REQUEST FOR A SPECIAL PROJECT 2021–2023

MEMBER STATE: Netherlands

Principal Investigator[1]: Jeanette Onvlee

Affiliation: KNMI

Address: Utrechtseweg 297,
P.O.Box 201, 3730 AE De Bilt
Netherlands

E-mail: onvlee@knmi.nl

Other researchers: ~80 researchers from the HIRLAM countries.

Project Title: HIRLAM-C phase 3 (2021-2022) special project

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** sehlam__________________	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2021	
Would you accept support for 1 year only, if necessary?	YES X☐	NO ☐

Computer resources required for 2021-2022: (To make changes to an existing project please submit an amended version of the original form.)		**2021**	**2022**	**2023**
High Performance Computing Facility	(SBU)	40 MSBU	40 MSBU	
Accumulated data storage (total archive volume)[2]	(GB)	22.000	22.000	

An electronic copy of this form must be sent via e-mail to: *special_projects@ecmwf.int*

Electronic copy of the form sent on (please specify date): 22/6/2020

Continue overleaf

[1] The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide an annual progress report of the project's activities, etc.

[2] If e.g. you archive x GB in year one and y GB in year two and don't delete anything you need to request x + y GB for the second project year.

Principal Investigator: Jeanette Onvlee

Project Title: HIRLAM-C phase 3 (2021-2022)

Extended abstract

The HIRLAM-C research programme is a continuation of the research cooperation of the previous HIRLAM projects. The members of HIRLAM-C are 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 forecasting system Harmonie, and its associated ensemble prediction system HarmonEPS, in particular to enhance their quality for the accurate prediction of high-impact weather at a target horizontal resolution of 0.5-2.5km. The Harmonie system is being developed within the IFS code framework in close cooperation with the Aladin consortium.

Following the past Hirlam practice, a Harmonie Reference system is being maintained on the ECMWF HPC platform. The emphasis in the requested HIRLAM-C phase 3 Special Project at ECMWF will be primarily on experimentation, evaluation and testing of the Harmonie Reference System, and its ensemble configuration HarmonEPS. The computational resources from the Special Project will be used mainly to experiment with newly developed model, assimilation and ensemble generation components and to evaluate their meteorological and technical performance in beta-releases, before releasing them as Reference. In-depth validation and intensive (pre-)operational testing of these developments will be carried out both in the member institutes and at ECMWF.

The HIRLAM-C programme was originally scheduled to run from January 2016 until December 2020, but in the past months it has been decided to extend the programme with two years until end 2022. The present request is for a Special Project for the duration of this extension of the HIRLAM-C period, so for a third phase for the years 2021-2022. Below, the main foreseen research activities on data assimilation, the forecast model, surface analysis and modelling, probabilistic forecasting and code efficiency and scalability (undertaken jointly with Aladin partners) from early 2021 until the end of 2022 are outlined.

A) Data assimilation developments:

The following activities are foreseen:

<u>A1: Enhanced use of high-density observations:</u>
At the beginning of the HIRLAM-C period, the Harmonie Reference system used as default a 3D-Var assimilation system, routinely assimilating conventional data and radiances from AMSU-A, AMSU-B and MHS. During the years 2016-2020 several types of spatially and temporally dense observations have been added to the assimilation system either operationally or optionally: radar radial wind and reflectivity volume data, GNSS ZTD delays, radiances from MetOp, FY-3C, FY-3D and IASI, SEVIRI, geostationary and polar atmospheric motion vectors, Mode-S EHS observations, winds from several scatterometers, ADM/Aeolus HLOS winds, and observations from private weather stations. In addition, much work has been done to optimize the impact of these high-resolution data through e.g. improved data quality control, more intelligent thinning and super-obbing strategies, bias correction, and the introduction of more appropriate (fine-scale) structure functions.

In 2021 and 2022 one aim will be to continue to improve the availability, quality control, bias correction, and use of existing data sources. It will be attempted e.g. to make better use of radar radial wind data and a wider use Mode-S aircraft data through enhanced quality control, improved preprocessing and wider data sharing. In order to take into account the footprint of satellite data with respect to the model resolution, the use of the so-called supermodding technique will be extended from scatterometers (Mile et al., 2020) to other types of remote sensing data such as radiances.

Secondly, work will continue to prepare for the introduction and assessment of several new types of observations: experiment with various approaches to assimilating all-sky radiances; continue with impact and quality control studies for GNSS slant delays; add a wider range of already existing but presently unused microwave observations; do preparations for new satellite platforms and instruments such as MTG and EPS-SG; study the use of polarimetric radar information; and try to make more boundary layer lidars and ceilometers available for use.

In the third place, crowd-sourced observations from e.g. amateur weather stations, smart phones, road networks and wind turbines, are increasingly being proven to add significant value to NMS observational networks, particularly in urban areas. These data may provide an important contribution especially to high resolution nowcasting model setups, provided that sufficient attention is paid to their collection, quality control and bias correction. Ways to better employ, homogenize and share such data will be considered. For quality control, machine learning techniques are being utilized and developed.

<u>A2: Development and optimization of flow-dependent assimilation: 4D-Var, hybrid 3/4DEnVar</u>

Until now, the impact of new high-resolution observations in the Harmonie 3D-Var configurations has generally been positive but limited in both size and duration. There are many indications that this is at least partly due to inherent limitations of the 3D-Var method. On the algorithmic side, one of the main aims for HIRLAM-C therefore has been the introduction of more sophisticated flow-dependent data assimilation methods. A computationally efficient multiple outer loop Harmonie 4D-Var system has been developed which can be run together with the full range of observations that is used in the operational Harmonie 3D-Var suites. With conventional observations plus AMSU/MHS radiances, 4D-Var in 3-hourly cycling has already been shown to outperform 3D-Var. The system is expected to enter operations from early 2021 onwards. In the coming years, impact and optimization studies for 4D-Var will continue for a wide range of observations. Experiments will be performed using e.g. different time windows and an extended set of control variables. Also studies will be done to further assess and improve 4D-Var's usefulness in the nowcasting range (e.g. using more frequent cycling and the overlapping windows technique developed by Yang et al. (2017)), as well its computational performance.

The next step forward will be to combine the strengths of ensemble and variational approaches, to the mutual benefit of the model analysis and its probabilistic forecasting ability. For the next few years, the primary aim in data assimilation algorithmic development is to continue to build and implement a flexible algorithmic framework for 3- and 4D ensemble variational assimilation (3/4DEnVar) for Harmonie, suitable for both assimilation and ensemble forecasting purposes. In the past years, a hybrid 3D-EnVar system and an LETKF system have been developed. Sensitivity studies to validate and find optimal settings for both these systems have taken place in the past few years. In the coming years, the EnVar and LETKF prototypes will be recoded in order to adhere to the new OOPS structure; also, the two methods will be intercompared with each other, and against 3- and 4D-Var, to assess their added value in both meteorological and computational terms.

<u>A3. Development of data assimilation suitable for the nowcasting range</u>

For nowcasting purposes, in the first years of HIRLAM-C experiments were performed with the 3D-Var system in rapid update cycling mode, at update frequencies of 1h and less, using observations such as radar, GNSS and locally processed AMV's with very short cutoff times. Spin-up effects have been studied, as well as the impact of these rapidly updating systems with the respect to less frequent updating with a more sophisticated technique such as 3-hourly 4D-Var. The overlapping windows or continuous assimilation technique developed by Yang et al. (2017) has enabled a more efficient use of the most recent observations, increased ensemble size, earlier delivery and less jumpiness between ensemble runs, while still permitting relatively large observation windows. A start has been made with the use of continuous 3D-Var assimilation in nowcasting mode: using nowcasting ensembles consisting of suites with base times consecutively shifted over 10-15 minutes, with short cutoff times and with 1-2-hourly cycling of frequently available observations. In the coming two years, the assessment of such nowcasting ensembles will be continued with both 3- and 4D-Var. Also, it will be considered how to best tackle the issue of correlated observations in this time range.

At high resolutions, it becomes increasingly important for the analysis system to be able to correct for position and phase errors of fine-scale atmospheric features in the model as compared to satellite and radar imagery. Variational assimilation methods are not well versed in handling such non-additive errors. In earlier years, a so-called field alignment (FA; Geijo 2012) technique has been developed, by which displacement errors of the model first guess field relative to a radar wind or reflectivity field can be identified and corrected before a normal 3- or 4D-Var takes place. FA is mainly beneficial in the first 3-6 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 (Geijo and Escriba 2018) has been shown to achieve a better handling of analysis increments and a faster balancing of the model. In experiments, FA and VC appear to act in a complementary manner to enhance model performance in the nowcasting range. In the coming two years, the FA and VC techniques will be combined and tested in pre-operational mode over longer periods.

A cloud initialization (CI) technique has recently been introduced to permit the use of NWC SAF cloud products (cloud type, mask and microphysics) from MSG satellite imagery and polar satellites to adjust 3D humidity in variational model analyses and nowcasts (Gregow, 2017). In the coming years, research will be done on the performance and further optimization of this scheme in combination with 3-and 4D-Var, and with the FA-VC scheme.

B) Atmospheric Forecast model

The following activities are foreseen:

<u>B1: Studies to eliminate systematic model errors for cloud and boundary layer behaviour.</u>

In the past years many studies have been performed with the aim to understand and eliminate certain systematic model biases, for (low) clouds, fog and visibility, the timeliness and initiation of severe convection, the description of stratiform precipitation and open cell convection, in surface fluxes and under winter-time stable boundary layer conditions. Several clear avenues of improvement have been identified for improving the representation for convection, low clouds and fog, which will be pursued in the coming years. One element playing a role here has been the poor description of evapotranspiration of vegetation in the physiographic database, leading to too great drying out of soil in the growing season and too little triggering of summer convection. An improved new version of this database has been implemented to solve this problem. Adaptations have also been made in the turbulence, shallow convection and cloud schemes to improve the triggering of supercell formation and reduce the bias for low clouds.

One issue which has been addressed but which remains to be completed in the coming years is the overprediction of the extent and depth of fog, particularly over sea. Detailed well-observed case studies have shown that the fog evolution in the model 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 present ICE3 bulk microphysics scheme, with its fixed climatological values for CCN over sea, land and urban areas, is unable to adequately represent the swift CCN evolution occurring during fog formation. However, it has been shown that fog problems can still be reduced significantly by a reduction of CCN densities in combination with a change in cloud emissivity. The impact of these changes on other aspects of the model is still under investigation. A more permanent 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 experimentation in 2021.

In the past years, several other shortcomings in the ICE3 microphysics scheme have been detected, for which possible solutions have been prepared which will need to be tested more thoroughly in the coming two years. Hopefully these developments will positively affect the model description of supercooled liquid water, snow and freezing drizzle, open cell convection, and the occurrence and evolution of small-scale showers. It will also need to be assessed how the new second-moment LIMA scheme behaves in comparison with ICE-3 under such circumstances.

A key goal of HIRLAM-C has been to achieve a more realistic, internally consistent treatment of the radiation-cloud-microphysics-aerosol description in the model. In the past years, aerosol parametrizations have been developed for radiation and clouds, for a variety of aerosol types (sea salt, hydrophilic black carbon, organic matter, sulphate, nitrate and ammonium) and including aerosol wet deposition. A main target for the next two years will be to combine the more realistic initialization of CCN and the aerosol parametrizations with the second-moment microphysics scheme LIMA, which allows evolution of the CCN concentrations in the model, and compare the performance of this combination with that of the present model for a wide range of weather conditions. As the computational cost of both the aerosol parametrizations and the LIMA scheme is considerably higher than in the present model, attention will also need to be paid to achieving greater computational efficiency for the new algorithms.

Another remaining target is to improve model stable boundary layer behaviour, in particular for low-wind stable winter conditions over a heterogeneous surface. This will be done through e.g. assessment of the influence of increasing model vertical resolution in the boundary layer, alternative turbulence formulations, experimentation with the choice of lowest model level height, the treatment of displacement height in the surface scheme (e.g. by the application of a roughness sub-layer at the top of the canopy (e.g Raupach et al. 2012, Hartmann and Finnigan 2007)) and alternative diagnostic formulations for levels below the lowest model level.

Finally, it is aimed to assess the ECRAD scheme from ECMWF in Harmonie, in comparison to the existing default Harmonie radiation scheme and the ACRANEB2 scheme developed by LACE.

<u>B2: Preparations for increased vertical and horizontal resolution</u>

At present, the Harmonie forecast system is typically run operationally at 2.5km horizontal resolution and with 65 layers in the vertical. In the past few years, much work has been done to prepare the model for use at higher vertical (~90 levels) and horizontal resolution (~300-750m). On the basis of the experiences with such model configurations for a wide variety of domains, a recommended setup for Harmonie at resolutions of 500-750m has been formulated, which is being implemented in the Reference system: dynamics settings, manner of nesting, the treatment of horizontal diffusion, etc. In the coming years, more attention will be given to the (limits of) applicability of the physics parametrizations for turbulence and shallow convection. In cooperation with ALADIN partners, several approaches for (quasi-)3D formulations for turbulence and radiation will be investigated. Model performance will be compared with that of LES models, field studies and high-resolution observation networks. In regions with steep orography, the impact of use of sub-grid orographic parametrizations for radiation and momentum at increased spatial resolution remains to be assessed in depth. In urban areas, various existing but yet unexplored options of the urban TEB (town energy budget) surface module will be studied. Where available, appropriate (local) high-resolution physiographic information will be introduced for an accurate description of the surface at hectometric resolutions.

The aim is to develop data assimilation and ensemble setups suitable for these high resolution models. It will therefore be investigated which (nowcasting) data assimilation setups and which observation types are most suitable to initialize the frequently run hectometric-scale models. For high-resolution ensembles, the so-called overlapping window technique developed by Yang et al. (2017) will be explored. In all of these studies, it will be considered what will be the optimal way to spend available computational resources, seeking the best balance between horizontal and vertical resolution, domain size, ensemble size and model complexity.

C) Surface analysis and modelling

The following activities are foreseen:

<u>C1: Enhanced use of satellite surface observations with more advanced surface assimilation</u>
For the surface analysis, the Harmonie Reference System presently uses a simple OI approach to assimilate conventional screen level observations. It is aimed to use a much wider range of relevant remote sensing surface observations in the surface analysis; however, this requires application of a more advanced assimilation method. In the past few years, a new surface assimilation framework has been constructed consisting of a set of Simplified Extended Kalman Filters (SEKFs) for soil, snow, lakes and sea ice assimilation. The primary goals for the coming few years will be to implement this set of SEKF's operationally in the course of 2021, and then to start progressively exploiting a much greater variety of remote sensing observations, from e.g. ASCAT, SMOS, MODIS, CryoClim, Sentinel and H-SAF products for soil moisture, LAI, surface temperature, lake water temperature and ice fraction assimilation, sea ice and snow cover and depth. Experiments to assess the potential of these satellite surface observations are already ongoing, and will be continued at greater intensity in the coming two years.

In parallel, some research efforts will focus on the exploration of more powerful future assimilation methods such as EnKF and particle filters. For experimentation, an EnKF is already being run with AMSR2 and SMOS radiance observations. In the coming years, the experimentation with EnKF is likely to intensify. Also, ideas will need to be developed, together with ALADIN partners, how to achieve a more consistent coupled atmosphere-surface ensemble assimilation system.

<u>C2: Introduction of more advanced surface model descriptions</u>
For surface modelling, the Surfex system of parametrizations is used.
Operational experiences have shown that the quality of the physiographic databases used in the surface modelling is of critical importance to model performance. In Surfex v8.1, a new and spatially more detailed (300m resolution) database called ECOCLIMAP Second Generation (ECOCLIMAP-SG) has become available. This database, derived from ESA/CCI observations, in many aspects is an improvement over its predecessor. However, the significant differences between the two databases have made it necessary to assess carefully for which part of the surface modules retuning was appropriate when switching to ECOCLIMAP-SG. In particular many parameters of the soil scheme have had to be adapted for optimal use with ECOCLIMAP-SG, with new settings for e.g. stomatal resistance, heat capacity, and the vegetation roughness formulation. The new database will be introduced in the Harmonie Reference system in July 2020. Also after its introduction, some necessary local adaptations in ECOCLIMAP-SG will still need to be made in the coming years.

Activities in the past years have focussed on developing a better handling of vegetation- and snow-covered surfaces, a more accurate description of forests, extensions of the lake model with ice and snow over ice parametrizations, and the development of a sea ice model, a sub-grid scale orographic parametrization for radiation and a wind farm parametrization. Much effort has also been spent on assessing more sophisticated multi-layer soil and snow schemes which have been developed by the Surfex community. All of these new modules have individually been clearly shown to be beneficial. Now they need to be tested, implemented and possibly retuned in combination with ECOCLIMAP-SG, with each other, and also in relation to the new set of SEKF assimilation schemes. The implementation and joint testing of the new modules and the corresponding assimilation algorithms will be the main focus of surface developments for 2021.

Following this, studies can take place with new but as yet unexplored options within Surfex, such as prognostic LAI from the A-gs vegetation scheme or various options in the TEB town scheme. New roughness formulations will be considered, in order to take better into account fluxes for heterogeneous surfaces (e.g. roughness layer theory, Hartmann and Finnigan 2007). Experiments are also foreseen in coupling Harmonie/Surfex with an ocean model (NEMO), wave models (WaveWatch3, WAM) and hydrological models (TRIP) using OASIS couplers, but these will probably not take place in the context of this special project.

D) Probabilistic forecasting

Most HIRLAM institutes presently run operational or pre-operational versions of the ensemble configuration of the convection-permitting Harmonie system, HarmonEPS. HarmonEPS, is a flexible system with a range of possibilities to describe uncertainties in different parts of the NWP system.

To account for the uncertainties in the initial conditions, presently EDA is used by default to account for initial uncertainty by perturbing the observations and running separate assimilation cycles for all members. Activities on optimization of the EDA setup will continue. However, in addition, initial condition perturbations can also be generated with a hybrid 3DEnVar (in the B-matrix, socalled Brand perturbations (Frogner et al. 2019)) and an LETKF system. In a close cooperation between EPS and data assimilation staff, a detailed inter-comparison will be carried out between the hybrid 3DEnVar/Brand and the LETKF systems, and their impact on both data assimilation and EPS performance. Also, both systems will be compared against 3D-Var and 4D-Var, in terms of both data assimilation and ensemble performance aspects.

The continuous cycling or overlapping windows approach originally developed at DMI (Yang 2018) has shown to be a powerful instrument to enable increased (double or even triple) ensemble size to ~20-30 members, as well as earlier delivery of ensemble products and less jumpiness between ensemble runs. It will become the default way to run HarmonEPS ensembles by end 2020. Experiments in the coming years will focus more on applying and assessing the continuous cycling approach to sub-km resolution nowcasting ensembles.

For the representation of model error, both the SPPT scheme (Buizza et al, 1999) for perturbation of tendencies and the SPP scheme for perturbing parameters are available. Both have shown some positive impact, and it appears that added value can be gained by combining them. The SPG pattern generator by Tsyrulnikov and Gayfulin (2017) presently is the default in HarmonEPS. In the SPPT scheme, presently only total physics tendencies are used. For the SPP scheme, in the past two years ~15 sensitive parameters in the microphysics, radiation, dynamics and turbulence have been identified, introduced and assessed. Both for SPPT and SPP, studies have been done on assessing optimal spatial and temporal scales and various stochastic pattern generators.

For the present SPPT setup, there are several ideas on how it might be improved, e.g. by considering partial rather than only total tendencies. In the coming years, these options will be studied further, and tendency diagnostics will be refined accordingly. An assessment will be made by the physics team of ways through which more truly stochastic physics formulations may be achieved on the longer term. More physics parameters will be included and assessed in the SPP scheme, also some from the dynamics (e.g. SLHD). The SPP scheme will also need to be updated to the new physics which is expected to become available soon (e.g. LIMA, ECRAD, aerosol formulations), and will be extended with relevant parameters from the new soil, vegetation and snow schemes which will be implemented in Cy46. Furthermore it will be studied how correlated parameters can best be handled.

For the surface, perturbations can be made in surface analysis fields (Bouttier et al, 2015) for e.g. SST, albedo, LAI, soil temperature and soil moisture. In the coming years, studies will be continued to determine optimum temporal and spatial scales for surface perturbations. Experiments will be done with new surface fields to perturb, such as snow depth, sea ice extent and stomatal resistance. Also the SPP scheme will need to be updated with relevant parameters from the new soil/vegetation and snow schemes which will be implemented in Cy46.

Lateral boundary conditions in HarmonEPS can be perturbed using the ensemble of the nesting model, IFS ENS (default) or by SLAF (Scaled Lagged Average Forecasting from ECMWF high resolution forecasts). The option of clustering of IFS ENS members is available, but has shown little consistent added value. In the coming two years, the work on LBC perturbations will be given relatively low priority.

In addition to testing perturbation techniques as described above, there are structural changes that also need testing, such as the balance between the number of ensemble members, the horizontal and vertical resolution, and the size of the area. These are important questions to answer to get the most out of the available operational computer resources in the members states. Work on this has already started and will be continued in the coming years. Also, the optimal balance between the different types of perturbations is something to investigate further.

E) Code efficiency and scalability

An important task to achieve is the optimization of code efficiency and scalability, with a view to use the model effectively on very massively parallel hardware platforms. Efforts on this will continue to be intensified in the coming years.

Already ongoing activities within HIRLAM and our ALADIN partners to adapt scientific algorithms towards greater computational efficiency and/or scalability will be continued. Examples of these are e.g. the development of EnVar data assimilation techniques (which are both faster and more scalable than 3- or 4D-Var); the use of continuous assimilation and non-linear spectral grids in sub-km resolution nowcasting ensembles; the trend to avoid computations in spectral space as much as possible; and the study of alternative (non-spectral) Helmholtz solvers and multi-grid approaches in the model dynamics.

Investigations into the use of single precision in the forecast model and the 3D-Var analysis have already shown that most model components can be "safely" run in single precision, yielding practically identical meteorological results and ~40% reduction run time reductions. There are still some components to be investigated in this respect, in particular 4D-Var.

In terms of parallellization, code optimization efforts will be continued. The Barcelona Supercomputing Center (BSC) has made a full profiling and analysis of Harmonie performance on the ECMWF platform. The recommendations from this assessment will be worked out in the coming year. One of the things to be tackled is the OpenMP thread parallelization, which appears to be underutilizing the potential of many-thread processors. Re-analysis experiences have shown the potential of significantly speeding up the screening and minimization on certain platforms (including ECMWF) by means of a better placement of tasks. Future work will also include a deeper assessment and optimization of the parallelization in new model components, such as 4D-Var, the aerosol parametrizations and initialization from CAMS, and the new surface model and assimilation components. will of the probably remains needed.

Assessments of Harmonie performance on new architectures such as ARM and AMD are already being done. To prepare the model for use on an increasing variety of computer architectures, HIRLAM and ALADIN are following the strategy of "separation of concerns" initiated by ECMWF. We will need to gain basic knowledge of some of the underlying concepts in developments related to this in the IFS environment, such as the use of CLAW and the development and use of domain specific languages, e.g. in the context of projects like ESCAPE-2. Work will be done to prepare the model for use with the ATLAS library and to study the new MultIO server as possible replacement of the existing IO server. Collaboration with external partners such as the Barcelona Supercomputing Center is sought in order to achieve an analysis of the code efficiency on a variety of computer architectures, including mixed CPU/GPU platforms.

Duration of the project and estimated resource requirements:

The duration of the HIRLAM-C programme has been prolonged with two years until 31-12-2022. The present request is for special project resources in the last two years (1-1-2021 until 31-12-2022) of the programme.

For testing and tuning of the deterministic Harmonie 4D-Var system at ECMWF at 2.5km horizontal resolution and 90 vertical levels over the DMI domain, runtime costs amount to ~20000 HPCF units per experiment day. The estimated needs for the testing of the deterministic Reference system are:

- pre-release technical tests: 12 months in total
- parallel validation: 12 months total
- pre-operational impact and sensitivity tests evaluating individual components: 12 months
- debugging, problem detection and fixing activities: 12 months
- real time trunk suite, 12 months in total

So in total roughly 60 months or 60 * 30 * 20000 units = 36 million HPCF units are estimated to be required per year for testing and experimentation with the deterministic Harmonie Reference System at ECMWF in the coming years.

To test HarmonEPS, an ensemble of 10 members is run over the MetCoop domain with one forecast per day running to +36h, and three +6 forecasts per day to keep the cycling. This costs ~120 000 HPCF units per day. Typically, updates are tested for 3 weeks in summer and 3 weeks in winter, so 42 days per update. With 2 tests for each perturbation type per year (initial, LBCs, surface, model and structural) we end up with:
120 000 HPCF units * 42 days * 10 experiments = 50.4 million HPCF units per year.

In total therefore, the testing of the Harmine and HarmonEPS configurations require ~ 86 million HPCF units per year. A considerable amount of these total requirements will be covered partly through explicit contributions from member states to a dedicated Hirlam SBU pool supplementing the special project resources, partly through direct billing to the member state HPCF quotas. For the Hirlam-C phase 3 special project, we apply for 2021 and 2022 for 40 million HPCF units/year, and a data storage of 22,000 GB, most of latter on temporal storage (ECTMP).

References:

Bouttier, F., L. Raynaud, O. Nuissier, and B. Ménétrier, 2015: Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX. Quart. J. Roy. Meteor. Soc., 142, 390–403, doi:https://doi.org/10.1002/qj.2622.

Buizza, R., M. Miller, and T. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. Quart. J. Roy. Meteor. Soc., 125, 2887–2908, doi:https://doi.org/10.1002/qj.49712556006.

Frogner, I.-L. et al., 2019. HarmonEPS: the Harmonie ensemble prediction system, Wea. & Forec. Vol 34, p.1909-1937, DOI: 10.1175/WAF-D-19-0030.1

Geijo, C., 2012: Assimilation of radar reflectivity data using the field alignment technique, Hirlam Newsletter 59, 10-20

Gregow, E., 2017, Harmonie – MSG cloud data assimilation experiments, ALADIN-HIRLAM Newsletter 10, pp 22-29, http://www.umr-cnrm.fr/aladin/meshtml/NL10-final.pdf

Hartmann, I., and Finnigan, J., 2007. A simple unified theory for flow in the canopy and roughness sublayer, Boundary Layer Meteorology, 123, 339-363

Mile, M., Randriamampianina, R., Marseille, G.-J., and Stoffelen, A., 2020. Scatterometer data assimilation with the supermodding method in high resolution limited-area model: a proof of concept, submitted to Q. J. R. Meteorol. Soc.

Raupach, M., 1994, Simplified expressions for vegetation roughness length and zero-plane displacement as functions of canopy height and area index, Boundary layer Meteorology, 71, 211-216

Tsyrulnikov M., and Gayfulin, D, 2017. A limited area spatio-temporal stochastic pattern generator for simulation of uncertainties in ensemble applications, Meteor. Zeitschrift 26, 549-566, DOI: 10.1127/metz/2017/0815

Yang, X., 2018. Sub-km Harmonie and on-demand setup for storm forecast, ALADIN-HIRLAM Newsletter 10, pp.35-39, http://www.umr-cnrm.fr/aladin/meshtml/NL10-final.pdf

Yang, X., Feddersen, H., Hansen Sass, B., Sattler, K., 2017: "Construction of a continuous mesoscale EPS with time lagging and assimilation on overlapping windows", ALADIN-HIRLAM Newsletter 8, pp.112-118, http://www.umr-cnrm.fr/aladin/meshtml/NL8-final.pdf