Developments on a machine learning downscaling system for storm surge in the Northern Adriatic Sea

Rodrigo Campos Caba[1], Lorenzo Mentaschi[1], Nadia Pinardi[1], Jacopo Alessandri[1], Paula Camus[2], Massimo Tondello[3], Andrea Mazzino[4], and Franceso Ferrari[4]

[1]Department of Physics and Astronomy, University of Bologna, Bologna, Italy.
[2]Departamento de Ciencias y Técnicas del Agua y del Medio Ambiente, University of Cantabria, Santander, Spain.
[3]HS Marine SrL., Noventa Padovana, Italy.
[4]Department of Civil, Chemical and Environmental Engineering, University of Genoa, Genoa, Italy.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

INTRODUCTION

- The accurate prediction of storm surges remains a critical challenge for coastal communities worldwide. Traditional approaches have primarily relied on dynamical models, requiring significant computational resources and time to produce downscaling (Anderson et al., 2021).
- In recent years, the emergence of machine learning (ML) algorithms has revolutionized the field of met-ocean downscaling and prediction. ML algorithms excel in capturing complex patterns and relationships within vast datasets, enabling more rapid and scalable prediction capabilities compared to dynamical models (Adeli et al., 2023).
- In this work, developments of a ML system for coastal downscaling of storm surge in the Northern Adriatic Sea, trained with the output from a dynamical model, are presented.

DYNAMICAL DOWNSCALING

- Long-term (1987-2020) high-resolution (50 m nearshore, Figure 1) hydrodynamic simulations with the SHYFEM (Micaletto et al., 2021) model.
- Barotropic (BT) and baroclinic (BC) configurations.
- Two different atmospheric databases (wind and mean sea level pressure fields): ERA5 and a dynamical downscaling (3 km horizontal resolution) of the CFSR model (University of Genoa, UniGe).
- Three simulated long-term databases have been obtained: BT-ERA5, BC-ERA5 and BC-UniGe.
- During the performance evaluation (Figure 2) two customized versions of the Mean Absolute Deviation were introduced, MADp and MADc: $MADp = \overline{|S_{prc} - O_{prc}|}$ $\quad MADc = \overline{|S - O|} + MADp$

S_{prc} and O_{prc}: simulation (S) and observation (O) percentile values.

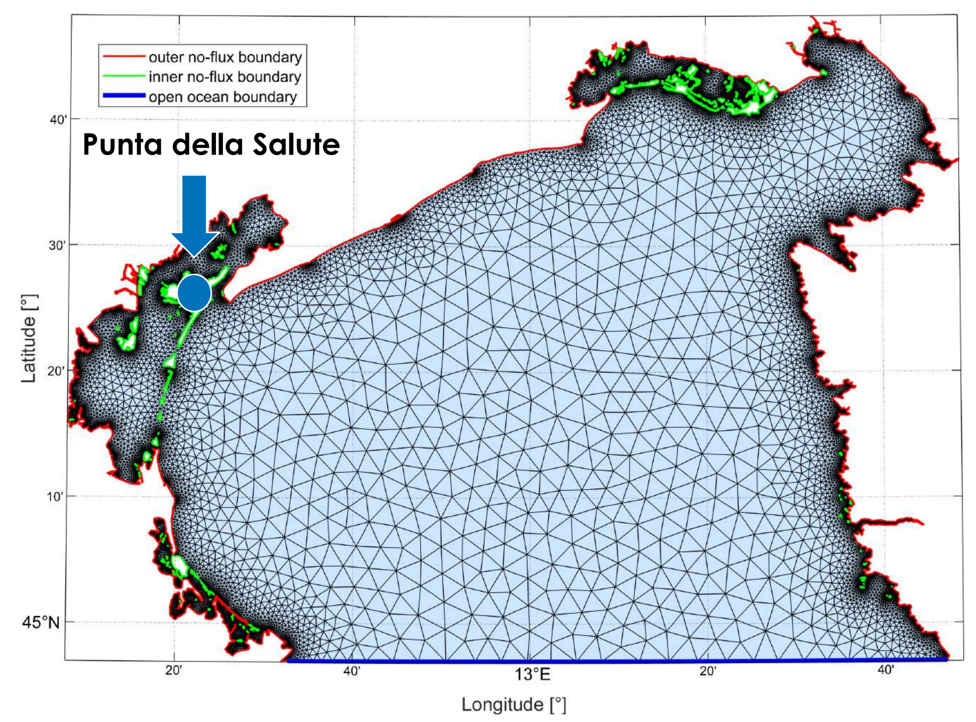

Figure 1: Unstructured grid used for numerical simulations.

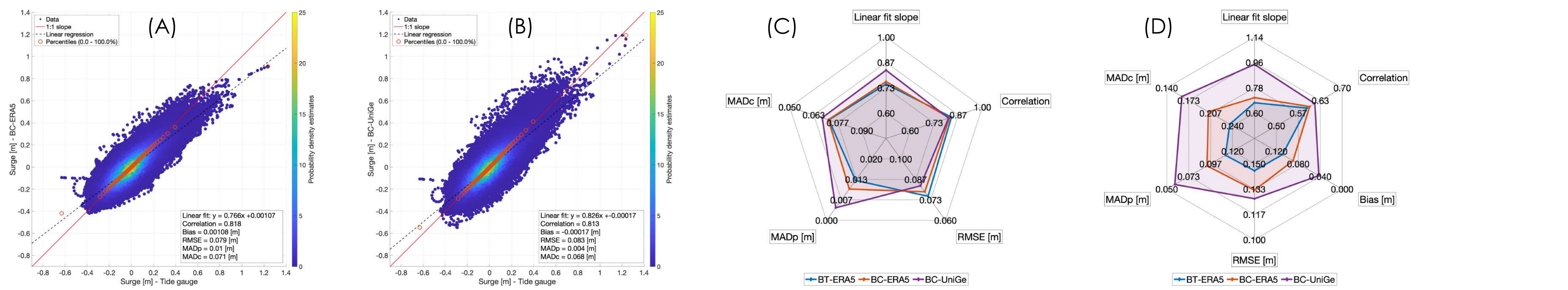

Figure 2: Performance evaluation of the dynamical downscaling in Punta della Salute. (A) and (B): Scatter plots for BC-ERA5 and BC-UniGe, respectively. (C): Radar chart for the total amount of data. (D): Radar chart for surge values above the 99th percentile of the cumulative distribution. In radar charts for RMSE, Bias, MADp and MADc a reverse axis is used, this ensures that simulations covering a larger area on each metric represent a better performance.

MACHINE LEARNING DOWNSCALING

- **Predictand (target):** Storm surge time-series from dynamical downscaling (just baroclinic simulations).
- **Predictors (features):** Sea level non-tidal residuals (Copernicus Sea Physics Reanalysis), tides (FES2014), wind and mean sea level pressure fields (ERA5).
- Principal Component Analysis (PCA) is applied to reduce the spatial dimensionality of predictors. Time series of each predictor is reconstructed considering the first three principal components.
- In the training period 80% of the data is used, and the 20% remaining is for testing. The data is organized in a way that allows extreme events to be included during training.
- **Algorithms:** Multivariate Linear Regression (MLR), Long Short-Term Memory (LSTM) network.
- **Location**: Punta della Salute.

RESULTS MACHINE LEARNING DOWNSCALING

- Better performance is observed when the predictand (target) is the output from BC-ERA5. Under these conditions, the performance of MLR is comparable to more sophisticated methods such as LSTM.
- However, when the target corresponds to the output from BC-UniGe, the algorithms exhibit decreased precision, particularly in the representation of the highest values.
- This limitation is slightly mitigated with the use of LSTM.

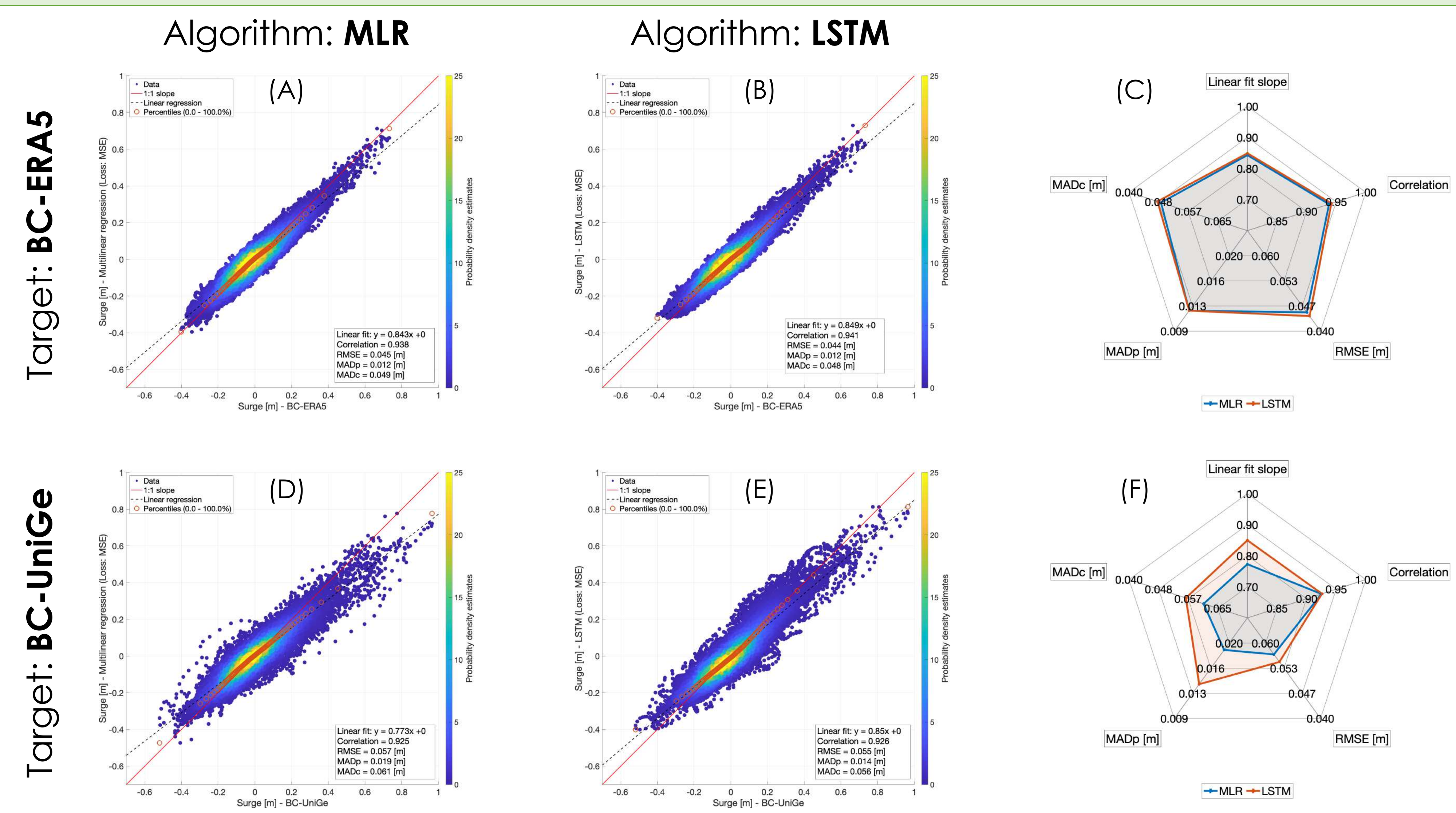

Figure 3: Performance evaluation of ML downscaling. (A) and (B): Scatter plots for MLR and LSTM using BC-ERA5 as target, respectively. (C): Radar chart for MLR and LSTM using BC-ERA5 as target. (D) and (E): Scatter plots for MLR and LSTM using BC-UniGe as target, respectively. (F): Radar chart for MLR and LSTM using BC-UniGe as target.

CURRENT WORK

- Tests for possible improvements with LSTM, mainly when BC-UniGe is used as target.
- Application of algorithms on different locations along the coast of the domain showed on Figure 1.

REFERENCES

- Adeli, E., Sun, L., Wang, J., and Taflanidis. (2023). An advanced spatio-temporal convolutional recurrent neural network for storm surge predictions. Neural Computing and Applications 35:18971-18987. https://doi.org/10.1007/s00521-023-08719-2
- Anderson, D.L., Ruggiero, P., Mendez, F.J., Barnard, P.L., Erikson, L.H., O'Neill, A.C., Merrifield, M., Rueda, A., Cagigal, L., and Marra, J. (2021). Projecting climate dependent coastal flood risk with a hybrid statistical dynamical model. Earth's Future, 9, e2021EF002285. https://doi.org/10.1029/2021EF002285.
- Micaletto, G., Barletta, I., Mocavero, S., Federico, I., Epicoco, I., Verri, G., Coppini, G., Schiano, P., Aloisio, G., and Pinardi, N. (2021). Parallel implementation of the SHYFEM model. Geoscientific Model Development. https://doi.org/10.5194/gmd-2021-319