REQUEST FOR A SPECIAL PROJECT 2026–2028

MEMBER STATE:	ITALY
Principal Investigator[1]:	FRANCESCO BARBARIOL
Affiliation:	Italian National Research Council – Institute of Marine Sciences (CNR-ISMAR)
Address:	Arsenale – Tesa 104, Castello 2737/F, Venice, Italy
Other researchers:	Luigi Cavaleri, Luciana Bertotti, Alvise Benetazzo, Christian Ferrarin, Marco Bajo
Project Title:	Enhancing efficiency and accuracy of coastal storm surge and wave forecasts by means of ensemble clustering and wind forcing correction.

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.)	2026	
Would you accept support for 1 year only, if necessary?	YES X	NO ☐

Computer resources required for project year:		**2026**	**2027**	**2028**
High Performance Computing Facility	[SBU]	5,000,000	5,000,000	5,000,000
Accumulated data storage (total archive volume)[2]	[GB]	25000	25000	25000

EWC resources required for project year:		**2026**	**2027**	**2028**
Number of vCPUs	[#]	70	70	70
Total memory	[GB]	350	350	350
Storage	[GB]	25000	25000	25000
Number of vGPUs[3]	[#]	2	2	2

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 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.

Principal Investigator: FRANCESCO BARBARIOL

Project Title: Enhancing efficiency and accuracy of coastal storm surge and wave forecasts by means of ensemble clustering and wind forcing correction.

Extended abstract

Motivation

Protecting civil and economic interests from storm surges and stormy sea states along the coastlines is particularly important, even more in a context of changing climate and increasing sea level.

The first component of an early-warning system for coastal hazards is a numerical weather prediction system capable of providing timely and accurate weather and marine forecasts. In enclosed or semi-enclosed basins as the Adriatic Sea, largely surrounded by land orography at the northern end of the Mediterranean Sea, the weather forecasts from global systems as the IFS from ECMWF can be affected by a systematic underestimate of the surface winds that are used to force hydrodynamic and spectral wave numerical models, especially in coastal areas. These are employed to enhance space and time resolution of forecasts in local/regional areas, also including processes that cannot be resolved at global scale. To improve local/regional wind forecasts, and therefore the accuracy of surge and wave forecasts, correction strategies have been explored and tested in the past, which basically statistically downscale the global wind forecasts thanks to observations as those provided by satellite scatterometers. These approaches have used, e.g., the linear scaling of wind speed or the Quantile-Quantile matching of probability distributions of surface winds, including speed and direction variability of the model-minus-observation bias (Cavaleri and Sclavo, 2006; Barbariol et al., 2022; Benetazzo et al., 2022). Other approaches can be implemented as well, like learning-based approaches that can be trained to account for, for instance, the spatial patterns of surface wind on the sea, their seasonal variability or the different model-minus-observation bias of winds blowing from or to the coast.

Besides, to account for the uncertainty of atmospheric forecasts and provide probabilistic-based marine forecasts of surge and waves, ensemble winds are used. However, the high-resolution downscaling of surge and wave ensemble forecasts is computationally demanding and sometimes not feasible for local/regional meteorological centres when the full ensemble is accounted for. In the past, ensemble reduction strategies have been proposed to downscale the IFS ensemble at regional scale (Molteni et al., 2001), using clustering algorithms that allow to select a reduced number of representative members. Dealing with a reduced number of forcings and numerical simulations though maximizing the information provided by the ensemble could make local/regional forecasts of surge and waves more efficient.

Tasks

1. **SEA SURFACE WIND CORRECTION**

   The statistical correction of global IFS winds on the surface of enclosed basins (Mediterranean and Adriatic Sea, in particular) will be enhanced by using learning-based methods. A rationale for correction is to let the method learn how the model-minus-observation error in wind speed and direction depends upon the location (distance from the coast, basin area, etc) and wind regime. After training, the method is applied to correct forecast winds. This analysis will rely on a multi-year dataset of IFS wind speed forecasts (U10, V10 components of the 10-meter wind speed) to be compared to an observational dataset of surface wind speed components from satellite scatterometers. Surge and wave numerical simulations will then be run to verify the improvement in model accuracy achieved by using corrected wind forcing with respect to uncorrected wind forcing.

2. **MODEL ENSEMBLE REDUCTION**

   The ensemble reduction of sea surface wind speed will be explored by means of clustering algorithms, building upon the methodology of Molteni et al. (2001). Pre-processing based on Machine Learning (e.g., using Principal Component Analysis or Auto-Encoders) will be tested in order to reduce the dimension of the input dataset and enhance clustering performance. The aim is to obtain a reduced wind ensemble, composed of a small number (around ten) of representative members, that resembles the characteristics of the original ensemble (e.g., mean, spread and probability distribution). The reduced wind ensemble will be used to force the surge and wave models and obtain a reduced wave ensemble, which will be compared to the reduced ensemble obtained by clustering the full wave ensemble obtained by using all the wind ensemble members. Different marine storm events will be tested in order to account for different predictabilities (small or large spreads), thus verifying the enhancement in surge and wave forecast efficiency, preserving a proper level of forecast accuracy.

Technical characteristics of the code and data used

The numerical system used to simulate the storm surge and waves is the SHYFEM-WAVEWATCH III (SHYFEM-WW3) modelling framework, which couples the SHYFEM hydrodynamic model and the WW3 spectral wave model, both operating on unstructured computational grids. The SHYFEM hydrodynamic model (System of HydrodYnamic Finite Element Modules - https://github.com/SHYFEM-model/shyfem) is an open-source model created and developed at CNR-ISMAR for the modelling of marine hydrodynamics (Umgiesser & Bergamasco, 1993). In the past, it has been used in numerous works all over the world, with particular regard to the Mediterranean and Adriatic basins. The model solves the shallow water equations with a semi-implicit numerical scheme and an unstructured computational grid, with spatial resolution ranging from a few km in the open sea to 500 m near the coast. The model has been coupled to the WW3 wave model. WW3 (https://github.com/NOAA-EMC/WW3) is a third-generation spectral numerical model for the simulation of wind-waves. It is developed by an international scientific community, which periodically implements the most recent research results on marine wave dynamics. The model numerically solves the spectral action density balance equation, expressed in terms of directional spectra (WW3DG, 2019). Several parameterizations of the source terms allow to calculate the generation of waves by the wind, the decay from dissipative processes (breaking), the nonlinear wave-wave interactions and the transformation of the waves near the coast (including the effects of shoaling, refraction and seabed friction).

The correction of wind and clustering of the ensemble will be performed with scripting in the python language with toolbox and packages for data analysis. Some of these packages (e.g., pytorch library for machine learning) benefit from the use of GPUs that can be granted by the access to the European Weather Cloud resources.

The data used for the project are the IFS-ECMWF wind speed and mean sea level pressure forecasts of the deterministic and ensemble systems, over the Mediterranean Sea area.

References

Barbariol et al, (2022). *Wind-wave forecasting in enclosed basins using statistically downscaled global wind forcing*. Frontiers in Marine Science

Benetazzo et al, (2022). *Correction of ERA5 Wind for Regional Climate Projections of Sea Waves*. Water

Cavaleri and Sclavo, (2006). *The calibration of wind and wave model data in the Mediterranean Sea*. Coastal Engineering

Molteni et al, (2001). *A strategy for high-resolution ensemble prediction. I: Definition of representative members and global-model experiments*. Q. J. Royal Meteorological Society.

Umgiesser and Bergamasco (1993). *A staggered grid finite element model of the Venice Lagoon*. Finite Elements in Fluids, Pineridge Press, Barcelona

The WAVEWATCH III Development Group (WW3DG). *User manual and system documentation of WAVEWATCH III version 5.16*. Tech. Note 329, NOAA/NWS/NCEP/MMAB, College Park, MD, USA, 326 pp. + Appendices.