ENHANCING STORM SURGE DOWNSCALING: A COMPARATIVE STUDY OF MACHINE LEARNING AND DYNAMICAL MODELING IN THE NORTHERN ADRIATIC SEA

Rodrigo Campos-Caba[1,2], Jacopo Alessandri[6], Paula Camus[3], Andrea Mazzino[4,5], Francesco Ferrari[4,5], Ivan Federico[6], Michalis Vousdoukas[7,8], Massimo Tondello[9], Giovanni Coppini[6], and Lorenzo Mentaschi[1,2,6]

1.- Department of Physics and Astronomy (DIFA), University of Bologna.
2.- Interdepartmental Research Centre for Environmental Sciences (CIRSA), University of Bologna.
3.- Departamento de Ciencias y Técnicas del Agua y del Medio Ambiente, University of Cantabria.
4.- Department of Civil, Chemical and Environmental Engineering, University of Genoa.
5.- Istituto Nazionale di Fisica Nucleare, Sezione di Genova.
6.- CMCC Foundation – Euro-Mediterranean Center on Climate Change.
7.- Department of Marine Sciences, University of the Aegean.
8.- MV Coastal and Climate Research Ltd.
9.- HS Marine SrL.

4th International Workshop on Waves, Storm Surges, and Coastal Hazards
September, 2025

STORM SURGE DOWNSCALING

Effective storm surge prediction is vital for safeguarding coastal communities and enhancing disaster preparedness.

While machine learning is increasingly used for storm surge downscaling, direct comparisons with high-resolution models—especially for extreme events—are still rare.

This study compares advanced dynamic modeling with ML techniques (ML emulators) to improve storm surge time-series reconstruction in the northern Adriatic Sea.

STORM SURGE

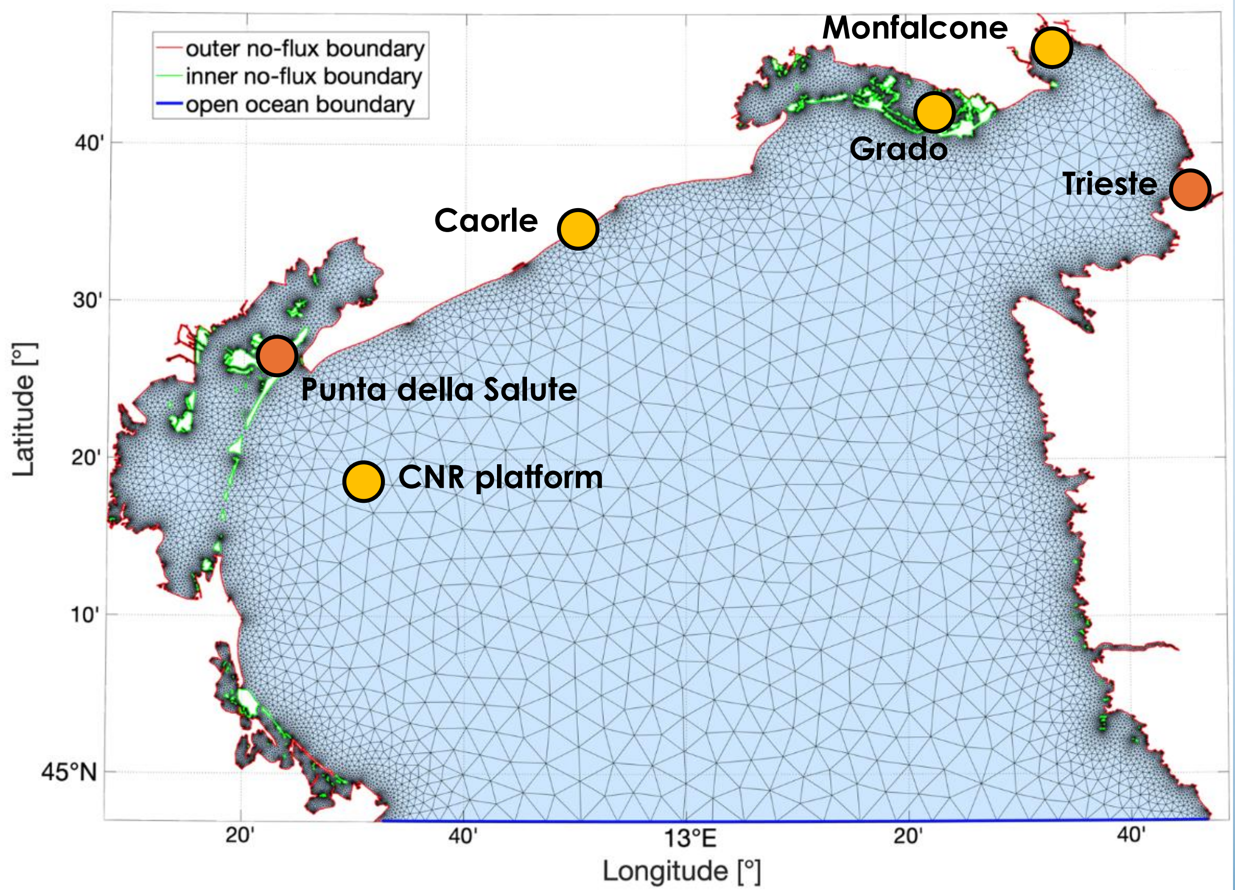
MACHINE LEARNING DOWNSCALING

Predictand (target): Storm surge time-series from observations in Punta della Salute and Trieste (consistent and long-term hourly observed data).

Predictors (features): Sea level height (Copernicus Sea Physics Reanalysis), tides (FES2014), wind and mean sea level pressure fields (ERA5).

In the **training period** 80% (28 years) of the data is used. The remaining 20% (6 years) was split into 3 years for **validation** (to select the best run of each ML emulator, which were executed 40 times, considering the slope of the linear fit) and 3 years for **testing** (evaluate performance of the ML model).

Benchmark: High-resolution model (50 m resolution at the coastline) generated through numerical simulations (dynamic downscaling) in the northern Adriatic Sea (Campos-Caba et al., 2024). The numerical simulations were carried out using the SHYFEM-MPI model (Umgiesser et al., 2004; Micaletto et al., 2022).

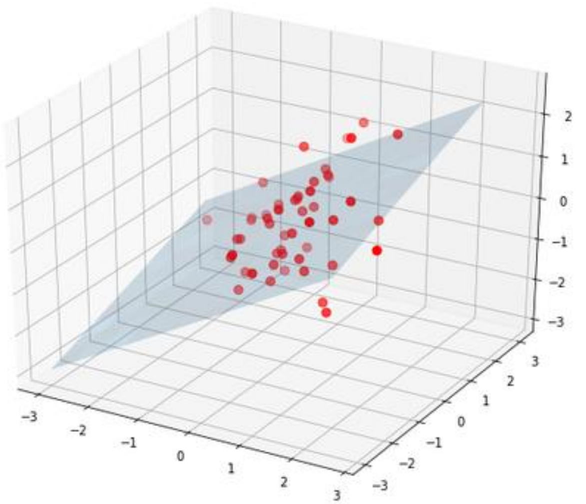


Unstructured grid used for the dynamic downscaling and locations for validation.

STORM SURGE

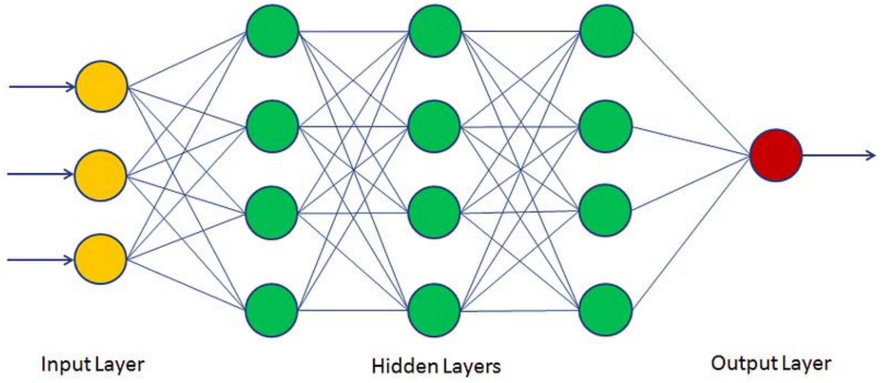
MACHINE LEARNING DOWNSCALING

MULTIVARIATE LINEAR REGRESSION (MLR)



Source: www.medium.com

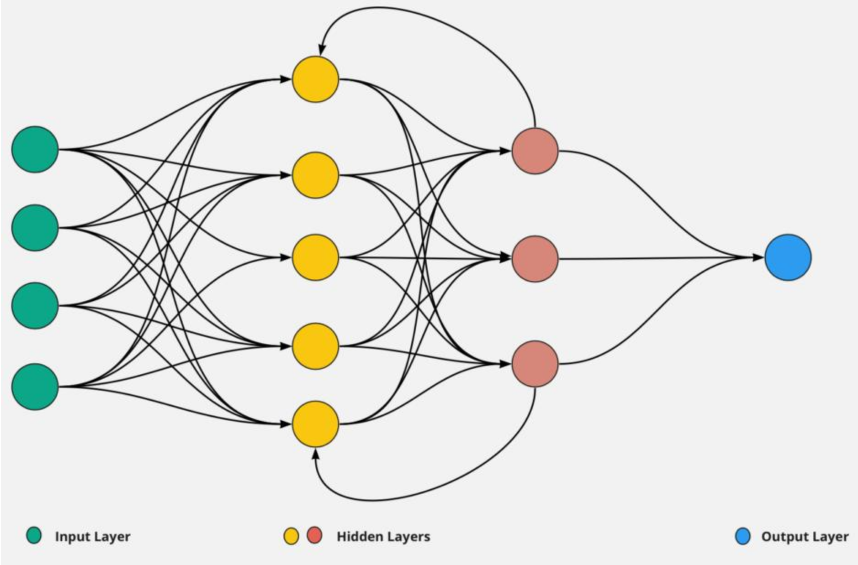
MULTILAYER PERCEPTRON (MLP)



Source: www.machinelearninggeek.com

RECURRENT NEURAL NETWORK (RNN)
HYBRID RECURRENT NEURAL NETWORK (RNNh)

LONG SHORT-TERM MEMORY NETWORK (LSTM)
HYBRID LONG SHORT-TERM MEMORY NETWORK (LSTMh)



Source: www.dataaspirant.com

Hybrid emulators incorporate both, a Neural Network (NN) layer and a linear layer, combining their outputs to enhance predictive accuracy.

STORM SURGE

MACHINE LEARNING DOWNSCALING

As loss functions (function that quantifies the difference between the predicted outputs and the actual target values) both MSE and MADc^2 (Campos-Caba et al., 2024), were applied across all models.

Also considered for performance analysis

$$MADp = \overline{|S_{prc} - O_{prc}|}$$

$$MADc = \overline{|S - O|} + MADp$$

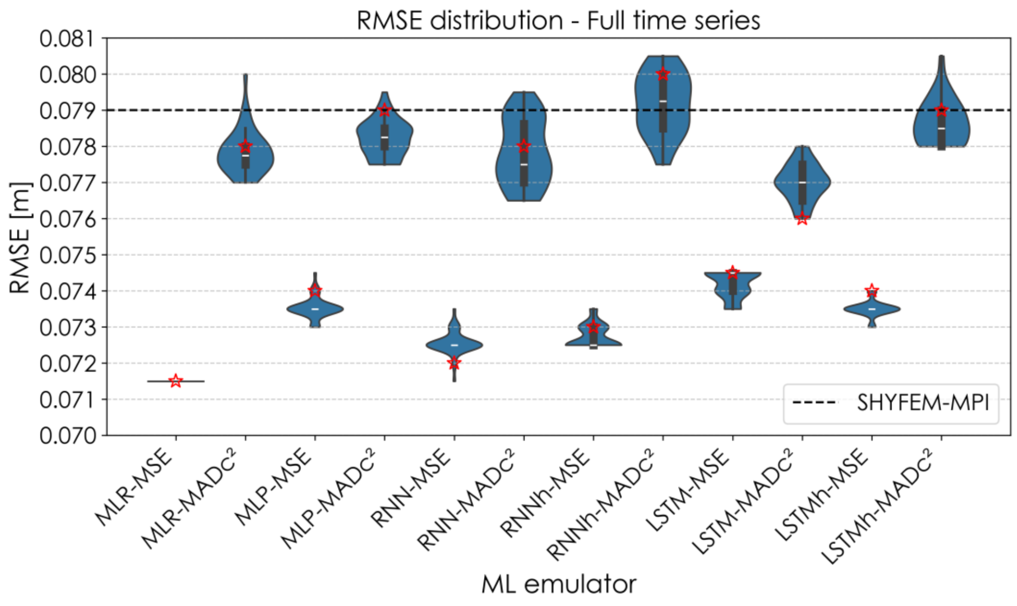
Only quadratic loss functions were considered due to their differentiability and smooth gradient profile, which creates an optimization landscape that is better suited for gradient descent.

ML emulator configuration	Loss function	Acronym
MLR model	MSE	MLR-MSE
MLP model		MLP-MSE
RNN model		RNN-MSE
RNNh model		RNNh-MSE
LSTM model		LSTM-MSE
LSTMh model		LSTMh-MSE
MLR model	MADc^2	MLR- MADc^2
MLP model		MLP- MADc^2
RNN model		RNN- MADc^2
RNNh model		RNNh- MADc^2
LSTM model		LSTM- MADc^2
LSTMh model		LSTMh- MADc^2

Each model was executed 40 times.

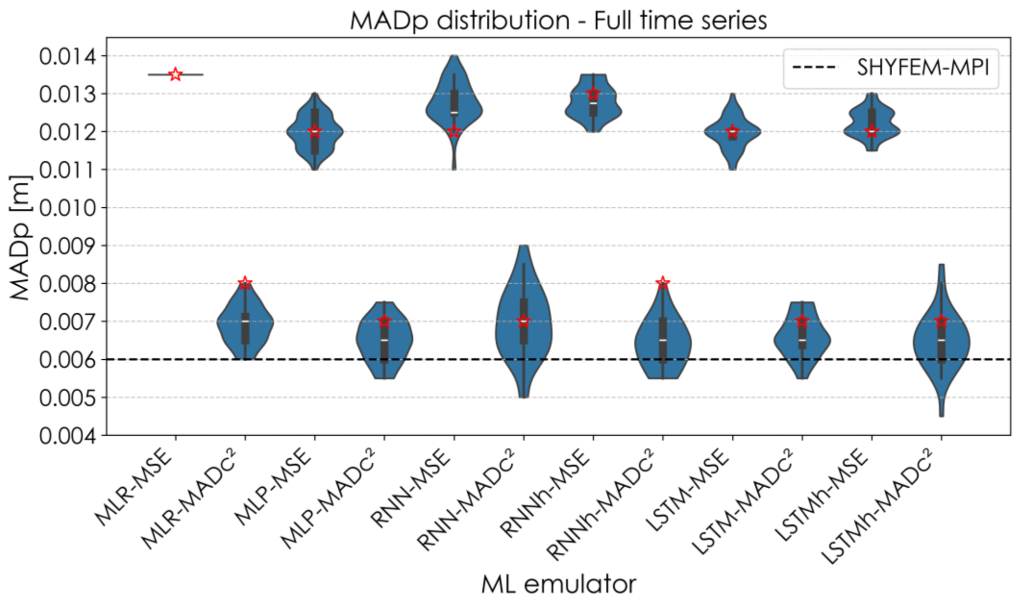
STORM SURGE

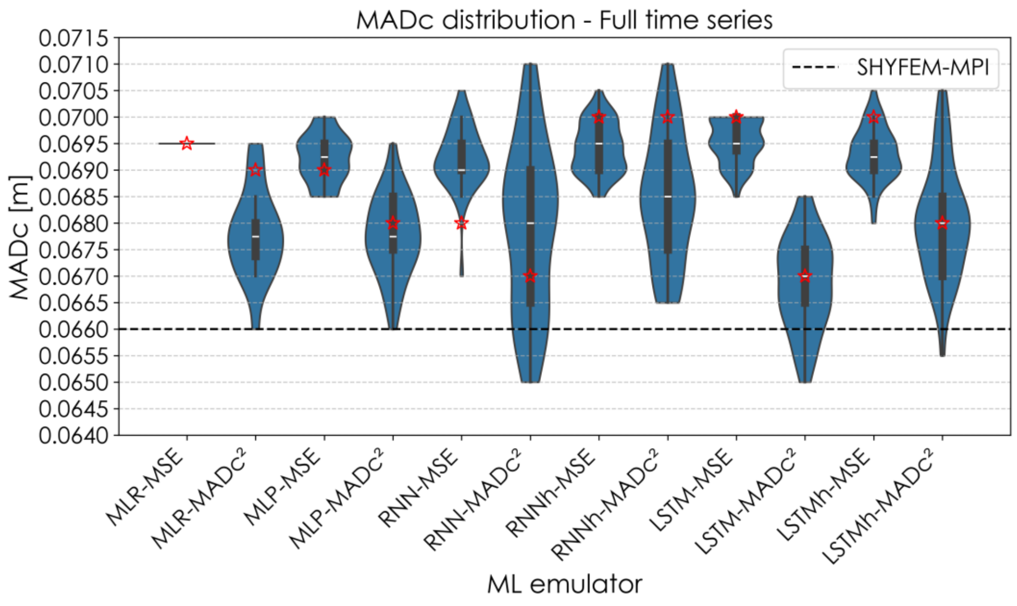
MACHINE LEARNING DOWNSCALING



Violin plots for the total amount of data:

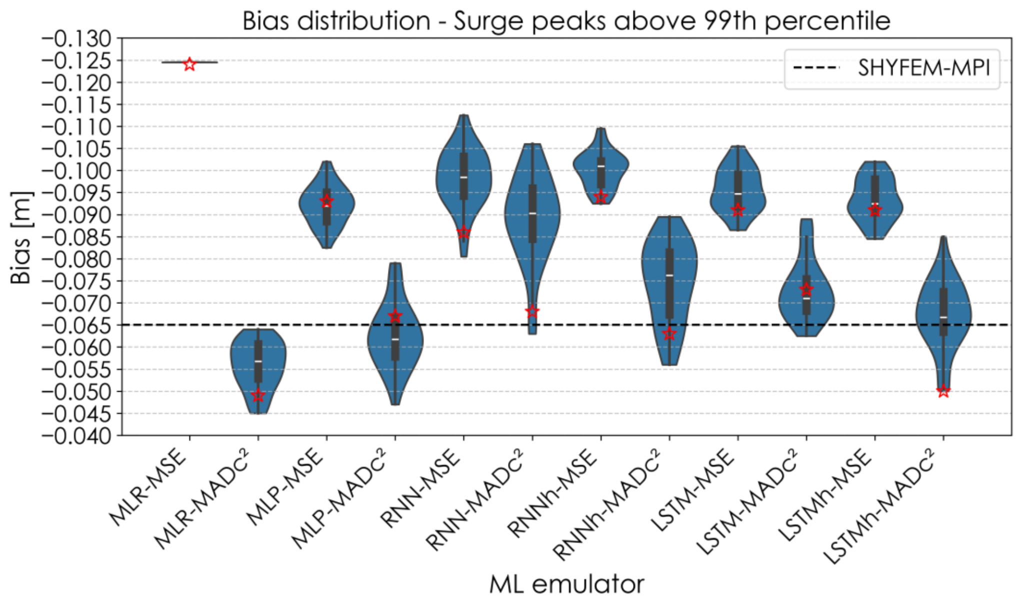
- Values were obtained for each run at each location and subsequently averaged.
- **Values below the dashed black line**: indicate better performance of the ML emulators compared to SHYFEM-MPI.
- **Values above the dashed black line**: indicate lower performance of the ML emulators compared to SHYFEM-MPI.
- Red star: represents the mean value across locations obtained for the best run during the validation period.





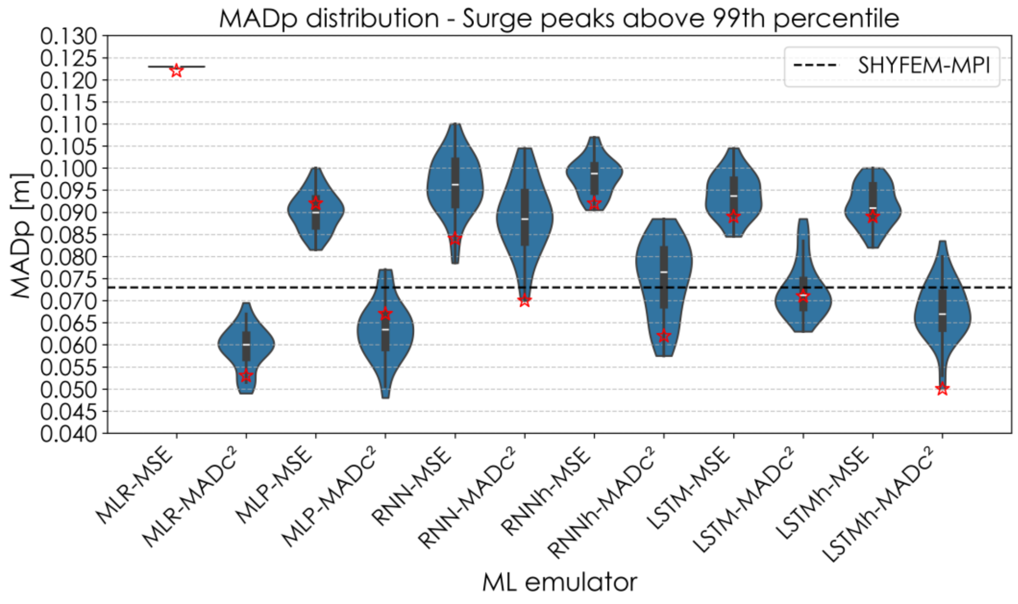
STORM SURGE

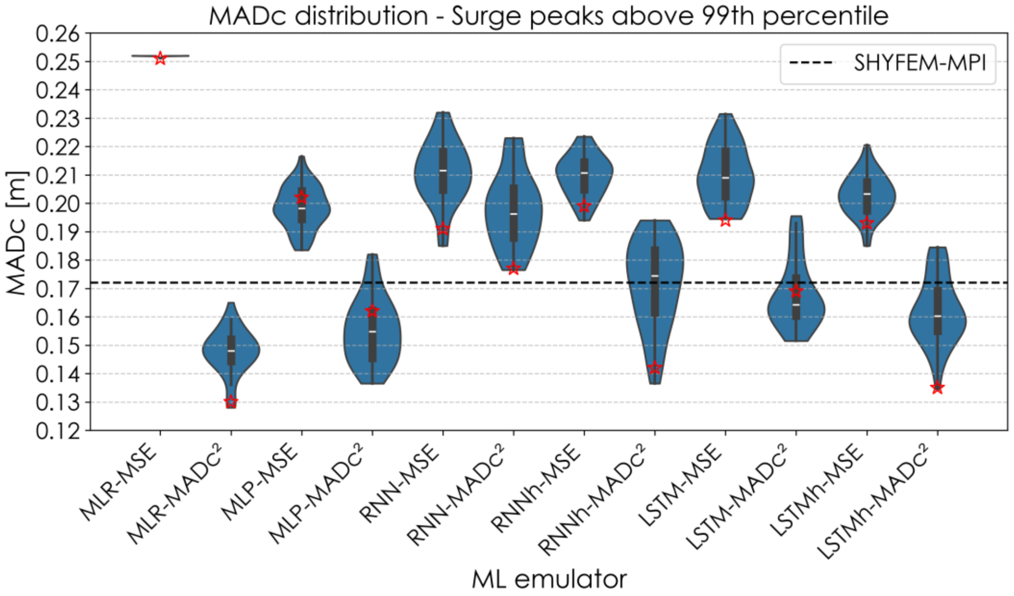
MACHINE LEARNING DOWNSCALING



Violin plots for surge peaks above the 99th percentile:

- Values were obtained for each run at each location and subsequently averaged.
- **Values below the dashed black line**: indicate better performance of the ML emulators compared to SHYFEM-MPI.
- **Values above the dashed black line**: indicate lower performance of the ML emulators compared to SHYFEM-MPI.
- Red star: represents the mean value across locations obtained during the validation period (validation-based selection).





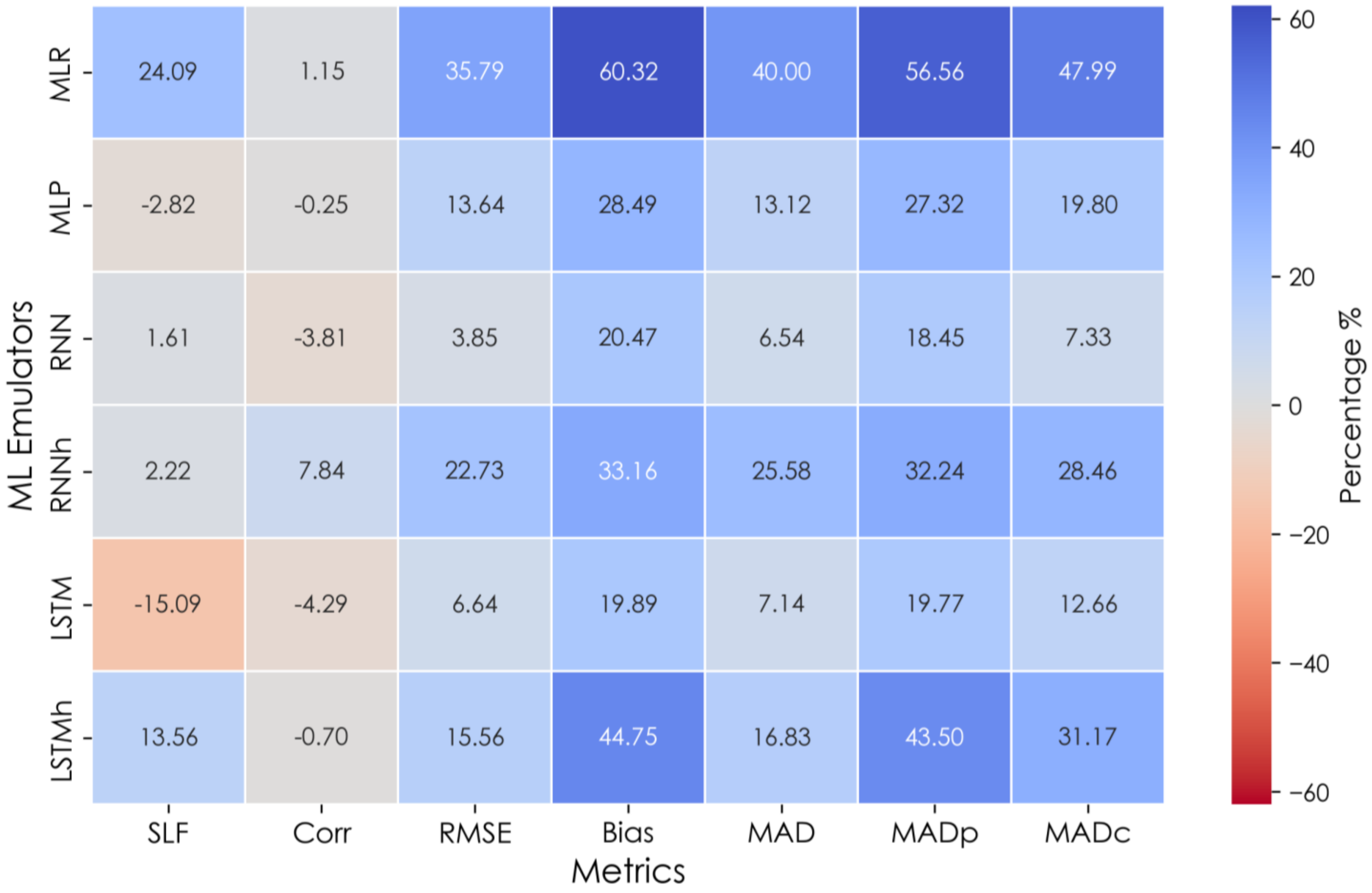
STORM SURGE

MACHINE LEARNING DOWNSCALING

Heat map:

- Averaged values obtained for the selected run represented by a red star in violin plot for the surge peaks above the 99th percentile.
- **Positive values (cold colors):** improvements using the $MADc^2$ loss function.
- **Negative values (warm colors):** decrease in performance with $MADc^2$ loss function.

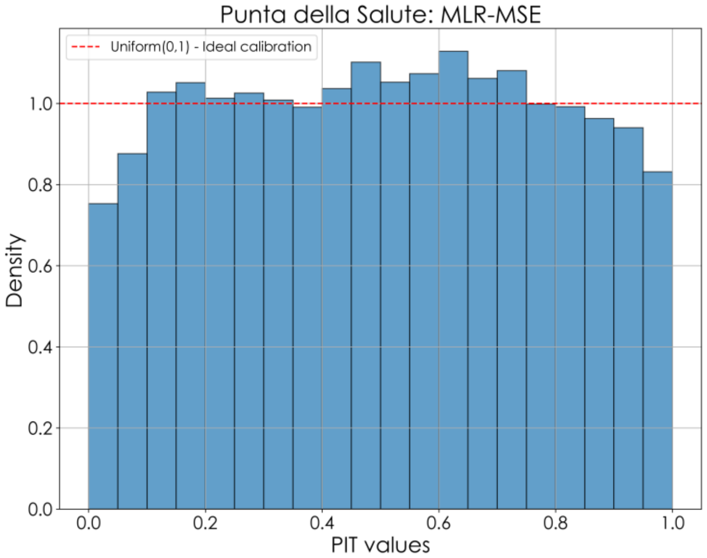
Mean percentage variation in performance metrics for surge peaks above the 99th percentile.

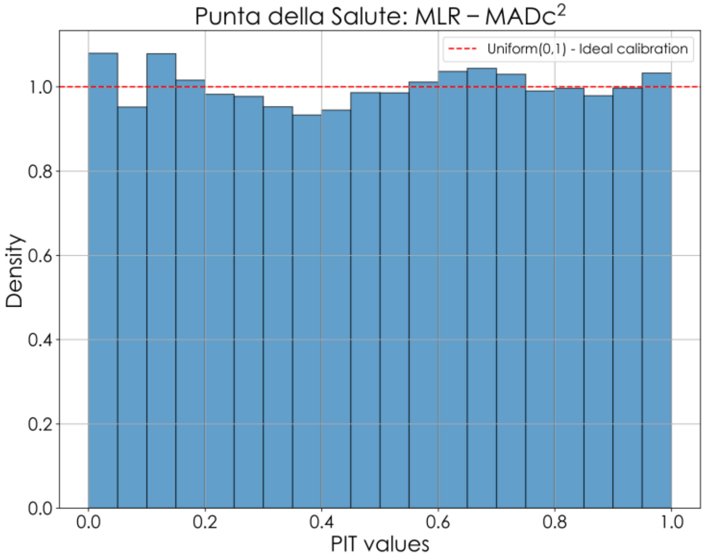


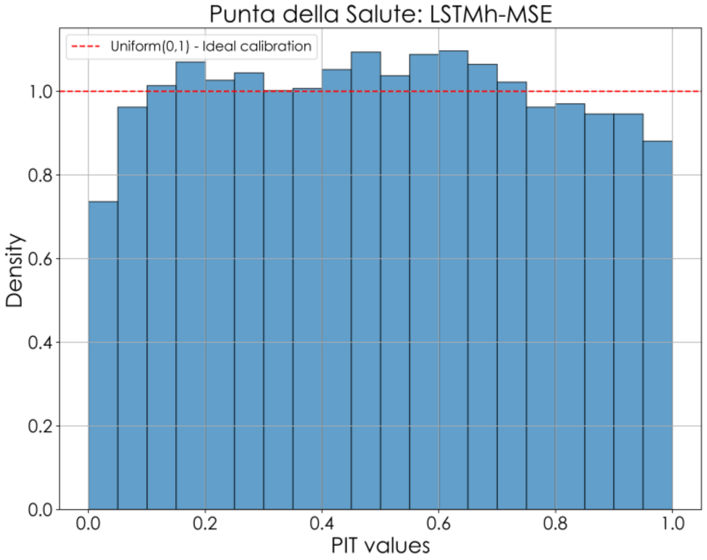
STORM SURGE
MACHINE LEARNING DOWNSCALING

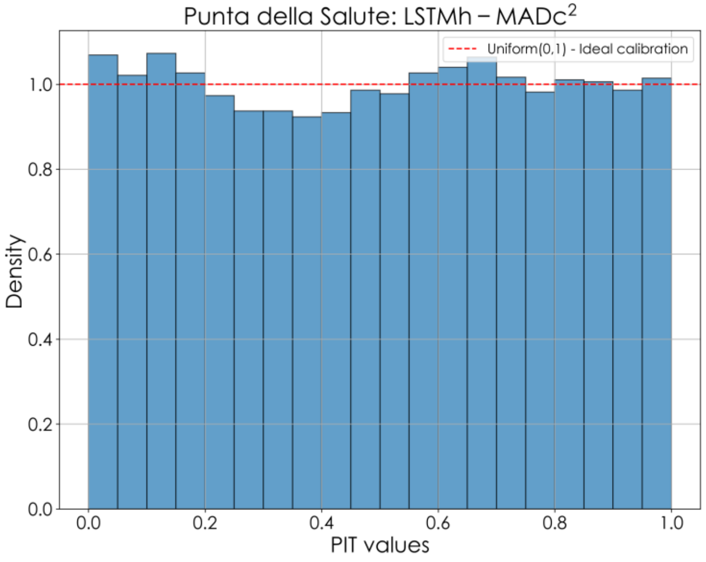
Calibration: Probability Integral Transform (PIT) histogram for full time-series

What does the PIT histogram represent?
Each PIT value corresponds to the position of an observation relative to the emulator's predictive distribution. If the emulator is perfectly calibrated, then PIT values should be uniformly distributed in [0,1].



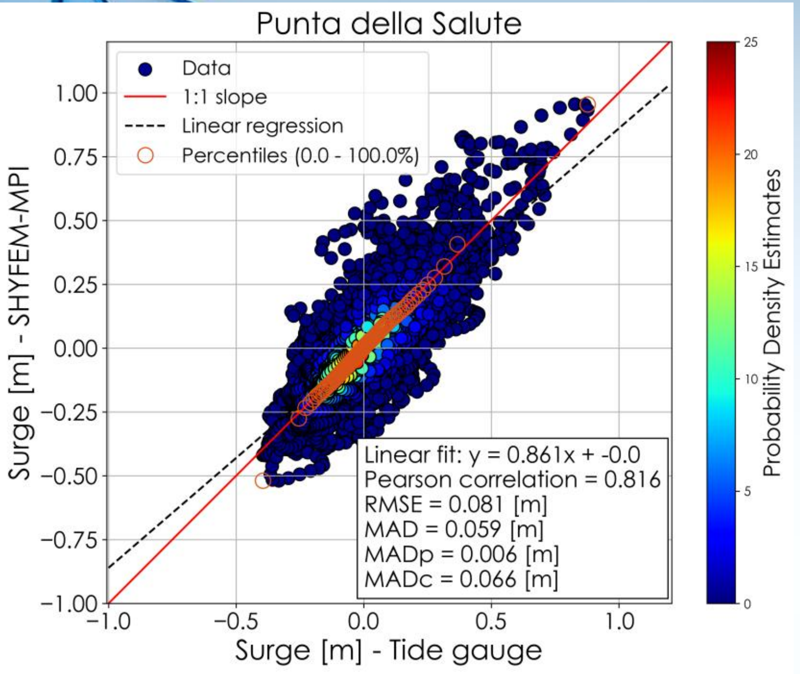


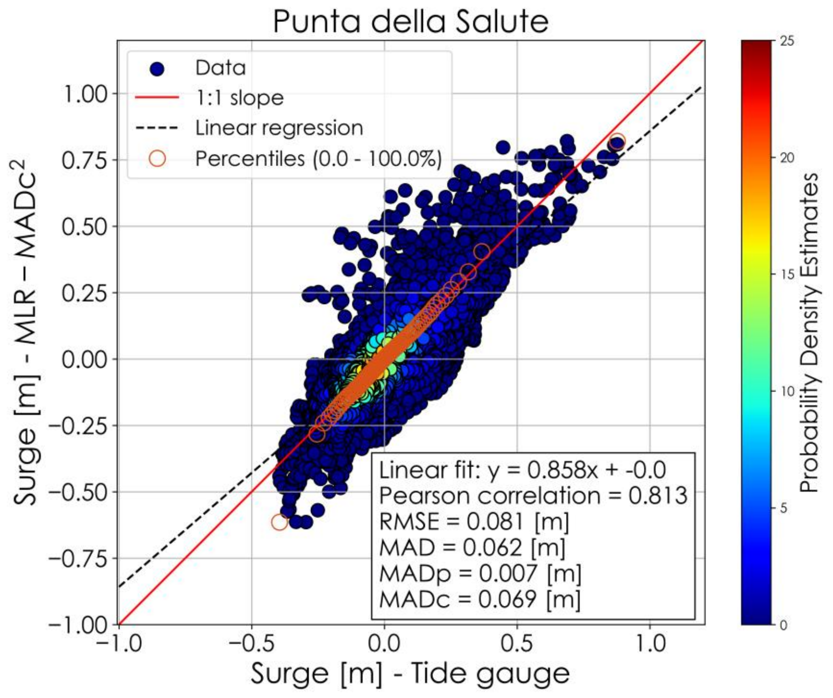


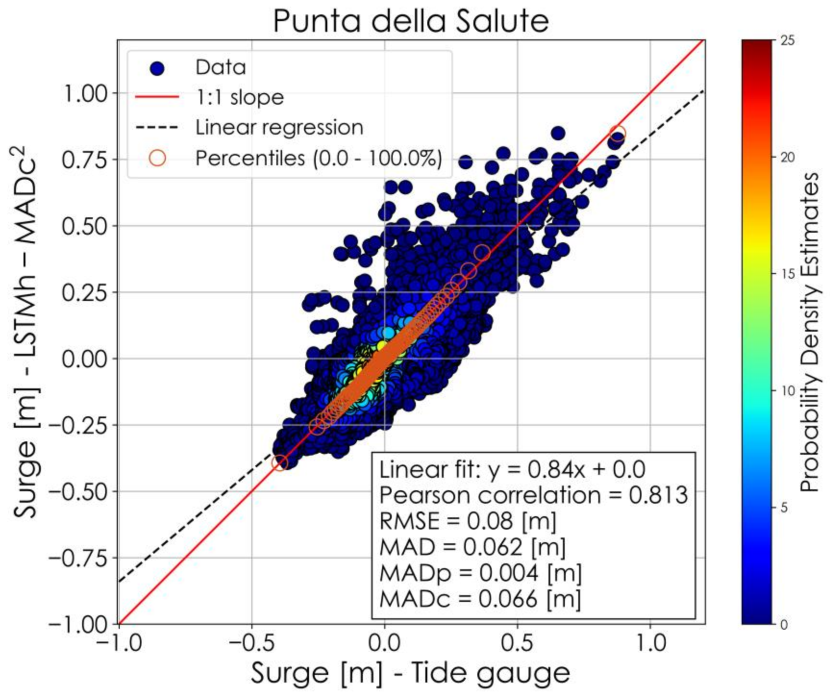


STORM SURGE PREDICTION

MACHINE LEARNING DOWNSCALING





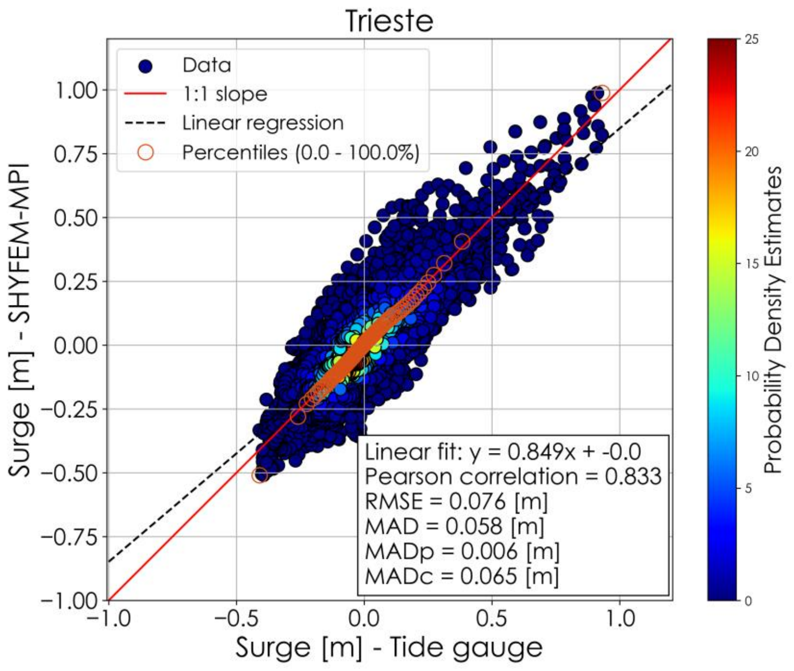


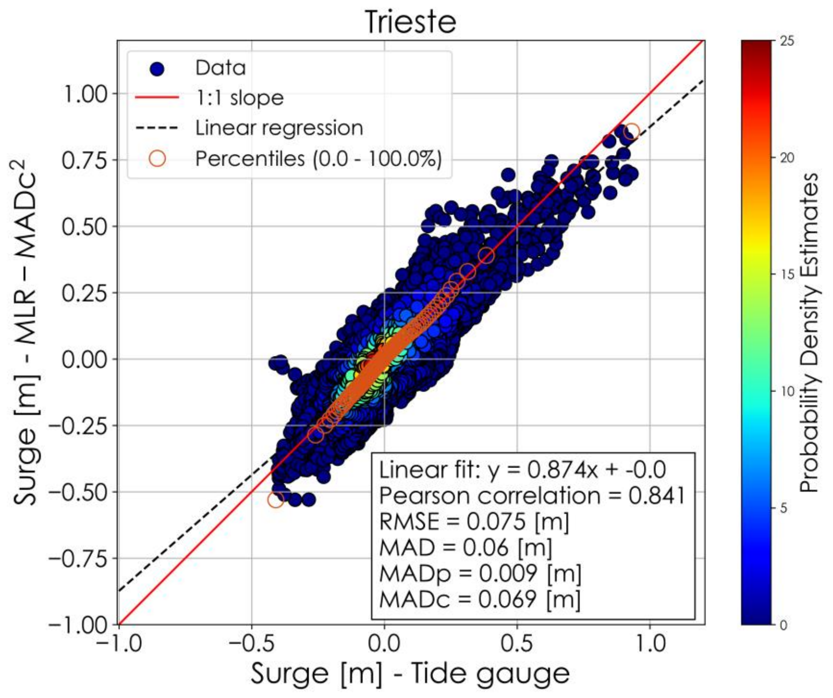
Scatter plots:

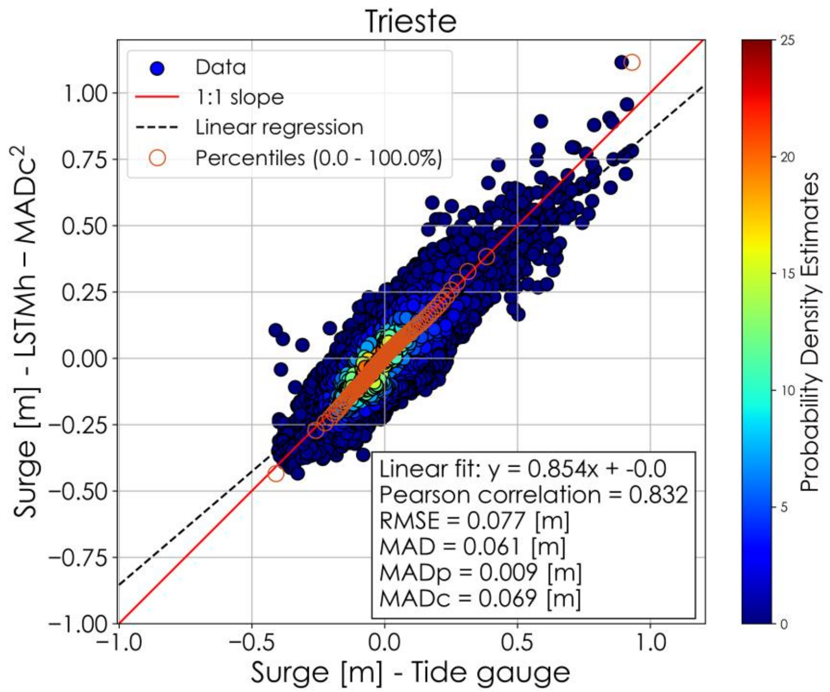
- Performance of ML emulators selected during validation period.

STORM SURGE PREDICTION

MACHINE LEARNING DOWNSCALING







Scatter plots:

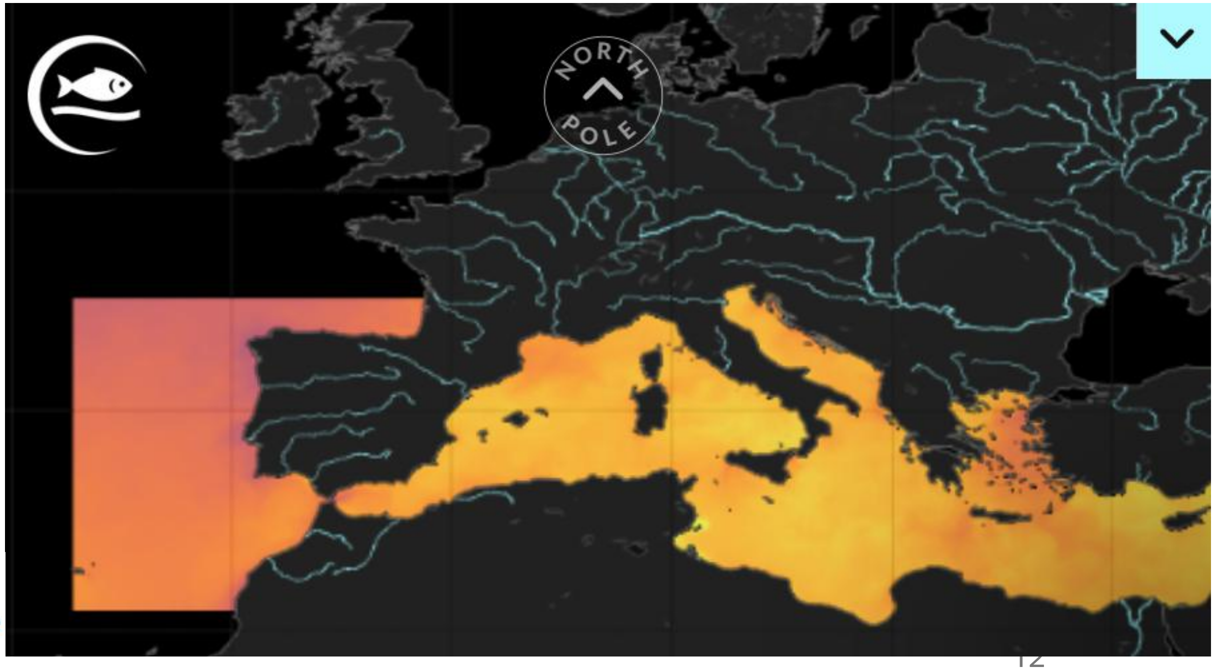
- Performance of ML emulators selected during validation period.

MACHINE LEARNING DOWNSCALING

APPLICATION OF TRAINED MODELS FOR NOVEMBER 2022 STORM SURGE EVENT

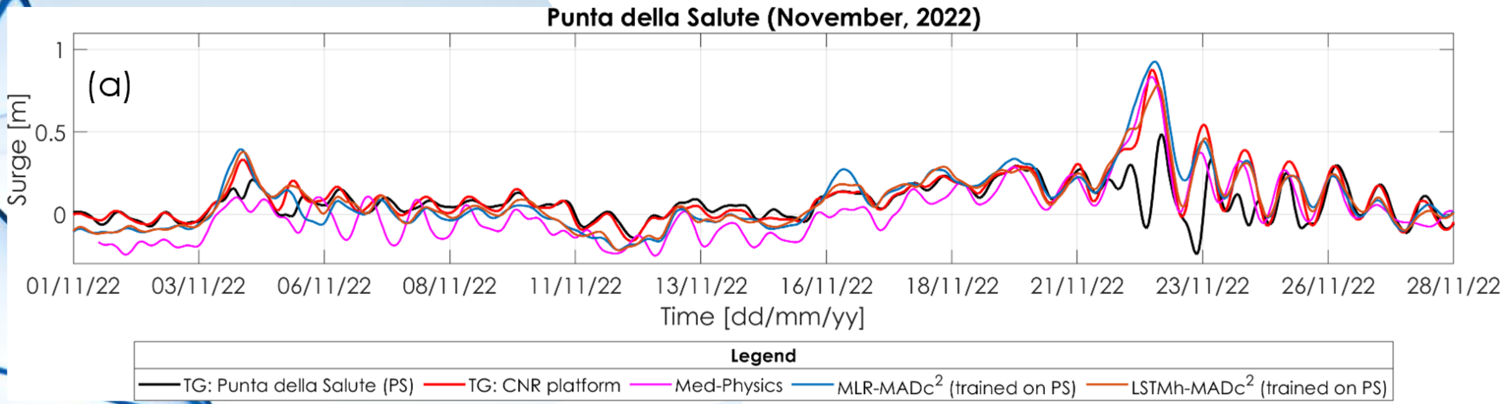
On 22 November 2022, an extreme storm surge event occurred in the northern Adriatic Sea, producing severe damages on its coastline (Mel, et al., 2023).

BENCHMARK: Mediterranean Sea Physics Analysis and Forecast (Med-Physics)

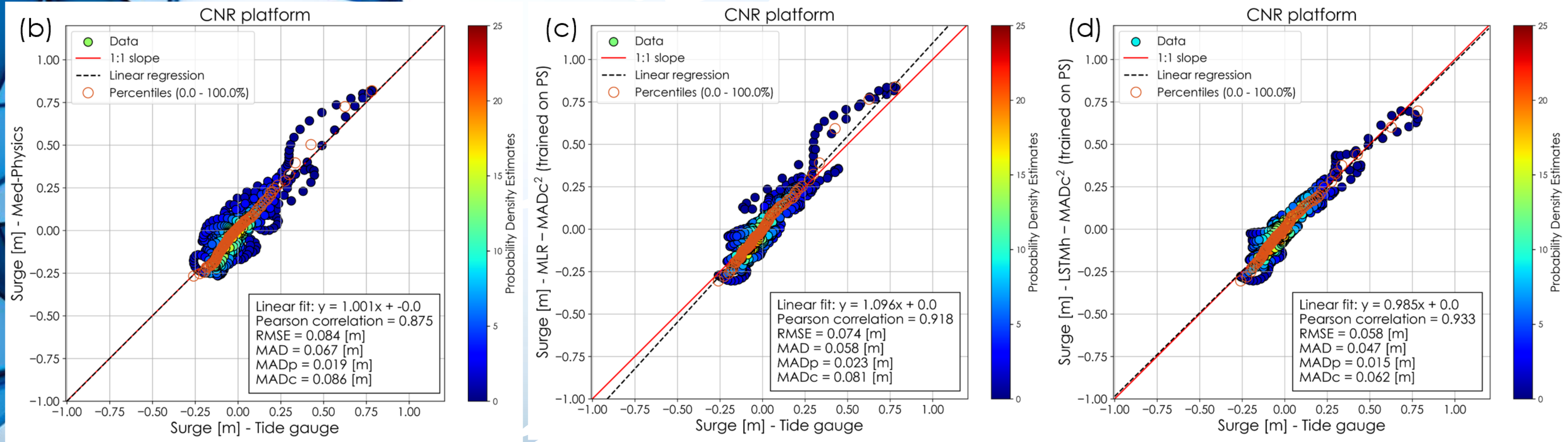


MACHINE LEARNING DOWNSCALING

APPLICATION OF TRAINED MODELS FOR NOVEMBER 2022 STORM SURGE EVENT



Venice and the surrounding urban settlements have been protected from flooding thanks to the operation of the MoSE (Mel et al., 2023).



(**a**) Time series of observed and predicted storm surges; (**b–d**) Scatter plots comparing observed values (at the CNR platform) with predictions from: (**b**) Med-Physics model, (**c**) MLR-MADc^2 trained on Punta della Salute, and (**d**) LSTMh-MADc^2 trained on Punta della Salute.

STORM SURGE DOWNSCALING

KEY POINTS

Machine learning emulators, even relatively simple ones, can match or surpass dynamic downscaling for storm surge time series reconstruction.

The custom MADc^2 loss function greatly enhances prediction of extreme surges.

While complex models tend to perform better, applying MADc^2 even improves simple emulators like linear regression.

MADc^2 trained emulators also proved robust in generalization, retaining strong skill during the November 2022 event.

This work highlights machine learning as a powerful, efficient tool for storm surge time-series reconstruction.

- Dynamic downscaling: 1-year simulation → 36 h on CMCC Zeus supercomputer (36 cores).
- ML emulators: run on a laptop GPU (NVIDIA RTX™ 3000 Ada Generation Laptop GPU with 8 GB) 20–60 s (MLR/MLP), 3–7 min (LSTM emulators).

REFERENCES

Campos-Caba, R., Alessandri, J., Camus, P., Mazzino, A., Ferrari, F., Federico, I., Vousdoukas, M., Tondello, M., and Mentaschi, L. (2024). Assessing storm surge model performance: what error indicators can measure the model's skill?, Ocean Sci., 20, 1513–1526, https://doi.org/10.5194/os-20-1513-2024.

Micaletto, G., Barletta, I., Mocavero, S., Federico, I., Epicoco, I., Verri, G., Coppini, G., Schiano, P., Aloisio, G., & Pinardi, N. (2022). Parallel Implementation of the SHYFEM Model. Geosci. Model Dev., 15, 6025–6046. https://doi.org/10.5194/gmd-2021-319

Mel, R., Coraci, E., Morucci, S., Crosato, F., Cornello, M., Casaioli, M., Mariani, S., Carniello, L., Papa, A., Bonometto, A., & Ferla, M. (2023). Insights on the extreme storm surge event of the 22 November 2022 in the Venice Lagoon. Journal of Marine Science and Engineering. https://doi.org/10.3390/jmse11091750

Umgiesser, G., Canu, D. M., Cucco, A., & Solidoro, C. (2004). A finite element model for the Venice lagoon: Development, set up, calibration and validation. J. Marine Syst., 123–145. https://doi.org/10.1016/j.jmarsys.2004.05.009

ENHANCING STORM SURGE DOWNSCALING: A COMPARATIVE STUDY OF MACHINE LEARNING AND DYNAMICAL MODELING IN THE NORTHERN ADRIATIC SEA

Rodrigo Campos-Caba, Jacopo Alessandri, Paula Camus, Andrea Mazzino, Francesco Ferrari, Ivan Federico, Michalis Vousdoukas, Massimo Tondello, Giovanni Coppini, and Lorenzo Mentaschi

Correspondence to:
Rodrigo Campos-Caba (rodrigo.camposcaba@unibo.it)
Lorenzo Mentaschi (lorenzo.mentaschi@unibo.it)

4th International Workshop on Waves, Storm Surges, and Coastal Hazards
September, 2025