Supplement of Nat. Hazards Earth Syst. Sci., 26, 3683–3706, 2026
https://doi.org/10.5194/nhess-26-3683-2026-supplement

Natural Hazards and Earth System Sciences
Open Access

Supplement of

Storm surge dynamics in the northern Adriatic Sea: comparing AI emulators with high-resolution numerical simulations

Rodrigo Campos-Caba et al.

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

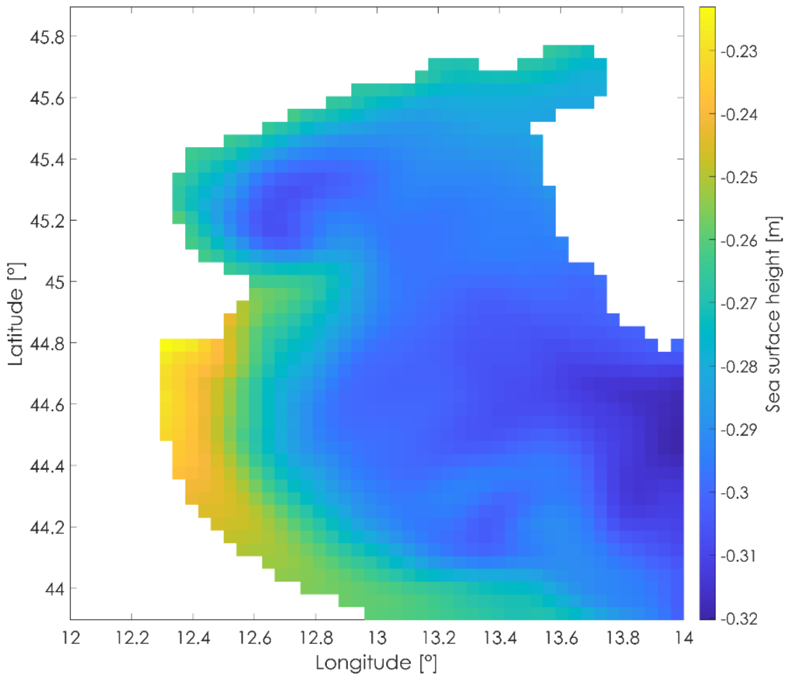

Figure S1: Spatial domain of predictors, exemplified by sea surface height at a random time from Med-MFC database.

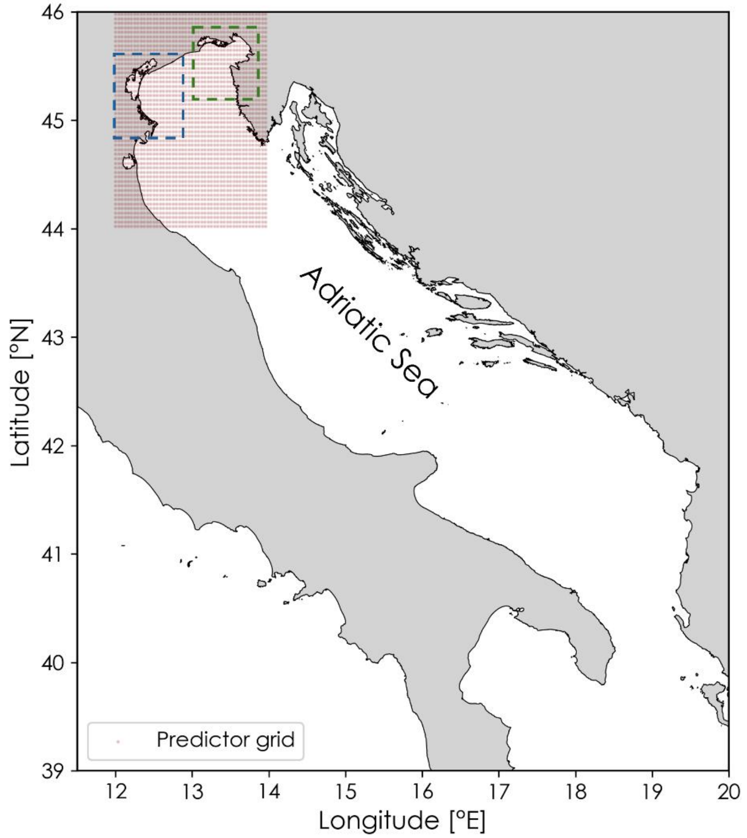

Figure S2: Predictor domains evaluated during the domain-selection procedure. The dashed blue and green boxes denote the regional domains tested for Punta della Salute and Trieste, respectively, while the red dots indicate the final predictor domain retained for the experiments.

Table S1: Validation-set MADc obtained with the selected run of the MLR-MSE emulator using 7 principal components, for the different predictor-domain sizes evaluated during domain selection. Regional domain refers to the domains enclosed by the dashed blue (Punta della Salute) and green (Trieste) boxes, whereas larger domain corresponds to the final predictor domain indicated by the red dots in Figure S2.

	Regional domain	Larger domain
Trieste	0.098 m	0.096 m
Punta della Salute	0.086 m	0.085 m
Average	0.092 m	0.091 m

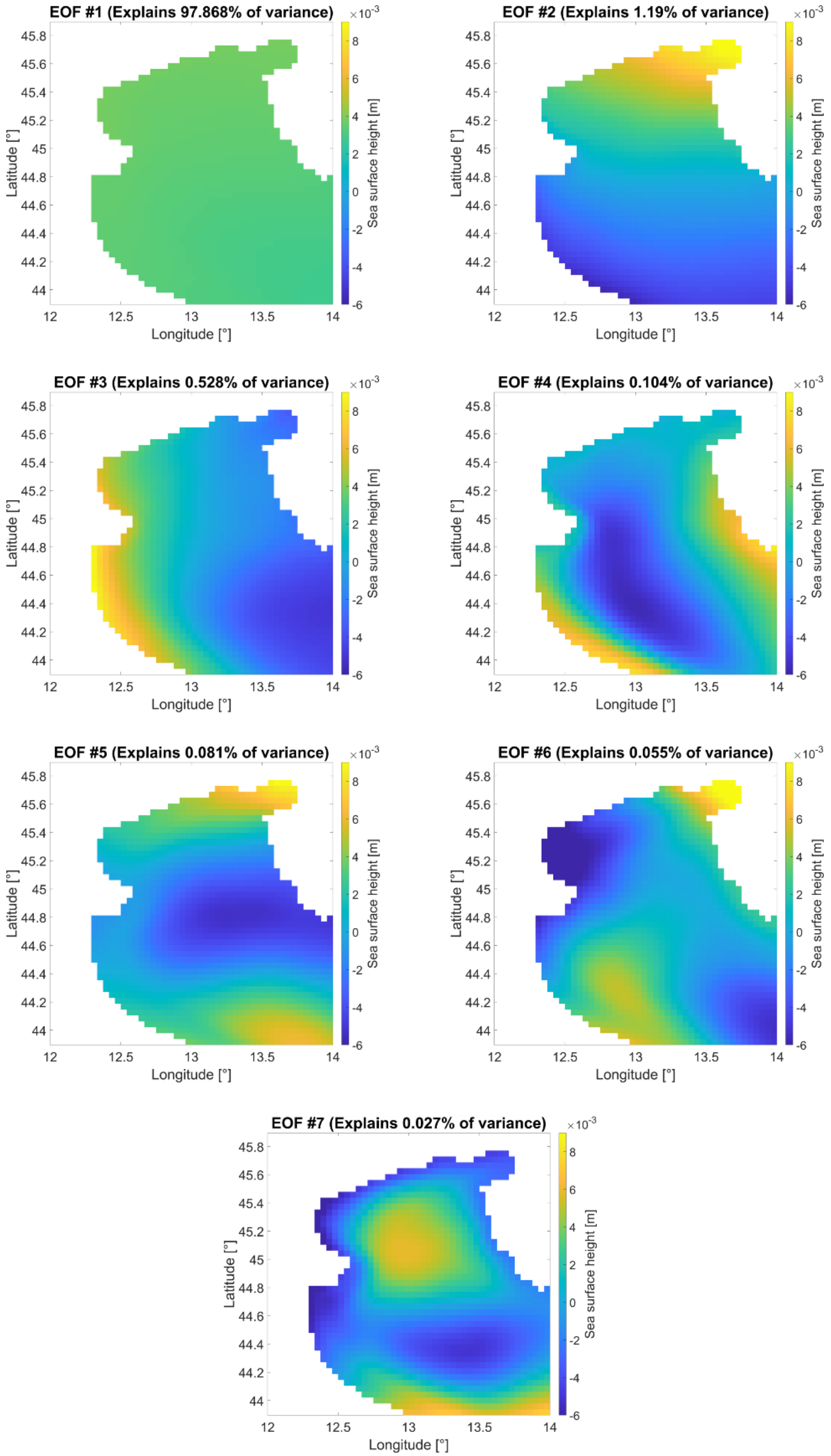

Figure S3: First seven Empirical Orthogonal Functions (EOFs) for sea surface height.

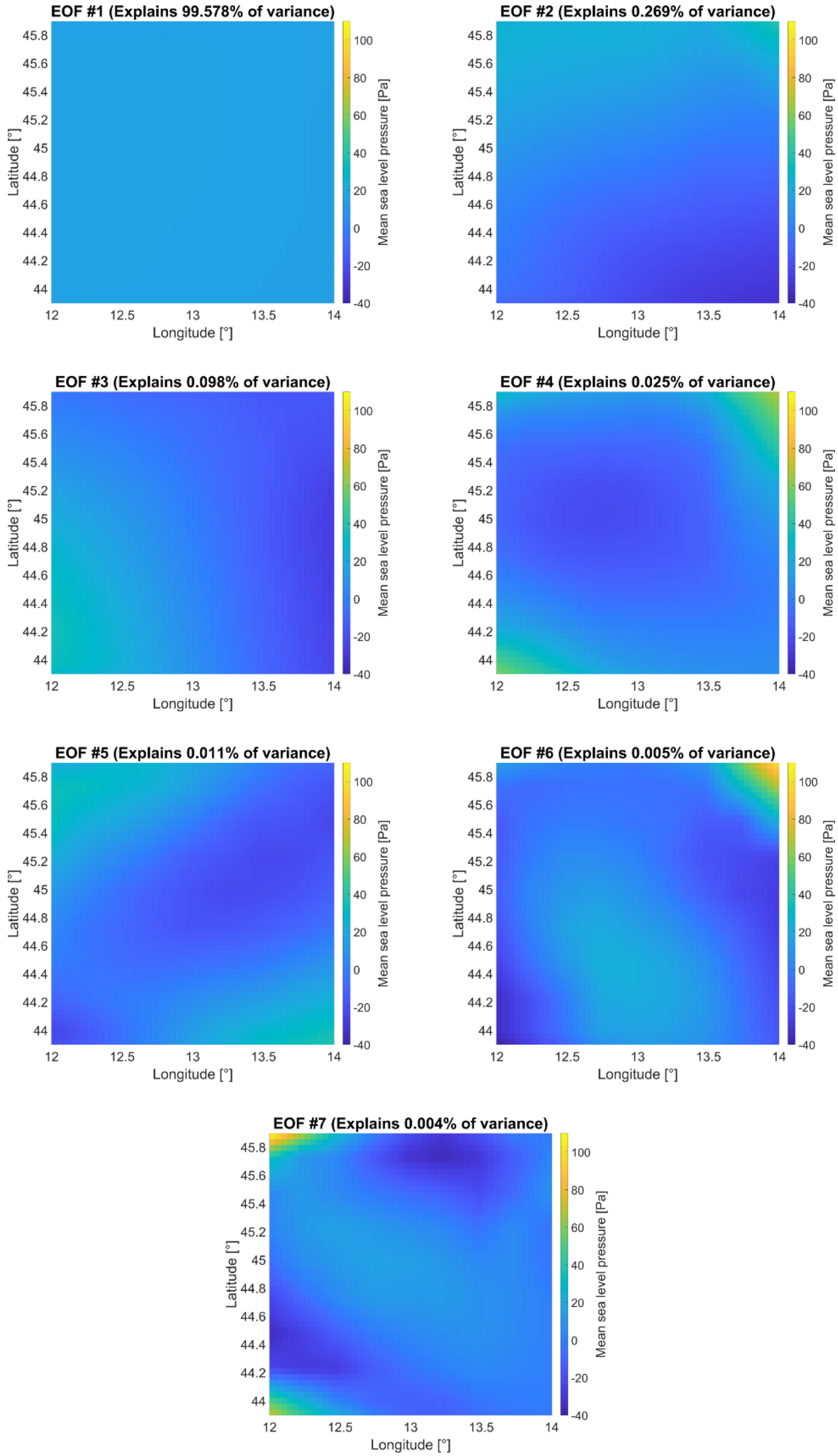

Figure S4: First seven Empirical Orthogonal Functions (EOFs) for mean sea level pressure.

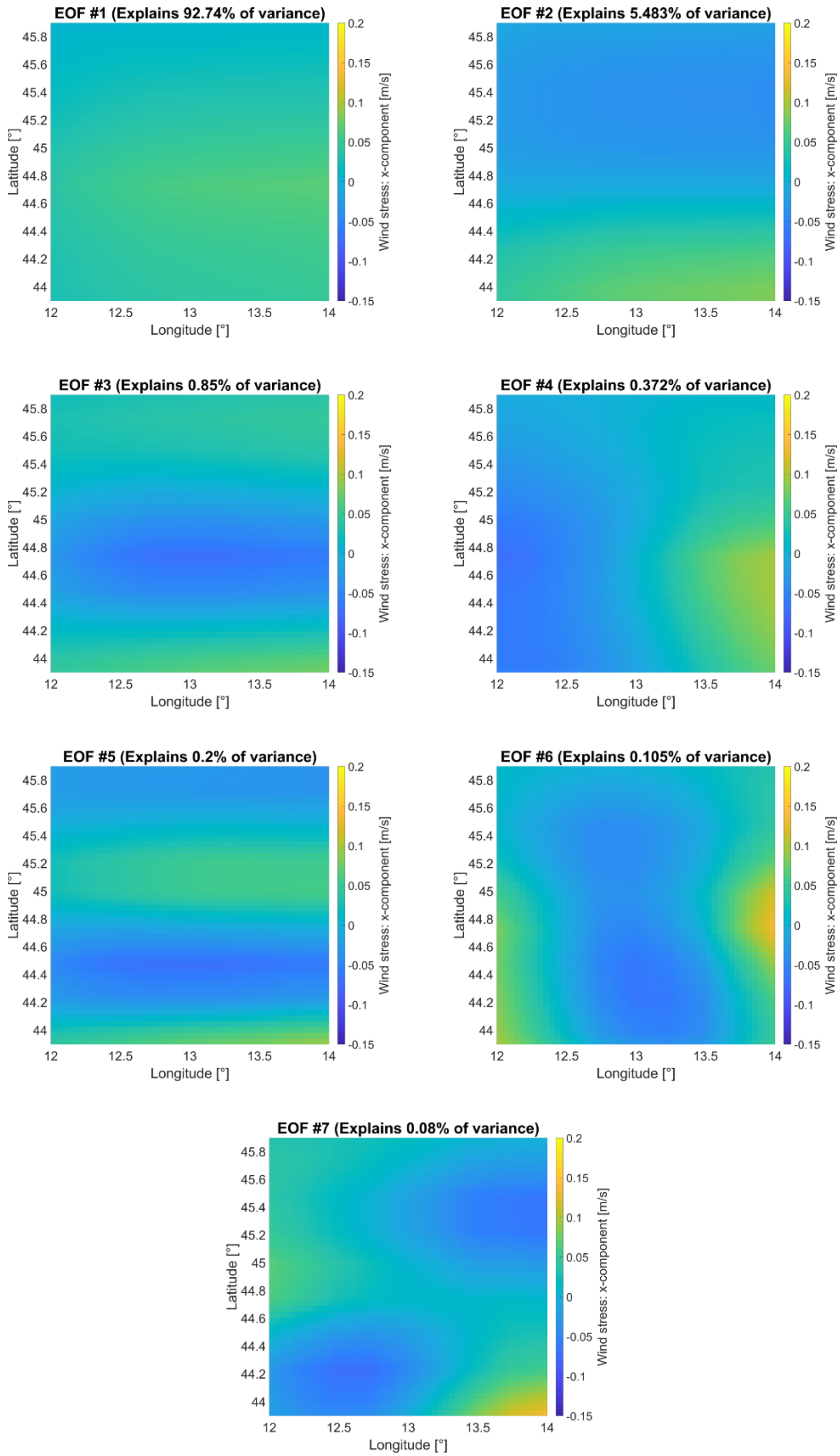

Figure S5: First seven Empirical Orthogonal Functions (EOFs) for x-component of wind stress.

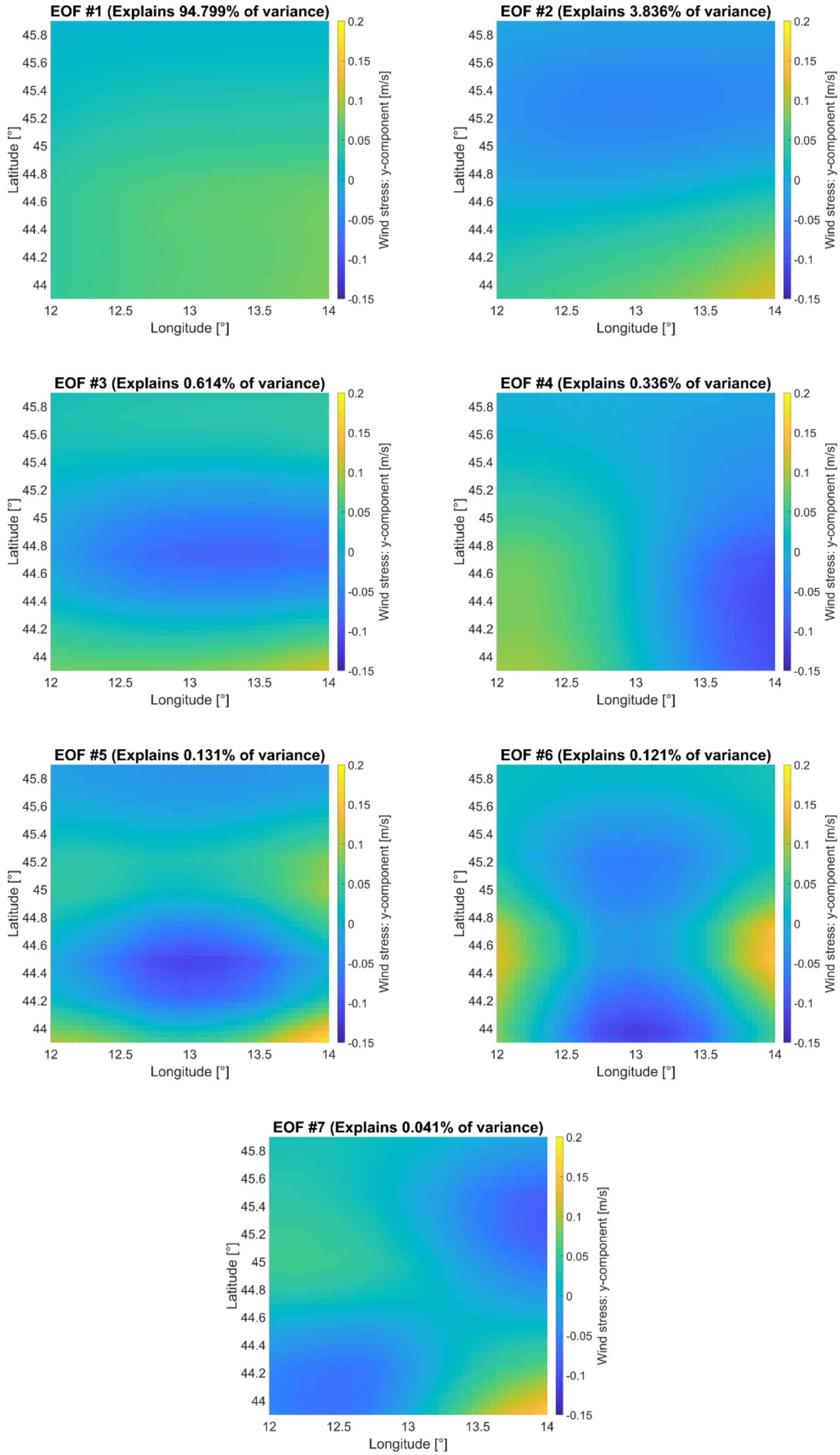

Figure S6: First seven Empirical Orthogonal Functions (EOFs) for y-component of wind stress.

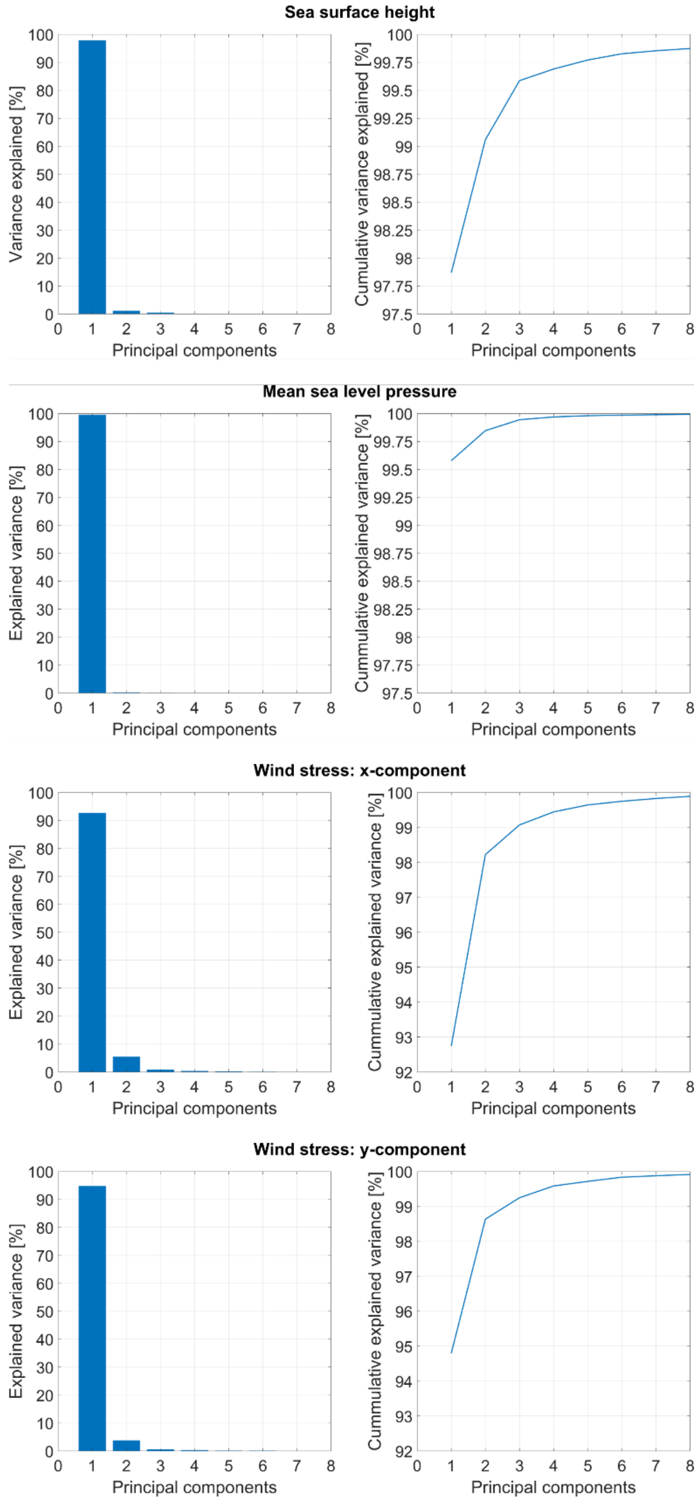

Figure S7: Explained variance of each principal component and cumulative explained variance for the predictors.

Table S2: Validation-set MADc obtained with the selected run of the MLR-MSE emulator using 3, 5, and 7 principal components (PCs).

	3 PCs	5 PCs	7 PCs
Trieste	0.098 m	0.098 m	0.096 m
Punta della Salute	0.087 m	0.086 m	0.085 m
Average	0.093 m	0.092 m	0.091 m

Table S3: Performance metrics of the LSTM-MSE emulator for the comparison between the full-time series + TBPTT approach and the sliding window (look-back) formulations, evaluated for surge peaks exceeding the 99th percentile and averaged across locations. Look-back configurations are defined by the number of hours included in the input window. The best-performing method for each metric is highlighted in bold.

Method	**Slope**	**Pearson correlation**	**RMSE [m]**	**Bias [m]**	**MAD [m]**	**MADp [m]**	**MADc [m]**
Full-time + TBPTT	0.742	**0.725**	**0.121**	**-0.058**	**0.099**	**0.058**	**0.157**
Look-back 3 hours	**0.755**	0.652	0.158	-0.083	0.123	0.088	0.210
Look-back 12 hours	0.623	0.660	0.142	-0.091	0.115	0.090	0.204
Look-back 24 hours	0.747	0.691	0.151	-0.099	0.121	0.096	0.217
Look-back 36 hours	0.608	0.599	0.162	-0.102	0.131	0.099	0.229

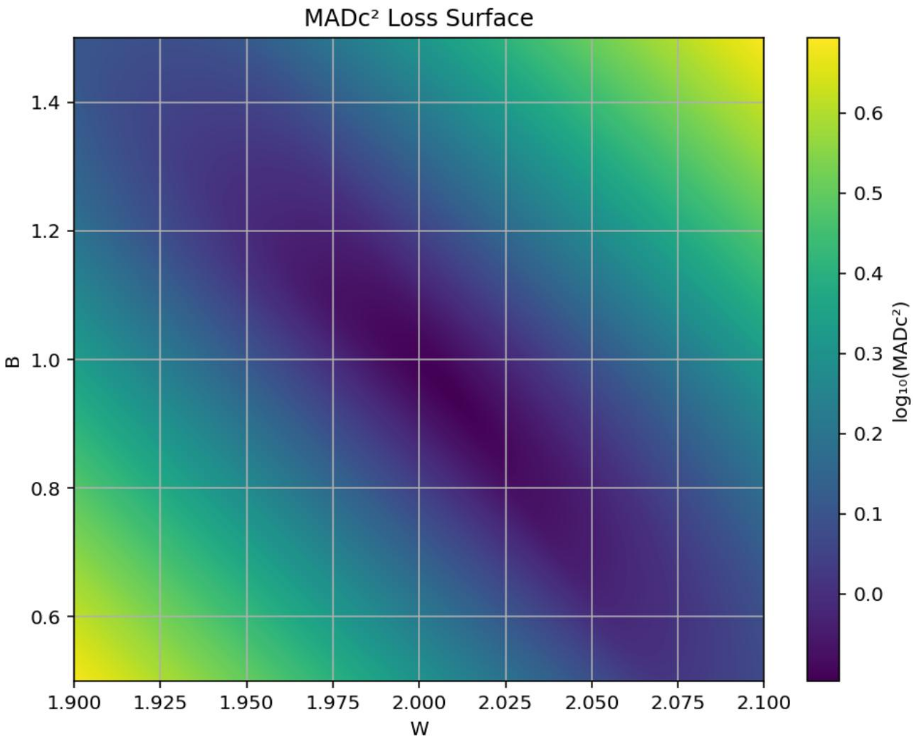

Figure S8: Synthetic MADc^2 loss surface for the univariate linear regression model $\hat{y} = Wx + B$. The surface is computed over a grid of weight (W) and bias (B) values around the true coefficients ($W = 2, B = 1$). Colors represent $log_{10}(MADc^2)$. The unique minimum near the true parameters illustrates the smooth and well-behaved optimization landscape in this simplified setting.

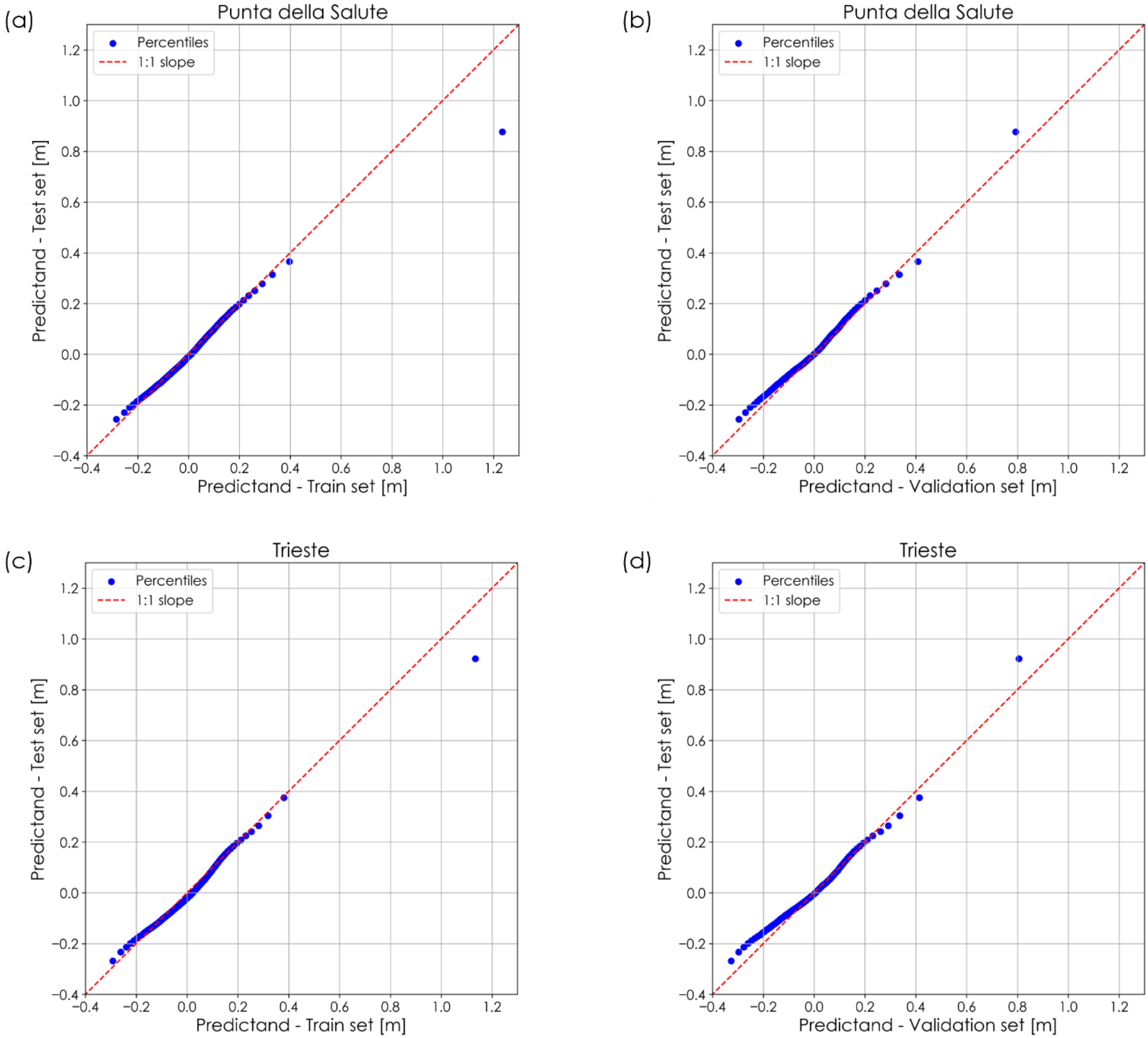

Figure S9: Percentiles comparison between training and test sets (a and c) and validation and test sets (b and d).

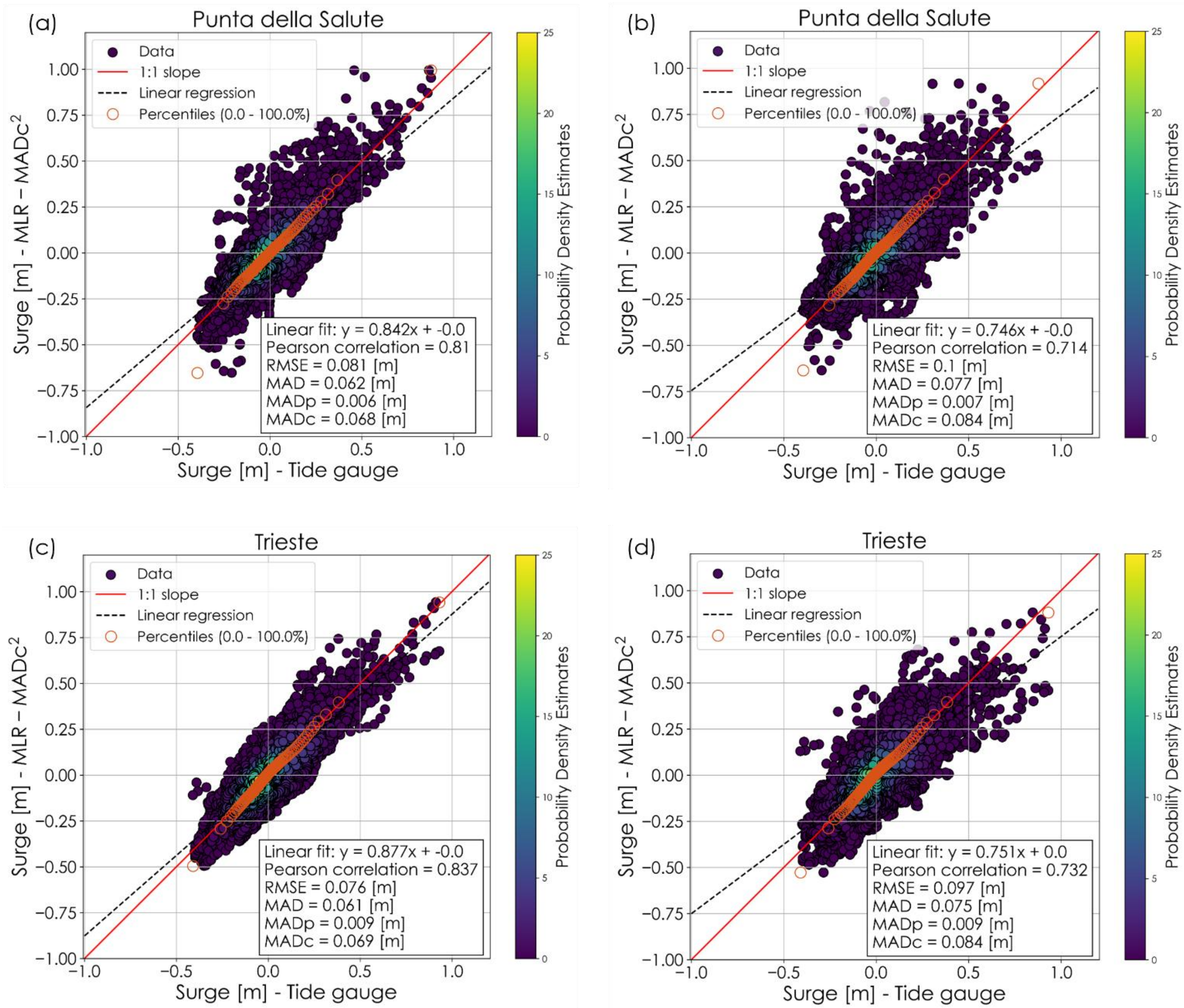

Figure S10: Comparison of ML emulator performance using (a, c) SSH-informed predictors (basin-scale sea surface height, tides, and atmospheric fields) and (b, d) atmospheric-only predictors (mean sea level pressure and wind stress components). Panels (a) and (b) correspond to Punta della Salute, while panels (c) and (d) correspond to Trieste. Although atmospheric-only configurations exhibit reduced correlation and increased error relative to SSH-informed models, they retain substantial skill in reproducing surge variability and higher-percentile behavior.

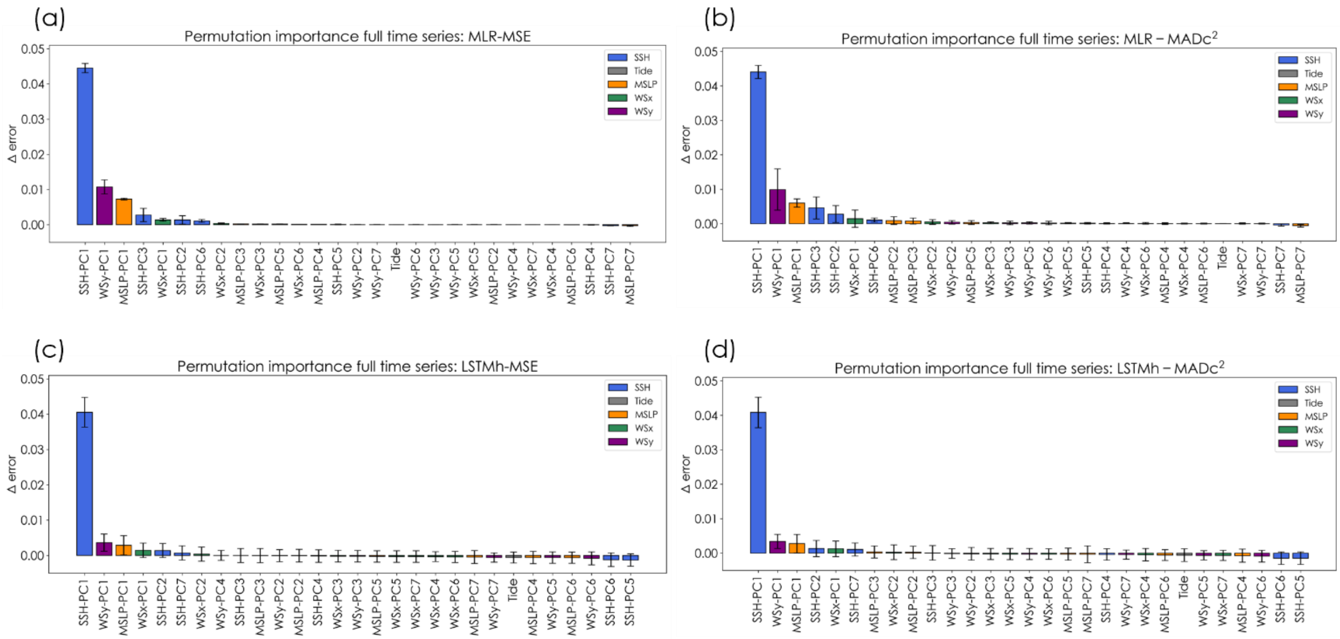

Figure S11: Permutation importance averaged by locations for full time series. (a) MLR-MSE, (b) MLR-MADc^2, (c) LSTMh-MSE, and (d) LSTMh-MADc^2.

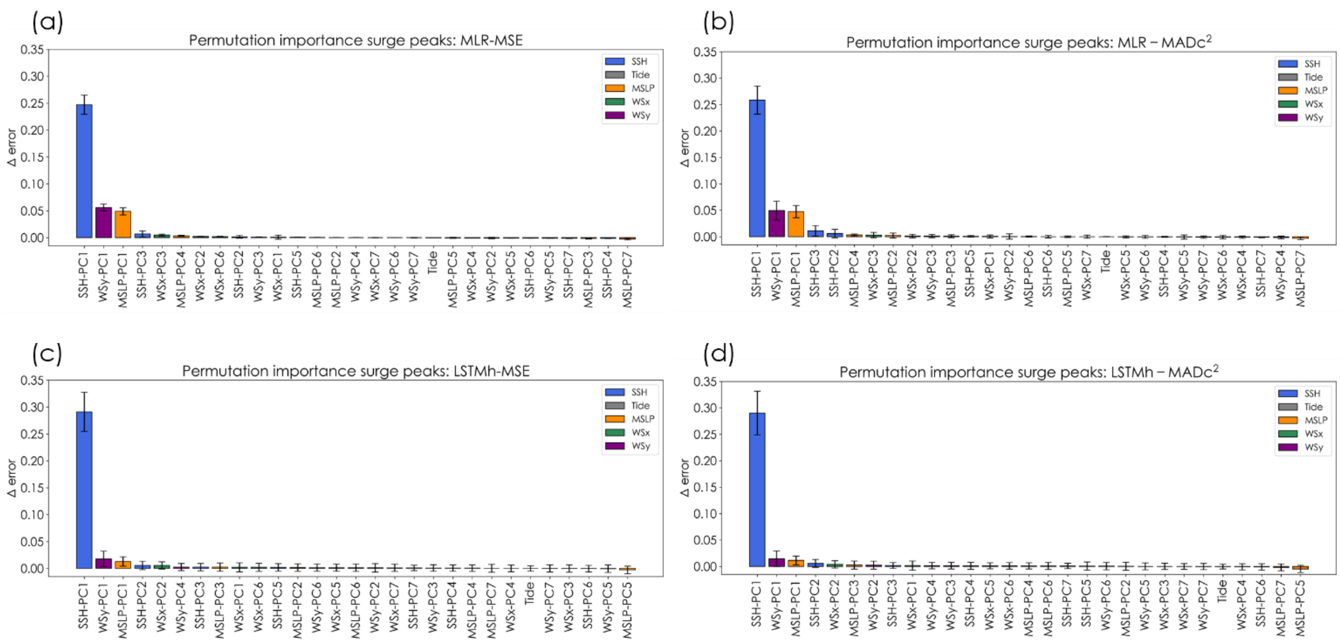

Figure S12: Permutation importance averaged by locations for surge peaks above 99th percentile. (a) MLR-MSE, (b) MLR-MADc^2, (c) LSTMh-MSE, and (d) LSTMh-MADc^2.