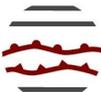


Newsletter n°2

28 February 2022



*A moment of relaxation during the ACCORD/MUSC WW
Eric, Ana, Martina, Laura, Piotr are walking on the fields of Viikki
across Helsinki University agricultural faculty*

ACC  RD

A Consortium for CO₂ 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.accordion-nwp.org/>

Contents

Edito, *Claude Fischer, PM*

ACCORD events: second half of 2021 and 2022 *Patricia Pottier, CSS*

Networking: ACCORD directory of staff and calendar *Patricia Pottier, CSS*

Practices of meteorological quality assurance within ACCORD *Carl Fortelius*

Verification of the lightning probability forecast scheme based on AROME-BG *Boryana Tsenova, Konstantin Mladenov, Milen Tsankov, Andrey Bogatchev*

The analog-based method application to gridded data post-processing *Iris Odak Plenković, Irene Schicker, Markus Dabernig, Alex Kann, Aitor Atencia*

AROME-MAROC CY41T1 Quality Assessment using the simulated satellite images in two MSG channels WV6.2 and IR10.8 *Meryem BOUCHAHMOUD, Abderrahim MOUHTADI, Siham SBII, Ilias KACIMI*

HOOFF - a preprocessing tool for OPERA weather radar data files *Peter Smerkol*

The “46t1 e-suite” running at Meteo-France: content and evaluation *CNRM/GMAP (NWP Research group) and DIROP (Operations group) collective article*

The road to operational CY43T2 in Austria *Christoph Wittmann, Florian Meier, Stefan Schneider, Clemens Wastl and Florian Weidle*

Summary of NWP related activities in Slovenia in 2021 *Jure Cedilnik, Benedikt Strajnar, Neva Pristov, Jana Čampa*

Recent NWP activities at Météo Algérie : Report on the new HPC and the newcomers training *Walid CHIKHI, Nour El Isslam KERROUMI, Zakaria BENGHABRIT, Mohamed MOKHTARI, Abdenour AMBAR*

NWP activities in Romania *Alina Dumitru, Simona Taşcu, Mirela Pietrişi, Alexandra Crăciun*

Recent ACCORD Related Activities at TSMS *Yelis Cengiz, Alper Güser, Mustafa Başaran*

Summary of ACCORD autumn surface working week *Patrick Samuelsson*

Introducing a roughness-sublayer in the vegetation-atmosphere coupling of HARMONIE-AROME *Metodija M. Shapkalijevski, Samuel Viana, Aaron Boone, Quentin Rodier, Patrick Le Moigne and Patrick Samuelsson*

Algorithmic amelioration of the deficiencies in the screen level parameters forecast based on a dynamical downscaling approach *Martin Dian, Mária Derková, Martin Petráš*

Improving model performance in stable situations by using a pragmatic shift in the drag calculations - XRISHIFT *Mariken Homleid*

MUSC working week and ACCORD visits - Helsinki, November 2021 *Laura Rontu, Eric Bazile, Piotr Sekula, Ana Šljivić, Wim de Rooy*

The Representation of Turbulent Kinetic Energy in HARMONIE-AROME at Hectometer Scale *Mirjam Tjihuis, Natalie Theeuwes, Jan Barkmeijer*

Hectometric-Scale Experiments at Met Éireann *Colm Clancy, James Fannon, Eoin Whelan*

Hectometric experiments with the HARMONIE-AROME system at IMO *Bolli Pálmason, Guðrún Nina Petersen, Sigurdur Thorsteinsson, Xiaohua Yang*

The ACCALMIE project: coordinated approach for chemistry and aerosols in Météo-France's models, inline and offline *Vincent Guidard, Quentin Libois*

EPS research and development in RC-LACE in 2021 *Clemens Wastl, Martin Belluš, Gabriella Szépszó*

Evaluation of HARMONIE-AROME cycle 43h2.1 at AEMET *Javier Calvo, Joan Campins Pau Escriba, Daniel Martin, Gema Morales, Jana Sanchez Arriola y Samuel Viana*

Edito

Claude Fischer, ACCORD Programme Manager

Since the publication of the first ACCORD Newsletter on 5 October 2021, the consortium activity has continued at all levels. The ACCORD Assembly on 8 December has adopted the 2021 progress reports including the realized manpower figures for the first semester. It has approved the proposed work plans for 2022, and discussed the next steps in order to pave the way for an ACCORD-based bid proposal to the Destination Earth Project (DestinE). Although this effort is somehow “on the edge of ACCORD itself”, it is probably useful and fair to mention that quite a number of teams and staff have been busy with discussing the concrete content of the on-demand LAM Digital Twin offer that they’d like to propose in DestinE. In this respect, the Assembly clearly expressed its motivation that the results of DestinE shall be shared by all ACCORD Members, should the bid be successful. The ACCORD Management Group (MG) has been, in several aspects, involved in the DestinE discussions as well. The MG also has, among other tasks, defined its guidelines regarding the preparation of the funded actions in link with Working Weeks and scientific visits to be supported by the consortium budget. These guidelines have been presented to the LTMs and will be used in the construction of the 2022 costed Detailed Action Plan (DAP).

In this Newsletter 2 (NL2), several articles refer to common ACCORD efforts, such as the first thematic working weeks, consortium-wide questionnaires to better understand how the many different teams work, results from the first ACCORD-supported visits. The pandemic situation has remained an important factor in our scientific organization, forcing quite often to use video-meeting facilities as the only possible solution. Nevertheless, we acknowledge the first in-person meeting for working together on MUSC, as well as a few scientific visits that could take place. Their outcome is reported in several contributions to NL2. Like in NL1, a number of articles refer to Meteorological Quality Assurance and status reporting on national NWP systems. These contributions are found in the first part of NL2. The other thematic contributions refer to surface modelling, very high resolution modelling (hectometric or in link with LES), EPS. One specific contribution describes a starting project aimed at building an aerosol-chemistry library of codes shared by all teams in the MF Research Dept. Certainly a number of our scientific areas are not represented in this NL2, however their absence does not mean they would not matter anymore (!). The strategic and quite urgent issue of adapting our codes in preparation of new HPC architectures will be a major topic in 2022 and beyond. Modernization of our working practices regarding code management is another topic where we expect to actively explore solutions, and make significant progress this year.

We are now (already!) heading towards the venue of the 2nd ACCORD All Staff Workshop (ASW), likely to be held in hybrid mode with ARSO as the hosting institute (Ljubljana, 4-8 April). The planning is in good progress and information on its organization and agenda is in preparation at the time of writing this editorial. The ASW will be a turning point in the scientific coordination of the year, as it will give the kick-off for discussing the preparation of the Rolling Work Plan for 2023 (RWP), which will then become a main task for the ACCORD MG, the work package co-leads, in liaison with the LTMs and their staff, over the spring and summer.

Awaiting for these future steps, with hopefully a brighter perspective for organizing in-person meetings in 2022 despite the challenging conditions of the every-day international context, I thank the contributors to NL2 and wish all readers a good time consulting the newsletter.

Claude Fischer.

Post Scriptum.

A few words about the elaboration of our ACCORD newsletters. For the concrete writing of an article, please refer to the editorial guidelines, accessible at:
<http://www.accord-nwp.org/?Recommendations-templates>

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

NL3 is expected for after the summer break, however contributors can post their material at any time during the year to the PM and the CSS. Do also not hesitate to encourage scientific contributions by young scientists (PhDs, post-docs etc.) or technical contributions (codes, porting, optimization etc.).

ACCORD events: second half of 2021 and 2022

Patricia Pottier, ACCORD Consortium Scientific Secretary

1 Introduction

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

2 After summer 2021

Governance meetings:

- CNA meeting on 13 September 2021
- [2nd LTM meeting](#) on 4 October 2021
- [2nd STAC meeting](#) on 15 November 2021
- Bureau meeting on 19 November and 29 November
- [regular meetings of the MG](#) (every two weeks)

Although not an ACCORD event, but with the participation of many ACCORD colleagues, the [EWGLAM meeting](#), initially planned in Brussels took place as a fully online event, between 27 September and 1 October 2021.

Working Days and Working Weeks

- MUSC WW, Helsinki (Finland): Develop a common MUSC cycle (same MASTER), based on cy46t1 export, for the 3 CSC, 29 November - 3 December.
- Autumn Surface Working Week 2021, 18-22 October 2021, more information on [ACCORD wiki dedicated page](#)
- Joint 2021 LACE Data Assimilation Working Days and ALADIN Data Assimilation basic kit (DAsKIT) Working Days, 22-24 September 2021, remote meeting/Ljubljana, [more information](#)
- EPS on-line meeting on Calibration and user-oriented approaches, 7 Oct 2021, more information on [ACCORD wiki EPS dedicated page](#)
- 1st meeting with LTSR (Local Team System Representative), 20 October 2021, [information on system pages on ACCORD wiki](#)

Thematic meetings

- Transversal activities on future software infrastructure, more information on [ACCORD wiki SPTR dedicated page](#)
- WG on Very High Resolution Modeling (VHR-MOD): [meetings information](#)
- WG on Machine Learning (ML): [meetings information](#)
- WG on Physics Interoperability, currently in strong link with SPTR
- DA Research Teams and Support Teams meetings, more information on the [ACCORD wiki dedicated pages](#)
- Surface monthly meetings, more details on the [ACCORD wiki pages](#)

3 Events planned in 2022

All Staff Workshop

The big event in 2022 will be the [2nd All Staff Workshop on 4-8 April in Ljubljana](#). Please [register](#) before the end of February 2022. Taking into account the Covid situation, we plan for an **hybrid** meeting (still to be confirmed by our Slovenian hosts), with participants in Ljubljana (hotel or ARSO, to be confirmed) and remote participants.



Governance bodies meetings

- The [Management Group](#) resumed their every other Friday morning meetings at the beginning of January. Additional meetings are planned on the Fridays when found necessary for the MG agenda (preparation of DAP2022, of RWP2023, reporting of RWP2022 ..).
- [3rd LTM meeting](#) (DAP2022, changes in CMR) on 31 January 2022
- Bureau meeting to analyse the bid proposal for DestinE and propose the questions for PAC (the Bureau initially planned for the 21 February is postponed after the DestinE ITT publication)
- PAC, tentatively mid- March, convened by the Bureau
- Extraordinary Assembly or remote voting (about codes for DE): end of March
- Bureau meeting (questions about modernization of working practices)
- joint PAC & STAC in spring (to treat policy or organisational questions in link with modernisation of working practices)
- STAC in Spring (review progress on modernisation, update of work plan, impact of DE on work plans)
- Bureau meeting (preparation of the Assembly)
- Assembly meeting: on-line meeting at the end of June/beginning of July or 27 June in Bologna, date and place to be decided by the Bureau
- STAC in autumn
- PAC in autumn if convened by the Assembly
- End of the year Assembly: on-line meeting or 5 or 8 December in Darmstadt, date and place to be decided by the Bureau

Working Days and Working Weeks

The WD and WW which might be organised in 2022 are currently discussed within the MG. From the comprehensive list below, some already have been agreed and a few might eventually be considered rather for 2023:

- harp training course, 16-18 February 2022, on-line: up-to-date information about the event on [the ACCORD MQA wiki](#)
- EPS Working Week, 25-29 April 2022, Innsbruck (Austria), more information on the [ACCORD wiki dedicated page](#)
- DA WW on high resolution data assimilation and nowcasting and 4D-Var, 25-29 April 2022, Budapest (Hungary): more information on the [ACCORD wiki dedicated page](#)
- NWP SURFEX training week, 9-13 May 2022, Budapest (Hungary), more information on the [ACCORD wiki dedicated page](#)
- DA Code training days, 16-20 May 2022, Toulouse (France)
- WW on user friendly harp, May or June, date and place t.b.d
- DA WW dedicated to Support Teams and OOPS, 20-24 June 2022, place t.b.d.: more information on the [ACCORD wiki dedicated page](#)
- LACE working days & DAsKIT (DA-ST2) meeting in September, date and place t.b.d
- DA research teams working week in autumn, date and place t.b.d.
- ALARO WW / code adaptation, refactoring, data structures, end of summer/autumn
- Harmonie-Arome CSC code refactoring WW, end of summer/autumn
- DA research teams working week in autumn, date and place t.b.d.
- Cloud-radiation-aerosol ww to update/merge code branches into a common scientific base code, Helsinki (Finland), date t.b.d.
- DAVAĬ training for users, on-line, date t.b.d.
- ACCORD source forge training, autumn or winter, webinar, date t.b.d. (after Assembly decision on the forge)
- DAVAĬ contributors-developers WW, date and place t.b.d.
- Tech support visits for GIT transfer of knowledge, date and place t.b.d.
- ACCORD user (kick-off) meeting, date and place t.b.d.
- WW on code adaptation and related tools, date and place t.b.d.
- 3D physics group meeting, autumn, date and place t.b.d.
- Surface Working Week, in autumn, web-based, date t.b.d

Networking: ACCORD directory of staff and calendar

Patricia Pottier, ACCORD Consortium Scientific Secretary

1 Introduction

In [the 1st ACCORD Newsletter](#), various aspects of the newborn ACCORD consortium were presented: governance bodies, strategy, rolling work plans, canonical configurations and operations, networking tools and actions (logo, domain, website, wiki, mailing lists, on-line meetings, ...). Recently, an online directory of teams and bodies was added to the ACCORD website to ease contacts between ACCORD people. The time zones and online calendar can also be useful for preparing ACCORD events and meetings.

2 Directory of teams and bodies

The [ACCORD directory of staff](#) is available on the ACCORD website. With its search engine, quickly find the contact details of an ACCORD colleague, the composition of a team (select country/NMS or family), the members of the committees and bodies (PAC, STAC, LTM, MG).

3 Online calendar

ACCORD now has its [Google calendar](#) with governance meetings (Assembly, PAC, STAC, LMT), Management Group (MG) meetings, Working Days, Working Weeks, thematic WG meetings, ...

4 Time zones

The 26 ACCORD Members belong to [different time zones](#) and the time difference between 2 members can be quite big (bigger in winter) even if the majority belongs to the Central European Time (UTC+1) and changes for Central European Summer Time (UTC+2) in summer. When planning for ACCORD video-conferences, the agenda generally uses CET/CEST. The maps below show the differences with the local time, an illustration that some meetings can result in quite unusual working hours !

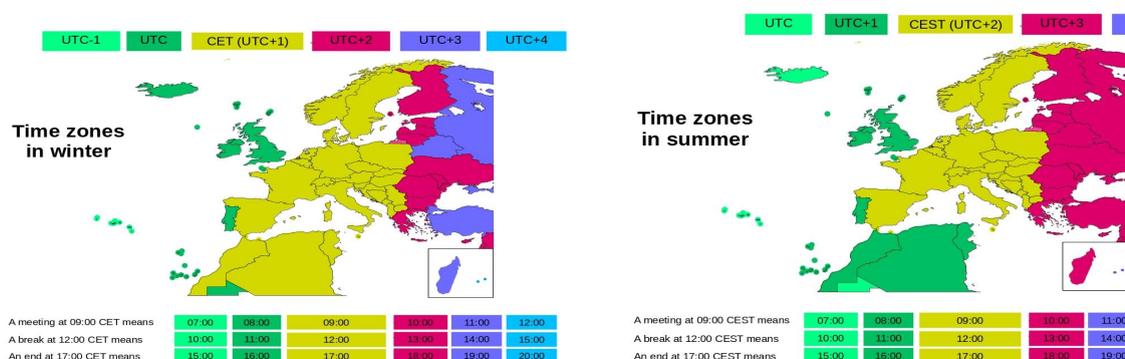


Figure 1: Time zones in winter (left) and summer (right)

Practices of meteorological quality assurance within ACCORD

Carl Fortelius

1 Introduction

The meteorological services that form the ACCORD consortium and the corresponding local teams all are faced with the need to verify their forecasts produced, and to ensure the realism and correct working of forecasting systems used in the production. Individual systems, as well as the forecasts, reflect local conditions, needs, and priorities, and may vary considerably between the services, e.g. with respect to climatological regime, domain size, resolution, target range, post processing, data assimilation, observation usage, or ensemble forecasting. It is therefore natural for methods and practices of meteorological quality assessment to differ between local teams according to the local needs and priorities.

In order to map the methods and practices in use, a questionnaire was distributed among the local teams, inviting them to describe the procedures of their respective teams or groups regarding forecast verification, collection of user of feedback, forecasting system monitoring, and case studies. The present article summarizes the main findings extracted from the received replies, covering 18 institutes out of 26 ACCORD partners. To a large extent, it reflects the impression formed by one interpreter, but for the interested reader further details, including the raw answers, are available in the section of meteorological quality assurance on the ACCORD wiki.

The questionnaire is described in Section 2, and the main results are presented in Section 3. A summary and conclusions are given in Section 4

2 The survey

The survey was conducted by distributing a questionnaire to all local teams in ACCORD, asking them to describe their respective practices. Respondents were invited to identify themselves in their reply, but reference to any particular team will be avoided in the present report.

The query started with an introductory part, where respondents were asked to describe their operational suites of deterministic forecasts and ensembles, giving details about domains, forecasts ranges, and use of data assimilation. The following sections dealt with data and methods for verification of deterministic forecasts and ensembles, gathering of forecaster feedback, reporting and use of verification results, diagnostic studies of data assimilation system and forecast model, and case studies.

In the 18 replies, 24 forecasting suites were described as operational. Out of these, 13 suites make use of upper air data assimilation (3D-VAR). LAM-ensembles are produced or used operationally by 14 services: four services use ensembles from A-LAEF, 4 use ensembles from MetCoOp and six from other sources. Only one rapid update cycling suite was termed operational.

3 Results

3.1 Forecast verification

As a rule, operational forecasts are subject to regular verification. The main use of verification is in identification of weaknesses in the forecasting system/forecast and in providing guidance for development of the forecasting system. Almost equally important is the evaluation of new cycles. 11 out of 18 replies indicate that differences are tested for statistical significance, when evaluating cycles. Providing information to users is not often mentioned as an application of forecast verification.

It is common practice to compile verification reports regularly, but about one third of the institutes compile reports only on demand. Monthly and seasonal summaries are the most common, but periods ranging from daily to annual are in use, as well. Reports may be publicly accessible, but in the majority of cases, 70-85%, access is confined within the institute itself, or shared within a sub group of the ACCORD community.

Observations Surface observations and balloon soundings from reporting weather stations are by far the most widely used data for verification, and often the only one. Met. reports (METAR) from air fields are mentioned by three respondents. A variety of other data sources, including earth observation data from satellites and radars as well as crowd sourced observations, are also in use, but are typically specific to individual institutes. The observations are mostly subject to a basic quality control, consisting of a gross error check and removal of outliers.

Filtering Filtering of observations and forecasts in time and space is often applied in order to separate the data into sub sets of interest. Twelve out of 13 respondents filter deterministic forecasts of near-surface variables according to forecast lead time and/or time of validity. 11 respondents filter according to time of analysis. Eleven out of 14 respondents verify over multiple geographical domains, while 8 and 6 respondents filter according to orographic height or type of surface, respectively. Four respondents do not apply spatial filtering. In the case of upper air forecasts, mainly temporal filtering is applied, as only one respondent out of 9 carries out verification over several geographical domains.

For ensembles of near surface variables, all the 7 respondents filter according to forecast lead time, while 4 respondents filter by time of analysis and/or time of validity. Three out of 6 respondents filter according to geographic area and/or orographic height. Filtering for upper air variables is carried out much in the same way as for near-surface variables: 6, 3, 3 out of six respondents filter, respectively, according to forecast lead time, time of validity, or time of analysis. One out of 9 respondents verifies upper air ensembles over multiple domains.

Software Table 1 summarizes software used in forecast verification. Most common are software developed either jointly by members of the ACCORD community, (harp, Monitor, LAEF-verification), or else locally by the answering team. The open source verification package WRFT/Verif is also used for surface variables. One institute reports using observation screening data from data assimilation for verification of upper air forecasts.

Metrics Tables 2 and 3 summarize, respectively, point metrics and spatial metrics in use for deterministic forecasts of near-surface parameters and upper air parameters. For an overview of metrics, visit the website of the WWRP/WGNE Joint Working Group on Forecast Verification Research. Point measures of systematic and random difference are the most commonly used metrics, although categorical measures are also in use, and applied mainly for near-surface variables. Non-local measures such as FSS (Roberts et al., 2008, Roberts, 2008), SAL (Wernli et al., 2008, Wernli et al. 2009), SLX (Sass, 2021), or accounting for predicted values in

Table 1: Software used for forecast verification. harp, Monitor, and LAEF verification are maintained by ACCORD partners, WFRT/verif is an open-source utility available at <https://github.com/WFRT/verif>

Forecast type (answers)	harp	Monitor/ WebGraf	LAEF verification	WFRT/ Verif	screening (obs-fg)	local development	none
Det. sfc (15)	6	8	1	2		6	
Det. upper (13)	4	7	1		1	2	2
Ens. sfc (7)	2		1	1		1	2
Ens. upper (7)	2		1			1	3

the neighbourhood of observation locations (SO-NF) when calculating metrics, are applied almost exclusively to forecasts of precipitation, although two instances of applying spatial methods to upper-air fields are reported, as well.

Table 2: Point metrics used to verify deterministic forecasts. Measures of systematic and random error include mean bias, mean absolute error, rms-error, error standard deviation. Categorical measures include measures such as Hit rate, false alarm ratio, frequency bias, various skill scores, etc.

	systematic and random error	categorical verification	correlation coefficient	answers
Surface	12	8	1	14
Upper air	8	2		11

Table 3: Spatial metrics used to verify deterministic forecasts. The heading "field verif." covers subjective evaluation, spectral analysis, and rms-difference of fields

	FSS	SAL	SLX	SO-NF	field verif.	answers
Surface, precipitation only	3	4	1	1	3	11
Upper air	1				1	7

Table 4, summarizing metrics used to verify ensemble forecasts shows, that these forecasts are evaluated almost exclusively based on a point by point comparison. Non local methods, that would be be less sensitive to errors in location or timing of events are hardly in use.

3.2 Forecasting system diagnostics

Data assimilation and observation usage The number of observations and their geographical coverage are typically monitored. Every one out of 13 respondents monitor the number of observations in different categories, and 12 respondents monitor their geographical coverage. 12 out of 13 respondents collect statistics of the departures of background and analysis from the observations, while 10 respondents compile quality control summaries describing the number of observations tested, used, or rejected. Five respondents use measures like the degrees of freedom of signal (DFS) to monitor the impact of different observation types. Seven respondents monitor the convergence of the minimization of cost functions in variational data assimilation, or similar diagnostics. Finally, six respondents produce maps of analysis increments.

Forecast model Monitoring in terms of diagnostic measures is less common for the forecast model than for the data assimilation and observation usage. Only six respondents follow one or more of the diagnostics listed in Fig. 1. Surface fluxes and near surface vertical profiles at selected locations are the favoured diagnostics, used by four of the respondents. Two respondents out of the six monitor energy norms or surface fluxes averaged

Table 4: Metrics used to verify ensemble forecasts. Probabilistic point measures include scores such as CRPS, Rank histogram, Brier score/skillscore, ROC, reliability-scores, etc. CRPS, HR, FAR with up scaling

	Point metrics			Spatial metrics		answers, point/spatial
	spread-skill	mean bias	Probabilistic scores	CRPS, HR, FAR with upscaling	CRPS, bias, stdv	
Surface	6	3	6	1		8/5
Upper air	6	1	5		1	7/3

over selected domains, while power spectra fo selected variables or tendencies of state variables over selected domains are watched by 1 respondent each.

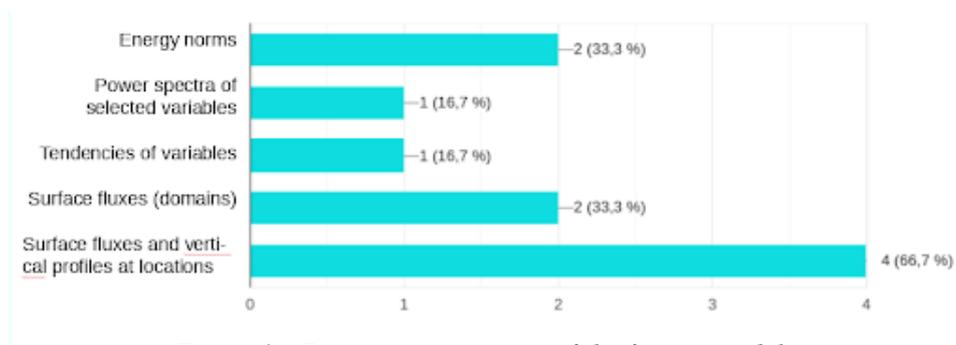


Figure 1: Diagnostic measures of the forecast model.

Reporting Results of the monitoring are reported at meeting of the local team or production syndicate in all cases. Eight out of 12 respondents produce written reports in consortium newsletters or similar publications, while 5 respondents mention reporting in research papers.

3.3 Case studies

In-depth analyses of forecasts and forecasting systems in selected cases of interest are performed by 15 respondents. Often, but not always, failed forecasts of high-impact weather are chosen for study. Selection of cases may be triggered by forecasters, and most of the cases relate to heavy precipitation, especially when associated with flooding. Poor visibility and extreme temperature or high wind speeds are mentioned, as well. At least one of the answering institutes maintain an archive of interesting cases. Case studies typically include detailed comparison of forecasts from different sources to observations collected especially for the purpose. In many cases, the events are re-forecasted using many different configurations of the forecasting system, with varying model options and choices of observation usage. Results of case studies are typically reported at internal meetings or in newsletter articles, but about half the respondents mention reporting in research articles.

3.4 Collection of user feedback

Teams were asked to describe the collection of user feedback by selecting a set of suggested alternatives. The replies indicate that meetings between users and developers are the most favoured means of communication. Evaluation forms filled in by forecasters are used by 6 respondents, and likewise 6 respondents produce written feedback reports in newsletters or similar.

4 Summary and conclusions

Based on the input from 18 out of 26 teams forming the ACCORD community, verification, feedback from duty forecasters, and ability to forecast selected cases of significance are all important factors steering the development of forecasting systems and deciding the acceptance of new suites.

The vast majority of verification is carried out by point by point comparison of forecasted parameters to data reported by surface weather stations and ships. Upper-air fields are less closely watched, and nearly exclusively verified against balloon soundings. The use of so-called spatial verification methods, extending the comparison between model and data over relevant neighbourhoods, are not widely used, and mostly applied to precipitation.

Results of forecast verification are typically displayed in near real time and collected in reports that are shared locally within institutes, or among sub groups of ACCORD. Partners might benefit from a wider distribution, facilitating the comparison of different local implementations of CSCs.

Several software packages for verification are developed and used within the consortium. Not all packages are capable of probabilistic verification or spatial verification methods, or can ingest or utilize spatially distributed fields of data to support such methods. As recognized in the ACCORD strategy document for 2021-2025, partners share a common need for a verification system designed specifically to evaluate forecasts at the scale of convective weather systems in all their aspects. Such a system would apply measures taking full account of the limited predictability at these scales, and be capable of tapping all available and relevant data sources.

5 References

Roberts NM, Lean HW, Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.* 136: 78–97, 2008.

Roberts NM, Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorol. Appl.* 15: 163–169, 2008.

Sass, Bent, A scheme for verifying the spatial structure of extremes in numerical weather prediction: Exemplified for precipitation. *Meteorological Applications*. 28. 10.1002/met.2015, 2021.

Wernli, H., M. Paulat, M. Hagen, and C. Frei, SAL—A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.*, 136, 4470–4487, 2008.

Wernli, H., Hofmann C, and M. Zimmer, Spatial Forecast Verification Methods Intercomparison Project: Application of the SAL Technique. *Wea. Forecasting*, 24, 1472– 1484, 2009.

WWRP/WGNE Joint Working Group on Forecast Verification Research, Forecast Verification - Issues, Methods and FAQ, https://www.cawcr.gov.au/projects/verification/verif_web_page.html, visited on Feb 15, 2022

Verification of the lightning probability forecast scheme based on AROME-BG

Boryana Tsenova, Konstantin Mladenov, Milen Tsankov, Andrey Bogatchev

1 Introduction

Particle charging in thunderstorms is associated with collisions between riming graupel and vapour grown ice crystals (Reynolds et al., 1957; Takahashi, 1978; Jayaratne et al., 1983; ...). The majority of lightning activity occurs in cloud regions where graupel particles are present, followed by those with snow and hail (Dotzek et al., 2001; Bruning et al., 2007). Lightning probability forecast scheme over Bulgaria takes into account the integrated graupel mixing ratio r_g forecasted by AROME-BG between model levels 35 and 15 (or between 2756 m and 10306 m). Its performance is evaluated for the warm half year 2021 using ATDnet lightning data (Anderson and Klugmann, 2014). The operational AROME-BG model configuration (actually based on cy43t2) at NIMH is the following: the integration domain is covering Bulgaria, with a horizontal resolution of 2.5 km, 60 vertical levels, a time step of 60 s and a forecast range of 36 h. It runs four times daily - at 00, 06, 12 and 18 UTC and uses the ALADIN-BG output for initial and boundary conditions. As a first approach, cases with $r_g > 0$ are considered as probable for lightning occurrence. Due to the uncertainty of ATDnet accuracy over the region of Bulgaria, lightning data and forecasted graupel mixing ratios were considered on resolutions of 5x5 km and 10x10 km with flash rate for one and three hours, as well on a daily base using upscaling neighborhood method, with the aim to evaluate the accuracy of the forecast for the precise location and time respectively. Two daily model runs are considered – at 06 and 18 UTC. Commonly used skill-scores in meteorological forecasts are used as evaluation metrics – probability of detection (POD) and false alarm rate (F).

2 Results

Lightning probability forecast is evaluated at a diurnal, monthly and spatial base. Cases with lightning were considered as bins from the corresponding mesh with at least one detected flash. As an example, Figure 1 presents the spatial distribution of lightning density on a grid with resolution of 5 km for the period April-September 2021 over the considered region (left panel) as well the spatial distribution of the cases with at least one detected flash for the same period with the same spatial resolution (right panel). These last values are used to evaluate the ability of the forecasted by AROME-BG graupel mixing ratio to predict lightning activity.

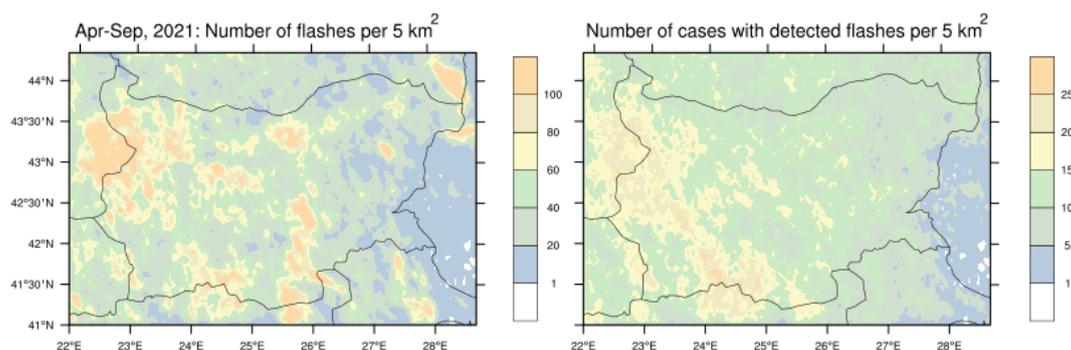


Figure 1: Number of flashes per 5 km² detected between April and September 2021 (left panel) and the corresponding number of cases with at least one detected flash on the grid with a resolution of 5 km (right panel) for the same period

In Figure 2 the diurnal distributions of the cases with at least one detected by ATDnet flash over the considered region on grids with a resolution of 5x5 km and 10x10 km, with frequencies of one and three hours for the period April-September 2021 are shown.

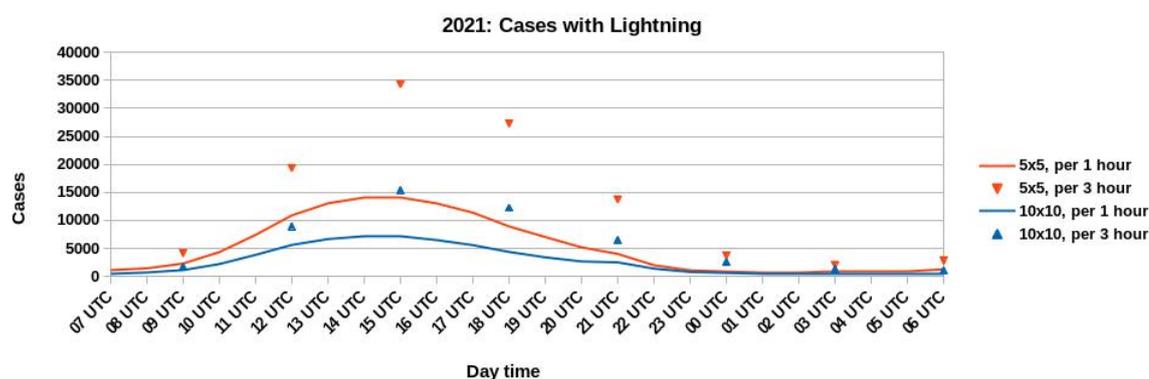


Figure 2: Diurnal distributions of the cases with at least one detected by ATDnet flash over the considered region on a grid with a resolution of 5x5 km (red) and 10x10 km (blue) with a frequency of one (line) and three hours (triangle) for the period April-September 2021.

Figures 3 and 4 show the diurnal distributions respectively of the probability detection (POD) and false alarm rate (F) for the different forecast resolutions and time frequencies for the two forecast runs – at 06 and 18 UTC. The monthly values of POD and F are shown in Figures 5 and 6.

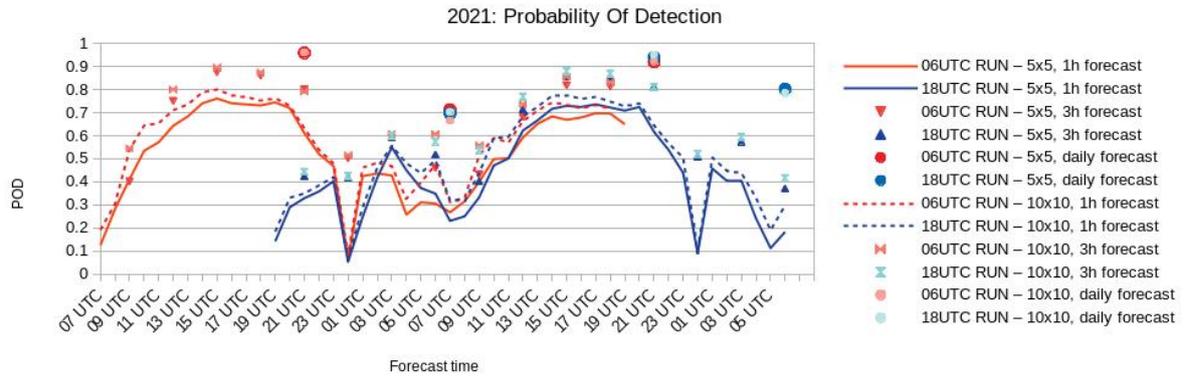


Figure 3. Diurnal distribution of the probability of detection (POD) for the period April-September 2021.

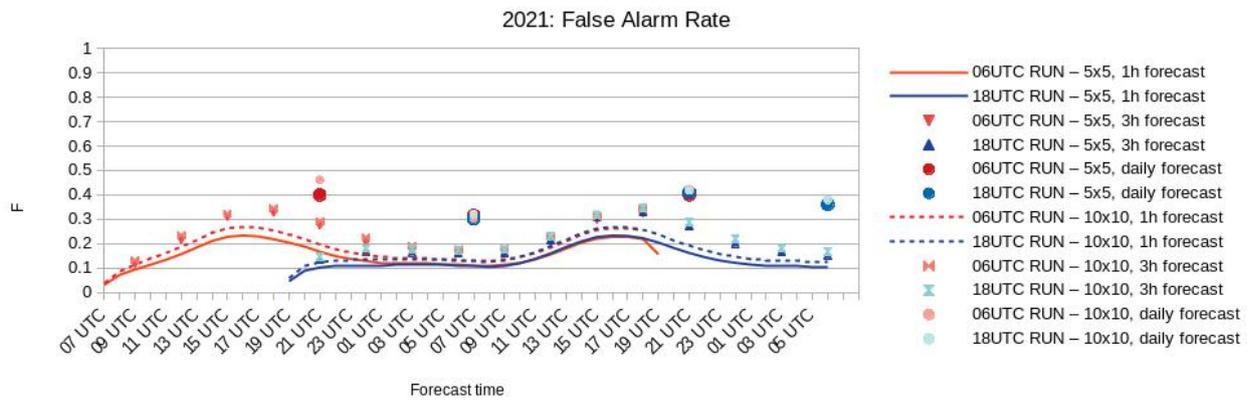


Figure 4. Diurnal distribution of the false alarm rate (F) for the period April-September 2021.

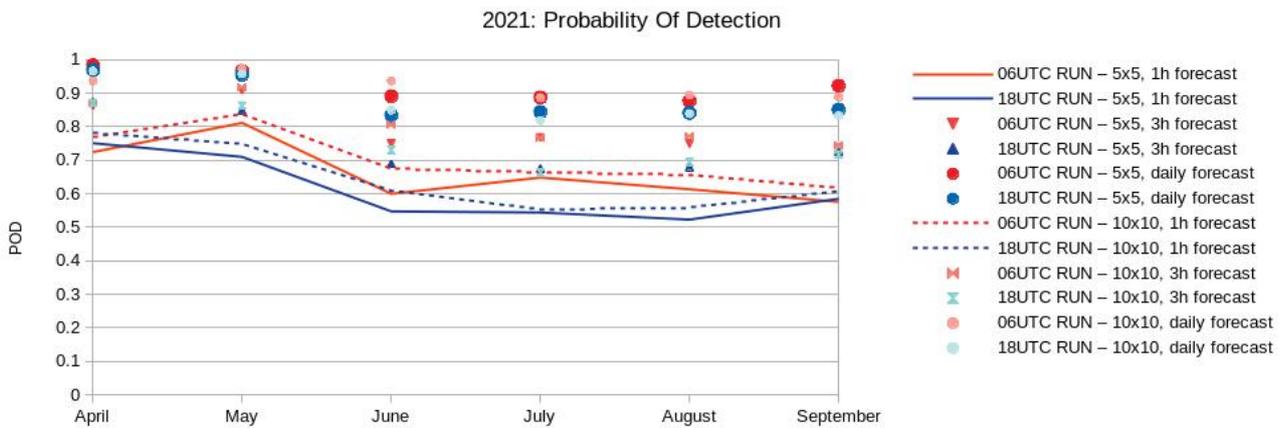


Figure 5. Monthly probability of detection (POD)

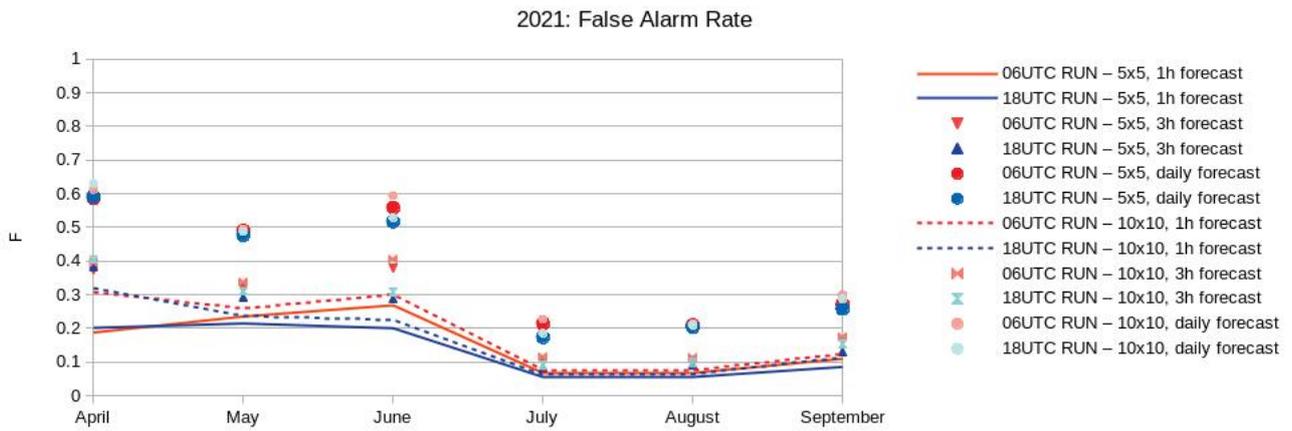


Figure 6. Monthly false alarm rate (F)

In Figures 3 and 4 are presented the diurnal distributions respectively of the probability detection (POD) and false alarm rate (F) for the different forecast resolutions and time frequencies for the period between April and September 2021 for the two forecast runs – at 06 and 18 UTC. The monthly values of POD and F are shown in Figures 5 and 6.

Figure 7 shows the monthly spatial distribution of POD and F for the model runs at 06 and 18 UTC and the detected by ATDnet flash density for the months April to September 2021.

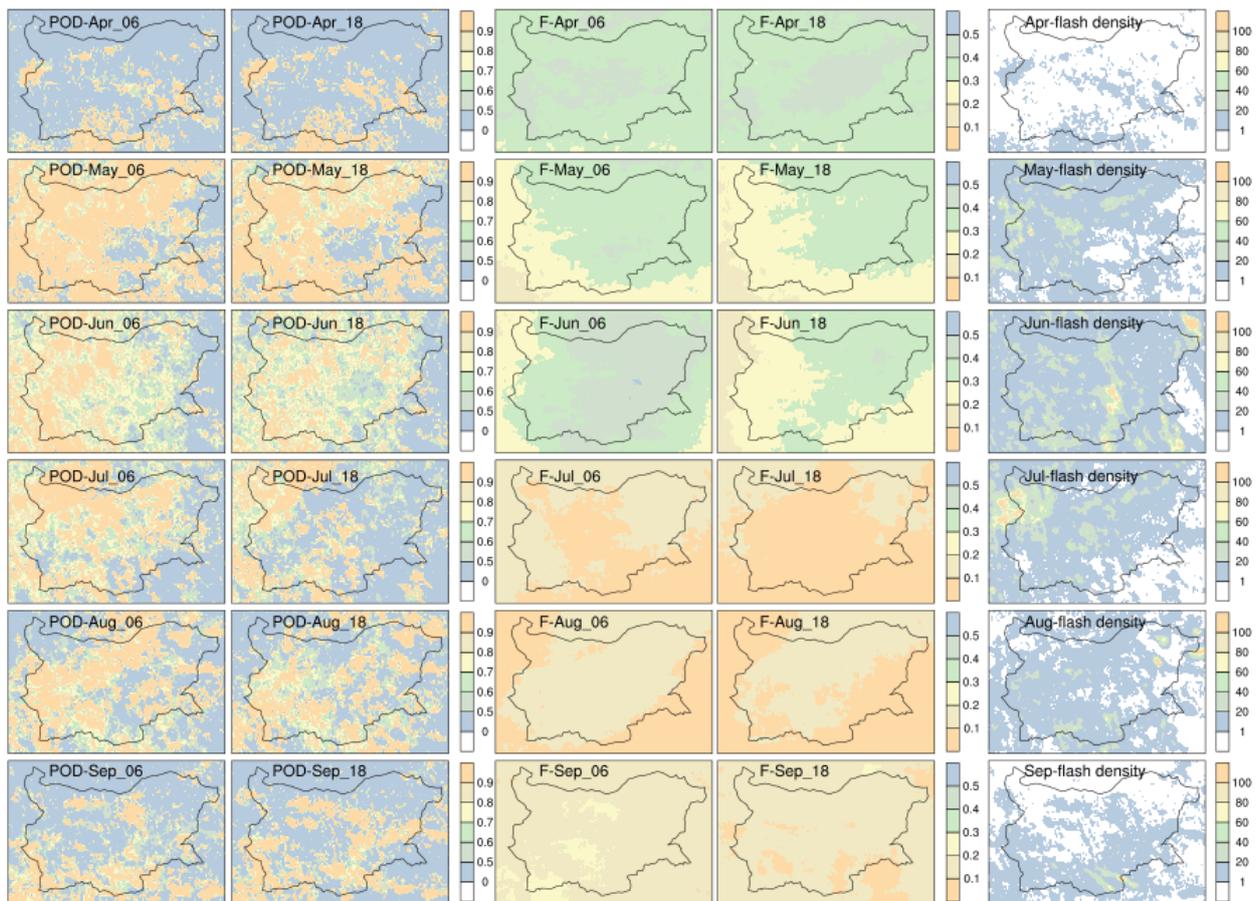


Figure 7. Spatial distribution on a grid of 5x5 km of probability of detection (POD) and false alarm rate (F) for the model runs at 06 and 18 UTC (denoted respectively with “_06” and “_18”) and the detected by ATDnet flash density for the months April to September 2021 used to evaluate model forecast with a time frequency of 3 hours

Results show that the decreases of forecast spatial resolution and time frequency lead to improvement of forecast probability of detection (POD) and to a slight deterioration of its false alarm rate (F) and the impact of forecast time frequency is more pronounced. Probability of detection (POD) is better during the day hours in comparison to night hours, which shows that lightning activity is better forecasted and with less overestimations during the day, when in principle most flashes are detected. False alarm rate (F) are less during the night hours. There are no significant differences in diurnal distribution between POD and F for the two different model runs at 06 and 18 UTC. Regarding spatial distribution, POD is higher for the first third of the warm half-year of 2021 (April and May), while F - in July, August and September; June is with lowest skill-scores. Spatial distribution of POD is different for the two model runs and the 06 UTC run is with higher values of probability of detection. In April, May and September, the probability of detection (POD) is high over the regions with detected lightning activity for the two model runs, but false alarm rate (F) have also high values (especially in April) over the whole considered domain. In June and July (the months with highest lightning activity over Bulgaria) skill-scores give different results for the two model runs with POD and F higher for 06 UTC run; better scores are obtained in western part of Bulgaria, while the forecast for lightning activity over the Black sea (eastern part of the domain) is worse.

The main conclusion from the study is that graupel mixing ratio taken from AROME-BG could be used as a tool to forecast lightning probability with a relatively high performance. The relatively low performance over the sea could be due to different reasons, as model performance or not most appropriate considered model levels over water pool. Such investigations have to be done in the future.

3 References

- Anderson, G. and Klugmann, D., (2014), European lightning density using ATDnet data, *Nat. Hazards Earth Syst. Sci.*, 14, 815–829, 2014
- Bruning E.C., Rust W.D., Schuur T.J., MacGorman D.R., Krehbiel P.R., Rison W., (2007), Electrical and polarimetric radar observations of a multicell storm in TELEX. *Monthly Weather Review*, 135 (7), 2525
- Dotzek N., Holler H., Theyer C., Fehr T., (2001), Lightning evolution related to radar-derived microphysics in the 21 July 1998 EULINOX supercell storm. *Atmospheric Research*. 56 (1-4), 335
- Jayarathne, E.R., Saunders, C.P.R., Hallett, J., (1983). Laboratory studies of the charging of soft-hail during icecrystal interactions. *Q. J. Royal Meteorol. Soc.* 109, 609 – 630.
- Reynolds, S., Brook, M., Gourley, M.F., (1957). Thunderstorm charge separation. *J. Met.* 14, 426 – 436.
- Takahashi, T., (1978). Riming electrification as a charge generation mechanism in thunderstorms. *J. Atmos.*, 1536 – 1548

The analog-based method application to gridded data post-processing

Iris Odak Plenković, Irene Schicker, Markus Dabernig, Alex Kann, Aitor Atencia

1 Introduction

Analogies between, for example, similar past forecasts, measurements, or analyses are a potentially useful tool when the training dataset is long enough, thus enabling an adequate identification of true analogs. Thus, reducing the number of degrees of freedom in the matching procedure makes this method an excellent candidate for point-based post-processing, where NWP input can be deterministic or ensemble forecast. Previously, the point-based analog approach was thoroughly tested as a deterministic approach (Odak Plenkovic et al., 2018) and applied to calibrate the A-LAEF ensemble (Odak Plenkovic et al, 2020).

However, accurate forecasts at remote locations are used to drive many user-specific applications (e.g., road temperature forecasts along an entire roadway, soil temperature forecasts for agriculture, wind speed for windfarms). For that reason, besides the point-based post-processing for the measuring sites, there is also an increasing demand for gridded products.

The latter is a direct motivation for the development of tools needed for using an analog-based method to produce gridded output based on an analysis. The purpose of this newsletter is to present recent developments achieved during the ACCORD stay at ZAMG, which took place November 22nd – December 17th, 2021, including some practical details. During this stay, the algorithms for two analog-based experiments that produce gridded products are developed and followed by preliminary results.

2 Data and algorithms

Previously used algorithms were not developed for large datasets, so optimization was the first necessary step. The scripts are adjusted to work in Python3 and were parallelized. Using the .h5 format in I/O in the `read_hdf` function accelerated the process. Due to the time constraints, datasets prepared during the previous stays are used. As input to the analog methods, the control member of the ECMWF ensemble forecast is used as a raw forecast. The gridded INCA analysis fields (Haiden et al., 2011) is used as a “ground truth”, similar to the observations in the point-based analog approach. For testing the novel gridded analog algorithm within a decent amount of time, the INCA wind speed analyses are bilinearly interpolated onto the ECMWF forecast grid. This will be changed once the algorithms are properly tested and validated.

After the algorithms optimizations, several experiments were performed, including two distinctive approaches, Point-by-Point and Field-wise, that will be explained in more detail afterward. All experiments in this work use wind speed and direction variables as predictors, normalized by standard deviation but no additional predictor-weighting strategy is currently applied.

The analog ensemble (AnEn) consists of 10 INCA wind speed values corresponding to 10 best-matching analogs for the ECMWF control member. The analyzed INCA wind speed values are also used as observed values in the verification procedure. Since the years used for training and a year as indented validation data set in the following experiments sometimes differ, they will be explicitly stated for each experiment further on.

Point-by-Point approach

The first approach is the simplest transfer from point-based to gridded products: treating every grid point as an independent location. The quality of the analog is thus evaluated by the following metric:

$$\|F_t A_{t'}\| = \sum_{i=1}^{N_A} \frac{w_i}{\sigma_{fi}} \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2}, \quad (1)$$

where F_t is the current NWP deterministic forecast at a given grid point, valid at the future time t , whereas $A_{t'}$ is an analog at a given point with the same forecast lead time, but valid at a past time t' . The N_A is the number of predictors used in the search for analogs, w_i are the weights corresponding to a particular predictor, normalized with the standard deviation of the time series of past forecasts of a given variable at the same grid point is σ_{fi} . The \tilde{t} is equal to half the number of additional times over which the metric is computed (the half of the time window of any specified width), therefore $F_{i,t+j}$ and $A_{i,t'+j}$ are the values of the forecast and the analog in the time window for a given variable, respectively. Analogs are found independently for every forecast time and grid point, narrowing the search around the particular time of a day by a time window. The example of probabilistic forecast produced by this approach is shown in Figure 1. To emphasize the fact that grid points are treated independently, only the markers at the exact grid point locations are used here. The latter also provides better insight into the domain size and horizontal resolution for all the experiments provided in this work. The standard practice, however, would be using contour function for displaying such forecasts. It can be noted that, even though the Point-by-Point approach is a good starting point, it might produce noisy forecasts (as in Frediani et al., 2017). Additionally, since every grid point is treated separately, the method itself is still computationally slow.

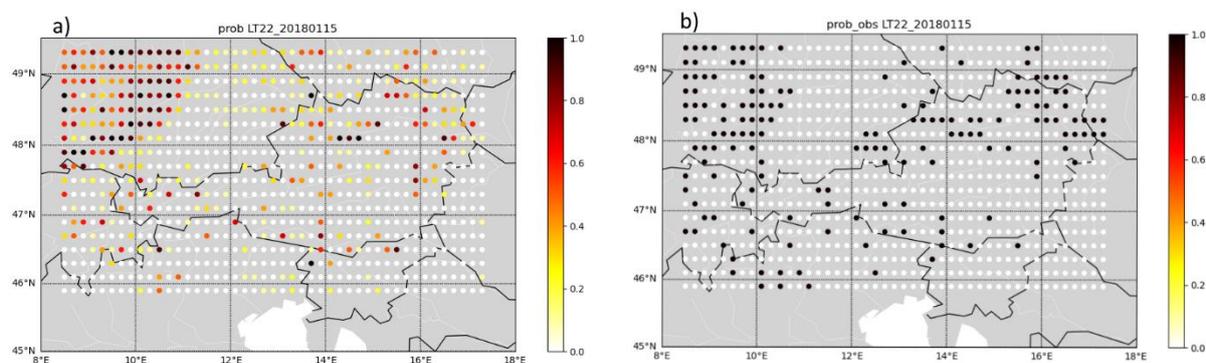


Figure 1. The example of probabilistic forecast that wind speed will exceed 5 m/s predicted by Point-by-Point analog-based approach for January 15th, 2018 (lead time 22 h) and the domain used for all the experiments (a) and the corresponding observations derived from INCA analysis and bilinearly interpolated to ECMWF grid (b).

Field-wise approach

Similarly, to the previous approach, the distance metric (1) is calculated independently for every forecast time and grid point using ECMWF forecasts, narrowing the search around the particular time of a day by a time window. However, in order to determine the best match (i.e., best analog), the calculated metric is averaged over an entire field, instead of each grid point separately. In other words, one can compare an average error on the entire field and use the mean value to choose the most similar fields, selecting the 10 lowest values. The gridded INCA wind speed analyses corresponding to those historical timestamps are used and they include all grid points.

For these experiments, for fairer comparison with the Point-by-Point approach, only the INCA wind speed values bilinearly interpolated to ECMWF grid are used as members of analog ensemble forecast (as in Figure 1). However, such a constraint is not generally required. Moreover, the potential benefit of this approach is the possibility of using different grid setups, e.g., differing in horizontal resolution and/or exact location of grid points. Additionally, since the Field-wise approach is much more computationally quicker, several experiments are executed.

3 Results

As already mentioned, several different analog-based experiments are executed and compared, using INCA wind speed analysis as observed values. The experiments are for one winter (January) and one summer month (July), producing consistent results. For that reason, only the results for January are shown in this work.

The main aim of any kind of NWP model post-processing is to improve the results of the raw model. All analog-based experiments show an improvement compared to the raw ECMWF forecasts. The latter can be noticed if, for example, the lead-time performance measured by RMSE for raw ECMWF (Figure 2) is compared to AnEn RMSE values for the ensemble mean in Figures 3b and 5.

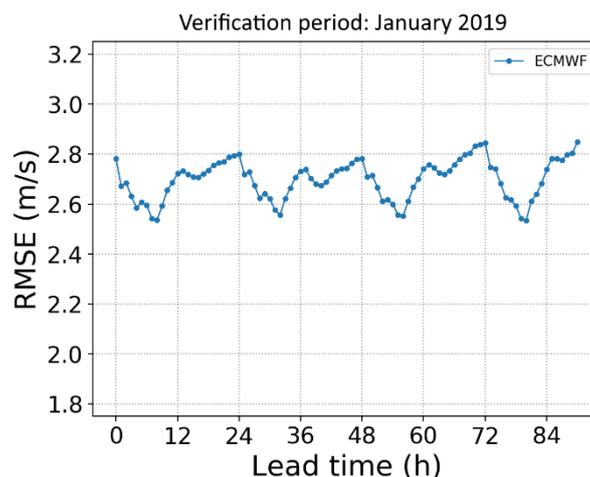


Figure 2. The RMSE values for ECMWF control ensemble member forecast at different lead times during January 2019. The forecasts are verified using INCA analysis wind speed values on the entire domain.

The Point-by-Point and Field-wise analog approaches are compared, taking into consideration the simplicity and effectiveness of execution (e.g., computational speed) as well as success in overall performance (e.g., better verification scores). Even though the Point-by-Point approach is much more similar to point-based application and thus was much easier to implement, the algorithm is still relatively slow, whereas the Field-wise approach is much faster. For example, for a month-long Point-by-Point algorithm that uses a 1-year-long training dataset, the execution lasted ~5 h. On the other hand, to produce month-long Field-wise-approach-based forecasts that use a 2-years-long dataset for training needed only ~40 min to finish using the same machine. The difference in execution is probably due to averaging the distance metric, reducing the number of times that the Field-wise approach needs to open/close INCA files, etc. For these practical reasons, it made sense to compare the Field-wise approach with longer training to the Point-by-Point approach with shorter training (Figure 3). The results for January 2019 are comparable for these two approaches, measured by CRPS and RMSE-spread plot (Figure 3). One can notice that differences among approaches slightly increase with the lead time. Also, the Field-wise approach seems to be less prone to underdispersiveness.

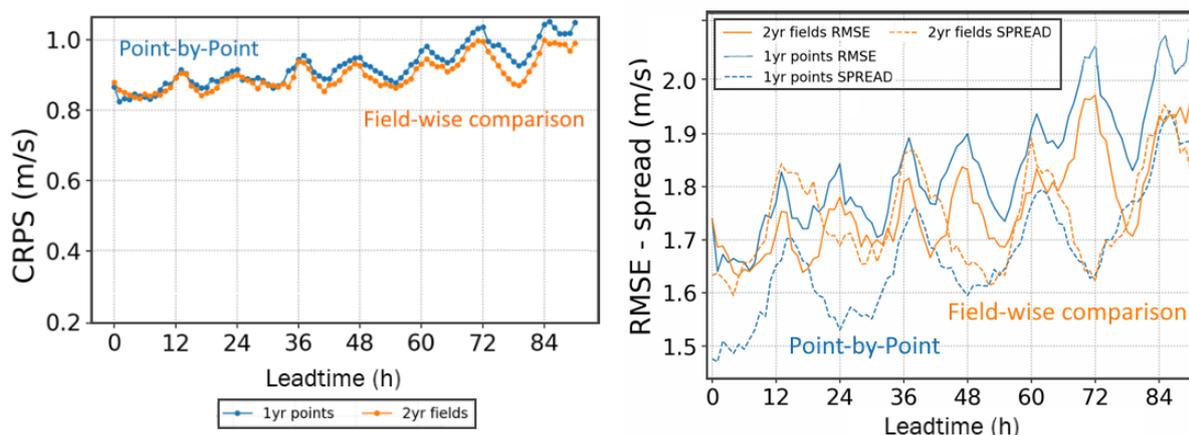


Figure 3. The CRPS (a) and RMSE-spread (b) for the Point-by-Point analog-based approach (using a 1-year-long training dataset) is compared to the Field-wise approach (using a 2-years-long training dataset) during January 2019. All forecasts are verified using INCA analysis wind speed values on the entire domain.

Since the Field-wise approach with 2-years-long training is used for inter-comparison with the Point-by-Point approach, the sensitivity to training length is also examined. The results show that error measured by AnEn mean RMSE is smaller for the longer training, as expected, whereas the spread remains similar (Figure 4).

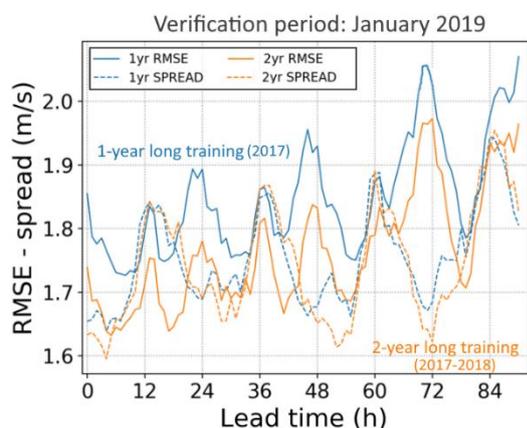


Figure 4. The intercomparison of the Field-wise analog-based method approach that uses a 1-year-long training dataset (2017) and the one that uses a 2-years-long training dataset (2017-2018) using the RMSE-spread plot. The results are calculated for January 2019. All forecasts are verified using INCA analysis wind speed values on the entire domain.

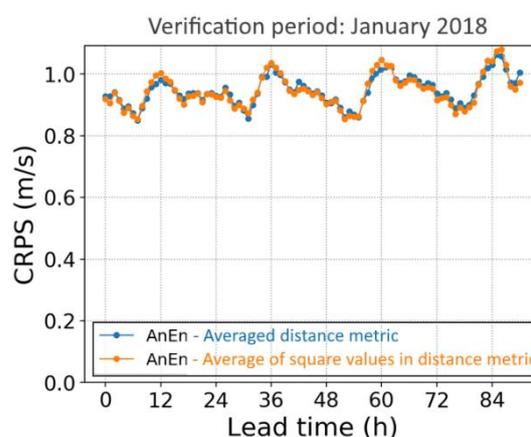


Figure 5. The Field-wise analog-based experiment for which the average distance metric on the entire field is used to define best matching analogs is compared to the experiment for which the squares of distance metric are averaged, weighing the larger values at the grid points more. The comparison is made using the CRPS measure and shown for January 2018. All forecasts are verified using INCA analysis wind speed values on the entire domain.

While implementing the Field-wise approach the question arose if it would be better to weigh the larger distance metric (at a particular grid point) values more, before averaging the values for the entire field and using it for choosing the best analogs. To answer if such modification could improve the result

further, the distance metric values at grid points are squared and then averaged. The difference metric across the field would thus be treated more like “root-mean-square-difference” than like “mean bias”. The results however show very few differences among these two experiments (Figure 5). For that reason, the more simplified approach that uses only averaging is adopted.

4 Discussion and future work

To summarize, the tools needed for using an analog-based method to produce gridded output are successfully developed. Two distinctive approaches are tested, Point-by-Point and Field-wise approach, generating comparable results when using training datasets of the same length. The preliminary results also show that the Field-wise experiment with 2-years-long training seems to produce the best result, gets very close to the Point-by-Point experiment with 1-year-long training, and is computationally less demanding.

In addition to these experiments, future work might include an additional method for choosing the best analogs in order to simplify the information. For instance, one can identify objects (also in Frediani et al., 2017), use principal components (PC, as in Xavier and Goswami, 2007) or empirical orthogonal functions (EOF; similarly as a point-based application in Barnett and Preisendorfer, 1978; also Zorita and Storch, 1998). Additional calibration might also be done using ensemble model output statistics (EMOS). In addition to these ideas, methods such as quantile mapping and rank-weighted best-member dressing (Hamill and Scheuerer, 2018) or Schaake shuffle (as in Scheuerer and Hamill, 2018) can also be considered.

Acknowledgments

The authors would like to thank our colleagues in the ACCORD and RC-LACE consortiums for their support in our work.

5 References

Barnett, T.P. and R.W. Preisendorfer, 1978: Multifield Analog Prediction of Short-Term Climate Fluctuations Using a Climate State vector. *J. Atmos. Sci.*, 35, 1771–1787, [https://doi.org/10.1175/1520-0469\(1978\)035<1771:MAPOST>2.0.CO;2](https://doi.org/10.1175/1520-0469(1978)035<1771:MAPOST>2.0.CO;2)

Frediani, M.E., T.M. Hopson, J.P. Hacker, E.N. Anagnostou, L. Delle Monache, and F. Vandenberghe, 2017: Object-Based Analog Forecasts for Surface Wind Speed. *Mon. Wea. Rev.*, 145, 5083–5102, <https://doi.org/10.1175/MWR-D-17-0012.1>

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., Gurber, C., 2011: The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region. *Weather and Forecasting*. 26. 166-183. <https://doi.org/10.1175/2010WAF2222451.1>.

Hamill, T.M. and M. Scheuerer, 2018: Probabilistic Precipitation Forecast Postprocessing Using Quantile Mapping and Rank-Weighted Best-Member Dressing. *Mon. Wea. Rev.*, 146, 4079–4098, <https://doi.org/10.1175/MWR-D-18-0147.1>

Odak Plenković I, Delle Monache L, Horvath K, Hrastinski M (2018) Deterministic Wind Speed Predictions with Analog-Based Methods over Complex Topography. *Journal of applied meteorology and climatology* 57:2047-2070.

Odak Plenković I, Schicker I, Dabernig M, Horvath K, Keresturi E (2020) Analog-based post-processing of the ALADIN-LAEF ensemble predictions in complex terrain. *Quarterly Journal of the Royal Meteorological Society* 146:1842– 1860.

Scheuerer, M. and T.M. Hamill, 2018: Generating Calibrated Ensembles of Physically Realistic, High-Resolution Precipitation Forecast Fields Based on GEFS Model Output. *J. Hydrometeor.*, 19, 1651–1670, <https://doi.org/10.1175/JHM-D-18-0067.1>

Xavier, P.K. and B.N. Goswami, 2007: An Analog Method for Real-Time Forecasting of Summer Monsoon Subseasonal Variability. *Mon. Wea. Rev.*, 135, 4149–4160, <https://doi.org/10.1175/2007MWR1854.1>

Zorita, E. and H. von Storch, 1999: The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. *J. Climate*, 12, 2474–2489, [https://doi.org/10.1175/1520-0442\(1999\)012<2474:TAMAAS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2)

AROME-MAROC CY41T1 QUALITY ASSESMENT USING THE SIMULATED SATELLITE IMAGES IN TWO MSG CHANNELS WV6.2 AND IR10.8

Meryem BOUCHAHMOUD, Abderrahim MOUHTADI, Siham SBII, Ilias KACIMI

1 Introduction

The NWP Moroccan suite is mainly based on the non-hydrostatic model AROME-MOROCCO running on 2.5 km of resolution and 90 vertical levels. This work aims to assess its reliability using two MSG channels IR10.8 and WV6.2. The used data correspond to an hourly forecasts produced by AROME-MAROC cy41t1 during the whole year of 2019. Throughout this work, two spatial interpolation methods are used, the nearest model grid point and the inverse distance method.

2 FORECAST EVALUATION

2.1 MATERIALS

For this study, we worked with AROME-MAROC data that was provided on an hourly basis, over an area covering from latitude 20.5 degrees to 37.0 degrees and from longitude -0.25 degrees to -17.5 degrees. During this work, we limited our study to the daily running products from midnight.

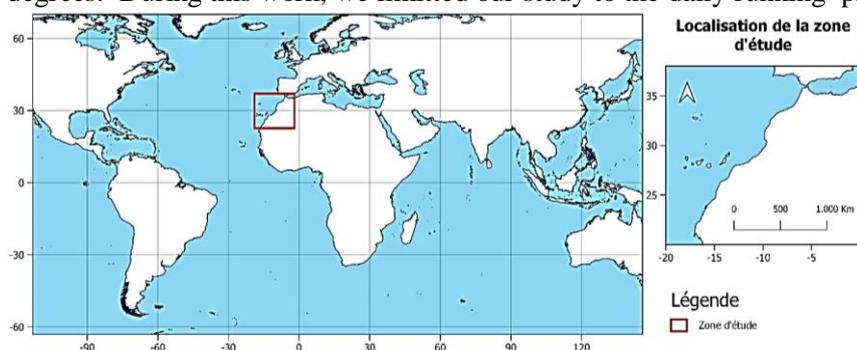


Figure 1: study area

As observational data, we used the Second Generation Meteosat data that was retrieved from the EUMETSAT platform, on data store in the High Rate Collection SEVIRI Level 1.5 Image Data-MSG-0 degree. The data used corresponds to the last 15 minutes for every hour and is available on a 3km resolution.

2.2 METHODS

The downloaded MSG data has been processed to retrieve only the two above-mentioned channels and convert the radiances to temperature. After, we sampled the AROME MAROC data on the MSG data grid by two different approaches: the nearest neighbour approach and the inverse of the distances.

2.3 EVALUATION FRAMEWORK

In the present study, the forecast verification mainly based on two scores BAIS and RMS.

$O_t - F_t$ is the difference between observed and forecasted brightness temperatures at the moment: (t).

$$RMS = \sqrt{\frac{1}{n} \sum_{t=1}^n (O_t - F_t)^2}$$

To assess the reliability of the forecasted satellite images in the two channels IR10.8 and WV6.2, the study was carried out according to different periods (hourly, monthly and seasonally).

3 RESULTS

3.1 NEAREST NEIGHBOR RESULTS

As shown in figure 1, the model AROME tends to overestimate both the top surface temperatures of objects as well as the water vapor content in WV6.2 channel. Also, daytime behaviour is marked by a maximum overestimation between 8:00 a.m. and 10:00 a.m. and a minimum in the afternoon for both channels. As for the RMS, it consolidates the overestimation of the afternoon for both channels.

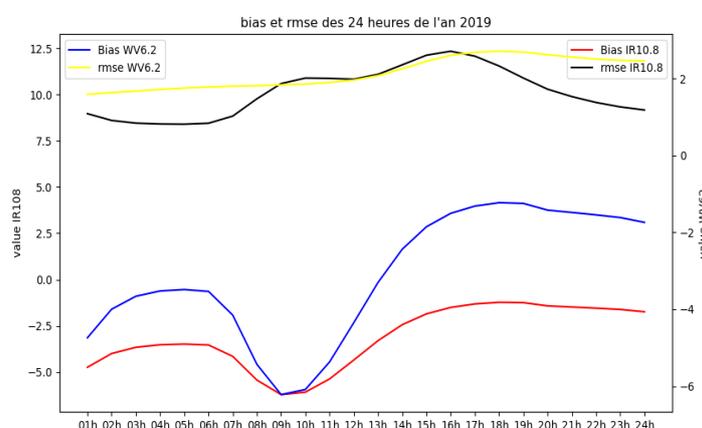


Figure 2: Biais and RMS for 24h of the year 2019

Almost the same observation is obtained when studying the hourly behaviour on a seasonal basis (Fig2, and Fig3). The shape of the curves confirms the diurnal behaviour of the model which is materialized by a peak of overestimation around 10:00 a.m. (-6 to -7 ° K) and a relatively less marked overestimation in the afternoon (-3 to -2 ° K), for both channels. The RMS (10.8 and WV6.2) show relatively large amplitudes especially in the afternoons.

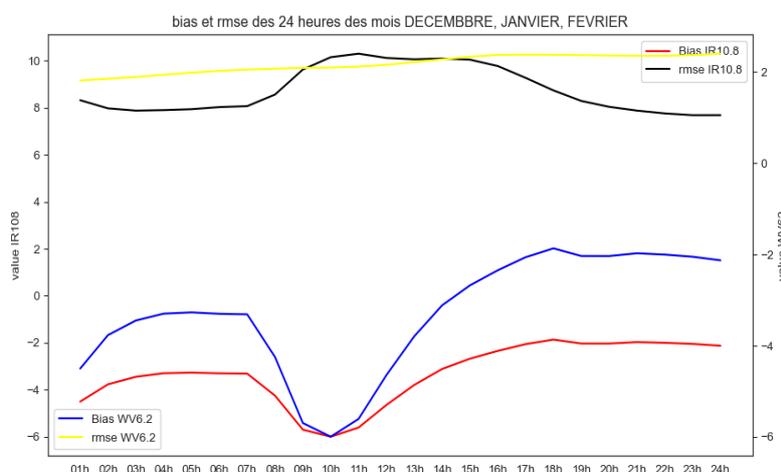


Figure 3: Biais and RMSE for December, January and February 2019

By comparing the seasons, we see that the RMS of the season (June-July-August) presents greater amplitudes than the other seasons in the afternoon around 15 ° K. This is due to the presence of unstable situations during the summer season, which give rise to the formation of convective patterns. These unstable clouds do not have a significant horizontal extension. These are usually isolated clouds that do not occupy a large area and therefore, during the interpolation, there would be a high probability of not have the same radiance value.

On the contrary, during the season (December-January-February) the amplitude of the RMS presents the lowest values (between 8 and 10 ° K) in comparison with the other seasons. The winter season is known by the meteorological disruptions that interest Morocco, so more cloudiness and therefore the deviations of the forecast from the observation will be less intense and therefore the probability of having values close to radiance is relatively high during the interpolation.

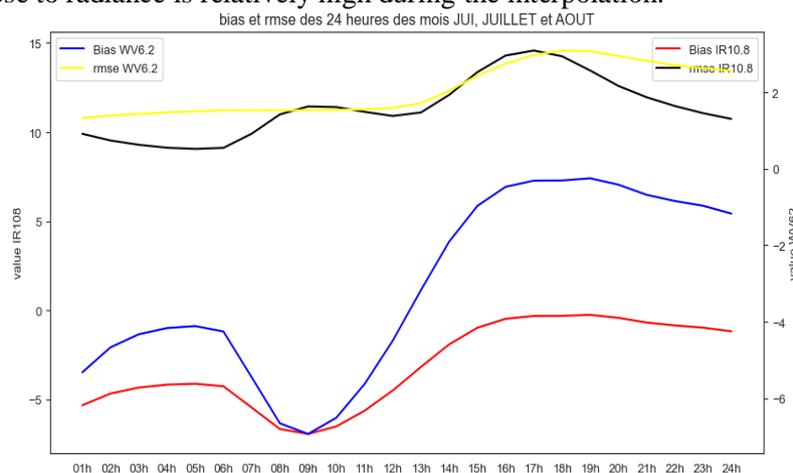


Figure 4: Bias and rms June, July and August

Regarding the monthly analysis of the bias and RMS, we noted some fluctuations from a month to another (Figure 5) which means that AROME MAROC tends to overestimate the temperatures of the upper surfaces of objects throughout the year. As for the hourly average of the deviation from the forecast of the WV6.2 product for the months (blue curve, figure 5), it oscillates between -0.4 and + 0.5K. This result shows that for the predicted satellite image corresponding to channel 6.2, AROME Maroc underestimates (in general) the water vapour content from late spring to autumn (positive values: May-June-July-August September and October); and overestimates this content during the winter period and the beginning of spring (negative values: November- December- January- February- March and relatively April).

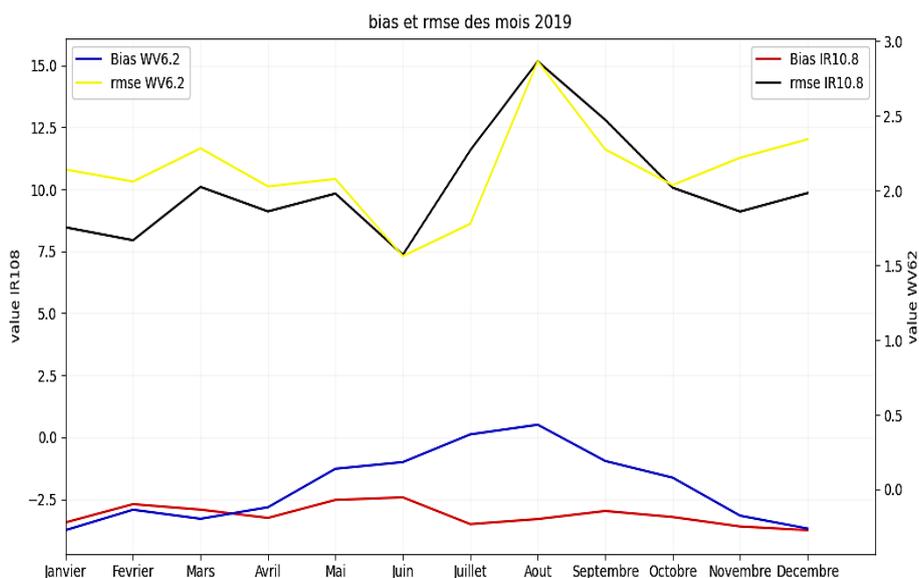


Figure 5: BIAS and RMS per month

2.5 INVERSE DISTANCE WEIGHT RESULTS

For the inverse distance approach (IDW), we took into consideration the 4 most closed neighbours points. The results obtained are similar to the nearest neighbour with slightly less values of Bias and RMS

3 CONCLUSION

During this study, and in order to assess the reliability of AROME-MAROC cy41t1, a special emphasize was done on the simulated satellite images in two channels WV6.2 and IR10.8. The results show a diurnal behaviour materialized by relatively important errors between 8:00 a.m. to 10:00 a.m. and 2:00 p.m. to 4:00 p.m. The model has tendency to overestimate the radiances of the upper surface of objects, this overestimation is seen for the channel IR10.8. And for channel WV6.2, the overestimation or underestimation are relatively small and may be acceptable. This study also showed that the overestimation of radiances is much greater in convective situations.

HOOF - a preprocessing tool for OPERA weather radar data files

Peter Smerkol, Slovenian Environmental Agency

1 Introduction

OPERA (Operational Programme for the Exchange of Weather Radar Information) [1] is a programme of EU-METNET which provides 3D data from weather radars in Europe in HDF5 and BUFR formats. The provided data is quality controlled and organized in a standardized format called ODIM (OPERA Data Information Model) [2].

Although the files obtained from OPERA are standardized, there can still be differences in the file structure and metadata between files coming from different countries, as countries have different ways of organizing data within the ODIM format. This makes it challenging to input the files for different countries automatically into the ALADIN code for assimilation of data into the model. Specifically, the data from files is input into the ALADIN preprocessing package BATOR.

HOOF (Homogenization Of OPERA Files) tool was originally devised to solve this problem, by homogenizing the internal structure of all obtained OPERA HDF5 files to the level that is required in order to input the files into the BATOR package (section 2). Recently, a new version of HOOF was developed, which additionally provides dealiasing of radial wind measurements (section 3) and superobing of data (section 4).

HOOF is written in the Python3 language and is controlled via a namelist file. It can be run in the terminal to process all files in a specified folder, or can be run from a graphical user interface to process one file at a time and at the same time graphically inspect the processed results.

2 Homogenization

The ODIM standard defines various groups in the HDF5 structure for specific data and metadata. Generally, a `/dataset` group represents one radar scan, defined by its start date and elevation angle. Each quantity (such as reflectivity, radial wind, PPI, etc.) belonging to a scan is then written into separate `/dataset/data` groups and the corresponding indexes obtained from quality control into `/dataset/quality` groups. The relevant metadata for scan and quantity (such as elevation angle, number of rays, wavelength, etc.) is written into accompanying `/what`, `/where` and `/how` groups.

For radars from different countries, this general structure is only loosely adhered to, because of different scanning strategies, different saved metadata, different types of data measured, and so on.

In HOOF, you can specify which type of quantities and metadata are retained after homogenization, and the algorithm arranges the retained data into groups in the same way for all radars. It first arranges scans according to start dates and elevation angles, and finds the quantities belonging to the scans. The general structure of the homogenized file is shown in figure (1).

In this way, by specifying the necessary quantities and metadata, you can ensure that the BATOR package will successfully read the data it needs for assimilation into the ALADIN model.

3 Dealiasing

Because radial wind velocity is measured via phase difference of the electromagnetic wave between two radar bursts, which is defined on the interval $[-\pi, \pi]$, the resulting measurements are aliased on an interval $[-v_{ny}, v_{ny}]$, where v_{ny} is the velocity measured at the maximum phase, called the Nyquist velocity. All velocities that are bigger than the Nyquist velocity (in absolute value) are transformed back to the interval $[-v_{ny}, v_{ny}]$. That means that the true velocity v is related to the measured velocity v_{obs} by:

$$v = v_{obs} + 2nv_{ny}, \quad (1)$$

where n is an unknown integer number, called the Nyquist multiplier, which has to be determined for each data point.

In HOOF, the dealiasing is done with the torus mapping method, devised in [3]. In this method, we assume a linear wind model, where the radial wind velocity is expressed with the azimuth and elevation angles:

$$v_m = (u \sin \alpha + v \cos \alpha) \cos \vartheta, \quad (2)$$

where u and v are the zonal and meridional wind components, α is the azimuth angle and ϑ is the elevation angle.

For observations at constant distances from radar and constant elevation angles, this model would give a sinusoidal curve in azimuth. However, because of aliasing, the curve becomes discontinuous. This can be avoided by mapping the curve onto a torus (see figure (2)).

The curve mapped onto the torus is a parametric curve:

$$F(\varphi) = \left(\left[R + \frac{v_{ny}}{\pi} \sin \left(\pi \frac{v_{obs}}{v_{ny}} \right) \right] \sin \alpha, \left[R + \frac{v_{ny}}{\pi} \sin \left(\pi \frac{v_{obs}}{v_{ny}} \right) \right] \cos \alpha, \frac{v_{ny}}{\pi} \cos \left(\pi \frac{v_{obs}}{v_{ny}} \right) \right), \quad (3)$$

where R is the torus radius satisfying $R > v_{ny}/\pi$.

From this, we can express:

$$D = \frac{\partial F_3}{\partial \alpha} = -au + bv, \quad (4)$$

$$a = \cos \alpha \cos \vartheta \sin \left(\pi \frac{v_{obs}}{v_{ny}} \right), \quad (5)$$

$$b = \sin \alpha \cos \vartheta \sin \left(\pi \frac{v_{obs}}{v_{ny}} \right). \quad (6)$$

The a , b and D coefficients can be calculated from each measured point in a chosen dataset, and the unknown wind model components can then be calculated by minimizing the equation (4) with respect to u and v .

For the same chosen dataset, the Nyquist multiplier n can then be found by minimizing the equation (1) with respect to n .

In HOOF, data is divided into datasets of points coming from the same height layer (with a default height of 100 m). This is done to try to maximally satisfy the assumption of the linear wind model.

Examples of results of the HOOF dealiasing can be seen in figures (3) and (4). The procedure works well in a large percent of cases, and an algorithm for rejecting the wrongly dealiased data is in development.

4 Superobing

Superobing is a procedure where many (close) observation points are combined into one point (a superobservation), which has a value that best represents all the combined points. This is done in order to reduce the number of points that are assimilated into the ALADIN model, thus reducing the memory requirements and calculation time.

In HOOF, the superobing is done as in the *preopera.py* Python script, used by the HIRLAM community, with small corrections in the code that allow for bins of (almost) arbitrary size in the ray direction.

For reflectivity points, the superobed value is set to the average if there are enough (> 30%) good quality (> 0.7) rainy (> 12 dB) points. Otherwise, if there is at least one good quality dry point, the value is set to the minimum value (-32 dB), otherwise the point is discarded.

For radial wind points, the superobed value is set to the average if there are enough good quality points with a small standard deviation (< 10 *m/s*), otherwise the point is discarded. All the values quoted can be changed via the namelist.

Some examples of the results of the superobing procedure are shown in figure (5).

5 Conclusions

The upgraded version of HOOF now contains all procedures needed to use the reflectivity and radial wind data provided in OPERA HDF5 files, but it has to be noted that the dealiasing and superobing procedures still has to be properly validated (currently ongoing work).

The current version of HOOF is available on the ACCORD wiki page [4], and the full documentation is being written and will be available soon on the same page.

For all comments and questions, I am available via email: peter.smerkol@gov.si

References

- [1] <https://www.eumetnet.eu/activities/observations-programme/current-activities/opera/>
- [2] https://www.eumetnet.eu/wp-content/uploads/2021/07/ODIM_H5_v2.4.pdf
- [3] G. Haase, T. Landelius 2004, *Dealiasing of Doppler radar velocities using a torus mapping*, J. Atmos Oceanic Technol. 21, 1566-1573.
- [4] <https://opensource.umr-cnrm.fr/projects/accord/wiki/HOOF>

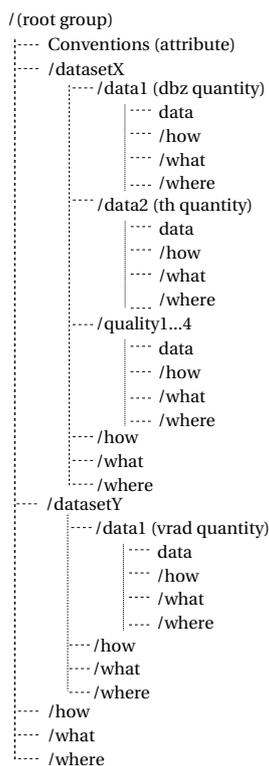


Figure 1: General structure of the HOOF homogenized output file, prepared for input into the BATOR package.

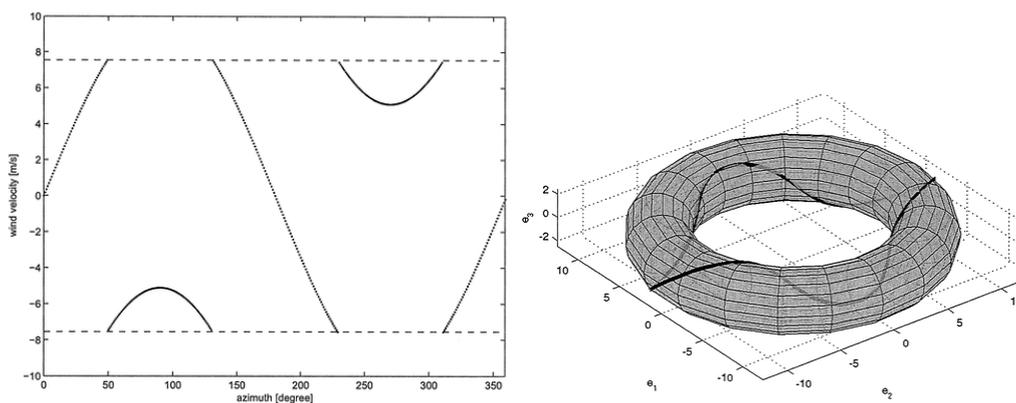


Figure 2: Left: Curve of aliased linear wind model expressed with azimuth at constant distance from radar and elevation. Right: The same curve mapped onto a surface of a torus.

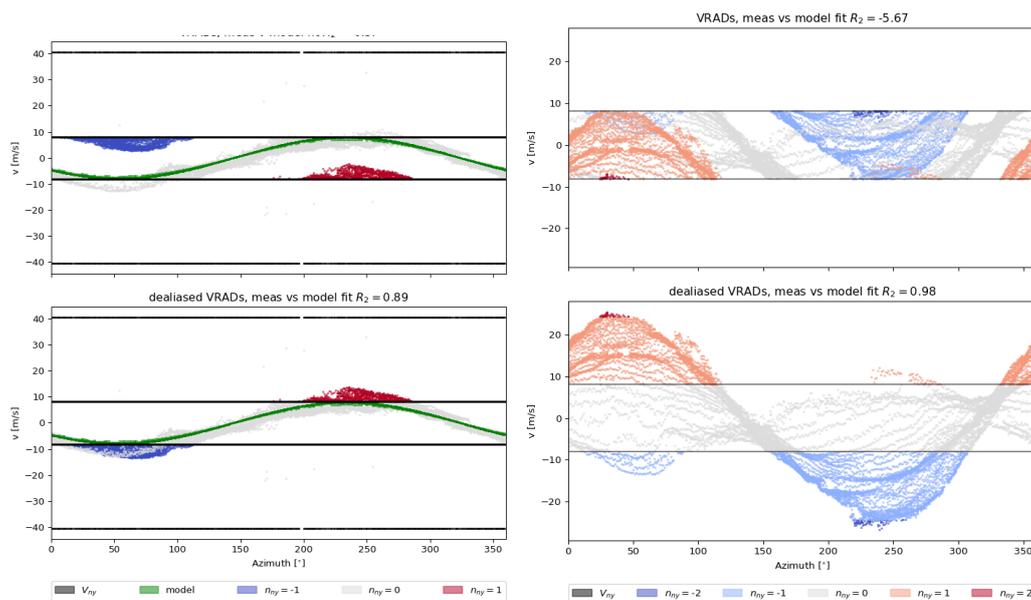


Figure 3: Velocity-azimuth plots of dealiasing results. Left: Result for a dataset from a single height sector. Right: Result for combined datasets from one radar elevation.

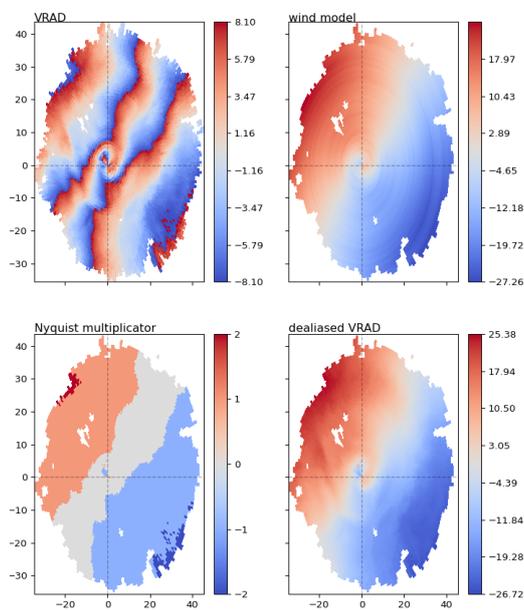


Figure 4: Geographical plots for data from one radar elevation. Top left: Aliased data. Top right: Wind model. Bottom left: Nyquist multiplier. Bottom right: Dealiasied data.

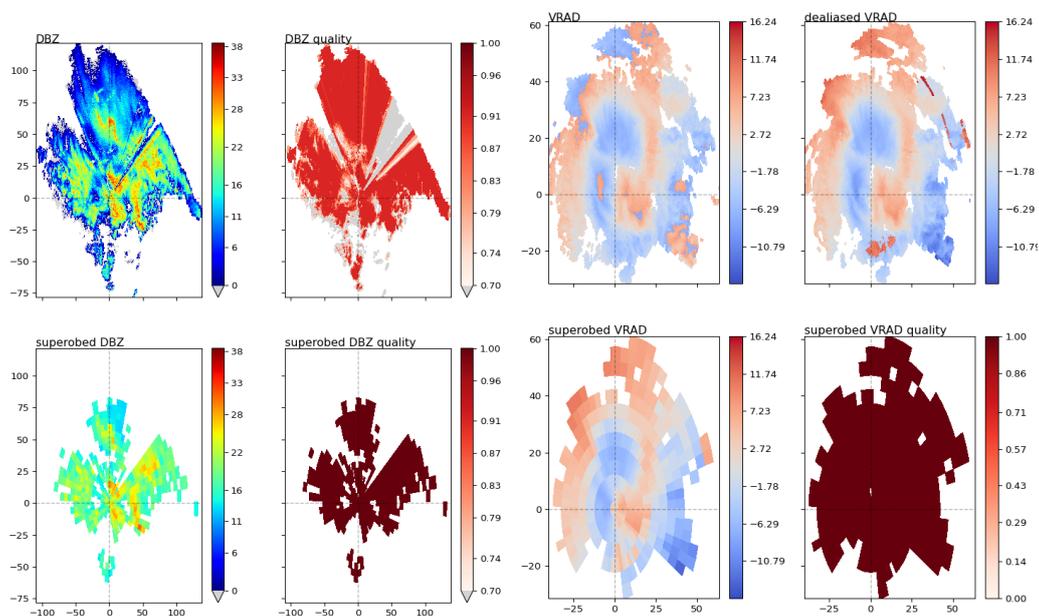


Figure 5: Geographical plots of superobing results for reflectivity data (left four plots) and radial wind data (right four plots). For reflectivity, top left plot shows original data, top right original quality of data, bottom left the superobed data and the bottom right the new superobed quality. For radial wind data, top left plot shows original aliased data, top right dealiased data, bottom left superobed dealiased data and bottom right the new superobed quality.

The “46t1 e-suite” running at Meteo-France: content and evaluation

CNRM/GMAP (NWP Research group) and DIROP (Operations group) collective article

1. Introduction

This article presents the main aspects of the “46t1 e-suite” that has been developed at Meteo-France in preparation of the operational NWP chain from 2022. It is the first e-suite since the two new HPC facilities ATOS Bull Sequana XH2000 *belenos* and *taranis* were set into operations in early 2021, which offer a gain of computing power by a factor ~ 5 compared to the previous HPC. In this e-suite, the resolution of many systems increase: ARPEGE-EPS, AROME-EPS and the AROME-overseas. As a result, for the first time at Meteo-France, the resolutions of the EPS reach the resolutions of their deterministic counterparts, whose forecasts are then embedded in their respective EPS. Another key aspect of the “46t1 e-suite” is an in-depth renovation of the ARPEGE physics and the implementation of the Tiedtke-Bechtold (Tiedtke (1989), Bechtold et al. 2008, Bechtold et al. 2014) scheme for the parameterisation of convection. Many other changes have been introduced, in the ARPEGE and AROME assimilation and physics, which lead eventually to large and significant improvement of forecast scores. The “46t1 e-suite” has been running since October 2021 for all the models and it is foreseen to be implemented into operations before summer 2022. A reforecast of the whole chain covering a one-year period has been run, in order to prepare an up-to-date learning database for the statistical post-processing before the operational implementation.

Table 1: General description of the “46t1 e-suite” running at Meteo-France,

	Configuration	Resolution	Comments
ARPEGE	Global domain, 4D-VAR assimilation (2 minimisations: T224C1/T499C1), forecast+102h (00, 06, 12, 18 UTC)	T1798C2.2 L105 (from 10m to 0.1hPa)	Background error covariances given by ARPEGE EDA
ARPEGE EDA	Global domain, 50 members	T499C1 L105 (same as ARPEGE)	
ARPEGE EPS	Global domain, 34 perturbed members + ARPEGE forecast, forecast+102h (00, 06, 12, 18 UTC)	Same as ARPEGE	Initial conditions given by ARPEGE analysis + perturbations from ARPEGE EDA + singular vectors
AROME-France	Large France domain, 3D-VAR assimilation forecast+51h (00, 03, 06, 09, 12, 15, 18, 21 UTC)	1.3km L90	Coupled to ARPEGE at the boundaries
AROME-IFS	forecast+51h (00, 06, 12, 18 UTC)	1.3km L90	Coupled to IFS at the lateral boundaries, Atmospheric analysis from IFS

AROME EDA	Same domain as AROME	3.25km L90 (same as AROME-France)	Coupled to ARPEGE EDA at the boundaries
AROME EPS	Same domain as AROME, 16 perturbed members, forecast +51h (03, 09, 15, 21 UTC)	Same as AROME-France	Coupled to ARPEGE EPS at the boundaries Initial conditions given by AROME analysis + perturbations from AROME EDA
AROME-NWC	Same domain as AROME, 3D-VAR assimilation, short cutoff, forecast+6h every 1hour,	Same as AROME-France	Coupled to ARPEGE at the boundaries
AROME overseas	5 domains in the tropics: Indian Ocean, Polynesia, New Caledonia, French Caribbean, French Guyana	1.3km L90	Coupled to IFS at the lateral boundaries, Atmospheric analysis from IFS, Surface analysis from ARPEGE

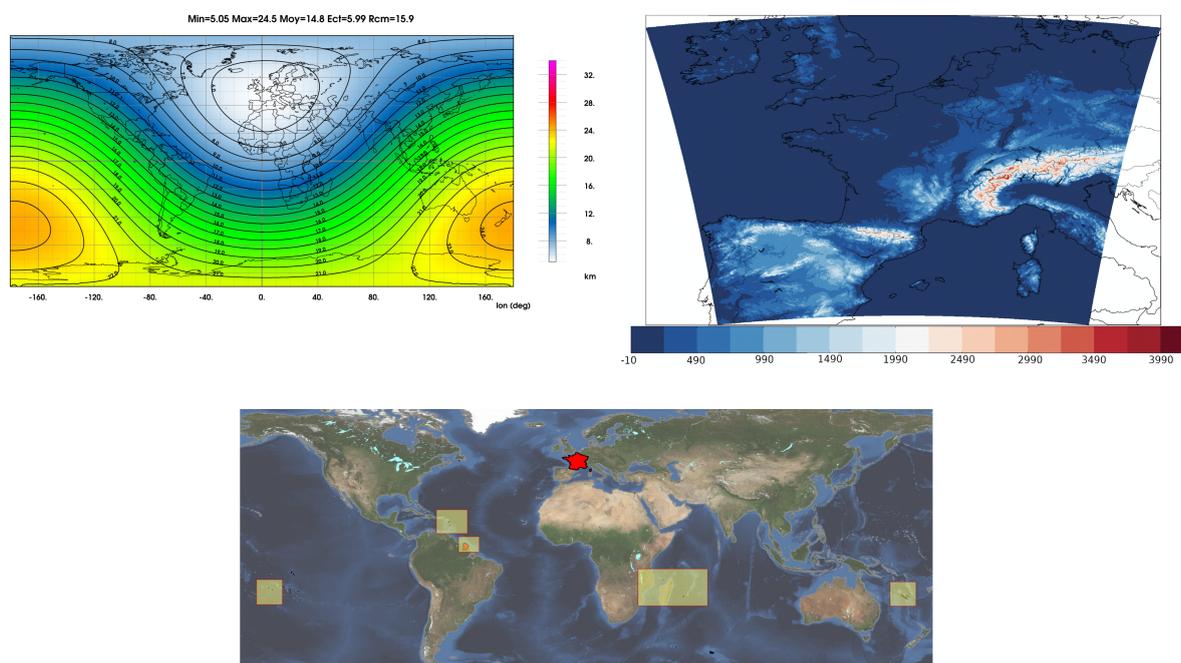


Figure 1: Equivalent resolution (in km) of ARPEGE T1798C2.2 configuration (upper-left panel), AROME-France domain at 1.3km resolution (upper-right panel) with the orography (m), and the 5 of AROME-overseas domains (yellow boxes in the bottom panel).

2. ARPEGE global system

1. ARPEGE physics

Regarding ARPEGE physics, four fundamental changes have been implemented in the 46t1 e-suite:

- The deep convection scheme that has been used since the origins of ARPEGE, albeit with some evolutions, has been replaced by the Tiedtke convection scheme, from the ECMWF IFS model. The resulting improvement in the quality of the model in the intertropical zone is very important.
- An evolution of the flux parameterisation at the ocean interface developed at CNRM (ECUME). A new technique for adjusting the coefficients to the observations of the measurement campaigns at sea allows a much better representation of these flows. The resulting improvement in ARPEGE is significant at all latitudes.
- The replacement of the radiation code used for the visible part of the spectrum by the RRTM scheme, which enables the McIca solver to be activated, which calculates more accurately the overlapping effects of the different cloud layers. This change, and especially the resulting model settings, is a first step towards the use of the EcRad (Hogan and Bozzo, 2018) modular radiation code shared by several NWP models.
- The activation of the 1D-version of the GELATO sea-ice model, integrated into SURFEX. The sea-ice surface temperature is now a prognostic variable, instead of being derived from a climatology and remaining constant during a forecast. Surface temperature variations are now much more realistic, both spatially and temporally.

These changes of the ARPEGE physics apply to every component of the chain where forecasts are run: in the 4D-VAR, in the ARPEGE EDA, and in the 102 hours forecasts. The only exception is half of the ARPEGE EPS perturbed members, which use a different convection scheme (see Section *ARPEGE ensemble prediction* below).

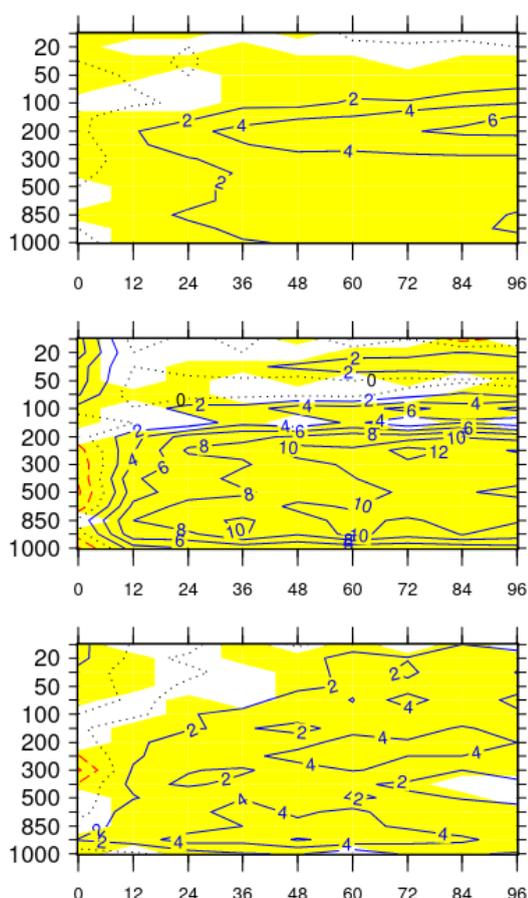


Figure 2: Normalised scores of ARPEGE deterministic forecasts (blue lines, in % of improvement of the 46t1 e-suite compared to the operational chain) of wind speed compared to radiosondes, north to 20°N (upper panel), in the intertropical zone (middle panel) and in south the 20°S (bottom panel). The x-axis is the forecast lead time, and the y-axis is the isobaric height (in hPa). The few red lines indicate a degradation of the 46t1 e-suite compared to the operational chain. The differences are significant in the yellow zones. The scores have been computed from July 2021 to January 2022.

2. ARPEGE assimilation

The ARPEGE 4D-VAR assimilation has benefited from many important changes, either from the development of the capacity to assimilate new kinds of observations, from a better use of existing observations, and from improvements in the assimilation algorithms. The list of major changes in the 46t1 e-suite compared to the operational ARPEGE 4D-VAR may be summarised as:

- the capacity to assimilate microwave radiances in precipitating conditions, using a Bayesian restitution of 1D profile and then 4D-VAR; this procedure is applied to 5 MHS/ATMS instruments,
- the implementation of an analysis of snow height by the assimilation of surface in-situ measurements,
- activation of MESCAN structure function of background error correlations in the surface analysis,
- a revision of horizontal screening of satellite observations, that allows for better spatially distributed observations,
- a new tuning of the error variances from AEOLUS spaceborne lidar winds and of scatterometer winds,
- assimilation of the data from falling radiosondes at 15 sites in Europe,
- the implementation of a 2D observation operator for the assimilation of GNSS radio-occultation data,
- assimilation of the data from the satellites that were not assimilated before: microwave radiances from MWHS-2/FY-3D, infrared radiances from GOES-16, winds from HY-2B and HY-2C scatterometers,

- an update of the list of ground-based in-situ stations, buoys, radiosondes and of GNSS reception.

These changes have been implemented also in ARPEGE-EDA. In addition, two changes were introduced in the EDA:

- individual inflation factors are computed and applied specifically for each ensemble member (instead of a single set of factors which used to be the same for all members). Inflation allows effects of model errors, which are accumulated during each forecast step of the data assimilation cycling, to be represented in the ensemble, leading to more realistic features in the ensemble spread and in the spatial structures of perturbations. The individual inflation allows the dynamics of each member's perturbation to be better adjusted during the cycling. This change was found to have positive impacts on both deterministic and ensemble forecasts.
- the wavelet formulation of background error covariances is now hybrid in a scale-dependent way : in small scale bands, a larger weight is given to flow-dependent ensemble covariances compared to climatological covariances, whereas it is the reverse in large scale bands. This replaces the previous hybrid formulation that relied on a vertical-dependent combination of flow-dependent perturbations (having a larger weight in the troposphere) and climatological perturbations (having a larger weight in the stratosphere), prior to the computation of wavelet covariances. The scale-dependent hybrid formulation allows vertical correlations to be made more robust and more continuous throughout the different atmospheric layers, and the new wavelet covariance computation is also faster in elapse time.

3. ARPEGE ensemble prediction

For the first time since its implementation, the vertical and horizontal resolution of ARPEGE-EPS are the same as the deterministic ARPEGE forecast system. This means that the ARPEGE-EPS global ensemble forecast will have an accuracy of 5 km over France and 24.5 km in the antipodes (Figure 1), and a vertical discretization of 105 levels from 10m to 0.1 hPa. The perturbed forecasts of ARPEGE-EPS are thus fully representative of the uncertainty of the forecasting process.

The physical modelling of the processes in the ensemble forecasting system has also been deeply revised. Half of the forecasts use the new ARPEGE deep convection scheme (Tiedtke-Bechtold, 1989), that offers a significant improvement for the global numerical forecast process. The remaining forecasts will be based on an already used convection scheme, PCMT (Piriou et al, 2007). In addition, the modelling error representation evolves: the set of ten physical packages used so far will be replaced by a more comprehensive approach using stochastic perturbations of several parameters of the parameterisation schemes. Each forecast will therefore randomly use a slightly different setting, sampling the possible values of the physical parameters of the model.

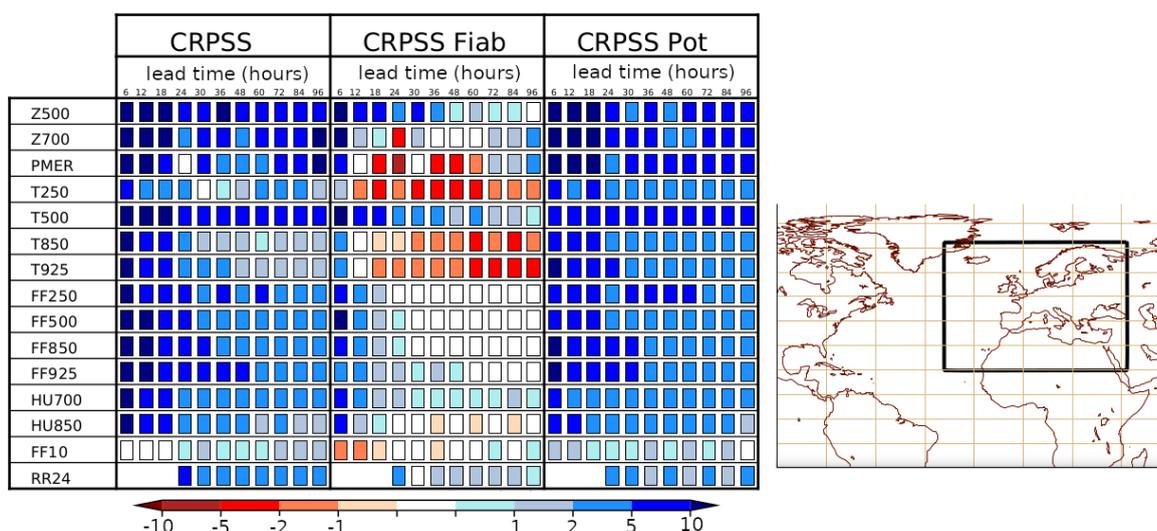


Figure 3: Continuous Ranked Probability Skill Scores (CRPSS, left array) and their Hersbach (2000) decomposition (reliability “CRPSS Fiab” in the middle array + resolution “CRPSS Pot in the right array) of the ARPEGE-EPS “46t1 e-suite” compared to the operational version of ARPEGE-EPS (left panel), on the Europe-Atlantic domain (black box in the right panel), for different variables (rows) and at different lead times. The CRPSS skill scores are expressed in % (blue colours for improvement of the e-suite, red colours for degradation). The reference is surface observations for FF10 and RR24, and the ARPEGE analysis for the other parameters.

A ten-months evaluation of the new ARPEGE-EPS configuration shows a clear and large improvement of the probabilistic scores compared to the operational ARPEGE-EPS version. Figure 3 illustrates such behaviour, for several weather variables over a Europe-Atlantic domain. The CRPSS of the new ARPEGE-EPS version has a clear and significant improvement for all parameters. These very positive results are also valid over the other regions of the globe. The Hersbach (2010) decomposition of the CRPSS into reliability and resolution components underlines that the probabilistic resolution has largely increased. The deterioration of the reliability component of the CRPSS for some variables (in particular T250 hPa or T850 hPa) may be due to a degradation of the bias combined with a decrease, more or less important, of the dispersion of the ensemble, which is consistent with the new method for ARPEGE-EPS model perturbations.

Since the ARPEGE-EPS members are now fully consistent with the ARPEGE unperturbed forecast, the production line of the e-suite has been made consistent (Table 1): 4 times a day, the ARPEGE unperturbed forecast is now embedded in the ARPEGE-EPS, which makes an ensemble of 34 perturbed members and an unperturbed member.

4. New ARPEGE user-oriented diagnostics

In order to serve the final users of meteorological forecasts, some diagnostics are developed out of the ARPEGE forecast members. These diagnostics are computed on the model grid, gaining full benefit of the prognostic variables and model physics at high resolution. Then they are interpolated on a latitude-longitude grid to match the users’ requirements. Such diagnostics generally serve the domains of general forecasting, aeronautics, energy, road transport, among others.

In the ARPEGE 46t1 e-suite, some new diagnostics have been developed:

- CIN, a new formulation of CAPE and lightning density,
- cloud heights (for aeronautics),
- clear-air turbulence (for aeronautics).

Besides, many existing diagnostics in the deterministic ARPEGE forecasts have been extended to the ARPEGE-EPS members, and probabilistic outputs have been developed.

3. AROME regional system over France

1. AROME physics

Forecasters have reported for some years that AROME underestimated precipitation in summer daytime convection situations without synoptic forcing. Studies were carried out in 2018 showing the sensitivity of these situations to the numerical diffusion settings applied to the AROME hydrometeors prognostic variables. Following some complementary work using idealised simulations, a promising solution was found in the AROME transport scheme, and more precisely in the choice of the method used to interpolate the model fields in this process. Semi-academic tests have indeed shown that the transport scheme was responsible for an artificial creation of mass on the cloud fields (by nature much less 'smooth' than a wind or temperature field, or even quite binary in the case of small convective cells). The choice of mass conservative interpolators in the semi-lagrangian advection scheme allowed the model to be re-tuned by replacing the semi-lagrangian numerical diffusion applied to the cloud fields by new semi-lagrangian weight calculations (Malardel and Ricard, 2015) as already done for specific moisture. The tests showed that the model was greatly improved in the problematic situations reported by the forecasters (Figure 4), without being degraded in the general case.

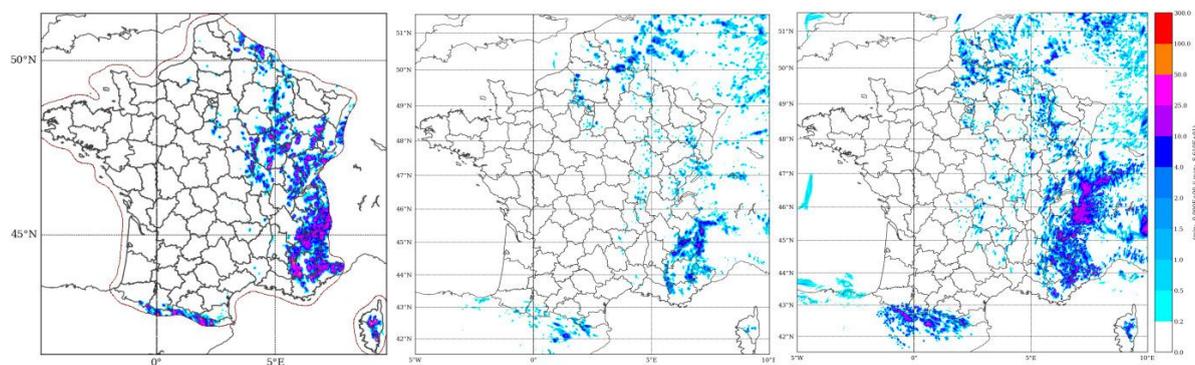


Figure 4: AROME accumulated precipitation forecasts during 24-hours on the 25th July 2018, for the operational version (middle panel), the 46t1 e-suite version (right panel), compared to an observed reference (left panel, radar estimated accumulated rainfall).

In addition to this new setting, the new version of AROME benefits, like ARPEGE, from the new version of the ECUME ocean flow parameterisation.

These changes of the AROME physics apply to every component of the chain where forecasts are run: in the 3D-VAR, in the AROME EDA, in the 51 hours AROME-EPS forecasts (for the unperturbed AROME forecast and for the 16 perturbed members), in AROME-NWC and in AROME-overseas.

2. AROME assimilation

The AROME 3D-VAR assimilation has benefited from many important changes, either from the development of the capacity to assimilate new kinds of observations, from a better use of existing observations, and from improvements in the assimilation algorithms. The list of major changes in the 46t1 e-suite compared to the operational AROME 3D-VAR may be summarised as (some are common with ARPEGE):

- the use of a new radar product, called “SERVAL”, for assimilation, and a consistent in-depth revision of the observation operator, of the quality control and of the horizontal and vertical screening of observations,
- a revision of horizontal screening of satellite observations, that allows to keep more observations in the assimilation,
- assimilation of GNSS radio-occultation data, using a 2D observation operator,
- assimilation of scatterometer winds at a higher horizontal resolution and addition of the data from ASCAT-C, which overall provides 4 times more data than in the operational AROME chain,
- assimilation of the data from falling radiosondes at 15 sites in Europe,
- an update of the list of ground-based in-situ stations, buoys, radiosondes and of GNSS reception.

These changes have been implemented also in AROME-EDA. In addition in the EDA:

- implementation of a surface snow height analysis with stochastic perturbations,
- assimilation of the OPERA European radars and of new satellite data (all these were already assimilated in the AROME 3D-VAR),
- a fundamental change in the representation of model error: activation of a SPPT scheme (same as AROME-EPS) instead of inflation.

3. AROME ensemble prediction

For the first time since its implementation, the horizontal resolution of AROME-EPS is the same as the deterministic AROME forecast system, reaching 1.3km. The vertical resolution remains identical for AROME and AROME-EPS. The perturbed forecasts of AROME-EPS are thus fully representative of the uncertainty of the forecasting process.

AROME-EPS benefits from the upgrades of the other NWP systems, with improved initial conditions from AROME 3D-VAR and AROME-EDA, a new AROME physics package and better ARPEGE-EPS boundary conditions.

The performances of this high-resolution AROME-EPS have been evaluated over several months and show significant improvements, in particular for 10-metre wind and precipitation forecasts (Figure 5). The analysis of several cases of heavy precipitating events also indicates a more accurate prediction of rainfall intensity.

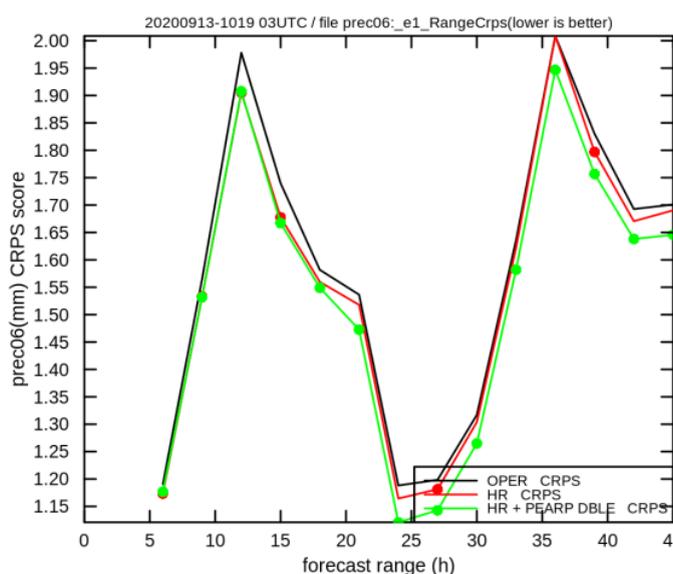


Figure 5: CRPS of AROME-EPS 6-hours accumulated precipitation forecasts as a function of forecast lead time: operational version (black line), high-resolution AROME-EPS (red line) and high-resolution AROME-EPS coupled with the new ARPEGE-EPS (green line).

Since the AROME-EPS members are now fully consistent with the AROME unperturbed forecast, the production line of the e-suite has been made consistent (Table 1): 4 times a day, the AROME unperturbed forecast is now embedded in the AROME-EPS, which makes a consistent ensemble of 16 perturbed members and an unperturbed member.

4. New AROME user-oriented diagnostics

As for ARPEGE, in order to serve the final users of meteorological forecasts, some diagnostics are developed out of the AROME forecast members. In the AROME 46t1 e-suite, some new diagnostics for convection have been developed: CIN, a new formulation of CAPE and lightning density. Figure 6 illustrates the lightning diagnostic for ARPEGE and AROME.

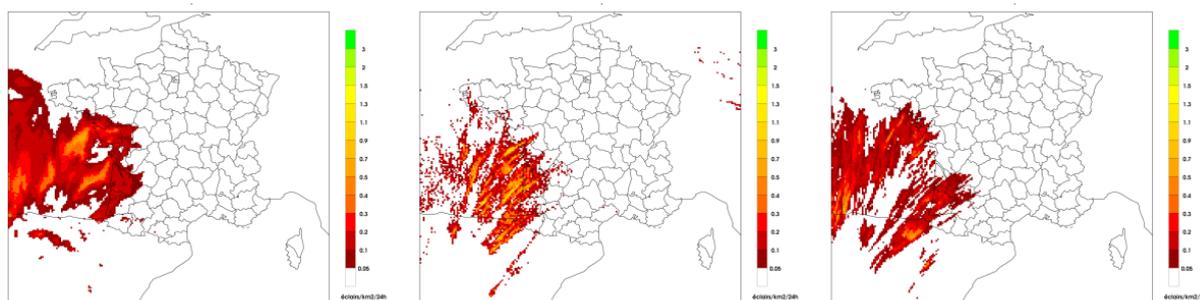


Figure 6: Lightning valid 5 May 2020 at 00UTC, from an ARPEGE +24h forecast (left panel), the Meteorage observations (middle panel) and an AROME +24h forecast (right panel).

4. AROME regional system overseas

Besides the improvements in AROME physics, the AROME overseas in 46t1 e-suite benefit also from an increase of resolution up to 1.3 km. This is particularly an improvement for the representation of rainfall accumulation over tropical islands, as illustrated in Figure 7.

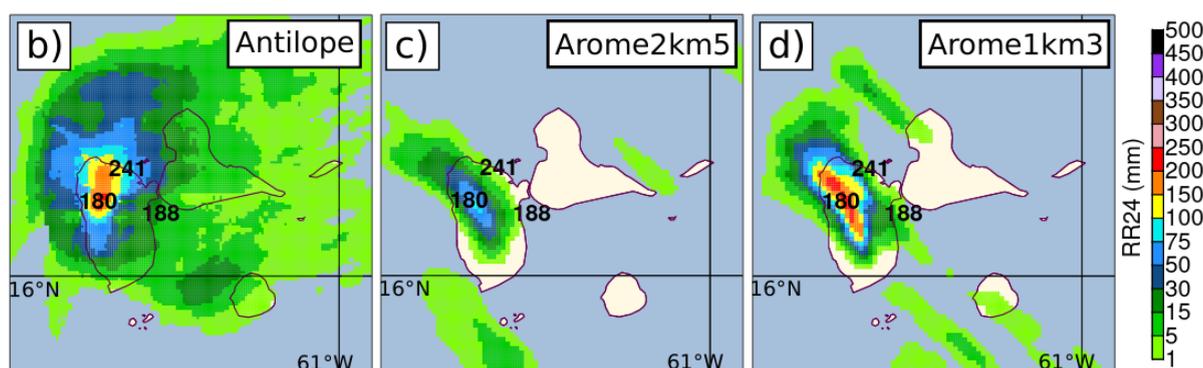


Figure 7: Spatial distribution of surface precipitation over Guadeloupe for the 24h-period from 03 February 2021 06 UTC to 04 February 2021 06 UTC, represented for (b) Antilope, (c) Arome operational at 2.5km and (d) Arome 46t1 e-suite at 1.3km. The numbers stand for some significant rain gauge observations for three locations Sainte-Rose, Petit-Bourg and Pointe-Noire, respectively.

The AROME overseas in 46t1 e-suite are computed with single precision (32b) instead of double precision (64b), which saves some computing cost (about 40%) without degrading the forecast performance. The preparation of single precision computations required extensive tests and evaluation of the model, which could be finalised for this e-suite.

5. Conclusion and perspectives

Allowed by the results of intense research and development efforts and by the increase of computing power, the 46t1 e-suite offers an improved NWP chain at the global and regional scales. It also marks the first year of a “full ensemble” prediction chain.

By the time we are writing this article, this chain is running as an e-suite and its operational implementation is foreseen by mid-2022.

We are also preparing a next e-suite, whose main novelties are expected to be the implementation of a 3DnVAR in AROME-France (instead of the present 3D-VAR) using OOPS, and the implementation of AROME ensemble prediction systems overseas.

6. References

Bechtold, P., Koehler, M., Jung, T., Leutbecher, M., Rodwell, M., Vitart, F. and Balsamo, G. (2008). Advances in predicting atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Q. J. R. Meteorol. Soc.*, 134, 1337-1351.

Bechtold, P., N. Semane, P. Lopez, J.-P. Chaboureau, A. Beljaars, and N. Bormann (2014). Representing equilibrium and non-equilibrium convection in large-scale models. *J. Atmos. Sci.*, 134, 1337-1351.

Hersbach, H., (2000). Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, *Wea. Forecasting*, 15(5), pp 559-570.

Hogan, R. J., & Bozzo, A. (2018). A flexible and efficient radiation scheme for the ECMWF model. *Journal of Advances in Modelling Earth Systems*, 10, 1990– 2008. <https://doi.org/10.1029/2018MS001364>

Malardel, Sylvie & Ricard, Didier. (2014). An alternative cell-averaged departure point reconstruction for pointwise semi-Lagrangian transport schemes. *Quarterly Journal of the Royal Meteorological Society*. 141. 10.1002/qj.2509.

Piriou, J.-M., J.-L. Redelsperger, J.-F. Geleyn, J.-P. Lafore and F. Guichard, An Approach for Convective Parameterization with Memory (2007). Separating Microphysics and Transport in Grid-Scale Equations, *J. Atmos. Sci.*, pp. 4127- 4139.

Tiedtke, M., (1989). A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models, *Mon. Wea. Rev.*

The road to operational CY43T2 in Austria

Christoph Wittmann, Florian Meier, Stefan Schneider, Clemens Wastl and Florian Weidle

1 Introduction

In December 2021, the AROME based forecast systems AROME-Aut, C-LAEF and AROME-RUC operated at ZAMG were finally switched from cy40t1 to cy43t2. Although the initial implementation of cy43t2 and first tests took already place in spring 2020, the period until its operational implementation was unusual long compared to previous upgrades. A series of experimental suites and hindcasts tests had to be implemented before a final setup was found that could outperform the cy40t1 based versions. The reason for this rather long phase was a combination of exaggerated ambitions to include new elements (e.g. B-matrix update, orography) into the upgrade in combination with some cycle specific problems occurring during e-suite runs. The following sections therefore briefly summarize the major problems (section 3 and 4) and developments (section 2) on top of the cy43t2 export version. Section 5 in addition presents an overview about a series of test runs performed with SURFEX using different options and physiographic input data sets.

Table 1: Operational model systems operated by ZAMG

	AROME-Aut	C-LAEF	AROME-RUC
Model version	cy43t2bf11	cy43t2bf11	cy43t2bf11
Resolution	2.5km	2.5km	1.2km
Members	1	16 + 1	1
Levels (lowest/highest)	90 (5m / 35km)	90 (5m / 35km)	90 (5m / 35km)
Time step	60s	60s	30s
Area (grid points)	Alpine area (600x432)	Alpine area (600x432)	Austrian area (900x576)
Orography / physiography	GMTED2010 ECOCLIMAP 1	GMTED2010 ECOCLIMAP 1	GMTED2010 ECOCLIMAP 1
Initial conditions	3DVAR / OI	Ens 3DVAR / Ens OI	3DVAR / OI
LBC model	ECMWF HRES	ECMWF ENS	AROME-Aut
LBC update	1h	1h	1h
Surface scheme	SURFEX 8.0	SURFEX 8.0	SURFEX 8.0
Starting times	00, 03, ... 21 UTC	00, 03, ... 21 UTC	00, 01, ..., 22, 23 UTC
Cycle interval	3 hours	3 hours	1 hour
Forecast range	60 hours	60 hours / 48 hours	12/25 hours
Hardware	HPE Apollo 8600 (ZAMG)	Cray XC40 (ECMWF)	HPE Apollo 8600 (ZAMG)

2 Local developments on top of CY43T2_bf11

Several local developments were integrated into the system upgrades on top of the cy43t2bf11 export version. Some of them were already implemented in the local cy40t1 version, some of them were coded explicitly for cy43t2. The following list contains short descriptions together with a note whether and when it should be available in a common t-cycle:

- **2m diagnostics:** To improve the screening level temperature forecast performance additional options were coded into SURFEX for the L_{CANOPY}=T. case (N_{2M}=2, N_{MTG}=1,...5) and the L_{CANOPY}=F. case (N_{2M}=3 or 4). The modifications have already been reported in the ACCORD Newsletter #1 (Meier et al, 2021). The options N_{2M}=3 and N_{2M}=4 were prepared as a contribution for CY48T2.
- **Sublimation tuning for graupel and snow:** During wintertime upslope snowfall events, an exaggerated luv/lee gradient can be observed in the AROME forecasts. It results in underestimated snow amounts in lee valleys, in particular during events with strong upper air flows directed towards the mountains. The sublimation of falling graupel and snow was identified as one of the main reasons for that and this fact led to the implementation of namelist-tunable factors to influence the sublimation process in the rain_ice routine. The modification (L_{REDS}PG in NAMPARAR) was prepared as a contribution for CY48T2.
- **Lightning diagnostics:** The parameter “SURFDIAGFLASH” was added as a CFU field (activated via L_{FLASH}=T.) for AROME to diagnose the lightning density per square km. The method is taken from McCaul (2009) and uses the graupel flux near the -15 degree Celsius isotherm in a cloud as a proxy for the occurrence of lightning. This type of lightning diagnostics is available starting from CY47.
- **Extended stratus diagnostics:** Parameters indicating the occurrence of low stratus situations in combination with a) the potential for the building of anthropogenic snow and b) “LVP - low visibility procedure” situations around airports (=combination of thresholds for low visibility and ceiling) were asked from several users. Thus, “index” typed parameters for stratus, anthropogenic snow potential and LVP were implemented as pseudo-historic variables, i.e. counting the occurrence of certain atmospheric conditions at every time step between two output steps.
- **Updraft helicity:** A new variable “UPDRAFT_HELICITY” was added as a fullpos field. It is a measure that is often used for the detection of supercells and rotating updrafts. The method implemented follows Kain et al. 2008. Updraft helicity is calculated as an integral of the w component of wind and the horizontal vorticity over a layer defined via NAMFPC parameters. The new parameter was prepared as a contribution for CY48T2.
- **Wind farm parametrisation:** A version of the Fitch et al. 2012 and Volker et al. 2015 respectively wind farm parametrisations was included. They provide additional tendencies for U, V and in case of Fitch et al. also TKE to AROME physics if activated which can improve mainly hub height winds, but also 10m wind and precipitation. The offshore turbine types implemented into HARMONIE by KNMI were also integrated on top of several dozens of onshore types. The scheme works each timestep during integration and requires two additional input files, which require a once per domain pre-processing step with a python routine linked to epygram. So far only input for Central European area based on several sources can be provided. The parametrisation was prepared as a contribution to C48T2.
- **Latent Heat Nudging:** The latent heat nudging scheme was locally coded for cy43t2 and is currently used in the AROME-RUC nowcasting version to nudge 5min rainfall analyses and forecasts from

the INCA nowcasting scheme, but could be also extended to other gridded rainfall products like OPERA composite. It adds temperature and moisture tendencies to AROME physics based on the difference in observed and modelled 5min precipitation during integration. An interpolation of the observations to a model file is required as pre-processing step. The scheme was prepared as a contribution for CY48T1.

- FDDA-Nudging of surface stations: A scheme to assimilate on top of 3D-Var 10min observations of T2m, RH2m and 10m wind into AROME-RUC by nudging was locally coded firstly into cy40t1 and also put to cy43t2. Observations are interpolated to the model grid and weighted according to distance and height difference following the approach of Liu et al. 2006. To fit the grid point distribution within AROME physics the relaxation scheme was split into two steps. Tests show reduced bias of the assimilated observations during the first hour, but more testing and tuning is needed.
- Flexible threshold for Doppler wind first guess check: As the quality of the national radar networks is varying over Europe regarding sensor types and Nyquist velocities, the rejection limit for screening which was constant before was made flexible via namelist switch keeping the default value as it is already available for other observed parameters RBGQC in NAMCOSJO. For Austria a significant error inflation and also application of lower rejection limit and VARQC is used. The modification is available from CY48T1 on in the common code.
- Surface perturbations: For the convection permitting ensemble system C-LAEF an adaptation of the Météo-France surface perturbation scheme (Bouttier et al., 2016) has been implemented. It contains a separate perturbation technique for constant and prognostic surface fields. Seasonal/constant fields (vegetation index, vegetation heat coefficient, leaf area index, land albedo, land roughness length) are taken from the unperturbed control run and then perturbed with different seeds in each member. Prognostic fields (soil moisture, soil temperature, snow depth, sea surface fluxes) are taken from the surface analysis (CANARI) and are then perturbed with different seeds in each member which means that those parameters are cycled in each member.
- Stochastic physics: The model perturbation scheme in the operational C-LAEF version is a combination of a tendency (shallow convection, microphysics, radiation) and parameter (turbulence) perturbation scheme (Wastl et al., 2021).
- Modified cloud diagnostics: The partial and total cloudiness was found (by forecasters) to be too binary, i.e. an exaggerated u-shape of modelled cloud covers compared to observations can be observed. This is true in particular for high clouds. Thus, an adaption of the cloud diagnostics was therefore included in acnebpart.

3 Global radiation in connection with the Ororad scheme

The orographic radiation code has been redesigned for cy43t2 by colleagues from Météo-France. In contrast to the original implementation (cy40t1) the necessary geometrical parameters (shadowing, slope, sky view) are now calculated on the basis of the model topography and no more on external topographical datasets (DEM / digital elevation model) as it was the case for cy40t1. But the advantage that everything can be computed directly within the model (parameters are included into the PGD file) comes along with a lower precision of the geometrical parameters in cy43t2 because of the coarser model resolution compared to a highly resolved DEM. The impact on the orographic shadowing

(mountains are producing shadows in the valley) is very strong during winter months and was found problematic during the evaluation of cy43t2 in comparison to cy40t1.

To reproduce the results of cy40t1 as its best, several approaches to generate more realistic “shadows” in the valleys during wintertime have been tested. The approach finally chosen for cy43t2 is the calculation of the geometric shadow (which is done by searching of higher points in the surroundings) on the basis of the maximum topography and not on the mean orography (=default version coded in surfex/SURFEX/ horizon_rog.F90). By doing so, the sunshine duration is reduced significantly in the narrow Alpine valleys (see Figure 1). This adaptation needs to be done with care and the sunshine duration in such areas has to be verified with station measurements in order to not overestimate the orographic shadowing. However, in future with higher model resolutions this problem will be reduced.

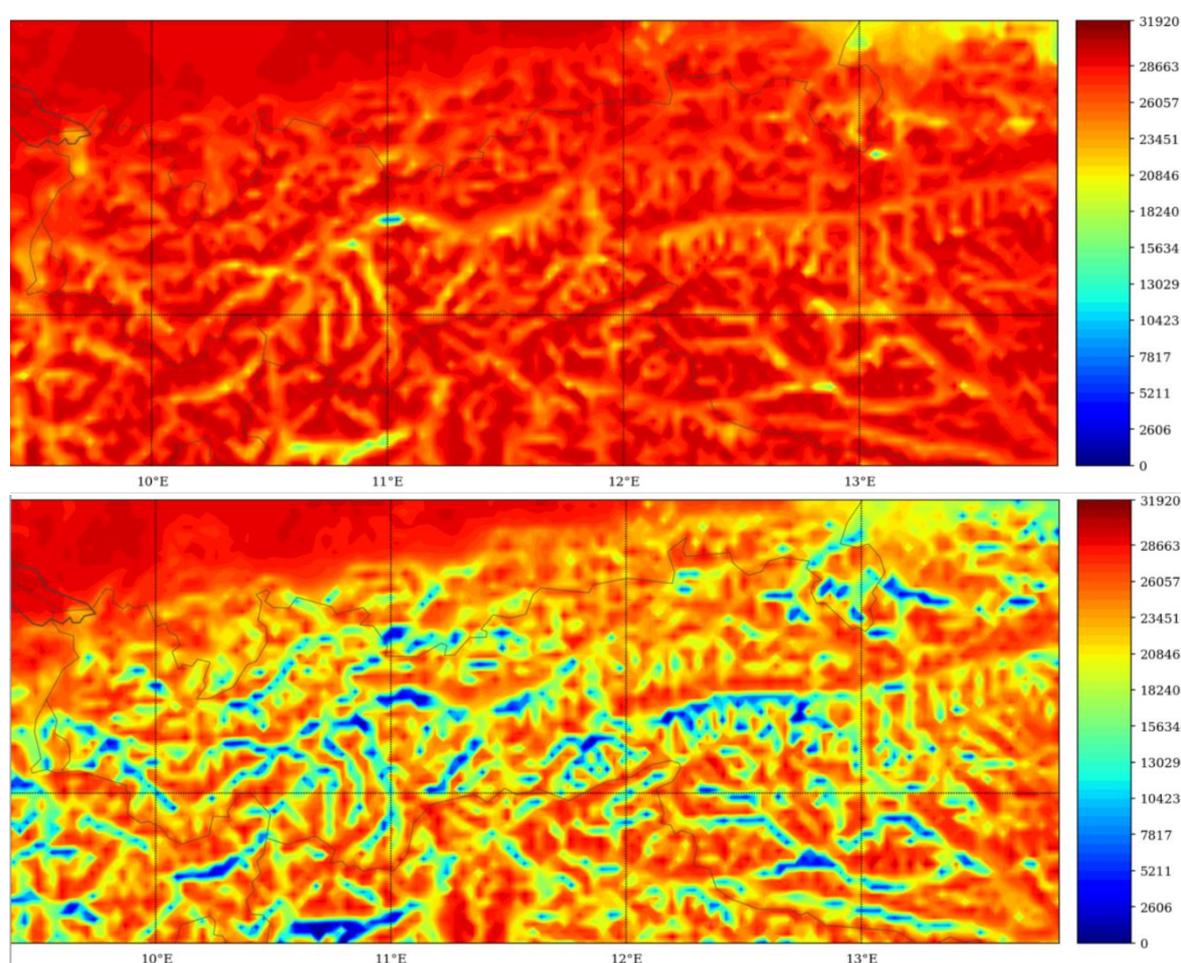


Figure 1: Daily sunshine duration ([s] since start of the model) in the western parts of Austria (Alps) for a test case in January 2022. The upper figure shows the orographic shadowing impact based on the mean orography, the lower figure based on the maximum orography.

4 Problems related to OI/MAIN and 3DVAR

In summer 2020 an esuite of AROME-Aut using cy43t2 was implemented at ZAMG. The namelists were adapted in a way to reproduce operational cy40t1 results as close as possible. The model orography was updated from a rather coarse GTOPO to higher resolved GMTED (ZSFILTER=1). Accordingly also the B-Matrix was updated with a climatologic C-LAEF EDA version. In autumn 2020 cases with

very low dew point temperatures in the first forecast hours were observed at some stations located in Alpine Valleys. Further investigation showed that 3D-Var introduced a shallow very dry layer close to the surface in certain weather situations. A number of experiments have been carried out leading to the conclusion that the chosen value for REDNMC in the minimization namelist was responsible to cause the assimilation to remove lot of humidity close to the surface. The REDNMC parameter was finally reduced from 1.2 to 0.5 for cy43t2.

This change removed a large part of the dry bias (relative humidity) close to the surface. We believe that the change in orography towards a more detailed topography (from filtered GTOPO2 to GMTED with filter option 1) and the different structure of the new B-Matrix, which was implemented in cy43t2, caused this issue by giving too much weight to stations with a relative high vertical displacement to model orography. This caveat of a more detailed orography was accounted for by reducing the weight of observations in 3D-VAR.

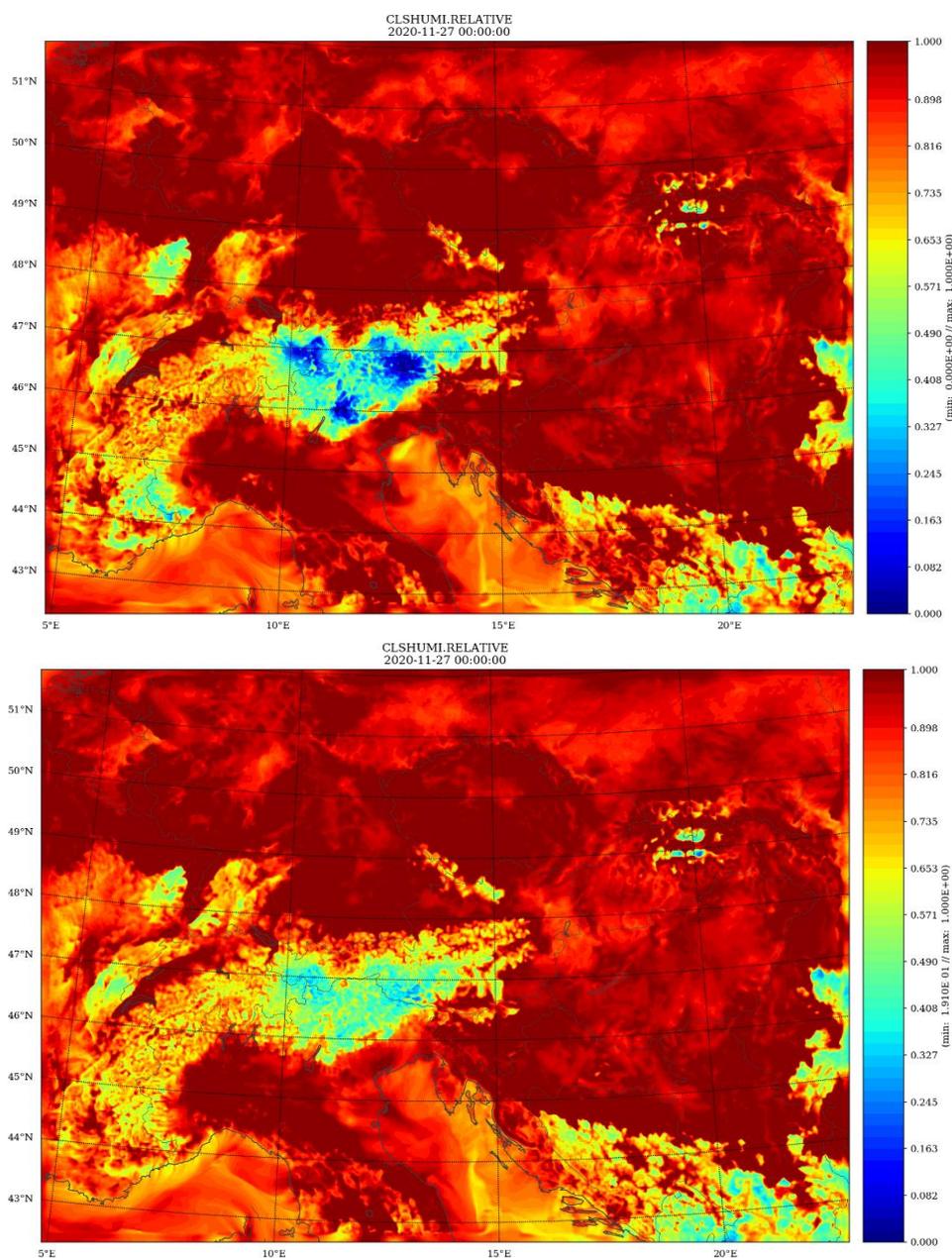


Figure 2: CLSHUMI.RELATIVE at initial time by using $rednmc=1.2$ (top) and $rednmc=0.5$ (bottom). Areas with relative humidity close to zero over the Alps are removed by reducing $rednmc$.

In the beginning we had also a very strong dry bias that turned out to be simply caused by a wrong naming for the OIMAIN namelist. Initially, we kept the naming “OPTIONS.nam” as it was used in the cy40t1 versions, but it turned out that the naming should have been changed to EXSEG1.nam for cy43t2. This trivial mistake led to the situation that the namelist was actually ignored and by default OIMAIN tries to read an ASCAT soil moisture file not provided. In consequence the model does not crash, but causes wrongly negative WG1 and WG2 soil moisture increments (see Figure 3). These led the model moisture drift. After renaming the namelist correctly the soil moisture increments got reasonable. Furthermore a bug in SST for points without nature tile but land sea mask indicating some land (misfit of SURFEX tiles and land sea mask) was fixed with the help from Météo France surface assimilation team. The SST is relaxed towards T2m values in that case now. In the cy43t2bf version SST would stay constant at those points leading after long time to quite wrong values and instability in the model’s dynamics. In our case the Venice area with complex coast line and parts of Dalmatian Coast were especially affected. In Météo France operational version and later cycles this bug is already fixed.

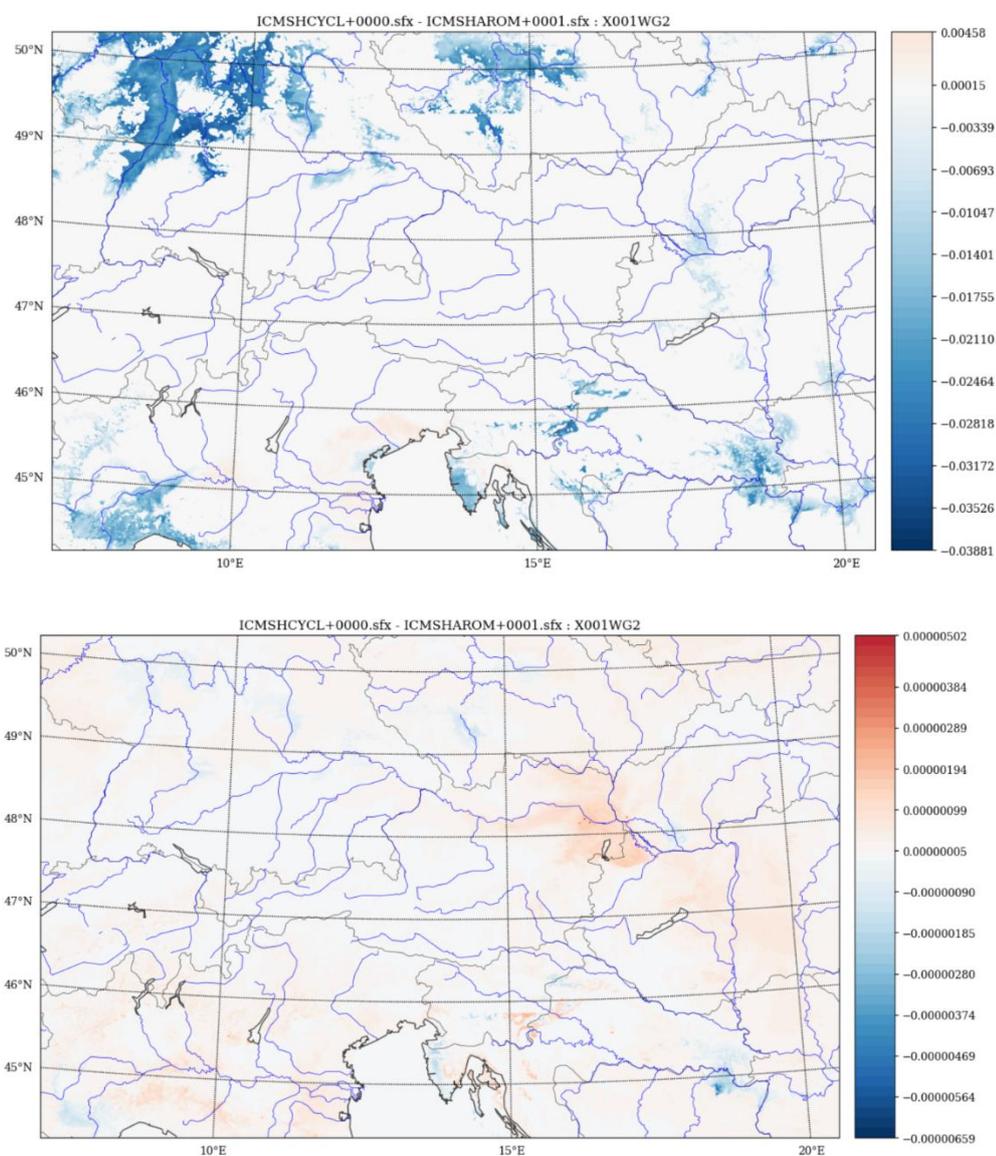


Figure 3: Wrongly negative WG2 increments in case of wrong OIMAIN namelist name (top) leading to significant model drying in cy43t2 vs corrected version (bottom).

5 SURFEX Tests

Until now, ZAMG uses ECOCLIMAP_I_GLOBAL_V1.6 with 1 patch and climatological LAI in its operational AROME configuration. During the switch to the new cycle, it was investigated if other land cover and LAI options would be preferable. For this purpose, three different land cover data sets, 3 different LAI sources, two patch options (either 1 or 12) and the canopy scheme have been investigated in several combinations (see table 2).

The land cover data sets are ECOCLIMAP_I_GLOBAL_V1.6, ECOCLIMAP_II_EUROP_V2.5 and a combination of the Austrian LISA data set (<https://www.landinformationssystem.at/#/lisa/overview>) and Urban Atlas (<https://land.copernicus.eu/local/urban-atlas>). For the latter, the PGD climate file is modified so the distribution of the covers fits to LISA/UrbanAtlas.

The LAI is either described by the climatological values provided in SURFEX, prognostic (CPHOTO=NIT) or taken from Copernicus Global Land (<https://land.copernicus.eu/global/products/lai>). The latter is interpolated to the model grid and added to the surface file using the PERTSURF executable.

All experiments are started on February 1st, 2019 at 00UTC. To avoid a cold start, soil temperature and moisture, vegetation and TEB-related temperatures (roof, road, wall) are interpolated from the operational AROME of this day to the initial surface file. No data assimilation is applied. Forecasts are computed up to +24 hours and the 24-hour forecast is used as initial file for the next run, so the soil is cycled throughout the experiments which are computed until June 30th, 2019.

Table 2: Overview of experiments

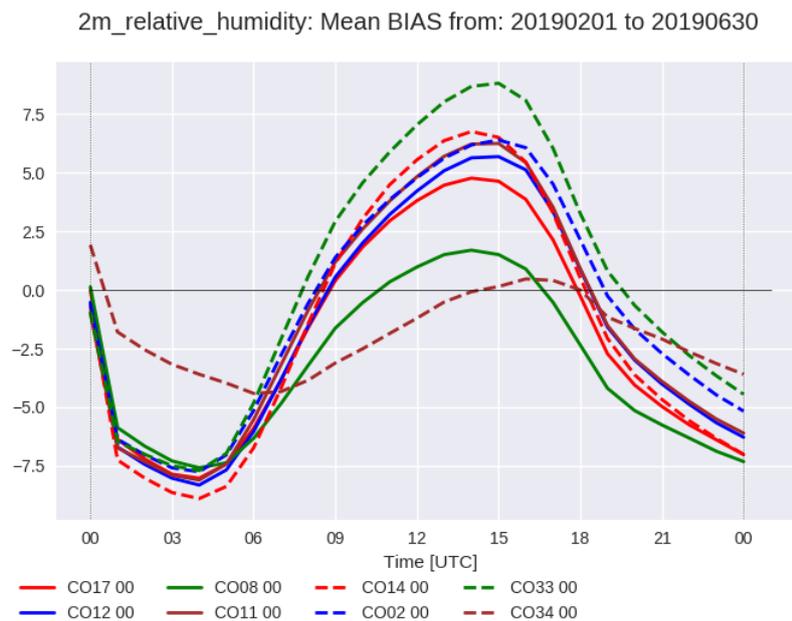
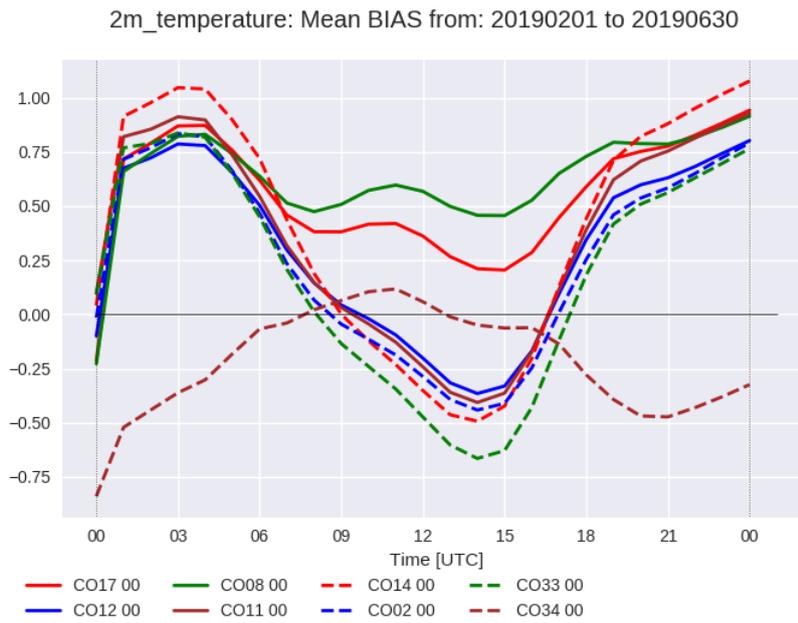
Experiment	Land cover data	LAI	Patches	Canopy	comment
CO12	ECOCLIMAP 1	Climate	1	T	=reference
CO11	ECOCLIMAP 2	Climate	1	T	
CO02	ECOCLIMAP 1	Copernicus	1	T	
CO14	LISA + UrbanAtlas	Climate	1	T	
CO33	ECOCLIMAP 2	Copernicus	1	T	
CO17	ECOCLIMAP 1	Climate	12	T	
CO08	ECOCLIMAP 1	Prognostic	12	T	
CO03	ECOCLIMAP 2	Climate	12	T	ZSFILTER=1
CO06	ECOCLIMAP 2	Climate	12	T	
CO15	ECOCLIMAP 2	Prognostic	12	T	
CO34	ECOCLIMAP 2	Prognostic	12	F	

All experiments use the force-restore soil scheme, the D95 snow scheme, TEB and a 2.5km grid. Forecasts are validated against ~250 stations in Austria and precipitation analyses.

Regarding land cover data, there is no obvious difference between the two ECOCLIMAPs and LISA in T2M forecast performance. The operational setting with ECOCLIMAP_I has the smallest diurnal cycle in the bias (see Figure 4). When comparing climatological versus Copernicus LAI, the latter results on average in colder forecasts especially during daytime. Depending on the stations used for comparison, both features can be beneficial. But none of the configurations is significantly better than others on

average and none is optimal for all regions within the model domain so with one patch, the configuration chosen so far is used also for CY43.

Investigating 12 patches (with the main goal to compute prognostic LAI with CPHOTO=NIT) shows a clear impact. Forecasted T2M is higher, which is beneficial to reduce the cold bias during daytime, but several configurations are far too warm, resulting in a bias of more than +8K at 12UTC (not shown). This is due to an issue with the canopy scheme which is not working properly with 12 patches. Experiment CO34 without the canopy scheme does not show this problem. This experiment has the smallest diurnal cycle in bias and is a strong indicator that the interpolation scheme to compute T2M has a stronger impact on forecast quality than the surface description in the model. This feature will be investigated in more detail in the future.



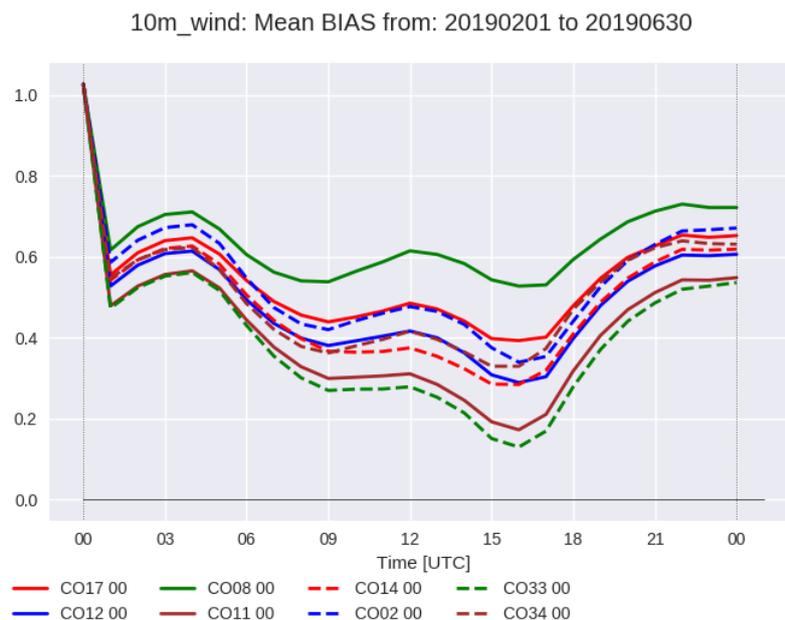


Figure 4: Bias of the 2m air temperature, relative humidity and 10m windspeed averaged for all stations in Austria and all forecasts from February to June 2019.

6 Outlook

The current operational model systems at ZAMG (AROME-Aut, C-LAEF and AROME-RUC) are independent systems, each relying on its own assimilation scheme for atmosphere and surface. This is done for some good reasons (independence, redundancy, etc.) but requires significant maintenance efforts. In particular, the separation between the deterministic system (AROME-Aut implemented at ZAMG HPC) and the ensemble system (C-LAEF at ECMWF HPC) which are meanwhile both running on the same resolution and domain makes sense from data security/availability point of view, but is very demanding from a computational point of view. Further, a significant increase of computational resources can be expected in the near future for ZAMG’s operational system (Upgrade ECMWF HPC in 2022, ZAMG HPC in 2023). Based on these conditions, ZAMG follows the plan to 1) upgrade the ensemble system C-LAEF from 2.5km to approx. 1km within the next years, 2) remove duplication by replacing deterministic AROME-Aut by the C-LAEF control run and 3) implement a central and separated surface assimilation scheme that is available to deliver surface initial conditions for the future C-LAEF 1km version and AROME-RUC.

7 References

Bouttier, F., Raynaud, L., Nuissier, O. and Ménétrier, B. (2016). Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX, Quart. J. Roy. Meteor. Soc. 142, 390-403, <https://doi.org/10.1002/qj.2622>

Fitch, A. C., J. B. Olson and J. K. Lundquist (2012): Parametrization of Wind Farms in Climate Models. J. of Climate, 26,17, 6439-6458, <https://doi.org/10.1175/JCLI-D-12-00376.1>

Kain et al. (2008). Some Practical Considerations Regarding Horizontal Resolution in the First Generation of Operational Convection-Allowing NWP, Weather and Forecasting, 23(5), 931-952.

Liu, Y., A. Bourgeois, T. Warner, S. Swerdlin and J. Hacker (2006): Implementation of observation-nudging based on FDDA into WRF for supporting AFEC test operations. 6th WRF conference.

Meier, F. et al (2021). Adapting the screening level diagnostics to improve AROME temperature forecasts in Alpine areas, ACCORD Newsletter #1, available from <http://www.umr-cnrm.fr/accord/IMG/pdf/accord-nl1.pdf>

McCaul, E. W., Jr., Goodman, S. J., LaCasse, K. M., & Cecil, D. J. (2009). Forecasting Lightning Threat Using Cloud-Resolving Model Simulations, *Weather and Forecasting*, 24(3), 709-729.

Volker, P. J. H., J. Badger, A. N. Hahmann and S. Ott (2015): The Explicit Wake Parametrisation V1.0: a wind farm parametrisation in the mesoscale model WRF. *Geosci. Model Dev.*, 8, 11, 3715–3731, <https://doi.org/10.5194/gmd-8-3715-2015>

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

Summary of NWP related activities in Slovenia in 2021

Jure Cedilnik, Benedikt Strajnar, Neva Pristov, Jana Čampa

1. Overview

This contribution describes main activities at the Slovenian Environment Agency (ARSO). First part is devoted to maintenance of various model configurations (the newest is high-resolution NWP suite with hourly data assimilation), preparation of tools for their verification and optimising them at different hardware/software platforms. Research and development work (second part) is mainly done in the field of data assimilation. The impact of radar reflectivity from OPERA in the ALARO model over central Europe was investigated. The net drying effect is observed in the current set-up, results can be improved with the better method for usage of “undetected” observations. First step using SODAR RASS measurements at the location of Krško (SE Slovenia) was done. In the experiment these data were properly prepared and passively assimilated.

2. Operational and development applications

An overview of our (pre-) operational model set-ups at the Slovenian Environment Agency is presented in table 1.

Table 1: Main characteristics of three NWP systems at ARSO

	aosruc04ec	aos01ruc	seemhews
Model code version	cy43t2	cy43t2	cy43t2
Resolution	4.4 km	1.3 km	2.5 km
Levels	87	87	87
Grid points	432 x 432	589 x 589	1429 x 1141
Initial conditions	CANARI,3DVAR	CANARI,3DVAR	CANARI,3DVAR
Initialization	none/SCC	none/SCC	none/SCC
Physics	ALARO	ALARO	ALARO
Dynamics	hydrostatic	NH	NH
Time step	180 s	60 s	90 s
Boundaries	ECMWF HRES	ECMWF HRES	ECMWF HRES
Forecast length	72/36 hours	36 hours	72 hours
Cycle interval	3 hours	1 hour	3 hours
Frequency of output	1 hour	5 minutes for selected fields, otherwise 1 hour	1 hour
Initial times	00, 03, 06, ... UTC	Every hour	00 and 12 UTC
Computing site	ARSO (SGI ICE - forman)	ARSO (SGI ICE - ventus)	cca/ccb@ECMWF

Status	operational	pre-operational	operational (not yet TC)
Observations	SYNOP + AWS, AMDAR/MODE-S MRAR/EHS, AMV, TEMP, SEVIRI, AMSU-A/MHS/IASI, ASCAT/OSCAT, ZTD EGVAP(passive.)		
		& radar reflectivity	& SEEMHEWS surface observations (ongoing)
Cut-off	2h15min	35min	9h15min

The 1.3 km system runs in RUC mode are relatively stable - the model is running round the clock with almost constant load on a separated dedicated part of the cluster (roughly 1/3 of total HPC is used for that). Final validation is still ongoing.

The 2.5 km system (a large south-eastern Europe region) is running at ECMWF on cca/ccb and is expected to be declared TC2 when the migration to Bologna will be finished. Daily model outputs are computed and provided to the SEE-MHEWS-A project members via CIP (common visualisation platform). Continued validation and further optimization are also planned. One important goal would be to extend the observation data set by use of a higher number of local ground observations (such as those exchanged by participating partners and ECMWF, see Wetterhall et al. (2022) for more information) .

Only few modifications have been implemented in the main operational 4.4 km system, one of them is post-processing of a few variables to flight levels (figure 1).

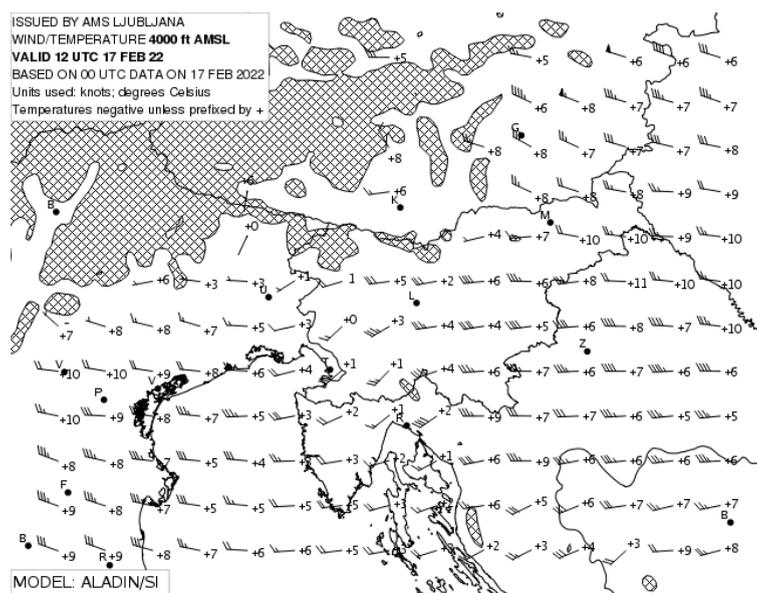


Figure 1: New product for aviation: wind and temperature at flight level high

Tools for validation and verification

The HARP verification package was installed and scripts for deterministic verification were prepared, including preparation of scorecards. Shiny web application is used to browse operational and experiment verification results.

The observation dataset for verification was extended with automatic weather stations from Slovenia and bilaterally-exchanged data from Austria and Italy on hourly scale. These are merged with the European Synop dataset and prepared in observation format readable by the HARP verification package (vobs). This enables more reliable verification of precipitation compared to previously-used Harmonie Monitor verification package.

Our aim is to produce daily verification products, but first the quality of our local implementation has to be confirmed (suspicion of some bugs).

Miscellaneous on porting, migration and hardware issues

The local HPC environment has been upgraded and divided to two separate clusters: one of them being in continuous production mode (round the clock running of the 1.3 km domain) and the other for the remaining operational tasks and research.

Furthermore, the model has been ported to VEGA - an EuroHPC pre-exascale system in Maribor. The full NWP suite migration is being carried out. Once finished it will serve as a backup of our system and as a platform for running experimental suites.

Additionally, initial preparations have started to port the SEEMHEWS system from cca in Reading to the new aa machine in Bologna.

3. Data Assimilation activities

Most of the research and development work is done in data assimilation. Two topics are presented below, others include testing of various methods for dealiasing of radial winds, implementation of the torus mapping method into HOOOF (see separate contribution on HOOOF), successful porting and test of the OOPS executable. Quite some effort was needed to organise an external preprocessing of Slovenian GNSS data, which will be included in the E-GVAP dissemination.

1. Assimilation of radar reflectivity

The impact of radar reflectivity from OPERA in the ALARO model over central Europe was investigated. Earlier experience suggested that the current implementation of Bayesian inversion resulted in the net drying effect. While earlier studies of the effect (stay S. Panežić in Prague) included tuning of the neighbourhood size and prescribed observations error, this effort was focused on use of dry observations and their effect on humidity fields. A major concern is proper use of “undetected” observations, to which a detection threshold, deduced from observed data, is assigned and used to compute Bayesian weights. Two alternative scenarios were evaluated in separate experiments: the first only allows wet-to-wet or wet-to-dry comparisons so comparison only takes place if either the model background or observation value is above a given rain threshold (e.g. 13 dBZ); the second one uses alternative averaging of humidity profiles in the neighbourhood in case drying is needed. Instead of only averaging the fully dry ones, all the profiles below the rain threshold are used in the Bayesian

inversion. As a separate experiment, recent operational modifications by Meteo France (on top of the export version cy43t2, Maud Martet) were evaluated. The impact was studied over August 2020 and for two resolutions used at ARSO (4.4 and 1.3 km).

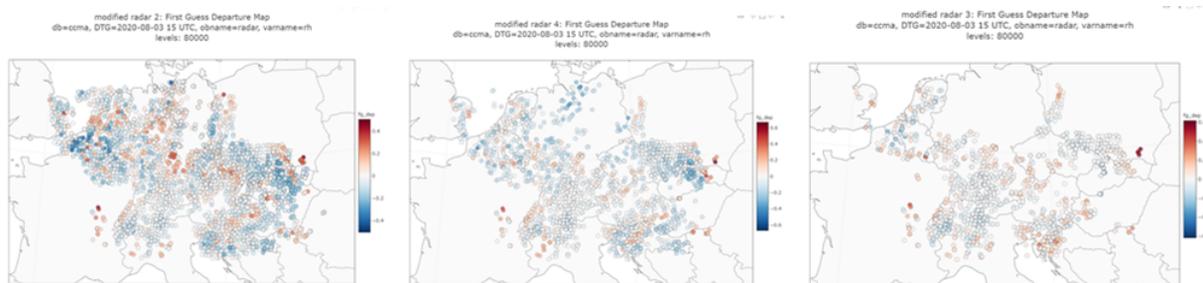


Figure 2: Observation-minus-background departures of relative humidity for assimilated reflectivity points at 800 hPa. Default 4.4 km ALARO setup (left), only wet pixels (middle) and modified averaging during the drying cases (right).

The different setups largely influence (reduce) the number of assimilated observations. While the reduction from “wet-only” run is quite logical, a further reduction in the ”modified drying” experiment is a result of more rejections from the QC, where the same sign of increment in reflectivity and humidity is required. Objective verification was focused on precipitation, Synop reports and automatic weather stations in Slovenia, Austria and Italy were used.

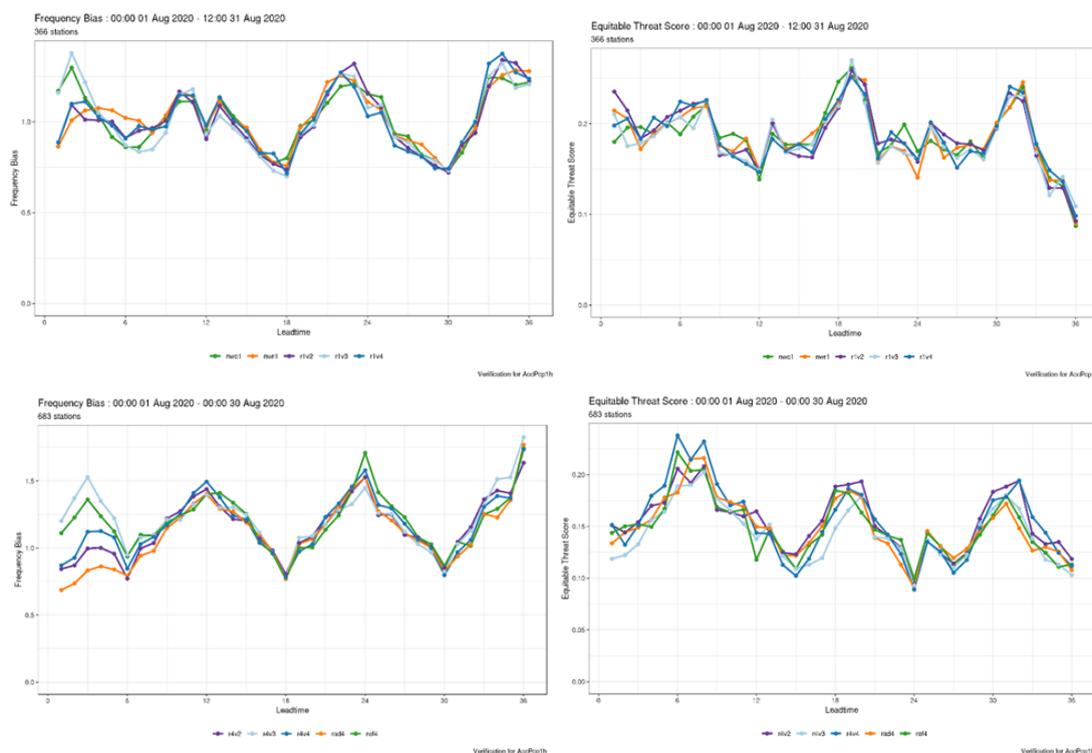


Figure 3: Frequency bias (left) and ETS score (right) for 2 mm threshold of hourly precipitation amount in 1.3 km (top row) and 4.4 km model suite (bottom). The reference without reflectivity DA is in green and default DA setup in orange. The dark blue is the “wet-only” experiment and the light blue is the “modified drying” experiment. The purple line represents operational setup at MF

Looking at several precipitation thresholds and accumulation times and other variables such as cloudiness it is concluded that the “modified MF setting” experiment outperforms the default one and that the “wet only” experiment is improving that one mainly in terms of bias and is expected to be used in the first operational setup. The “modified drying” experiment is the only one that leads to additional moistening with respect to the reference run.

2. Ab-initio studies of SODAR RASS in 3D-Var - observation monitoring

SODAR RASS measurements at the location of Krško (SE Slovenia) were passively assimilated over a one month period in the operational set-up environment at 4.4 km resolution. Another such instrument has been acquired at a location close to Ljubljana and the general goal is to assimilate both. The SODAR RASS performs measurements of the vertical profile of wind and virtual temperature up to a few hundred metres above ground.

As an initial step, the SODAR RASS measurements were coded individually as point aircraft observations and fed into the model. The temperature time series plots of measurements and model are shown on Figure 4, and the first results for OMG statistics are on Figure 5. Notice that the comparison is slightly flawed due to the reason that the instrument actually measures virtual temperature, hence the bias in the OMG statistics on the order of 1-2 K - this would lead to an estimation of mixing ratio being around 10 g/kg which seems completely sound. Other than that, the statistics seem promising for active assimilation. No significant deterioration or systematic bias was observed related to different heights of observation.

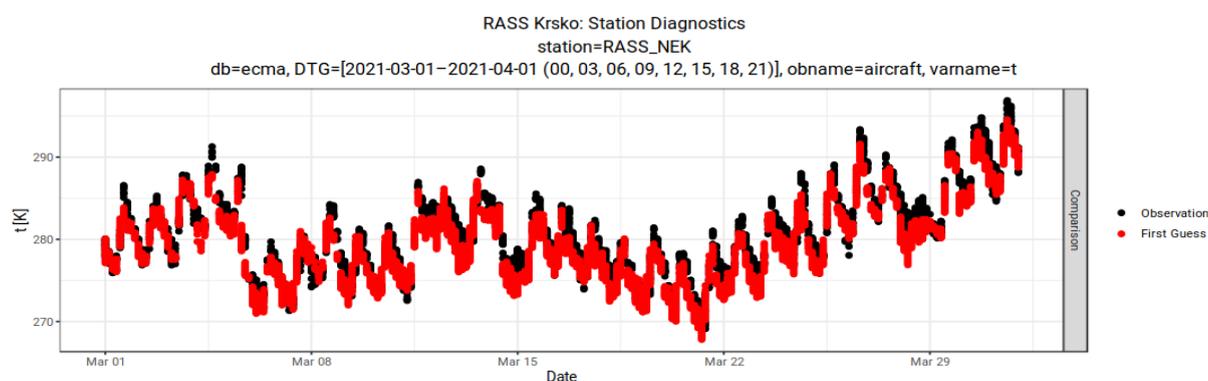


Figure 4: Time-series of temperature for observation [black] and model [red] for all available heights (up to a few hundred metres) for the SODAR RASS measurements for the location of Krško (SE Slovenia). Notice the systematic bias coming from comparison of observed virtual temperature and “ordinary” temperature from the model.

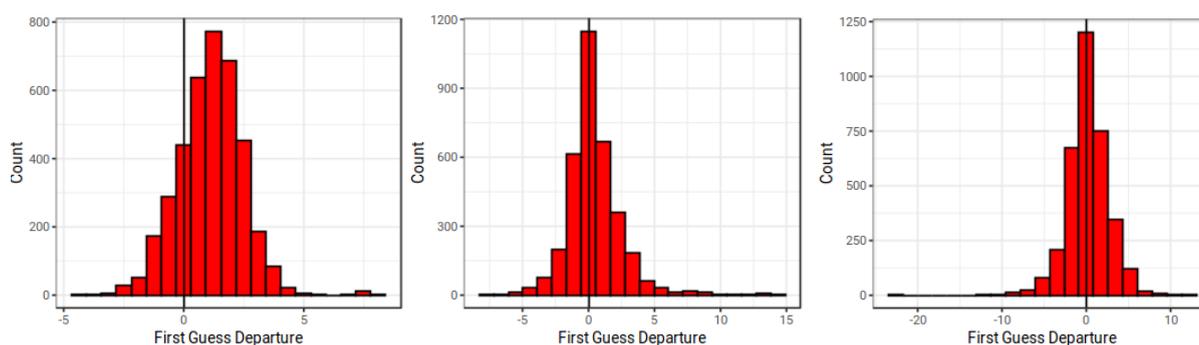


Figure 5: *OMG histograms for the SODAR RASS measurements for all available heights for location of Krško (SE Slovenia) during the one month period: left image is for temperature, centre for u wind and right for v wind. Notice the positive bias in the temperature histogram due to the same reason as explained above.*

4. Outlook

The future NWP development activities at ARSO will focus on operational implementation of the hourly ALARO-RUC NWP system and its further tuning and preparation of various tailored products for users, such as probability from lagged hourly model runs. We plan to further exploit the additional supercomputing resources (computer VEGA, part of Euro-HPC) to back-up operations and conduct experiments, mainly in the DA area (new observations, EnVAR as alternative assimilation algorithm). With these resources, ARSO plans to contribute to the DestinE-Extreme project.

5. References

Fredrik Wetterhall, Umberto Modigliani, Bojan Kasic, 2022: ECMWF’s support for the SEE-MHEWS-A project, Newsletter No. 170 - Winter 2021/22

Recent NWP activities at Météo Algérie : Report on the new HPC and the newcomers training

Walid CHIKHI ⁽¹⁾, Nour El Isslam KERROUMI ⁽¹⁾, Zakaria BENGHABRIT ⁽¹⁾ Mohamed MOKHTARI ⁽¹⁾, Abdenour AMBAR ^(1,2)

⁽¹⁾ Numerical weather prediction department, Office National de la Meteorologie, Algiers, ALGERIA.

⁽²⁾ Faculty of Earth Sciences and Territory Planning, Sciences and Technologies University USTHB, Algiers, ALGERIA.

1 Introduction

Since November 2020, the NWP team of Météo Algérie has been hugely affected by an incident occurred at the computing system (IBM-HPC). For this reason, a kind of minimum service has been provided to our forecasters by keeping only a part of operational suite running for more than a year. Recently, a new HPC platform has been acquired by the IT direction, namely "HPC-Fennec", allowing to restore the entire operational suite, and to resume research activities within the team.

In this paper we'll briefly describe the characteristics of the new HPC and summarize the activities of newcomers within our NWP team as well as some perspectives.

2 HPC-Fennec Description

Meteo-Algeria has just replaced its old HPC "IBM-HPC" following the incident and a loss of more than half of its clusters. The old HPC was equipped with 26 computing nodes and 416 processors allowing it to reach a theoretical computing power of 10Tf. Below are the characteristics of the new HPC acquired :

Table 1 : Characteristics of the new Algerian HPC-Fennec.

HPC-Fennec Characteristics	Operational	Development
Nodes	20	08
Sockets	20 x (Intel 4316 Silver)	08 x (Intel 4214 Silver)
Number of Cores	20	12
Number of AVX-512 FMA Units	02	01
Max Turbo Frequency	3.40 GHz	3.20 GHz
Theoretical performance	87.04 Tflops	9.8 Tflops
Total Theoretical performance	96.84 Tflops	
Real Performance	81.81 Tflops	9.21 Tflops
Total Real Performance	91.02 Tflops	

3 Local training report

The purpose of this training was to familiarize with the different configurations used in Météo-Algerie, from the compilation and installation of the code until the post-processing of the outputs. As a practical example, we have installed the export version of Cycle 43 on the new HPC.

1. Compilation and installation of Cycle 43 :

In order to realize this work, three (03) files have been made available to us by our supervisors which are :

- GMKPACK : Is a procedure to create an environment and to compile binaries , in this training we have used the version 6.6.4 (gmckpack.6.6.4d.tar.gz).
- Code of the Cy43 (local_cy43t2.tar.gz).
- GMKFILE.x : The configuration file, it contains the architecture for Linux platforms with the compiler and different libraries paths used.

- Compiler and prerequisite libraries:

The compilation of the following libraries and model is done with INTEL compilers (icc,ifort). Below are the needed libraries for the compilation of the NWP models :

- lapack, blas , zlib
- auxlibs-installer (rgb, bufr, dummies, eclite, gribex)
- HDF5 , NETCDF et NETCDF-Fortran
- grib_api.

- GMKPACK :

The GMKPACK is a procedure that has been written by GCO and Ryad El Khatib (Météo France) in order to create an environment to compile and make binaries from Arpege, Aladin, ODB, ... or others libraries. In this training we have used the version 6.6.4, below are the steps followed for its compilation :

- We created two directories named “Pack” and “GMKPACK” in our home path.
- We have set the different environment variables in our .bash_profile and sourced it.
- Finally, we compiled our gmckpack by simply running the scripts «*build_gmkpack*».

At the end of the compilation, the following message has been displayed on our terminal :

```
To create new packs with gmckpack you will need to setup a configuration file,
defining compilers, options, etc.
You can setup this file manually or use a an automatic assitant.
Do you want to run the configuration file maker assistant now (y) or later [n] ?
```

You can then create the **gmckfile.x** by answering [y], and you start give your paths librairies, or you can use an old **gmckfile.x** (already created on your environment) by answering [n]. In both cases, your file should be placed here :

```
/home/${user}/GMKPACK/gmckpack_support/arch/
```

- Installation and compilation of CY43T2bf10:

a/ Creation of a new pack 43T2bf10 : In the directory "/home/\${user}/pack/", we run the following command line **to create a new pack:**

```
gmckpack -r 43t2 -b bf10 -a -l INTEL -o x -p ?
```

A new folder "43t2_bf10.01.INTEL.x" is created which contains the following sub-folders :

- "bin": a folder to contain all binaries that will be generated.
- "lib": a folder to contain all the compiled libraries.
- "src": a folder to contain the source code of the model.
- "ics_usr" : *the compilation script for all the binaries.*

b/ Importation of the CY43T2bf10 source code : we have to copy manually and extract the source code from "local_cy43t2.tar.gz" to "43t2_bf10.01.INTEL.x/src/local" folder.

c/ Compilation of CY43T2bf10 :

- First, we compile the source code by setting the keys below and running the script :

```
ICS_POSTPONE_ABORT = yes
ICS_ICFMODE = full
```

-Second, we switch-off the same keys and run again the ics script:

```
ICS_POSTPONE_ABORT = no
ICS_ICFMODE = off
```

A total of 74 binaries have been generated, including "**MASTERODB**". Hereafter the list of all the generated binaries:

ADDSURF	GOBPTOUT	odbmd5sum.x
adjust_seqnos.x	hcat	odbsql.x
ATSTPROG	ICE_GRB	ODBTOOLS
BATOR	INVERSION	OI_MAIN
BLEND	IOASSIGN	PERTSURF
BLENDSUR	LECBDAP	PGD
bufr2odb.x	LFITools	PINUTS
CHECK_LIMITS	MANDALAY	PREGPSSOL
CLUST	MASTER911	PREP
COMBI	MRGVARBC	PROCADRE
CONVERT_ECOCLIMAP	MSE_SURFEX	PROGRID
create_odb.x	split_timeslot_bufr_data.x	QSCAT_25TO50
DATEGRIB	SST_NETCDF	QSCAT_BUFR
dcagen.x	SURFEX	QSCAT_DCONEQC
DDHC	SXPOST	QSCAT_FILTER
MASTERODB	TEST_BIPER	UNHOLO
ETESTADJ	TEST_EZONES	RTTOV_CONV_COEF
FCQODB	TEST_GPCOU	SFXTOOLS
FEDIACOV	TEST_TWOTRUNC	simulobs2odb.x
FESTAT	TESTADJ	SODA
fscheduler.x	split_bufr_data.x	

2. Case study and visualization of outputs :

After having compiled the pack and generated the binaries, the next step was to run a simulation with one of the operational configurations. For this, a preparation of the environment was carried out (scripts adaptation, deployment of different constant and namelists... etc).

In order to manipulate and adapt to different configurations, we have chosen to run a simulation using **AROME** coupled with the operational **ALADIN** (Dynamic adaptation). The output have been post-processed and different parameters have been extracted and potted (Fig01).

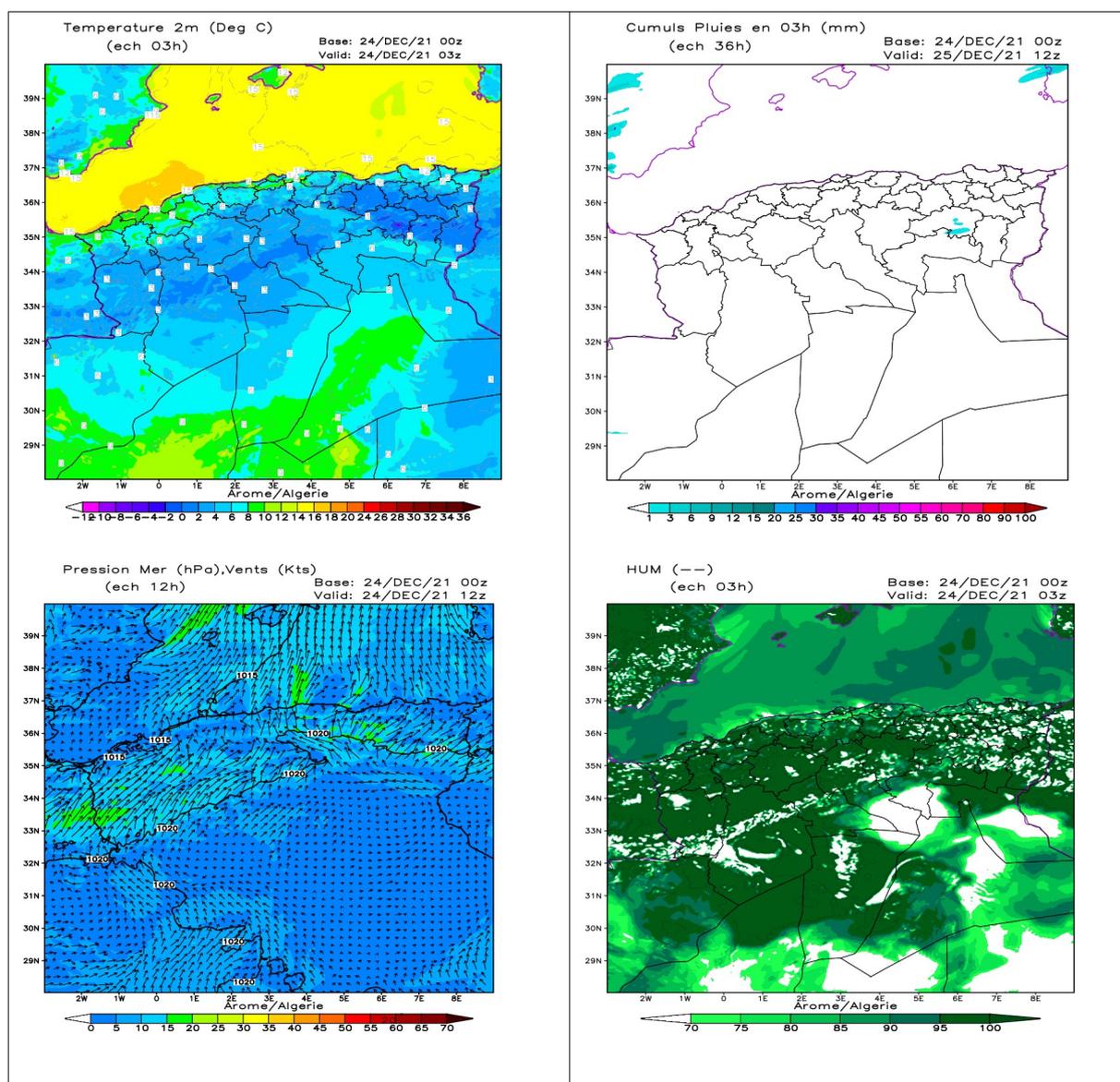


Fig 01 : T2m, Precipitations, MSLP and humidity AROME outputs for December 24th, 2021.

4 Conclusion

This training falls within the framework of the newcomers integration within the NWP team. They have been introduced to all the daily tasks at Météo Algérie and the different research axis among the ACCORD consortium. The main programming basics and tools have been widely experienced by the newcomers : Linux basics, Fortran coding, GMKPACK, models compilation and installation, preparation of clim files, running forecasts , post-processing of outputs, visualization tools, etc ..

5 Perspectives

For the newcomers, the plan is to steer them toward specific topics as follow :

- Mr CHIKHI Walid : will work on diagnostics on AROME and system (he's been already designed as Local Team System Team Representatives) ;
- Mr BENGHARBIT Zakaria : will work on diagnostics on AROME ;
- Mr KERROUMI Nour El Islam : will work on surface processes.

Regarding the computing resources recovery, we have started to bring back into service many tasks that have been interrupted and we have also set some short-term objectives as follows:

- Deploy the control system prepared as part of the twinning project last year on of the new HPC.
- Compilation of the export version of Cy46 and validation of the diagnostics for an operational use (visibility – lightening – Capeshear).
- Restore of the pre-operational assimilation suite that have been interrupted due to the HPC incident.
- Elaboration of new configuration AROME-Assim large domain (3km ; 90 levels), with an extension towards the north in order to take into account more observations available on southern Europe.

NWP activities in Romania

Alina Dumitru, Simona Taşcu, Mirela Pietrişi, Alexandra Crăciun

1. Introduction

In this study, the performance of the ALADIN/ALARO model was analysed for experiments at higher horizontal resolution than the current operational version. In order to prepare the future operational setup, the statistical validation of the results is a necessary step.

2. Experiments

The results presented in this study are obtained from four experiments. All configurations use 60 vertical levels with 21 levels below 850 hPa, but they differ in horizontal resolution or domain size. Exp 1 is based on the current operational configuration, running at 6.5 km resolution, with ALARO-0 baseline physical package. Exp 2 and exp 3 have the same resolution (4 km), but differ in the size of the domain, while exp 4 has finer resolution, 2.5 km. Experiments exp 2, exp 3 and exp 4 are based on the ALARO-1vb physical package. A representation of the model domains is shown in Figure 1 and a summary of the experiments is presented in Table 1. A preliminary preparation for these experiments was introduced in the previous issue of this newsletter [1]. Previously encountered problems when running exp 4 in e001 configuration were solved.

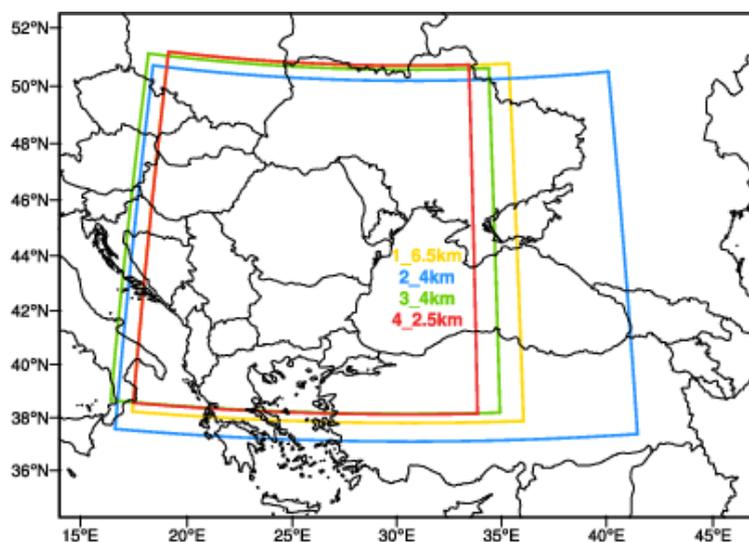


Figure 1: Representation of the integration domain for each experiment.

Table 1: Description of experiments.

Experiment	exp 1	exp 2	exp 3	exp 4
Horizontal resolution	6.5 km	4 km	4 km	2.5 km
Number of points on OX	240	600	450	640
Number of points on OY	240	432	400	640

3. Forecast validation

In order to analyse the model performance, validation scores were computed for the summer season of the year 2021 (June, July and August), for forecasts provided by 00 UTC runs. The forecast range is 78 hours for experiments exp 1, exp 2 and exp 3, while for exp 4, only 30 hours forecasts are provided, because of the computer time required to obtain these finer resolution experiments on the current platform. Observational data are used from 157 national meteorological stations in Romania. The meteorological parameters under evaluation are: 2 m temperature, 10 m wind speed, mean sea level pressure, 2 m relative humidity and 6-hour accumulated precipitation.

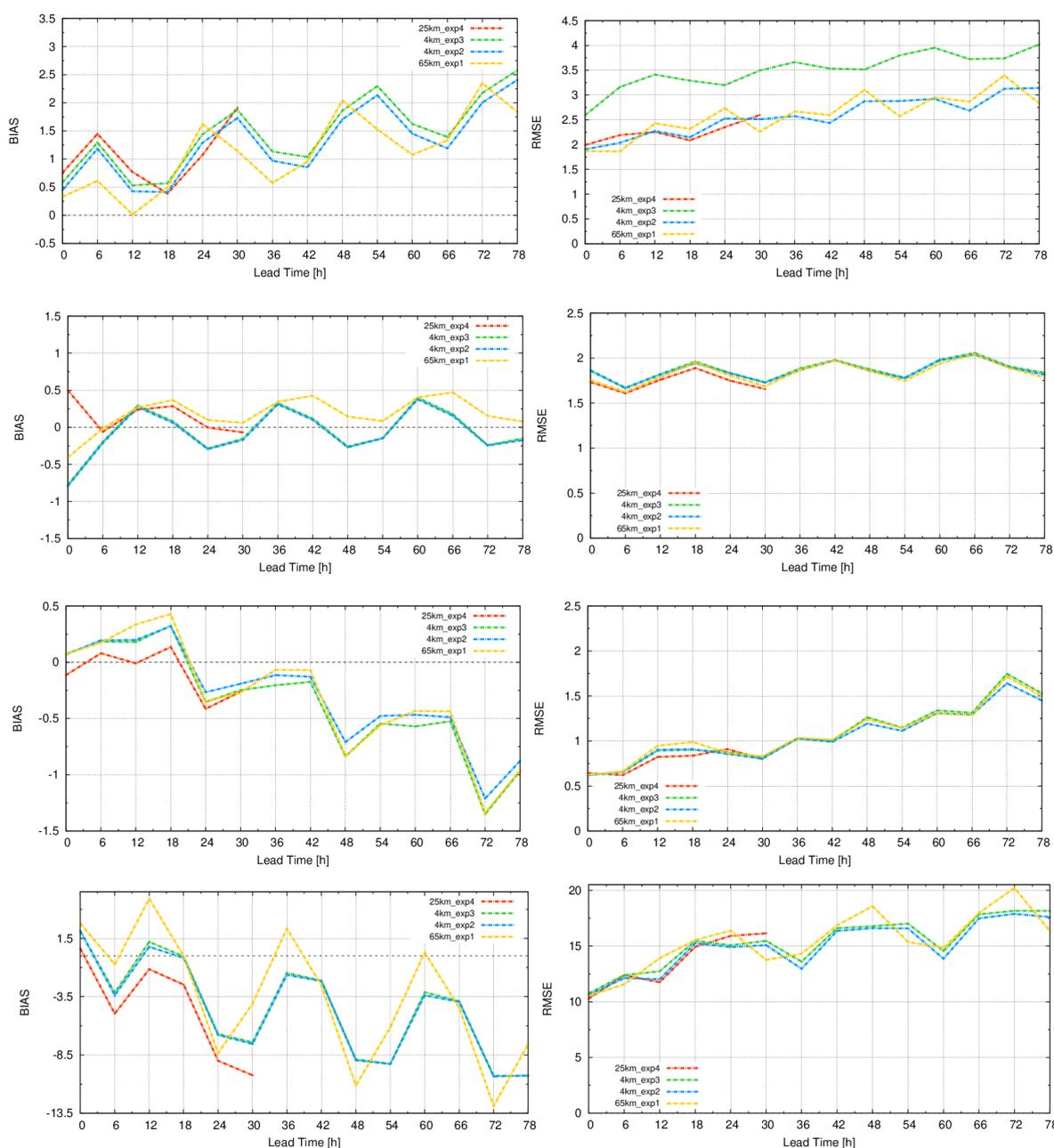


Figure 2: BIAS (left column) and RMSE (right column) for 2 m temperature (first row), 10 m wind speed (second row), MSLP (third row) and 2 m RH (last row); summer 2021.

We can observe different behaviour of the model forecast, depending on each parameter. If we look at the BIAS scores (Figure 2 - left column): for 2 m temperature, it seems the operational version

performs best; the overall tendency of the 6.5 km model to overestimate the wind speed in the second part of the day is replaced in the 4 km experiments with more visible variations between daytime and nighttime patterns of the score. For MSLP, we can see most improvement in the 2.5 km resolution, while for 2 m relative humidity, higher resolution determines a decrease in the peak values of the score.

As it concerns RMSE (Figure 2 - right column), the differences between the experiments are less visible. However, there is an indication that exp 4 introduces a bigger variance of the 2 m temperatures forecast errors. Also, higher values of the score are smoothed for 2 m humidity in higher resolutions.

Precipitation forecast verification

Furthermore, the precipitation forecast was put under closer attention, since it is expected that going to higher resolution will result in better representation of precipitation, which sometimes the current operational version fails to capture, especially for higher quantities that occur in the convective season. From a meteorological point of view, the summer of 2021, for which all the results are shown, was characterised by significant amounts of precipitation.

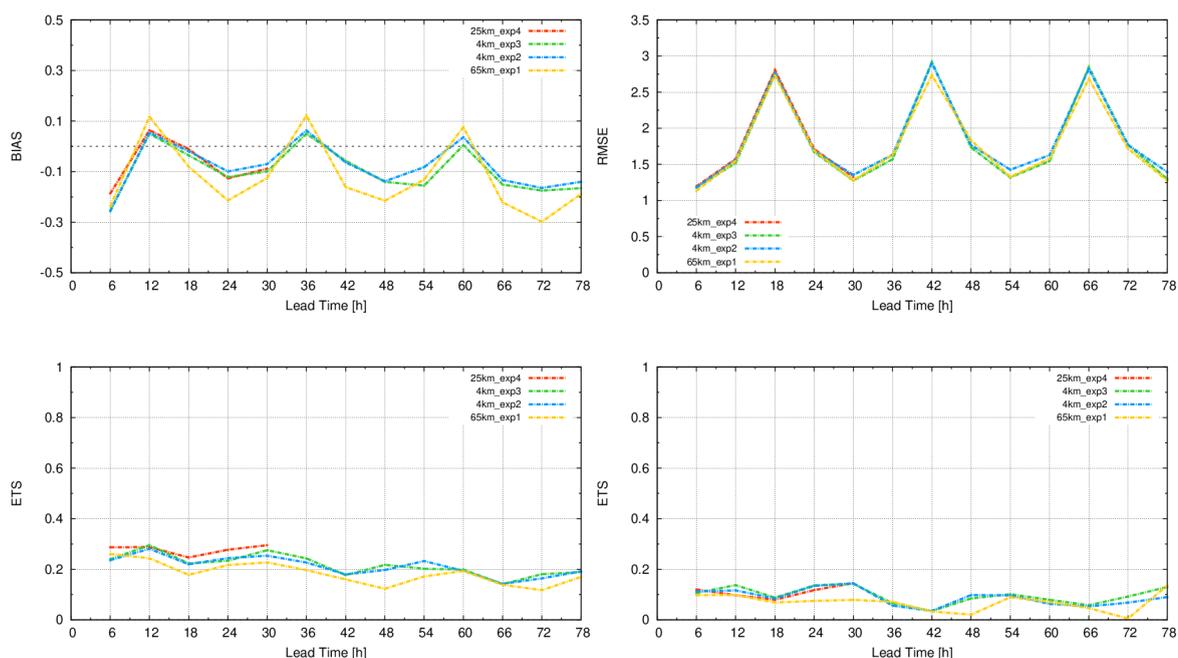


Figure 3: Precipitation forecast scores: upper panel - BIAS (left) and RMSE (right); lower panel - ETS for 0.1 mm (left) and 10 mm (right) thresholds; summer 2021.

BIAS, RMSE and ETS (Equitable Threat Score) scores are presented in Figure 3 for the 6-hour accumulated precipitation forecast. Overall, we can see a reduction of BIAS in higher resolution experiments when compared to the operational version. The maximum values of the BIAS score are smoothed in higher resolution forecasts. In addition to this, there is a slight increase of the errors in the operational version after longer integration, which is not present in higher resolution experiments. Very similar RMSE values were obtained in all experiments.

The ETS score was computed considering several precipitation thresholds. Here, it is shown for two of them: 0.1 mm (on the left) and 10 mm (on the right). We can observe some indication of a slight

improvement of the score (a perfect score being 1) in higher resolutions, especially for exp 4 (2.5 km) computed for the 0.1 mm threshold.

Taking into consideration the complexity of this meteorological parameter, we can sometimes miss important information regarding model performance if we consider only classical verification scores, such as BIAS, RMSE and ETS. For this purpose, the use of standard probabilistic metrics such as the BRIER score was considered.

The BRIER score applied to neighborhoods of forecast grid points have been evaluated using the HIRA (High Resolution Assessment) framework [2]. HIRA, which is included in the MET (Model Evaluation Tools) packages, was developed using surface observation stations and provides the method for evaluating numerical models in the neighborhood of point observations [3]. The BRIER score was computed for the same period (June - August 2021), using four neighborhood sizes (2 x 2, 3 x 3, 4 x 4 and 5 x 5) and different hourly precipitation thresholds (≥ 0.1 , ≥ 0.5 , ≥ 1.0 , ≥ 2.0 , ≥ 4.0 , ≥ 8.0 mm/h). This study considers all 30-h ALARO configurations forecasts initialized at 00 UTC. As computed in MET, the BRIER score is the mean squared probability error, therefore smaller values are better [2].

In Figure 4 the BRIER score is shown for two different neighborhood sizes (2 x 2 and 5 x 5 neighborhood point observation) for the threshold of precipitations higher than 0.1 mm/h. Thus, four comparisons are made. On the left side, 2 x 2 neighborhoods are considered for: exp 1 (~ 13 km) compared to exp 2, exp 3 (~ 8 km) and exp 4 (~ 5 km). On the right side, a 5 x 5 neighborhood: exp. 1 (~ 32.5 km) operational version compared to exp. 2 and exp. 3 (~ 20 km) and exp 4 (~ 12.5 km).

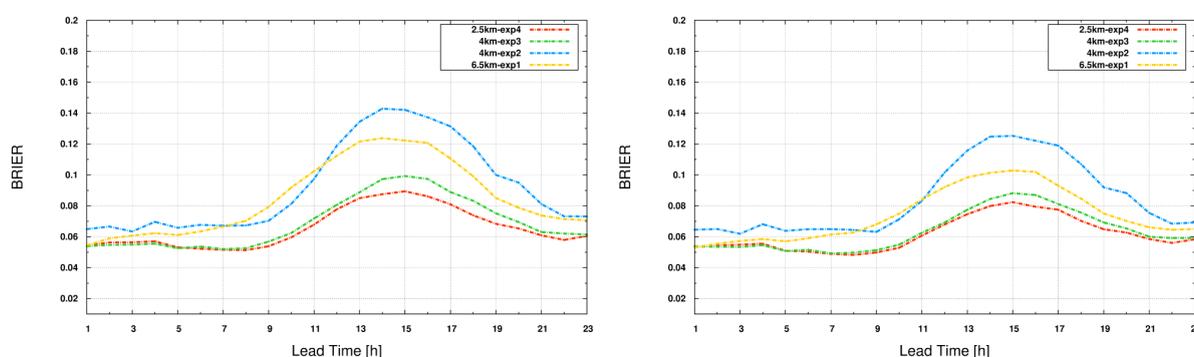


Figure 4: BRIER score for hourly precipitation threshold ≥ 0.1 mm, 2 x 2 (left) and 5 x 5 (right) neighborhood verification observation; summer 2021.

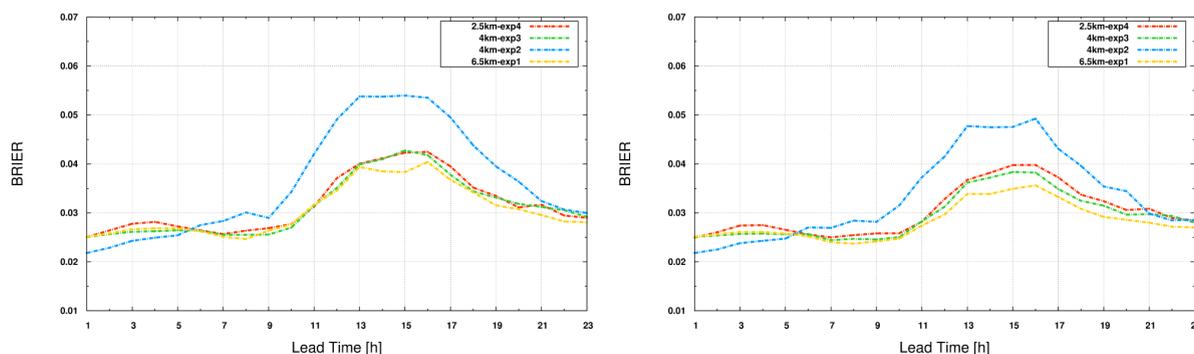


Figure 5: BRIER score for hourly precipitation threshold ≥ 2.0 mm, 2 x 2 (left) and 5 x 5 (right) neighborhood verification observation; summer 2021.

For lower precipitation threshold, it seems that lower values of the score are obtained for exp 3 and exp 4, with a more visible difference increasing during daytime and a peak around 15 UTC. It is visible that there is an increase of the BRIER score values starting at 10 UTC, when usually the convective activity is triggered. Yet, the exp 4 forecast provides a more accurate solution during the 10 UTC to 19 UTC period. The overall pattern for each experiment remains the same when increasing the neighborhood size, but the score values are slightly reduced. Having bigger neighborhood size, it means that more model grid points are included.

When we consider higher hourly precipitation threshold (≥ 2.0 mm), the increase in resolution does not determine significant changes, the values of the score for exp 1, exp 3 and exp 4 being very similar in most forecast ranges.

We can observe in both Figure 4 and 5 that there are visible differences between the exp 2 and exp 3 forecast scores, although they have the same horizontal resolution. An explanation for this result may come from the fact that weather patterns develop differently in each domain, having different boundary conditions [2]. We can see from figure 1 that exp 1, exp 3 and exp 4 have similar domain boundaries and sizes, while the integration domain for exp 2 is extended to the East.

4. Conclusion

An evaluation of forecast performance using statistical scores for surface parameters was done, with closer focus on the precipitation forecast, for a three-month period, during the summer season of 2021. Standard statistical scores showed that for several parameters there is some improvement coming with the higher resolution experiments, while for the others this conclusion is not so visible in the current analysis. An additional element, the forecast neighborhood, was introduced to measure the skill in a model configuration. The general behaviour of the BRIER score shows that for lower precipitation thresholds, exp 4 performs better than the other experiments for all neighborhood sizes. Yet, for higher precipitation thresholds, it seems that the differences between the experiments are very small. The results indicate where we can expect some improvements when upgrading the current operational version to a higher resolution one. Further investigation will continue for specific weather events.

Acknowledgments:

"Model Evaluation Tools (MET) was developed at the National Center for Atmospheric Research (NCAR) through grants from the National Science Foundation (NSF), the National Oceanic and Atmospheric Administration (NOAA), the United States Air Force (USAF), and the United States Department of Energy (DOE). NCAR is sponsored by the United States National Science Foundation."

References

- [1] Dumitru A., S. Taşcu, A. Crăciun, M. Neașu (2021). NWP activities in Romania, ACCORD NL 1, October 2021. <http://www.umr-cnrm.fr/accord/IMG/pdf/accord-nl1.pdf>
- [2] Mittermeier, M. P. (2014). A Strategy for Verifying Near-Convection-Resolving Model Forecasts at Observing Sites, *Weather and Forecasting*, 29(2), 185-204.

[3] Gotway J.H., K. Newman, T. Jensen, Brown, B., R. Bullock, and T. Fowler, 2018: The Model Evaluation Tools v8.0 (METv8.0) User's Guide. Developmental Testbed Center. Available at: http://www.dtcenter.org/met/users/docs/users_guide/MET_Users_Guide_v8.0.pdf. 407 pp.

Recent ACCORD Related Activities at TSMS

Yelis Cengiz, Alper Güser, Mustafa Başaran

1 Introduction

The recent developments and activities on numerical weather prediction at TSMS are summarized in this article. One of these activities is the calculation of B-Matrix for AROME-Turkey and the other one is the adaptation of A-LAEF for Turkish domain.

2 Calculation of B-matrix for AROME-Tr

At TSMS, the application of assimilation into AROME model continues in the frame of DAsKIT program. A 3 hourly surface assimilation cycling has been set up by using AROME model. The features of the 3 hourly rapid update cycle (RUC) were shown in the Table 1. The system defined below is running on eflow in the pre-operational mode.

Table 1: The Features of AROME-Tr RUC Configuration

Model version	cy43t2_bf10
Resolution	1.7 km
Levels	72
Boundaries	ECMWF-IFS
Starting times	00 UTC, 06 UTC, 12 UTC and 18 UTC
Cycle interval	3 hours
Forecast length	+24 hours
Surface assimilation method	CANARI-OI MAIN
Observation pre-processing tool	SAPP
Observations digested by BATOR	Synop observations

To continue the work of assimilation with upper air, the B-Matrix for AROME-Tr was calculated by ensemble downscaling method for the period of 20210811-20210909 on ECMWF cca with the support of ZAMG and SHMU. Firstly, the input data necessary for 903 configuration was retrieved from MARS database by using the script provided by Martin Bellus. 4 members of ensemble forecast were retrieved for 00 UTC and 12 UTC network times. This script creates the files which include surface gridpoint data, upper air gridpoint data and spectral data.

Subsequently, the scripts, which were provided by ZAMG, to obtain LBC data, to calculate forecast differences and to calculate the B-Matrix were run respectively. 903 configuration was run to generate the lbc files for AROME-Tr domain by using the mentioned input data. After that 3 hourly forecasts were run for each member and for each network time and 3 hourly forecast differences were calculated as follows: MEMBER1-MEMBER2 and MEMBER3-MEMBER4.

In total 120 forecast differences were obtained. The diagnostics of B-Matrix were shown in Figure 1.

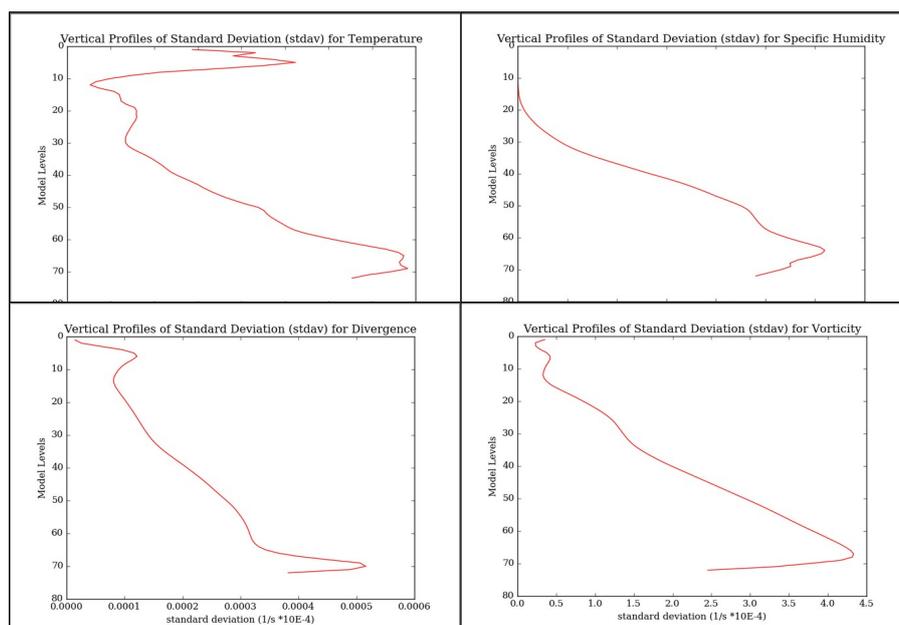


Figure 2: Diagnostics of B-Matrix of AROME-Tr. (Vertical Profiles of Standard Deviation for Temperature on top left, Vertical Profiles of Standard Deviation for Specific Humidity on top right, Vertical Profiles of Standard Deviation for Divergence on bottom left and Vertical Profiles of Standard Deviation for Vorticity on bottom right)

3 A-LAEF Activities at TSMS

Since 1 December 2020, the post-processing data of A-LAEF (Belluš, 2020) has been utilized at TSMS by the kind support of Martin Belluš. 16 members (+ 1 control member) of A-LAEF are disseminated to TSMS. The horizontal resolution of the data, which is received by TSMS, is 4.5 km and the vertical resolution is 60 levels. The A-LAEF outputs are visualized on an interactive web page which was prepared by TSMS. The user of the web page can interactively select the needed parameter such as spaghetti plots for upper air parameters, ensemble mean plots for surface parameters. Besides the user also can set the thresholds for the probability maps. An example of the interactive web page is shown in the Figure 3.

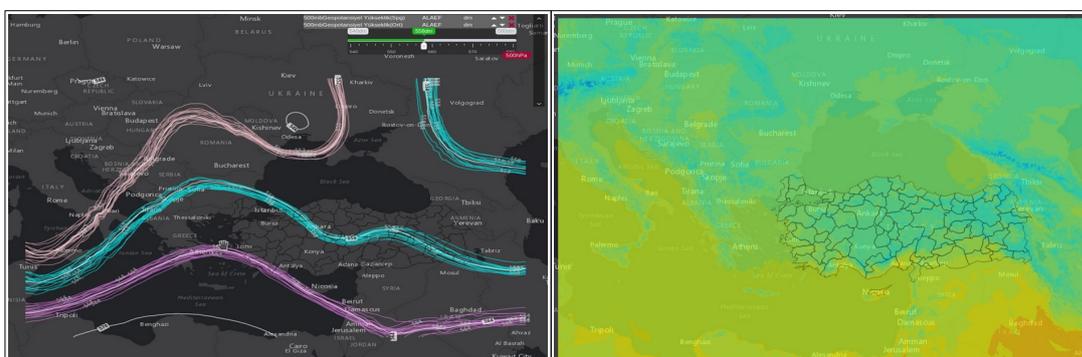


Figure 3: Spaghetti plot of 500 geopotential height with at 500 hPa (on the left) and Ensemble Mean plot of 2 meter temperature (on the right).

4 Acknowledgement

We thank ZAMG colleagues: Clemens Wastl, Florian Meier and SHMU colleague: Martin Bellus for their guidance and technical support. We also would like to thank Maria Monteiro for her support in DAsKIT activities.

5 References

Belluš, M.: A-LAEF ecFlow TC-2 Suite (ECMWF) - update 06/2020. Technical report. ECMWF. 2020.

Summary of ACCORD autumn surface working week

Patrick Samuelsson (ACCORD Surface AL, SMHI)

1. Introduction

When it was time for the first ACCORD Surface Working Week (WW) last October we were still very much affected by the pandemic era, so it became natural to arrange it as a web-based seminar week. After being through a long training period with web-based meetings with many participants, like two All Staff workshops and two EWGLAM meetings, we could proceed quite confident that it should work fine also this time. And it did! On and off some 30 people were connected. Thanks to Patricia we had recording capability through the BlueJeans system. Thus, all presentations are available both as pdf:s and as recordings. Please visit the [WW wiki page](#) for all information.

We divided the week into a number of surface subjects with morning and afternoon sessions spread out over the week. We were presented with a variety of surface subjects and an impressive surface activity in ACCORD. Here follows some background perspective and a few comments per such subject or session.

2. Sessions

1. Urban - Town Energy Balance

Currently, in those operational NWP systems using SURFEX and TEB, a quite basic setup of TEB is applied. Development and studies with TEB shows that more advanced TEB options can be beneficial also for NWP. But while approaching these more advanced descriptions it is also important to understand what characterises different urban and city environments over Europe. The session included studies of atmospheric-urban feedbacks over London, TEB behaviour over Toulouse and HARMONIE-AROME/TEB comparison with urban observations in Amsterdam. For example, we learned that we should critically examine the physiography settings over airport areas (Blagnac example) in our model to avoid potentially big biases. Rafiq also gave a climate perspective on urban areas from his work with the IPCC report.

The discussion included the aspects of difficulties in validating our NWP-urban behaviour due to lack of observations or at least lack of access to observations. Strategies on how to handle this will be further discussed in the Urban thematic group led by Rafiq.

2. Physiography and ECOCLIMAP

Historically we have only been able to relate to quite a limited number of alternatives for our physiography input to our NWP modelling systems. However, in the last few years we have seen a quite substantial increase in different physiography databases, and very much so at European level. In addition the world of Machine Learning opens even more doors for alternatives, combining e.g. physiography maps with satellite information, something we will see more and more. Also our strive towards hectometric model resolutions forces us to find new alternatives to the current global physiography maps we have for SURFEX and ECOCLIMAP. Besides these exciting perspectives, in

the near future, we will still rely heavily on the ECOCLIMAP maps kindly provided by the SURFEX team in Toulouse. However, all physiographic information comes with limitations and a couple of presentations dealt with this as was already indicated in the Urban session.

The session included: a presentation on how the SLIM project delivers algorithms, acquisition or modelling chains to produce global surface land information mapping. A presentation on how high-resolution physiography is applied in SURFEX offline with the aim to achieve more details in near-surface conditions for Vienna. A presentation that explains problems with PGD processing for certain domains. And a presentation on application of Machine Learning in identification of vegetation for land-cover mapping.

The PGD problem related to ECOSG processing was discussed further. Marie explained how some hard-coded urban properties need to be considered when defining domains. Also the problem of undefined parameter values at high-latitudes was discussed. HIRLAM will look more into how to tackle this since they are concretely affected by the problem at the moment.

3. Snow-related development

Snow is a fascinating and challenging aspect of surface modelling. The presence of snow changes drastically the surface characteristics with respect to e.g. insulation of the soil beneath, surface properties with respect to albedo, roughness and heat capacity and a source of water storage. SURFEX provides different levels of complexity in snow modelling where the operational one is still represented by the most simple bulk-layer model. Apart from the processes themselves we also need to correct the simulated snow pack evolution with help of observations in data assimilation. For years we have utilised the SYNOP in-situ snow depth observations but now we also take steps towards remote sensing products.

The session included: two presentations related to application of SURFEX/Crocus for detailed snow modelling in the French and Slovenian Alps regions. The French one with focus on avalanche forecasts and the Slovenian one also including hydrological aspects of snow melt. And two presentations on how satellite products of snow extent are now implemented in the HARMONIE-AROME CSC.

Regarding snow avalanche modelling in an NWP-style framework we can probably conclude that it requires many detailed aspects with respect to resolution, forcing, processes and observations for assimilation. Therefore, operational snow avalanche warning systems are still very much based on more empirical in-situ methods.

4. New development and experiences with multi-layer surface physics

Since 20-30 years we have relied, and are still relying on, Force-Restore style description of the surface in our own NWP community. This, in combination with OI surface assimilation, has served us well but we also know that it has its limitations with respect to more complex surface characteristics, like forests, and when the diurnal cycle is not well defined. Process complexity development, computer resources and observations provide now the opportunity to go towards more physically sound descriptions of the surface. In ACCORD we have the possibility to utilise well developed options for multi-layer processes in SURFEX including the 14-layer diffusion soil scheme, the 12-layer Explicit snow scheme and the MEB processes for explicit canopy.

Three presentations were given where different aspects of these multi-layer combinations were discussed. One of them also included a new development related to roughness sublayer processes over tall vegetation. A fourth presentation included the step towards dynamic treatment of the vegetation through the A-gs processes where it was indicated that the annual cycle of LAI may not always be correctly represented by the A-gs processes.

The multi-layer physics show good potential but as with all steps towards new combinations of operational setups there are aspects which are still not satisfactorily simulated. Thus, more work is needed before we can replace the current Force-Restore working horse.

The optimal number of ISBA patches was discussed from different perspectives. Operational setups now use one or two patches, where two patches represent separation of forest and open land. However, the quite specific characters of bare soil, as part of open land, may be better represented if we go for three patches.

5. Surface data assimilation

In surface data assimilation we have quite a variety of algorithms being developed within the ACCORD community. For a long time CANARI-OI has been the working horse for all of us but now we see first steps beyond that on an operational level. Depending on our NWP environments and aims different pathways are now taken beyond CANARI-OI. On the algorithm side these are EPS-connected OI, SEKF and EnKF setups. On the surface analysis side we see titanlib/gridpp as an alternative to CANARI. For surface data assimilation observations we have relied entirely on SYNOP observations for a long time but now we see development and applications related to satellite products and radiances and crowdsource data.

From an algorithm perspective, three presentations were given with respect to SEKF developments and applications and one with respect to EnKF. Both Hungary and Norway are at or near test-operational setups including SEKF with the use of SYNOP observations. From Austria we saw a combination of SEKF-LAI development. One presentation included how pySurfex and titanlib/gridpp are used in a now-casting environment of the HARMONIE-AROME CSC. From the satellite perspective we had one presentation including assimilation of microwave brightness temperature and one on the assimilation of surface-sensitive microwave observations over land currently evaluating different methods to estimate surface emissivity.

The discussion was partly dedicated to the problem on how to define representative innovations for patches which have no observations representing them, more specifically here the combination of SYNOP observations and forest patch. Ideas on how to proceed with testing were presented.

6. DAsKIT (Data Assimilation starters kit) surface assimilation

The DAsKIT activities do formally belong to the upper-air data assimilation plans in ACCORD but since there is also an ingredient of surface OI assimilation here and specific questions related to that was on the table we saw this WW as a good opportunity to present problems and discuss how to proceed. Concrete suggested solutions were reached during the session. Please refer to the [DA ST2 activities on local implementation of a DA Starters KIT](#) for details.

7. A few separate items

The intention was to provide a session on ocean, air-sea/ice coupling, waves and lakes however this specific week did not attract so many presentations on these subjects. In the end we were still happy to enjoy one contribution on the status of coupled HARMONIE-AROME-OASIS-WW3 activities in Ireland. These subjects will be covered later in a specific arrangement.

Ekaterina Kurzeneva gave an overview on HIRLAM surface activities and also announced a discussion that took place the week after on the subject of atmosphere-surface coupled data

assimilation. These activities are formally arranged under the [Research Team 9](#) of the upper-air data assimilation area.

Introducing a roughness-sublayer in the vegetation-atmosphere coupling of HARMONIE-AROME

Metodija M. Shapkalijevski¹, Samuel Viana², Aaron Boone³, Quentin Rodier³, Patrick Le Moigne³, and Patrick Samuelsson¹

¹SMHI, Sweden

²AEMET, Spain

³Météo-France, France

1 Introduction

In the past 40 years, many studies showed that the standard flux-gradient relations, which are based on the similarity theory, are not fully applicable over a surface covered with high roughness elements (e.g. Thom et al. 1975; Raupach 1979; Garratt 1980; Cellier and Brunet 1992; Simpson et al. 1998; Mölder et al. 1999; Shapkalijevski et al. 2016). This is due to the canopy-induced turbulent mixing as a consequence from the interaction between the roughness elements and the atmospheric flow - a process that has not been explicitly treated in the standard surface similarity formulations. The effects of the so called roughness sublayer (RSL) on the surface-atmosphere coupling became more important since the lowest atmospheric level in NWP systems has been placed closer to the canopies (vegetation, urban), entering the layer (RSL) where the standard flux-gradient coupling relationships need a revision. This in turn affects the effective surface fluxes of momentum, energy and gases between the canopy and the atmosphere.

The present work aims at integrating the most advanced theory of the RSL into the SURFEX model (v8.1), adding increased physical details in the classical similarity theory over a vegetated surface. We have therefore developed a code of the RSL parameterization based on the Harman and Finnigan's RSL theory (Harman 2012; Harman and Finnigan 2008; Harman and Finnigan 2007), and have implemented it as a subroutine within the SURFEX framework. By doing so, we account for the altered turbulent exchange of momentum and energy in the coupling layer between the canopy and the atmosphere.

To be more efficient, robust and consistent with the RSL implementation in SURFEX, we have established a direct collaboration between the SURFEX developers (especially Multi-energy Balance (MEB) module developers), Meso-NH experts, and HARMONIE-AROME developers. To do so, we have used the ACCORD scientific visit opportunity to accomplish a visit to Météo-France (Toulouse).

Details of the code and the technical implementation in SURFEX shall be presented in a separate technical documentation. In what follows, preliminary results on the effects of the incorporated RSL parameterization in SURFEX on vegetation-atmosphere exchange properties shall be investigated. Since the latter defines the lower boundary forcing in atmospheric models, a further investigation of the RSL effects on the atmospheric boundary-layer state in HARMONIE-AROME and Meso-NH shall be provided and documented.

2 Implementation

We have incorporated the RSL parameterization in SURFEXv8.1 as independent optional subroutine and tested it *offline*, as well as *online*, as coupled to HARMONIE-AROME (from cy43h2.1.1) and to Meso-NH-v5.5.0. The RSL parameterization modifies the soil-vegetation-atmosphere transport (SVAT) of energy and momentum mainly by modifying the computation of the drag coefficients of momentum and heat (Harman 2012). The theoretical concept for the RSL implementation here was similar to the procedure presented in Harman (2012), as well as in more recent RSL implementations in GCM (Bonan et al. 2018) and WRF (Lee et al. 2020). Details

of the implementation procedure and modified subroutines currently can be found [here](#), and also at the 43rd EWGLAM conference.

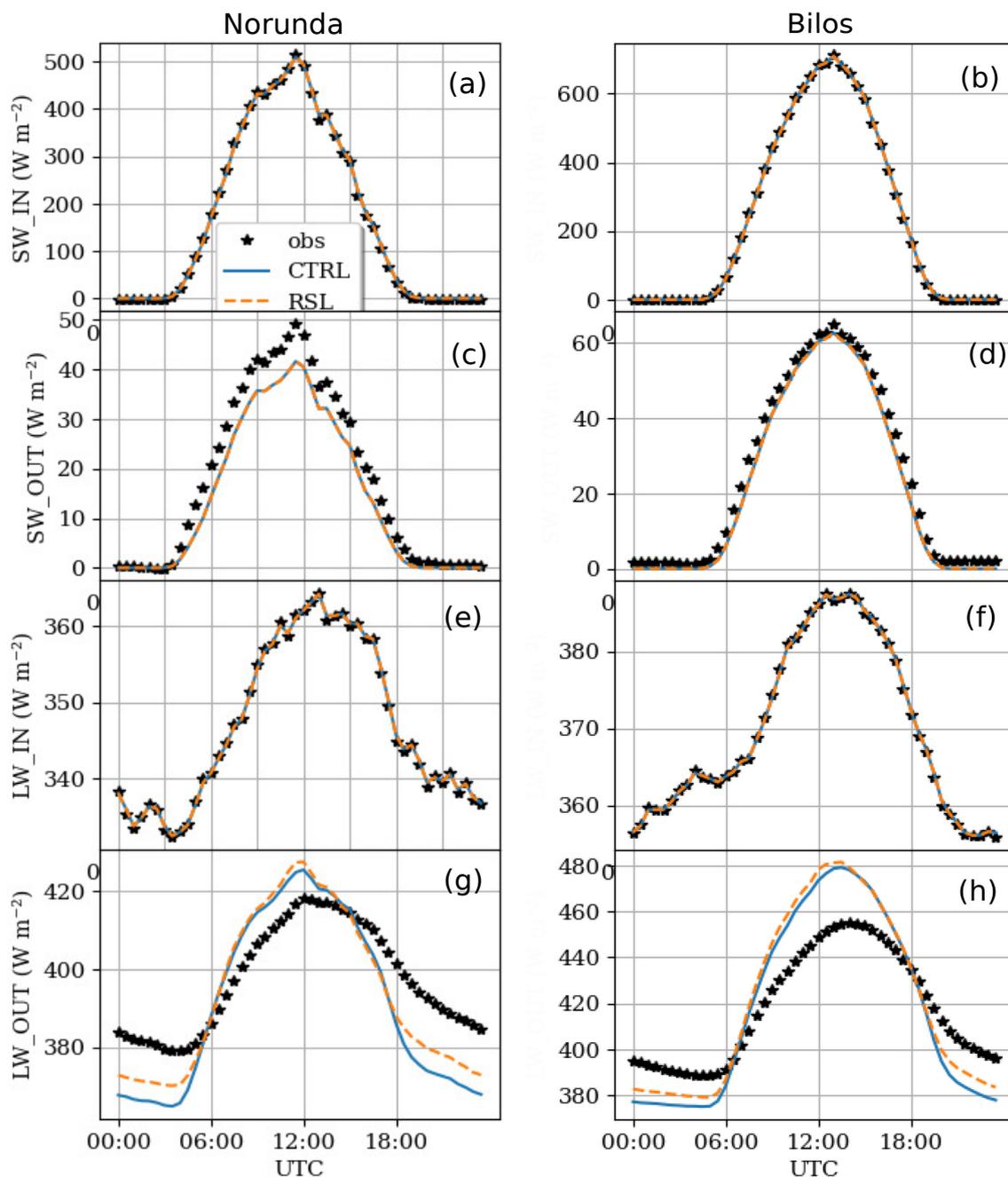


Figure 1: Diurnal cycle of the observed versus the modelled (using ISBA-MEB) 30-min. averaged radiation fluxes above the canopy over the Norunda (a, c, e, g) and Bilos (b, d, f, h) observatories. Black symbols represent the observed radiation fluxes; the orange dashed line and blue solid line show the SURFEX offline outputs with and without the RSL, respectively.

The SVAT coupling in SURFEXv8.1 is done through the traditional “composite” energy balance in ISBA (Noilhan and Planton 1989), or through the more advanced multi-energy balance (ISBA-MEB) scheme (Boone et al. 2017; Napoly et al. 2017). Thus, the RSL subroutines can be activated in each of these options. By setting `LRSL = TRUE/FALSE` in the SURFEX namelist, the RSL parameterization is activated/deactivated automatically. Only when activated, the ISBA or ISBA-MEB scheme, uses ‘dynamical’ roughness lengths and

displacements heights within the RSL parameterization. Also, when activated, a new output for the roughness lengths is set (X001Z0M and X002Z0M, for patch 1 (e.g. unforested) and patch 2 (forested), respectively). Given that the RSL collapses into traditional MOST when displacement height tends to zero, the implementation was done for both forested and unforested patches. Our aim is, by taking care of the SURFEX coding principles and standards, to prepare the RSL parameterization code for an optional usage in the next official version of SURFEX (v9).

3 Application and validation

3.1 Offline case studies

So far, we have set three realistic SURFEX offline case experiments (with and without the RSL parameterization) in forested areas, where long term flux-gradient measurements are available. This was done to further explore the canopy and the roughness effects on turbulent exchange properties between the vegetation and the atmosphere by validating the model results against tower flux-gradient observations. The case studies were developed at three locations: *i*) **Norunda** in Sweden, *ii*) **Bilos** in Salles (France) (both as a part of the Integrated Carbon Observation System - **ICOS**), as well as *iii*) **BERMS** (Boreal Ecosystem Research and Monitoring Sites) in Canada. All three ecosystems are similar (Boreal needleleaf evergreen plant functional type in SURFEX) with varying average canopy height between 10 and 29 meters, and similar canopy density/sparsity (e.g. leaf area index (LAI) between 1 and 3). 30 minutes average meteorological observational data (wind speed, temperature, humidity, radiation) has been used to force the simulations at 58 meters for Norunda case, 15.6 m for Bilos case, and 37 meters for BERMS case, respectively (heights above the ground surface). Observed fluxes and gradients above the canopy have been used to validate the simulation performance (see Figs. 1 - 3).

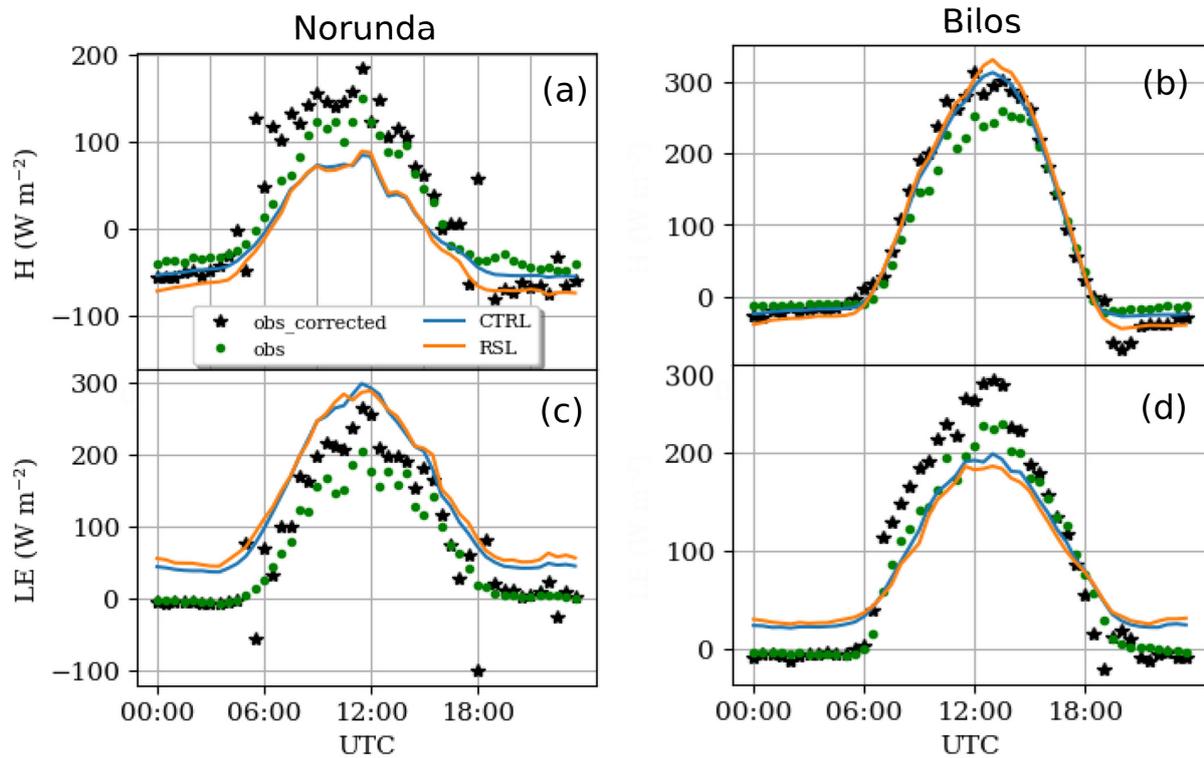


Figure 2: Diurnal cycle of the observed versus the modelled (using ISBA-MEB) 30-min. averaged sensible (H) and latent (LE) fluxes above the canopy at the Norunda (a, c) and Bilos (b, d) observatories. Observed fluxes (green solid dots represents the eddy co-variance technique applied on raw data) are corrected (black symbols) using the Bowen ratio method (Blanken et al. 1997).

Here we mainly present results from the simulations at Norunda and Bilos, as these sites are within the Met-CoOp and France domains of HARMONIE-AROME. They correspond to the average diurnal cycles during the period of the simulations: 8-31 August, 2019 for Norunda and 1 June - 9 September, 2021 for Bilos. Prior to these periods, a short spinup period of 7 days for Norunda and 45 days for Bilos was run (a subject of continuous data availability). The initial state of the soil temperatures was obtained from the ICOS observations, while the initial soil humidities were established by assuming the soil moisture proportional to the field capacity. Although soil moisture observations are available for both sites, these were not used in the current simulations to avoid uncertainty in relation to the unknown soil textures. Some additional sensitivity shall be provided in the future however, but for the purpose of validation the RSL contribution in the SURFEX SVAT scheme (the aim of the current study), this type of sensitivity is irrelevant.

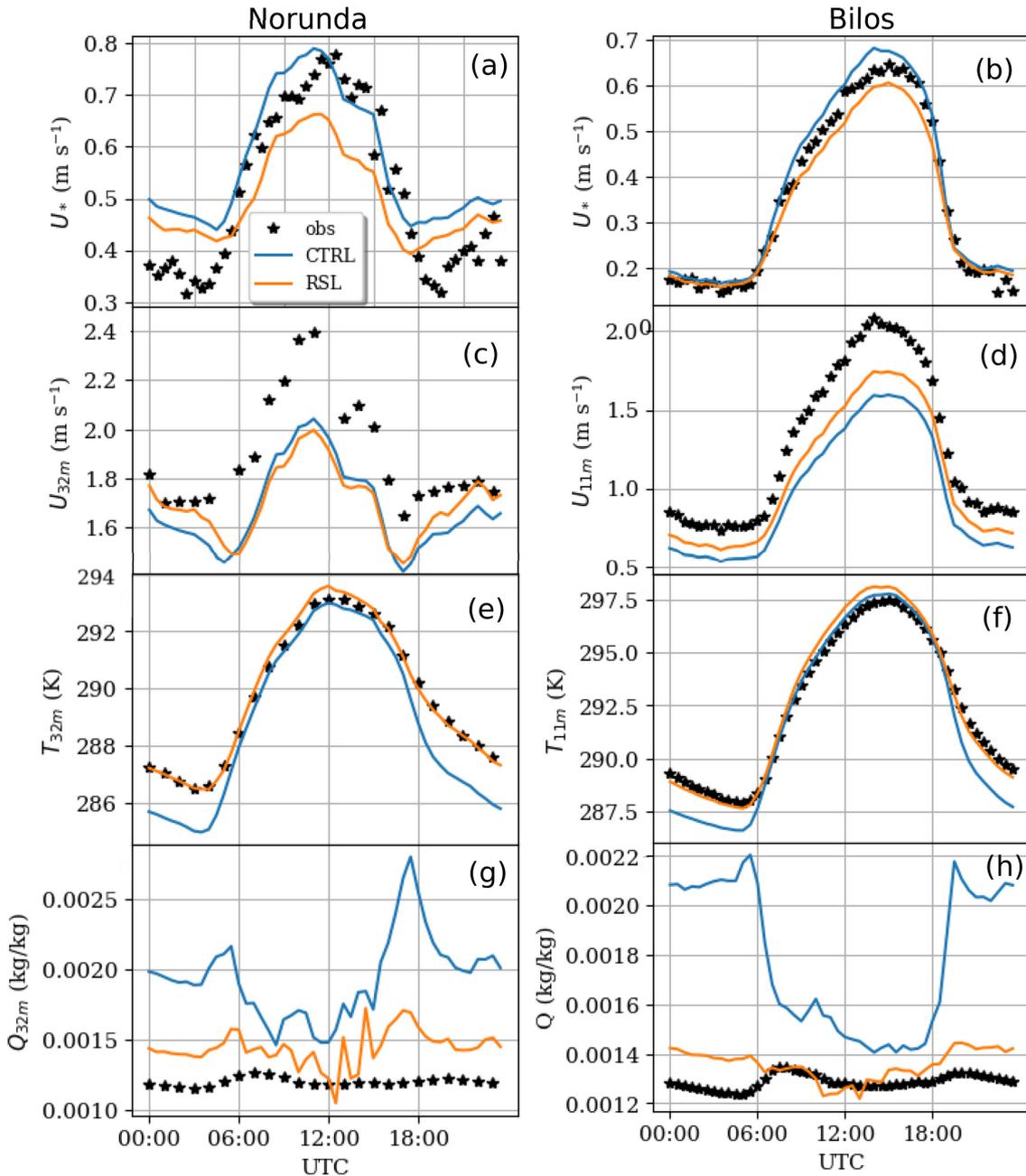


Figure 3: Diurnal cycle of the observed versus the modelled (using ISBA-MEB) 30-min. friction velocity (u_*), modulus of the wind speed (U), temperature (T) and specific humidity (Q) above the canopy at Norunda observatory (a, c, e, g) and at Bilos observatory (b, d, f, h).

Radiation forcing of the model above the canopy, as incoming short-wave (\downarrow SW) and long-wave (\downarrow LW) energy, resulted in reasonable simulation of the outgoing radiation partitions (\uparrow SW and \uparrow LW) and thus the net available energy overall. More precisely however, there was a slight overestimation of the outgoing long-wave radiation during daytime, as well as its underestimation during nighttime (up to 20 W m^{-2}) (Fig. 1g and 1h). The latter was followed by a larger underestimation of the daytime sensible heat (H) between the canopy and the forcing level over Norunda (up to about 70 W m^{-2} on average), while the latent heat (LE) partitioning was overestimated for about 50 W m^{-2} (Fig. 2a – 2c). H over Bilos was simulated well, but LE was still unrealistically overestimated during nighttime and underestimated during daytime (Fig. 2b – 2d). The unsatisfactory results for the LE, in general, require some additional investigation, especially the unrealistic positive moisture flux during nighttime (Fig. 2c – 2d). However, more important for this study is that the activation of the RSL effects in the vegetation-atmosphere coupling did not show strong influence on both H and LE, except the slight improvement in the modelled nighttime cooling (Fig. 2a and 2b). A larger improvement due to the inclusion of the RSL parameterization is found on the modelled dynamics within the roughness sublayer, as represented by the friction velocity u_* , and the wind speed just above the canopy (Fig. 3a – 3d). This led to a clear improvement in the diagnostically calculated wind speed at Bilos, while at Norunda the improvement due to the RSL parameterization was mainly during nighttime; decrease of friction velocity and increase of wind speed. The worst model dynamics in general over the Norunda can be related to the presence of more complex boundary-layer dynamics. For instance, the diurnal cycle of wind speed at 58 m had a secondary maximum during nighttime, suggesting the presence of low-level jets (LLJ). While these LLJs were already absent (or largely decreased) at the level of 32 m, the interpolated wind speed at this level showed their existence. Meanwhile, the temperature (Fig. 3e and 3f) and specific humidity diagnostics (Fig. 3g and 3h) above the canopy were equally improved at both sites due to the RSL presence.

Preliminary results showed similar behaviour for both ISBA and ISBA-MEB with respect to RSL inclusion, but more consistent comparison shall be additionally provided.

3.2 Online simulations

HARMONIE-AROME (Bengtsson et al. 2017) Several simulations with and without the RSL parameterization in the vegetation-atmosphere coupling in full three-dimensional HARMONIE-AROME setup have been conducted. The simulations were done over the IBERIA and METCOOP domains to illustrate different topographical and climatological characteristics. The ten days of simulations for each of the domains had spinup time of ten days. A summary of the simulations is shown in Table 1.

Table 1: HARMONIE-AROME simulation setups

Domain	IBERIA	METCOOP
System version (ISBA-FR)	cy43h2.1.1	cy43h2.1.1
System version (ISBA-DIF/MEB)	pre-cy46h1	pre-cy46h1
Simulation time (after spinup)	10 days	10 days
Simulation date(s)	20190101 - 20190110	20210415 - 20210425

The presence of the RSL parameterization in vegetation-atmosphere coupling of HARMONIE-AROME systematically reduced the scatter in the diagnosed wind speed at 10 m when compared to standard meteorological observations (at stations throughout the domains) (Fig. 4). More precise validation of the HARMONIE-AROME flux-gradient output above tall vegetation (e.g. forests) cannot be provided by using observations from these standard meteorological stations. Flux-gradient observations from towers within forests should be used instead (e.g. Norunda or Bilos observatory; work in progress).

Meso-NH (Lac et al. 2018) To test the implementation RSL parameterization and its effects on the online SURFEX vegetation-atmosphere coupling in Meso-NH, we have set a simple and well studied experiment over the largely vegetated La Réunion Island in the tropics, a French department in the Indian Ocean. This has

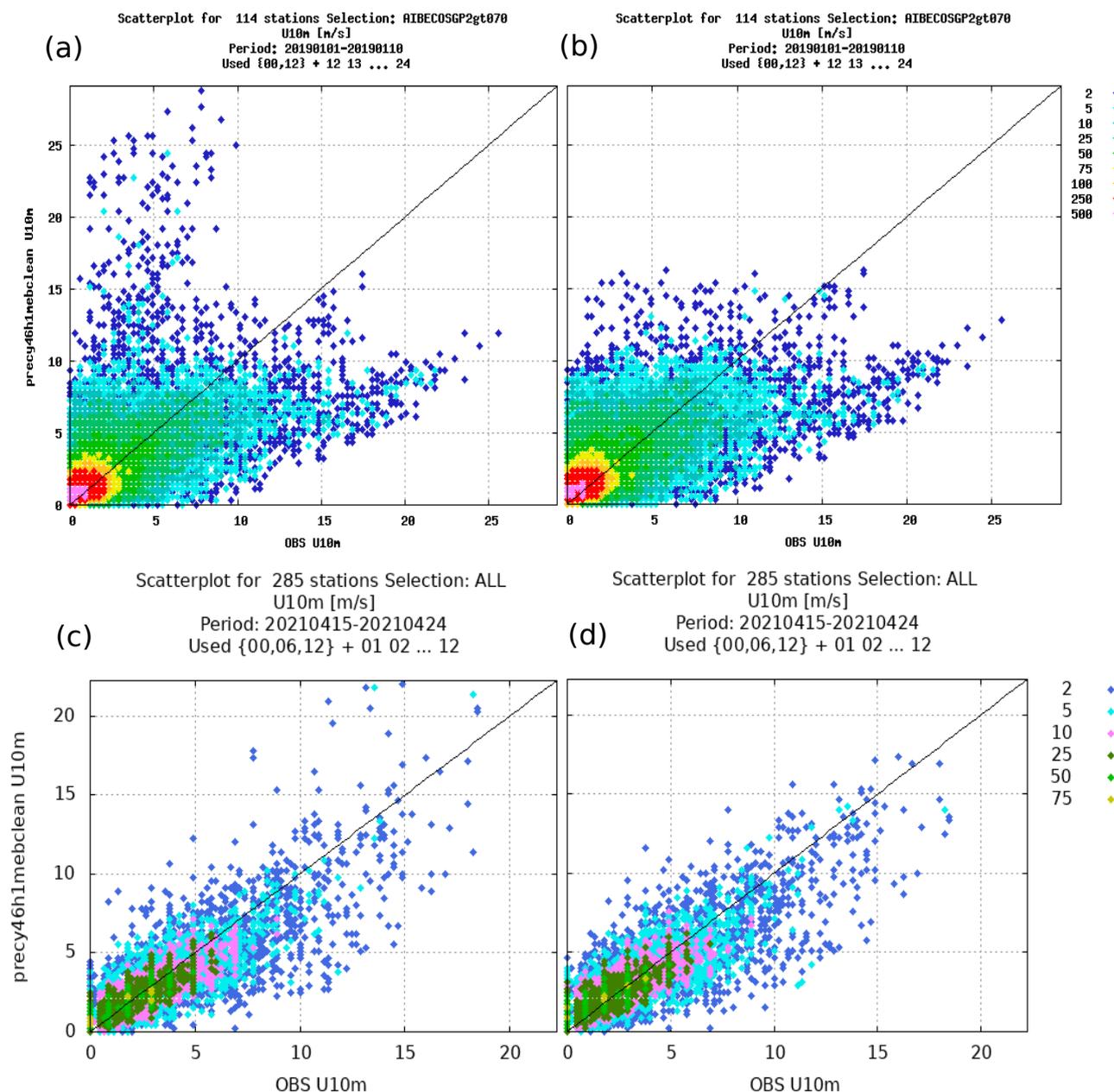


Figure 4: Scatterplot of the modelled by HARMONIE-AROME and the observed wind speed at 10 m over the IBERIA ((a)-(b)) and METCOOP/Sweden ((c)-(d)) domains without ((a)-(c)) and with the RSL parameterization ((b) and (d)), respectively. The scatterplots for the IBERIA simulation correspond to a selection of stations where the fraction of forest is larger than 70%.

been done on the basis of ISBA-FR approach. An example of the qualitative differences of the near-surface meteorology in ISBA-FR with and without the RSL parameterization are shown in Fig. 5. Decreased wind speeds in the field are result of the stability dependent roughness lengths, which in the composite ISBA is a constant fraction of the vegetation height only.

3.3 Discussion and outlook

The surface-atmosphere coupling in forested areas, in the case of the SURFEX offline runs is done at levels where forcing data is available (e.g. tower observation in the above cases). This (forcing) data is usually set

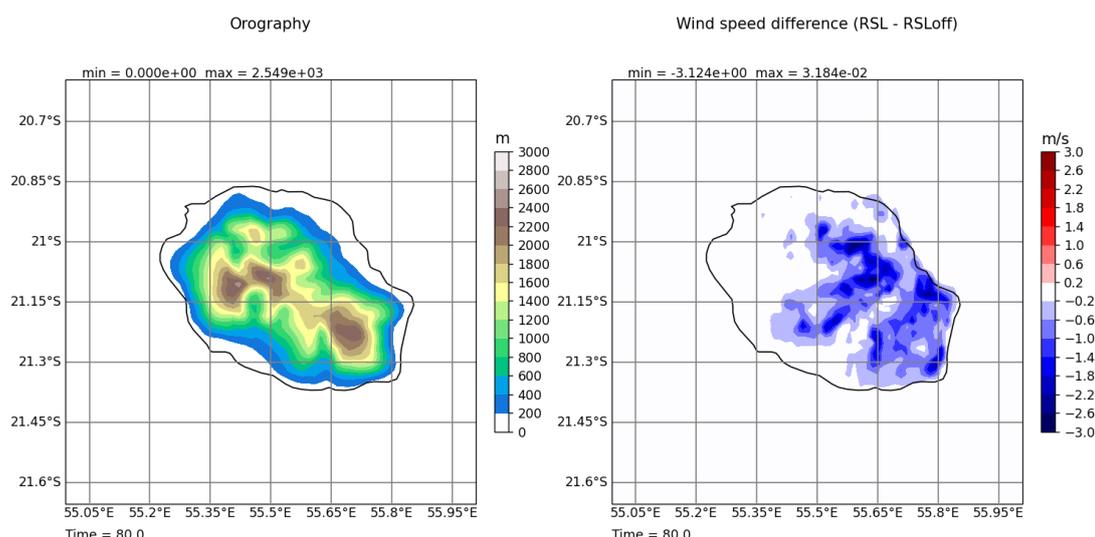


Figure 5: The absolute difference of the wind-speed field at lowest model level over the La Réunion Island, as diagnosed by the vegetation-atmosphere coupling scheme with and without RSL parameterization.

(well) above the canopy to meet the assumptions required by the surface theory (e.g. constant vertical flux in height). However, for the online SURFEX simulations, the surface-atmosphere coupling in forested areas (e.g. PATCH2) is done at a height set at the lowest atmospheric level (z_n), which is independent from the surface cover properties (heterogeneity) and thus *a priori* assumed that the meteorological conditions at z_n are similar to those above the canopy top in reality; then, the vegetation-atmosphere coupling is virtually done between the z_n and the displacement height (the zero vertical coordinate for the atmospheric model). Due to the assumptions made here, a point validation of online SURFEX simulations over the presented case studies (e.g. HARMONIE-AROME) are expected to vary from the results presented based on the offline SURFEX output.

This is exactly the purpose of the authors' current activity. The aim is to set up full 3D HARMONIE-AROME simulations for the given time periods, August 2019 and summer 2021, over the Swedish and the Spanish domains, respectively. Similar analysis as presented here for the offline cases shall be provided for the extracted grid points of the online NWP output at Norunda and Bilos, as well as its validation against the tower observations. The results of the offline and the online simulations shall be compared and further discussed.

4 Conclusions

We studied the consistency of the RSL implementation in both ISBA and ISBA-MEB vegetation-atmosphere coupling schemes in SURFEX and HARMONIE-AROME. We also investigated the RSL effects on the surface fluxes and diagnostics in SURFEX offline and online (HARMONIE-AROME, Meso-NH) experiments by using high-quality flux-gradient observations above canopies.

The preliminary results showed consistent implementation of the RSL parameterizations in SURFEX offline and online experiments. The SURFEX offline validation showed satisfactory agreement in representing the dynamics and the thermodynamics above the canopy at Norunda. The presence of the RSL parameterization decreased the somewhere-present bias between the observed and modelled fluxes and gradients in both ISBA and ISBA-MEB coupling, with a larger improvement in representing the diurnal cycle of momentum fluxes and the diagnostic wind speed above the canopy. Additionally, the online HARMONIE-AROME simulations showed consistent reduction in diagnostic wind-speed scatter above Iberian and Scandinavian forests.

Additional online HARMONIE-AROME and Meso-NH simulations with and without the RSL parameterization over Sweden and France/Spain should provide a valuable information about the differences of the offline and online SVAT schemes in SURFEX and thus the NWP performance in simulation near-surface exchange processes above forested areas.

Acknowledgments

We would like to acknowledge the ACCORD consortium for supporting our visit at Météo-France, but also thank the hosts, Aaron, Patrick and Quentin, for organizing our working conditions there in pandemic times. We would also like to acknowledge the help by Mölder Meelis (Lund University, and PI at Norunda observatory), as well as Adrien Napoly (MeteoFrance) for the data provision at Norunda and Berms observatories, respectively.

References

- Bengtsson, L., U. Andrae, T. Aspelién, Y. Batrak, J. Calvo, W. de Rooy, E. Gleeson, B. Hansen-Sass, M. Homleid, M. Hortal, et al. (2017). “The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system”. In: *Monthly Weather Review* 145.5, pp. 1919–1935.
- Blanken, P., T. A. Black, P. Yang, H. Neumann, Z. Nesic, R. Staebler, G. Den Hartog, M. Novak, and X. Lee (1997). “Energy balance and canopy conductance of a boreal aspen forest: partitioning overstory and understory components”. In: *Journal of Geophysical Research: Atmospheres* 102.D24, pp. 28915–28927.
- Bonan, G. B., E. G. Patton, I. N. Harman, K. W. Oleson, J. J. Finnigan, Y. Lu, and E. A. Burakowski (2018). “Modeling canopy-induced turbulence in the Earth system: a unified parameterization of turbulent exchange within plant canopies and the roughness sublayer (CLM-ml v0)”. In: *Geoscientific Model Development* 11.4, pp. 1467–1496.
- Boone, A., P. Samuelsson, S. Gollvik, A. Napoly, L. Jarlan, E. Brun, and B. Decharme (2017). “The interactions between soil–biosphere–atmosphere land surface model with a multi-energy balance (ISBA-MEB) option in SURFEXv8 – Part 1: Model description”. In: *Geoscientific Model Development* 10.2, pp. 843–872. DOI: [10.5194/gmd-10-843-2017](https://doi.org/10.5194/gmd-10-843-2017).
- Cellier, P. and Y. Brunet (1992). “Flux-gradient relationships above tall plant canopies”. In: *Agricultural and Forest Meteorology* 58.1, pp. 93–117. DOI: [10.1016/0168-1923\(92\)90113-I](https://doi.org/10.1016/0168-1923(92)90113-I).
- Garratt, J. R. (1980). “Surface influence upon vertical profiles in the atmospheric near-surface layer”. In: *Quarterly Journal of the Royal Meteorological Society* 106.450, pp. 803–819. DOI: [10.1002/qj.49710645011](https://doi.org/10.1002/qj.49710645011).
- Harman, I. N. (2012). “The role of roughness sublayer dynamics within surface exchange schemes”. In: *Bound.-Lay. Meteorol.* 142, pp. 1–20. DOI: <https://doi.org/10.1007/s10546-011-9651-z>.
- Harman, I. N. and J. J. Finnigan (2007). “A simple unified theory for flow in the canopy and roughness sublayer”. In: *Bound.-Lay. Meteorol.* 123, pp. 339–363. DOI: <https://doi.org/10.1007/s10546-006-9145-6>.
- (2008). “Scalar concentration profiles in the canopy and roughness sublayer”. In: *Bound.-Lay. Meteorol.* 129, pp. 323–351. DOI: <https://doi.org/10.1007/s10546-008-9328-4>.
- Lac, C., J.-P. Chaboureau, V. Masson, J.-P. Pinty, P. Tulet, J. Escobar, M. Leriche, C. Barthe, B. Aouizerats, C. Augros, et al. (2018). “Overview of the Meso-NH model version 5.4 and its applications”. In: *Geoscientific Model Development* 11.5, pp. 1929–1969.
- Lee, J., J. Hong, Y. Noh, and P. A. Jiménez (2020). “Implementation of a roughness sublayer parameterization in the Weather Research and Forecasting model (WRF version 3.7. 1) and its evaluation for regional climate simulations”. In: *Geoscientific Model Development* 13.2, pp. 521–536.
- Mölder, M., A. Grelle, A. Lindroth, and S. Halldin (1999). “Flux-profile relationships over a boreal forest — roughness sublayer corrections”. In: *Agricultural and Forest Meteorology* 98-99, pp. 645–658. DOI: [10.1016/S0168-1923\(99\)00131-8](https://doi.org/10.1016/S0168-1923(99)00131-8).
- Napoly, A., A. Boone, P. Samuelsson, S. Gollvik, E. Martin, R. Seferian, D. Carrer, B. Decharme, and L. Jarlan (2017). “The interactions between soil–biosphere–atmosphere (ISBA) land surface model multi-energy balance (MEB) option in SURFEXv8 – Part 2: Introduction of a litter formulation and model evaluation for local-scale forest sites”. In: *Geoscientific Model Development* 10.4, pp. 1621–1644. DOI: [10.5194/gmd-10-1621-2017](https://doi.org/10.5194/gmd-10-1621-2017).

- Noilhan, J. and S. Planton (1989). “A Simple Parameterization of Land Surface Processes for Meteorological Models”. In: *Monthly Weather Review* 117.3, pp. 536–549. DOI: [10.1175/1520-0493\(1989\)117<0536:ASPOLS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117<0536:ASPOLS>2.0.CO;2).
- Raupach, M. (1979). “Anomalies in flux-gradient relationships over forest. Boundary-Layer Meteorol”. In: *Bound.-Lay. Meteorol.* 16, pp. 467–486. DOI: <https://doi.org/10.1007/BF03163564>.
- Shapkalijevski, M., A. F. Moene, H. G. Ouwensloot, E. G. Patton, and J. V.-G. de Arellano (2016). “Influence of Canopy Seasonal Changes on Turbulence Parameterization within the Roughness Sublayer over an Orchard Canopy”. In: *Journal of Applied Meteorology and Climatology* 55.6, pp. 1391–1407.
- Simpson, I., G. Thurtell, H. Neumann, G. Den Hartog, and G. C. Edwards (1998). “The Validity of Similarity Theory in the Roughness Sublayer Above Forests”. In: *Bound.-Lay. Meteorol.* 87, pp. 69–99. DOI: [10.1023/A:1000809902980](https://doi.org/10.1023/A:1000809902980).
- Thom, A. S., J. B. Stewart, H. R. Oliver, and J. H. C. Gash (1975). “Comparison of aerodynamic and energy budget estimates of fluxes over a pine forest”. In: *Quarterly Journal of the Royal Meteorological Society* 101.427, pp. 93–105. DOI: [10.1002/qj.49710142708](https://doi.org/10.1002/qj.49710142708).

Algorithmic amelioration of the deficiencies in the screen level parameters forecast based on a dynamical downscaling approach

Martin Dian, Mária Derková, Martin Petráš (SHMU)

1 Introduction

An experimental convection-permitting setup of the ALARO CSC, denoted ALA2, is running regularly at SHMU in the dynamical downscaling mode. It proved to perform better compared to the operational ALADIN/SHMU (ALARO-1vB) runs in some cases, namely in convective situations it is appreciated by our forecasters.

However, during summer 2021 a serious deterioration of the ALA2 2 m temperature forecast has been noticed. It was traced down to the surface fields originating from a driving model ARPEGE. CANARI optimal interpolation configuration has been utilized to correct surface and soil temperature and moisture that seemed to be sufficient to improve the screen level parameters forecast, as shown in this contribution.

2 ALARO 2 km setup

In parallel to the operational ALARO/SHMU 4.5 km/L63 operational configuration, an experimental convection-permitting ALARO version with 2 km horizontal resolution and 87 vertical levels, denoted ALA2, is also daily applied at SHMU (Simon et al., 2021). It uses ALARO-1vB physics settings for upper air, and ISBA scheme for surface physics (i.e. no SURFEX). ALA2 has been operated in the dynamical adaptation mode without data assimilation, having the digital filter initialisation activated to increase the numerical stability and reduce the initial noise. Effectively, two identical ALA2 versions are running regularly at SHMU, their difference is only the initial and lateral boundary data. The primary version (further in the text referred to as ALA2_A) is driven by the ARPEGE global model using an 1-hour coupling frequency, and the initial file is an ARPEGE analysis (the “zero” coupling file). Later, another ALA2 version has been set up, which is driven by the ECMWF global model deterministic run with a 3-hourly coupling frequency (further in the text referred to as ALA2_E). ALA2_E uses A-LAEF analysis from the control run as the initial conditions. The details of ALA2 setups are listed in Table 1.

3 Problems with the 2 m temperature forecast

Description of the problem

During summer 2021 it was reported that ALA2_A forecast of 2 m temperature is often overestimated, in particular the maximum temperature forecast during the hot days. This issue is illustrated on Figure 1, where the forecasts of models run at SHMU (see the Figure legend for more details) for an arbitrarily chosen date for two stations (Dudince and Trencin) are plotted against the observations (violet colour). The ALA2_A forecast for the 3rd day is 2-4 degrees warmer than the measurements. Such a behaviour has been confirmed by the means of the objective verification scores of 2 m

temperature and relative humidity for a two weeks period 23/07-05/08/2021 against 95 automatic weather stations (AWS) of SHMU, see Figures 2 and 3. Scores of ALA2_A are in light blue. Such 2 m temperature overestimation was not present in the ALA2_E runs (black curve), nor in the operational ALADIN/SHMU (4.5 km/L63, CANARI & DF bending, coupling with ARPEGE; yellow line). Moreover, 2 m temperature scores of ALA2_A were usually not worse compared to operational ones throughout the year (not shown).

The ALA2_A problem was traced down to the surface field of the initial state. We recall that the ALA2_A initial state is provided by ARPEGE via e927 configuration, where surface fields are internally interpolated from SURFEX to ISBA variables. There seems to be an inconsistency between SURFEX and ISBA fields that is not fully tackled by e927, as illustrated on Figure 4. The differences in the surface and soil temperature and moisture between the initial fields from ARPEGE interpolated via ee927 into the target 2 km resolution and the CANARI analysis are shown for an arbitrarily chosen warm date. We presume such differences are not fully explained by change of horizontal resolution and/or driving and target models topographies.

Note that there was no investigation of the particular synoptic situations in which the forecast failures described above occurred.

Table 1: ALARO 2 km/L87 setups @ SHMU

Label	ALA2_A	ALA2_E
Model CSC version	ALARO-1vB	
Code version	CY43t2_bf11	
Resolution	2.0 km	
Levels	87	
Area	1024 x 768 km	
Initial conditions	ARPEGE analysis -> CANARI	A-LAEF CNTRL memb. analysis
Upper air data assimilation	none - dynamical downscaling	
Initialisation	Digital Filter	
Boundaries	ARPEGE, a' 1 hour	ECMWF, a' 3 hours
Surface scheme	ISBA	
Starting times	00, 12 UTC	00, 12 UTC
Integration lengths	78, 72 hours	81, 81 hours

Utilization of the CANARI surface analysis

Simple technical solution was proposed to correct surface fields using CANARI analysis. The solution is labeled "technical" as there is no full data assimilation cycling applied. The procedure involves 2 m temperature and relative humidity SYNOP observations from the GTS and additionally national AWS data from OPLACE¹ that are utilized to correct the 12 h first guess surface and deep soil temperature and moisture of ALA2_A. Then, the four fields that are modified by CANARI (SURFTEMPERATURE, PROFTEMPERATURE, SURFRESERV.EAU, PROFRESERV.EAU) are replaced in the INIT file, that is an ARPEGE short cut-off analysis represented as the LBC0 file, using local utilities *farecdel* and *fa_copy*. We recall that this procedure is applied on the production runs only, no long cut-off assimilation is utilized. In other words - in the proposed algorithm only the four surface fields listed above are cycled, other surface fields are interpolated from Arpege using ee927 configuration. Such a strategy proved to be sufficient to reduce the issue with the 2 m parameters forecasts while it enables us to profit from the fresh upper-air fields provided by ARPEGE 4D-Var. The scores of a 2 weeks long test, +72 hours runs from 00 and 12 UTC for the 2m temperature and relative humidity are shown on Figures 2 and 3, where the new setup of ALA2_A version, denoted

¹ OPLACE: RC LACE common operational database, Trojakova et. al 2019

ALA2_C, is plotted in green. The scores are displayed from 23/07/2021, while CANARI has been introduced on 18/07/2021. The impact of CANARI analysis on the forecast for individual points is also shown on Figure 1 for Trencin and Dudince stations (light blue lines). This new ALA2_C setup has been used regularly since 06/08/2021.

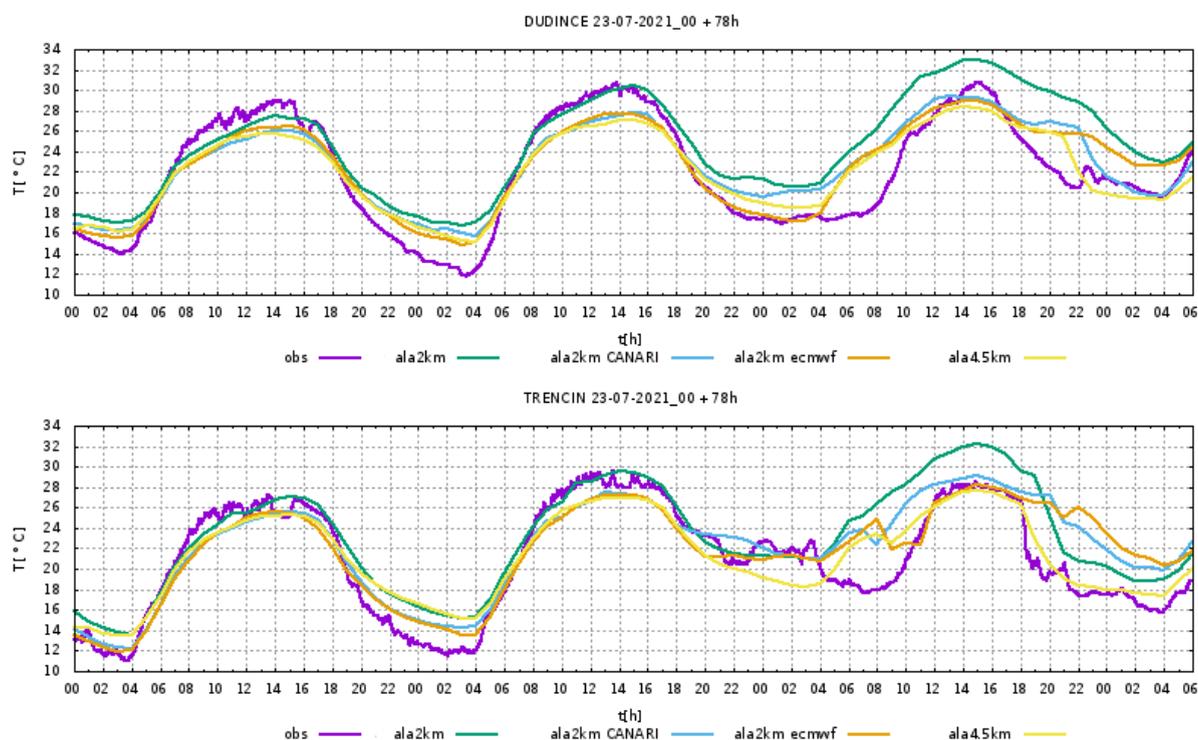


Figure 1: 2m temperature forecast for Dudince (top) and Trencin (bottom) stations. Base 23th July 2021 00 UTC run. Colour lines: violet - observations, green - ALA2_A without CANARI, light blue - ALA2_A with CANARI, orange - ALA2_E, yellow - operational ALARO/SHMU.

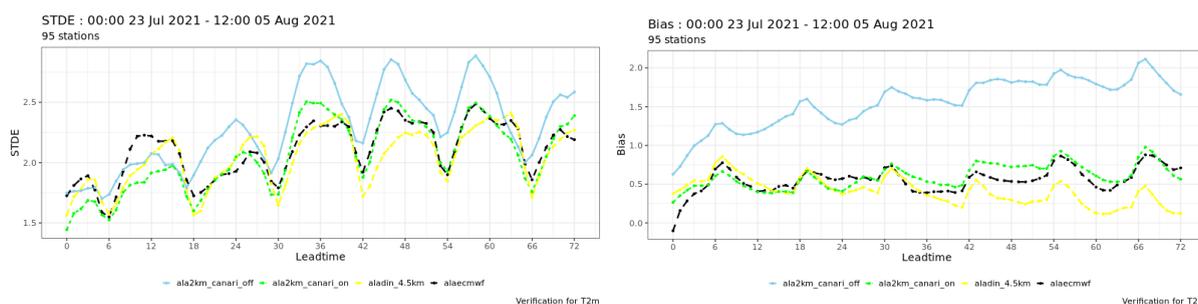


Figure 2: STDEV (left) and BIAS (right) scores of the height adjusted T2m for 00 and 12 UTC forecasts, period 23 July - 5 August 2021 from HARP verification. Operational ALADIN/SHMU is plotted in yellow, ALA2_E (ALA2 coupled to ECMWF with INIT from A-LAEF CNTRL memb) in black, ALA2_A (ALA2 without CANARI, coupled to ARPEGE) in light blue, ALA2_C (ALA2 with CANARI, coupled to ARPEGE) in green.

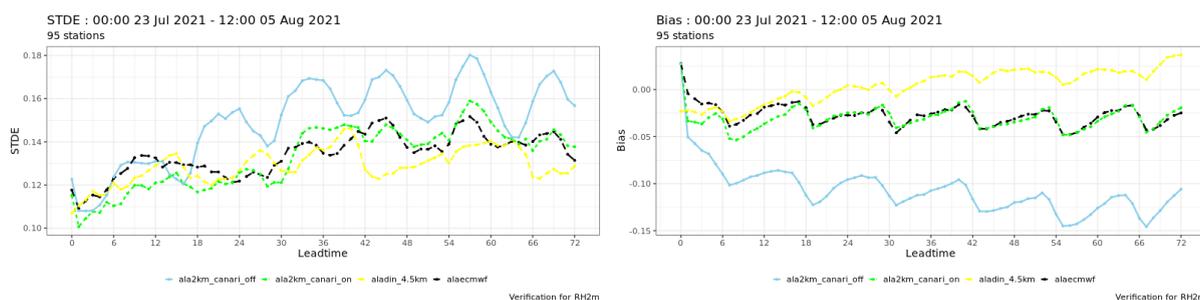


Figure 3: The same as the previous Figure, but for 2 m relative humidity.

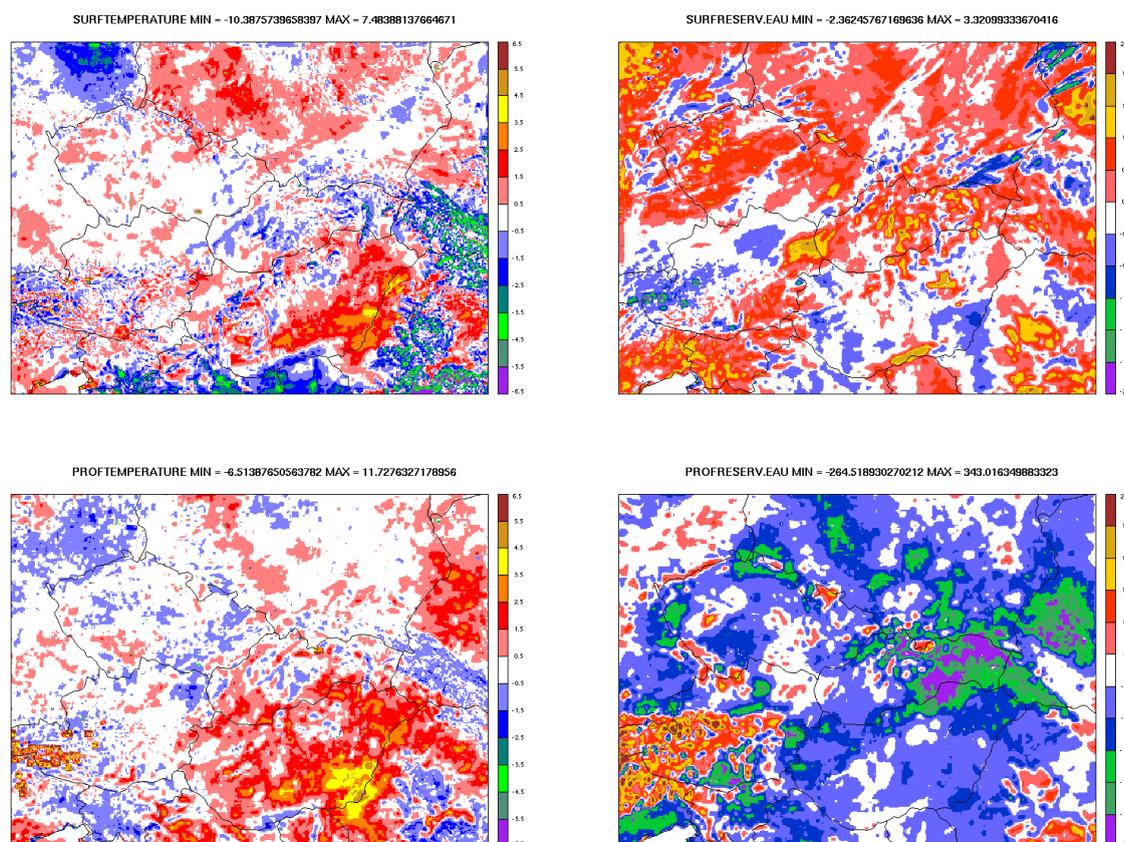


Figure 4: Differences between the ARPEGE initial fields interpolated into 2 km horizontal resolution grid of ALA2_A and corresponding CANARI analysis. Base time: 05-08-2021_00. Top left: surface temperature. Bottom left: mean soil temperature. Top right: surface soil moisture. Bottom right: total soil moisture.

4 Conclusions and Perspectives

The aim of this newsletter contribution was to describe and make others aware of the issue that has been noticed in runs of ALARO with ISBA applied in the dynamical adaptation mode during summer 2021.

To alleviate the deficiencies identified in the 2m temperature and relative humidity forecasts of the ALARO 2 km/L87 configuration run at SHMU the CANARI analysis for the surface parameters has been introduced. The utilized cycling approach is rather technical - the analyzed temperature and soil

moisture fields are replaced in the initial conditions file that is an ARPEGE analysis so that the upper air fields of the driving model analysis are preserved. Such a simple solution was sufficient to solve a pressing problem. However the full data assimilation cycle based on 3D-Var for convection-permitting setup of ALARO CSC is planned with the new HPC.

5 References

Simon, A., M. Belluš, K. Čatlošová, M. Derková, M. Dian, M. Imrišek, J. Kaňák, L. Méri, M. Neštiak and J. Vivoda, 2021: Numerical simulations of 7 June 2020 convective precipitation over Slovakia using deterministic, probabilistic and convection-permitting approaches. *IDŐJÁRÁS*, Vol. 125, No. 4, 571–607. DOI:10.28974/idojaras.2021.4.3

Trojáková, A., Mile, M. and Tudor, M., 2019: Observation Preprocessing System for RC LACE (OPLACE). *Adv. Sci. Res.* 16, 223–228. <https://doi.org/10.5194/asr-16-223-2019>.

Improving model performance in stable situations by using a pragmatic shift in the drag calculations - XRISHIFT

Mariken Homleid

1. Introduction

The lowest temperatures ever recorded in North Scandinavia were $-52.6\text{ }^{\circ}\text{C}$ in Lapland, Sweden, in 1966, $-51.5\text{ }^{\circ}\text{C}$ in Kittilä, Finland, in 1999 and $-51.4\text{ }^{\circ}\text{C}$ in Kautokeino, Norway, in 1886. Even with climate warming, there still may be occurrences of temperatures below $-30\text{ }^{\circ}\text{C}$ during winter. It has been a long-standing issue to forecast the lowest temperatures; the forecasts are often more than $10\text{ }^{\circ}\text{C}$ too warm. These problems are also reflected in summary scores of 2 m temperature where positive mean errors (ME) are seen in cold winters. This problem is associated with having a stable boundary layer. The lowest temperatures during stable summer nights are also overestimated on average. Too much modelled fog over the sea is also supposed to be a systematic shortcoming related to the performance in stable situations. Even if very low temperatures occur less frequently, due to climate change, we would like to improve the model performance in stable situations.

HARMONIE-AROME (Bengtsson et al. 2017) uses Surface Externalisée (SURFEX) externalized surface scheme (Masson et al. 2013). The surface flux calculations are controlled by the XRIMAX option. This option gives an upper limit to the estimated Richardson number (Ri). $\text{XRIMAX} = 0$ implies that the fluxes in stable conditions are based on drag coefficients for neutral conditions, and could be an obstacle for improving the performance in stable situations. Sensible heat fluxes that are too high prevent the surface temperatures from decreasing as they should. The XRIMAX setting affects all fluxes: sensible and latent heat fluxes, and momentum fluxes over all 4 tiles.

A surface boundary multi-layer scheme, with six additional levels between the lowest model level and the surface, were activated in the first HARMONIE-AROME runs with SURFEX. With the lowest level at 0.5 m, it was found that $\text{XRIMAX} = 0$ gave better and more realistic results than $\text{XRIMAX} = 0.2$ (Donier et al. 2012, Müller et al 2017). The surface boundary multi-layer scheme was switched off when going from one to two patches, triggering new experiments with $\text{XRIMAX} > 0$.

Experiments with $\text{XRIMAX} > 0$ (e.g. 0.2 or 0.5) showed potential for improved performance in cold stable situations at inland stations, but also showed unrealistic temperature drops, particularly at coastal stations in the north and at inland stations in mountainous regions where the roughness length was low and when snow was on ground. Some results of experiments with $\text{XRIMAX} = 0.2$ and 0.5 are presented in Sec. 2.

The potential improvements seen with $\text{XRIMAX} > 0$ and the drawbacks with $\text{XRIMAX} = 0$, inspired us to search for pragmatic solutions to improve the performance in stable situations by letting XRIMAX exceed 0, but with a reduced impact. The pragmatic solutions involve modifications in the calculation of the drag coefficients as functions of the Richardson number and the roughness length. Sec. 3 gives some equations from the SURFEX Scientific Documentation (Le Moigne 2018). The question of how the Richardson number is estimated is not addressed here. A minor change to reduce the sensitivity of the wind speed is however introduced, by increasing the threshold of the wind speed in the calculations from 1 m/s to 2 m/s.

The final setup is presented in Sec. 4, followed by some examples and summary results from parallel experiments. Clear improvements are seen in some situations. On average the change leads to a significant reduction in the 2 m temperature forecast during the cold season. This setup was chosen for operational runs of MetCoOp.

2. Experiments with XRIMAX>0

XRIMAX is set differently in different operational weather forecasting centers. For example, in Météo France it is set to 0.2. Inspired by colleagues at Météo France, several experiments with XRIMAX > 0 were evaluated in the HIRLAM consortium. The experiments have shown potential for improved performance in cold stable situations, at inland stations, but also uncover many examples of unrealistic temperature drops, particularly at coastal stations, but also at inland stations.

Time series of 2 m temperature and 10 m wind speed from an inland station and a coastal station for the end of January 2019 are shown in Fig. 1a and b to illustrate the potential improvements and problems. Results from three parallel model experiments on the AROME-Arctic domain are included; with XRIMAX = 0, 0.2 and 0.5. Results for the preoperational AROME-Arctic with a setup similar to the experiments are also provided (XRIMAX = 0 as well).

- While the observed temperature at the inland station of Kilpisjärvi was below -30°C several times during the final 12 days of January 2019, the preoperational AROME-Arctic and the reference forecasts with XRIMAX = 0 were seldom below -25°C and on average were 4.6°C too warm.
- Both XRIMAX = 0.2 and XRIMAX = 0.5 led to significant improvements to the temperature forecasts at Kilpisjärvi during this time period, with XRIMAX = 0.5 performing better.
- XRIMAX > 0 also has an effect on near-surface wind speeds, leading to an increase in 10 m wind speed. At Kilpisjärvi the overestimation during weak wind situations worsened, but we also saw increased wind speeds in situations when the operationally forecasted wind speed was too low.
- While most inland stations benefit from XRIMAX > 0, there are many examples of degradations at coastal stations. The surface temperatures at the coast are influenced by the Sea Surface Temperature, sometimes more in reality than modelled. At Torsvåg, a lighthouse on the Northern coast of Norway, the forecasts with XRIMAX = 0 were already too cold. XRIMAX = 0.2 and XRIMAX = 0.5 led to much lower temperatures in some situations, on average about 3°C lower than observed.
- XRIMAX > 0 has an effect on near surface wind speeds, also at coastal stations, but less so than at inland stations.

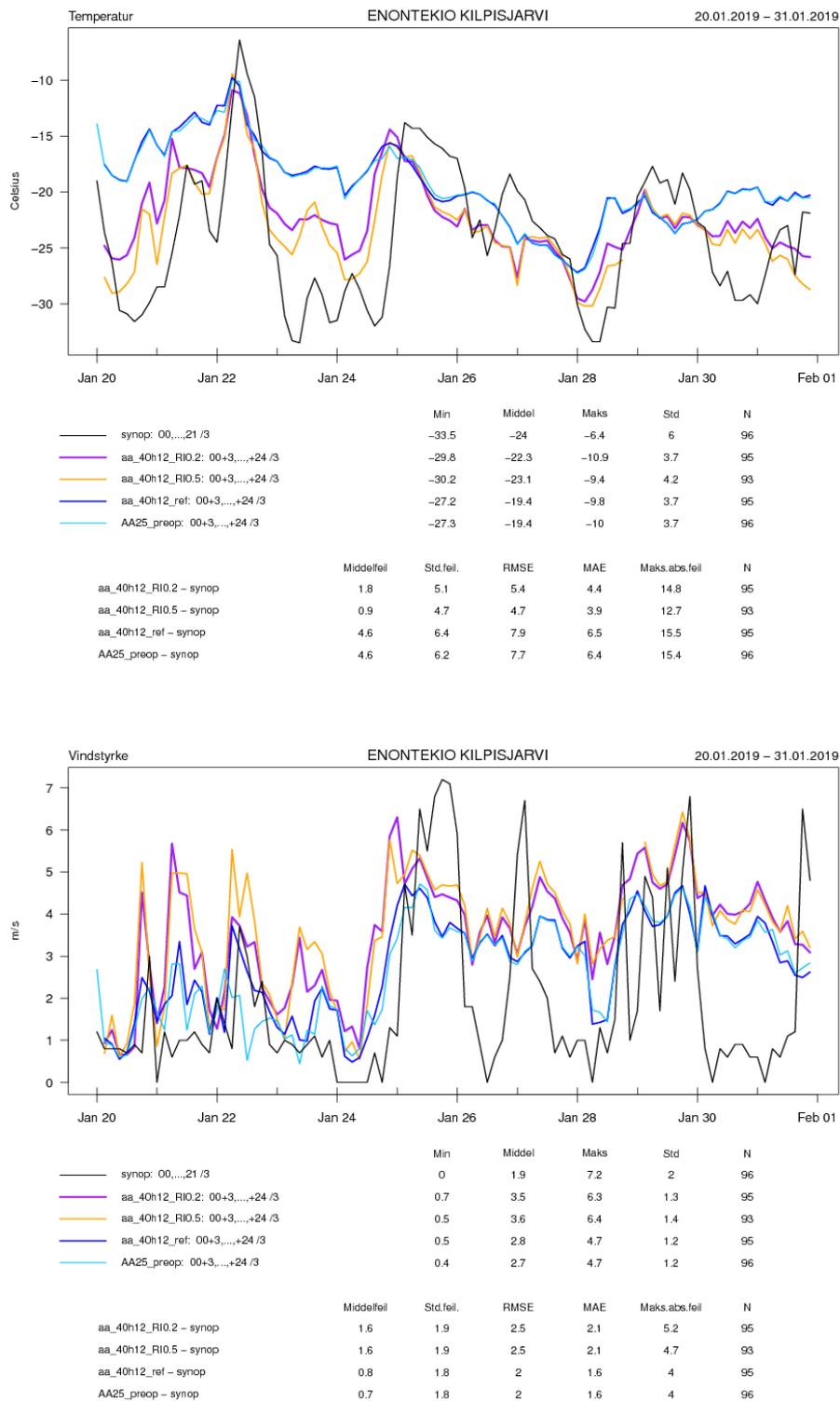


Figure 1a: Time series of 2 m temperatures (upper) and 10 m wind speed (lower); observations (black), forecasts: AROME-Arctic (cyan), REF (blue), EXP with XRIMAX = 0.2 (purple) and EXP with XRIMAX = 0.5 (orange) at Kilpisjärvi at the border between Finland, Norway and Sweden.

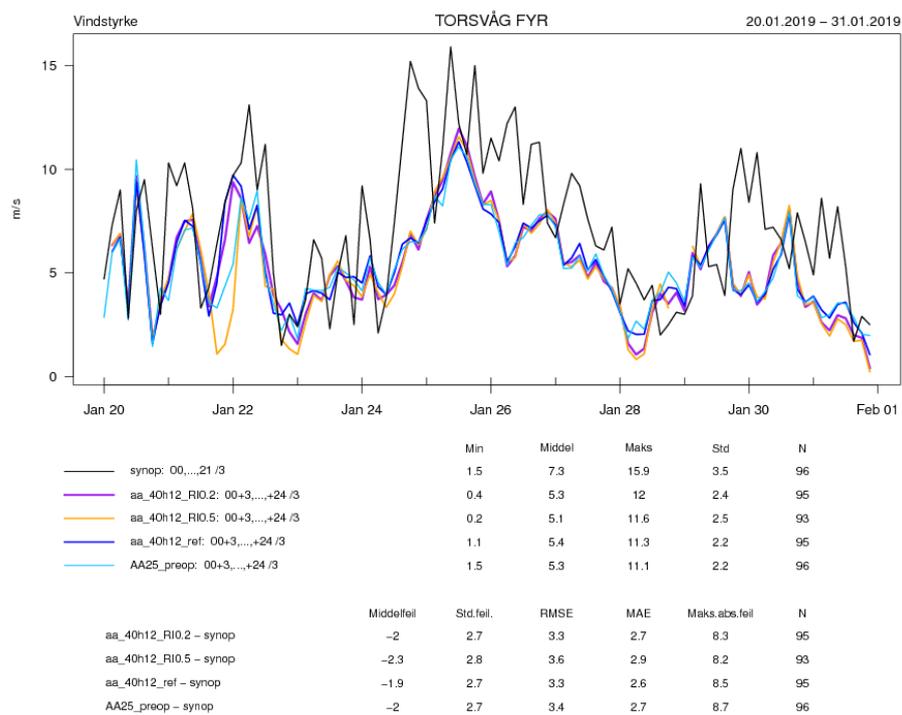
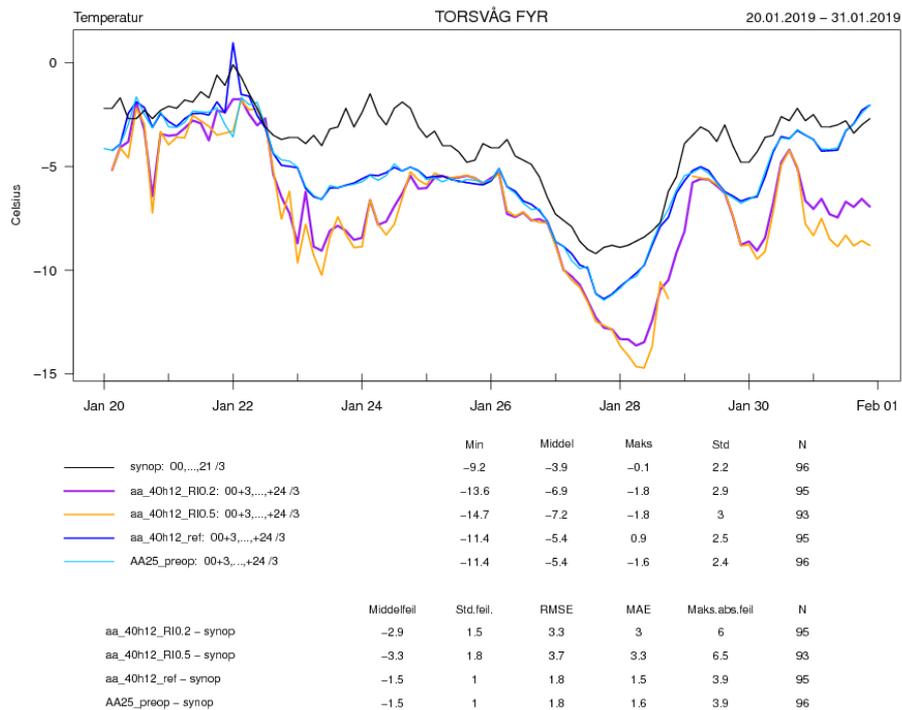


Figure 1b: Time series of 2 m temperature (upper) and 10 m wind speed (lower); observations (black) and forecasts; AROME-Arctic (cyan), REF (blue), EXP with XRIMAX = 0.2 (purple) and EXP with XRIMAX = 0.5 (orange) at Torsvåg lighthouse.

Summary statistics confirm the results demonstrated in the time series above. AROME-Arctic and REF with $XRIMAX = 0$ have a clear positive bias in 2 m temperature forecasts averaged over 16 inland stations in the Northern part of Norway, Fig 2, left. The bias is significantly reduced with $XRIMAX = 0.2$ and close to 0 with $XRIMAX = 0.5$. The quality is also improved, measured by SDE.

The problematic performance demonstrated at Torsvåg is typical in the coastal region. With $XRIMAX = 0$ the performance is quite good, but $XRIMAX > 0$ leads to significant degradation in 2 m temperature forecasts averaged over 33 coastal stations in Northern Norway, Fig 2, middle. The experiment with $XRIMAX = 0.5$ has a negative ME that increases with lead time. The quality is also degraded, measured by SDE.

When averaging over all 80 stations in the Northern part of Norway for those 12 "cold winter days" at the end of January 2019, we also see some degradation in summary scores for 2 m temperature in experiments with $XRIMAX > 0$, Fig 2, right.

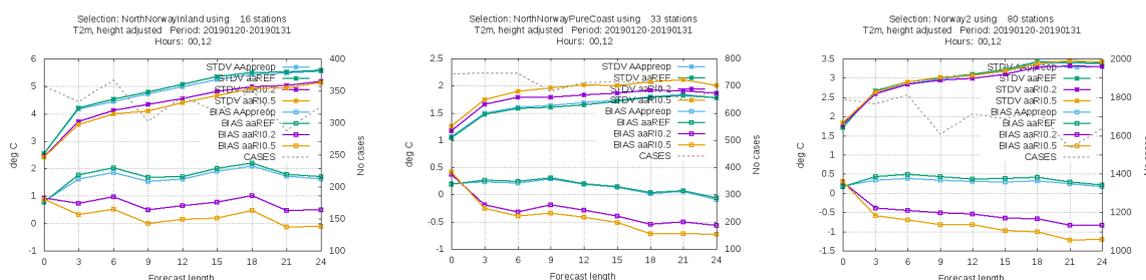


Figure 2: Mean Error (ME) and Standard Deviation of Error (SDE) of 2 m temperature forecasts as a function of lead time: AROME-Arctic (cyan), REF (green), EXP with $XRIMAX = 0.2$ (purple) and EXP with $XRIMAX = 0.5$ (orange) averaged over 16 inland (left), 33 coastal (middle) and 80 (right) stations in the Northern part of Norway for 20 - 31 January 2019.

We searched for a pragmatic solution to "do better" than $XRIMAX = 0$ in stable situations, but first some reflections:

- MétéoFrance runs AROME-France with $XRIMAX = 0.2$, but at the finer resolution of 1.3 km grid spacing in the horizontal, and 90 vertical levels with the lowest level 5 m above the ground. AROME-Arctic and MetCoOp ensemble system (MEPS) runs with 2.5 km grid spacing in the horizontal and 65 vertical levels with the lowest one at ~12 m above the ground. Would the sensitivity to positive $XRIMAX$ be less at finer resolution? A few experiments at finer vertical resolution indicate lower sensitivity (not shown here).
- The largest improvements are seen in stable situations, with $Ri \gg 0$, and with clear skies in reality and in the model.
- The effect on 2 m temperatures of missing clouds in the model or too small impact of e.g. thin ice clouds on the radiation seemed to be amplified with $XRIMAX > 0$.
- At coastal stations, some of the unrealistic temperature drops were associated with the weakly stable or close to neutral conditions.

3. Dependency of fluxes on the Richardson number

A few equations from the SURFEX scientific documentation (Le Moigne, 2018) are given here to explain the dependency of the surface fluxes of sensible heat and momentum on surface roughness and the stability (estimated by the Richardson number) through the drag coefficients.

The Richardson number, Ri, is defined as the ratio between potential and kinetic energy of the surface layer:

$$Ri = \frac{gz}{U^2} \left(\frac{\theta_{v_a} - \theta_{v_s}}{\bar{\theta}_v} \right) \quad (2.24)$$

where g is the gravity, U is the wind speed at height z, θ_{v_a} , θ_{v_s} and $\bar{\theta}_v$ are virtual potential temperatures at height z, at the surface and mean between surface and height z.

The sensible heat flux is proportional to the temperature difference between the surface and the atmosphere:

$$H = \rho_a c_p C_H V_a (T_s - T_a) \quad (4.183)$$

where c_p is the specific heat, ρ_a , V_a and T_a are air density, wind speed and temperature at the lowest model level, T_s is the surface temperature and C_H is the drag coefficient for heat.

The drag coefficients for momentum C_D and heat C_H are given by

$$C_D = C_{DN} F_m \quad (4.208)$$

$$C_H = C_{DN} F_h \quad (4.209)$$

with

$$C_{DN} = \frac{k^2}{[\ln(z/z_0)]^2} \quad (4.210)$$

where k is the Von Karman constant, z_0 is the roughness length for momentum and z is the height. z_{0h} is the roughness length for heat. The relationship $z_{0h} : z_0$ is typically 1:10.

$$F_m = 1 - \frac{10Ri}{1 + C_m \sqrt{|Ri|}} \quad \text{if } Ri \leq 0 \quad (4.211)$$

$$F_m = \frac{1}{1 + \frac{10Ri}{\sqrt{1+5Ri}}} \quad \text{if } Ri > 0 \quad (4.212)$$

$$F_h = \left[1 - \frac{15Ri}{1 + C_h \sqrt{|Ri|}} \right] \times \left[\frac{\ln(z/z_0)}{\ln(z/z_{0h})} \right] \quad \text{if } Ri \leq 0 \quad (4.213)$$

$$F_h = \frac{1}{1 + 15Ri \sqrt{1+5Ri}} \times \left[\frac{\ln(z/z_0)}{\ln(z/z_{0h})} \right] \quad \text{if } Ri > 0 \quad (4.214)$$

The drag coefficient for evaporation differs from the drag coefficient for heat only over the sea.

XRIMAX is the upper limit of Ri - the estimated Richardson number - in the drag calculations. In the code: if $Ri > XRIMAX$, $Ri = XRIMAX$. $XRIMAX = 0$ implies "neutral" drag coefficients are to also be used in situations which are stable according to the estimated Ri.

Examples of drag coefficients for momentum and heat as a function of Ri are shown in Fig. 3, for some selected values of roughness. It is worth noting the high sensitivity of the drag coefficients to Ri for values of Ri close to 0. Already small positive values of Ri lead to significant reductions in the drag coefficients, and thereby reductions in the surface fluxes for heat, moisture and momentum.

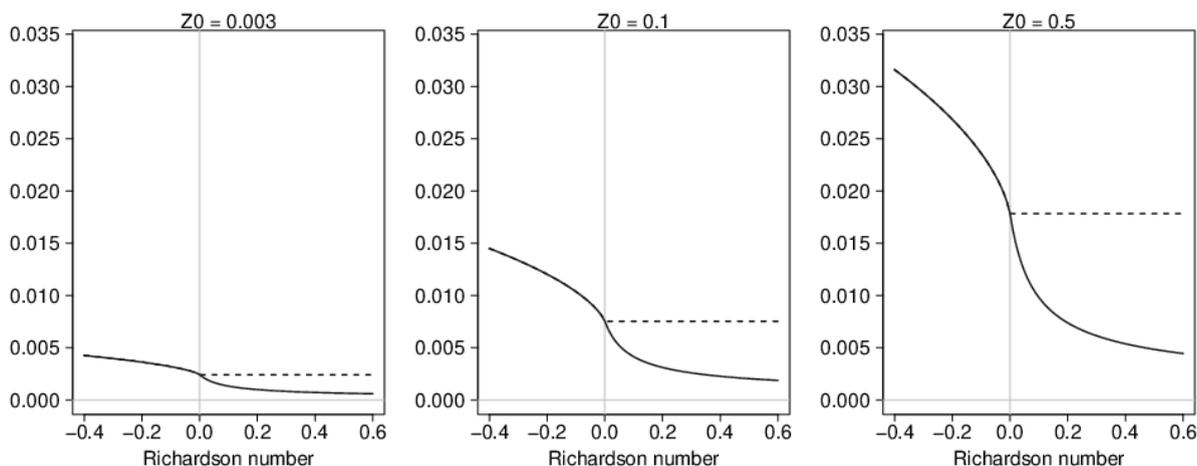


Figure 3a: Drag coefficient for momentum as a function of Richardson number (Ri) for $z_0 = 0.003$ m (left), $z_0 = 0.1$ m (middle) and $z_0 = 0.5$ m (right). The dashed lines indicate how the values under neutral conditions ($Ri = 0$) are used for $Ri > 0$ when $XRIMAX = 0$.

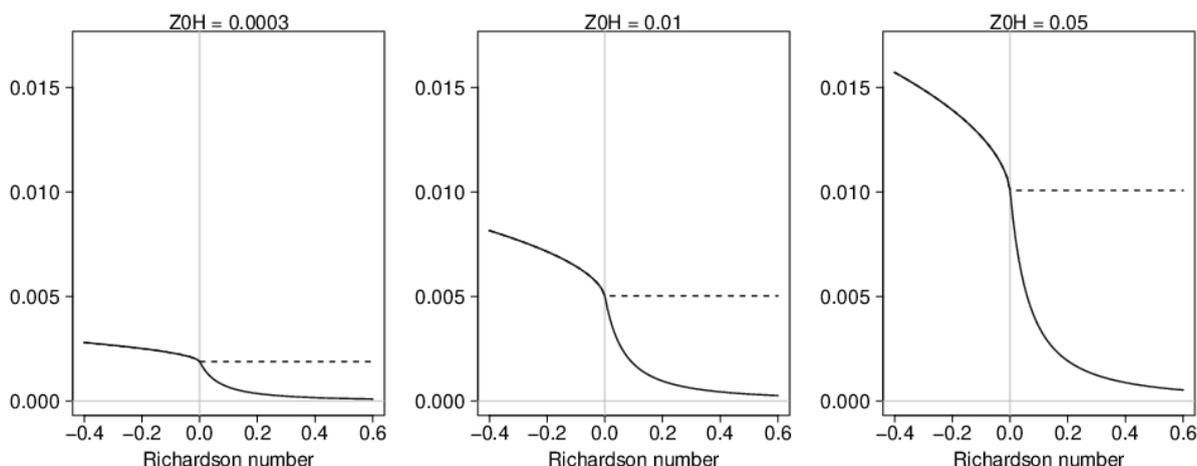


Figure 3b: Drag coefficient for heat as a function of Richardson number (Ri) for $z_{0h} = 0.0003$ m (left), $z_{0h} = 0.01$ m (middle) and $z_{0h} = 0.05$ m (right). The dashed lines indicate how the values under neutral conditions ($Ri = 0$) are used for $Ri > 0$ when $XRIMAX = 0$.

4. Pragmatic solutions to "do better" in stable situations

First, we tried to moderate the rapid decrease in the drag coefficients as a function of Ri in stable situations ($Ri > 0$) by modifying some coefficients in eqs. 4.212 and 4.214, to obtain drag coefficients between the ones obtained default with $XRIMAX > 0$ and the ones for neutral conditions (with $XRIMAX = 0$). As an example, the effects of modifying the coefficient with the value **15** in eq. 4.214, (from **15** to **5** or **3**) are shown for $z_{0h} = 0.0003$ m in Fig. 4. Corresponding modifications were also suggested for the drag coefficients calculations for momentum. Experiments with various combinations of coefficients showed that it was possible to reduce the sensitivity of the drag coefficients to the estimated Ri, and thereby reduce the negative results in some situations, but this is difficult without also reducing the positive results in other situations.

Motivated by the fact that the deterioration occurred mainly in situations with relatively small values of Ri, we tried another pragmatic solution by letting the drag coefficients remain "neutral" for small positive values of Ri. We suggest following the default calculations shifted horizontally by XRISHIFT for $Ri > XRISHIFT$, and neutral drag coefficients for $0 > Ri > XRISHIFT$. This is illustrated in Fig. 4, where $XRISHIFT = 0.1$.

A minor change to reduce the sensitivity to the wind speed in the drag calculations was also introduced, by increasing the threshold of the wind speed in the calculations from 1 m/s to 2 m/s, by increasing XVMODFAC from 0.1 to 0.2.

The effect of the suggested set-up was evaluated in experiments on the METCOOPB domain for March 2019 and August 2020. The corresponding time series for Kilpisjärvi and Torsvåg for 1 - 10 March 2019 are shown in Fig. 5.

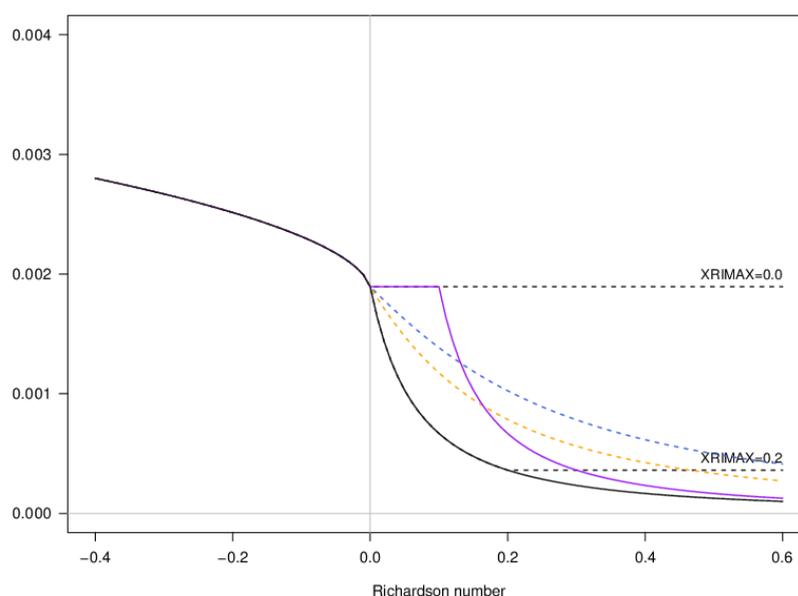


Figure 4: Drag coefficient for heat as a function of Richardson number for $z_{0h} = 0.0003$ m (roughness of snow). Black dashed lines show values used when $XRIMAX = 0$ and $XRIMAX = 0.2$. Blue and orange dashed lines are obtained by modifying a coefficient in the drag calculations in Eq. (4.214) from **15** to **5** (orange) or **3** (blue). The purple line illustrates the effect of setting $XRISHIFT = 0.1$.

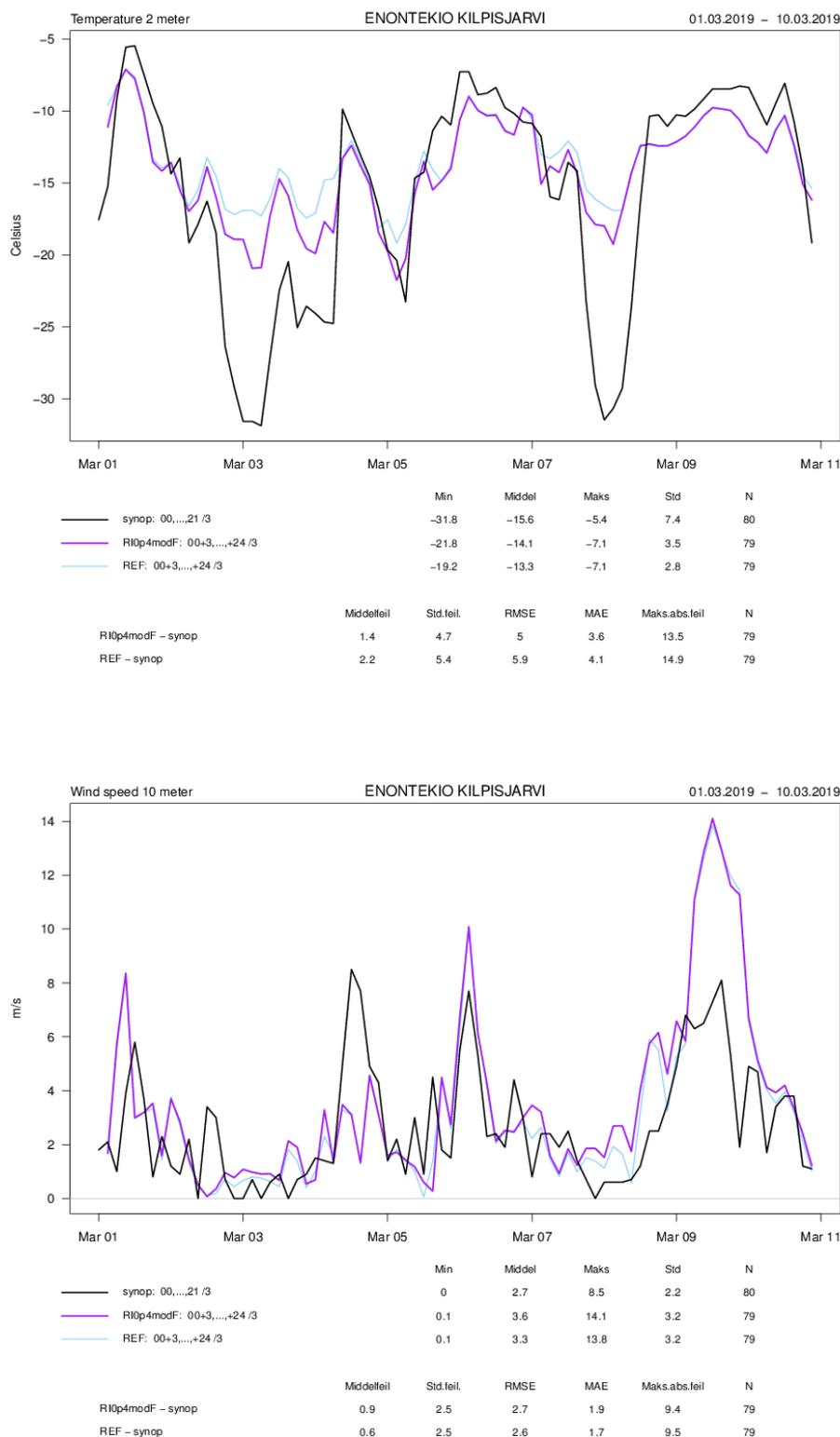


Figure 5a: Time series of 2 m temperature (upper) and 10 m wind speed (lower); observations (black), forecasts: REF (cyan), EXP with XRIMAX = 0.4, XRISHIFT = 0.1 and XVMODFAC = 0.2 (purple) from Kilpisjärvi at the border between Finland, Norway and Sweden.

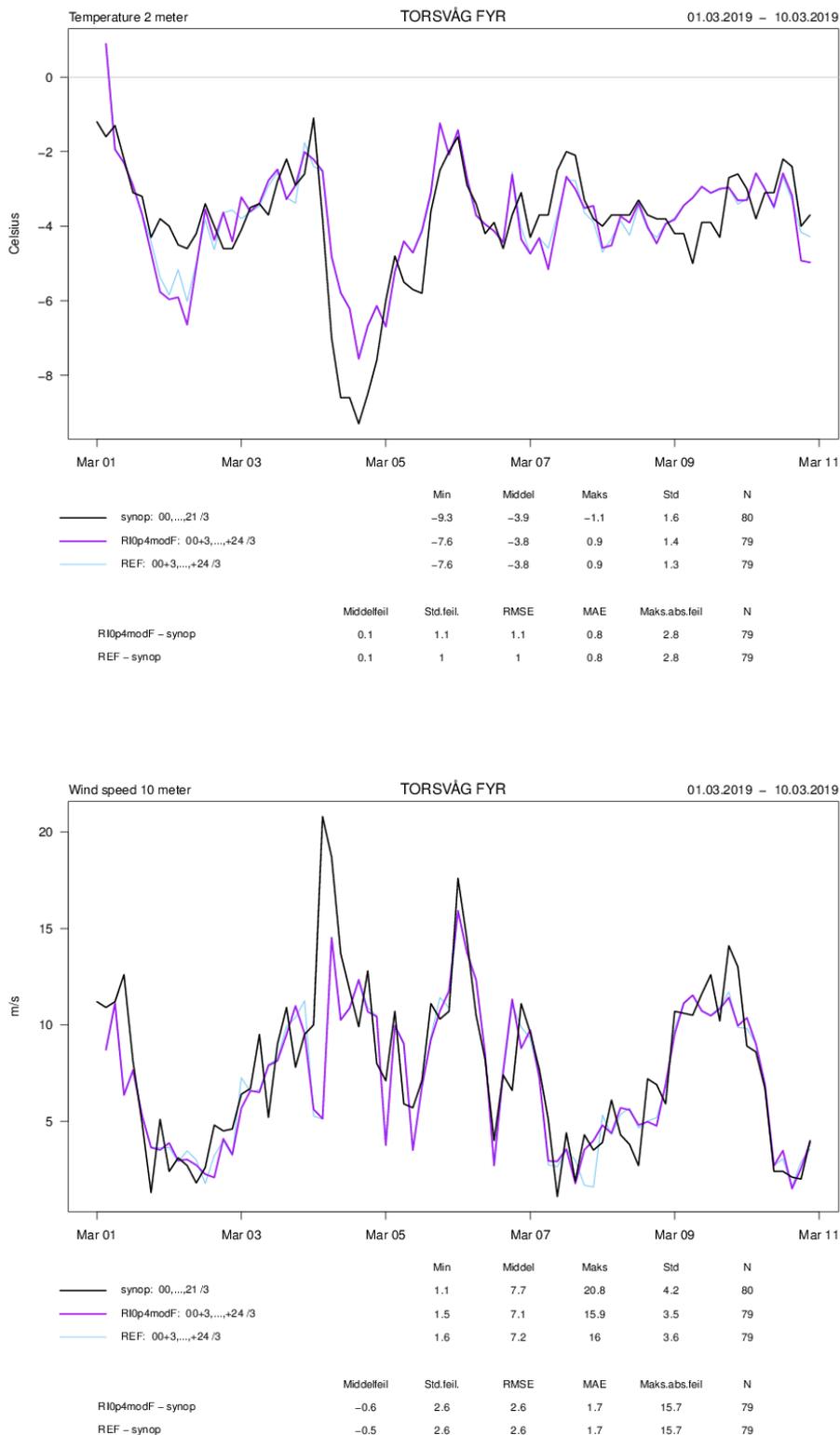


Figure 5b: Time series of 2 m temperature (upper) and 10 m wind speed (lower); observations (black), forecasts: REF (cyan), EXP with XRIMAX = 0.4, XRISHIFT = 0.1 and XVMODFAC = 0.2 (purple) from Torsvåg lighthouse.

The effect of $XRIMAX = 0.4$ and $XRISHIFT = 0.1$ is less than the effect of $XRIMAX = 0.2$, as expected. We see small improvements under stable situations, and less degradation in situations close to neutral stability. But still, the degradation in some situations is large: on average, the setup with $XRIMAX = 0.4$, $XRISHIFT = 0.1$ and $XVMODFAC = 0.2$, leads to an undesired reduction of the 2 m temperatures, particularly in winter, see summary results for March 2019 in Fig. 6.

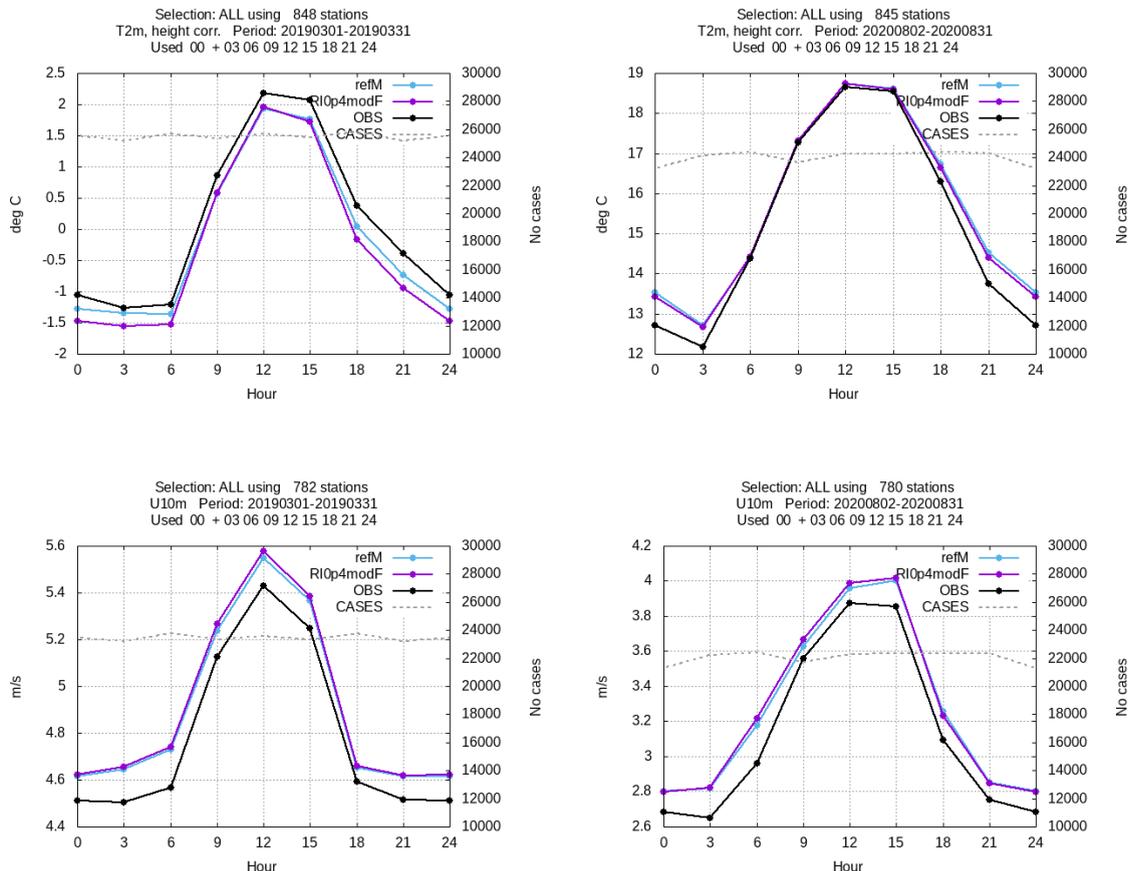


Figure 6: Daily variation in 2 m temperature (upper) and 10 m wind speed (lower) observations (black) and forecasts: REF (cyan) and EXP with $XRIMAX = 0.4$ and $XRISHIFT = 0.1$ (purple) averaged over ~800 stations 1 - 31 March 2019 (left) and 2 - 31 August 2020 (right).

Encouraged by the improvements under stable situations, and despite the overall reduction in temperature in winter, the setup with $XRIMAX = 0.4$, $XRISHIFT = 0.1$ and $XVMODFAC = 0.2$ was introduced in MEPS_preop_cy43 from 4 November 2020. MEPS was upgraded to cycle 43 in March 2021.

5. Discussion and summary

To improve the performance in the stable boundary layer, several experiments with the limitation of the Richardson number, $XRIMAX$, > 0 have been run with the HARMONIE-AROME canonical configuration of the ALADIN-HIRLAM NWP system. The first experiments with $XRIMAX > 0$ (e.g. 0.2 or 0.5) showed the potential of improved performance in cold and stable situations, at inland stations, but also showed many examples of unrealistic temperature drops, at coastal stations in the north but also at inland stations, particularly in mountainous regions with low roughness, and when snow was on ground. The effect on 2 m temperatures of missing clouds in the model or too small impact of the clouds on the radiation seemed to be amplified with $XRIMAX > 0$. At coastal stations, some of these unrealistic temperature drops were associated with the weakly stable (or, close to neutral) boundary layer.

AROME-France is run with $XRIMAX = 0.2$. Are the problems with $XRIMAX = 0.2$ in MEPS on the MetCoOp domain related to the fact that AROME-France runs at a finer resolution, 1.3 km grid spacing in the horizontal, and 90 vertical levels with the lowest level 5 m above the ground? MEPS runs with 2.5 km grid spacing in the horizontal and 65 vertical levels with the lowest one at ~ 12 m. Would the sensitivity to positive Ri be less at finer resolution? A few experiments with finer vertical resolution indicated less sensitivity (not shown here).

Also, the largest problems occur in the winter season, when there is snow on ground, and they seem to increase for high latitudes. $XRIMAX = 0.2$ was also tested for use in CARRA, but considered to result in more problems than benefits.

The potential improvements seen with $XRIMAX > 0$ and the problems with continuing with $XRIMAX = 0$, inspired us to search for pragmatic solutions to improve the performance in very stable situations by letting $XRIMAX$ exceed 0, without such serious degradation in less stable situations. The pragmatic solutions were sought in modifications to the calculation of the drag coefficients as functions of the Richardson number and the surface roughness. The final set-up introduced with the update of MEPS to cycle 43 in March 2021 is presented here. We see clear improvements in some situations, but still see some problems. On average, the set-up with $XRIMAX = 0.4$, $XRIHIFT = 0.1$ and $XVMODFAC = 0.2$, led to a significant reduction of the 2 m temperatures.

6. References

- Lisa Bengtsson et al.: The HARMONIE-AROME Model Configuration in the ALADIN-HIRLAM NWP System. *Mon. Wea. Rev.*, **145**, 1919-1935, <https://doi.org/10.1175/MWR-D-16-0417.1>, 2017.
- Donier, S. Y., S. Faroux, and V. Masson: Evaluation of the impact of the use of the ECOCLIMAP2 database on AROME operational forecasts. Météo-France Tech. Rep., 89 pp. Available online at http://www.umr-cnrm.fr/surfex/IMG/pdf/test_eco2_arome.pdf, 2012.
- Masson, V. et al.: The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes. *Geosci. Model Dev.*, **6**, 929-960, doi:10.5194/gmd-6-929-2013, 2013.
- Malte Müller et al.: AROME-MetCoOp: A Nordic convective-scale operational weather prediction model. *Wea. Forecasting*, **32**, 609-627, doi:10.1175/WAF-D-16-0099, 2017.
- Patrick Le Moigne: SURFEX V8.1 main scientific documentation. Available at umr-cnrm.fr/surfex/, 2018.

7. Acknowledgements

To Emily Gleeson and Ekaterina Kurzeneva for encouraging the efforts to summarise these results.

To Patrick Samuelsson, Danijel Belušić and the ACCORD and HIRLAM Surface Teams for inspiring discussions.

To the HIRLAM System for providing nice surroundings for HARMONIE-AROME experiments.

MUSC working week and ACCORD visits Helsinki, November 2021

Laura Rontu, Eric Bazile, Piotr Sekula, Ana Šljivić, Wim de Rooy

1 Working week

A working week on MUSC, the single-column ACCORD model, was arranged in Helsinki 15-19 November 2021. The aim of the workshop was to build a common platform for MUSC (Modèle Unifié, Simple Colonne) use across the consortium. This means, in addition to the common source code, ready-made test experiments with namelists, initial and forcing data as well as tools for preparing input and analysing the output.

Experiences of MUSC use for model-observation intercomparison experiments, code development and testbed applications were shared. The meeting was held in hybrid form, built on the agenda and materials on the wiki page <https://hirlam.org/trac/wiki/Meetings/Physics/MUSCWW21>. Presentations and results of the working week are available via this wiki page.

Participants from Croatia (Martina Tudor, Ana Šljivić), Finland (Laura Rontu), France (Eric Bazile), Ireland (Eoin Whelan), Norway (Teresa Valkonen), Poland (Piotr Sekula) were present in Helsinki. More participants attended the online sessions also from Iceland and Spain. After 20 months of remote work during the pandemic, the possibility of face-to-face communication in the international team of colleagues was welcomed by the participants.

The starting point of the working week was the ARMCu experiment without SURFEX (à la mitraillette) using prescribed surface fluxes. The task was to merge the modifications done in Toulouse on cy46t1_op1_MUSC and the CY46T1_bf.06++, used by Harmonie for CY46H1. Before having a truly unique code for the 3CSC based on cy46t1, a cy46h_MUSC was created with some updated routines from cy46t1_op1_MUSC. For ALARO a ARMCu namelist was created to be used with cy46t1_op1_MUSC. However, in parallel, a merge between the modified physics routines used by Harmonie-Arome in cy46h was done with cy46t1_op1_MUSC and successfully tested on ARMCu without SURFEX. This is the first prototype for the common code. However, a more ideal case, such as FIRE and Astex-Lag, must be validated for the 3CSC before suggesting modifications to the common cycle.

The graphic tool Atlas, developed by Romain Roehrig for automatic comparison with LES, was successfully installed and used by Eoin Whelan. This tool was also used to create Figure 1. It shows the cloud fraction of the different setups of the ARMCu experiments, built during the working week. Cloud fraction is one possible variable for evaluation of the model performance. Heat and momentum fluxes in the boundary layer are more important for the model quality, but look less illustrative.

The first plot on the upper row is the reference from the Meso-NH LES. The next two figures show AROME output on land and sea points. Note that land-sea mask can only change the results in case parametrizations (such as microphysics) use it to modify aerosols or the cloud condensation nuclei. Among the 8 experiments, the AROME results are closest to the reference LES.

In the middle row, results of ALARO experiments without SURFEX are shown: reference (left), experiments B (middle) and B2 (right). They were run by Martina Tudor and Bogdan Bochenek, using namelists developed during the workshop. All three ALARO experiments show somewhat different results, whose analysis is ongoing.

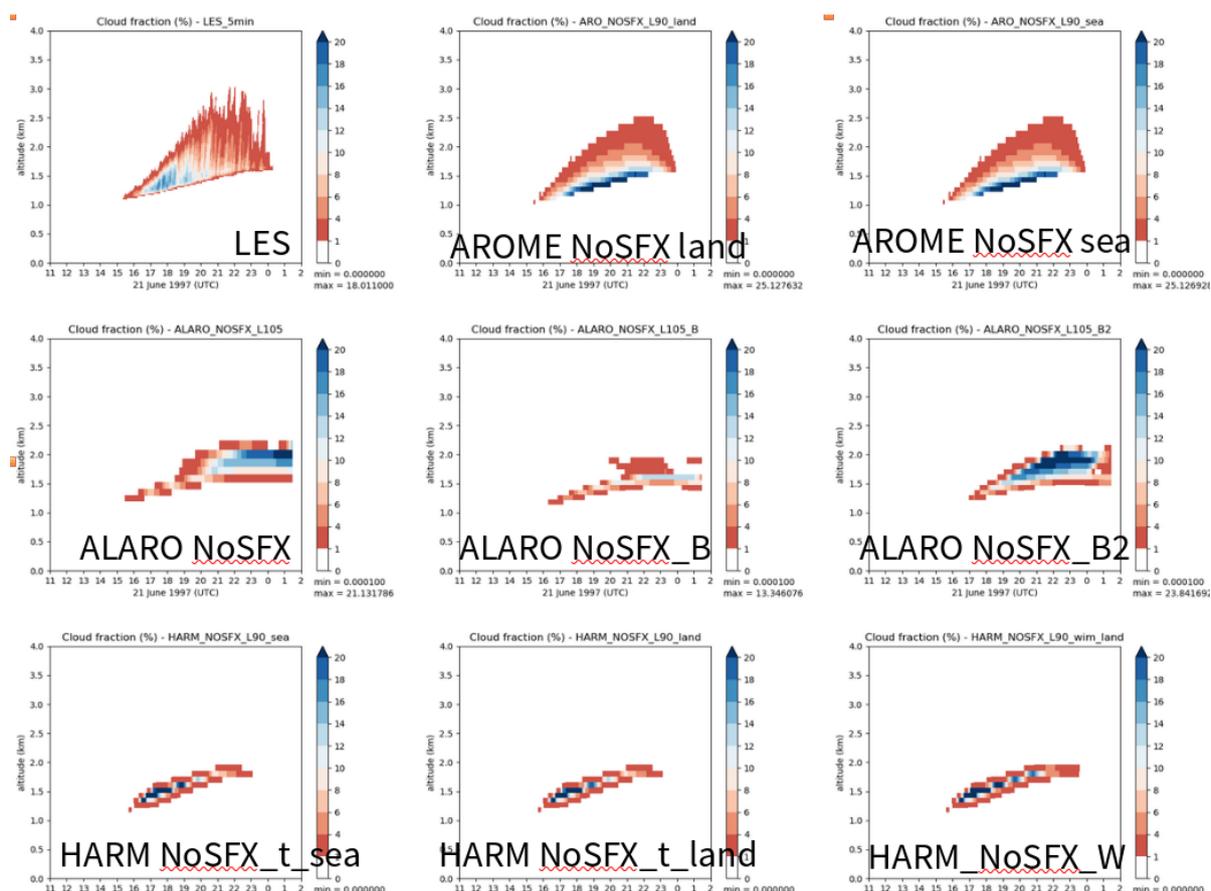


Figure 1: Cloud fraction [%] for the ARMCU case of 21 June 1997. x-axis: time, y-axis: height [km].

The bottom row shows Harmonie-Arome comparison: the first two plots are obtained using the prototype of the common MUSC code (merge of t- and h-branches, compiled with gmckpack, run by Eric Bazile using namelist from 46h1) at land and sea points. The third plot represents the cy46h_MUSC result obtained by Wim de Rooy. Cloud fraction results in the bottom row are sufficiently similar, which suggests that the merge of 46h and 46t forecast model codes was done correctly. The Harmonie-Arome result is also consistent with the results by de Rooy et al. 2022, where a detailed discussion can be found.

To summarize: the working week was really efficient and a success of collaborative work for the common MUSC.

2 ACCORD scientific visits

Development work in the framework of the ACCORD rolling workplan task PH6 on cloud-aerosol-radiation interactions was continued, using MUSC as tool for updating and testing new code. Ana Šljivić (18 November - 2 December) focused on importing aerosol updates by Ján Mašek for ACRANE2 radiation scheme to cy43h2. Piotr Sekula (18 November - 9 December) analyzed the sensitivity of cloud-precipitation microphysics to the introduction of near-real-time aerosol data. This was done in the framework of HARMONIE-AROME cy46h1 updated by Daniel Martín.

MUSC experiments were set up for two cases at a location of Harmaja lighthouse (60.109 N, 24.951 E), off the Southern coast of Finland close to Helsinki. A case of the 23rd of February 2021 represents Saharan dust

intrusion in winter conditions. A case of the 19th of April 2021 represents clear and clean conditions. The first case is interesting for sensitivity studies related to cloud-aerosol microphysics developments while the second offers a possibility to study aerosol impact on radiation, treated by different radiation schemes. Initial data for MUSC experiments were picked from full HARMONIE-AROME experiments that were performed to provide atmospheric and surface state as well as near-real-time and climatological aerosol input. The aim of the experiments was to test sensitivities and develop the parametrizations, not a model-observation intercomparison. Experiments were run within development versions of cy43h2 (cy43_rontu_aerad by L. Rontu) and cy46h1 (pre-CY46h1_nrtar by D. Martin).

During the visits, MUSC experiments were set up in various environments, input and namelists prepared, test runs done and new tools for analysing the results built and used. For technical reasons, in some local environments/compiler the setups remained unsuccessful. Detailed reports of the results and work done will be provided elsewhere, a few highlights are shown below (Sections 2.1 and 2.2).

2.1 Testing of aerosol parametrizations in ACRANEB2 scheme

The aim was to include modifications made by Ján Mašek (cy46t1) to cy43h2 and later to cy46h1. Development work was done within the branch cy43_rontu_aerad (from <https://git.hirlam.org/users/rontu/Harmonie>) in two experiments: aerad43h2 and aerjan43h2. In the first, the reference branch settings were used, including unmodified ACRANEB2 (Geleyn et al., 2017, Mašek et al., 2016). In the second, modified ACRANEB2 and a new subroutine acraneb_aer_550.F90 were included together with other needed modifications.

MUSC experiments were run for the clean clear-sky case of the 19th of April 2021. Four sources of aerosols were tried:

- az - zero aerosols
- at - Tegen climatological aerosol optical depths
- am - climatological mass mixing ratios (MMRs)
- an - near real time MMRs.

ACRANEB2 modset contains a new subroutine acraneb_aer_550.F90 (provided by Ján Mašek), which converts aerosol load given as layer optical depth at 550 nm (ZAOD) to aerosol SW and LW broadband optical properties (layer optical depth ZATAU, single scattering albedo ZASSA and asymmetry factor ZAASY) of the mixture. The main difference between original and modified experiments is that in the original experiments input to ACRANEB2 is ZAOD, and in the modified ones input to ACRANEB2 are ZATAU, ZAASY and ZASSA.

In case when Tegen aerosols are on, ZAOD is obtained using radaer.F90, which distributes the original Tegen aerosol optical depth (2D field) to the model levels using an exponential profile. For the modified ACRANEB2, ZATAU, ZAASY and ZASSA can be made from ZAOD by acraneb_aer_550.F90.

In the experiments with MMRs (am and an), subroutine aeroport.F90 calculates aerosol optical properties based on MMRs, which arrive from m-climate files (2D) or from Copernicus Atmosphere Monitoring Service (CAMS) via boundaries (3D). Aerosol inherent optical properties from ECMWF (Bozzo et al., 2020, Rontu et al., 2020) are used. In the original settings, aeroport.F90 diagnoses ZAOD from MMRs (climatological or near-real-time) for schemes that use aerosol optical depths as input (original ACRANEB2 or the default IFS radiation scheme). In the advanced way, aeroport.F90 calculates ZATAU, ZAASY and ZASSA of the aerosol mixture for the modified ACRANEB2, which is capable of using these fields directly.

There is no difference in the radiation for experiments with zero and Tegen aerosols (not shown). The difference between the old and the advanced way of using the radiation scheme is evident only for experiments with MMRs (Figure 2). For the net shortwave radiation fluxes the differences are always negative. This means that the new aerjan43h2 experiments have more net SW radiation than aerad43h2 experiments. Differences are greater for

the near-real-time MMR case. The difference is zero after the sunset ($\approx 11^{th}$ forecast hour, initial time is 6 UTC).

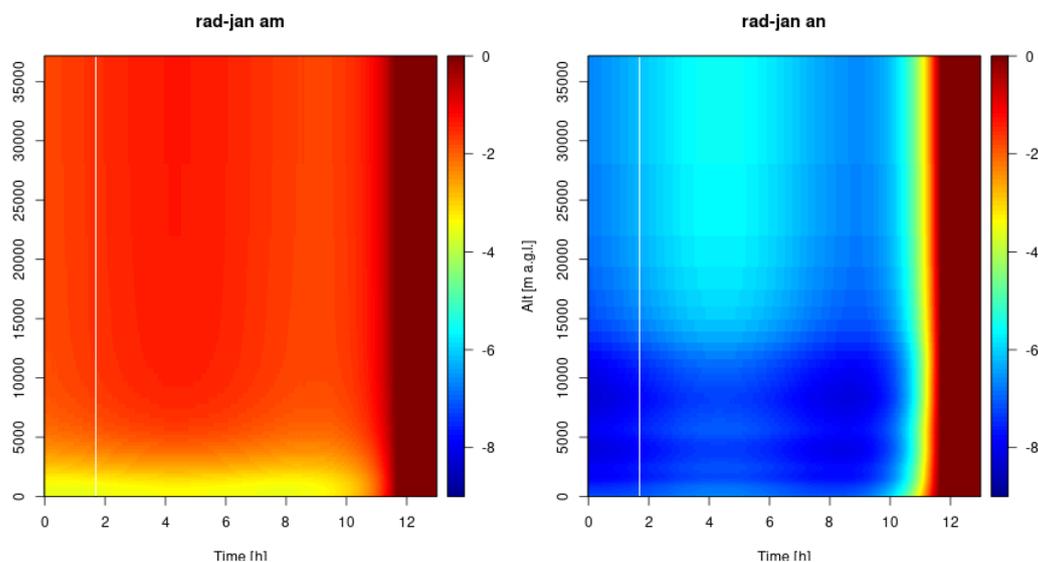


Figure 2: Differences in shortwave radiation between aerad43h2 and aerjan43h2 for climatological MMRs (left) and for near-real-time MMRs (right)

2.2 Testing aerosol impact on cloud-precipitation microphysics

The sensitivity tests of cloud-precipitation microphysics were done for the HARMONIE-AROME model version cy46h1, updated by Daniel Martín. Atmospheric model used radiation scheme from IFS cycle 25 (Foucault-Morcrette). and near-real time aerosol data. A case of the 23rd of February 2021, which represents Saharan dust intrusion, was studied in detail due to the significant role of the dust in ice nucleation within clouds. With the aim of analyzing aerosol impact on cloud-precipitation microphysics and radiation, a reference model configuration (provided by Daniel Martin) was prepared. 10 different configurations were added, in which only one or paired variables were changed in comparison to the reference configuration.

Analysis of the model output confirmed the presence of dust particles in the vertical profile used in the MUSC simulation. The highest value of mass mixing ratio (MMR) of all analyzed species, about $0.1 \mu\text{g}/\text{kg}$, was observed for dust particles with optical diameter in range from 0.9 to $20 \mu\text{m}$ at the altitude of 3000 m a.g.l. . The total amount of dust in the column was about $5 \text{ mg}/\text{m}^2$ that was a sixth part of the total estimated aerosol mass of $28 \text{ mg}/\text{m}^2$. The estimated MMR values for the atmospheric column over the Harmaja lighthouse were obtained from Copernicus Atmosphere Monitoring Service data (<https://apps.ecmwf.int/datasets/data/cams-nrealtime/levtype=ml/>) via a three-dimensional HARMONIE-AROME experiment.

The strongest differences in the radiation budget from the results from reference model were observed for two tested configurations. For the first configuration, the increase of supersaturation over 100 m height from default value $0.08\text{E-}2$ to $0.2\text{E-}2$ affected the increase the Cloud Droplet Number Concentration (Figure 3). In the second case, the activation of ice nuclei caused a sudden change of specific cloud ice during the first 30 minutes of the forecast. After half an hour of the forecast, the model stabilized (Figure 4). In both cases, the downwelling shortwave radiation flux (SWDN) at the surface reduced about 10% compared to the reference simulation. Note that in the cloudy, foggy and dusty winter-time conditions, the magnitude of the simulated SWDN was around $100 \text{ W}/\text{m}^2$ before noon. The significant increase of ice concentration after the activation of ice nuclei seems to be too strong, due to this fact further studies of this phenomenon are planned.

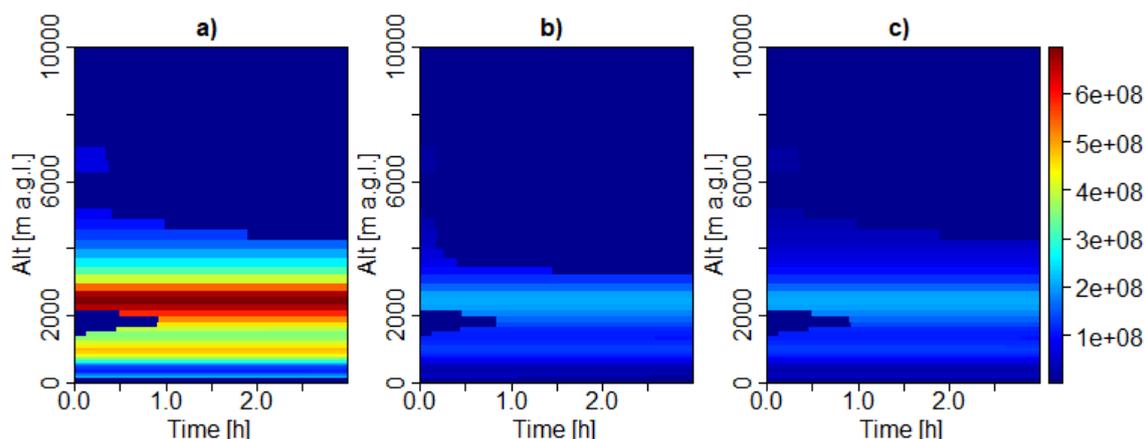


Figure 3: The Cloud Condensation Nuclei vertical profiles for a) configuration with increased value of supersaturation, b) activated ice nuclei processes and c) reference model configuration at 23 February 2021.

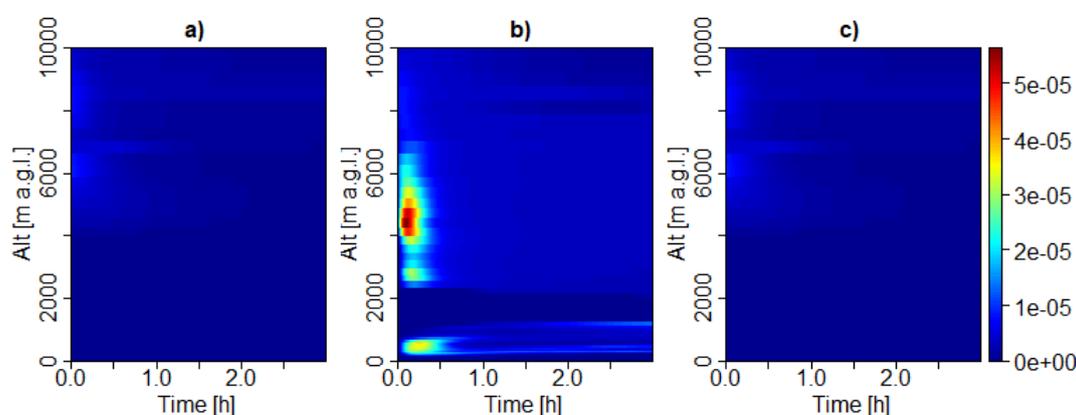


Figure 4: The specific cloud ice vertical profiles for a) configuration with increased value of supersaturation, b) activated ice nuclei processes and c) reference model configuration at 23 February 2021.

2.3 Comparisons between cy43 and cy46 experiments

Afterwards, the results from an experiment aerjan43h2 were compared with those from aerjan46h1. Both experiments included updated acraneb2 radiation scheme. Presently a more limited subset of aerosol-radiation parametrizations and aerosol sources is available within the cycle 46 than in cycle 43. Also, for example the surface parametrizations and surface description in these cycles differ. Initial atmospheric profiles were picked from corresponding preliminary full model experiments in cy43h2 and cy46h2, which makes them slightly different. Because of this, the results of these experiments are not strictly comparable but may give an idea of the magnitude of the possible differences. Figure 5 shows that the differences in downwelling short-wave radiation between experiments in cy43 and cy46 varied from -4 to 8 W/m^2 , i.e. about ± 1 to 2% of the downward shortwave flux at the surface during the first three hours (around midday from 9 to 12 UTC) of the MUSC experiments in the April clean clear-sky case. Differences in the downwelling longwave flux at surface were within less than 0.5 W/m^2 (not shown).

The experiments represented cases of no aerosol, default Tegen aerosol and near-real-time aerosol, that were used as input for the IFS radiation (default) and acraneb2 schemes. Additional experiments (aa and ae) were run within cy46. In aa, optical properties of aerosol mixture for updated acraneb2 were prepared and used in the same way as in experiment (an) in cy43 (see Section 2.1). In this case, the difference between cy43 and cy46 results was close to zero as in the case with no aerosol included (overlapping cyan and orange lines). Other differences between experiments remain to be studied in detail. In the February 2021 cloudy and dusty case

(results not shown), the corresponding differences between experiments were an order of magnitude greater. Significant differences in downwelling long-wave radiation were also found in this case.

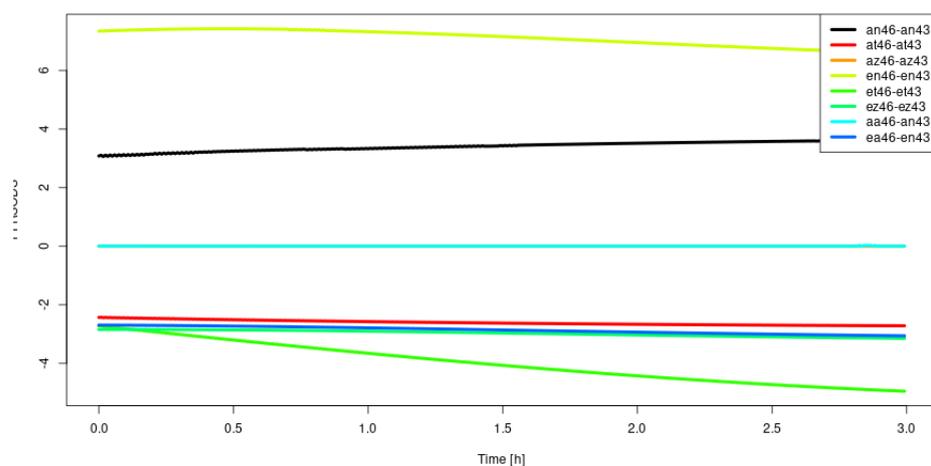


Figure 5: Differences of downwelling short-wave radiation at the surface level (unit W/m^2) between the cy46 and cy43 in experiments using near-real-time aerosol mass mixing ratio input for acraneb2 (an and aa) and IFS radiation schemes (en and ea), in experiments using Tegen AOD at 550 nm (at and et) and in experiments using no aerosol input (az and ez).

Acknowledgements

Thank you for contributing to the workshop and supporting the visits: Reima Eresmaa, Emily Gleeson, Daniel Martin, Ján Mašek, Yann Seity. Funding of the exchange visits was supported by the ACCORD management and by Krakow university for Piotr Sekula.

3 References

- Bozzo, A., Benedetti, A., Flemming, J., Kipling, Z. and Rémy, S.: An aerosol climatology for global models based on the tropospheric aerosol scheme in the Integrated Forecasting System of ECMWF. *Geoscientific Model Development*, **13**, 1007–1034, doi:10.5194/gmd-13-1007-2020, 2020.
- Geleyn, J.-F., Mašek, J., Brožková, R., Kuma, P., Degrauwe, D., Hello, G. and Pristov, N.: Single interval longwave radiation scheme based on the net exchanged rate decomposition with bracketing. *Q.J.R. Meteorol. Soc.*, **143**, 1313–1335, doi:10.1002/qj.3006, 2017.
- Mašek, J., Geleyn, J.-F., Brožková, R., Giot, O., Achom, H. O. and Kuma, P.: Single interval shortwave radiation scheme with parameterized optical saturation and spectral overlaps. *Q.J.R. Meteorol. Soc.*, **142**, 304–326, doi:10.1002/qj.2653, 2016.
- Rontu, L., Gleeson, E., Martin Perez, D., Pagh Nielsen, K. and Toll, V.: Sensitivity of Radiative Fluxes to Aerosols in the ALADIN-HIRLAM Numerical Weather Prediction System. *Atmosphere*, [205], doi:10.3390/atmos11020205, 2020.
- de Rooy, W.C., Siebesma, P., Baas, P., Lenderink, G., de Roode, S.R., de Vries, H., van Meijgaard, E., Meirink, J.F., Tijm, S. and van 't Veen, B.: Model development in practice: A comprehensive update to the boundary layer schemes in HARMONIE-AROME cycle 40. *Geosci. Model Dev.*, **15**, 1513–1543, doi:10.5194/gmd-15-1513-2022, 2022.

The Representation of Turbulent Kinetic Energy in HARMONIE-AROME at Hectometer Scale

Mirjam Tjihuis, Natalie Theeuwes, Jan Barkmeijer

1 Introduction

Continuously increasing computer power allows for numerical weather predictions at higher and higher resolution. With the computer capacity increasing even further, weather forecasts can be made to run at sub-kilometric scales. This makes forecasts at city scale possible, in order to improve predictions of urban heat, flooding and poor air quality. For example, Lean et al. (2019) used the UK Met Office unified model to show that fine-scale details can be modelled like parks and rivers being colder than their surroundings and Ronda et al. (2017) modelled the impact of large water bodies on urban heat at 100-m grid spacing. Furthermore, the wind field at night can be modelled in more detail, improving the representation of nocturnal processes like low-level jets, katabatic winds and fog formation (Kealy 2019). For example, Boutle et al. (2016) found that using high resolution (333m) can improve fog forecasts. Additionally, regions with complex orography can be modelled in more detail at higher resolutions (Yang 2019; Valkonen et al. 2020).

However, increasing the resolution also comes with challenges. With increasing resolution, we enter the so called grey zone of turbulence (Wyngaard 2004). The grey zone is the range of grid spacings where the length scales of the features are similar to the grid spacing. Moist convection takes place on scales around 1 to 10 km, turbulence in the convective boundary layer takes place on scales around 100-1000m. At sub-km scales, the grid spacing is too fine to have the whole range of sizes of e.g. turbulent eddies within the grid cell, which is an assumption of the current parameterisations. At the same time, the grid spacing is too large to resolve everything. To study the grey zone of turbulence Large-Eddy Simulation models can be employed (Kealy 2019). In this study we will use the Dutch Atmospheric Large-Eddy Simulation model (DALES, Heus et al. (2010)). In this grey zone there will be a resolution for which the resolved and subgrid turbulence are equally important and it is interesting to determine how a weather forecasting model handles this situation.

This work focuses on the main question: What is the potential of running HARMONIE-AROME for very high resolution (down to 50 m) for a specific fair-weather case over the Netherlands? To answer this question, three objectives are defined:

- To compare different resolutions and domain sizes;
- To investigate different set-ups for the convection and turbulence schemes;
- To compare the high-resolution HARMONIE-AROME with a Large-Eddy Simulation (DALES).

2 Methods

2.1 Case description

In this study, we will focus on the 14th of September 2020. We will focus on the area around Cabauw, the Netherlands. There was a high pressure system located over central Europe, resulting in a warm autumn day with maximum temperature of 28 °C measured in Cabauw. This day was chosen for two reasons. Firstly, it was a very calm day, with a clear convective boundary layer during the day. This makes this day very suitable to investigate how well convection and turbulence are resolved with increasing model resolution. Secondly, this day was part of a measurement campaign in the Ruisdael project¹. Therefore, additional observations (e.g., radiosondes) are available.

2.2 Model description and experimental setup

In this study we used HARMONIE-AROME cycle 43h2.1.1 with a default grid spacing of 2.5 x 2.5 km and 65 vertical levels (Bengtsson et al. 2017). The model uses hourly boundary conditions from the operational model of ECMWF (European Centre for Medium-range Weather Forecasts ²).

We want to investigate the performance of HARMONIE-AROME at sub-km scale with minimal adaptations to the default settings. For the best representation and for more complex cases, additional changes can be considered or might even be necessary to have a stable forecast (e.g. changing the spectral truncation, thus using a cubic or quadratic grid instead of a linear grid (Malardel et al. 2016)).

For the high-resolution experiments we use horizontal resolutions of 2500m, 500m, 100m and 50m and we use multiple nesting. This means that the simulation at 500m resolution receives its boundaries from the simulation at 2500m. The simulations at 100m receive their boundaries from the 500m simulation, etc. This nesting procedure makes sure that the boundaries match the new resolution as much as possible. Before starting these experiments, we checked the nesting procedure. To this end, two short experiments were performed. These experiments are 6h forecasts with a cold start, which means that they start from ECMWF fields. The first experiment has the default settings and domain NETHERLANDS, which is 800x800 grid points in size. The second experiment has a smaller domain of 400x400 grid points and is nested in the control experiment. This means that the experiment with the smaller domain gets its boundaries from the control experiment. Apart from this, the experiments are exactly the same. As expected, we found that both forecasts are the same, which means that the nesting procedure works properly.

For the main analysis, 24h forecasts are used, starting at 14 September, 00 UTC. These forecasts have a warm start, meaning that we used analyses that have been computed previously as initial conditions. This has the advantage that the 24h forecasts do not include any spin-up effects. An overview of the settings of the forecasts is given in Table. 1. Details about the implementation of the methods can be found in the appendix. When increasing the resolution, the timestep is lowered with the same factor. The choice for the number of grid points for high-resolution runs is a balance between having a sufficiently large domain and not having too time-consuming simulations. To examine the spin-up of the boundary conditions, a 100m forecast is performed for both an 800x800 gridpoints domain and a 400x400 gridpoints domain. At 50m resolution an error occurred when a domain of 400x400 gridpoints was used. With these settings the necessary interpolations (in `prepare_pgd`) could not be performed. Thus an 800x800-gridpoint domain was used instead.

For the sub-grid motions HARMONIE-AROME uses parameterisations for convection and turbulence. Details about these schemes can be found in Lenderink and Holtslag (2004), Bengtsson et al. (2017), and De Rooy et al. (2022). We performed experiments in which the convection scheme was turned off completely (`noconv`). This

¹<https://ruisdael-observatory.nl>

²www.ecmwf.int

Table 1: Overview of the 24h runs

Experiment name	NLON/NLAT	GSIZE (m)	TSTEP (s)	Nested in	Physics option
2500m	800x800	2500	75	ECMWF	Default
500m	400x400	500	15	2500m	Default
100m small domain	400x400	100	3	500m	Default
100m noconv small domain	400x400	100	3	500m	noconv
100m	800x800	100	3	500m	Default
100m noconv	800x800	100	3	500m	noconv
100m noconv noturb	800x800	100	3	500m	noconv noturb
50m	800x800	50	1	100m	Default
50m noconv	800x800	50	1	100m noconv	noconv

is done by defining the local surface buoyancy flux positive (thus towards the surface). This makes sure that the boundary layer is always classified stable, thus the convection scheme is not used. Additionally, we performed one experiment in which both the convection and turbulence parameterisations are turned off (noconv_noturb).

2.3 DALES

If the resolution is sufficiently high, most of the turbulence can be explicitly resolved. The turbulence is well resolved with grid sizes of order 10 m (Boutle et al. 2014) and the turbulence is mainly resolved with grid sizes between 10 m and 100 m (R. Honnert et al. 2011). Large-Eddy Simulations (LES) are designed for these grid sizes and resolve most of the turbulence. This type of simulations is computationally too demanding to use for operational weather forecasts (Kealy 2019). Nevertheless, it is very useful to compare the turbulence in the high-resolution HARMONIE-AROME runs to a small-scale LES, to determine how well turbulence is handled by HARMONIE-AROME at the high resolutions. Therefore, two simulations with DALES are performed, with resolutions of 100 m and 25 m. These resolutions are chosen as 100 m is the same as the HARMONIE-AROME experiments for which different set-ups are tested, and 25 m is such a high resolution that more turbulence can be resolved. Our DALES simulations are forced with data from the HARMONIE-AROME 2500 m simulation. In contrast to the HARMONIE-AROME simulations, the DALES simulations have a homogeneous land surface and periodic boundary conditions. The domain size is roughly 30 km x 30 km for the run at 100 m resolution and 15 km x 15 km for the run at 25 m resolution.

2.4 Observations

The results from HARMONIE-AROME and DALES will also be compared to observations. We used measurements from the observational site in Cabauw, the Netherlands. Specifically, we used the covariances of the wind components as measured with sonic-anemometers along a 200-m tower at 3, 60, and 180 m height. In addition, measurements from radiosondes are available from De Bilt at 0, 8, 12, 16 UTC for 14 and 15 September.

2.5 Calculations

This sections gives a brief description of the computations that were needed to compare the data from HARMONIE-AROME, DALES and the observations.

The total turbulent kinetic energy (TKE) is the sum of the TKE resolved by the model and the subgrid or parameterised TKE. The resolved TKE is calculated from the variances in u, v and w.

$$TKE = 0.5(\sigma_u^2 + \sigma_v^2 + \sigma_w^2)$$

The variances in u , v and w in HARMONIE-AROME are determined over a 15 km x 15 km area around Cabauw. In DALES, the variances are determined over the whole domain (thus, 30 km x 30 km for the run at 100 m resolution and 15 km x 15 km for the run at 25 m resolution). The equation is also used to calculate the total TKE from the observations. In DALES, σ_w^2 and the parameterised turbulence are at the cell edges, whereas σ_u^2 and σ_v^2 are at the cell centres. We linearly interpolated σ_w^2 and the parameterised turbulence to be able to combine them with σ_u^2 and σ_v^2 .

To normalise the TKE profiles, the boundary-layer height and convective velocity scale are calculated. Additionally, the Obukhov length is calculated to better understand the differences between HARMONIE-AROME and DALES. The boundary layer height (h) is defined as the height with the maximum potential temperature gradient. As the observations at Cabauw do not reach the boundary layer top, the boundary layer height is calculated from the radiosondes that were launched in De Bilt. The convective velocity scale (w_*) is calculated as:

$$w_* = \left(\frac{gh}{\theta_v} (\overline{w'\theta'_v})_s \right)^{1/3} .$$

Herein, g is the acceleration due to gravity, h is the boundary layer height, θ_v is the virtual potential temperature and $(\overline{w'\theta'_v})_s$ is the buoyancy flux at the surface. The virtual potential temperature, θ_v , is directly available from the DALES output and it is calculated from the potential temperature and specific humidity for HARMONIE-AROME and for the observations. The surface buoyancy flux $(\overline{w'\theta'_v})_s$ is calculated from the kinematic sensible and latent heat flux and the specific humidity. It should be noted that there is a slight mismatch in the height of variables used. The fluxes are at the surface, but the potential temperature and specific humidity are at the first model level.

3 Results and Discussion

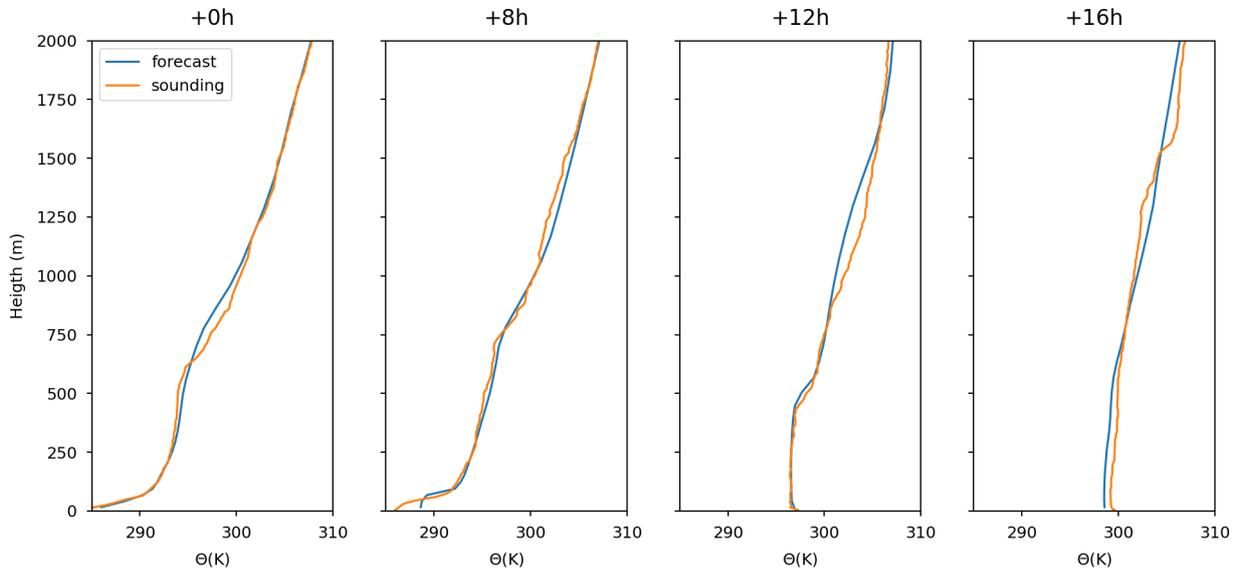


Figure 1: Average vertical profiles of potential temperature as observed by radiosondes launched at De Bilt and as simulated by HARMONIE-AROME (experiment 2500m) for 0, 8, 12 and 16 UTC on the 14th of September 2020.

3.1 Case description and model validation

Before discussing the different runs shown in Table 1 in detail, we compare the default run with the radiosonde observations (Fig. 1). The radiosonde is subject to drift during the ascend. However, the first 2 km, in which we are interested, is measured in about 6 min. The horizontal distance that the radiosonde travels in this 6 min is smaller than the size of our 2500 m grid cells. Therefore, it is justified to compare the observations of the radiosonde with the profile of one HARMONIE-AROME grid cell from our 2500m experiment. As mentioned before, the 14th of September 2020 was a warm day, with temperatures up to 28 degrees in Cabauw. There was a light breeze from the East-southeast. The sensible and latent heat flux show a clear daily cycle with a maximum around 12 UTC. Compared to the latent heat flux, the sensible heat flux is relatively small, with a Bowen ratio of around 0.2. The small sensible heat flux, together with subsidence and a strong inversion above the boundary layer, results in a well-mixed boundary layer at 12 UTC with a depth of around 500 m (Fig. 1). The boundary-layer development measured by the radiosondes is captured well by HARMONIE-AROME. The most noticeable difference is the overestimation of the near-surface temperature by HARMONIE-AROME at 8 UTC.

The remainder of this work focuses on 12 UTC. At this time there is a clear well-mixed boundary layer (Fig. 1) and the sensible heat flux is still high, thus we expect well developed turbulent motions. In general, vertical profiles are calculated as an average over an area of 15km x 15km around Cabauw. This area is chosen such that it does not contain very large differences in land use, but still covers sufficient points to take a meaningful average, also with the 2500m resolution.

3.2 High resolution runs with default physics

In this section, the four runs with different resolution are compared (all with the default settings regarding turbulence and convection). The impact of higher resolution immediately becomes clear when looking at maps of the temperature at the lowest model level in HARMONIE-AROME at ~ 12 m (Fig. 2). At 2500m, there is a difference in temperature between the more build-up areas and the more rural areas, but further details are lacking. When going to the 500m resolution, features like the river Waal ($\sim 5.83^\circ N$) and the lakes ($\sim 52.02^\circ N$, $4.7^\circ E$) become clearly visible. At even higher resolution (100 m and 50 m), the level of detail increases further, however the step from 100 m towards 50 m grid spacing hardly produces any extra details. The main reason is the grid resolution of the land-use map (ECOCLIMAP Second Generation) is 300m, preventing the presentation of features smaller than this resolution. Regarding the area-average vertical profiles, the different resolutions result in practically the same profile (not shown).

The impact of the higher resolution is also clearly visible in the maps of vertical velocity (Fig. 3). These maps show the vertical velocity at model level 57, which corresponds to a height of roughly 250m. At 12 UTC, this is halfway the boundary layer height. With increasing resolution, the size of the updrafts and downdrafts decreases and their intensity increases. This is line with the results of Lean et al. (2019) and Beare (2014). For the 500m resolution and especially for the 100m and 50m resolution, it is clear that the updrafts and downdrafts are organised in rolls along the wind. For the 50m resolution, the complete domain is plotted. Differences between the 100m and 50m run at the sides of the plotted area are therefore likely the result of boundary effects in the 50m run. The impact of the boundaries is investigated in more detail for the 100m run and discussed in section 3.4.

With higher resolution, smaller motions can be resolved that have to be parameterised at the coarser resolutions. Thus, with increasing resolution, one would expect that the amount of resolved TKE increases, whereas the amount of parameterised TKE decreases (Beare 2014). However, we found minimal differences in resolved and parameterised TKE between the different resolutions. The fact that hardly more turbulence is resolved when increasing the resolution is caused by the convection scheme. The mass flux scheme transports heat and momentum from the surface upwards. This reduces the instability of the profile, independent of the resolution. As the instability of the profile is taken away, the physics of the model will not result in (much) turbulent

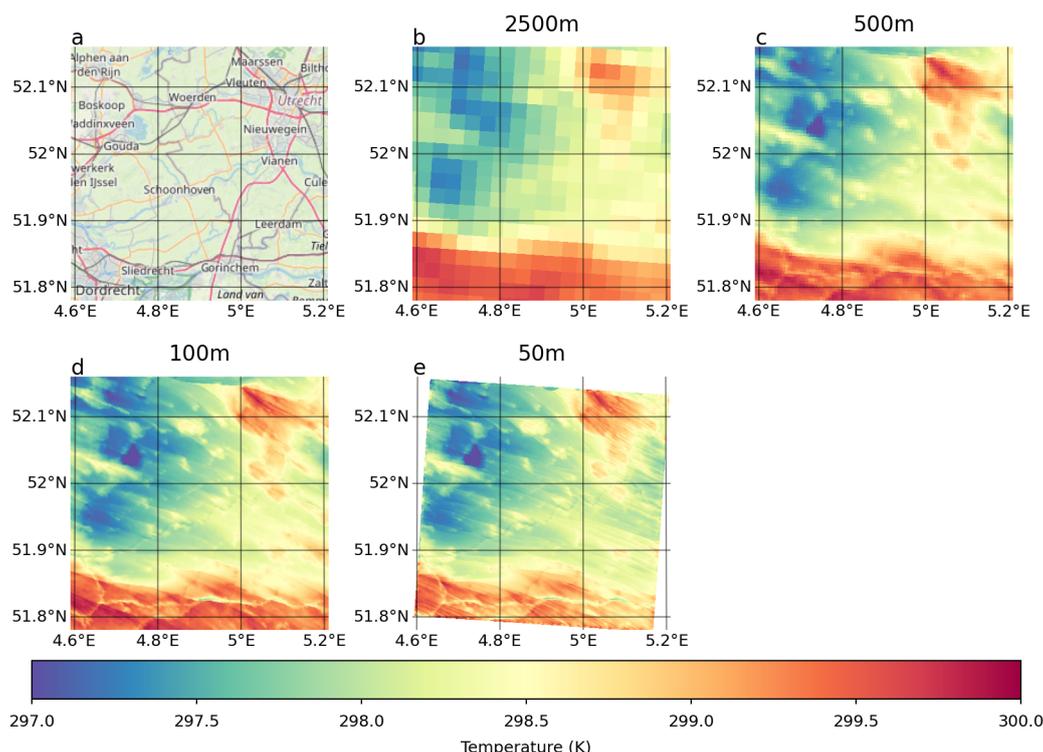


Figure 2: a) OpenStreetMap in a domain of roughly 40km x 40km around Cabauw (www.openstreetmap.org). b-e) Maps of the absolute temperature for experiments 2500m, 500m, 100m and 50m (as described in Table. 1). These maps show the absolute temperature in a domain of roughly 40km x 40km around Cabauw at the lowest model level in HARMONIE-AROME (~12m) at 12 UTC.

motions anymore. So even though a higher resolution simulation might be able to resolve more turbulent motions, the presence of the convection scheme causes the runs at different resolutions to be very similar. This raises the question what happens if the convection scheme is disabled, which will be discussed in the next section.

3.3 Sensitivity to the convection and turbulence schemes

When the convection scheme is disabled, the parameterised TKE is reduced and the resolved TKE is increased (Fig. 4). The total TKE in the noconv runs is larger than default. This is likely because the turbulence scheme is not adapted to the higher resolution. This scheme is developed for the 2500m resolution, where the vertical gradients and turbulent fluxes are much larger in the vertical than in the horizontal direction. At higher resolution, this is no longer true and a 3D turbulence scheme would be better.

What is also remarkable is the high amount of resolved TKE around the boundary-layer top in HARMONIE-AROME in the simulations with the default settings. Around the boundary-layer top, the wind profile (not shown) shows a strong increase in wind speed. As the TKE profiles are an average over 15 km x 15 km, a large range of wind speeds occurs over the area, causing the high velocity variance.

Disabling the convection scheme also influences the vertical velocity field (Fig. 3). The along-wind rolls that are visible in the default run are much less pronounced and less straight in the noconv run. The vertical motions without the convection scheme are also more intense.

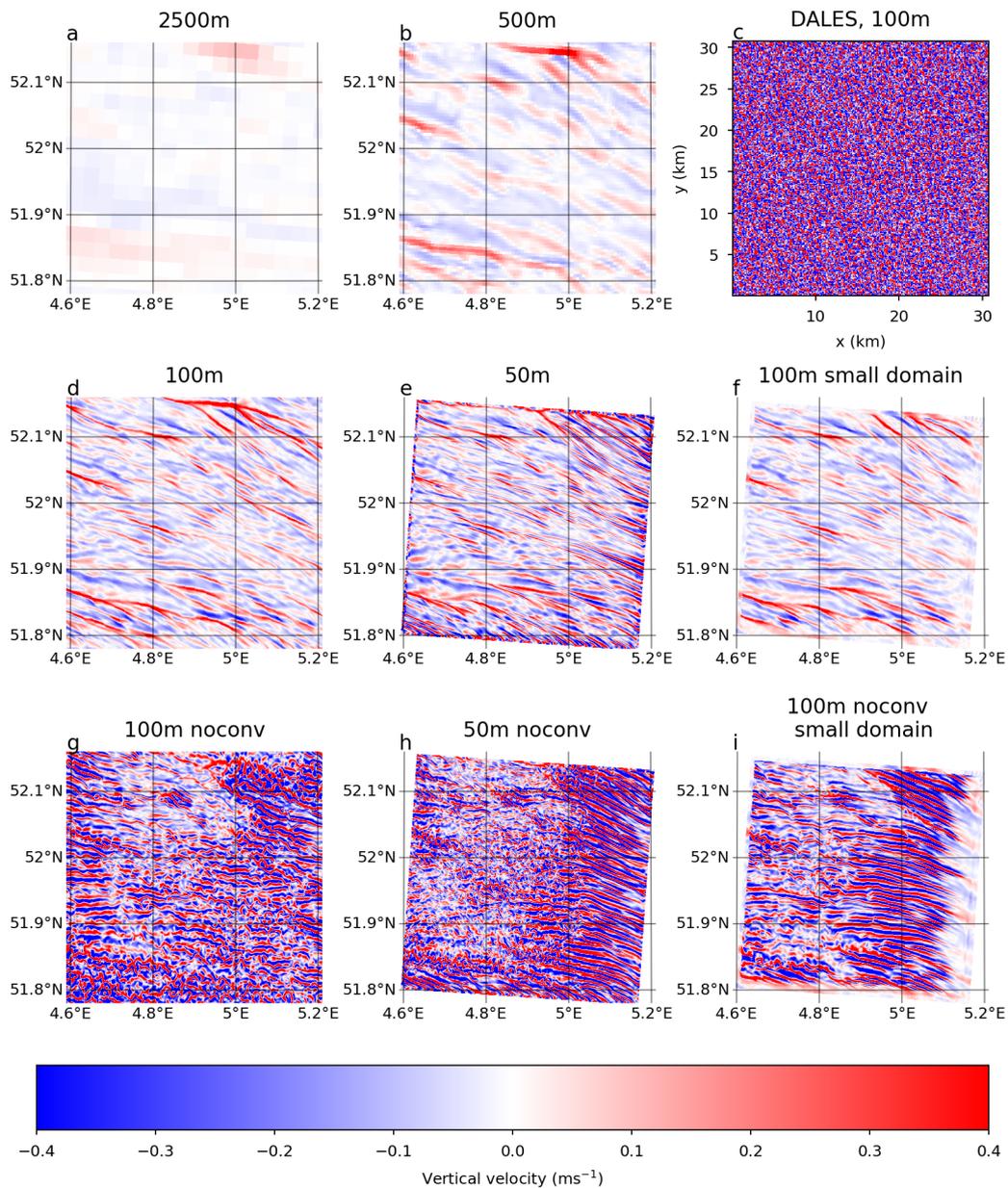


Figure 3: Maps of the vertical velocity for the HARMONIE-AROME simulations (as described in Table. 1) 2500m (a), 500m (b), 100m (d), 50m (e), 100m small domain (f), 100m noconv (g), 50m noconv (h), 100m noconv small domain (i), and the DALES simulation with 100m resolution (c). These maps are at 250m height (\sim half the boundary layer) at 12 UTC.

As mentioned in section 2, we also performed a simulation with the turbulence scheme disabled (not shown). The potential temperature profile of this run shows a very low surface temperature and with a profile typical for night-time stable boundary layer. This profile can be explained by the missing turbulence scheme. At the surface, the first transport of heat is done by the smallest eddies. These are sub-grid, also at 100m grid spacing. By turning off the turbulence scheme completely, these eddies are also missing. Therefore, the heat from the surface is not mixed with the air above anymore. Therefore, the temperature profile changes only very little over time and there is no mixed layer formed. We can therefore conclude that switching the turbulence scheme off completely causes an unrealistic simulation, because of the missing smallest eddies.

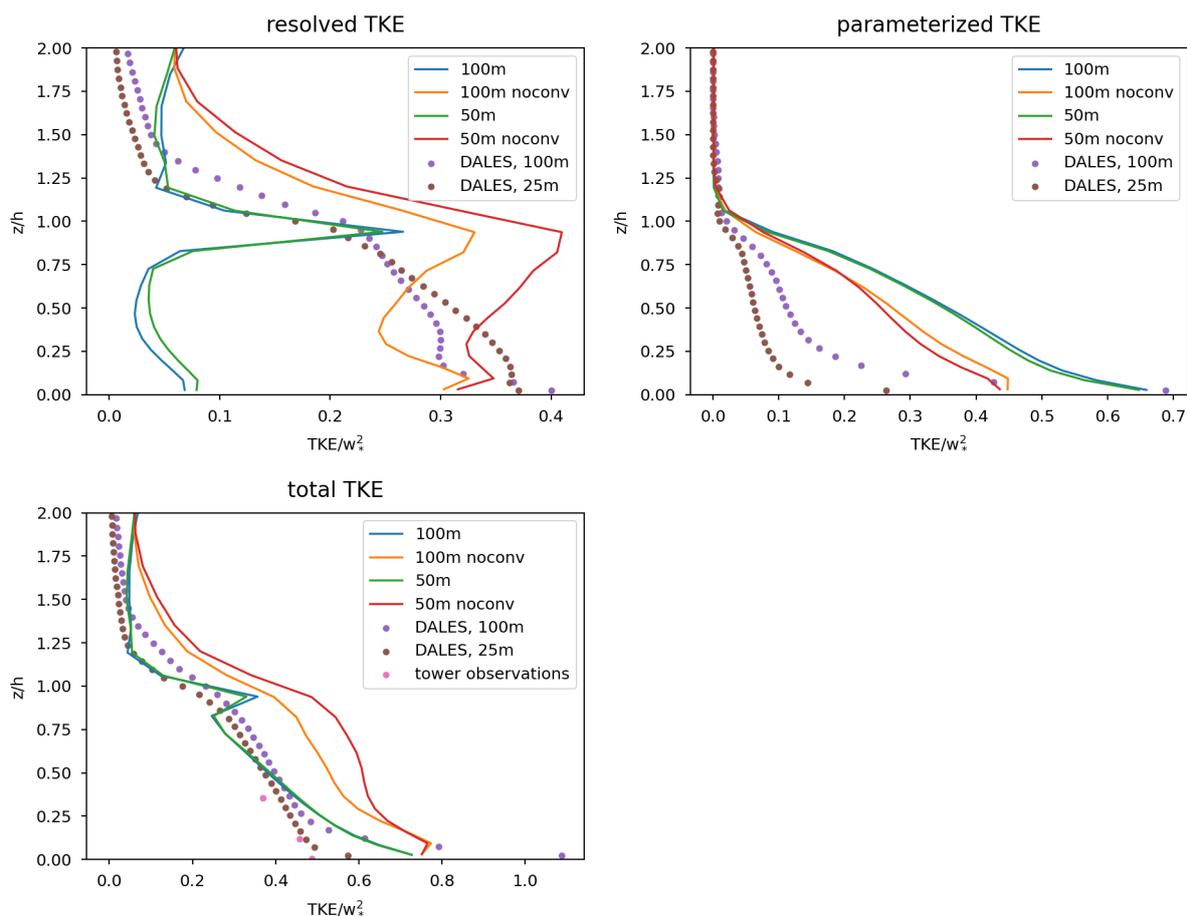


Figure 4: Normalized vertical profiles of resolved, parameterised and total TKE at 12 UTC for HARMONIE-AROME (100m, 100m noconv, 50m and 50m noconv), DALES (100m and 25m) and from the observations.

3.4 Impact of the boundaries

Apart from the different convection and turbulence settings, the 100m forecasts with and without convection were also made on a domain of 400x400 grid points, to study the impact of the boundaries. The impact of the boundaries is clearly visible in the vertical velocity fields (Fig. 3). As the wind is from the east-southeast, the impact of the boundaries is visible on the eastern side of the domain. When a small domain is used, the vertical velocity is limited to ~ 0.2 m/s and structures are all at least 1–2 km wide directly downwind of the inflow boundary (Fig. 3f). Further away from the boundaries (Fig. 3d), vertical velocities larger than 0.4 m/s and structures narrower than 1 km occur. The impact of the boundaries is much larger in the run without convection (Fig. 3i). Here, clear along-wind rolls occur in the first half of the 400x400 domain. After that, the structures become less straight. In the case of the 100m run without convection, the boundaries are taken from a simulation that does not only differ in horizontal resolution, but also in the convection scheme. It is clear that it takes only a short distance for the model to adapt to a higher resolution. But it takes much longer for the model to adapt to the change in convection scheme, in other words, it takes longer before the model can develop the turbulence by itself. The impact of the boundaries is investigated before by Lean et al. (2019) using the UK met office model. They found a region with no overturning close to the boundary and roll structures afterwards, which extend about 15 km into the domain. This is similar to what we find in the 100 m domain without convection (Fig. 3i). The impact of the boundaries is case dependent, as several factors determine if

rolls are formed after the inflow boundary and how far they extend into the domain: the wind speed, stability and model numerics (Lean et al. 2019). Boutle et al. (2014) also found that roll structures are formed close to the inflow boundaries, as it takes time before the model can fully develop turbulent structures.

3.5 Comparison with DALES

Firstly, the vertical profiles of temperature are presented (Fig. 5). In general, all profiles are similar. In the first part of the profile, the default run is closest to the DALES runs. Whereas the inversion height of the run without convection is closest to the DALES runs.

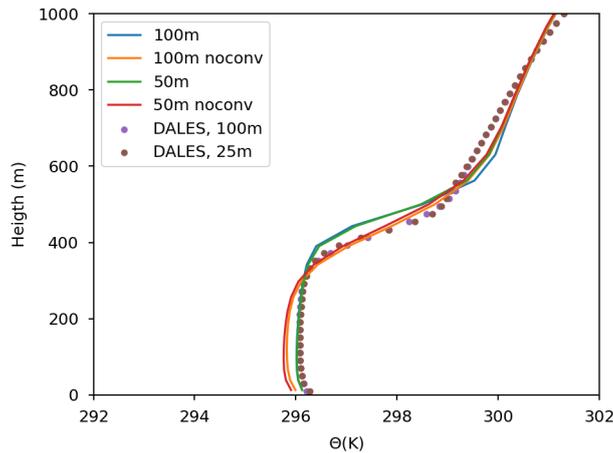


Figure 5: Vertical profile of the potential temperature for HARMONIE-AROME (100m, 100m noconv, 50m and 50m noconv) and DALES (100m and 25m) at 12 UTC.

The vertical velocity field of the DALES run shows very small updrafts and downdrafts with a cell structure (Fig. 3c). Because the resolution of 100 m is coarse for an LES model, the cell structures are not fully resolved. Both HARMONIE-AROME runs show roll structures, although these are moving towards cell structures in the run without convection. These linear features are horizontal convective rolls (HCRs)(Lemone 1973). Miao and Chen (2008) showed that these HCRs generally occur when $-h/L < 25$, where h is the boundary-layer height and L the Obukhov length. Comparing the HARMONIE-AROME runs with the DALES run shows that indeed, $-h/L$ is smaller than 25 for the HARMONIE-AROME runs (~ 11) and larger than 25 in DALES (~ 30). Thus it is in line with the work of Miao and Chen (2008) that HCRs are formed in HARMONIE-AROME and not in DALES. From the observations, $-h/L$ is also larger than 25 (~ 60). The formation of HCRs in HARMONIE-AROME might have different causes. It can be a consequence of way in which the model calculates (HARMONIE-AROME is a semi-lagrangian model). Furthermore, in HARMONIE-AROME, the surface has different land use types, whereas DALES has a homogeneous surface. These differences in surface impact the friction velocity and thus $-h/L$. Previous work showed that the formation of HCRs is more likely over urban areas, where the friction velocity is large (Miao and Chen 2008).

The vertical profiles of TKE also show differences between HARMONIE-AROME and DALES. The parameterised TKE in DALES is much lower than in HARMONIE-AROME (Fig. 4). In the lowest layers, the amount of parameterised TKE is strongly increasing in the DALES runs. This is because in these layers, close to the surface, the dominant eddies are very small. At 100m resolution, these eddies are not (completely) resolved.

The resolved TKE in DALES is similar in size as the resolved TKE in the HARMONIE-AROME run without convection. But the shape of the profile is different. Therefore, the resolved TKE is split into the contribution of the u and v component of the wind and the w component (Fig. 6). The resolved u and v component of TKE has a similar profile in the DALES and HARMONIE-AROME noconv run. On the other hand, the profile

of the w component is very different. The DALES run shows a typical profile with the maximum variance in w around 1/3 of the boundary-layer height (Stull 1988; Beare 2014). From this, the question arises whether a higher vertical resolution in HARMONIE-AROME is needed to resolve the TKE better. Experiments with increased vertical resolution should therefore be part of further research.

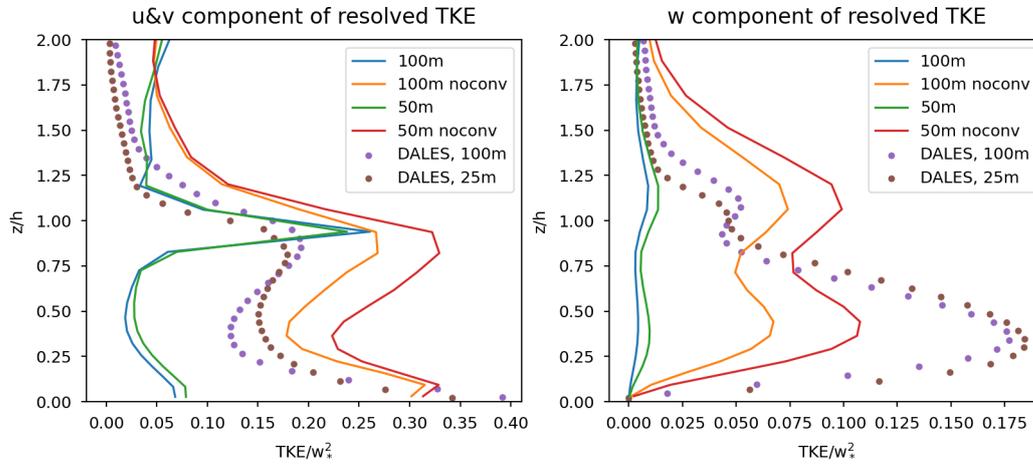


Figure 6: Normalized vertical profiles of the u and v component of resolved TKE (left panel) and the w component of resolved TKE (right panel) at 12 UTC for HARMONIE-AROME (100m, 100m noconv, 50m and 50m noconv) and DALES (100m and 25m).

Lastly, the total TKE. The profiles of total TKE almost overlap for the default 100m run, the DALES runs and the observations in the largest part of the boundary layer (Fig. 4). The run without convection seems to overestimate the TKE. Differences in total TKE occur close to the surface. What is striking is the high amount of resolved TKE in the DALES run of 100m resolution, resulting also in a very high amount of total TKE. This might be related to the overestimation of the sensible heat flux in DALES, causing too much turbulence close to the surface. Also shown is the 25m run of DALES. As expected, compared to the 100m DALES run, the 25m run has more resolved TKE, less parameterised TKE and a similar total TKE.

Additionally, the results of the 50m runs with and without convection are shown in figures 3, 4, 5, and 6. The vertical velocity field of the 50m run without convection shows along wind rolls in the first part of the domain, similar to the 100m noconv run on the small domain. In the 100m noconv small domain run, low vertical velocities (around 0.2 m/s) were found in the first ~ 5 km after the inflow boundary. In the 50m noconv run there is not such a region with low vertical velocities directly downwind of the inflow boundary. This is because the noconv 50m run is nested in the 100m noconv run, where the vertical velocities are already much higher than in the default runs.

The parameterised TKE in the 50m noconv run is almost the same as in the 100m noconv run. However, the resolved TKE increases even further, which makes the total TKE clearly too high (Fig. 4). This shows that with the current set-up part of the TKE is resolved and parameterised, and is taken into account twice. This shows that at these resolutions, when turning off the convection scheme, additionally changes to the turbulence scheme are needed.

4 Conclusions and Outlook

4.1 Main conclusions

In this study, we investigated the potential of HARMONIE-AROME at high-resolution (down to 50m) for a clear-sky and calm day in the Netherlands (14 September, 2020). For this day, the default HARMONIE-AROME settings were suitable, meaning that no instability issues were encountered. Forecasts at four resolutions (2500m, 500m, 100m, 50m) showed that the level of spatial detail increases when going to a higher resolution, as features like rivers become visible in the surface temperature maps and the up- and downdrafts become smaller in size and more intense.

However, as long as the convection scheme is left on, there is hardly more turbulence resolved. When the convection scheme is turned off, more TKE is resolved and less TKE is parameterised. The total amount of TKE is slightly too high when the convection scheme is disabled.

This study does not go into the question at which resolution the convection scheme is no longer needed. It became clear, however, that at 100m a substantial part of the TKE can be resolved when the convection scheme is disabled. As hardly any more turbulence is resolved as long as the convection scheme is left on, it is preferable to (gradually) turn the convection scheme off at high resolution. The vertical velocity fields in the HARMONIE-AROME simulations with the convection scheme show clear along wind rolls. When the convection scheme is turned off, the rolls become less pronounced, less straight, and more cellular. With DALES we found very different, cellular structures in the vertical velocity field.

4.2 Minor conclusions

- When both the convection and turbulence scheme are disabled, the temperature profile is too cold at the surface and too stable. This shows that the turbulence scheme is essential to model the smallest eddies that take care of the first transport from the surface to the atmosphere.
- The model boundaries influence the structure of the up and downdrafts near the boundaries. Especially when a simulation without convection scheme is nested in a simulation with convection scheme, there are clear rolls in the vertical velocity field near the inflow boundary. These rolls are roughly present in the first 15 km of the domain for this particular case.

4.3 Recommendations

Previous research already showed that the spatial details that high resolution simulations can provide can be very beneficial for city-scale forecasts, modelling nocturnal processes and modelling regions with complex orography. In this study, we focused on vertical velocity and turbulent kinetic energy. We found that the structures in the vertical velocity field clearly differ between the HARMONIE-AROME simulation with convection, without convection and the DALES simulation. We hypothesize that the structures formed in the run without convection are more realistic than the rolls in the run with convection, as the run with the convection scheme on is clearly dominated by the convection scheme. It is therefore recommended to turn the convection scheme off and test if the convection scheme can be turned off already at 500m. Additionally, it should be investigated if the convection scheme can still be turned off when cloud formation occurs. Furthermore, we saw a clear effect of the boundaries on the structures formed in the vertical velocity field. It is therefore recommended to investigate if more information can be supplied to the model at the boundaries, to reduce the spin-up distance. Additionally, the turbulence scheme might be adapted to allow HARMONIE-AROME to resolve even more of the TKE. For example, the length scale could be made dependent on the grid size and stability. Or one could

use a different parameterisation that mainly produces turbulence close to the surface with a length scale equal to the grid length (Rachel Honnert et al. 2021).

To get a much broader insight in the advantages and challenges of running HARMONIE-AROME at high-resolutions, the next step is to run HARMONIE-AROME daily at a high resolution. By running HARMONIE-AROME daily a large range of conditions will be covered. This would allow for a much more elaborate comparison with DALES and with observations and will test the numerical stability of the model setup under different weather conditions.

Along with the increase in horizontal resolution, we recommend to increase the vertical resolution. We found that the variance in vertical velocity is not well modelled by HARMONIE-AROME. With increased vertical resolution, the vertical motions can possibly be better resolved. For comparison, the number of vertical levels below 2000m in HARMONIE-AROME is now 27, while in DALES it is 73.

Although the model runs stably with the defaults settings for this case, we recommend trying other settings that might increase the efficiency and decrease the computation time, e.g. use of single precision. The case we invested is a very calm day, thus it is not guaranteed that the defaults settings also result in stable simulations for more dynamic cases. Furthermore, the 24h forecast at 50m resolution, currently takes more than two days, which makes these simulations not suitable for any practical use, only for research purposes. Possible settings include using a cubic or quadratic grid instead of the default linear grid (Malardel et al. 2016) and using single precision instead of the default double precision (Váňa et al. 2017).

Lastly, we recommend to have a broader look than just the temperature profiles, vertical wind and TKE that are shown here. Structures of other variables (temperature, moisture) are bigger than for vertical velocity (R. Honnert et al. (2011)), thus these should be resolved for a larger part.

Acknowledgements

We want to thank Wim de Rooy and Geert Lenderink for their valuable suggestions and for making the necessary adaptation to the HARMONIE-AROME source code. We thank Fred Bosveld for providing Cabauw data and Stephan de Roode for providing the radiosonde data. And we thank Bart van Stratum for his valuable suggestions and support.

Appendix - Detailed run settings

This appendix contains a detailed description of which changes are made in which scripts.

For all runs, the following changes are made in the config_exp.h script: ANAATMO=none, ANASURF=none, OBSMONITOR=no and WRITEUPTIME = 1. These changes make sure that no data assimilation is performed and that we have hourly output data. Additionally, the 100m and 50m forecasts are performed without IO_server, therefore IO_server = no in config_exp.h and nproc_io=0 in Env_submit.

For the nesting, the following changes are made in config_exp.h: HOST_MODEL='aro', HOST_SURFEX=yes, BDLIB=the experiment to nest in, BDSTRATEGY = same_forecast.

The adaptations to the domain (as described in Table. 1) are implemented in harmonie_domains.pm. For the default 2500m run, the standard domain NETHERLANDS (centered at Cabauw) is used. The domains for the high-resolution runs are also centered at Cabauw.

The physics options are implemented as follows:

- noconv: this option is implemented by putting the local buoyancy flux variable to a fixed, positive number (ZKHFVL = 10 in vdfhghtnhl.F90), after which the the model executable MASTERODB should be recomputed;

- noconv noturb: for this experiment, the same changes are made as for the noconv experiment. Additionally, the turbulence scheme is disabled by putting LTURB to FALSE in the script `harmonie_namelist.pm`.

5 References

- Beare, R.J. (2014). “A length scale defining partially-resolved boundary-layer turbulence simulations”. In: *Boundary-layer meteorology* 151.1, pp. 39–55.
- Bengtsson, L., U. Andrae, T. Aspelien, Y. Batrak, J. Calvo, W. de Rooy, E. Gleeson, B. Hansen-Sass, M. Homleid, M. Hortal, et al. (2017). “The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system”. In: *Monthly Weather Review* 145.5, pp. 1919–1935.
- Boutle, IA, JEY Eyre, and AP Lock (2014). “Seamless stratocumulus simulation across the turbulent gray zone”. In: *Monthly Weather Review* 142.4, pp. 1655–1668.
- Boutle, IA, A Finnenkoetter, AP Lock, and H Wells (2016). “The London Model: forecasting fog at 333 m resolution”. In: *Quarterly Journal of the Royal Meteorological Society* 142.694, pp. 360–371.
- De Rooy, W.C., P. Siebesma, P. Baas, G. Lenderink, S. de Roode, H. de Vries, E. van Meijgaard, J.F. Meirink, S. Tijm, and B. van ’t Veen (2022). “Model development in practice: A comprehensive update to the boundary layer schemes in HARMONIE-AROME.” In: *Geoscientific Model Development Discussions*. in press.
- Heus, T, CC Van Heerwaarden, Harmen JJ Jonker, AP Siebesma, S Axelsen, K Dries, O Geoffroy, AF Moene, D Pino, SR De Roode, and J. Vilà-Guerau de Arellano (2010). “Formulation of the Dutch Atmospheric Large-Eddy Simulation (DALES) and overview of its applications”. In: *Geoscientific Model Development* 3.2, pp. 415–444.
- Honnert, R., V. Masson, and F. Couvreux (2011). “A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometeric scale”. In: *Journal of the Atmospheric Sciences* 68.12, pp. 3112–3131.
- Honnert, Rachel, Valéry Masson, Christine Lac, and Tim Nagel (2021). “A theoretical analysis of mixing length for atmospheric models from micro to large scales”. In: *Frontiers in Earth Science*, p. 537.
- Kealy, J.C. (2019). “Probing the ‘grey zone’ of NWP—is higher resolution always better?” In: *Weather* 74.7, pp. 246–249.
- Lean, H.W, J.F. Barlow, and C.H. Halios (2019). “The impact of spin-up and resolution on the representation of a clear convective boundary layer over London in order 100 m grid-length versions of the Met Office Unified Model”. In: *Quarterly Journal of the Royal Meteorological Society* 145.721, pp. 1674–1689.
- Lemone, M.A. (1973). “The structure and dynamics of horizontal roll vortices in the planetary boundary layer”. In: *Journal of Atmospheric Sciences* 30.6, pp. 1077–1091.
- Lenderink, G. and A. Holtslag (2004). “An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers”. In: *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography* 130.604, pp. 3405–3427.
- Malardel, S., N. Wedi, W. Deconinck, M. Diamantakis, C. Kühnlein, G. Mozdzyński, M. Hamrud, and P. Smolarkiewicz (2016). “A new grid for the IFS”. In: *ECMWF Newsletter* 146.23-28, p. 321.
- Miao, S. and F. Chen (2008). “Formation of horizontal convective rolls in urban areas”. In: *Atmospheric Research* 89.3, pp. 298–304.
- Ronda, RJ, GJ Steeneveld, BG Heusinkveld, JJ Attema, and AAM Holtslag (2017). “Urban finescale forecasting reveals weather conditions with unprecedented detail”. In: *Bulletin of the American Meteorological Society* 98.12, pp. 2675–2688.
- Stull, R.B. (1988). *An introduction to boundary layer meteorology*. Vol. 13. Springer Science & Business Media.
- Valkonen, T., P. Stoll, Y. Batrak, M. Kølitzow, T.M. Schneider, E.E. Stigter, O.B. Aashamar, E. Støylen, and M.O. Jonassen (2020). “Evaluation of a sub-kilometre NWP system in an Arctic fjord-valley system in winter”. In: *Tellus A: Dynamic Meteorology and Oceanography* 72.1, pp. 1–21.

- Váňa, F., P. Düben, S. Lang, T. Palmer, M. Leutbecher, D. Salmond, and G. Carver (2017). “Single precision in weather forecasting models: An evaluation with the IFS”. In: *Monthly Weather Review* 145.2, pp. 495–502.
- Wyngaard, John C (2004). “Toward numerical modeling in the “Terra Incognita””. In: *Journal of Atmospheric Sciences* 61.14, pp. 1816–1826.
- Yang, X. (2019). “TAS, an operational forecast model at hectometric scale”. In: *ALADIN-HARMONIE Newsletter* 12, pp. 42–49.

Hectometric-Scale Experiments at Met Éireann

Colm Clancy, James Fannon, Eoin Whelan

1 Introduction

This article summarises some of the key results from a series of experiments with HARMONIE-AROME at sub-kilometre horizontal resolutions, and with an increased number of vertical levels. This is a topic of considerable interest across the consortium with numerous recent contributions; for example Antoine et al. (2021), Chikhi et al. (2021), Suárez-Molina and Calvo (2021), Yang (2018).

Currently at Met Éireann, and elsewhere, operational suites use a horizontal resolution of 2.5 km and 65 vertical levels. In this work domains with 500 m and 750 m resolution were tested, along with the 90 levels from Météo France. Cycle 43h2.1.1 was used for all of these experiments. A more comprehensive technical note detailing this work is available from Met Éireann.

2 Experiment Configurations

An initial phase of experiments on Irish domains were carried out to test configuration settings. These consisted of single 36-hour forecasts beginning from 0000 UTC on the 22nd of February 2021. On the night of the 22nd/23rd, a low pressure system to the west of the country brought strong winds and heavy rain to the south-west of Ireland in particular.

2.1 Single Precision

Building upon work carried out at ECMWF and Météo-France, the option to run HARMONIE-AROME in single precision (SP) was made available as part of Cycle 43h2.1 (see Vignes, 2019; HIRLAM, 2020). This has been largely motivated by ECMWF's extensive research in this area, which has demonstrated that SP IFS offers a $\sim 40\%$ runtime saving, relative to double precision (DP), with no significant degradation in forecast quality (Váně et al., 2017), culminating in the migration to operational SP forecasts as part of ECMWF's Cycle 47r2 upgrade (Lang et al, 2021).

The SP option in HARMONIE-AROME had previously been investigated as part of pre-operational testing of Cycle 43h2.1 at Met Éireann (Bessardon et al., 2020), with initial results suggesting an approximate 30% forecast runtime saving and largely neutral impact on surface verification scores. Given this substantial saving, it was deemed appropriate to investigate using SP in this project. Early tests at 2.5 km and 500 m resolution showed savings of 30-40% in the runtimes and very similar results in the SP and DP simulations, with the SP forecasts exhibiting a small positive MSLP bias relative to DP. These results are consistent with previous studies (Feddersen, 2021). Therefore, it was decided to use SP by default¹ in all of the high-resolution experiments. No significant issues regarding the stability of the SP forecasts were encountered.

¹Note that SP was used for Forecast, with other model components retained in DP. This is now available through the "dual" precision option in Cycle 43h2.2

2.2 Dynamics settings

The main aim of initial single-forecast testing was to find a stable configuration for a 500 m-resolution domain with 90 vertical levels. The quadratic or cubic grids were targeted: given their truncation of the spectral resolution, these have the advantage of both saving computational cost and potentially being more stable, albeit with a formal reduction in accuracy. In addition to the time-step, a key factor to be addressed is the appropriate level of diffusion necessary for higher-resolution forecasting: increasing diffusion helps to control noise. However, too much diffusion will over-smooth the solution and potentially counteract any benefits of increased resolution.

Concern for the dynamics at very high resolutions generally revolves around “complex orography” and “steep slopes”. Irish geography is fairly benign in these respects, and so testing may not push the dynamics to the limit. Put another way, a configuration that is found to be sufficient for Ireland may not be adequate for other regions. So in addition to the Irish tests with the 22nd of February 2021 case, experiments were run on 500 m-resolution domains over the Canary Islands and Greenland, using cases previously studied by DMI and AEMET. Unsurprisingly, a case of extreme wind over the south of Greenland during the 27th-28th of December 2017, with an observed mean wind-speed of 46 m/s, proved the greatest challenge for stability. This required a predictor-corrector scheme with a 10s time-step.

For the tests over the Canary Islands (25th-26th of February 2018), as well as all of the Irish domain experiments, a time-step of 15s was found to be sufficient for stability using the SETTLS scheme (default in the HARMONIE-AROME CSC) on both quadratic and cubic grids. Extensive testing of the spectral horizontal diffusion showed that the namelist settings REXPDH=4 and RDAMP*=1, on a cubic grid, were a sensible (perhaps over-cautious) choice. Full details of these experiments may be found in the technical note, available from Met Éireann.

2.3 Boundary coupling

When running a 2.5 km-resolution HARMONIE-AROME forecast with lateral boundary conditions (LBC) from the IFS, only the dynamics fields are used at all coupling times by default. The cloud water species - ice (I) and liquid (L) - as well as the hydrometeors - graupel (G), rain (R), snow (S) - are only coupled at the initial time. However it is possible to include them at all times, as long as they exist in the LBC files. Figure 1 shows the effect of the additional coupling of cloud water and hydrometeors in terms of 24-hour rainfall accumulation on the large IRELAND25 domain (the operational domain at Met Éireann until 2018). The impact of the change of coupling would seem to be insignificant, occurring (not surprisingly) close to the boundaries far from our region of interest, the island of Ireland.

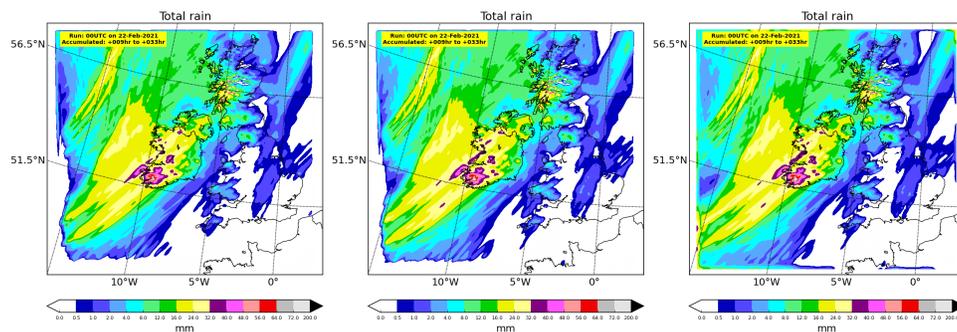


Figure 1: Forecasted 24-hour accumulated rainfall for the 22nd of February case using the IRELAND25 domain with (left to right) default IFS LBC coupling; coupling I/L; coupling all I/L/G/R/S.

It is a different matter when we have a geographically-small domain, unavoidable when we move to higher resolutions. An example is shown for the smaller 500×500 domain covering the east of Ireland at 500 m

resolution in Fig. 2; for comparison the 2.5 km reference is shown zoomed into the region on the left panel. Now the dry regions “close” to the boundaries are much more significant, and the boundary spin-up is clearly an issue of concern: it would suggest that full coupling is essential for small domains.

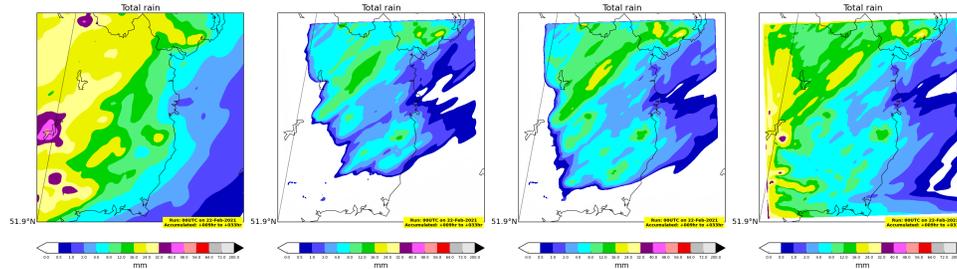


Figure 2: Forecasted 24-hour accumulated rainfall for the 22nd of February case. Left to right: 2.5 km forecast zoomed into the high-resolution domain region; 500 m domain with default IFS LBC coupling, with coupling of I/L, and with coupling of all I/L/G/R/S.

2.4 Domain size

Results in the previous section showed the sensitivity of rainfall amounts to the LBC coupling. Figure 3 shows snapshots of total cloud cover forecasts at 500 m resolution where all experiments use the default coupling, but the domain size is progressively increased. The domain in the third panel, where we begin to see some convergence in the predicted cloud structure, has horizontal dimensions 720×600. When run in single precision, a simulation on this domain costs roughly the same as double precision simulation on a 500×500 domain (the left panel of Fig. 3).

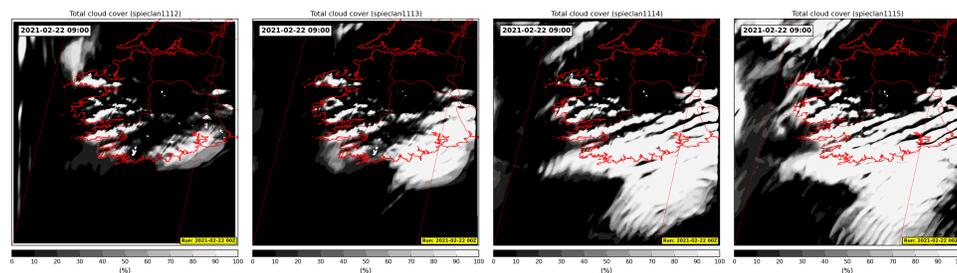


Figure 3: Forecasted total cloud cover during the 22nd of February case, 500 m resolution using IFS LBC with default coupling. Domain size increases from left to right, with the region shown corresponding to the smallest 500×500 domain on the left.

3 Results of Cycling Experiments

Longer cycling experiments were run in order to generate verification statistics on accuracy, and to test for any nonlinear instabilities which might arise. Simple upper-air blending was used, rather than any more complicated data assimilation. Three-hour cycling was used, with a 33-hour forecast at 0000 UTC and a 24-hour forecast at 1200 UTC. Two test periods were considered: the 3rd to 16th of February 2020, which contained Storms Ciara and Dennis, and the 10th to 24th of April 2021, a calmer, more stable period with a number of fog events.

3.1 Domains at 500 m

A number of 500 m-resolution domains were initially chosen for longer cycling tests. These are shown in Fig. 4. The two smaller domains have dimension 720×600 , following the results shown in Section 2.4 above. As was seen there, larger domains are preferable, and so a third 500 m domain, of size 1200×1200 was also tested (red domain in Fig. 4). Due to the cost of this large domain, the experiment was run for just 7 days for the February case only, from the 3rd to 10th of February 2020. Full IFS LBC coupling (i.e. I, L, G, R, and S) was used.

Results can be compared for stations in the overlapping region of the three domains visible in Fig. 4. Scores of MSLP and 10 m wind-speed were found to be quite comparable. For 2 m temperature bigger differences were seen, with stations near the domain boundaries showing very different cloud cover (as seen in Fig. 3) and dominating the errors.

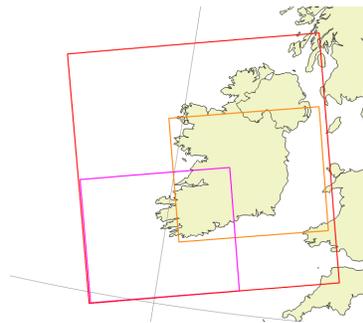


Figure 4: Larger test domains, of dimension 720×600 (SWIRL500W2, purple, and EIRL500W2, orange) and 1200×1200 (LIRL500, red).

The domain size and boundary effects are most clearly visible in precipitation forecasts. Figure 5 shows time-series for the three experiments. While no experiment is perfect, the green curve (EIRL500W2 domain in Fig. 4) stands out from the 10th of February, when it seems to significantly miss the observed precipitation. Figure 6 shows sample maps of 12-hour accumulations during this period, along with the observed amounts. Again, we see clearly how the EIRL500W2 domain is suffering near its south-west boundary.

Note that in all of these experiments IFS LBC were used with full coupling of I/L/G/R/S. We can instead nest within a 2.5 km HARMONIE-AROME forecast, also with I/L/G/R/S coupling. Figure 7 compares the total

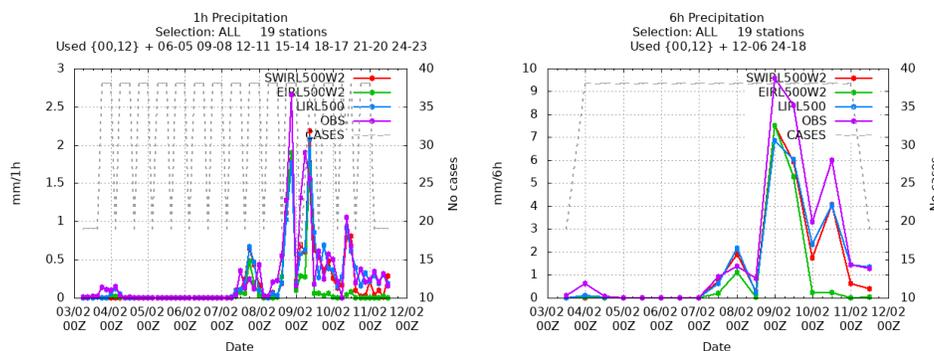


Figure 5: Averaged time-series for 1-hour (left) and 6-hour (right) precipitation for 500 m experiments during February 2020. The experiments use the domains shown in Fig. 4.

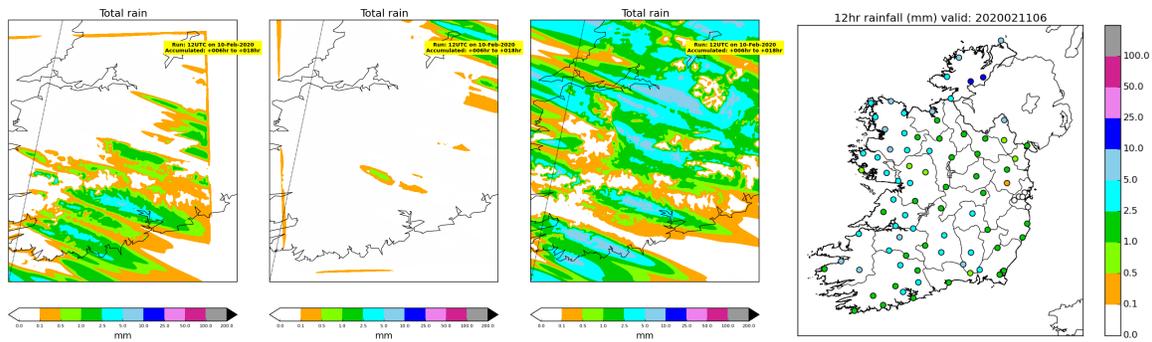


Figure 6: Forecasted 12-hour accumulated precipitation valid at 06 UTC on the 11th of February 2020, using the 12 UTC cycle on the 10th. The snapshots (from left) show the overlap regions from experiments with domains SWIRL500W2, EIRL500W2 and LIRL500; all shown in Fig. 4. Right: observations.

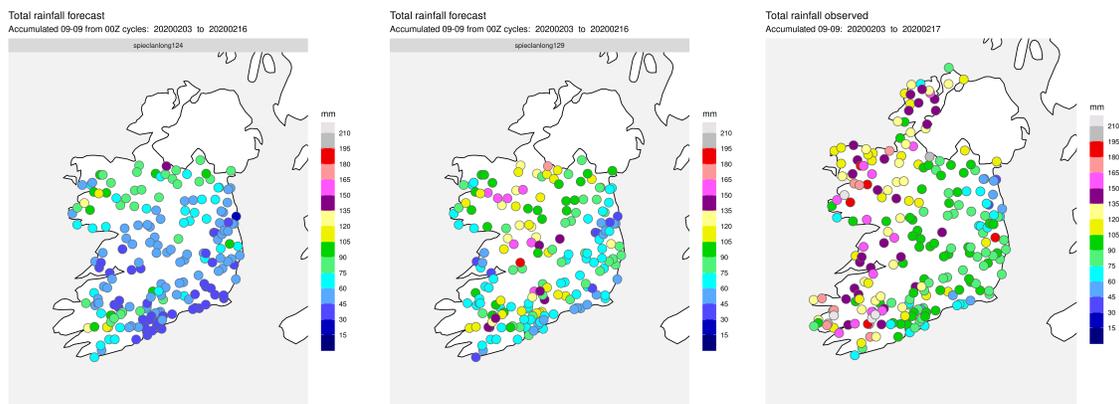


Figure 7: Accumulated rainfall for the February test period, using the +9 to +33 hour forecasts from the 0000 UTC cycles to sum the daily 09-09 amounts. Experiments at 500 m resolution on EIRL500W2 domain (see Fig. 4) using LBC from IFS (left) and HARMONIE-AROME (middle). Right: observed accumulations.

forecasted rainfall accumulations over the two-week February period between experiments with LBC from IFS and HARMONIE-AROME. While all the boundary regions in the nested experiment (middle panel) are still too dry compared with observed amounts, they show significant improvement from when IFS LBC are used.

3.2 Comparison of various resolutions

As illustrated in Sections 2.4 and 3.1, the domain size can have a major impact on the performance of the sub-kilometre resolution simulations, particularly in relation to precipitation. This highlights the need for a “large” domain, in the geographic sense, i. e. we want our region of interest to be far away from the boundaries. The use of the large LIRL500 domain in Fig. 4, however, is prohibitively expensive from an operational perspective. As a compromise, we also consider a 800×800 domain with 750 m resolution, IRL750, which matches the same region as LIRL500. Previous testing at Met Éireann has suggested a 30s time-step is sufficient for stability at this resolution, compared to a 15s time-step at 500 m, which ultimately results in the 750 m runs costing approximately 80% of the smaller-domain 720×600 500 m simulations.

Table 1 describes experiments for comparison for the February 2020 test period. In addition to the 2.5 km references, and the 750 m mentioned above, the 500 m domain is kept but only when nested within a reference

Table 1: Details of cycling experiments. The 2.5 km domain, IRELAND25, can be seen in Fig. 1. The sub-kilometre resolution domains are shown in Fig. 4: IRL750 matches the region of LIRL500. Full coupling of the LBC cloud and hydrometeors is used for the sub-kilometre experiments. For the spectral diffusion, all experiments use $RRDXTAU=123$ and $REXPDH=4$. The 2.5 km experiments () use the HARMONIE-AROME default of $RDAMP=20$ for all variables except for $RDAMPPD=200000$. The higher-resolution use the same $RDAMP$ on all. ECOCLIMAP SG is used in all experiments.*

Exp	Domain	Resolution	Vertical	Time-step	Grid	RDAMP	LBC
104	IRELAND25	2.5 km	65	75	Q	20*	IFS
106	IRELAND25	2.5 km	90	75	Q	20*	IFS
144	IRL750	750 m	90	30	C	10	IFS
149	IRL750	750 m	90	30	C	10	HAR
129	EIRL500W2	500 m	90	15	C	1	HAR

HARMONIE-AROME. As discussed in Section 3.1, this mitigates the problem of dry boundaries somewhat.

Point verification of these experiments for a selection of near-surface parameters is shown in Fig. 8. It is immediately evident that increased horizontal resolution does not necessarily yield better results, in terms of these metrics anyway. On the other hand, looking at 10 m wind-speed (bottom row of Fig. 8) suggests that the increase in number of vertical levels has a large impact. From the time-series (lower right panel), we can see that this impact comes particularly from the higher wind-speeds, which tend to be over-predicted by the 65-level experiment (in red).² Once we increase to 90 levels (green), many of the remaining results are comparable with those experiments at sub-kilometre horizontal resolution. The results also appear to suggest that it is preferable to choose the 750 m resolution on the larger domain, rather than a smaller domain at 500 m.

In terms of precipitation, there is a notable increase in the total rainfall accumulation over the February 2020 period when coupling the IRL750 run to HARMONIE-AROME LBCs as opposed to IFS (see Fig. 9), in line with the behaviour over the EIRL500W2 domain observed in Section 3.1. Rainfall totals are also increased slightly when switching from 65 to 90 levels at 2.5 km resolution. A study of the MÉRA reanalysis (Whelan et al., 2018) found a general dry bias in the mountainous west of the country, so an increase here is welcome.

It can be seen in Fig. 8 that HARMONIE-AROME coupling tended to increase the wind-speed bias and RMSE due to an over-prediction of the strongest wind-speeds in particular. However, it should be noted that this may be an inherited bias from the host model which had 65 vertical levels.

3.3 Impact on fog

The impact of increased vertical and horizontal resolution on HARMONIE-AROME fog forecasts over Ireland was also assessed for cases of interest in the two test periods, yielding four fog cases in total (06/02/2020 and 13-15/04/2021). For each case the 1200 UTC forecast on the preceding day was considered in the analysis. As discussed in the previous Newsletter (Clancy et al., 2021), operational HARMONIE-AROME guidance at Met Éireann frequently over-predicts fog over Ireland and the surrounding seas.

Focussing on the April 14th case, sample cloud base height forecasts comparing 65 and 90 vertical levels, with 2.5 km resolution, are given in Fig. 10. The default 65 level simulation indicates substantial regions with low cloud base (<100m, given by red in the maps) in the midlands/north of Ireland and in the Irish sea. These low visibility regions are significantly reduced in size in the 90 level run, with the sample cloud cross-sections indicating a notable reduction in the sea fog in particular. The corresponding 0600 UTC IR MSG satellite image and observations are shown in Fig. 11 for comparison; while some fog is present, it is not as widespread as suggested by HARMONIE-AROME, although it appears that the increase in vertical resolution helps to remove

²Note that the lowest model level is approximately 12 m and 5 m for 65 and 90 vertical levels, respectively. As 10 m wind speed is a diagnostic field, i.e. extrapolation/interpolation to 10 m must be carried out, it may be worth considering these differences when interpreting the results at different vertical resolutions.

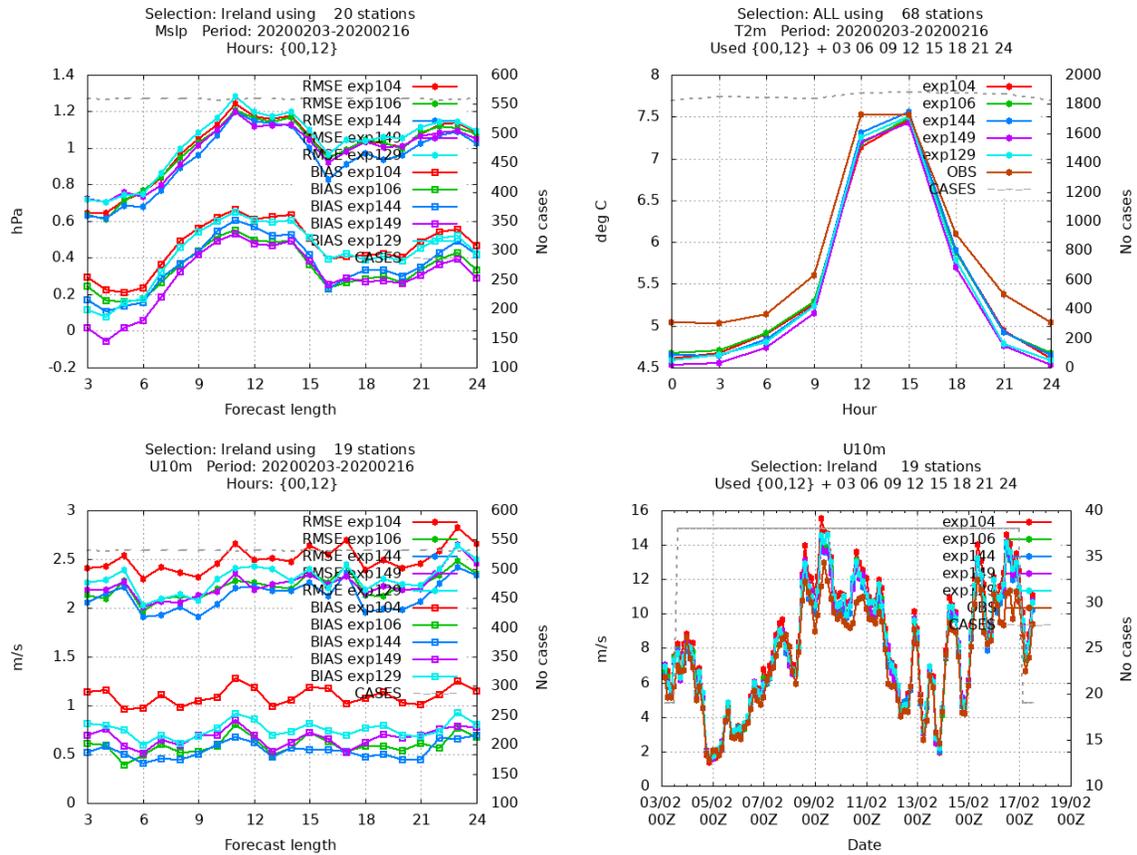


Figure 8: Point verification of the experiments listed in Table 1. Top: MSLP RMSE/bias (left) and 2 m temperature daily cycle (right). Bottom: 10 m wind-speed, RMSE/bias (left) and averaged time-series (right)

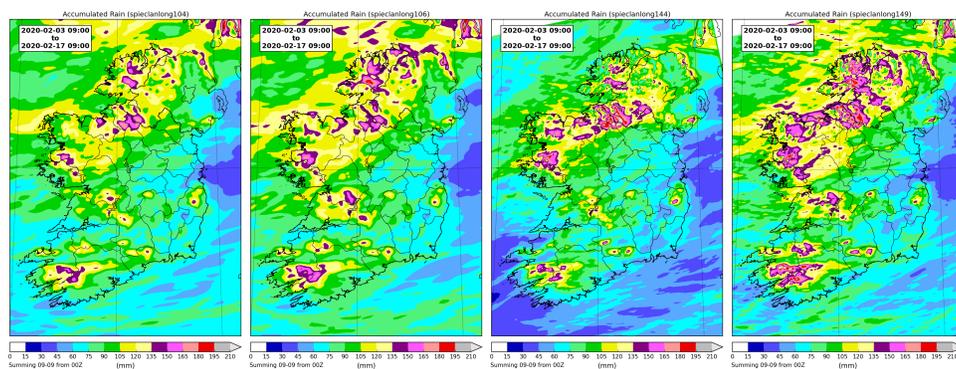


Figure 9: Accumulated rainfall for the February period, using the +9 to +33 hour forecasts from the 0000 UTC cycles to sum daily 09-09 amounts. Left to right: experiments 104, 106, 144, and 149, as described in Table 1.

some of the erroneously-forecast fog. This reduction in fog extent in the 90 level runs is consistent across each of the four fog cases considered. Both land and sea fog appear to be reduced by the increase in vertical resolution, particularly the latter.

The role played by increased horizontal resolution has also been assessed for the 750 m and 500 m resolution

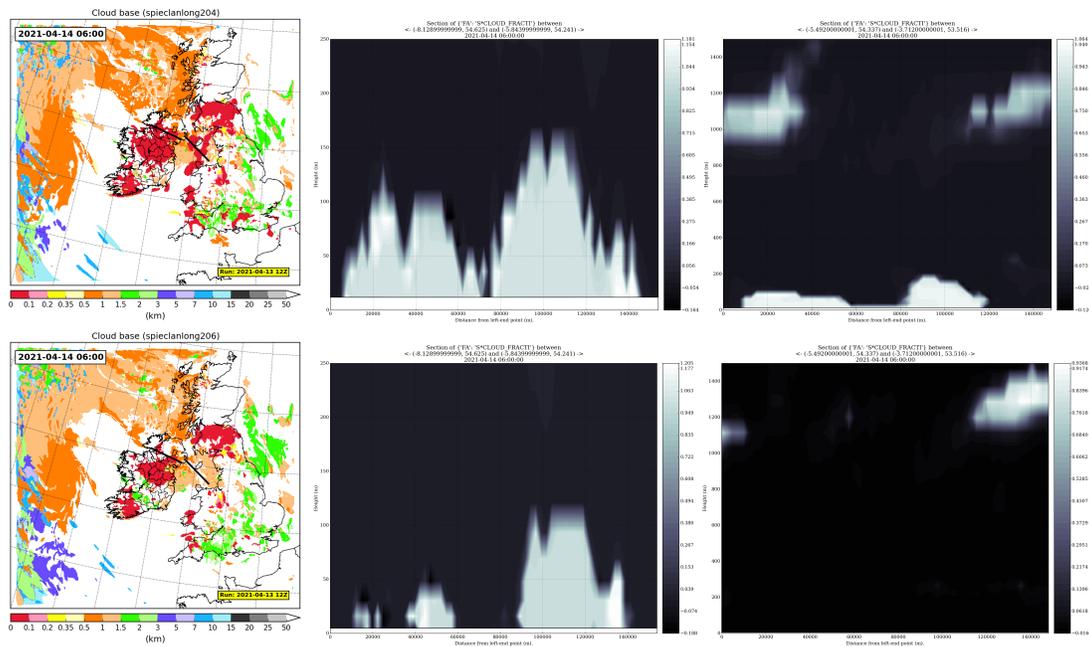


Figure 10: Left: Cloud base height forecasts over the IRELAND25 domain valid at 0600 UTC on April 14th 2021, using the 1200 UTC forecast on April 13th. Two transects, one over land and sea, are indicated by the black lines. Corresponding cloud cross-sections for (middle) the transect over Ireland, between 0 and 250 m above the surface, and (right) the transect over the Irish Sea, between 0 and 1500 m above the surface. The 65 and 90 level experiments are above and below, respectively, with configurations given by 104 and 106 in Table 1.

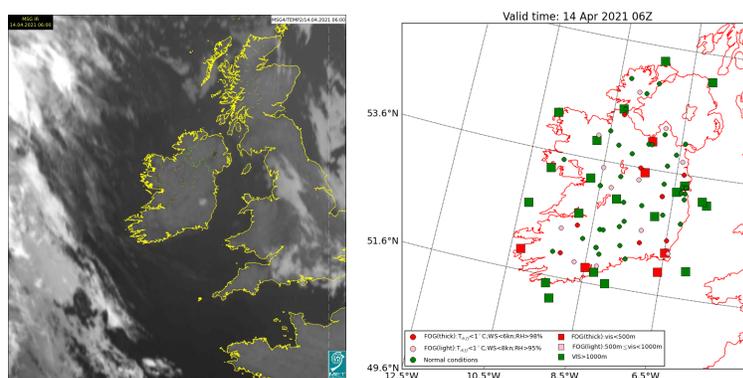


Figure 11: MSG IR satellite image (left) and observations (right) valid at 0600 UTC on April 14th 2021.

experiments. Total fog frequency is compared by constructing visibility distributions over the entire forecast length, as illustrated in Fig. 12, for all cases available. Here the 2.5 km and 750 m runs have been regridded onto the EIRL500W2 grid using the "regrid" transformation in harp's "read_forecast" function, and all gridpoints over land and sea are considered in the histograms. Mixed behaviour can be observed over the four cases, with the higher-resolution runs leading to slightly less, similar, or significantly more fog relative to the 2.5 km run with 90 levels. It is also interesting to note the impact of switching from the smaller EIRL500W2 domain to the larger IRL750 domain, which tends to reduce fog frequency overall, again highlighting the importance of domain size effects.

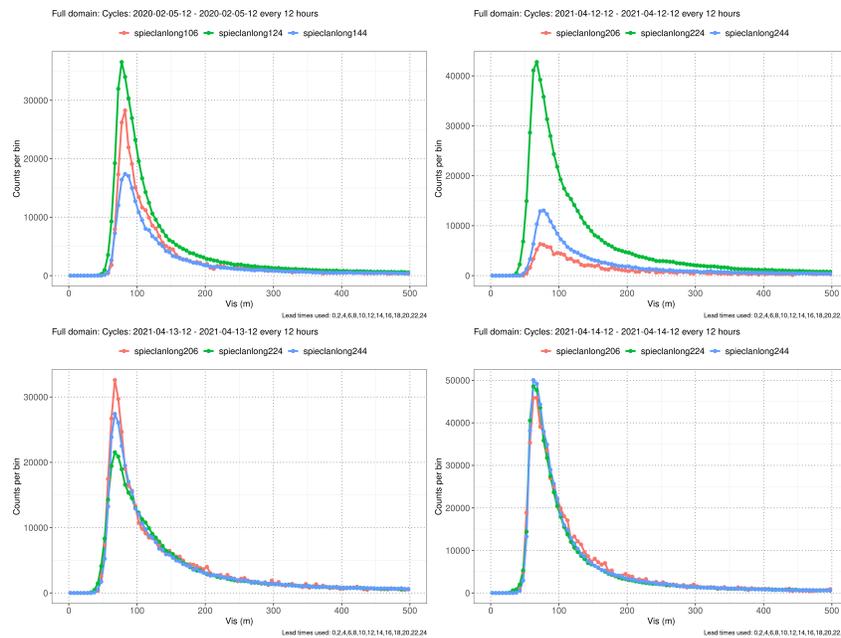


Figure 12: Visibility histograms from the 1200 UTC forecasts on (from top left to bottom right) 5th February 2020 and 12th-14th April 2021. The red, green, and blue lines represent 2.5 km, 500 m, and 750 m resolution experiments, respectively. Experiment configurations are the same as 106, 129, and 144 in Table 1, respectively, except that full coupling with IFS LBCs are used for the EIRL500W2 run. Lead times 0-24, in steps of 2 hours, are used to construct the distributions, which are zoomed into the the lowest visibility range.

4 Summary

The suite of higher horizontal and/or vertical resolution experiments discussed in this article has highlighted a number of important aspects to consider when developing a prospective sub-kilometre scale HARMONIE-AROME model for operational guidance. LBC coupling and the domain size are both found to have a strong influence on forecast performance, particularly in relation to precipitation. The increase in vertical resolution to 90 levels appears to be beneficial overall, improving both MSLP and wind-speed scores over Ireland while also reducing fog prevalence over the small number of cases available in this study.

While our results indicate that it is preferable to use a large 750 m resolution domain as opposed to a smaller domain at 500 m, it is difficult to draw conclusions on the overall benefit of the 750 m runs relative to the 2.5 km reference at the same vertical resolution. In terms of point verification at least, increasing horizontal resolution does not simply lead to better averaged scores. However, there are of course limitations to using point verification at such resolutions, and objective assessment using spatial methods will be considered further. Preliminary results using CRPS scores over a common neighbourhood for the 2.5 km and 750 m simulations, based on the methodology described in Mittermaier (2014), are found to generally corroborate the point verification results.

Forecasts on domains of much higher resolution are likely to be of more obvious benefit for extremes, rather than for general conditions. Additional efforts will therefore be made to next assess the performance of the high-resolution models in extreme weather events, with initial work suggesting roughly similar wind-speed and precipitation forecasts for the 750 m and 2.5 km runs during Storms Ciara and Dennis in February 2020.

One should also keep in mind that Ireland is relatively flat with little complex orography, particularly in the east of the country, and as such the benefit of increased resolution to capture local orographic features may be minimal. This is in marked contrast with sub-kilometre experiments over Greenland and the Faroe Islands

carried out at DMI, in which the high-resolution model demonstrates a clear value over its 2.5 km counterpart for storm forecasting in regions of complex orography (Yang, 2019).

Finally, we note some technical issues encountered during this work. Tests with small domains to the south-west of Ireland (such as the purple in Fig. 4) failed in the Prepare_pgd stage when attempting to use ECOCLIMAP SG; version 2.5_plus had to be used instead. This was traced to an interpolation issue when a domain does not have “enough” land points, and is currently under investigation within the community.

The use of single precision was discussed in Section 2.1. In general this was successful and yielded satisfactory results. However, one problem encountered here related to nesting within another AROME experiment: it was found that the host experiment had to be in double precision. This issue remains to be investigated.

5 Acknowledgements

Computing resources for this work were provided through a Special Project at ECMWF.

6 References

Antoine S., R.Honnert and Y.Seity, Improvement of fog forecast at hectometric scales in AROME, ACCORD Newsletter 1, October 2021.

Bessardon G., C.Clancy, C.Daly, R.Darcy, E.Gleeson, A.Hally and E.Whelan, Met Éireann Updates, ALADIN-HIRLAM Newsletter 14, January 2020.

Chikhi W., A.Ambar and M.Mokhtari, Testing visibility diagnostics in AROME at high resolution, ACCORD Newsletter 1, October 2021.

Clancy C., E.Gleeson, and E.McAulfield, Met Éireann’s Physics Testing in HARMONIE-AROME CY43, ACCORD Newsletter 1, October 2021.

Feddersen H., Running HarmonEPS ensemble members in single precision. 1st ACCORD All Staff Workshop 2021. http://www.umr-cnrm.fr/accord/IMG/pdf/feddersen_accord_asw2021_singprecmb.rs.pdf

HIRLAM, Harmonie 43h2.1 Release Notes, 2020. <https://hirlam.org/trac/wiki/ReleaseNotes/harmonie-43h2.1>

Lang S.T.K., A.Dawson, M.Diamantakis, P.Dueben, S.Hatfield, M.Leutbecher, T.Palmer, F.Prates, C.D.Roberts, I.Sandu and N.Wedi, More accuracy with less precision, Q.J.R. Meteorol. Soc., 2021

Mittermaier, M., A Strategy for Verifying Near-Convection-Resolving Model Forecasts at Observing Sites, Weather and Forecasting, 29(2), 185-204, 2014

Suárez-Molina D. and J.Calvo, Very high-resolution experiments at AEMET, ALADIN-HIRLAM Newsletter 16, February 2021

Váně F., P.Düben, S.Lang, T.Palmer, M.Leutbecher, D.Salmond, and G.Carver, Single Precision in Weather Forecasting Models: An Evaluation with the IFS, Mon. Wea. Rev., 145(2), 495–502, 2017

Vignes O., Single precision in cycle 43. Joint 29th ALADIN Workshop & HIRLAM ASM 2019. http://www.umr-cnrm.fr/aladin/IMG/pdf/sp_cy43_olev_asm2019.pdf

Whelan E., E.Gleeson, and J.Hanley, An Evaluation of MÉRA, a High-Resolution Mesoscale Regional Reanalysis, J. Appl. Meteor. Climatol., 57, 2179–2196, 2018

Yang X., Sub-km HARMONIE and on-demand setup for storm forecast, ALADIN-HIRLAM Newsletter 10, Jan 2018

Yang X., TAS, an operational forecast model at hectometric scale, ALADIN-HIRLAM Newsletter 12, Jan 2019

Hectometric experiments with the HARMONIE-AROME system at IMO

Bolli Pálmason, Guðrún Nína Petersen, Sigurdur Thorsteinsson, Xiaohua Yang

1 Introduction

Iceland has a complex orography with steep mountains and narrow valleys and fjords as well as a rather extreme wind climate. As high wind speeds and wind gusts are the most common natural hazard in Iceland it is important to be able to forecast winds as well as possible. In order to achieve that it is clear that the orography and physiography needs to be properly described.

The Icelandic Meteorological Office (IMO) has since 2018 run a HARMONIE-AROME hectometric experiment for Iceland, inbedded in an operational 2.5 km horizontal resolution run, see Figure 1. The 2.5 km run is termed *harmonie* and the hectometric run *hm750m*. The *hm750m* has semi-operational status and had in the beginning a forecast length of 24 hours which has now been extended to 36 hours. In general the hectometric experiment improves on both wind and later, with tuning, temperature forecasts. Regarding precipitation, the forecasts were too dry close to the boundaries but with changes in the boundary conditions they have improved significantly. Here we describe the experiment, some verifications and results from cases as well as some future plans.

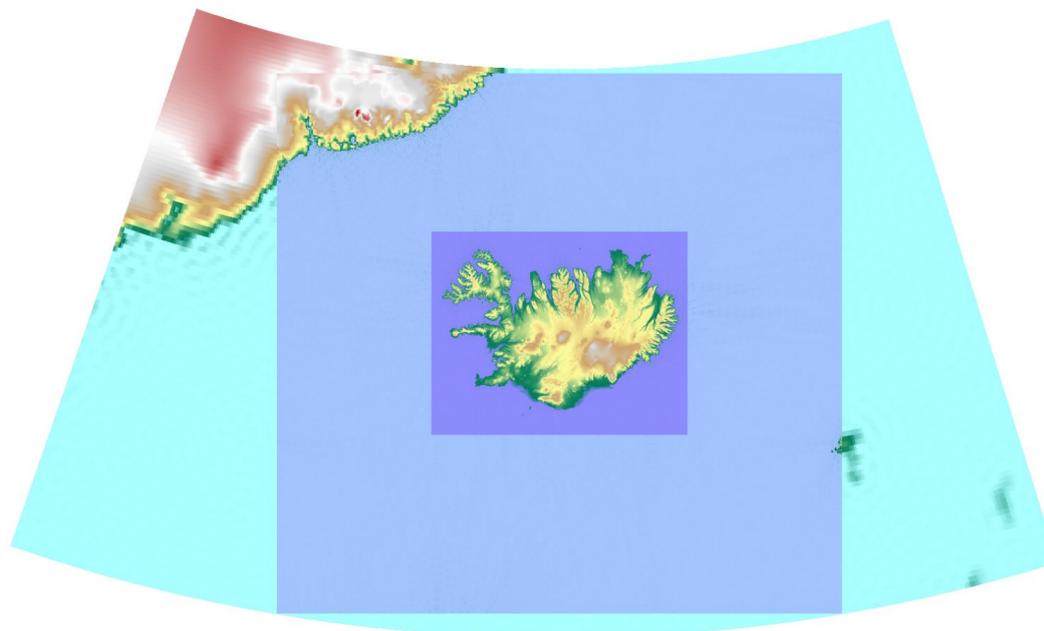


Figure 1: The Icelandic HARMONIE-AROME model domains, the *harmonie* domain in light blue and the *hm750m* domain in dark blue. Outermost area shown is from ECMWF *HRES*. The model orography is shown at their present horizontal resolution.

2. Design of parallel experiment

a. Model setup

In addition to the operational HARMONIE-AROME Iceland-Greenland domain (*IGB*, 2.5 km horizontal resolution) run jointly with the Danish Meteorological Office (DMI, Yang et al., 2018), IMO has since 2018 run two extended parallel forecasts for domains centred on Iceland. Both runs are performed with the HARMONIE-AROME model configuration, but based on different versions of the model system and have different horizontal resolution. All setups have 65 vertical model levels extending up to 10 hPa. The main characteristics of the model setup are summarised in Table 1.

Table 1: Setup of the HARMONIE-AROME modelling systems in use at IMO.

Component	harmonie	hm750m	IGB
Forecast model	cy38h1.2	cy40h1.1.1	cy40h1.1
Surface data assimilation	Optimal interpolation	Optimal interpolation	Optimal interpolation
Observation usage, SYNOP	T2m and RH2m	T2m and RH2m	T2m and RH2m
Upper-Air Data assimilation	Blending	Blending	3D-Var
Observation usage, upper-air	NA	NA	Conventional and satellites
Lateral boundaries	6-hour lagged HRES hourly coupled	harmonie coupled same analysis time	6-hour lagged HRES hourly coupled
Horizontal resolution	2.5 km	0.75 km	2.5 km
Timestep	60 s	30 s	75 s
Grid	Linear 500x480	Quad. 750x600	Cubic 1080x1280
Cycle interval	6 hours	6 hours	3 hours
Forecast launched	4 times per day	4 times per day	8 times per day
Forecast length	66 hours	36 hours	66 h (4x) / 51 h (4x)
Surface roughness over snow	0.001 m	0.003 m	0.001 m
XRIMAX	0.0	3.0	0.25
Snow emissivity	0.98	0.9	0.98

b. Experiment design

Our models are run on the ECMWF computer system (*cca/ccb*) every 6 hours. *hm750m* starts when *harmonie* has reached 36 hours forecast step and thus can provide the lateral boundary condition (LBC) for the hectometric experiment. The main difference between the two experiments, *harmonie* and *hm750m*, is the horizontal resolution although there are a few other differences, stated in Table 1:

- The runs apply different model cycles
- At the lateral boundaries *hm750m* applies *harmonie* coupled, while *harmonie* applies 6-hourly lagged *HRES* forecasts
- Tuning in *hm750m* dynamics is similar as in Xiaohua Yang's (DMI) hectometric runs over Greenland hectometric domains (Yang, 2018; 2019)
- *hm750m* has different values of RIMAX and emissivity over snow, adjusted after some testing to improve the negative temperature bias in wintertime
- *hm750m* has higher values for the surface roughness of momentum (z_0) over snow to improve on the positive wind speed bias in wintertime

In September 2021, after a series of case testing, changes were made to *hm750m* to couple all hydrometeors (ice/liquid/rain/snow/graupel, I/L/R/S/G) for LBC. This resulted in a much more realistic precipitation close to the borders of the domain, see later section. The usage of *HRES* as LBC for *hm750m* has been tested, but needs further examination.

To evaluate the relative quality of the analyses and subsequent forecasts from the different parallel experiments, we verify the forecasts against SYNOP observations within the model domain.

c. Verifications

The verification is carried out for mean sea level pressure, 2-m temperature and humidity, 10-m wind and precipitation. Figure 2 shows a scatter plot of 27, 30, 33 and 36 hour forecasts of 2-m temperature against observations for *harmonie*, *IGB*, *hm750m* and *HRES* for September 2021. The mean absolute error (MAE), the root mean square error (RMSE), mean error (ME) and correlation are shown in the upper left corner of each panel. In general, there is a negative temperature bias. *hm750m* has the best performance of the four model setups, as seen in a tighter cloud of points, the lowest ME and RMSE and the highest correlation.

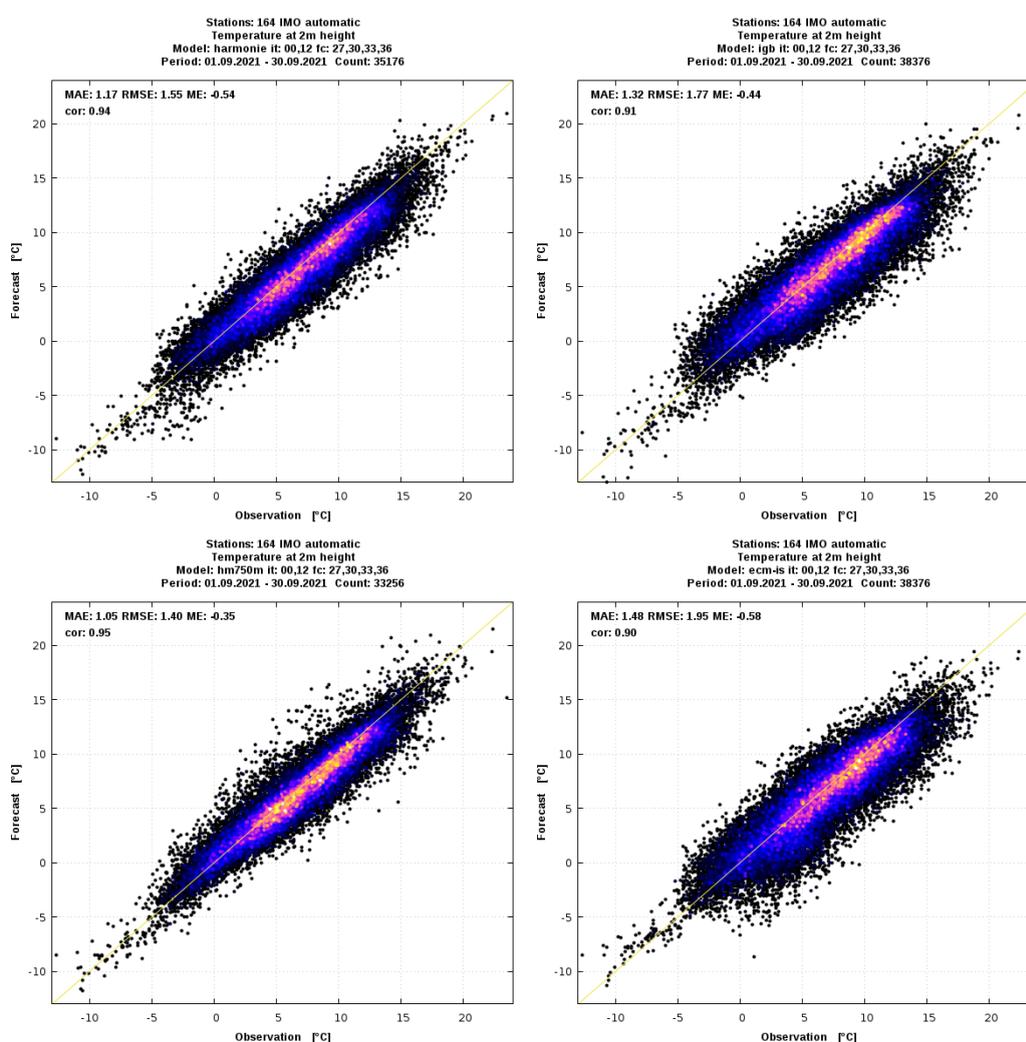


Figure 2: Scatter plots of observed (x-axis) and forecasted (27,30,33,36 hours, initial time 00 and 12 UTC, y-axis) 2-m temperature for September 2021. Top left: *harmonie*, top right: *IGB*, bottom left: *hm750m* and bottom right: *ECMWF HRES*.

Similar results are seen for 10-m wind speed, see Figure 3. In general the 2.5 km resolution runs, *harmonie* and *IGB*, overpredict wind speed slightly while *HRES* underpredicts, especially for high wind speeds. *hm750m* has a slight positive ME, the lowest RMSE and the highest correlation.

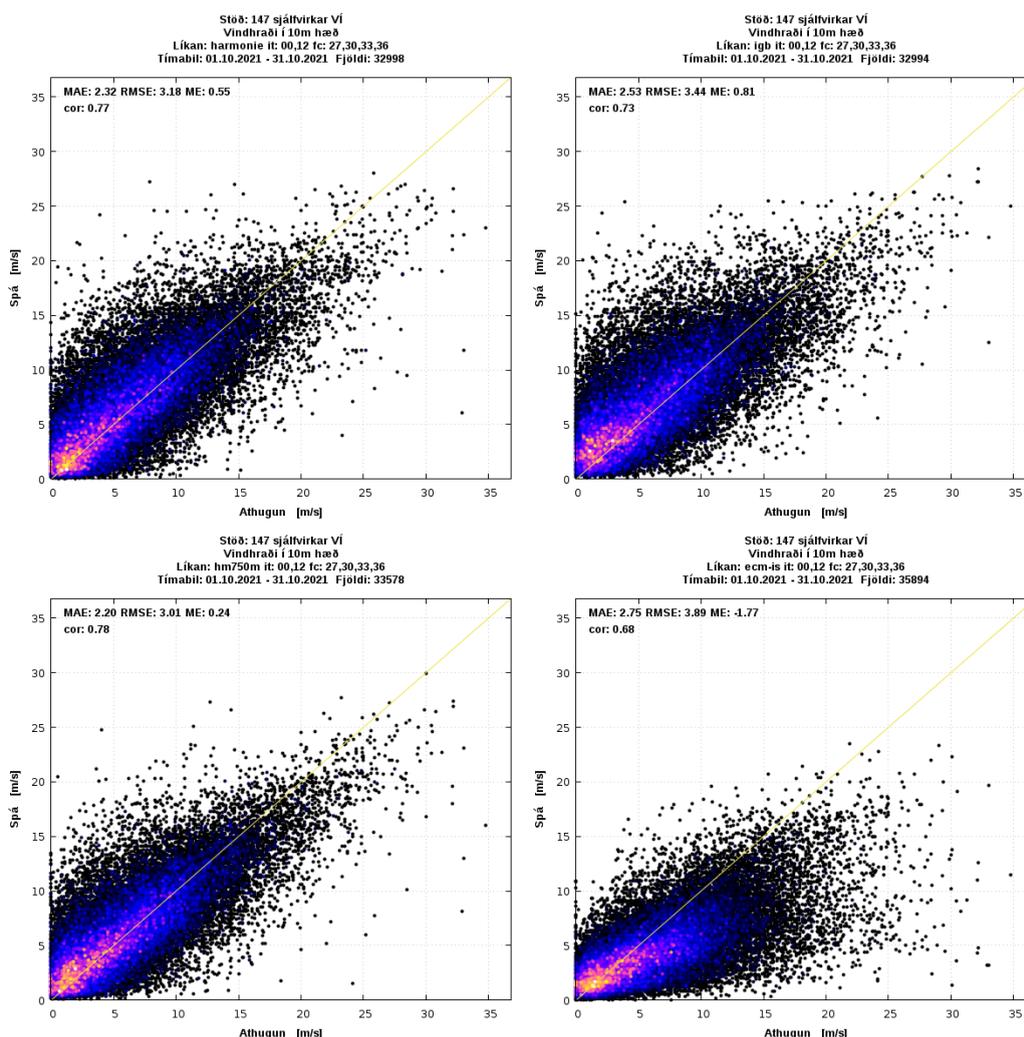


Figure 3: Scatter plots of observed (x-axis) and forecasted (27,30,33,36 hours, initial time 00 and 12 UTC, y-axis) 10-m wind speed for October 2021. Top left: *harmonie*, top right: *IGB*, bottom left: *hm750m* and bottom right: *HRES*.

Figure 4 shows the bias (ME) and RMSE verification scores for temperature, wind speed and relative humidity averaged over July 2021 and January 2022, for forecasts initialised at 00 and 12 UTC and forecast steps 6–36 hours. Again it can be seen that *hm750m* performs best for temperature and wind speed and good RMSE score are seen for humidity, but slightly negative bias compared to *harmonie*, that is applied as LBC.

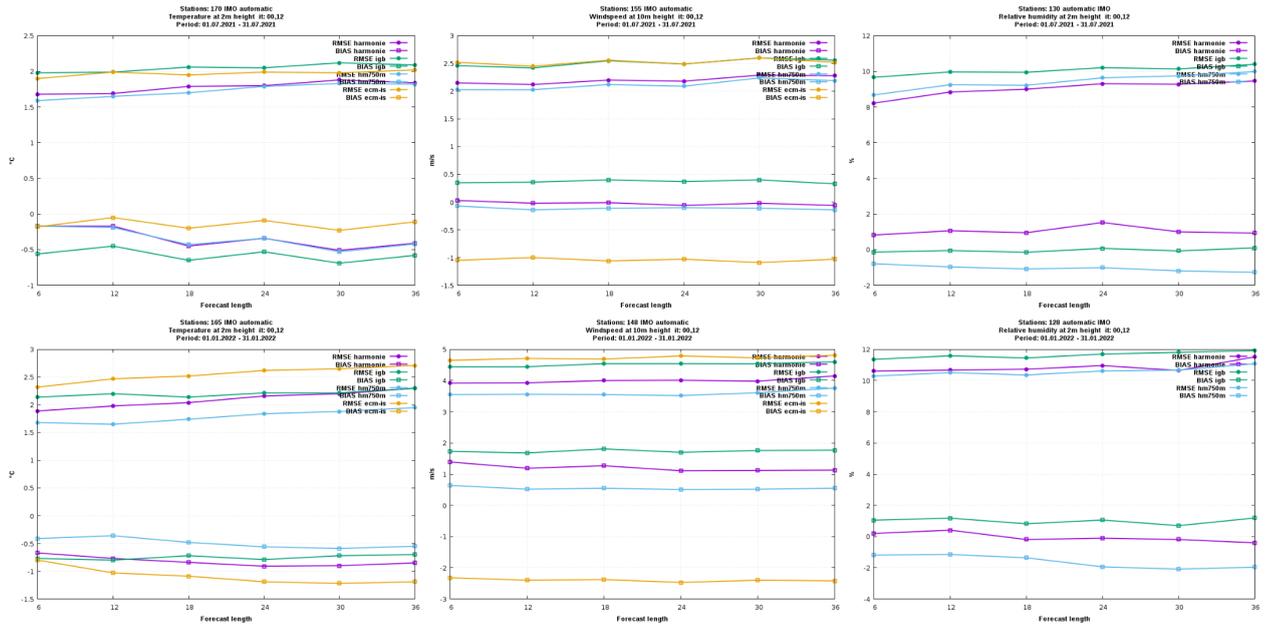


Figure 4: RMSE (circles) and BIAS (squares) as functions of forecast length for left: 2-m temperature, centre: 10-m wind speed, right: 2-m relative humidity. The forecast runs are purple: harmonie, green: IGB, blue: hm750m and yellow HRES. Upper panel: July 2021. Lower panel: January 2022

A few skill scores for 10-m wind speed thresholds are shown in Figure 5. While *hm750m* has similar POD as the 2.5 km resolution runs, *harmonie* and *IGB*, the FAR is lower, especially for wind speed above 12 m/s. Also, *hm750m* has a BIAS score close to 1 for wind speed up to 16 m/s, while the other two HARMONIE-AROME runs have a higher score indicating overprediction and *HRES* approaches zero for the highest wind speeds, indicating a significant underprediction.

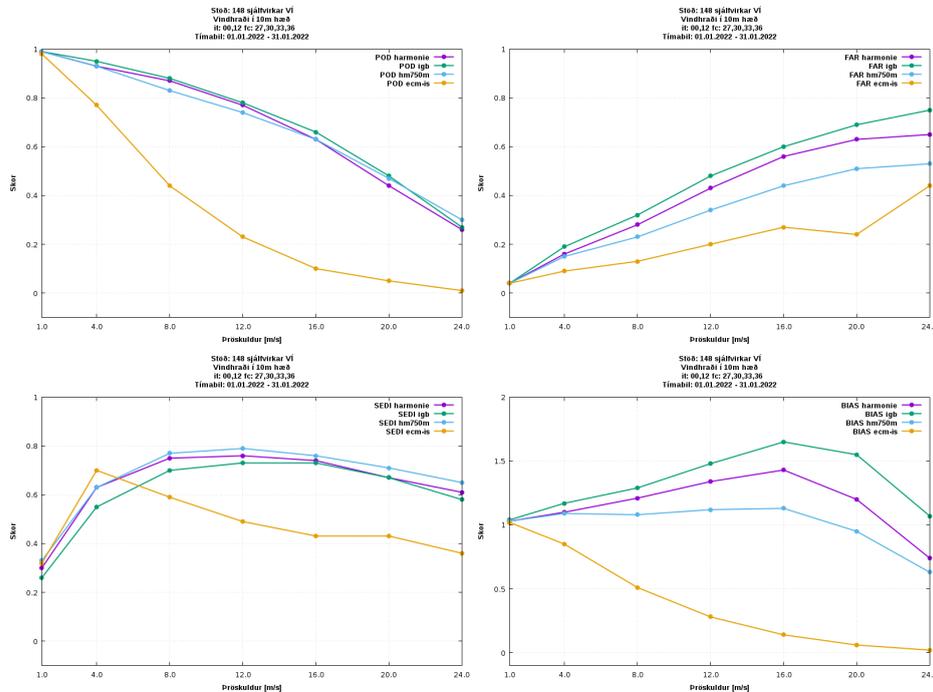


Figure 5: Skill scores for 10-m wind speed thresholds in January 2022. Top left: probability of detection (POD), top right: false alarm rate (FAR), bottom left: symmetric extremal dependency index (SEDI) and bottom right: bias skill score (BIAS). The forecast runs are purple: harmonie, green: IGB, blue: hm750m and yellow HRES.

d. Case studies

A series of single forecast experiments were conducted using HARMONIE-AROME cycle 43h2.1 at 750 m horizontal resolution. The aim of the experiments was to improve precipitation close to the boundary. This was done by including all hydrometeors in the LBC data. This was tested for both LBC from *harmonie* (hydrometeors I/L/R/S/G) as well as for *HRES* LBC (I/L/R/S). Furthermore, the SLHD diffusion was turned off, as recommended by Clancy et al. (2021) and Seity et al. (2021).

The focus of the experiment was on a real case from 28 September 2021, looking at forecasts valid at 12 UTC. A cyclone moved westward along the north coast of Iceland, with lowest centre pressure of 967 hPa. It brought precipitation to the northern part of the country, mainly as snowfall, see Figure 6. At 12 UTC the low was just north of northeast Iceland, inside the *harmonie* domain but just north of the lateral boundaries of the *hm750m* domain. *harmonie* operational experiment simulated the precipitation associated with the system realistically.

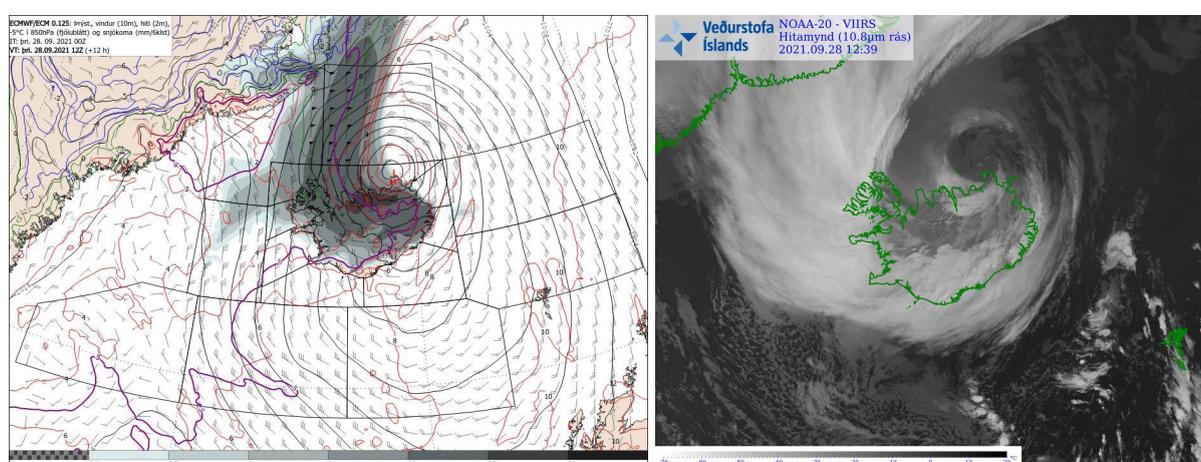


Figure 6. Left: ECMWF HRES forecast +12h, valid 28 September 2021 at 12 UTC, showing 6 h accumulated snowfall as grey shading. Right: an infrared satellite image from 28 September 2021 at 12:31 UTC.

A *hm750m* experiment, with HARMONIE-AROME Cy43h2.1, coupled with *harmonie* LBC without hydrometeors, initial time at 06 UTC, valid at 12 UTC, had a precipitation band with similar orientation and intensity over land but was systematically missing precipitation over the sea, close to the northern boundary, as well as in the northwesternmost part of the country, see Figure 7, left. However, a run with all hydrometeors showed little difference over land, while over the sea to the north the precipitation was in better agreement with *harmonie*, *IGB* and *HRES* (Figure 7, right).

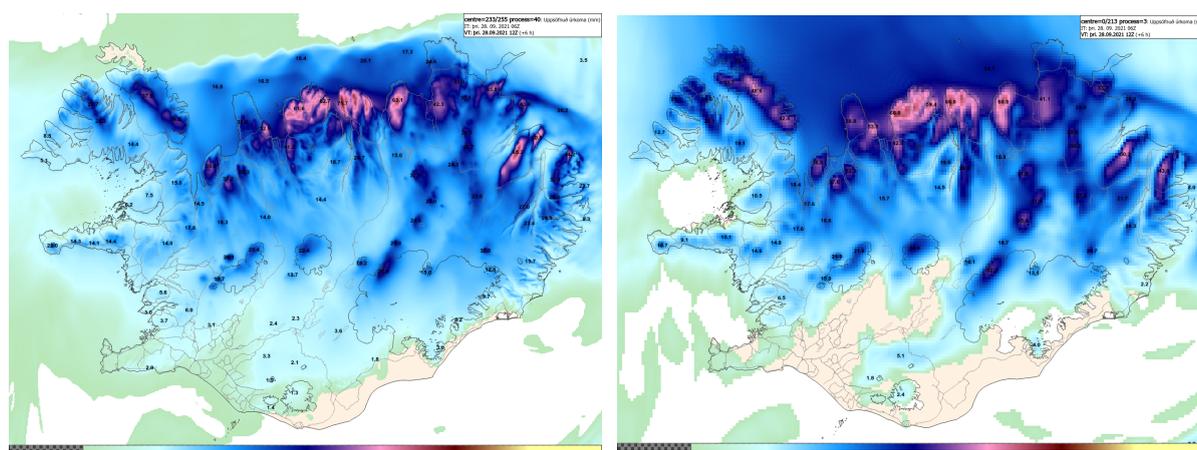


Figure 7: 6 hour accumulated precipitation in *hm750m*, 6 hour forecast valid 28 September 2021 at 12 UTC (*HARMONIE-AROME Cy43h2.1*). Left: No hydrometeors coupling and right: with hydrometeors coupling.

Due to the fact that *hm750m* in general outperforms *harmonie* and *IGB*, the IMO forecasters would like for the forecast length to be extended but would also like it to arrive earlier in house. However, as currently the *hm750m* forecast cycle has to wait until *harmonie* has reached the forecast step of 36 h there are limits to how early the hectometric run is accessible. It was therefore also tested if *hm750m* could be run with *HRES* as LBC. An experiment was conducted applying *HRES* as LBC, with all available hydrometeors (I/L/R/S). Figure 8 shows a comparison of 12 hour accumulated precipitation in *HRES* and 6 hour accumulated precipitation in *hm750m*, both valid at 12 UTC. While the precipitation over land is realistic in *hm750m* there is a strange local maxima close to the northern boundary, neither seen in *HRES*, *IGB*, *harmonie* nor *hm750m* coupled with *harmonie*. The use of *HRES* as LBC for hectometric experiments thus needs some further investigation into LBC matters.

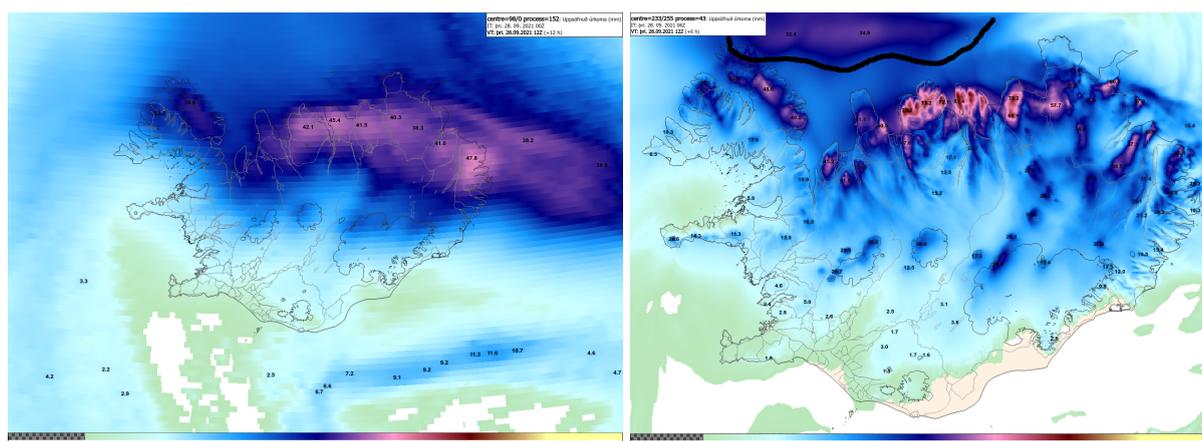


Figure 8. Accumulated precipitation, valid 28 September 2021 12 UTC. Left: *HRES*, 12 hour accumulated precipitation and right: *hm750m* (*HARMONIE-AROME Cy43h2.1*) with *HRES* as LBC and all available hydrometeors activated, 6 hour accumulated precipitation. Note the local maxima close to the northern boundary, marked with black line, neither seen in *HRES* nor *hm750m* coupled to *harmonie*.

3. Conclusions and future plans

The functionality of different components of the hectometric HARMONIE-AROME setups have been investigated along with their performance. Our confidence in this setup has increased with each month of verifications and different weather situations. For almost every calendar month we see improved scores with the 750 m setup compared to other models we use, but of course no precipitation verification is shown here - keeping in mind that point verification of precipitation will be very unfavourable to the much higher horizontal resolution of the hectometric run. However, despite the lack of precipitation close to the borders before the coupling of all hydrometeors, the experience of our forecasters is that “precipitation shadows” on the lee sides of mountains are more realistic in *hm750m* than the coarser model setups.

With availability of new HPCs, both at ECMWF in Bologna and at United Weather Centre - West (UWC-W) in Iceland, we plan to increase the size of the domain and test 90 vertical levels. Further tests are needed to fix issues we see with direct *HRES* coupling with all hydrometeors as well as other possible issues with the LBC. The gain in timeliness of the hectometric experiment using *HRES* compared with *harmonie* as LBC is of importance to our forecasters. We also plan to start tests soon with *cy43h1.2* and high resolution ECOCLIMAP-SG physiography for the domain - currently it's using ECOCLIMAP-II with local modifications.

2 References

- Colm Clancy, James Fannon and Eoin Whelan (2021): HARMONIE-AROME at Hectometric Scale. Met Éireann, NWP Note 2022/02.
- Xiaohua Yang (2018). Sub-km HARMONIE and on-demand setup for storm forecast. ALADIN-HIRLAM Newsletter 10, January 2018.
- Xiaohua Yang (2019): TAS, an operational forecast model at hectometric scale. ALADIN-HIRLAM Newsletter 12, January 2019.
- Xiaohua Yang, Bolli Palmason, Kai Sattler, Sigurdur Thorsteinsson, Bjarne Amstrup, Mats Dahlbom, Bent Hansen Sass, Kristian Pagn Nielsen and Guðrún Nína Petersen (2018): IGB, the Upgrade to the Joint Operational HARMONIE by DMI and IMO in 2018. ALADIN-HIRLAM Newsletter 11, August 2018.
- Yann Seity, Salomé Antoine, Rachel Honnert and Benoit Vié (2021): First ACCORD All Staff Workshop 12 - 16 April 2021.

The ACCALMIE project: coordinated approach for chemistry and aerosols in Météo-France's models, inline and offline.

Vincent Guidard, Quentin Libois

1 Introduction

Aerosols and the chemical composition of the atmosphere are essential elements of natural and anthropogenic atmospheric processes, both for surface fluxes and through their various interactions with each other and with other compartments of the Earth system. In addition to their life cycle, such as the movement of volcanic ash or desert dust, aerosols interact with gaseous compounds, with water droplets and ice crystals in clouds, and with atmospheric radiation. Aerosols thus have direct (absorption and scattering of radiation), semi-direct (thermal impact of radiation absorption) or indirect (interactions with microphysics) effects. The chemical composition of the atmosphere also has an impact on atmospheric radiation. It is therefore essential to take aerosols and gas composition into account when representing the Earth system.

In fact, atmospheric composition is at the heart of many of Météo-France CNRM's research activities. It is also present in the operational chains developed or co-developed in the laboratory. In particular, it is taken into account in the various atmospheric models used at the CNRM, which represent it explicitly or are based on climatologies. Some of these models are global (ARPEGE NWP, ARPEGE-climate and CNRM-ESM climate, MOCAGE CTM), others are regional (AROME NWP, ALADIN-climate, AROME in climate mode, CNRM-RCSM, MOCAGE CTM, Meso-NH research community). The modelling of aerosols and gaseous compounds in these various models has been developed at different times, generally in response to specific scientific or technical issues. These developments have therefore taken place without any real coordination between models. Some developments took place in the framework of theses or post-doctorates, but were only minimally maintained afterwards, due to a lack of expertise on the subject. The time scales for developments and their outcomes are not always the same from one CNRM group to another, each being subject to timetables sometimes imposed by actions external to the CNRM (joint IFS ARPEGE cycles, CMIP exercises, etc.). The strength of the permanent staff (which alone allows for medium- and long-term monitoring) has not always been sufficient to initiate rationalisation and real sharing of expertise between the applications and groups. Nevertheless, there have been some reflections in recent years aimed at synthesising what exists. The new scientific challenges to be met (modelling of an earth system, among others) to respond to both research and operational problems, on climate, process knowledge or real-time numerical forecasting, in a context that is not very favourable in terms of permanent human resources, have led the CNRM management to launch a preliminary project to define the outlines and methods for making progress on these challenges.

2 State of the Art at CNRM

As mentioned in the introduction, the modelling of atmospheric composition at the CNRM is based on different models and meets various objectives. First of all, the MOCAGE chemistry-transport model allows the modelling of aerosols and chemical compounds in the troposphere and stratosphere. This model is used for global and regional operational applications of air quality and atmospheric composition, as well as in research for climate change impact studies. As MOCAGE is a chemistry-transport model, it needs meteorological information (wind, pressure, temperature, humidity, cloud

variables) to be provided by an external model (which can be an NWP or climate model). The chemical scheme used is a combination of the RACM (tropospheric) and REPROBUS (stratospheric) schemes, named RACMOBUS, allowing to represent about 100 species and 400 chemical reactions. The interactions between primary aerosols and gases leading to the formation of secondary inorganic aerosols are described by the ISORROPIA module. Research is underway to use the parameterisation of Spracklen et al. (2011) for secondary organic aerosols, as well as the inclusion of the SSH-aerosols module (developed by CERE) which allows different levels of complexity in the representation of secondary aerosols. Emission fluxes (dynamic or from inventories) and deposition fluxes are computed in a tool integrated to MOCAGE called SUMO, upstream of the MOCAGE run, from the fields of the forcing model and surface properties. If the forcing model is ARPEGE or AROME, the surface fields come from calculations made in SURFEX.

An interactive, i.e. 'on-line', aerosol scheme called TACTIC (Tropospheric Aerosols for Climate In CNRM) was introduced and validated in ARPEGE-Climat and ALADIN-Climat in the early 2010s (Michou et al. 2015; Nabat et al. 2015b). TACTIC is largely derived from the aerosol scheme introduced in IFS in the framework of the European projects GEMS, then MACC and MACC-II (Morcrette et al. 2009). The development of TACTIC is the result of close collaboration with Jean-Jacques Morcrette and has benefited greatly from the fact that the ARPEGE, IFS and ARPEGE-Climat codes share a broad base. This collaboration continued in the framework of the CAMS₄₃ project, then CAMS2₃₅. Aerosols are taken into account by means of prognostic variables (between twelve and eighteen depending on the configuration), according to a sectional distribution based on their size and their hydrophobic or hydrophilic character. This scheme thus makes it possible to represent the main types of aerosols and their impacts on radiation, clouds and climate. It has been used in particular in CNRM-ESM2-1 for the last CMIP6 exercise (Michou et al. 2020), as well as for numerous studies on a regional scale in the Mediterranean (Drugé et al. 2019, Nabat et al. 2020) or in southern Africa (Mallet et al. 2019). The dynamic emission fluxes of primary aerosols are calculated in the atmospheric model from information coming from SURFEX. Other emission fluxes are provided by global inventories provided in the framework of international projects such as CMIP6. In addition, ARPEGE-Climat also includes a gas chemistry scheme (REPROBUS), which allows 63 chemical species with 168 chemical reactions in the stratosphere and troposphere above 560 hPa to be represented interactively. There is also a linear version that allows the simple representation of stratospheric ozone evolution.

The Meso-NH model can be coupled to the ORILAM model for the description of inorganic and secondary organic aerosols and their evolution. Surface exchange fluxes (desert dust and sea salt emission, deposition) are provided by SURFEX. A coupling of the MEGAN biogenic emission model with SURFEX has been performed in the past.

The AROME model has benefited from the developments made in Meso-NH by the connection of ORILAM in AROME. This configuration is called AROME-dust, and has been validated mostly for the representation of desert dust

Some pioneering work was carried out from 2011 onwards to bring together the RACM and REPROBUS schemes in a library called SUGAR (*Système Unifié pour les Gaz et Aérosols Réactifs*, Unified system for reactive gases and aerosols), which therefore allows the RACMOBUS approach. This library is used in the operational configurations of MOCAGE CTM and IFS (MOCAGE) since 2013. However, the use of this library was not considered relevant for climate models at that time.

3 Towards a common library

The desire to create a common aerosol-chemistry library (BAC, *Bibliothèque Aérosols Chimie*) stems from the fact that aerosols and chemistry are treated separately in the various models used at CNRM, even though some of the processes modelled are identical. Converging towards a common library would make it possible to access all the existing tools at CNRM from all the models. This will facilitate the pooling of skills that are sometimes confined to one model. This should also facilitate the dissemination of advances to partners and their contributions to our common research. This library would include a single interface to access all existing routines concerning the physical chemistry of aerosols. The different atmospheric models would then call this interface (several times per time step if necessary) and it would redirect to the appropriate routines according to the options chosen by the user (in namelist for example). In order for the routines contained in the library to continue to be used outside of BAC (for example via IFS), they must be able to be integrated without their call structures being modified. This single interface will thus cover needs from the simplest to the most complex, and must also cover all spatial resolutions, from LES to climate scales. Interfaces have already been set up in IFS to pool calls to the different chemistries available. It will be useful to see if what has been done in this framework can be used as a starting point for writing the LAC interface. Around this library, everything concerning surface fluxes (emissions and deposits) will be included directly in SURFEX. The call to this unique interface will require the coding of new corresponding interfaces in each of the models. These interfaces will make it possible to convert the quantities of the host model into those recognised by BAC. They will also ensure that the outputs of BAC are translated into useful information for the radiative and microphysical schemes, and vice versa. Finally, all exogenous data (aerosol optical properties, reaction rates, physico-chemical parameters, etc.) will be gathered in the same directory, and if possible in netcdf files that can be modified without recompiling the code.

Surface fluxes in SURFEX

Aerosol emission processes at the surface are currently represented both in SURFEX and in atmospheric modules that use the SURFEX surface fields to parametrise these emissions (e.g. TACTIC). The aim is to bring together everything related to emissions and deposition in SURFEX, and for all models to make explicit use of SURFEX. This will involve copying certain routines into SURFEX. As SURFEX is now common to all models (except IFS), this will allow for a pooling of emission schemes. For the processes included in SURFEX that depend on aerosol fluxes (snow albedo in Crocus, lake turbidity, etc.), it will be necessary to convert the information coming out of BAC into useful information for these processes.

Chemistry-Aerosols library (BAC)

The BAC library must bring together all the routines that describe the physicochemical evolution of aerosols and gases. Knowing the thermodynamic state of the atmosphere and its chemical composition at time t , it must be able to estimate the chemical composition at time $t+dt$, where dt is the dynamic time step of the model. It must therefore include all the chemical reactions that can take place in the atmosphere, which describes the physico-chemical ageing of aerosols, as well as the processes that allow aerosols to be formed from gases, and vice versa. On the other hand, microphysical changes in atmospheric composition (nucleation, leaching, etc.) will be treated outside BAC when the microphysical scheme allows it. Otherwise, an ad hoc parametrisation will allow this to be done directly in BAC. Whatever the refinement of the model in terms of the representation of the chemistry and aerosols, BAC will be called. The BAC interface must therefore be able to handle a wide range of information, from a few prognostic aerosols to dozens of chemical species. It is at the level of this interface that the selection of the routines that are actually called will be made. It will thus be useful to define an aerosol table and a gas table whose "species" dimensions will depend on the choice made in

the host model. The host model will pass an aerosol table and a gas table as input to BAC. It will also have to transmit a table of values specifying to which species each "species" index in this table corresponds. This type of vector exists in IFS-ARPEGE under the name of GFL structure, which is a modular structure allowing to manage additional fields to the usual thermo-dynamic and microphysical fields of these models.

To avoid activating species one by one, which would require endless namelists, it will be possible to define beforehand some configurations corresponding to levels of complexity relevant to various applications. For example, aerosols and chemistry should be activated separately. One could also imagine activating only desert dust, a dozen or so primary aerosols or all possible aerosols, a few key gases or, on the contrary, a whole predefined list of gases. It would then be appropriate to group individual routines into groups of routines within BAC. It is also possible that all the routines concerning the species received as input are activated, but this would make it impossible to distinguish between species that are chemically evolving and those that would only be transported, for example. These details will be decided at the time of implementation of the project.

It will be necessary to consider the order in which the physical routines are called. BAC can be called several times, before or after radiation and microphysics, among others. Making a choice as close as possible to what already exists will limit the differences between the current model configurations and those using BAC.

4 Implementation, governance and schedule

Structure of the project

a. Definition

A large part of the work will be based on a phase allowing all CNRM teams to participate in the project to identify the processes that already exist and that need to be collected either in SURFEX or in BAC. This will be followed by the definition of the interfaces of the host models with SURFEX on the one hand and BAC on the other. The time invested in this definition phase should make it possible to gain in efficiency during the implementation phase. All the teams involved in the project will be represented, in order to find the broadest possible consensus on the technical choices. This definition phase is essential and will consume a significant part of the forces dedicated to this project.

This part will take place between February and September 2022.

b. Implementation

This very software-intensive phase will see coding along three axes: surface fluxes in SURFEX, the constitution of BAC and the construction of host model interfaces with BAC and SURFEX. Unit tests will be carried out during implementation to ensure consistency with existing schemes. The surface fluxes and BAC will be coded first in a transversal way to the host models. An interface routine will be proposed in BAC and the calls and interfaces to BAC and surface flows will be adapted in each of the host models. The SURFEX documentation will be updated as a result of the developments and a BAC documentation will be written and maintained.

This part will take place between July 2022 and August 2023

c. Validation

The validation of the surface fluxes developed in SURFEX will be done offline to start with, in order to compare them with the various flows calculated in the pre-existing tools of the project. The BAC calls and the models with online chemistry-aerosols will then be validated. Case studies will be defined to compare the original versions of the models with the version using LAC. Documentation of these validation steps will be prepared.

This part will take place between November 2022 and January 2024.

Internal organisation at CNRM

This definition of the project highlights the need for the project to work closely with the SURFEX team. A focal point in the project team (Joaquim Arteta) will be in permanent dialogue with a member of the SURFEX team (Marie Minvielle). In addition, as the development phases in the atmospheric models are fairly extensive in time, the technical leadership of these developments is provided by a pair of experts (Béatrice Josse and Pierre Nabat).

From the point of view of project management, the pair that leads the project will hold project monitoring meetings to ensure that the project is progressing nominally, to identify any stumbling blocks and to liaise with the technical facilitation. The pair will present the progress of the project to the CNRM Steering Committee at its request. Presentations will be made at the CNRM General Assembly.

Links with partners outside CNRM

The work carried out at CNRM on atmospheric composition is largely carried out in collaboration with external partners. The transition of our tools to BAC should in no way complicate our relations with these partners. On the contrary, it should facilitate the inclusion of tools of various origins. To this end, the two pilots will interface with external partners who are interested in the project, without being able to provide a work force for its implementation. These are in particular the French labs LAERO and LACy. ECMWF also has a special role, as a majority of the developments carried out in the project will require modifications in the common IFS-ARPEGE code. The goals of the project and its progress will be presented at the IFS-ARPEGE coordination meetings. Finally, the ACCORD partners also share the IFS-ARPEGE code and have started projects on aerosols and chemistry in AROME (AROME-dust, PH6.1 work-package, etc.). The ACCORD project manager has been contacted to harmonise ACCALMIE development plans as much as possible with the ACCORD work-plans. The Meso-NH community is also interfacing with the project. The resources available internally at the CNRM are very limited, in particular because of other internal actions. But the partners, from LAERO in particular, are available to help the project progress and eventually feed the library with other schemes.

EPS research and development in RC-LACE in 2021

Clemens Wastl, Martin Belluš, Gabriella Szépszó

1 Introduction

Currently, three ensemble systems (A-LAEF, C-LAEF, AROME-EPS) are running in full operational mode within RC-LACE. Major upgrades in the operational systems have been undertaken in 2021 for C-LAEF in Austria (cy43t2, surface perturbation scheme, 3h assimilation cycle) and AROME-EPS in Hungary (cy43t2, 1h coupling). The next upgrade of the common A-LAEF system (e.g. new model cycle, assimilation cycle, etc.) is planned with the new ECMWF HPC in 2022.

Several research and development activities have been made in the area of ensemble prediction within RC-LACE in 2021. For the common LACE A-LAEF system a new Mediterranean Sea domain has been set-up for ocean model coupling and some new visualizations and EPS maps have been developed which are displayed on the common webpage. For the C-LAEF system in Austria a new surface perturbation scheme has been developed and tested, the parameter perturbation scheme (SPP) has been extended and new visualisations (EPSgrams, etc.) have been set-up. EPS related work in Hungary has been partly devoted to the preparation of the operational upgrades and partly to further testing the local perturbations.

2 Operational systems upgrades

A-LAEF:

The A-LAEF (ALARO - Limited Area Ensemble Forecasting) system runs as a time critical application (TC2) at the ECMWF HPCF since July 2020. It is a sequel to the former ALADIN-LAEF system developed within the RC LACE cooperation in 2011. New A-LAEF system has significantly increased horizontal and vertical resolution (4.8 km/60 L), and involves new perturbation techniques. The A-LAEF scripting system has been re-written from scratch, utilizing ecFlow workflow package. The ensemble of surface data assimilation ESDA (Belluš et al., 2016) simulates initial condition uncertainty, together with the upper-air spectral blending by digital filter initialization (Derková and Belluš, 2007). The model errors are simulated by four different ALARO-1 parameterization configurations for turbulence, microphysics, deep and shallow convection, and radiation. The advantage of used ALARO-1 physics is their seamless functionality on the horizontal scales between 2 and 10 km (Termonia et al., 2018). Therefore, the A-LAEF system can simulate the uncertainty on meso-synoptic scales. A stochastic perturbation of physics tendencies for the surface prognostic fields (Wang, Belluš, Weidle, 2019) further enhances model spread obtained by the multi-physics. The ensemble comprises 16 perturbed members and one control run coupled to the ECMWF ENS in lagged mode. It is available twice a day (00 and 12 UTC runs) for the next 3 days. The A-LAEF multi-GRIBs containing model fields for surface and pressure levels are regularly disseminated via ECPDS to the RC LACE members and partners - Croatia, Czech Republic, Romania, Slovakia, Slovenia, (Poland via SHMU), and Turkey.

During 2021, several important upgrades took place in the A-LAEF operations. A big MSEA domain for ocean models coupling NEMO and SHYFEM has been implemented. One can see the A-LAEF integration domain and all its post-processing domains in Figure 1. The coupling system was upgraded twice. Firstly, to adopt the IFS/ENS cy47r2 input data containing 137 vertical levels instead of former 91 (and some changes in the availability of the parameters needed for c901 configuration). Secondly,

to the IFS/ENS cy47r3 inputs with mostly the meteorological impact. The OBS backup used in ESDA assimilation was implemented, utilizing newly generated GTS data files prepared at SHMU and uploaded hourly to the OPLACE database (this was successfully used in the operations twice, since the implementation). The obsoul_merge tool, used not only in the A-LAEF assimilation, was upgraded in order to add whitelisting possibilities, zipped files support, and most importantly strict formatting mandatory since cy46. The incremental DFI setup in spectral blending procedure was coded in order to eliminate the occasional spin-up issues, but never used operationally due to increased SBUs consumption and overall small impact. Furthermore, regular upgrades and enhancement of the ecFlow tasks took place, and latest version of ECMWF software packages (eccodes, cdo, and ecFlow) were utilized as they became available.

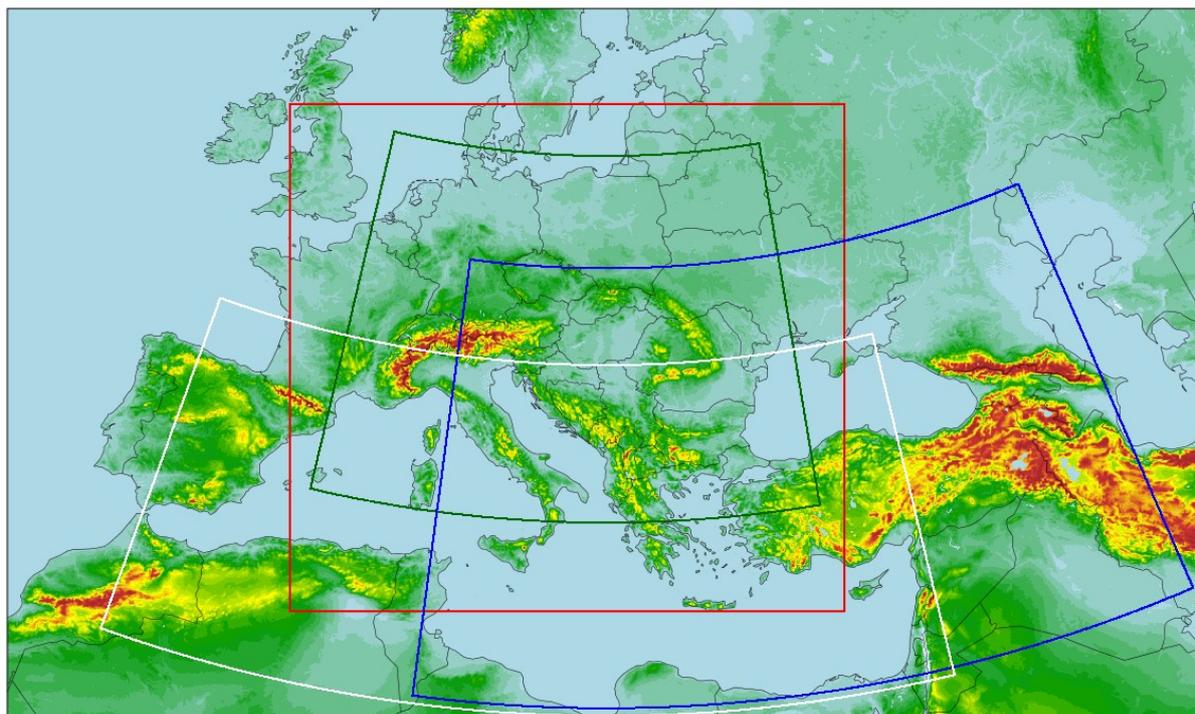


Figure 1: The A-LAEF integration domain with model orography, and all post-processing subdomains (LACE - red, CZ - green, MSEA - white, and TR - blue).

C-LAEF:

The C-LAEF system is running in full operational mode (TC2 application) at the EMCWF HPC since November 2019. A major upgrade of the system has been made on December 6th 2021 when the model cycle has been changed from cy40t1 to cy43t2. With this upgrade also the number of runs per day has been increased from 4 to 8 which goes along with an upgrade of the assimilation cycle from 6 hours to 3 hours. Due to limited computer resources only the 00 and 12 UTC runs are kept long with lead times of 60 and 48 hours respectively, while the short intermediate runs (3 hours lead time) are just to keep the assimilation cycle running. Extensive tests and verification (Figure 2) have shown that especially for situations with high convective activity and for weather systems with rather small scale and fast changing features, the higher frequency of assimilations has a positive impact on the model scores.

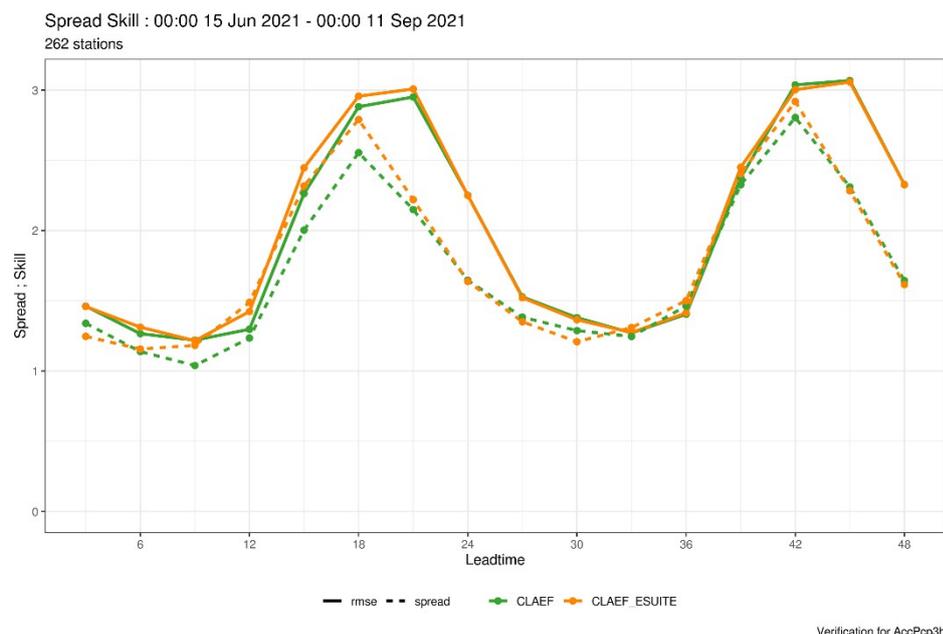


Figure 2: 3h accumulated precipitation verification for summer 2021. The green lines show ensemble spread (dashed) and RMSE (full) for the operational C-LAEF system (cy40t1), while the orange lines refer to a C-LAEF Esuite based on cy43t2.

The perturbation scheme in the operational C-LAEF version comprises perturbations of initial conditions (observation perturbations near the surface and in the upper air; ensemble JK), lateral boundary conditions (coupling with different ECMWF-ENS members) and a combination of tendency and parameter perturbations for the representation of model error (Wastl et al., 2021). Long-term verification of the operational C-LAEF suite has shown that the system is suffering from a significant lack of ensemble spread, especially for surface parameters like T2m and RH2m. To improve the situation the surface perturbation scheme of Météo France (Bouttier et al., 2016) has been implemented and tested in a C-LAEF Esuite during a winter and summer test month. In this scheme (activation by switch LPERTSURF) several surface parameters are perturbed stochastically at the beginning of each model integration. This means that the output file of the surface assimilation (CANARI in our case) is perturbed by the external routine pertsurf.F90. In case of Météo France the surface analysis of the unperturbed control (member 00) is used in all members and this analysis is then perturbed by LPERSTURF with different seeds in each member. When we verified first runs with this scheme and compared it with the operational C-LAEF (without surface perturbation) we found a significant reduction of spread in the parameters near the surface. It turned out that this comes from using the same surface analysis in all members of the e-suite. In the operational C-LAEF version the surface analysis is made separately for each member with own observation perturbation of T2m and RH2m. To overcome this problem we have modified the Météo France LPERTSURF scheme accordingly:

- Seasonal or constant fields (vegetation index, vegetation heat coefficient, leaf area index, land albedo, land roughness length) are taken from the unperturbed control run and are then perturbed with different seeds in each member (same as Météo France does)
- Prognostic fields (soil moisture, soil temperature, snow depth, sea surface fluxes) are taken from the surface analysis (CANARI) and are then perturbed with different seed in each member; this means that those parameters are cycled in each member

By doing so we can increase the spread of surface parameters like T2m or RH2m significantly compared to the Météo France method. Figure 3 shows the impact of the new surface perturbation scheme on the T2m temperature for a test phase in November 2021. Spread is significantly increased over the whole forecasting range, while RMSE is smaller for most lead times. This results in a much better spread/skill ratio.

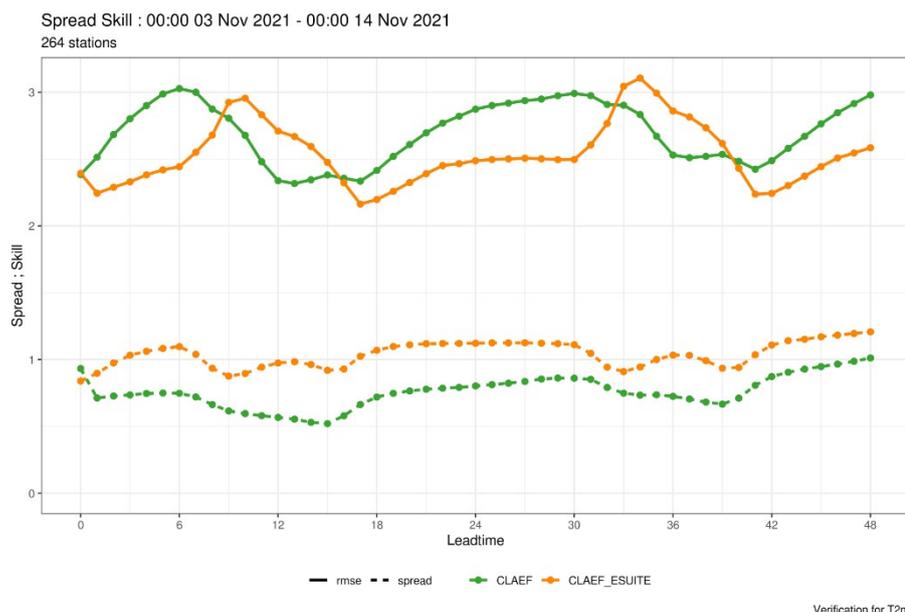


Figure 3: T2m verification for a test phase in November 2021. The green lines show ensemble spread (dashed) and RMSE (full) for the operational C-LAEF system, while the orange lines refer to a C-LAEF Esuite with additional surface perturbations.

The ECMWF-ENS coupling file production for C-LAEF and AROME-EPS has been reorganized and optimized in May 2021. Since this time the coupling files are produced under the Time-critical option 3 at ECMWF for a large common coupling domain (covering the C-LAEF and AROME-EPS domain) with the 903 configuration which is much more effective than the previous methodology with separate production (Table 1).

Table 1: Reorganization of ECMWF ENS coupling file production for RC-LACE.

LACE_EF	47R1	Next Config - 47R2	Comments
00/12	11 Members - 3-hourly to STEP=12 47 levels	17 Mem. - hourly to STEP=18 60 levels	New config ~6 times more expensive than existing one.
06/18	11 Members - 3-hourly to STEP=60 47 levels	17 Mem. - hourly to STEP=72 60 levels	But cost compensated by new job organisation.

AROME-EPS:

Lateral boundary conditions as well as atmospheric initial condition for the Hungarian AROME-EPS are provided by ECMWF ENS. The coupling frequency was 3 hours until May 2021. ECMWF upgraded the IFS model cycle from 47r1 to 47r2 in May which brought two changes in the Hungarian LAMEPS:

1. The number of the vertical levels in the raw boundary conditions increased from 91 to 137.
2. A higher, 1-hour coupling frequency was applied in AROME-EPS.

ECMWF provided test data before operational introduction of the new cycle and update of the LBC production. We conducted some case studies to check (1) the impact of more vertical levels and (2) the joint impact of more levels and higher coupling frequency. Three test days were selected to compare the forecasts with the operational LBCs and the higher resolution

test versions: 30 March 2021 was an uneventful, anticyclonal day, while on 13-14 April 2021 a complex front system crossed the territory of Hungary.

Significant difference was not noticed between the results of the two coupling settings, the 1-hour coupling shows a slight improvement (Figure 4). The positive impact of the higher vertical resolution of LBCs is visible only in the forecasts longer than 24 hours.

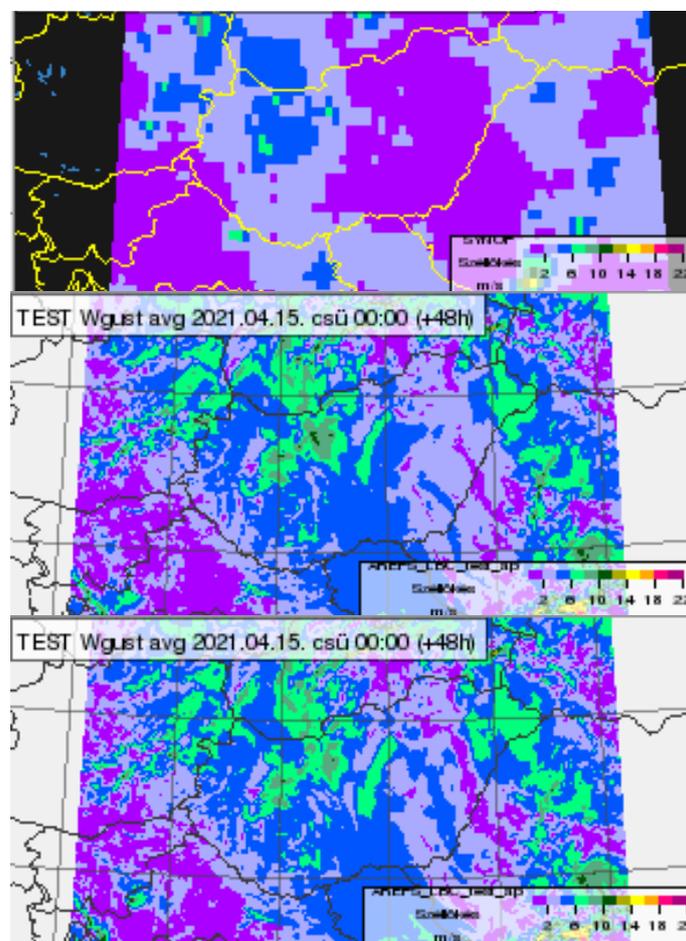


Figure 4: Wind gust at 0 UTC on 15 April 2021 based on SYNOP measurements (top), ensemble mean of AROME-EPS forecasts started at 0 UTC on 13 April 2021 with hourly (middle) and 3 hourly (bottom) coupled LBCs.

3 Research and Development

A-LAEF:

In 2021, the popularization of A-LAEF operational products became our main goal and motivation. Several hundreds of different EPS maps for various parameters and products are prepared from each ensemble run and made available for the forecasters and public (public maps, restricted for small Slovak domain and selected group of parameters, are available at: <https://www.shmu.sk/produkty/nwp/alaef/>). New products were introduced, like precipitation phase maps, and detailed EPSgrams. Next to the standard box-plot graphics, they contain also the daily MIN/MAX, SUM values, and quantiles for the precipitation. All these products are prepared on local HPC at SHMU, using custom parallelized Perl/R programs. The A-LAEF EPSgrams for selected towns are available on RC LACE portal too. The reliability and availability of A-LAEF products raised confidence in the system among the forecasters and

public. In day-to-day confrontation with the operational deterministic ALADIN forecast at SHMU (4.5 km), the benefit of A-LAEF ensemble became obvious. This was even more pronounced during the periods with strong winds or sudden temperature changes, and confirmed by the long-term objective verification scores using HARP (Figure 5). We made as well a detailed study of the convective precipitation case over Slovakia from July 2020 (Simon et al., 2021), where the different approaches to simulate this event were tested, including the A-LAEF system. The year 2021 was also quite rich in terms of extreme weather events and difficult to predict situations. Among all of them, we show here two most striking precipitation cases, where A-LAEF system clearly manifested its benefit.

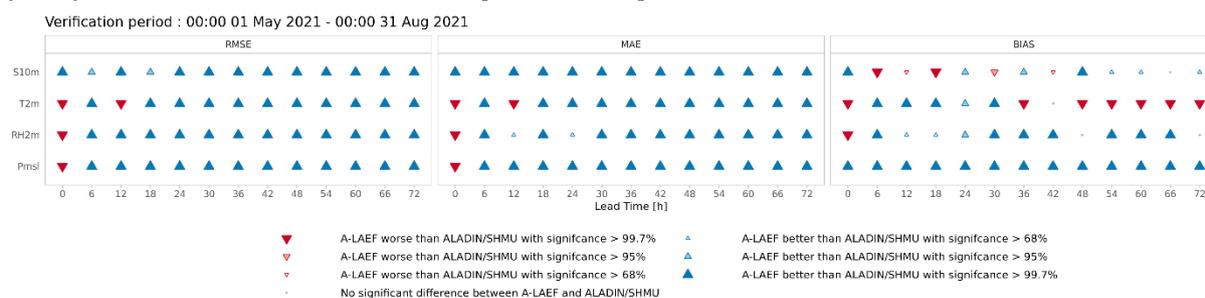


Figure 5: HARP verification scores for wind speed (S10m), temperature (T2m), relative humidity (RH2m) and mean sea level pressure (Pmsl), showing a relative comparison between A-LAEF EPS MEAN and ALADIN/SHMU deterministic forecast for convective season 2021. Blue triangles (A-LAEF is better), red triangles (ALADIN/SHMU is better), with the size representing the significance. (HARP verification tool was implemented at SHMU by Martin Petráš.)

Catastrophic floods in Germany (July 13-15, 2021)

After several episodes of heavy rain, the cyclonic weather system (Bernd) caused persistent or recurring heavy rainfall. The central parts of Germany were affected locally, but the west of Rhineland-Palatinate and the southern half of North Rhine-Westphalia were largely affected. As a result, small rivers and flash floods began to expand locally. Unfortunately, in addition to immense property damage, over 160 people lost their lives. The A-LAEF system successfully captured the precipitation event (Figures 6 and 7), with well-localized patterns (even with the high probabilities of extreme precipitation amounts, which is not usual in these situations and was not present in other ensembles).

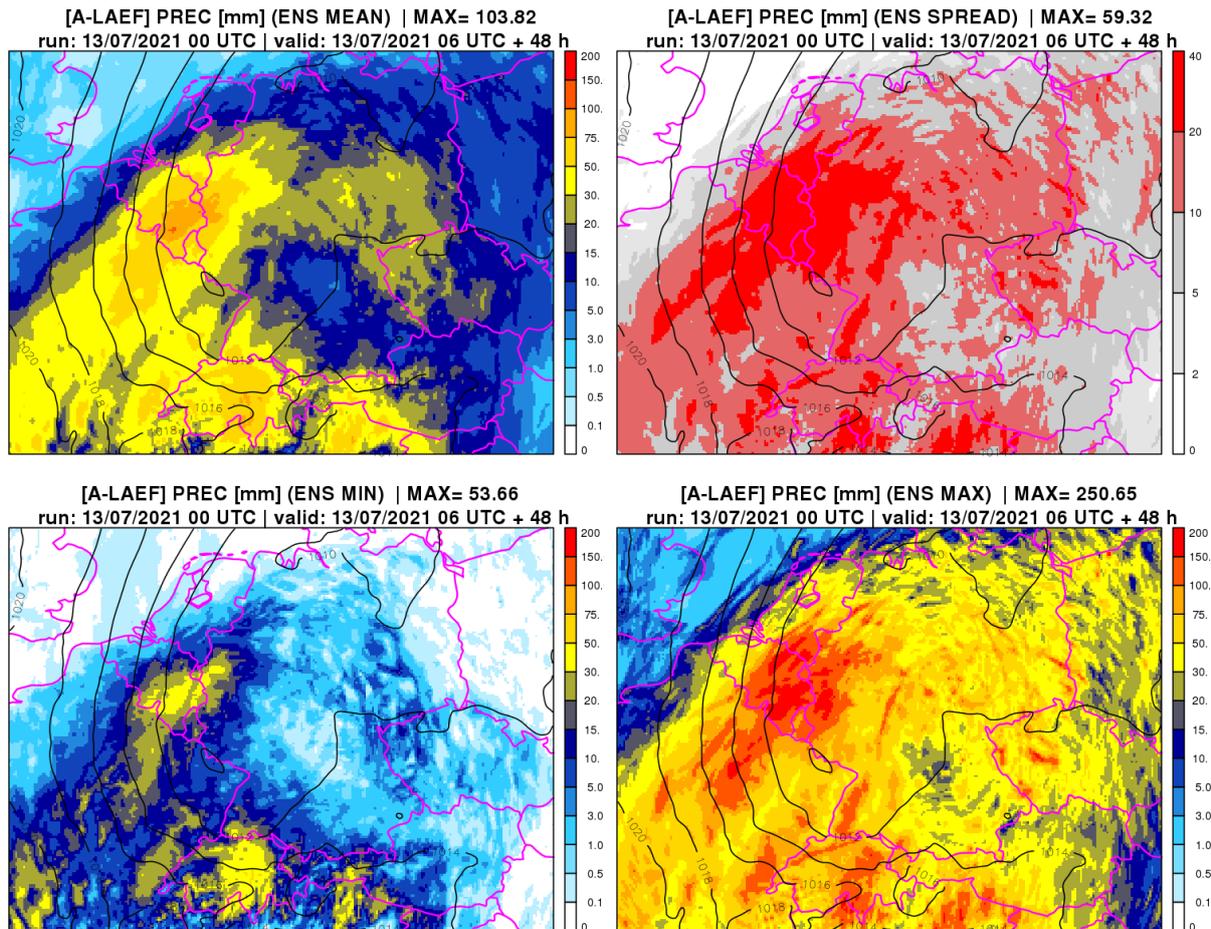


Figure 6: The A-LAEF 48-hourly precipitation forecast (06 to 06 UTC) from July 13, 2021 00 UTC over the affected area (EPS MEAN – top left, SPREAD – top right, EPS MIN – bottom left, EPS MAX bottom right).

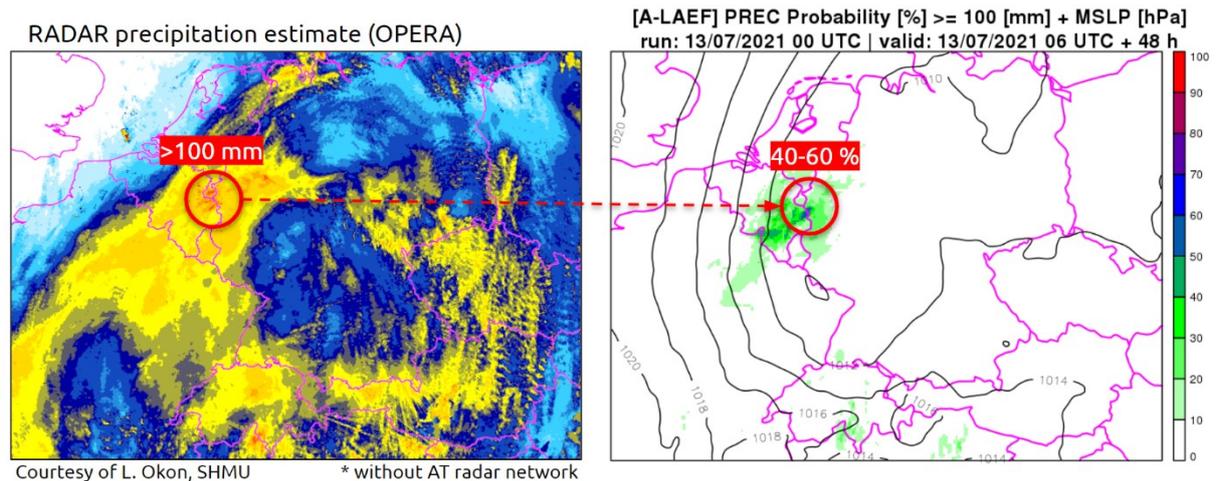


Figure 7: The 48-hourly precipitation amounts estimated by OPERA radar network (July 13, 06 UTC to July 15, 06 UTC – as the forecast in Figure 6) compared to the A-LAEF probabilities for ≥ 100 mm threshold.

Record precipitation totals in Italy (October 4, 2021)

A European record was broken in Northern Italy (Liguria region), where more than 740 mm of rain fell within the 12 hours period on October 4, 2021, causing floods, landslides, and several rivers broke their banks. There was also 178 mm of rainfall measured in just 1 hour

(Urbe Vara Superiore), and over 900 mm in 24 hours (Rossiglione). The localization of precipitation maxima with high probabilities for extreme thresholds was spot-on by the operational A-LAEF forecast 24, and even 48 hours ahead (Figures 8 and 9). The ensemble mean of 24-hourly precipitation accumulation was higher in comparison with the operational deterministic run at SHMU, and significantly better in predicting the exact location of the maxima (Figure 10). Because, the operational version of A-LAEF system was coupled to the cy47r2 IFS/ENS that time, the forecast was later recreated with the new cy47r3 couplings (available at ECMWF as a test suite) using A-LAEF special project SBUs. The results were even better, regarding the improved localization of the event by the unperturbed control run, while the ensemble statistics like MEAN and probabilities for given thresholds were almost identical (not shown).

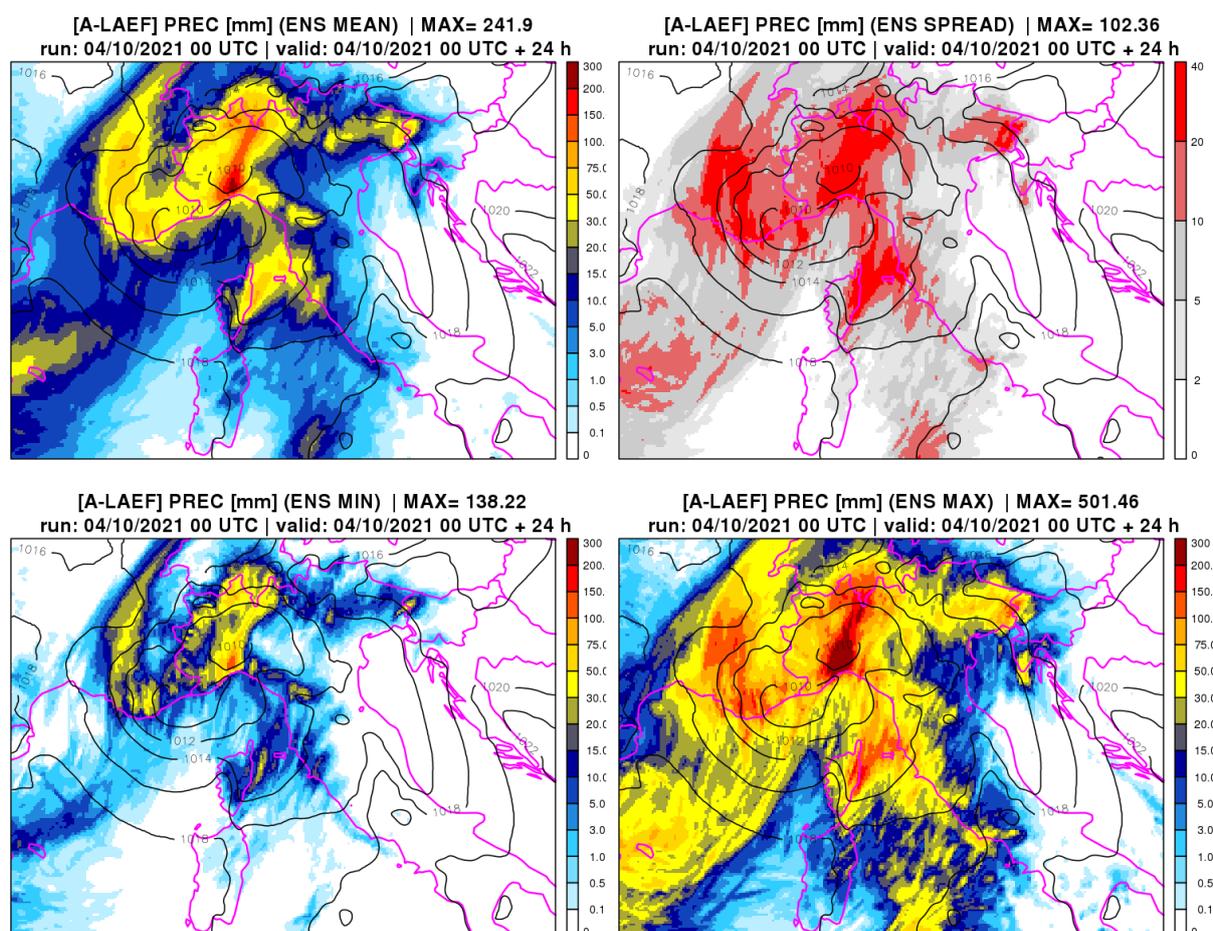


Figure 8: The A-LAEF 24-hourly precipitation forecast from October 4, 2021 00 UTC over the affected area (EPS MEAN – top left, SPREAD – top right, EPS MIN – bottom left, EPS MAX bottom right).

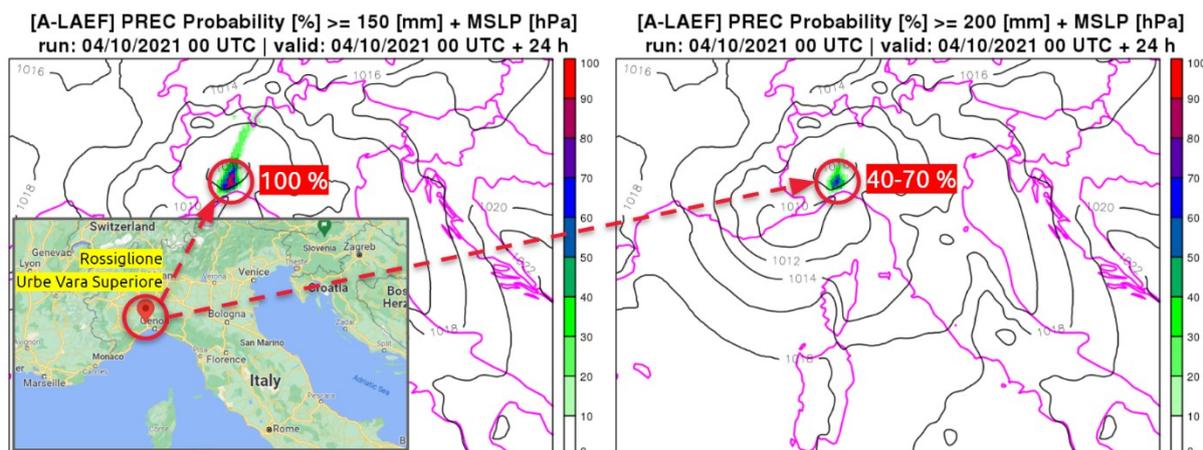


Figure 9: The exact location of record precipitation measurements in Italy compared to the A-LAEF probabilities for ≥ 150 mm/24 h threshold (left) and ≥ 200 mm/24 h threshold (right).

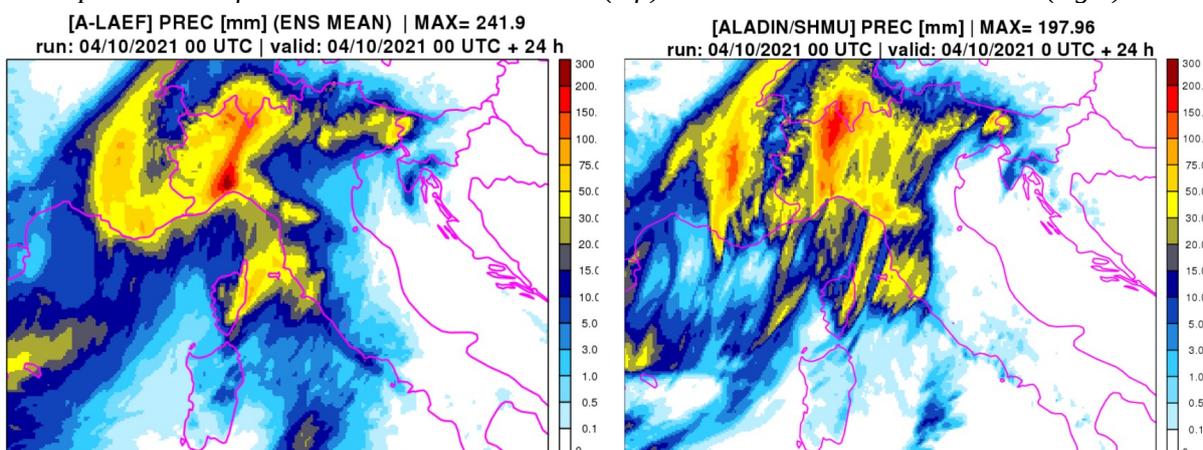
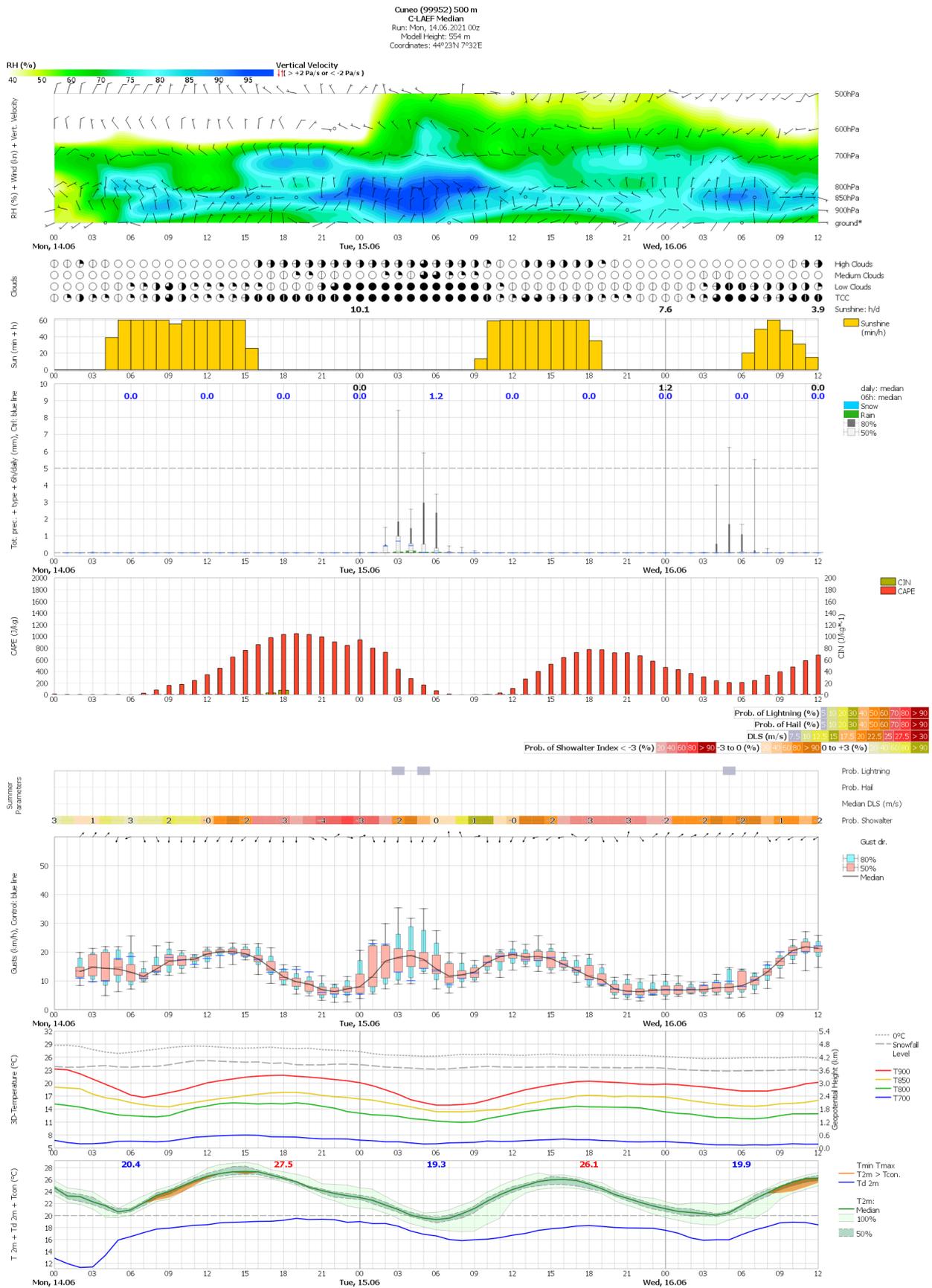


Figure 10: The A-LAEF EPS MEAN 24-hourly precipitation forecast from October 4, 2021 00 UTC over the affected area (left) compared to the same forecast of the operational ALADIN/SHMU deterministic model.

On the development side, the export version of A-LAEF scripting system for an easy installation under standard users at ECMWF HPCF was prepared and distributed. An online training for RC LACE partners was organized at the same time. This should improve the overall awareness of the system among the partners, and serve as a base tool for the A-LAEF Special Project experiments, as well as for our further research and development activities.

C-LAEF:

Strong focus on EPS related work at ZAMG in 2021 has been put on the development and production of EPS maps and EPSgrams for forecasters and customers. In a close co-operation between model developers and forecasters needs and ideas were evaluated and specific probabilistic products and maps were developed. These products comprise classical probability maps such as probability of exceeding certain thresholds, extreme members, median, etc. but also EPSgrams for more than 250 points in the Alps (Figure 11). All these maps and EPSgrams are produced with the software Visual Weather. The Visual Weather template for the C-LAEF EPSgrams has been shared within RC-LACE partners and such EPSgrams will be published for selected points on the RC-LACE webpage very soon. With the operational implementation of these EPS products, the acceptance and usage of C-LAEF within the routine work of the forecasters has significantly increased.



Some work in 2021 has also been dedicated to the model error representation in C-LAEF. Model error in the operational C-LAEF version is currently represented by a hybrid stochastic perturbation scheme, where perturbations of tendencies in shallow convection, radiation and microphysics are combined with parameter perturbations in the turbulence scheme. To increase the physical consistency in the perturbation scheme, the SPP (stochastically perturbed parametrizations; Ollinaho et al., 2017) approach has been applied to the radiation, microphysics and shallow convection scheme in C-LAEF as well. The new implementation with 11 perturbed parameters (2 in radiation, 2 in microphysics, 4 in turbulence, 1 in diffusion, 1 in surfex, 1 in shallow convection) is currently tested and verified in a C-LAEF Esuite. The perturbations are produced by the newly implemented SPG pattern generator (Tsyrlunikov and Gayfulin, 2017). However, a lot of tuning considering the perturbation scale and range has to be made before final operationalization. It is also planned to add some more parameters (especially in microphysics) to the perturbation list.

AROME-EPS:

The current operational AROME-EPS is downscaling of the first 11 members of ECMWF-ENS without data assimilation. Experiments are ongoing to introduce some local perturbations with ensemble of data assimilations (EDA). AROME-EDA is based on the 3-hour assimilation cycle of AROME/HU (Tóth et al., 2021). The experiment began with a cold start from the AROME/HU analysis and the perturbations were added to the SYNOP, AMDAR, TEMP and GNSS ZTD data. The integration domain, the spatial resolution, the applied model version were identical to AROME/HU configuration. AROME was driven by the control and the first 10 members of ECMWF ENS, from 4 runs a day during the spinup period. After 15-day spinup, the forecast experiments with 10+1 members started at 0 UTC with 24-hour lead time, because we concluded that EDA mainly affects the first 6-18 hours of the forecast. The results are shown for a summer period and a spring case study (Jávorné-Radnóczy et al., 2022).

July in 2019 was rainy with some heavy precipitation incidents, overestimation of temperature and wind in the operational AROME-EPS forecasts. Looking at the control members (dashed lines in Figure 12), data assimilation largely improves the initial condition for 2-meter temperature and mean sea level pressure, but degrades that of the 10-meter wind speed (the reason for the latter is not understood yet). The initial spread increases with the EDA perturbations and it still has impact after 24 hours. The forecast quality mainly enhances in the first 18 hours. Precipitation RMSE gets clearly better in the afternoon. However, too much rain is predicted in mornings by EDA in contrast to the downscaled run, which underestimated the precipitation amount. Brier scores indicate for all thresholds that EDA makes the forecast better (Figure 12 for 3-hourly precipitation exceeding 10 mm).

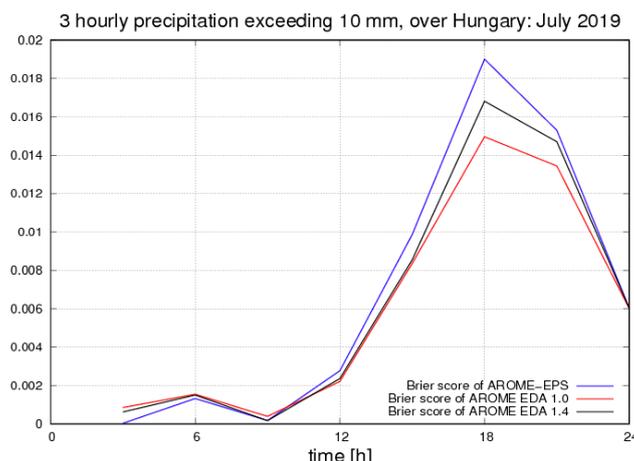


Figure 12: Brier score for 3-hourly precipitation sum exceeding 10 mm (right) in July 2019. Operational AROME-EPS (blue), AROME-EDA with perturbations multiplied by 1.0 (red) and 1.4 (black) are shown.

We verified the upper air forecasts against TEMP observations at levels of 925, 850 and 700 hPa. The comparison shows an improvement in the initial temperature, wind speed and humidity with applying EDA (Figure 13 for relative humidity at 850 hPa). The impact on the forecast quality is mostly neutral to positive. The spread increased with EDA but this impact basically diminishes after 12 hours (except humidity).

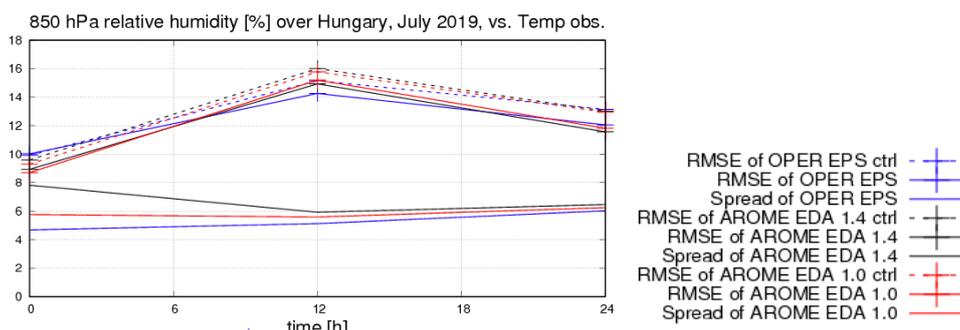


Figure 13: Ensemble spread (–) and RMSE (+) of ensemble mean for relative humidity at 850 hPa in July 2019. Operational AROME-EPS (blue), AROME-EDA with perturbations multiplied by 1.0 (red) and 1.4 (black) and RMSE of the corresponding control members (dashed) are shown.

On 25th April 2020, a cold front passed over the Northeast of Hungary accompanied by robust flow in the upper atmosphere, and weaker wind near the surface in the afternoon. Due to the downward mixing of strong wind, precipitation fell over this area. Both the operational AROME/HU and AROME-EPS members underestimated the maximum wind gust. The data assimilation brought the initial conditions closer to the measurements (see Figure 14 for Vásárosnamény which is located in the critical northeastern area). A growth of the initial spread caused by EDA perturbations can be also seen in the meteogram.

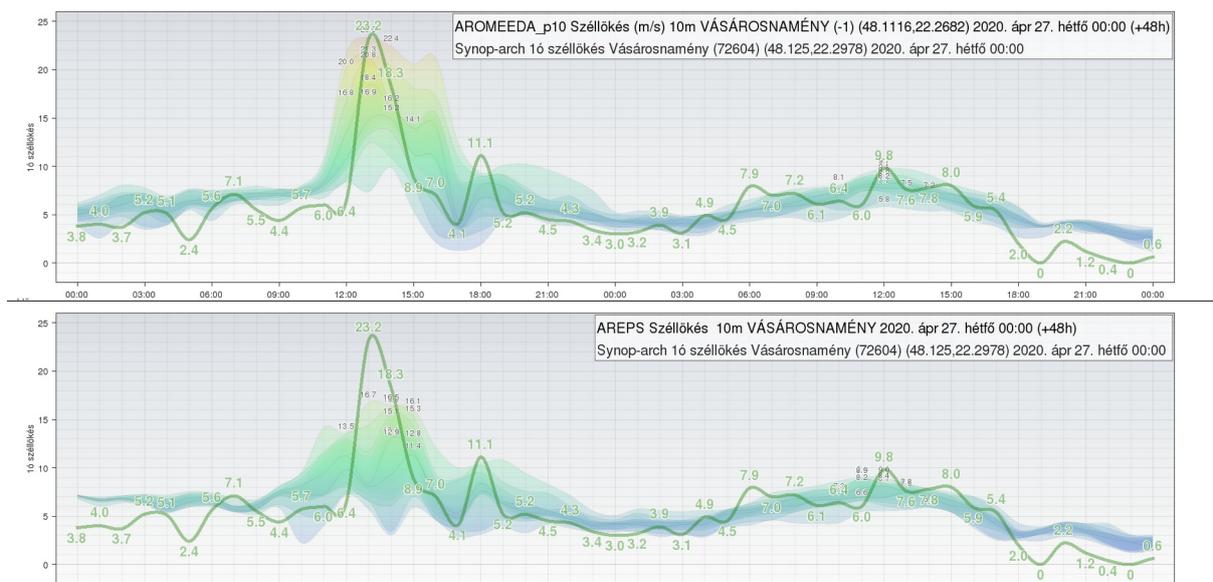


Figure 14: Hourly wind gust forecasts of operational AROME-EPS (bottom) and AROME-EDDA (top) run initialized at 0 UTC on 25 April 2020 and corresponding SYNOP observations (green line and values) for Vásárosnamény lying in the Northeast corner of Hungary. The range of the ensemble members is shaded.

4 References

- Belluš, M., Y. Wang, F. Meier, 2016: Perturbing surface initial conditions in a regional ensemble prediction system. *Mon. Wea. Rev.* 144, 3377-3390, <https://doi.org/10.1175/MWR-D-16-0038.1>.
- Bouttier, F., Raynaud, L., Nuisser, O. and Ménétrier, B. (2016) Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX. *Quart. J. Roy. Meteor. Soc.* 142, 390-403, <https://doi.org/10.1002/qj.2622>
- Derková, M., M. Belluš, M., 2007: Various applications of the blending by digital filter technique in the ALADIN numerical weather prediction system. *Meteorologický časopis* 10, 27–36.
- Jávorné-Radnóczi, K., Szépszó, G., Tóth, B., 2022: Short range probabilistic forecasts at Hungarian Meteorological Service: evaluation of operational AROME-EPS and test of Ensemble of Data Assimilations. *Quarterly Journal of Hungarian Meteorological Service (Időjárás)*, under revision.
- Ollinaho, P., Lock, S. J., Leutbecher, M., Bechtold, P., Beljaars, A., Bozza, A., Forbes, R. M., Haiden, T., Hogan, R. J. and Sandu, I. (2017). Towards process-level representation of model uncertainties: Stochastically perturbed parametrisations in the ECMWF ensemble. *Quart. J. Roy. Meteor. Soc.* 143, 408–422, <https://doi.org/10.1002/qj.2931>
- Simon, A., M. Belluš, K. Čatlošová, M. Derková, M. Dian, M. Imrišek, J. Kaňák, L. Méri, M. Neštiak and J. Vivoda, 2021: “Numerical simulations of 7 June 2020 convective precipitation over Slovakia using deterministic, probabilistic and convection-resolving approaches”, *IDŐJÁRÁS*, Vol. 125, No. 4, DOI:10.28974/idojaras.2021.4.3 (pp. 571–607).
- 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., 2018: The ALADIN System and its

Canonical Model Configurations AROME CY41T1 and ALARO CY40T1. *Geosci. Model Dev.*, Vol 11, 257-281, <https://doi.org/10.5194/gmd-11-257-2018>.

Tóth, H., Homonnai, V., Mile, M., Várkonyi, A., Kocsis, Zs., Szanyi, K., Tóth, G., Szintai, B., Szépszó, G., 2021: Recent developments in the data assimilation of AROME/HU numerical weather prediction model. *Quarterly Journal of Hungarian Meteorological Service (Időjárás)* 125, 4, 521–553, DOI: 10.28974/idojaras.2021.4.1.

Tsyrlunikov, 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.

Wang, Y., M. Belluš, F. Weidle, et al., 2019: Impact of land surface stochastic physics in ALADIN-LAEF. *Quarterly Journal of the Royal Meteorological Society*, 1–19, <https://doi.org/10.1002/qj.3623>.

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>

Evaluation of HARMONIE-AROME cycle 43h2.1 at AEMET

Javier Calvo, Joan Campins, María Díez, Pau Escribà, Daniel Martín, Gema Morales, Jana Sánchez-Arriola y Samuel Viana

1 Introduction

After a thorough evaluation, AEMET's HARMONIE-AROME deterministic suite has been upgraded to Cycle 43h2.1. A description of the main characteristics of the version and of the local changes compared to the HARMONIE-AROME reference version is done. Finally, an objective evaluation is performed compared to cycle 40h1.1 and the main meteorological implications of the new version are outlined.

2 Operational set up

The deterministic operational suite is based on HARMONIE-AROME which is run at 2.5 km 4 times per day with a forecast length of 72 hours over 2 geographical domains (Fig. 1). 3DVar analysis with 3hr cycle includes radar reflectivities, ATOVS, IASI, GNSS ZTD, ASCAT wind and AMDAR humidity observations. IFS humidity enters in the blending process (LSMIX) with the ECMWF forecasts. Upper air analysis includes 2 m temperature and relative humidity.

SAPP preprocessing is used for conventional observations. Radar data comes from OPERA using BALRAD preprocessing and including Spanish, Portuguese and French radars. The control of the HARMONIE-AROME operational suite is based on ecfLOW.

The operational system has been migrated to a new ATOS computer system composed of two clusters each with 140 computed nodes mounted on Bull Sequana X440 A5 chassis. Each node with 2 AMD EPYC™7742 processors (64 cores). The peak performance of the system is 1350 TFlops.

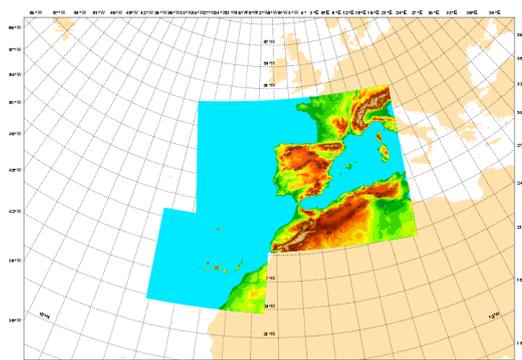


Figure 1: AEMET's operational domains

3 Major changes in cycle 43h2.1

The major changes of cycle 43h2.1 compared to cycle 40h1.1 are

- New physiographic data base: ECOCLIMAP Second Generation
- New clay and sand database (SOILGRIDS)
- SURFEX 8.1 version (Le Moigne *et al.*, 2018)
- 2 patches for Nature tile and disabling Surface Boundary Layer Scheme
- Modified values of minimum stomatal resistance R_{sm}
- Increase maximum Richardsson number from 0.0 to 0.2
- HARATU turbulence scheme updates mixing at the top of PBL
- EDMF and microphysics optimizations
- Decrease soil moisture increments in the OI assimilation scheme
- New blacklisting procedure based on ECMWF's and redesign of the assimilation scripting system.

The major local changes compared with the HARMONIE-AROME reference system are

- Increase roughness increasing heterogeneity of open land patch (FAKETREES)
- Orographic roughness parametrization OROTUR enabled (Rontu, 2006)

4 Meteorological impact

Parallel experiments comparing cycle 43h1.2 with cycle 40h1.1 have been conducted for the periods: 1aug/17sep2020(AS20), 1oct/30nov2020(ON20), 1dec2020/28feb2021(DEF21) and 1mar/21jun2021(MAM21). The forecast length for these tests is 24 hr. The overall impact of the model upgrade can be seen in Table 1. There is a general improvement with some concern for the dew point temperature. The impact on the upper level variables is small.

Table 1: Summary of the verification results comparing cycle 43h1.2 with cycle 40h1.1. Filled triangles mean 90% confidence. For the categorical scores (FF10> and Ppt12>) there is no significance test but a subjective evaluation of the objective scores

	Area: Spain-Portugal			
	AS20	ON20	DEF21	MAM21
MSLP	▲	△	△	△
T 2m	▲	■	△	▲
10m wind	▲	▲	▲	▲
10m gust	▲	▲	▲	▲
FF10> 10m/s	▲	▲	▲	▲
Td 2m	▼	▼	■	■
CC	■	■	△	△
Ppt12>3 mm	△	■	△	△
Ppt12>10 mm	△	▽	△	■

2m dew point temperature

Concerning the degradation seen in the dew point temperature, it is due to the first forecast lengths as can be seen in Fig. 1. This degradation is not seen in relative humidity.

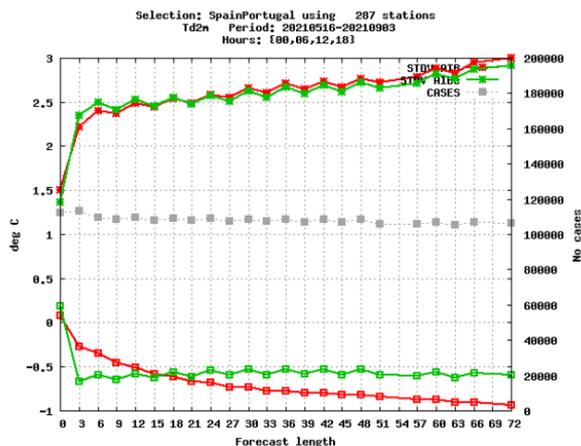


Figure 2: STDV and Bias of 2 m dew point temperature function of the forecast length for the period 16 May–3 Sep 2022. Cy43h2.1 in green and cy40h1.1 in red.

10m wind

The upgrade to ECOCLIMAP Second Generation and the discontinuation of the canopy scheme produces and increase of the wind bias. The overestimation of 10 m wind was already significant in cycle 40h1.1. This bias is not as clear for other HARMONIE-AROME domains. The wind is improved in the new version mainly due to the increase of roughness on the open land patch (FAKETREES) and activation of an orographic roughness parametrization (OROTUR). The impact on 10 m wind can be seen in Figures 3 and 4.

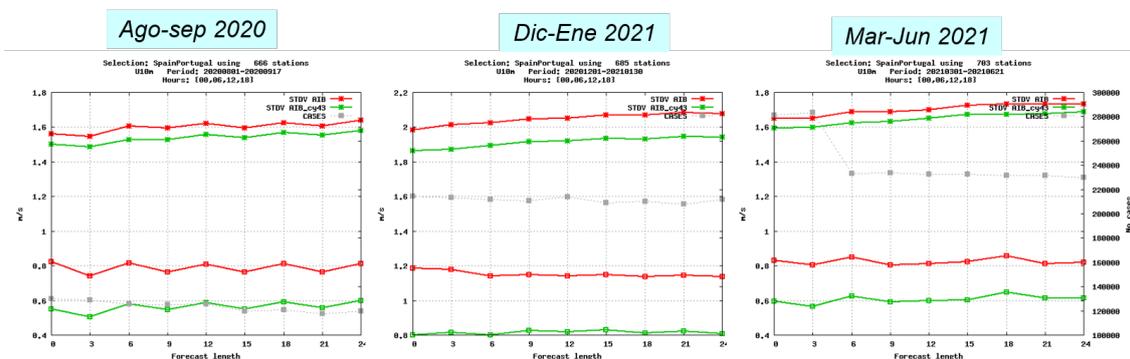


Figure 3: STDV and Bias of 10 m wind function of the forecast length for the periods described in the text. Cy43h2.1 in green and cy40h1.1 in red.

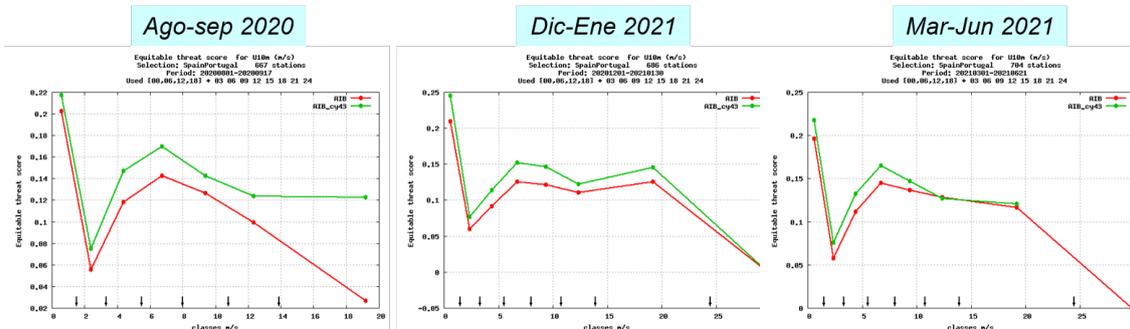


Figure 4: Categorical verification of 10 m wind: Equitable Threat Score function of the observed velocity. Cy43h2.1 in green and cy40h1.1 in red.

It is interesting to note that the impact of the activation of the orographic parameterization is seen not only in mountain regions (Fig. 5). The improvement on the Canary Islands domain is not so clear with

a reduction of the wind bias but with a deterioration of the winds above 10 m/s (not shown). Probably, a revision of the use of OROTUR for this region is need.

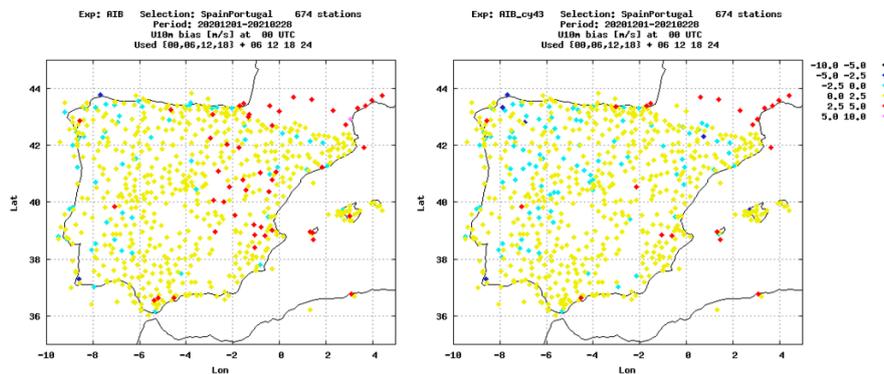


Figure 5: Spatial distribution of the 10 m wind bias for the period 1dec2020-28feb2021. Blue colors indicate underestimation and yellow and red colors overestimation. Cy43h2.1 (right) and cy40h1.1 (left).

10m wind gusts

There is an improvement of the 10 m wind gusts with a bias close to zero (Fig. 6). Note that we use a rafagosity factor of 2.5 instead of the reference 3.5 in both Cy43h2.1 and cy40h1.1. This reduction in the rafagosity factor produces an underestimation of the gusts for the strongest large scale wind events as can be seen in the categorical verification of the winter period (Fig. 7).

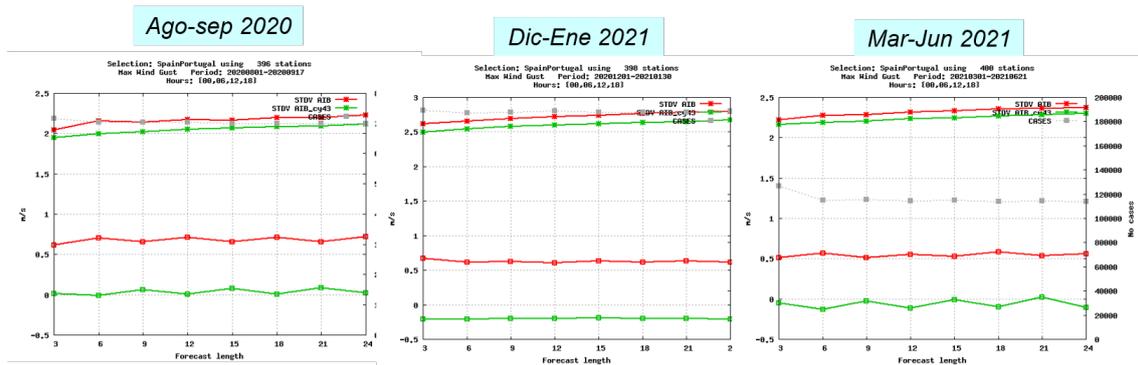


Figure 6: STDV and Bias of 10 m wind gusts function of the forecast length for the periods described in the text. Cy43h2.1 in green and cy40h1.1 in red.

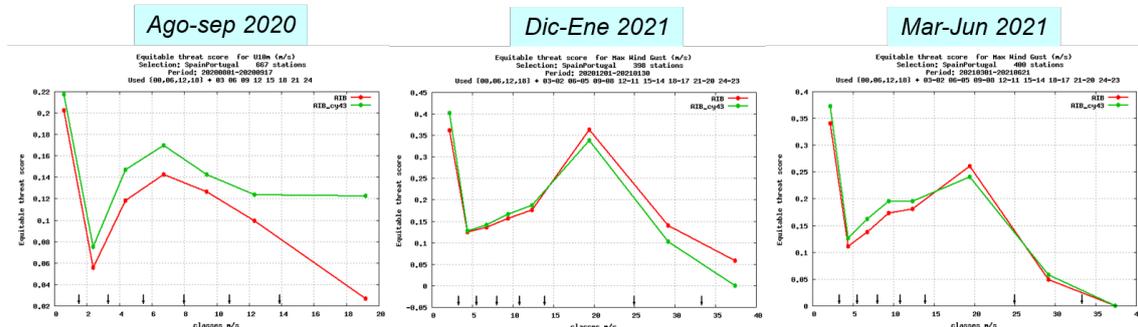


Figure 7: Categorical verification of 10 m wind gusts: Equitable Threat Score function of the observed velocity Cy43h2.1 in green and cy40h1.1 in red.

The problem of overestimation of wind gusts in convection events is still present in the new version as can be seen in the scatterplots of the summer period (Fig. 8)

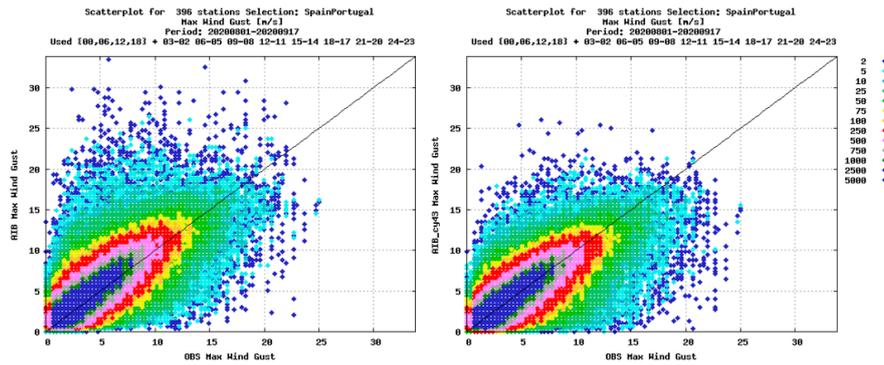


Figure 8: Forecast-observation plot of 10 m wind gusts for the period 1ago-17sep2020. Cy43h2.1 (right) and cy40h1.1 (left).

Low clouds/Fog over sea

The new version increases low clouds and fog over sea (Fig. 9). This is a serious problem for other HARMONIE-AROME domains but not so much over Spain because we also have cases with clear underestimation of low clouds and fog especially over the Mediterranean sea.

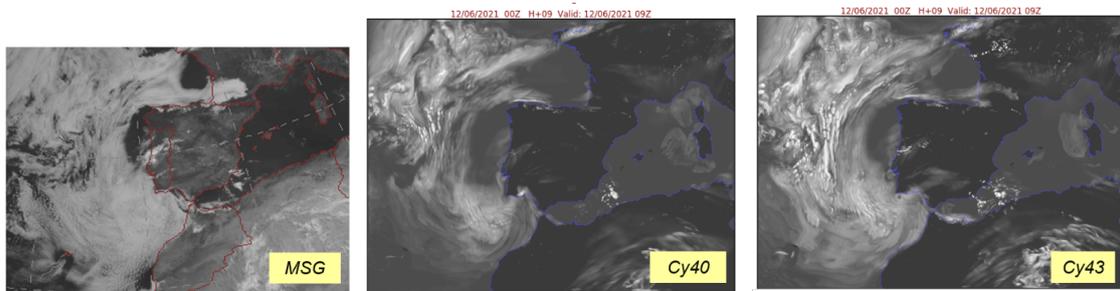


Figure 9: Simulated satellite images on 12/06/2021 at 09 utc compared with Meteosat Visible Image.

Fog over land

The new version decreases the frequency of fog as can be seen from the categorical verification of visibility on Fig. 10 for the winter period, the season when radiation fog usually occurs over the Iberian Peninsula. Although from the subjective evaluation of the forecasts it seems that the model gives a good indication of the spatial distribution of the fogs, from the point verification we can see that the errors are still big and only for the North Meseta reasonable ETS values are obtained (Fig. 11). Overall conclusion is that the cycle change implies some degradation on the fog forecasts. This seems mainly due to the change of the maximum Richardson number because the use of the RISHIFT parameterization (Homleid, 2022) produces results close to the cycle 40h1.1 ones.

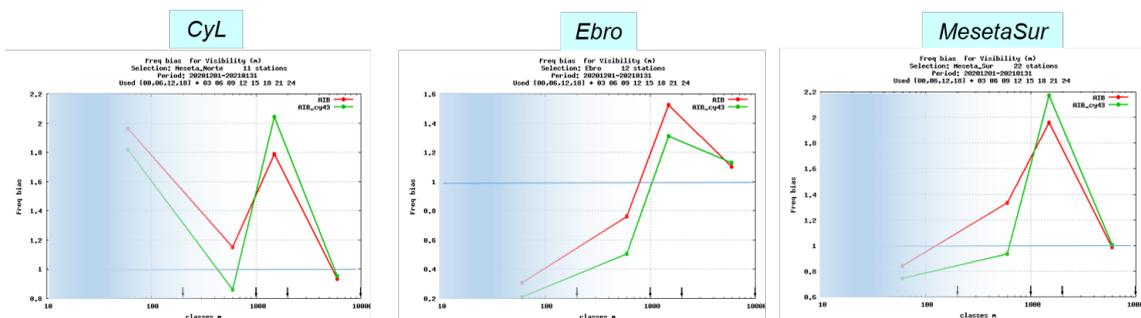


Figure 10: Categorical verification of visibility for the winter period for three climatological regions: North Meseta (CyL), South Meseta and Ebro valley. Frequency bias function of observed visibility. Cy43h2.1 in green and cy40h1.1 in red.

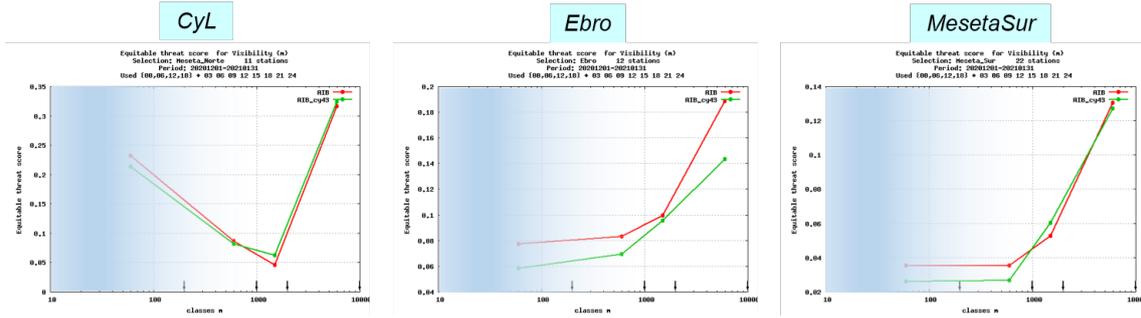


Figure 11: As Fig 10, Equitable Threat Score function of observed visibility

Precipitation

The model upgrade implies an improvement of the precipitation forecasts as can be seen in Fig.12 but without big changes in model climatology (Figs. 13 and 14).

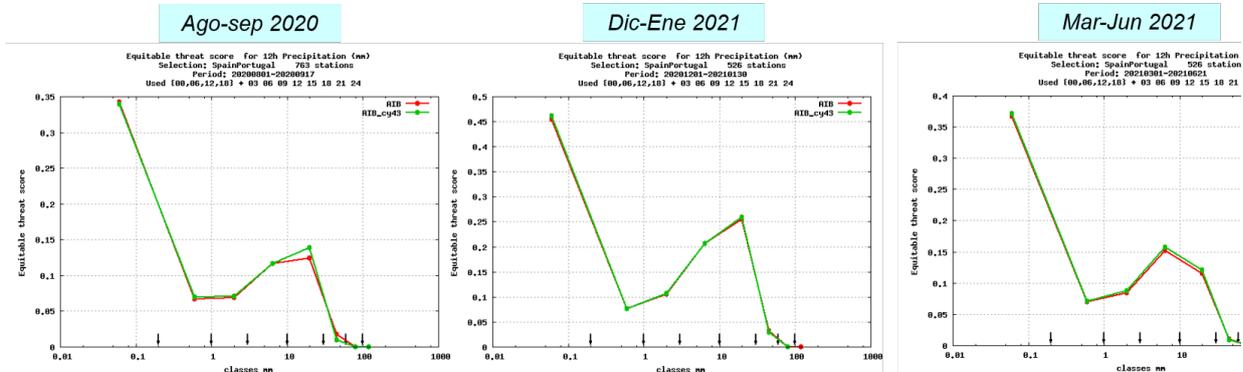


Figure 12: Categorical verification of precipitation accumulated in 12 hr. Equitable Threat Score function of the observed rain gauges precipitation. Cy43h2.1 in green and cy40h1.1 in red.

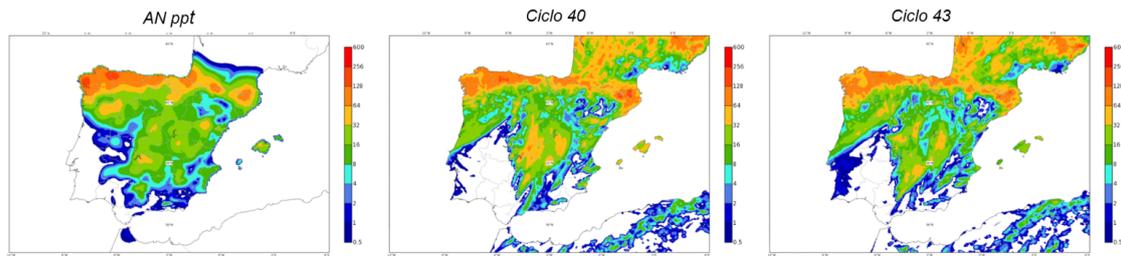


Figure 13: Precipitation accumulated in one month (august 2020). Analysis of precipitation (left), cy40h1.1 (center) and Cy43h2.1 (right). Note that the analysis is based only in rain gauges with much lower resolution, very few data over Portugal and no data over France.

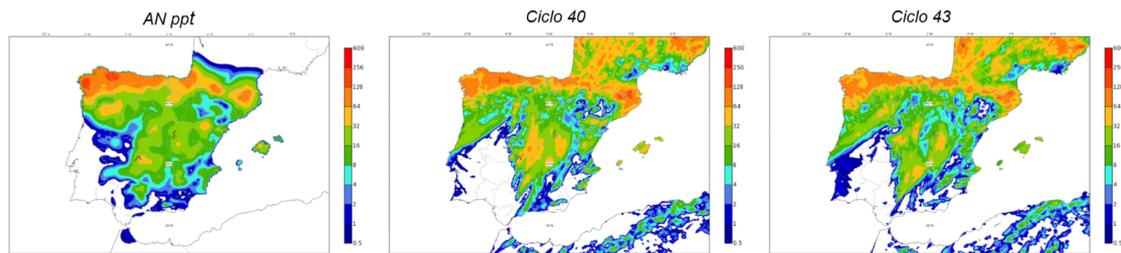


Figure 14: As Fig. 13 for December 2020

Lightning intensity

The density of electrical discharges is estimated from the vertically integrated graupel. The HARMONIE-AROME reference computation is not used but a particular tuning to try to simulate the cloud-earth discharges observed from the AEMET observation network. Due to the changes in the microphysics that imply a reduction of the proportion of graupel, a reduction of the discharge density is seen in the new version (Fig. 15). We think that the modifications on the hydrometeors partition come from updates on cycle 40h1.2 (Ivarsson, 2017). Besides, there has been an update of the AEMET observational network, so a new tuning in the algorithm for the discharges estimation seems to be needed.

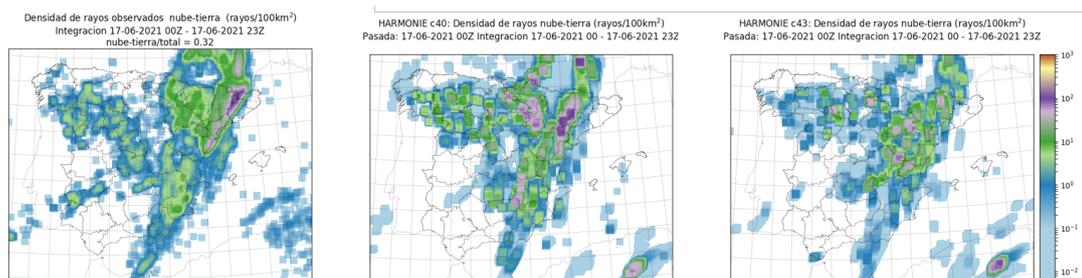


Figure 15: Comparison of discharge density in 24 hr on 17/06/2021. Observation (left), cy40h1.1 (center) and Cy43h2.1 (right). Thanks to José A. Sosa

5 Conclusions

The deterministic operational suite was upgraded to cycle 43h2.1 on 7th September 2021. The operations have been moved to a new computer system that increases AEMET’s computing power at least 6 times.

The upgrade from cycle 40h1.1 to cycle 43h2.1 implies a general improvement in most forecast variables. Bigger improvement is found in 10 m wind that comes mainly from the activation of the orographic roughness parameterization OROTUR and enhancement of roughness on open land by FAKETREES. Wind gusts also improved but there is some concern about the underestimation of the gusts for the more intense events associated with large scale lows. There is some increase in the cloud cover amount, especially for low clouds over sea. Nevertheless, there is some decrease on the frequency of radiative fogs over land which implies a slight degradation of its prediction. Finally, there is an improvement on precipitation forecasts and there is a decrease on the maxima of the density of electrical discharges.

6 References

- Homleid, M. (2022). Improving model performance in stable situations by using a pragmatic shift in the drag calculations - XRISHIFT . *ACCORD Newsletter*, 2.
- Ivarsson, K.-I. (2017), Recent development of cloud microphysics (ICE3) within MetCoOp. *ALADIN-HIRLAM Newsletter*, 9.
- Le Moigne and coauthors (2018): *SURFEX V8.1 main scientific documentation*. Available at https://www.umr-cnrm.fr/surfex/IMG/pdf/surfex_scidoc_v8.1.pdf
- Rontu, L. (2006) A study on parametrization of orography-related momentum fluxes in a synoptic-scale NWP model. *Tellus A: Dynamic Meteorology and Oceanography*, 58:1, 69-81, doi: 10.1111/j.1600-0870.2006.00162.x
- Rooy, Wim de, et al. (2020). Improved parametrization of the boundary layer in Harmonie-Arome (focusing on low clouds). *ALADIN-HIRLAM Newsletter*, 15.