REQUEST FOR A SPECIAL PROJECT 2012–2014

MEMBER STATE: The Netherlands…………………….………………………

Principal Investigator[1]: Jeanette Onvlee, Hirlam-B programme manager ……………

Affiliation: KNMI…………………………………………………………

Address: Wilhelminalaan 10, PO.Box 201, 3730 AE De Bilt, the Netherlands

…………………………………………………………………….……

…………………………………………………………………….……

E-mail: onvlee@knmi.nl…………………………………………….…………

Other researchers: Hirlam programme team (present project leaders: Ulf Andrae (SMHI), Mariano Hortal (AEMET), Trond Iversen (met.no), Magnus Lindskog (SMHI), Laura Rontu (FMI), Xiaohua Yang (DMI))

Project Title:

Hirlam-B project
…………………………………………………………………….……

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SPsehlam** ___________________	
Starting year: (Each project will have a well defined duration, up to a maximum of 3 years, agreed at the beginning of the project.)	2011	
Would you accept support for 1 year only, if necessary?	YES ☐	NO XX☐

Computer resources required for 2012-2014: (The maximum project duration is 3 years, therefore a continuation project cannot request resources for 2014.)	**2012**	**2013**	**2014**
High Performance Computing Facility (units)	3750000	4500000	
Data storage capacity (total archive volume) (gigabytes)	10000	12000	

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):

………………20/4/2011………………

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.

Principal Investigator: Jeanette Onvlee……………………………………………………

Project Title: HIRLAM-B project……………………………………………………

Extended abstract

It is expected that Special Projects requesting large amounts of computing resources (500,000 SBU or more) should provide a more detailed abstract/project description (3-5 pages) including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used. The Scientific Advisory Committee and the Technical Advisory Committee review the scientific and technical aspects of each Special Project application. The review process takes into account the resources available, the quality of the scientific and technical proposals, the use of ECMWF software and data infrastructure, and their relevance to ECMWF's objectives. - Descriptions of all accepted projects will be published on the ECMWF website.

The HIRLAM-B Programme, which has started in January 2011 and will end in December 2015, is a continuation of the research cooperation of the previous HIRLAM projects. The expected members of HIRLAM-B are the national meteorological institutes in Denmark, Estonia, Finland, Iceland, Ireland, Netherlands, Norway, Spain and Sweden.

Within HIRLAM-B, research efforts are focussed on the development, implementation and further improvement of a non-hydrostatic meso-scale analysis and forecast system, called Harmonie, and on of a short-range mesoscale ensemble prediction system suitable for severe weather. The Harmonie system is being developed in a code cooperation with the Aladin consortium. For the original Hirlam model, a final version 7.4 has been released at the end of 2010. The Hirlam system will be maintained for a while (e.g. for ensemble forecasting purposes), but is expected to be gradually replaced by Harmonie.

Following the past Hirlam practice, a Reference system is being maintained on the ECMWF HPC platform for both the Hirlam and Harmonie models. The emphasis in the HIRLAM-B Special Project at ECMWF is primarily on the evaluation and testing of the Harmonie Reference System. The quality of the Reference System is of paramount importance to all Hirlam members. The Special Project computational resources will be used mainly to experiment with newly developed model components and evaluate their meteorological and technical performance in beta-releases, before releasing them as Reference.

The sections below describe the main research activities intended in the coming years. In-depth validation and intensive (pre-)operational testing of all of these developments will be carried out both in the member institutes and in the ECMWF environment. In addition to the development of the deterministic model, research is also done on ensemble forecasting at scales of ~10km horizontal resolution down to convection-permitting ensembles. For these developments, separate special project resources have been requested.

<u>Data assimilation:</u>

At present Harmonie uses as default a 3D-Var assimilation system; additionally, there is a (still very basic) 4D-Var research system. A first priority for upper-air data assimilation is to refine the present 3D-Var system with respect to applications at horizontal resolutions of 1-5 km. These refinements will initially focus on tuning of present observation usage, and exploitation of new types of mesoscale observations, in particular radar. For use of the 3D-VAR in rapid update cycling mode for nowcasting purposes, model spin-up is a challenge that needs to be assessed and addressed. Optimisation of initialisation methods as well as of cycling procedures for moist parameters related to moist physics will be important.

The study of hybrid ensemble data assimilation systems (until recently carried out with the Hirlam model) will be transferred to Harmonie. Several hybrid ensemble assimilation techniques are under development: (a) introduction of flow dependent background error covariances by applying a Hybrid 3D-Var/ETKF variational data assimilation technique, following ideas of Lorenc (2003), (b) application of ensemble data assimilation (Berre et al., 2007) and (c) air-mass dependent background error covariances (Montmerle and Berre, 2010), to distinguish between balances in rainy and dry areas.

4D-Var has proven very successful at the synoptic scale in introducing flow-dependency and in extracting information from complex observations types, both topics being very important for km-scale data assimilation. Despite the potential problems of 4D-VAR in handling non-linearities and achieving good scalability, its potential strengths are considered to be of great enough interest to continue exploring it further. Thus, in parallel to the various approaches to be explored for enhancing 3D-VAR, the Harmonie 4D-VAR system will be further extended and tested, initially, on horizontal resolutions around ~5 km, but later also at higher resolutions. The 4D-VAR developments will include the introduction of the ECMWF simplified physics package with its relatively advanced moist physics, as well as tuning and evaluation of the weak constraint digital filter initialisation and the application of multiple outer loops.

Concerning the use of observations, first priority in the upper air analysis will be given to the assimilation of radar radial wind and reflectivity observations. Considerable resources will be devoted to pre-processing, quality control, tuning and assimilation of these radar data. The aim is to achieve that radar data from most HIRLAM countries can be ingested into Harmonie during the first half of 2011; following that, the required quality control will be put in place and observation impact studies will be initiated, starting with single observation experiments. At somewhat lower priority, several other remote sensing sources of spatially dense observations will be introduced and their potential explored. In particular ground-based GPS, Mode-S wind data and ASCAT winds are considered to be promising. In the third place, it is aimed to improve the use of observations types that already are available in Harmonie 3D-Var and 4D-Var and that have traditionally also been used for HIRLAM synoptic scale assimilation: conventional observations and measurements from satellite-based instruments (AMSU, MHS, AMV, IASI). We need to tune and further optimize the handling of these data with regard to quality control, thinning, bias correction and error characteristics.

Surface data assimilation has clearly shown to be beneficial for Harmonie. An extended utilization of various types of surface observations is needed to improve the characterisation of surface conditions. The analysis of all surface aspects relevant to NWP specifically requires the use of remote sensing surface observations, and optimal use of these observations in turn requires application of more advanced assimilation methods than the presently available OI. For soil data assimilation, an EKF has been developed which is presently being implemented and evaluated in the HARMONIE system. Research on surface observation usage will mainly focus on exploiting a greater variety of remote sensing observations, such as ASCAT and SMOS products for soil moisture, GLOBSNOW observations for snow assimilation and MODIS data for lake water temperature and ice fraction assimilation.

Dynamics

The main emphasis in the dynamics research is on increasing the accuracy and efficiency of the dynamics code, with a particular view to prepare the model for future use at very high resolutions and on massively parallel computer systems. Additional issues to be considered are the physics-dynamics interaction on km-scales, and boundary condition treatment, domain size and nesting considerations.

Experiments with a VFE formulation developed in a cooperation between HIRLAM and ECMWF has shown that this discretization results in an accurate and stable scheme, provided that a change in vertical coordinate to a height-based coordinate is introduced. This adaptation is being implemented and will be tested in 2012.

The semi-implicit semi-Lagrangian scheme in the present Harmonie dynamics offers the possibility of using long time steps, but its efficiency is seriously hampered by the need to use, in some circumstances, a predictor-corrector scheme for stability reasons. It is aimed to make the code more stable, thus allowing to dispense with the need for a predictor-corrector scheme. A second disadvantage of the semi-Lagrangian technique is that it is not designed to conserve neither the mass of dry air nor the proportion of any other component of the atmosphere. Improvements of the semi-Lagrangian method to better conserve mass will be sought which do not involve a large increase of the expense of the model. Some initial successful tests have already been done changing the interpolations used in the semi-Lagrangian treatment of the continuity equation, which has led to improved conservation of dry air mass. As a next step it will be considered to enhance the mass conservation for individual atmospheric components in the semi-Lagrangian treatment of the mixing ratio equations.

The question has been raised if a spectral formulation is still appropriate when model mesh sizes approach scales of several 100m. In order to check whether there are limitations in the spectral technique at such very high resolutions, for example at or near steep orographic slopes, the HARMONIE model will be run in LES mode at resolutions down to tens of meters over areas where there is an orographic data base of enough resolution.

The present semi-implicit time-stepping scheme introduces dispersion in the treatment of short linear gravity waves. Tests will be done on the impact of replacing the scheme with a semi-analytical one which should be free of this effect.

An important issue to be studied is what should be the necessary changes to the current Harmonie dynamics (spectral and VFE, semi-Lagrangian, semi-implicit) in order to sustain good performance on HPCs on the longer term, in terms of a good balance between scalability, accuracy and stability. Issues for which alternatives to the present dynamics will be considered, are e.g.:

- Compare Semi-Lagrangian and Eulerian advection at high (sub-km) resolution
- Keep spectral solver but compare spectral technique for the computation of derivatives against local methods (finite elements, finite volume, finite differencing at high order)
- Compare spectral and grid-point Helmholtz solvers
- Compare staggered versus non-staggered grids
- Compare semi-implicit vs. explicit treatment of gravity and acoustic waves in horizontal

These alternative dynamics options should be introduced and tested in a conservative, stepwise and modular way, keeping the present solution until the alternative has proven its value.

Nesting experiments performed during 2010 have shown that a single-nested configuration, even with a significant jump in resolution from the host model (ECMWF at 16km resolution) to the nested model (Harmonie at 2.5km and HIRLAM at 3km), is of better forecast quality than a double-nested system involving an intermediate HIRLAM or ALADIN model. These experiments will be evaluated further, and where necessary follow-up experiments will be defined. The MetOffice Unified Model employs a variable resolution in the boundary zone, which stepwise becomes coarser as its outer boundaries are approached. In principle this approach allows for an even greater contrast in resolution between host and nested model. It will be considered if such a solution may be desirable when going to sub-km scale mesh sizes.

<u>Physics parametrizations</u>:
Harmonie contains two branches of upper air parametrizations: AROME physics with explicit deep convection, and ALARO with partially parametrized convection. Physics developments will principally focus on improving the package most suitable for use at high resolutions, AROME, applying HIRLAM and ALARO experiences where relevant. For mesh sizes of several km, several processes which need to be parameterized at coarser scales can be assumed to be at least largely resolved, and thus can be described explicitly, such as deep convection and orographic gravity waves. At higher resolutions, of ~1km or less, also shallow convection and turbulence become partially resolved.

Future models of such high spatial detail are likely to require 3-dimensional parametrizations of unresolved processes as well as the introduction of some form of stochasticity in the physics. A 3D physics-dynamics interface will have to be developed to allow these developments. Investigations are being done on the use of 1+2D and 3D parametrizations for turbulence and radiation, respectively. The relation between the horizontal diffusion applied by the model dynamics (Semi-Lagrangian or conventional horizontal diffusion) and the parametrized vertical diffusion will be studied under different flow conditions. Moisture-conserving turbulence parametrizations in two and three dimensions are being developed based on the present TKE-scheme of AROME. Tilted array modelling of surface radiation will be applied to study the importance of transient and local cloud shadow effects for surface radiation fluxes, and the possible impact of a 3-dimensional treatment of the radiation parametrization. In addition to stochastic physics, a new approach using cellular automata is being used to test the impact of introducing stochasticity at various stages of development of organized convection.

The challenge of weather prediction under (very) stable boundary layer conditions is of great practical importance especially in at the northern latitudes, where cold temperatures and development of low clouds and fog are frequent. It is planned to make a concerted effort on improving the parametrization of stable boundary layer effects by an interdisciplinary team. Attention will be paid to the following aspects:

- Turbulence parametrizations of the (long-living very) stable boundary layer. The applicability of new theoretical concepts suggested by Zilitinkevich et al. (2008) will be studied.
- Interaction between waves and turbulence in the boundary layer. QNSE formulations (Sukoriansky et al 2005) of this will be studied.
- The decoupling between upper air and near-surface layer in the model as compared to reality. The influence of model vertical resolution should be studied, as well as the near-surface sublevel approach in the canopy scheme of SURFEX, and the optimal definition of the lowest model level height.
- Cloud microphysics and radiation interactions, treatment of the long-wave radiation near the surface. Cooling at the surface and at the top of fog/stratus.
- Influence of surface properties - snow, ice, vegetation. Test of a multiple energy balance formulation for snow-forest-soil and its relations to the surface-layer sublevel approach.
- Study the relative role and interaction of the prognostic parametrizations and data assimilation for the

stable boundary layer.

The present microphysics scheme in Harmonie is ICE3. An alternative to be studied is the development of a more advanced, second-moment microphysics scheme, which treats the number concentration of cloud condensation nuclei in a prognostic manner. A second-moment scheme would permit a physically realistic way of taking into account indirect aerosol effects.

A comprehensive radiation inter-comparison study has been started, in which several ways are considered to further improve the radiation treatment. The impact of using a more advanced clear-sky radiation transfer parametrization will be studied. Another aspect to be addressed is the handling of cloud-radiation interactions in the model. It is aimed to improve the treatment of radiation-surface interactions, in particular the effects of slopes and vegetation. The impact of the aerosol direct effects on radiation will be studied, and an estimate will be made of the possible importance of ozone in Harmonie.

An aspect deserving further study is the influence of the size of the model domain on the forecast quality of high-resolution models. Past experience indicates that the use of small model domains may prevent or hinder a good representation of mesoscale features by the model. This needs to be examined more thoroughly.

For surface modelling, the focus will be on improving the description of Northern, Arctic and Antarctic conditions in Harmonie. Key issues are the handling of snow, ice, forest, lakes and sea in the forecast model and the surface data assimilation. For an accurate surface description, a close connection between the surface model and surface data assimilation is of crucial importance. Surface assimilation should provide an initial state for the relatively swiftly evolving variables at or below the land and water surface, such as mean lake temperature or soil moisture. For the description of constant or more slowly evolving surface properties - orography, land use, vegetation – on the other hand, detailed physiographic information is required both by the parametrizations and the surface assimilation. In Harmonie/SURFEX, the starting point of physiographic information on vegetation, land use, water bodies etc. is the ECOCLIMAP database. During recent years, significant developments of high-resolution surface physiographic databases have taken place, mostly based on satellite information. The potential of these additional sub-km scale data sources will be studied.

Code efficiency and scalability

An important aspect to consider in the coming years is the optimization of code efficiency and scalability, with a view to use on very massively parallel hardware platforms. Comprehensive profiling of the code at the introduction of every new cycle will be necessary to clearly establish which parts of the code are the most limiting factors in terms of efficiency and scalability.

The main bottleneck for scalability, in Harmonie as in most other forecast models, is the need for I/O to read initial data and to write out forecast fields at required intervals. The use of intermediate formats like FA within Harmonie is less efficient than the I/O handling as it is done in the global IFS, which uses a Fields DataBase (FDB). This is something to be tackled at high priority.

From the point of view of scalability, the 4D-Var code is one of the potentially most troublesome issues in the future. The main problem, as shown by studies at ECMWF, is the interpolation from high to low resolutions in incremental 4D-Var; this could be improved by allowing both the outer and the inner loops within the same execution of the code, avoiding thereby the writing of intermediate files. Another option would be use of a weak constraint 4D-Var formulation with sub-windows running in parallel. Ultimately, if these solutions would be insufficient, the decision might need to be taken to abandon the 4D-Var concept altogether, in favour of e.g. an ensemble Kalman filter approach. Ensemble methods by definition are practically perfectly scalable.

In terms of parallelization, several existing and potential future bottlenecks can be identified. The greatest problem at the moment is the lack of a proper shared-memory parallelization of the SURFEX surface scheme. This should be largely solved with code adaptations introduced in 2011. The best way to improve the parallelization of a computer code on the longer term is to restrict as much as possible the need for communications among processors. In the context of the semi-implicit (SI) semi-Lagrangian (SL) Harmonie code, communications are unavoidable for applying SL interpolations and to perform spectral transforms. The efficiency of the communications needed for the SL scheme can be improved with the use of "on demand" SL communications.

It is intended to benchmark the Harmonie model on as massively parallel machines as are available to the consortium. These benchmarks should be done not only for the model as a whole, but also for the system with different "bricks" removed (e.g. physics, I/O, …).

Duration of the project and estimated resource requirements:

The duration of the requested special project is to be from 1 January 2011 until 31 December 2013. For the final years of the Hirlam-B programme, 2014 and 2015, a new proposal for special project resources is likely to be submitted at that time, with updated scientific and technical goals.

The computational costs of these Harmonie experimentation and validation activities will likely extend well beyond the requested project resources (see the specification below). A pool of national resources at ECMWF has therefore been created by the HIRLAM institutes to supplement the special project resources.

For testing and tuning of the Harmonie system at ECMWF at 2.5km horizontal resolution, runtime costs amount to ~15000 SBU per experiment day. The estimated needs for Reference system testing are (assuming a level of activities similar to previous years):

- pre-release technical tests: 4 months in total (4 “alpha” and “beta”software releases, each with approximately two-week technical testing and debugging)
- parallel validation: 4 months total (2 month-long tests per major release for approximately 2 release versions)
- pre-operational impact and sensitivity tests evaluating individual components: 10 months
- debugging, problem detection and fixing activities: 4 months

So in total roughly 22 months or 22 * 30 * 15000 units = 10M HPCF units are estimated to be required per year for testing and experimentation with the Harmonie Reference System at ECMWF in the coming years. A considerable amount of the above will be covered by utilizing allocated member service national resources, partly through explicit contribution from member states to a dedicated spsehlam pool, partly through direct billing to the member state SBU quotas (as has been the case in the past). For the Hirlam-B special project, we apply for 2012 for 3.75 million HPCF units, and a data storage of 10,000 GB.