Execution of Rolling ACCORD A Consortium for COnvection-scale modelling Research and Development Work Plan 2023

ACCORD Management Group and Support Team

1. Executive summary

As in the previous year, this scientific reporting has been prepared by the ACCORD Management Group, as a companion document to the Rolling Work Plan for 2023 available here:

http://www.umr-cnrm.fr/accord/IMG/pdf/rwp2023_adopted.pdf

2. Summary of ACCORD activities in 2023 on Management

2.1. Scientific Management and organization of the MG

An important achievement in 2023 was the nomination of our two new Area Leaders, Metodija Shapkaljevski (SMHI) for Physics and Benedikt Strajnar (ARSO) for Data Assimilation. Our two new MG colleagues started to work in April for ACCORD with, as a first major task, to fulfil their role in the preparation of next year's RWP (RWP2024). They also very swiftly joined the definition of some of the ACCORD funded actions for 2023 (DAP2023).

The position of the Consortium Scientific Secretary (CSS) also had to be restaffed as Patricia Pottier was leaving on 1 August for retirement. Météo-France offered to reopen the position within its own recruitment procedures, and a mixed MF-ACCORD selection panel eventually was formed to interview Mrs Anne-Lise Dhomps. On the warm recommendations by the panel, both MF (as the employer) and the ACCORD Assembly (via the Bureau) confirmed the recruitment and the nomination. It might be stressed in this reporting that Anne-Lise and Patricia organized a period of overlapping in May-June, both showing a high level of flexibility with respect to their other, professional and personal duties.

Update regarding the R2O white paper.

At the level of scientific management tasks, one important aspect was the continued efforts by the MG to trigger and develop the proposals based on the "R2O white paper". The white paper identified three pillars where improvement across ACCORD was highly desirable: common work environment and common testing, to organise user feedback and documentation. While a first achievement regarding the common work environment is on very good track (the ACCORD source code forge, as well as the technical unit testing tool "***davai***"), more elaboration on how ACCORD can organise common testing of integrated configurations (including some flavour of MQA) will be necessary in 2023-2024. The

organisation of user feedback will start in the autumn 2023, with the start of the work between MG and the User Representatives ("UR": use the questionnaire, organise the first steps with the URs, agree on a first time table for user feedback etc.). The efforts on documentation will very largely depend on the successful application of a potential candidate for the position of the Documentation Officer ("DO").

Update regarding AI and Machine Learning.

The WG-ML has finalised its portfolio with proposals of ML-topics in the beginning of 2023. The portfolio was presented to STAC (STAC-5 on 25 May). Since then, discussions and planning on ML topics in relation to NWP have taken fairly diverse directions,

- at the level of the teams and with the ACCORD/MG: a fair large number of tasks, within several WPs of the RWP2024, include AI/ML algorithms (about 18 R&D tasks in DA, in Surface, in Physics, in EPS).
- at the level of MG alone: the ACCORD/MG has taken up the question of how ACCORD as a consortium could engage into AI/ML efforts, for the benefits of all its members. While some main tracks of R&D seem to emerge, in line with previous recommendations by STAC-5, several important questions remain fairly open (like how to organise data for training, or how to organise infrastructure - some of these questions being also dependent on how ACCORD could interface with other initiatives as explained in the next bullet).
- in connection with European-level initiatives involving ECMWF and Member States: ECMWF will start a ML Pilot Project involving its Member States and funded by ECMWF budget. European NHMSs, beyond the geographical scope of ACCORD, are preparing a European-wide framework program to be hosted under the EUMETNET organisational facilities. The question of how ACCORD can interface with these initiatives, for a mutual benefit, progressively raises as these programs take shape (*on the whole, the existence of that many initiatives illustrates the highly entropic nature of this thematic across meteorological institutes, in the wake of results published by research labs involving big-tech companies since the autumn 2022*).

AI-based tools and ML for NWP also are important aspects to discuss in the preparation of the 2026-2030 strategy of ACCORD. Assuming the main goal of ACCORD remains to elaborate and make available to all consortium members state-of-the-art numerical trolls for high-resolution, (high) impact based and severe weather forecasting, how should ACCORD be engaged in R&D and maintenance of AI/ML-based tools besides the core efforts the consortium organizes around the physical-based "traditional" NWP codes ?
This question, along with its several fine-grain declinations, will be addressed in the continued preparation of the next phase strategy (to start with, by STAC with MG).

Participation of MG to the preparation of the next phase ACCORD strategy.

As approved by the Assembly on 26 June, the MG has started to elaborate on the high-level questions for the next phase strategy, and will work together with STAC on proposing guidelines to hand over to the future talk teams. Joint meetings between STAC and MG are planned for 3 (online) and 25-26 (Brussels, RMI) October 2023.

2.2. Information and communication

Newsletter 4 was published in June 2023, the intention is to publish Newsletter 5 in the beginning of 2024, with contributions expected by mid-December. The All Staff Workshop was held in March in Tallinn (Estonia). A very few cold facts about the ASW attendance:

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

The ASW program and material remains available at the ACCORD website ACCORD (follow "Events" in the menu). ACCORD was represented at this year's EWGLAM conference (Reykjavik, September) by about 25 presentations in both plenary and side sessions.

Visits to member institutes. Tunisia.

The MG visited the Tunisian NHMS (INM) in May. The visit was an unprecedented opportunity for MG members to meet in-person our Tunisian NWP colleagues as a whole team, discuss with them in an informal context and as well as listen to a number of presentations by INM staff. Our INM colleagues presented their fields of interest and their local work in NWP and in link with the ACCORD RWP. They introduced their main national and international projects and they presented use cases of critical weather forecasts (cases of severe Mediterranean storms and heavy precipitation for instance). The MG presented some of their current initiatives such as to organise user feedback. The MG was very impressed by the high quality of the presentations by INM staff, their enthusiastic participation in the discussions and, last but not least, their very warm welcoming to their institute and to their country.

Group photo of INM representatives and staff with the ACCORD/MG during the visit in Tunis, in the premises of INM.

Iceland.

On 28 September, the ACCORD/MG met the NWP team of IMO (Icelandic Met Office) besides the EWGLAM/C-SRNWP conference, in order to learn more about the staff composing the team, their fields of interest and their organisation.

The IMO team explained their most common meteorological hazards (on the top ranking, very strong low level winds) and their NWP configurations, either shared (Hirlam, UWC-West) or national (especially a 750m grid spacing instance of the ACCORD models). The MG was impressed by the very close connection of the NWP team with the forecasters, as well as by the dissemination of meteorological fields as weather charts for the general public and in relation to alert indexes. These charts and alert stages are widely communicated to the public who understands them. IMO has regular meetings with the Icelandic public security authorities (roads, pollutants etc.) whenever needed. It was agreed that the IMO team, via its LTM, would contact the relevant MG member in order to further discuss their implication in the ACCORD RWP (perhaps by envisaging scientific visits). Another striking feature was the importance for the team in Reykjavik, of a thorough scripting system shared with other partners.

Group photo of IMO representatives and staff with the ACCORD/MG during the visit in Reykjavik, in the premises of IMO. Remote participants were Anne-Lise and Alexandre.

2.3. On the edges of the consortium, link with other organizations at a scientific and technical level

ECMWF.

The ACCORD Integration Leader, the PM, the Area Leaders for SPTR and System, all participate in the IFS/Arpege coordination meetings, held about three times per year (twice online, once in-person). The coordination meeting is the place where scientific and technical choices for the IFS/Arpege codes are being discussed and agreed upon, in connection with other projects like hybrid24 (ECMWF) or code adaptation to the EuroHPC machines (Destination Earth). These meetings also are the main place where the content and timing of

cycles are discussed. A specific item on the agenda regularly is preserved for the LAM (ACCORD) partners to present their suggested items.

On specific aspects of R&D, one might notice the rather good relationships between dynamics experts across ECMWF/MF/ACCORD, certainly triggered by some former ACCORD team staff having joined ECMWF and others having visited the Center on scientific visits.

Destination Earth.

Several ACCORD/MG members also are directly involved in DEODE work package management (**SPTR**, **DYN**), others are involved part time on specific tasks and deliverables, and some STAC members participate in WP management on DEODE side. This cross-involvement currently ensures a fairly flexible, if not agile, coordination and exchange of information between the two projects. The DEODE core management team meets weekly, sometimes joined by the ACCORD/PM (who also acts as Legal Officer for DEODE, to handle IPR issues).

In the autumn 2023, the way how to handle subsisting IPRs in ACCORD's background IPRs will be discussed with ECMWF, as well as the possibility to use specific DEODE Deliverables for another purpose than the provision of Services in Destination Earth (can ACCORD use and co-develop scripts with DEODE teams, under what conditions ?). The outcome of these discussions, led by the ACCORD/PM (for ACCORD), will be presented to the ACCORD Bureau and Assembly in due time. It has potentially consequences on choices in our RWP (System) and on how to handle DEODE manpower at the intersection of the ACCORD/RWP (as discussed by the ACCORD Assembly at the end of 2022).

3. Activity report per Area

3.1. Strategic program: Transversal software developments (SPTR)

- ***overview of the activity in the area***

The work in this area has significantly speeded up since the start of the DEODE project. As in 2022, the focus of the work lies on enabling the ACCORD forecast models to utilize the compute power provided by GPUs. In the RWP2023, this is split up in two parts: (i) adapting the code layout (of the forecast model) to allow for different pieces of the model to run on different architectures, and (ii) the actual porting of individual pieces of the model to GPU.

Regarding the first part, significant progress was made for the ALARO CSC, with a thorough refactoring of the physics parameterizations control routines. The refactored code has entered the common cycle 49t1. As expected, the refactoring of the two other CSC's (AROME-Meteo-France and HARMONIE-AROME) has started, but is progressing at a slower rate. Nevertheless, an important step was taken in 2023 by moving the physics parameterizations of AROME-Meteo-France and HARMONIE-AROME to an external repository PHYEX, which will make the porting of these parameterizations to GPUs easier.

Besides the refactoring of the ACCORD physics parameterizations, a substantial effort was made to improve the flexibility of the code with respect to the granularity of the parallelism. Where CPUs benefit from a coarse granularity, with different cores treating relatively large gridpoint blocks, GPUs benefit from a finer granularity, with as much parallelism exposed to the GPU as possible. A very flexible mechanism was devised and implemented for the

ARPEGE global model, and this mechanism has also been adopted and tested for the ALARO CSC. This mechanism relies on scripted source-to-source generation of GPU-enabling code and on smart (GPU-aware) data structures.

Regarding the porting of individual pieces of the model to GPU, significant progress has been made in most sub-topics. The LAM spectral transforms have been integrated with the global spectral transforms (the open-source ***ectrans*** package), and a version was built that runs on AMD GPUs. Both single- and double-precision are supported and a benchmark has been run using up to 32 GPUs. While the initial GPU performance is not very impressive (Figure 3.1.1), possible optimizations have been identified and will be tested in an upcoming Cray/AMD Hackathon.

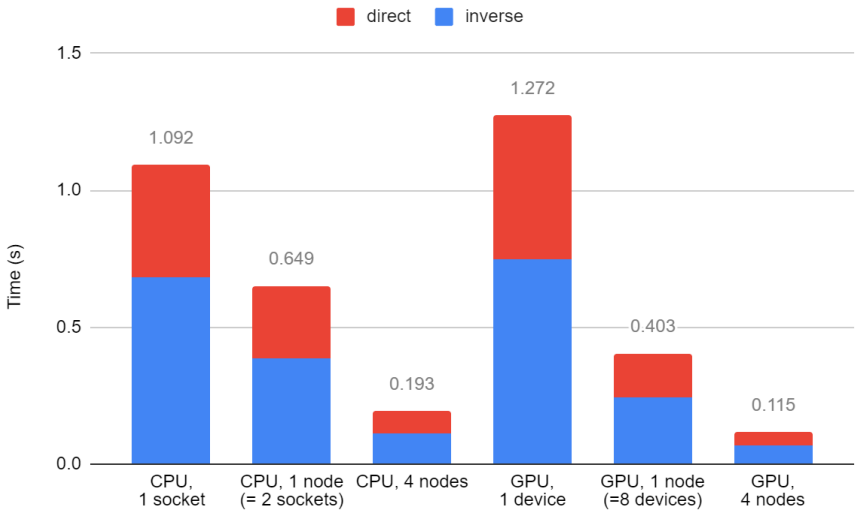

Figure 3.1.1: Timings of LAM spectral transforms on AMD CPUs and AMD GPUs. The grid dimensions are 1536 x 1024 x 100, with a linear spectral truncation. Transforms include 3 scalar fields, 2 wind fields and derivative calculations. (courtesy of D. Degrauwe)

For the GPU-porting of gridpoint calculations, the strategy is taken to rely on the source-to-source translation tool Loki developed at ECMWF. During the first half of 2023, a significant amount of time was spent to get familiar with this tool, and a training for ACCORD people will be/was given early October. In the second half of 2023, work started to use Loki to generate a GPU version of the ACRANEB2 radiation scheme. While several issues were encountered due to the immaturity of the compilers for AMD GPUs, we succeeded to come up with a source-to-source translation recipe that generates a working version.

Regarding the other topics in the SPTR area of the RWP2023: work is planned to start in November on Atlas/FVM; no progress was made regarding the adaptation of MultIO for ACCORD/LAM; access to various architectures is quite well, with NVIDIA GPUs being available at ECMWF, AMD GPUs being available on Lumi through the Destination Earth programme, and several ACCORD partners having NEC Vector-Engine powered HPC's.

- ***highlights from 2023***

As apparent from the activities overview, progress is made on various elements of porting the ACCORD forecast model to GPUs. A definite highlight of 2023 is that the different pieces of the puzzle start to fit together: the smart data structures and parallel granularity mechanism from ARPEGE, the refactoring of the ALARO physics control routine, the porting of the

ACRANEB2 radiation parameterization to AMD GPUs and the porting of the LAM spectral transforms to AMD GPUs. This finally culminated in the first GPU-enabled ALARO forecast running on the Lumi HPC. While benchmarking and profiling still need to be done, this definitely marks an important step.

- ***perspectives and priorities for 2024***

- Although anticipated already in 2022, the refactoring of AROME-MeteoFrance and HARMONIE-AROME is lagging a bit behind. Efforts on this topic will be increased.
- Now a proof-of-concept of the GPU-porting strategy is available with the porting of the ACRANEB2 scheme; this strategy can be applied to other parameterizations and other gridpoint calculations, maybe even to the calculation in spectral space.
- The porting to GPU of some parts of the forecast model still has to start, most notably the semi-Lagrangian advection scheme.

3.2. Towards modelling at (sub-)km resolution (HR)

- ***overview of the activity in the area***

VHR modelling encompasses activity in a wide range of Areas, with dynamics and physics probably still playing leading roles. However, EPS/predictability and MQA also address VHR modelling and specific developments like the implementation of spatial methods in "***harp***" likely will enlarge the range of suitable tools for validation. Many teams have defined VHR configurations, a few others are on the verge of doing so, and sharing knowledge seems like a priority task for 2023. As the topic grows in interest and priority, scientific tasks more and more are hosted within the relevant thematic areas.

- ***organisation of the work within the consortium***

VHR modelling is currently organized at three different levels in the consortium: a WG-VHR has met until spring 2023 and it is planned that it would resume its meetings in late autumn (chaired by the PM); a dedicated Working Week was organized in February (SMHI, see below); VHR modelling tasks are described in the RWP2023 in a dedicated WP ("**HR**") however also spread across various tasks in thematic WPs.

In addition, several ACCORD teams regularly meet to discuss VHR configurations under the lead of the WP2 management in the framework of DEODE.

- ***highlights from 2023***

A dedicated Working Week was organized for ACCORD teams on 14-16 February in the premises of SMHI. A major aim of this WW was to spread knowledge about VHR configurations, how to define and use them, with a focus on teaching teams and newcomers. The WW in-person attendance was of about 40 participants to the sessions, about 20 to the practical exercises, and about 30-50 colleagues connected online to the talks. The program was organised so as to allow presentations in the morning (hybrid format) and practical work in the afternoons with experts and newcomer staff working together and sharing knowledge. The exercises also were used for on-the-spot debriefs.

A few snapshots from the talks follow, with illustrations.
Petra Smolikovà (CHMI) introduced a wide range of dynamics, physics, and domain definition choices suitable for VHR model settings. In her very pedagogical introductory talk,

she also explained methods and tools to evaluate the model results. She gave examples of model numerical cost depending on options to ensure stability, like in this slide:

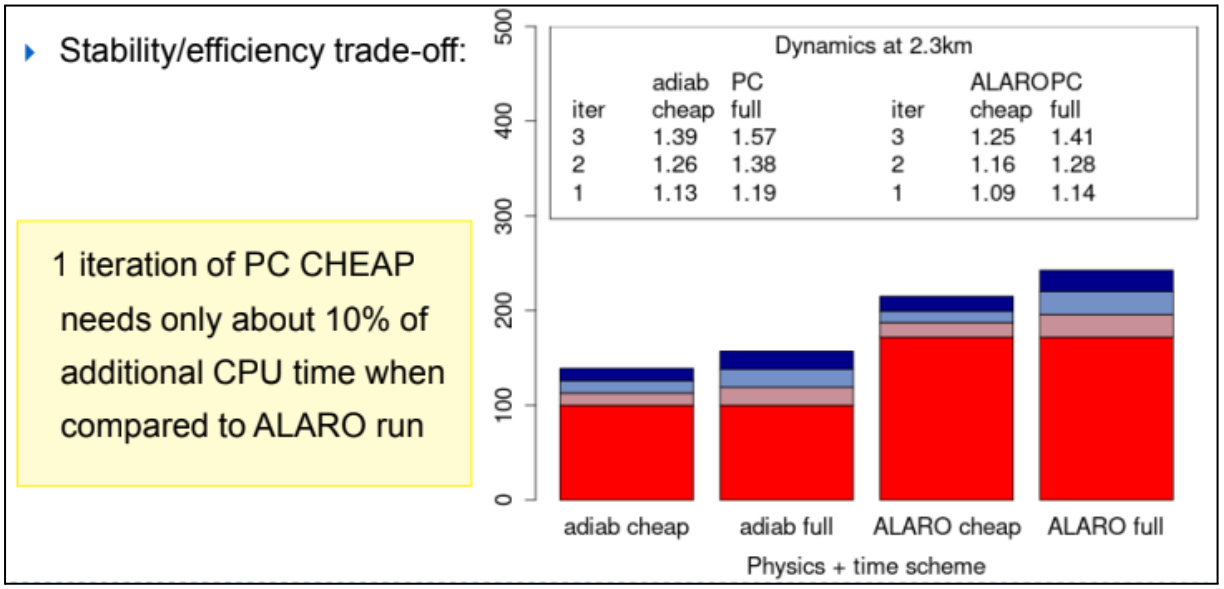

Fig. 3.2.1. Pedagogical examples of numerical cost of several ACCORD model configurations with increasing use of options to ensure stability. (courtesy by P. Smolikovà)

Carl Fortelius (FMI) gave an overview of a wide range of observations suitable for model validation at the scale of km and hm grid point resolution. His presentation was organized according to the meteorological feature of interest, like in this illustration for precipitations:

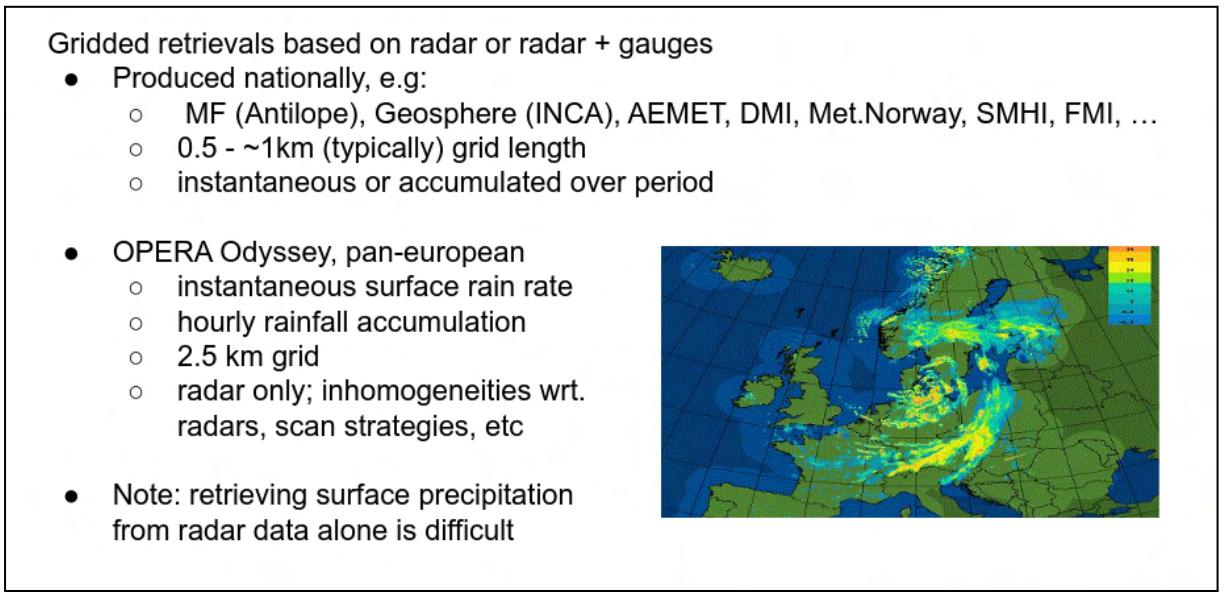

Fig. 3.2.2. Examples of gridded data that could be used for validating results of a VHR model configuration. (courtesy by Carl Fortelius)

Several use cases were shown by participants, like for instance the mesoscale storm “Kiira” which hit the southern coast of Finland on 12 August 2017 (case presented by Erik Gregow, FMI). An illustration comparing radar accumulated rain with two Harmonie-Arome model configurations is shown below. The use case was complemented by other observations and a variety of sensitivity results from Harmonie-Arome (spin-up, boundary conditions, resolution 2.5km - 750m - 500m):

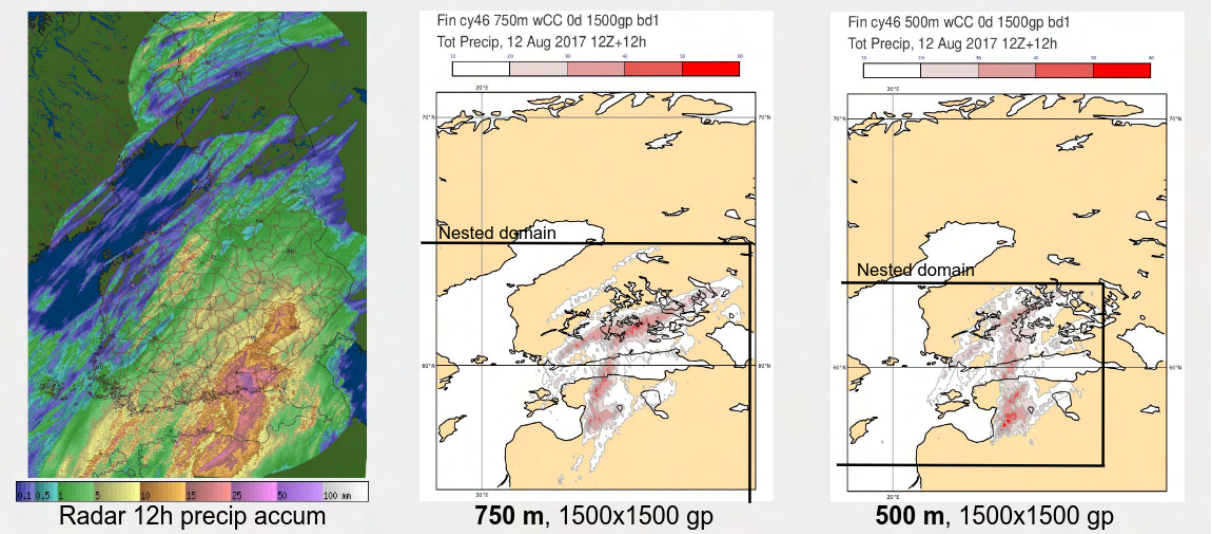

Fig. 3.2.3. Example of comparison of 12h accumulated rain from two Harmonie-Arome configurations (750m, 500m) versus 12h accumulated rain rates from radar network. On the whole, a large enough model domain seemed beneficial to better capture the patterns and intensity of precipitations, and some additional gain of realism was observed in the coastal area with the highest resolution. (courtesy by Erik Gregow)

- ***perspectives and priorities for 2024***

The last WG-VHR meeting in spring 2023 was devoted to a complete overhaul of the "**HR**" WP in the RWP. The choice was made to let any specific, topical R&D work enter the relevant WP of a thematic Area (such as 3D aspects of turbulence in **PH1**, microphysics aspects in **PH3**, numerical stability studies in **DY1** and so on). This is a small, however significant shift in the organization of VHR R&D in our consortium, as this means that the teams will be much more organized under the lead of the Area Leaders and co-leads of WPs. The new organization of the **HR** WP now reads like:

- HR1: studying the scientific added value (use cases, link with observations)
- HR2: studying the impact of surface conditions (computation or improvement of PGD files)
- HR3: studying the numerical stability aspects (on the experimental side)
- HR4: preparing a shared experimental environment on ECMWF's ATOS platform

The WG-VHR will be "called back to life" at the end of the autumn hopefully, in order to (1) re-assess its own objectives as a WG, (2) focus on sharing of experience and (3) evaluate how the above list of tasks actually could be implemented.

The DEODE WP2 will continue to work until the end of April 2024 (end of phase 1).

3.3. Dynamics (DY1 to DY3)

- ***overview of the activity in the area***

Our general goal is to provide a stable and accurate framework for our ACCORD model. The on-demand time scheme is one of those ideas that intends to increase accuracy with more iterations only when required so limiting the overall computational cost, the results are very promising and require further validation.

Another interesting development is the reduction of coupling files by writing only the frame data of the coupling files (by frame we mean the border of the domain, the one that is used by the coupling algorithm). This is not so easy because we need to save the data in

gridpoint space instead of spectral space with the consequence that the overall data reduction is not satisfactory. We will need to use more aggressive compression options in order to significantly reduce the overall size.

Work on new discretization methods leading to more conservative models by construction have continued. A number of new dynamics options were modified, phased and tested:

- NVDVAR=5, W : a new version of the vertical divergence variable and its grid point counterpart, including additional terms related to orography.
- LGWFRIC: a new sponge layer treated implicitly. This new treatment improves model stability by minimising the reflection of fast waves at the top of the model.
- LSIPRA: An additional semi-implicit parameter for hydrostatic surface pressure to stabilise the model on very high relief (e.g. the Himalayas) where non-linear pressure residuals are important.

Those options developed in the past few years have been improved and tested on problematic cases such on the Svalbard domain with interesting results though not solving all the oscillations.

A recent version of the new FVM dynamic kernel developed at ECMWF has been downloaded. This year we set up 3D orographic test cases over the Alps. We also developed vertical levels close to those used by ACCORD but in the *z*-coordinate following the terrain. The goal is to have a fair comparison in terms of cost and stability with the ACCORD dynamics.

Finite elements on the vertical : A new version of vertical finite elements (VFE) has been coded to make this option more stable. It is available in cy48T1. Tests were carried out this year, and a phasing is currently underway to adapt these modifications to the cy49T0 cycle in preparation for CY49T1.

Improved interpolation in the semi-lagrangian schemes : New options have been developed, coded and integrated into the code in 48T1_op1. The first option NQMGMV modifies GMV[1] interpolations, enabling linear interpolations (which are currently not possible), or "extended" cubic interpolations. The other option NQMGF takes care of GFL[2] and enables the same "extended" cubic interpolations. This extension of cubic interpolations concerns the quasi-monotonous option, as well as thresholding to 0 for GFL. Impact studies of these options are in progress.

Improved interpolation, mass conservation: The aim is to tackle some of the weaknesses of the semi-Lagrangian scheme, notably the treatment by a filter that conserves mass and avoids values beyond the limits, which happens all the more when the interpolations are of high order. These ideas have begun to be coded. Work is also in progress on pre-conditioning the point-of-origin search in AROME, and on improving the COMAD scheme on the vertical.

Finally a lot of testing and tuning at very high resolution have been performed in the context of the DE_330 project. The goal is to be able to have a stable configuration at 500m and 200m. We believe that we can state that the 500m configuration is secure now, using more iterations, more stable SI setting, and more relaxation in the upper layers. Also the

[1] GMV fields contain surface pressure, horizontal wind components, temperature, as well as the additional variables representing vertical divergence and departure from hydrostatic 3D pressure in NH

[2] GFL fields contain specific humidity, cloud and precipitating water species, any other advected variable

hectometric operational or near-operational configuration with HARMONIE-AROME or AROME gave us more confidence in our capability to have a working model at those resolutions.

- ***organisation of the work within the consortium***

The dynamic work is divided between different services: Météo-France, CHMI, SHMU, Met-Eireann, RMI. Collaboration exists in many respects since the skills on different aspects of the code are shared between different services. In the framework of the Destination Earth project, a more integrated collaboration on specific tasks will take place.

- ***highlights from 2023***

This year the activity has been strongly impacted by the DE_330 project whose goal is to provide an on demand configuration at 500m resolution and 200m. One of the important aspects for this project is to have a stable model at 500m (the 200m resolution will be more for further project phases).
Another important key research and development concerns our recruitment for working on the FVM new core and the work that has started to learn how this new system works and to start making modifications to the code to be able to compare and benchmark it to ACCORD. This is a completely new working environment intensively using python and domain specific language (DSL) that translates python high level routines into usable optimised low_level langage. This could represent our future working methods and we should anticipate those coding techniques.
Finally it is worth mentioning the continuation of the work on the formulation of Euler equations as the increment of hydrostatic primitive equations. The aim is to add "non-hydrostaticity" gradually and omit it where numerical stability is questionable (with vertical or time from start dependency). In parallel, control parameters are introduced in the linear model enabling its modification after the linearization from the full model. These degrees of freedom are introduced with the aim of improved stability similarly to how it was done previously for the SITR/SITRA parameters. This work has not progressed as fast as what was expected this year, but this is a very innovative approach that deserves to be highlighted.

- ***perspectives and priorities for 2024***

The DE_330 will still be a priority in 2024, the goal will be to go further down towards 200m or even finer resolution. The stability will be a challenge there and we should not be afraid to reduce the truncation even if the actual resolution will be slightly modified by this parameter.
The exploration and comparison of the FVM model remains a priority since this could be the next dynamical core of the IFS system and our collaboration with ECMWF could be modified depending on the model we use in the future.
The continuation of the improvement of the historical spectral core will remain our everyday task since it has been proven to be robust. The spectral aspect for example, that could be seen as a problematic feature as regards the scalability and data exchange of the spectral transforms can now be seen as a strength because of the possibility to reduce the spectral truncation to improve stability with moderate impact on the quality.

3.4. Data Assimilation (DA1 to DA4, DA6, DA7)

- ***Overview of the activity in the area***

During the first two quarters of 2023, a total of 187.25 person months was reported as contribution to data assimilation (DA) work, which is 41.4% of all expected. The two largest parts are implementation of existing (**DA3,** 55 person months) and developments of new observations (**DA4,** 57 person months). The distinction between "existing" and "new" observation types in the work plan became unclear after a few years; the newly proposed work packages for 2024 therefore divide observational packages by type (in-situ, ground-based remote sensing, satellite data). Considerably less person months are devoted to algorithmic developments (76.75 person months in total for packages **DA1, DA2** and **DA6**), including work on initial implementation of data assimilation (former DAsKIT). The algorithmic developments concentrated around implementation of DA algorithms under OOPS, in particular the implementation of 3D/4D ensemble variational (EnVar) assimilation. This work has matured and provides very encouraging results which has allowed implementation in the e-suite at Météo-France. Optimizations of pre-operational 4D-Var setups were carried out by several HIRLAM members.

- ***Organisation of the work within the consortium***

The DA area collaboration was reorganised during 2023, based on previous experience with topical working groups (originally research and support teams). The idea was to share the responsibility for coordination of DA packages between the co-chairs and to keep the sub-teams manageable in size. These teams are now fully aligned to the nine new RWP2024 packages, and the related teams are now called research and development (RD) teams. The teams, which are expected to meet twice a year, also have their own Wiki spaces (https://opensource.umr-cnrm.fr/projects/accord/wiki/Data_Assimilation) and other means of communication. The goal is to review the progress periodically and efficiently spread information on relevant ongoing issues.
The members also agreed a common reporting practice within a shared document (https://docs.google.com/document/d/11HIS8WrfpxvDcZaVQXVEDhcmDDmvmgX6frrM1mxmLv4/edit), with a goal to minimise the periodical reporting (to do it once by staff members), while taking into account the current reporting practices and requirements of management groups in the sub-consortia (HIRLAM, LACE, Météo-France). This procedure makes the accomplished work and progress more transparent and visible and contributes to easier identification of staff working on the same topics across the consortium. Another building block to enhance the collaboration in the consortium are the working weeks (WW). With the goal to offer each staff member an opportunity to attend the WW once per year and keeping the groups reasonable in size (<20 participants), 4 working weeks were organised in 2023. Each of those covered topics managed by a few RD teams and the agenda was coordinated between ALs and the co-chairs. In the future, starting with 2024, 3 WWs per year are planned, with one hosting institute from LACE, HIRLAM, and other ACCORD (DAsKIT)+MF. On top of these activities, regular meetings with the DAsKIT group are taking place every other month in order to review, discuss and assist their implementation progress.

- ***Highlights from 2023***

The teams have progressed with the majority of the planned activities in 2023. The algorithmic developments concerned among others the pre-operational implementation of

4D-Var (non-OOPS version) in HIRLAM, implementation and phasing of ensemble-based algorithms (3D/4DEnVar and hybrid EnVar in OOPS), application of field alignment for sub hourly DA prototypes, implementation of initialisation techniques and cycling strategies as well as handling of host model information (Jk, BlendVar). Related to this, diagnostic tools to assess DA performance are more systematically gathered and shared within the consortium. Observation handling progress concerns a refined use of traditionally used observation types (satellite, radar, aircraft, radiosondes, GNSS-derived data etc.) and exploitation of various crowd-sourced data (netatmo, smartphones) and other novel observation types (microlinks, balloons, GNSS-like observations from alternative sources). Extensions in coverage and quality enhancements of Mode-S data from EMADDC has allowed for further impact studies and applications. Several adaptations towards nowcasting (rapid-refresh, RUC) have been applied, including work with refinements of cloud ingest initialization and investigation of spin-up and short observation cut-off. Three areas with significant developments and achievements are highlighted in a bulleted list below: application of 3DEnVar in OOPS, progress with use of Mode-S and radar data. Apart from that, a large progress has been made with use of satellite data at several centres, including enhanced use of low-peaking channels for microwave sensors, new instruments in operations for MetCoOp, treatments of reflections over ice and snow, developments of footprint operators for scatterometers and radiances, and ongoing activities in order to apply all-sky (cloudy, rainy areas) assimilation for microwave, infrared and visible radiances. These activities are expected to provide a significant outcome in 2024. Several preparatory activities for forthcoming satellite missions (sensors IASI-NG, MWS, FCI, IRS, LI), including OSSEs, were also carried out, to be continued when new data is available.

- 3DEnVar in OOPS: The migration to OOPS DA codes enabled first-operation ready configuration of 3DEnVar at Météo-France which has been running as a real-time E-suite in cy48t1 since September 2023 (Brousseau, Vogt et al.). This was made possible by previous validation of OOPS 3D-Var and phasing of EnVar from cy46t1, including tuning of localization. Long-period impact studies have confirmed major positive impacts of 3DEnVar compared to 3D-Var, both for average scores for all parameters (figure) and over specific case studies (high precipitation events, winter storms and fog). Studies on the ensemble size were conducted and the ensemble size of AEARO (AROME EDA at MF) has been increased to 50 members, resulting in a clear positive impact on the forecast quality. Small negative humidities originating from SPPT perturbations were shown to cause crashes during heat waves in 2022 and a correction was applied to avoid them. An Incremental Analysis Update (IAU) approach has been applied, in order to add analysis increments smoothly during the forecast. Initial implementation and experiments with 3DEnVar in their DA systems are also carried out at GeoSphere Austria and Met Norway.

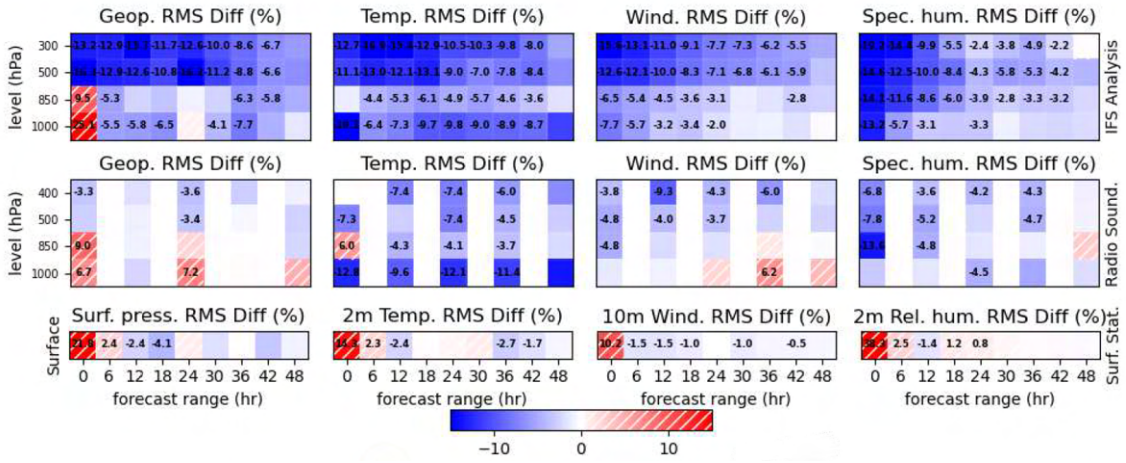

Figure 3.4.1. Percentage of reduction of RMSE for 3DEnVar wrt. 3DVar in AROME-France for different parameters over 3 months (September-November 2022). References are IFS analyses, radiosonde and surface stations. (courtesy by V. Vogt)

- Impact of Mode-S observations: The study carried out by HIRLAM in the MetCoOp domain (Lindskog et al.) focused on impact of Mode-S data using 3DVar assimilation. Mode-S EHS datafrom EMADDC was shown to contribute to compensating for huge loss of AMDAR data during the pandemic period and the quality of EMADDC temperature data was demonstrated to have improved substantially during the last years. The impact on forecasts was however neutral overall with impact for individual active weather situations. At Meteo France (Vivien Pouret et al.), the Mode-S EHS data were evaluated in 3DEnVar, and with this assimilation approach, a strong and positive impact was demonstrated (figure). Changes to treat duplicated data were necessary. The use of Mode-S data is planned also in ARPEGE, using a dedicated adaptive data weighting technique.

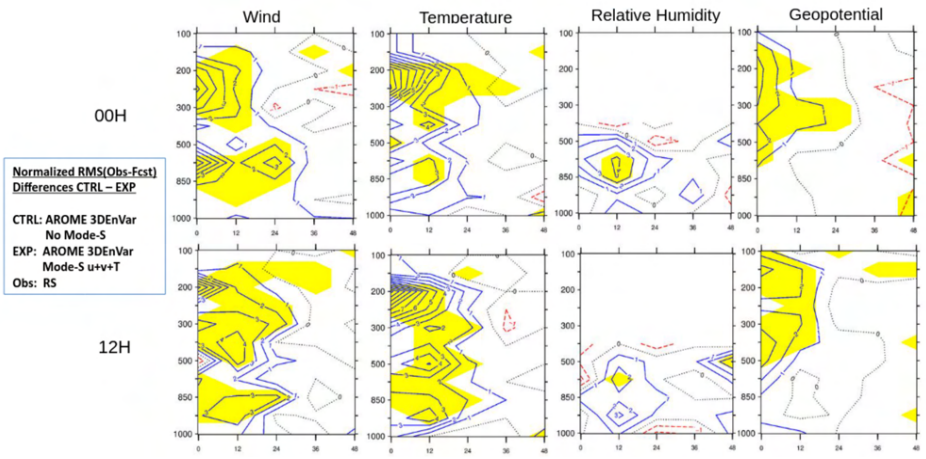

Figure 3.4.2. Normalised RMS differences between experiments with and without Mode-S EHS data in 3DEnVar AROME-France over 1 month period. Improvements in blue, degradations in red. (courtesy by V. Pourret)

- Advances in use of radar observations: In HIRLAM (Ridal et al.), optimal settings for first guess departure limits, observation errors and super observations for radial winds

were studied, using the radar data from ODE. A solution for wind dealiasing (Smerkol et al.) was implemented and made available as part of Python radar preprocessing software HOOF. As several tests in different setups ALARO CSC (Panežić et al.) showed a drying effect from reflectivity assimilation, a number of proposals for mitigations was proposed and coded, including increase of observation error for dry observations and restrictions on use of dry (undetect) observations. At Météo-France (Martet et al.), a direct assimilation of reflectivities in AROME-France in the framework of 3D-EnVar, with hydrometeors in the control variable, demonstrated a positive impact on forecast, particularly in the first few hours.

- ***Perspectives and Priorities for 2024***

The following tasks will have special attention in 2024 (new package names with respect to 2023):

- ➢ Continue exploitation of alternative observation (personal weather stations, smartphones, etc.) by paying attention to relevant quality control techniques (**DA1**).
- ➢ Continue the implementation of radar data, including direct reflectivity assimilation, wind dealiasing and migration to the new OPERA volume data hub (Nimbus) (**DA2**).
- ➢ Refine and enhance assimilation of satellite products (all-sky radiances, low-peaking channels, lightning data, hydrometeors in control vector, footprint operators), bias correction procedure, and prepare for assimilation of new satellite products (**DA3**).
- ➢ Operational implementation of 3DEnVar under OOPS at Météo-France; pursue associated evolutions and experimentations with OOPS of 3D-Var and 3DEnVar by other ACCORD members (**DA5/DA6**).
- ➢ Continue porting/implementation of LAM 4D-Var under OOPS (**DA5**).
- ➢ Pursue experimentations and evaluations of the 3D/4DEnVar and hybrid EnVar formulation under OOPS, including the assimilation of high frequency observations every 15 minutes (**DA6**).
- ➢ Continue exploring DA procedures, settings, and frameworks appropriate for nowcasting applications (**DA7/DA8**).
- ➢ Continue development and sharing of common diagnostics tools for monitoring of DA functionality, overall optimization of DA systems (**DA8**).
- ➢ Identify and explore machine-learning approaches in DA (**DA9**).
- ➢ Continue supporting the “catching-up” implementation process by all ACCORD members.

3.5. Physics parameterizations (PH1 to PH8, PH10)

- ***overview of the activity in the area***

The ACCORD physics parameterization section has been reorganised to better address the goals set in the interoperability roadmap for physics parameterization within the ACCORD community. In that respect, within the RWP2023, the Physics core parameterization work sets along three main thematic areas: turbulence and convection (**PH1**), radiation (**PH2**), microphysics and clouds (**PH3**). However, other important aspects, such as interfaces between the physics and the dynamics (**PH8**), and the atmosphere-surface coupling (**PH7**) has been emphasised. Other commonly important activities have been related to the development of new model output and postprocessing parameters (**PH5**), cloud-aerosol-radiation (**PH6**) interactions, as well as further developing the common single column model (**PH4**). Within

the fast-growing activity in AI for Meteorology, the possibilities for emulation and/or improvement of existing physics-based parameterizations by stochastic and ML approaches in our systems has been also explored (**PH10**).

The scientific activity in the ACCORD model physics during the first half year of 2023 has been mainly focused on developing the parameterizations for NWP systems with resolutions towards hectometric scales, including near-real time aerosols as an input and forcing, and more advance model diagnostics (see Section *Transversal activities*). The ACCORD Physics management activities during this period have been focused on developing the RWP2024, among others ACCORD relevant documents (Section 2), but specific effort has been made to coordinate and organize the preparation for the Physics Working Weeks in Sodankylä, Finland (see next section).

- ***Transversal activities***

A large amount of work was devoted to prepare the physics parameterization in NWP system setups with very high resolutions ($\Delta_{x,y}$ = 750 m, $\Delta_{x,y}$ = 500 m, $\Delta_{x,y}$ = 200 m). This involved developing and testing scale-adaptive (or scale-aware) convection schemes (e.g. deep convection in ALARO, shallow-convection and cloud scheme in HARMONIE-AROME), parameterization of the 3D-turbulent effects in areas with a steep orography (based on Goger et al. 2008), as well as searching for a balance between the accepted computational cost and the forecast precision, while keeping the numerical stability of the system as less affected as possible. Several numerical experiments were conducted over different domains with a variety of surface-orographic and meteorological conditions using AROME and HARMONIE-AROME (e.g. over an urban area -Paris-, over high and moderate mountains -Alps, Innsbruck as illustrated in Fig. 3.5.1-, over a sea-coastal area under storm conditions -Baltic sea-south Finland, see Fig. 3.2.3 in Sect. 3.2-, over Ireland, over the Netherlands -Cabauw-).

500m	dt	level	Shal XCMF	Goger	Rayt	SITRA	SIPR	NSITER	CHEAP	FULL	Cost
GL6J	20s	90	0.065	no	360s	50	60000	2	F	T	x4.7
GLMG	20s	90	0.065	no	360s	50	90000	1	F	T	x3.5
GLOR	20s	90	0.03	yes	360s	50	90000	1	F	T	x3.8

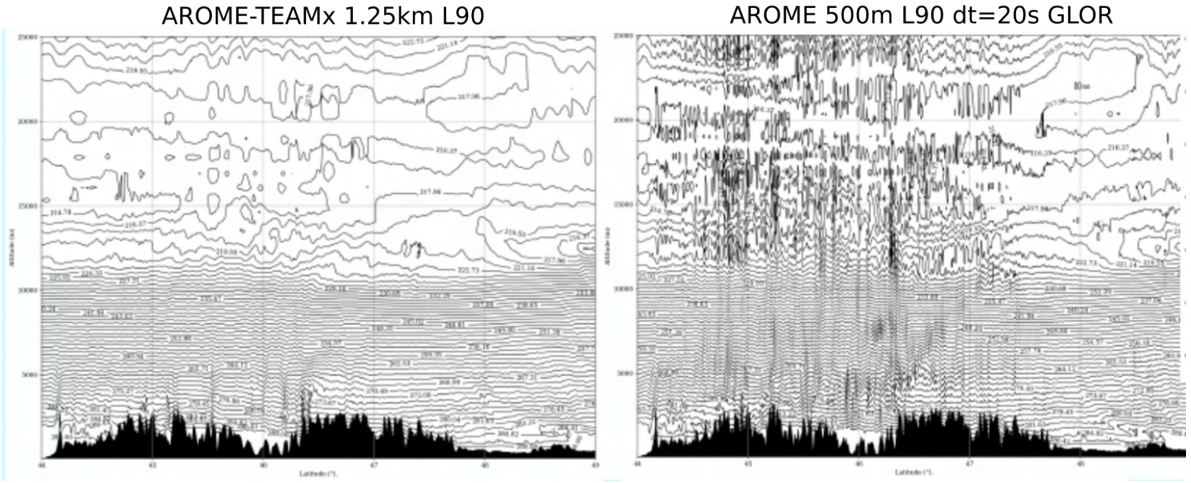

Figure 3.5.1: An example of stable high resolution ($\Delta_{x,y}$ = 500 m, 90 levels) AROME setups with the pseudo 3D turbulence effects (table), and a comparison of the temperature over the

Innsbruck-area cross section on 18.08.2022 (strong precipitation case) against the control case of AROME ($\Delta_{x,y}$ = 1.25 km, 90 levels) (courtesy by Eric Bazile).

Additional cases were conducted to validate the latest setup of the TKE-based mixing length formulation in ALARO (TOUCANS scheme), but also to validate the horizontal diffusion scheme and resolved vs. subgrid TKE in a multiscale environment (Fig. 3.5.2).

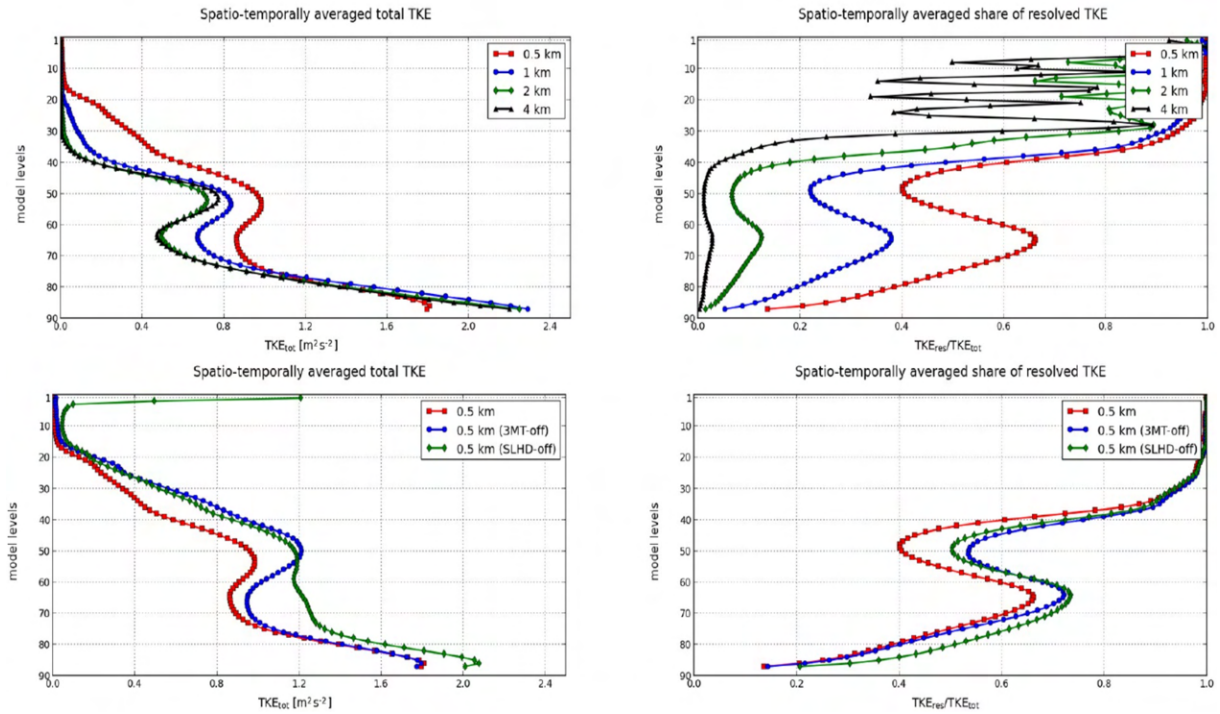

Figure 3.5.2: (upper panels) Spatio-temporally resolved TKE, subgrid TKE, total TKE and share of resolved TKE across resolutions for the convective case 21.6.2018. (averaged between 12 and 18 UTC over a sub-domain with dimensions 256 km x 128 km). (lower panels) Different experiments at 500 m but without deep convection scheme (3MT-off) and Semi-Lagrangian Horizontal Diffusion (SLHDoff). as well as with SLHD being replaced by full spectral diffusion of operational AROME-France (courtesy by Mario Hrastinski).

Based on this work we are now able to adjust and use the parameterization suitable for hectometric scale and steep orography up to 500 m horizontal grid spacing. There are still challenges when going to 200 m grid mesh due to dynamics instability (work in progress). It is worth noting that increased an numerical instability is expected when increasing the number of vertical model levels in the boundary layer and the troposphere for the default dynamics setting (e.g. HARMONIE-AROME). To stabilize the numerical instability one needs to decrease the timestep (dt) or/and increase NSITER. An alternative option is to utilize the vertical coordinate with a lowest atmospheric level at around 10 m (MetCoOp setup). For the current 2.5 km HARMONIE-AROME with 90 levels, this allows keeping large timesteps (dt=75 s) and NSITER=1, thus the same computational efficiency.

The new optimized version of ecRad is already running in Arome CY48T3. Significant savings especially in Arpege already are noted (20% for the whole model when using McICA) and that's before testing the new gas optics ecCKD (in progress), which should give substantial further savings. The plan further includes its implementation in cy49t0. However, for some reason yet to be understood, the new TripleClouds scheme was found slower than the new McICA version even though the former has been optimized (work in progress) *(courtesy by Peter Ukkonen and Ryad El Khatib).*

The strong cooperation between the CSCs in the field of cloud-aerosol-radiation (CAR) interactions led to encouraging results in model performance when replacing Tegen and CAMS aerosols climatology with near-real-time (NRT) aerosols. There is an improvement at the low clear-sky-index (CSI) end of the spectrum, with the large negative biases being greatly reduced. However, short-wave (SW) radiation now seems to be a little overestimated on clearer days. Convective precipitation amounts seem to be increased when NRT aerosols are used. The minimum value of CDNC was also tested, as well as sea-salt emissions and a series of fog cases. This CAR work is currently being described in a written manuscript with an aim of peer-reviewed publication.

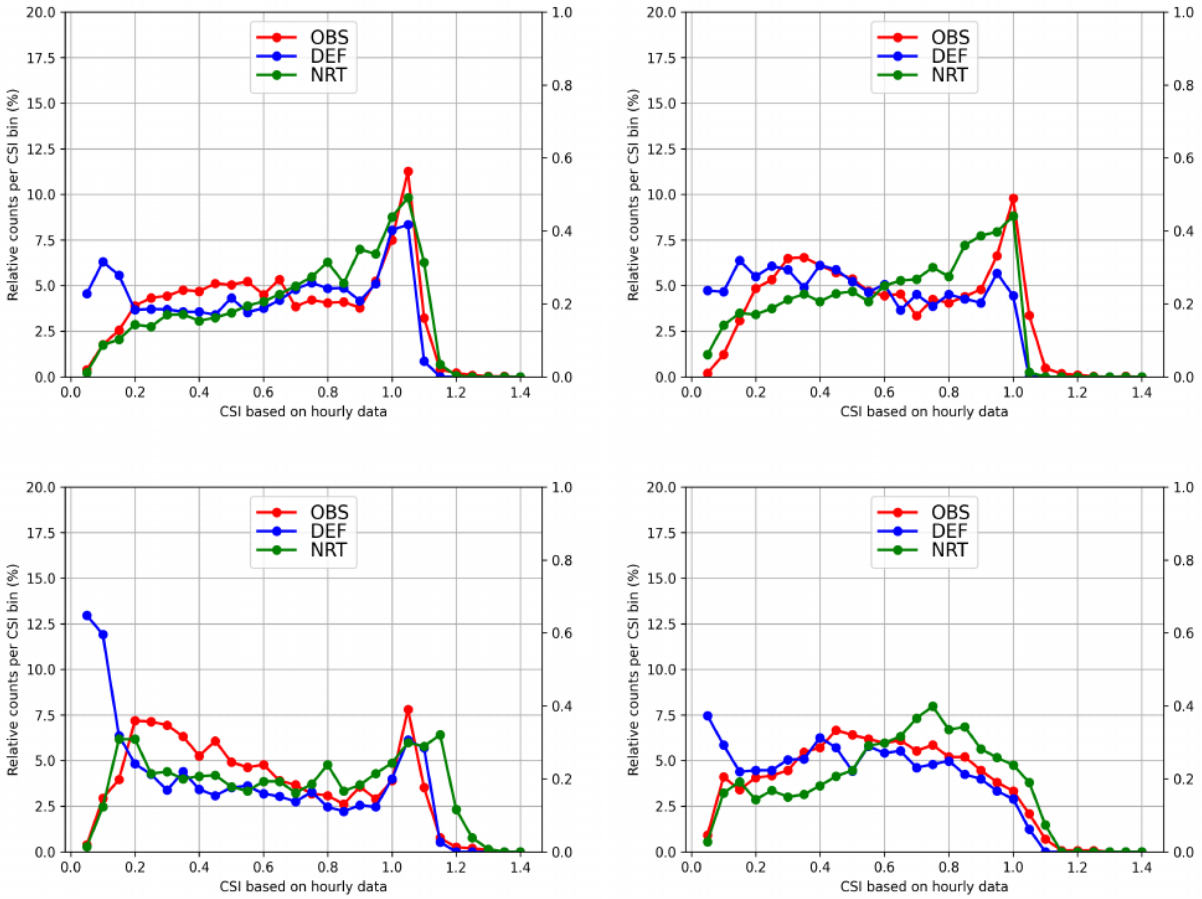

Figure 3.5.3: One-dimensional clear-sky index (CSI) for Spring, Summer, Autumn and Winter 2 week periods. There is a noticeable improvement in low CSI cases but in general there is less cloud when NRT aerosols are used (courtesy by Emily Gleeson).

The work on establishing a common single column model framework for parametrization validation has been advancing by planning and organising a common CSCs hands-on workshop as part of the Physics Sodankylä working week (18-22.09.2023[3]), next to the activity in the surface-atmosphere coupling in a stable boundary layer. Therefore, the stable boundary layer idealized cases (GABLS1 and GABLS4) have been selected to be tested on the Environment for MUSC simulations (EMS) python-based software to serve as a tutorial studies for developing additional and new cases of interest (e.g. Sodankylä case). This work has gained momentum and ongoing tests are in progress.

Some work regarding enhanced model output and postprocessing parameters has also been reported. For instance, new diagnostics of deep convection for aviation sector has been included in full-pos, diagnostics of lightning (index), rain-snow limit, sea ice, type of

[3] https://opensource.umr-cnrm.fr/projects/accord/wiki/Physics_WW_Sept2023

precipitation, visibility, have been revised, as well as computation of eclipses in AROME/ARPAGE for photovoltaic have been tested. Furthermore, work has been conducted to post-process an output from high-resolution simulation for special cases of freezing rain.

The surface-atmosphere coupling in the stable boundary layer and the large temperature bias in stable stratification has been a long-standing issue, and thus a topic of discussion during the Physics Sodankylä working week. There has been agreement among the participants to document and publish the number of suggestions and the consequent developments towards more robust surface-atmosphere coupling in statically stable stratification. Additionally, progress have been reported during the EWGLAM meeting towards inclusion of the roughness-sublayer parameterization[4] in HARMONIE-AROME, but also in ACCORD-NWP SURFEX repository (so other CSCs can use this development).

Finally, the Review manuscript "*Perspectives towards stochastic and learned-by-data turbulence in Numerical Weather Prediction*" (by Metodija Shapkalijevski) has been revised and resubmitted to Weather and Forecasting journal, and at the moment of writing these lines it has been accepted for publication within the journal.

- ***Short summary of highlights for 2023 per CSC***

➢ **AROME**

- Two new AROME configurations have been built around Paris and Marseille at 500 m horizontal resolution, with 120 vertical levels (the lowest one at 2.5 m). Numerous tests of surface scheme options (EcoSG, soilgrid, Garden, OpenStreetMap...) and physics/dynamics tunings have been performed. These two new AROME versions are scheduled to become operational by the time of the Paris 2024 Olympics event.

- Within the TEAMx project, several tests have been performed over Austria at 500m and 200m with AROME to evaluate the impact of a preliminary version of the pseudo 3D turbulence with the “Goger term”. Detailed evaluation for several parameters T2m, 10m wind and precipitations is done and shared with the Austrian colleagues

- PHYEX (“externalization of MesoNH Physics”) entered CY49T1, and will ensure a closer connection between MesoNH and AROME developments in atmospheric physics as well as possibilities of new tests in AROME thanks to the access to all MesoNH namelists switches. A more flexible LIMA microphysics scheme will for instance be available in this new package.

➢ **ALARO**

- ALARO physics parameterization code refactoring, including new apl_alaro, code reorganisation into parameterization centred blocks

- usage of CAMS aerosols in radiation, as replacement for Tegen monthly means or using near real time CAMS aerosol fields

- implementation and testing of the Lopez microphysics scheme within the ALARO microphysics shows that the Lopez evaporation could be a useful alternative to the

[4] https://drive.google.com/drive/folders/1-feg8V0Ca6HFp2A6l9zc4lV-rhbGZkd6

current Kessler type evaporation. The version with Abel-Boutle distribution for rain was developed as an alternative to the Marshall-Palmer

- The alternatives to the autoconversion process in ALARO were implemented and tested leading to the dismissal of Khairoutdinov and Kogan (2000), while Chaboureau and Pinty (2006) was found useful and implemented in the common code
- additional diagnostic fields were coded in fullpos to be diagnosed with ALARO: Storm relative helicity; Updraft helicity; Updraft and downdraft track; Calculation of both MUCAPE and MLCAPE
- TOUCANS turbulence scheme options and further developments under revision
- considerable effort has been made in coupling ALARO with SURFEX, especially concerning computation of fields needed to describe surface impact on turbulence

➢ **HARMONIE-AROME**

- New physics components added to CY46 (but not default settings): NRT aerosols, ICE-T, Wind farm parametrization, scale-aware shallow convection, radiation added to Vfld files, OCDN2 added to rain_ice.F90, on/off switch for shallow convection scheme
- Continued to strengthen our links with HCLIM as regarding testing physics options (e.g. RFRMIN(24), ECUMEv6), Ph.D. & MSc students working with KNMI, MetNo and really strengthening the physics team
- Testing of new components has started: NRT aerosols and ICE-T have large impacts on clouds and radiation - lots of tuning options - a difficult task. Turbulence/Shallow Convection: Wim - PhD, Marvin, and students are growing expertise in this area so that we have a larger pool of experts in this area
- DDH working in CY46, also possible to use DDH output to create MUSC atmospheric files, progress with a common MUSC version, code refactoring
- VHR: Workshop at KNMI and subsequent paper, testing of increased horizontal and vertical resolution, 90L, LBC coupling

- ***perspectives and priorities on transversal topics for 2024***

We have started to apply parameterizations suitable for hectometric scale and steep orography, however there are still many challenges that should be overcomed in the next period. While the HR-WG (Section 3.2) focuses on studying the applicability and added value of the developed NWP systems towards hectometric scales, the ACCORD physics group will continue to provide valuable contributions to these systems by trying to resolve some of the main raiming issues (with respect to physics parameterization). For instance:

1) When and how should the shallow-convection schemes be deactivated, and how should they be used in the “gray zone” (**PH1**)?
2) How to properly compute the horizontal gradients when moving between spectral and grid space (e.g. 3D effects in turbulence, **PH1**)?

3) How (often) should a radiation scheme be used, and can we use SPARTACUS on a coarser resolution for assessing 3D effects? (**PH2**)

4) How can we establish dynamics stability at VHR (e.g. 200m) resolution over steep orography? More generally, what is the optimal and satisfactory balance (tuning) between model resolution, numerical stability, and computation efficiency? (**PH8, PH1**).

Many of these aspects will be discussed on the coming 3D-turbulence working week in Toulouse (4-8.12.2023), during which internal (ACCORD) and external competence is expected to gather (e.g.. Brigitta Goger, Ivan Bastak, Joan Cuxart, Richard Forbes).

Additionally, the availability and possibility of using near-real-time (NRT) aerosols as an input in our NWPs has been beneficial for the forecast. There is a need of using the NRT aerosols in more dynamic/prognostic mode (advection/transport, removal by deposition) (**PH6**). Moreover, we are starting to use more complex microphysics (more prognostic hydrometeors, LIMA) to account for more realistic precipitation forecasts (**PH2, PH6**). With respect to atmosphere-surface coupling, there are already ongoing activities on identifying and possibly improving surface-atmosphere coupling parameterizations in very stable thermal stability (low wind speed) (**PH7**), as well as improving the model diagnostics (**PH5**). The first (preliminary) results of the new surface-atmosphere coupling in urban areas are expected during 2024 (**PH7**). A serious effort has been given to advance the development of the common single column model (MUSC, **PH4**) towards a comprehensive tool for physics parameterization intercomparison within the ACCORD's NWPs. The currently-established working momentum, in that respect, is expected to strengthen during the next period when a fully-devoted MUSC working week is planned (early spring 2024).

3.6. Surface analysis and modelling (SU1 to SU6)

- ***overview of the activity in the area***

The ACCORD strategy for 2021-2025 divides the surface activities in three main topics: surface model, physiography and data assimilation. In this section we present some specific progress in these topics during the last year.

Surface model:
In LACE the work on debugging ALARO-CSC with SURFEX continued delivering fixes for known problems (except TEB). Compatibility problems with the old ISBA scheme were addressed (FACZ0 scaling, evolution of prognostic snow albedo and setting of surface thermic coefficient) and new features were introduced into SURFEX (roughness length averaging enabling inclusion of orographic component, snow impact on roughness length via snow height and TOUCANS screen level interpolation after SURFEX). Finally, the related code was prepared for the CY49T1 (to be updated based on the review comments).

For the HARMONIE-AROME CSC, the testing of the ISBA-DIFF scheme vs ISBA-FR continues, with the purpose of running ISBA-DIF operationally. Currently the main problems are: (i) early snow melting due to the joint temperature profile in the soil under snow-covered and snow-free parts of the grid box, (ii) too strong heat flux within the soil, (ii) dry bias. The TEB scheme is continuously tested against town observations (Paris, Turku). SICE scheme evaluations continue with use of different observations, including ice mass. Testing of ECUME/ECUME6 schemes against flux observations of the field campaign near Barbados

using the C. DeMott method showed that ECUME gives less latent heat flux, which is more realistic. The C. DeMott method suggests comparisons in the phase space, which allows evaluation of the flux calculation algorithm independently of other parts of the model and weather conditions. Other activities: further testing of the FAKETREE option, removal of unrealistically thin snow in spring in ISBA-FR.

For the AROME CSC at Météo-France, considerable efforts have been invested to verify the SURFEX version v8.1 behaviour in cy48t, since an earlier version of SURFEX was used in older cycles. It has been confirmed that results from older cycles can now be reproduced with cy48t for both AROME and Arpege setups. In addition the multi-layer physics options of SURFEX, especially the diffusion soil scheme and the explicit snow scheme, has been evaluated in cy48t AROME-France experiments. Problems with snow cover in the Alps and issues with summer diurnal cycle of T2m and Q2m have been identified. These issues will be taken into account in further tests and development.

Physiography:
In LACE the importance of the tree height as a tuning variable for roughness and 10m wind speed has been emphasised. Together with the rest of the ACCORD surface community an agreement on how to allow useful tuning parameters for tree height has reached. There is also ongoing work at Geosphere Austria on development of fine scale cover maps over Europe using ML approach.

In HIRLAM, activities started to use the ML methods for producing very fine resolution physiography maps (60m) for Europe. Activities to improve ECOCLIMAP SG for the polar regions (Iceland, Greenland, Svalbard) are ongoing. Validation experiments are ongoing based on optimisation of the PGD algorithms for ECOCLIMAP SG. In addition alternative input file format of NetCDF has been introduced which leads to considerably reduced file sizes.

Setups with ECOCLIMAP SG for the AROME CSC at Météo-France have confirmed that the need for increased roughness over open land is needed (known as the FAKETREE option in SURFEX).

Data assimilation:
LACE: Testing of satellite products in SEKF is ongoing. Assimilation of superficial soil moisture (ASCAT H08-SM-OBS-2) was tested and is technically feasible for operations, but more scientific effort is needed since results are not yet satisfactory. Pre-processing, reprojections of SCATSAR-SWI were performed. LAI observations based on Sentinel-2 data were tested in SEKF with prognostic LAI in SURFEX with positive impact on T2m. Tuning of CANARI OI was investigated for better performance of BlendVar systems and to remove T2m forecast jumpiness that arises due to too large deep soil reservoir increments in CANARI. CANARI first guess check of snow height was made dependent on the station height to avoid rejection of mountain stations (not yet in common code).

In HIRLAM further testing and improving of snow DA is ongoing: (i) bug correction in the temperature profile initialization after DA in the multilayer snow scheme, (ii) testing of different approaches for SWE/SD using model snow density, (iii) testing of different approaches for patches. Comparison of several SIC (sea ice concentration) products showed high discrepancy of them, especially in the ice margin zone. Recommendation is to use the uncertainty information from SIRANO SIC product to develop a perturbation method. Results of comparison of SEKF and EnKF for Arome-Arctic showed slight improvement for humidity

scores with EnKF. First attempts were taken to use a ML approach for an observation operator in assimilation of Sentinel-1 data for soil moisture.

For the AROME CSC at Météo-France, a 2D-Variational approach in the OOPS framework has been introduced as a replacement for the currently operational 2D-OI surface analysis. The EDA system is utilised in a 2D-EnVar version of this implementation. Tests so far show that the CANARI-OI results can be reproduced although with slight degradation during the first few forecast hours. Further tests and development are ongoing. A diffusion soil version of the vertical OI soil assimilation scheme has been introduced which will be utilised in validations of the multi-layer surface physics setups including surface data assimilation.

- ***organisation of the work within the consortium***

We have continued with the agreement among surface leaders to arrange one surface working week per semester at ACCORD level, the spring one in-person and the autumn one online only. Last November the online version gathered some 15-30 participants attending seven different sessions. This year, in May, the in-person one was hosted by SMHI in Norrköping, Sweden, where 19 people participated at site and 15 by remote connection. The next online working week will happen November 20-24 2023.

In addition to the ACCORD overall working weeks, a couple of specialised WW have been arranged, one on ML in physiography last October at FMI (Helsinki) and one dedicated to multi-layer surface physics this October at Météo-France in Toulouse. A WW on TEB/urban aspects will take place at KNMI (Netherlands) in December.

The online monthly meetings on the third Tuesday of every month have continued during the year. They are appreciated since they give the opportunity to bring up detailed presentations and discussions in between the working weeks.

Hard and successful work has been invested during the year to reach a common ACCORD SURFEX version in our NWP Github branches, the IAL one and the SURFEX-NWP one. For the first time we will reach a truly common version in cy49t.

- ***highlights from 2023***

Assimilation of soil moisture

Matjaž Ličar (SI) performed an inline data assimilation experiment of satellite based superficial soil moisture observations (ASCAT H08–SM-OBS-2) with SURFEX. Raw data were downscaled and pre-processed from 25 km resolution to 1 km in GeoSphere Austria. Assimilation of SSM was performed using SODA/SEKF with 4 control variables (TG1, TG2, WG1, WG2) twice a day. It was proven that due to the near real time nature of the satellite observations, the use of such a product is feasible in an operational setting. The neutral or even deteriorated results of the experiments are however not yet satisfactory. More details in RC LACE stay report.

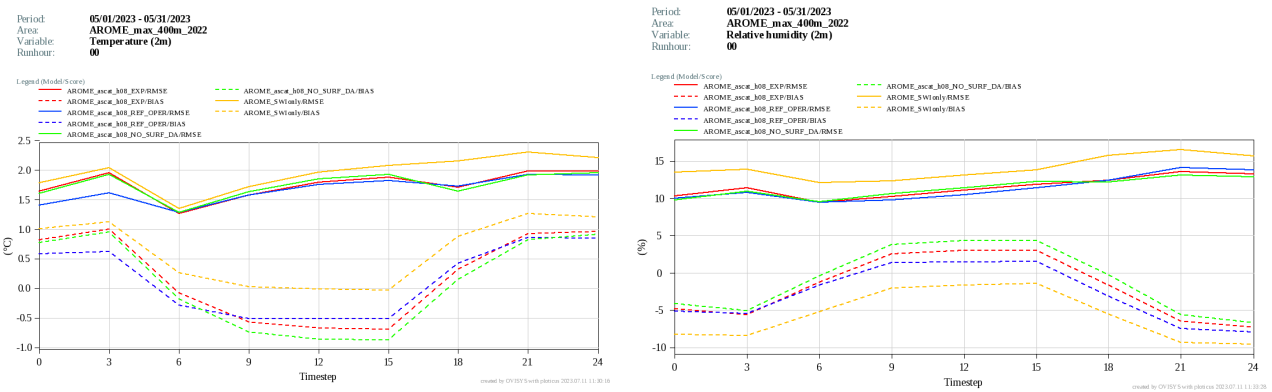

Figure 3.6.1: Bias (dashed) and RMSE (solid) of T2m (left) and RH2m (right) with respect to synop observations as a function of lead time for operational AROME reference (blue), case with no surface DA (green), experiment with satellite SSM DA (red for σ_{obs}=0.4, yellow for σ_{obs}=0.1) (courtesy by Matjaž Ličar)

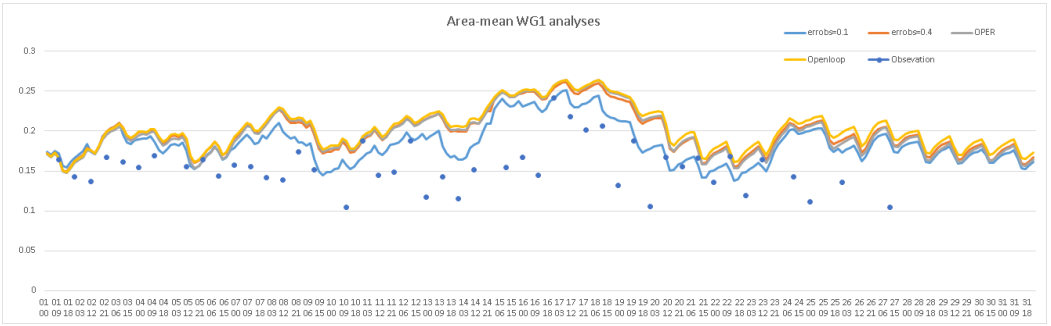

Figure 3.6.2: Analyses of surface soil moisture layer WG1 for SSM DA experiments (blue for σ_{obs}=0.1, red for σ_{obs}=0.4), experiment with no surface DA (yellow), operational reference experiment (grey) and satellite SSM observations (points). Averaged for 116 locations. (courtesy by Matjaž Ličar)

Progress with surface multi-layer physics

The replacement of the still operational Force-Restore system with the multi-layer physics package including 14-layer diffusion soil scheme, 12-layer explicit snow scheme and Multi-Energy Balance for the forest vegetation has continued during the year. However, the replacement of the ECMWF HPC from cca to Atos slowed down the HARMONIE-AROME progress considerably. Now we see more consortium-wide efforts towards multi-layer surface physics and the recent WW in Toulouse is a good example of how we tackle common problems together with very concrete steps forward.

Towards high-resolution and high-quality physiography for Europe

Land cover in ACCORD is operationally based on ECOCLIMAP 1st (1 km) and 2nd (ESA-CCI land cover 300 m) generations. The 2nd generation one (ECOSG) comes with parameter maps (LAI, albedo, …). As a complement to these, Met Éireann has continued the ML efforts to reach a higher resolution (60 m) more precise land cover map over Europe and a first prototype version now exists and is under testing. Below is an example of how the

current ECOSG looks in the surroundings of Dublin (left) compared to the new high-resolution version, ECO SG plus (right).

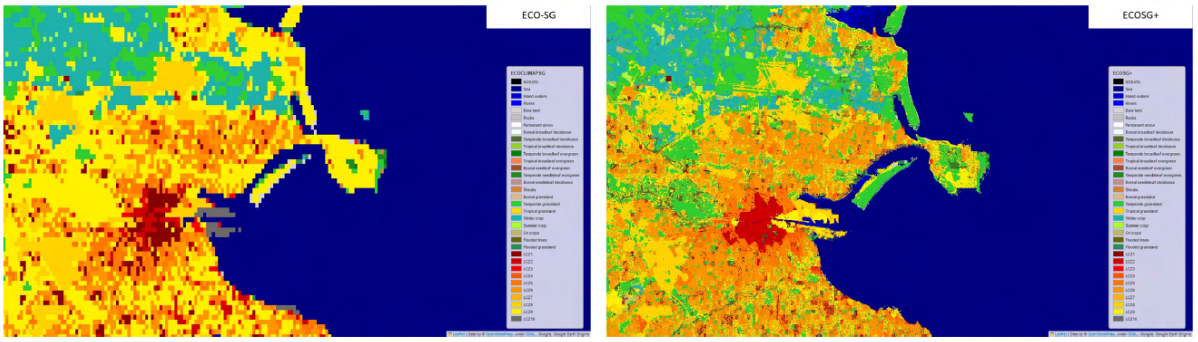

Figure 3.6.3: ECO SG (left) and ECO SG plus (right) - a target map, based on the decision tree. (Met Eireann)

A couple of important NWP contributions to SURFEX

A problem with excess spring snow melt in the 12-layer Explicit snow scheme was identified in HARMONIE-AROME validation experiments at high latitudes. The problem was confirmed by AROME-France simulations including the Alps region. Earlier the problem has also been identified in climate simulations, e.g. by the Arpege climate group. We have agreed that a fully proper solution would mean a separation of soil columns beneath snow and non-snow surfaces. However, before we find manpower into this effort a pragmatic solution has been implemented which handles the core of the problem and allows further experiments. It has been identified that turbulent exchange coefficients in the surface layer have been suffering from approximations which are valid for relatively high lowest model levels but which are not valid when the lowest model approaches close to the surface, on the order of a few metres. A more proper solution is now implemented.

- ***perspectives and priorities for 2024***

Perspectives with respect to the three main topics surface model, physiography and data assimilation:

Surface model:
Within LACE, the ALARO-CSC with SURFEX activities will focus on activation of some more advanced options interesting for NWP and climate applications (e.g., diffuse ISBA, ISBA-ES snow scheme, FLAKE model, etc.). Furthermore, debugging the TEB option (problem with snow) will be continued. The robustness of ALARO with SURFEX in climate mode is planned to be tested. Finally, first steps towards the operational implementation shall be made. AROME-based activities will include validation and further improvement of individual components (DIF, ES, MEB), and their combination, in the NWP system. The stable regime issues will be investigated, within both surface and boundary layer, aiming to improve the model performance.

HIRLAM will continue testing and development of ISBA-DIFF+ES+MEB+SEKF+pysurfex, also with separate components; action towards improved 2m diagnostics; stable boundary layer activities.

Physiography:
Getting towards exploitation of ECOCLIMAP-SG physiography data with ALARO-CMC in LACE. Continuation of work on development of fine scale cover maps over Europe using ML approach.

In HIRLAM the ambition is to further develop ML physiography and to continue activities on optimization of PGD.

Continued testing of ECOCLIMAP-SG in the AROME CSC.

Data assimilation:
LACE:

- Implementation and validation of OI analysis (SYNOP) in coupled ALARO-SURFEX, further tuning of soil moisture initialization.
- Assimilation of satellite based observations (SWI, LAI) within SEKF in AROME/SURFEX.
- Implementation and validation of snow analysis (SYNOP).

HIRLAM:

- Continue with snow DA testing.
- Continue with EnKF and ML for observation operators.
- Use uncertainty from SIC products for EPS.
- Use observations from SIMBA boys over lakes for verification.

Météo-France:

- Continued development and testing of the 2D-Variational approach in the OOPS framework.
- Continued work on assimilation of SEVIRI land surface temperature in AROME.
- Testing and further development of the diffusion soil version of the vertical OI soil assimilation.

3.7. Ensemble forecasting and predictability (E6 to E12)

- ***overview of the activity in the area***

The work in Ensemble forecasting and predictability can naturally be divided into three subjects,

- Calibration of ensemble forecasts and development of user-oriented products
- migration and preparation of ensemble prediction systems
- perturbations

which are further subdivided in the work packages E6-E12 (starting at E6 and not E1, because the tasks in the former CSC-dependent work packages E1-E5 have been included in the new work packages E8-E12).
Application of AI and machine learning is increasingly being investigated for **calibration** and, more generally, enhancement of ensembles. The hope is that machine learning methods might help overcome some of the limitations imposed by the computational cost of running ensemble prediction systems with our traditional NWP models.

The relatively limited use of EPS after its introduction in operations might not have matched the cost of running an operational EPS. One of the reasons being that many users find it complicated and time consuming to utilise the full potential of EPS, and consequently we **develop products** that support forecasts of high-impact weather as well as products tailored to activities that depend critically on different aspects of the weather, such as aviation, road weather, dispersion modelling and renewable energy.
Migration and preparation of EPS take up many resources, both human and computational. In 2022-23 migration from ECMWF's old Cray HPC to the new Atos has been a challenge, but has led to successful operationalisation of both the A-LAEF and C-LAEF ensemble systems (see, e.g., the presentations http://www.umr-cnrm.fr/accord/IMG/pdf/lace_eps_wastl_04_2023.pdf and http://www.umr-cnrm.fr/accord/IMG/pdf/a-laef_accord_asw_2023.pdf from the 3rd ACCORD All-Staff Workshop in Tallinn in March 2023). The successful upgrade should also be mentioned of AROME-EPS at The Hungarian Meteorological Service to use ensemble data assimilation for initial conditions (see http://www.umr-cnrm.fr/accord/IMG/pdf/accord_asw_mqa_boglarkatoth.pdf from the 3rd ACCORD All-Staff Workshop).
The work on **perturbations** has focused mostly on model perturbations, particularly stochastic parameter perturbations (SPP). Recently, stochastic perturbations of surface parameters and dynamics have been investigated (see highlights below), as has flow-dependent perturbations of microphysics parameters, using cloud fraction as perturbation weight.

- ***organisation of the work within the consortium***

An ACCORD EPS working week was held in Oslo in April. Work included further development of SPP, surface perturbations, including an update of soil moisture perturbations where the perturbations are constrained by the soil wetness index, dynamics perturbations and EnVar, see the wiki page https://opensource.umr-cnrm.fr/projects/accord/wiki/Oslo_24-28_April_2023 .
In June an online meeting on the use of machine learning in EPS was held. The two presentations included (1) a report from the ACCORD machine learning working group with focus on EPS and (2) ideas for machine learning enhanced ensembles, including use of generative ML models to increase the ensemble size conditioned on a conventional NWP ensemble, and use of ML to enhance SPP in Harmonie-Arome; for slides and a recording, see https://opensource.umr-cnrm.fr/projects/accord/wiki/Use_of_machine_learning_for_EPS .

- ***highlights from 2023***

The work on perturbations has been continued and extended to dynamics. SLWIND was added to the stochastically Perturbed Parameterizations (SPP) scheme. This option perturbs the midpoint wind along the trajectory in the semi-Lagrangian advection scheme used to compute a refined position for the origin point of the trajectory (only for the Coriolis term). The formula for the wind used is

$$V = 0.5*RW2TLFF*(V(F)+V(O)) + (1-RW2TLFF)*V(M)$$

where F is final, O is origin and M is midpoint along the trajectory. The perturbation option sets RW2TLFF=0.5 (as opposed to RW2TLFF=1 if unperturbed) and adds a random rotation with uniform distribution and zero mean angle to the V(M) wind. SLWIND improves the probabilistic skill of the forecasts when tested over two two-week periods, one in summer and one in winter. No bias-issues are seen from perturbing SLWIND in these experiments. Figure 3.7.1 shows the spread and skill for 12h accumulated precipitation for the summer period tested.

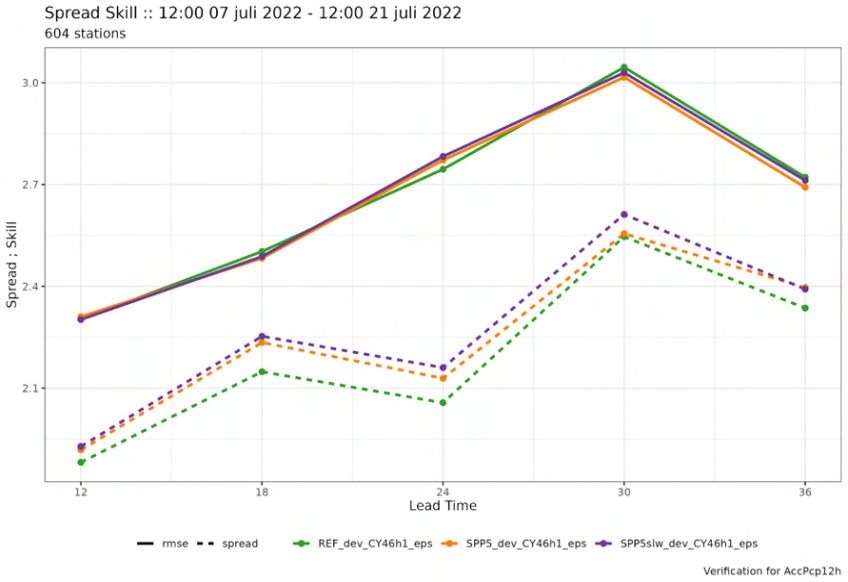

Fig. 3.7.1. Spread (dashed) and skill (solid) for 12h accumulated precipitation for 7-21 July 2022. In green is a reference run without SPP, in orange with SPP without SLWIND, in purple SPP with SLWIND (courtesy of Inger-Lise Frogner).

More technical improvements to SPP include the possibility of namelist driven pattern correlations and the possibility to set individual length and time scales for the parameters. Another important aspect is bringing HARMONIE-AROME SPP (and surface perturbation changes) into cy49t1. A problem with running SPP in single precision has been solved. As the patterns evolved with time in single and double precision, they diverged. The changes amounts to keeping spectral fields (in SPECTRAL_ARP types only) in double precision, not only in the stochastic pattern generator solver (which was already done), but also to avoid copying single → double → single precision between calls to the solver.

On the evolution towards new configurations we note that a C-LAEF e-suite has been running in 1 km horizontal resolution since May 2023 on the ECMWF HPC. The e-suite includes 16+1 members running forecasts in single precision; it uses cy46t1 with 3h assimilation cycle, I/O server and SPP perturbations. The control member uses EnVar assimilation (based on 51 members). Verification scores look good (see Fig. 3.7.2 below) – for many parameters better than the operational C-LAEF system. The C-LAEF 1km data is shared between Austrian and Slovenian colleagues, and operationalization is planned for the end of 2024.

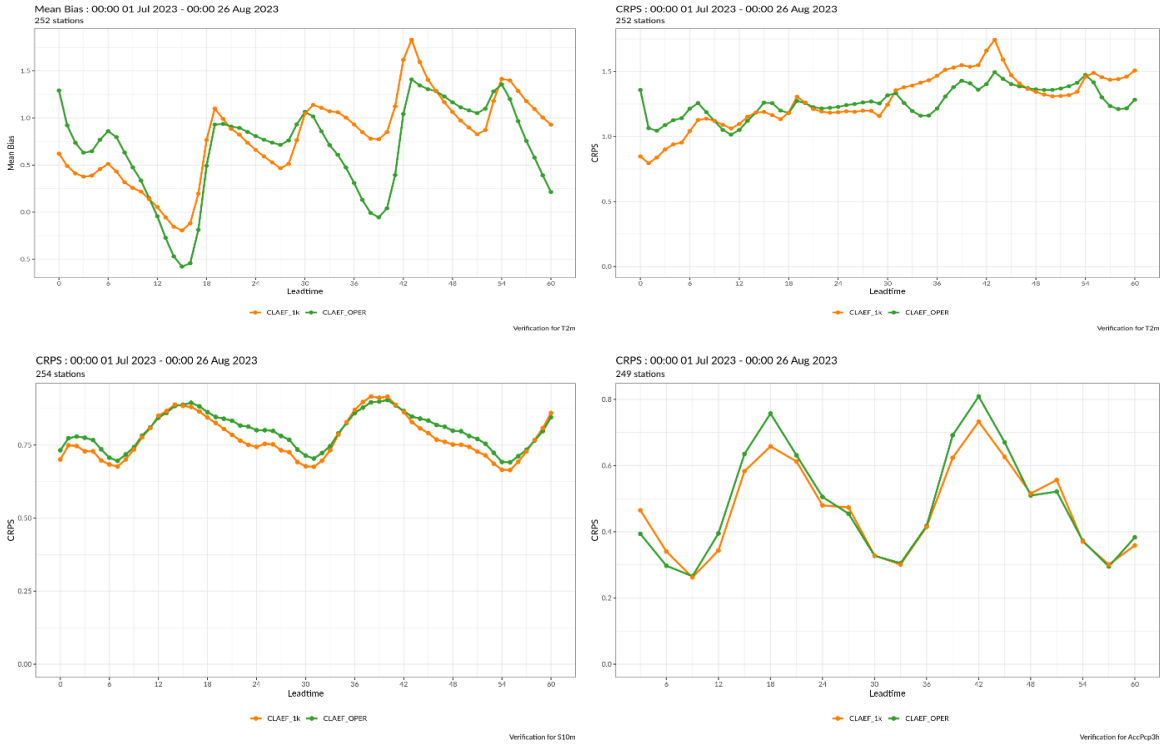

Fig. 3.7.2. Mean bias (upper left) and CRPS (upper right) of 2m temperature; CRPS of 10m wind speed (lower left) and 3h accumulated precipitation (lower right) of C-LAEF OPER (green) and C-LAEF 1km (orange) for the period July-August 2023 (courtesy of Clemens Wastl).

The latest Météo-France Arome-EPS e-suite is based on cy48t1 and integrates coupling to Arome 3DEnVar analysis, upgrades in Arome physics, an increase of ensemble size from 17 to 25 members (including the control member) and forecasts run in single precision. The performance of this e-suite is overall positive, with a noticeable impact from the 3DEnVar at early ranges. The operational declaration is planned early 2024.

AI/Machine Learning has been applied to summarise Arome-EPS precipitation forecasts in a reduced number of representative scenarii. Given a set of typical precipitation patterns (equivalent to weather regimes), each member is attributed to the closest pattern, and all members assigned to the same pattern form a scenario. To improve computational efficiency and avoid drawbacks of gridpoint metrics (double penalty), the approach is entirely performed in the latent space of a convolutional auto-encoder. A 4-year Arome-EPS archive has been used for training, and the method has been objectively evaluated over one year. An experimental production of this clustering has been set up in June for subjective evaluation by Météo-France forecasters; see example below in Fig. 3.7.3.

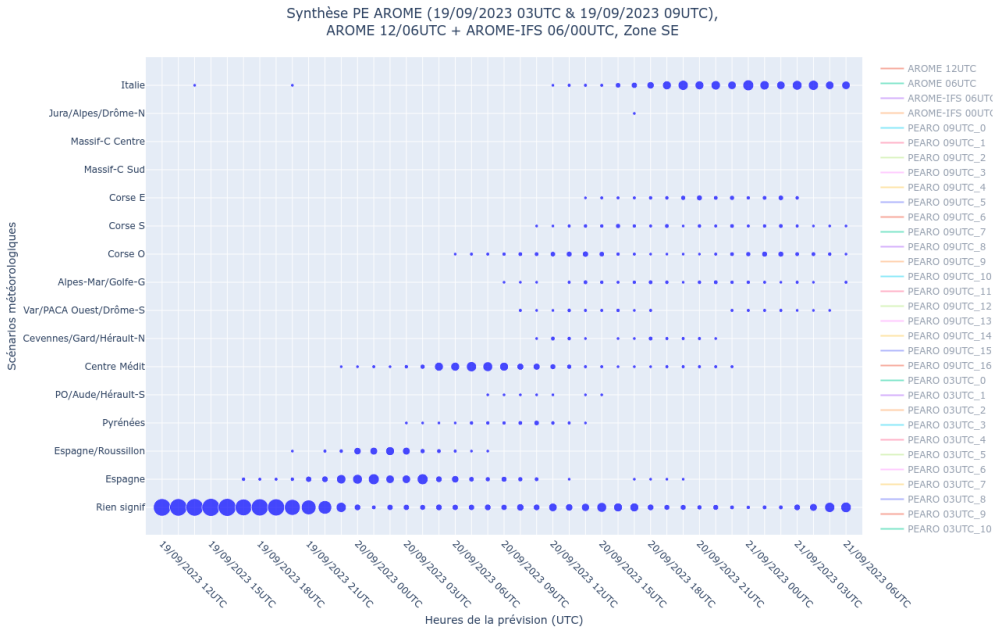

Fig. 3.7.3. Clustering visualisation product. Typical precipitation patterns (~ regimes) are displayed along the ordinate axis, lead times are in abscissa. The size of blue dots is proportional to the number of members assigned to each pattern. Forecasts of individual members inside each cluster can be visualised by clicking on the blue dots (courtesy of Laure Raynaud).

Another example of the use of AI is the development toward large NWP-AI ensembles. By leveraging the capability of generative AI to produce realistic images, a generator of Arome-EPS forecasts has been developed based on a state-of-the-art GAN architecture. Generated samples of 2-m temperature and 10-m wind are physically consistent and are mostly indistinguishable from true Arome forecasts. A ‘guided’ generation, conditioned on available Arome-EPS members, has then been proposed so that the generator can be used to enrich current ensembles at a lower computational cost. AI-enhanced ensembles are shown to improve most probabilistic scores (see Fig. 3.7.4), and compare well with a very large Arome ensemble. Extension to the generation of precipitation forecasts is ongoing.

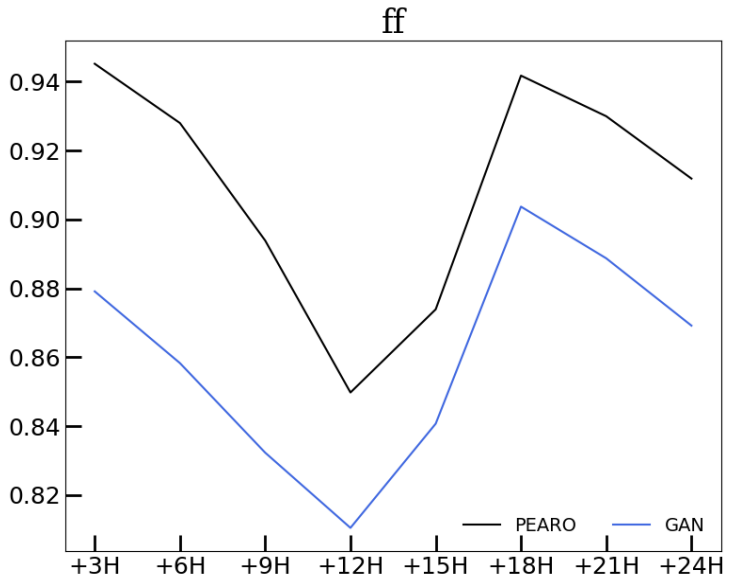

Fig. 3.7.4. CRPS of 10m wind speed as a function of lead time. Operational Arome-EPS (16 members, black) and AI-enhanced ensemble (120 members, blue; courtesy of Laure Raynaud).

The recently upgraded common LACE EPS, A-LAEF based on Alaro (see https://www.ecmwf.int/en/elibrary/81316-mesoscale-ensemble-prediction-system-laef) has generally performed well. Compared to global models it clearly shows the benefit of a limited-area ensemble system in higher resolution. An example is the flood event in Italy on 15-16 May 2023 where floods and landslides in the Emilia-Romagna region in Italy caused eight deaths and forced thousands to evacuate their homes. Figure 3.7.5 below shows A-LAEF precipitation forecasts for the event.

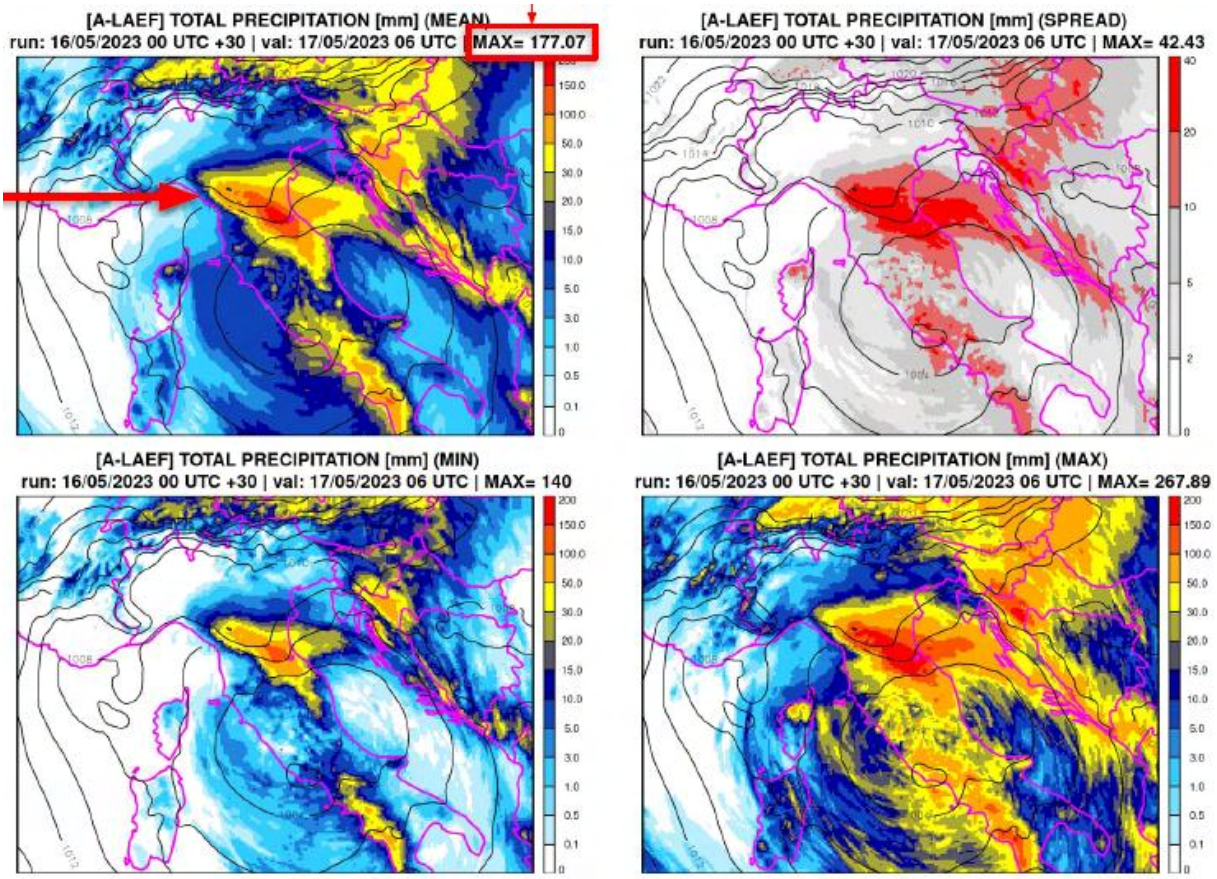

Fig. 3.7.5. 24h accumulated precipitation from 06UTC on 16 May to 06UTC on 17 May 2023. Ensemble mean (upper left), ensemble spread (upper right), ensemble minimum (lower left) and maximum (lower right) based on A-LAEF 00UTC on 16 May (courtesy of Martin Bellus).

- ***perspectives and priorities for 2024***
    - Work will continue on stochastic parameter perturbations (SPP), including more upper-air parameters, surface parameters and dynamics perturbations.
    - The URANIE platform will be used more to test parameter sensitivity prior to inclusion in SPP.
    - Use of flow-dependent SPP will be explored further.
    - Investigation of the use of AI/ML for generation and postprocessing of ensembles is expected to increase throughout the consortium.
    - If the legal constraints are resolved, a common script system for EPS research and development based on the script system developed for DestinE/DEODE will be developed.
    - New developments will be merged into the common cy49t1.

3.8. Meteorological quality assurance and verification (MQA1 to MQA3)

- ***overview of the activity in the area***

Activities of the MQA area include research and development of methods and code for forecast verification and model diagnosis, as well as application of such methods for assessing the meteorological quality of operational forecasts and new developments, and facilitating communication between users and developers of ACCORD forecasting systems. Probabilistic and deterministic forecasting at high or very high resolution, in particular related to clouds and precipitation, are at the focus, together with forecasting extreme events in general. Facilitating the efficient use of currently under-used data such as measurements from space borne platforms, surface radars, surface radiation networks, citizen networks, et c., is arguably the most urgent task within the area.

The activities of the area fall into three work packages, focusing on:

1. the common verification system "harp" (MQA1)
2. new metrics and methods for verification and quality assurance (MQA2)
3. forecast verification, system evaluation, and error attribution (MQA3)

Out of the 17 FTEs contributed to the area during the first half of the year, these packages have received 19, 16, and 65 %, respectively, as shown in Fig. 3.8.1.

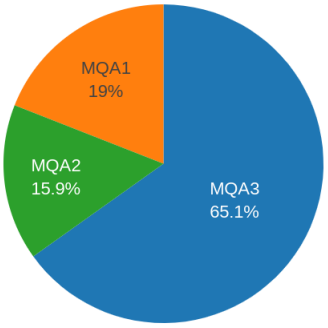

Fig. 3.8.1 Division of contributed labour among the work packages of MQA. First half of 2023.

Many activities within the area are intimately linked to other areas of the RWP. Thus, identifying and investigating systematic errors connects to activities of **upper air physics and surface interaction,** while development and application of new methods and metrics for verification naturally links with **ensemble prediction and sub km-scale modelling**. Application of methods applied to **data assimilation** opens vast opportunities for enhancing the data pool that can be used for verification. Methods employing **machine learning and artificial intelligence** have been applied for quality control of observations and in the context of verifying forecasts of convective storms by means of cloud tracking. Significant development of tools and methods for meteorological quality assurance making use of the common verification environment "harp" have taken place in the frame of **Destination Earth** work package DT_extremes.

- ***organisation of the work within the consortium***

Local teams are the principal unit for monitoring and evaluating operational installations and e-suites. Concerted trials of new cycles, however, are organised at the level of CSCs, as well. At the level of the consortium, a working week and a scientific visit, both focussing on the use of spatial methods and supporting the use of satellite data and radar data in the frame of harp, are features of the current year, and a series of bi-monthly harp community meetings have been initiated for the purpose of organising collaboration and exchanging of experiences.

At the level of the project management work has started to organise the interaction between users and developers of ACCORD systems as outlined in the "R2O white paper". A body of user representatives, tasked with the collection and reporting of user feedback, and conveying the response of ACCORD back to colleagues at their institutes, is being formed.

Efforts and achievements within the area are monitored in a series of annual progress reports, collecting contributions from all staff members twice in a year. The report is available on the ACCORD wiki:
https://opensource.umr-cnrm.fr/projects/accord/wiki/Meteorological_Quality_Assurance

- ***highlights from 2023***

Out of the progress in the area, two examples are singled out for special mention here, one comes from the domain of developing verification methods, and outlines recent progress in developing harp, while the other one describes a study investigating the consequences of different physics options related to clouds.

Harp is increasingly being used for verification of operational forecasts and experiments. Teams applying harp at the moment or in the near future include: UWC-W and MetCoOp, Algeria, Austria, Belgium, Croatia, Hungary, Poland, and Turkey. Applications rely on local developments for configuration and control of data flow and verification tasks. Scripting based on or similar to the common *accord-verif-scripts* are in use within UWC-W, Algeria, and Croatia.

Harp version 0.2.0 is being tagged this autumn under the leadership of **Andrew Singleton (Met Norway)**, involving a large refactoring of the harp code and re-organization of the harp packages, and merging of updates from the harpSpatialTeam development branch (**Alex Deckmyn, RMI, and the harpSpatialTeam**). The new version makes the handling of multidimensional data and the writing of methods for "geolists" much more straightforward. Additionally C++ functions have been written for the fast computation of a number of summary statistics of geolits, and the reading of grib data has been improved, enabling a distinction between instantaneous and accumulated parameters with the same shortName. The computation of the fractions skill score for ensembles has been added to *harpSpatial* with the new function *ens_fss()*. Some adaptation to currently used R-scripts will be necessary when updating to the new version.

A **working week on spatial verification in harp** was held at DMI on 20-24 February. Fortelius, ACCORD NL4). Out of space born data sources, brightness temperature from MSG SEVIRI are used in Austria and in the frame of UWC-W. Cloud water path from MSG SEVIRI, and scatterometer winds from HY-2B/2C 25 km are likewise used within UWC-W, while the NASA-IMERG precipitation retrievals are used in Spain. Precipitation analyses involving ground based radars have been used in Norway and Denmark. A series of regular

meetings was initiated at the working week, aiming at coordinating the acquisition and use of data supporting spatial methods in harp.

Developments in the frame of **DestinE DT_extremes** include implementation of the "Panelification tool" in harp (see below), as well as a functionality to generate sqlite-files out of grib-files, thus eliminating the need for intermediate vobs-files. Additionally, code for the purpose of managing standardised verification tasks within the project have been developed in a branch forked out of the common *accord-verif-scripts* repository (**Samuel Viana, AEMET**). All of these developments need to be merged into the main branch as soon as possible.

One of the metrics of spatial verification is fraction skill score (FSS). It is a metric done for a range of thresholds and window sizes. Comparing a range of models or model versions with/without developments can be overwhelming. To tackle this issue, the so-called **panelification tool** has been developed at GeoSphere Austria. It was and continuously is developed in python in the first place. The panelification tool allows a quick understanding which model (version) performed best, displaying the ranking of a range of point verification scores and FSS for several thresholds and percentiles (see Figure 3.8.2). A description of which information a panel contains is displayed in Figure 3.8.3.

So far, this tool was not available in harp. But now the functionality of the panelification tool has been made available within harp **(Polly Schmederer, GeoSphere Austria)**. So far the changes are still locally at GeoSphere. They are going to be shared on harphub.

In order to make this functionality work within the harp framework the following steps have been taken:

- The development version of harpSpatial has been pulled.
- scores were added to the list of point based scores within harp spatial: mean absolute error (MAE), root mean squared error (RMSE), Rpearson.
- FSS using percentiles as thresholds was added to the spatial score. Additional to the standard FSS using precipitation amounts as thresholds.
- the possibility of separating re-gridding (and plotting) domain and verification (sub-) domain,
- also returning the re-gridded fields for plotting and not just the scores.
- ranking of the models for each score.
- plotting of the observation and forecast fields along with the summary scores and their ranks.

An example of how a plot done with the harp panelification tool looks is displayed in *Figure 3.8.2.* below. *Figure 3.8.3.* describes the information contained in one panel.

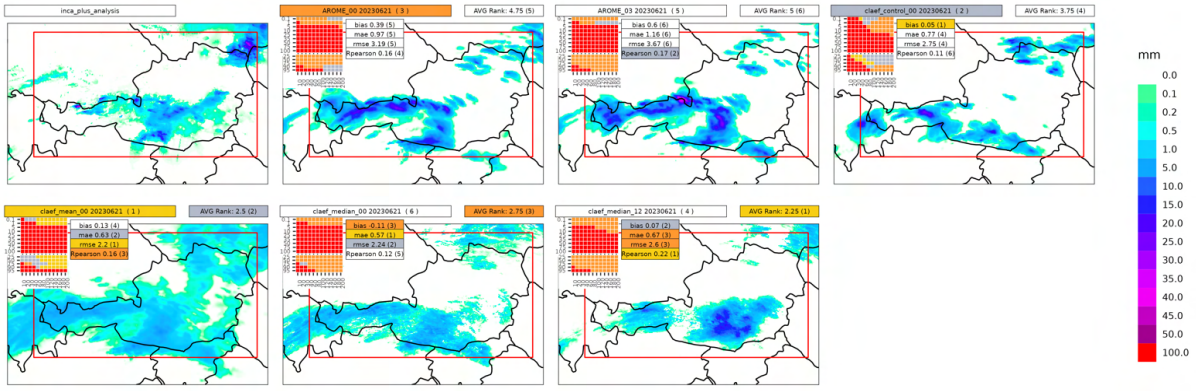

Figure 3.8.2: Example of a panelification plot done with harp showing 1h accumulated precipitation at verification time 21.06.2023 17 UTC. The models AROME (00 UTC run), AROME (03 UTC run), C-LAEF control (00 UTC run), claef mean (00 UTC run), C -LAEF median (00 UTC run) and C-LAEF median (12 UTC run) (beginning at the 2. from top left) are verified against INCA plus analysis (top left). (courtesy by Polly Schmederer - GeoSpere Austria)

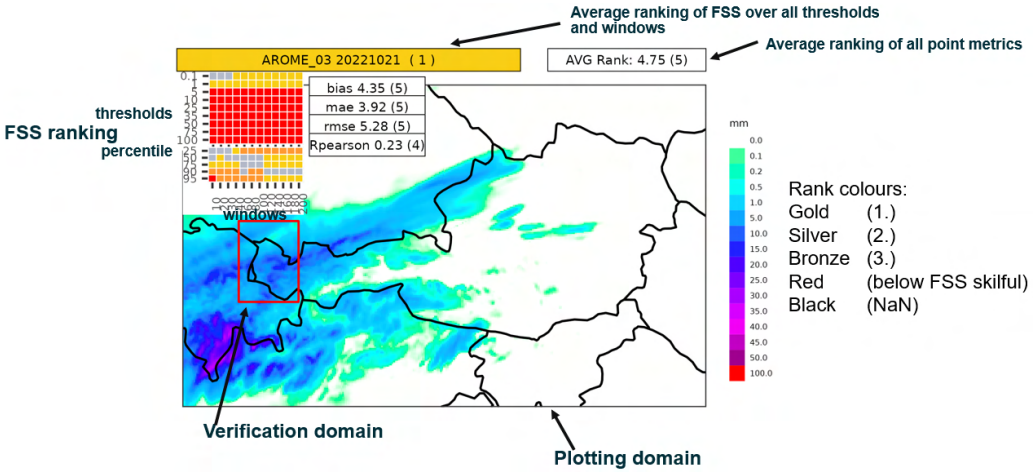

Figure 3.8.3: Description of information displayed in one panel. (courtesy by Polly Schmederer - GeoSpere Austria)

A new scheme for subgrid variance of the saturation departure has been tested in AROME/HU. In AROME there are two options available to compute the subgrid variance of the saturation departure. The older version (LOSIGMAS=.F.) computes this variance diagnostically, while the new version (LOSIGMAS=.T.) computes this prognostically using the Turbulent Kinetic Energy scheme. The two schemes produce a significantly different cloud cover: the diagnostic formulation simulates smooth cloud fields while the prognostic formulation tends to produce a cloud cover value with two categories ("zero or one"), with lower average values. The operational 2.5 km AROME version of OMSZ applies the diagnostic formulation. Forecasters gave feedback that in some cases during winter 2022/2023 AROME/HU significantly overestimated the cloud cover and consequently underestimated 2 metre temperature. Modification of LOSIGMAS from .F. to .T. for these cases improved forecasts of cloud cover and temperature. After these successful case studies two longer periods were tested by **Balázs Szintai (OMSZ)**, one in summer (2022-08-01–2022-08-31) and one in winter (2022-12-17–2023-01-17).

For all the experiments the tuning parameter VSIGQSAT (regulating the width of the PDF) was set to the default value of 0.02. For the summer period, additionally, SAL verification was performed for precipitation using radar data. The impact of the modification varies for different variables. For cloud cover and surface incoming radiation a significant reduction of the positive bias can be observed, however, RMSE scores deteriorate (Fig. 3.8.4), which could be attributed to the more discrete cloud cover field. For convective precipitation, a deterioration can be observed as the new version simulates too strong convective cells (Fig. 3.8.5). For other variables the modification results in mostly neutral impact. Sensitivity experiments are planned with larger values of VSIGQSAT (e.g. 0.06). (0.5 PM)

Further studies have been conducted to test the impact of the different settings of the subgrid variance of the saturation departure. Sensitivity of the forecasts towards the VSIGQSAT parameter has been investigated, and the two time periods (summer and winter) used in previous studies have been re-run with [VSIGQSAT=0.06] setting. This resulted in slightly improved scores for most parameters, however, incoming solar radiation significantly degraded, especially in winter (Fig. 3.8.6). Consequently, the decision was taken that further tests will be conducted with the [VSIGQSAT=0.02] setting. As the previous summer period (August 2022) was very dry, another period (July 2021) was chosen to investigate the impact on summer convection. For 2 metre temperature a slight improvement in RMSE during daytime can be noticed (Fig. 3.8.7). For convective precipitation a slight degradation was observed. For other parameters the modification had mostly neutral impact. After careful analysis of the results the decision was taken that the [LOSIGMAS=T, VSIGQSAT=0.02] is worth to be tested on a longer summer period as a parallel suite with involvement of forecasters in the model evaluation on a daily basis. The parallel suite was started at the end of June and will be run until mid September. (0.5 PM)

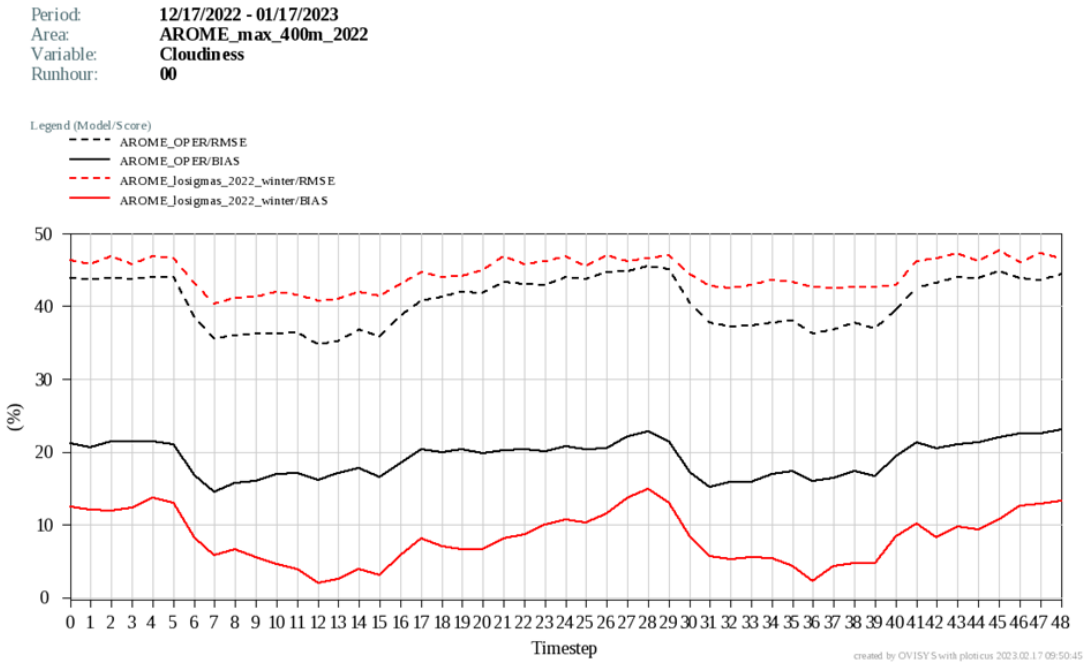

Figure 3.8.4: Bias (solid line) and RMSE (dashed line) of AROME forecasts as a function of lead time from 17 December 2022 to 17 January 2023. Black: operational AROME/HU (using LOSIGMAS=.F.); red: experiment using LOSIGMAS=.T. (courtesy by Balázs Szintai - OMSZ)

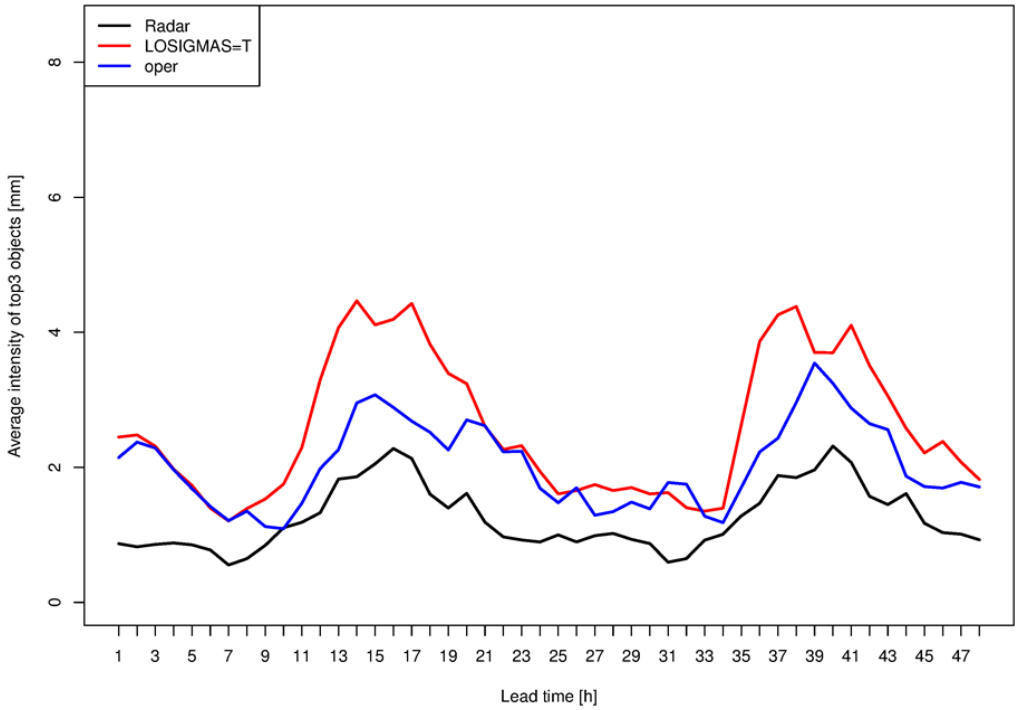

Figure 3.8.5: Average intensity of the three strongest convective precipitation objects (based on SAL object definition) as a function of lead time for August 2022. Blue: operational AROME/HU (using LOSIGMAS=.F.); red: experiment using LOSIGMAS=.T.; black: radar data (courtesy by Balázs Szintai - OMSZ)

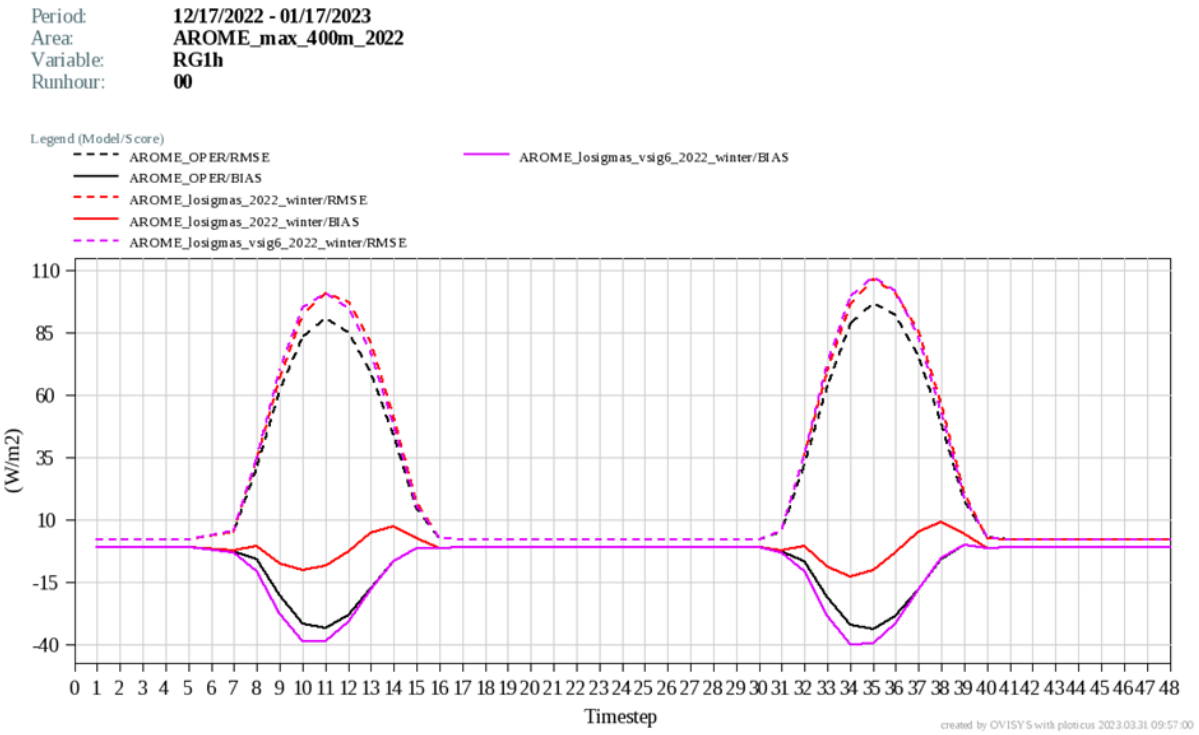

Figure 3.8.6: Bias (solid line) and RMSE (dashed line) of AROME forecasts for incoming solar radiation as a function of lead time from 17 December 2022 to 17 January 2023. Black: operational AROME/HU (using LOSIGMAS=.F.); red: experiment using [LOSIGMAS=.T., VSIGQSAT=0.02]; purple: experiment using [LOSIGMAS=.T., VSIGQSAT=0.06]. (courtesy by Balázs Szintai - OMSZ)

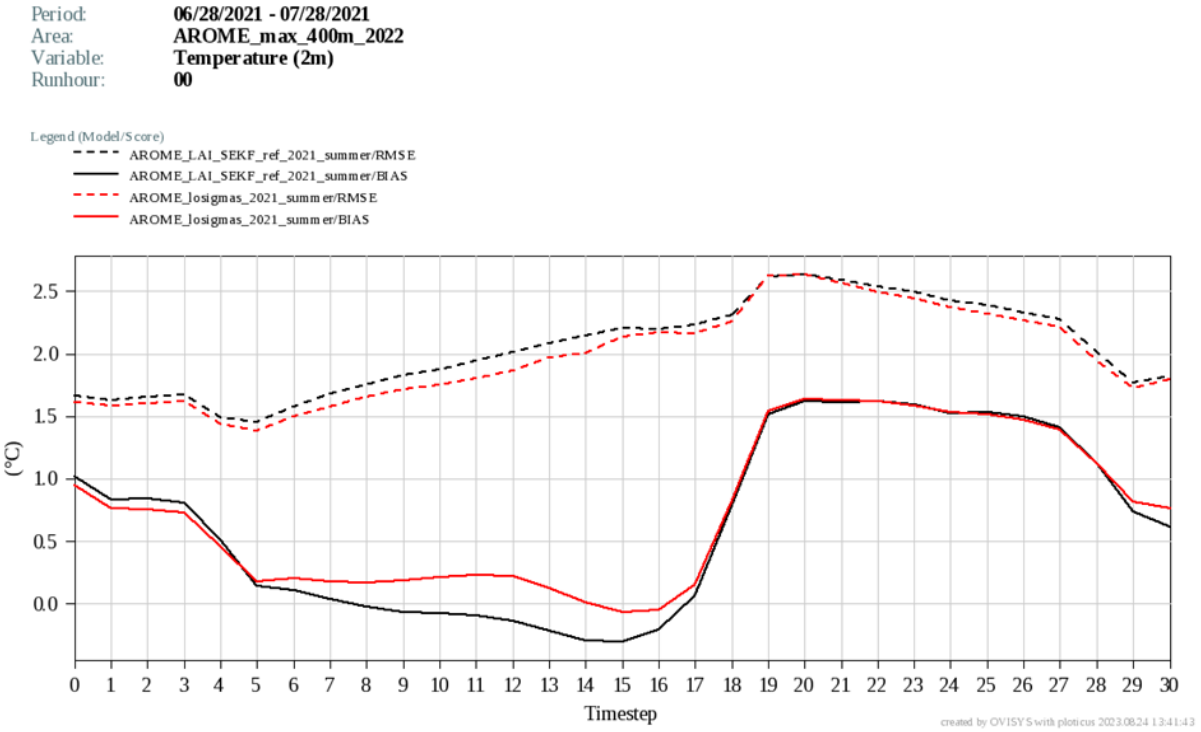

Figure 3.8.7: Bias (solid line) and RMSE (dashed line) of AROME forecasts for 2 metre temperature as a function of lead time from 28 June to 28 July 2021. Black: operational AROME/HU (using LOSIGMAS=.F.); red: experiment using [LOSIGMAS=.T., VSIGQSAT=0.02]. (courtesy by Balázs Szintai - OMSZ)

- ***Perspectives and priorities for 2024***

Efforts in support of assessing forecasts down to regional or local scales will remain a first priority within the area in 2024, with facilitating the acquisition and use of spatially-distributed fine scale data from space borne instruments as a central goal. Actions need to be initiated in preparation of merging locally-implemented solutions into a common infrastructure for verification, featuring paths to observational data and scripting to control a variety of verification tasks. In parallel, standardised procedures for ensuring the meteorological quality and robustness of new components and export versions of CSCs need to be worked out and implemented.

Implementing the consortium-wide collection and treatment of user feedback as outlined in the "R2O white paper" will call for substantial efforts of MG and experts during the year. Collection of feedback from the body of user representatives has to be initiated, and practices for dealing with the collected feedback worked out and put in place.

Along with the increase of model resolution, comes the need to evaluate forecasts specific to built-up regions. It can be foreseen that data from citizen networks and other unconventional, or previously unused, sources will be of major importance, underpinning the need to develop methods for acquisition and quality control of such observations. In the medium term, new approaches to verification, recognizing the enormous heterogeneity of the urban landscape, need to be developed.

Methods of machine learning and artificial intelligence are currently applied to cloud tracking and to quality control of observations. In addition to these lines of development, one might

consider applying machine learning to the verification problem itself. Can deep learning algorithms add value to pure statistical methods, when it comes to determining how “good” is a given forecast? Or, do we need to develop new metrics, such as measures of physical consistency, in order to evaluate forecasts provided by data driven engines?

3.9. Technical code and system development (SY1 to SY4)

- ***overview of the activity in the area***

This area reflects the efforts divided in four work packages covering the Code Optimization efforts (**SY1**), the Maintenance and development of the Harmonie Reference System (**SY2**), the Revision of the Harmonie scripting systems (**SY3**), and the activities Towards a more common working environment (**SY4**).

The code optimization activities (**SY1**) have the main objective to identify and overcome bottlenecks for code computational performance. To discover these blocking points in the code performance, and in its scalability, code profiling and benchmarking techniques have been used.

Due to code refactoring tasks (**SPTR1**) and future machine learning solutions, the approach of analysing the performance of code on traditional architectures is taking less prominence and fewer resources in recent years. Hybrid, coprocessor-based systems offer clear improvements in economic and energy costs. In the case of alternative data driven models, they also take advantage of this hybridisation of HPC, being tremendously efficient in reducing computational cost. A significant work was realized by MF to assess the portability and the performance of refactored code components on a vector accelerator HPC (at CHMI).

The Arome and Harmonie-Arome CSCs have the capability to compile and execute the codes in single precision (SP), double precision (DP) and dual precision (2 binaries one single and other with double precision are compiled and at the same time, allowing use of them in a mixed precision simulation suite, i.e. the scripting allows to select which binary that is used in certain process in the model suite). The use of SP computations is currently explored for the operational EPS suites of some ACCORD partners showing a 30-40% gain in computational speed.

Examples of the use of SP in pre-operational environments are the UWC-W EPS at 2 km horizontal resolution and 90 vertical levels and the C-LAEF EPS of RC-LACE running on ECMWF Atos-Machine. MF are testing in e-suite configuration the Arome forecasts in SP (deterministic and EPS, based on CY48T1). The URANIE sensitivity and uncertainty analysis platform is also being used to systematically assess the impact of SP,

The implementation of new more costly algorithmics in the modelling systems created the necessity to profile, analyse and optimise them for operational usage. This is the case of 4DVar data assimilation technique, that Harmonie-Arome has implemented as a feature in harmonie-43h2 for the first time, improved in harmonie-46h1 version and also has been integrated in CY48T1 common code.

Another important activity related to the usability of the code is the use of software containers to facilitate the deployment of the NWP code in different computational architectures. For Harmonie-Arome, the containerized version allows to compile the code using GCC+OpenMPI and including ecFlow. In addition, the MUSC single column model has been containerized also, both of them for CentOS 8, and are available in https://github.com/Hirlam/HarmonieContainers.

The maintenance and development of the Harmonie Reference System (**SY2**) required a substantial effort from the HIRLAM System Group. During 2023, the new Harmonie-Arome release candidate will be based on CY46T1 and it is being developed as dev-CY46h1. This version has been regularly merged with the Harmonie cy43h2 codes to maintain the same level of development of cy43 both in scripts and codes. The meteorological quality assurance for releasing harmonie-46h1.1 reference version is under test and will include not only deterministic validation but also EPS over the three areas covered by MetCoOp, UWC-West and AEMET.

Since 2021, the HIRLAM community began prototyping the use of GitHub as a source code management (SCM) tool for all public and private code. Some experience has been gained in this cloud-based SCM, allowing the establishment of new working practices that clearly improve the quality and management of software developments. Using the pull request mechanism and GitHub actions allows the testing of the code contributions before being integrated in the repos. In the process of maintaining code documentation as close as possible with each code version a migration of code documentation from the hirlam.org wiki to GitHub and it is accessible in https://hirlam.github.io/HarmonieSystemDocumentation/dev/.

Continuing the training on git, GitHub GUI and the new associated workflows the HIRLAM System Team has prepared some talks in the different working weeks that have taken place during 2023.

Discussions between the HIRLAM System Team, the ACCORD SYS/AL and the ACCORD Integration Leader did take place in order to exchange experience and good practice when using the *github.com* environment for collaborative work on the codes. During the integration phase of CY49T1, the ACCORD source code forge environment based on github has been used as the main tool for integrating all partners' code contributions. In addition, several webinars have been organized at ACCORD level in order to explain to the teams how to use this new environment.

The Harmonie Scripting System Review (**SY3**) has some stand-by tasks pending from the decision of a possible ACCORD common scripting solution and the separation between common codes and scripting following the multi-repository philosophy. Despite this, a cleanup of the Harmonie scripts has been done, especially in the context of the OOPS implementation in harmonie-46h1 and a preliminary under development feature that will allow sub-hourly cycling.

Also, for compiling OOPS code a new compilation strategy, cmake has been tested and will be part of the harmonie-46h1 and future releases. Using this more standardised tool for general code building, with broad community support, allows us to reduce circular code dependencies, faster and parallel compilation, and better handling of code interfaces.

One of ACCORD's goals is to increase collaboration between partners. This has been described in three work packages **COM2.1** (Code generation and maintenance: regular maintenance and evolutions, official releases), **COM2.T** Code generation and maintenance: Transition to new work practices and environment and **SY4** (Towards a more common working environment: explore practical choices, prototyping, scripting). The COM activities will be described in more detail in the next section Organization of work within the consortium. In connection to these plans, the testing tool “DAVAÏ” has been further developed and is now accessible on the ECMWF Atos machine (see in next section for the details).

Thanks to the DEODE project a discussion about the minimum requirements and functionalities that common scripting and DEODE model engines have been initiated. The primary goal is to deliver a system for the DestinE engine workflow with a clear ambition that this can be of further usage within ACCORD.

- ***organisation of the work within the consortium***

In the work package **COM2.1** (Code generation and maintenance: regular maintenance and evolutions, official releases), the progress on the coordination aspects and the work environment for planning timing and content of the code cycles, allowed the building of a CY49 common cycle in March 2023. In this year, the implementation of new work practices for the development of cycles, using systematic testing of contributions (Davai) and the use of the ACCORD forge on github for contributions to CY49T1. The use of the forge has led to a significant improvement of the workflow: easier to publish changes to the code via git commands from anywhere, making them public and visible (within the consortium), activation of revisions, better validated branches etc. As a result of this improvement more than 100 contributions from all families have been managed to CY49T1, using the ACCORD forge and Davai to assess their impact.
The workflow will be re-evaluated for future cycles taking into account the various needs of developers and integrators, as well as new tools to improve it.

The technical validation of codes is currently based on a set of unit tests implemented as a tool called DAVAÏ. As a reminder, the DAVAÏ test system allows any version of the code to be tested (new development or the result of a merger, for example). The steps are as follows: obtaining the codes to be tested, creation of executables, execution of sets of integrated or elementary test cases (ideally representative of canonical configurations, such as IFS, ARPEGE, AROME, ALARO, HARMONIE-AROME), automatic comparison of the results with the reference results, and simple visualisation of these results on the CIBULAI web platform. This validation is a crucial step in the process of integration and validation of code changes. For this reason, the test system has been maintained, adapting the test bench to the evolutions of the CSCs, and introducing new tests. To facilitate this maintenance, an interface

is available on the Meteo France and ECMWF platforms that allows users to run the DAVAÏ tool independently of the Olive scripts. This will allow the implementation of tests for different CSCs and a more complete set of tests to ensure the quality of the code to be developed.

For this purpose, a DAVAÏ Contributors-Developers Working Week was organised in November 2022 and a second edition will take place in October 2023.
To ensure the DAVAÏ implementation capability, common computational resources are available thanks to an ECMWF SPFRACCO Special Project dedicated to the technical validation (testing) of the contributions to a T-cycle.

- ***highlights from 2023***

During 2023, there are several achievements that will help us move forward in developing more reliable and efficient NWP codes for use in ACCORD Member Institutes. The reduction of computational costs, through the use of single precision calculations, the implementation of new algorithms such as 4DVar and OOPS/EnVar, the use of the ACCORD forge on GitHub and a more portable and stable DAVAÏ code testing tool on the Atos HPC ECMWF, are improvements that will lead to a more productive code integration process.

Also, as part of this improvement, knowledge transfer from the forge and DAVAÏ has been ensured by conducting tutorials.

This more collaborative environment will contribute to the increasingly efficient integration of the latest scientific advances and code solutions in the upcoming cycles CY49T0, CY49T1, and beyond including cycles common with ECMWF.

During the final part of the year the interaction with DE_330 (informally known as DEODE) will continue to ensure the use of the scripting developed following the Intellectual Property Rights (IPR) agreements signed with the ECMWF and the European Commission.

- ***perspectives and priorities for 2024***

The establishment of the ACCORD forge was the first step in opening up the possibility to collaborate on other parts of the modelling chain such as the scripting systems although a proactive role with the DEODE project could help in the possibility or reuse of the modelling engine as modular components of a future common solution. The new working practices introduced by the use of the forge will facilitate the construction of the common CY50 cycle and the CY50T1 version.

Regarding code testing, the portability of DAVAÏ to ECMWF HPC and other supercomputers will allow to increase code testing capabilities by introducing tests covering different domains such as CSC configurations, single precision computations and other compilation contexts such as clustering tools and compilation systems. In this context, DAVAÏ for workstation will be ported to perform toy-tests.

The Hirlam family is organizing a dedicated code phasing effort for CY49T1 with the aim of pushing their development branches from the Hirlam/Harmonie-Arome CSC repository fully toward the ACCORD repo. This effort is in association with the definition of a tiger team

aimed to promote the ACCORD working practices among scientists and find support for system activities in the scientists' area of expertise.

A key question is how to define a common ACCORD effort on the validation of a new version of code at a higher level of component integration than that realised in 2023. A working group will be established to discuss the definition of the export concept for all CSCs and to examine the technical requirements.

- ***Overview of timing of cycles and integration process***

The timeline below illustrates the construction process of the recent cycles. This table is regularly shared with the teams (at the ASW, in LTM meetings). Its detailed timing and content is discussed in the IFS-Arpege coordination meetings (all contributions) and in the ACCORD/MG meetings (regarding the LAM code contributions).

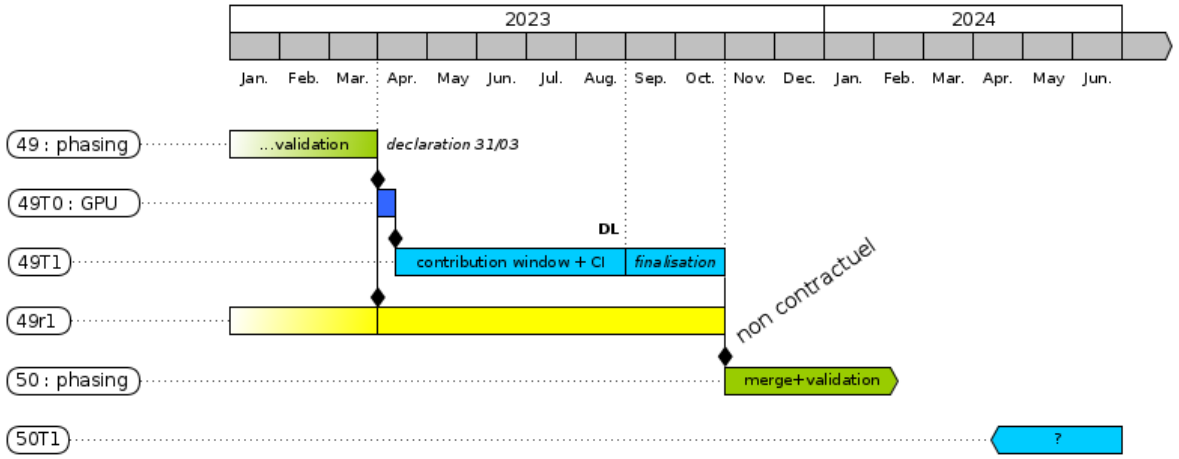

Figure 3.9.1. Timetable of the integration process and cycles over 2023-2024. (courtesy by A.Mary)