REQUEST FOR AN EMI R&D PROJECT 2027–2029

Country/Organisation: ……………………………ITALY…………………….….

Principal Investigator[1]: ………………………… Dr. Stefano Federico (cm4)………………….……..

Affiliation: ... National Research Council of Italy, Institute of atmospheric sciences and climate….(CNR-ISAC)……….

Address: via del Fosso del Cavaliere 100, 00133 Rome
…………………………………………

………………………………………………………………………………………….…………

Other researchers: ………………….. Claudio Transerici (cmn)
Rosa Claudia Torcasio (it85)………………….……..

………………………………………………………………………………………….…………

Project Title: …... Improvement on Numerical Weather Prediction (NWP) at the short and medium range for high impact meteorological events…….

…………………………………………………………………………………………….………

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.	**SPITFEDE**	
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 X	NO ☐

Computer resources required for project year:		**2027**	**2028**	**2029**
High Performance Computing Facility	[SBU]	20000000	20000000	20000000
Graphics Processing Unit Cluster-A	[GBU]			
Graphics Processing Unit Cluster-B	[GBU]			
Accumulated data storage (total archive volume)[2]	[GB]	100000	100000	100000

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

Continue overleaf.

Principal Investigator: ………… Stefano Federico (cm4)………………

Project Title: ……… Improvement on Numerical Weather Prediction (NWP) at the short and medium range for high impact meteorological events…….

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).

Introduction

This project is a continuation of the special project SPITFEDE that we did at the ECMWF in the years 2024-2026, with the addition of new observations and methods. Indeed, the special project during the period 2024-2026 was focused on the lightning and radar reflectivity data assimilation, on winds data assimilation and on the GNSS-ZTD/STD data assimilation. Moreover, the project 2024-2026 followed another special project whose focus was on lightning and radar reflectivity data assimilation.

We have shown in the previous projects that data assimilation of different data sources can be very useful to improve the weather forecast over Italy and the central Mediterranean. We will continue this study adding new types of observations (winds and lightning from satellites), extending statistics and case studies and testing new data assimilation methods.

The forecast will be focused on the short and medium range. More in detail, the forecast of convective environments, which will be investigated through lightning data assimilation, will be restricted at the short-range (3-6h after the ending of the assimilations phase), because these environments have short prediction ranges. For larger scale storms, for which we will consider the assimilations of winds, the forecast range will be longer (up to 96-120 h, depending on the case study). For GNSS data assimilation the forecast range will be between lightning and winds (24-48h).

Initial and boundary conditions for the experiments detailed below are derived from the output of the European Centre for Medium-Range Weather Forecast-Integrated Forecasting System (ECMWF-IFS) operational forecast, to simulate an operational environment. In the case of winds data assimilation, we will use an ensemble approach. In this case, initial and boundary conditions will be derived from the operational ensemble forecast of ECMWF Ensemble Prediction System (ECMWF-EPS).

For the whole project, we use the Weather Research and Forecasting (WRF) model. The WRF model is a limited-area numerical weather prediction (NWP) and atmospheric simulation model designed for both scientific research and operational forecasting. Developed primarily in the United States by organizations such as NCAR, NOAA, and NCEP, it is open-source and widely used globally. A general-purpose data assimilation software, which can also be used with WRF, is developed at CNR-ISAC and will be applied, together with the WRFDA (WRF data assimilation system) software in this project. Further details on the CNR-ISAC software can be found at the web page: http://meteo.artov.isac.cnr.it .

Winds data assimilation into the WRF model

Winds are key to understanding cloud and storm processes and the coupling between water, heat, and atmospheric circulation, yet they remain one of the major gaps in the Global Observing System (GOS) (Baker et al., 2014). Indeed, wind observations are sparse and unevenly distributed, especially in cloudy regions, on the mesoscale, and in the vertical dimension.

In this activity we want to investigate the role of winds observation for two main satellites observing sources: WIVERN and ASCAT.

The WInd VElocity Radar Nephoscope (WIVERN; Battaglia et al., 2022) has been recommended to be launched for the 11th Earth Explorer mission. Earth Explorer missions are proposed by the scientific community to demonstrate how breakthrough technology can deliver an astounding range of scientific findings about our planet.

WIVERN, will carry a 94 GHz radar with a conically scanning 800 km swath. It will provide, for the first time, profiles of in-cloud winds and precipitation on a global scale, with a revisit time of 1.5 d at the equator. The implications of this unprecedented sampling capability have been recently discussed in Tridon et al. (2023) and Scarsi et al. (2025). WIVERN measurements will have a vertical resolution of approximately 600 m and an instantaneous footprint of less than 1 ×1 km^2.

Wind measurements will then be averaged along the scanning direction over 5 km or more to reduce measurement noise. WIVERN will make a bridge between the synoptic scale and the mesoscale thanks to the wide swath and the high spatial resolution.

A recent paper (Federico et al., 2026a) showed the potential of WIVERN data assimilation for the prediction of the Medicane Ianos trajectory, nevertheless results refer to only a case study and cannot be extrapolated in general. In this special project we want to focus on a second Medicane (Jolina, occurred in March 2026) and on a different type of Mediterranean storm (for example a derecho wind storm, to be defined during the project). The approach used to assimilate wind observations is that of the ensemble 3DVar (En3DVar), in which the background error matrix is aware of the meteorological conditions of the day. However, this requires carrying simulations including the whole ensemble. Other sensitivity tests to the background and observations error matrix, and to the data thinning will be also considered in this activity.

The Advanced Scatterometer (ASCAT) is one of the European instruments carried on MetOp and is used to determine information about the wind for use primarily in weather forecasting and climate research. Data from ASCAT will also find applications in other areas such as the monitoring of land- and sea-ice, snow cover and soil moisture.

ASCAT uses radar to measure the electromagnetic backscatter from the wind-roughened ocean surface, from which data on wind speed and direction can be derived. The measuring principle relies on the fact that winds over the sea cause small-scale disturbances of the sea surface, which modify its radar backscattering characteristics in a particular way. These backscattering properties are well known and are dependent on both the wind speed over the sea and the direction of the wind with respect to the point from which the sea surface is observed.

Surface winds are very important in Medicanes (Miglietta et al., 2025) and data assimilation of surface winds can play an important role for improving the prediction of these storms, including the trajectory and intensity. In this project we will use an En3DVar to study the predictability of two Medicanes: Ianos, occurred in mid-September 2020 and Jolina, occurred in March 2026.

The computational estimation of these experiments starts for a basic simulation of WRF lasting few days. The spatial resolution of the model will be around 4-5km and the number of grid points of the order of 500 in both WE and SN directions. The number of vertical levels is in the range 50-60. This simulation for a few days requires about 15,000 SBU. Considering that we will run a whole ensemble of 51 member (50 perturbed and 1 unperturbed) a total of 750,000 SBU-1,000,000 SBU is required for each ensemble simulation. Now at least four experiments will be considered, and different observing systems will also be considered (no ASCAT and WIVERN DA, one data source data assimilation, both sources data assimilation for the common case study). In addition, sensitivity tests to changing the background and observations error matrices are envisaged as in Federico et al. (2026a). All in all, this activity will require about 30,000,000 SBU.

It is finally noticed that, as WIVERN is not flying, pseudo-observations will be used for this instrument; these pseudo-observations will be generated following Federico et al. (2026a).

GNSS atmospheric water vapour assimilation

Atmospheric water vapour is a fundamental element of weather forecasting, nevertheless its accurate observation is difficult, mainly because of its high spatiotemporal variability.

A reliable way to estimate water vapour is through Global Navigation Satellite Systems (GNSS) satellites, whose signals are received from ground-based stations allowing to estimate Zenith Total Delay (ZTD). ZTD can be easily related to Precipitable Water Vapour (PWV).

The GNSS observation at the zenith uses delays from different directions to improve the estimate of the delay in the vertical direction, thus strengthening the solution. This process, however, reduces the number of observations available for each GNSS receiver. Anyway, the high quality of GNSS observations in the zenith direction has become a reference for other instruments and has been widely assimilated in NWP models worldwide.

Estimating the delay in different directions poses challenges for the convergence of the solution and for errors in the retrieved slant total delay (STD). Similarly, while the assimilation of the zenith delay for GNSS receivers is well-established, the assimilation of the delay in the inclined directions remains largely unexplored.

In this activity, GNSS information along slant paths is assimilated into the WRF model. Two approaches are shown: the first considers the assimilation of the delay along slant paths, the second the assimilation of precipitable water vapour along slant paths.

In the first approach, the assimilation of STD is done using tropospheric gradients in the East and North directions as in Torcasio et al. (2026).

The second approach considers PWV data assimilation along slant paths (PWVS). In this case, the STD signal is converted in precipitable water vapour and assimilated in WRF.

In this project the two methods will be used to predict the precipitation and water vapour at the medium range (24-48h) with a specific focus on the short-range. A period of two months will be considered, and the forecast will be updated every 6h (4 simulations per day). In addition, comparison with a control simulation, without the use of GNSS observations and with a simulation using GNSS-ZTD will be explored. The spatial resolution of the model is of the order of 2-3 km and 50-70 vertical levels will be used. Sensitivity tests to data thinning will be also considered.

The NMC method (Parrish and Derber, 1992) will be used for the computation of the background error matrix. This method calculates statistics considering the forecast differences between 24 h and 12 h prediction over a certain period. Two separate calculations for the two months will be done. For each month two simulations per day starting one 12h after the other lasting 24h each are needed.

As for the computational requirement, based on experience with ECMWF-HPC, one simulation will require between 15,000 and 20,000 SBU, depending on the final configuration used. Considering that 4 simulations per day are required, the period of two months, the different configurations of the model, and that errors that always occur (about 10% of the SBU requested) a total of 20,000,000 SBU is estimated for this activity.

Assimilation of MTG-LI into the WRF model

Many studies are available in literature on the improvement of precipitation forecast at the short-range (3-6h) through lightning data assimilation (LDA). At CNR-ISAC many experiments of LDA have been done over Italy since more than 10 years. Lightning has been assimilated mostly by nudging, considering the method of Fierro et al. (2012) both in the Regional Atmospheric Modeling System (RAMS) and WRF models. The positive impact of lightning data assimilation on the precipitation forecast has been evaluated alone or in conjunction with other data sources for some case studies or for longer periods (e.g. Federico et al., 2019; Federico et al., 2026b). The impact of LDA on precipitation prediction over the sea (Torcasio et al. 2021) and on lightning forecast (Federico et al. 2024) have also been investigated. Along with nudging, also the 3DVar method was applied for LDA (e.g. Federico et al. 2022; Torcasio et al., 2023).

Results show a positive and significant impact of LDA on precipitation prediction, especially for the first 3h of forecast, the main drawback being the increase in false alarms.

All the above-mentioned experiments were performed assimilating lightning data acquired from ground-based networks.
Recently, we started to assimilate lightning data provided by the Lightning Imager (LI) onboard the Meteosat Third Generation (MTG) satellite. Data from MTG-LI are available starting from July 2024. In this project we will assimilate lightning from MTG-LI into the WRF model considering different experiments. All experiments will be set up in a Very Short-term Forecast (VSF) mode. Simulations will last 12h, with the first 6h for spin-up and data assimilation and the last 6h of each run as forecast phase. To simulate an operational-like environment, LDA will stop half an hour before the beginning of the forecast phase.
The first experiment will consider MTG-LI flashes as they are received from the sensor, without any adjustment. A second experiment will assimilate MTG-LI flashes after a parallax error correction is applied. To recover at least in part the issue of increased false alarms after LDA, a third experiment will consider the suppression of spurious convection. To do this, MTG-LI flashes need to be assimilated by 3DVar. All experiments will consider a two-months period: a fall month and a summer month. This is because in Italy summer and fall are the two seasons most affected by flashes (Federico et al., 2026b). The model grid will have a 3km spatial horizontal resolution a 650x650 grid points in the horizontal direction and 55 vertical levels. Since every simulation will produce a 6h forecast, four simulations are needed to cover a whole day for each experiment. To assess the improvement given by LDA, simulations with LDA will also be compared with control runs, without LDA. Furthermore, the 3DVar method requires the calculation of a proper background error matrix. This calculation will be done by means of the NMC method (Parrish and Derber, 1992). NMC method calculates statistics considering the forecast differences between 24 h and 12 h prediction over a certain period. Two separate calculations for the two months will be done. For each month two simulations per day starting one 12h after the other lasting 24h each are needed.
To achieve the whole experiment, it is estimated that about 8 millions of SBU are needed. Indeed, to simulate 1 experiment we need about 6000 SBU (estimate based on past experience with ECMWF-HPC). We need to run 3 different experiments for a period of 2 months. Four simulations are required to accomplish the forecast of a whole day. Considering also that errors always occurs and, from past experience, 10% of the SBU assigned are lost by errors, a total of 10,000,000 SBU is envisaged for this activity.

Storage requested

Up to now we discussed the computational resources required for this project. As for the storage we have a long experience with the usage of ECFS, which will be used in this project too. A total space of 100 TB is required to manage the large datasets produced in the project. Part of the post-elaboration will be accomplished in our computational facility in Rome.

References

Baker, W. E., Atlas, R., Cardinali, C., Clement, A., Emmitt, G. D., Gentry, B. M., Hardesty, R. M., Källén, E., Kavaya, M. J., Langland, R., Ma, Z., Masutani, M., McCarty, W., Pierce, R. B., Pu, Z., Riishojgaard, L. P., Ryan, J., Tucker, S., Weissmann, M., and Yoe, J. G.: Lidar-Measured Wind Profiles: The Missing Link in the Global Observing System, B. Am. Meteorol. Soc., 95, 543– 564, https://doi.org/10.1175/BAMS-D-12-00164.1, 2014.

Battaglia, A., Martire, P., Caubet, E., Phalippou, L., Stesina, F., Kollias, P., and Illingworth, A.: Observation error analysis for the WInd VElocity Radar Nephoscope W-band Doppler conically scanning spaceborne radar via end-to-end simulations, Atmos. Meas. Tech., 15, 3011–3030, https://doi.org/10.5194/amt-15-3011-2022, 2022.

Fierro, A.O., Mansell, E., Ziegler, C., MacGorman, D., 2012. Application of a lightning data assimilation technique in the WRFARW model at cloud-resolving scales for the tornado outbreak of 24 May 2011. Mon. Weather Rev. 140, 2609–2627.

Federico, S., Torcasio, R.C., Avolio, E., Caumont, O., Montopoli, M., Baldini, L., Vulpiani, G., Dietrich, S., 2019. The impact of lightning and radar reflectivity factor data assimilation on the very short-term rainfall forecasts of RAMS@ISAC: Application to two case studies in Italy. Nat. Hazards Earth Syst. Sci. 19, 1839–1864.

S. Federico, R.C. Torcasio, S. Dietrich, Improvement of quantitative precipitation forecast at the short range through lightning data assimilation, in Precipitation Science: Measurements, Remote Sensing, Microphysics, and Modeling, Edited by: Silas Michaelides, Elsevier 2022, Pager 661-668, ISBN: 978-0-12-822973-6, https://doi.org/10.1016/B978-0-12-822973-6.00026-3

Federico, S., Torcasio, R.C., Popova, J., Sokol, Z., Pop, L., Lagasio, M., Lynn, M.H., Puca, S., Dietrich, S., 2024. Improving the lightning forecast with the WRF model and lightning data assimilation: results of a two-seasons numerical experiment over Italy. Atmos. Res. 304, 107382. ISSN 0169-8095. https://doi.org/10.1016/j. atmosres.2024.107382.

Federico, S., Torcasio, R. C., Transerici, C., Montopoli, M., Cambiotti, C., Manconi, F., Battaglia, A., and Pourshamsi, M.: Assimilating WIVERN windpseudo-observations in WRF model: an application to the outstanding case of the Medicane Ianos, Weather Clim. Dynam., 7, 165–183, https://doi.org/10.5194/wcd-7-165-2026, 2026a.

Stefano Federico, Rosa Claudia Torcasio, Claudio Transerici, Eugenio Realini, Xiangyang Song, Giovanna Venuti, Forecasting convective precipitation over northern Italy: A comparison of lightning and GNSS-ZTD data assimilation, Atmospheric Research, Volume 331, 2026b, 108687, ISSN 0169-8095, https://doi.org/10.1016/j.atmosres.2025.108687.

Miglietta, M. M., and Coauthors, 2025: Defining Medicanes: Bridging the Knowledge Gap between Tropical and Extratropical Cyclones in the Mediterranean. Bull. Amer. Meteor. Soc., 106, E1955–E1971, https://doi.org/10.1175/BAMS-D-24-0289.1.

Parrish, D.F.; Derber, J.C. The National Meteorological Center's Spectral Statistical Interpolation analysis system. *Mon. Weather Rev.* **1992**, *120*, 1747–1763.

Scarsi, F. E., Battaglia, A., Maahn, M., and Lhermitte, S.: How to reduce sampling errors in spaceborne cloud radar-based snowfall estimates, The Cryosphere, 19, 4875–4892, https://doi.org/10.5194/tc-19-4875-2025, 2025.

Torcasio, R.C., Federico, S., Comellas Prat, A., Panegrossi, G., D'Adderio, L.P., Dietrich, S., 2021. Impact of lightning data assimilation on the short-term precipitation forecast over the Central Mediterranean Sea. Remote Sens. 13, 682. https://doi.org/10.3390/rs13040682.

Torcasio, R.C., Papa, M., Del Frate, F., Dietrich, S., Toffah, F.E., Federico, S., 2023. Study of the intense meteorological event occurred in September 2022 over the Marche Region with WRF Model: impact of lightning data assimilation on rainfall and lightning prediction. Atmosphere 14, 1152.

Torcasio, R.C., Realini, E., Crespi, M. et al. Assimilation of anisotropic GNSS observations in the WRF model: preliminary results of the NEW-ARGENT project.Rend. Fis. Acc. Lincei (2026). https://doi.org/10.1007/s12210-025-01399-1

Tridon, F., Battaglia, A., Rizik, A., Scarsi, F. E., and Illingworth, A.: Filling the Gap of Wind Observations Inside Tropical Cyclones, Earth and Space Science, 10, e2023EA003099, https://doi.org/10.1029/2023EA003099, 2023.