

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 July 2019 – June 2020

Project Title: HIRLAM-C 2d phase (2019-2020) 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 2019

Expected end date: 31 December 2020

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)	32 MSBU	32 MSBU	32 MSBU	32 MSBU (by 15 June)
Data storage capacity	(Gbytes)	20.000	20.000	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)
- improvement of the model behaviour for low clouds, fog and stable boundary layer conditions
- increasing the range and impact of high-resolution and remote sensing data to be assimilated (esp. all-sky radiances, OPERA radar data, crowd-sourced observations and satellite surface observations)
- a more sophisticated and consistent description of the radiation-cloud-microphysics-aerosol interaction and winter stable boundary layer conditions
- a more sophisticated surface analysis and modelling system.
- definition of and experimentation with sub-km resolution nowcasting (ensemble) setups

Summary of problems encountered (10 lines max)

No problems worth mentioning. Excellent support from ECMWF as usual.

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

In the period until end 2020, the main priorities will be:

- bringing major developments like 4D-Var and the overhaul of the surface analysis and modelling system to operational status
- the continued development of ensemble assimilation techniques
- increased use of crowd-sourced and satellite surface observations
- bring developments which improve forecast model behaviour for fog and low clouds into operations
- prepare for the introduction of aerosol parametrizations in combination with near-real-time aerosol analyses from CAMS
- the development and introduction of high-resolution nowcasting ensemble setups
- and achieving enhanced computational efficiency, scalability and portability.

List of publications/reports from the project with complete references

Andrae, U., et al., 2020: “A continuous EDA-based ensemble in MetCoOp”, ALADIN-HIRLAM Newsletter no. 14, p. 189, January 2020, available from <http://www.cnrm-game-meteo.fr/aladin/>

Bremnes, J.B., 2020: [Ensemble Postprocessing Using Quantile Function Regression Based on Neural Networks and Bernstein Polynomials](https://doi.org/10.1175/MWR-D-19-0227.1). *Mon. Wea. Rev.*, **148**, 403–414, <https://doi.org/10.1175/MWR-D-19-0227.1>

Fortelius, C., et al., 2020: “Selected NWP-related research at FMI”. ALADIN-HIRLAM Newsletter no.14, p.58, January 2020. Available from <http://www.cnrm-game-meteo.fr/aladin/>

Frogner, I., et al., 2019: “HarmonEPS—The HARMONIE Ensemble Prediction System”. *Wea. Forecasting*, **34**, 1909–1937, <https://doi.org/10.1175/WAF-D-19-0030.1>

Geijo, C., 2019: “Modeling flow-dependent covariances with Gaussian Integrals”, ALADIN-HIRLAM Newsletter, no.13, p.41, August 2019, available from <http://www.cnrm-game-meteo.fr/aladin/>

Emily Gleeson, et al., “HARMONIE-AROME Single-column Tools and Experiments”, Under review for ASR 2020

Ivarsson, K.I., 2019: “Progress in turbulence, shallow convection and cloud microphysics in MetCoOp”, ALADIN-HIRLAM Newsletter, no.13, p.32, August 2019, available from <http://www.cnrm-game-meteo.fr/aladin/>

Ivarsson, K.I., 2019: "A Score for Probability Forecasts of Binary Events Based on the User's Cost–Loss Relations", *Weather and Forecasting*, 02-2020 sid 113-127 (<https://doi.org/10.1175/WAF-D-19-0013.1>)

Hamalainen, K., 2019: <https://journals.ametsoc.org/doi/full/10.1175/MWR-D-19-0184.1?journalCode=mwre>

Kangas, M., 2019: <https://www.geosci-model-dev.net/12/3481/2019/>

Lenderink, G., et al., 2019: <https://iopscience.iop.org/article/10.1088/1748-9326/ab214a/pdf>

Lindskog, M. and Landelius, T., 2019. Short-Range Numerical Weather Prediction of Extreme Precipitation Events Using Enhanced Surface Data Assimilation. *Atmosphere*, 10, 587; doi:10.3390/atmos10100587

Müller, A., Gillard, M., Nielsen, K.P., Piotrowski, Z., 2019: Definition of novel Weather & Climate Dwarfs, *arXiv:1908.07040 [cs]* (2019), <https://arxiv.org/abs/1908.07040>

Nipen, T., et al., 2019: Adopting citizen observations in operational weather prediction, *Bull. Am. Meteorol. Soc.*, 1011E43-E57 <https://doi.org/10.1175/BAMS-D-18-0237.1>

Randriamampianina, R.; Schyberg, H.; Mile, M. Observing System Experiments with an Arctic Mesoscale Numerical Weather Prediction Model. *Remote Sens.* 2019, 11(8), 981; <https://doi.org/10.3390/rs11080981>

Rontu, L., Pietikäinen, J.-P., and Martin Perez, D.: “Renewal of aerosol data for ALADIN-HIRLAM radiation parametrizations”, *Adv. Sci. Res.*, 16, 129–136, <https://doi.org/10.5194/asr-16-129-2019>, 2019.

Rontu, L. et al. “Sensitivity of Radiative Fluxes to Aerosols in the ALADIN-HIRLAM Numerical Weather Prediction System”. *Atmosphere* 2020, 11, 205. <https://doi.org/10.3390/atmos11020205>

Samuelsson, P., Kourzeneva, E., de Vries, J., and Viana, S., 2020: “HIRLAM experience with ECOCLIMAP Second Generation.” ALADIN-HIRLAM Newsletter no.14, p.154, January 2020. Available from <http://www.cnrm-game-meteo.fr/aladin/>

Sanchez-Arriola, J., et al. 2020: “A major update in AEMET operational suite including radar reflectivity, ASCAT data, humidity in the large scale blending and assimilation of screen level variables in 3D-Var”, ALADIN-HIRLAM Newsletter no.14, p.128, January 2020. Available from <http://www.cnrm-game-meteo.fr/aladin/>

Sass, B.H., 2019: “SLX verification scheme (Structure of Local eXtremes)”, ALADIN-HIRLAM Newsletter, no.13, p.37, August 2019, available from <http://www.cnrm-game-meteo.fr/aladin/>

Schonach, D., 2019: “Ingestion of clouds into the MetCoOp Nowcasting model”, ALADIN-HIRLAM Newsletter, no.13, p.18, August 2019, available from <http://www.cnrm-game-meteo.fr/aladin/>

Stoll, PJ, Valkonen, TM, Graversen, RG, Noer, G., 2020: “A well-observed polar low analysed with a regional and global weather-prediction model”. *QJR Meteor. Soc.* 2020 1-28. <https://doi.org/10.1002/qj.3764>

van Stratum, B.J.H., et al., 2020: “Wind turbine parameterisation in Harmonie-Arome”, ALADIN-HIRLAM Newsletter no. 14, January 2020. Available from <http://www.cnrm-game-meteo.fr/aladin/>

Summary of results of the current year (from July 2019 to June 2020)

The HIRLAM-C research programme (January 2016 - December 2020) 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. 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. 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. From the summer of 2019 until summer 2020, a large amount of testing efforts (and most of the special project resources for 2019 and 2020) at ECMWF have been devoted to the assessment of many new model components in Harmonie Cy43h2.1. For experimentation during the remainder of 2020, use will be made of the dedicated SBU pool provided by HIRLAM members on top of the special project resources.

A) Data assimilation

A1: Optimal use of (high-density) atmospheric observations:

At present, Harmonie 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, SEVIRI water vapour observations, Mode-S data, AMV's and scatterometer winds. All these data types show positive impact after continuous efforts on 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.

This year, work has continued to improve and homogenize the preprocessing of various types of observations. Many HIRLAM services are testing the ability of (and possible need for adaptations in) the Scalable Acquisition and Prep-Processing (SAPP) system to handle standard as well as non-conventional data types (e.g. Mode-S, crowd-sourced data). A more homogeneous way of handling OPERA data is needed to permit use of radar data from a wider range of countries. In this context, the ALADIN, LACE and HIRLAM partners have decided to use the HOOF homogenization tool developed by LACE for radar homogenization, and the PrepOpera algorithm developed by HIRLAM for super-obbing of the radar data. The preprocessing of Mode-S data at KNMI has been adapted to derive temperature data with greater accuracy. The quality of these data now has become similar to those of AMDARs and radio sondes, and bias correction of Mode-S temperature data no longer appears to be necessary.

The work on the assimilation of radiances from Metop-C, FY-3C and FY-3D, including the tuning of VarBC, has been successful, and several services have started to use these data operationally. Options to use more low-peaking channels and emissivity atlases have been investigated, with mixed success. Similarly, preparations have been made to include, and assess the impact of, scatterometer winds from Metop-C and HY2B. Real-time Aeolus Horizontal Line of Sight (HLOS) observations showed good impact in Harmonie, despite their relatively small data volume. Work has been done to derive a bias correction for these data. To take into account the footprint of scatterometer data relative to the effective model resolution, an improved supermodding approach has been implemented and optimized in the 3D- and 4D-Var observation operator for ASCAT winds. As a next step, the supermodding approach will be applied to radiances.

The assimilation (and bias correction) of surface pressure and screen level parameters in the atmospheric analysis is under investigation, as well as an assessment of the potential use of descent radiosonde data. The results look promising over several domains. Crowd-sourced data from e.g. private weather stations, smart phones, wind turbines, and road observations, have already demonstrated their potential for the verification and postprocessing of Harmonie, provided that sufficient attention is paid to their quality control (see e.g. fig.1): the greater spatial density of such data with respect to standard meteorological observing networks permits better identification of fine-scale weather features, in particular

also over urban areas. Experiments are presently ongoing to assess the added value which they may have when being assimilated directly in the analysis systems Harmonie setups. Investigations on further optimization of the quality control of crowd-sourced data with machine learning techniques are continuing.

In the context of Copernicus re-analyses for Europe and the Arctic, a European setup, CERRA, has been prepared for operations and been shown to be an improvement with respect to its predecessor UERRA, The CARRA (Copernicus Arctic Regional Re-Analysis) setup for the Arctic has entered operations. Experimentation is continuing with several options aimed at improving the standard 3D-Var setup: EDA with observational or random B-matrix perturbations, and large scale mixing through a Jk operator. These studies will later be re-evaluated for 4D-Var.

A2: Development, operationalization and optimization of flow-dependent data assimilation methods.

There are two main objectives in this area: prepare 4D-Var for operational use, and, for the longer term, continue the development and start deployment of (hybrid) ensemble assimilation methods.

Longer runs and case studies have been performed with multi-incremental 4D-Var with different combinations of high-resolution observations, in order to test the optimal settings for this system. SMHI has built in the option to run 4D-Var with a larger extension zone (200-300km) in the minimization than default, to better represent horizontal error correlations in the data assimilation. This larger E-zone is only applied in the minimization, and not for the forecast model, where it is not needed. Tests have shown that a larger extension zone in the minimization step gives more realistic results.

With a relatively limited set of observations, 3-hourly multiple outer loop 4D-Var has already shown a better performance than 3-hourly 3D-Var, at affordable cost. The use of 4D-Var has now been enabled for the full set of observations which can be used in the Harmonie 3D-Var, and impact assessment experiments with adding progressively all these observation types are ongoing for several domains.

In near-real-time runs of multi-incremental 4D-Var in nowcasting mode, hourly 3D-Var, hourly 4D-Var and 4D-Var with overlapping windows are now being intercompared. Two new adaptations, a correction for a recently discovered wrong use of LBC data and the introduction of control of LBC at the beginning of the time window, should further improve the 4D-Var performance in the first few hours of the forecast. It is planned to implement 4D-Var in Cy43h2.2 and make it available for operations by the end of 2020. Later, it is intended to test 4D-Var also at high (750m) resolution.

The MPI parallization of the code has been improved. Efforts on a more effective OpenMP parallization are ongoing. Studies on the use of 4D-Var in mixed precision will start soon.

On a longer term, an ensemble assimilation system is envisaged. Two alternative ensemble assimilation methods, an LETKF and a hybrid EnVar system, have been developed for this and introduced as options for experimentation. The EnVar system can make use both of observation perturbations and of perturbations in the B-matrix (so-called B_{rand} perturbations). For LETKF, longer inter-comparison experiments against operational 3D-Var are ongoing, with mostly positive results so far. Preliminary results for a short period look promising for hybrid EnVar with both LETKF and B_{rand} perturbations, but this work needs further elaboration. Both setups will be upgraded to the latest Harmonie pre-operational cycle (Cy46h). Following this, an inter-comparison will be performed to test hybrid EnVar/ B_{rand} and LETKF against each other and against standard 3D- and 4D-Var.

A3. Optimization of data assimilation setups for the nowcasting range

For nowcasting, one would like to assimilate the latest observations fast and frequently. However, limits are put on this by model spin-up in the first hour or two of the forecast. Above, the experiments to test whether 4D-Var configurations may be better able to handle spin-up have already been mentioned. The continuous or overlapping window cycling strategy used in COMECS (Yang, 2018) permits efficient use of the most recent observations in an ensemble in combination with relatively long assimilation windows. For 3D-Var, DMI has set up and begun testing a nowcasting version of its COMECS Harmonie ensemble system with six parallel suites with base times consecutively shifted by 10 minutes, short cut-off times and with hourly or two-hourly assimilation cycling of frequently available observations. A start has been made with creating a setup for sub-hourly cycling also in a 4D(En)Var framework.

At high spatial and temporal resolutions, it becomes increasingly important for the analysis system to be able to correct for position and phase errors of fine-scale atmospheric features. Present variational methods are not well versed in handling such non-additive errors. In earlier years, the field alignment (FA) method has been developed, by which displacement errors of e.g. the modelled precipitation or wind field with respect to radar 3D data can be identified and corrected before 3D-Var takes place. FA is mainly

beneficial in the nowcasting range, showing e.g. a significant reduction of errors in wind fields in the first 6 hours of the forecast. To reduce the problems of model imbalance in the nowcasting range, a new initialization formulation has been developed. The introduction of this so-called variational constraints (VC) method (Gejjo and Escriba, 2018) has been shown to achieve a better handling of analysis increments and a faster balancing of the model. VC and FA together appear to act in a complementary manner to enhance model performance in the nowcasting range.

A good description of the 3D structure of clouds is an essential ingredient for nowcasting. The cloud initialization (CI) method was developed to permit a more sophisticated use of NWC SAF cloud products (cloud type, mask and microphysics) from MSG and polar satellites to adjust 3D humidity in model analyses and nowcasts (Gregow, 2017). The CI scheme has been shown to have a good impact on clouds, but causes a slight degradation in precipitation and still suffers from some spin-up. It will be studied whether the latter can be reduced by applying CI before, rather than after, the assimilation.

B) Atmospheric forecast model

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

Coordinated experimentation has been performed in the past two years with the aim to understand the (multiple) causes of observed poor model behaviour for convection and low clouds: the timeliness and initiation of deep convection, the description of stratiform precipitation and open cell convection, too persistent fog in the model, a negative cloud frequency bias and a too high model cloud base in fog situations. Several clear avenues of improvement have been identified for improving the representation for convection, low clouds and fog.

One cause of problems with the model triggering convection too little and too late has turned out to be the unrealistic description of climatological evapotranspiration in ECOCLIMAP2.2, in combination with a too strong impact of the surface assimilation. A new physiographic database, ECOCLIMAP-SG, offers a significant improvement in this respect.

The initiation of supercell formation in the early evening sometimes does not happen with the present HARATU turbulence scheme. This appears to be caused by too much dry air being mixed by HARATU from above into the boundary layer, which may then have become too dry to permit convective initiation. A set of changes in the cloud and turbulence schemes, involving reduced mixing from the cloud layer, changes in the free atmospheric length scale, and several adaptations affecting the mixing of momentum and moisture near inversions, led to a strong and positive impact on the representation of deep convection and precipitation, but also of low clouds.

In stable boundary conditions, fog has long remained a significant problem. Generally, fog was shown to build up too quickly, have a too large extent (especially over sea), be too dense (too much cloud water), and cool too much over sea. Fog layers in the model already start to cool strongly when the fog layer is still quite shallow, much more than observed. An increase in model vertical resolution in the boundary layer did not alleviate this problem. Detailed well-observed case studies showed that fog evolution is extremely sensitive to the amount and evolution of cloud condensation nuclei (CCN), and to the way in which the cloud emissivity is formulated in the long-wave radiation parametrization. The best solution will be to combine the use of near-real-time CCN density values derived from CAMS, with the use of the aerosol parametrizations and the new second-moment microphysics scheme LIMA which permits CCN densities to evolve. This combination will become available for testing in Cy46h1 in 2021. In the present ICE3 microphysics scheme, fixed values for CCN densities are assumed over sea, land and urban areas. Using this scheme, fog problems can be reduced strongly by a drastic reduction of the fixed values of CCN density, in agreement with observations, in combination with a change in cloud emissivity in the long-wave radiation parametrization. Due to the thus diminished overprediction of cloud water in the model, less fog occurs over sea, and it no longer quickly grows and spreads (e.g. fig.2). The proposed changes in CCN density values also result in removing the present negative bias in SW radiation, and in much better behavior of precipitation in e.g. the Norwegian coastal and mountainous areas. However, their impact on other aspects of the model still needs to be carefully checked. Work on that has started.

A third weak point is the balance between “cold” and “warm” hydrometeors in mixed-phase clouds and the way in which they interact with the development of convection. In both 1D and 3D experiments, it was shown that the ICE3 microphysics scheme lets far too much solid precipitation form in the form of graupel and too little in snow. This may lead to an underestimation of cold pool formation. Elements of the Thompson microphysics scheme have been implemented in the standard ICE-3 scheme. This leads to a more realistic division of precipitation between snow and graupel, more supercooled liquid water and less ice, and

more production of snow and freezing drizzle. It also provided a more realistic description for the stratiform precipitation which is formed from the dispersion of the anvil in the dying-out phase of the deep convection life cycle, and a better representation of open cell convection and precipitation in mountainous areas.

Another long-standing problem has been the underestimation of temperature minima at night during very stable boundary layer conditions. This behaviour is very sensitive to the value adopted for the maximum Richardson number. Optimizing this value leads to more accurate night time temperature minima under stable to very stable conditions, and improvements have also been seen for fog.

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

Work has continued on quantifying the benefits (and costs) of using aerosol parametrizations in combination with near-real-time aerosol information from CAMS on clouds and radiation. Experiments are being carried out with cloud-radiation-aerosol interactions for a variety individual aerosol types, using prescribed daily CAMS global 137 level aerosol fields. Cy43h2 has been adapted to include NRT 3D mixing ratio aerosol fields from CAMS in the cloud and microphysics schemes, and in the hlradia radiation scheme: three types of sea salt, hydrophilic black carbon, organic matter, sulfate, nitrate and ammonium. Aerosol wet deposition has been included.

For the radiation scheme, CAMS mass mixing ratios for the individual aerosol types are combined with prognostic model specific humidity fields for each time step, grid point and level, with prescribed inherent optical properties. From these, e.g. the optical depth AOD and transmission are calculated for the aerosol mixture. Sensitivity studies have been done for the simple hlradia scheme, showing clear impact on short-and longwave radiative fluxes. Work remains to be done to include the inherent optical properties and mass mixing ration CAMS fields also in the radiation schemes IFSradia and ACRANEB2.

Rather than assuming cloud condensation nuclei concentrations to be constant, the microphysics scheme has been updated to calculate CCN concentrations from the CAMS mixing ratio fields assuming log-normal size distributions. These are then used in the auto-conversion, cloud droplet sedimentation and cloud liquid collision processes. One month of verification (September 2019) has been carried out with and without aerosol parametrizations in the radiation and microphysics schemes. The ETS of precipitation and the bias in RH2m improved, at the cost of a degradation in T2m bias as function of forecast length. This may be due to the fact that the dependence of the mass extinction of hydrophilic species on humidity was originally ignored. Including this dependence bring the global radiation and direct normal irradiance in closer agreement to observations. Studies over Norway have shown that including the aerosol forcing from NRT CAMS fields leads to a much improved representation of coastal and mountain precipitation, reducing the customary coastal dry bias and the wet bias over the mountains significantly.

The use of aerosol information in this way increases the cost of Cy40h runs with 14%. This will need to be confirmed for Cy43h2.1 as well, and work will need to be done with the aim to optimize the aerosol code further (e.g. by not using the CAMS fields for every time step).

B3. High (sub-km) resolution modelling

At present, the Harmonie forecast system is generally operationally run at 2.5km horizontal resolution and with 65 layers in the vertical. Efforts to investigate and optimize model performance at higher horizontal (300-750m) and vertical (90L) resolution are ongoing. A workshop on high resolution modelling and nowcasting was held in Dec 2019 in Las Palmas. The main outcome of this meeting was a recommended setup of the model for sub-km resolution and for nowcasting: dynamics settings, manner of nesting, vertical resolution increase, and the use of continuous (overlapping windows) assimilation. Several staff will start work on implementing this setup in the second half of this year. A critical aspect for all partners is the quality of the surface characterization (orography and physiography), and the need to optimize this using local high-resolution databases. It will require a significant effort to do this consistently across European borders. Some issues with climate generation and the use of sub-hourly output remain to be solved in the short term. More efforts will be required to e.g. arrange for sub-hourly input (observations and boundaries), and to determine the proper scales of the B-matrix corresponding to the model resolution. The use of single precision in high resolution models is still under investigation.

Tornadoes are a phenomenon which seems to be occurring increasingly in Europe. At 2.5km resolution, Harmonie is not able to adequately represent such fine-scaled features. However, in 1km resolution runs over the US, it was shown that Harmonie is able to represent the typical “hook echo” structure which is characteristic for the larger-scale flow around tornadoes. Updraft helicity, which is known to be a good indicator for tornado occurrence, will be included as a parameter in the model post-processing.

C) Surface analysis and modelling

C1: Improving the sophistication of surface model components

For surface modelling, the Surfex system of parametrizations is used in Harmonie.

A critical aspect for surface modelling is the quality of the surface characterization, and the need to optimize this at sub-km scales using local high-resolution databases in a consistent manner. In Surfex v8.1, a new and spatially more detailed (resolution 300m) version of the physiographic database, ECOCLIMAP-Second Generation, has become available. ECOCLIMAP-SG contains physiographic information which has been more directly obtained from ESA/CCI satellite observations. Much effort has been done in the past year to assess the strengths and weaknesses of ECOCLIMAP-SG and the consequences of applying it for model performance. The main characteristics of ECOCLIMAP-SG in comparison to its predecessor ECOCLIMAP-v2 have been described in Samuelsson et al. (2020). The description of coastal and urban areas and the yearly cycle of evapotranspiration are clearly represented in a more realistic manner in ECOCLIMAP-SG, although some weaknesses have been identified as well, which partly remain to be addressed. It has therefore been decided to adopt ECOCLIMAP-SG as the new physiographic database in Cy43h2. The significant differences between the old and new databases made it necessary to assess very carefully for which aspects of the surface model parameters retuning was appropriate, and to optimize those parameters to use in combination with ECOCLIMAP-SG (e.g. modified heat capacity, maximum Richardson number, stomatal resistance and the vegetation roughness formulation). In addition, ECOCLIMAP-SG can still benefit from quite many local adaptations, and introducing these will require a significant effort also in the coming years.

The ECOCLIMAP-SG validation efforts over various have highlighted the issue that present Monin-Obukhov similarity theory falls short in describing the complexity of heterogeneous canopies. In order to deal with this, alternative diagnostic formulations for screen level parameters and flux formulations from roughness sublayer theory (Hartmann and Finnigan 2007) are being investigated.

The new Soilgrids sand/clay database and updated Global Lake Database (GLDB) have been tested successfully and will become default in the summer of 2020.

Operationally, use is still made of a 3-layer fore-restore soil scheme and a simple snow module. Experiments with the new many-layer diffusion soil, extended snow and snow-over-vegetation schemes have shown the added value of these more advanced schemes. However, their introduction needs to be done in combination with more advanced schemes for soil and snow assimilation, which are still being worked on (see section C2 below). This combined introduction is in preparation, and is planned to be achieved in 2021.

The sea ice scheme SICE has been extended with sea ice drift. Updates have been made in the FLAKE lake model and the GLDB lake database. An orographic parametrization for turbulence, OROTUR, has been implemented. Also, a parametrization describing the creation of turbulent wakes in and behind wind farms over sea, based on the method of Fitch et al. (2012), has been introduced and validated successfully against floating lidar and platform observations (fig.3). Also, required metadata on wind turbine and farm characteristics have been collected for most wind farms present and planned on the North Sea until 2030.

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

One of the main objectives in HIRLAM-C is to replace the very unsophisticated OI surface analysis system with more advanced assimilation methods, in combination with using a wider range of satellite surface observations. The focus in the past year has been to finalize preparations for the introduction of a set of simplified extended Kalman Filters (SEKF's) for soil, snow, and sea ice. As described above, these SEKF's will be used in combination with the new diffusion soil and extended snow schemes. The introduction of the SEKF assimilation schemes has been delayed to 2021 due to the need for retuning surface model components caused by the introduction of ECOCLIMAP-SG. Also, for the soil SEKF several aspects are still under study, such as a new set of control variables. Also, the sensitivity of the surface analysis to e.g. settings for cycling and assimilation window lengths is being checked. For the horizontal spatialization, the new gridpp OI scheme will replace the present CANARI scheme, because of its more modern and compact code base, and its proven ability to handle a variety of crowd-sourced data from e.g. private weather stations.

In parallel, some research efforts have focused on the exploration and development of more powerful envisaged future assimilation methods such as EnKF and particle filters. For experimentation, an EnKF is already being run with AMSR-2 and SMOS radiance observations. In the coming years, the design of a coupled surface-atmosphere data assimilation system will be considered.

In addition to conventional near-surface observations, crowdsourced data and satellite observations will progressively be introduced. Positive impact has been seen from the assimilation of crowd-sourced data from private weather stations (Bremnes et al. 2019) and smart phones (Hintz et al. 2019). Presently, the bias correction of these data is still being investigated. Experiments with assimilating a variety of snow products from e.g. H-SAF and CryoClim showed positive impact, primarily during the melting season. Impact studies have already been performed or are ongoing involving ASCAT, SMOS, MODIS, and H-SAF retrieval products of soil, vegetation, sea surface and inland waters properties and snow- and ice-covered surfaces. These will be added progressively to the surface assimilation after the introduction of the SEKF schemes. Following this, the aim is to also include raw satellite radiances and Sentinel-S SAR data in the analysis of soil moisture, soil temperature and snow.

D) Probabilistic forecasting:

A scientific description of the HarmonEPS system, its various implementations and its options in detail, has been published (Frogner et al, 2019). Work has continued on assessing a variety of perturbation methods in model physics, surface, initial and boundary conditions for HarmonEPS. Also, the HarmonEPS system has been upgraded from Cy40h to Cy43h2.

Problems have arisen in the SPPT model perturbations. When using only SPPT perturbations, experiments showed appreciable impact and spread; however, when combining them with other perturbation types, their impact becomes negligible. Many experiments have been, and are being, done to discover the cause of this. For stochastic parameter perturbations (SPP), work has continued to further increase the number of physical parameters, and to study the effect of using different distributions and spatial and temporal scales for different parameters. However, due to the need to solve the SPPT issue mentioned above, this work has lately been temporarily stalled.

The continuous cycling approach originally developed at DMI (Yang 2018) has shown to be a powerful instrument to enable increased (double or even triple) ensemble size, as well as earlier delivery of ensemble products and less jumpiness between ensemble runs. It is gradually being introduced in most operational HarmonEPS setups. In the coming year, more detailed inter-comparisons will be made for the three mechanisms available for generating initial condition perturbations: EDA (presently default), LETKF and hybrid EnVar/Brand, after the hybrid EnVar and LETKF development branches have been ported to the most recent pre-operational cycle, Cy46. Also, the optimal number of ensemble members using EDA will continue to be studied.

In the past months, the work on surface perturbations has been intensified by new staff members. Experiments with perturbing surface fields like SST, soil moisture and LAI have shown to lead to improved spread for near-surface parameters, but to have little impact at higher levels in the atmosphere.

A high-priority target is to achieve affordable high (sub-km) resolution nowcasting ensembles, which are computationally highly demanding considering the tight time constraints for nowcasting. One important element for achieving this which has been experimented with extensively, is the continuous cycling approach. Another element which has been studied is the replacement of the linear grid by quadratic or cubic grids, which are computationally 30-50% more efficient, at the loss of some effective resolution.

E) Code efficiency, scalability and portability

Various code adaptations have been made to permit use of single precision for Harmonie-Arome, on top of the work done by ECMWF (Vana et al.) and Meteo-France (Marguinaud) for the global models. The option of single versus double precision is being tested in several services. Both the Reference 3D-Var and forecast model can be run mostly in single precision, which has so far been tested successfully for almost all available observation types. For most variables, verification scores are indistinguishable between single and double precision, with the exception of a positive bias in mean sea level pressure increasing with forecast length and a slight tropospheric cooling in the single precision runs. The cause for this bias and cooling behaviour is still under investigation. 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). Work has started to test the 4D-Var system in mixed precision.

Efforts have continued to improve the computational efficiency of 4D-Var through a better use of MPI and OpenMP parallelization. The Barcelona Supercomputing Center has been performing a general assessment of the computational efficiency of Harmonie. Both at the code and compilation level, a number of clear potential gains have been identified. In particular, the OpenMP implementation in Harmonie is

confirmed to be suboptimal. This first assessment will be followed up with an optimization of the efficiency of Harmonie on the existing and new ECMWF platforms.

Given the increasing heterogeneity of HPC architectures, it is important to establish model performance on a range of different architectures. At KNMI a first investigation has been made on the performance of Harmonie on a cluster of ARM chips, the advantage of ARM architectures being their high energy efficiency. Harmonie could be made to work on an ARM cluster without any recoding, at the cost of somewhat greater time-to-solution but also significantly less energy-to-solution.

DMI staff have managed successfully to build the necessary components of the Harmonie suite onto a container, with a view to test the possibility to run Harmonie in the “cloud”. Independently, staff from Met Norway managed to port the MUSC single column model code onto a virtual machine. With such steps it is aimed to make the Harmonie system more portable and usable to a wider community for research purposes.

Figures:

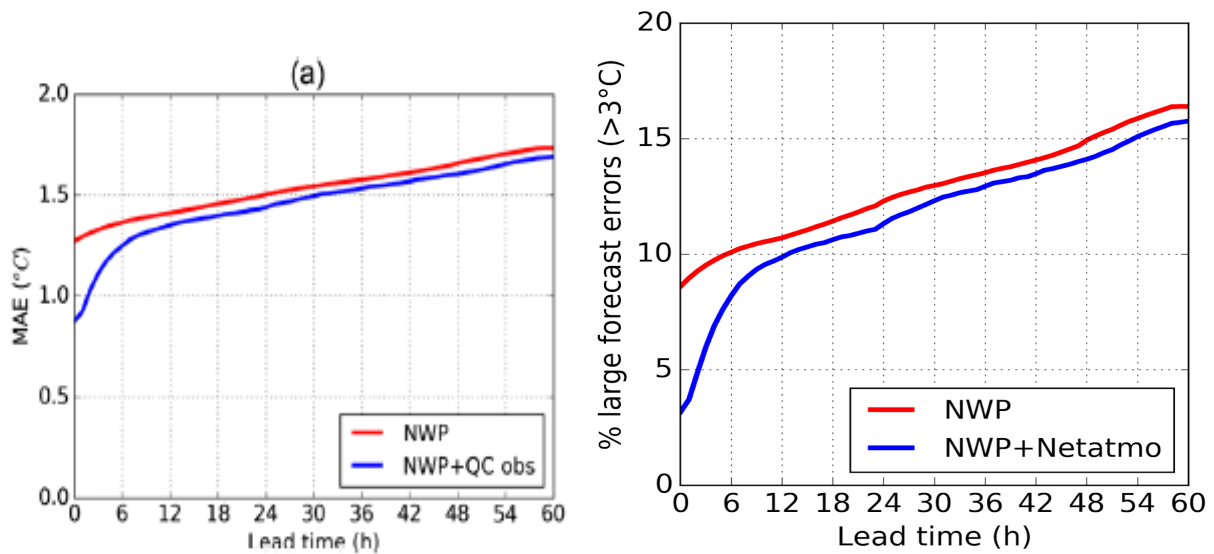


Fig. 1: Impact of the inclusion of quality-controlled Netatmo private weather station data in the postprocessing of Harmonie forecasts of T2m for Norway (taken from Nipen et al. 2019). Normally, observations from the observation network of Met Norway are used for statistical adjustment of Harmonie forecasts and thus obtaining a so-called “gridded truth”. In the experiments by Nipen et al., quality-controlled Netatmo data have been added to those of Met Norway’s own network. The left figure shows the mean absolute T2m error as a function of lead time, the right figure the percentage of large forecast errors (>3 degrees) vs lead time. The blue curves represent the model results when quality-controlled Netatmo are included in the postprocessing of Harmonie forecasts, on top of the observations from the Met Norway surface network only (red curves). The Netatmo data provide clear added value in the post-processing throughout the forecast range, but particularly in the first 6 hours of the forecast. Experiments in which they are assimilated directly into Harmonie are presently underway.

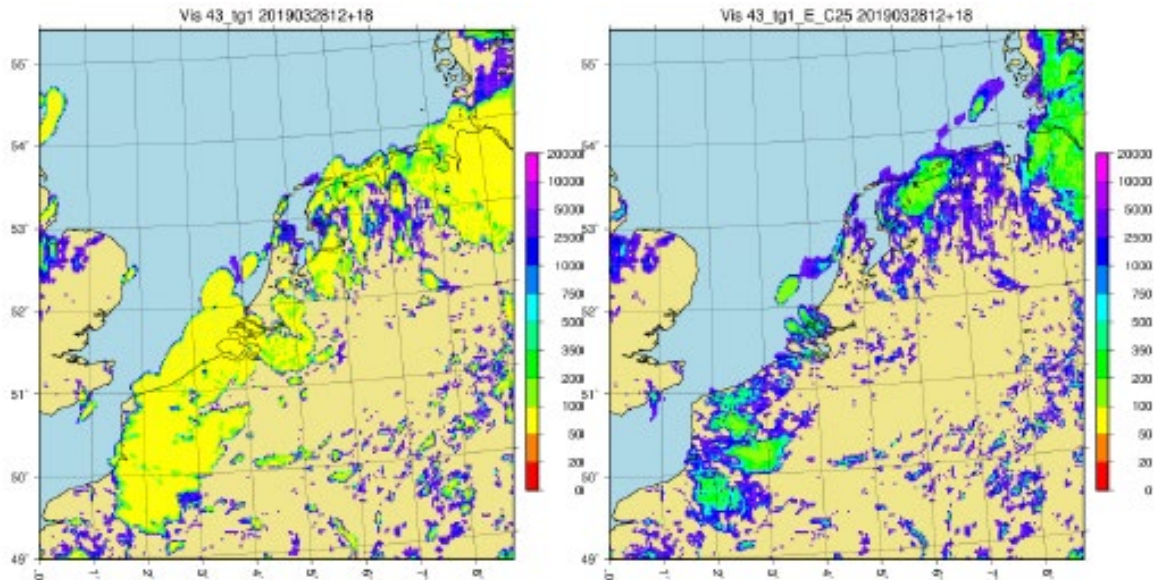


Fig.2: Improvement of fog forecasts through changes in cloud emissivity and CCN density as seen in a testcase on 28 March 2019 over the North Sea. In this test case, the model showed the rapid formation of dense fog over a large area above the North Sea, which did not occur in reality. Both the thickness and extent of the fog in the model grew fast, in contrast to what was seen in satellite imagery. The left figure shows the +18h visibility forecast from the Harmonie pre-operational model Cy4tg1, in which the erroneous fog over the North Sea is clearly present (yellow-coloured structure). The right figure shows the same forecast but with reduced cloud emissivity in the long-wave radiation parametrization and a reduction in the CCN density in the model. In this configuration the fog formation over sea is far less extensive, and it no longer grows quickly in horizontal extent and thickness after its original formation, in agreement with observations.

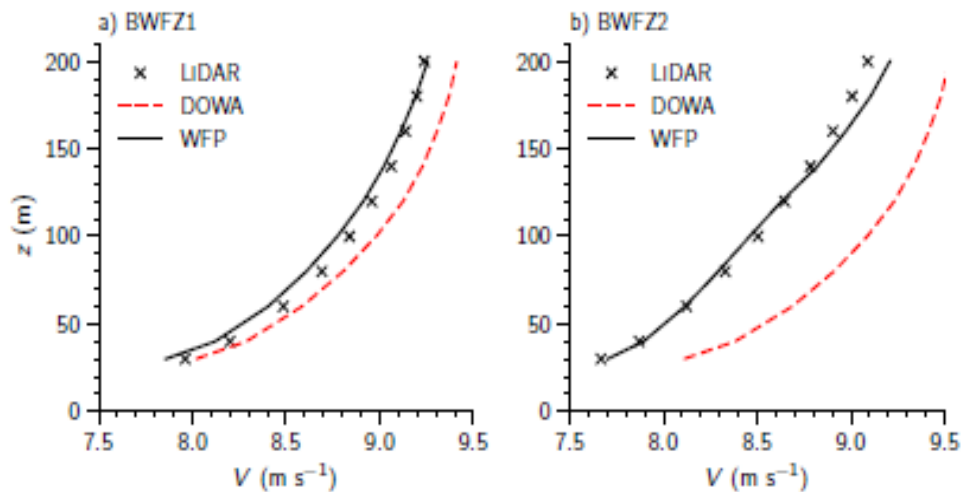


Fig. 3: Verification of the impact of the Fitch et al. 2012 wind farm parametrization (van Stratum et al. 2019) on Harmonie boundary layer wind profiles in the area around the Borssele wind farm. The left and right picture show the vertical wind profiles from the model for in comparison to observations from two nearby lidars BWFZ1 and BWFZ2, downstream of the wind farm in cases of the usually south-westerly wind flow for that area. The red dotted curve represents profiles from the model without wind farm parametrization, the black curve the profiles for the model including wind farm parametrization under south-westerly flow conditions. The crosses represent the lidar observations. Use of the wind farm parametrization clearly gives a more realistic representation of wind profiles in the boundary layer downstream of the wind farm, and more accurate energy predictions (not shown here). The spatial impact of the parametrization on the wind field can extend well beyond the immediate vicinity of the wind farm (order several tens of km, not shown here).