REQUEST FOR AN EMI R&D PROJECT 2027–2029

Country/Organisation: National Observatory of Athens, Greece

Principal Investigator[1]: Dr. Vassilis Amiridis

Affiliation: National Observatory of Athens

Address: Lofos Nymphon , PO Box 20048, 11810, Athens

Other researchers:
Thanasis Georgiou (ageorgiou@noa.gr)
Kyriakos Malisovas (kmalisovas@noa.gr)
Vassilis Spyrakos (v.spyrakos@noa.gr)
George Roditis (groditis@noa.gr)

Project Title: Redefining dust mass in numerical weather prediction

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.)	2027	
Would you accept support for 1 year only, if necessary?	YES ☒	NO ☐

Computer resources required for project year:		**2027**	**2028**	**2029**
High Performance Computing Facility	[SBU]	70M	70M	70M
Graphics Processing Unit Cluster-A	[GBU]	0	0	0
Graphics Processing Unit Cluster-B	[GBU]	0	0	0
Accumulated data storage (total archive volume)[2]	[GB]	0	0	0

[1] The Principal Investigator will act as contact person for this EMI R&D 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.

EWC resources required for project year:		2027	2028	2029
Number of vCPUs	[#]	0	0	0
Total memory	[GB]	0	0	0
Storage	[GB]	0	0	0
Number of vGPUs[3]	[#]	0	0	0

Continue overleaf.

Principal Investigator: Dr. Vassilis Amiridis

Project Title: Redefining dust mass in numerical weather prediction

Extended abstract

All EMI R&D 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 EMI R&D 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 EMI R&D Project reports and the use of ECMWF software and data infrastructure will also be considered in the evaluation process.

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

1. Summary

Atmospheric dust mass loading remains one of the most poorly constrained quantities in aerosol modelling, with model intercomparisons showing order-of-magnitude disagreement in emitted, transported and deposited dust. This proposal requests HPC resources to produce a revised, observationally constrained estimate of dust mass load, built on two recent developments: new forward operators for dust optical properties that account for non-spherical particle shape, and an emission retrieval system based on a performant transport model with an exact adjoint. We will constrain dust emissions against EarthCARE observations under the new optical model, derive a revised dust mass field, and evaluate the consequences for modelled cloud condensation and ice nucleating particle populations using contemporary cloud and aerosol retrievals.

2. Scientific background and motivation

Dust emissions are not measured directly but inferred through modelling, and the associated uncertainties propagate into every downstream dust field. Results from the AEROCOM intercomparison show a standard deviation of dust emissions of around 1500 Mt per year against a median of the same order (Huneeus et al., 2011 [1]), with the largest disagreement over southern hemisphere source regions, indicating that the emission mechanisms themselves are unevenly understood across the globe.

Part of this spread originates not in the emission scheme but in the optical assumptions used to compare models with observations, since models are constrained with assimilated optical observations (such as Aerosol Optical Depth). We have recently developed forward operators for dust optical properties (extinction, optical depth) that treat dust as non-spherical. These operators make dust substantially more extinctive than the spherical approximation commonly assumed, which means a model requires considerably less dust mass to reproduce a given observed optical signal. In our assimilation experiments, two analyses that are indistinguishable in observation space correspond to dust mass loadings differing by roughly 46% depending on the optical model used (Figure 1). A large fraction of the apparent inter-model mass spread may therefore be an artefact of optical-model choice rather than genuine disagreement about emissions. This work was supported by an earlier special project with account name spgramir.

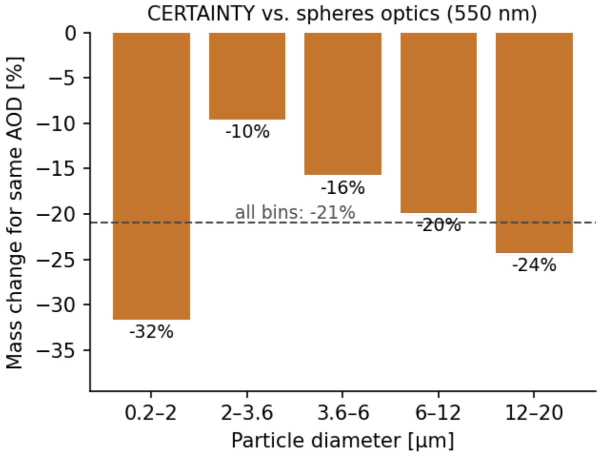

Figure 1: Mass change required to reproduce the same aerosol optical depth when changing from the traditional spherical operator to the new one. Each bar represents one dust size bin of the GOCART aerosol model, as it is implemented in WRF-CHEM.

This matters well beyond the aerosol optical budget. Aerosol-cloud interactions scale with particle mass and number rather than optical depth: dust is a significant source of cloud condensation nuclei (CCN) and one of the most important atmospheric ice nucleating particle (IN) species. A systematic revision of dust mass load therefore propagates directly into the modelled CCN/IN budget and the cloud radiative and microphysical response. If the optical correction implies that current models carry the wrong dust mass to match observations, then the dust contribution to cloud processes could also be correspondingly misrepresented.

The objective of this work is to close that gap by constraining dust emissions using the new retrieval technique under the new optical model, deriving a revised estimate of atmospheric dust mass load, and quantifying how that revision changes the modelled CCN, IN and cloud fields.

3. Methodology and workflow

The workflow couples an ensemble meteorological forecast, a new adjoint-capable dust transport model, an emission retrieval step assimilating EarthCARE observations, and a forecast-and-verification step targeting cloud-relevant quantities. The general idea is based on Yumimoto et al. 2008 [2], albeit with new space-borne observations, the advanced optical model and the new transport model implementation.

3.1 Transport and adjoint model. The core of the system is an offline dust transport model implemented in JAX[1], handling emission, advection, diffusion dry deposition and gravitational settling of dust, driven by externally supplied meteorology. The JAX implementation provides an exact adjoint by automatic differentiation at no additional development cost, and is performant enough to make iterative emission optimisation computationally feasible. This removes the central obstacle in adjoint-based emission retrieval, namely the construction and maintenance of a hand-coded adjoint. Using the adjoint, an optimisation algorithm similar to 4D-Var (based on Solvik et al. 2025 [3]) is implemented to constrain the dust emissions of the model.

3.2 Ensemble meteorology. Because the transport and the emission retrieval both depend on the meteorological state, we first generate an ensemble of regional meteorological forecasts with WRF, initialised and bounded by the ECMWF IFS ensemble. This propagates the operational meteorological uncertainty into our regional domain, which covers the northern part of Africa, as well as North Atlantic, areas dominated by dust emission and transport.

3.3 Background error from the dust ensemble. Driving the transport model with each meteorological member produces an ensemble of dust and emission forecasts. The spread of this ensemble is used to construct the background error covariance B for the assimilation. The augmented state variable contains both an emission flux linear coefficient for each dust emitting grid cell of the model, as well as the mass mixing ratios transported, since we cannot assume there are no transport and meteorological errors. This gives a flow-dependent, meteorologically consistent estimate of background error rather than a static prescribed covariance.

3.4 Emission retrieval. Using the new forward operators, EarthCARE observations are assimilated to optimise the emission field such that the resulting forecast reproduces the observations. The optimisation minimises the observation-space misfit through the model adjoint, returning a revised emission estimate consistent with both the observations and the non-spherical optical treatment. Applying this across multiple

[1] https://docs.jax.dev/en/latest/index.html

dust source regions and events yields a set of constrained emission estimates. We will additionally investigate whether the retrieved increments exhibit systematic, parameterisable structure that could be expressed as an explicit correction transferable to other models.

3.5 Cloud-relevant forecast and verification. Using the revised state estimate, we run forecasts and examine the resulting CCN and IN populations and cloud properties. Verification draws on contemporary retrievals of cloud and aerosol properties, with the specific products to be selected from currently available sources (for example PACE and EarthCARE CPR). This step tests whether a revised dust mass load produces a detectable change in modelled cloud-active particle populations and whether that change is consistent with observed cloud properties.

4. Computational requirements

Based on the previous project `spgramir`, we have determined that one 30-member WRF-CHEM ensemble at 27km horizontal resolution required approximately 250k SBUs for 1 week of simulation. Refining from 27 km to 20 km is a factor of about 1.35 in grid spacing, and the cost scales by roughly the cube of this factor once the required timestep reduction and overheads are included, giving a factor of about 2.5 and approximately 0.6M SBU per ensemble-week.

The core experiments comprise an ensemble meteorological forecast with WRF initialised and bounded by the ECMWF IFS ensemble, an ensemble dust transport run to construct the background error covariance, and control and data assimilation experiments across representative weeks and multiple dust source regions. At approximately 0.6M SBU per ensemble-week, these core ensemble experiments amount to approximately 24M SBU per year across the planned cases, regions and configurations.

The emission retrieval runs an iterative optimisation loop around the JAX transport model and its automatic-differentiation adjoint. The per-iteration cost is low relative to a coupled forecast, since the transport model is light and the JAX auto differentiation capabilities provide the tangent linear and adjoint models at no significant extra cost. The system is under active development and is expected to be operational by autumn, ahead of the project start, but it has so far been exercised only in an idealised setting, so the number of iterations required for convergence against real EarthCARE observations is not yet characterised. Rather than budget a speculative figure, we treat the retrieval as drawing on the headroom within the annual envelope. Once the core ensemble experiments and the cloud-resolving verification are accounted for, the remaining allocation is reserved for retrieval experiments. Should the retrieval prove more demanding than this reserve allows, we will compromise on ensemble resolution to stay within the allocation. Should the reverse prove true, we will expand the target period of the ensemble to increase the statistical robustness of our results. Characterising the per-retrieval cost is an explicit year-one objective that will help plan the next 2 years of the project.

The final stage tests whether the revised dust mass load produces a detectable change in cloud-active particle populations, using limited-area cloud-resolving simulations at convection-permitting resolution over selected cases. We estimate the cost by extrapolation from the ensemble benchmark. A single member at 20 km costs approximately 20,000 SBU per simulated week, one thirtieth of the 0.6M ensemble-week cost. Refining to a convection-permitting 4 km grid is a factor of 5 in resolution, scaling the cost by roughly the cube of this factor, approximately 125, and giving about 2.5M SBU per case-week. Running a small set of cases, each as a paired control and revised-dust simulation over a few representative events, we budget approximately 15M SBU per year, concentrated in years two and three once the constrained emission estimates are available.

Given our experience with the previous special project, we expect that the working datasets can comfortably fit inside the available HPCPERM and SCRATCH filesystems without the need for archival storage.

In summary, the known components are the core ensemble experiments at approximately 24M SBU per year and the cloud-resolving verification at approximately 15M SBU per year, together with spin-up, testing and storage overhead at approximately 10M SBU per year. Against an annual envelope of approximately 70M SBU, this reserves on the order of 21M SBU per year for the emission retrieval experiments. The request is therefore approximately 70M SBU per year, sustained at a stable level over three years for a total of approximately 210M SBU.

5. Acknowledgements

Parts of the past work mentioned (forward operators) are funded from the CERTAINTY Horizon Europe programme under Grant Agreement No 101137680 and are supported with computational resources from ECMWF through the special project DUSTAR and its extension (account spgramir). The new transport model is funded by the ESA-GLADIS project (Contract No. 4000147846/25/I/AG).

6. References

[1] N. Huneeus *et al.*, “Global dust model intercomparison in AeroCom phase I,” *Atmospheric Chem. Phys.*, vol. 11, no. 15, pp. 7781–7816, Aug. 2011, doi: 10.5194/acp-11-7781-2011.

[2] K. Yumimoto, I. Uno, N. Sugimoto, A. Shimizu, Z. Liu, and D. M. Winker, “Adjoint inversion modeling of Asian dust emission using lidar observations,” *Atmospheric Chem. Phys.*, vol. 8, no. 11, pp. 2869–2884, Jun. 2008, doi: 10.5194/acp-8-2869-2008.

[3] K. Solvik, S. G. Penny, and S. Hoyer, “4D-Var Using Hessian Approximation and Backpropagation Applied to Automatically Differentiable Numerical and Machine Learning Models,” *J. Adv. Model. Earth Syst.*, vol. 17, no. 4, p. e2024MS004608, 2025, doi: 10.1029/2024MS004608.