SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

Reporting year	Jan 2021 – June 2021
Project Title:	HIRLAM-C 3d phase (2021-2022) Special Project
Computer Project Account:	spsehlam
Principal Investigator(s):	J. Onvlee
Affiliation:	KNMI
Name of ECMWF scientist(s) collaborating to the project (if applicable)	…………………………………………………………… ……………………………………………………………
Start date of the project:	1 January 2021
Expected end date:	31 December 2022

Computer resources allocated/used for the current year and the previous one

(if applicable)
Please answer for all project resources

		Previous year		**Current year**	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)			40 MSBU	4.7 MSBU (by 20 June)
Data storage capacity	(Gbytes)			20.000	20.000

Summary of project objectives (10 lines max)

The main areas of attention are:

- introduction and optimization of flow-dependent assimilation techniques (4D-Var, 3/4DEnVar)
- increasing the range and impact of high-resolution and remote sensing data to be assimilated (esp. all-sky radiances, crowd-sourced observations and satellite surface observations)
- improvement of the model behaviour for fog, convection initiation and stable boundary layer conditions
- a more sophisticated description of the radiation-cloud-microphysics-aerosol interaction and stable boundary layer conditions
- introduction of a more sophisticated surface analysis and modelling system.
- development of hectometric resolution nowcasting (ensemble) setups.
- Computational efficiency and scalability

Summary of problems encountered (10 lines max)

Running experiments in Reading rather slow at the moment, which is giving delays in the pre-operational test experiments which have started last May. Excellent support from ECMWF as usual.

Summary of plans for the continuation of the project (10 lines max)

In the period until end 2022, the main priorities will be:

- bringing several major developments like the overhaul of the surface analysis and modelling system, and the introduction of aerosols parametrizations and near-real-time CAMS aerosol data to operational status
- the continued development of ensemble assimilation techniques and assimilation in the nowcasting range
- increased use of very high resolution crowd-sourced data and satellite surface observations
- alleviating forecast model deficiencies, with focus on fog, stable conditions, and surface characteristics
- the development of hectometric scale and nowcasting ensemble setups, with focus on urban aspects
- and achieving enhanced computational efficiency, scalability and portability.

List of publications/reports from the project with complete references

Jan Barkmeijer , Magnus Lindskog, Nils Gustafsson, Jelena Bojarova, Roohollah Azad, Isabel Monteiro, Pau Escribà, Eoin Whelan, Martin Ridal, Jana Sánchez-Arriola, Ole Vignes, Roel Stappers, Roger Randriamampianina, 2021: HARMONIE-AROME 4D-Var. *ALADIN-HIRLAM Newsletter* 16, p.20, www.umr-cnrm.fr/aladin/IMG/pdf/nl16.pdf

Evgenia Belova, Sheila Kirkwood, Peter Voelger, Sourav Chatterjee, Karathazhiyath Satheesan, Susanna Hagelin, Magnus Lindskog & Heiner Körnich, 2021: Validation of Aeolus winds using ground-based radars in Antarctica and in northern Sweden , *AMT special issue: Aeolus data and their application (AMT/ACP/WCD inter-journal SI),* under review, DOI:10.5194/amt-2021-54

Evert I. F. de Bruijn, Fred C. Bosveld, Siebren de Haan, and Albert A.M. Holtslag, 2021: Opportunistic Sensing with Recreational Hot-Air Balloon Flights, *BAMS*, E273–E278, doi.org/10.1175/BAMS-D-19-0285.1

Chen, J., et al., 2021: Quality control and bias adjustment of crowdsourced wind speed observations. *QJRMS* (under review).

J. Díaz-Fernández, P. Bolgiani, D. Santos-Muñoz, L. Quitián-Hernández, M. Sastre, F. Valero, J.I. Farrán, J.J. González-Alemán, M.L. Martín, 2021: Comparison of the WRF and HARMONIE models ability for mountain wave warnings. *Atmospheric Research* (under review)

Kasper S. Hintz, Conor McNicholas, Roger Randriamampianina, Hywel Williams, Bruce Macpherson, Marion Mittermaier, Jeanette Onvlee-Hooimeijer, Balázs Szintai, 2021: Crowd-sourced observations for short-range numerical weather prediction: Report from EWGLAM/SRNWP Meeting 2019, *Atm.Sci.Lett.* https://doi.org/10.1002/asl.1031

Lindskog, M., A. Dybbroe, R. Randriamampianina. 2021: Use of Microwave Radiances from Metop-C and Feng Yun-3 C/D Satellites for a Northern European Limited-area Data Assimilation System, *Adv. Atmos. Sci.*, http://www.iapjournals.ac.cn/fileDQKXJZ/journal/article/dqkxjz/newcreate/AAS-2020-0326.pdf

Daniel Martin-Perez, 2020: Update of the Use of CAMS Aerosols in Harmonie-Arome. *ALADIN-HIRLAM Newsletter* 15, p.53, www.umr-cnrm.fr/aladin/IMG/pdf/nl15.pdf

Kristian Pagh Nielsen, Laura Rontu, Emily Gleeson, 2021: Uncertainties in Numerical Weather Prediction Chapter 9 – Radiation. Edited by Haraldur Ólafsson and Jian-Wen Bao. Pages 237-264, Elsevier.

L. Quitián-Hernández, P. Bolgiani, D. Santos-Muñoz, M. Sastre, J. Díaz-Fernández, J.J. González-Alemán, J.I. Farrán, L. Lopez, F. Valero, M.L. Martín, 2021: Analysis of the October 2014 subtropical cyclone using the WRF and the HARMONIE-AROME numerical models: Assessment against observations, *Atm. Res*. 105697, ISSN 0169-8095, https://doi.org/10.1016/j.atmosres.2021.105697.

de Rooy, W.C., Siebesma, P., Baas, P., Lenderink, G., de Roode, S., de Vries, H., van Meijgaard, E., Meirink, J.-F., Tijm, S., and van ’t Veen, B., 2021: Model development in practice: A comprehensive update to the boundary layer schemes in HARMONIE-AROME. *Geophys. Model Dev.* (under review).

D. Suárez-Molina and J. Calvo, 2021: Very high-resolution experiments at AEMET. *ALADIN-HIRLAM Newsletter* 16, p.106, www.umr-cnrm.fr/aladin/IMG/pdf/nl16.pdf

Walsh, E., Bessardon, G., Gleeson, E., and Ulmas, P., 2021: Using machine learning to produce a very high resolution land-cover map for Ireland, Adv. Sci. Res., 18, 65–87, https://doi.org/10.5194/asr-18-65-2021.

Xiaohua Yang and Henrik Feddersen, 2021: Benefit of early delivery ASAP data to LAM forecasts, *ALADIN-HIRLAM Newsletter* 16, p.45, www.umr-cnrm.fr/aladin/IMG/pdf/nl16.pdf

Summary of results (from January 2021 to June 2021)

The HIRLAM-C research programme (January 2016 - December 2022) is a research cooperation of the national meteorological institutes of Denmark, Estonia, Finland, Iceland, Ireland, Lithuania, Netherlands, Norway, Spain and Sweden, with Meteo-France as associated member. Since January 2021, the HIRLAM-C cooperation forms part of the wider ACCORD consortium. Within HIRLAM-C, research efforts are focussed on the development, implementation and further improvement of the mesoscale analysis and forecast system Harmonie, and its associated ensemble prediction system HarmonEPS. A Harmonie Reference system is being maintained on the ECMWF HPC platform. The computational resources for the HIRLAM-C Special Project at ECMWF are primarily used for experimentation with, and evaluation of the performance of, newly developed elements for this Reference System. A set of pre-operational tests for Cy43h2.2 has recently started and will continue this summer, among others with full ensemble tests. Below, the main R&D and testing activities in the fields of data assimilation, the atmospheric forecast model, surface analysis and modelling, ensemble forecasting and code efficiency during the past year are outlined.

A) Data assimilation

A1: Optimal use of (high-density) atmospheric observations:
At present, Harmonie still uses as default a 3D-Var assimilation system, which assimilates conventional data and cloud-free radiances from AMSU-A, AMSU-B, MHS and IASI. Additionally, several types of spatially and temporally dense data can be assimilated optionally: radar radial wind and reflectivity volume data, GNSS ZTD, GPS-RO, SEVIRI water vapour observations, Mode-S data, AMV's and scatterometer winds, and data from private weather stations. Continuous efforts are being made to further enhance the impact of these data types through improved data quality control, the application of more intelligent thinning or super-obbing strategies, enhanced variational bias correction, and careful tuning and optimization of the observation statistics and structure functions.

Work has continued to add more observations. These include satellite data from new platforms or instruments like MetOp-C, MWHS2 and ATMS, but also a greater use of Mode-S aircraft data and new sources of surface observations. Good results have been achieved with the inclusion of MetOp-C scatterometer data using the supermodding approach. Supermodding is now also being applied to radiances. Recently added satellites have improved the data coverage for radiances especially in the morning. Several ways are being investigated to further enhance the impact of these data, e.g. by alterations in varBC and in the use of low-peaking channels.

Promising results have been obtained in studies of the assimilation of surface pressure and screen level parameters in the upper air analysis has continued in several services, and in an assessment of the potential use of BUFR descent radiosonde data. Experiments are continuing with various methods of quality control and settings for varBC and observation error characteristics.

Investigations on quality control and bias correction of crowd-sourced data (private weather stations, smart phones, wind turbines, road observations) are gradually widening. Various quality control algorithms have been developed for the handling of private weather station data for an increasing variety of parameters, most of which have been made available via open source repositories (e.g. Chen et al. 2021). One issue which deserves to be studied more is the impact of thinning approaches used to avoid the use of correlated observations. The revision in the processing of smart phone data made by DMI staff has made it possible to "depersonalize" these data and thus avoid privacy constraints on their use, but this has made bias correction more complicated. It is being investigated how the processing of these data can be further improved. In general, pre-processing of data from moving platforms is still quite challenging (Hintz et al. 2021).

A serious restriction of most third-party datasets is that they are surface-based only. It will be desirable to also obtain and make greater use of third-party ground-based boundary layer profiling, e.g. making more use of the increasing amount of airport or urban lidar and ceilometer data.

A2: Development, operationalization and optimization of flow-dependent data assimilation methods.
On the algorithmic side, efforts have remained focussed on two objectives: 1) Bring 4D-Var to operational status in Cy43h2.2, and 2) Further development of the existing hybrid 3/4D EnVar ensemble assimilation methods, and introducing them within the OOPS code framework.

A proposed version of 4D-Var is presently undergoing final pre-operational testing. This final 4D-Var setup, described in the paper by Barkmeijer et al. (2021), involves a 3h rather than a 2h time window, as this will admit significantly more satellite observations. Experiments have shown that 4D-Var is clearly superior to 3D-Var in clean tests, and is able to make use of significantly more data from e.g. the recently added microwave satellite radiances and satellite winds (e.g. see fig.1). Currently, attention is focussing on optimization needed to make the (computationally more demanding) 4DVar operationally feasible (e.g. by (selective) use of single precision, or by the use of quadratic rather than linear grids), extend it with not-yet-operational observation types (e.g. GNSS STD, SEVIRI, CRiS, …), and to test its usefulness in nowcasting setups in combination with the overlapping windows technique and/or the continuous data assimilation approach developed at ECMWF. Also experiments are still continuing on different configurations for the control of lateral boundary conditions, the use of a large extension zone and the use of model tendency as control vector.

A Harmonie 3D-Var version within the OOPS framework is running properly with conventional data, producing very similar results to standard 3D-Var. This OOPS-based system will be expanded and tested with the use of non-conventional data, and with the existing 4D-Var and hybrid 3/4D EnVar schemes in the coming 1-2 years. Aspects like the optimal ensemble size, localisation length, use of full or hybrid EnVar, and spinup issues will be studied. Intercomparisons between the different methods will be done, both from a data assimilation and from an ensemble perspective.

A3. Optimization of data assimilation setups for the nowcasting range

For nowcasting, a main goal is to move towards sub-hourly cycling and assess the challenges for assimilation on timescales of ~10min cycling. Quite many script changes are needed for this and work on them has started. Problems with correlated observations become more pronounced at higher temporal and spatial resolution, and for 4D-Var especially, temporal correlations are an issue to be tackled. For the handling of spatially correlated observations, error modelling approaches should be considered as a more sophisticated alternative to the standard thinning or superobbing practices. A dedicated nowcasting branch will be set up in the near future for testing different configurations. The experiments in the nowcasting branch should include the cloud initialization approach (Gregow 2017), the overlapping windows ensemble technique (Yang, 2018), and the Field alignment / variational constraint (FA/VC) nowcasting method applied to radar observations (Geijo 2012, 2019), but also the continuous assimilation approach as used by ECMWF. Different options for improving the initialisation and reducing spinup will be compared in this branch. In the cloud initialization technique, the use of satellite observations has been restricted to the neighbourhood of the cloud top. Fog and low cloud cases will be used to test the ability of this adjusted method to adequately remove erroneous low clouds and fog in the model. Verification will be carried out not only for lead times but also for time of delivery.

One relevant issue is how to best take the large scales from the nesting model into account in the nowcasting range. For this, there presently are the LSMIX and Jk cost function approaches available. Ole Vignes has derived an alternative model cost term treatment, by combining nesting and nested model background into a single background cost term; first tests have shown the performance of this “mixed” cost term approach to be better for the short range (up to ~12h) in the summer, less clear results were obtained for a winter period. These investigations will be continued.

B) Atmospheric forecast model

B1. Studies to eliminate systematic model errors for clouds and boundary layer behaviour:

A major long-standing forecast has been the prediction of fog under stable boundary conditions, especially over sea: fog generally building up too quickly, having a too large extent (especially over sea), too much cloud water, and starting a too quick and too strong cooling over sea.

Physics experiments in 2020 have provided a better understanding of the way in which a too high long wave emissivity and the assumption of constant cloud droplet number concentrations in the ICE3 microphysics scheme have contributed to this model behaviour. It was shown that with an adapted LW emissivity formulation (developed by Kristian Nielsen), in combination with a much lower and homogeneous value of Nc in ICE3, fog problems could already be mitigated substantially, although not completely over sea. Also, these changes led to a better representation of precipitation from shallow convective clouds, and a strong reduction in the cloudy SW radiation bias. This year, it turned out (both from observations and on theoretical grounds) that it would be more appropriate to assume a profile for the cloud condensation nuclei number density, increasing from ~50 to 200 in the first 1000m. The set of changes with this profile is presently being tested over three domains.

A more definitive solution is expected from the use of the second moment scheme LIMA and the introduction of aerosol parametrizations in combination with near-real-time aerosol initialization from CAMS. However, there are still many aspects to experiment with and tune this setup. One of the things that is being considered additionally, is the need for a parametrization describing the fraction of aerosols which are active in condensation. Measurements of droplet size distributions show that shallow fog has larger droplets than deeper well-developed fog, cooled from the top by turbulence. Because of this, it is being considered to make the activation of aerosols TKE and vertical velocity dependent. HIRLAM participation in the model inter-comparison experiments for the SOFOG3D IOP's may be very useful for the tuning of this aerosol activation.

A long-standing weakness of the Harmonie model has been its poor representation of open cell convection features observed in clouds, and the precipitation associated with such clouds. With the removal of a bug in the evaporation/melting of rain/snow, and reducing the too strong momentum mixing by the shallow convection scheme, a more realistic organization of open cell cloud structures could be observed. In the case that additionally also the new ECUME-6 sea surface scheme is used, which produces more realistic surface fluxes over sea, the model cloud cover and precipitation behavior in coastal areas also becomes much more realistic. These changes have all been implemented and are undergoing testing.

One of the parameters which probably contributes to a negative bias in cloud cover which has been seen recently, is the eddy turnover time in the EDMF shallow convection scheme. 1-D experiments suggest that this can be improved by making the eddy turnover time dependent on height. Experimentation with this is ongoing.

B2. Improved description of the cloud-radiation – microphysics- aerosol interaction:

Presently, 11 types of aerosol can be retrieved from near-real-time (NRT) CAMS data. Cases and 10-day verification results have shown that a consistent use of NRT aerosol in the model physics has significant (beneficial) impact on model snowfall amounts, dust attenuation of SW radiation, fog and precipitation. It was further demonstrated that the introduction of a threshold in the cloud droplet radius for the activation of auto-conversion and of the vertical velocity in the cloud droplet sedimentation may be required to improve precipitation forecasts. At the moment, a consistent data flow is being set up for the use of near-real-time aerosol information from CAMS into the different relevant parts of the model physics (radiation, cloud and microphysics schemes). Later this year, more extensive experimentation will be done with the LIMA second moment scheme as alternative to the 1-moment ICE3 microphysics. Both the meteorological and computational performance of LIMA will be assessed in more detail.

B3. Sub-km resolution modelling

Many HIRLAM services are experimenting by now with 500-750m horizontal resolution model setups. Some of these setups have reached operational status already, and gradually efforts are being extended to smaller scales, to discover the limitations which the present dynamics and physics parametrizations may have there. Fig.2 shows an example of a 150m resolution setup over the Faroer Islands (with steep topography), which was shown to provide a realistic representation of local conditions and remain numerically stable even in a case of Bf 12 force winds.

Tests with model domains at 500m resolution and 90 vertical levels have shown issues with climate generation files when running over sea-dominated domains. These studies have also demonstrated the benefits of coupling the hydrometeors with those of the nesting model at the lateral boundaries. In sub-km resolution Harmonie runs, boundary layer cloud streets are often represented too strongly and smoothly. The introduction of stochastic elements in the physics parametrizations may help to disrupt the too smooth model behavior and locally trigger deep convection. These ideas will be pursued further.

C) Surface analysis and modelling

C1: Improving the sophistication of surface model components

A critical aspect for surface modelling is the quality of the surface characterization (orography and physiography). Especially at sub-km resolution, there is a clear need to improve over the existing ECOCLIMAP-Second Generation global database by means of local high-resolution databases and/or satellite data in a consistent manner. A serious limitation of ECOCLIMAP-SG is that it assumes each 300m pixel to be homogeneous, corresponding to only the dominant land-use type for that pixel area. In reality, in many landscapes this assumption is clearly not valid, leading to biases in the assumed surface vegetation roughness. Temporary fixes to counter this database weakness have been prepared and tested on various domains (e.g. fig.3). However, in the light of ambitions to move to hectometric scales, options should be explored to use the highest possible resolution satellite data from e.g. Sentinel-2 in combination with machine learning to improve physiographic input data for different patches in the model, as a follow-up on exploratory studies on this by MetEireann.

The studies on the behaviour of ECOCLIMAP-SG in high-resolution setups also have given a renewed impetus to explore a more realistic treatment of the canopy layer (both vegetation and urban) in the determination of surface roughness lengths. SMHI and AEMET staff have introduced, and have begun evaluating, a roughness sublayer into the model which accounts for the effects of canopy-induced turbulent mixing (Harman and Finnigan 2007). This parametrization includes the concept of displacement height, and a more advanced computation of the roughness length z_{0m} as a function of leaf area index (vegetation sparseness) and atmospheric stability.

In the past two years, many tests have been done with the new multi-layer diffusion soil, extended snow and snow-over-vegetation schemes (DIFF,ES, MEB) and the new SEKF assimilation scheme for soil and snow. These schemes have now all been combined and are being run pre-operationally for the Arome-Arctic domain. This setup uses assimilation of screen level conventional and additional private weather station data. Assimilation of snow surface temperature and soil moisture will be added at a later stage. The new setup has shown a positive impact on model 2m biases and snow behaviour over the Arctic area. The next step this year will be to validate the combined impact of the surface model and assimilation changes in various other Harmonie domains.

A new parametrization for surface fluxes over the ocean, ECUME-6, has been activated. ECUME-6 in general leads to stronger fluxes of heat and moisture than its predecessor. Overall, the switch to ECUME-6 leads to more accurate near-surface values and low-tropospheric profiles for temperature and humidity, a lower bias in mean sea level pressure, and somewhat more fog over both land and sea.

Progress has been made in testing the performance of coupled atmosphere-ocean-wave model setups over Arctic and Irish domains, with neutral to slightly positive results. SMHI and KNMI have continued work on creating and assessing the coupling of the model with local hydrological models.

C2. Enhanced use of satellite surface observations in combination with more advanced surface assimilation

The new SEKF assimilation scheme will enable the increased use of satellite surface products. The first thing to be tested in this context, is the assimilation of satellite snow extent, which has been shown to be important in the cold season in areas where SYNOP snow cover stations are sparse or missing. To avoid errors in the screening/thinning and interpolation of raw satellite data, the so-called "snow barrel" method has been developed by FMI for the assimilation of satellite snow extent data. Presently, the data flow in the Harmonie assimilation system is being adjusted to permit satellite snow extent assimilation with the barrel approach.

The second new data type to be included in the surface assimilation is the assimilation of OSI-SAF sea ice surface temperatures, in order to improve SICE temperature forecasts on or near sea ice. For this purpose, use is made of a so-called bias-aware EKF, which is better able to take into account the SICE model temperature bias under certain situations. Normally, in an EKF, a bias correction is done at analysis time; but in SICE, it was seen that for "fast" variables like air temperature, a bias was very quickly regained during the model forecast. In the bias-aware EKF, the bias correction is applied incrementally during the first three hours of the model forecasts, as a flux term in the surface energy balance. The general performance of the new sea ice data assimilation scheme is being tested against independent MODIS observations.

Thirdly, it is aimed to activate lake assimilation, after which satellite (ASCAT, SMOS, MODIS and H-SAF) retrievals of soil, vegetation, sea surface and inland waters properties and snow- and ice-covered surfaces will be added progressively to the surface assimilation. The data flow for including these data is being prepared.

After the introduction of the SEKF soil and snow assimilation, it is intended to move stepwise towards a more advanced EnKF setup which can be used as a basis for a strongly coupled atmosphere-surface data assimilation system and the assimilation of satellite surface radiances. A rough road map for this has been developed. A restructuring of the surface assimilation is planned with the aim to permit the use of various assimilation algorithms (SEKF, EKF, EnKF) independently from specific tiles and/or patches (which is presently not the case).

D) Probabilistic forecasting:

In the EPS area, recent work has mainly focussed on the best use of model perturbations. In 2020 many experiments have been done with perturbations arising from model tendencies (SPPT) versus parameter perturbations (SPP). SPP perturbations were shown to have good impact and SPPT perturbations offered very little added value. In the past months, work has concentrated on documenting the SPPT vs SPP experiments (Frogner et al., 2021) and further sensitivity studies with SPP behaviour in low cloud and fog conditions. At present, 11 parameters in the microphysics, radiation, convection and turbulence schemes can be perturbed. The four most influential parameters impacting on the spread of low cloud forecasts have been identified: the stable condition length scale, the saturation limit sensitivity for condensation, the threshold cloud thickness used in the shallow/deep convection decision, and the asymptotic free atmospheric length scale. In summer, for cloud variables the SPP ensemble performance gives better spread but some degradation in screen-level variables; in winter the SPP perturbations letter to better spread but also clearly higher RMSE. A paper on this work has recently been submitted to MWR. The main focus in the SPP activities for the remainder of the year will be to experiment with alternative parameter pdf assumptions and study the behaviour of correlated parameters. Initial experiments using the same perturbation pattern for two correlated parameters resulted in less bias and better spread. Also, attempts are being made to reduce the computational cost of SPP, e.g. by applying SPP only every hour instead of every time step.

Earlier, it has been seen that ensemble members tend to be drier than the control run. It has been shown that this is at least largely due to soil wetness perturbations which were insufficiently constrained in cases where the soil moisture in reality was above field capacity or below the wilting point. It is being considered how a more effective constraint for these perturbations can best be introduced.

Now that the hybrid Envar/Brand and LETKF codes have been ported to Cy46h, in the coming year also more detailed inter-comparisons will be made of the performance of the three mechanisms available for generating initial condition perturbations: EDA, LETKF and hybrid EnVar/Brand. Also, a study will be done on the optimal number of ensemble members using EDA.

At DMI, the impact has been studied of making ensemble forecasts in single precision (SP), while keeping the control run in double precision. SP ensemble member forecasts were ~30% faster than DP runs on the DMI Cray. The meteorological differences between SP and DP forecasts for a single member appeared to be quite small compared to differences between individual perturbed members, and SP and DP runs generally performed equally well in terms of meteorological scores. However in view of two crashes under winter conditions, the numerical stability of SP runs remain to be investigated further. Also, some differences were seen in vertical humidity profiles between SP and DP members which need to be studied more deeply.

E) Code efficiency and portability

Various tests have been carried out for Harmonie-Arome in single vs double and mixed precision, not just for the forecast model but also for the various parts of the 3- and 4D-Var data assimilation and ensemble systems, see the DA and EPS sections above. When going from double to mixed precision, typical run time reductions of ~35-40% are found, mainly because data volumes get smaller (memory cache can hold larger arrays, data involved in MPI passing is halved).

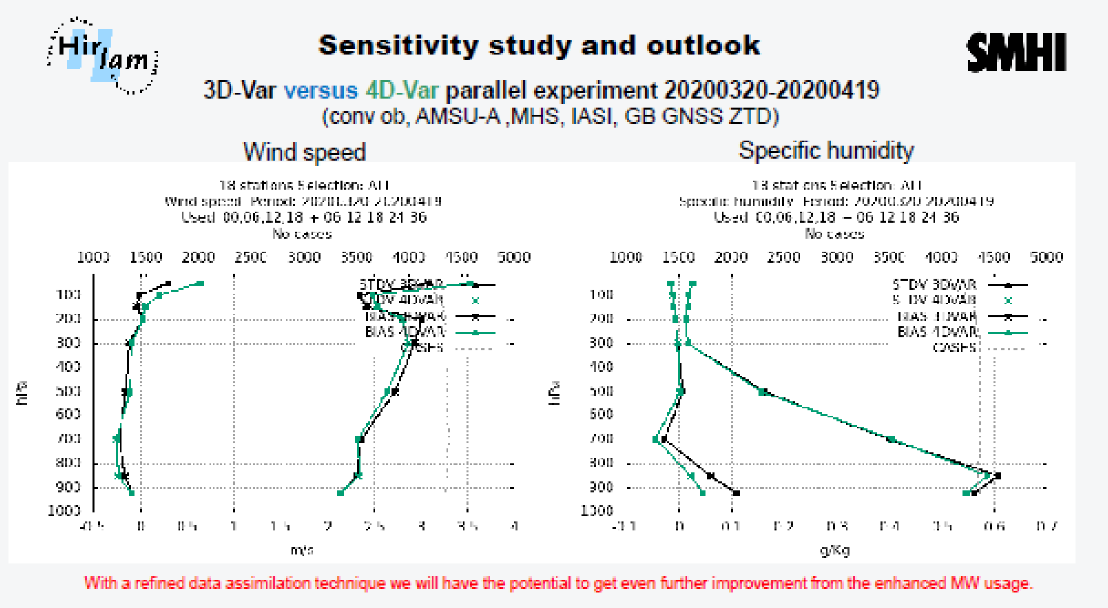
In collaboration with ECMWF, staff at DMI have been working on improving the computational efficiency of a part of the ECRAD radiation scheme by ~40%, using machine learning. This will be implemented and tested in Harmonie-Arome as well.

The Barcelona Supercomputing Center has been performing several optimizations in the Harmonie-Arome system, both at the code and the compilation level, and is also working to develop a computational performance monitoring system based on BSC standard open source tools. This system will be used in the future for routine monitoring of the Harmonie main develop branch at ECMWF, and for the testing of the consequences of new code changes on computational performance and cost.

Given the increasing heterogeneity of HPC architectures, it is important to establish model performance on a range of different architectures. A “containerized” version of the model has been made to permit easier porting to such alternative platforms for benchmarking purposes. In the context of a procurement process for obtaining a new HPC for the joint operational use of the services of DMI, IMO, KNMI and MetEireann, benchmarks have been performed on a variety of systems involving CPU, GPU, AMD and ARM architectures.

On the short term, an OpenAcc directive approach will be used in order to make the various components of the ACCORD LAM models, including Harmonie-Arome, ready for use on GPU and mixed CPU-GPU architectures. Further code refactoring will also be needed to introduce ATLAS data structures into the ACCORD LAM models. In the wider ACCORD team, good progress has been made already with the ATLAS refactoring and several parts of the code have already been made GPU-ready. New HIRLAM staff will join these efforts in the near future.

Figures:



Sensitivity study and outlook

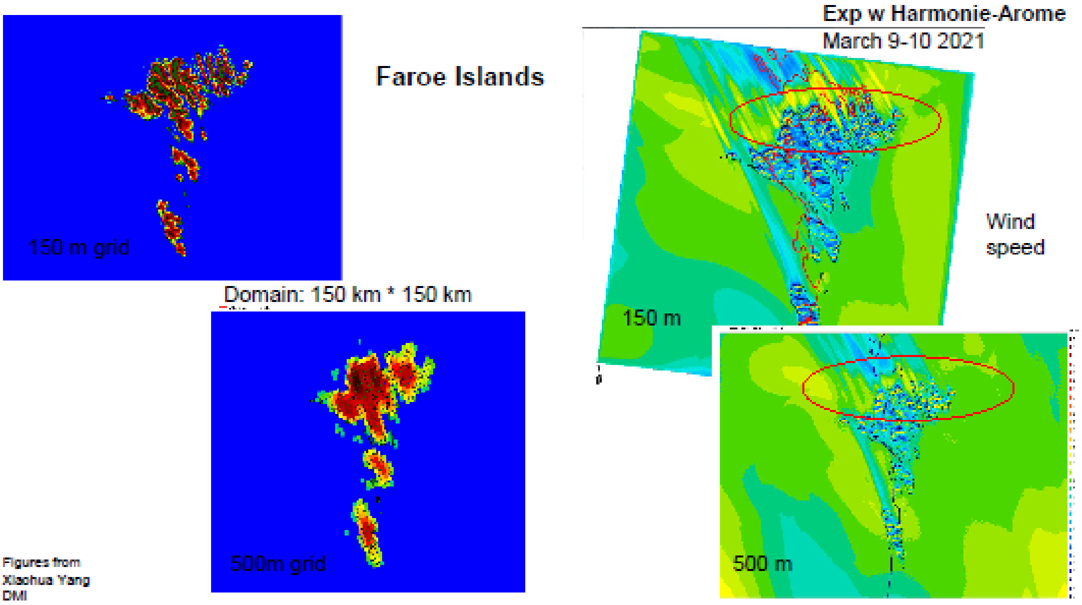
3D-Var versus 4D-Var parallel experiment 20200320-20200419
(conv ob, AMSU-A ,MHS, IASI, GB GNSS ZTD)

Wind speed

Specific humidity

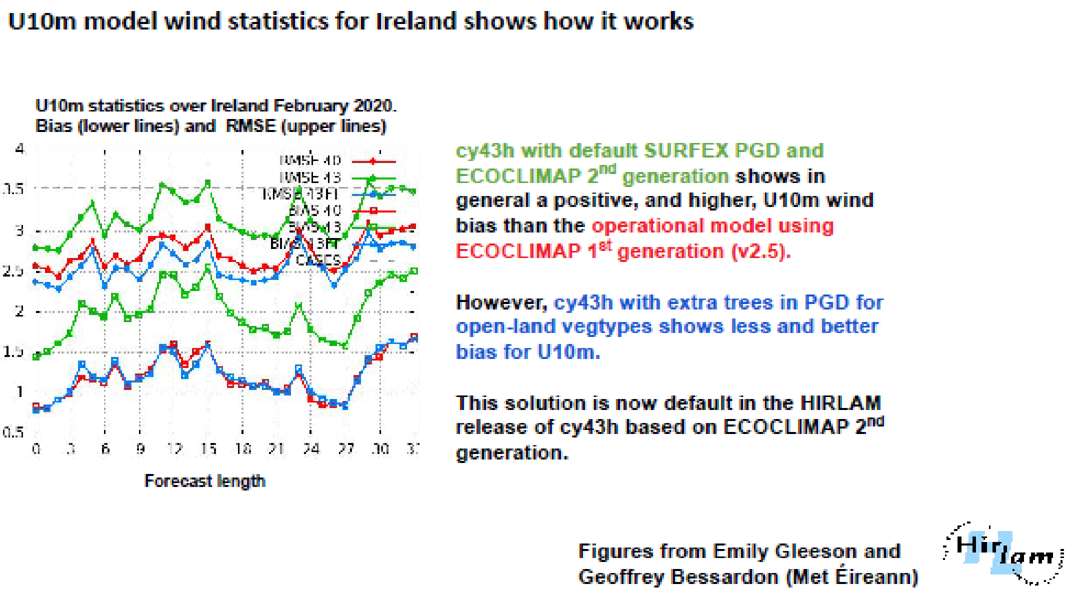
With a refined data assimilation technique we will have the potential to get even further improvement from the enhanced MW usage.

Fig. 1 An example of a 1-month experiment comparing the impact of assimilating the operational data sources plus new microwave radiance data in 3- and 4D-Var. The figure shows standard deviation and bias for vertical profiles of wind speed (left) and specific humidity (right), verified against radiosonde data. 4D-Var (green curves) generally outperforms 3D-Var (black) in this and other similar experiments.



Figures from
Xiaohua Yang
DMI

Fig. 2: Example of a 150m and a 500m resolution model setup for the Faroer Islands. On the left side is shown the complex orography of the island group, on the right side a case study for March 9 2021, when hurricane-force winds struck the islands. Verification against available local observations showed that this extreme event was well captured by the 150m model, but did not reach the actual code red wind warning levels in the 500m and 2.5km resolution model setups.


U10m model wind statistics for Ireland shows how it works

cy43h with default SURFEX PGD and ECOCLIMAP 2nd generation shows in general a positive, and higher, U10m wind bias than the operational model using ECOCLIMAP 1st generation (v2.5).

However, cy43h with extra trees in PGD for open-land vegtypes shows less and better bias for U10m.

This solution is now default in the HIRLAM release of cy43h based on ECOCLIMAP 2nd generation.

Figures from Emily Gleeson and Geoffrey Bessardon (Met Éireann)

Fig. 3: Experiment demonstrating the degradation in near-surface wind and roughness caused by the use of only the dominant vegetation type patch in an ECOCLIMAP-Second Generation gridbox, and how this can be alleviated by inclusion of extra tree cover in the physiographic database (PGD), which realistically should have been there. The figure shows the verification of u10 standard deviation and bias in an experiment over Ireland. There is a clear positive wind bias caused by the introduction of ECOCLIMAP Second Generation in Cy43h2 when using the default Surfex PGD setup (green curves), as compared to Cy40 using ECOCLIMAPv2.5 (red). This bias is practically eliminated by the introduction of “fake trees” in the PGD (blue).