REQUEST FOR A SPECIAL PROJECT 2025–2027

MEMBER STATE:	Norway
Principal Investigator[1]:	Harald Schyberg
Affiliation:	The Norwegian Meteorological Institute
Address:	Henrik Mohns Plass 1 0371 Oslo
Other researchers:	Jostein Blyverket (MET Norway), Åsmund Bakketun (MET Norway), Eivind Støylen (MET Norway), Stephanie Guedj (MET Norway), Benjamin Menetrier (MET Norway), Xiaohua Yang (DMI), Emy Alerskans (DMI), Søren Thorsen (DMI), Jelena Bojarova (SMHI), Abhishek Lodh (SMHI)
Project Title:	Regional reanalysis demonstrators for CopernIcus Climate Change Service Evolution

To make changes to an existing project please submit an amended version of the original form.)

If this is a continuation of an existing project, please state the computer project account assigned previously.	**SP** ……..	
Starting year: (A project can have a duration of up to 3 years, agreed at the beginning of the project.)	2025	
Would you accept support for 1 year only, if necessary?	YES ☐	NO ✅

Computer resources required for project year:		**2025**	**2026**	**2027**
High Performance Computing Facility	[SBU]	208.3M	124.8M	-
Accumulated data storage (total archive volume)[2]	[GB]	210 700	244 000	-

EWC resources required for project year:		**2025**	**2026**	**2027**
Number of vCPUs	[#]			

[1] The Principal Investigator will act as contact person for this Special Project and, in particular, will be asked to register the project, provide annual progress reports of the project's activities, etc.
[2] These figures refer to data archived in ECFS and MARS. 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 etc.
[3]The number of vGPU is referred to the equivalent number of virtualized vGPUs with 8GB memory.

Total memory	[GB]			
Storage	[GB]			
Number of vGPUs[3]	[#]			

Continue overleaf.

Principal Investigator:	Harald Schyberg
Project Title:	Regional reanalysis demonstrators for CopernIcus Climate Change Service Evolution

Extended abstract

All Special Project requests should provide an abstract/project description including a scientific plan, a justification of the computer resources requested and the technical characteristics of the code to be used. The completed form should be submitted/uploaded at https://www.ecmwf.int/en/research/special-projects/special-project-application/special-project-request-submission.

Following submission by the relevant Member State the Special Project requests will be published on the ECMWF website and evaluated by ECMWF and its Scientific Advisory Committee. The requests are evaluated based on their scientific and technical quality, and the justification of the resources requested. Previous Special Project reports and the use of ECMWF software and data infrastructure will also be considered in the evaluation process.

Requests exceeding 10,000,000 SBU should be more detailed (3-5 pages).

Regional reanalysis demonstrators for CopERnIcus climate change Service Evolution (CERISE)

Overview and goals of the project

This application is for computer resources for carrying out research and for running regional reanalysis demonstrators which are candidates for producing future regional reanalysis products in the Copernicus Climate Change Service (C3S). This is part of the Horizon-Europe project CopERnIcus climate change Service Evolution (CERISE) coordinated by ECMWF (see https://cerise-project.eu/).

In general the CERISE project aims to enhance the quality of the C3S reanalysis and seasonal forecast portfolio, with a focus on land-atmosphere coupling. The project plan focuses on supporting the evolution of C3S by improving the C3S climate reanalysis and seasonal prediction systems and products towards enhanced integrity and coherence of the C3S Earth system Essential Climate Variables. CERISE will develop new and innovative coupled land-atmosphere data assimilation approaches and land initialisation techniques to pave the way for the next generations of the C3S reanalysis and seasonal prediction systems. These developments will include innovative work on observation operators using Artificial Intelligence to ensure optimal data fusion integrated in coupled assimilation systems. They will enhance the exploitation of Earth system observations over land surfaces, including from the Copernicus Sentinels and from the European Space Agency Earth Explorer missions, moving towards an all-sky and all-surface approach. CERISE Research and Innovation will bring the C3S tools beyond the state-of-the-art in the areas of coupled land-atmosphere data assimilation, observation operators, and land initialisation methodologies.

This application is specifically for the parts of CERISE developing, implementing and running regional land and coupled reanalysis demonstrators, where sufficient HPC resources are not available otherwise. This part of the project is carried out by three of the partners in the CERISE consortium, the Norwegian Meteorological Institute (MET Norway), the Swedish Meteorological

and Hydrological Institute (SMHI) and the Danish Meteorological Institute (DMI), and the resources applied for here will be shared by scientists and developers of these three institutes.

Methodology, scientific plan

In the part of CERISE covered by this application we will design, develop, produce and deliver a set of regional scale reanalysis prototypes. The prototypes will both include standalone land as well as coupled land-atmosphere three-dimensional reanalyses for regional domains. It will in particular focus on improved consistency and long-term trends of land-atmosphere ECVs. Proof of concept of coupled data assimilation methods for Earth System reanalysis at regional pan-European and pan-Arctic scale will be available by the end of the project, where prototype systems are candidates for operational implementation in the next phase of Copernicus (Copernicus 3, post-2027). The regional land reanalysis prototypes will provide consistent long-term surface initial conditions for the coupled regional reanalysis stream initialisation. Initialisation of the coupled reanalysis streams from dedicated offline land reanalysis will enable to take account of the long-term evolution of natural and anthropogenic changes and avoid discontinuities in the slow-evolving components in land surface variables (e.g. root zone soil moisture) between reanalysis production streams. The regional land reanalysis prototypes will in addition add value with resolution and detail and will exploit the evolution of satellite observations and novel land surface assimilation methodology. In particular, for the Arctic domain where the reanalysis relies heavily on satellite observations, the approach will target improved accuracy for surface and near-surface quantities. This will be relevant as input for Arctic climate process studies (including glacier mass balance, permafrost melt, related surface carbon fluxes such as CO2 and CH4). The regional reanalysis prototypes will be run over limited time slices to produce demonstration reanalysis datasets intended to demonstrate feasibility and allow evaluation. All data created in CERISE will be archived at, and managed by, ECMWF with easy access to all participants in CERISE as well as on request to others.

Unified land surface data assimilation

The Land data assimilation systems evolution in CERISE will be centered on the development of unified multivariate ensemble-based land data assimilation. The objective is to move away from the collection of univariate land assimilation approaches currently used in the C3S systems, and to enhance the consistency across the components. Furthermore, using an ensemble approach will enhance the quantification of uncertainties within the Land data assimilation systems.
The ensemble based approach will be based on a flavor of the ensemble Kalman filter (EnKF). It will be utilized both in the land and the coupled demonstrators. In the land demonstrators we will explore perturbation methods of the forcing variables input to the land surface model and perturbation methods specific for the multi-layer soil and snow schemes. This perturbation methodology will also be used in the coupled demonstrators. For the land demonstrators the focus will be on including satellite observations sensitive to surface soil moisture, snow water equivalent and/or snow extent.

The unified land data assimilation system will also be implemented in the coupled demonstrators. In the coupled demonstrators the unified system will be used to assimilate screen-level variables and SYNOP snow depth observations.

Coupled data assimilation

For the coupled regional reanalysis systems, we will investigate different degrees of coupled land-atmosphere assimilation using weakly coupled and different combinations of outer loop coupling. Here we will identify optimal degrees of coupling required for consistent initialisation of the surface-atmosphere variables. It will include code infrastructure developments to enable exchange of information between the different data assimilation systems used for the atmosphere and the surface and to improve the coupling modularity. This will Improve the exploitation of surface-sensitive observations by coupling the land surface and the atmospheric observation operators. These developments will enable simultaneous analysis of land surface and atmospheric variables in the coupled assimilation system, moving towards an all sky and all surface approach. Specifically, a land surface observation operator for passive microwave satellite sensors will be developed using machine learning. This land surface operator will be coupled to the atmospheric observation operator RTTOV (Radiative Transfer for TOVS), with potential benefits for the microwave scientific community. CERISE will focus on the exploitation of low frequency passive microwave observations in the coupled data assimilation system using existing sensors such as SMOS and AMSR2 in preparation of the Copernicus Sentinel mission CIMR. In parallel, we will investigate the assimilation of derived skin temperature from satellite radiances in the atmospheric four-dimensional variational assimilation (4D-Var) Extended Control Variable. This to constrain the land surface variables, including snow temperature, soil temperature, and sea ice temperature. This development will follow the ongoing work for ocean-atmosphere coupled assimilation conducted at ECMWF and will be extended to land surface for the regional demonstrators. Preliminary assessment will be conducted using standard NWP evaluation tools with particular attention to snow covered areas. The coupled assimilation systems developed will be used as input to the reanalysis prototypes demonstration.

In the coupled regional demonstrators we will in parallel develop 4D-Var and 3D-EnVar variational assimilation schemes. For the 4D-Var, infrastructure developments will bring the multi-layer surface physics used in CARRA2 (2nd Copernicus Arctic Regional Reanalysis) into the 4D-Var outer loop. In addition we will bring the ensemble based land surface data assimilation developments into the 4D-Var outer loop. Developments of 3D-EnVar will be done in the Object-Oriented Prediction System (OOPS). In 3D-EnVar a more explicit coupling strategy will be explored, for example by extending the control vector with surface temperature.

Description of the demonstrators

CARRA-Land-Pv2 (Copernicus Arctic Regional Reanalysis, Land Prototype version 2) will be an Arctic offline surface prototype reanalysis for two up to five-year time periods (within the same time slices as for CARRA-Land-Pv1, developed in 2024 and not covered in this application) at a resolution of 2.5 km or finer. It will use forcing from time slices of the operational CARRA2 pan-Arctic dataset that will become available in the course of the project. It will demonstrate added value to description of Arctic surface, snow and soil quantities relative to the surface data provided from CARRA2. This prototype will integrate the unified land surface data assimilation system outlined above.

CARRA3-Pv1 (3rd Copernicus Arctic Regional Reanalysis, Prototype version 1) will be a coupled regional reanalysis system for a pan-Arctic domain, candidate for implementation, in the next phase of Copernicus, as CARRA3. It will be produced for a full year post 2015 at a spatial resolution of 3.75 km. It will take onboard unified ensemble land data assimilation in SURFEX and outer loop coupling in HARMONIE-AROME. It will also integrate skin temperature assimilation over some surfaces (e.g. snow/ice).

CERRA2-Pv1 (2nd Copernicus European Regional Reanalysis, Prototype version 1) will be a coupled regional ensemble reanalysis system for a pan-European domain. It will be produced for a full year post 2015 at a spatial resolution of 3.75 km (5.5 km for ensemble members). The prototype will evolve on CARRA3-Pv1 and aims at identifying relevance for European domain coupled reanalysis for future Copernicus phases.

Overview description and characteristics of the NWP system to be used

The developments in the project will be based on versions of the HARMONIE-AROME NWP system, which is a km-scale Limited Area NWP system which is targeted to model convective scale phenomena. HARMONIE-AROME CSC is one of the three Canonical System Configurations jointly developed by the ACCORD consortium in the IFS modeling environment. The regional coupled reanalysis products are based on the HARMONIE-AROME system with SURFEX (https://climate.copernicus.eu/copernicus-arctic-regional-reanalysis-service and https://climate.copernicus.eu/copernicus-regional-reanalysis-europe-cerra). They will use domain boundary conditions from the C3S operational global reanalysis products of ERA5 and ERA6 when it becomes available towards the end of the project.

The base source code for the CERISE regional demonstrators will be the CARRA2 system, which will start production in Q2 2024. CARRA2 source code is based on the HARMONIE-AROME CSC Cycle 46. The CARRA2 system has been highly optimized to run as efficiently as possible on the Atos HPC. The demonstrators in CERISE will benefit from the CARRA2 specific changes to the HARMONIE-AROME CSC, such as running high-resolution (2.5 km) on a large pan-Arctic grid and adaptations in observation usage for reanalysis. We will also benefit from the work done on running both the land surface model and the coupled model run forecasts in single precision. The first demonstrator (CARRA3-Pv1) will be based on the same Cycle as CARRA2 with CERISE developments on top. For the second coupled demonstrator (CERRA2-Pv1) we will port the relevant CARRA3-Pv1 changes to HARMONIE-AROME CY49.

Estimation and justification of HPC resource usage

The estimated HPC resource usage is based on the current CARRA2 system. We will be running at a coarser resolution, but with ensembles for the surface (CARRA3-Pv1) and more upper-air ensemble members (20 + 1 deterministic) for (CERRA2-Pv1), compared to CERRA (11 members). This is needed for the unified land surface data assimilation system, and the coupled assimilation methodology which is state-of-the-art for the CERISE demonstrators. It will in addition provide uncertainty estimates for the regional reanalysis demonstrators. The total SBU usage and storage for the three demonstrators is listed in Table 1. How this will be spread throughout the project period and combined with SBUs for testing and experiments is described in Table 2. The delivery date for the CARRA2-Land-Pv1 is 31st August 2025 and for CARRA3-Pv1 it is 31st December 2025. So for 2025 the HPC resources will be used to deliver these datasets. The remaining resources, 94.8M SBUs, will be used in 2026 to deliver the CERRA2-Pv1 demonstrator.

Table 1: HPC resources for the demonstrators

	Surface ensemble	**Coupled ensemble**	**Time-period**	**Total SBUs**	**Forcing**	**Model output**	**Total storage**
CARRA-Land-Pv2	Yes	No	Two up to 5 year periods	140.0M	45.7TB	100TB	145.7TB

CARRA3-Pv1	Yes	No	1 year	18.3M	0	65TB	65TB
CERRA2-Pv1	No	Yes	1 year	94.8M	0	44TB (Det) + 10TB*20 (EPS)	244TB

Table 2 shows the total SBU cost (demonstrators + development/testing). In addition to the cost of producing the demonstration data sets, we estimate around 30% additional SBU resources for developments and testing per year during the project period.

Table 2: Total HPC resources per year, demonstrators + testing

	2025	**2026**	**2027**	**Due date**
CARRA-Land-Pv2	140M SBUs	-	-	31. August 2025
CARRA3-Pv1	18.3M SBUs	-	-	31. December 2025
CERRA2-Pv1	-	94.8M SBUs	-	Due 30. September 2026
Testing	50M SBUs	30M SBUs		
Total	**208.3M SBUs**	**124.8M SBUs**	-	