to the number of BMUs, consequently the forecast is constant then jumpy when changing patterns. The method is reliably applicable only where extensive measurements are done (alternatively, one could train SOM on an ocean or wave model) and in this form it is useful on very limited areas due to limited number of patterns.

4 References

Voronoi G (1908) Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Journal für die Reine und Angewandte Mathematik 133:97–178

Kohonen T (2001) Self-organizing maps, 3rd edn. Springer, New York

Mihanović H, Cosoli S, Vilibić I, Ivanković D, Dadić V, Gačić M (2011) Surface current patterns in the northern Adriatic extracted from high-frequency radar data using self-organizing map analysis. J Geophys Res 116:C08033

Kalinić H, Mihanović H, Cosoli S, Vilibić I (2015) Sensitivity of self-organizing map surface current patterns to the use of radial versus Cartesian input vectors measured by high-frequency radars. Comput Geosci 84:29–36

Kalinić H, Mihanović H, Cosoli S, Tudor M, Vilibić I (2017) Predicting ocean surface currents using numerical weather prediction model and Kohonen neural network: a northern Adriatic study. Neural Comput & Applic (2017) 28 (Suppl 1):S611–S620

Vilibić, I. et al. Self-Organizing Maps-based ocean currents forecasting system. Sci. Rep. 6, 22924; doi: 10.1038/srep22924 (2016).

Data assimilation progress at Météo Algérie

Mohand Ouali AIT MEZIANE[1], Ghiles CHEMROUK[1]

(1)Prévision Numérique du Temps, Office Nationale de La Météorologie,, 1 Avenue Mohamed Khemisti BP 153 Dar el Beida-16011- Alger. Algérie.

1 Introduction

In the first part of this work, we present the state of the data pre-processing and data assimilation at Météo Algerie. In the second part, we show the first results of current work about monitoring of observations for GPS ground and TEMP.
As part of the twinning program between Météo France and Météo Algérie, a program of monitoring of observations has been initiated. With the support of Météo France experts and our colleagues in ACCORD consortium, we have increased the number of observations used in our DA system and extract information from ODB's.
We have performed an experience of ALADIN-3DVar assimilation using cy43t2 with different types of observations: GPS ground, BUOY, SYNOP, TEMP and AMDAR. Further, we present some results for GPS ground and TEMP.

2 Data processing status

The next table shows the different observation types processed.
For surface observations, we do processing of SYNOP (main, intermediate, and non-standard), SYNOP ship, and BUOY. At upper air, we do processing of TEMP and AMDAR. For satellite data, we do processing of ASCAT winds and SATOB.

DATA	**SYNOP**				**BUOY**	**GPS ground**	**TEMP**	**AMDAR**	**ASCAT**	**SATOB MSG wind**
	Main	Inter	Non standard	Ship						
Time	hourly			hourly	hourly	hourly	hourly	hourly	hourly	hourly
Header	SM, ISM	SI, ISI	SN, ISN	ISS	IO	ISX (EGRR)	IUS, IUX, IUK	IUA	-ISXX (EHDB) -W_XX (12.5km , 25.0km)	IUVA, IUVD, and IUVE
Format	BUFR, ASCII	BUFR, ASCII	BUFR, ASCII	BUFR	BUFR	BUFR, ASCII	BUFR, ASCII	BUFR, ASCII	BUFR	BUFR
Tools	Bufr_dc , bash script	Bufr_dc , bash script	Bufr_dc , bash script	Bash script	Bash scrip	ecCodes python program	ecCodes bash script	EcCodes python program	EcCodes python program	Bash script
Number	**3237.25**			/	**480.51**	**25580,07**	/	/	/	/

3 Status of data assimilation

Tables below show the AROME-3DVar and ALADIN-3DVar configurations over Algeria domain.

Table 1: ALADIN cy43t2 (3DVar) configuration

Model version	ALADIN-3DVar
Resolution	8 Km, 450 * 450 grid points
Levels	70
Type of initialisation	First ARPEGE coupling file
Initial conditions	ALADIN Analysis 3DVar
Boundaries	ARPEGE
Surface scheme	no
Cycle interval	6 hours
Forcast range	72h at 00h, 12h

Table 2: AROME cy43t2 (3DVar) configuration

Model version	ALADIN-3DVar
Resolution	3 Km, 500 * 500 grid points
Levels	41
Type of initialisation	First ALADIN coupling file
Initial conditions	ALADIN Analysis 3DVar
Boundaries	ARPEGE
Surface scheme	no
Cycle interval	6 hours
Forcast rang	48h at 00h, 12h

Table 3 shows the different configurations of data assimilation at Météo Algérie. We point out that our suites were in pre-operational phase before the breakdown of our HPC. After that, in order to continue our work we've rebuilt the ALADIN-3DVar and AROME-3DVar cy43t2 at Météo France HPC (belenos).

Table 3: Status of data assimilation

Configurations	Status	Cycle	Machine	Cycling
ALADIN 3DVar	Pre-operational	Cy43t2	Local HPC	06 hours
	Tested	Cy43t2	Belenos (MF)	06 hours
	Tested	Cy43t2	Belenos (MF)	03 hours
AROME CANARI OIMAIN	Tested	Cy40t1 and 43t2	Local HPC	06 hours
	Not yet	Cy43t2	Belenos	06 hours
AROME 3DVar	Pre-operational	Cy43t2	Local HPC	06 hours
	Tested	Cy43t2	Belenos	R00 hour

4 Monitoring of observations

We have performed an experience of assimilation using ALADIN-3DVar with different observations during two weeks: from 16-03-2021 to 30-03-2021, with 06 hours forecast networks: 00h, 06h, 12h and 18h.
In surface we used SYNOP, SYNOP SHIP, GPS ground and BUOY. In the following we present some results of GPS ground.

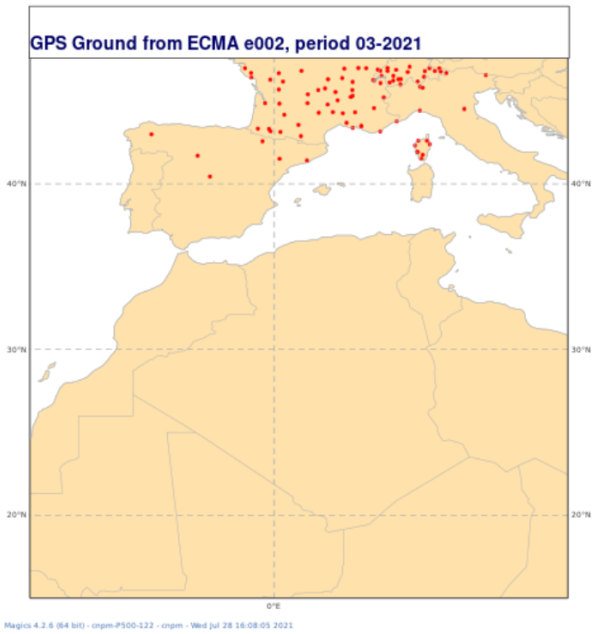


Figure 1: Location of GPS ground observation

The figure 1 shows the geographic location of GPS ground stations used in our experience. We ALADIN Algeria domain.

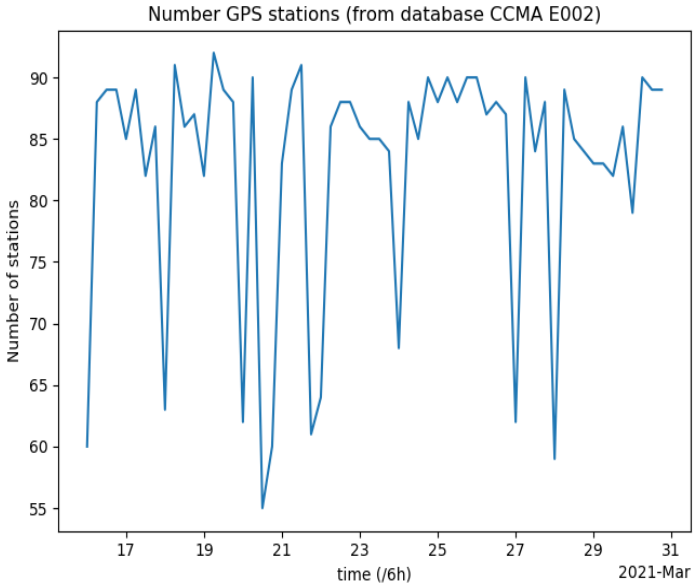


Figure 2: Daily number of GPS station used

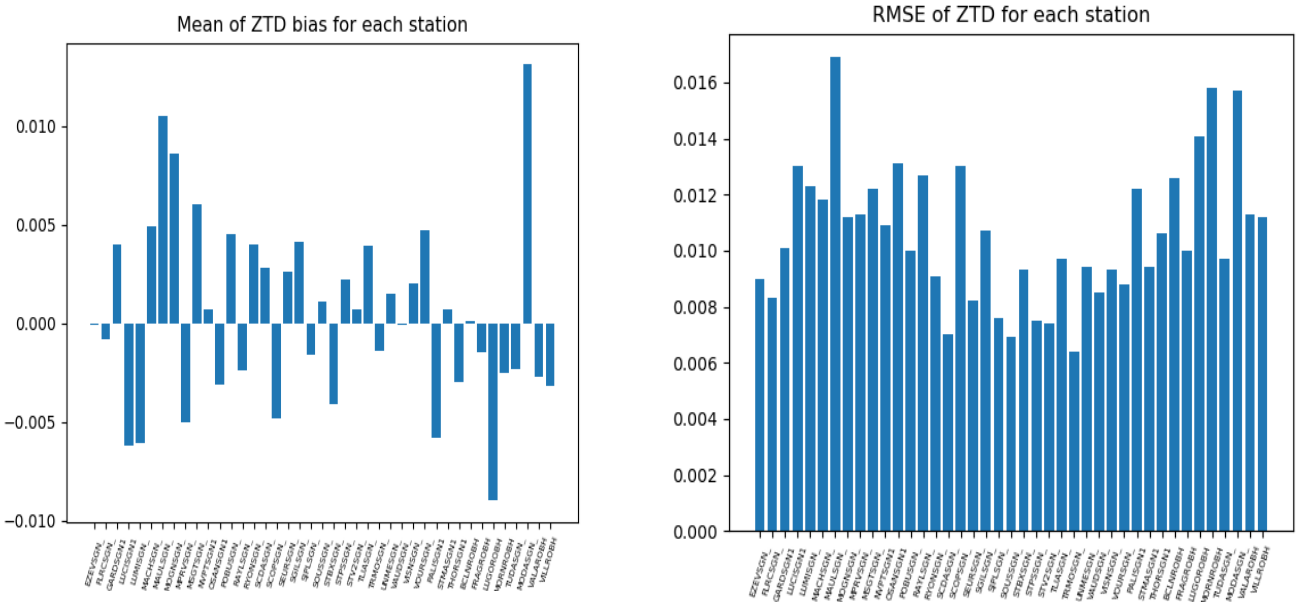


Figure 3: Mean of ZTD bias (on the left) and ZTD RMSE (on the right) for each station

The number of ground GPS stations in the screening step varies between 55 and 93.

mean : 82.73, std : 10.06 Total Number : 4963.
min : 55.00 max : 92.00.

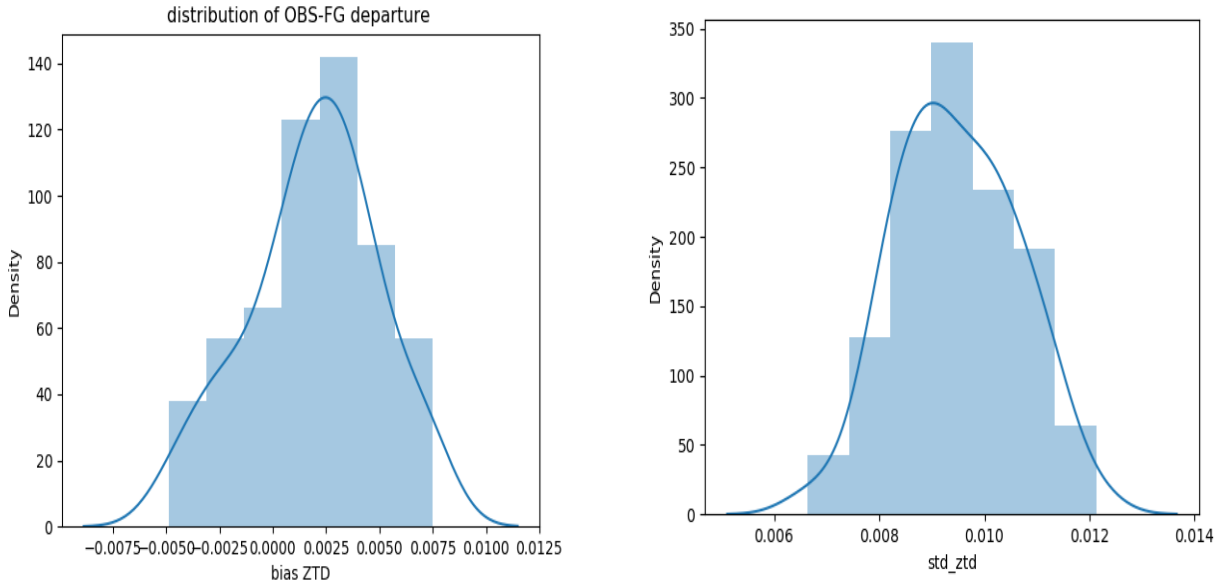


Figure 4: Distribution of OBS - FG departure

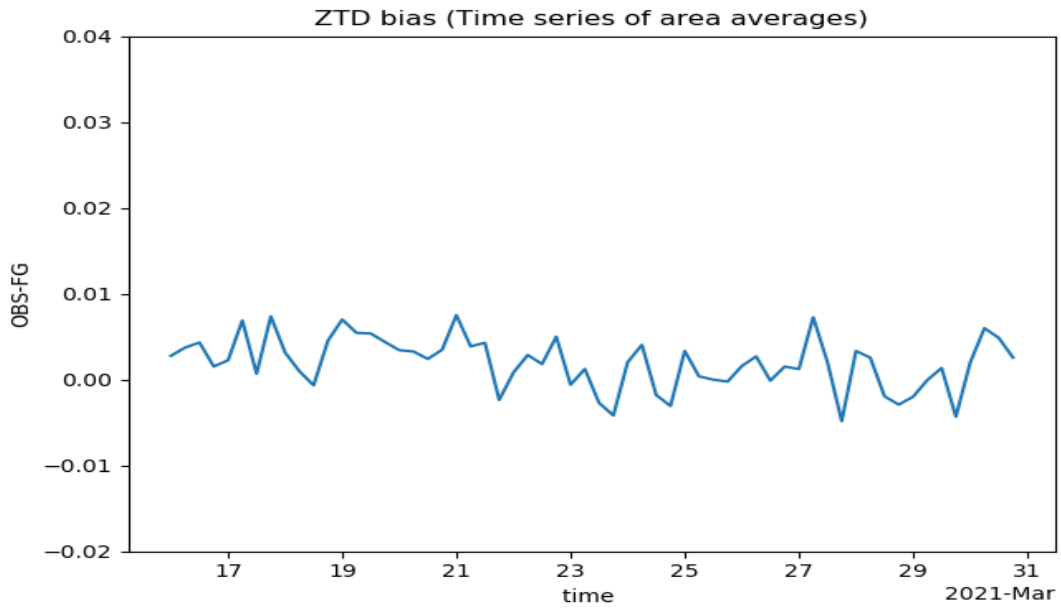


Figure 5: Time series of ZTD bias

The figure 5 on the left shows the error distribution of atmospheric zenith total delay (ZTD), the figure on the right shows the distribution of standard deviation of atmospheric zenith total delay.

Here, we present the monitoring part of TEMP observation. In the table below we show a list of TEMP stations over ALADIN Algeria domain.

Station id	Station Name
60390	DAR-EL-BEIDA
60571	BECHAR
60680	TAMANRASSET
60630	IN SALAH
60656	TINDOUF
7645	NIMES-COURBESSAC
16080	MILANO/LINATE
16245	PRATICA DI MARE
16320	BRINDISI
16429	TRAPANI/BIRGI
16113	S. P. C. MOLINEL
16045	UDINE/RIVOLTO
16046	Mezzanego

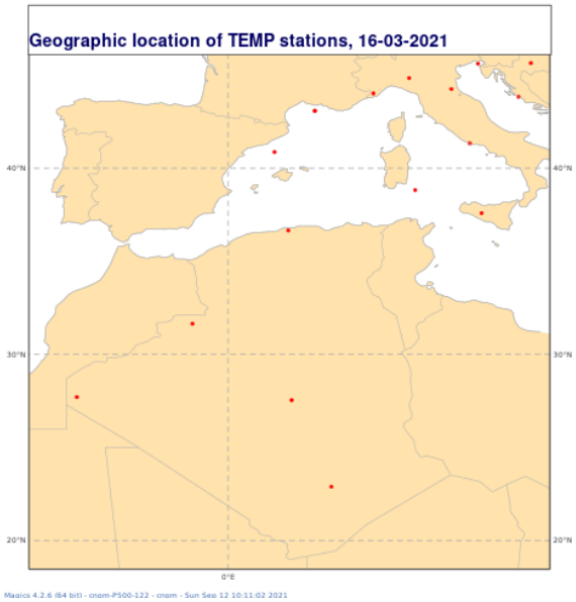
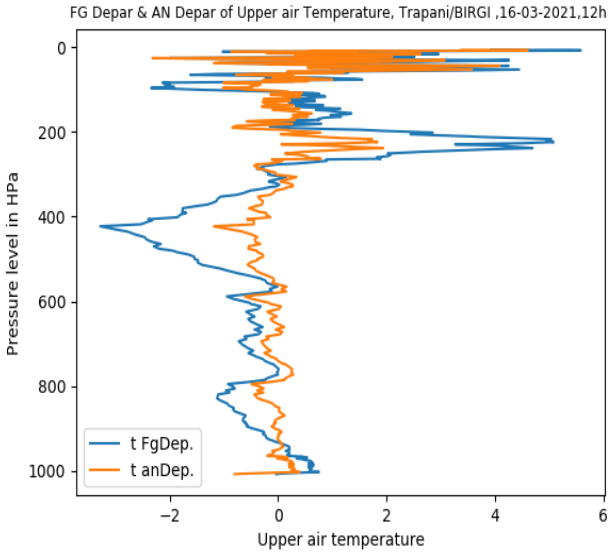
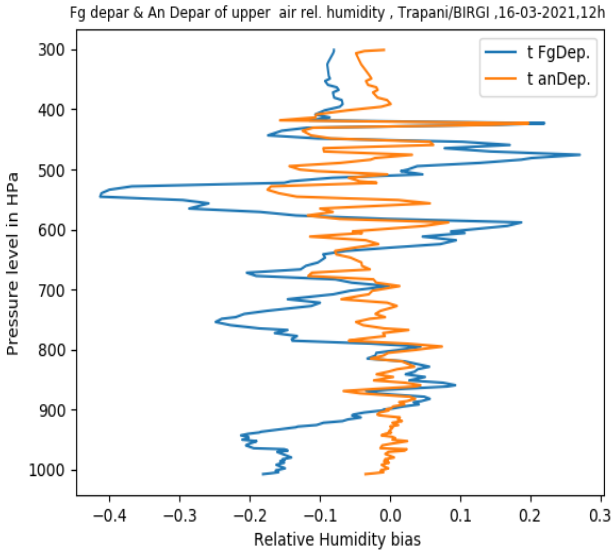


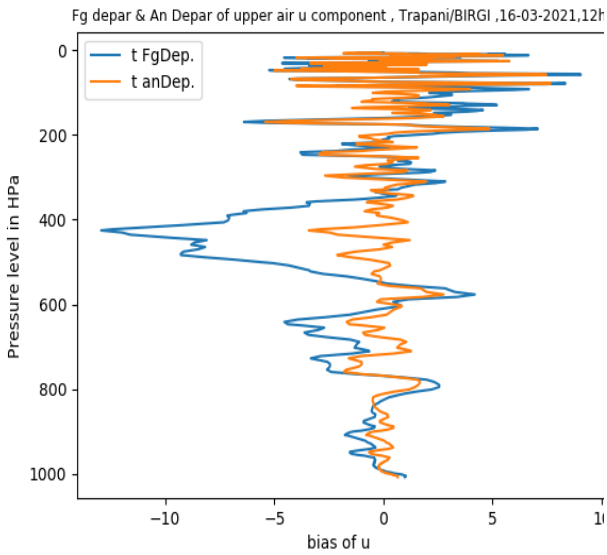
Figure 6: Location of TEMP stations over ALADIN Algeria domain

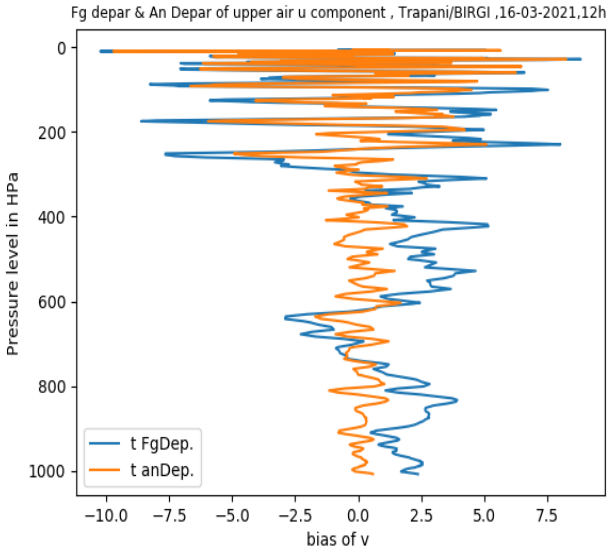




fg_depar : mean = 0.2272 , std = 1.6196
an_depar: mean = 0.1928 std = 0.8814

fg_depar : mean = -0.0845, std = 0.1164
an_depar : mean = -0.0226, std = 0.0501





fg_depar: mean = -0.4399, std = 3.5601
fg_depar: mean = 0,1902, std = 1,8607

an_depar : mean = 0.7509, std = 3.3915
an_depar : mean = -0,0515, std = 2.3596

Conclusion

We have presented the state of our preprocessing system and the different assimilation suites: ALADIN-3DVar and AROME-3DVar. We've increased the number of observations assimilated such as SYNOP ship, BUOY and GPS ground.
In the last part we presented the monitoring of observation. The first results of upper air observation: TEMP, shows a positive impact in the atmospheric state (obs - analysis departures, obs - first-guess departures). For GPS ground, we evaluated different statistics: the BIAS and RMSE of ZTD for available stations over ALADIN domain.

Acknowledgement
I would like to thank Mr Bruno Lacroix, expert of Météo France, for his support. I also would like to thank our colleagues in the ACCORDS consortium : Mr Idir Dehmous, Ms Alena Trojáková and Mrs Maria Monteiro for their support in our work.

References

Máté, M., Patrik, B., and Szabolcs Rózsa, 2019:The use of GNSS zenith total delays in operational AROME/Hungary 3D-Var over a central European domain. 12, 1569–1579.

Hdidou, F.Z., Smoumia, M., Patrick, M., Jean-François, M., Hassanae, E and Zaineb, D, 2019:Impact of variational assimilation of ground-based GNSS zenith total delay into AROME-Morocco model.https://doi.org/10.1080/16000870.2019.1707854,

DAsKIT progress report

Maria Monteiro, Ouali Aitmeziane, Haythem Belghrissi, Andrey Bogatchev, Yelis Cengiz, Lesley De Cruz, Alex Deckmyn, Idir Dehmous, Fatima Hdidou, Wafa Khalfaoui, João Rio, Zahra Sahlaoui, Malgorzata Szczęch-Gajewska, Boryana Tsenova

1. Introduction

The purpose of this report is to depict the status and progress of the coordinated effort of the ACCORD consortium (former ALADIN, LACE, HIRLAM consortia) towards the implementation of operational Data Assimilation (DA) activities in countries where the initial model state is not obtained by use of locally available atmospheric observations. The reporting period ends in March 2021.

This effort of the ACCORD consortium appears as a legacy of the ALADIN Strategic Core Programme known as 'DAsKIT' [1], which has become a Working Package (WP) of the annual ACCORD Rolling Working Plan (RWP), and its focus is now on building of local DA capacities. It combines the synergies from ten countries, namely: Algeria, Belgium, Bulgaria, Estonia, Morocco, Poland, Portugal, Romania, Tunisia and Turkey.

The report has five parts dealing with: Section 1, this introduction; Section 2, the progress registered since last year; Sections 3 and 4, the on-going local and global activities; and finally, Section 5, relating the expected goals for the WP in the near future.

2. DAsKIT status by country

During the past three years of duration, the activities and working material of the DAsKIT Programme has been registered in the ALADIN website (see http://www.umr-cnrm.fr/aladin/spip.php?rubrique74). In 2020, an article was published in the joint ALADIN-HIRLAM newsletter [1] where the Programme was introduced and the DA starters KIT was described.

Since then, two important achievements for the shared effort included:

i) the synchronization of the working code's cycle in all partners (the local implementation code version is now CY43T2); and,

ii) the upgrade of the DAsKIT set - an AROME based testbed over the Mainland Portuguese geographical domain (PT2) with Iberian surface observations - to a combined solution of surface (by Optimal Interpolation, OI) plus upper-air (by 3-Dimensional Variational, 3D-Var) DA.

Synopsis on the countries status

The status of the participating countries is illustrated by Table 1. In order to share efforts in an optimised way, an *a priori* list of tasks was proposed to the participating countries (also identified as working packages on the ALADIN annual working plan) that should guide them in its local work. The list included the local implementation of the original DAsKIT set, which was created to solve the surface AROME DA cycling by OI in an affordable way.

In Table 1, the columns represent different RWP tasks while the rows represent its status in the country.

Table 1: Synopsis on the DAsKIT status by country (*)

	Data Acquisition			Data pre-proc		Data monitor		BATOR	Surface DA		B-matrix		3D-Var DA legacy		Combined DA		Validation tool	
Country	GTS SYNOP	GTS TEMP	GTS AMDAR	HOME MADE	UNIVERSAL	OTHER	OBSMON	CONV. DATA	CYCLED	OPER	SPIN UP	EDA	CYCLED	OPER	CYCLED	OPER	HOME MADE	HARP
DZ																		
BE																		
BG																		
MA																		
PL																		
PT																		
TN																		
TK																		

(*) The country is indicated by ISO Alpha Code 2. The column top cells are 'blue' if a task is in its final form or 'light-blue' if the task is in an intermediate step; the row cells are 'green' if the task is complete or 'light-green' if the task is complete in a sub-optimal way. The lack of local computer power, meaning the country is using a remote computer, is neglected here.

It is pleasant to note that every country succeeded in getting conventional data in-doors and even to create home-made pre-processing tools. Besides, at least four countries have also implemented universal pre-processing tools.

Furthermore, it deserves to remark that all the countries succeed in implementing the OBSMON tool for diagnostic and data monitoring purposes, although the tool is still used in a sub-optimal way. To counter this observation and what concerns BATOR, all countries got effective on its usage.

Moving to the algorithms and in what concerns the model surface, it can be seen that all the countries were able to cycle a surface DA scheme. Noticeably, the Belgium team has already implemented it into operations, based on the DAsKIT set.

Concerning upper-air, many countries already computed a first B-matrix although a few have yet to proceed with its computation. Moreover, a couple of countries already computed a second, Ensemble

Data Assimilation (EDA) based B-matrix and half of the countries are now able to locally cycle a 3D-var scheme.

Regarding the use of a validation and verification tool, it can also be seen that all the countries already had some home-made validation tool although, at this moment the Belgium team is being able to use the universal package HARP on a regular basis (Poland also has installed it in-doors, waiting to reach operational status).

By looking at this synopsis, it becomes obvious that some effort has to be allocated towards the implementation of local upper-air algorithms and, in particular, to support the local computation of B-matrices.

3. Local work highlights

In this Section, some of the on-going local activities are highlighted.

Surface data assimilation

From the surface DA cycling implementations, it deserves to be emphasized the operational setup of Belgium, which is illustrated in Figure 1. It consists on a 3-hour cycling of the CANARI_OI-MAIN algorithm created from the surface DAsKIT set [1] for the local AROME-Belgium model with a 1.3km of horizontal resolution using GTS and local SYNOP data from local pre-processing, and as coupling model a local version of ALARO-Belgium at 4.0 km horizontal resolution.

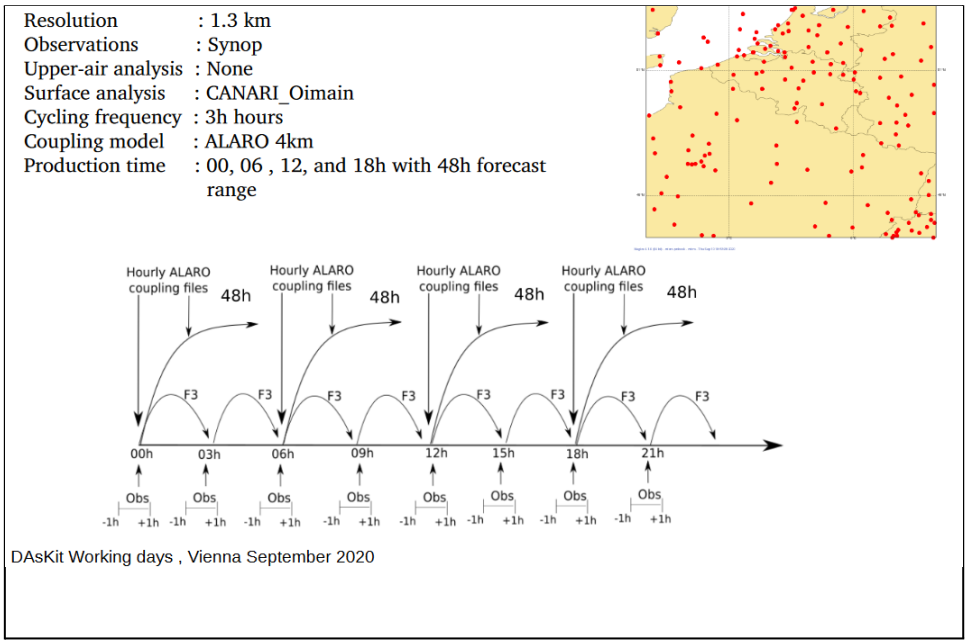


Figure 1: Schematic representation of the 3-hour operational surface DA cycling system implemented at the Royal Meteorological Institute (RMI) in Belgium for AROME-Belgium at 1.3km resolution, using local and GTS surface observations (SYNOP) under WMO BUFR template.

In addition, it is worth mentioning the improved scores obtained by the Turkish team with the local surface DA implementation (also created from the original DAsKIT set) as displayed in Figure 2. The improvement is noticeable by comparing with last year's preliminary validation results for the 2-metre

parameters. This advance was due to a fertile tuning study on the analysis and surface schemes, previously discussed with Météo-France (to further details, read point 5 of "Recommendations and actions" in [2]).

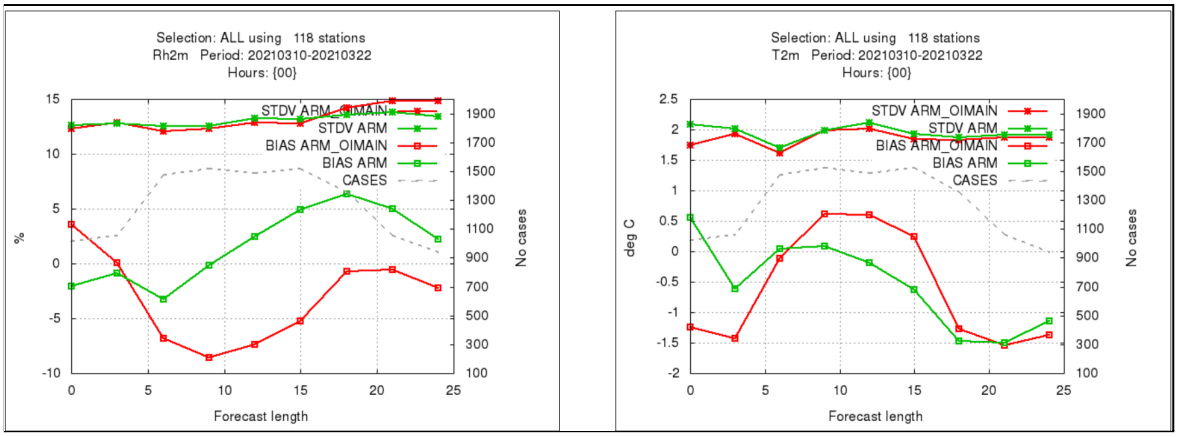


Figure 2: Standard deviation and bias along the lead time, of 2-metre relative humidity (%, left) an 2-metre temperature (deg C, right) of AROME-Turkey (1.7 km, IFS coupling model) forecasts, with the different initialisations: dynamical adaptation (green line); and from a surface DA (red line).

Finally, a comment is noted on the preliminary results of the Portuguese surface DA system implemented at ECMWF. It was created from the same DAsKIT set and uses locally shared Iberian surface observations. For instance, it was shown that - for a Winter period - the spatially averaged initial model state of the 2-metre relative humidity becomes closer to the observations than its reference (Root Mean Square Error, RMSE, as metric), which is the actual operational initial conditions of AROME-Portugal, locally known as "AROME-PT2", obtained by dynamical adaptation.

Combined data assimilation

Concerning the combined DA cycling implementations, results from two countries are illustrated in Figures 3 and 4.

Figure 4 illustrates some of the results obtained with the local combined DA solution implemented for ALARO-Belgium with a 4.0 km resolution. For comparison, the values of RMSE for 2-metre temperature forecasts, with assimilation, are shown together with those from the dynamical adaptation of ALARO-Belgium at 4.0 km (red line) and at 1.3 km resolutions (blue line). Following the magenta line in the plots (model integration with assimilation) we can see a clear added value is due to the assimilation of observations on ALARO-Belgium at 4.0 km (which is not seen just by increasing the resolution). Furthermore, it was concluded that the same level of implementation efforts of a combined DA solution for AROME-Belgium at 1.3 km (with several conventional observations) seems to have a neutral impact over the 2-metre parameters scores when compared to the surface DA solution.

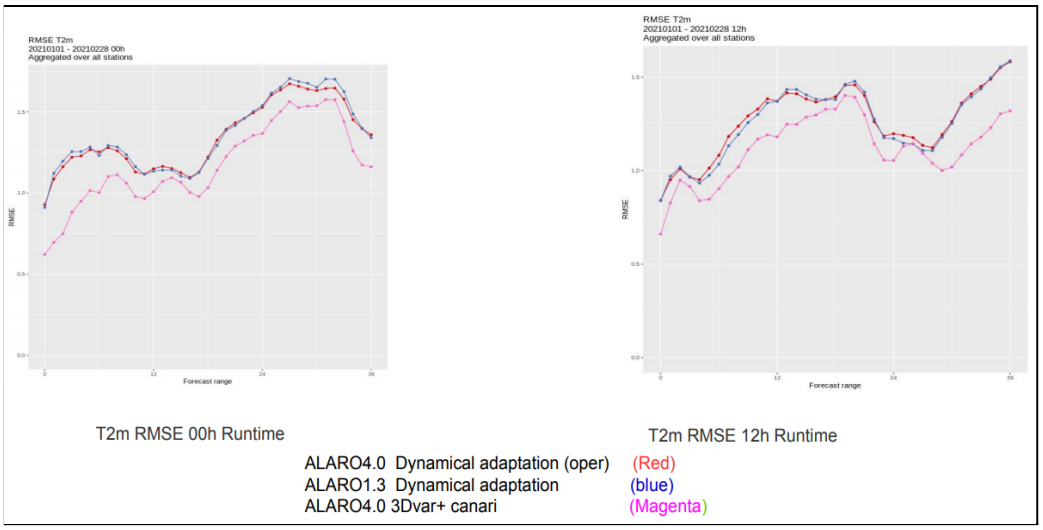


Figure 3: RMSE along the lead time, of 2-metre temperature (ºC) of ALARO-Belgium at 4.0 km resolution, when the initialisation is done through a combined DA solution, using just conventional observations, on a Winter period: starting at 00 UTC (felt panel) and starting at 12 UTC (right panel).

In addition, Figure 4, shows some preliminary results obtained with the implementation of the combined DA solution at ECMWF for AROME-Portugal. It is possible to see from these preliminary diagnostics (observations minus guess (OMG) *vs*. observations minus analysis (OMA)), that, for 2-metre relative humidity - during a Summer period - the analysis departures to the observations become smaller than its initial guess departures which seems to indicate a successful implementation.

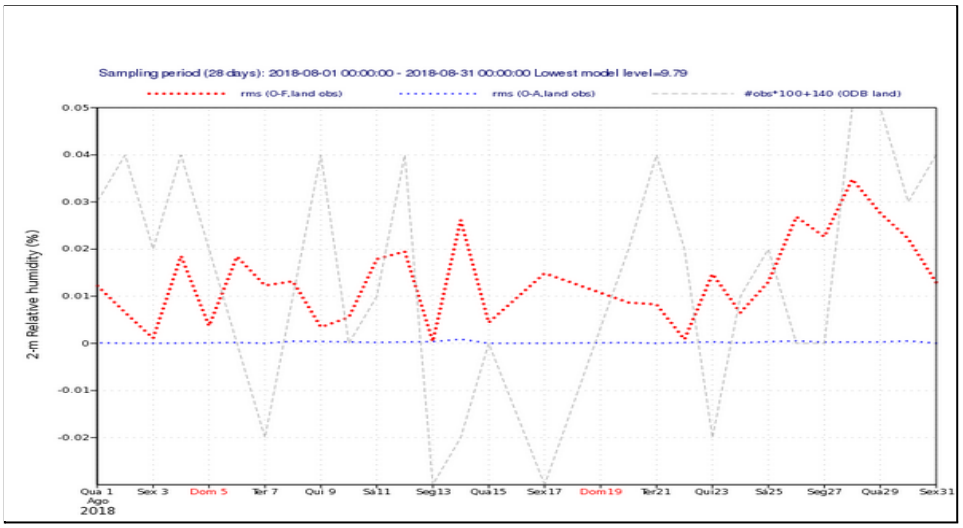


Figure 4: Time series of departures to observations, of 2-metre relative humidity (%) of AROME-Portugal fields, during a one month Summer period: first guess departures (red line); and analysis departures (blue line).

Data monitoring

It starts to be obvious that the increased teams maturity is driving a new DA mind-set, in particular the need to implement good practices. In this way, data monitoring studies in view of the optimal use of the local observations (not illustrated here) are a common concern. However, as mentioned in Section 2, just a few countries started regular data assimilation monitoring activities. Yet, there is much work to be done here. Examples of new observation types at play are: Ground-based GNSS, Mode-S and radar volumetric data.

Local optimisation studies

Finally, we would like to mention the recent research investment some countries have put on their local DA systems optimisation, to overcome the sparseness or other singularities of the local observational networks: by assimilating new types of observations, as it is the case of Morocco [3], [4]; and by adding a Jk blending component in their DA algorithm, as it is the case of Tunisia.

4. Further progress: scripts and communication

From the experience of working together, further outcomes from this group came out of the original Programme, therefore two announcements on further progress take place in this Section: the first concerns the way to share the DA workflow and still keep it affordable; and the second concerns a recent enhancement of the communication platforms.

Sharing the DA workflow

In this section, an on-going exercise is described, on how the DA workflow can be shared in the near future. During last year, two scripting systems that solve DA algorithms at CY43T2, have been running on ECMWF platforms, both based on ecFlow (the ECMWF scheduler; see summarised info at www.umr-cnrm.fr/aladin/IMG/pdf/daskit_scripts_setting.pdf). One is kept by Portugal and results from an extension of the Slovenian scripting system to the AROME model configuration. The other, called 'NodeRunner', is the Belgium scripting system and results from a re-furnishing of their original operational system. The former is a Python interface which uses the concept of Python objects to define 'tasks' and 'families' while the latter, being also a Python interface to ecFlow, uses the concept of external initial files to generate the different experiment settings.

Keeping in mind that the two systems are available under the same computing environment, a coordination effort took place to analyse both systems in order to understand which one possibly got the best design, so as to ensure an optimised way of how to share the DA workflow among DAsKIT participants. In parallel, a query was launched over the participating countries aiming to understand its needs in terms of scripts. It was possible to conclude that all countries were willing to, or at least open to, moving to ecFlow. Therefore, a discussion was set up about the best way to share the workflow and is illustrated by the diagram on Figure 5. In the fathermath, a first announcement can take place here on the DA workflow sharing: a new version of 'NodeRunner' scripting system is being prepared with the concepts of Python objects brought from the Slovenian scripting system. Besides, there are some expectations that this system may be extended with the tasks for B-matrix computation.

The diagram in Figure 5 tries to illustrate the evolution of the balance between 'shared effort' vs. 'degrees of freedom to implement a DA solution', which is behind the decision to create the new version of 'NodeRunner'. In this context, one should understand the "shared effort" as the joint effort to develop and maintain a common tool, hence supported by all participating countries; and 'degrees of freedom to local implementation' as a function directly related to the local affordability. In the Figure, the shadowed areas in 'green' and 'blue' represent the amounts of shared effort and

affordability, respectively. In this way, starting from the original DAsKIT set which was created as a set of shell scripts (panel a)), we can understand that with this sharing solution, the countries have full freedom on their local implementations. However, if instead of sharing a shell scripting system, we share ecFlow scripts and the experiment definition files (panel b)), then, we see the shared effort increases (we have more and more complex tools to maintain), but the workflow sharing form is still affordable to the countries.

In this situation: i) we increase the overall sharing effort which is more than compensated by splitting it over a larger pool of countries (as with those LACE countries, that keep the same scripting philosophy); ii) the countries can grasp the opportunity to rely on new standards, libraries and practices; and iii) the countries can still read and understand the basic DA workflow without much effort and eventually adapt it to their local scripting system.

At a higher sharing level, by considering the Python objects and suites' generator (panel c)), we see that the countries can still implement the DA workflow in a modular way and, with a certain amount of freedom in accordance with their actual resources. However, the local investment increases since some expertise on Python may now be required (in any case the workflow is still readable, even if one country is not able to follow the latest progress).

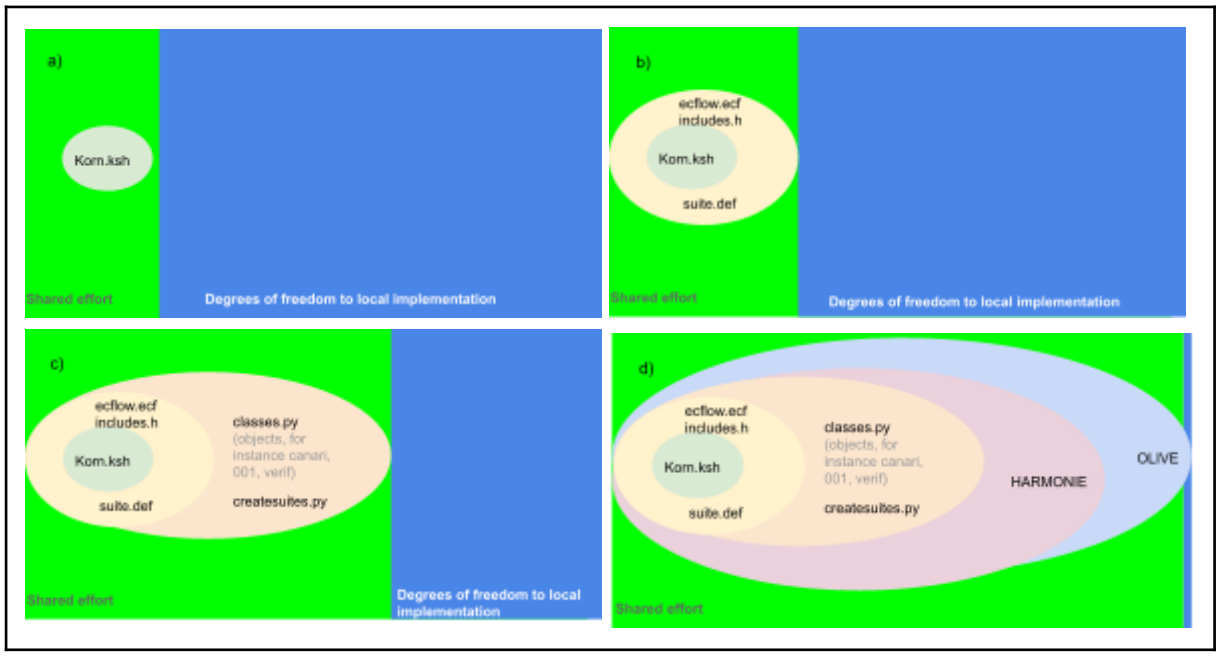


Figure 5: Diagram on the evolution of the balance between 'shared effort' and 'degrees of freedom to local implementation' of a common (affordable) tool to share the DA workflow, when the form of the tool relies on: a) a set of Korn shell scripts; b) a set of ecFlow files (scripts+includes); c) a set of Python objects and a suites generator (including the ecFlow set); and on the ideal case, d) a complete system, like OLIVE (Météo-France).

The complexity of this sharing level seems to be a turning point on the sharing form decision, since increasing the scripting complexity even further, may prevent the readiness comprehension of the workflow.

Progressing in this way we see that ideally, every country could work with the OLIVE environment (Météo-France) in a fully shared way, but this compels access to the same level of resources, which is out of reach at the moment.

A second announcement can take place here, which is the creation of the DAsKIT wiki page (available at http://212.175.180.89/daskit/index.php/Main_Page) by the Turkish Meteorological Service to enhance communication among operational DA newcomers.

5. Summary

The following list of achievements can be enumerated since last year:

1. every country knows how to use OBSMON under T-codes (ODB created with T-codes);

2. every country learnt how to process BUFR SYNOP coming from their local SAPP data pre-processing systems;

3. the Belgium pre-processing tool (POP-RMI) was implemented in Tunisia, as a proof of concept for the countries which do not have access to SAPP;

4. the working code cycle was synchronized in all countries;

5. the DAsKIT testbed was upgraded to the combined DA at CY43T2; and, finally,

6. the first local operational implementation of a DAsKIT surface DA system was put in operations.

For the next short-term period, our support team aims to implement new working dynamics, with less frequent status reports but more dedicated actions, and also getting gradually involved in further ACCORD DA activities.

The following activities are foreseen:

1. the local implementation of a combined DA solution of surface plus upper-air (3D-Var) in more countries;

2. an investment on the workflow sharing based on ecFlow scripts plus Python objects and suites generator, with a possibility to extend it to B-matrix computation;

3. the enrichment of the DAsKIT wiki page contents and eventually the inclusion of some training material. Finally,

4. it starts to make sense to see participating countries getting involved with further ACCORD topics, like those already proposed by the DA Area Leader: OBSMON development, a verification tool that may support all types of observations, etc…).

Acknowledgements
To Alena Trojakova for supporting the DAsKIT work in many ways. To Jean-François Mahfouf and Pierre Brousseau for the fruitful discussions on surface and upper-air aspects of the DAsKIT set. To Benedikt Strajnar for sharing the ARSO DA scripts at ECMWF. Claude Fischer and Roger Randriamampianina accommodated the continuity of the DAsKIT efforts in the ACCORD working plan.

6. References

[1] Monteiro, M., Aitmeziane, O., Belghrissi, H., Bogatchev, A., Cengiz, Y., Deckmyn, A., Dehmous, I., Hdidou, F., Khalfaoui, W., Kilonko, M., Rio, J., Sahlaoui, Z., Sezer, M., Szczęch-Gajewska, M., Tsenova, B., 2020, *The DAsKIT programme: status and plans*, ALADIN-HIRLAM NL 15, June 2020.

[2] Monteiro, M., 2020, Summary report on DAsKIT video-conference, 25 March 2020, ALADIN Core Programme on Data Assimilation (from http://www.umr-cnrm.fr/aladin/IMG/pdf/daskit_summary_report_25032020.pdf).

[3] Sahlaoui Z, Mordane S, Wattrelot E, Mahfouf J-F, 2020, Improving heavy rain forecasts by assimilating surface precipitation in the convective scale model AROME: A case study of the Mediterranean event of November 4, 2017. Meteorol Appl.; 27:e1860. https://doi.org/10.1002/met.1860.

[4] Fatima Zahra Hdidou, Soumia Mordane, Patrick Moll, Jean-François Mahfouf, Hassnae Erraji & Zaineb Dahmane, 2020, Impact of the variational assimilation of ground-based GNSS zenith total delay into AROME-Morocco model, Tellus A: Dynamic Meteorology and Oceanography, 72:1, 1-13, DOI: 10.1080/16000870.2019.1707854.

Data assimilation progress in RC LACE

Benedikt Strajnar and RC LACE DA Team

1 Overview

The article presents a yearly overview of data assimilation (DA) activities and developments in the RC LACE part of the ACCORD consortium up to spring 2021. The presented material was contributed by members of local data assimilation teams in the partner countries. The operational status is discussed in the first section, and this is followed by a summary of several ongoing activities in the upper-air and surface data assimilation.

1 Operational status

The RC LACE national weather institutes are currently operationally running 8 different NWP systems that are initialized by DA, as summarized in the Table 1. A large variety of different global and regional observations are used. The utilized assimilation algorithms are either 3D-Var, blending or BlendVar, in combination by the OI for surface analysis, applied in 6- or 3-hourly assimilation cycling. Those systems, together with two ensemble systems (A-LEAF, C-LAEF) which are not covered in this short review, present a backbone for members' operational products. Most of the centers now upgraded their suites to model version cy43 (Hungary and Slovakia switched in the beginning of 2021). The nowcasting scale is approached by two separate operational suites at hourly time resolution: the AROME-RUC (rapid update cycle) in Austria and the analysis-only system called VarCanPack in the Czech Republic (details in Table 2).

DA	AUSTRIA AROME	CROATIA ALARO	CZECH REP. ALARO	HUNGARY ALARO	HUNGARY AROME	SLOVAKIA ALARO	SLOVENIA ALARO
Resol.	2.5L90, 600 x 432	4.0L73 480 x 432	2.3L87-NH 1069 x 853	8L49	2.5L60	4.5L63	4.4L87 432 x 432
Cycle	40t1	38t1_bf8	43t2ag	cy43t2bf11	cy43t2bf11	cy43t2bf11	43t2_bf10
LBC	IFS 1h (lagged)	IFS 3h (lagged)	ARP 3h	IFS 3h (lagged)	IFS 1h (lagged)	ARP 3h	IFS 1h/3h (lagged)
Method	OI_main MESCAN + 3d-Var	OI + 3D-Var	OI + BlendVar	OI + 3D-Var	OI_main + 3D-Var	OI + DF Blending	OI + 3D-Var
Cycling	3h	3h	6h	6h	3h	6h	3h
B matrix	Downscaled LAEF 11 km	NMC method	EDA	EDA	EDA	-	Downscaled ECMWF ENS
Initialization	No (SCC)	No (SCC)	IDFI in production, SCC	DFI	No	No	No (SCC)
Observations	Synop + AS Amdar Geowind Temp ASCAT, Snowgrid/MODIS snowmask., Mode-S EHS	Synop Amdar/MRAR Geowind Temp Seviri	Synop + AS (soil) Amdar/MRAR/EHS AMV/HR, Profiler, ASCAT, Temp Seviri,	Synop + AS Amdar Geowind Temp, Seviri AMSUA/MHS	Synop + AS GNSS ZTD Amdar/Mode-S MRAR Temp	Synop + AS	Synop + AS Amdar/MRAR/EHS Geowind Temp Seviri AMSUA/MHS/IASI ASCAT/OSCAT E-GVAP ZTD (passive)

Table 1: Operational DA for NWP systems run by RC LACE countries.

DA	AUSTRIA AROME-RUC	CZECH REP. VarCanPack
Resol	1.2 L90 900 x 576	2.3L87-NH 1069 x 853
Cycle	40t1	43t2pt_op1
LBC	AROME 1h	-
Method	OI_main MESCAN + 3d-Var + LHN + FDDA	3DVAR + OI
Cycling	1h	-
B matrix	Static EDA + differences of the day	EDA
Initialization	IAU	-
Observations	Synop + AS, Amdar/MRAR/EHS national, EHS EMADDC, Geowind, Temp, Seviri, AMSUA/MHS/HIRS/ATMS/IASI (+ Metop-C), ASCAT, GNSS ZTD (Austria), GPSRO (OPLACE), Radar RH/Dow, INCA + AS at hig.freq., MODIS snowmask	Synop + AS, Amdar/MRAR/EHS, Geowind/HRWIND,Profiler, ASCAT, Seviri

Table 2: Operational DA for NWP-based nowcasting systems at hourly scale run by RC LACE countries.

2 Upper-air assimilation

2.1.1 Spin-up control in RUC

Especially in the hourly cycled systems such as AROME-RUC, run at the Austrian meteorological service (ZAMG), the spin-up may have a significant contribution on the quality of analyses and forecasts. To avoid initial shocks, the analysis can be applied incrementally to the first guess in time, using a procedure known as incremental analysis update (IAU). Several applications of IAU were tested as shown in Fig. 1. When IAU is applied to the hourly first guess, the solution is very smooth but probably inaccurate. One the other hand, a cycled experiment without IAU (or even the downscaled “open loop” forecast) results in increased spin-up of up to an hour. The solution for operations thus includes the application of IAU in two steps: first, an analysis, computed as combination of a first guess trajectory with observations, is added backwards in time during the repeated first guess integration between -45 minutes and the analysis time; in the following step the resulting state at analysis time is pushed towards the analysis over a very short time in the forecast (over first 7.5 minutes). This is seen as good compromise between accuracy and balance.

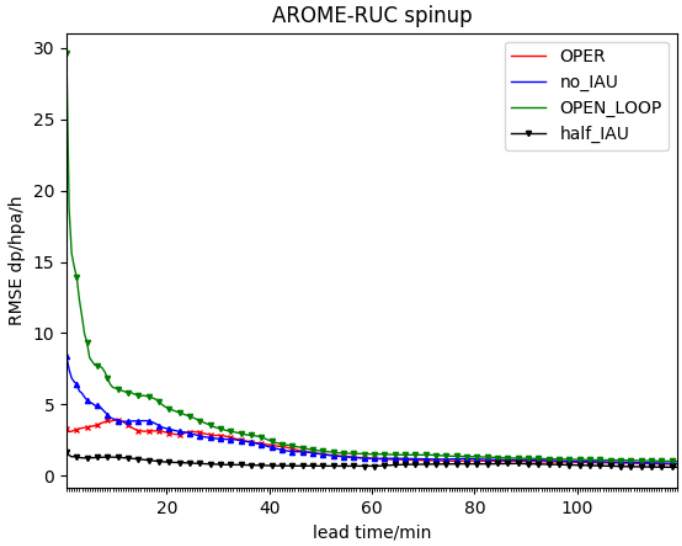


Figure 1: Root-mean-square of pressure tendency over the first 2 hours of AROME-RUC 1.2 km forecast. The runs correspond to downscaled run from AROME 2.5km (green line), cycled forecast without IAU (blue), IAU applied during the hourly guess (black) and the operational setting, explained in the text (red).

2.1.2 Tuning of the B-matrix for a new vertical geometry

At the Hungarian meteorological service, significant efforts were invested into preparation of a new B-matrix for a 90-level model setup by running a full EDA experiment based on the previously calculated B-matrix from the downscaled global EDA. Significant spin-up in precipitation was observed in the first hours of forecasts in assimilation runs (Fig. 2) and several experiments over winter and summer period were prepared in order to investigate this effect. The experiments included a-posteriori Desrozier's tuning of observation and background error (including a separate tuning for humidity), switching on or off the Canopy scheme, use of GNSS data assimilation and the stratospheric vertical balance. The experiments continued in 2021 but a satisfactory solution was not yet reached.

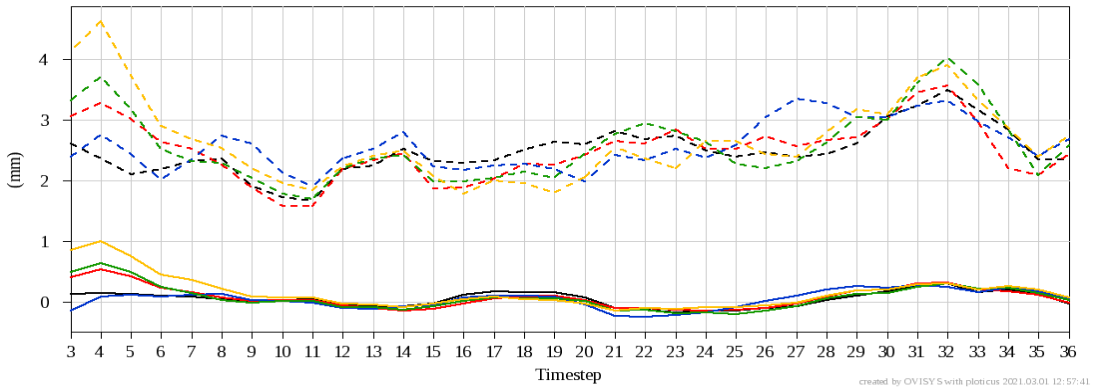


Figure 2: Bias (solid line) and RMSE (dashed line) of precipitation forecasts in the 0 UTC runs from 9 to 31 July 2020. Black: AROME/HU at 60 levels; blue: dynamical adaption run at 90 levels; yellow, red and green: differently tuned assimilation runs at 90 levels.

2.1.3 Radar data assimilation

Radar reflectivity and radial wind observations are seen as an essential component of a mesoscale DA system especially with increased analysis cycle frequency. It received more attention in LACE over the last years by several stays and regular videoconferences, mostly focused on application of radar assimilation in order to enhance the knowledge of the underlying humidity pseudo-observation algorithm and to investigate suitability of the reflectivity observation operator to be used with ALARO physics. The effect of graupel in the ALARO microphysics scheme was checked and while providing very similar precipitation structures, a small reduction of random error in reflectivity innovations was observed. The screening procedure for radar reflectivity was also investigated. One unresolved issue is assimilation of dry reflectivity observations with zero first guess departure which can only occur when model is dryer than the radar observation (i.e. below detection threshold) at a given location and elevation. It would be preferable to discard those observation from assimilation, but by default they keep being assimilated and produce pseudo-humidity columns with resulting positive or negative increments. An impact study with radar reflectivity was carried out in Slovenia within the operational 4.4 km and 1.3 km experimental RUC setting over a summer and winter period. The radar reflectivity was mainly shown to decrease the frequency bias of precipitation and to some extent the equitable threat score for smaller thresholds (Fig. 3). For selected cases, a too aggressive drying effect was observed for convective cases and this needs further attention. An initial study for assimilation of radial winds was also performed, using only radar sites with wind-optimized scans such as those from the German networks. For most of central Europe, an intermediate step is needed to dealias raw wind measure-

ments. A torus mapping method to accomplish this is under development as part of a stand-alone pre-processing software meant for easier handling and sharing within members.

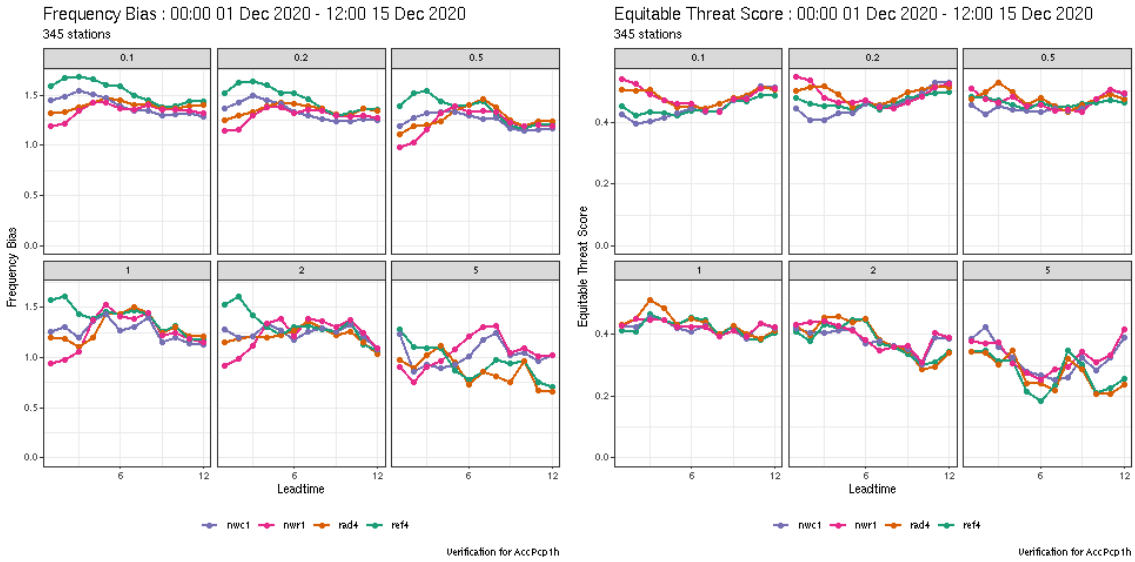


Figure 3: Frequency bias (left) and ETS score (right) of hourly precipitation forecast. Runs are Slovenian prototype hourly RUC (denoted nwc1), RUC + radar reflectivity (nwr1), operational ALADIN 4.4 km (ref4) and ALADIN 4.4 km + radar reflectivity (rad4). Period is 1-15 December 2020.

2.1.4 Increased use of Mode-S data

The Covid-19 epidemic caused a temporary drop in the amount of AMDAR observations, the data type with very high impact in DA systems run by LACE countries. To prevent drop in quality of forecasts, EMADDC of KNMI made an expedited progress in collection in distributing additional Mode-S observations. Their service now provides data from the MUAC area and additionally Denmark, Austria, Slovenia and Romania, and is on the way to ingest more data over the LACE area. The data was immediately shared/implemented with OPLACE, and some countries started to benefit from this very large additional data counts. The observation-minus-background error statistics of the European Mode-S EHS data (EHS_EU) have been checked for the two periods in October 2020, as illustrated in Fig. 4, separately by type and area. The overall EHS_EU wind statistics are now comparable with AMDARs. The separation by areas showed somewhat larger temperature standard deviation difference for the Romanian airspace which was reported back to EMADDC. Progress is visible also on use of local Mode-S data sources (Hungary), although the situation with decreased data counts is not ideal for evaluation of these data sets.

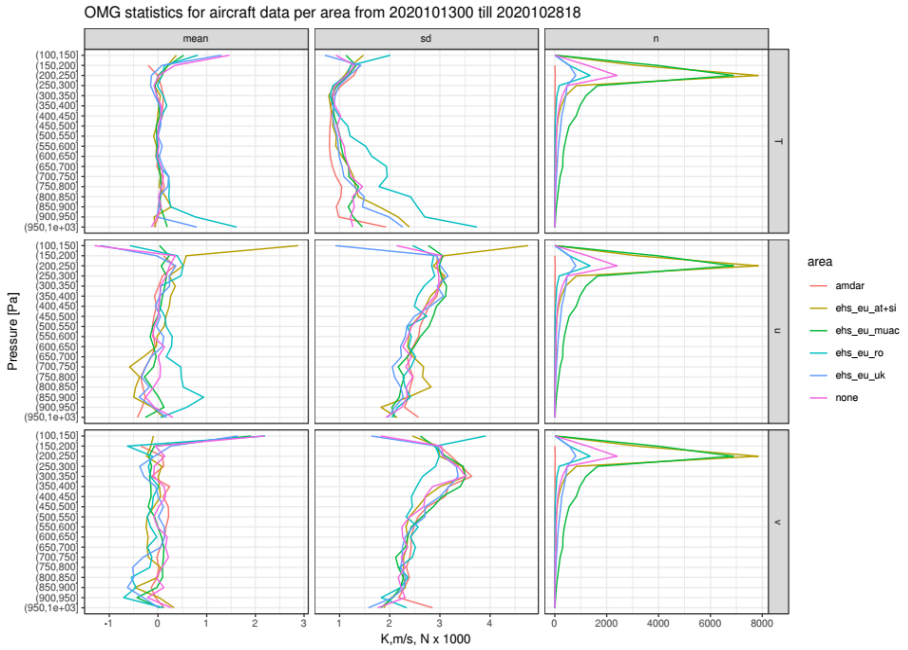


Figure 4: Vertical profile of aircraft innovations separated by geographical areas. BIAS (left), STD (middle) and number of innovations for all AMDAR data (red), MUAC area (green), UK area (blue), Romanian area (light blue), Austrian and Slovenian area (brown) for the period of 13-28 October 2020.

2.1.5 Extended use of mid-tropospheric AMVs

An impact study was carried out in Hungary in order to evaluate atmospheric motion winds from the MSG in AROME/HU over two periods with frequent weather fronts and convection. In the assimilation, AMVs were used in both long and short forecasts as the timelines of the data allows it. The AMVs were generated from IR, VIS and WV channels. Only a small fraction of the data was active in the assimilation of initial experiments with lots of blacklisted data (above QI 85%). Most of the active wind vectors were located between 300 and 250 hPa and a few measurements were used between 1000 and 850 hPa. Experiments showed very small, mostly neutral impact of the AMV data in both periods for the surface parameters (temperature, humidity, wind, pressure). In the convective period a small and rather positive effect can be seen for the surface wind gusts. The impact was also supported by differences in the SEDI verification score of 24-hour precipitation amounts. Impact studies continued in 2021 where additional mid-troposphere observations were added, based on good departure statistics in passive assimilation. These experiments showed additional small improvements.

3 Developments for surface data assimilation

Apart from tuning the currently operational OI in some LACE countries, an increased attention was given to validation of simplified extended Kalman filter (SEKF) and related surface observations, especially in Hungary, where several periods and interesting weather cases were studied over the Hungarian domain, using AROME cy40t1 and SURFEX 7.3. A 3-hourly data assimilation cycle was started two weeks before the events. Forcings required by the offline SURFEX run were coming from AROME inline forecasts at 9 m (radiation, precipitation, wind, humidity, pressure). A typical problem in AROME/HU that the minimum temperature is usually overestimated and maximum temperature is underestimated in long-live anticyclonic cases was addressed. More accurate night-time analysis and forecast was provided by the SEKF experiment, as shown on a case in Fig. 5. The 2 m temperature shows large improvement for the nighttime hours and the large warm bias in the nights was reduced by SEKF with respect to the operational OI-MAIN (Fig. 6). However, the daytime forecasts over

Hungary do not differ spectacularly from the OI-MAIN ones. Afternoon 2 m temperatures are sometimes overestimated due to too dry soil in the SEKF. Further experiments involve tuning of the observations and background error magnitude as well as perturbations to estimate the Jaccobians for SEKF analysis. A parallel suite is planned at both the current and improved (kilometer-scale) horizontal resolutions.

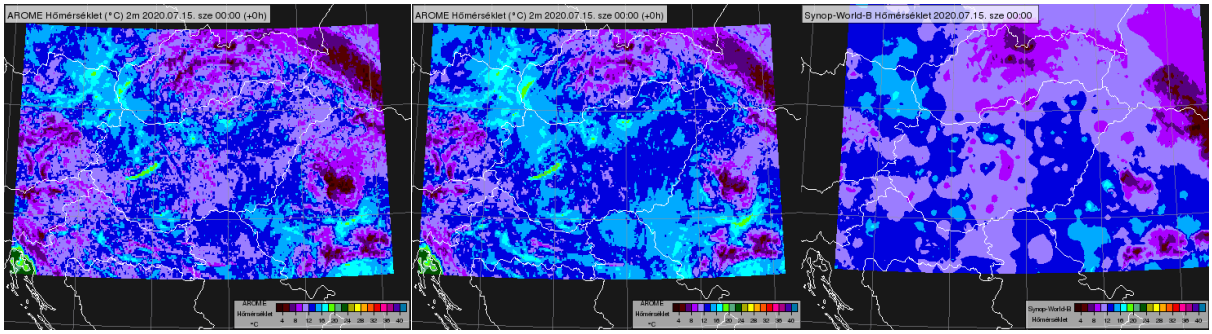


Figure 5: 2 m temperature analysis in AROME cy43 with OI-MAIN and SEKF, observations at 0 UTC on 15 July 2020.

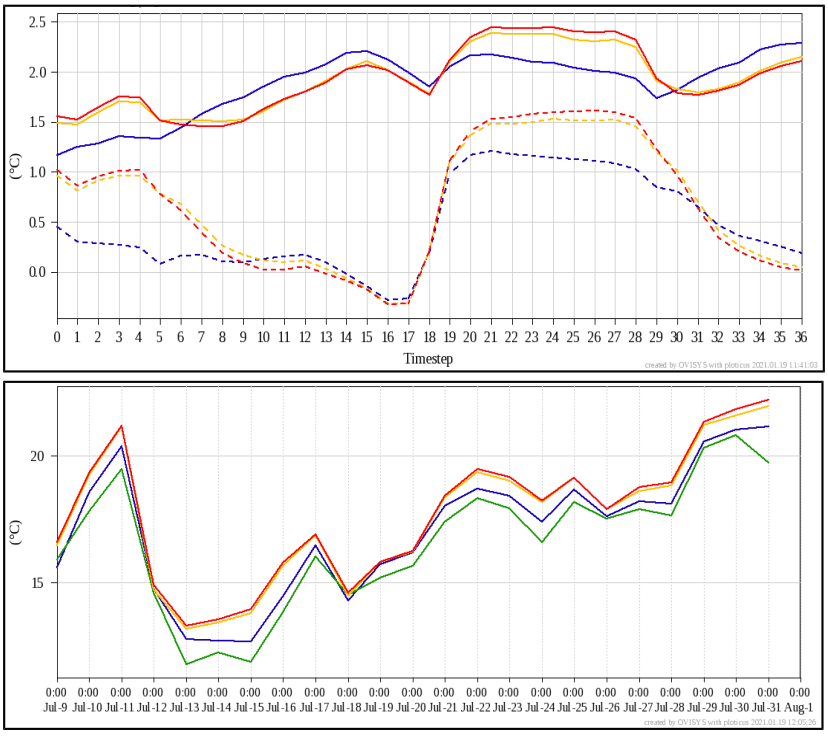


Figure 6: Evolution of 2 m temperature analysis and 12 h forecast in the 0 UTC runs and observations (green) from 9 July to 31 July 2020. Blue: cy43 with SEKF, orange: cy43 with OI-MAIN, red: cy40 with OI-MAIN.

4 Outlook

Effective use of observations stays the top priority of LACE DA group. One of the outstanding goals is to consolidate current and previous work on radar reflectivity and radial in order to increase the number of operational applications with enhanced success on the convective scale. Other important

observations with already much investment but rare operational applications are the GNSS-derived data. Use of other data such as wind profilers, atmospheric motion vectors, as well as new data types such as GNSS slant delays and microwave telecommunication link delays will also be enhanced and developed. In parallel, Additionally, the SEKF for soil assimilation within the SURFEX land model will be further validated with AROME model and progressively explored also in combination with ALARO model physics (provided that validation of coupling with SURFEX will be completed). This also includes sensitivity studies with additional observations for surface and soil. Finally, a progressive familiarization with the OOPS code infrastructure is foreseen, not only to replace the current 3D-Var with its OOPS equivalent in the relatively near future, but also to be able to utilize more advanced algorithms such as ensemble variational (EnVar) assimilation.

EPS research and development in RC LACE in 2020

Clemens Wastl, Martin Belluš, Katalin Jávorné Radnóczi

1 Introduction

Three ensemble systems within RC LACE (A-LAEF, C-LAEF, AROME-EPS) are currently running in full operational mode. They are running very stable and smoothly and the usage of EPS at the meteorological centres is constantly increasing. Also the scientific work is progressing well, new developments (multi domain fullpos in A-LAEF, stochastic perturbation scheme in C-LAEF and AROME-RUC Peps in Austria, EDA in Hungary, etc.) and EPS applications (visualizations, maps, EPSgrams) have been made in the RC LACE countries in 2020.

2 Operational systems development and upgrades

Table 1: Table with operational systems running by RC LACE member states.

System	Description	HPC
A-LAEF	Common RC LACE EPS with 4.8 km horizontal resolution based on ALARO-1 physics.	cca/ccb
C-LAEF	Convection-permitting EPS of Austria with 2.5 km horizontal resolution based on AROME physics.	cca/ccb
AROME-EPS	Convection-permitting EPS of Hungary with 2.5 km horizontal resolution based on AROME physics.	local

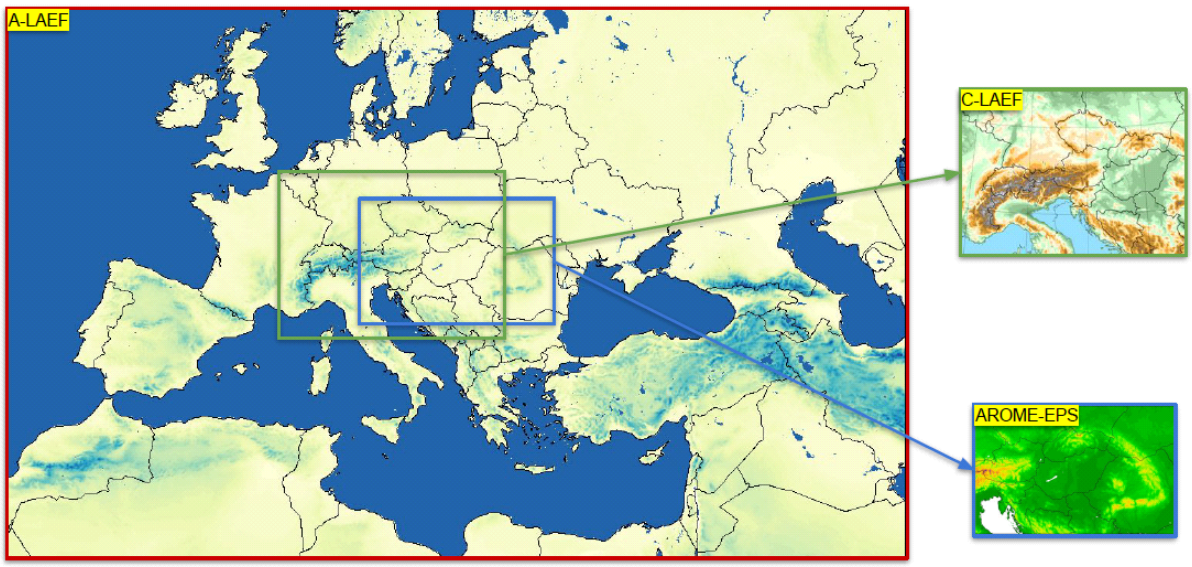


Figure 1: Domains of operational EPSs in RC LACE.

Table 2: Table with technical characteristics of the operational systems running by RC LACE member states.

	A-LAEF	C-LAEF	AROME-EPS
CMC	ALARO	AROME	AROME
Code version	cy40t1	cy40t1	cy43t2
Horizontal resolution	4.8 km	2.5 km	2.5 km
Vertical levels	60	90	60
Runs per day	2	4	1
Forecast length	+72h (00/12 UTC)	max. +60h (00 UTC)	+48h (00 UTC)
Members	16+1	16+1	10+1
Assimilation cycle	yes (12h)	yes (6h)	-
Coupling	ECMWF ENS (6h)	ECMWF ENS (3h)	ECMWF ENS (1h)
IC perturbation	ESDA [surface], spectral blending/DFI [upper-air]	ESDA [surface], EDA, Ensemble-JK [upper-air]	-
Model perturbation	ALARO-1 multi-physics + surface stochastic physics (SPPT)	hybrid stochastic scheme comb. of parameter and tendency perturbations	-
LBC perturbation	ECMWF ENS (c903)	ECMWF ENS (c903)	ECMWF ENS (c903)

A-LAEF

For A-LAEF following changes have been made in the operational suite in 2020:

- Addition of unperturbed control run (member=00)
- Prolongation of forecast to 72 hours
- Approval of time-critical 2 application by ECMWF on July 22, 2020
- Implementation of ECPDS dissemination of A-LAEF GRIB files to the user's destination (SK, SI, RO, TR, CZ + PL via local dissemination from SHMU)
- MSLP smoothing by spectral Gaussian filter within in-line fullpos (for all 16+1 members)
- Implementation of multi-domain off-line LATLON fullpos for Turkey and Czech Rep. (Figure 2)
- Conversion of LATLON FA files to multi-GRIBs and dissemination to Turkey and Czech Rep.
- New post-processed fields in A-LAEF Lambert multi-GRIBs (snow line, 0-isotherm height,
- @H100 - T, U, V, RH)
- Implementation of parallel ecflow suite (e-suite) at ECMWF for testing new A-LAEF upgrades
- and modifications
- Automatic download of VOBs files from ECMWF (for the HARP verifications)
- Preparation of probabilistic A-LAEF maps (Figure 3)
- Preparation of user interfaces for public website and SHMU intranet (public URL: http://www.shmu.sk/produkty/nwp/alaef/)

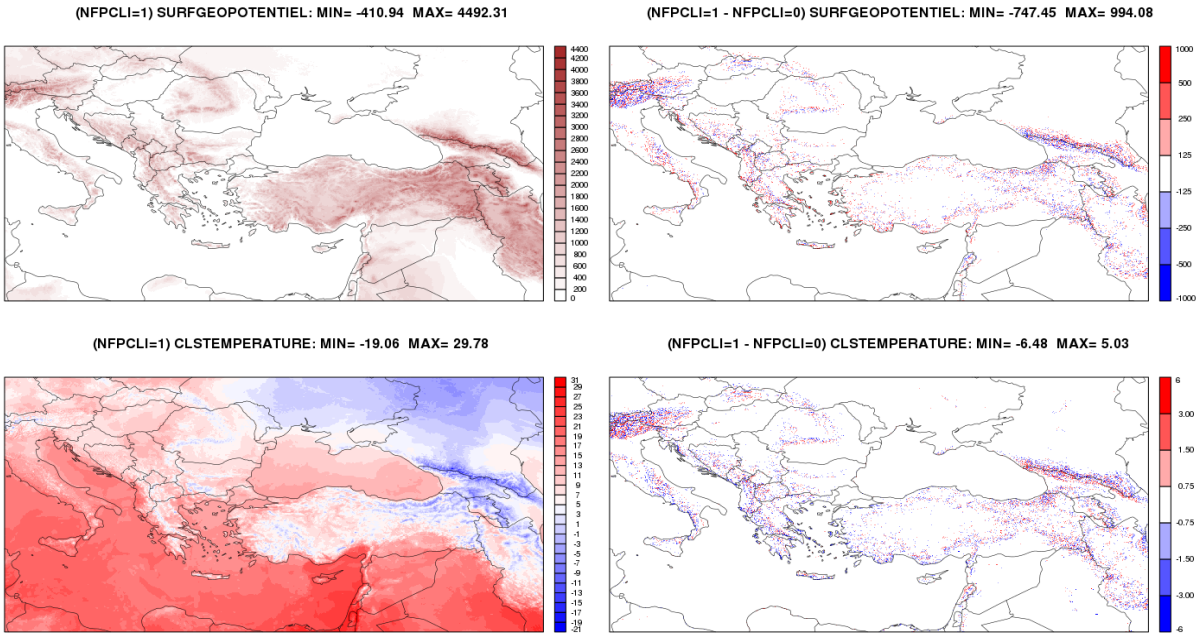


Figure 2: The impact of using CLIM files in fullpos containing the fitted orography for target LATLON domain of Turkey (NFPCLI=1) instead of using the interpolated orography from the model (NFPCLI=0). Surface geopotential (top), temperature at 2 m (bottom), with the respective fields to the left and the difference to the right.

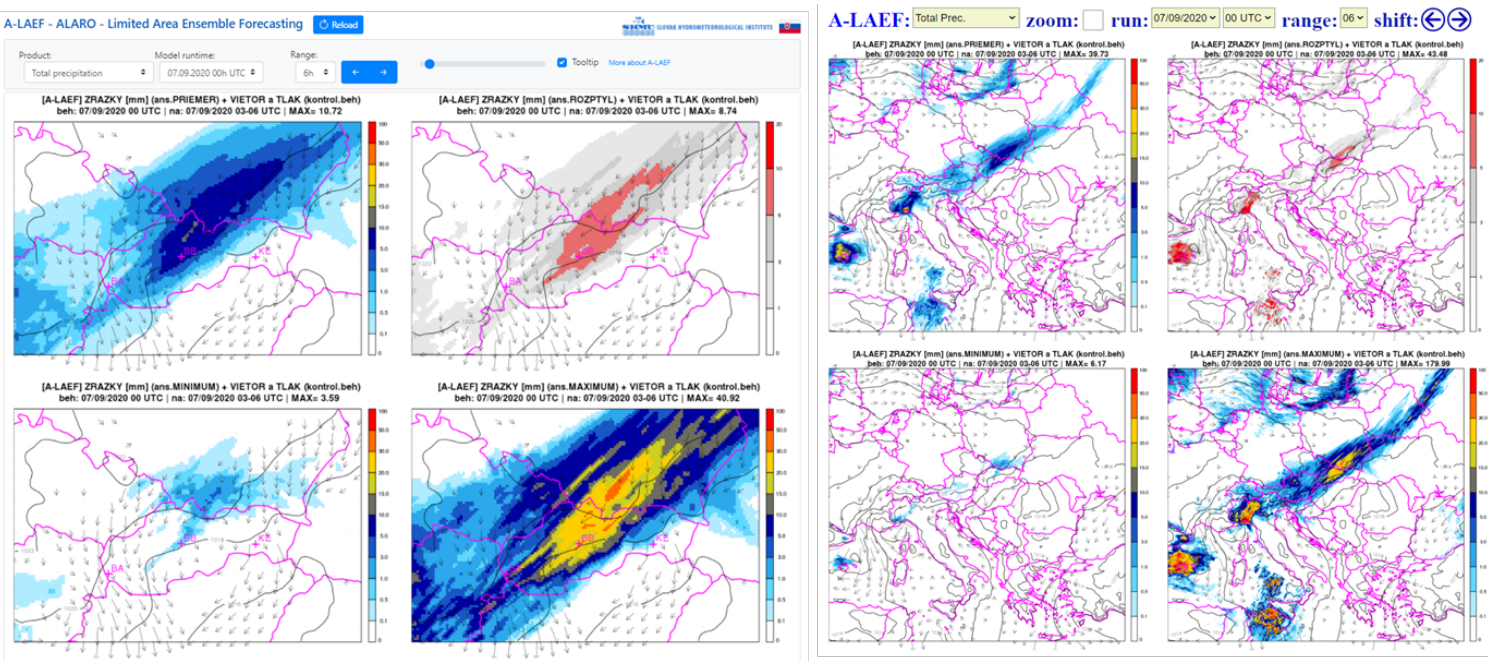


Figure 3: Example of the operational probabilistic A-LAEF system maps.

A bootstrap verification by HARP shows that the operational A-LAEF system outperforms local operational deterministic model ALADIN/SHMU running at similar spatial resolution (4.5 km/63 L). The primary verification scores during a winter month in 2020 indicate significant improvement for most of the parameters (except MSLP). The scores are worse for the initial time due to perturbations involved in the ensemble system, but for the later lead times A-LAEF ensemble mean is better, with high significance (Figure 4). The superiority of the A-LAEF ensemble system is even more pronounced during the summer convective season (not shown).

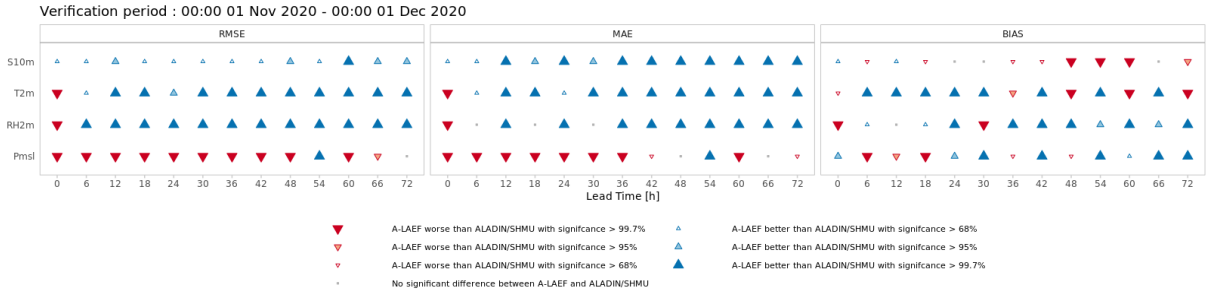


Figure 4: HARP bootstrap method comparing scores between A-LAEF (EPS mean) and ALADIN/ SHMU (operational deterministic model running at SHMU) for the period from 1st Nov to 1st Dec 2020 (both 00 and 12 UTC runs), for wind speed at 10 m (S10m), temperature (T2m) and relative humidity (RH2m) at 2 m, and mean sea level pressure (Pmsl). Observation data from approximately 96 SHMU AWS stations were used.

C-LAEF

For C-LAEF following research and development has been made in 2020:

- Implementation of new cy43t2 at the ECMWF HPCF
- Set-up of a complete C-LAEF e-suite with cy43t2 including adaption of scripting system
- New observations (GNSS, Mode-S, satellite data) were tested in C-LAEF EDA
- Investigation of 2m temperature and humidity problem in C-LAEF
- Extension of C-LAEF SPP scheme by additional perturbations in physics parametrizations
- Time lagged EPS out of AROME-RUC was created using neighbourhood methods
- Development and operational production of EPS maps and EPSgrams with Visual weather

First results with new observations (GNSS, satellite data, Mode-s) showed a problem with T2m and RH2m during the day with too much soil moisture (Figure 5). This was caused by a positive precipitation bias in the first forecast hours. It turned out that the problem comes from the variational bias correction of GNSS data (too long adaptation time). After adaptation of this, the scores improved significantly.

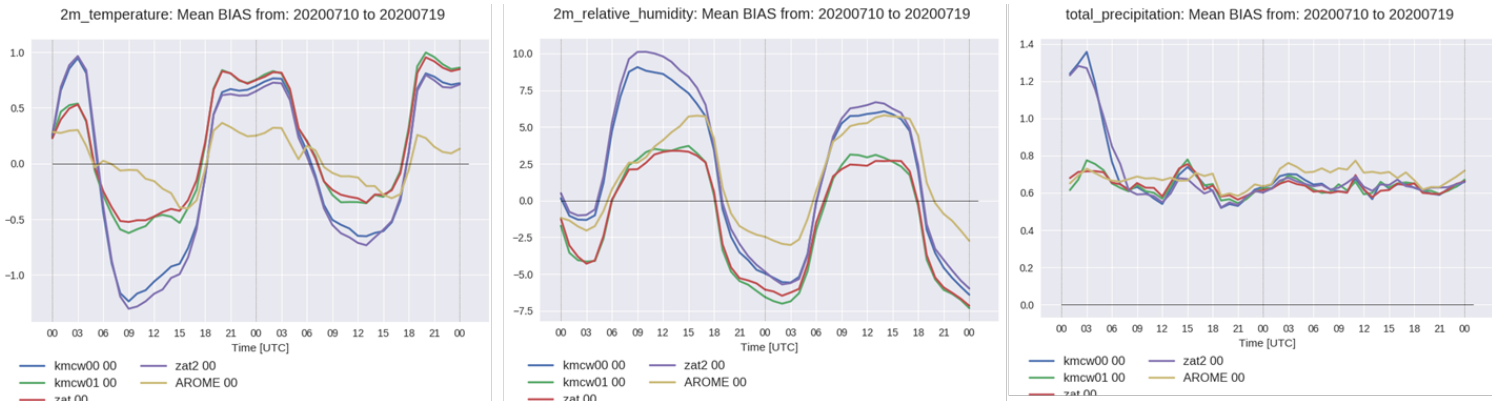


Figure 5: Verification of different sets of "new observations" for 1 week (July 10 – July 19, 2020). BIAS of 2-meter temperature (left), 2-meter relative humidity (middle) and precipitation (right).

The verification of the operational C-LAEF data showed a strong temperature and relative humidity bias in the night, especially strong in Alpine valleys during clear nights in winter. The problem was identified to come from the Canopy scheme. Because of this some test were made and a new 2m diagnostics with weights depending on stability and orography was set up (Figure 6, separate newsletter article about this topic).

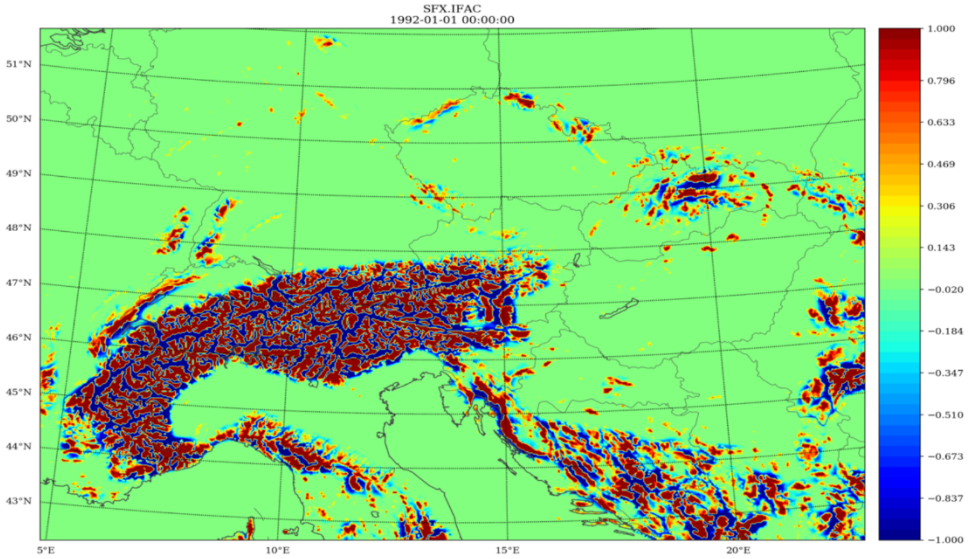


Figure 6: IFAC (inversion factor) for the AROME-Aut/C-LAEF domain with values varying between -1 (valley) to 1 (mountain tops) and values around 0 for flatland areas.

The operational C-LAEF uses a combination of tendency perturbations (shallow convection, microphysics, radiation) and parameter perturbations (turbulence) for representation of model error. To increase the physical consistency within the ensemble it is planned to extend the stochastic parameter perturbation scheme (SPP) to other physics parametrizations such as microphysics, shallow convection and the radiation scheme. For this purpose the code from HIRLAM (cy43t2) was adapted and some new key parameters were added to the perturbation scheme. In this context also the SPG pattern generator (Tsyrulnikov and Gayfulin, 2017) was implemented and tested (Figure 7).

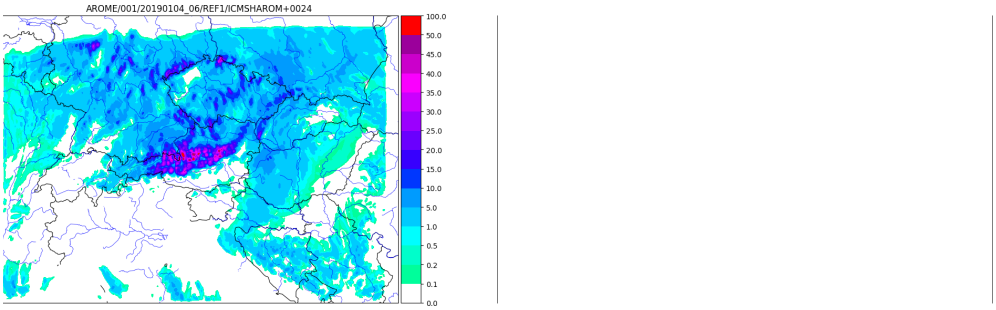


Figure 7: Sensitivity tests on sublimation of snow/graupel in wintertime in modified C-LAEF perturbation scheme (reference left, experiment with reduced sublimation middle, difference between EXP and REF right).

AROME-EPS

For AROME-EPS following research and development has been made in 2020:

- AROME-EPS runs operationally at OMSZ since February 4, 2020
- Comparison of AROME-EPS and its predecessor, the 8 km resolution ALARO-EPS (Figure 8)
- Experiments with ensemble data assimilation (EDA) in AROME-EPS (Figure 9)

A comparison of ALARO-EPS and AROME-EPS for a longer period (June 2019 to January 2020) shows that wind gust forecasts are better in all seasons in AROME-EPS (Figure 8 for July), however, AROME-EPS is more underdispersive than ALARO-EPS. Also the average wind improved usually, except for some convective situations. T2m, RH2m, total cloudiness are better during the day, but get

worse during the night. Precipitation in AROME-EPS is better in winter and autumn, and a little bit worse in summer (Figure 8 for July) than in ALARO-EPS.

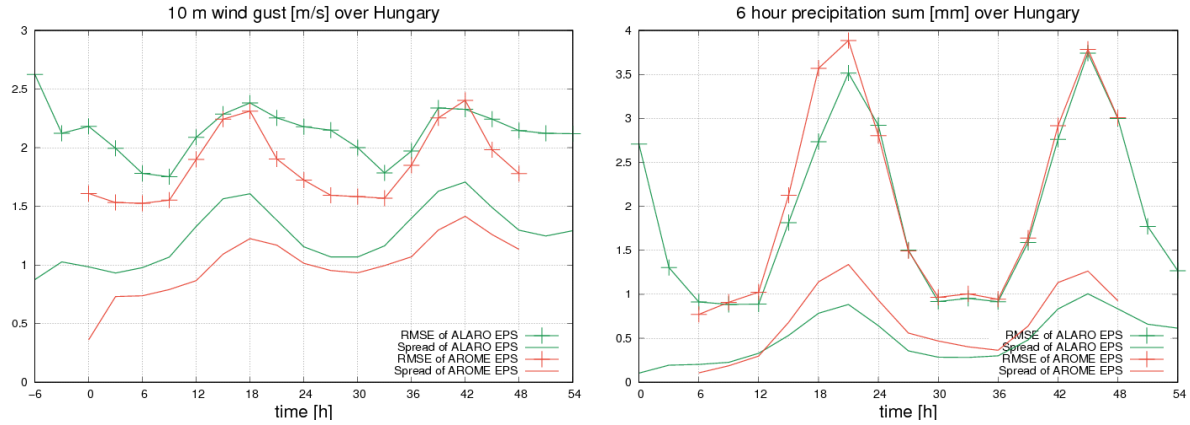


Figure 8: RMSE of ensemble mean (symbols) and EPS spread (solid) of ALARO-EPS (green) and AROME-EPS (red) for wind gust (left), 6-hourly precipitation sum (right) in July 2019.

For testing EDA in AROME-EPS, a 3-hourly assimilation cycle identically to the operational deterministic AROME analysis (same domain, resolution, etc.) was set up with a 3-hourly coupling to ECMWF-ENS. Conventional and GNSS ZTD measurements were used and the observation perturbations were executed offline before the surface assimilation and after screening. Furthermore, a perturbation scaling was added (multiplication of perturbations with real number). The AROME-EPS version was tested for 3 periods of about 1 month each (1 in winter 2020, 2 in summer) and compared to the operational AROME-EPS (Figure 9).

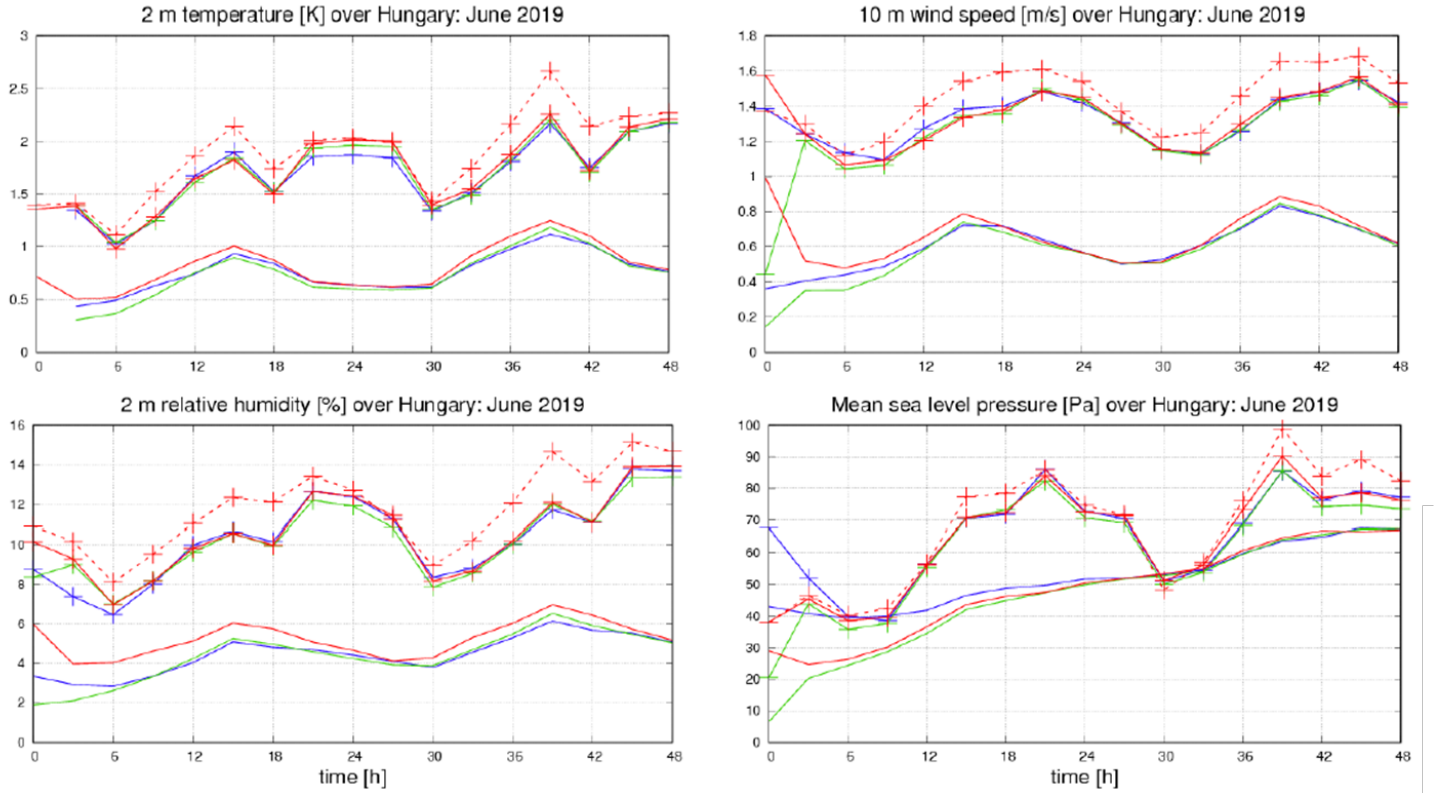


Figure 9: RMSE of ensemble mean (symbols) and spread (solid) of surface parameters in the experiments from 28 May to 19 June 2019. Operational EPS (blue), EDA with (red) and without perturbations (green), control member of EDA (dashed).

Adding EDA to AROME-EPS can improve forecasts of surface parameters, especially for wind, temperature, mean sea level pressure and precipitation in the summer season. Humidity and cloudiness seem to have higher error values at the beginning of the forecasts with EDA, probably caused by assimilation problems. Some tuning of perturbations still has to be done before a final operationalization.

3 References

Belluš M., 2020: New high-resolution ensemble forecasting system A-LAEF. Meteorologický časopis, 23 (2), pp. 75-86, ISSN 1335-339X (in Slovak)

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

Plenković, I. O., I. Schicker, M. Dabernig, K. Horvath, E. Keresturi, 2020: Analog-based post-processing of the ALADIN-LAEF ensemble predictions in complex terrain, Quarterly Journal of the Royal Meteorological Society, 146: 1842– 1860, https://doi.org/10.1002/qj.3769

Taşcu S., M. Pietrişi, C. Wittmann, F. Weidle and Y. Wang, 2020: Forecast skill of regional ensemble system compared to the higher resolution deterministic model. IDŐJÁRÁS / Quaterly Journal of the Hungarian Meteorological Service, 124 (3). pp. 401-418. ISSN 0324-6329

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, DOI: 10.1127/metz/2017/0815.

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

Doctoral theses:

Iris Odak Plenković successfully defended her doctoral thesis “Wind speed prediction using the analog method over complex topography” in July, at Faculty of Science, Department of Geophysics, University of Zagreb.

Endi Keresturi successfully defended his doctoral thesis “Initial condition perturbations in a convective scale ensemble prediction system” in February 2021, at the Faculty of Science, Department of Geophysics, University of Zagreb.